

Python Data Visualization Tutorial This is the code example for the the post 9 Data Visualization Techniques You Should Learn in Python . The post is a Python data visualization tutorial in which we mainly learn how to use Python Seaborn to create 8 different plots. In the last post we are using the Python package ptitprince to create a Raincloud plot in Python.

Installing ptitprince using pip

pip install ptitprince

pip install ptitprince

```
Collecting ptitprince
  Downloading ptitprince-0.2.7.tar.gz (12 kB)
  Preparing metadata (setup.py) ... done
Requirement already satisfied: matplotlib in /usr/local/lib/python3.10/dist-packages (from ptitprince) (3.7.1)
Requirement already satisfied: numpy>=1.20 in /usr/local/lib/python3.10/dist-packages (from ptitprince) (1.25.2)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from ptitprince) (1.11.4)
Collecting seaborn==0.11 (from ptitprince)
  Downloading seaborn-0.11.0-py3-none-any.whl (283 kB)
    283.1/283.1 kB 4.3 MB/s eta 0:00:00
Requirement already satisfied: pandas>=0.23 in /usr/local/lib/python3.10/dist-packages (from seaborn==0.11->ptitprince) (2.0.3)
Requirement already satisfied: contourpy>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (1.2.1)
Requirement already satisfied: cycler>=0.10 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (0.12.1)
Requirement already satisfied: fonttools>=4.22.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (4.51.0)
Requirement already satisfied: kiwisolver>=1.0.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (1.4.5)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (24.0)
Requirement already satisfied: pillow>=6.2.0 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (9.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in /usr/local/lib/python3.10/dist-packages (from matplotlib->ptitprince) (2.8.2)
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=0.23->seaborn==0.11->ptitprince) (2023.4)
Requirement already satisfied: tzdata>=2022.1 in /usr/local/lib/python3.10/dist-packages (from pandas>=0.23->seaborn==0.11->ptitprince) (2024.1)
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->matplotlib->ptitprince) (1.16.0)
Building wheels for collected packages: ptitprince
  Building wheel for ptitprince (setup.py) ... done
  Created wheel for ptitprince: filename=ptitprince-0.2.7-py3-none-any.whl size=10656 sha256=ec1009a4eec7336d7e23433f783aa05bc885f42e6a8534adb0d08
  Stored in directory: /root/.cache/pip/wheels/0e/43/31/e76a3bf61865543f076a9d9eb027a740caefb379424ecba4e8
Successfully built ptitprince
Installing collected packages: seaborn, ptitprince
  Attempting uninstall: seaborn
    Found existing installation: seaborn 0.13.1
    Uninstalling seaborn-0.13.1:
      Successfully uninstalled seaborn-0.13.1
Successfully installed ptitprince-0.2.7 seaborn-0.11.0
```

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')
```

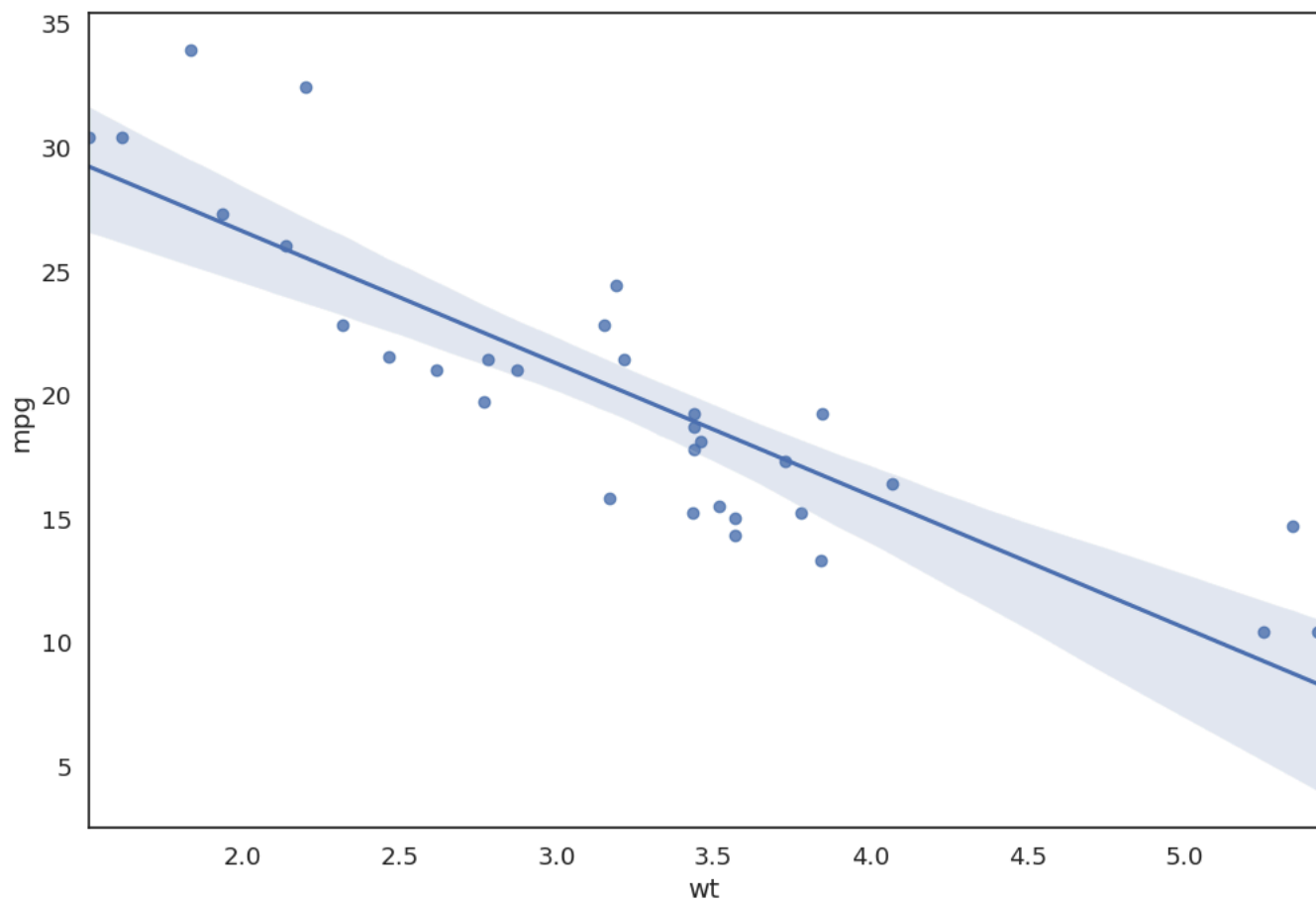
```
# Optional but changes the figure size
fig = plt.figure(figsize=(12, 8))
```

