UNIVERSIDADE FEDERAL DE OURO PRETO

Estudo de Padrões de Execuções Musicais em Programações de Estações de Rádio

Alexandre Pereira Norberto
Universidade Federal de Ouro Preto

Orientador: Álvaro Rodrigues Pereira Júnior

Dissertação submetida ao Instituto de Ciências Exatas e Biológicas da Universidade Federal de Ouro Preto para obtenção do título de Mestre em Ciência da Computação

Ouro Preto, outubro de 2015

Estudo de Padrões de Execuções Musicais em Programações de Estações de Rádio

Alexandre Pereira Norberto
Universidade Federal de Ouro Preto

Orientador: Álvaro Rodrigues Pereira Júnior

N824e Norberto, Alexandre Pereira.
Estudo de padrões de execuções musicais em programações de estações de rádio [manuscrito] / Alexandre Pereira Norberto. - 2015.
138f.: il.: color; grafs; tabs.

Orientador: Prof. Dr. Álvaro Rodrigues Pereira Júnior.

Dissertação (Mestrado) - Universidade Federal de Ouro Preto. Instituto de Ciências Exatas e Biológicas. Departamento de Computação. Programa de Pós-Graduação em Ciência da Computação.
Área de Concentração: Ciência da Computação.

1. Recuperação da informação. 2. Mineração de dados (Computação). 3. Música. 4. Rádio. 5. Televisão. I. Pereira Júnior, Álvaro Rodrigues. II. Universidade Federal de Ouro Preto. III. Titulo.

CDU: 004.62:78

Catalogação: www.sisbin.ufop.br

MINISTÉRIO DA EDUCAÇÃO E DO DESPORTO
Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Biológicas – ICEB
Programa de Pós-Graduação em Ciência da Computação

Ata da Defesa Pública de Dissertação de Mestrado

Ao primeiro dia do mês de outubro de 2015, às 10 horas no Salão Nobre da Escola de Minas, reuniram-se os membros da banca examinadora composta pelos professores: **Prof. Dr. Álvaro Rodrigues Pereira Júnior (presidente e orientador), Profa. Dra. Amanda Sávio Nascimento e Silva, Prof. Dr. Ivair Ramos Silva, Profa. Dra. Geane Carvalho Alzamora e Prof. Dr. Rodrygo Luis Teodoro Santos**, aprovada pelo Colegiado do Programa de Pós-Graduação em Ciência da Computação, a fim de arguirem o mestrando **Alexandre Pereira Norberto**, com o título **"Estudo de padrões de execuções musicais em programações de estações de rádio"**. Aberta a sessão pelo presidente, coube ao candidato, na forma regimental, expor o tema de sua dissertação, dentro do tempo regulamentar, sendo em seguida questionado pelos membros da banca examinadora, tendo dado as explicações que foram necessárias.

Recomendações da Banca:

(X) Aprovada sem recomendações

( ) Reprovada

( ) Aprovada com recomendações: _____

Banca Examinadora:

Prof. Dr. Álvaro Rodrigues Pereira Júnior

Profa. Dra. Amanda Sávio Nascimento e Silva

Prof. Dr. Ivair Ramos Silva

Profa. Dra. Geane Carvalho Alzamora

Prof. Dr. Rodrygo Luis Teodoro Santos

Prof. Dr. Luiz Henrique de Campos Merschmann
Coordenador do Programa de Pós-Graduação em Ciência da Computação
DECOM/ICEB/UFOP

Ouro Preto, 01 de outubro de 2015.

Dedico este trabalho a Deus*, aos meus pais* Daniel *e* Maria Aparecida*, a minha irmã* Gisele*, a minha namorada* Fernanda *e aos meus amigos.*

Resumo

Por ser um meio de comunicação em massa, o rádio apresenta grande importância comercial e política, tal como a televisão, porém com características diferentes. O rádio investe em um estilo de programação com forte componente musical, reforçando uma parceria rentável com a indústria fonográfica. Por outro lado, a indústria fonográfica, juntamente com os artistas do ramo musical e estudiosos da área da comunicação, têm interesse em saber como os produtos (músicas) estão sendo veiculados e consumidos pelas pessoas através dos meios de comunicação, além de como esses veículos interagem para a divulgação desses produtos e de seus conteúdos. Ao se fazer análises na programação de uma rádio, pode-se extrair informações muito interessantes, como por exemplo, se existe um estilo musical predominante e qual seria o padrão de execuções musicais em sua programação. Existe também a hipótese de que os locutores de rádio podem observar acontecimentos externos, como a repercursão de determinados artistas em outras mídias, como aparições em programas de televisão (TV), o que pode impactar na quantidade de execuções de músicas daqueles artistas e na maneira que os locutores definem suas programações musicais nas rádios. Nesse sentido, este trabalho investiga a seguinte hipótese: "A presença de artistas em programas de TV de grande audiência impacta positivamente no número de execuções desses artistas em rádios?". Para isso, foi desenvolvida uma metodologia de análise da possível influência da mídia televisiva sobre a forma com que as programações das rádios são definidas. Embora alguns artistas da música com quem conversamos acreditam que a hipótese seja verdadeira, nossa investigação mostra que estatísticamente a hipótese não pode ser confirmada, sendo então rejeitada.

Abstract

Being a means of mass communication, radio has great commercial and political importance, such as television, but with different characteristics. Radio invests in a programming style with strong musical component, reinforcing a profitable partnership with the music industry. On the other hand, the music industry, along with the artists of the music business and scholars of the field of communication, is interested in how the products (songs) are being served and consumed by people through the media, and how these vehicles interact to the disclosure of these products and their contents. When doing analysis in radio programming, it is possible to find very interesting information, such as if there is a predominant musical style and what is the standard of musical performances on their programming. There is also the possibility that the broadcasters can observe external events such as the repercussion of certain artists in other media such as participations on television programs (TV), which can impact on the amount of executions of those artists and the way that broadcasters define their musical programs on the radio. In this sense, this work investigates the following hypothesis: "Does the presence of music artists in TV programs with high audience impact positively on the number of executions of these artists on radio?". In order to investigate this hypothesis, it has been developed a methodology for analyzing the possible influence of television media on the way the programming of radios are set. Although some music artists we have chatted claim the hypothesis might be true, our investigation shows that statistically the hypothesis cannot be confirmed, being then rejected.

Declaração

Esta dissertação é resultado de meu próprio trabalho, exceto onde referência explícita é feita ao trabalho de outros, e não foi submetida para outra qualificação nesta nem em outra universidade.

Alexandre Pereira Norberto

Agradecimentos

Agradeço a Deus, por me possibilitar essa e todas as conquistas de minha vida.

Agradeço a meus pais, Daniel e Maria Aparecida, a minha irmã Gisele e a minha namorada Fernanda, por toda força, carinho e amor.

Agradeço a toda minha família, pelo aconchego.

Agradeço a todos os meus amigos, pelos auxílios, companheirismo e bons momentos.

Agradeço ao professor Álvaro, meu orientador.

Agradeço à professora Amanda e ao professor Luiz, que auxiliaram no desenvolvimento deste trabalho.

Agradeço aos demais professores, pelos ensinamentos.

Agradeço ao Radialize e à SoundView pelo fornecimento dos dados para a realização deste trabalho.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Por fim, agradeço a todos que me ajudaram direta ou indiretamente neste trabalho.

Muito Obrigado.

Sumário

Lista de Figuras **xii**

Lista de Tabelas **xiii**

1 Introdução **1**

1.1 Objetivos 3

1.2 Justificativa 3

1.3 Definição do problema 5

1.4 Organização do texto 6

2 Fundamentação teórica **7**

2.1 A indústria da música 7

2.2 O rádio e a sua grade de programação 9

2.3 A ferramenta Radialize e o serviço da SoundView 10

2.4 Tarefas e técnicas de mineração de dados 12

2.4.1 Predições numéricas 13

2.4.2 Mineração de regras de associação 14

2.5 Teste de hipóteses 15

2.5.1 Teste de Friedman 16

2.5.2 Teste de Wilcoxon Pareado 17

3 Trabalhos relacionados **19**

3.1 Trabalhos associados à área da computação 19

3.2 Trabalhos associados à área da comunicação 21

4 Análise de coocorrência e correlação **24**

4.1 Metodologia de análise de coocorrência e correlação 25

4.1.1 Preparação das bases transacionais 26

4.1.2 Comparação das listas de coocorrências 29

4.2 Resultados experimentais de coocorrência e correlação 31

4.2.1 Análise para artistas 31

4.2.2 Análise para músicas 34

4.2.3 Análise para gêneros musicais 36

4.3 Limitações do estudo de coocorrência e correlação 38

4.4 Conclusões 40

5 Metodologia de análise da influência da TV nas programações de rádios **41**

5.1 Levantamento de artistas e músicas para monitoramento 43

5.2 Reconhecimento por áudio das músicas em rádios 44

5.3 Escolha dos programas de TV 45

5.4 Monitoramento de participações em programas de TV 45

5.5 Seleção e cruzamento dos dados de rádio e TV 46

5.6 Realização da tarefa de predição 47

5.6.1 Métodos de médias simples 47

5.6.2 Métodos de regressão 48

5.6.3 Métodos de preenchimento de valores faltantes (*Imputation*) . . . 49

5.7 Comparação dos resultados de predição com dados reais 50

6 Resultados das análises de influências nas programações de rádios 52

6.1 Configurações para as análises 52

6.2 Análises de influências dos programas de TV 56

6.3 Análises de influências de outros eventos 64

6.4 Limitações do estudo 69

7 Conclusões e trabalhos futuros 71

A Tabela de audiências dos programas monitorados 75

B Tabelas de taxas de acerto dos preditores avaliados 77

B.1 Resultados para os métodos de preenchimento de valores faltantes 78

B.2 Resultados para os métodos de regressão 79

C Tabelas de taxas de acerto dos preditores para semanas/dias em análise 81

C.1 Resultados para os métodos de preenchimento de valores faltantes 82

C.2 Resultados para os métodos de regressão 102

Referências Bibliográficas 123

Lista de Figuras

2.1 Arquitetura alto nível dos componentes do Radialize. 11

4.1 Etapas da metodologia de análise de coocorrência e correlação entre itens da programação das rádios. 25

4.2 Processo de geração de bases transacionais 27

4.3 Processo de geração da lista de coocorrências 29

5.1 Fluxograma da metodologia de análise de influência da TV sobre execuções musicais nas rádios. 42

6.1 Execuções da música Todo Mundo, de Gaby Amarantos (part. Monobloco) no período da Copa do Mundo de Futebol de 2014. 65

6.2 Execuções da música País do Futebol, de MC Guimê (part. Emicida) no período da Copa do Mundo de Futebol de 2014. 66

6.3 Lançamento da música Indescritível, da dupla João Bosco e Vinícius. . . 67

6.4 Lançamento da música Ele Não Vai Mudar da dupla João Neto e Frederico. 68

6.5 Lançamento da música Implorando pra Trair de Michel Teló. 69

Lista de Tabelas

2.1 Opções de Testes Estatísticos 16

4.1 Características das bases transacionais geradas 28

4.2 Similaridade entre listas (Artistas) 32

4.3 Similaridade Jaccard entre listas (Artistas) 33

4.4 Novas correlações encontradas entre artistas, que não estavam presentes na Last.fm 34

4.5 Similaridade entre listas (Músicas) 35

4.6 Similaridade Jaccard entre listas (Músicas) 35

4.7 Similaridade entre listas (Gêneros Musicais) 36

4.8 Similaridade Jaccard entre listas (Gêneros Musicais) 37

4.9 Novas correlações encontradas entre gêneros musicais, que não estavam presentes na Last.fm 37

6.1 Avaliação das predições por preenchimento de valores faltantes 54

6.2 Avaliação das predições por regressões 54

6.3 Análise sobre as taxas de variação relativas à primeira metade do ano, segunda metade do ano e ano inteiro (Análise por semanas) 57

6.4 Análise sobre as taxas de variação relativas às músicas em ascensão, estabilidade e queda (Análise por Semanas) 58

6.5 Análise sobre as taxas de variação relativas à primeira metade do ano, segunda metade do ano e ano inteiro (Análise por dias) 59

6.6 Análise sobre as taxas de variação relativas às músicas em ascensão, estabilidade e queda (Análise por dias) 60

6.7 Análise para os programas da Rede Globo: Fantástico e Domingão do Faustão 60

6.8 Resultados dos testes de Wilcoxon relativos às análises da primeira metade do ano, segunda metade do ano e ano inteiro (Análise por semanas) . . . 62

6.9 Resultados dos testes de Wilcoxon relativos às músicas em ascensão, estabilidade e queda (Análise por Semanas) 63

6.10 Resultados dos testes de Wilcoxon relativos à primeira metade do ano (Análise por dias) 63

6.11 Resultados dos testes de Wilcoxon relativos às análises para os programas: Fantástico e Domingão do Faustão 64

A.1 Audiências dos programas de TV monitorados 75

B.1 Avaliação dos algoritmos de preenchimento de valores faltantes 78

B.2 Avaliação dos algoritmos de regressão 79

C.1 Resultados dos algoritmos de preenchimento de valores faltantes 82

C.2 Resultados dos algoritmos de regressão 102

Capítulo 1

Introdução

Com a invenção e posterior popularização do rádio, as formas de experiência musical sofreram mudanças consideráveis, como afirma Trotta (2009). A transmissão de músicas em larga escala através do rádio promoveu um intercâmbio entre músicos e ouvintes de diferentes localidades. Além de transmitir músicas, o rádio detém um grande potencial comunicativo, tornando-se uma mídia sonora com alto poder de alcance e informação, como afirma Sanchotene et al. (2010). Com a internet, este meio de comunicação ganhou uma nova forma de transmissão que potencializou seu alcance ainda mais. Antes, uma emissora de rádio poderia ser acessada apenas por um aparelho receptor de ondas hertzianas, mas agora ela também pode ser acessada por um computador, através de uma URL (do inglês, *Uniform Resource Locator*) na *web*.

Segundo Prata (2008), com o surgimento e constante crescimento da *web*, empresas foram criando suas próprias páginas para divulgarem seus produtos e serviços. As rádios também adotaram essa prática, inicialmente como forma de apresentação da empresa, locutores e anúncios publicitários. Posteriormente, as rádios passaram a fazer suas transmissões também pela internet, permitindo assim que fossem ouvidas tanto pelo aparelho de rádio quanto pelo computador, podendo alcançar um público ainda maior. Como destaca Bufarah Junior (2003), uma emissora de rádio na internet ultrapassa os limites regionais, ganhando um caráter global e facilitando sua audição em diversos lugares do mundo.

Aizenberg et al. (2012) afirmam que as estações de rádios na internet, por meio de suas listas de reprodução, representam uma fonte prontamente disponível e abundante de sinais humanos na música, que são essencialmente conjuntos coesos de faixas rela-

cionadas. Pachet et al. (2001) relatam que existe uma lógica nas programações das rádios e, normalmente, as músicas não são escolhidas de forma aleatória. Os locutores de rádio, em geral, têm um grande conhecimento das músicas que eles executam, e tal conhecimento é precisamente o que dá ao programa o seu toque característico.

Portanto, ao se fazer análises na programação de uma rádio, podem-se extrair informações muito interessantes, como por exemplo, se existe um estilo musical predominante e qual seria o padrão de execuções musicais em sua programação. Além disso, existe a hipótese de que os locutores de rádio podem observar acontecimentos externos, como a repercursão de determinados artistas[1] em outras mídias, como aparições em programas de televisão (TV), o que pode impactar na quantidade de execuções de músicas daqueles artistas e na maneira que os locutores definem suas programações musicais nas rádios.

A fim de se obter um melhor entendimento das características e padrões apresentados nas programações das rádios, primeiramente foi realizado um estudo preliminar para verificar a correlação entre itens (artistas, músicas e gêneros musicais) que coocorrem nas programações das rádios, sendo possível estudar a coesão do conteúdo musical nas programações definidas pelos especialistas. Este estudo abriu portas para se estudar acerca da seguinte hipótese: "A presença de artistas em programas de TV de grande audiência impacta positivamente no número de execuções desses artistas em rádios?"

Tendo em vista a hipótese exposta acima, foi desenvolvida como parte principal deste trabalho uma metodologia de análise da possível influência da mídia televisiva sobre a forma com que as programações das rádios são definidas pelos especialistas (locutores ou DJs). Esse estudo busca mensurar o impacto da presença de artistas em programas de TV sobre as execuções de suas músicas nas rádios. A metodologia desenvolvida compreende processos que vão desde a coleta e monitoramento de dados em programações de estações de rádio e em programas de TV, até o cruzamento e análise desses dados, com o propósito de se identificar padrões de execução nas rádios provocado por influência da TV. Além disso, foram realizadas análises mais específicas para identificar variações no padrão de execuções de músicas, que pudessem estar associadas a acontecimentos de determinados eventos, como por exemplo lançamentos de músicas e músicas que se tornaram temas, seja de novelas ou mesmo da Copa do Mundo de Futebol da FIFA, que ocorreu no Brasil em 2014.

[1] Usamos de forma genérica o termo "artistas", que pode se referir a cantores, bandas ou intérpretes musicais em geral.

Por meio da metodologia proposta de análise de influências entre mídias (TV e rádio) e pelos estudos de padrões de execuções realizados neste trabalho, espera-se que o conteúdo musical das programações das rádios, bem como a maneira com que ele é definido e apresentado pelos especialistas, sejam melhor compreendidos.

1.1 Objetivos

O objetivo geral deste trabalho é apresentar evidências favoráveis ou contrárias à hipótese de que a participação de artistas musicais em programas populares de televisão influenciam positivamente na forma que as músicas dos artistas são executadas em rádios.

Os objetivos específicos deste trabalho são:

- Realizar um estudo de coocorrência de músicas em execução em programas de rádios, visando gerar conhecimento sobre a forma que as músicas de artistas são executadas em rádios, bem como sobre como as músicas, os artistas e os estilos musicais se relacionam entre si;
- Desenvolver uma metodologia para análise de padrões de execuções musicais em programações de rádios, a partir da participação de artistas em programas de TV, no intuito de descobrir evidências favoráveis ou contrárias à hipótese do trabalho;
- Realizar um estudo de caso a partir do monitoramento real de programas de TV e rádios brasileiras, a fim de permitir que a metodologia desenvolvida seja aplicada, de fato, para apresentar evidências favoráveis ou contrárias à hipótese levantada;
- Avaliar o impacto de eventos como a Copa do Mundo de Futebol e lançamentos oficiais de músicas na programação das rádios, como um complemento à análise de impacto da participação de artistas em programas de TV na programação das rádios.

1.2 Justificativa

Por ser um meio de comunicação em massa, o rádio apresenta grande importância comercial e política, tal como a televisão, porém com características diferentes. De acordo com Moreira (1999), o rádio consolida sua agilidade para informar e investe em um estilo

de programação com forte componente musical, reforçando uma parceria rentável com a indústria fonográfica.

Por outro lado, a indústria fonográfica, juntamente com os artistas do ramo musical e estudiosos da área da comunicação, têm interesse em saber como os produtos (músicas) estão sendo veiculados e consumidos pelas pessoas, através dos meios de comunicação e, mais que isso, como esses veículos interagem para a divulgação desses produtos e de seus conteúdos.

Acredita-se que os meios de comunicação sofram algum tipo de influência uns dos outros. No entanto, estudos que apresentam estes tipos de análises, com detalhes que comprovem a existência dessa influência, sendo capazes inclusive de mensurá-la, até então são desconhecidos. Um dos possíveis motivos da escassez de estudos nessa direção, pode ser associado à dificuldade em se obter dados sobre os conteúdos veiculados nos meios de comunicação, principalmente nas rádios, uma vez que o número de estações de rádio é muito alto.

O presente trabalho explora esse universo das programações das rádios e da veiculação dos conteúdos musicais, pois faz uso de dados, bem como dos serviços de coleta, fornecidos pela SoundView (ver Seção 2.3), uma empresa que monitora ocorrências de músicas em rádios brasileiras na *web*, tornando possível a realização desse estudo. Como consequência deste estudo, artistas serão capazes de avaliar com maior precisão o quanto eles devem investir com suas respectivas assessorias de imprensa para conseguir espaços nas agendas dos programas de TV.

Em nossos estudos, realizamos entrevistas com algumas pessoas do meio artístico-musical, que afirmaram acreditar que, de forma geral, a hipótese levantada no trabalho é verdadeira. Ou seja, foi possível observar que o meio artístico-musical acredita que a participação dos artistas em programas de TV com boa audiência possa influenciar positivamente no número de execuções desses artistas em programas de rádio.

Portanto, este estudo se propõe a analisar dados até então pouco estudados, provenientes do conteúdo musical veiculado nas emissoras de rádio, com o intuito de identificar padrões e influências (impactos) de outros meios de comunicação, como a televisão, sobre as execuções musicais nas rádios. Com isso, espera-se que os resultados deste trabalho possam fornecer, tanto para a indústria fonográfica quanto para estudiosos da comunicação, subsídios para uma melhor compreensão sobre como as rádios definem suas programações e apresentam seu conteúdo musical aos ouvintes, bem como sobre a forma com que as participações de artistas musicais em programas de TV influenciam

as execuções em rádios.

1.3 Definição do problema

Para conduzir a principal investigação deste trabalho, sobre a influência de mídia televisiva nas programações musicais das rádios, foram analisados valores esperados de execuções de músicas nas rádios (tendo como referência padrões de execuções anteriores) após participações de artistas, associados a essas músicas, em programas de televisão. Logo, almeja-se responder à seguinte questão de pesquisa:

$\boldsymbol{QP_0}$: Existe impacto nas execuções de músicas em rádios provocado pela participação de artistas em programas de televisão?

Para investigar empiricamente a questão de pesquisa QP_0, foram elaboradas as seguintes hipóteses:

Hipótese Nula (H_0): Não existe diferença entre os valores de execuções esperados e os observados.

Hipótese Alternativa (H_1): Existe diferença entre os valores de execuções esperados e os observados.

A rejeição da hipótese nula seria um indício de que houve significativa diferença entre o conjunto de dados esperados e o conjunto de dados observados, o que sugere um desvio no padrão de execuções. Ao se realizar essa análise para um número grande de instâncias, a aceitação da hipótese fica fortalecida, aumentado-se as evidências a favor da hipótese alternativa. Isso nos remete a formular uma nova questão de pesquisa, a saber:

$\boldsymbol{QP_1}$: Os valores de execuções observados são maiores que os esperados?

Caso a resposta para essa questão seja afirmativa, é uma evidência de que existe impacto positivo, possivelmente causado pela participação de um artista em um programa de TV, uma vez que os valores observados de execuções superam as expectativas. Caso a resposta encontrada seja negativa, é um indicativo de que simplesmente não há aumento das execuções pois, intuitivamente, não se acredita que as participações dos artistas na TV possam provocar impactos negativos nas execuções de suas músicas nas rádios.

1.4 Organização do texto

O restante deste texto está organizado da seguinte forma. No Capítulo 2, é apresentada a fundamentação teórica, na qual são descritos conceitos importantes associados ao conteúdo desta dissertação. No Capítulo 3, é apresentada a revisão bibliográfica, contendo trabalhos relacionados com este e com abordagens a partir da perspectiva de duas áreas do conhecimento diferentes: ciência da computação e comunicação. No Capítulo 4, é apresentada uma análise sobre a coocorrência e correlação entre itens executados nas programações de rádios. No Capítulo 5, é apresentada a metodologia elaborada para análise de influência da mídia televisiva sobre as programações das rádios. Já no Capitulo 6, são apresentados os resultados da análise de influência entre mídias (TV e Rádio) realizada. Por fim, no Capítulo 7, são apresentadas as conclusões e as direções para trabalhos futuros.

Capítulo 2

Fundamentação teórica

Este capítulo apresenta a fundamentação teórica, onde são descritos conceitos importantes associados ao conteúdo deste trabalho. A Seção 2.1 apresenta como a música, enquanto produto, é comercializada, distribuída e consumida na chamada Indústria da Música. A Seção 2.2 apresenta características do rádio e de sua grade de programação. Na sequência, a Seção 2.3 apresenta a empresa SoundView e a arquitetura alto nível do serviço de monitoramento de músicas nas rádios da ferramenta Radialize. Já a Seção 2.4 apresenta as tarefas e técnicas de mineração de dados utilizadas por este trabalho para a identificação de padrões nas programações de rádios. Por fim, na Seção 2.5, são apresentados os testes de hipóteses utilizados para validar as hipóteses levantadas neste trabalho.

2.1 A indústria da música

De acordo com Genes et al. (2012), no começo do século XX, a Economia da Música era composta basicamente por produtores de fonógrafos e gramofones, fabricantes de discos de matéria plástica, a indústria de instrumentos musicais e os artistas e consumidores de música. Com o passar dos anos, surgem os primeiros selos independentes, responsáveis por popularizar gravações de *blues* e *jazz*, impulsionando a indústria fonográfica em sua ascensão.

Com a popularização do rádio e do cinema sonoro na década de 1930, abriu-se um novo mercado para a comercialização da música, ampliando a indústria do entretenimento. As canções eram reproduzidas nas rádios em diversos programas, enquanto o

cinema trazia os musicais às telas de exibição. Com isso, a popularizarização de cantores e intérpretes da época era alavancada e os grandes estúdios passaram a se tornar também importantes agentes da cadeia. A música ganhava novos meios de distribuição e de consumo. A televisão, que foi popularizada após a Segunda Guerra Mundial, também se tornou responsável pela disseminação das músicas ao transmitir programas musicais e apresentações de artistas populares da época.

Com a evolução técnica da indústria da música ao longo dos anos, surgiram diferentes integrantes em sua cadeia produtiva, como por exemplo, a indústria de equipamentos de som e gravação, além dos novos produtores e fornecedores de disco. A criação dos *walkmans* e CD *players* possibilitou uma maior facilidade em se ouvir músicas, fazendo com que essa atividade passasse a complementar as demais atividades do dia a dia das pessoas. Os CDs começaram a ser comercializados em diversos lugares como supermercados, bancas de jornal e livrarias, ganhando novos meios de distribuição e comercialização e atingindo um público cada vez maior. Segundo Prestes Filho (2004), já no final do século XX, a internet se popularizou e as redes de compartilhamento de arquivos contribuíram para a distribuição desenfreada de arquivos de música e também, como consequência, para a proliferação da pirataria na indústria musical.

A partir do melhoramento dos processos de criação e distribuição dos arquivos digitais, o crescente número de usuários em redes de compartilhamento e o aumento da velocidade de transferência de dados pela internet, como afirmam Genes et al. (2012), percebem-se então três possibilidades para o usuário consumir música: vendas *on-line*, *downloads* e *streaming*. As vendas *on-line* proporcionam às gravadoras a disponibilização de faixas sem custos adicionais de distribuição. Desse modo, identifica-se um possível mercado baseado em taxas de adesão ou patrocínios, vantajoso para usuários com maior frequência de consumo. As vendas *on-line* de músicas se tornaram mais eficientes no começo da década de 1990. Entretanto, de acordo com Prestes Filho (2004), o que mostra maior possibilidade de sucesso no mercado não é a venda de álbuns completos mas a comercialização de faixas individuais. Com a evolução tecnológica, a distribuição de músicas por *streaming* vem se tornando um recurso cada vez mais adotado por usuários, permitindo pensar nesse modo como substituto à venda de CDs de hoje, pelo fato de dar acesso a conteúdos individuais sem a necessidade do *download*.

Portanto, o modelo de negócios que mais ganhou força na internet nos últimos anos foi o *streaming* pois, apesar da facilidade em se adquirir músicas através da internet (por meio de *downloads*), muitas vezes o consumidor não quer adquirir uma versão digital da música, uma vez que pode ter acesso imediato e grátis a esse conteúdo.

Websites como o YouTube[1], assim como rádios *on-line*, permitem que os internautas ouçam músicas sob demanda, sem a necessidade de realizar o *download* do conteúdo e, na maioria dos casos, de forma gratuita, tornando-se rapidamente muito populares entre os ouvintes de música. Tornaram-se, assim, poderosas ferramentas *on-line* de acesso às músicas, uma vez que, no caso de *websites* como o YouTube, os internautas passaram a postar os mais diversos tipos de faixas musicais, desde lançamentos até gravações extremamente raras, atraindo cada vez mais novos usuários. Já no caso das rádios *on-line*, o conteúdo musical é definido por especialistas (locutores e DJ's informados sobre o cenário musical mundial), podendo ser consumido por pessoas de diferentes localizações geográficas.

2.2 O rádio e a sua grade de programação

O rádio, enquanto veículo de comunicação, busca sempre estar o mais próximo possível dos ouvintes, sendo o parceiro de todas as horas, principalmente a partir do momento em que se tornou portátil e pôde acompanhar os ouvintes a qualquer lugar, como afirma Medeiros (2009). Portanto, a linguagem utilizada pelos comunicadores no rádio busca uma forma de aproximação e uma interação com o público, estreitando os laços como em uma conversa cara a cara.

A presença de um locutor é fundamental durante uma transmissão sonora radiofônica tradicional, uma vez que trata-se de um veículo de comunicação basicamente sonoro. Logo, o locutor se torna uma figura importante para que o ouvinte possa associar sua voz a uma rádio em meio às demais estações. Dessa forma, o papel do locutor também é de quebrar o silêncio entre uma música e outra durante o fluxo de transmissão, marcando o início e o fim das mesmas.

Outro elemento característico do rádio é a sua grade de programação, que surge como uma espécie de roteiro onde devem ser encaixados os diversos programas produzidos pela emissora, obedecendo ao fluxo contínuo de transmissão e a uma apresentação de forma "sequencial e cronológica", como destaca Barbosa Filho (2003).

Ainda de acordo com Medeiros (2009), as vinhetas são elementos radiofônicos que marcam o início e o fim de um programa de rádio ou produto radiofônico (vinhetas de abertura e encerramento). Elas geralmente são curtas, possuindo de cinco a dez segundos

[1]http://www.youtube.com

de duração e, assim como o locutor, são também um elemento radiofônico característico que identifica um determinado programa de rádio, bem como a emissora à qual o ouvinte está sintonizado. As "chamadas"são elementos radiofônicos muito parecidos com as vinhetas, onde o objetivo é a divulgação institucional. As "chamadas"também são curtas e são utilizadas para anunciar um programa ou a programação do dia seguinte.

Os blocos de comerciais também entram na programação das rádios como um elemento característico, principalmente quando se trata das chamadas "rádios comerciais". A publicidade no rádio foi responsável por impulsionar as vendas de aparelhos receptores e, portanto, por aumentar o número de estações e sua cobertura. Inicialmente, os comerciais eram conhecidos como "taxa de transmissão", sendo cobrados aos programas que iam ao ar. Com o passar do tempo, começaram a ser pagos pelos seus patrocinadores.

No início das transmissões em frequência modulada (FM) no Brasil, a programação das rádios se tornou restrita à veiculação de músicas. O bloco musical na grade de programação é característico principalmente nas rádios conhecidas como "vitrolão", típicas das transmissões em FM. Conhecidos também como "Programação Musical", os blocos musicais podem ser considerados como um gênero de entretenimento e são caracterizados por uma sequência de músicas dispostas em diferentes programas definidos por estilos musicais, por exemplo, ou muitas vezes agem como "um grande painel musical", geralmente determinados pela tendência musical adotada por cada emissora. Segundo Barbosa Filho (2003), os blocos de programação são constituídos por séries de músicas que definem a estrutura da programação musical. Normalmente, essas sequências musicais têm duração em torno de seis a doze minutos (podendo existir horários em que a programação ocorre sem interrupções, no intuito de fixar audiência) e estão envolvidas por blocos: de anúncios comerciais, de chamadas (textos institucionais), de serviço, jornalísticos e de entretenimento.

2.3 A ferramenta Radialize e o serviço da SoundView

Este trabalho fez uso de informações sobre o conteúdo musical executado nas programações das rádios, disponibilizadas pela ferramenta Radialize[2] e pela empresa SoundView[3]. De acordo com Rezende (2014), o Radialize é uma ferramenta de descoberta de conteúdos veiculados em rádios nacionais e estrangeiras na *web*, possuindo funcionali-

[2]http://www.radialize.com.br
[3]http://www.soundview.com.br

dades para apresentar e recomendar itens, como rádios, artistas, músicas e gêneros. A partir do projeto Radialize, foi criada a empresa SoundView, que tem como objetivo monitorar ocorrências de propagandas e músicas em rádios brasileiras na *web*, fornecendo informações importantes tanto para anunciantes quanto para artistas, como por exemplo, o número de vezes que uma campanha publicitária foi executada nas rádios (para verificar se as rádios cumprem o que foi contratado) e a popularidade de uma música lançada.

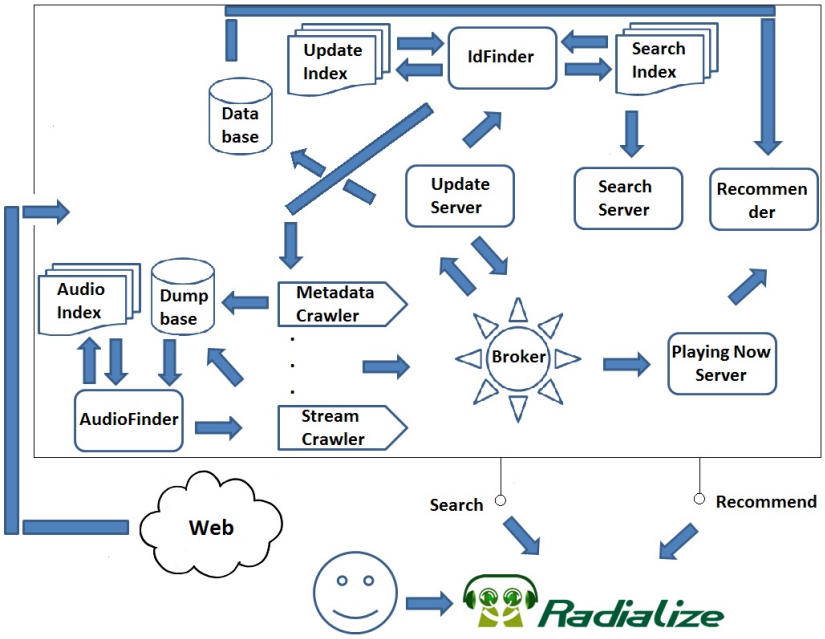

Figura 2.1: Arquitetura alto nível dos componentes do Radialize.

