**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS**

**NÚCLEO DE EDUCAÇÃO A DISTÂNCIA**

**Pós-graduação *Lato Sensu* em Ciência de Dados e Big Data**

**Rafael Santos Carvalho**

**Predição da arrecadação de tributos: um comparativo entre o *Facebook™ Prophet* e as redes neurais *Long Short-Term Memory***

**Varginha, Minas Gerais**

**2021**

**Rafael Santos Carvalho**

**PREDIÇÃO DA ARRECADAÇÃO DE TRIBUTOS: UM COMPARATIVO ENTRE O *FACEBOOK™ PROPHET* E AS REDES NEURAIS *LONG SHORT-TERM MEMORY***

Trabalho de Conclusão de Curso apresentado ao Curso de Especialização em Ciência de Dados e Big Data como requisito parcial à obtenção do título de especialista.

**Varginha, Minas Gerais**

**2021**

**LISTA DE FIGURAS**

[Figura 1 – Valores e proporções da arrecadação por fonte de receita 8](#_Toc68879640)

[Figura 2 – Download e tratamento de dados do CAGED 11](#_Toc68879641)

[Figura 3 – Métodos de atualização de valores a valor presente 12](#_Toc68879642)

[Figura 4 – Descrição da arrecadação nominal dos tributos estaduais 13](#_Toc68879643)

[Figura 5 – Arrecadação nominal do ICMS do RS 13](#_Toc68879644)

[Figura 6 – Arrecadação nominal do IPVA do RS 13](#_Toc68879645)

[Figura 7 – Arrecadação nominal do ITCD do RS 14](#_Toc68879646)

[Figura 8 – Descrição da arrecadação real dos tributos estaduais 15](#_Toc68879647)

[Figura 9 – Arrecadação real do ICMS do RS 15](#_Toc68879648)

[Figura 10 – Arrecadação real do IPVA do RS 16](#_Toc68879649)

[Figura 11 – Arrecadação real do ITCD do RS 16](#_Toc68879650)

[Figura 12 – Boxplot da arrecadação real do ICMS do RS 17](#_Toc68879651)

[Figura 13 – Boxplot da arrecadação real do IPVA do RS 17](#_Toc68879652)

[Figura 14 – Boxplot da arrecadação real do ITCD do RS 18](#_Toc68879653)

**LISTA DE ABREVIATURAS E SIGLAS**

CAGED – Cadastro Geral de Empregados e Desempregados

ICMS – Imposto sobre Circulação de Mercadorias e Serviços

IGP-M – Índice Geral de Preços - Mercado

IPVA – Imposto sobre a Propriedade de Veículos Automotores

ITCD – Imposto sobre Transmissão Causa mortis e Doação

LDO – Lei de Diretrizes orçamentárias

LOA – Lei Orçamentária Anual

LRF – Lei de Responsabilidade Fiscal

LSTM – *Long Short-Term Memory*

PIB – Produto Interno Bruto

RS – Rio Grande do Sul

**SUMÁRIO**

[1. Introdução 6](#_Toc68857711)

[1.1. Contextualização 6](#_Toc68857712)

[1.2. O problema proposto 7](#_Toc68857713)

[2. Coleta de dados 9](#_Toc68857714)

[3. Tratamento de dados 9](#_Toc68857715)

[4. Análise e Exploração dos Dados 12](#_Toc68857716)

[4.2. Os dados nominais 12](#_Toc68857717)

[4.3. Os dados reais 14](#_Toc68857718)

[5. Links 17](#_Toc68857719)

[6. Referências 18](#_Toc68857720)

1. Introdução
   1. Contextualização

Todo ente, seja ele município, estado federado ou a própria União possui o dever de realizar suas despesas de maneira harmônica com suas receitas, dever que, caso não cumprido, pode levar o ente a problemas fiscais como inadimplência de suas dívidas com fornecedores, atrasos dos pagamentos de aposentadorias, pensões e até mesmo dos salários do próprio funcionalismo. Além da própria crise fiscal que se origina da execução de despesas em montante superior ao das receitas, pode ainda o gestor público responder por crime de responsabilidade, conforme preceitua os artigos 10 e 11 da Lei 1.079/1950:

“CAPÍTULO VI

DOS CRIMES CONTRA A LEI ORÇAMENTÁRIA

Art. 10. São crimes de responsabilidade contra a lei orçamentária:

[...]

