

```
In [ ]: #importando as bibliotecas
import pandas as pd
import os
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [ ]: # Exibindo os dados de cada ano
files = [file for file in os.listdir(r'C:/Users/kpaul/OneDrive/Data Analytics/Python/Project/EGP/dados')]
for file in files:
    print(file)

finbra_2017.csv
finbra_2018.csv
finbra_2019.csv
finbra_2020.csv
finbra_2021.csv
```

```
In [ ]: #Juntando todos os dados para criar um data base unico
path = r'C:/Users/kpaul/OneDrive/Data Analytics/Python/Project/EGP/dados'
all_data = pd.DataFrame() # criando um database vazio

for file in files:
    current_data = pd.read_csv(path+'/'+file, sep=';', encoding='latin1')
    all_data = pd.concat([all_data , current_data])
print(all_data)
```

	Instituição	Cod.IBGE	UF	População	\
0	Governo do Estado do Mato Grosso	51	MT	3305531	
1	Governo do Estado do Mato Grosso	51	MT	3305531	
2	Governo do Estado do Mato Grosso	51	MT	3305531	
3	Governo do Estado do Mato Grosso	51	MT	3305531	
4	Governo do Estado do Mato Grosso	51	MT	3305531	
...	
14770	Governo do Estado de Roraima	14	RR	631181	
14771	Governo do Estado de Roraima	14	RR	631181	
14772	Governo do Estado de Roraima	14	RR	631181	
14773	Governo do Estado de Roraima	14	RR	631181	
14774	Governo do Estado de Roraima	14	RR	631181	
					Coluna \
0	Despesas Empenhadas				
1	Despesas Empenhadas				
2	Despesas Empenhadas				
3	Despesas Empenhadas				
4	Despesas Empenhadas				
...	...				
14770	Inscrição de Restos a Pagar Processados				
14771	Inscrição de Restos a Pagar Processados				
14772	Inscrição de Restos a Pagar Processados				
14773	Inscrição de Restos a Pagar Processados				
14774	Inscrição de Restos a Pagar Processados				
					Conta Identificador da Conta \
0	Despesas Exceto Intraorçamentárias		siconfi-cor_TotalDespesas		
1	01 - Legislativa		siconfi-cor_TotalDespesas		
2	01.031 - Ação Legislativa		siconfi-cor_TotalDespesas		
3	01.032 - Controle Externo		siconfi-cor_TotalDespesas		
4	01.122 - Administração Geral		siconfi-cor_TotalDespesas		
...		
14770	26.782 - Transporte Rodoviário		siconfi-cor_TotalDespesas		
14771	27 - Desporto e Lazer		siconfi-cor_TotalDespesas		
14772	27.811 - Desporto de Rendimento		siconfi-cor_TotalDespesas		
14773	27.812 - Desporto Comunitário		siconfi-cor_TotalDespesas		
14774	Despesas Intraorçamentárias		siconfi-cor_TotalDespesas		
	Valor Ano				
0	16808633403	2017			
1	687932086,3	2017			
2	422143,98	2017			
3	47728708,12	2017			
4	582546800,9	2017			
...			
14770	129834,1	2021			
14771	195426,26	2021			
14772	66825	2021			
14773	128601,26	2021			
14774	5821304,29	2021			

[74379 rows x 9 columns]

In []: all_data.to_csv(path + "/all_data.csv", index = False)

In []: #Exibindo as primeiras Linhas dos dados
all_data.head()

	Instituição	Cod.IBGE	UF	População	Coluna	Conta	Identificador da Conta	
0	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	Despesas Exceto Intraorçamentárias	siconfis.cor_TotalDespesas	16808
1	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01 - Legislativa	siconfis.cor_TotalDespesas	6879
2	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.031 - Ação Legislativa	siconfis.cor_TotalDespesas	42
3	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.032 - Controle Externo	siconfis.cor_TotalDespesas	4772
4	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.122 - Administração Geral	siconfis.cor_TotalDespesas	5825

In []: `all_data['Valor'] = all_data['Valor'].str.replace(',', '.'). #Substituindo a vírgula
all_data['Valor'] = all_data['Valor'].astype(float) #Convertendo a coluna valor em float
all_data['Cod.IBGE'] = all_data['Cod.IBGE'].astype(str) # Convertendo a coluna Cod.IBGE para string
all_data.rename(columns={'Coluna': 'Tipo de Despesa'}, inplace = True) #Renomeando as colunas`

