This notebook is an exercise in the Data Visualization course. You can reference the tutorial at this link.

Now it's time for you to demonstrate your new skills with a project of your own!

In this exercise, you will work with a dataset of your choosing. Once you've selected a dataset, you'll design and create your own plot to tell interesting stories behind the data!

Setup

Run the next cell to import and configure the Python libraries that you need to complete the exercise.

In [ ]:

```
import pandas as pd
pd.plotting.register_matplotlib_converters()
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
print("Setup Complete")
```

The questions below will give you feedback on your work. Run the following cell to set up the feedback system.

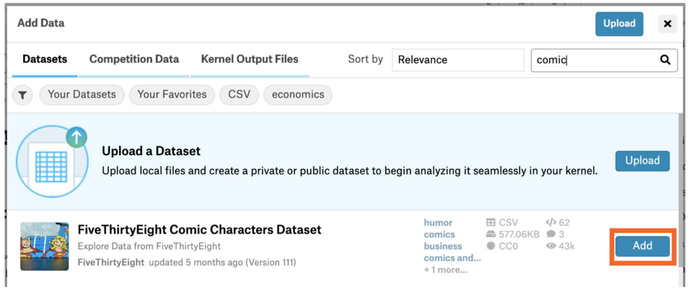
In [ ]:

```
# Set up code checking
from learntools.core import binder
binder.bind(globals())
from learntools.data_viz_to_coder.ex7 import *
print("Setup Complete")
```

Step 1: Attach a dataset to the notebook

Begin by selecting a CSV dataset from Kaggle Datasets. If you're unsure how to do this or would like to work with your own data, please revisit the instructions in the previous tutorial.

Once you have selected a dataset, click on the **[+ Add Data]** option in the top right corner. This will generate a pop-up window that you can use to search for your chosen dataset.



Sort by Relevance

comic

Upload a Dataset

Upload local files and create a private or public dataset to begin analyzing it seamlessly in your kernel.

FiveThirtyEight Comic Characters Dataset

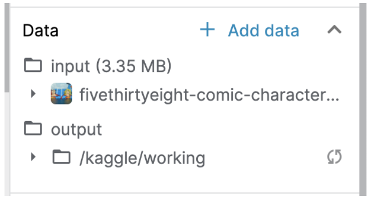
Explore Data from FiveThirtyEight

FiveThirtyEight updated 5 months ago (Version 111)

humor comics business comics and... + 1 more...

CSV 577.06KB CC0 62 3 43k

Once you have found the dataset, click on the **[Add]** button to attach it to the notebook. You can check that it was successful by looking at the **Data** dropdown menu to the right of the notebook -- look for an **input** folder containing a subfolder that matches the name of the dataset.



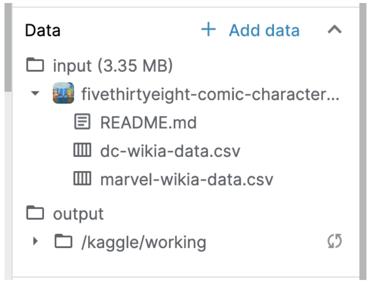
input (3.35 MB)

fivethirtyeight-comic-character...

output

/kaggle/working

You can click on the carat to the left of the name of the dataset to double-check that it contains a CSV file. For instance, the image below shows that the example dataset contains two CSV files: (1) **dc-wikia-data.csv**, and (2) **marvel-wikia-data.csv**.



Data + Add data

input (3.35 MB)

fivethirtyeight-comic-character...