A Figura 2.1 ilustra, em alto nível, as operações e os componentes da arquitetura do Radialize. A ilustração apresenta dois tipos de coletores: os coletores que utilizam informações de metadados (*Metadata Crawlers*) e os coletores que utilizam informações de áudio (*Stream Crawlers*). Os coletores de metadados extraem informações disponibilzadas pelos *players* das rádios da *web*, como o nome da música, nome do artista e *timestamp* de execução. Quando não existe esse tipo de informação, então são utilizados os coletores de áudio, que são responsáveis por extrair o áudio das rádios e enviar segmentos para o componente *AudioFinder*, que por sua vez irá verificar se um cojunto de segmentos representa a ocorrência de uma música conhecida.

O componente *Broker* é responsável por enviar os dados recebidos dos coletores para os componentes *PlayingNowServer* e *UpdateServer*. O *PlayingNowServer* é responsável por manter informações a respeito de quais artistas e músicas estão em execução em

cada estação de rádio. O *UpdateServer* é responsável pela atualização da base de dados do Radialize e também pela atualização dos índices utilizados, realizando a inserção de novos artistas, novas músicas e novas rádios.

O *IdFinder* é responsável por encontrar os respectivos identificadores de mídias no Radialize. Para isto, são mantidos dois índices textuais de maneira a dar suporte ao trabalho desse componente: *update index* e *search index*. O *update index* é utilizado pelo *UpdateServer* com a finalidade de descobrir se uma nova mídia (artista e música) existe na base de dados. Já o *search index* é responsável por responder a consultas realizadas pelos usuários. O *AudioFinder*, por sua vez, é responsável por reconhecer e encontrar os identificadores das mídias, mantendo para isso, um índice de arquivos de áudio.

O índice de áudio contém todas as mídias fornecidas previamente ao Radialize. Novas mídias são indexadas no momento em que são identificadas pelos coletores de metadados, recebendo um identificador único atribuído pelo *UpdateServer*. O *Recommender* é responsável por calcular e ordenar as recomendações que serão enviadas aos usuários. A ordenação dos itens recomendados é baseada na relevância e quais desses itens estão sendo executados em alguma rádio no momento. Tal informação é fornecida pelo componente *PlayingNowServer*. Caso nenhum item esteja sendo executado, a ordem inicial dos itens é mantida. Rádios recomendadas são executadas conectando o *player* do Radialize diretamente ao servidor de transmissão de áudio usado pela rádio.

2.4 Tarefas e técnicas de mineração de dados

A Mineração de Dados é uma etapa no processo de Descoberta de Conhecimento em Bases de Dados que consiste na aplicação de algoritmos específicos que extraem padrões a partir dos dados, como afirmam Fayyad et al. (1996). A Mineração de Dados é composta por diferentes tarefas, sendo as mais comuns: Descrição, Classificação, Estimação, Predição, Agrupamento e Associação. Mais adiante são apresentadas as tarefas utilizadas neste trabalho, que são: Regressão (ver Seção 2.4.1) e Associação (ver Seção 2.4.2).

É importante ressaltar que uma tarefa de mineração não é o mesmo que uma técnica de mineração. Uma tarefa de mineração consiste na especificação de algo que se está buscando nos dados, como regularidades e padrões interessantes a serem encontrados

(por exemplo, padrões de compras de clientes em uma base de registros de compras de um supermercado). A técnica de mineração consiste na elaboração/utilização de métodos específicos que nos auxiliem a realizar a tarefa de mineração, como descobrir os padrões que nos interessam.

2.4.1 Predições numéricas

Segundo Camilo and Silva (2009), os métodos de predição visam estimar um possível valor de uma determinada variável, analisando-se os valores das demais. As predições numéricas têm como objetivo realizar previsões de valores para variáveis contínuas, discretas ou mistas. Os métodos mais conhecidos para predições numéricas são as regressões, porém, alguns métodos de classificação também podem fazer predições. As técnicas de regressão modelam o relacionamento de variáveis independentes (chamadas preditoras) com uma ou mais variáveis dependentes (chamadas respostas). As variáveis preditoras são os atributos dos registros, e a resposta é o que se quer predizer.

Regressão Linear

As regressões são chamadas de lineares quando a relação entre as variáveis preditoras e a resposta segue um comportamento linear nos parâmetros. Nesse caso, é possível criar um modelo no qual o valor de y é uma função linear de x, como por exemplo $y = b + wx$. É possivel utilizar o mesmo princípio para modelos com mais de uma variável preditora.

Regressão Não-Linear

Nos modelos de regressão não-linear, a relação entre as variáveis preditoras e a resposta não segue um comportamento linear nos parâmetros. Por exemplo, a relação entre as variáveis pode ser modelada como uma função polinomial. No entanto, para esses casos (Regressão Polinomial), é possível realizar uma conversão pra uma regressão linear. Outros modelos também podem ser encontrados na literatura, como: *Logistic Regression*, *Poisson Regression* e *Log-Linear Models*.

2.4.2 Mineração de regras de associação

De acordo com Camilo and Silva (2009), essa é uma das técnicas mais conhecidas de mineração de dados, devido ao problema da Análise da Cesta de Compras (*Market Basket Analysis*). Consiste em identificar o relacionamento entre os itens mais frequentes em um determinado conjunto de dados, e permite obter resultados do tipo: SE um cliente compra leite e pão TAMBÉM compra manteiga. Essa construção recebe o nome de Regra de Associação (*Association Rule*). Ou seja, uma regra de associação é um padrão da forma $X \rightarrow Y$, onde X e Y são conjuntos de valores (itens comprados por um cliente, sintomas apresentados por um paciente, etc.).

Portanto, tomando como exemplo a Análise da Cesta de Compras, descobrir regras de associação entre produtos comprados por clientes numa mesma compra pode ser útil para a tomada de decisões estratégicas, como melhorar a organização das prateleiras para facilitar as compras dos clientes, ou realizar promoções para induzir os clientes a comprarem mais produtos em conjunto.

Neste trabalho, foi utilizada a técnica de mineração de regras de associação no Capítulo 4, para se extrairem conjuntos de itens frequentes (artistas, músicas e gêneros musicais) que coocorrem nas programações de rádios e, assim, verificar a existência de correlação entre esses itens.

Mineração de Itens Frequentes (Frequent Itemsets Mining)

Apresentada por Agrawal et al. (1993), essa técnica pode ser dividida em duas etapas. Na primeira etapa, conjuntos de itens frequentes (*Frequent Itemsets*) são gerados, respeitando-se um valor mínimo de frequência (*Support*) para os itens. Na segunda etapa, as regras de associação são geradas pela mineração desses conjuntos de itens frequentes gerados. Para garantir resultados válidos, os conceitos de suporte (*Support*) e confiança (*Confidence*) são utilizados em cada regra produzida. A medida de suporte indica o percentual de registros (dentre todo o conjunto de dados) que se encaixam nessa regra. Já a confiança mede o percentual de registros que atendem especificamente à regra, por exemplo, o percentual de quem compra leite e pão e também compra manteiga.

Para uma regra ser considerada forte, ela deve atender a um certo grau mínimo de suporte e confiança. Um dos mais tradicionais algoritmos de mineração utilizando a estratégia de itens frequentes é o Apriori, apresentado por Ye and Chiang (2006).

2.5 Teste de hipóteses

Na Estatística, hipóteses são formuladas sobre os parâmetros das distribuições de probabilidades do fenômeno estudado, sendo que tais hipóteses são submetidas a testes para evidenciar a validade das mesmas. Dessa forma, de acordo com Da Rosa (2009), os testes estatísticos surgem como regras de decisões, ligadas a um fenômeno da população, que nos permitem definir, com o auxílio de amostras, se determinadas hipóteses podem ser rejeitadas ou não.

A hipótese a ser testada é designada por Hipótese Nula (H_0). Além disso, existe também a Hipótese Alternativa (H_1), que se trata da conclusão obtida quando a hipótese nula é rejeitada. Quando se formula uma decisão sobre H_0, podem ocorrer dois tipos de erros: erro tipo I e erro tipo II. O primeiro consiste em rejeitar a hipótese nula quando a mesma é verdadeira, enquanto o segundo consiste em aceitá-la quando a mesma é falsa. A esses erros estão associadas probabilidades denotadas por α e β, respectivamente. Desse modo, essas probabilidades ficariam assim: $\alpha = \text{P(Rejeitar } H_0 | H_0 \text{ é Verdadeira)}$ e $\beta = \text{P(Não Rejeitar } H_0 | H_0 \text{ é Falsa)}$.

Quando apenas α (conhecida também como nível de significância) é utilizada, como é usual, está sendo realizado um Teste de Significância. Como o valor α entra no processo de determinação de aceitação ou rejeição de H_0, a condição de objetividade da prova exige que o nível de significância seja estabelecido, sendo os valores 0,05 e 0,01 mais comumente utilizados para α, como afirmam Câmara and Silva (2001).

Após definir as hipóteses e o nível de significância, é necessário escolher o teste estatístico adequado, levando-se em conta fatores como: o tipo dos dados que estão sendo estudados (contínuo, categórico: ordinal ou nominal); a distribuição desses dados (Distribuição Normal ou Distribuição Não Normal); e os tipos de amostras estudadas (Independentes ou Dependentes). Os testes estatísticos podem ser divididos em dois grandes grupos: paramétricos e não-paramétricos. Os testes paramétricos são usados quando a distribuição dos dados é conhecida (Distribuição Normal). Já os testes não-paramétricos são usados para os casos em que a distribuição dos dados não é conhecida. A Tabela 2.1 exibe as opções de testes estatísticos mais comumente usados na prática para comparação de médias e medianas, de acordo com Campos (2001).

No caso deste trabalho, o teste estatístico mais adequado para investigar empiricamente a questão de pesquisa QP_0 (ver Seção 1.3) foi o teste de Friedman, porque

Tabela 2.1: Opções de Testes Estatísticos

Testes Estatísticos			
Paramétricos		**Não-Paramétricos**	
Independentes	**Pareados**	**Independentes**	**Pareados**
2 amostras	**2 amostras**	**2 amostras**	**2 amostras**
Teste t (Student)	Teste t (Student)	Mann-Whitney T. da Mediana x^2(2 x 2) Proporções Exato (Fischer)	Wilcoxon T. dos sinais MacNemar Binomial
>2 amostras	**>2 amostras**	**>2 amostras**	**>2 amostras**
Análise de variância	Análise de variância	Kruskal-Wallis Mediana (m x n) x^2(m x n) Nemenyi	Cochran Friedman

Fonte: Campos (2001)

deseja-se saber se k amostras correspondentes pertecem à uma mesma população. Além disso, segundo Nascimento et al. (2012) e Siegel and Castellan (1988), o uso desse teste é adequado para dados que seguem uma escala ordinal e a distribuição dos mesmos não é conhecida, como é o caso dos dados analisados neste trabalho. Já para investigar a questão de pesquisa QP_1 foi utilizado o teste de Wilcoxon, pois ele é indicado para verificar diferenças entre amostras pareadas, indicando qual das amostras é maior ou menor, uma vez que esse teste considera tanto a magnitude relativa das amostras quanto a direção das diferenças existentes entre elas.

2.5.1 Teste de Friedman

O teste de Friedman é um teste não-paramétrico utilizado para comparar dados amostrais pareados, verificando-se a hipótese de que as k amostras relacionadas provêm de uma mesma população. Como destaca Guimarães (2014), nesse tipo de análise observa-se o mesmo grupo de indivíduos sob cada uma das k condições, ou então, formam-se conjuntos de indivíduos homogêneos entre si, sendo estes alocados aleatoriamente a cada

uma das condições.

A resposta do teste depende de como os dados são organizados em uma tabela de k colunas e n linhas, sendo essa etapa muito importante, uma vez que depende dessa organização a interpretação do resultado do teste. Portanto, o fator comparado principal deve ser colocado nas colunas e os dados são colocados nas linhas da tabela.

Segundo Campos (2001) o teste de Friedman não utiliza os dados numéricos diretamente, mas sim os postos ocupados por eles na tabela, após a ordenação por valores ascendentes desses dados, que é realizada separadamente em cada uma das amostras e não em conjunto.

Como demonstra Guimarães (2014), considerando a estrutura da tabela citada anteriormente, com k colunas e n linhas, para cada linha atribui-se postos de 1 a k, calculando-se em seguida a estatística X_r^2 de Friedman dada pela Equação 2.1:

$$X_r^2 = \frac{12}{nk(k+1)} \sum_{i=1}^{k} R_i^2 - 3n(k+1) \tag{2.1}$$

onde R_i é a soma dos postos na coluna i.

De acordo com Campos (2001), a proposta do teste de Friedman considera que, se as diversas amostras provêm de uma mesma população, ou seja, se elas sao consideradas estatisticamente iguais (hipótese nula H_0), a distribuição dos postos nas diversas colunas será mais ou menos equivalente, de maneira que a soma dos postos em cada coluna será aproximadamente igual. Já a hipótese alternativa H_1 seria de que as amostras não pertenceriam à mesma população, sendo consideradas diferentes e havendo diferenças também entre as somas das diversas colunas.

2.5.2 Teste de Wilcoxon Pareado

O teste de Wilcoxon Pareado (do inglês, *Wilcoxon Matched-Pairs Signed-Rank Test*) é um teste não-paramétrico para verificar diferenças entre amostras pareadas, utilizando informações sobre a direção das diferenças entre os pares. Desse modo, de acordo com Siegel and Castellan (1988), é possível definir se uma amostra é maior que a outra, baseando-se no resultado do teste. Portanto, ele pode ser considerado um teste mais

poderoso, pelo fato de considerar tanto a magnitude relativa das amostras quanto a direção das diferenças existentes entre elas.

A princípio, são calculados os valores numéricos da diferença entre cada par, sendo possíveis as seguintes condições: aumento (+), diminuição (-) e igualdade (=). Após calcular todas as diferenças entre os valores obtidos para cada par, as mesmas são ordenadas pelo seu valor absoluto, sem considerar o sinal, substituindo-se os valores originais pelo posto que ocupam na escala ordenada. No teste de Wilcoxon, ao invés de considerar apenas o sinal das diferenças entre os pares (variação para mais ou para menos), leva-se em conta também a magnitude do aumento ou da diminuição.

Esse teste faz a comparação entre as amostras tendo como referência as medianas, verificando se a diferença entre as medianas é igual a 0, por meio do uso de *rankings* (posição dos dados das amostras) de diferenças entre os valores pareados. Pelo fato de se basear nas medianas ao invés das médias, esse teste pode ser considerado mais robusto, se comparado a outros métodos que analisam as médias dos dados. Enquanto as médias podem ser amplamente afetadas pela distribuição dos dados, por ruidos ou *outliers*, as medianas são menos afetadas, se tornando melhores indicadores para a tendência central dos dados.

As hipóteses nula e alternativa formuladas para o teste de Wilcoxon são as seguintes:

H_0: A diferença entre as medianas das amostras pareadas é igual a 0

H_1: A diferença entre as medianas das amostras pareadas é diferente de 0 (se bilateral), ou > 0, < 0 (se unilateral)

Neste trabalho, o teste de Wilcoxon é utilizado para os casos em que a hipótese nula do teste de Friedman (ver Seção 2.5.1) é negada, sugerindo que existe uma diferença entre as amostras. Nesses casos, deseja-se saber qual das amostras apresenta maiores ou menores valores.

Capítulo 3

Trabalhos relacionados

Neste capítulo, são apresentados os trabalhos que se relacionam com este, a partir da perspectiva de duas áreas do conhecimento: ciência da computação e comunicação. Na Seção 3.1, são mostrados os trabalhos que envolvem a computação para realizar estudos sobre músicas e rádios. Na Seção 3.2, são exibidos os trabalhos que também fazem uso dos mesmos objetos de estudo, porém com foco e embasamento na área da comunicação.

3.1 Trabalhos associados à área da computação

Pachet et al. (2001) apresentam uma abordagem de classificação baseada na análise de similaridade e correlação de artistas e músicas que coocorrem em *playlists* de estações de rádio e em álbuns de coletâneas musicais. Com isso, foi introduzida uma forma de se extrairem automaticamente similaridades entre músicas e artistas. A técnica proporciona uma matriz de distâncias para conjuntos arbitrários de itens e os experimentos foram conduzidos em diversas bases de músicas e artistas. Os resultados preliminares mostram que a técnica consegue extrair similaridades entre itens, em pequenas bases de dados.

Kuo et al. (2005) propõem um modelo para recomendação de músicas com base em emoções, onde consultas expressadas por emoções retornarão músicas com características relacionadas às emoções definidas na consulta. O processo de construção do modelo inclui a extração de características de músicas de filmes, a detecção de emoções no segmento do filme e a descoberta de associações entre características de músicas e emoções, através de um grafo de afinidade. O resultado experimental mostra que o método proposto atinge,

na média, 85% de precisão.

Neubarth et al. (2012) demonstram como a mineração de regras de associação pode ser aplicada para se descobrirem relações entre duas ontologias da música *folk*: gênero e região. O método consiste em um modelo de regras e medidas de avaliação para identificar diferentes categorias de associações entre gêneros e regiões. Dadas as associações entre pares de características de conteúdos de músicas (gêneros e regiões), o método identifica combinações de conteúdo-região, conteúdo-gênero e região-gênero.

Liu et al. (2008) apresentam uma abordagem não-supervisionada para agrupamento de músicas, a partir de características extraídas de seus sinais de áudio. Desse modo, as músicas são divididas em grupos, onde o número de grupos pode ser determinado automaticamente pela minimização estimada *Rand Index*. Assim, a informação contida em cada grupo formado pode servir como entrada para o aumento de acurácia na identificação de características chaves em músicas, graças à forte correlação entre músicas em um grupo específico.

Mahajan et al. (2012) utilizam implementações dos algoritmos Apriori e *Sequential Pattern Mining* (SPAM), para buscar padrões sequenciais frequentes em uma base de dados contendo bilhões de classificações de usuários em músicas (Yahoo! *Music*). A partir desses padrões descobertos, é possível saber quais sequências de músicas são frequentemente ouvidas e em qual ordem elas são recomendadas, além de possibilitar a predição de sequências de classificações em músicas de usuários com gostos semelhantes.

Aizenberg et al. (2012) apresentam um modelo probabilístico para recomendadores de músicas. O modelo de filtragem colaborativa, ou do inglês *Collaborative Filtering* (CF), utiliza métodos para analisar padrões de escuta, ou outros tipos de retornos, de diversos usuários, para que assim, sejam estabelecidas similaridades entre usuários e itens. Para demonstrar sua capacidade de modelar e estabelecer padrões através de faixas, artistas e rádios, o modelo foi aplicado em uma coleção de metadados, coletados através de 4147 estações de rádios na internet, em um período de 15 dias. O artigo mostra que, quando devidamente analisadas, estações de rádio formam uma excelente coleção de sinais musicais gerados por pessoas, permitindo a modelagem de faixas e artistas, relacionar faixas e acompanhar as tendências populares. Como resultados, são apresentados os importantes benefícios que o modelo CF traz para os sistemas de recomendação.

Koenigstein et al. (2011) também apresentam um modelo para recomendação de músicas, que é baseado em uma matriz de fatoração e incorpora uma análise temporal

de classificação dos usuários e as tendências populares. No artigo, é descrito um método para compartilhamento de informações entre diferente itens de mesma taxonomia, que pode atenuar um efeito causado pelas técnicas de *Collaborative Filtering*, conhecido como *cold-start*, onde o consumo de músicas está inclinado para alguns artistas mais populares, tornando insuficiente a classificação de artistas menos conhecidos.

Lidy and Rauber (2006) apresentam um método para traçar perfis de estações de rádio, onde é possível obter qual é o foco das mesmas. É utilizada a técnica *Self-Organizing Maps* para organizar as programações das estações de rádio em um mapa bidimensional. A partir do mapa gerado, são derivados perfis visuais das rádios, considerando o que é mais predominante na programação, como por exemplo, músicas de determinados gêneros, propagandas ou participações de ouvintes por telefone. Com isso, o usuário tem a possibilidade de selecionar as rádios que mais se adequam ao seu gosto.

Ragno et al. (2005) mostram uma forma de inferir semelhanças entre objetos, de áudio ou vídeo, com base na proximidade de suas ocorrências. Na técnica, é construído um grafo onde as músicas representam os nós e as arestas contêm pesos que representam a proximidade entra as músicas. Através dessa técnica, dada uma música ou um artista, pode-se gerar uma *playlist* com as músicas relacionadas.

Todos os trabalhos citados se relacionam com este, pois apresentam estudos e análises sobre músicas e estações de rádios na internet, sob o ponto de vista da computação. No entanto, o foco do presente trabalho é a análise dos padrões das músicas executadas nas rádios, com a finalidade de se entender como são definidas as programações musicais e como estas sofrem influência de acontecimentos externos e de outras mídias, como TV e internet. Esse foco não foi observado em trabalhos relacionados.

3.2 Trabalhos associados à área da comunicação

Kischinhevsky (2011) faz um mapeamento dos impactos da automação das operações nas emissoras em Frequência Modulada (FM), com foco no Rio de Janeiro, diante dos cortes de gastos e das equipes, ocorridos nos últimos anos. Além disso, é colocada em discussão a hipótese da perda das características e da relevância simbólica do rádio aberto, ao se substituir profissionais por computadores. O autor afirma que o rádio AM/FM vem perdendo espaço entre as audiências jovens com mais acesso às novas tecnologias, e isso traz um impacto direto sobre as receitas publicitárias e sobre as estruturas de produção

e conteúdos radiofônicos.

Janotti Junior (2006) propõe uma análise midiática da música popular massiva, levando em consideração aspectos técnicos, econômicos e plásticos inscritos no formato de música. Para isso, é necessario compreender a Indústria do Entretenimento, especialmente a Indústria Fonográfica, seus agentes e consumidores, como um dispositivo midiático que deixa marcas nos produtos da música popular massiva, para compreensão do papel ocupado pelas canções na cultura e comunicação contemporâneas.

Macedo et al. (2011) analisam a grade de programação das seguintes emissoras de rádio: Rádio Nacional FM, Rádio Senado, Rádio Câmara e Rádio Justiça. O objetivo é avaliar se os gêneros de programação predominantes em cada emissora estão em concordância com sua identidade política/institucional, para saber se a programação de uma emissora de rádio institucional reflete sua identidade política. Foi concluído que a tradição do rádio educativo, criado na década de 1960, permanece, porém com uma nova abordagem, cujo foco é a educação para a democracia e cidadania, baseado no poder unilateral do Estado de decidir o que o cidadão deve saber.

Sanchotene et al. (2010) buscam compreender como os ouvintes de duas rádios comunitárias de Santa Maria - RS, a Caraí FM e a Comnorte, recebem os conteúdos transmitidos e como se estabelece a relação entre os moradores das comunidades e essas emissoras de rádio. Foi utilizada a metodologia dos grupos focais para a discussão e análise da recepção das programações e, entre os resultados, foi observada uma participação limitada por parte dos ouvintes.

Dubber (2013) investiga o que aconteceu com o rádio em meio a um ambiente midiático caracterizado pelas tecnologias digitais, reavaliando também o que ele sempre representou. O autor defende uma abordagem para o estudo do rádio por meio da ecologia midiática, capaz de reconhecer a periodização da história dos meios e de considerar a forma e o discurso do rádio como uma negociação entre potencialidades e efetividades.

da Silva et al. (2012) investigam o conceito de "Ecologia da Mídia", o qual tem como base fundamental a análise dos meios de comunicação, como estes são inseridos em um meio ambiente e como é o impacto da introdução dessa nova tecnologia para a sociedade, propondo o estudo desses comportamentos. Com base nesses conceitos, foi realizado um estudo de conteúdo da rodada final do Campeonato Brasileiro de Futebol de 2011, buscando fazer uma correlação entre a transmissão televisiva e a da internet. Chegou-se à conclusão de que por mais que as novas tecnologias tenham mais poder no cenário

dessa ecologia da mídia, todos os meios possuem suas peculiaridades e preferências com o público na transmissão do fato esportivo.

Prata (2008) estuda os novos gêneros e as novas formas de interação na webradio, que foi criada a partir de transformações tecnológicas e que aponta para um novo modelo de radiofonia. Para o estudo, foi realizada uma pesquisa qualitativa em cima de 30 emissoras de rádio divididas em três grandes grupos: hertzianas, hertzianas com presença na internet e webradios. A partir da pesquisa, percebeu-se que os gêneros antigos se reconfiguram nos novos, usando elementos já conhecidos e outros inusitados, como os signos textuais e imagéticos. Também observou-se que os produtores e comunicadores de webradios ainda não detêm domínio suficiente das novidades e recursos advindos com a internet. Por fim, conclui-se que os novos gêneros da webradio são: *chat*, e-mail, endereço eletrônico, enquete e fórum, nascidos genuinamente em meio digital.

Todos os trabalhos citados se relacionam a este, pois apresentam estudos, realizados a partir da perspectiva do campo da comunicação, que utilizam como objetos de estudo: músicas, rádios, meios de comunicação em massa e interação entre esses meios. Porém, esses trabalhos apresentam análises qualitativas, enquanto o presente trabalho utiliza dados monitorados de programações de rádios para aplicar análises quantitativas sobre padrões e impactos nas execuções musicais.

Capítulo 4

Análise de coocorrência e correlação

Neste capítulo, são apresentados os resultados de uma investigação para estudar a coesão do conteúdo musical das programações de rádio, ou seja, averiguar se existe correlação entre os itens (artistas, músicas e gêneros musicais) executados nessas programações. Basicamente, este estudo permite avaliar se a coocorrência de itens que são executados nas programações das rádios resulta em uma forte correlação entre esses itens, bem como se é possível descobrir novas correlações que podem não ser tão fáceis ou diretas de serem estabelecidas.

A motivação de tal investigação parte do pressuposto de que existe uma lógica nas programações das rádios e, normalmente, as músicas não são escolhidas de forma aleatória, pois os locutores ou DJs, em geral, possuem um grande conhecimento das músicas que eles executam, escolhendo cuidadosamente o sequenciamento e a relação entre elas para evitar a quebra de identidade do programa, como explicam Pachet et al. (2001).

Esse é um estudo preliminar, podendo inclusive apresentar uma fraca relação com o objetivo central deste trabalho. Porém, ele foi realizado para se obter um melhor entendimento sobre as características e padrões das programações das rádios, bem como para tentar compreender a forma como os componentes musicais estão conectados. As análises e descobertas demonstradas neste capítulo servem para complementar os resultados e contribuições apresentadas por este trabalho como um todo.

Este capítulo está organizado da seguinte forma. A Seção 4.1 apresenta a metodologia utilizada para a realização da análise apresentada neste capítulo. Em seguida, a Seção 4.2 mostra e discute os resultados experimentais dessa análise, apresentando as conclusões

sobre este estudo preliminar.

4.1 Metodologia de análise de coocorrência e correlação

Nesta seção, apresenta-se a metodologia elaborada para o estudo de correlação de itens, que abrange o processo de construção das bases transacionais utilizadas na análise realizada (Seção 4.1.1), a extração dos conjuntos de itens frequentes a partir das bases transacionais (Seção 4.1.2) e a forma como é realizada a comparação entre esses conjuntos e as listas de itens similares fornecidas por um importante serviço de recomendação e escuta de músicas. A Figura 4.1 ilustra o fluxo das etapas da metodologia desenvolvida.

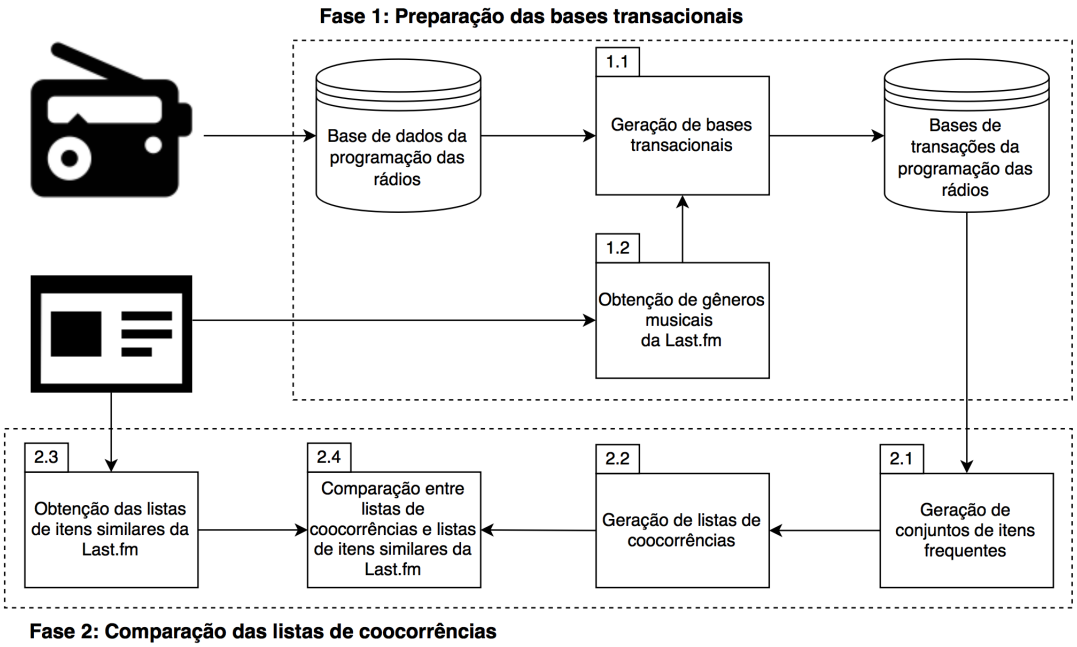

Figura 4.1: Etapas da metodologia de análise de coocorrência e correlação entre itens da programação das rádios.

As etapas da metodologia são divididas em duas fases. Na primeira fase (Preparação das bases transacionais), foram utilizadas técnicas de mineração de regras de associação para extração de conjuntos de itens frequentes (músicas, artistas e gêneros musicais) a partir de bases de transações que foram construídas utilizando-se dados da programação de estações de rádio, fornecidos pelo Radialize e pela SoundView (ver Seção 2.3). Na segunda fase (Comparação das listas de coocorrências), foi verificado o grau de similaridade entre os itens contidos nos conjuntos frequentes minerados e aqueles contidos nas

listas de itens similares extraídas da Last.fm[1]. Com isso, foi possível verificar se a coocorrência de itens (músicas, artistas e gêneros musicais) em programações de rádios indica a existência de correlação (similaridade) entre eles, tendo como referência a Last.fm. Mais detalhes sobre as etapas da metodologia, ilustradas na Figura 4.1, são descritos nas seções subsequentes.

4.1.1 Preparação das bases transacionais

Como mostra a Figura 4.1, a primeira fase da metodologia começa com a coleta de metadados de programações de estações de rádio. Essa análise utilizou os metadados coletados a partir dos serviços do projeto Radialize e da empresa SoundView (ver Seção 2.3). É importante ressaltar que esses metadados são públicos e podem ser coletados através da extração de dados nas páginas de transmissão das rádios pela *web*, pois geralmente são informados metadados, como os nomes do artista e da música no momento em que ela está sendo executada na rádio.

Para o presente estudo, a tarefa de coleta não foi realizada de fato, uma vez que foi utilizada a base de dados que havia sido previamente coletada e fornecida pelos serviços do Radialize e da SoundView. Os dados fornecidos correspondem a registros de execuções musicais relativos ao período de janeiro de 2011 a abril de 2014, correspondendo a 58.247 artistas, 335.676 músicas e 2413 rádios, de um total de 18.624.318 de registros. Cada registro contém os seguintes campos: código identificador da rádio, código identificador do artista, código identificador da música, nome do artista, nome da música, *timestamp* do início da execução da música e *timestamp* do fim da execução da música.

Após a carga dos dados na base de programação das rádios, o processo se inicia com a etapa 1.1 (Geração de bases transacionais), onde essa base é transformada em bases transacionais, de maneira que cada transação representa a programação de uma rádio, sendo que seus itens podem ser artistas, músicas ou gêneros musicais contidos na programação daquela rádio. Filtros de número de ocorrências mínimas de itens nas programações das rádios foram utilizados na geração das diferentes bases transacionais. Desse modo, a partir de um número mínimo de ocorrências definido no filtro, um item (por exemplo, artista) é incluído em uma transação apenas se ocorrer na programação da rádio esse número mínimo de vezes estabelecido no filtro. Ao longo deste trabalho serão apresentados resultados de experimentos realizados a partir de bases geradas utilizando-

[1]http://www.lastfm.com

se diferentes valores de frequência para os filtros.

A Figura 4.2 ilustra o processo de geração das bases transacionais a partir da base de metadados de programações de rádios. Para artistas, músicas e gêneros musicais foram geradas bases contendo diferentes quantidades de transações. Filtros de número de ocorrências mínimas de itens nas programações das rádios foram utilizados na geração das diferentes bases. Desse modo, a partir de um número mínimo de ocorrências definido no filtro, um item (por exemplo, artista) é incluído em uma transação apenas se ocorrer na programação da rádio esse número mínimo de vezes estabelecido no filtro. Ao longo deste trabalho, serão apresentados resultados de experimentos realizados a partir de bases geradas utilizando-se diferentes valores de frequência para os filtros.