```
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

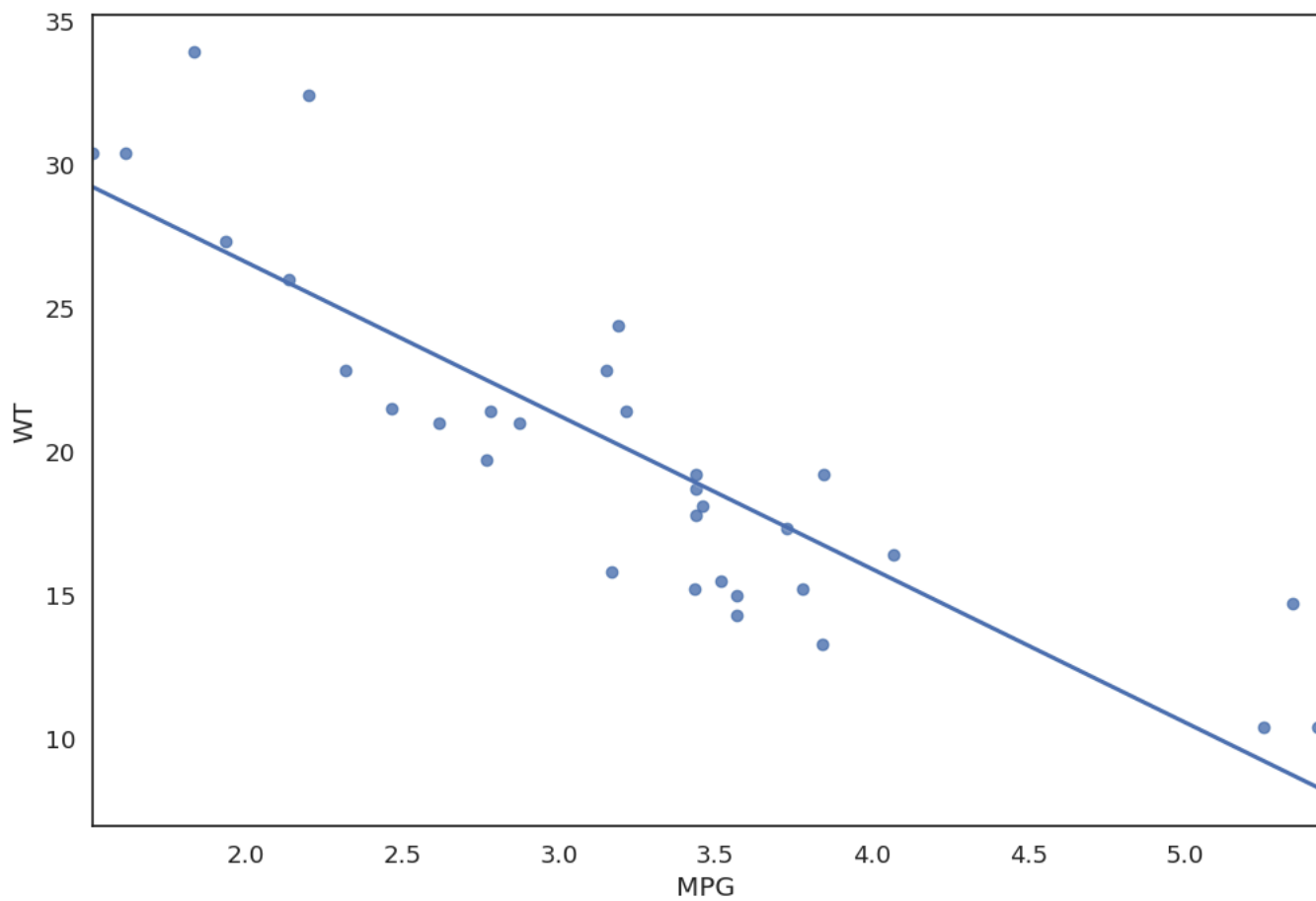
In the next example we are going to ORDET JAG INTE HITTA the seaborn plot a bit. First we are going to remove the confidence interval but we are also going to change the labels on the x-axis and y-axis.

```
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')]
```

