

Curso de Ciência de Dados – Unipê
Disciplina: Programação Aplicada à CD
Professor: Thyago Maia
Aluno: Huguemberg Mariano

Atividade: Plotar um gráfico a partir de um dataset qualquer utilizando o Jupyter Notebook.



Ideia proposta: Analisar os dados da Formula 1, identificando o piloto que obteve o melhor desempenho após as mudanças no regulamento de motores da categoria, passando do antigo motor V8 para os mais econômicos e menos poluentes motor V6 a parti de 2014.

Atividade prática com Python: o arquivo Jupyter com o script usado pode ser acessado no repositório do GitHub: <https://github.com/huguemberg/>

JupyterPrática_F1 Last Checkpoint 28 minutos atrás (autosaved)

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3 (ipykernel)

```
In [ ]: ### Atividade prática relativa ao segunda prova
### Programação - Thiago Nêta
### HUGUENBERG MARIANO

# Fonte: https://github.com/hakube/formula1-dataset

Análise de Dados da Formula 1 de 1950 a 2020

In [ ]: import pandas as pd
import numpy as np

from matplotlib import pyplot as plt
from matplotlib import rcParams
import seaborn as sns

In [ ]: #Importando o arquivo csv e visualizando as 5 primeiras linhas

df = pd.read_csv("formula1_data_races.csv", sep=";")
df.head(5)

In [ ]: #Verificando o tipo de cada coluna

df.dtypes

In [ ]: # Filtrando o primeiro colocado de cada corrida a partir do ano de 2014

filtro_p1 = df.position == "1"
filtro_14a20 = df.year >= 2014
dfp1 = df.loc[filtro_p1 & filtro_14a20]
dfp1

In [ ]: # Agrupando o vencedores por piloto, assim temos o total de vitórias dos pilotos na F1

dfp1.groupby(['driver']).position.count().sort_values(ascending=False)

## Fazendo o gráfico

In [ ]: # Estabeleceu o tamanho e qual o estilo do gráfico

%matplotlib inline
rcParams['figure.figsize'] = 6,8
sns.set_style('whitegrid')

In [ ]: # Informamos qual a posição dos dados no gráfico e fazemos a plotagem dele

dfp1.groupby(['driver']).position.count().sort_values(ascending=True).plot(kind='barh')
plt.title("PILOTOS COM MAIS VITÓRIAS NA F1 COM MOTORES HÍBRIDOS (V6) ENTRE 2014 A 2020")
plt.xlabel('Quantidade de vitórias')
plt.ylabel('Pilotos')

In [ ]: dfp1.groupby(['driver']).position.count().sort_values(ascending=True)

In [ ]: # Os dados analisados mostram o número de vitórias alcançadas pelos pilotos na F1, desde 2014, quando houve a mudança dos
# motores, saindo do antigos V8 para o modelo V6, estes são mais econômicos e menos poluentes.
# O Gráfico em questão demonstra ainda um domínio no período do piloto Lewis Hamilton, chegando a ter mais de 50% de vitórias
# nas corridas disputadas.
```

Parte 2/3

JupyterPrática_F1 Last Checkpoint 32 minutos atrás (autosaved)

File Edit View Insert Cell Kernel Widgets Help Notebook saved Trusted Python 3 (ipykernel)

```
In [ ]: ### Atividade prática relativa ao segunda prova
### Programação - Thiago Nêta
### HUGUENBERG MARIANO

# Fonte: https://github.com/hakube/formula1-dataset

Análise de Dados da Formula 1 de 1950 a 2020

In [1]: import pandas as pd
import numpy as np

from matplotlib import pyplot as plt
from matplotlib import rcParams
import seaborn as sns

In [2]: #Importando o arquivo csv e visualizando as 5 primeiras linhas

df = pd.read_csv("formula1_data_races.csv", sep=";")
df.head(5)

Out[2]:
```