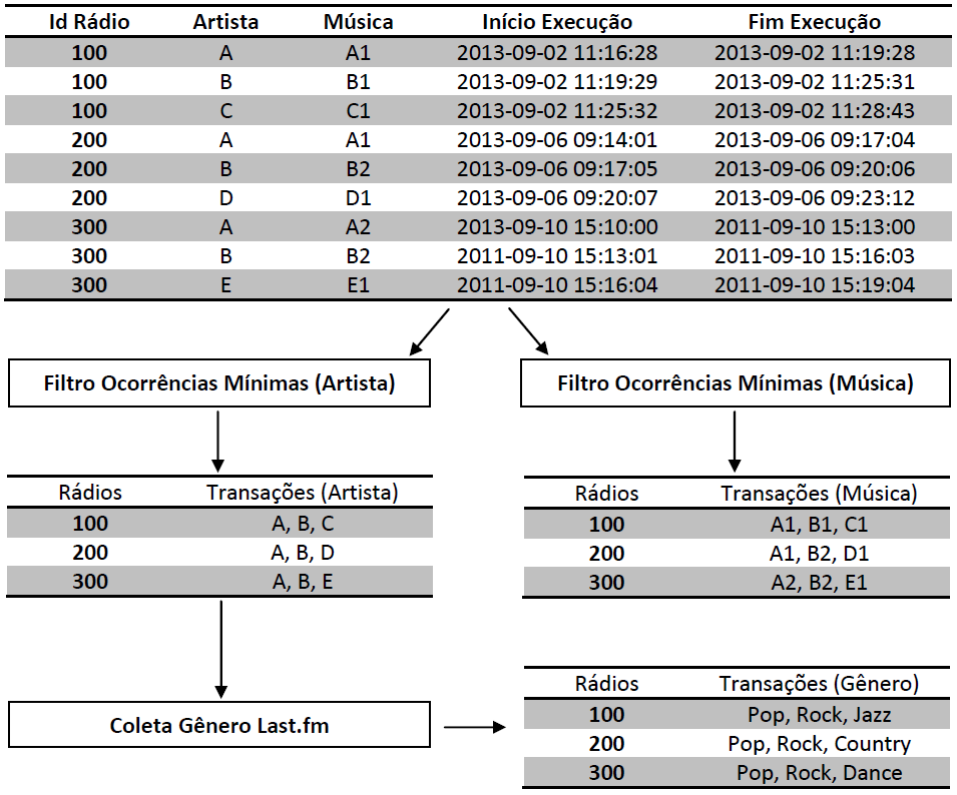

Figura 4.2: Processo de geração de bases transacionais

Como a base de metadados de programações de rádios não contém informações explícitas sobre os gêneros musicais presentes nas programações das rádios, essas informações foram coletadas a partir da Last.fm para se gerar as bases transacionais de gêneros musicais, como ilustra a etapa 1.2 (Obtenção de gêneros musicais da Last.fm) na Figura 4.1. As bases transacionais de gêneros musicais foram geradas a partir das bases transacionais de artistas, onde cada artista presente em uma transação foi substituído

pela primeira *tag* da lista de *tags* relacionadas àquele artista na Last.fm. Da lista de *tags* de cada artista obtida da Last.fm por meio de sua API foi utilizada apenas a primeira *tag*, pois essas *tags* são definidas por usuários e aquelas mais bem classificadas, ou seja, que estão no topo da lista, geralmente, representam os principais gêneros musicais daquele artista.

Tabela 4.1: Características das bases transacionais geradas

Artistas			
Filtro	**Nº Trans.**	**Qtde. Méd.**	**Desv. Pad.**
Sem filtro	2402	351,98	259,55
3	2031	277,08	187,91
5	1775	246,11	165,59
10	1514	214,72	139,86
Músicas			
Filtro	**Nº Trans.**	**Qtde Méd.**	**Desv. Pad.**
Sem filtro	2413	895,80	822,52
3	1891	585,54	421,11
5	1632	494,98	339,46
10	1325	399,22	265,69
Gêneros Musicais			
Filtro	**Nº Trans.**	**Qtde Méd.**	**Desv. Pad.**
Sem filtro	2374	143,86	94,69
3	2001	177,24	72,24
5	1747	105,63	65,19
10	1484	92,81	56,76

A Tabela 4.1 apresenta as características das bases transacionais geradas para artistas, músicas e gêneros musicais. Essa tabela apresenta os filtros que dizem respeito ao número mínimo de ocorrências de um item nas programações das rádios (1ª coluna). A tabela apresenta também, para essas bases transacionais, o número de transações (2ª coluna), o número médio de itens das transações (3ª coluna) e o desvio padrão do número de itens das transações (4ª coluna).

Como pode ser visto, os filtros do número mínimo de ocorrências de itens foram

variados (sem filtro, 3, 5 e 10). Nesse caso, quanto maior é o valor do filtro, menor é o número de rádios cuja programação atende ao critério estabelecido e, consequentemente, menor é o número total de transações da base gerada (lembrando que cada transação representa o conteúdo executado na programação de uma rádio).

4.1.2 Comparação das listas de coocorrências

A partir das bases transacionais de artistas, músicas e gêneros musicais, foram extraídos os conjuntos de itens frequentes, sendo a etapa 2.1 (Geração de conjuntos de itens frequentes) da metodologia, utilizando-se para isso a implementação do algoritmo Apriori Ye and Chiang (2006), disponível no pacote arules Hahsler et al. (2009) do *software* R Venables et al. (2002). Neste trabalho foram testados os seguintes valores de suporte mínimo: 5%, 10%, 15% e 20%. Além disso, foram extraídos somente conjuntos frequentes de tamanho 2, que posteriormente foram processados para gerar as listas de itens (músicas, artistas e gêneros musicais) que coocorrem nas programações das rádios, tratando-se da etapa 2.2 (Geração de listas de coocorrências).

A Figura 4.3 ilustra o processo de geração da lista de coocorrências a partir dos conjuntos frequentes extraídos pelo algoritmo Apriori. Por exemplo, os conjuntos frequentes {A, B}, {A, C} e {A, D} indicam que, na programação das rádios, o item A coocorre com os itens B, C e D.

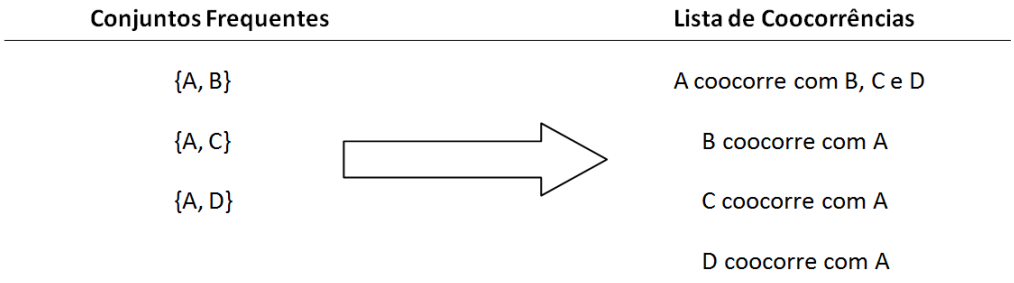

Figura 4.3: Processo de geração da lista de coocorrências

Após a geração das listas de coocorrências, realiza-se a comparação das mesmas com as listas de itens similares obtidas da Last.fm, como ilustrado nas etapas 2.3 (Obtenção de itens similares da Last.fm) e 2.4 (Comparação entre listas de coocorrências e listas de itens similares da Last.fm) da metodologia. A base de dados de itens similares da Last.fm foi escolhida por ser de um grande domínio público, conhecida como uma base robusta

por ter sido gerada por uma plataforma de serviços musicais consolidada e reconhecida, com um volume extenso de usuários em todo o mundo.

Para calcular a similaridade entre o conjunto de itens que coocorrem com o item j (obtidos da lista de coocorrências) e o conjunto de itens similares ao item j (obtidos da lista da Last.fm) foi desenvolvida a métrica M_j apresentada na Equação 4.1. Essa métrica retorna o percentual de itens que coocorrem com j e também estão presentes na lista de similaridades da Last.fm para o mesmo item j. Desse modo, a métrica é definida como a cardinalidade do conjunto resultante da interseção entre o conjunto A (itens que coocorrem com j) e o conjunto B (itens similares ao item j), dividido pela cardinalidade do conjunto A.

$$M_j = \frac{|A \cap B|}{|A|} \tag{4.1}$$

Por exemplo, suponha que a lista de coocorrências contenha a informação de que o artista x coocorre nas programações das rádios com os artistas w, y e z, ou seja, $A = \{w, y, z\}$. Por outro lado, a lista de artistas similares ao artista x (segundo a Last.fm) é dada pelo conjunto $B = \{y, k, t, v, m, n\}$. Sendo assim, a similaridade entre essas listas de acordo com a métrica proposta é $M_x = 1/3$, uma vez que temos apenas um artista em comum nos conjuntos A e B e três artistas contidos no conjunto A.

Além da métrica de similaridade elaborada (Equação 4.1), utiliza-se também o coeficiente de similaridade de Jaccard Jaccard (1901), também conhecido como índice Jaccard, como uma alternativa adicional para se fazer a análise de similaridade entre as listas de coocorrências e as listas de itens similares da Last.fm. O cálculo do índice de Jaccard é demonstrado na Equação 4.2.

$$J_j = \frac{|A \cap B|}{|A \cup B|} \tag{4.2}$$

Seguindo o exemplo utilizado para a Equação 4.1, o resultado do índice de Jaccard entre A e B, para o artista x, é $J_x = 1/8$. Nesse caso, estamos levando em consideração também o tamanho da lista de artistas similares ao artista x, de acordo com a Last.fm. No entanto, os tamanhos entre as listas vindas da Last.fm e as listas de coocorrências das rádios podem apresentar tamanhos muito diferentes, não representando bem os

resultados dos índices.

Portanto, além do próprio índice de Jaccard, também utiliza-se nas análises o quociente da divisão entre os índices de Jaccard obtidos e os tamanhos das listas de coocorrências correspondentes, demonstrado na Equação 4.3.

$$MJ_j = \frac{J_j}{|A|} \tag{4.3}$$

Aproveitando os exemplos dados para as Equações 4.1 e 4.2, o resultado da divisão entre J_x e A, para o artista x, é $MJ_x = (1/8)/3$.

A seguir apresenta-se a justificativa para o uso das três métricas nas análises. A primeira métrica (Equação 4.1) é utilizada para se saber a porcentagem dos itens das listas de coorrências (A) que estão contidos nas respectivas listas de itens similares da Last.fm (B). Já o índice de Jaccard (Equação 4.2) é usado para comparar a lista A com a lista B, verificando o grau de similaridade entre as duas listas. Por fim, temos a última métrica elaborada (Equação 4.3) que serve para tentar manter mais justos os resultados de Jaccard, pois com a aplicação dos filtros as listas de coocorrências ficam cada vez menores em relação às listas de itens similares da Last.fm, impactando nos resultados obtidos por Jaccard, podendo tornar os resultados tendenciosos.

4.2 Resultados experimentais de coocorrência e correlação

Nesta seção, são apresentadas as análises de similaridade feitas para artistas (Seção 4.2.1), músicas (Seção 4.2.2) e gêneros musicais (Seção 4.2.3).

4.2.1 Análise para artistas

Em relação aos resultados da análise para artistas, a Tabela 4.2 apresenta os valores médios de similaridade obtidos por meio da métrica de similaridade apresentada anteriormente, na Seção 4.1.2. Nessa tabela, a primeira coluna indica o tipo de filtro utilizado na geração da base de dados transacional (ver Seção 4.1.1). As demais colunas mostram a similaridade média entre as listas de coocorrências e a lista de itens similares da Last.fm (Sim.) e o tamanho médio das listas de coocorrências (Tam. List.) para três

valores diferentes de suporte mínimo: 5%, 10% e 15%.

Como pode ser visto na Tabela 4.2, para o suporte mínimo de 5% foram obtidas listas de coocorrências com tamanhos médios entre 81,5 e 186,3. Já os resultados dos cálculos de similaridade entre as listas de coocorrências e a lista de itens similares da Last.fm ficaram entre 0,23 e 0,34. Nota-se que, à medida em que o filtro de ocorrências mínimas de artistas aumenta, o tamanho médio das listas de coocorrências diminui e o valor médio de similaridade aumenta. O tamanho médio das listas diminui porque quanto maior é o filtro, menor é a quantidade de itens que compõem cada transação da base de dados transacional. Já o valor médio de similaridade aumenta porque as listas de coocorrências ficam mais restritas, permanecendo somente itens que estão mais correlacionados.

Tabela 4.2: Similaridade entre listas (Artistas)

	Suporte 5%		**Suporte 10%**		**Suporte 15%**	
Filtro	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**
Sem filtro	0,23	186,3	0,39	49,8	0,59	11,6
3	0,27	138,3	0,47	32	0,72	6
5	0,30	117,5	0,49	24,8	0,66	3,2
10	0,34	81,5	0,55	15,3	0,83	1,5

Para o suporte mínimo de 15%, maior suporte onde foi possível se obter conjuntos de itens frequentes das bases transacionais de artistas, foram obtidas listas de coocorrências com tamanhos médios entre 1,5 e 11,6. Os resultados dos cálculos de similaridade entre essas listas e as listas de itens similares da Last.fm ficaram entre 0,59 e 0,83, apresentando assim valores médios maiores do que aqueles valores obtidos para o suporte mínimo de 5%. Ao analisar esses números percebe-se que, além do filtro de ocorrências mínimas, quando se aumenta o suporte mínimo, as listas de coocorrências também ficam mais restritas, fazendo com que seus itens tendam a estar cada vez mais correlacionados.

Os resultados apresentados na Tabela 4.2 mostram que com o suporte mínimo e filtro de ocorrências mínimas baixos, o tamanho das listas de coocorrências geradas pode ser grande e os itens podem não estar muito correlacionados. Por outro lado, quanto maiores são esses filtros, menores são as listas de coocorrências e mais correlacionados são os itens nelas contidos, fazendo com que, geralmente, os valores médios de similaridade entre essas listas e a lista de itens similares da Last.fm sejam cada vez maiores.

A Tabela 4.3 apresenta, para a análise de artistas, os valores médios dos índices de Jaccard obtidos entre as listas de coocorrências e as listas de itens similares da Last.fm. A primeira coluna representa o tipo de filtro utilizado na geração da base de dados transacional. As demais colunas apresentam o índice de Jaccard médio entre as listas de coocorrências e as listas de itens similares da Last.fm (Jac.) e o resultado médio da divisão entre cada índice Jaccard obtido e o tamanho da lista de coocorrências correspondente (Jac. List.). Os valores de suporte utilizados foram os mesmos: 5%, 10% e 15%.

Pode-se observar na Tabela 4.3 que, para o suporte de 5%, à medida em que o filtro de ocorrências mínimas de artistas aumenta, o índice médio de Jaccard aumenta também, juntamente com o quociente médio entre Jaccard e o tamanho da lista de coocorrências. Porém, quando se aumenta o suporte (10% e 15%), o índice médio de Jaccard vai diminuindo, ao passo que o quociente entre Jaccard e o tamanho da lista continua aumentando.

Tabela 4.3: Similaridade Jaccard entre listas (Artistas)

	Suporte 5%		**Suporte 10%**		**Suporte 15%**	
Filtro	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**
Sem filtro	0,1216	0,0014	0,0871	0,0034	0,0408	0,0057
3	0,1373	0,0019	0,0797	0,0043	0,0281	0,0070
5	0,1437	0,0021	0,0697	0,0045	0,0172	0,0065
10	0,1549	0,0026	0,0585	0,0052	0,0099	0,0083

Esse comportamento do índice de Jaccard diminuir ocorre pois quanto maiores os valores de suporte e os filtros aplicados, menores ficam as listas de coocorrências, se distanciando do tamanho das listas de itens similares da Last.fm. Por outro lado, quando se observa o quociente médio da divisão do Jaccard pelo tamanho da lista de coocorrências, percebe-se que o mesmo vai aumentando à medida que os filtros aplicados vão ficando mais rigorosos, conforme também observado pela métrica utilizada na Tabela 4.2.

A Tabela 4.4 apresenta algumas novas correlações entre artistas que foram encontradas, ou seja, artistas que coocorrem nas programações das rádios e que não são considerados similares pela Last.fm. As configurações utilizadas para essa análise foram as seguintes: suporte mínimo do Apriori de 10% e filtro de ocorrências mínimas de 10 itens. Embora não sejam considerados similares pela Last.fm, foi observado que esses

artistas compartilham características, como por exemplo o estilo musical, representando conhecimento novo descoberto através da metodologia apresentada neste artigo, e que não foi descoberta pela Last.fm.

Tabela 4.4: Novas correlações encontradas entre artistas, que não estavam presentes na Last.fm

Artista	Novas Correlações
Alan Jackson	Carrie Underwood
Rihanna	Enrique Iglesias, Taylor Swift, Maroon 5, One Republic, Jason Mraz, Coldplay
Beyonce	Taio Cruz, Black Eyed Peas, Maroon 5
The Rolling Stones	Steve Miller Band, Tom Petty
3 Doors Down	Red Hot Chili Peppers, U2, Green Day
Green Day	Kings Of Leon, Train, U2

Embora não sejam considerados similares pela Last.fm, foi observado que esses artistas compartilham características, como por exemplo o estilo musical, representando conhecimento novo descoberto através da metodologia apresentada neste capítulo, e que não foi descoberta pela Last.fm.

4.2.2 Análise para músicas

A Tabela 4.5 apresenta a mesma estrutura da Tabela 4.2, porém os dados são relativos à análise de similaridade realizada para músicas. Percebe-se que para o suporte mínimo de 5% foram obtidas listas de coocorrências com tamanho médio entre 58,6 e 225,5. As similaridades médias entre essas listas e as listas de itens similares da Last.fm ficaram entre 0,20 e 0,35. Nota-se que esses resultados seguem o mesmo padrão de comportamento daqueles apresentados nas análises para artistas, ou seja, à medida em que o filtro de ocorrências mínimas de músicas aumenta, o tamanho médio das listas de coocorrências diminui e, geralmente, o valor médio de similaridade entre essas listas e as listas de itens similares da Last.fm aumenta.

Para o suporte mínimo de 10%, maior suporte onde foi possível se obter conjuntos de itens frequentes das bases de transações de músicas, foram obtidas listas de coocorrências

Tabela 4.5: Similaridade entre listas (Músicas)

	Suporte 5%		**Suporte 10%**	
Filtro	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**
Sem filtro	0,20	225,5	0,42	12,9
3	0,25	149,2	0,47	6,2
5	0,29	110,8	0,45	5
10	0,35	58,6	0,60	3

com tamanho médio entre 3 e 12,9. Os resultados de similaridade média entre essas listas e as listas de itens similares da Last.fm ficaram entre 0,42 e 0,60, apresentando assim valores médios maiores do que aqueles obtidos para o suporte mínimo de 5%. Assim como para as bases de dados de artistas, quando se aumenta o suporte mínimo, as listas de coocorrências ficam mais restritas, fazendo com que seus itens tendam a estar cada vez mais correlacionados.

A Tabela 4.6 segue o mesmo formato da Tabela 4.3, porém apresenta resultados relativos à analise para músicas. Essa tabela demonstra que, tanto para o suporte de 5% quanto para o suporte de 10%, à medida em que o filtro de ocorrências mínimas de artistas aumenta, o índice médio de Jaccard vai diminuindo, ao passo que o quociente entre Jaccard e o tamanho da lista continua aumentando.

Tabela 4.6: Similaridade Jaccard entre listas (Músicas)

	Suporte 5%		**Suporte 10%**	
Filtro	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**
Sem filtro	0,0724	0,0005	0,0135	0,0016
3	0,0640	0,0007	0,0088	0,0018
5	0,0568	0,0009	0,0085	0,0017
10	0,0439	0,0012	0,0050	0,0023

Observa-se, nesse caso, o mesmo comportamento identificado na análise para artistas (ver Tabela 4.3), pois as listas de coocorrências ficam cada vez menores do que as listas de itens similares da Last.fm. Porém, quando se observa o quociente médio da divisão do Jaccard pelo tamanho da lista de coocorrências, nota-se que os resultados aumentam

à medida que os filtros aplicados ficam mais rigorosos.

4.2.3 Análise para gêneros musicais

A Tabela 4.7 apresenta a mesma estrutura das Tabelas 4.2 e 4.5, porém os dados são relativos à análise de similaridade para gêneros musicais. Observa-se que para o suporte mínimo de 5% foram obtidas listas de coocorrências com tamanhos médios entre 32 e 51,1. As similaridades médias entre essas listas e as listas de itens similares da Last.fm ficaram entre 0,15 e 0,21. Nota-se que esses resultados seguem o mesmo padrão de comportamento dos resultados apresentados nas análises para artistas e músicas, onde ao se aumentar o filtro de ocorrências mínimas de itens, o tamanho médio das listas de coocorrências diminui e, geralmente, a similaridade média entre essas listas e a da Last.fm aumenta.

Tabela 4.7: Similaridade entre listas (Gêneros Musicais)

	Suporte 5%		**Suporte 10%**		**Suporte 15%**		**Suporte 20%**	
Filtro	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**	**Sim.**	**Tam. List.**
Sem filtro	0,15	51,1	0,19	32,2	0,26	21,6	0,38	15,3
3	0,18	40,5	0,24	24	0,43	17,1	0,58	9,9
5	0,21	35,9	0,28	22,8	0,45	13,7	0,56	7,9
10	0,18	32	0,41	17,8	0,54	10,5	0,66	5,1

Para o suporte mínimo de 20%, maior suporte onde foi possível se obter conjuntos de itens frequentes das bases de transações de gêneros musicais, foram obtidas listas de coocorrências com tamanhos médios entre 5,1 e 15,3. Os resultados de similaridade média entre essas listas e as listas da Last.fm ficaram entre 0,38 e 0,66, apresentando assim valores médios maiores do que aqueles obtidos para o suporte mínimo de 5%. Assim como para as bases de dados de artistas e de músicas, quando se aumenta o suporte mínimo, as listas de coocorrências ficam mais restritas, fazendo com que seus itens tendam a estar cada vez mais correlacionados.

A Tabela 4.8 segue o mesmo formato das Tabelas 4.3 e 4.6, porém apresenta resultados relativos à analise para gêneros musicais. A Tabela 4.8, mostra que quando o

suporte vai aumentando, até certo ponto (de 5% a 15%), o índice de Jaccard também vai aumentando. Mas a partir do suporte de 15% com filtro de 5, o índice de Jaccard médio começa a cair. Já o quociente médio entre Jaccard e o tamanho da lista continua sempre aumentando.

Tabela 4.8: Similaridade Jaccard entre listas (Gêneros Musicais)

	Suporte 5%		**Suporte 10%**		**Suporte 15%**		**Suporte 20%**	
Filtro	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**	**Jac.**	**Jac. List.**
Sem filtro	0,0702	0,0018	0,0759	0,0026	0,0918	0,0040	0,0973	0,0066
3	0,0730	0,0025	0,0896	0,0035	0,0980	0,0076	0,0754	0,0110
5	0,0758	0,0031	0,0962	0,0043	0,0890	0,0080	0,0659	0,0108
10	0,0844	0,0023	0,0983	0,0072	0,0891	0,0101	0,0567	0,0131

Os resultados das análises para gêneros musicais apresentam características semelhantes às análises de resultados para artistas e para músicas, variando-se apenas os valores apresentados.

Tabela 4.9: Novas correlações encontradas entre gêneros musicais, que não estavam presentes na Last.fm

Gênero	**Novas Correlações**
Thrash Metal	Rock
Glam Rock	Rock, Pop, New Wave, Progressive Rock
Rap	Pop, Soul
Heavy Metal	Rock, New Wave, Progressive Rock
Grunge	Classic Rock, New Wave, Punk Rock
House	Rock, Pop, Soul, Hip-Hop, Alternative, Indie, Rnb

A Tabela 4.9 segue a mesma estrutura da Tabela 4.4 (Seção 4.2.1), apresentando algumas novas correlações encontradas entre gêneros musicais, ou seja, gêneros que coocorrem nas programações das rádios e que não são considerados similares pela Last.fm.

As configurações utilizadas para essa análise foram as seguintes: suporte mínimo do Apriori de 10% e filtro de ocorrências mínimas de 10 itens. Como exemplos, temos coocorrências como Thrash Metal e Rock e também Heavy Metal e Rock, mostrando assim que a metodologia proposta é capaz de descobrir conhecimento novo a respeito da correlação entre gêneros musicais.

Analisando os resultados obtidos nesta seção, percebe-se que, quanto maior é a restrição, aumentando-se o filtro de ocorrências mínimas de itens nas programações das rádios e o suporte mínimo para geração de conjuntos de itens frequentes, melhores ficam os resultados da métrica de similaridade. Isso ocorre pois com filtros mais rigorosos aplicados, os itens coocorrentes resultantes tendem a ser mais correlacionados. Portanto, conclui-se que a informação de coocorrência de itens musicais em estações de rádio tem grande valor para extração de conhecimento sobre a correlação entre os itens.

4.3 Limitações do estudo de coocorrência e correlação

Nesta seção, são discutidas as limitações das análises realizadas, baseando-se nas categorias de ameaças à validade apresentadas por Wohlin et al. (2012). Desse modo, para cada categoria, são identificadas as possíveis ameaças à validade do estudo, bem como as medidas tomadas para reduzir essas ameaças. A seguir, são listadas as ameaças levantadas.

Validade interna: Uma ameaça à validade interna do estudo é a escolha do algoritmo para extração dos conjuntos de itens frequentes e o suporte mínimo utilizado. Para tentar minimizar esse risco, foi utilizado o Apriori, que é um algoritmo clássico para extração de conjuntos de itens frequentes, tendo bom reconhecimento da comunidade científica. Além disso, foram utilizados variados valores de suporte para minimizar o risco da escolha de apenas um valor de suporte, que poderia não ser o adequado.

Validade externa: Uma das ameaças identificadas para a validade externa foi a escolha correta das rádios a serem utilizadas para as coletas de dados de programações musicais e do *website* a ser utilizado para a obtenção dos conjuntos de itens musicais similares (Last.fm). Além disso, garantir que esses dados coletados representem a realidade das programações das rádios e também das correlações entre os itens musicais.

Em relação às rádios utilizadas no estudo, conforme apresentado na Seção 4.1.1 , as coletas dos dados foram realizadas de forma automática por serviços de terceiros.

Sendo assim, pode acontecer de determinadas rádios terem suas programações monitoradas mais regularmente do que outras. Portanto, em uma tentativa de se extrairem os melhores conjuntos de dados de rádios para as análises que melhor refletissem a realidade de suas programações, foram selecionados de maneira automática os dados de programações das rádios que foram coletados mais regularmente.

Já o *website* Last.fm, escolhido como fonte para obter os conjuntos de itens musicais similares, não garante que esses itens sejam realmente correlacionados. Porém, esse serviço foi escolhido por ser uma das principais plataformas de recomendação de músicas no mundo, contando com a colaboração de diversos usuários para as classificações dos itens musicais, gerando uma rica fonte de conhecimento coletivo, principalmente no que tange à similaridade entre artistas.

Validade de construção: identificou-se como uma ameaça de construção o cenário em que itens (músicas, artistas e gêneros musicais) iguais, provenientes de fontes distintas (rádios e Last.fm), fossem considerados diferentes por apresentarem alterações no título de uma fonte para a outra. Para tentar resolver esse tipo de problema, foi aplicada a distância de edição, através do algoritmo de Levenshtein, para a comparação dos títulos dos itens. Esse algoritmo permite inserções, remoções e substituições para fazer a comparação entre cadeias de caracteres Navarro (2001).

Validade de conclusão: uma ameaça à validade de conclusão do estudo é em relação à métrica utilizada para avaliar a similaridade entre os conjuntos de itens frequentes, extraídos das rádios, e os conjuntos de itens similares, extraídos da Last.fm. Embora tenha sido elaborada por este trabalho, não sendo uma métrica solidamente utilizada por demais estudos, foi a maneira mais coerente encontrada para se extrair uma taxa que representasse o grau de interseção entre os conjuntos de itens, para o problema em análise. Esta métrica é baseada no coeficiente de similaridade de Jaccard, que também foi usado nas análises. Porém o índice de Jaccard considera em seu denominador a união entre o conjunto de itens frequentes e a lista de itens similares da Last.fm, resultando assim em índices muito pequenos, uma vez que a interseção é dividida pela união dos conjuntos.

4.4 Conclusões

Com base nos resultados experimentais obtidos, nas análises realizadas neste capítulo, observou-se que é possível encontrar correlação entre os itens que coocorrem nas programações das rádios. Também verificou-se que, quanto mais restrito é o conjunto de itens analisado (aumentando-se o filtro de ocorrências mínimas de itens e o suporte do Apriori), maior é a similaridade entre as listas de coocorrências de itens das programações das rádios e as listas de itens similares produzida pela Last.fm. No entanto, novas correlações entre itens foram descobertas, ou seja, foram encontrados itens que coocorrem nas programações das rádios de forma muito frequente, mas que não são considerados similares pela Last.fm.

Capítulo 5

Metodologia de análise da influência da TV nas programações de rádios

Neste capítulo, é apresentada a metodologia de análise de influência desenvolvida com o objetivo de estudar a possível influência da mídia televisiva sobre a maneira com que as programações musicais das rádios são definidas. Mais precisamente, busca-se mensurar o impacto da exposição de artistas em programas de TV de alta audiência sobre as execuções de suas músicas nas rádios. Com esse propósito, a metodologia desenvolvida compreende processos que vão desde o monitoramento e a coleta de dados em programações de estações de rádio e programas de TV, até o cruzamento e análise desses dados. A Figura 5.1 apresenta o fluxograma da metodologia desenvolvida, bem como os processos que estão envolvidos e como eles se relacionam.

O monitoramento de dados relativos a artistas e músicas em programações de rádios começam pelo processo manual de "Levantamento de artistas e músicas para monitoramento"(ver Seção 5.1), onde foram selecionados os artistas e suas músicas a serem adicionados ao índice de monitoramento de artistas e músicas, que por sua vez foi utilizado pelo componente de "Reconhecimento por áudio das músicas em rádios"(ver Seção 5.2) para a identificação automática das execuções das músicas (juntamente com os artistas) a serem monitoradas nas programações das rádios. Os dados das execuções identificadas foram armazenados na "Base de dados da programação das rádios"para posterior análise em conjunto com os dados de ocorrências de artistas e músicas em programas de TV.

O monitoramento de dados relativos a participações de artistas musicais em pro-

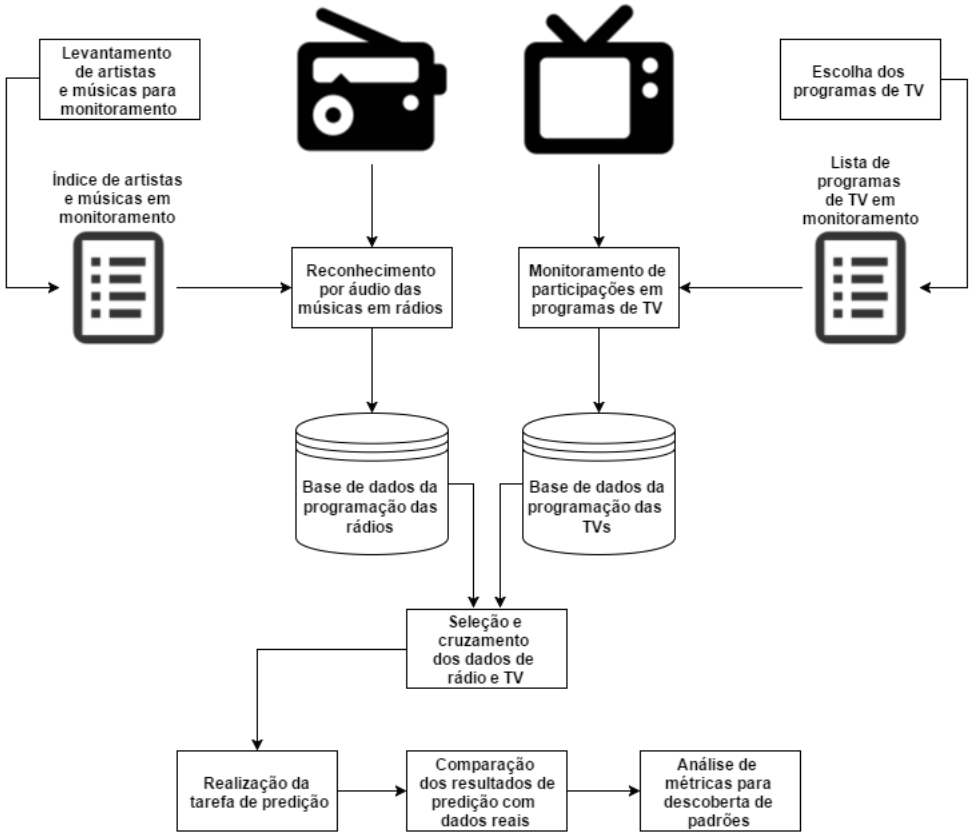

Figura 5.1: Fluxograma da metodologia de análise de influência da TV sobre execuções musicais nas rádios.

gramas de TV, que foram realizadas paralelamente ao processo de coleta de dados de artistas e músicas em programações de rádios, começam pelo processo manual de "Escolha dos programas de TV"(ver Seção 5.3) onde foram selecionados os programas de TV de alta audiência a serem adicionados ao índice de monitoramento de programas de TV, que por sua vez foi utilizado no processo manual de "Monitoramento de participações em programas de TV"(ver Seção 5.4) para controle dos programas de TV a serem acompanhados. Os dados de ocorrências coletados são armazenados na "Base de dados da programação das TVs"para posterior análise em conjunto com os dados de execuções de artistas e suas músicas em programações de rádios.