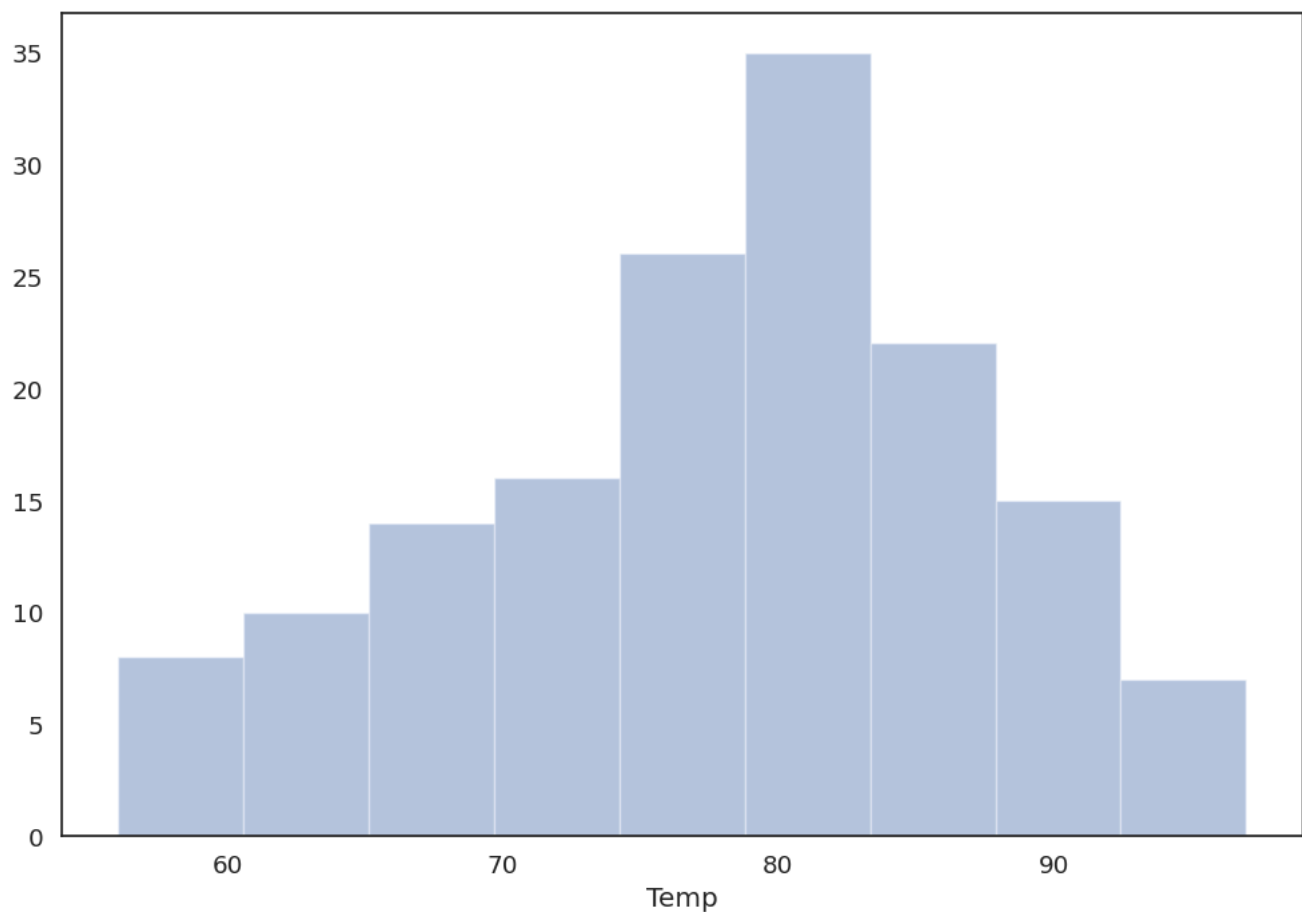


2 Histogram in Python using Seaborn

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 have 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'>