2 - Exceder ou transportar, sem autorização legal, as verbas do orçamento;

[...]

10 - captar recursos a título de antecipação de receita de tributo ou contribuição cujo fato gerador ainda não tenha ocorrido;

[...]

CAPÍTULO VII

DOS CRIMES CONTRA A GUARDA E LEGAL EMPREGO DOS DINHEIROS PÚBLICOS:

Art. 11. São crimes contra a guarda e legal emprego dos dinheiros públicos:

1 - ordenar despesas não autorizadas por lei ou sem observância das prescrições legais relativas às mesmas; [...]” (Brasil, 1950).

Já as Leis de Diretrizes Orçamentárias (LDOs) surgem para orientar a elaboração das Leis Orçamentárias Anuais (LOAs), estas que, por sua vez, determinam o orçamento do ente para determinado ano. Conforme a Lei de Responsabilidade Fiscal (LRF) cabem às LDOs:

“Art. 4º A lei de diretrizes orçamentárias atenderá o disposto no [§ 2o do art. 165 da Constituição](http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm#art165§2) e:

I - disporá também sobre:

a) equilíbrio entre receitas e despesas;

b) critérios e forma de limitação de empenho, a ser efetivada nas hipóteses previstas na alínea b do inciso II deste artigo, no art. 9o e no inciso II do § 1o do art. 31;

c) (VETADO)

d) (VETADO)

e) normas relativas ao controle de custos e à avaliação dos resultados dos programas financiados com recursos dos orçamentos;

f) demais condições e exigências para transferências de recursos a entidades públicas e privadas;

II - (VETADO)

III - (VETADO)

§ 1o Integrará o projeto de lei de diretrizes orçamentárias Anexo de Metas Fiscais, em que serão estabelecidas metas anuais, em valores correntes e constantes, relativas a receitas, despesas, resultados nominal e primário e montante da dívida pública, para o exercício a que se referirem e para os dois seguintes.

§ 2o O Anexo conterá, ainda:

I - avaliação do cumprimento das metas relativas ao ano anterior;

II - demonstrativo das metas anuais, instruído com memória e metodologia de cálculo que justifiquem os resultados pretendidos, comparando-as com as fixadas nos três exercícios anteriores, e evidenciando a consistência delas com as premissas e os objetivos da política econômica nacional;

III - evolução do patrimônio líquido, também nos últimos três exercícios, destacando a origem e a aplicação dos recursos obtidos com a alienação de ativos;

IV - avaliação da situação financeira e atuarial:

a) dos regimes geral de previdência social e próprio dos servidores públicos e do Fundo de Amparo ao Trabalhador;

b) dos demais fundos públicos e programas estatais de natureza atuarial;

V - demonstrativo da estimativa e compensação da renúncia de receita e da margem de expansão das despesas obrigatórias de caráter continuado.

§ 3o A lei de diretrizes orçamentárias conterá Anexo de Riscos Fiscais, onde serão avaliados os passivos contingentes e outros riscos capazes de afetar as contas públicas, informando as providências a serem tomadas, caso se concretizem.

§ 4o A mensagem que encaminhar o projeto da União apresentará, em anexo específico, os objetivos das políticas monetária, creditícia e cambial, bem como os parâmetros e as projeções para seus principais agregados e variáveis, e ainda as metas de inflação, para o exercício subsequente.” (Brasil, Lei Complementar 101/2000 – Lei de Responsabilidade Fiscal)

Percebe-se, portanto, que os gestores públicos estão sujeitos a diversos controles legais que visam a evitar que as despesas superem as receitas, estabelecendo metas, bem como a avaliação do cumprimento das metas estabelecidas em exercícios anteriores.

Sendo assim, a projeção das receitas é de suma importância para elaboração das Leis de Diretrizes Orçamentárias e das Leis Orçamentárias Anuais, bem como para a avaliação contínua pela Administração Pública da execução das despesas, com vistas a equilibrá-la à projeção de receitas, evitando *déficits* fiscais e infrações à Lei de Responsabilidade Fiscal.