Após todo o processo de aquisição de dados, é realizada a "Seleção e cruzamento dos dados de rádio e TV"(ver Seção 5.5) a fim de realizar a verificação do impacto das ocorrências de artistas na TV sobre suas execuções nas rádios. Após a seleção e cruzamento dos dados, primeiramente ocorre a "Realização da tarefa de predição"(ver Seção 5.6), com o intuito de analisar e selecionar os métodos de predição mais adequados

para se observar o comportamento das execuções das músicas dos artistas nas rádios. Com os métodos selecionados, realiza-se a "Comparação dos resultados de predição com dados reais"(ver Seção 5.7) para verificar a expectiva das predições (acima, igual ou abaixo) em relação aos valores reais de número de execuções. Por fim, realiza-se a "Análise de métricas para descoberta de padrões"para observar se existe padrão de comportamento das execuções dos artistas nas rádios, provocado pela aparição dos mesmos em programas de TV.

5.1 Levantamento de artistas e músicas para monitoramento

Para monitorar as execuções de uma determinada música de um artista nas rádios era necessário encontrá-lo e cadastrá-lo previamente em um índice de monitoramento utilizado pelo componente de reconhecimento por áudio em rádios, para que assim fosse possível a identificação do exato momento e em qual rádio houve a execução da música. Para isso, foi preciso acompanhar informações relacionadas ao cenário musical nacional, em diferentes veículos de comunicação (rádio, TV e *web*), na tentativa de manter o índice de monitoramento atualizado com os lançamentos de artistas e músicas, além de manter no índice casos específicos a serem estudados.

Uma das dificuldades encontradas neste trabalho foi a descoberta, em tempo hábil, de músicas interessantes a serem inseridas no índice de monitoramento das rádios, como por exemplo, a descoberta de uma música antes de seu lançamento oficial. Na tentativa de contornar esse problema, a *web* foi o meio de comunicação mais utilizado para se buscar informações relativas a lançamentos musicais e notícias em geral envolvendo o cenário musical nacional, devido à maior disponibilidade e agilidade de disseminação das informações proporcionada por ela. Portanto, foram constantemente acompanhadas as últimas notícias divulgadas em *websites* especializados em música, como: Terra Música[1], G1 Música[2], UOL Música[3], Vagalume[4] e Paradão de Sucessos[5]. Além de *websites*, as redes sociais também serviram como uma importante fonte de informações, uma vez que os artistas as utilizam para a divulgação de seus trabalhos.

[1]http://musica.terra.com.br
[2]http://g1.globo.com/musica
[3]http://musica.uol.com.br
[4]http://www.vagalume.com.br
[5]http://www.paradaodesucessos.com

Com o intuito de monitorar informações sobre lançamentos musicais de maneira automática e ágil, foram desenvolvidos coletores para os serviços: Facebook e YouTube. Por meio do coletor do Facebook, monitorava-se as postagens das páginas oficiais dos artistas, ou de fã-clubes para o caso dos artistas que não tivessem uma página oficial. Por meio do coletor do YouTube, monitorava-se as postagens dos canais oficiais dos artistas.

As página oficiais dos artistas e de fã-clubes, no Facebook e no YouTube, foram buscadas e descobertas manualmente. Como semente para ambos os coletores, foi criada uma lista com 27 palavras-chave relacionadas ao meio fonográfico no intuito de tentar encontrar postagens importantes, contendo informações sobre lançamentos musicais. Alguns exemplos das palavras-chave utilizadas são: lançamento, música nova, *single*, *teaser*. As postagens coletadas eram exibidas em uma interface gráfica, para serem avaliadas manualmente quanto à relevância da informação contida, e assim, os lançamentos descobertos serem cadastrados rapidamente no índice de artistas e músicas em monitoramento, para então serem monitorados em rádios, conforme o passo seguinte (Seção 5.2).

5.2 Reconhecimento por áudio das músicas em rádios

A identificação da execução de uma música em uma rádio ocorre por meio do reconhecimento do áudio da mesma no *streaming* transmitido pela rádio. Após a identificação do momento de execução da música, são coletados os seguintes dados: nome da rádio, nome do artista, nome da música, *timestamp* do início da execução da música e *timestamp* do fim da execução da música.

As coletas de dados de artistas e músicas nas programações das rádios ocorreram por meio do uso da infraestrutura e dos serviços disponibilizados pela ferramenta Radialize e pela empresa SoundView. A apresentação da ferramenta e da empresa, bem como detalhes sobre o funcionamento dos serviços de coleta de dados nas rádios, se encontram na Seção 2.3.

5.3 Escolha dos programas de TV

Para as coletas de dados de ocorrências de artistas e músicas em programas de TV, primeiramente foram selecionados os programas semanais com participações musicais de maior audiência (ver Apêndice A) das cinco principais emissoras da televisão brasileira: Globo, Sistema Brasileiro de Televisão (SBT), Record, Bandeirantes (Band) e Rede TV. Para se saber a audiência dos programas de TV, foi utilizado o *website* RD1[6], que apresenta dados de audiências e notícias relacionadas à televisão brasileira. Dessa forma, os programas escolhidos para as coletas foram:

- **Globo**: Estrelas, Caldeirão do Huck, Altas Horas, Esquenta, Domingão do Faustão, Fantástico;
- **SBT**: Programa Raul Gil, Domingo Legal, Eliana, Programa Sílvio Santos, De Frente com Gabi;
- **Record**: Programa da Sabrina, Legendários, Domingo Show, O Melhor do Brasil, Hora do Faro, Domingo Espetacular;
- **Band**: Pânico na Band;
- **Rede TV**: Sábado Total, Ritmo Brasil, Frank Aguiar e Amigos.

Os programas selecionados foram inseridos no índice de monitoramento que consta apenas da lista dos programas de cada canal de TV em monitoramento.

5.4 Monitoramento de participações em programas de TV

Estando os programas de TV identificados e registrados no índice de monitoramento, o próximo passo consiste no monitoramento da apresentação de cada programa, com o objetivo de registrar participações de artistas. Cada participação de um artista em um programa de TV era registrada na base de dados de programação das TVs, por meio de uma interface gráfica. Os seguintes dados eram registrados: nome do artista, nome

[6]http://rd1.ig.com.br

do programa, data do programa e, se disponível, os nomes das músicas que o artista cantou.

O processo é totalmente manual, realizado de duas formas diferentes: assistindo a parte musical do programa, seja ao vivo ou por uma gravação disponível na *web*, ou acompanhando o *website* do programa. A coleta automática se tornou inviável, pois não foram identificados padrões de como tais informações são apresentadas pelos *websites* dos programas, tornando a coleta manual a alternativa mais simples e confiável.

O mesmo processo aplicado no monitoramento de participações de artistas em programas de TV pode ser facilmente estendido para o monitoramento de outros eventos envolvendo artistas e suas músicas, como acontecimento de shows, músicas que são escolhidas como temas de algo (no caso deste trabalho, temas de novelas e da Copa do Mundo de Futebol), entre outros eventos.

5.5 Seleção e cruzamento dos dados de rádio e TV

Para analisar a influência de programas de TV nas programações musicais das rádios, primeiramente são selecionadas participações de artistas em programas de TV, armazenadas na base de dados de programações das TVs. Em seguida, para cada participação de artista em um programa, foi extraído da base de dados de programação das rádios o somatório de execuções, agrupando por dias e por semanas, de todas as músicas monitoradas do artista em questão, dentro do período de três semanas (para a análise por semanas) e de três dias (para a análise por dias) posteriores e anteriores à ocorrência do artista no programa de TV, com o objetivo de, posteriormente, se avaliar o impacto da participação do artista na TV sobre os programas de rádio.

Para analisar uma música individualmente, no lugar de se extrair o somatório de execuções de todas as músicas do artista, extraiu-se o somatório apenas da música que continha o maior número de execuções, considerando que esta seria a música de trabalho do artista. Para a extração do somatório de execuções das músicas do artista, foram consideradas apenas as rádios, bem como as músicas, que foram coletadas regularmente, ou seja, que apresentaram pelo menos uma execução por semana/dia dentro do período analisado. Também foram selecionadas somente as rádios com o perfil musical do artista analisado, ou seja, rádios que continham em suas programações execuções de músicas do artista.

5.6 Realização da tarefa de predição

Após realizar o cruzamento dos dados, extraindo o somatório de execuções nas rádios de artistas que participaram em programas de TV monitorados, resta verificar a possível influência da aparição desses artistas na TV sobre a forma de execução de suas músicas nas rádios. A hipótese é que esse número de execuções, em uma situação normal, apresente um comportamento previsível, porém, aconteça algum aumento em função da participação do artista na TV. Para averiguar a hipótese, foram utilizados diferentes métodos para a predição do número de execuções de artistas nas rádios, na semana/dia imediatamente anterior, e também posterior, à participação do artista em um programa de TV. Desse modo, foi analisado e comparado o comportamento das execuções do artista nas rádios antes e depois de sua participação na TV.

O objetivo de utilizar métodos de predição nessa análise é observar a expectativa dos valores preditos em relação ao valor real de execuções. Caso o valor predito seja o mesmo ou muito próximo em relação ao valor real, significa que, levando em consideração as outras semanas/dias, o comportamento das execuções na semana/dia analisada pode ser considerado normal, ou seja, previsível. Caso contrário, pode ser um indicativo de que houve um comportamento inesperado no padrão de execuções do artista, provocado pela ocorrência do mesmo em algum programa de TV.

A seguir, serão apresentados os métodos de predição propostos para emprego na metodologia de análise elaborada. Os métodos apresentados na Seção 5.6.1 se baseiam na média de crescimento/queda das execuções para fazer as predições. Já na Seção 5.6.2, são descritos os métodos de regressão linear e não-linear para predizer os números de execuções nas semanas/dias a serem analisados, utilizando para isso dados de execuções passadas. Por último, na Seção 5.6.3 são descritos os métodos de preenchimento de valores faltantes, que fazem predições com base tanto nos dados passados quanto nos dados futuros de execuções.

5.6.1 Métodos de médias simples

O primeiro método proposto representa a taxa média de crescimento/queda de execuções musicais entre semanas/dias, que é calculada da seguinte maneira. Determina-se um intervalo de semanas/dias que antecedem a semana/dia a ser analisada, e então, calcula-se a taxa de variação de execuções de uma semana/dia para a outra dividindo-se o

número de execuções da semana/dia posterior pelo número de execuções da semana/dia imediatamente anterior. Esse processo é realizado sucessivamente até se chegar na última semana/dia do intervalo estabelecido. Por fim, calcula-se a média aritmética simples das taxas de variação de execuções obtidas para o intervalo, obtendo-se a taxa média de crescimento/queda das execuções entre as semanas/dias, no período especificado. Ressalta-se que a semana/dia em análise não entra nesse processo de cálculo da taxa média de crescimento/queda das execuções.

Quando a taxa média de crescimento/queda apresenta valores acima de 1, significa que as execuções estão com tendência de crescimento. Caso os valores forem iguais a 1, representa que as execuções estão estáveis. Já valores abaixo de 1 representam que as execuções têm uma previsão de queda.

Após obter a taxa média de crescimento/queda das execuções entre semanas/dias, a mesma é utilizada para predizer o número de execuções da semana/dia em análise, da seguinte forma. Toma-se o número de execuções da semana/dia anterior à semana/dia em análise e multiplica-se pela taxa média de crescimento obtida. Desse modo, o valor encontrado representa a predição do número de execuções da semana/dia analisada, considerando a taxa média de crescimento de execuções entre semanas/dias utilizadas como referência.

O segundo método elaborado segue a mesma proposta do método de taxa média de crescimento/queda de execuções, apresentado anteriormente, porém utiliza ponderação no cálculo da taxa média de crescimento das execuções entre semanas/dias, onde o peso da primeira semana/dia do intervalo utilizado é de 1,3, incrementando-se 0,30 no peso para cada semana/dia subsequente, de maneira que as semanas/dias mais próximas à semana/dia de ocorrência de um artista na TV (que será predita) recebem os maiores pesos.

5.6.2 Métodos de regressão

Além dos métodos mais simples apresentados anteriormente, elaborados na tentativa de predizer o número de execuções da semana/dia a ser analisada, foram utilizadas implementações de algoritmos próprios para esse tipo de tarefa. Para isso, foram utilizados os seguintes métodos: análise de regressão linear, utilizado como um modelo linear típico para predições numéricas (ver Seção 2.4.1); e *Random Forest*, implementado por Breiman (2001) para problemas de classificação e regressão, sendo que, para esse último, as

florestas de regressão utilizadas são para modelos de regressão múltipla não-linear (ver Seção 2.4.1).

Ambos os métodos foram utilizados a partir de implementações disponíveis no *software* R, nas quais passa-se como parâmetro, para as funções dos respectivos métodos, um conjunto de dados utilizados para o treinamento dos modelos e para a realização das predições. As funções, por sua vez, retornam os valores que foram preditos.

Ressalta-se que esse conjunto de dados utilizado como treinamento dos modelos de predição representa o número de execuções das semanas/dias anteriores à semana/dia a ser analisada. Já as funções, retornam os valores preditos para as semanas em análise, com base nos modelos treinados.

5.6.3 Métodos de preenchimento de valores faltantes (Imputation)

Por fim, foram utilizados métodos de preenchimento de valores faltantes (*Imputation*), na tentativa de se buscar resultados de predição melhores, uma vez que utilizam valores passados e futuros para o treinamento do modelo de predição, podendo assim alcançar uma taxa de acerto maior do que os demais métodos citados anteriormente, que por sua vez utilizam apenas valores passados para realizar as predições.

Foram utilizadas implementações de métodos de preenchimento de valores faltantes, disponíveis no pacote yaImpute do *software* R, onde podem ser encontrados diversos métodos para a busca de valores vizinhos mais próximos (kNN) para usar como referência no preenchimento dos valores faltantes em um conjunto de dados. Para todos os métodos disponíveis no pacote, exceto o que é baseado em uma matriz de proximidade *Random Forest*, o conceito de proximidade é definido com base na distância euclidiana ponderada, $dist(p, q, W) = [(p-q)^T W(p-q)]^{1/2}$, onde W é a matriz de pesos. A principal diferença entre esses métodos é como o espaço de busca dimensional é construído (como a matriz de pesos é definida). Os métodos, disponíveis no pacote, utilizados por este trabalho foram os seguintes: raw, euclidean, mahalanobis, msn e randomForest. Crookston et al. (2008) apresentam mais detalhes sobre o pacote yaImpute, bem como o funcionamento dos métodos citados disponíveis nele.

O método raw se baseia na distância dos valores não transformados, ou crus, de variáveis X, ao passo que o método euclidean usa variáveis X normalizadas para definir a distância. O método mahalanobis transforma o espaço de busca pelo inverso da matriz de covariância de variáveis X, antes de computar as distâncias. Para o método

msn, a distância é computada em um espaço canônico projetado. Por fim, no método randomForest, os valores são considerados similares quando eles tendem a um mesmo nó terminal em um conjunto de árvores de classificação e regressão.

5.7 Comparação dos resultados de predição com dados reais

Para verificar a variação entre os números reais de execuções e os números preditos, foi utilizada a Equação 5.1, onde o número real (NR) é divido pelo número predito (NP), obtendo-se como resultado uma taxa de variação (T) entre esses números.

$$T(NR, NP) = \frac{NR}{NP} \tag{5.1}$$

Caso o valor de T seja 1, significa que o número real é exatamente o mesmo que o número predito, e portanto, manteve-se o padrão de execuções que era esperado pelo preditor. Se T for maior que 1, indica que o número real é maior do que o número predito, superando a expectativa do preditor. Caso contrário, se o valor de T for menor que 1, representa que o número real foi menor do que o número predito, indicando que a expectativa do preditor foi maior do que a realidade.

Após a obtenção das taxas de variação (T) entre os valores reais e os preditos do número de execuções nas rádios, depois de cada participação de artistas em programas de TV, foram analisados os valores de T obtidos para observar se houve alguma mudança significativa no padrão das execuções, causada pela participação do artista na TV, ocasionando "erros"nas predições realizadas.

Como o volume de dados é grande, foram utilizados dois tipos de análises para se obter medidas que proporcionassem uma visão geral dos dados. Na primeira forma de análise, calculam-se as médias simples e ponderadas dos valores de T, observando em seguida, se os valores das médias obtidas tendem a ser próximos a 1 ou não, para verificar a previsibilidade dos padrões de execuções. Os pesos utilizados nas médias ponderadas foram: audiência do programa em que o artista apareceu, número de execuções de músicas do artista na semana/dia analisada, número de rádios que executaram o artista na semana/dia analisada.

O segundo tipo de análise realizada foi utilizando testes de hipóteses (ver Seção 2.5) para verificar se os números de execuções preditos pertencem a uma mesma população que os números de execuções reais (hipótese nula H_0), ou seja, os testes avaliam se existe ou não uma diferença significativa entre um conjunto de valores e o outro. Para testar a hipótese nula (H_0), foi utilizado o teste de Friedman (ver Seção 2.5.1). Em caso de negação da hipótese nula para Friedman, foi o utilizado o teste de Wilcoxon Pareado (ver Seção 2.5.2) para verificar qual das amostras comparadas é maior ou menor. Ambos os testes foram aplicados por meio do *software* R, que conta com implementações dos mesmos. Essas análises são adequada para este estudo, pois é preciso saber formalmente (baseando-se em testes de hipóteses) se os valores reais sofreram alterações após as participações dos artistas na TV, ficando distantes dos valores preditos, e em caso de alterações, quais foram (aumento ou diminuição).

Capítulo 6

Resultados das análises de influências nas programações de rádios

Neste capítulo, são apresentadas as análises sobre os dados coletados das programações das rádios, realizadas com o intuito de investigar possíveis influências de determinados eventos sobre as execuções, nas rádios, das músicas que estão relacionadas a esses eventos. Como foi explicado na Seção 5.4, um evento é qualquer acontecimento, externo às rádios, que envolva um determinado artista (ou uma música do artista) e que possa impactar no padrão de execuções das músicas (ou de apenas uma determinada música) do artista nas rádios.

Este capítulo está organizado da seguinte forma. A Seção 6.1 apresenta as configurações e preparação dos dados para a realização das análises. Já na Seção 6.2, são apresentadas as análises de influências das ocorrências de artistas em programas de TV sobre as execuções de suas músicas nas rádios. Por fim, na Seção 6.3 são realizadas as análises de influências de outros eventos (lançamentos de músicas e músicas tema) sobre as execuções de músicas nas rádios.

6.1 Configurações para as análises

Todas as análises apresentadas neste capítulo foram feitas sobre o somatório de execuções (agrupando por dias e por semanas) das músicas (ou apenas a música de trabalho) dos artistas nas rádios. Vale destacar que foram consideradas apenas as rádios, bem como as músicas, que foram coletadas regularmente, ou seja, que apresentaram pelo menos

uma execução por semana/dia dentro do período analisado. Também foram selecionadas somente as rádios com o perfil musical do artista analisado, ou seja, rádios que continham em suas programações execuções de músicas do artista.

O monitoramento de execuções de artistas e músicas em programações de rádios para as análises foi iniciado em janeiro de 2014 e prosseguiu até novembro do mesmo ano, excluindo-se o mês de junho devido a uma falha nas coletas de dados de rádios por áudio. Ao longo desse tempo, foram coletados 3.087.917 registros de execuções de 177 artistas e de 313 músicas diferentes, para um total de 2.207 rádios tradicionais brasileiras monitoradas.

As coletas de dados de ocorrências de artistas e músicas em programas de TV monitorados ocorreram no mesmo período das coletas em rádios, ou seja, de janeiro de 2014 a novembro do mesmo ano. Foram cadastrados dados de participações de 352 artistas e de 408 músicas diferentes, distribuídos pelos 21 programas monitorados, com um total de 1128 apresentações dos artistas nos programas.

Para a realização das análises de influência da TV nas programações das rádios, foi feito o cruzamento dos dados de ocorrências de artistas nos programas de TV monitorados (ver Seção 5.3) com os dados de execuções de músicas desses artistas nas rádios (ver Seção 5.5), seguindo a metodologia apresentada no Capítulo 5. Após o cruzamento desses dados, foram utilizados métodos para predição de valores de execuções nas semanas/dias a serem analisadas (ver Seção 5.6). Os resultados das predições foram comparados com os números de execuções reais (ver Seção 5.7) para se avaliar a taxa de acerto dos preditores (selecionando-se para a análise aqueles que apresentam melhores resultados de predição) e para identificar possíveis influências da TV que alterassem o padrão das execuções nas rádios, causando erros nas predições. As tabelas completas com as taxas de acerto de execuções dos preditores, para cada ocorrência de artista na TV, estão disponíveis no Apêndice C.

Para avaliar a taxa de acerto dos preditores, foram utilizados dados de execuções em períodos em que os artistas não tiveram qualquer registro de participações em programas de TV, de forma que os resultados dos modelos não sofressem influência de dados provenientes de semanas/dias em análise, que poderiam ser tendenciosos.

Conforme apresentado na Seção 5.6, foram testados diversos métodos como preditores, que por sua vez foram divididos em três classes: métodos de médias simples, métodos de regressão e métodos de preenchimento de valores faltantes. A Tabela 6.1 apresenta os resultados das médias e os respectivos desvios padrão para as taxas de

variação dos valores reais sobre os valores preditos (taxa de acerto) para os métodos de preenchimento de valores faltantes.

Tabela 6.1: Avaliação das predições por preenchimento de valores faltantes

Semanas						
	Median	**MAL**	**RAW**	**EC**	**MSN**	**RF**
Média	1,0084	0,9962	0,9962	0,9962	0,9962	1,0188
Desvio Padrão	0,0394	0,0586	0,0586	0,0586	0,0586	0,0708
Dias						
	Median	**MAL**	**RAW**	**EC**	**MSN**	**RF**
Média	0,9930	1,0255	1,0255	1,0255	1,0255	1,0255
Desvio Padrão	0,1435	0,0674	0,0674	0,0674	0,0674	0,0674

Já a Tabela 6.2 apresenta os resultados das médias e os desvios padrão para as taxas de variação dos valores reais sobre os valores preditos para os métodos de regressão: linear (RL), Random Forest (RF), além dos métodos de médias simples elaborados por este trabalho, Taxa Média de Crescimento/Queda de Execuções (TME) e Taxa Média Ponderada de Crescimento/Queda de Execuções (TMPE).

Tabela 6.2: Avaliação das predições por regressões

Semanas				
	TME	**TMPE**	**RL**	**R. Forest**
Média	0,9106	0,9086	1,0213	0,9306
Desvio Padrão	0,2235	0,2100	0,2404	0,2185
Dias				
	TME	**TMPE**	**RL**	**R. Forest**
Média	0,8772	0,8845	0,9572	0,9036
Desvio Padrão	0,2156	0,2116	0,1835	0,2280

Observa-se que os métodos que apresentaram melhores resultados foram os de preenchimento de valores faltantes, pois as médias ficaram próximas a 1 e os desvios padrão

foram pequenos, indicando que as taxas de acerto foram maiores do que as taxas obtidas pelos métodos de regressão. No entanto, não é possível assumir que apenas os métodos de preenchimento de valores faltantes são suficientes para as análises, uma vez que esse tipo de método leva em consideração valores futuros, ou seja, de períodos após a participação do artista em um programa de TV, podendo gerar ruído nos resultados. Portanto, são utilizados nas análises subsequentes também os resultados do melhor método de regressão, além do melhor método de preenchimento de valores faltantes.

Nesse sentido, o método de regressão que apresentou média mais próxima a 1 e o menor desvio padrão foi o RL. Já os métodos de preenchimento de valores faltantes MAL, RAW, EC e MSN apresentaram resultados melhores e iguais entre si, pelo fato de compartilharem um modelo de predição semelhante, variando apenas o cálculo para os pesos das matrizes de pesos utilizadas nas predições. Nesse caso, poderia se escolher qualquer um dentre esses métodos, sendo o escolhido o método EC, que utiliza distância Euclidiana para o cálculo dos pesos das matrizes. Portanto, apenas os resultados desses dois métodos selecionados foram utilizados nas análises neste capítulo. As tabelas com todas as taxas de variação dos preditores avaliados se encontram no Apêndice B.

Nas análises, foram verificadas as taxas de variação entre os valores reais e os valores preditos para as execuções posteriores às participações de artistas em programas da TV monitorados, observando assim se houve alguma alteração no padrão de execuções das músicas nas rádios após a aparição dos artistas na TV. Para observar essas diferenças nas taxas e tirar as conclusões sobre o estudo, foram utilizadas médias simples e ponderadas, além de testes de hipóteses (ver Seção 5.7), para investigar as hipóteses levantadas no início do trabalho (ver Seção 1.3).

Foram realizadas análises a partir de diversas dimensões a saber: por músicas de trabalho e músicas agrupadas, pelo ano inteiro e por partes do ano (primeira e segunda metade), por número de execuções por dias e por semanas, além de músicas em ascensão, estabilidade e queda. Essa variedade de análises se justifica para investigar a existência de influências que podem não estar evidentes em uma visão mais geral, mas podem ser encontradas através de visões mais específicas.

Nas análises de lançamentos de músicas e de músicas utilizadas como temas, foi possível estudar apenas parte dos casos, devido à dificuldade de descoberta, em tempo hábil, de músicas relacionadas a esses eventos para inserção das mesmas no índice de monitoramento nas rádios, conforme relatado na Seção 5.1.

6.2 Análises de influências dos programas de TV

Nesta seção são apresentadas as análises da influência de participações de artistas em programas de TV sobre as execuções de suas músicas nas rádios. Os resultados apresentados correspondem a análises realizadas sobre as taxas de acerto dos preditores Euclidean (EC) e Regressão Linear (RL), ou seja, as taxas de variação (ver Seção 5.7) entre os valores reais e preditos de execuções nas rádios (agrupadas por semanas e por dias).

As Tabelas 6.3, 6.4, 6.5 e 6.6 apresentam os resultados das médias simples (2ª e 5ª colunas) e os respectivos desvio-padrões (3ª e 6ª colunas), bem como os resultados dos testes de Friedman (4ª e 7ª colunas). Esses resultados são relativos às taxas de variação dos valores reais de execuções nas rádios sobre os valores preditos, posteriores às participações dos artistas em qualquer programa de TV monitorado. Os resultados apresentados nessas tabelas estão divididos em "Músicas Geral", onde são contabilizadas as execuções de todas as músicas dos artistas; e "Música Trabalho"onde são contabilizadas apenas as execuções de uma música de cada artista, a chamada música de trabalho, que é a música que o artista divulga como seu principal lançamento em um dado período de tempo.

As Tabelas 6.3 e 6.4, apresentam resultados relativos a análises realizadas por semanas, ou seja, os números de execuções de músicas nas rádios foram agrupados por semanas, enquanto as Tabelas 6.5 e 6.6 apresentam os resultados para as análises realizadas por dias. A Tabela 6.3 apresenta as taxas de variação relativas ao ano de 2014, divididas em: primeira metade do ano, segunda metade do ano e ano todo. Desse modo, os resultados podem ser analisados a partir de cenários diferentes durante o ano.

Nota-se que os resultados obtidos para o método EC demonstram que não houve alterações significativas entre os valores reais e os preditos, pois as médias ficaram bem próximas a 1 e os resultados do teste de Friedman foram altos, o que significa que a expectativa das predições, tendo como referência os valores das outras semanas, está na média em concordância com a realidade das execuções. Já os resultados apresentados pelo método RL (Tabela 6.3) foram contrários à hipótese, pois o teste de Friedman mostra que houve variação entre os valores reais e os preditos. Porém os valores reais ficaram, na média, abaixo dos valores preditos, uma vez que o valor-p, que é uma estatística de teste, ficou abaixo de 0,05 em todos os casos. Desse modo, de acordo com esses resultados, não foi observada variação no padrão de execuções esperado que

pudesse indicar a influência positiva da aparição dos artistas nos programas de televisão sobre suas execuções nas rádios. De fato, houve uma variação negativa.

Tabela 6.3: Análise sobre as taxas de variação relativas à primeira metade do ano, segunda metade do ano e ano inteiro (Análise por semanas)

Primeira Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0045	0,0743	0,4733	1,0036	0,0814	0,8111
RL	0,9597	0,1303	1,56E-08	0,9869	0,1546	0,0011
Segunda Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0041	0,0831	0,3763	1,0122	0,0799	0,1011
RL	0,9588	0,2300	5,89E-05	0,9588	0,2300	5,19E-05
Ano Inteiro						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0043	0,0777	1,0000	1,0071	0,0806	0,2254
RL	0,9593	0,1816	2,2E-16	0,9742	0,1922	1,13E-08

Por sua vez, a Tabela 6.4 apresenta os resultados das taxas de variação de músicas classificadas, com base na taxa média de crescimento/queda de execuções (ver Seção 5.6.1), como: em ascensão, em estabilidade e em queda. Desse modo, os resultados podem ser analisados a partir de comportamentos de músicas em fases de execuções diferentes. Assim como os resultados apresentados na Tabela 6.3, percebe-se que os resultados exibidos na Tabela 6.4 indicam que não existiu uma variação no padrão de execuções esperado, mesmo para as músicas em diferentes fases de execução (em ascensão, estabilidade e queda), pois nos casos de hipótese nula negada, para o método RL, a variação foi contrária, sendo os valores preditos maiores que os reais.

Os resultados apresentados adiante, nas Tabelas 6.5 e 6.6, seguem o mesmo padrão das análises realizadas por semanas, porém as análises foram realizadas com dados de

Tabela 6.4: Análise sobre as taxas de variação relativas às músicas em ascensão, estabilidade e queda (Análise por Semanas)

Músicas em Ascensão						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0049	0,0522	1,0000	1,0103	0,0570	0,0038
RL	0,9758	0,2402	3,42E-03	0,9804	0,2359	0,0054
Músicas em Estabilidade						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	0,9926	0,0251	0,2568	0,9981	0,0359	0,4795
RL	0,8925	0,1111	0,0046	0,8971	0,1227	0,0081
Músicas em Queda						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0052	0,0888	0,7449	1,0067	0,0933	1,0000
RL	0,9573	0,1514	1,44E-08	0,9772	0,1701	5,97E-06

execuções por dias, conforme metodologia apresentada no Capítulo 5. A Tabela 6.5 apresenta os resultados das taxas de variação, divididas em: primeira metade do ano, segunda metade do ano e ano todo. Assim como nas análises por semanas, através dos resultados apresentados na Tabela 6.5, não foi observado um impacto positivo sobre as execuções nas rádios mediante a participação dos artistas nos programas de TV, para as análises por dias.

A Tabela 6.6, por sua vez, mostra os resultados das taxas de variação de músicas classificadas como: em ascensão, em estabilidade e em queda. Para essa análise por dias e por fases de execuções das músicas, os resultados do teste de Friedman ficaram altos (acima de 0,05) para todos os casos e métodos, indicando portanto que não houve uma interferência aparente que aumentasse as execuções das músicas, causada pela participação dos artistas em programas de TV.

A seguir, é apresentada uma análise de influência da participação de artistas em

Tabela 6.5: Análise sobre as taxas de variação relativas à primeira metade do ano, segunda metade do ano e ano inteiro (Análise por dias)

Primeira Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0177	0,0667	0,2230	1,0017	0,0564	0,8575
RL	0,9730	0,9953	0,0076	0,8933	0,4186	0,0349
Segunda Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0351	0,1104	0,7815	1,0351	0,1104	0,7815
RL	1,1471	0,7552	0,3173	1,1471	0,7552	0,3173
Ano Inteiro						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0225	0,0804	0,3763	1,0109	0,0756	1,0000
RL	1,0441	0,9056	0,1655	0,9970	0,5906	0,3490

programas de televisão com maiores audiências, com o objetivo de tentar validar a hipótese ao menos para participação de artistas em programas de grande audiência. Portanto, foram selecionadas as participações nos dois principais programas da Rede Globo de Televisão: Fantástico e Domingão do Faustão. Para esses casos, as análises se deram também por meio de observações de médias aritméticas ponderadas das taxas de variação do número de execuções reais sobre o número de execuções preditas. As ponderações utilizadas foram: audiência do programa de TV, número de execuções de músicas do artista na semana/dia analisadas, número de rádios que executaram a(s) música(s) do artista na semana/dia analisadas.

A Tabela 6.7 apresenta os resultados (por semanas e por dias) das médias simples (1ª coluna), desvios padrão (2ª coluna) e das médias para cada ponderação: número de execuções das músicas (3ª coluna), número de rádios que executaram as músicas (4ª coluna) e audiência do programa de TV (5ª coluna). Além disso, são apresentados também os resultados do teste de Friedman (6ª coluna).

Tabela 6.6: Análise sobre as taxas de variação relativas às músicas em ascensão, estabilidade e queda (Análise por dias)

Músicas em Ascensão						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0290	0,1013	1,0000	1,0272	0,1095	0,7630
RL	0,9086	0,9674	0,8815	1,0568	0,4352	0,6617
Músicas em Estabilidade						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	0,9856	0,0298	0,5637	0,9926	0,0294	0,6547
RL	1,0425	0,1944	1,0000	0,9375	0,2007	0,4795
Músicas em Queda						
	Músicas Geral			**Música Trabalho**		
	Média	**D. Pad.**	**Fried.**	**Média**	**D. Pad.**	**Fried.**
EC	1,0231	0,0718	0,2230	1,0045	0,0546	0,7055
RL	1,1332	0,8844	0,0895	0,9595	0,7131	0,1521

Tabela 6.7: Análise para os programas da Rede Globo: Fantástico e Domingão do Faustão

Semanas						
	Média	**D. Pad.**	**Nº Mús.**	**Nº Rád.**	**Aud.**	**Fried.**
EC	1,0038	0,0555	1,0184	1,0127	1,0085	0,6374
RL	1,0094	0,1324	0,9829	1,0004	1,0143	0,0010
Dias						
	Média	**D. Pad.**	**Nº Mús.**	**Nº Rád.**	**Aud.**	**Fried.**
EC	1,0246	0,1036	1,0125	1,0198	1,0286	0,5930
RL	1,1784	0,7522	0,9625	0,9893	1,2450	0,3458

Mesmo para essa análise mais específica para programas de maiores audiências, de acordo com os resultados obtidos, não foram encontradas evidências que pudessem in-

dicar um padrão de impacto positivo nas execuções após os artistas participarem desses programas. Os resultados, tanto para as médias simples quanto para as médias ponderadas, ficaram próximos a 1. Em apenas um caso a hipótese nula para Friedman foi negada, na qual para os valores de execuções por semanas, preditos pelo método RL, variam dos valores reais, deixando o valor-p abaixo de 0,05. Este caso será analisado mais adiante neste capítulo, através do teste de Wilcoxon.

