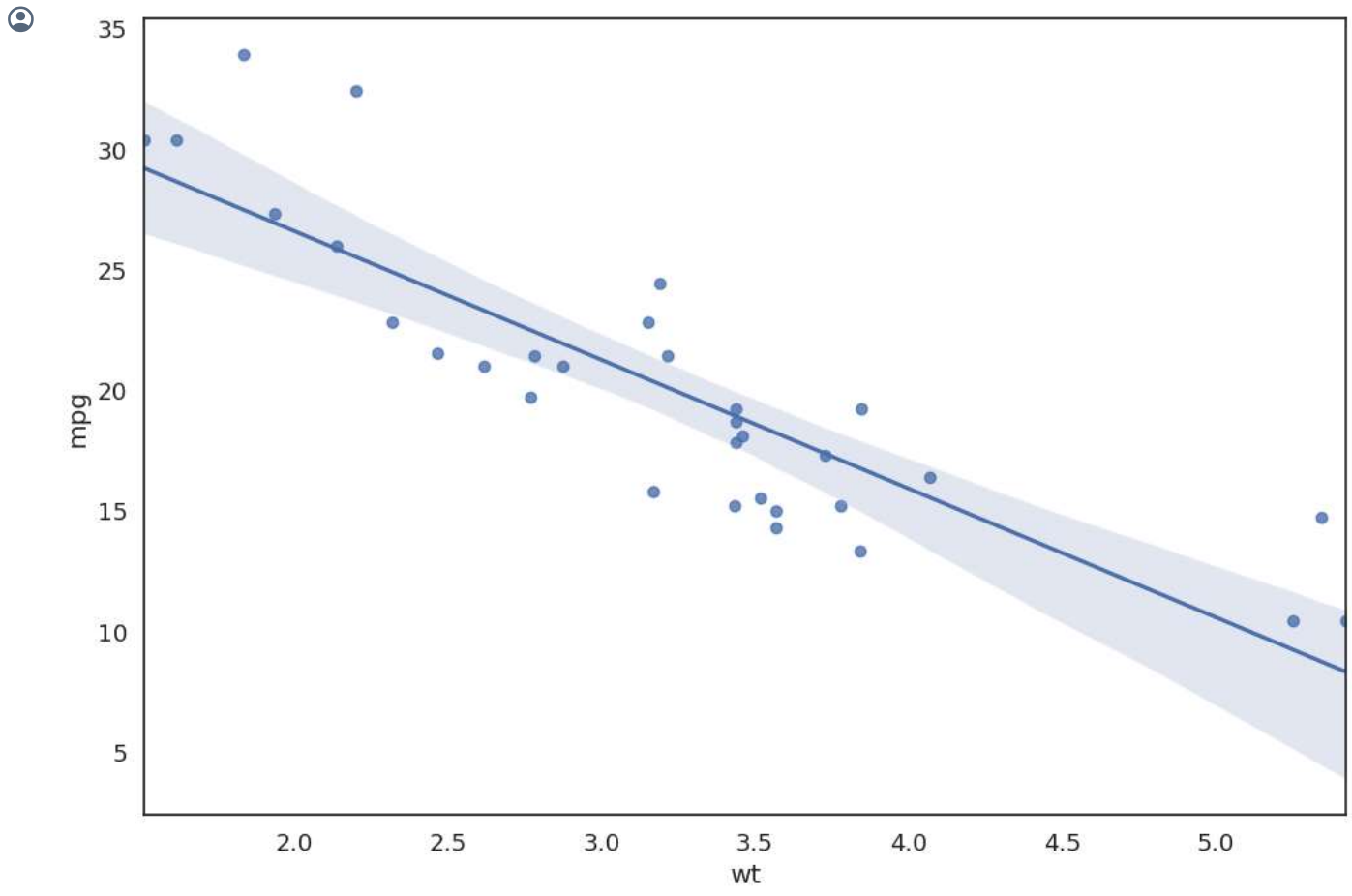
1 Scatter Plot in Python

In the first Python data visualization example we are going to create a simple scatter plot. As previously mentioned we are going to use Seaborn to create the scatter plot

```
%matplotlib inline

import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns

# Suppress warnings
import warnings
warnings.filterwarnings('ignore')


# Optional but changes the figure size
fig = plt.figure(figsize=(12, 8))

df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv')

# Change the size of the font
sns.set(font_scale=1.2)

# Set white backgrounds on the plots
sns.set_style("white")

ax = sns.regplot(x="wt", y="mpg", data=df)
```