In []: `all_data.shape #verificando o numero de linhas e de colunas do dataframe`

Out[]: `(74379, 9)`

In []: `all_data`

Out[]:

	Instituição	Cod.IBGE	UF	População	Tipo de Despesa	Conta	Identificador da Conta
0	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	Despesas Intraorçamentárias	siconfin-cor_TotalDespesas
1	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01 - Legislativa	siconfin-cor_TotalDespesas
2	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.031 - Ação Legislativa	siconfin-cor_TotalDespesas
3	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.032 - Controle Externo	siconfin-cor_TotalDespesas
4	Governo do Estado do Mato Grosso	51	MT	3305531	Despesas Empenhadas	01.122 - Administração Geral	siconfin-cor_TotalDespesas
...
14770	Governo do Estado de Roraima	14	RR	631181	Inscrição de Restos a Pagar Processados	26.782 - Transporte Rodoviário	siconfin-cor_TotalDespesas
14771	Governo do Estado de Roraima	14	RR	631181	Inscrição de Restos a Pagar Processados	27 - Desporto e Lazer	siconfin-cor_TotalDespesas
14772	Governo do Estado de Roraima	14	RR	631181	Inscrição de Restos a Pagar Processados	27.811 - Desporto de Rendimento	siconfin-cor_TotalDespesas
14773	Governo do Estado de Roraima	14	RR	631181	Inscrição de Restos a Pagar Processados	27.812 - Desporto Comunitário	siconfin-cor_TotalDespesas
14774	Governo do Estado de Roraima	14	RR	631181	Inscrição de Restos a Pagar Processados	Despesas Intraorçamentárias	siconfin-cor_TotalDespesas

74379 rows × 9 columns



In []:

```
#agora vamos pegar só as linhas que contem a conta 06 - Segurança Pública
df_filtrado = all_data.query("Conta == '06 - Segurança Pública'")
df_filtrado.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 666 entries, 26 to 14723
Data columns (total 9 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   Instituição      666 non-null    object  
 1   Cod.IBGE         666 non-null    object  
 2   UF               666 non-null    object  
 3   População        666 non-null    int64   
 4   Tipo de Despesa  666 non-null    object  
 5   Conta             666 non-null    object  
 6   Identificador da Conta 666 non-null    object  
 7   Valor             666 non-null    float64 
 8   Ano               666 non-null    int64   
dtypes: float64(1), int64(2), object(6)
memory usage: 52.0+ KB
```

In []: df_filtrado.sample(20)

Out[]:

	Instituição	Cod.IBGE	UF	População	Tipo de Despesa	Conta	Identificador da Conta	
2228	Governo do Estado de Pernambuco	26	PE	9410336	Despesas Pagas	06 - Segurança Pública	siconfi-cor_TotalDespesas	2.71264
13502	Governo do Estado de Roraima	14	RR	605761	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	3.78960
13151	Governo do Estado do Pará	15	PA	8690745	Despesas Empenhadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	3.27181
11993	Governo do Estado do Rio de Janeiro	33	RJ	17366189	Inscrição de Restos a Pagar Processados	06 - Segurança Pública	siconfi-cor_TotalDespesas	1.48890
8178	Governo do Estado do Pará	15	PA	8513497	Despesas Empenhadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	2.85185
2492	Governo do Estado do Pará	15	PA	8272724	Despesas Empenhadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	2.71163
5260	Governo do Estado do Maranhão	21	MA	6954036	Despesas Empenhadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	1.84964
1828	Governo do Estado de Rondônia	11	RO	1787279	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	7.79523
5210	Governo do Estado do Amapá	16	AP	829494	Inscrição de Restos a Pagar Processados	06 - Segurança Pública	siconfi-cor_TotalDespesas	1.69896
6518	Governo do Estado do Ceará	23	CE	9075649	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	2.69889
3850	Governo do Estado do Rio Grande do Sul	43	RS	11377239	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	5.29086
2631	Governo do Estado da Paraíba	25	PB	3996496	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	1.33036
1720	Governo do Estado de Santa Catarina	42	SC	7164788	Despesas Liquidadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	2.35131
13600	Governo do Estado de Roraima	14	RR	605761	Despesas Pagas	06 - Segurança Pública	siconfi-cor_TotalDespesas	3.78324
9794	Governo do Estado do Paraná	41	PR	11348937	Despesas Empenhadas	06 - Segurança Pública	siconfi-cor_TotalDespesas	3.86201