	position	number	driver	car	laps	time	points	grand_prix	year	season lap	
0	1	77	Valtteri Bottas	BOT	Mercedes	71.0	130.55.738	25.0	AUSTRIA	2020	no
1	2	16	Charles Leclerc	LEC	Ferrari	71.0	+2.700	18.0	AUSTRIA	2020	no
2	3	4	Lando Norris	NOR	McLaren Renault	71.0	+5.4810	16.0	AUSTRIA	2020	yes
3	4	44	Lewis Hamilton	HAM	Mercedes	71.0	+5.8800	12.0	AUSTRIA	2020	no
4	5	55	Carlos Sainz	SAINZ	McLaren Renault	71.0	+6.3030	10.0	AUSTRIA	2020	no

```
In [3]: #Verificando o tipo de cada coluna

df.dtypes

Out[3]:
```

	position	number	driver	car	laps	time	points	grand_prix	year	season lap
position	object	int64	object	object	object	float64	float64	object	int64	object

```
In [4]: # Filtrando o primeiro colocado de cada corrida a partir do ano de 2014

filtro_p1 = df.position == "1"
filtro_14a20 = df.year >= 2014
dfp1 = df.loc[filtro_p1 & filtro_14a20]
dfp1

Out[4]:
```

	position	number	driver	car	laps	time	points	grand_prix	year	season lap	
0	1	77	Valtteri Bottas	BOT	Mercedes	71.0	130.55.738	25.0	AUSTRIA	2020	no
20	1	44	Lewis Hamilton	HAM	Mercedes	71.0	122.55.683	25.0	STYRIA	2020	no
40	1	44	Lewis Hamilton	HAM	Mercedes	70.0	136.12.473	25.0	HUNGARY	2020	yes
60	1	44	Lewis Hamilton	HAM	Mercedes	52.0	120.01.283	25.0	GREAT BRITAIN	2020	no
80	1	33	Max Verstappen	VER	Red Bull Racing Honda	52.0	119.41.893	25.0	70TH ANNIVERSARY	2020	no
...
2726	1	44	Lewis Hamilton	HAM	Mercedes	44.0	151.43.021	25.0	JAPAN	2014	yes
2740	1	44	Lewis Hamilton	HAM	Mercedes	53.0	131.50.744	25.0	RUSSIA	2014	no
2760	1	44	Lewis Hamilton	HAM	Mercedes	56.0	140.04.785	25.0	UNITED STATES	2014	no
2780	1	6	Nico Rosberg	ROS	Mercedes	71.0	130.02.056	25.0	BRASIL	2014	no
2800	1	44	Lewis Hamilton	HAM	Mercedes	55.0	130.02.019	25.0	ABU DHABI	2014	no

130 rows x 10 columns

```
In [5]: # Agrupando o vencedores por piloto, assim temos o total de vitórias dos pilotos na F1

dfp1.groupby(['driver']).position.count().sort_values(ascending=False)

Out[5]:
```

driver	count
Lewis Hamilton	73
Nico Rosberg	20
Sebastian Vettel	14
Max Verstappen	10
Valtteri Bottas	9
Daniel Ricciardo	7
Charles Leclerc	3
Kimi Räikkönen	1
Pierre Gasly	1
Sergio Perez	1

Name: position, dtype: int64

```
Fazendo o gráfico

In [6]: # Estabelece o tamanho e qual o estilo do gráfico

%matplotlib inline
rcParams['figure.figsize'] = 6,8
sns.set_style('whitegrid')

In [7]: # Informamos qual a posição dos dados no gráfico e fazemos a plotagem dele

dfp1.groupby(['driver']).position.count().sort_values(ascending=True).plot(kind='barh')
plt.title("PILOTOS COM MAIS VITÓRIAS NA F1 COM MOTORES HÍBRIDOS (V6) ENTRE 2014 A 2020")
plt.xlabel('Quantidade de vitórias')
plt.ylabel('Pilotos')

Out[7]:
```