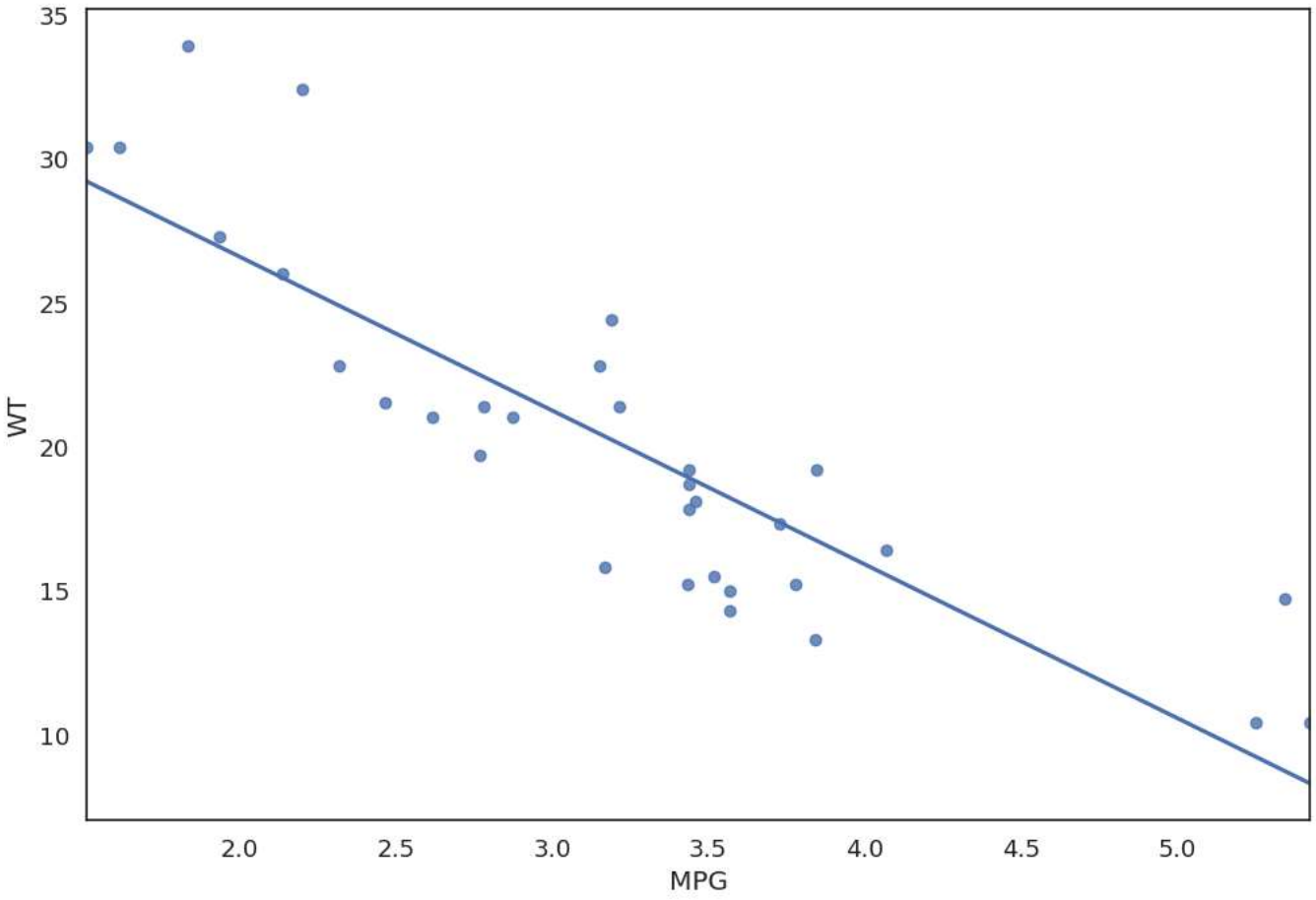


Changing the Labels on a Seaborn Plot

In the next example we are going to ORDET JAG INTE HITTAR the seaborn plot a bit. First we are going to remove the confidence intervall but we are also going to change the bales on the x-axis and y-axis.

```
fig = plt.figure(figsize=(12, 8))
ax = sns.regplot(x="wt", y="mpg", ci=False, data=df)

# Set white backgrounds on the plots
sns.set_style("white")

# Change axis labels
ax.set(xlabel="MPG", ylabel="WT")
```

```
[Text(0.5, 0, 'MPG'), Text(0, 0.5, 'WT')]
```