* 1. O problema proposto

Para avaliar continuamente se as despesas programadas poderão efetivamente ser executadas sem gerar *déficit* fiscal, os gestores públicos necessitam de ferramentas que permitam predizer a arrecadação futura para tomarem decisões acerca da continuidade, suspensão ou até mesmo o cancelamento da execução de despesas. Geralmente as administrações tributárias realizam estimativas trimestrais e, quando necessitam de uma previsão em uma janela menor de tempo, utilizam previsões mensais. Muitas destas estimativas têm como objetivo ajustar as despesas a curto/médio prazo, bem como dar andamento a medidas burocráticas de aumento de receita, como solicitação de crédito suplementar, emissão de títulos da dívida pública e aumento de tributos, que necessitam de maior tempo para surtirem efeito, uma vez que estão quase sempre sujeitas ao processo legislativo.

Este trabalho visa encontrar métodos de predição em períodos mais curtos que os tradicionais, permitindo que gestores públicos prevejam a arrecadação em curtíssimo prazo (diariamente), permitindo tomada de decisões em caráter emergencial, principalmente no tocante à execução das despesas discricionárias, visando ao equilíbrio das contas públicas dos entes federados.

Para isso, serão utilizados dois algoritmos de predição de séries temporais (Facebook™ Prophet e redes neurais Long Short-Term Memory – LSTM) para analisar os dados diários de arrecadação tributária de três tributos estaduais do Rio Grande do Sul (Imposto sobre a Propriedade de Veículos Automotores – IPVA, Imposto sobre Transmissão Causa mortis e Doação – ITCD e Imposto sobre Circulação de Mercadorias e Serviços – ICMS) e estimar a arrecadação em datas futuras. Segundo dados do Portal da Transparência[[1]](#footnote-1) do Estado do Rio Grande do Sul, em 2020 estes três tributos foram responsáveis por 63,53% da arrecadação estadual.



Figura 1 – Valores e proporções da arrecadação por fonte de receita

Em um segundo momento, dados econômicos do estado do Rio Grande do Sul e nacionais serão agregados ao *dataset* dos dados de arrecadação de ICMS, permitindo avaliar a eficácia destas novas variáveis preditoras para a redução do erro dos modelos.

1. Coleta de dados

Os dados referentes à arrecadação diária de tributos do estado do Rio Grande do Sul foram coletados do Receita Dados[[2]](#footnote-2), portal de dados da Secretaria Estadual de Fazenda do estado do Rio Grande do Sul. Já os dados do IGP-M para atualização dos valores dos demais *datasets* a valor presente (quando necessário) foram coletados do *IpeaData*[[3]](#footnote-3).

Quando da agregação de novos dados ao *dataset* de arrecadação, os dados serão também obtidos do *IpeaData*, sendo os dados do PIB trimestral do Rio Grande do Sul e de admissões e demissões mantidos na mesma escala que os coletados, uma vez que os dados de emprego, por não se tratarem de valores monetários, não estão sujeitos a alterações com o passar do tempo e os dados do PIB trimestral do estado já estão expressos em valores reais.

Os métodos que realizam as coletas de dados estão na classe *DownloadDados*, dentro de arquivo de mesmo nome, e abrangem leitura de arquivos CSV (arrecadação mensal, PIB mensal do Brasil e dados de admissões e demissões), leitura de arquivos do Microsoft Excel (PIB trimestral do RS) e *webscrapping* (IGP-M mensal).

1. Tratamento de dados

Uma vez que os valores coletados são nominais, estes serão ajustados a valor presente com base no IGP-M, índice que melhor reflete as alterações gerais de preço de mercado, tanto para consumidores (no caso do governo quando adquire produtos e serviços), quanto para produtores (no caso dos contribuintes, que pagam os tributos – principalmente o ICMS – sobre bases de cálculo que geralmente estão correlacionadas ao valor de mercado do bem, produto ou serviço). Sendo assim, a atualização dos valores de arrecadação pelo IGP-M permite a verificação da arrecadação real tanto pela ótica dos contribuintes quanto pela ótica do governo. Caso os valores não fossem atualizados a valor presente, os modelos de aprendizado de máquina teriam que lidar, além das variações reais da arrecadação, com as variações da inflação, o que provavelmente aumentaria o erro médio das predições.