	Instituição	Cod.IBGE	UF	População	Tipo de Despesa	Conta	Identificador da Conta	
5074	Governo do Estado do Piauí	22	PI	3212180	Inscrição de Restos a Pagar Não Processados	06 - Segurança Pública	siconfisicor_TotalDespesas	1.49149
6269	Governo do Estado do Ceará	23	CE	8963663	Inscrição de Restos a Pagar Não Processados	06 - Segurança Pública	siconfisicor_TotalDespesas	4.75012
13840	Governo do Estado de Sergipe	28	SE	2265779	Despesas Empenhadas	06 - Segurança Pública	siconfisicor_TotalDespesas	8.92986
3197	Governo do Estado do Paraná	41	PR	11516840	Inscrição de Restos a Pagar Processados	06 - Segurança Pública	siconfisicor_TotalDespesas	5.13838
668	Governo do Estado de Santa Catarina	42	SC	7075494	Despesas Liquidadas	06 - Segurança Pública	siconfisicor_TotalDespesas	2.25521

In []: # verificando se há dados nulos em uma das colunas
`(df_filtrado.isnull().sum()/len(df_filtrado)*100).sort_values(ascending=False)`

Out[]:

Instituição	0.0
Cod.IBGE	0.0
UF	0.0
População	0.0
Tipo de Despesa	0.0
Conta	0.0
Identificador da Conta	0.0
Valor	0.0
Ano	0.0

dtype: float64

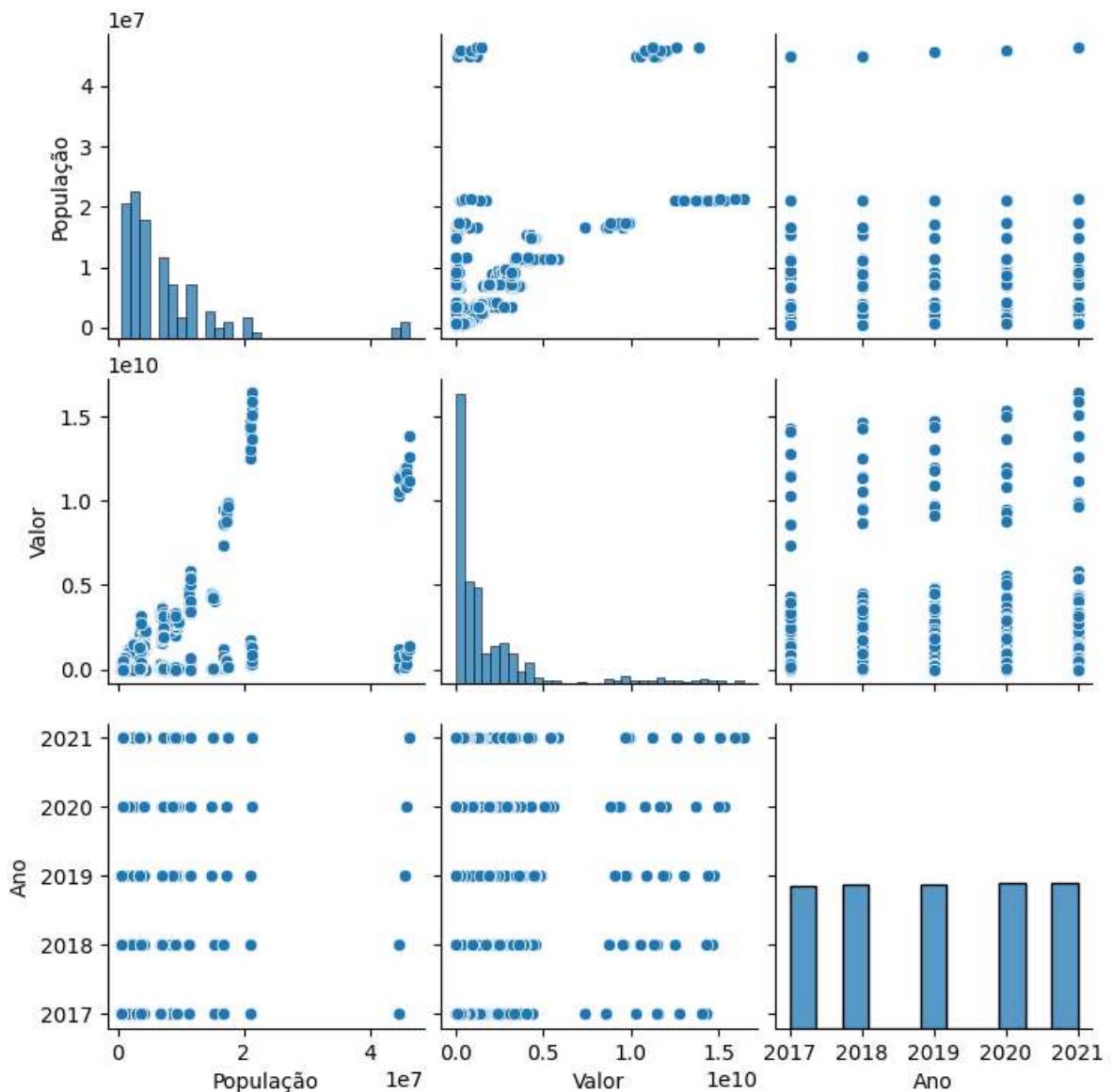
In []: #analise descritiva no das variaveis quantitativas