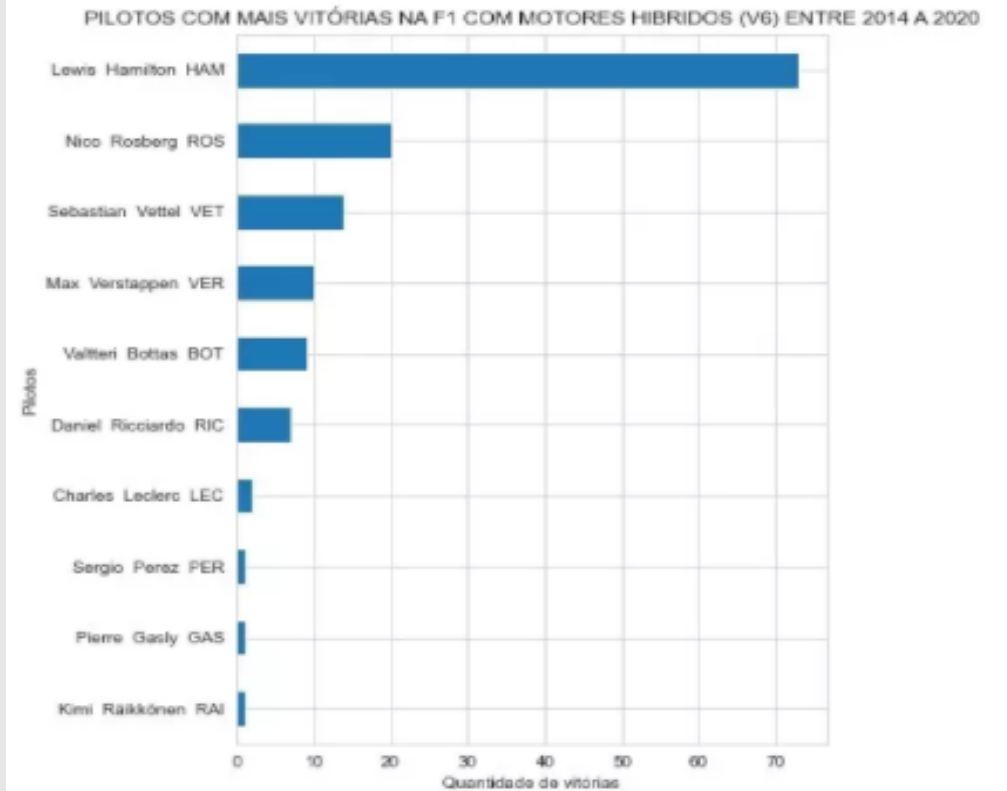
```
In [ ]: dfp1.groupby(['driver']).position.count().sort_values(ascending=True)

In [ ]: # Os dados analisados mostram o número de vitórias alcançadas pelos pilotos na F1, desde 2014, quando houve a mudança dos
# motores, saindo do antigos V8 para o modelo V6, estes são mais econômicos e menos poluentes.
# O Gráfico em questão demonstra ainda um domínio no período do piloto Lewis Hamilton, chegando a ter mais de 50% de vitórias
# nas corridas disputadas.
```

Curso de Ciência de Dados – Unipê
Disciplina: Programação Aplicada à CD
Professor: Thyago Maia
Aluno: Huguemberg Mariano

Ideia proposta: Analisar os dados da Formula 1, identificando o piloto que obteve o melhor desempenho após as mudanças no regulamento de motores da categoria, passando do antigo motor V8 para os mais econômicos e menos poluentes motor V6 a parti de 2014.

Resultado Final



Rank de pilotos da F1 com mais vitórias na era de motores híbridos V6 (2014 a 2020)

Fonte: <https://github.com/hakube/formula1-dataset>

Os dados analisados mostram o número de vitórias alcançados pelos pilotos na F1, desde 2014, quando houve a mudança dos motores, saindo do antigo V8 para o modelo V6, este mais econômico e menos poluente.

O Gráfico em questão demonstra um domínio no período do piloto Lewis Hamilton, chegando a ter mais de 50% de vitórias nas corridas disputadas.