Já os valores do PIB nacional mensal foram ajustados pelo IGP-M da mesma forma que os dados de arrecadação, uma vez que da maneira que foram coletados estavam expressos em termos nominais.

Na coleta de dados do PIB nacional, foram necessários pequenos ajustes na exibição das datas (ajustes realizados nos métodos que fazem o *download* dos dados), que traziam dados do mês de outubro como sendo de janeiro. Já na coleta de dados de emprego do CAGED (admissões e demissões), além do ajuste das datas foi necessário o preenchimento de alguns campos que retornaram *NaN* (base de dados antiga), este preenchimento foi realizado com base no *download* de um terceiro *dataset* que continha os dados do saldo mensal de contratações, permitindo deduzir o valor de admissões ou demissões que estivessem sem valores válidos com base na subtração do saldo pelo campo (admissão ou demissão) que contivesse valores válidos para determinado mês.

@staticmethod  
def download\_dados\_emprego():  
 *""" Baixa os dados de emprego do IPEA Data """* # Baixa os dados de saldo de empregos para preencher campos que vierem sem valores válidos  
 url\_saldo\_caged\_antigo = 'http://www.ipeadata.gov.br/ExibeSerie.aspx?oper=exportCSVBr&serid272844966=272844966&serid272844966=272844966'  
 df\_saldo\_caged\_antigo = pd.read\_csv(url\_saldo\_caged\_antigo, sep=';')  
 df\_saldo\_caged\_antigo['Data'] = df\_saldo\_caged\_antigo['Data'].astype(str).str.replace('.', '/').replace('0000', '').str.replace('/1$', '/10')  
 df\_saldo\_caged\_antigo = df\_saldo\_caged\_antigo.drop(df\_saldo\_caged\_antigo.columns[2], axis=1)  
 df\_saldo\_caged\_antigo.columns = ['Data', 'Saldo']  
   
 # Dados antigos do CAGED (Admissões e Demissões)  
 url\_caged\_antigo = 'http://www.ipeadata.gov.br/ExibeSerie.aspx?oper=exportCSVBr&serid231410417=231410417&serid231410418=231410418'  
 df\_antigo = pd.read\_csv(url\_caged\_antigo, sep=';')  
 df\_antigo['Data'] = df\_antigo['Data'].astype(str).str.replace('.', '/').replace('0000', '').str.replace('/1$', '/10')  
 df\_antigo = df\_antigo.drop(df\_antigo.columns[3], axis=1)  
 df\_antigo.columns = ['Data', 'Admissoes', 'Demissoes']  
 datas\_nan = df\_antigo['Data'][np.isnan(df\_antigo['Demissoes'])==True] # Datas cujo valor de demissões é NaN  
 saldos\_nan = df\_saldo\_caged\_antigo[df\_saldo\_caged\_antigo['Data'].isin(pd.DataFrame(datas\_nan)['Data'])] # Saldos em datas cujo valor de demissões é NaN  
 df\_antigo['Demissoes'][np.isnan(df\_antigo['Demissoes'])==True] = df\_antigo['Admissoes'][np.isnan(df\_antigo['Demissoes'])==True]-saldos\_nan['Saldo']   
 df\_antigo['Demissoes'] = df\_antigo['Demissoes'].astype(int)  
   
 # Dados novos do CAGED (Admissões e Demissões)  
 url\_caged\_novo = 'http://www.ipeadata.gov.br/ExibeSerie.aspx?oper=exportCSVBr&serid2096725334=2096725334&serid2096725335=2096725335'  
 df\_novo = pd.read\_csv(url\_caged\_novo, sep=';')  
 df\_novo['Data'] = df\_novo['Data'].astype(str).str.replace('.', '/').replace('0000', '').str.replace('/1$', '/10')  
 df\_novo = df\_novo.drop(df\_novo.columns[3], axis=1)  
 df\_novo.columns = ['Data', 'Admissoes', 'Demissoes']  
   
 df\_final = df\_novo.iloc[::-1].append(df\_antigo.iloc[::-1]).reset\_index(drop=True)  
   
 for i in range(0, len(df\_final)):  
 ano = df\_final['Data'].str.split('/')[i][0]  
 mes = df\_final['Data'].str.split('/')[i][1]  
 df\_final['Data'].loc[i] = mes+'/'+ano  
   
 return df\_final

Figura 2 – Download e tratamento de dados do CAGED

Após o tratamento dos dados antigos e novos foi realizada a junção dos dois *dataframes*, retornando um único Pandas *dataframe* como resultado.