Out[]:

	População	Valor	Ano
count	6.660000e+02	6.660000e+02	666.000000
mean	7.740456e+06	1.848780e+09	2019.007508
std	9.060356e+06	2.991035e+09	1.414725
min	5.142290e+05	4.689493e+04	2017.000000
25%	2.786588e+06	5.664501e+07	2018.000000
50%	4.001667e+06	8.455845e+08	2019.000000
75%	9.410336e+06	2.271558e+09	2020.000000
max	4.628933e+07	1.639873e+10	2021.000000

In []: sns.pairplot(df_filtrado)

Out[]: <seaborn.axisgrid.PairGrid at 0x1cedb81d870>



```
In [ ]: fig, ax = plt.subplots()

ax.bar(df_filtrado['Ano'], df_filtrado['Valor'])

plt.xlabel('Ano')
plt.ylabel('Valor')
plt.title('Venda Total por Ano')

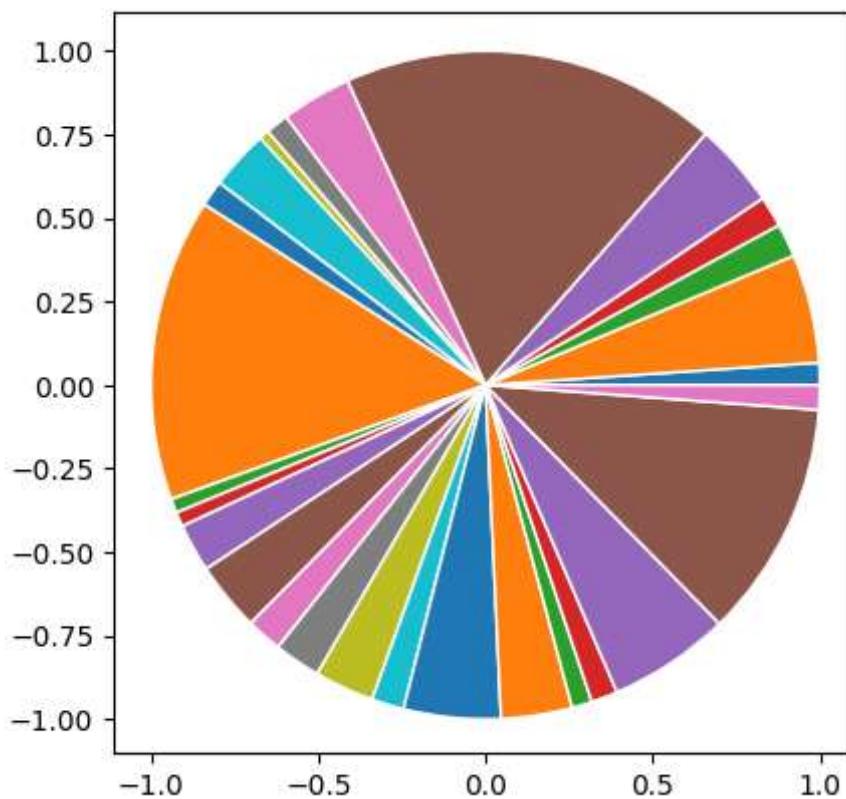
plt.show()
```



```
In [ ]: X = df_filtrado.groupby('Instituição')['Valor'].sum()
colors = plt.get_cmap('Blues')

fig, ax = plt.subplots()
ax.pie(X,
        wedgeprops={"linewidth": 1, "edgecolor": "white"}, frame=True)
```

```
Out[ ]: ([<matplotlib.patches.Wedge at 0x1cefec7e740>,
<matplotlib.patches.Wedge at 0x1cefec7ec20>,
<matplotlib.patches.Wedge at 0x1cefec7f100>,
<matplotlib.patches.Wedge at 0x1cefec7f5e0>,
<matplotlib.patches.Wedge at 0x1cefec7fac0>,
<matplotlib.patches.Wedge at 0x1cefec7ffa0>,
<matplotlib.patches.Wedge at 0x1cefecb04c0>,
<matplotlib.patches.Wedge at 0x1cefecb09a0>,
<matplotlib.patches.Wedge at 0x1cefecb0e80>,
<matplotlib.patches.Wedge at 0x1cefecb1360>,
<matplotlib.patches.Wedge at 0x1cefec7e710>,
<matplotlib.patches.Wedge at 0x1cefecb1cf0>,
<matplotlib.patches.Wedge at 0x1cefecb21d0>,
<matplotlib.patches.Wedge at 0x1cefecb26b0>,
<matplotlib.patches.Wedge at 0x1cefecb2b90>,
<matplotlib.patches.Wedge at 0x1cefecb3070>,
<matplotlib.patches.Wedge at 0x1cefecb3550>,
<matplotlib.patches.Wedge at 0x1cefecb3a30>,
<matplotlib.patches.Wedge at 0x1cefecb3f10>,