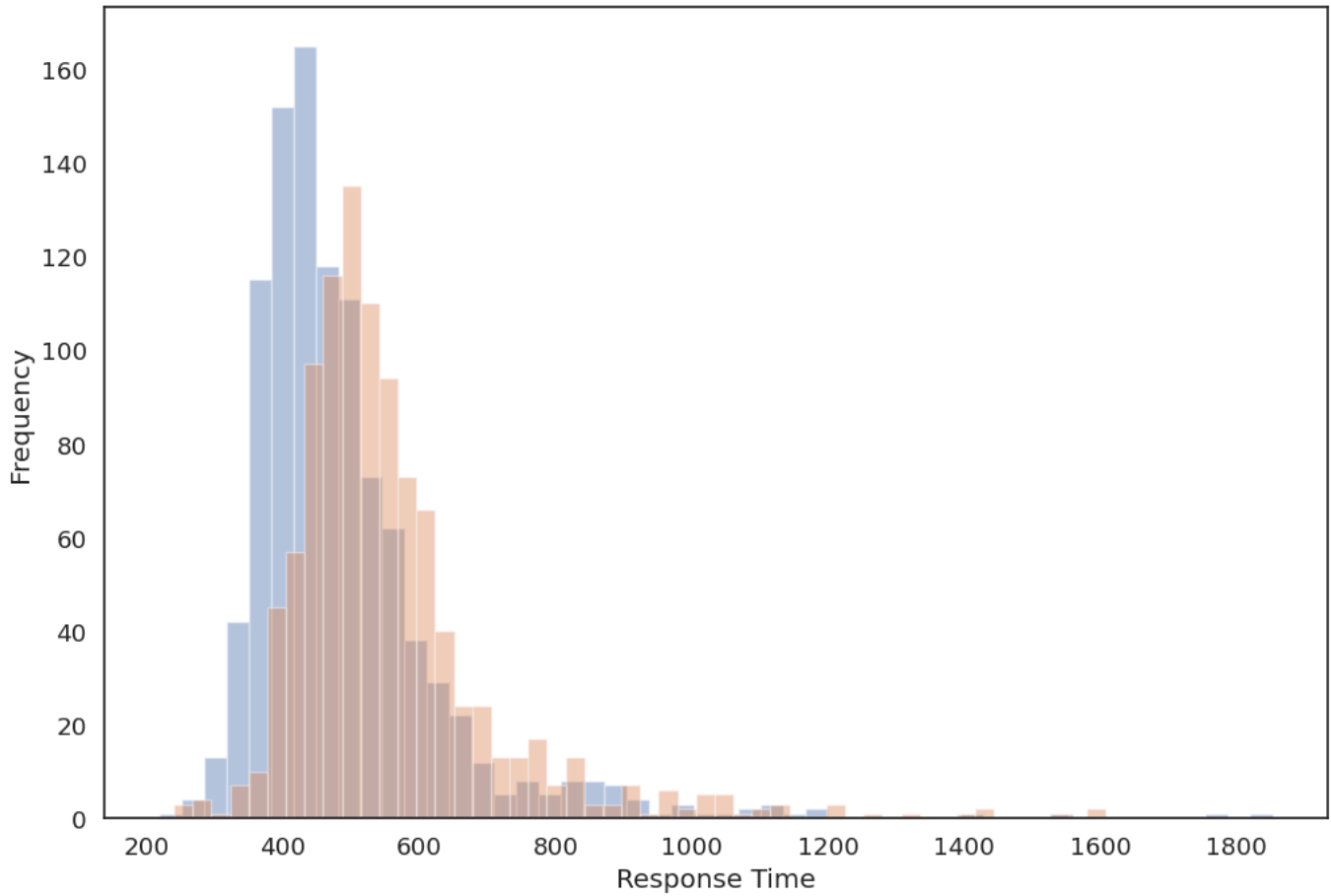
Grouped Histogram in Seaborn Python

```
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