Já as atualizações dos valores a valor presente são realizadas pelos métodos *corrige\_inflacao* e *corrige\_inflacao\_pib* da classe *CorrigeValores* do arquivo *Util.py*, que verificam o período a que se referem os dados do PIB e da arrecadação e multiplicam pelo número índice, trazendo os valores a valor presente da última data disponível da série histórica do IGP-M.

@staticmethod  
def corrige\_inflacao(pd\_tributo, pd\_inflacao):  
 *""" Dado um dataframe com os números-índice para correção e outro dataframe com os valores de tributo,  
 corrige estes trazendo-os ao valor presente da última data disponível no dataframe de valores de tributo."""* ultimo\_mes\_inflacao = pd\_inflacao.loc[0, 'Mes/Ano']  
 ultimo\_numero\_indice = pd\_inflacao.loc[0, 'Numero Indice']  
  
 for indice, linha in pd\_tributo.iterrows():  
 mes\_atual = linha['Data'].strftime("%m/%Y")  
  
 ''' Nos meses anteriores à última data publicada para a inflação, o valor do tributo será  
 multiplicado pela razão entre o último número-índice e o número índice correspondente ao mês seguinte ao do  
 tributo. '''  
 if datetime.strptime(mes\_atual, '%m/%Y').date() < datetime.strptime(ultimo\_mes\_inflacao, '%m/%Y').date():  
 numero\_indice\_atual = pd\_inflacao[pd\_inflacao['Mes/Ano'] == mes\_atual]['Numero Indice'].item()  
 elif datetime.strptime(mes\_atual, '%m/%Y').date() >= datetime.strptime(ultimo\_mes\_inflacao, '%m/%Y').date():  
 numero\_indice\_atual = ultimo\_numero\_indice  
  
 pd\_tributo.loc[indice, 'Valor'] = pd\_tributo.loc[indice, 'Valor'] \* ultimo\_numero\_indice / numero\_indice\_atual  
  
 return pd\_tributo  
  
@staticmethod  
def corrige\_inflacao\_pib(pd\_pib, pd\_inflacao):  
 *""" Dado um dataframe com os números-índice para correção e outro dataframe com os valores do PIB Nominal,  
 corrige estes trazendo-os ao valor presente da última data disponível no dataframe de valores de PIB."""* ultimo\_mes\_inflacao = pd\_inflacao.loc[0, 'Mes/Ano']  
 ultimo\_numero\_indice = pd\_inflacao.loc[0, 'Numero Indice']  
  
 for indice, linha in pd\_pib.iterrows():  
 mes\_atual = linha['Data']  
  
 ''' Nos meses anteriores à última data publicada para a inflação, o valor do tributo será  
 multiplicado pela razão entre o último número-índice e o número índice correspondente ao mês seguinte ao do  
 tributo. '''  
 if datetime.strptime(mes\_atual, '%m/%Y').date() < datetime.strptime(ultimo\_mes\_inflacao, '%m/%Y').date():  
 numero\_indice\_atual = pd\_inflacao[pd\_inflacao['Mes/Ano'] == mes\_atual]['Numero Indice'].item()  
 elif datetime.strptime(mes\_atual, '%m/%Y').date() >= datetime.strptime(ultimo\_mes\_inflacao, '%m/%Y').date():  
 numero\_indice\_atual = ultimo\_numero\_indice  
  
 pd\_pib.loc[indice, 'PIB'] = pd\_pib.loc[indice, 'PIB'] \* ultimo\_numero\_indice / numero\_indice\_atual  
  
 return pd\_pib

Figura 3 – Métodos de atualização de valores a valor presente

1. Análise e Exploração dos Dados
   1. Os dados nominais

No primeiro momento serão analisados os dados nominais dos tributos estaduais, com o objetivo de comparar com os dados reais, estes que serão efetivamente utilizados para treinamento e predição nos modelos. Para tanto, seguem abaixo os dados descritivos das séries temporais da arrecadação nominal dos três tributos estaduais:

|  |  |  |  |
| --- | --- | --- | --- |
| **Medida** | **ICMS** | **IPVA** | **ITCD** |
| **Quantidade** | 3.167,00 | 3.119,00 | 3.097,00 |
| **Média** | 104.326.511,37 | 9.263.437,50 | 1.709.255,15 |
| **Desvio Padrão (σ)** | 145.973.777,35 | 18.156.530,23 | 2.582.832,55 |
| **Mínimo** | -153.507,11 | -167,22 | -1.375,75 |
| **25%** | 16.454.912,34 | 1.829.823,52 | 612.259,54 |
| **50%** | 39.408.950,31 | 4.662.392,00 | 1.184.296,94 |
| **75%** | 145.005.742,05 | 9.745.511,49 | 2.011.134,15 |
| **Máximo** | 1.164.392.552,98 | 294.336.676,91 | 62.862.554,05 |

Figura 4 – Descrição da arrecadação nominal dos tributos estaduais

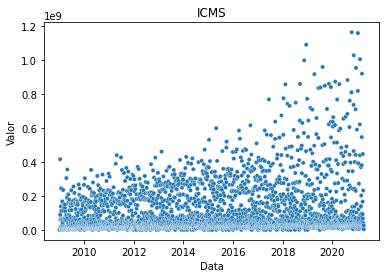


Figura 5 – Arrecadação nominal do ICMS do RS

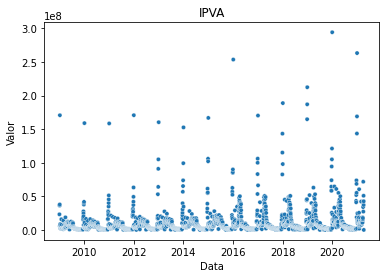


Figura 6 – Arrecadação nominal do IPVA do RS

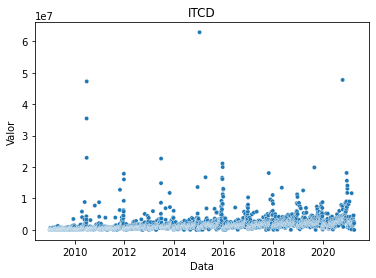


Figura 7 – Arrecadação nominal do ITCD do RS

Denota-se que o ICMS, além de possuir um maior valor médio nominal de arrecadação, é o tributo que possui maior variância (σ²), o que certamente aumentará o valor do erro médio dos modelos que serão desenvolvidos. Outro aspecto a ser comentado é a existência de valores negativos, mesmo que em menor valor absoluto, nas séries. A existência de arrecadação negativa se deve ao fato de que dos valores computados já estão deduzidas as restituições e benefícios fiscais, sendo os meses com valores negativos, portanto, aqueles em que as restituições e/ou benefícios fiscais excederam a arrecadação bruta. Observando os gráficos, percebe-se visualmente que todas as séries possuem tendência de crescimento nominal, o que já era esperado por conta da inflação. Outra característica interessante é a sazonalidade da arrecadação do IPVA nos primeiros meses do ano devido aos pagamentos à vista e das primeiras parcelas. No ITCD percebem-se alguns picos de arrecadação, mas não fica visualmente claro se há alguma sazonalidade.

* 1. Os dados reais

Atualizando os valores a valor presente, fica observado o aumento da magnitude dos novos números. Como a variância é, como também dito na análise dos valores nominais, fator que impacta diretamente o desempenho do modelo, o aumento desta medida, decorrido da atualização, aumentará o desafio para que os modelos prevejam a arrecadação com grau aceitável de erro.

|  |  |  |  |
| --- | --- | --- | --- |
| **Medida** | **ICMS** | **IPVA** | **ITCD** |
| **Quantidade** | 3.167,00 | 3.119,00 | 3.097,00 |
| **Média** | 176.546.450,23 | 15.716.591,77 | 2.767.067,95 |
| **Desvio Padrão (σ)** | 225.599.315,66 | 29.589.472,15 | 4.523.596,03 |
| **Mínimo** | -209.539,42 | -356,90 | -3.004,87 |
| **25%** | 27.993.106,87 | 3.130.045,03 | 1.216.229,59 |
| **50%** | 68.630.690,51 | 8.435.535,40 | 2.003.726,62 |
| **75%** | 259.310.821,37 | 17.177.296,69 | 3.063.960,79 |
| **Máximo** | 1.560.701.121,06 | 421.701.485,08 | 113.094.527,20 |

