

ETL: Manipulando dados com Pandas

Pandas

Prof. Dr. Diego Bruno

Education Tech Lead na DIO

Doutor em Robótica e *Machine Learning* pelo ICMC-USP



Leitura de dados

Prof. Dr. Diego Bruno

Pandas



Leitura dos dados

```
[ ] import pandas as pd
url = 'https://raw.githubusercontent.com/Muralimekala/python/master/Resp2.csv'
df1 = pd.read_csv(url)
# Dataset is now stored in a Pandas Dataframe
```

```
[ ] import pandas as pd
url = 'https://raw.githubusercontent.com/Muralimekala/python/master/Salaries.csv'
sf = pd.read_csv(url)
#salaries CSV
sf.head()
```

Tabela gerada

	Id	EmployeeName	JobTitle	BasePay	OvertimePay	OtherPay	Benefits	TotalPay	TotalPayBenefits	Year	Notes	Agency	Status
0	1	NATHANIEL FORD	GENERAL MANAGER-METROPOLITAN TRANSIT AUTHORITY	167411.18	0.00	400184.25	NaN	567595.43	567595.43	2011	NaN	San Francisco	NaN
1	2	GARY JIMENEZ	CAPTAIN III (POLICE DEPARTMENT)	155966.02	245131.88	137811.38	NaN	538909.28	538909.28	2011	NaN	San Francisco	NaN
2	3	ALBERT PARDINI	CAPTAIN III (POLICE DEPARTMENT)	212739.13	106088.18	16452.60	NaN	335279.91	335279.91	2011	NaN	San Francisco	NaN
3	4	CHRISTOPHER CHONG	WIRE ROPE CABLE MAINTENANCE MECHANIC	77916.00	56120.71	198306.90	NaN	332343.61	332343.61	2011	NaN	San Francisco	NaN
4	5	PATRICK GARDNER	DEPUTY CHIEF OF DEPARTMENT,(FIRE DEPARTMENT)	134401.60	9737.00	182234.59	NaN	326373.19	326373.19	2011	NaN	San Francisco	NaN

Obrigado!

Prof. Dr. Diego Bruno