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

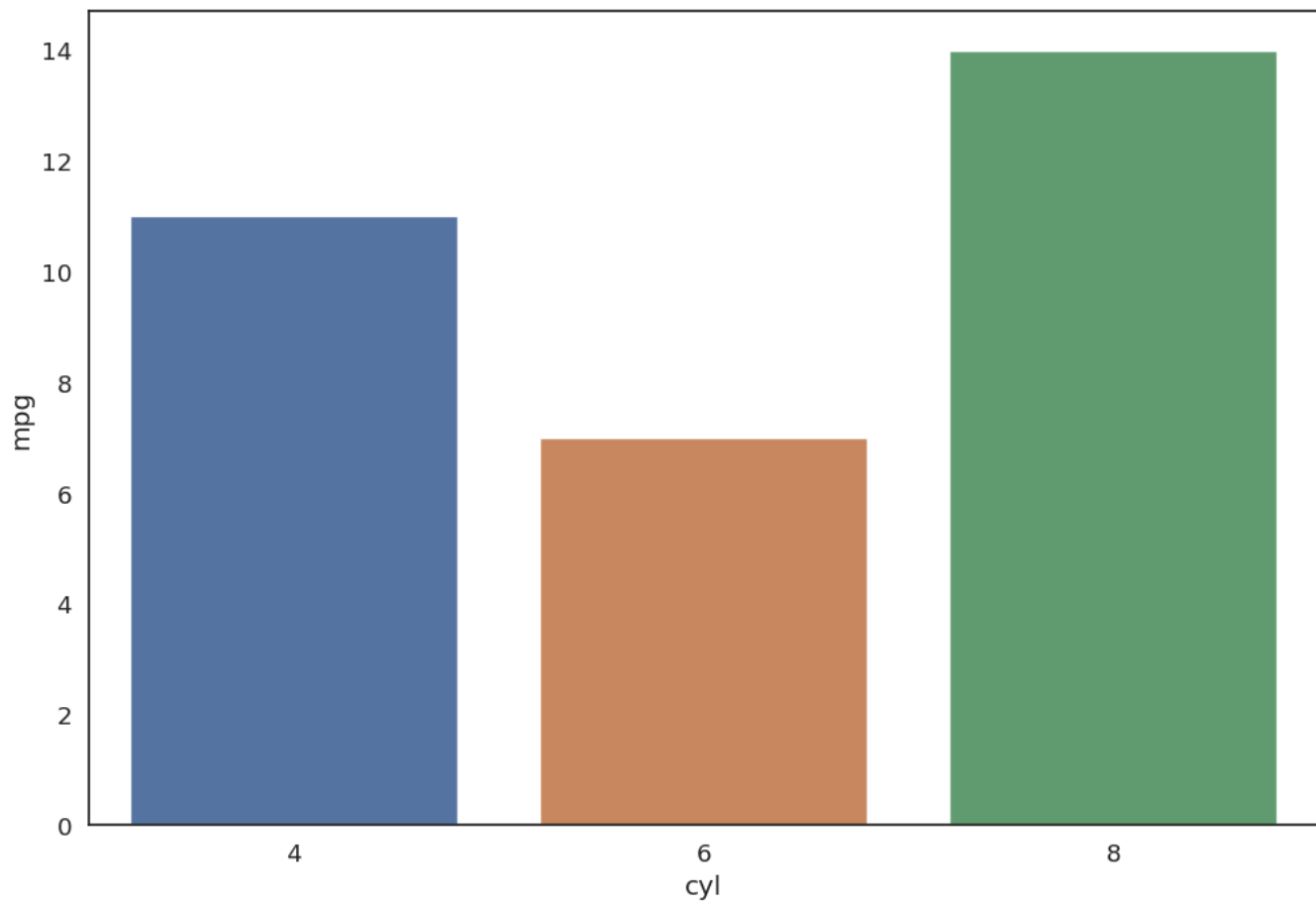
```
df_grpd = df.groupby("cyl").count().reset_index()
```