Como pôde ser visto nos resultados apresentados até aqui, em determinados casos, principalmente relacionados ao método de predição RL, o valor-p obtido pelo teste de Friedman foi muito abaixo de 0,05, sendo portanto, rejeitada a hipótese nula e aceitando-se a hipótese alternativa de que existe uma diferença entre as duas amostras analisadas. Para esses casos, é interessante saber qual das amostras apresenta valores maiores que a outra, pois se for observado que os valores reais são maiores que os preditos, pode ser um indicativo de que a expectativa do preditor foi superada por influência de programas de TV.

Para verificar qual das amostras analisadas é maior, foi utilizado o teste de Wilcoxon Pareado (ver 2.5.2), no qual são comparadas as duas amostras e é obtido um valor-p, assim como no teste de Friedman, que indica se a hipótese nula é aceita ou rejeitada. Como o teste de Wilcoxon se baseia na mediana dos dados, caso a hipótese nula seja rejeitada, observa-se qual das amostras apresenta mediana maior, para se identificar a amostra maior.

As Tabelas 6.8, 6.9, 6.10, e 6.11 apresentam os resultados dos testes de Wilcoxon para os casos em que a hipótese nula foi rejeitada pelo teste de Friedman. Essas tabelas seguem a seguinte estrutura: as colunas nomeadas como "Wilcox"apresentam os resultados (valor-p) dos testes de Wilcoxon, as colunas nomeadas como "Med. R."apresentam as medianas dos conjuntos de valores reais de execuções, enquanto as colunas nomeadas como "Med. P."apresentam as medianas dos conjuntos de valores preditos.

A Tabela 6.8 apresenta os resultados do teste de Wilcoxon para as análises por semanas, realizadas para a primeira metade do ano, segunda metade do ano e ano inteiro. Os resultados sugerem que a hipótese nula foi negada também para o teste de Wilcoxon, ou seja, existe uma diferença entre os conjuntos de valores reais e os conjuntos de valores preditos. Ao observar as medianas, nota-se que os conjuntos de valores preditos são maiores do que os conjuntos de valores reais, mostrando que as expectativas do preditor ficaram acima dos valores reais. Portanto, não é possivel afirmar que há impacto positivo

Tabela 6.8: Resultados dos testes de Wilcoxon relativos às análises da primeira metade do ano, segunda metade do ano e ano inteiro (Análise por semanas)

Primeira Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	6,96E-08	237,5	323	0,0003	209	224
Segunda Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	1,47E-05	65	68	9,11E-08	19	68
Ano Inteiro						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	2,2E-16	168	269	6,00E-10	119	200

de participações de artistas em programas de TV nas execuções nas rádios.

Por sua vez, a Tabela 6.9 apresenta os resultados do teste de Wilcoxon para as análises realizadas por semanas, para as músicas em ascensão, estabilidade e queda. Os resultados da Tabela 6.9 demonstram que a hipótese nula também foi negada para esses casos, aceitando-se que existe diferença entre as amostras. Os valores das medianas mostram, novamente, que os valores preditos estão acima dos valores reais.

A Tabela 6.10 apresenta os resultados do teste de Wilcoxon para as análises por dias, realizadas para a primeira metade do ano. O resultado para música de trabalho indica que a hipótese nula foi aceita, ou seja, diferentemente do teste de Friedman, o teste de Wilcoxon sugere que não há fatores que comprovam a existência de diferenças (maior ou menor) entre os valores reais e preditos.

Por fim, a Tabela 6.11 apresenta os resultados do teste de Wilcoxon para as análises sobre os programas de maior audiência da Rede Globo de Televisão: Fantástico e Domingão do Faustão. Assim como nos demais casos, os valores de medianas mostram que os valores preditos são maiores que os reais.

Tabela 6.9: Resultados dos testes de Wilcoxon relativos às músicas em ascensão, estabilidade e queda (Análise por Semanas)

Músicas em Ascensão						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
EC	-	-	-	0,0522	58	56,5
RL	1,42E-02	19	58	0,0015	18,5	64
Músicas em Estabilidade						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	0,0141	538,5	1161,5	0,0156	144	389
Músicas em Queda						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	1,83E-09	193,5	304	1,39E-06	161,5	227,5

Tabela 6.10: Resultados dos testes de Wilcoxon relativos à primeira metade do ano (Análise por dias)

Primeira Metade do Ano						
	Músicas Geral			**Música Trabalho**		
	Wilcox.	**Med. R.**	**Med. P.**	**Wilcox.**	**Med. R.**	**Med. P.**
RL	0,0133	50	51	0,1403	38	43

Tabela 6.11: Resultados dos testes de Wilcoxon relativos às análises para os programas: Fantástico e Domingão do Faustão

	Semanas		
	Wilcox.	**Med. R.**	**Med. P.**
RL	0,0006	260,5	375,5

Para todos os casos dos testes de Wilcoxon apresentados anteriormente, pode-se observar que não há evidências de que os valores reais de execuções tenham sofrido alguma influência dos programas de televisão, pois mesmo quando a hipótese nula é rejeitada e aceita-se a hipótese de que existe uma diferença entre os valores reais e preditos, ao se comparar as medianas, nota-se que os conjuntos de dados preditos apresentam valores maiores que os conjuntos de dados reais.

6.3 Análises de influências de outros eventos

Nesta seção, são apresentadas análises realizadas por meio de observações de acontecimentos de eventos que possam influenciar no comportamento das execuções de determinadas músicas nas rádios. Dois eventos identificados que causaram impacto nas execuções das músicas em rádios foram os seguintes: o lançamento de uma música específica e a Copa do Mundo de Futebol.

As análises foram realizadas sobre gráficos que apresentam o número de execuções da música por dia, agrupadas por regiões do Brasil (Sudeste, Centro-Oeste, Norte, Sul e Nordeste), sendo que para cada região está destacado (entre parênteses, nos gráficos que serão apresentados adiante) o número de rádios diferentes que executaram a música, dentro do período de tempo estipulado para a análise.

Em relação à Copa do Mundo de Futebol, músicas que possuem letras com a temática voltada para o futebol ou que de alguma forma se tornam temas que embalam o evento, geralmente atingem uma disseminação maior, em diversas mídias de massa, durante o período em que acontece o evento. A seguir, são apresentados dois exemplos desse efeito no rádio.

O gráfico da Figura 6.1 apresenta o comportamento das execuções da música Todo Mundo da cantora Gaby Amarantos (com participação do Monobloco) durante o período

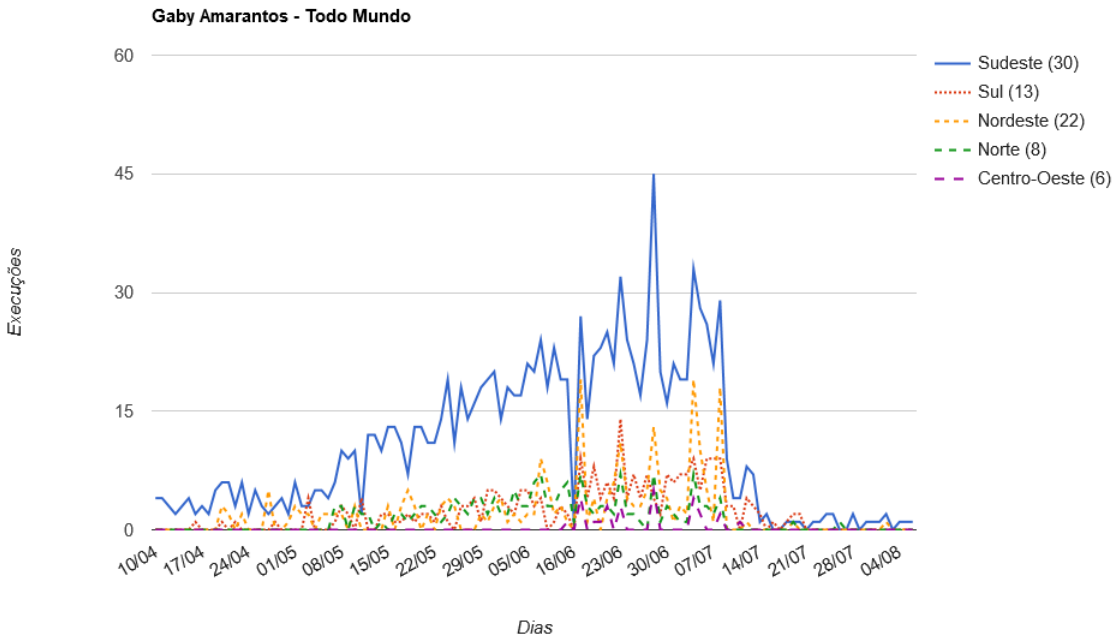

Figura 6.1: Execuções da música Todo Mundo, de Gaby Amarantos (part. Monobloco) no período da Copa do Mundo de Futebol de 2014.

em que ocorreu a Copa do Mundo de Futebol de 2014, que compreende o período de 12 de junho a 13 de julho de 2014. Essa música foi utilizada como um dos temas do evento, pois além de possuir na letra palavras relacionadas ao futebol, foi também alavancada pela parceria com uma grande empresa de refrigerantes que foi uma das patrocinadoras oficiais dessa Copa do Mundo.

Como pode ser observado no gráfico, destaca-se um aumento nas execuções da música começando no mês de maio (por volta do dia 08/05), antes mesmo do início da Copa do Mundo, alcançando valores maiores durante a ocorrência do evento em questão (de 12/06 a 13/07) e caindo novamente após o término do mesmo. Os maiores picos de execuções observados (23/06, 30/06 e 07/07) ocorreram nos dias dos jogos da seleção brasileira de futebol. A queda definitiva das execuções (a partir de 09/07) se deu após a eliminação da seleção brasileira pela seleção da Alemanha no dia 08/07, com o surpreendente placar de 7 a 1. Já a queda brusca nas execuções (próximo ao dia 16/06) representa um ruído nos dados devido a uma falha na coleta dos dados das programações das rádios.

O gráfico da Figura 6.2 apresenta o aumento nas execuções da música País do Futebol

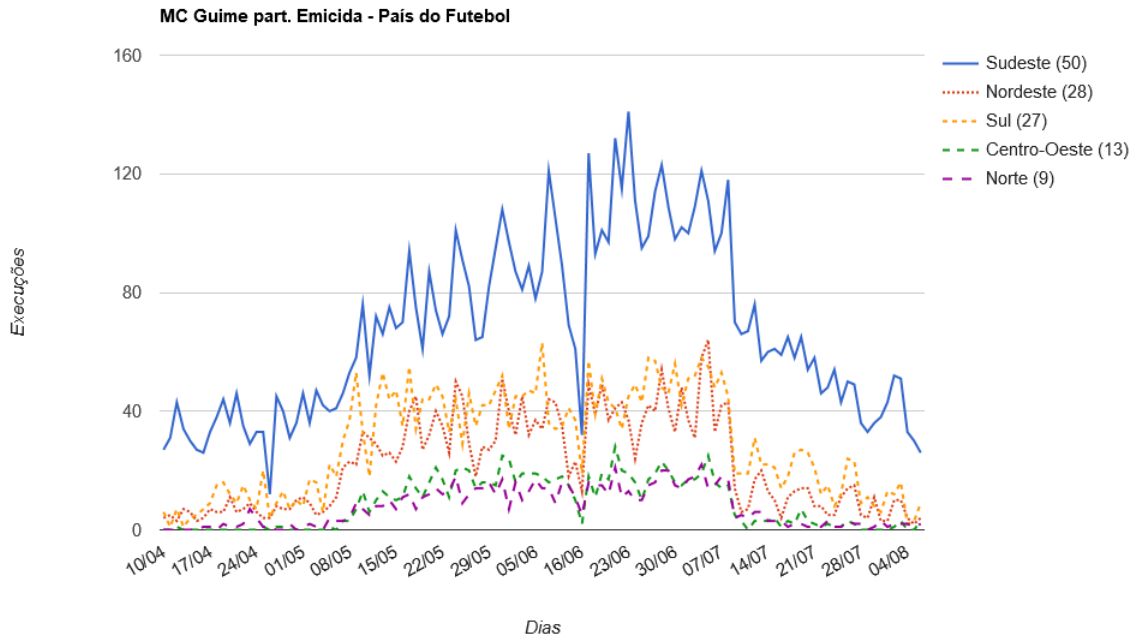

Figura 6.2: Execuções da música País do Futebol, de MC Guimê (part. Emicida) no período da Copa do Mundo de Futebol de 2014.

do artista MC Guimê, com participação do artista Emicida, durante o período da Copa do Mundo de Futebol de 2014. Essa música tem como foco principal o futebol e a comemoração da Copa do Mundo de 2014, em que o Brasil foi sede. Ela também foi selecionada como tema de abertura da telenovela Geração Brasil, da Rede Globo de Televisão, que estreiou em maio de 2014 e durou até outubro do mesmo ano.

Também é possível observar um aumento nas execuções da música durante a época da Copa do Mundo, seguindo o mesmo comportamento verificado na música de Gaby Amarantos analisada anteriormente. Após o término da Copa do Mundo o número de execuções volta a cair, não sendo observado algum impacto pelo fato da música ser tema de abertura de uma telenovela.

Já em relação a lançamentos de músicas, existe um empenho dos artistas e das gravadoras na divulgação do material de trabalho a ser lançado, que é largamente disseminado nas diferentes mídias, sejam elas de massa ou digitais. No caso das rádios, mesmo que não exista efetivamente uma ação promocional de divulgação, pode acontecer de rádios executarem o conteúdo em lançamento para acompanharem outras rádios que também

estejam fazendo o mesmo, satisfazendo os ouvintes e mantendo-se em constante atualização com o universo musical. A seguir, apresenta-se o comportamento das execuções nas rádios de algumas músicas antes, durante e após o dia de lançamento oficial.

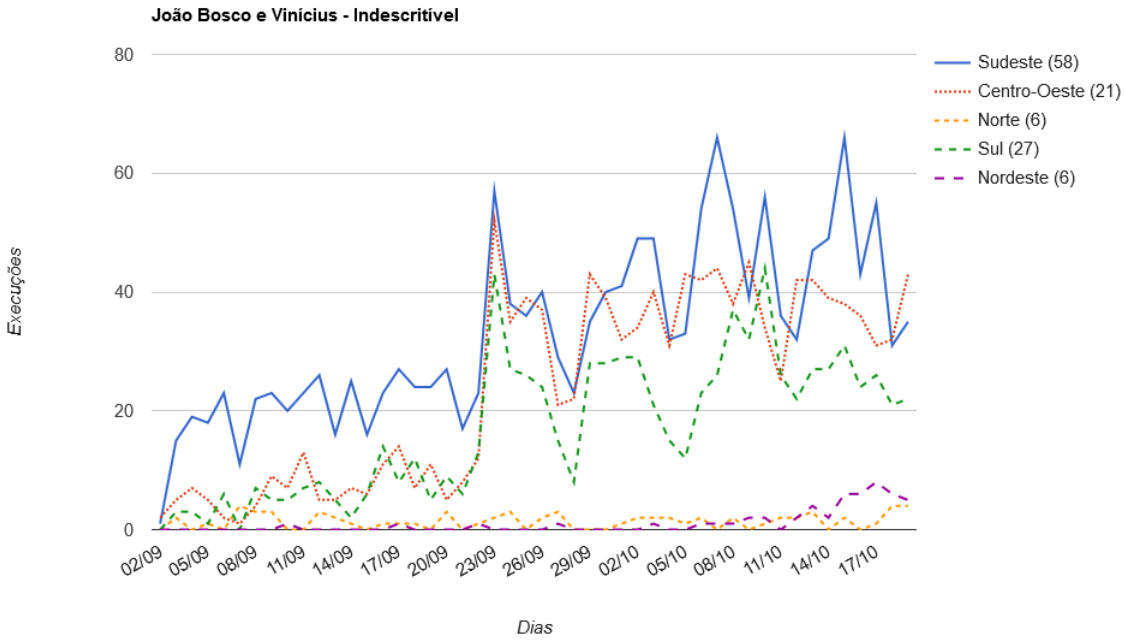

Figura 6.3: Lançamento da música Indescritível, da dupla João Bosco e Vinícius.

O gráfico da Figura 6.3 mostra o comportamento das execuções nas rádios da música Indescritível da dupla sertaneja João Bosco e Vinícius, durante os dias em torno do lançamento oficial. Observando os dados das regiões Sudeste, Centro-Oeste e Sul, nota-se que, antes do lançamento oficial da música, as rádios já estão executando-a, porém poucas vezes ao dia se comparado com o número de execuções por dia após o lançamento. Já no dia do lançamento (23/09) ocorre um pico no número de execuções e esse valor se mantém alto mesmo após o lançamento, sendo praticamente mais do que o dobro do que era antes do lançamento oficial.

Em mais um exemplo do impacto provocado pelo lançamento de uma música nas rádios, o gráfico da Figura 6.4 exibe o comportamento das execuções da música Ele Não Vai Mudar, da dupla sertaneja João Neto e Frederico, antes, durante e após o seu lançamento oficial. Analisando principalmente a região Sudeste, no dia do lançamento

da música, ocorre um salto evidente no número de execuções, indo de menos de 60 para mais de 180, e caindo novamente nos dias posteriores ao lançamento; porém os números de execuções permanecem significativamente maiores do que os apresentados antes do lançamento.

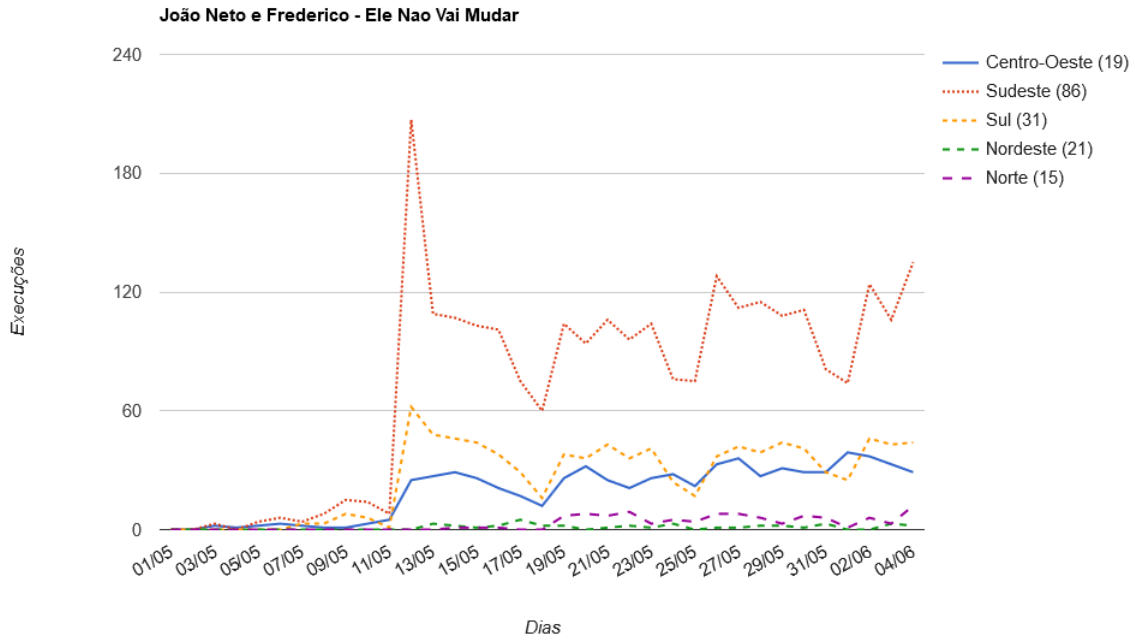

Figura 6.4: Lançamento da música Ele Não Vai Mudar da dupla João Neto e Frederico.

Em mais um caso, o gráfico da Figura 6.5 mostra o comportamento das execuções da música Implorando pra Trair, do cantor sertanejo Michel Teló durante seu lançamento oficial. Nas regiões Sudeste, Centro-Oeste e Sul destaca-se o impacto positivo do lançamento da música em suas execuções nas rádios. Após o lançamento, o número de execuções cai, mas continua maior do que antes do lançamento, seguindo o mesmo padrão dos demais exemplos apresentados.

Em todos os gráficos apresentados a análise foi focada principalmente na região Sudeste, pelo fato de ser a região com o maior número de rádios monitoradas. Em relação aos lançamentos musicais, os casos analisados representam apenas um estilo musical (Sertanejo Pop ou Sertanejo Universitário), não sendo possível afirmar que tal comportamento das execuções se repetiria no lançamento de músicas de outros artistas, com

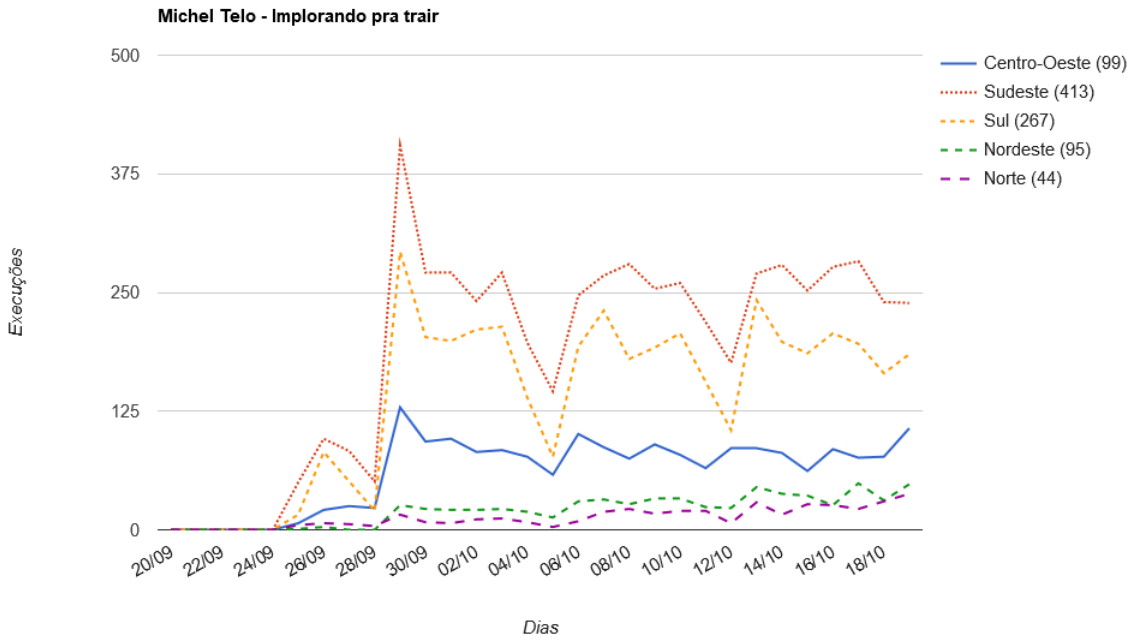

Figura 6.5: Lançamento da música Implorando pra Trair de Michel Teló.

outros estilos musicais.

6.4 Limitações do estudo

Nesta seção são discutidas as limitações das análises de influências realizadas, baseando-se nas categorias de ameaças à validade apresentadas por Wohlin et al. (2012), na qual são identificadas as possíveis ameaças à validade do estudo. Além das ameaças, são descritas as medidas tomadas para reduzir essas ameaças. A seguir, são listadas as ameaças levantadas.

Validade interna: Uma ameaça à validade interna do estudo é a escolha dos métodos de predição para predizer as execuções musicas nas semanas a serem analisadas. Para tentar minimizar esse risco, foram avaliados diversos métodos, selecionando para as análises aqueles que apresentam os melhores resultados de predição, a partir de uma análise preliminar realizada sobre os dados reais (ver Seção 6.1).

Validade externa: Assim como apresentado na Seção 4.3, a escolha das rádios

utilizadas é uma ameaça para a validade das análises realizadas neste capítulo, pois os dados foram coletados e fornecidos pelos serviços de terceiros, na qual as coletas podem acontecer de forma irregular, sendo algumas rádios mais monitoradas do que outras. Para tratar essa ameaça, foi realizado um pré-processamento para se extrair o melhor conjunto de rádios para as análises, filtrando de maneira automática as rádios cujos dados foram coletados mais regularmente.

Validade de construção: A escolha das músicas e artistas ideais a serem monitorados pode ser considerada outra ameaça à validade do estudo. Nesse caso, além dos artistas e músicas que já estavam no índice de monitoramento das rádios, inseridos pelos projetos fornecedores dos dados, foram utilizados artistas e músicas inseridos através do presente trabalho, utilizando o critério de situações interessantes a serem analisadas, como lançamentos musicais, além de artistas e músicas em evidência (para análise de impacto da TV sobre as rádios).

A escolha correta dos programas de TV a serem monitorados é mais uma ameaça identificada. Para esse caso, a fim de se minimizar os riscos, foram utilizados os programas semanais de maior audiência, contendo participações musicais, pertencentes às principais e mais populares emissoras da televisão brasileira.

A coleta de dados de participações de artistas nos programas foi realizada de forma manual, o que também representa uma ameaça. Para minimizar os riscos, esse dados foram coletados regularmente através do *website* oficial de cada programa. Além disso, os dados eram conferidos sistematicamente para se evitar erros.

Ressalta-se que o volume de dados utilizados nas análises foi grande, sendo um estudo bastante amplo. O monitoramento de artistas e músicas em programas de rádio e de TV se estendeu por todo o ano de 2014. Foram monitoradas 2.207 rádios nacionais, gerando 3.087.917 registros de execuções de 177 artistas e de 313 músicas diferentes. Já para a TV, foram registradas 1128 participações de 352 artistas, distribuídos pelos 21 programas monitorados. Portanto, com o período de tempo de monitoramento e a quantidade de dados gerados, foi possível realizar diversas análises, gerando resultados relevantes para o estudo.

Validade de conclusão: As métricas utilizadas para avaliar os impactos das partipações dos artistas em programas de TV representam uma ameaça à validade de conclusão das análises. Para tanto, utilizou-se testes de hipóteses para reforçar as conclusões das análises, aplicando-se formalidades estatísticas aceitas por estudiosos.

Capítulo 7

Conclusões e trabalhos futuros

Neste trabalho, foram utilizados metadados de programações musicais de estações de rádio na internet, com o intuito de verificar a existência de padrões de execução, analisando como essas programações são definidas e se sofrem influência de uma importante mídia em massa, a televisão.

Com relação às análises de coocorrências e correlação, o intuito foi investigar se o conjunto de informações contidas nas programações de rádios, geradas normalmente por DJs ou locutores que são especialistas na área da música, apresentam uma correlação lógica entre itens que coocorrem, podendo assim se tornar uma base de conhecimento importante para sistemas de recomendação de músicas. Para verificar se existe uma correlação entre os itens que coocorrem nas programações das rádios, foram utilizados como referência os conjuntos de itens similares retornados pela Last.fm.

As descobertas realizadas com essas análises são úteis tanto para evidenciar a existência da correlação entre os itens musicais das programações das rádios, quanto para melhorar o desempenho de sistemas de recomendação como a própria Last.fm, que tipicamente fazem uso de dados de similaridade para aumentar a qualidade da recomendação. Portanto, conclui-se que a metodologia proposta, no Capítulo 4, não somente é capaz de encontrar itens realmente similares a partir da comparação com a lista de similaridade da Last.fm, como também é capaz de identificar novos itens similares que não haviam sido identificados pela Last.fm. De forma geral, conclui-se também que os dados de programação de rádios são bastante ricos para a descoberta de conhecimento sobre correlações entre itens.

Como trabalho futuro para as análises de coocorrências e correlação, almeja-se reali-

zar análises em bases de metadados de programações de rádios visando identificar quais itens aumentam a probabilidade de ocorrência de outros itens. Outra sugestão de trabalho futuro é estender o trabalho para um maior número de rádios, visando descobrir ainda mais similaridades entre itens.

Com relação às análises de influências sobre as execuções musicais nas programações das rádios, o objetivo foi investigar a hipótese levantada de que a participação de um artista da música em um programa popular de televisão poderia causar um aumento nas execuções de suas músicas nas rádios, após a aparição do artista na TV.

Para realizar essa investigação, primeiramente, foi desenvolvida uma metodologia de análise que abrange desde o monitoramento e coleta dos dados tanto das programações das rádios quanto dos programas de TV, até o cruzamento e análise desses dados. Após o cruzamento dos dados de TV e rádio, verificou-se qual seria o impacto das participações de artistas da música em programas de televisão sobre as execuções de suas músicas nas rádios, por meio da verificação dos padrões das execuções antes e depois das participações nos programas de TV.

A melhor maneira encontrada para analisar o padrão das execuções e identificar situações anormais que pudessem indicar um impacto causado pela TV foi utilizando técnicas de predições numéricas. Foram avaliados diversos métodos de predição, porém os que apresentaram melhores resultados foram os algoritmos de preenchimento de valores faltantes, pelo fato de considerarem no treinamento dados anteriores e posteriores ao valor a ser predito, identificando mais precisamente o comportamento das execuções das semanas/dias em torno da semana/dia a ser analisada, para se fazer a predição.

Apesar do fato de os preditores baseados em preenchimento de valores faltantes terem sido os mais acurados, decidiu-se por realizar as análises também usando o melhor preditor baseado em regressão testado. Os preditores baseados em métodos de regressão usam apenas dados do passado. Dessa forma, a decisão foi tomada porque um possível impacto no rádio a partir da presença de artistas na TV poderia se estender até datas futuras, das quais o preditor por preenchimento de valor faltante poderia estar influenciado por um aumento em razão da própria participação do artista na TV. De fato, o estudo foi realizado para vários métodos de predição, no entanto, apenas o melhor preditor de cada classe supracitada foi escolhido para realização de uma análise mais fina e para apresentação dos resultados nesta dissertação.

Foram verificadas as taxas de acerto dos preditores relativas às execuções nas rádios posteriores às participações dos artistas na TV, verificando assim se houve alguma al-

teração no padrão dessas execuções. Para observar as diferenças entre os números de execuções reais e os preditos, possibilitando tirar as conclusões sobre o estudo, foram calculadas as médias simples e ponderadas dessas taxas, além de serem aplicados testes de hipóteses, utilizando para isso os testes de Friedman e Wilcoxon Pareado.

Os resultados obtidos nas análises realizadas demonstram que não foi encontrada qualquer evidência de que a participação de um artista da música em um programa de TV afete de forma positiva nas execuções de suas músicas nas rádios. Não é possível afirmar, através desses resultados, que tal impacto definitivamente não exista. Porém, os indícios contra a hipótese do trabalho são bastante fortes, ainda que generalistas.

Portanto, uma contribuição importante do trabalho é apresentar dados que servem para o artista valorizar o impacto de sua participação em um programa de TV. Ainda que a análise realizada dê suporte apenas para avaliar o impacto em rádios, a contribuição é relevante não apenas pela importância do meio rádio nos dias atuais, mas também para abrir portas para análises similares em outros meios, como o meio digital.

Quanto às análises de influências de outros eventos (lançamentos, Copa do Mundo e telenovelas) sobre as execuções das músicas nas rádios, observaram-se impactos positivos nas execuções das músicas em datas de lançamentos oficiais e durante o período da Copa do Mundo, para as músicas utilizadas como temas desse evento. Já em relação ao fato das músicas serem temas de abertura de telenovelas, não foi possível constatar qualquer impacto nas execuções em decorrência disso.

Este trabalho faz uma contribuição para o melhor entendimento sobre as programações das rádios, realizando análises sobre dados de difícil acesso e gerando informações importantes para a indústria fonográfica e estudiosos, que até então são escassas devido à falta de trabalhos que façam investigações mais detalhadas sobre esses tipos de dados.

Como trabalhos futuros, almeja-se realizar uma investigação semelhante, utilizando como referência a metodologia de análise de influência entre mídias desenvolvida neste trabalho, conforme já anteriormente sugere-se, para verificar se a participação de artistas musicais em programas de TV causa um impacto nas mídias digitais, como YouTube, Facebook, Twitter, entre outras.

Ainda como indicação de trabalhos futuros, é possível aplicar a metodologia para a análise a partir de dados mais robustos, como por exemplo, tentando-se descobrir os lançamentos mais rapidamente, de forma a ser possível monitorar um número maior de artistas e músicas. Também é possível aumentar e tornar constante o número de

rádios em monitoramento, que no presente trabalho foi variável no decorrer do período analisado, uma vez que dependia-se de terceiros para realizar o processamento no sinal das rádios, ficando assim essa análise restrita às necessidades de monitoramento de rádios dentro do perfil comercial do serviço terceiro.

Considerando um cenário onde algumas milhares de rádios estão em constante monitoramento, pode ser bastante interessante analisar o impacto da ocorrência de outros eventos, podendo compreender inclusive análises geograficamente descentralizadas. Por exemplo, seria possível avaliar o impacto da realização de um espetáculo em uma dada cidade em comparação com outras cidades, ou mesmo o impacto no estado. Poderia-se também avaliar o fluxo de crescimento no número de execuções de um lançamento (ou mesmo de um espetáculo, ou qualquer ação de marketing do artista), podendo compreender se há impacto na rádio a partir de lançamentos locais. Ou ainda, avaliar se um lançamento a nível nacional gera mais impacto em uma região ou em outra, ou se acontece primeiramente em uma região e logo tende a mudar para outra região.

Apêndice A

Tabela de audiências dos programas monitorados

Este apêndice apresenta a tabela com os programas de TV monitorados durante o estudo e suas respectivas audiências. A Tabela A.1 apresenta o nome do programa (1ª coluna), o canal a qual ele pertence (2ª coluna), o dia da semana que ele vai ao ar (3ª coluna) e a audiência do programa (4ª coluna).