<matplotlib.patches.Wedge at 0x1cefecf0430>,
<matplotlib.patches.Wedge at 0x1cefecf0910>,
<matplotlib.patches.Wedge at 0x1cefecf0df0>,
<matplotlib.patches.Wedge at 0x1cefecf12d0>,
<matplotlib.patches.Wedge at 0x1cefecf17b0>,
<matplotlib.patches.Wedge at 0x1cefecf1c90>,
<matplotlib.patches.Wedge at 0x1cefecf2170>,
<matplotlib.patches.Wedge at 0x1cefecf2650>],
[Text(1.099373129205974, 0.037131156457409296, ''),
Text(1.0703372771177677, 0.25372842413124175, ''),
Text(0.9910907648099838, 0.47722017550430695, ''),
Text(0.9398162075332214, 0.5716165638414208, ''),
Text(0.8251233844901228, 0.7274416817639509, ''),
Text(0.16087498952902984, 1.088172430152517, ''),
Text(-0.5583349739453117, 0.9477668789683928, ''),
Text(-0.6867853921157636, 0.8592588813485706, ''),
Text(-0.7287570451240958, 0.8239618736215873, ''),
Text(-0.8114955712699127, 0.7426135857963536, ''),
Text(-0.9011223882539231, 0.6308553252430753, ''),
Text(-1.0932172632327033, 0.12196727171663108, ''),
Text(-1.0253610918104856, -0.398289632555517, ''),
Text(-1.0069967531490411, -0.44267091518111873, ''),
Text(-0.9601982531258467, -0.5366743097019575, ''),
Text(-0.8496670787332101, -0.6986171020787948, ''),
Text(-0.7279401800249503, -0.8246836328588334, ''),
Text(-0.6193349898702726, -0.90907874814143, ''),
Text(-0.46497444718886216, -0.996894559851448, ''),
Text(-0.32142194963129533, -1.051992362279887, ''),
Text(-0.11036355463876135, -1.0944495812085167, ''),
Text(0.16919957233673452, -1.0869091520090657, ''),
Text(0.3179490521889147, -1.053047197523535, ''),
Text(0.39182427477624, -1.0278490831323797, ''),
Text(0.611858606544519, -0.9141274777608426, ''),
Text(0.997609634356257, -0.46343825634012503, ''),
Text(1.0992396748952544, -0.04089177345354142, '')])
```



In []: