

Análise dos dados de reembolso dos senadores - 2021

Willer Gomes Junior

```
In [1]: # Importando bibliotecas
import pandas as pd
import requests

# Configurando o matplotlib
%matplotlib inline

In [2]: # Carregar o arquivo csv

df = pd.read_csv('despesa_ceaps_2021.csv', sep=';', encoding = "cp1252", skiprows=1, decimal=",")
df.head()
```

Out[2]:

	ANO	MES	SENADOR	TIPO_DESPESA	CNPJ_CPF	FORNECEDOR	DOCUMENTO	DAT
0	2021	1	ACIR GURGACZ	Aluguel de imóveis para escritório político, c...	062.135.728-64	FERNANDO WALDEIR PACINI e ANA LUCIA DA SILVA S...	01/2021	05/01/202
1	2021	1	ACIR GURGACZ	Aluguel de imóveis para escritório político, c...	004.948.028-63	GILBERTO PISELO DO NASCIMENTO	001/21	06/01/202
2	2021	1	ACIR GURGACZ	Divulgação da atividade parlamentar	26.320.603/0001-64	INFORMANAHORA	10	25/01/202
3	2021	1	ACIR GURGACZ	Divulgação da atividade parlamentar	13.659.201/0001-47	LINHA PURPURA FOTO E VIDEO LTDA	44	07/01/202
4	2021	1	ACIR GURGACZ	Passagens aéreas, aquáticas e terrestres nacio...	17.872.428/0001-27	AMERICA VIAGENS E TURISMO LTDA - ME	VVIEWKR	21/01/202

```
In [3]: # verificar o tamanho e o tipo dos dados
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2328 entries, 0 to 2327
Data columns (total 11 columns):
ANO                2328 non-null int64
MES                2328 non-null int64
SENADOR            2328 non-null object
TIPO_DESPESA       2328 non-null object
CNPJ_CPF           2328 non-null object
FORNECEDOR         2328 non-null object
DOCUMENTO          2209 non-null object
DATA               2328 non-null object
DETALHAMENTO       1171 non-null object
VALOR_REEMBOLSADO  2328 non-null float64
COD_DOCUMENTO      2328 non-null int64
dtypes: float64(1), int64(3), object(7)
memory usage: 136.5+ KB
```

```
In [4]: df.describe()
```

Out[4]:

	ANO	MES	VALOR_REEMBOLSADO	COD_DOCUMENTO
count	2328.0	2328.000000	2328.000000	2.328000e+03
mean	2021.0	1.753007	1501.106491	2.155868e+06
std	0.0	0.707556	2846.116613	9.086718e+02
min	2021.0	1.000000	0.020000	2.154172e+06
25%	2021.0	1.000000	180.000000	2.155128e+06
50%	2021.0	2.000000	515.940000	2.155916e+06
75%	2021.0	2.000000	1665.070000	2.156605e+06
max	2021.0	3.000000	35000.000000	2.157603e+06

```
In [5]: # qual o total de reembolso
df['VALOR_REEMBOLSADO'].sum()
```

Out[5]: 3494575.9099999997

```
In [6]: # Quantos reembolsos foram solicitados por cada senador
df['SENADOR'].value_counts()
```

Out[6]:

CARLOS VIANA	111
JAQUES WAGNER	86
TELMÁRIO MOTA	82
MAJOR OLIMPIO	82
LUIS CARLOS HEINZE	78
...	
NEY SUASSUNA	3
STYVENSON VALENTIM	2
OMAR AZIZ	2
TASSO JEREISSATI	1
NAILDE PANTA	1

Name: SENADOR, Length: 75, dtype: int64

```
In [7]: # qual o total de reembolso por senador
df.groupby('SENADOR')['VALOR_REEMBOLSADO'].sum()
```

```
Out[7]: SENADOR
ACIR GURGACZ          66658.01
ALESSANDRO VIEIRA     48338.21
ANGELO CORONEL        44296.11
ANTONIO ANASTASIA     48141.14
CARLOS FÁVARO         58171.81
...
VENEZIANO VITAL DO RÊGO 35818.15
WELLINGTON FAGUNDES    15051.14
WEVERTON ROCHA        59713.48
ZENAIDE MAIA          35615.24
ZEQUINHA MARINHO      88828.98
Name: VALOR_REEMBOLSADO, Length: 75, dtype: float64
```

```
In [8]: # Ordenando o valor total de reembolso do maior para o menor.
df.groupby('SENADOR')['VALOR_REEMBOLSADO'].sum().sort_values(ascending=False)
```

```
Out[8]: SENADOR
TELMÁRIO MOTA          122736.19
MAILZA GOMES           122063.35
MARCIO BITTAR           99010.74
FERNANDO COLLOR        96000.00
CIRO NOGUEIRA          94654.76
...
NEY SUASSUNA           9501.33
OMAR AZIZ              6766.86
VANDERLAN CARDOSO      4324.55
STYVENSON VALENTIM     3153.27
NAILDE PANTA           1650.62
Name: VALOR_REEMBOLSADO, Length: 75, dtype: float64
```

```
In [9]: # Quais sao os maiores valores de reembolso
df.nlargest(5, 'VALOR_REEMBOLSADO')
```