For more about scatter plots

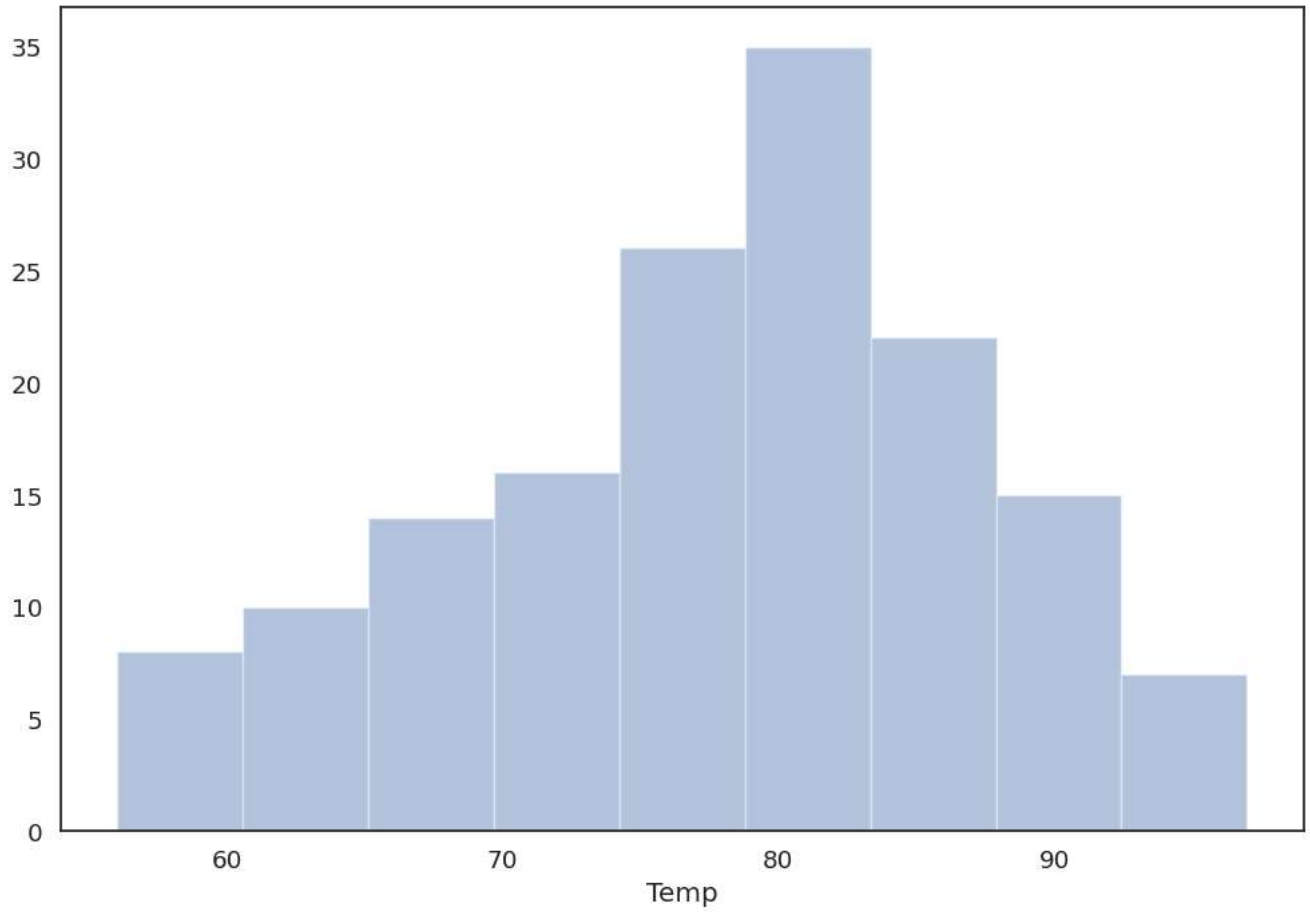
- How to make Scatter Plots in Python (YouTube Video)
- Expolratory Data Analysis with Pandas Scipy and Seaborn

2 Histogram in Python using Seaborn

In the next Python data visualization example we will create histograms. Histograms are fairly easy to create using Seaborn. In the first Seaborn histogram example we hve turned set the parameter *kde* to false. This so that we only get the histogram.

```
df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/airquality.csv')
fig = plt.figure(figsize=(12, 8))
sns.distplot(df.Temp, kde=False)
```