```
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

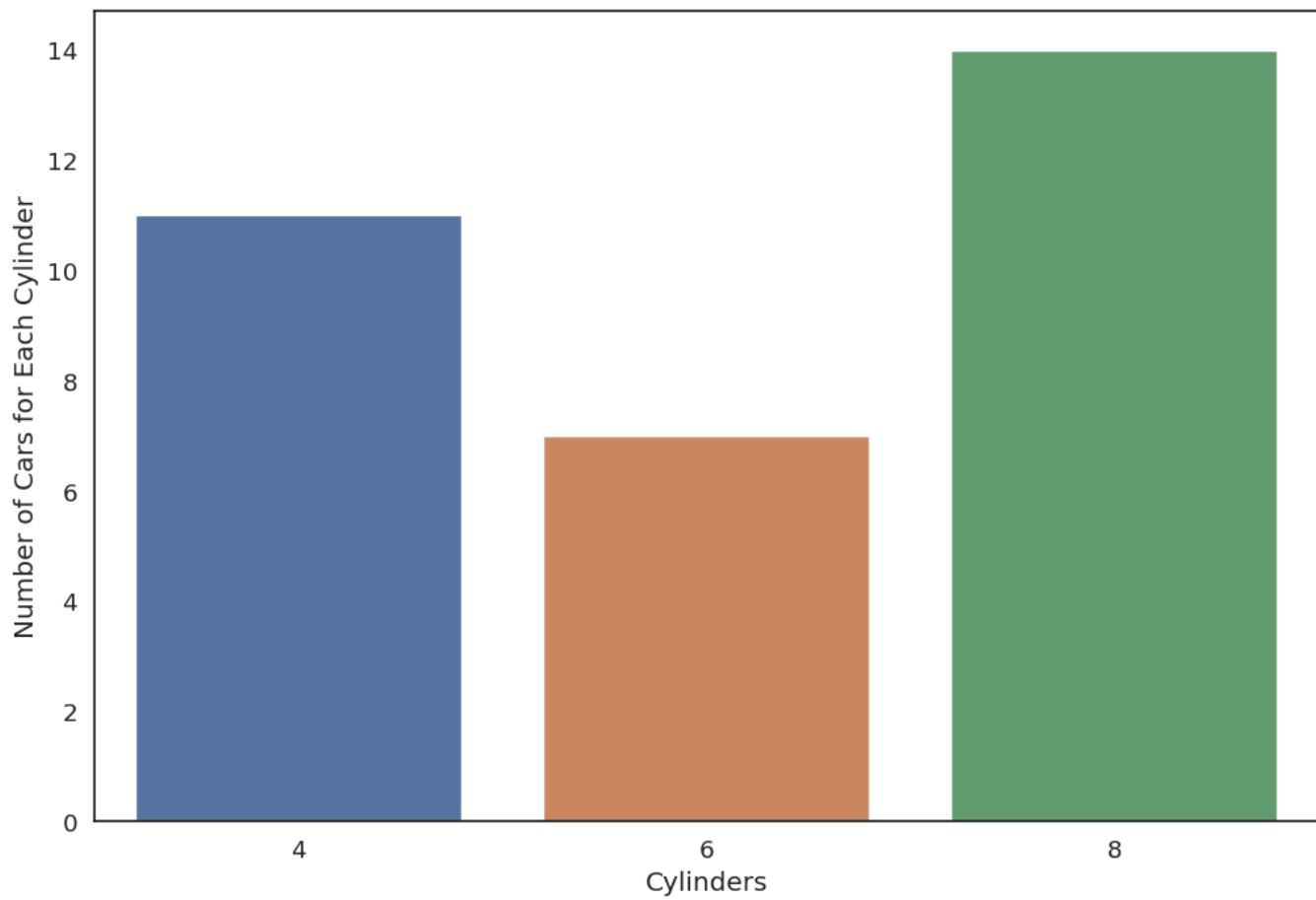
```
df = pd.read_csv('https://vincentarelbundock.github.io/Rdatasets/csv/datasets/mtcars.csv', index_col=0)
df_grpd = df.groupby("cyl").count().reset_index()

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')]
```