Figura 8 – Descrição da arrecadação real dos tributos estaduais

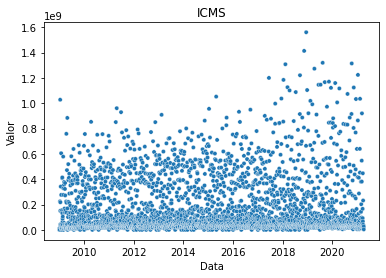


Figura 9 – Arrecadação real do ICMS do RS

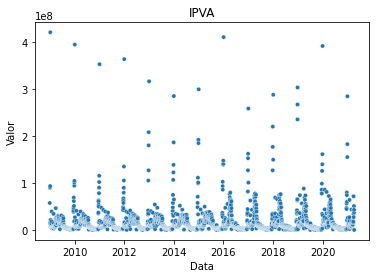


Figura 10 – Arrecadação real do IPVA do RS

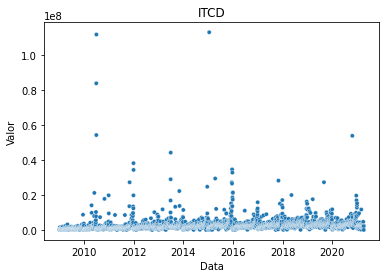


Figura 11 – Arrecadação real do ITCD do RS

Após a atualização dos valores, percebe-se que a tendência de aumento ficou reduzida em todas as séries quando em comparação com os valores nominais. Com a inflação eliminada dos dados, parte-se para uma análise mais aprofundada destes.

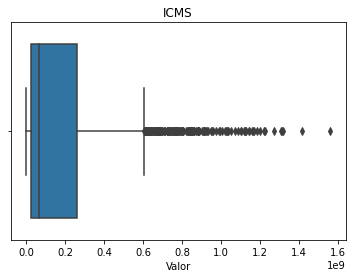


Figura 12 – Boxplot da arrecadação real do ICMS do RS

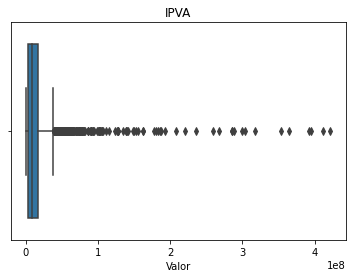


Figura 13 – Boxplot da arrecadação real do IPVA do RS

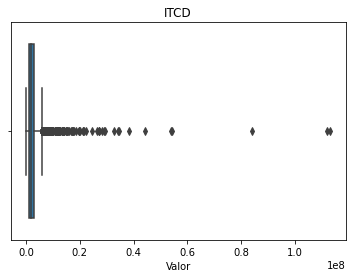


Figura 14 – Boxplot da arrecadação real do ITCD do RS

Pela amplitude dos *boxlots* é possível verificar a variância das séries temporais da arrecadação dos tributos. O ICMS, com variância bem superior ao demais, é o que possui maior intervalo interquartílico, seguido pelo IPVA e ITCD, com menores variâncias. O mesmo ocorre com os *outliers*, estando presentes em maior quantidade no ICMS do que nos demais tributos. Sendo valores que têm grande potencial de atrapalharem as predições, os *outliers* devem passar por tratamento para que suas importâncias sejam reduzidas. Além disso, cada modelo preditivo comporta-se de maneira distinta em relação aos tratamentos, devendo ser escolhido o melhor método para cada modelo.

1. Links
2. Referências

1. Disponível em: <http://www.transparencia.rs.gov.br/QvAJAXZfc/opendoc.htm?document=Transparencia.qvw &host= QVS%40QLVPRO06&anonymous=true>. Acesso em 7 de abril de 2021. [↑](#footnote-ref-1)
2. Disponível em: <http://receitadados.fazenda.rs.gov.br/Arquivos/Arrecadação Diária.csv >. Acesso em 8 de abril de 2021. [↑](#footnote-ref-2)
3. Disponível em: http://www.ipeadata.gov.br/ExibeSerie.aspx?stub=1&serid37796=37796&serid36482=36482. Acesso em 8 de abril de 2021. [↑](#footnote-ref-3)