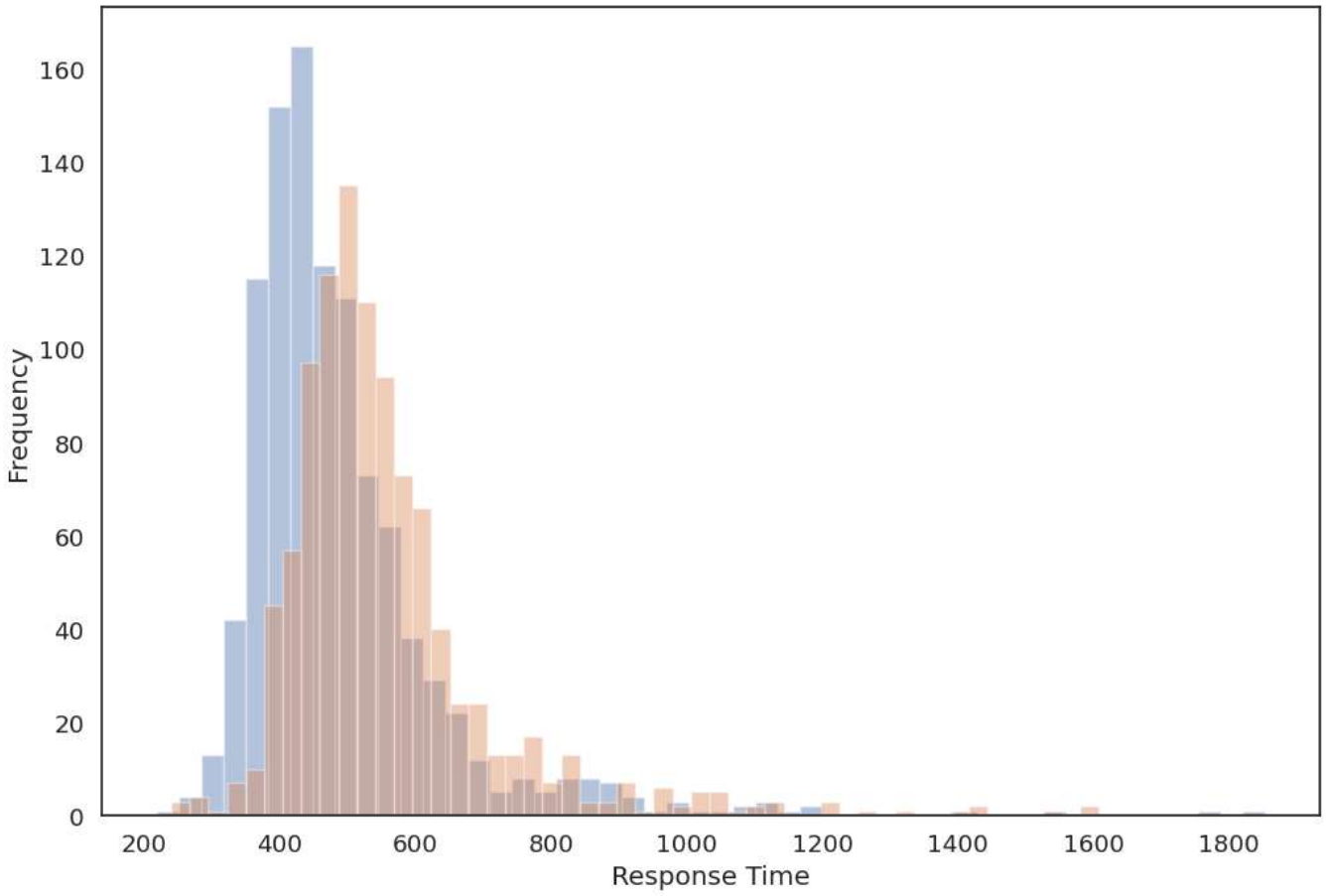
```
<Axes: xlabel='Temp'>
```



Grouped Histogram in Seaborn Python

```
df = pd.read_csv('https://raw.githubusercontent.com/marsja/jupyter/master/flanks.csv',
                 index_col=0)

fig = plt.figure(figsize=(12, 8))
for condition in df.TrialType.unique():
    cond_data = df[(df.TrialType == condition)]
    ax = sns.distplot(cond_data.RT, kde=False)

ax.set(xlabel='Response Time', ylabel='Frequency')
```

```
[Text(0.5, 0, 'Response Time'), Text(0, 0.5, 'Frequency')]
```