README.md

dc-wikia-data.csv

marvel-wikia-data.csv

output

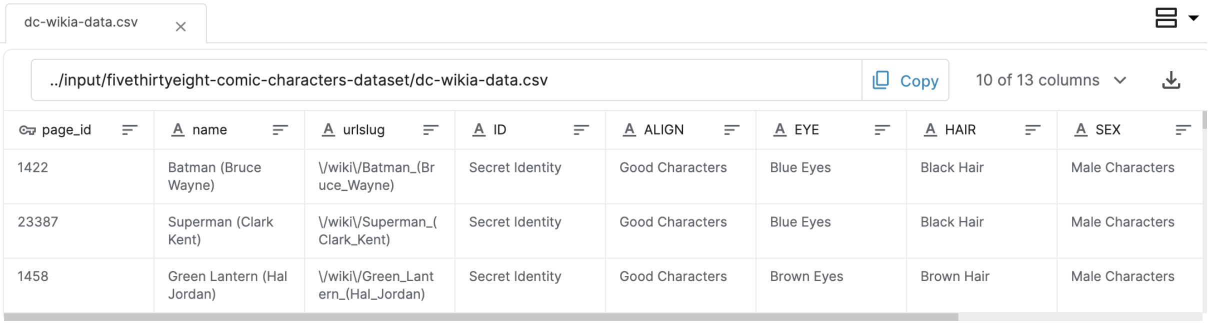
/kaggle/working

Once you've uploaded a dataset with a CSV file, run the code cell below **without changes** to receive credit for your work!

In [ ]:
```
# Check for a dataset with a CSV file
step_1.check()
```

Step 2: Specify the filepath

Now that the dataset is attached to the notebook, you can find its filepath. To do this, begin by clicking on the CSV file you'd like to use. This will open the CSV file in a tab below the notebook. You can find the filepath towards the top of this new tab.



dc-wikia-data.csv

../input/fivethirtyeight-comic-characters-dataset/dc-wikia-data.csv

Copy 10 of 13 columns

page_id	name	urlslug	ID	ALIGN	EYE	HAIR	SEX
1422	Batman (Bruce Wayne)	\/wiki\/Batman_(Bruce_Wayne)	Secret Identity	Good Characters	Blue Eyes	Black Hair	Male Characters
23387	Superman (Clark Kent)	\/wiki\/Superman_(Clark_Kent)	Secret Identity	Good Characters	Blue Eyes	Black Hair	Male Characters
1458	Green Lantern (Hal Jordan)	\/wiki\/Green_Lantern_(Hal_Jordan)	Secret Identity	Good Characters	Brown Eyes	Brown Hair	Male Characters

After you find the filepath corresponding to your dataset, fill it in as the value for `my_filepath` in the code cell below, and run the code cell to check that you've provided a valid filepath. For instance, in the case of this example dataset, we would set

```
my_filepath = "../input/fivethirtyeight-comic-characters-datase
t/dc-wikia-data.csv"
```

Note that **you must enclose the filepath in quotation marks**; otherwise, the code will return an error.

Once you've entered the filepath, you can close the tab below the notebook by clicking on the **[X]** at the top of the tab.

In [ ]:

```
# Fill in the line below: Specify the path of the CSV file to read
my_filepath = "../input/fivethirtyeight-comic-characters-dataset/dc-wik
ia-data.csv"
# Check for a valid filepath to a CSV file in a dataset
step_2.check()
```

Step 3: Load the data

Use the next code cell to load your data file into `my_data`. Use the filepath that you specified in the previous step.

In [ ]:

```
# Fill in the line below: Read the file into a variable my_data
my_data = pd.read_csv(my_filepath, low_memory=False)

# Check that a dataset has been uploaded into my_data
step_3.check()
```

After the code cell above is marked correct, run the code cell below without changes to view the first five rows of the data.

In [ ]:

```
# Print the first five rows of the data
my_data.head()
```

Step 4: Visualize the data

Use the next code cell to create a figure that tells a story behind your dataset. You can use any chart type (*line chart, bar chart, heatmap, etc*) of your choosing!

In [ ]:

```
# Create a plot
my_data.sort_values(by='APPEARANCES', ascending=False, inplace=True)
my_clean_data = my_data[0:11][['ALIVE', 'APPEARANCES']]

# Set the width and height of the figure
plt.figure(figsize=(10,6))

# Add title
plt.title("Awards per Author")

# Bar chart showing average arrival delay for Spirit Airlines flights by month
sns.barplot(x=my_clean_data['ALIVE'], y=my_clean_data['APPEARANCES'])

# Add label for vertical axis
plt.ylabel("N of Awards")
plt.xlabel("Author")
plt.xticks(rotation="90")

# Check that a figure appears below
step_4.check()
plt.show()
```

Keep going

Learn how to use your skills after completing the micro-course to create data visualizations in a **final tutorial**.

Have questions or comments? Visit the Learn Discussion forum to chat with other Learners.