

Analysis of the electoral accountability of the 2018 and 2020 elections

*Diego Cunha*¹

¹ Federal University of Bahia, Salvador, Bahia, Brazil

Abbreviated abstract: The electoral process has several stages that are very important for the execution of the entire process, which are the stages of voting, totalization, dissemination of results, electoral registration, candidacies, electoral logistics and accountability. This work aims to analyze one of the phases of the electoral process, the accountability phase of the 2018 federal election and the 2020 municipal election, in order to verify how much revenue and expenditure were obtained by parties and candidates in the northeast region.



Problem, Data, Previous Works

Analyzing expenses and income expenses in the 2018 and 2020 elections in the Northeast region as well as trying to understand how much the amount spent on a candidacy can result in a candidate's results in an election.

Data obtained from the data repository of the Superior Electoral Court.

- https://www.tse.jus.br/hotsites/pesquisas-eleitorais/prestacao_contas.html

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 282007 entries, 0 to 282006
Data columns (total 22 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   ANO_ELEICAO           282007 non-null int64
1   SQ_PRESTADOR_CONTAS  282007 non-null int64
2   SG_UF                 282007 non-null object
3   SG_UF                 282007 non-null object
4   NM_UF                 282007 non-null object
5   CD_CARGO              282007 non-null int64
6   DS_CARGO              282007 non-null object
7   SQ_CANDIDATO          282007 non-null int64
8   NR_CANDIDATO          282007 non-null int64
9   NM_CANDIDATO          282007 non-null object
10  NR_PARTIDO            282007 non-null int64
11  SG_PARTIDO            282007 non-null object
12  NM_PARTIDO            282007 non-null object
13  CD_TIPO_FORNECEDOR    282007 non-null object
14  CD_CNAE_FORNECEDOR    282007 non-null int64
15  DS_CNAE_FORNECEDOR    282007 non-null object
16  CD_ORIGEM_DESPESA     282007 non-null int64
17  DS_ORIGEM_DESPESA     282007 non-null object
18  DT_DESPESA           282007 non-null object
19  VR_DESPESA_CONTRATADA 282007 non-null float64
20  NR_MES                282007 non-null int64
21  ELEITO                282007 non-null int64
dtypes: float64(1), int64(10), object(11)
memory usage: 49.5+ MB
```

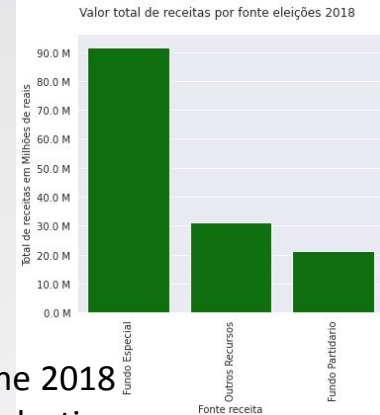
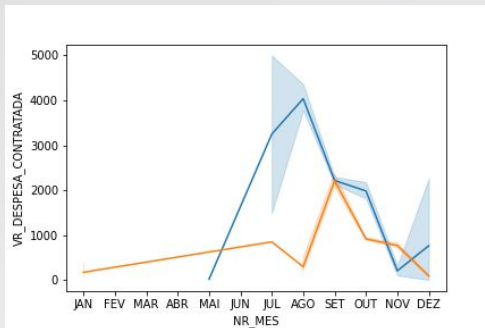
```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 60326 entries, 0 to 60325
Data columns (total 20 columns):
#   Column                Non-Null Count  Dtype
---  ---                ---
0   ANO_ELEICAO           60326 non-null int64
1   SQ_PRESTADOR_CONTAS  60326 non-null int64
2   SG_UF                 60326 non-null object
3   SG_UF                 60326 non-null object
4   NM_UF                 60326 non-null object
5   CD_CARGO              60326 non-null int64
6   DS_CARGO              60326 non-null object
7   SQ_CANDIDATO          60326 non-null int64
8   NR_CANDIDATO          60326 non-null int64
9   NM_CANDIDATO          60326 non-null object
10  NR_PARTIDO            60326 non-null int64
11  SG_PARTIDO            60326 non-null object
12  NM_PARTIDO            60326 non-null object
13  CD_FONTE_RECEITA      60326 non-null int64
14  DS_FONTE_RECEITA      60326 non-null object
15  CD_ORIGEM_RECEITA     60326 non-null int64
16  DS_ORIGEM_RECEITA     60326 non-null object
17  DT_RECEITA            60326 non-null object
18  VR_RECEITA            60326 non-null float64
19  ELEITO                60326 non-null int64
dtypes: float64(1), int64(9), object(10)
memory usage: 9.7+ MB
```

Data structured in rendering of accounts by election and by state of Brazil. All of this data has been merged into an expense dataset and another income dataset.

Methods

Initially, the files were unified into a single file referring to the election year. It was also necessary to make adjustments in some fields to prepare them for the next steps of data analysis.

CD_ORIGEM_DESPESA	DS_ORIGEM_DESPESA	VR_DESPESA_CONTRATADA
20140000	Publicidade por materiais impressos	R\$ 171.604.322,58
20120000	Serviços prestados por terceiros	R\$ 120.141.420,65
20220000	Produção de programas de rádio, televisão ou v...	R\$ 80.849.715,21
20110000	Publicidade por adesivos	R\$ 65.076.098,04
20800000	Atividades de militância e mobilização de rua	R\$ 63.321.759,61
20600000	Cessão ou locação de veículos	R\$ 36.167.531,95
20240000	Doações financeiras a outros candidatos/partidos	R\$ 35.187.139,20
20100000	Combustíveis e lubrificantes	R\$ 30.538.992,28
20010000	Despesas com pessoal	R\$ 26.568.802,09
20270000	Diversas a especificar	R\$ 16.877.153,77



For the analysis of the results we have four data sets available, two from the 2018 election and two from the 2020 election. Initially we know that the 2018 election was a federal election where the election for the positions of President, Governor, Senator, took place. Federal Deputy and State Deputy, and the 2020 election was a municipal election that took place for the positions of Mayor and Councilor.

Results and Conclusions

We can conclude that the execution of the current election process requires a large investment of public money, which makes us wonder if all this investment is really necessary. But I also believe that these funds are important to keep the electoral process in full operation, only needing an analysis for a better distribution of its resources, because as we can see the division of the Electoral Fund and the Party Fund always ends up favoring the largest parties politicians which favors to keep them always with the greatest representation in the congress and hardly changing this scenario.

For future work

- Further analysis
- Application of machine learning models
- detecting fraud in the rendering of accounts