Tabela A.1: Audiência dos programas de TV monitorados

Programa	Canal	Dia da semana	Audiência
Altas Horas	Globo	Sábado	12
Caldeirão do Huck	Globo	Sábado	13
De Frente com Gaby	SBT	Domingo	6
Domingo Espetacular	Record	Domingo	10
Domingo Legal	SBT	Domingo	5
Domingo Show	Record	Domingo	8
Domingo da Gente	Record	Domingo	5
Domingão do Faustão	Globo	Domingo	13
Eliana	SBT	Domingo	5
Esquenta	Globo	Domingo	10

Estrelas	Globo	Sábado	12
Fantástico	Globo	Domingo	19
Frank Aguiar e Amigos	RedeTV	Sábado	1
Hora do Faro	Record	Domingo	7
Legendários	Record	Sábado	8
O Melhor do Brasil	Record	Domingo	5
Programa da Sabrina	Record	Sábado	6
Raul Gil	SBT	Sábado	5
Ritmo Brasil	RedeTV	Domingo	1
Sai do Chão	Globo	Domingo	13
Sábado Total	RedeTV	Sábado	1

Apêndice B

Tabelas de taxas de acerto dos preditores avaliados

Este apêndice exibe as tabelas de resultados de predições de execuções de músicas de artistas nas rádios, em períodos que não foram registradas ocorrências dos mesmos em programas de televisão. As Tabelas B.1 e B.2 apresentam a data do período utilizado e o nome do artista (1ª coluna, 1º e 2º itens respectivamente). Também apresenta-se número de rádios que executaram as músicas do artista, seguido do número de execuções de todas as músicas do artista (nº de rádios – nº de execuções), agrupando-se por semanas e por dias (2º coluna, 1º e 2º itens respectivamente). As demais colunas apresentam as taxas de variação dos valores reais em relação aos valores preditos, para execuções por semanas e por dias (1º e 2º linhas respectivamente), utilizando-se os métodos de preenchimento de valores faltantes e de regressão que foram avaliados (ver Seção 6.1). Ressalta-se que em alguns casos não existem resultados para análises por dias, por falta execuções suficientes para as análises.

O restante deste apêndice está organizado da seguinte forma. A Seção C.1 apresenta as tabelas com os resultados de predições de execuções, para os métodos de preenchimento de valores faltantes. Por sua vez, a Seção C.2 apresenta as tabelas com os resultados de predições de execuções, para os métodos de regressão.

B.1 Resultados para os métodos de preenchimento de valores faltantes

Nesta seção apresenta-se, na Tabela B.1, os resultados das taxas de acerto de predições para execuções por semanas e por dias, utilizando-se os seguintes métodos de preenchimento de valores faltantes analisados: Median, RAW, Euclidean (EC), Mahalanobis (MAL), MSN e RandomForest (RF).

Tabela B.1: Avaliação dos algoritmos de preenchimento de valores faltantes

Data/Artista	Info. Nums.	Median	MAL	RAW	EC	MSN	RF
25/01/2014	13 – 60	0,9836	1,0000	1,0000	1,0000	1,0000	1,0000
Skank	- – -	-	-	-	-	-	-
26/01/2014	16 – 191	0,9409	1,0269	1,0269	1,0269	1,0269	1,0269
Naldo	16 – 23	0,8214	1,0455	1,0455	1,0455	1,0455	1,0455
26/01/2014	78 – 1452	1,0097	0,8800	0,8800	0,8800	0,8800	1,1843
Paula Fernandes	78 – 213	0,9953	1,1833	1,1833	1,1833	1,1833	1,1833
16/02/2014	76 – 1967	1,0867	1,0282	1,0282	1,0282	1,0282	0,9718
Luan Santana	76 – 303	1,0860	1,0448	1,0448	1,0448	1,0448	1,0448
23/02/2014	84 – 1698	0,9982	1,0720	1,0720	1,0720	1,0720	1,0720
Victor e Leo	84 – 250	1,0684	0,9615	0,9615	0,9615	0,9615	0,9615
02/03/2014	26 – 207	1,0000	0,9039	0,9039	0,9039	0,9039	0,9039
Leonardo	26 – 25	0,8621	1,0417	1,0417	1,0417	1,0417	1,0417
02/03/2014	63 – 1382	1,0088	0,9705	0,9705	0,9705	0,9705	0,9705
Cristiano Araujo	63 – 183	0,8433	1,0457	1,0457	1,0457	1,0457	1,0457
09/03/2014	16 – 41	0,9762	1,0513	1,0513	1,0513	1,0513	1,0513
Fernando e Sorocaba	16 – 6	1,2000	1,0000	1,0000	1,0000	1,0000	1,0000
06/07/2014	6 – 92	1,0110	1,0000	1,0000	1,0000	1,0000	1,0000
Ivete Sangalo	6 – 14	1,1667	0,9333	0,9333	0,9333	0,9333	0,9333

06/07/2014	19 – 389	1,0628	1,0264	1,0264	1,0264	1,0264	1,0264
Victor e Leo	19 – 59	1,0536	1,0000	1,0000	1,0000	1,0000	1,0000
27/07/2014	16 – 198	1,0154	1,0000	1,0000	1,0000	1,0000	1,0000
Skank	16 – 25	0,8333	1,0000	1,0000	1,0000	1,0000	1,0000

B.2 Resultados para os métodos de regressão

Nesta seção são apresentadas, na Tabela B.2, os resultados das taxas de acerto de predições para execuções por semanas e por dias, utilizando-se os seguintes métodos de regressão analisados: Taxa Média de Crescimento/Queda de Execuções (TME), Taxa Média Ponderada de Crescimento/Queda de Execuções (TMPE), Regressão Linear (RL) e Random Forest (R. Forest).

Tabela B.2: Avaliação dos algoritmos de regressão

Data/Artista	Info. Nums.	TME	TMPE	RL	R. Forest
09/02/2014	28 – 455	0,8800	0,8900	0,9900	0,9200
So pra Contrariar	28 – 59	0,7400	0,7500	0,8400	0,8000
09/02/2014	2 – 9	1,3300	1,3000	1,5000	1,5000
Thaeme e Thiago	- – -	-	-	-	-
09/03/2014	91 – 1990	0,9200	0,9200	0,9200	0,9100
Anitta	91 – 304	1,0100	1,0200	0,9200	0,9800
09/03/2014	83 – 1597	0,8300	0,8400	1,0200	0,8900
Luan Santana	83 – 257	0,9000	0,9200	0,8400	0,8500
16/03/2014	88 – 1210	0,7200	0,7400	1,3100	0,8300
Victor e Leo	88 – 181	0,7700	0,7900	1,0300	0,8700
23/03/2014	26 – 153	0,7700	0,7800	0,8400	0,8400
Leonardo	26 – 24	0,9100	0,9100	0,9200	0,9600

23/03/2014	68 − 1354	0,9700	0,9800	1,0800	1,0300
Cristiano Araujo	68 − 207	1,0000	1,0100	0,9800	0,9800
30/03/2014	16 − 39	0,9800	0,9800	0,8700	0,9300
Fernando e Sorocaba	16 − 7	1,2700	1,2600	1,4000	1,4000
04/05/2014	31 − 153	0,5600	0,5800	1,3500	0,6900
Ivete Sangalo	31 − 21	0,4600	0,4800	0,9100	0,5000
13/07/2014	3 − 18	0,5800	0,5700	0,5600	0,5500
Gaby Amarantos	- − -	-	-	-	-
27/07/2014	36 − 497	1,0400	1,0400	1,0300	1,0400
Ivete Sangalo	36 − 71	1,0100	1,0200	1,0100	0,9900
27/07/2014	49 − 702	0,7100	0,7100	0,7100	0,7000
Victor e Leo	49 − 94	0,6500	0,6400	0,6500	0,6400
17/08/2014	78 − 1056	0,9900	1,0000	0,9800	0,9700
Skank	78 − 154	0,9300	0,9300	1,0300	0,9700
14/09/2014	3 − 14	1,1900	1,1500	1,0800	1,0800
MC Guime	- − -	-	-	-	-
14/09/2014	3 − 14	1,1900	1,1500	1,0800	1,0800
MC Guime	- − -	-	-	-	-

Apêndice C

Tabelas de taxas de acerto dos preditores para semanas/dias em análise

Este apêndice exibe as tabelas de resultados de predições de execuções de músicas de artistas nas rádios, para todas as ocorrências de artistas em programas de TV. As Tabelas C.1 e C.2 apresentam a data de ocorrência do artista na TV (1ª coluna, 1º item), o nome do programa de TV (1ª coluna, 2º item), o nome do artista (1ª coluna, 3º item) e a música de trabalho do artista no período analisado (1ª coluna, 4º item), seguido do total de músicas monitoradas do artista (nome da música de trabalho – total de músicas). Também apresenta-se o número de rádios que executaram as músicas do artista, seguido do número de execuções de todas as músicas do artista (nº de rádios – nº de execuções), para todas as músicas e as músicas de trabalho, agrupando-se por semanas (2º coluna, a, 1º e 2º itens respectivamente)e por dias (2º coluna, 3º e 4º itens respectivamente). As demais colunas apresentam as taxas de variação dos valores reais em relação aos valores preditos, para análises de todas as músicas e de músicas de trabalho, para execuções por semanas (1º e 2º linhas) e por dias (3º e 4º linhas), utilizando-se os métodos de preenchimento de valores faltantes e de regressão. Ressalta-se que em alguns casos não existem resultados para análises por dias, por falta execuções suficientes para as análises.

O restante deste apêndice está organizado da seguinte forma. A Seção C.1 apresenta as tabelas com os resultados de predições de execuções, para os métodos de preenchimento de valores faltantes. Por sua vez, a Seção C.2 apresenta as tabelas com os

resultados de predições de execuções, para os métodos de regressão.

C.1 Resultados para os métodos de preenchimento de valores faltantes

Nesta seção são apresentadas, na Tabela C.1, os resultados das taxas de acerto de predições para execuções por semanas e por dias, utilizando-se os métodos de preenchimento de valores faltantes: Median, RAW, Euclidean (EC), Mahalanobis (MAL), MSN e RandomForest (RF).

Tabela C.1: Resultados dos algoritmos de preenchimento de valores faltantes

Prog./Artist.	Info. Nums.	Median	MAL	RAW	EC	MSN	RF
01/02/2014	57 − 1100	0,9857	0,9874	0,9874	0,9874	0,9874	0,9874
Altas Horas	57 − 993	1,0206	1,0020	1,0020	1,0020	1,0020	0,9970
Zezé di Camargo e Luciano	- − -	-	-	-	-	-	-
Teorias − 4	- − -	-	-	-	-	-	-
08/02/2014	17 − 248	1,0420	1,0206	1,0206	1,0206	1,0206	1,0206
Altas Horas	17 − 190	1,0270	1,0215	1,0215	1,0215	1,0215	1,0215
Daniel	- − -	-	-	-	-	-	-
Fale Um Pouco de Você − 3	- − -	-	-	-	-	-	-
22/02/2014	13 − 59	1,1132	1,0351	1,0351	1,0351	1,0351	1,0351
Altas Horas	13 − 25	1,2500	1,0870	1,0870	1,0870	1,0870	1,0870
Skank	- − -	-	-	-	-	-	-
Resposta − 7	- − -	-	-	-	-	-	-
08/03/2014	74 − 1771	1,0363	1,0178	1,0178	1,0178	1,0178	1,0178
Altas Horas	74 − 921	1,0823	1,0550	1,0550	1,0550	1,0550	1,0550
Paula Fernandes	- − -	-	-	-	-	-	-
Quem é − 6	- − -	-	-	-	-	-	-

08/03/2014	68 − 1301	1,0725	0,9566	0,9566	0,9566	0,9566	0,9566
Altas Horas	68 − 977	1,0748	0,9721	0,9721	0,9721	0,9721	0,9721
Bruno e Marrone	- − -	-	-	-	-	-	-
Você Me Vira a Cabeça − 4	- − -	-	-	-	-	-	-
29/03/2014	67 − 1498	1,0296	0,9993	0,9993	0,9993	0,9993	0,9993
Altas Horas	67 − 1065	1,0360	1,0019	1,0019	1,0019	1,0019	1,0019
Gusttavo Lima	- − -	-	-	-	-	-	-
Fui Fiel − 7	- − -	-	-	-	-	-	-
19/04/2014	74 − 1657	1,0441	1,0184	1,0184	1,0184	1,0184	0,9589
Altas Horas	74 − 1051	1,0616	0,9786	0,9786	0,9786	0,9786	0,9786
Paula Fernandes	- − -	-	-	-	-	-	-
Quem é − 6	- − -	-	-	-	-	-	-
19/04/2014	35 − 174	0,8325	1,0610	1,0610	1,0610	1,0610	1,0610
Altas Horas	35 − 119	0,8151	0,9520	0,9520	0,9520	0,9520	0,9520
João Neto e Frederico	- − -	-	-	-	-	-	-
Crime Perfeito − 4	- − -	-	-	-	-	-	-
19/04/2014	13 − 92	0,9293	1,1084	1,1084	1,1084	1,1084	1,1084
Altas Horas	13 − 84	0,9333	1,1507	1,1507	1,1507	1,1507	1,1507
César Menotti e Fabiano	- − -	-	-	-	-	-	-
Não Era Eu − 2	- − -	-	-	-	-	-	-
19/04/2014	14 − 250	1,0163	0,9506	0,9506	0,9506	0,9506	0,9506
Altas Horas	14 − 223	1,0276	0,9449	0,9449	0,9449	0,9449	0,9449
Fernando e Sorocaba	- − -	-	-	-	-	-	-
Gaveta − 2	- − -	-	-	-	-	-	-
26/04/2014	3 − 47	1,0930	1,0682	1,0682	1,0682	1,0682	0,8704
Altas Horas	3 − 47	1,0930	1,0682	1,0682	1,0682	1,0682	0,8704
Pixote	- − -	-	-	-	-	-	-
Uma Simples Carta − 1	- − -	-	-	-	-	-	-

26/04/2014	15 – 183	0,9337	0,9786	0,9786	0,9786	0,9786	0,9786
Altas Horas	15 – 166	0,9595	0,9651	0,9651	0,9651	0,9651	1,0573
Só pra Contrariar	- – -	-	-	-	-	-	-
Me Perdoa - Live – 2	- – -	-	-	-	-	-	-
10/05/2014	1 – 2	0,6667	1,0000	1,0000	1,0000	1,0000	1,0000
Altas Horas	1 – 2	0,6667	1,0000	1,0000	1,0000	1,0000	1,0000
Harmonia do Samba	- – -	-	-	-	-	-	-
Nova Paradinha – 2	- – -	-	-	-	-	-	-
09/08/2014	2 – 15	1,2500	1,0714	1,0714	1,0714	1,0714	1,0714
Altas Horas	2 – 15	1,2500	1,0714	1,0714	1,0714	1,0714	1,0714
Thaeme e Thiago	- – -	-	-	-	-	-	-
Cafajeste – 3	- – -	-	-	-	-	-	-
09/08/2014	2 – 11	0,7857	1,0000	1,0000	1,0000	1,0000	1,0000
Altas Horas	2 – 10	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Gaby Amarantos	- – -	-	-	-	-	-	-
Gaby Ostentação – 4	- – -	-	-	-	-	-	-
30/08/2014	19 – 257	1,0844	0,9590	0,9590	0,9590	0,9590	0,9590
Altas Horas	19 – 257	1,0844	0,9590	0,9590	0,9590	0,9590	0,9590
João Neto e Frederico	- – -	-	-	-	-	-	-
Ele Nao Vai Mudar – 4	- – -	-	-	-	-	-	-
06/09/2014	17 – 102	0,8160	0,9189	0,9189	0,9189	0,9189	0,9189
Altas Horas	17 – 102	0,8160	0,9189	0,9189	0,9189	0,9189	0,9189
Victor e Léo	- – -	-	-	-	-	-	-
O Tempo Não Apaga – 7	- – -	-	-	-	-	-	-
13/09/2014	45 – 541	1,0227	1,0056	1,0056	1,0056	1,0056	0,9890
Altas Horas	45 – 541	1,0227	1,0056	1,0056	1,0056	1,0056	0,9890
Zezé di Camargo e Luciano	- – -	-	-	-	-	-	-
Flores em Vida – 4	- – -	-	-	-	-	-	-

13/09/2014	12 − 128	0,9922	1,0667	1,0667	1,0667	1,0667	1,0667
Altas Horas	12 − 128	0,9922	1,0667	1,0667	1,0667	1,0667	1,0667
Gusttavo Lima	- − -	-	-	-	-	-	-
10 Anos − 7	- − -	-	-	-	-	-	-
27/09/2014	16 − 144	1,0141	1,0070	1,0070	1,0070	1,0070	1,0070
Altas Horas	16 − 144	1,0141	1,0070	1,0070	1,0070	1,0070	1,0070
Skank	- − -	-	-	-	-	-	-
Ela me deixou − 7	- − -	-	-	-	-	-	-
04/10/2014	2 − 16	1,0667	1,0000	1,0000	1,0000	1,0000	1,0000
Altas Horas	2 − 16	1,0667	1,0000	1,0000	1,0000	1,0000	1,0000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Cafajeste − 3	- − -	-	-	-	-	-	-
11/10/2014	5 − 79	1,0000	0,9518	0,9518	0,9518	0,9518	1,0533
Altas Horas	5 − 65	0,9848	1,0484	1,0484	1,0484	1,0484	1,0484
Michel Teló	- − -	-	-	-	-	-	-
Implorando pra trair − 5	- − -	-	-	-	-	-	-
11/10/2014	52 − 783	1,0051	0,9788	0,9788	0,9788	0,9788	0,9788
Altas Horas	52 − 783	1,0051	0,9788	0,9788	0,9788	0,9788	0,9788
Marcos e Belutti	- − -	-	-	-	-	-	-
Domingo de Manhã − 3	- − -	-	-	-	-	-	-
15/03/2014	74 − 1605	0,9391	0,9926	0,9926	0,9926	0,9926	0,9926
Caldeirão do Huck	74 − 873	0,9113	0,9479	0,9479	0,9479	0,9479	1,1009
Paula Fernandes	- − -	-	-	-	-	-	-
Quem é − 6	- − -	-	-	-	-	-	-
19/04/2014	1 − 3	1,5000	1,5000	1,5000	1,5000	1,5000	1,5000
Caldeirão do Huck	1 − 3	1,5000	1,5000	1,5000	1,5000	1,5000	1,5000
Seu Jorge	- − -	-	-	-	-	-	-
Amiga da Minha Mulher − 1	- − -	-	-	-	-	-	-

19/07/2014	5 – 18	1,2857	1,0588	1,0588	1,0588	1,0588	1,0588
Caldeirão do Huck	5 – 18	1,2857	1,0588	1,0588	1,0588	1,0588	1,0588
Michel Teló	- – -	-	-	-	-	-	-
Amiga Da Minha Irmã – 5	- – -	-	-	-	-	-	-
02/08/2014	45 – 602	1,0750	1,0203	1,0203	1,0203	1,0203	0,9465
Caldeirão do Huck	45 – 602	1,0750	1,0203	1,0203	1,0203	1,0203	0,9465
Zezé di Camargo e Luciano	- – -	-	-	-	-	-	-
Flores em Vida – 4	- – -	-	-	-	-	-	-
20/09/2014	5 – 10	0,7143	1,2500	1,2500	1,2500	1,2500	1,2500
Caldeirão do Huck	5 – 10	0,7143	1,2500	1,2500	1,2500	1,2500	1,2500
Michel Teló	- – -	-	-	-	-	-	-
Amiga Da Minha Irmã – 5	- – -	-	-	-	-	-	-
09/03/2014	67 – 1499	0,9603	1,0007	1,0007	1,0007	1,0007	1,0007
De Frente com Gaby	67 – 1063	0,9824	0,9981	0,9981	0,9981	0,9981	0,9981
Gusttavo Lima	76 – 240	0,9375	1,0345	1,0345	1,0345	1,0345	1,0345
Fui Fiel – 7	76 – 187	1,0108	0,9740	0,9740	0,9740	0,9740	0,9740
03/08/2014	45 – 602	1,0750	1,0203	1,0203	1,0203	1,0203	0,9465
De Frente com Gaby	45 – 602	1,0750	1,0203	1,0203	1,0203	1,0203	0,9465
Zezé di Camargo e Luciano	42 – 102	1,2439	1,0737	1,0737	1,0737	1,0737	1,0737
Flores em Vida – 4	42 – 102	1,2439	1,0737	1,0737	1,0737	1,0737	1,0737
23/03/2014	65 – 1244	0,9540	1,0970	1,0970	1,0970	1,0970	1,0970
Domingo Espetacular	65 – 1056	0,9724	1,1210	1,1210	1,1210	1,1210	1,1210
Anitta	81 – 260	1,1556	1,0078	1,0078	1,0078	1,0078	1,0078
Zen – 5	81 – 230	1,3068	1,0132	1,0132	1,0132	1,0132	0,9623
30/03/2014	74 – 1728	1,0183	0,9942	0,9942	0,9942	0,9942	0,9942
Domingo Espetacular	74 – 1083	1,0607	1,0084	1,0084	1,0084	1,0084	1,0084
Paula Fernandes	78 – 297	1,0722	1,0206	1,0206	1,0206	1,0206	1,0206
Quem é – 6	78 – 187	1,1615	1,0054	1,0054	1,0054	1,0054	1,0054

16/02/2014	15 − 281	1,0331	1,0036	1,0036	1,0036	1,0036	1,0036
Domingo Legal	15 − 263	1,0354	1,0000	1,0000	1,0000	1,0000	1,0000
Só pra Contrariar	25 − 63	1,1053	1,0000	1,0000	1,0000	1,0000	1,0000
Me Perdoa - Live − 2	25 − 60	1,1321	1,0000	1,0000	1,0000	1,0000	1,0000
13/07/2014	23 − 309	0,9450	1,0546	1,0546	1,0546	1,0546	1,0546
Domingo Legal	23 − 309	0,9450	1,0546	1,0546	1,0546	1,0546	1,0546
Lucas Lucco	25 − 51	1,1087	0,9808	0,9808	0,9808	0,9808	1,0200
Mozão − 1	25 − 51	1,1087	0,9808	0,9808	0,9808	0,9808	1,0200
06/04/2014	14 − 299	1,0993	1,0275	1,0275	1,0275	1,0275	0,9583
Domingo Show	14 − 266	1,0947	1,0076	1,0076	1,0076	1,0076	0,9708
Fernando e Sorocaba	16 − 57	1,2391	1,0556	1,0556	1,0556	1,0556	0,9194
Gaveta − 2	16 − 48	1,2000	1,0000	1,0000	1,0000	1,0000	1,0000
27/04/2014	67 − 1131	0,9577	1,0162	1,0162	1,0162	1,0162	1,0162
Domingo Show	67 − 755	0,9414	1,0371	1,0371	1,0371	1,0371	1,0371
Gusttavo Lima	76 − 190	0,9694	1,0270	1,0270	1,0270	1,0270	1,0270
Fui Fiel − 7	76 − 140	1,0219	0,9150	0,9150	0,9150	0,9150	0,9150
20/07/2014	2 − 14	1,2727	1,0769	0,9333	1,0769	1,0769	0,9333
Domingo Show	2 − 10	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Gaby Amarantos	- − -	-	-	-	-	-	-
Gaby Ostentação − 4	- − -	-	-	-	-	-	-
07/09/2014	19 − 233	0,9510	0,9957	0,9957	0,9957	0,9957	0,9957
Domingo Show	19 − 233	0,9510	0,9957	0,9957	0,9957	0,9957	0,9957
João Neto e Frederico	21 − 36	0,9474	1,0588	0,9474	1,0588	1,0588	0,9474
Ele Nao Vai Mudar − 4	21 − 36	0,9474	1,0588	0,9474	1,0588	1,0588	0,9474
23/02/2014	15 − 170	1,0366	1,0000	1,0000	1,0000	1,0000	1,0000
Domingo da Gente	15 − 72	1,1613	1,0141	1,0141	1,0141	1,0141	1,0141
Naldo	16 − 22	0,8148	1,0000	1,0000	1,0000	1,0000	1,0000
Caipifruta − 3	16 − 9	0,9000	1,0000	1,0000	1,0000	1,0000	1,0000

02/02/2014	26 − 358	0,9862	1,0000	1,0000	1,0000	1,0000	1,0000
Domingão do Faustão	26 − 358	0,9862	1,0000	1,0000	1,0000	1,0000	1,0000
Ivete Sangalo	27 − 58	1,0741	0,9831	0,9831	0,9831	0,9831	0,9831
Tempo de Alegria − 2	27 − 58	1,0741	0,9831	0,9831	0,9831	0,9831	0,9831
23/02/2014	74 − 1691	0,9713	0,9883	0,9883	0,9883	0,9883	0,9883
Domingão do Faustão	74 − 793	0,9814	1,0089	1,0089	1,0089	1,0089	1,0089
Paula Fernandes	78 − 305	1,1213	0,9839	0,9839	0,9839	0,9839	0,9839
Quem é − 6	78 − 147	1,1484	1,0068	1,0068	1,0068	1,0068	0,9363
16/03/2014	74 − 1605	0,9391	0,9926	0,9926	0,9926	0,9926	0,9926
Domingão do Faustão	74 − 873	0,9113	0,9479	0,9479	0,9479	0,9479	1,1009
Paula Fernandes	78 − 268	0,9371	0,9889	0,9889	0,9889	0,9889	0,9889
Quem é − 6	78 − 144	0,9000	0,9796	0,9796	0,9796	0,9796	0,9796
16/03/2014	26 − 303	0,9381	0,9681	0,9681	0,9681	0,9681	0,9681
Domingão do Faustão	26 − 303	0,9381	0,9681	0,9681	0,9681	0,9681	0,9681
Ivete Sangalo	27 − 55	1,1000	1,0185	1,0185	1,0185	1,0185	1,0185
Tempo de Alegria − 2	27 − 55	1,1000	1,0185	1,0185	1,0185	1,0185	1,0185
16/03/2014	65 − 1474	1,0265	0,9048	0,9048	0,9048	0,9048	0,9048
Domingão do Faustão	65 − 1231	1,0284	0,9200	0,9200	0,9200	0,9200	0,9200
Anitta	81 − 258	1,0118	0,9923	0,9923	0,9923	0,9923	0,9923
Zen − 5	81 − 227	1,1019	0,9870	0,9870	0,9870	0,9870	0,9870
16/03/2014	17 − 223	1,1675	1,0878	1,0878	1,0878	1,0878	1,0878
Domingão do Faustão	17 − 90	1,2162	1,0843	1,0843	1,0843	1,0843	1,0843
Thiaguinho	21 − 51	1,5000	1,2439	1,2439	1,2439	1,2439	1,2439
As Aparências Enganam − 4	21 − 19	1,5833	1,2667	1,2667	1,2667	1,2667	1,2667
16/03/2014	66 − 1174	0,9816	1,0731	1,0731	1,0731	1,0731	1,0731
Domingão do Faustão	66 − 883	0,9714	1,0888	1,0888	1,0888	1,0888	1,0888
Luan Santana	76 − 198	1,0154	0,9851	0,9851	0,9851	0,9851	0,9851
Tudo Que Você Quiser-Live − 5	76 − 151	1,0486	1,0272	1,0272	1,0272	1,0272	1,0272

16/03/2014	17 − 196	0,9655	0,9899	0,9899	0,9899	0,9899	0,9899
Domingão do Faustão	17 − 157	1,0261	1,0000	1,0000	1,0000	1,0000	1,0000
Daniel	23 − 34	0,9714	0,9714	0,9714	0,9714	0,9714	0,9714
Fale Um Pouco de Você − 3	23 − 25	0,9259	1,0000	1,0000	1,0000	1,0000	1,0000
23/03/2014	15 − 212	0,9258	1,0291	1,0291	1,0291	1,0291	1,0291
Domingão do Faustão	15 − 194	0,9238	1,0211	1,0211	1,0211	1,0211	1,0211
Só pra Contrariar	25 − 50	0,9804	1,0204	1,0204	1,0204	1,0204	1,0204
Me Perdoa - Live − 2	25 − 47	1,0000	1,0444	0,9592	1,0444	1,0444	0,9592
30/03/2014	20 − 135	1,0305	0,9926	0,9926	0,9926	0,9926	0,9926
Domingão do Faustão	20 − 75	0,9259	1,0000	1,0000	1,0000	1,0000	1,0000
Leonardo	26 − 21	0,8400	0,9130	0,9130	0,9130	0,9130	0,9130
Choro − 4	26 − 14	0,8750	1,0000	1,0000	1,0000	1,0000	1,0000
20/04/2014	65 − 929	0,8848	0,9431	0,9431	0,9431	0,9431	0,9431
Domingão do Faustão	65 − 735	0,9339	0,9633	0,9633	0,9633	0,9633	0,9633
Anitta	81 − 176	0,9670	1,1812	1,1812	1,1812	1,1812	1,1812
Zen − 5	81 − 140	1,1382	0,9524	0,9524	0,9524	0,9524	0,9524
04/05/2014	17 − 138	0,9020	0,9583	0,9583	0,9583	0,9583	1,0781
Domingão do Faustão	17 − 53	1,1277	0,9815	0,9815	0,9815	0,9815	0,9815
Thiaguinho	21 − 19	0,9500	1,0000	1,0000	1,0000	1,0000	1,0000
As Aparências Enganam − 4	21 − 11	1,5714	1,0000	1,0000	1,0000	1,0000	1,0000
11/05/2014	26 − 370	0,9814	0,9973	0,9973	0,9973	0,9973	1,0165
Domingão do Faustão	26 − 251	0,9544	1,0162	1,0162	1,0162	1,0162	1,0162
Ivete Sangalo	27 − 59	0,9833	1,0000	1,0000	1,0000	1,0000	1,0000
Amor que Não Sai − 2	27 − 38	0,8444	0,9500	0,9500	0,9500	0,9500	0,9500
27/07/2014	23 − 293	1,0616	0,9482	0,9482	0,9482	0,9482	0,9482
Domingão do Faustão	23 − 293	1,0616	0,9482	0,9482	0,9482	0,9482	0,9482
Lucas Lucco	25 − 40	0,9524	1,0526	1,0526	1,0526	1,0526	1,0526
Mozão − 1	25 − 40	0,9524	1,0526	1,0526	1,0526	1,0526	1,0526

24/08/2014	52 − 960	1,1176	1,0042	1,0042	1,0042	1,0042	1,0042
Domingão do Faustão	52 − 960	1,1176	1,0042	1,0042	1,0042	1,0042	1,0042
Marcos e Belutti	49 − 141	1,1463	0,9592	0,9592	0,9592	0,9592	0,9592
Domingo de Manhã − 3	49 − 141	1,1463	0,9592	0,9592	0,9592	0,9592	0,9592
05/10/2014	10 − 70	1,0000	0,8861	1,1475	0,8861	0,8861	0,8861
Domingão do Faustão	10 − 70	1,0000	0,8861	1,1475	0,8861	0,8861	0,8861
Ivete Sangalo	6 − 11	1,3750	1,0000	1,0000	1,0000	1,0000	1,0000
Amor que Não Sai − 2	6 − 11	1,3750	1,0000	1,0000	1,0000	1,0000	1,0000
09/03/2014	38 − 364	0,9239	1,0551	1,0551	1,0551	1,0551	1,0551
Eliana	38 − 230	0,9388	1,0952	1,0952	1,0952	1,0952	1,0952
Eduardo Costa	48 − 76	1,1692	0,9500	0,9500	0,9500	0,9500	0,9500
Enamorado − 4	48 − 44	1,2222	1,0476	1,0476	1,0476	1,0476	1,0476
20/04/2014	14 − 250	1,0163	0,9506	0,9506	0,9506	0,9506	0,9506
Eliana	14 − 223	1,0276	0,9449	0,9449	0,9449	0,9449	0,9449
Fernando e Sorocaba	16 − 40	0,9524	1,0256	1,0256	1,0256	1,0256	1,0256
Gaveta − 2	16 − 34	0,9444	1,0625	1,0625	1,0625	1,0625	1,0625
20/04/2014	10 − 66	0,8250	0,9167	0,9167	0,9167	0,9167	0,9167
Esquenta	10 − 42	0,7636	0,8235	0,8235	0,8235	0,8235	0,8235
Grupo Bom Gosto	17 − 16	1,4545	1,0000	1,0000	1,0000	1,0000	1,0000
Mo Love-Live − 2	17 − 10	1,6667	0,9091	0,9091	0,9091	0,9091	0,9091
13/07/2014	1 − 2	0,6667	1,0000	1,0000	1,0000	1,0000	1,0000
Esquenta	1 − 2	0,6667	1,0000	1,0000	1,0000	1,0000	1,0000
Psirico	- − -	-	-	-	-	-	-
Lepo Lepo − 1	- − -	-	-	-	-	-	-
03/08/2014	23 − 244	1,0041	0,8328	0,8328	0,8328	0,8328	1,2642
Esquenta	23 − 244	1,0041	0,8328	0,8328	0,8328	0,8328	1,2642
Lucas Lucco	25 − 38	1,0000	0,9500	0,9500	0,9500	0,9500	1,0556
Mozão − 1	25 − 38	1,0000	0,9500	0,9500	0,9500	0,9500	1,0556