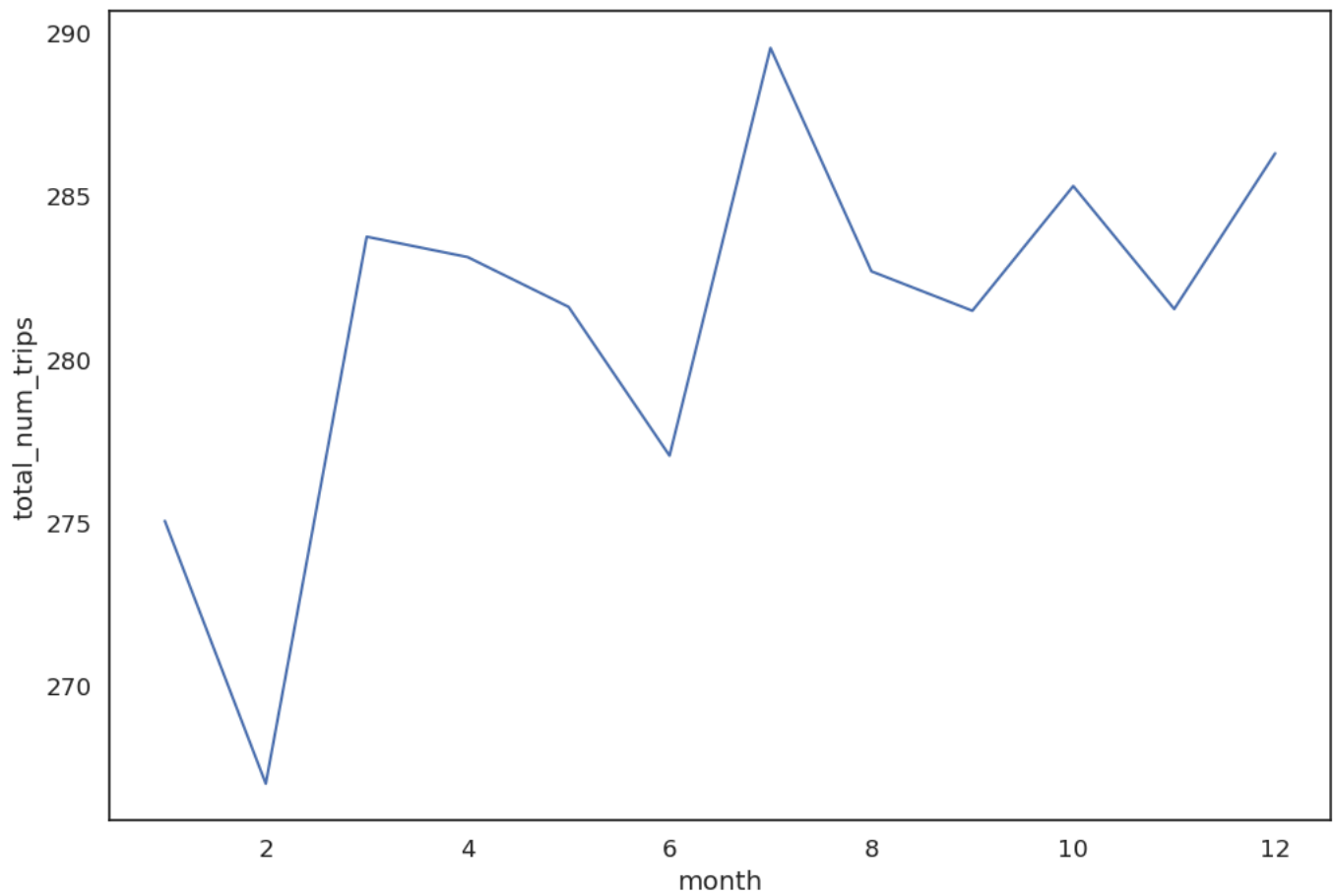
Time Series Plot in Seaborn

```
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'>



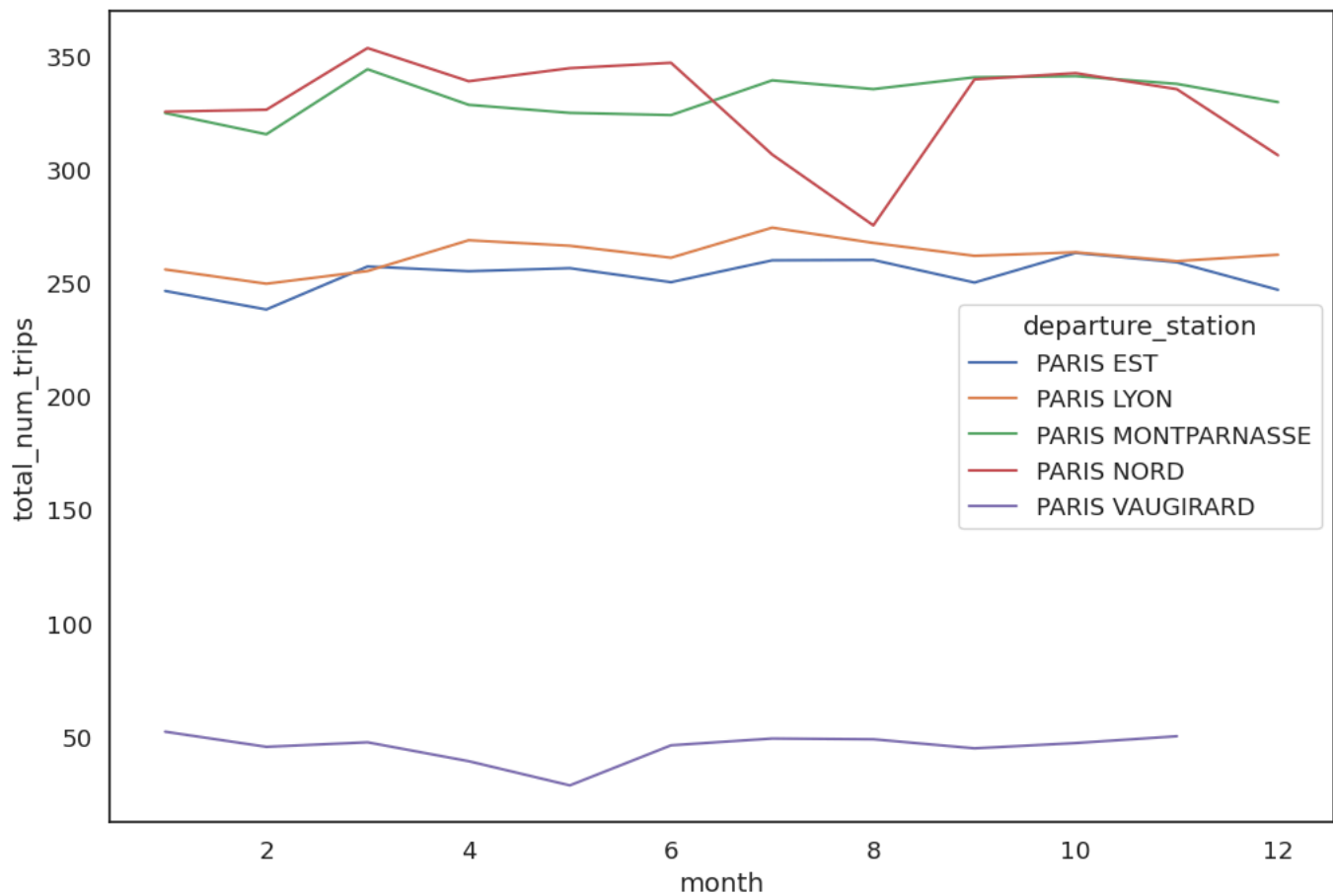
Grouped Time Series Plot in Seaborn