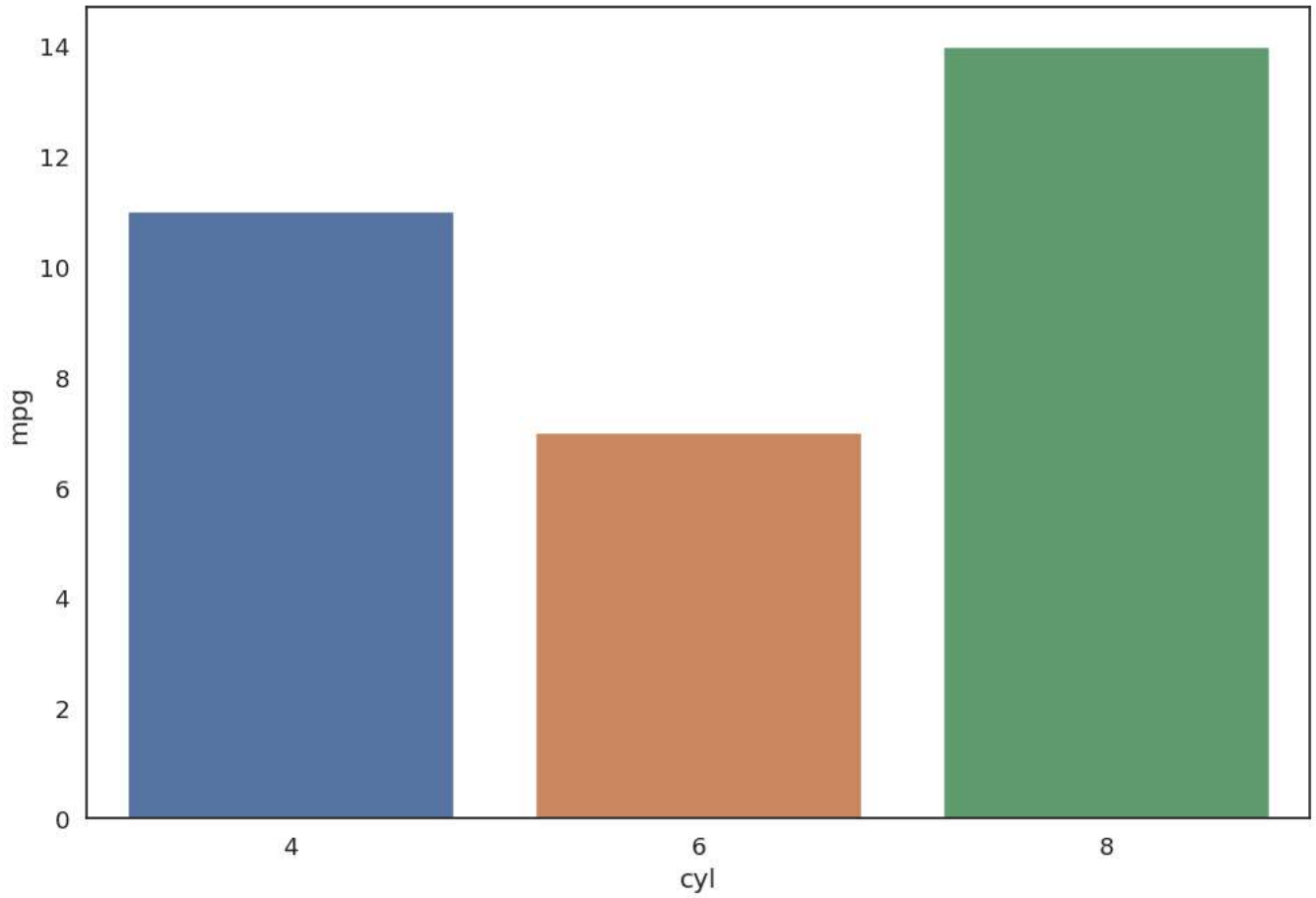


3 Bar Plots & Stack Bar Chart in Python using Seaborn

In this example we are starting by using Pandas groupby to group the data by the index column (the car models=r

```
df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv', index_col=0)

df_grpd = df.groupby("cyl").count().reset_index()

fig = plt.figure(figsize=(12, 8))


# Set white backgrounds on the plots
sns.set_style("white")

sns.barplot(x="cyl", y="mpg", data=df_grpd)
```

```
<Axes: xlabel='cyl', ylabel='mpg'>
```