03/08/2014	10 − 137	1,1913	1,0379	1,0379	1,0379	1,0379	1,0379
Esquenta	10 − 137	1,1913	1,0379	1,0379	1,0379	1,0379	1,0379
Ivete Sangalo	6 − 16	1,3333	1,2308	1,2308	1,2308	1,2308	1,2308
Amor que Não Sai − 2	6 − 16	1,3333	1,2308	1,2308	1,2308	1,2308	1,2308
10/08/2014	52 − 774	0,8079	0,9160	0,9160	0,9160	0,9160	0,9160
Esquenta	52 − 774	0,8079	0,9160	0,9160	0,9160	0,9160	0,9160
Marcos e Belutti	49 − 134	1,0075	0,9504	0,9504	0,9504	0,9504	1,0635
Domingo de Manhã − 3	49 − 134	1,0075	0,9504	0,9504	0,9504	0,9504	1,0635
24/08/2014	16 − 168	1,0435	1,0244	1,0244	1,0244	1,0244	1,0244
Esquenta	16 − 168	1,0435	1,0244	1,0244	1,0244	1,0244	1,0244
Skank	16 − 24	0,9600	0,9600	0,9600	0,9600	0,9600	0,9600
Ela me deixou − 7	16 − 24	0,9600	0,9600	0,9600	0,9600	0,9600	0,9600
08/02/2014	15 − 170	0,9551	1,0000	1,0000	1,0000	1,0000	1,0000
Estrelas	15 − 82	1,1389	1,0250	1,0250	1,0250	1,0250	1,0250
Naldo	- − -	-	-	-	-	-	-
Se Joga − 3	- − -	-	-	-	-	-	-
01/03/2014	12 − 237	1,1675	1,0000	1,0000	1,0000	1,0000	1,0000
Estrelas	12 − 167	1,2463	0,9940	0,9940	0,9940	0,9940	0,9940
Péricles	- − -	-	-	-	-	-	-
Se Eu Largar o Freio − 3	- − -	-	-	-	-	-	-
12/07/2014	5 − 16	1,1429	1,0000	1,0000	1,0000	1,0000	1,0000
Estrelas	5 − 16	1,1429	1,0000	1,0000	1,0000	1,0000	1,0000
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
12/07/2014	1 − 36	1,2000	1,0000	1,0000	1,0000	1,0000	1,0000
Estrelas	1 − 36	1,2000	1,0000	1,0000	1,0000	1,0000	1,0000
Valesca Popozuda	- − -	-	-	-	-	-	-
Beijinho no Ombro − 1	- − -	-	-	-	-	-	-

09/08/2014	19 − 241	0,8368	0,9377	0,9377	0,9377	0,9377	0,9377
Estrelas	19 − 241	0,8368	0,9377	0,9377	0,9377	0,9377	0,9377
João Neto e Frederico	- − -	-	-	-	-	-	-
Ele Nao Vai Mudar − 4	- − -	-	-	-	-	-	-
02/02/2014	7 − 81	0,9878	1,0946	1,0946	1,0946	1,0946	1,0946
Fantástico	7 − 70	1,0145	0,9211	0,9211	0,9211	0,9211	1,1111
Ana Carolina	10 − 15	0,8824	1,0714	1,0714	1,0714	1,0714	1,0714
Combustivel − 2	10 − 14	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
23/03/2014	82 − 1031	0,9416	1,0019	1,0019	1,0019	1,0019	1,0019
Fantástico	82 − 863	0,9578	1,0262	1,0262	1,0262	1,0262	1,0262
Victor e Léo	84 − 160	0,9639	1,0323	1,0323	1,0323	1,0323	1,0323
Na Linha do Tempo − 7	84 − 130	0,8966	1,0236	1,0236	1,0236	1,0236	1,0236
23/03/2014	65 − 1244	0,9540	1,0970	1,0970	1,0970	1,0970	1,0970
Fantástico	65 − 1056	0,9724	1,1210	1,1210	1,1210	1,1210	1,1210
Anitta	81 − 260	1,1556	1,0078	1,0078	1,0078	1,0078	1,0078
Zen − 5	81 − 230	1,3068	1,0132	1,0132	1,0132	1,0132	0,9623
30/03/2014	66 − 1094	1,0684	1,0186	1,0186	1,0186	1,0186	1,0186
Fantástico	66 − 795	1,0338	0,9803	0,9803	0,9803	0,9803	0,9803
Luan Santana	76 − 190	1,0674	0,9596	0,9596	0,9596	0,9596	0,9596
Tudo Que Você Quiser-Live − 5	76 − 141	1,0292	1,0217	1,0217	1,0217	1,0217	0,9338
13/04/2014	15 − 228	1,0909	1,0088	1,0088	1,0088	1,0088	1,0088
Fantástico	15 − 213	1,1576	0,9953	0,9953	0,9953	0,9953	0,9953
Só pra Contrariar	25 − 54	1,2857	0,9643	0,9643	0,9643	0,9643	0,9643
Me Perdoa - Live − 2	25 − 49	1,2895	0,9608	0,9608	0,9608	0,9608	0,9608
20/07/2014	5 − 18	1,2857	1,0588	1,0588	1,0588	1,0588	1,0588
Fantástico	5 − 18	1,2857	1,0588	1,0588	1,0588	1,0588	1,0588
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-

27/07/2014	5 − 17	1,2143	0,9444	0,9444	0,9444	0,9444	0,9444
Fantástico	5 − 17	1,2143	0,9444	0,9444	0,9444	0,9444	0,9444
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
03/08/2014	5 − 13	0,8667	1,0833	0,9286	1,0833	1,0833	0,9286
Fantástico	5 − 13	0,8667	1,0833	0,9286	1,0833	1,0833	0,9286
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
10/08/2014	5 − 11	0,7333	0,9167	0,9167	0,9167	0,9167	0,9167
Fantástico	5 − 11	0,7333	0,9167	0,9167	0,9167	0,9167	0,9167
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
17/08/2014	5 − 12	0,8000	1,0909	1,0909	1,0909	0,9231	1,0909
Fantástico	5 − 12	0,8000	1,0909	1,0909	1,0909	0,9231	1,0909
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
24/08/2014	5 − 14	1,0769	1,0000	1,0000	1,0000	1,0000	1,0000
Fantástico	5 − 14	1,0769	1,0000	1,0000	1,0000	1,0000	1,0000
Michel Teló	- − -	-	-	-	-	-	-
Amiga Da Minha Irmã − 5	- − -	-	-	-	-	-	-
12/10/2014	5 − 79	1,0000	0,9518	0,9518	0,9518	0,9518	1,0533
Fantástico	5 − 65	0,9848	1,0484	1,0484	1,0484	1,0484	1,0484
Michel Teló	- − -	-	-	-	-	-	-
Implorando pra trair − 5	- − -	-	-	-	-	-	-
12/10/2014	45 − 515	1,0300	1,0239	1,0239	1,0239	1,0239	1,0239
Fantástico	45 − 515	1,0300	1,0239	1,0239	1,0239	1,0239	1,0239
Zezé di Camargo e Luciano	42 − 79	1,1618	0,9875	0,9875	0,9875	0,9875	0,9875
Flores em Vida − 4	42 − 79	1,1618	0,9875	0,9875	0,9875	0,9875	0,9875

19/10/2014	17 − 114	1,2000	1,0179	1,0179	1,0179	1,0179	1,0179
Fantástico	17 − 114	1,2000	1,0179	1,0179	1,0179	1,0179	1,0179
Victor e Léo	19 − 19	1,5833	1,3571	1,3571	1,3571	1,3571	1,3571
O Tempo Não Apaga − 7	19 − 19	1,5833	1,3571	1,3571	1,3571	1,3571	1,3571
19/10/2014	5 − 83	0,9540	1,0506	1,0506	1,0506	1,0506	1,0506
Fantástico	5 − 71	0,9342	1,0923	1,0923	1,0923	1,0923	1,0923
Michel Teló	4 − 10	0,8333	1,0000	1,0000	1,0000	1,0000	1,0000
Implorando pra trair − 5	4 − 10	0,8333	1,0000	1,0000	1,0000	1,0000	1,0000
19/10/2014	6 − 16	0,8421	1,0667	1,0667	0,9412	0,9412	0,9412
Fantástico	6 − 16	0,8421	1,0667	1,0667	0,9412	0,9412	0,9412
Fernando e Sorocaba	- − -	-	-	-	-	-	-
Gaveta − 2	- − -	-	-	-	-	-	-
11/05/2014	5 − 14	0,9333	0,9333	0,9333	0,9333	0,9333	0,9333
Frank Aguiar e Amigos	5 − 14	0,9333	0,9333	0,9333	0,9333	0,9333	0,9333
Rio Negro e Solimões	- − -	-	-	-	-	-	-
Romântica − 1	- − -	-	-	-	-	-	-
27/04/2014	22 − 585	1,0227	0,9983	0,9983	0,9983	0,9983	1,0104
Hora do Faro	22 − 528	1,0252	1,0154	1,0154	1,0154	1,0154	1,0154
Marcos e Belutti	22 − 99	1,0761	1,0532	1,0532	1,0532	1,0532	0,9429
Domingo de Manhã − 3	22 − 92	1,1795	0,9892	0,9892	0,9892	0,9892	0,9892
03/08/2014	17 − 191	1,0324	0,8377	0,8377	0,8377	0,8377	0,8377
Hora do Faro	17 − 191	1,0324	0,8377	0,8377	0,8377	0,8377	0,8377
Victor e Léo	19 − 38	1,2258	1,0270	1,0270	1,0270	1,0270	1,0270
O Tempo Não Apaga − 7	19 − 38	1,2258	1,0270	1,0270	1,0270	1,0270	1,0270
28/09/2014	45 − 547	1,0399	1,0111	1,0111	1,0111	1,0111	0,9785
Hora do Faro	45 − 547	1,0399	1,0111	1,0111	1,0111	1,0111	0,9785
Zezé di Camargo e Luciano	42 − 74	0,9367	0,9737	0,9737	0,9737	0,9737	0,9737
Flores em Vida − 4	42 − 74	0,9367	0,9737	0,9737	0,9737	0,9737	0,9737

01/02/2014	74 − 1126	0,9030	0,9895	0,9895	0,9895	0,9895	0,9895
Legendários	74 − 655	0,9576	1,0599	1,0599	1,0599	1,0599	1,0599
Paula Fernandes	- − -	-	-	-	-	-	-
Não Fui Eu − 6	- − -	-	-	-	-	-	-
08/02/2014	2 − 9	1,2857	1,0000	1,0000	1,0000	1,0000	1,0000
Legendários	2 − 9	1,2857	1,0000	1,0000	1,0000	1,0000	1,0000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Sinto Saudade − 3	- − -	-	-	-	-	-	-
08/02/2014	17 − 237	1,0216	1,0000	1,0000	1,0000	1,0000	1,0000
Legendários	17 − 115	0,9829	0,9914	0,9914	0,9914	0,9914	0,9914
Thiaguinho	- − -	-	-	-	-	-	-
As Aparências Enganam − 4	- − -	-	-	-	-	-	-
15/02/2014	13 − 184	0,9840	1,0000	1,0000	1,0000	1,0000	1,0000
Legendários	13 − 170	0,9770	0,9884	0,9884	0,9884	0,9884	0,9884
César Menotti e Fabiano	- − -	-	-	-	-	-	-
Não Era Eu − 2	- − -	-	-	-	-	-	-
01/03/2014	35 − 347	1,2175	1,0176	1,0176	1,0176	1,0176	1,0176
Legendários	35 − 212	1,1042	1,0291	1,0291	1,0291	1,0291	0,9258
João Neto e Frederico	- − -	-	-	-	-	-	-
Crime Perfeito − 4	- − -	-	-	-	-	-	-
22/03/2014	38 − 340	1,0241	0,9855	0,9855	0,9855	0,9855	0,9855
Legendários	38 − 193	0,9602	1,0052	1,0052	1,0052	1,0052	1,0052
Eduardo Costa	- − -	-	-	-	-	-	-
Enamorado − 4	- − -	-	-	-	-	-	-
12/04/2014	63 − 1288	1,0455	0,9772	0,9772	0,9772	0,9772	0,9772
Legendários	63 − 1033	1,0477	1,0088	1,0088	1,0088	1,0088	0,9876
Cristiano Araújo	- − -	-	-	-	-	-	-
Maus Bocados − 3	- − -	-	-	-	-	-	-

19/04/2014	18 − 138	0,9079	0,9857	0,9857	0,9857	0,9857	1,1405
Legendários	18 − 92	0,8364	1,0222	1,0222	1,0222	1,0222	1,0222
Munhoz e Mariano	- − -	-	-	-	-	-	-
Pantera Cor de Rosa − 3	- − -	-	-	-	-	-	-
19/04/2014	22 − 621	1,0991	1,0615	1,0615	1,0615	1,0615	1,0615
Legendários	22 − 543	1,0838	1,0284	1,0284	1,0284	1,0284	1,0284
Marcos e Belutti	- − -	-	-	-	-	-	-
Domingo de Manhã − 3	- − -	-	-	-	-	-	-
26/04/2014	5 − 40	1,1429	1,0256	1,0256	1,0256	1,0256	1,0256
Legendários	5 − 22	0,9167	0,9565	0,9565	0,9565	0,9565	0,9565
Onze 20	- − -	-	-	-	-	-	-
Não Vai Voltar − 2	- − -	-	-	-	-	-	-
03/05/2014	28 − 220	0,9955	1,0427	1,0427	1,0427	1,0427	0,9483
Legendários	28 − 101	1,1222	0,9712	0,9712	0,9712	0,9712	0,9712
Turma do Pagode	- − -	-	-	-	-	-	-
Horário de Verão − 4	- − -	-	-	-	-	-	-
03/05/2014	5 − 15	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Legendários	5 − 15	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Rio Negro e Solimões	- − -	-	-	-	-	-	-
Romântica − 1	- − -	-	-	-	-	-	-
10/05/2014	65 − 833	0,9287	0,9619	0,9619	0,9619	0,9619	0,9619
Legendários	65 − 623	1,0349	0,9326	0,9326	0,9326	0,9326	0,9326
Anitta	- − -	-	-	-	-	-	-
Zen − 5	- − -	-	-	-	-	-	-
19/07/2014	45 − 663	1,1652	1,0425	1,0425	1,0425	1,0425	1,0425
Legendários	45 − 663	1,1652	1,0425	1,0425	1,0425	1,0425	1,0425
Zezé di Camargo e Luciano	- − -	-	-	-	-	-	-
Flores em Vida − 4	- − -	-	-	-	-	-	-

06/09/2014	2 − 6	0,5000	1,0000	1,0000	1,0000	1,0000	1,0000
Legendários	2 − 6	0,5000	1,0000	1,0000	1,0000	1,0000	1,0000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Cafajeste − 3	- − -	-	-	-	-	-	-
13/09/2014	6 − 19	0,8636	0,9500	0,9500	0,9500	0,9500	0,9500
Legendários	6 − 19	0,8636	0,9500	0,9500	0,9500	0,9500	0,9500
Fernando e Sorocaba	- − -	-	-	-	-	-	-
Gaveta − 2	- − -	-	-	-	-	-	-
20/09/2014	52 − 831	0,9858	0,9952	0,9952	0,9952	0,9952	0,9952
Legendários	52 − 831	0,9858	0,9952	0,9952	0,9952	0,9952	0,9952
Marcos e Belutti	- − -	-	-	-	-	-	-
Domingo de Manhã − 3	- − -	-	-	-	-	-	-
02/02/2014	19 − 97	1,0543	1,0104	1,0104	1,0104	1,0104	0,9700
O Melhor do Brasil	19 − 51	1,0625	1,0000	1,0000	1,0000	1,0000	1,0000
Michel Teló	23 − 7	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Humilde Residencia − 5	23 − 7	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
09/03/2014	35 − 299	1,0993	1,0455	1,0455	1,0455	1,0455	1,0455
O Melhor do Brasil	35 − 206	1,1638	0,9717	0,9717	0,9717	0,9717	0,9717
João Neto e Frederico	35 − 29	0,8529	0,9667	0,9667	0,9667	0,9667	1,1154
Crime Perfeito − 4	35 − 22	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
23/03/2014	66 − 1074	0,9471	0,9817	0,9817	0,9817	0,9817	0,9817
O Melhor do Brasil	66 − 811	0,9666	1,0201	1,0201	1,0201	1,0201	1,0201
Luan Santana	76 − 181	0,9330	1,0343	1,0343	1,0343	1,0343	1,0343
Tudo Que Você Quiser-Live − 5	76 − 138	0,9583	1,0147	1,0147	1,0147	1,0147	1,0147
30/03/2014	63 − 1318	1,0536	1,0178	1,0178	1,0178	1,0178	1,0178
O Melhor do Brasil	63 − 1046	1,0336	1,0019	1,0019	1,0019	1,0019	1,0019
Cristiano Araújo	63 − 203	1,0305	1,0305	1,0305	1,0305	1,0305	1,0305
Maus Bocados − 3	63 − 164	1,0061	0,9820	0,9820	0,9820	0,9820	0,9820

26/04/2014	65 − 985	1,0239	0,9899	0,9899	0,9899	0,9899	0,9899
Programa da Sabrina	65 − 763	1,0884	1,0381	1,0381	1,0381	1,0381	1,0381
Anitta	- − -	-	-	-	-	-	-
Zen − 5	- − -	-	-	-	-	-	-
26/07/2014	45 − 636	1,1357	0,9593	0,9593	0,9593	0,9593	0,9593
Programa da Sabrina	45 − 636	1,1357	0,9593	0,9593	0,9593	0,9593	0,9593
Zezé di Camargo e Luciano	- − -	-	-	-	-	-	-
Flores em Vida − 4	- − -	-	-	-	-	-	-
02/08/2014	2 − 9	0,6923	0,8182	0,8182	0,8182	0,8182	0,8182
Programa da Sabrina	2 − 8	0,8000	1,0000	1,0000	1,0000	1,0000	1,0000
Gaby Amarantos	- − -	-	-	-	-	-	-
Gaby Ostentação − 4	- − -	-	-	-	-	-	-
30/08/2014	2 − 19	1,7273	1,2667	1,2667	1,2667	1,2667	1,2667
Programa da Sabrina	2 − 19	1,7273	1,2667	1,2667	1,2667	1,2667	1,2667
Thaeme e Thiago	- − -	-	-	-	-	-	-
Cafajeste − 3	- − -	-	-	-	-	-	-
30/08/2014	1 − 11	1,0000	1,2222	1,2222	1,2222	1,2222	0,8462
Programa da Sabrina	1 − 11	1,0000	1,2222	1,2222	1,2222	1,2222	0,8462
Valesca Popozuda	- − -	-	-	-	-	-	-
Beijinho no Ombro − 1	- − -	-	-	-	-	-	-
20/09/2014	6 − 22	1,1579	0,9565	0,9565	0,9565	0,9565	0,9565
Programa da Sabrina	6 − 22	1,1579	0,9565	0,9565	0,9565	0,9565	0,9565
Fernando e Sorocaba	- − -	-	-	-	-	-	-
Gaveta − 2	- − -	-	-	-	-	-	-
27/09/2014	17 − 112	1,0275	1,0090	1,0090	1,0090	1,0090	1,0090
Programa da Sabrina	17 − 112	1,0275	1,0090	1,0090	1,0090	1,0090	1,0090
Victor e Léo	- − -	-	-	-	-	-	-
O Tempo Não Apaga − 7	- − -	-	-	-	-	-	-

08/02/2014	67 − 1849	1,0596	1,0027	1,0027	1,0027	1,0027	1,0027
Raul Gil	67 − 1258	1,0670	0,9836	0,9836	0,9836	0,9836	0,9836
Gusttavo Lima	- − -	-	-	-	-	-	-
Fui Fiel − 7	- − -	-	-	-	-	-	-
01/03/2014	38 − 453	1,1213	1,0203	1,0203	1,0203	1,0203	1,0203
Raul Gil	38 − 295	1,1569	1,0536	1,0536	1,0536	1,0536	1,0536
Eduardo Costa	- − -	-	-	-	-	-	-
Enamorado − 4	- − -	-	-	-	-	-	-
05/04/2014	63 − 1249	1,0122	0,9960	0,9960	0,9960	0,9960	0,9960
Raul Gil	63 − 1001	1,0030	0,9775	0,9775	0,9775	0,9775	1,0298
Cristiano Araújo	- − -	-	-	-	-	-	-
Maus Bocados − 3	- − -	-	-	-	-	-	-
13/09/2014	10 − 91	1,0706	1,0000	1,0000	1,0000	1,0000	1,0000
Raul Gil	10 − 91	1,0706	1,0000	1,0000	1,0000	1,0000	1,0000
Ivete Sangalo	- − -	-	-	-	-	-	-
Amor que Não Sai − 2	- − -	-	-	-	-	-	-
04/10/2014	2 − 16	1,0667	1,0000	1,0000	1,0000	1,0000	1,0000
Raul Gil	2 − 16	1,0667	1,0000	1,0000	1,0000	1,0000	1,0000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Cafajeste − 3	- − -	-	-	-	-	-	-
16/02/2014	2 − 6	0,7500	0,7500	0,7500	0,7500	0,7500	1,5000
Ritmo Brasil	2 − 6	0,7500	0,7500	0,7500	0,7500	0,7500	1,5000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Sinto Saudade − 3	- − -	-	-	-	-	-	-
02/03/2014	12 − 237	1,1675	1,0000	1,0000	1,0000	1,0000	1,0000
Ritmo Brasil	12 − 167	1,2463	0,9940	0,9940	0,9940	0,9940	0,9940
Péricles	16 − 27	1,0000	1,0385	1,0385	1,0385	1,0385	1,0385
Se Eu Largar o Freio − 3	16 − 26	1,1304	0,9630	0,9630	0,9630	0,9630	0,9630

100 Tabelas de taxas de acerto dos preditores para semanas/dias em análise

13/04/2014	66 − 955	1,0426	0,9795	0,9795	0,9795	0,9795	0,9795
Ritmo Brasil	66 − 696	1,0545	0,9574	0,9574	0,9574	0,9574	0,9574
Luan Santana	76 − 162	0,9643	1,0062	1,0062	1,0062	1,0062	1,0062
Tudo Que Você Quiser-Live − 5	76 − 125	0,9766	1,0331	1,0331	1,0331	1,0331	1,0331
20/04/2014	10 − 66	0,8250	0,9167	0,9167	0,9167	0,9167	0,9167
Ritmo Brasil	10 − 42	0,7636	0,8235	0,8235	0,8235	0,8235	0,8235
Grupo Bom Gosto	17 − 16	1,4545	1,0000	1,0000	1,0000	1,0000	1,0000
Mo Love-Live − 2	17 − 10	1,6667	0,9091	0,9091	0,9091	0,9091	0,9091
20/04/2014	12 − 109	0,9237	1,0187	1,0187	1,0187	1,0187	1,0187
Ritmo Brasil	12 − 63	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000
Péricles	16 − 20	1,2500	1,2500	1,2500	1,2500	1,2500	0,8000
Se Eu Largar o Freio − 3	16 − 12	2,0000	1,0000	1,0000	1,0000	1,0000	1,0000
27/04/2014	10 − 78	1,0263	0,9750	0,9750	0,9750	0,9750	0,9750
Ritmo Brasil	10 − 56	1,0769	1,0370	1,0370	1,0370	1,0370	0,9333
Grupo Bom Gosto	17 − 11	0,8462	1,0000	1,0000	1,0000	1,0000	1,0000
Mo Love-Live − 2	17 − 6	0,7500	1,0000	1,0000	1,0000	1,0000	1,0000
11/05/2014	66 − 649	0,9026	0,9863	0,9863	0,9863	0,9863	0,9863
Ritmo Brasil	66 − 412	0,8710	1,0000	1,0000	1,0000	1,0000	1,0000
Luan Santana	76 − 114	0,9048	0,9268	0,9268	0,9268	0,9268	0,9268
Tudo Que Você Quiser-Live − 5	76 − 78	0,8966	0,9398	0,9398	0,9398	0,9398	0,9398
21/09/2014	2 − 16	1,1429	1,0000	1,0000	1,0000	1,0000	1,0000
Ritmo Brasil	2 − 14	1,1667	1,0769	1,0769	1,0769	1,0769	1,0769
Gaby Amarantos	- − -	-	-	-	-	-	-
Gaby Ostentação − 4	- − -	-	-	-	-	-	-
19/10/2014	5 − 83	0,9540	1,0506	1,0506	1,0506	1,0506	1,0506
Ritmo Brasil	5 − 71	0,9342	1,0923	1,0923	1,0923	1,0923	1,0923
Michel Teló	4 − 10	0,8333	1,0000	1,0000	1,0000	1,0000	1,0000
Implorando pra trair − 5	4 − 10	0,8333	1,0000	1,0000	1,0000	1,0000	1,0000

02/02/2014	15 − 155	0,8708	0,9118	0,9118	0,9118	0,9118	0,9118
Sai do Chão	15 − 67	0,9054	0,9306	0,9306	0,9306	0,9306	0,9306
Naldo	16 − 22	0,8800	1,0000	1,0000	1,0000	1,0000	1,0000
Se Joga − 3	16 − 11	1,2222	1,1000	1,1000	1,1000	1,1000	0,8462
02/02/2014	68 − 1356	0,9516	0,9971	0,9971	0,9971	0,9971	0,9971
Sai do Chão	68 − 1055	0,9814	0,9991	0,9991	0,9991	0,9991	0,9991
Bruno e Marrone	73 − 232	0,9872	0,9872	0,9872	0,9872	0,9872	0,9872
Você Me Vira a Cabeça − 4	73 − 173	0,9505	1,0058	1,0058	1,0058	1,0058	1,0058
09/02/2014	74 − 1075	0,9496	0,9547	0,9547	0,9547	0,9547	0,9547
Sai do Chão	74 − 612	0,9623	0,9903	0,9903	0,9903	0,9903	0,9903
Paula Fernandes	78 − 160	0,9302	1,0127	1,0127	1,0127	1,0127	1,0127
Não Fui Eu − 6	78 − 83	0,9326	0,9765	0,9765	0,9765	0,9765	0,9765
01/02/2014	12 − 206	0,8512	0,9493	0,9493	0,9493	0,9493	0,9493
Sábado Total	12 − 147	0,8448	1,0138	1,0138	1,0138	1,0138	1,0138
Péricles	- − -	-	-	-	-	-	-
Se Eu Largar o Freio − 3	- − -	-	-	-	-	-	-
08/02/2014	2 − 9	1,2857	1,0000	1,0000	1,0000	1,0000	1,0000
Sábado Total	2 − 9	1,2857	1,0000	1,0000	1,0000	1,0000	1,0000
Thaeme e Thiago	- − -	-	-	-	-	-	-
Sinto Saudade − 3	- − -	-	-	-	-	-	-
08/03/2014	1 − 3	1,5000	1,0000	1,0000	1,0000	1,0000	1,0000
Sábado Total	1 − 3	1,5000	1,0000	1,0000	1,0000	1,0000	1,0000
Katinguelê	- − -	-	-	-	-	-	-
Ninguém Segura Esta Mulher − 1	- − -	-	-	-	-	-	-
05/04/2014	10 − 85	1,0241	0,9884	0,9884	0,9884	0,9884	0,9884
Sábado Total	10 − 60	1,0714	1,0526	1,0526	1,0526	1,0526	1,0526
Grupo Bom Gosto	- − -	-	-	-	-	-	-
Mo Love-Live − 2	- − -	-	-	-	-	-	-

11/10/2014	19 – 116	0,8529	1,0087	1,0087	1,0087	1,0087	1,0087
Sábado Total	19 – 116	0,8529	1,0087	1,0087	1,0087	1,0087	1,0087
João Neto e Frederico	- – -	-	-	-	-	-	-
Ele Nao Vai Mudar – 4	- – -	-	-	-	-	-	-

C.2 Resultados para os métodos de regressão

Nesta seção são apresentadas, na Tabela C.2, os resultados das taxas de acerto de predições para execuções por semanas e por dias, utilizando-se os métodos de regressão: Regressão Linear (RL) e Random Forest (R. Forest).

Tabela C.2: Resultados dos algoritmos de regressão

Prog./Artist.	Info. Nums.	RL	R. Forest
22/02/2014	14 – 68	0,9600	0,9300
Altas Horas	14 – 59	0,6700	0,7100
Skank	14 – 10	0,8300	0,8300
Sutilmente – 7	14 – 6	1,0000	0,8600
08/03/2014	79 – 1771	1,0000	1,0200
Altas Horas	79 – 921	1,1500	1,1500
Paula Fernandes	79 – 290	0,8900	1,0600
Quem é – 6	79 – 157	1,2500	1,1500
08/03/2014	74 – 1301	1,0700	1,0000
Altas Horas	74 – 977	1,1800	1,0100
Bruno e Marrone	74 – 204	0,7800	0,8800
Você Me Vira a Cabeça – 4	74 – 160	0,6600	0,9300

29/03/2014	77 − 1498	1,0400	1,0100
Altas Horas	77 − 1065	1,0100	1,0100
Gusttavo Lima	77 − 255	0,9900	1,0000
Fui Fiel − 7	77 − 178	0,9800	0,9600
19/04/2014	79 − 1657	0,9800	0,9600
Altas Horas	79 − 1051	1,0000	0,9800
Paula Fernandes	79 − 292	1,1200	1,0900
Quem é − 6	79 − 187	1,0900	1,1100
19/04/2014	35 − 174	0,8500	0,7700
Altas Horas	35 − 119	0,7600	0,7500
João Neto e Frederico	35 − 24	0,7500	0,7700
Crime Perfeito − 4	35 − 15	0,6800	0,6800
19/04/2014	25 − 92	0,8300	0,7400
Altas Horas	25 − 84	0,8200	0,7100
César Menotti e Fabiano	25 − 24	0,6200	0,8300
Não Era Eu − 2	25 − 22	0,4300	0,7900
19/04/2014	17 − 250	0,8900	0,8700
Altas Horas	17 − 223	0,8500	0,8900
Fernando e Sorocaba	17 − 44	0,7600	0,7600
Gaveta − 2	17 − 38	0,7900	0,7800
26/04/2014	5 − 47	0,6500	0,8400
Altas Horas	5 − 47	0,6500	0,8400
Pixote	5 − 8	0,7300	0,7300
Uma Simples Carta − 1	5 − 8	0,7300	0,7300
26/04/2014	26 − 183	1,0800	0,9000
Altas Horas	26 − 166	1,3300	0,9200
Só pra Contrariar	26 − 33	0,5600	0,6100
Me Perdoa - Live − 2	26 − 31	0,5600	0,6300

10/05/2014	1 – 2	1,0000	1,0000
Altas Horas	1 – 2	1,0000	1,0000
Harmonia do Samba	- – -	-	-
Nova Paradinha – 2	- – -	-	-
17/05/2014	24 – 158	0,9100	0,9300
Altas Horas	24 – 47	0,9600	0,9400
Thiaguinho	24 – 19	1,1200	1,0600
Ainda Bem - Live – 4	24 – 7	0,8800	0,8800
17/05/2014	50 – 280	0,9900	0,9200
Altas Horas	50 – 151	1,0300	0,9200
Jota Quest	50 – 32	1,0000	0,8200
Mandou Bem – 10	50 – 23	0,9200	0,7900
19/07/2014	2 – 8	1,1400	0,8900
Altas Horas	2 – 8	1,1400	0,8900
MC Guimê	- – -	-	-
Plaque de 100 – 1	- – -	-	-
09/08/2014	7 – 15	0,7600	0,7900
Altas Horas	7 – 15	0,7600	0,7900
Thaeme e Thiago	7 – 8	1,6000	1,3300
Cafajeste – 3	7 – 8	1,6000	1,3300
09/08/2014	2 – 10	0,7100	0,7100
Altas Horas	2 – 1	0,7100	0,7100
Gaby Amarantos	- – -	-	-
Gaby Ostentação – 4	- – -	-	-
30/08/2014	33 – 257	0,9400	0,9300
Altas Horas	33 – 257	0,9400	0,9300
João Neto e Frederico	33 – 62	1,1900	1,0500
Ele Nao Vai Mudar – 4	33 – 62	1,1900	1,0700

06/09/2014	40 − 102	0,8300	0,8300
Altas Horas	40 − 102	0,8300	0,8300
Victor e Léo	40 − 47	1,1500	0,9600
O Tempo Não Apaga − 7	40 − 47	1,1500	0,9600
13/09/2014	67 − 541	0,8500	1,0300
Altas Horas	67 − 541	0,8500	1,0300
Zezé di Camargo e Luciano	67 − 148	1,0600	1,1000
Flores em Vida − 4	67 − 148	1,0600	1,0900
13/09/2014	28 − 128	0,9000	0,9100
Altas Horas	28 − 128	0,9000	0,9000
Gusttavo Lima	28 − 41	0,8500	0,8700
10 Anos − 7	28 − 41	0,8500	0,8700
27/09/2014	75 − 144	1,0900	1,0800
Altas Horas	75 − 144	1,0900	1,0800
Skank	75 − 172	0,9500	1,0700
Ela me deixou − 7	75 − 172	0,9500	1,0700
04/10/2014	7 − 16	1,0500	1,0600
Altas Horas	7 − 16	1,0500	1,0600
Thaeme e Thiago	7 − 13	1,3000	1,3000
Cafajeste − 3	7 − 13	1,3000	1,3000
11/10/2014	6 − 65	1,3600	1,2700
Altas Horas	6 − 14	1,3600	1,2700
Michel Teló	6 − 1	0,3300	0,5000
Amiga Da Minha Irmã − 5	6 − 1	0,3300	0,3300
11/10/2014	100 − 783	0,9400	0,9400
Altas Horas	100 − 783	0,9400	0,9400
Marcos e Belutti	100 − 229	0,9400	0,9400
Domingo de Manhã − 3	100 − 229	0,9400	0,9400

15/03/2014	79 − 1605	0,9100	0,8800
Caldeirão do Huck	79 − 873	1,0800	1,0800
Paula Fernandes	79 − 268	0,9700	0,9400
Quem é − 6	79 − 144	1,2100	1,2100
19/04/2014	1 − 3	1,5000	1,5000
Caldeirão do Huck	1 − 3	1,5000	1,5000
Seu Jorge	- − -	-	-
Amiga da Minha Mulher − 1	- − -	-	-
19/07/2014	6 − 18	1,1700	1,1700
Caldeirão do Huck	6 − 18	1,1700	1,1700
Michel Teló	- − -	-	-
Amiga Da Minha Irmã − 5	- − -	-	-
02/08/2014	67 − 602	0,9700	0,9700
Caldeirão do Huck	67 − 602	0,9700	0,9700
Zezé di Camargo e Luciano	67 − 182	1,2100	1,2600
Flores em Vida − 4	67 − 182	1,2100	1,2600
20/09/2014	6 − 10	0,5000	0,5200
Caldeirão do Huck	6 − 10	0,5000	0,5000
Michel Teló	6 − 2	0,2900	1,0000
Amiga Da Minha Irmã − 5	6 − 2	0,2900	0,6700
09/03/2014	77 − 1499	0,9000	0,9000
De Frente com Gaby	77 − 1063	0,9400	0,9400
Gusttavo Lima	77 − 247	0,9000	0,9200
Fui Fiel − 7	77 − 187	0,8700	0,9400
03/08/2014	67 − 602	0,9700	0,9700
De Frente com Gaby	67 − 602	0,9700	0,9700
Zezé di Camargo e Luciano	67 − 182	1,2100	1,2600
Flores em Vida − 4	67 − 182	1,2100	1,2600