```
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'>



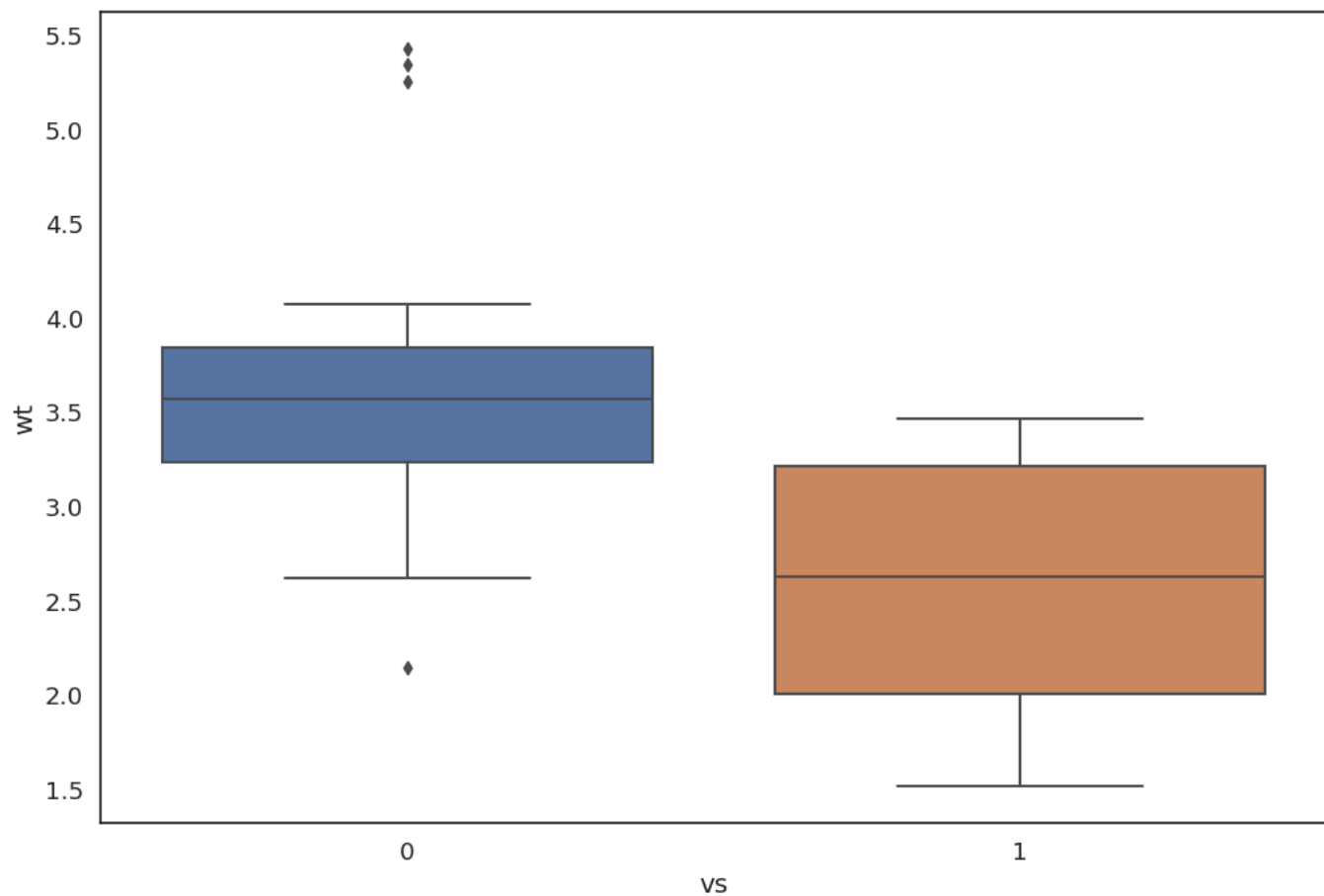
4 Box Plots in Python using Seaborn

```
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']])
```