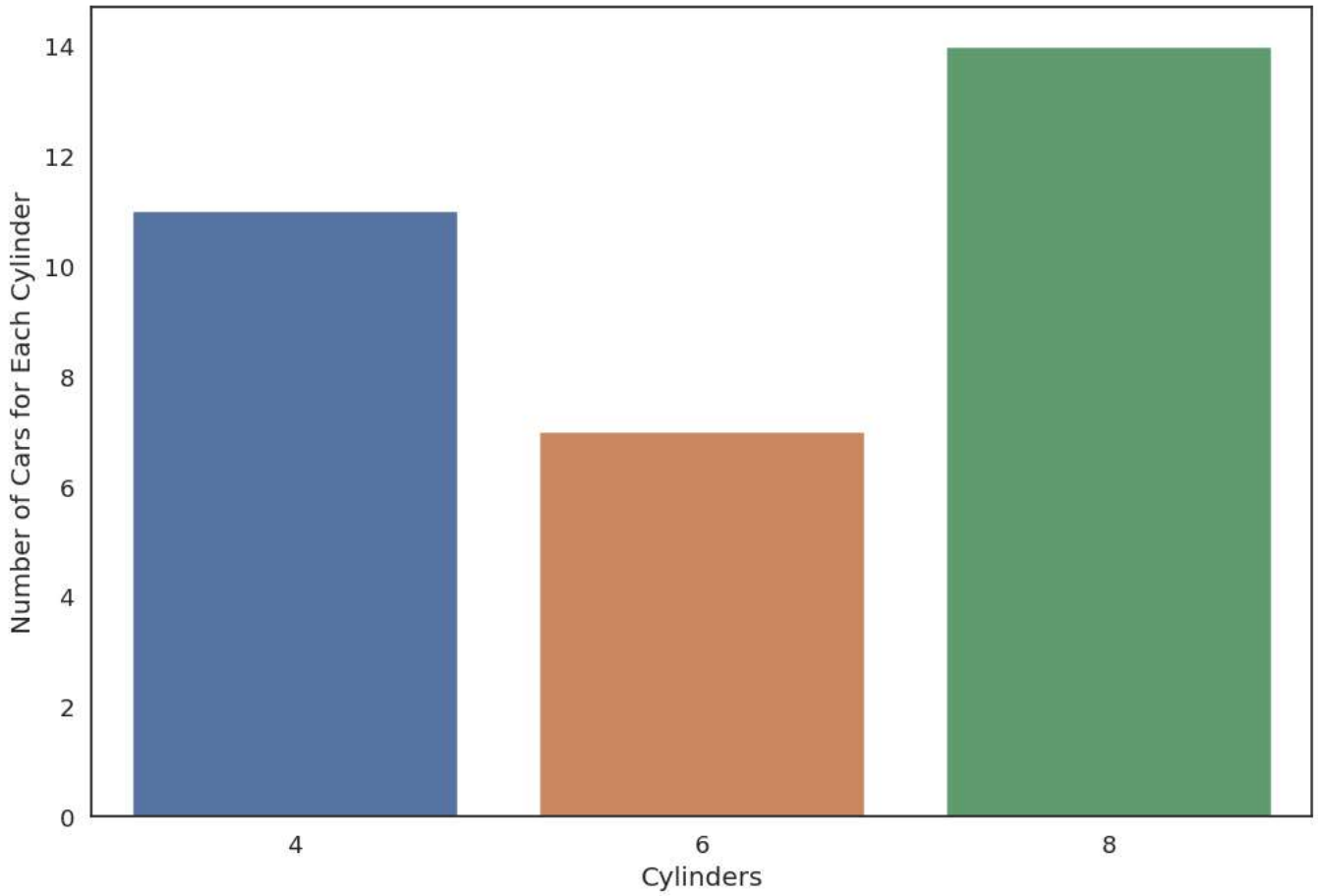


Setting the Labes of a Seaborn Bar Plot

In the next example we are going to change labels because the y-axis actually represents the count of cars in each cyl category:

```
df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv', index_col=0)

df_grpd = df.groupby("cyl").count().reset_index()

fig = plt.figure(figsize=(12, 8))
ax = sns.barplot(x="cyl", y="mpg", data=df_grpd)

# Set white backgrounds on the plots
sns.set_style("white")

ax.set(xlabel='Cylinders', ylabel='Number of Cars for Each Cylinder')
```

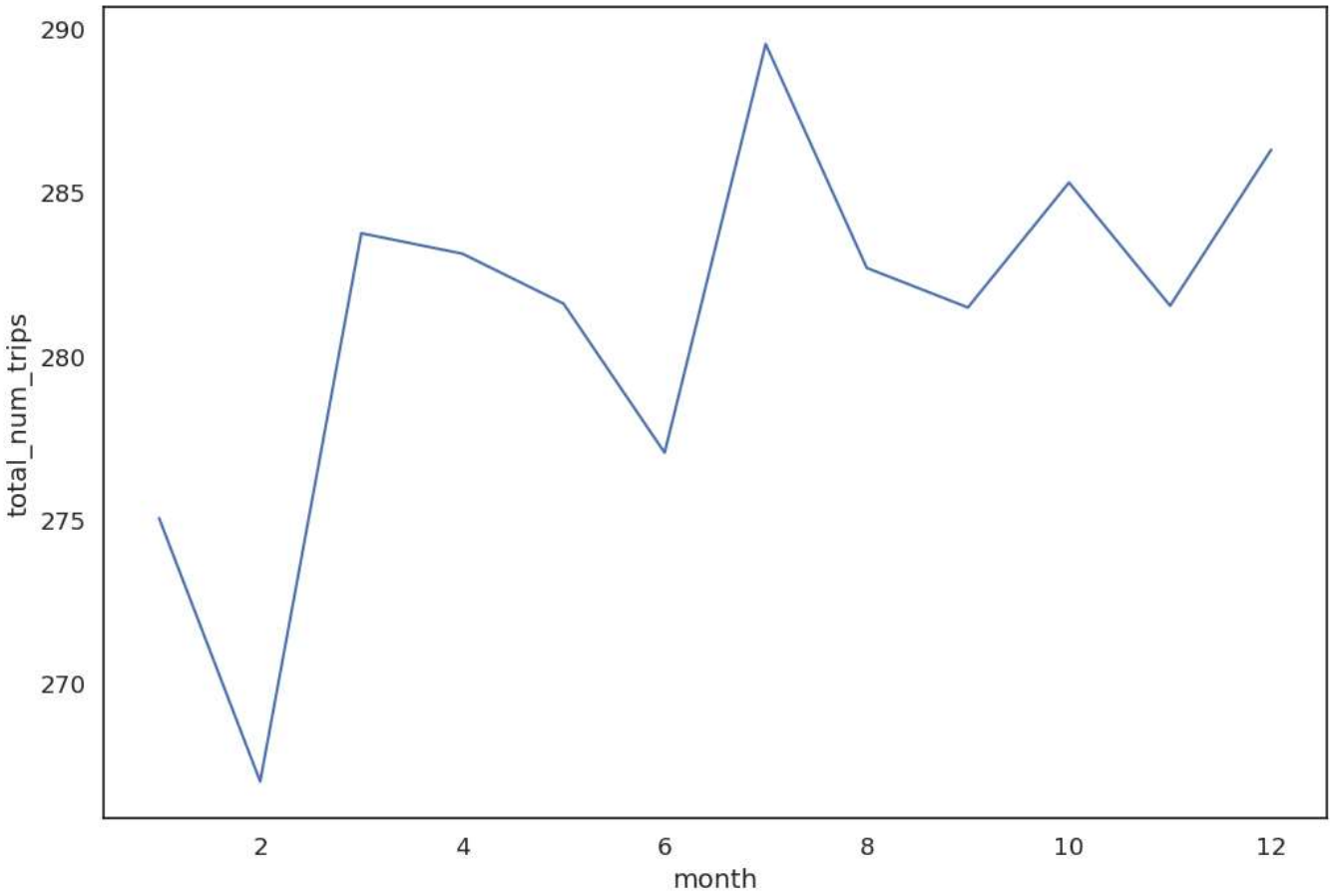
```
[Text(0.5, 0, 'Cylinders'), Text(0, 0.5, 'Number of Cars for Each Cylinder')]
```



Time Series Plot in Seaborn

```
import pandas as pd
import seaborn as sns

df = pd.read_csv("https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2019/2019-02-26/full_trains.csv")

fig = plt.figure(figsize=(12, 8))
sns.lineplot(x="month", y="total_num_trips",
             ci=None, data=df)
```

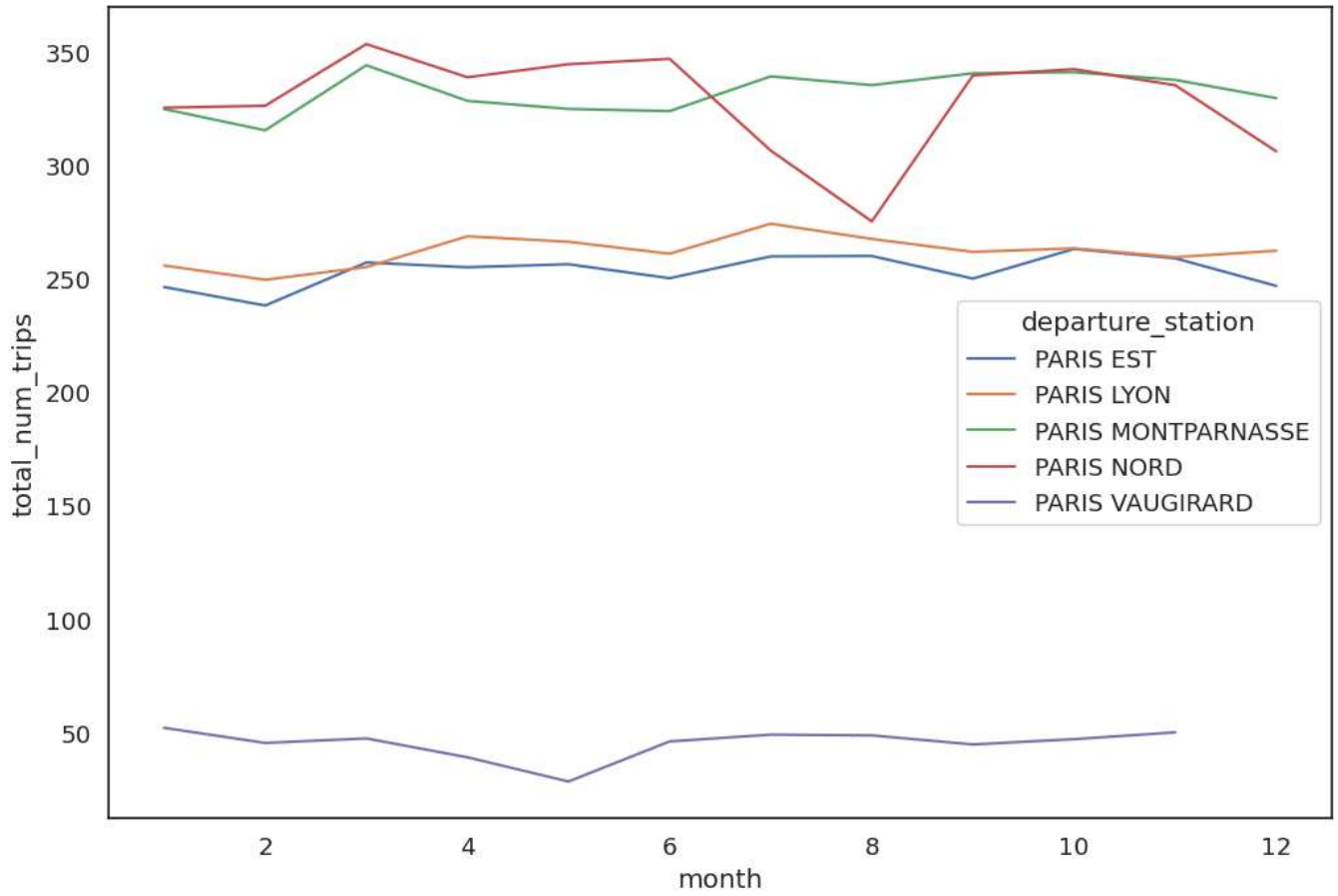
```
<Axes: xlabel='month', ylabel='total_num_trips'>
```



Grouped Time Series Plot in Seaborn

```
import pandas as pd
import seaborn as sns

df = pd.read_csv("https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2019/2019-02-26/full_trains.csv")

fig = plt.figure(figsize=(12, 8))
sns.lineplot(x="month", y="total_num_trips", hue="departure_station",
             ci=None, data=df[df.departure_station.str.contains('PARIS')])
```

```
<Axes: xlabel='month', ylabel='total_num_trips'>
```

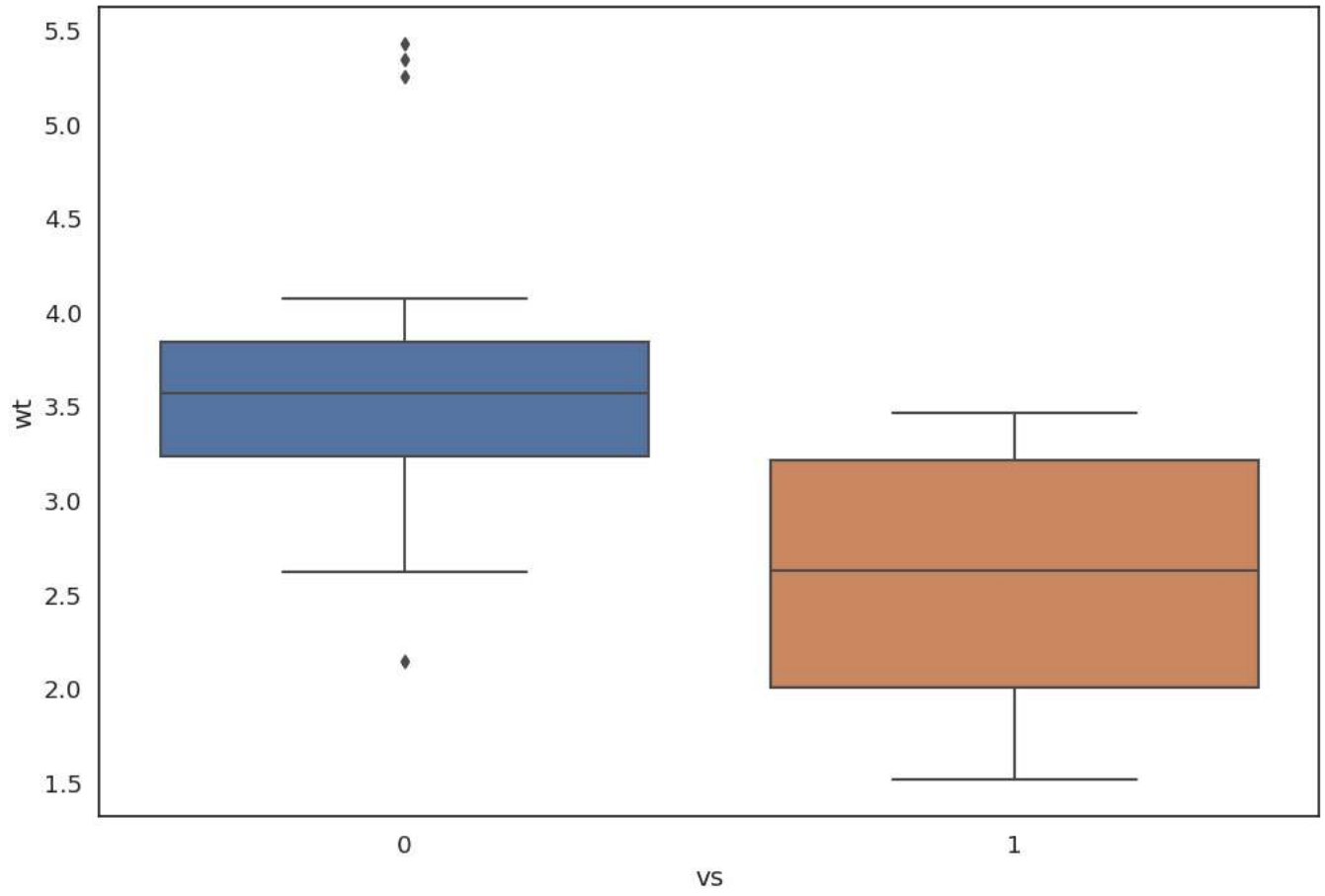


4 Box Plots in Python using Seaborn

In the next examples we are going to learn how to visualize data, in python, by creating box plots using Seaborn

```
import pandas as pd
import seaborn as sns

df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv', index_col=0)

fig = plt.figure(figsize=(12, 8))
sns.boxplot(x="vs", y='wt', data=df)
```

```
<Axes: xlabel='vs', ylabel='wt'>
```



5 Heat Map in Python using Seaborn

```
import pandas as pd
import seaborn as sns

df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv', index_col=0)

fig = plt.figure(figsize=(12, 8))
ax = sns.heatmap(df[['mpg', 'disp', 'hp', 'drat', 'wt', 'qsec']])
```

6 Correlation Matrix (Corrgram Corrplot)

We continue with an Python data visualization example in which we are going to use heatmap to create a correlation plot.


```
import numpy as np
import pandas as pd
import seaborn as sns


df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv',
                 index_col=0)

# Correlation matrix
corr = df.corr()

mask = np.zeros_like(corr)
mask[np.triu_indices_from(mask)] = True


fig = plt.figure(figsize=(12, 8))
sns.heatmap(corr, mask=mask, vmax=.3, center=0,
            square=True, linewidths=.5, cbar_kws={"shrink": .5})
```

```
<Axes: >
```