23/03/2014	84 − 1244	0,8100	0,8000
Domingo Espetacular	84 − 1056	0,8200	0,8200
Anitta	84 − 269	0,9600	0,9500
Zen − 5	84 − 238	1,0300	0,9900
30/03/2014	79 − 1728	1,0600	1,0600
Domingo Espetacular	79 − 1083	1,0900	1,1600
Paula Fernandes	79 − 297	1,1600	1,0800
Quem é − 6	79 − 187	1,4300	1,4500
16/02/2014	26 − 281	1,0000	0,9500
Domingo Legal	26 − 263	1,0000	0,9400
Só pra Contrariar	26 − 63	1,3700	1,0300
Me Perdoa - Live − 2	26 − 60	1,4600	1,0500
13/07/2014	55 − 309	0,8900	0,8800
Domingo Legal	55 − 309	0,8900	0,8800
Lucas Lucco	55 − 119	1,1900	1,0400
Mozão − 1	55 − 119	1,1900	1,0400
06/04/2014	17 − 33	0,9500	0,9300
Domingo Show	17 − 266	0,9500	0,9500
Fernando e Sorocaba	17 − 9	1,5000	1,5000
O Que Cê Vai Fazer? − 2	17 − 9	1,5000	1,5000
27/04/2014	77 − 1131	0,9800	0,8400
Domingo Show	77 − 755	0,9600	0,8200
Gusttavo Lima	77 − 194	0,8500	0,8500
Fui Fiel − 7	77 − 140	0,8300	0,8100
18/05/2014	23 − 143	1,1000	0,9800
Domingo Show	23 − 87	0,9700	0,8900
Daniel	23 − 22	1,0500	1,1600
Fale Um Pouco de Você − 3	23 − 13	1,0000	0,9300

20/07/2014	2 − 10	0,5000	0,5000
Domingo Show	2 − 4	0,5000	0,5000
Gaby Amarantos	- − -	-	-
Gaby Ostentação − 4	- − -	-	-
07/09/2014	33 − 233	0,9600	0,9500
Domingo Show	33 − 233	0,9600	0,9500
João Neto e Frederico	33 − 62	1,0200	1,0500
Ele Nao Vai Mudar − 4	33 − 62	1,0200	1,0500
23/02/2014	17 − 170	1,0200	1,0300
Domingo da Gente	17 − 72	1,1800	1,2600
Naldo	17 − 23	1,0500	1,1000
Caipifruta − 3	17 − 10	1,6700	1,2500
23/02/2014	79 − 238	0,8100	0,8100
Domingão do Faustão	79 − 793	0,8500	0,7500
Paula Fernandes	79 − 158	1,0100	0,9500
Se o Coração Viajar − 6	79 − 73	0,8800	0,8500
16/03/2014	79 − 1605	0,9100	0,8800
Domingão do Faustão	79 − 873	1,0800	1,0800
Paula Fernandes	79 − 268	0,9700	0,9400
Quem é − 6	79 − 144	1,2100	1,2100
16/03/2014	29 − 303	0,8500	0,8500
Domingão do Faustão	29 − 303	0,8500	0,8500
Ivete Sangalo	29 − 59	1,0200	1,0200
Tempo de Alegria − 2	29 − 59	1,0200	1,0200
16/03/2014	84 − 1474	0,8900	0,8800
Domingão do Faustão	84 − 1231	0,9100	0,9200
Anitta	84 − 268	0,9500	0,9200
Zen − 5	84 − 235	1,0200	0,9600

16/03/2014	24 − 223	1,0500	1,1000
Domingão do Faustão	24 − 90	1,2400	1,1800
Thiaguinho	24 − 52	1,5800	1,6800
As Aparências Enganam − 4	24 − 20	1,4300	1,6700
16/03/2014	77 − 1174	0,9200	0,8300
Domingão do Faustão	77 − 883	0,8800	0,8000
Luan Santana	77 − 199	0,7900	0,7500
Tudo Que Você Quiser-Live − 5	77 − 151	0,7700	0,8300
16/03/2014	23 − 196	1,1500	0,9600
Domingão do Faustão	23 − 157	1,2000	0,9900
Daniel	23 − 34	1,2600	0,9700
Fale Um Pouco de Você − 3	23 − 25	1,2500	0,9600
23/03/2014	26 − 212	0,9100	0,9000
Domingão do Faustão	26 − 194	0,9200	0,9200
Só pra Contrariar	26 − 50	0,9800	1,0600
Me Perdoa - Live − 2	26 − 47	1,0200	1,0400
30/03/2014	26 − 135	1,0800	0,9400
Domingão do Faustão	26 − 75	0,9600	0,8500
Leonardo	26 − 21	0,8100	0,7800
Choro − 4	26 − 14	0,8200	0,8200
20/04/2014	84 − 735	0,9100	0,8600
Domingão do Faustão	84 − 91	0,9300	0,8600
Anitta	84 − 166	1,1400	0,9400
Zen − 5	84 − 140	1,1500	0,9400
04/05/2014	24 − 138	0,9100	0,8600
Domingão do Faustão	24 − 53	1,0600	0,9300
Thiaguinho	24 − 21	0,9500	0,9500
As Aparências Enganam − 4	24 − 11	1,5700	1,5700

11/05/2014	29 − 119	1,0600	0,7700
Domingão do Faustão	29 − 251	1,0600	0,7700
Ivete Sangalo	29 − 24	1,1400	0,8600
Tempo de Alegria − 2	29 − 24	1,1400	0,8900
18/05/2014	64 − 468	1,1600	0,8500
Domingão do Faustão	64 − 402	1,1700	0,8200
Zezé di Camargo e Luciano	64 − 65	0,8300	0,7000
Teorias − 4	64 − 57	0,8500	0,7200
27/07/2014	55 − 293	0,9500	0,9000
Domingão do Faustão	55 − 293	0,9500	0,9100
Lucas Lucco	55 − 105	0,8600	0,8700
Mozão − 1	55 − 105	0,8600	0,8700
24/08/2014	100 − 960	1,1000	1,1200
Domingão do Faustão	100 − 960	1,1000	1,1200
Marcos e Belutti	100 − 284	0,9100	0,9300
Domingo de Manhã − 3	100 − 284	0,9100	0,9300
05/10/2014	26 − 70	0,8800	0,9100
Domingão do Faustão	26 − 70	0,8800	0,9100
Ivete Sangalo	26 − 34	1,0600	1,0600
Amor que Não Sai − 2	26 − 34	1,0600	1,0600
09/03/2014	51 − 364	0,8300	0,8100
Eliana	51 − 230	0,8800	0,8500
Eduardo Costa	51 − 80	1,1000	1,0400
Enamorado − 4	51 − 48	1,1200	1,1200
20/04/2014	17 − 250	0,8900	0,8700
Eliana	17 − 223	0,8500	0,8900
Fernando e Sorocaba	17 − 44	0,7600	0,7600
Gaveta − 2	17 − 38	0,7900	0,7800

20/04/2014	20 − 66	0,8800	0,8900
Esquenta	20 − 42	0,8900	0,8400
Grupo Bom Gosto	20 − 17	1,0600	0,8500
Mo Love-Live − 2	20 − 11	0,9200	0,7900
18/05/2014	77 − 1133	1,0600	0,9600
Esquenta	77 − 751	1,0600	0,9800
Gusttavo Lima	77 − 157	0,9800	0,8700
Fui Fiel − 7	77 − 100	0,8300	0,8300
13/07/2014	2 − 8	0,6700	0,6200
Esquenta	2 − 8	0,6700	0,6200
MC Guimê	- − -	-	-
Plaque de 100 − 1	- − -	-	-
13/07/2014	9 − 2	1,1900	1,2300
Esquenta	9 − 2	1,1900	1,2300
Psirico	9 − 10	0,9100	0,9100
Lepo Lepo − 1	9 − 10	0,9100	0,9100
03/08/2014	55 − 244	0,9100	0,8600
Esquenta	55 − 244	0,9100	0,8600
Lucas Lucco	55 − 82	0,7900	0,7500
Mozão − 1	55 − 82	0,7900	0,7600
03/08/2014	26 − 137	0,9800	0,9800
Esquenta	26 − 137	0,9800	0,9800
Ivete Sangalo	26 − 60	1,2200	1,2500
Amor que Não Sai − 2	26 − 60	1,2200	1,2500
10/08/2014	100 − 774	0,7600	0,7500
Esquenta	100 − 774	0,7600	0,7400
Marcos e Belutti	100 − 308	1,0100	0,9500
Domingo de Manhã − 3	100 − 308	1,0100	0,9500

24/08/2014	75 − 168	1,0500	1,0800
Esquenta	75 − 168	1,0500	1,0800
Skank	75 − 158	1,0300	0,9900
Ela me deixou − 7	75 − 158	1,0300	0,9900
21/09/2014	2 − 8	0,8000	0,8000
Esquenta	2 − 8	0,8000	0,8000
MC Guimê	- − -	-	-
Plaque de 100 − 1	- − -	-	-
01/03/2014	18 − 237	0,9700	0,9700
Estrelas	18 − 167	0,9600	0,9700
Péricles	18 − 30	0,9700	0,9400
Se Eu Largar o Freio − 3	18 − 29	1,0000	1,0000
12/07/2014	6 − 16	1,0000	1,0000
Estrelas	6 − 16	1,0000	1,0000
Michel Teló	- − -	-	-
Amiga Da Minha Irmã − 5	- − -	-	-
12/07/2014	6 − 36	0,9400	0,9500
Estrelas	6 − 36	0,9400	0,9500
Valesca Popozuda	6 − 18	1,8000	2,5700
Beijinho no Ombro − 1	6 − 18	1,8000	2,5700
09/08/2014	33 − 241	0,7300	0,7300
Estrelas	33 − 241	0,7300	0,7400
João Neto e Frederico	33 − 66	0,8700	0,8600
Ele Nao Vai Mudar − 4	33 − 66	0,8700	0,8600
23/03/2014	85 − 1031	1,0400	0,8600
Fantástico	85 − 863	1,0600	0,8500
Victor e Léo	85 − 160	1,0000	0,8400
Na Linha do Tempo − 7	85 − 130	0,9200	0,7900

23/03/2014	84 − 1244	0,8100	0,8000
Fantástico	84 − 1056	0,8200	0,8200
Anitta	84 − 269	0,9600	0,9500
Zen − 5	84 − 238	1,0300	0,9900
30/03/2014	77 − 1094	1,1400	1,0000
Fantástico	77 − 795	1,0600	0,9700
Luan Santana	77 − 191	0,9000	0,9700
Tudo Que Você Quiser-Live − 5	77 − 141	0,8300	0,8000
13/04/2014	26 − 228	0,9700	0,9800
Fantástico	26 − 213	0,9800	0,9900
Só pra Contrariar	26 − 54	0,8700	0,9600
Me Perdoa - Live − 2	26 − 49	0,8400	0,9800
20/07/2014	6 − 18	1,1700	1,1700
Fantástico	6 − 18	1,1700	1,1700
Michel Teló	- − -	-	-
Amiga Da Minha Irmã − 5	- − -	-	-
27/07/2014	6 − 17	1,5600	1,5600
Fantástico	6 − 17	1,5600	1,5600
Michel Teló	6 − 4	2,0000	2,0000
Amiga Da Minha Irmã − 5	6 − 4	2,0000	2,0000
03/08/2014	6 − 13	0,4600	0,8800
Fantástico	6 − 13	0,4600	0,8400
Michel Teló	6 − 3	1,5000	1,5000
Amiga Da Minha Irmã − 5	6 − 3	1,5000	1,5000
10/08/2014	6 − 11	0,7600	0,6700
Fantástico	6 − 11	0,7600	0,7000
Michel Teló	6 − 4	2,0000	2,0000
Amiga Da Minha Irmã − 5	6 − 4	2,0000	2,0000

17/08/2014	6 − 12	0,7600	0,7600
Fantástico	6 − 12	0,7600	0,7600
Michel Teló	6 − 2	0,6700	0,6700
Amiga Da Minha Irmã − 5	6 − 2	0,6700	0,6700
24/08/2014	6 − 14	1,3300	1,4100
Fantástico	6 − 14	1,3300	1,4100
Michel Teló	6 − 1	0,2000	0,3300
Amiga Da Minha Irmã − 5	6 − 1	0,2000	0,3300
12/10/2014	6 − 65	1,3600	1,2700
Fantástico	6 − 14	1,3600	1,2700
Michel Teló	6 − 1	0,3300	0,5000
Amiga Da Minha Irmã − 5	6 − 1	0,3300	0,3300
12/10/2014	67 − 515	1,0900	1,0300
Fantástico	67 − 515	1,0900	1,0300
Zezé di Camargo e Luciano	67 − 153	1,0800	1,0900
Flores em Vida − 4	67 − 153	1,0800	1,0900
19/10/2014	40 − 114	1,1500	1,1900
Fantástico	40 − 114	1,1500	1,1800
Victor e Léo	40 − 41	1,0500	1,0800
O Tempo Não Apaga − 7	40 − 41	1,0500	1,0800
19/10/2014	6 − 71	1,2700	1,1700
Fantástico	6 − 12	1,2700	1,1700
Michel Teló	6 − 4	4,0000	4,0000
Amiga Da Minha Irmã − 5	6 − 4	4,0000	4,0000
19/10/2014	13 − 16	0,8600	0,9100
Fantástico	13 − 16	0,8600	0,9100
Fernando e Sorocaba	13 − 10	2,0000	1,6700
Gaveta − 2	13 − 10	2,0000	1,6700

11/05/2014	6 − 14	0,8800	0,8300
Frank Aguiar e Amigos	6 − 14	0,8800	0,8800
Rio Negro e Solimões	6 − 2	0,5000	0,6700
Romântica − 1	6 − 2	0,5000	0,6700
27/04/2014	22 − 585	0,9800	0,9900
Hora do Faro	22 − 528	1,0100	1,0200
Marcos e Belutti	22 − 7	0,5800	0,6400
Domingo de Manhã − 3	22 − 6	0,6000	0,5500
18/05/2014	23 − 126	0,8000	0,7800
Hora do Faro	23 − 71	0,6800	0,7100
Munhoz e Mariano	23 − 16	0,9400	0,9400
Pantera Cor de Rosa − 3	23 − 11	0,8500	0,8500
03/08/2014	40 − 191	0,5800	0,7200
Hora do Faro	40 − 191	0,5800	0,7200
Victor e Léo	40 − 91	0,7500	0,7500
O Tempo Não Apaga − 7	40 − 91	0,7500	0,7500
21/09/2014	2 − 8	0,8000	0,8000
Hora do Faro	2 − 8	0,8000	0,8000
MC Guimê	- − -	-	-
Plaque de 100 − 1	- − -	-	-
28/09/2014	67 − 547	0,9800	0,9700
Hora do Faro	67 − 547	0,9800	0,9700
Zezé di Camargo e Luciano	67 − 138	0,9700	0,9600
Flores em Vida − 4	67 − 138	0,9700	0,9600
15/02/2014	25 − 184	0,9500	0,9800
Legendários	25 − 170	0,9600	0,9800
César Menotti e Fabiano	25 − 41	0,8700	0,8900
Não Era Eu − 2	25 − 40	0,9300	0,9500

01/03/2014	35 – 347	1,0900	1,0700
Legendários	35 – 212	0,9500	0,9700
João Neto e Frederico	35 – 50	1,3500	1,4300
Crime Perfeito – 4	35 – 22	1,4700	1,0000
22/03/2014	51 – 340	1,1400	0,9800
Legendários	51 – 193	1,2200	0,9300
Eduardo Costa	51 – 62	0,8200	0,8400
Enamorado – 4	51 – 34	0,7100	0,7600
12/04/2014	67 – 1288	1,0200	1,0100
Legendários	67 – 1033	1,0100	1,0000
Cristiano Araújo	67 – 194	0,9800	0,9600
Maus Bocados – 3	67 – 157	0,9500	0,9500
19/04/2014	23 – 138	0,8000	0,7800
Legendários	23 – 92	0,7600	0,7400
Munhoz e Mariano	23 – 18	1,0000	1,0000
Pantera Cor de Rosa – 3	23 – 13	0,8700	0,9300
19/04/2014	22 – 58	1,0400	1,0300
Legendários	22 – 543	0,9500	0,9500
Marcos e Belutti	22 – 12	0,9200	1,0000
I Love You-Live – 3	22 – 9	0,6400	0,8200
26/04/2014	6 – 40	0,9100	0,9100
Legendários	6 – 22	0,8300	0,7700
Onze 20	- – -	-	-
Não Vai Voltar – 2	- – -	-	-
03/05/2014	32 – 220	0,8600	0,8600
Legendários	32 – 101	0,9500	0,9200
Turma do Pagode	32 – 33	0,9200	0,9200
Horário de Verão – 4	32 – 17	0,8500	0,8500

03/05/2014	6 − 15	0,8000	0,7600
Legendários	6 − 15	0,8000	0,7600
Rio Negro e Solimões	6 − 1	0,3300	0,3300
Romântica − 1	6 − 1	0,3300	0,3300
10/05/2014	84 − 833	0,9600	0,9100
Legendários	84 − 623	0,9700	0,8700
Anitta	84 − 140	0,9900	0,9100
Zen − 5	84 − 103	0,8400	0,8200
17/05/2014	17 − 66	0,5800	0,5700
Legendários	17 − 24	0,8300	0,8300
Naldo	17 − 6	0,4600	0,4300
Caipifruta − 3	17 − 3	0,6000	0,7500
19/07/2014	67 − 663	1,1400	1,1700
Legendários	67 − 663	1,1400	1,1600
Zezé di Camargo e Luciano	67 − 179	1,2900	1,2900
Flores em Vida − 4	67 − 179	1,2900	1,2900
06/09/2014	7 − 6	0,9100	0,9500
Legendários	7 − 6	0,9100	0,9500
Thaeme e Thiago	7 − 8	0,8900	1,0000
Cafajeste − 3	7 − 8	0,8900	0,8900
13/09/2014	13 − 19	0,8800	0,9000
Legendários	13 − 19	0,8800	0,9000
Fernando e Sorocaba	13 − 8	0,6700	0,7300
Gaveta − 2	13 − 8	0,6700	0,7300
20/09/2014	100 − 831	0,9400	0,9400
Legendários	100 − 831	0,9400	0,9400
Marcos e Belutti	100 − 238	0,9000	0,8900
Domingo de Manhã − 3	100 − 238	0,9000	0,8900

18/10/2014	2 − 10	1,4300	1,4300
Legendários	2 − 10	1,4300	1,4300
Israel Novaes	- − -	-	-
Beijo Meu − 1	- − -	-	-
09/03/2014	35 − 299	1,0900	1,0300
O Melhor do Brasil	35 − 206	1,0300	1,0300
João Neto e Frederico	35 − 29	-4,8300	0,8800
Crime Perfeito − 4	35 − 22	1,0000	1,0000
23/03/2014	77 − 1074	1,0100	0,8800
O Melhor do Brasil	77 − 811	1,0900	0,9000
Luan Santana	77 − 181	0,8300	0,8200
Tudo Que Você Quiser-Live − 5	77 − 138	0,7800	0,8100
30/03/2014	67 − 1318	1,0600	1,0700
O Melhor do Brasil	67 − 1046	1,0300	1,0400
Cristiano Araújo	67 − 207	1,0500	1,0100
Maus Bocados − 3	67 − 167	1,0300	1,0800
26/04/2014	84 − 985	0,9200	0,7700
Programa da Sabrina	84 − 763	1,0700	0,9600
Anitta	84 − 136	0,8400	0,8100
Zen − 5	84 − 117	0,8500	0,8300
17/05/2014	32 − 184	0,8000	0,7800
Programa da Sabrina	32 − 68	1,0100	0,7800
Turma do Pagode	32 − 20	5,0000	0,7100
Horário de Verão − 4	32 − 11	-1,0000	1,0000
26/07/2014	67 − 636	0,8200	1,0000
Programa da Sabrina	67 − 636	0,8200	1,0000
Zezé di Camargo e Luciano	67 − 140	1,3100	1,0100
Flores em Vida − 4	67 − 140	1,3100	1,0100

02/08/2014	2 − 8	1,1400	1,1400
Programa da Sabrina	2 − 1	1,1400	1,1400
Gaby Amarantos	- − -	-	-
Gaby Ostentação − 4	- − -	-	-
30/08/2014	7 − 19	1,3400	1,3100
Programa da Sabrina	7 − 19	1,3400	1,3100
Thaeme e Thiago	7 − 11	1,5700	1,5700
Cafajeste − 3	7 − 11	1,5700	1,5700
30/08/2014	6 − 11	0,6800	0,5200
Programa da Sabrina	6 − 11	0,6800	0,5300
Valesca Popozuda	6 − 3	-0,5000	1,0000
Beijinho no Ombro − 1	6 − 3	-0,5000	1,0000
20/09/2014	13 − 22	0,6100	0,6900
Programa da Sabrina	13 − 22	0,6100	0,7000
Fernando e Sorocaba	13 − 5	0,5000	0,5000
Gaveta − 2	13 − 5	0,5000	0,5000
27/09/2014	40 − 112	1,0000	0,9700
Programa da Sabrina	40 − 112	1,0000	0,9800
Victor e Léo	40 − 32	0,7400	0,7300
O Tempo Não Apaga − 7	40 − 32	0,7400	0,7300
04/10/2014	2 − 7	1,0000	1,0000
Programa da Sabrina	2 − 7	1,0000	1,0000
Israel Novaes	- − -	-	-
Beijo Meu − 1	- − -	-	-
01/03/2014	51 − 453	1,0100	0,9200
Raul Gil	51 − 295	1,0300	0,9600
Eduardo Costa	51 − 73	0,9900	0,8700
Enamorado − 4	51 − 45	1,3200	1,0000

05/04/2014	67 – 1249	0,9900	1,0100
Raul Gil	67 – 1001	0,9900	1,0000
Cristiano Araújo	67 – 200	1,0300	0,9900
Maus Bocados – 3	67 – 163	0,9800	0,9600
13/09/2014	26 – 91	0,8700	0,8900
Raul Gil	26 – 91	0,8700	0,9000
Ivete Sangalo	26 – 34	0,8100	0,8500
Amor que Não Sai – 2	26 – 34	0,8100	0,8500
04/10/2014	7 – 16	1,0500	1,0600
Raul Gil	7 – 16	1,0500	1,0600
Thaeme e Thiago	7 – 13	1,3000	1,3000
Cafajeste – 3	7 – 13	1,3000	1,3000
18/10/2014	2 – 6	0,8600	1,0000
Raul Gil	2 – 6	0,8600	1,0000
MC Guimê	- – -	-	-
Plaque de 100 – 1	- – -	-	-
16/02/2014	2 – 6	0,8600	0,8600
Ritmo Brasil	2 – 6	0,8600	0,7500
Thaeme e Thiago	- – -	-	-
Sinto Saudade – 3	- – -	-	-
02/03/2014	18 – 237	0,9700	0,9700
Ritmo Brasil	18 – 167	0,9600	0,9700
Péricles	18 – 30	0,9700	0,9400
Se Eu Largar o Freio – 3	18 – 29	1,0000	1,0000
13/04/2014	77 – 955	0,9500	0,9100
Ritmo Brasil	77 – 696	0,9500	0,9200
Luan Santana	77 – 163	0,9100	0,8900
Tudo Que Você Quiser-Live – 5	77 – 125	0,9200	0,9100

20/04/2014	20 − 66	0,8800	0,8900
Ritmo Brasil	20 − 42	0,8900	0,8400
Grupo Bom Gosto	20 − 17	1,0600	0,8500
Mo Love-Live − 2	20 − 11	0,9200	0,7900
20/04/2014	18 − 109	1,0100	0,8400
Ritmo Brasil	18 − 63	1,3900	0,8300
Péricles	18 − 22	1,1600	1,0500
Se Eu Largar o Freio − 3	18 − 12	0,7500	0,8600
27/04/2014	20 − 78	0,8900	0,9300
Ritmo Brasil	20 − 56	1,2000	1,1100
Grupo Bom Gosto	20 − 14	0,9300	0,8200
Mo Love-Live − 2	20 − 8	0,7300	0,6700
11/05/2014	77 − 649	0,9700	0,8300
Ritmo Brasil	77 − 412	0,8900	0,7500
Luan Santana	77 − 114	0,9300	0,8300
Tudo Que Você Quiser-Live − 5	77 − 78	1,0000	0,7900
21/09/2014	2 − 14	1,0000	0,9100
Ritmo Brasil	2 − 2	1,0000	0,9100
Gaby Amarantos	- − -	-	-
Gaby Ostentação − 4	- − -	-	-
19/10/2014	6 − 71	1,2700	1,1700
Ritmo Brasil	6 − 12	1,2700	1,1700
Michel Teló	6 − 4	4,0000	4,0000
Amiga Da Minha Irmã − 5	6 − 4	4,0000	4,0000
08/03/2014	1 − 3	1,0000	1,0000
Sábado Total	1 − 3	1,0000	1,0000
Katinguelê	- − -	-	-
Ninguém Segura Esta Mulher − 1	- − -	-	-

05/04/2014	20 − 85	1,2400	1,0200
Sábado Total	20 − 60	1,2400	0,9800
Grupo Bom Gosto	20 − 17	0,5700	0,7100
Mo Love-Live − 2	20 − 12	0,6300	0,7500
17/05/2014	32 − 184	0,8000	0,7800
Sábado Total	32 − 68	1,0100	0,7800
Turma do Pagode	32 − 20	5,0000	0,7100
Horário de Verão − 4	32 − 11	-1,0000	1,0000
11/10/2014	33 − 116	0,8500	0,7900
Sábado Total	33 − 116	0,8500	0,8000
João Neto e Frederico	33 − 33	1,4300	0,7200
Ele Nao Vai Mudar − 4	33 − 33	1,4300	0,7300

Referências Bibliográficas

Agrawal, R., Imieliński, T. and Swami, A. (1993). Mining association rules between sets of items in large databases, *ACM SIGMOD Record*, Vol. 22, ACM, pp. 207–216.

Aizenberg, N., Koren, Y. and Somekh, O. (2012). Build your own music recommender by modeling internet radio streams, *Proceedings of the 21st international conference on World Wide Web*, International World Wide Web Conference '12, ACM, New York, NY, USA, pp. 1–10.

Barbosa Filho, A. (2003). *Gêneros radiofônicos: os formatos e os programas em áudio*, Paulinas.

Breiman, L. (2001). Random forests, *Machine learning* **45**(1): 5–32.

Bufarah Junior, Á. (2003). Rádio na internet: convergência de possibilidades, *Trabalho apresentado no Núcelo de Mídia Sonora, XXVI Congresso Anual em Ciências da Comunicação, Belo Horizonte/MG*, Vol. 2.

Câmara, F. G. and Silva, O. (2001). Estatística não paramétrica: testes de hipóteses e medidas de associação, *Departamento de Matemática-Universidade dos Açores, Ponta Delgada* .

Camilo, C. O. and Silva, J. C. (2009). Mineração de dados: Conceitos, tarefas, métodos e ferramentas, *Universidade Federal de Goiás (UFC)* pp. 1–29.

Campos, G. M. (2001). Estatística prática para docentes e pós-graduandos, *São Paulo: Faculdade de Odontologia de Ribeirão Preto* .

Crookston, N. L., Finley, A. O. et al. (2008). yaimpute: An r package for knn imputation, *Journal of Statistical Software* **23**(10): 1–16.

Da Rosa, J. M. C. (2009). Estatística ii (notas de aula), `http://www.est.ufpr.br/ce003/material/apostilace003.pdf`.

da Silva, B. C., Mantuan, B. and Marques, J. C. (2012). Ecologia da mídia: um estudo sobre suas interfaces na mediação do fato esportivo, *Trabalho apresentado no DT Estudos Interdisciplinares da Comunicação, XVII Congresso de Ciências da Comunicação na Região Sudeste, Ouro Preto/MG*.

Dubber, A. (2013). Repensando o rádio na era digital, *Significação-Revista de Cultura Audiovisual* **40**(39): 24–43.

Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996). From data mining to knowledge discovery in databases, *AI magazine* **17**(3): 37.

Genes, F., Craveiro, R. U. and Proença, A. (2012). Inovações tecnológicas na cadeia produtiva da música no século xxi, *Sistemas & Gestão* **7**(2): 174–191.

Guimarães, P. R. B. (2014). Estatística não paramétrica, `http://people.ufpr.br/~prbg/public_html/ce050/aluno%202014%20np.pdf`.

Hahsler, M., Grün, B., Hornik, K. and Buchta, C. (2009). Introduction to arules–a computational environment for mining association rules and frequent item sets, *The Comprehensive R Archive Network* .

Jaccard, P. (1901). *Distribution de la Flore Alpine: dans le Bassin des dranses et dans quelques régions voisines*, Rouge.

Janotti Junior, J. (2006). Por uma análise midiática da música popular massiva. uma proposição metodológica para a compreensão do entorno comunicacional, das condições de produção e reconhecimento dos gêneros musicais.

Kischinhevsky, M. (2011). Amigo ouvinte, o locutor perdeu o emprego: considerações sobre o processo de automação nas rádios fm do rio de janeiro, *Revista Eptic Online* **10**(3).

Koenigstein, N., Dror, G. and Koren, Y. (2011). Yahoo! music recommendations: modeling music ratings with temporal dynamics and item taxonomy, *Proceedings of the fifth ACM conference on Recommender systems*, RecSys '11, ACM, New York, NY, USA, pp. 165–172.

Kuo, F.-F., Chiang, M.-F., Shan, M.-K. and Lee, S.-Y. (2005). Emotion-based music recommendation by association discovery from film music, *Proceedings of the 13th annual ACM international conference on Multimedia*, ACM, pp. 507–510.

Lidy, T. and Rauber, A. (2006). Visually profiling radio stations., *The International Society for Music Information Retrieval*, pp. 186–191.

Liu, Y., Wang, Y., Shenoy, A., Tsai, W.-H. and Cai, L. (2008). Clustering music recordings by their keys., *ISMIR*, pp. 319–324.

Macedo, S. M., de Barros, A. T. and Bernardes, C. B. (2011). Identidade política e programação das rádios públicas: estudo comparativo de quatro emissoras federais, *Fronteiras-estudos midiáticos* **14**(1): 40–51.

Mahajan, S., Reshamwala, A., Sharma, N., Vineet, D., Sharma, A. and Shah, P. (2012). Prediction of yahoo! music sequences on user's musical taste, *Proceedings of the International Conference on Advances in Information Technology*, AIT, pp. 6–9.

Medeiros, M. (2009). Transmissão sonora digital: modelos radiofônicos e não radiofônicos na comunicação contemporânea, *Ciberlegenda* .

Moreira, S. V. (1999). Rádio@ internet, *Nélia R. del Bianco and Sônia Virgínia Moreira (eds), Rádio no Brasil: Tendências e perspectivas, Brasília: UNB, and Rio de Janeiro: Uerj* pp. 205–23.

Nascimento, A. S., Castor, F., Rubira, C. M. and Burrows, R. (2012). An experimental setup to assess design diversity of functionally equivalent services, *16th International Conference on Evaluation and Assessment in Software Engineering*.

Navarro, G. (2001). A guided tour to approximate string matching, *ACM computing surveys (CSUR)* **33**(1): 31–88.

Neubarth, K., Goienetxea, I., Johnson, C. and Conklin, D. (2012). Association mining of folk music genres and toponyms., *ISMIR*, pp. 7–12.

Pachet, F., Westermann, G. and Laigre, D. (2001). Musical data mining for electronic music distribution, *Web Delivering of Music, 2001. Proceedings. First International Conference on*, IEEE, pp. 101–106.

Prata, N. (2008). *Webradio: novos gêneros, novas formas de interação*, PhD thesis, Faculdade de Letras, UFMG.

Prestes Filho, L. C. (2004). *Cadeia Produtiva da Economia da Música*, Rio de Janeiro: Incubadora Cultural Gênesis PUC–Rio.

Ragno, R., Burges, C. J. C. and Herley, C. (2005). Inferring similarity between music objects with application to playlist generation, *Proceedings of the 7th ACM SIGMM international workshop on Multimedia information retrieval*, Management in Radiology '05, ACM, New York, NY, USA, pp. 73–80.

Rezende, D. D. d. (2014). *Um modelo de classificação supervisionada com rotulagem automática para reconhecimento de áudio de web rádios*, Master's thesis, Universidade Federal de Ouro Preto, UFOP, Ouro Preto, MG, Brasil.

Sanchotene, C., Zucolo, R. and Pires, J. (2010). No ar, a comunidade: um estudo de recepção a partir das rádios comunitárias de santa maria, *Intexto* **1**(22): 81–98.

Siegel, S. and Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences*, McGraw-Hill.

Trotta, F. (2009). Música e mercado: a força das classificações, *Contemporanea-Revista de Comunicação e Cultura* **3**(2).

Venables, W. N., Smith, D. M. and Team, R. D. C. (2002). An introduction to r.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B. and Wesslén, A. (2012). *Experimentation in software engineering*, Springer Science & Business Media.

Ye, Y. and Chiang, C.-C. (2006). A parallel apriori algorithm for frequent itemsets mining, *Software Engineering Research, Management and Applications, 2006. Fourth International Conference on*, IEEE, pp. 87–94.