Out[9]:

	ANO	MES	SENADOR	TIPO_DESPESA	CNPJ_CPF	FORNECEDOR	DOCUMENTO	D
1523	2021	2	MARCIO BITTAR	Contratação de consultorias, assessorias, pesq...	02.456.658/0001-38	IDÉIAS ESTRATÉGIA, PLANEJAMENTO E PROJETO LTDA	110	12/02/2
1528	2021	3	MARCIO BITTAR	Contratação de consultorias, assessorias, pesq...	02.456.658/0001-38	IDÉIAS ESTRATÉGIA, PLANEJAMENTO E PROJETO LTDA	111	09/03/2
1646	2021	3	NELSINHO TRAD	Contratação de consultorias, assessorias, pesq...	37.195.146/0001-40	Home Mix Produção e Assessoria em Radiodifusão...	00007115	23/03/2
490	2021	1	EDUARDO BRAGA	Contratação de consultorias, assessorias, pesq...	33.055.400/0001-53	L COELHO SERRA	21	05/01/2
491	2021	2	EDUARDO BRAGA	Contratação de consultorias, assessorias, pesq...	33.055.400/0001-53	L COELHO SERRA	24	03/02/2

```
In [10]: df.nsmallest(5, 'VALOR_REEMBOLSADO')
```

Out[10]:

	ANO	MES	SENADOR	TIPO_DESPESA	CNPJ_CPF	FORNECEDOR	DOCUMENTO	D
2158	2021	1	VANDERLAN CARDOSO	Aluguel de imóveis para escritório político, c...	40.432.544/0436-28	NET CLARO S.A	NaN	10/01/2
666	2021	1	FABIANO CONTARATO	Contratação de consultorias, assessorias, pesq...	28.133.619/0001-93	S/A A GAZETA	0073489	06/01/2
227	2021	1	CARLOS VIANA	Locomoção, hospedagem, alimentação, combustíve...	17.895.646/0001-87	UBER DO BRASIL TECNOLOGIA LTDA	NaN	31/01/2
682	2021	1	FABIANO CONTARATO	Locomoção, hospedagem, alimentação, combustíve...	17.895.646/0001-87	UBER DO BRASIL TECNOLOGIA LTDA.	NaN	19/01/2
675	2021	1	FABIANO CONTARATO	Locomoção, hospedagem, alimentação, combustíve...	17.895.646/0001-87	UBER DO BRASIL TECNOLOGIA LTDA.	NaN	08/01/2

```
In [11]: # quantidade de reembolso por tipo de despesa
df['TIPO_DESPESA'].value_counts()
```

```
Out[11]: Aluguel de imóveis para escritório político, compreendendo despesas concernen
tes a eles.
656
Locomoção, hospedagem, alimentação, combustíveis e lubrificantes
641
Passagens aéreas, aquáticas e terrestres nacionais
423
Aquisição de material de consumo para uso no escritório político, inclusive a
quisição ou locação de software, despesas postais, aquisição de publicações,
locação de móveis e de equipamentos.      235
Divulgação da atividade parlamentar
198
Contratação de consultorias, assessorias, pesquisas, trabalhos técnicos e out
ros serviços de apoio ao exercício do mandato parlamentar
169
Serviços de Segurança Privada
6
Name: TIPO_DESPESA, dtype: int64
```

```
In [12]: # quanto foi desembolsado por cada tipo de despesa
df.groupby('TIPO_DESPESA')['VALOR_REEMBOLSADO'].sum().sort_values(ascending=F
alse)
```

```
Out[12]: TIPO_DESPESA
Contratação de consultorias, assessorias, pesquisas, trabalhos técnicos e out
ros serviços de apoio ao exercício do mandato parlamentar
955474.38
Aluguel de imóveis para escritório político, compreendendo despesas concernen
tes a eles.
715743.48
Locomoção, hospedagem, alimentação, combustíveis e lubrificantes
598429.14
Passagens aéreas, aquáticas e terrestres nacionais
571463.47
Divulgação da atividade parlamentar
470466.02
Aquisição de material de consumo para uso no escritório político, inclusive a
quisição ou locação de software, despesas postais, aquisição de publicações,
locação de móveis e de equipamentos.      176741.90
Serviços de Segurança Privada
6257.52
Name: VALOR_REEMBOLSADO, dtype: float64
```

```
In [ ]:
```