

Importando o pandas

```
import pandas as pd
```

Lendo o csv

```
sal = pd.read_csv('Salaries.csv', sep=",")
```

```
/home/luiz/projects/python/notebook/jupyterenv/lib/python3.8/site-packages/IPy
exec(code_obj, self.user_global_ns, self.user_ns)
```

Pegando as cinco primeiras linhas do csv

```
sal.head()
```

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefit
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	Na
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	Na
2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	Na
3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	Na

Trazendo as informações do projeto

```
File "/tmp/ipykernel_3461/3775428509.py", line 1
    Trazendo as informações do projeto
    ^
SyntaxError: invalid syntax
```

SEARCH STACK OVERFLOW

```
sal.info()
```



```
3   BasePay          148041 non-null float64
4   OvertimePay      148646 non-null float64
5   OtherPay         148646 non-null float64
6   Benefits         112487 non-null float64
7   TotalPay         148646 non-null float64
8   TotalPayBenefits 148646 non-null float64
9   Year             148646 non-null int64
10  Notes             0 non-null float64
11  Agency            148646 non-null object
12  Status            38115 non-null object
dtypes: float64(7), int64(2), object(4)
memory usage: 14.7+ MB
```

Qual é o "BasePay" médio?

```
sal.BasePay.mean()

66327.2409237306
```

Qual é a maior quantidade de "OvertimePay" no conjunto de dados?

```
sal.OvertimePay.max()

245131.88
```

Qual é o cargo de JOSEPH DRISCOLL? Nota: use todas as maiúsculas, caso contrário você pode obter uma resposta que não coincida (há também um Joseph Driscoll com minúsculas).

```
sal.loc[sal["EmployeeName"] == "JOSEPH DRISCOLL"]["JobTitle"]

24    CAPTAIN, FIRE SUPPRESSION
Name: JobTitle, dtype: object
```

Quanto JOSEPH DRISCOLL ganha (incluindo benefícios)?

```
sal.loc[sal["EmployeeName"] == "JOSEPH DRISCOLL"]["TotalPayBenefits"]

24    270324.91
```

Qual o nome da pessoa paga mais baixa (incluindo benefícios)? Você percebe algo estranho sobre o quanto ele ou ela é paga?

```
menorSal = sal["TotalPayBenefits"].min();  
sal.query('TotalPayBenefits == @menorSal')
```

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits
110530	110531	David P Kucia	Police Officer 3	NaN	0.0	0.0	-33.89

Qual foi a média (média) BasePay de todos os funcionários por ano? (2011-2014)?

```
sal.loc[sal['Year'].between(2011,2014)].groupby("Year")["BasePay"].mean()
```

```
Year  
2011    63595.956517  
2012    65436.406857  
2013    69630.030216  
2014    66571.407565  
Name: BasePay, dtype: float64
```

Quantos títulos de trabalho únicos existem?

```
sal["JobTitle"].unique();
```

Quais são os 5 principais empregos mais comuns?

```
sal["JobTitle"].value_counts().head()
```

```
Transit Operator    7036  
Special Nurse      4389  
Registered Nurse   3736
```

```
sal["JobTitle"].str.len().corr(sal["TotalPayBenefits"])
```

```
-0.036929089358550604
```