

✓ Continuous Variables - Scatterplot - Exercise

```
import pandas as pd
import matplotlib.pyplot as plt
```

```
from google.colab import drive
import os
```

```
drive.mount('/content/drive')
os.chdir('/content/drive/MyDrive/')
for item in os.listdir():
    print(item)
print("-----")
os.chdir('/content/drive/MyDrive/cloud/GitHub/AdvDataViz/Notebooks/')
for item in os.listdir():
    print(item)
print("-----")
notebooks = "/content/drive/MyDrive/cloud/GitHub/AdvDataViz/Notebooks"
print(os.listdir(notebooks))
print("-----")
```

```
file = "student_performance.csv"
file_path = os.path.join(notebooks, file)
with open(file_path, "r") as f:
    contents = f.read()
```

```
➡ Mounted at /content/drive
learningStore
healthyCar
startup
cloud
Artificial Intelligence
```

```
-----
03 Matplotlib - Exercise.ipynb
02 Matplotlib.ipynb
01 Python_Pandas.ipynb
04 Continuous Variables - Histogram .ipynb
05 Continuous Variables - Histogram - Exercise .ipynb
07 Continuous Variables - Boxplot - Exercise .ipynb
03 Matplotlib - Exercise Solutions.ipynb
05 Continuous Variables - Histogram - Exercise Solutions.ipynb
06 Continuous Variables - Boxplot.ipynb
08 Continuous Variables - Scatterplot.ipynb
07 Continuous Variables - Boxplot - Exercise Solutions.ipynb
09 Continuous Variables - Scatterplot - Exercise Solutions.ipynb
09 Continuous Variables - Scatterplot - Exercise .ipynb
10 Categorical Variables - Bar_Pie.ipynb
12 Seaborn.ipynb
11 Pandas Data Visualization.ipynb
13 Seaborn - Exercise .ipynb
Top 50 US Tech Companies.csv
13 Seaborn - Exercise Solution.ipynb
15 Custom Modules.ipynb
14 Functions.ipynb
churn.csv
student_performance.csv
myplotlib.py
```




employee_attrition_.csv
heart-disease.csv

['03 Matplotlib - Exercise.ipynb', '02 Matplotlib.ipynb', '01 Python_Pandas.ipynb', '04 Cont

Dataset: Student Performance

```
#df = pd.read_csv("student_performance.csv")
df = pd.read_csv(file_path)
```

```
df.head()
```

	gender	race/ethnicity	parental level of education	lunch	test preparation course	math score	reading score	writing score
0	female	group B	bachelor's degree	standard	none	72	72	74
1	female	group C	some college	standard	completed	69	90	88
2	female	group B	master's degree	standard	none	90	95	93
3	male	group A	associate's degree	free/reduced	none	47	57	44
4	male	group C	some college	standard	none	76	78	75

Next steps: [Generate code with df](#) [View recommended plots](#) [New interactive sheet](#)

Use math score, reading score and writing score for the exercises.

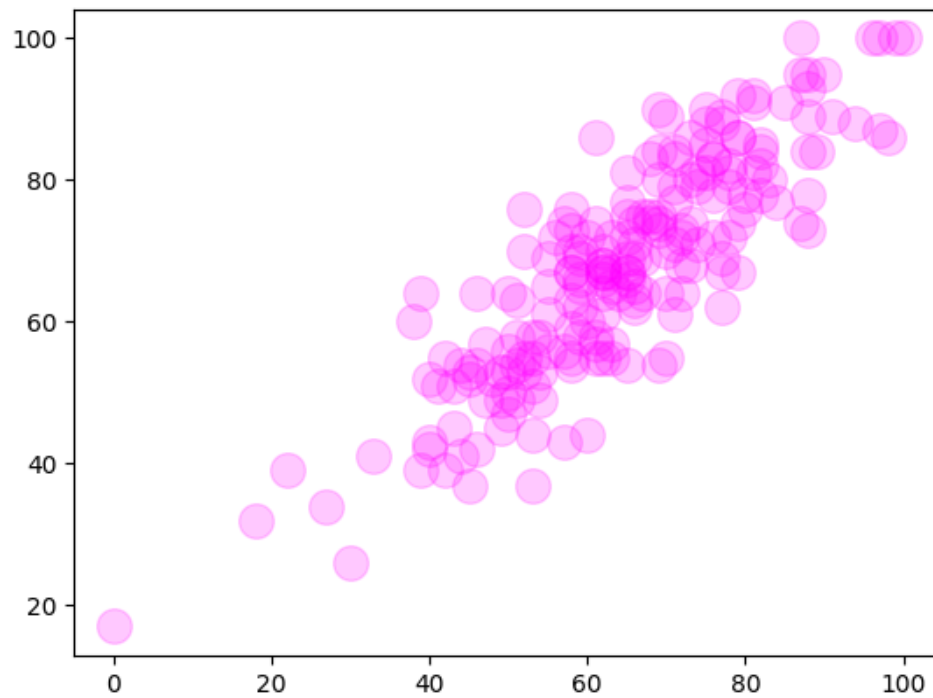
```
gender = df["gender"][:200].map({"male":0, "female":1})
```

```
math_score = df["math score"][:200]
reading_score = df["reading score"][:200]
writing_score = df["writing score"][:200]
```

- 1.) Make a scatterplot of "math_score" and "reading_score". Set the alpha to .2, the size to 200, and set its color to "fuchsia".

```
fig, ax = plt.subplots()
```

```
ax.scatter(x=math_score, y=reading_score, alpha=.2, s = 200, c = "fuchsia");
```



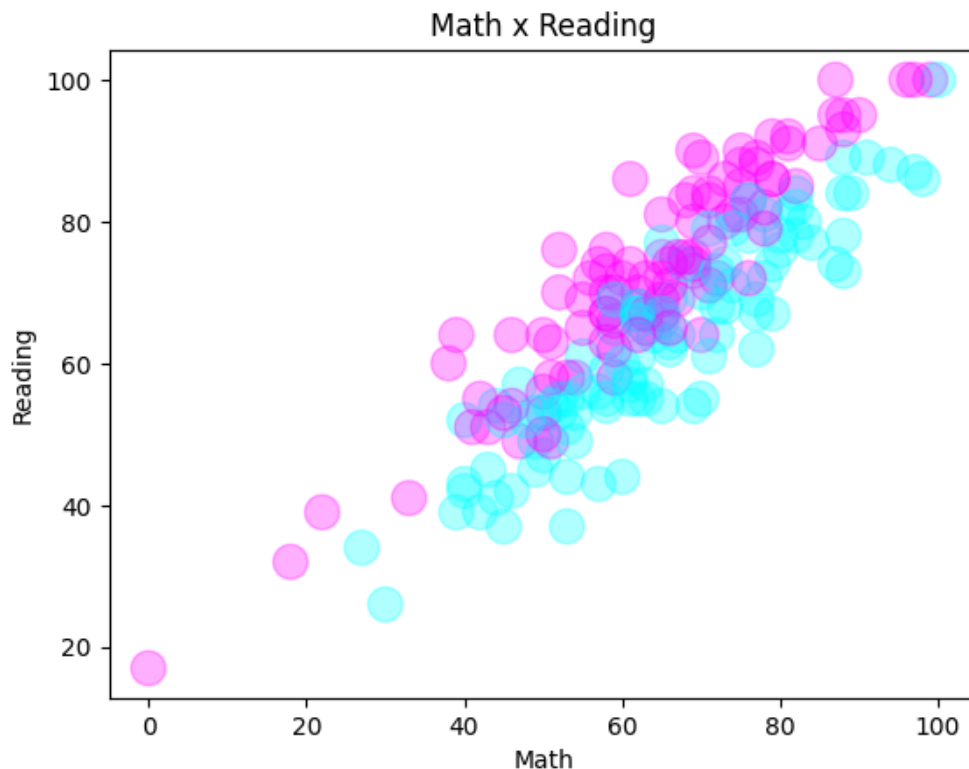
- 2.) Create a scatterplot of "math_score" and "reading_score". Set the alpha to .3 and the size to 200. Set the color to the "gender" column and set the colormap (cmap) to "cool". Set the xlabel to "Math", the ylabel to "Reading", and the title to "Math x Reading".

[Colormaps](#)

```
fig, ax = plt.subplots()

ax.scatter(math_score, reading_score, alpha=.3, s = 200, c = gender, cmap="cool")

ax.set_xlabel("Math")
ax.set_ylabel("Reading")
ax.set_title("Math x Reading");
```



✓ 3.) Create a figure with two axes side by side. Set the figsize to (15,5).

In the left axis, display a scatterplot of "math_score" and "reading_score". Set the alpha to .3 and the size to 300. Choose a custom color for the plot and set its title to "Math x Reading".

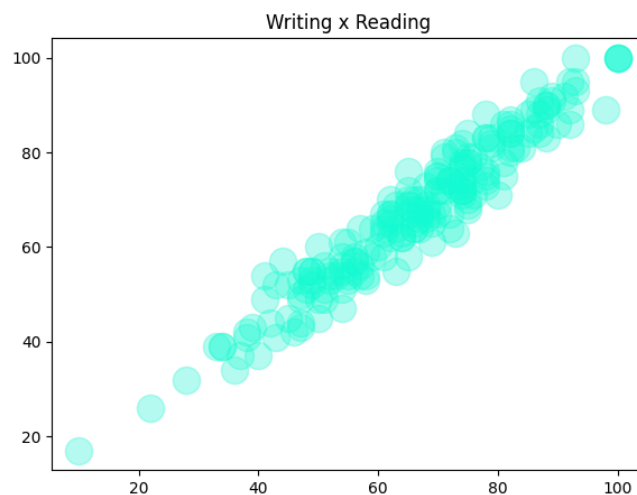
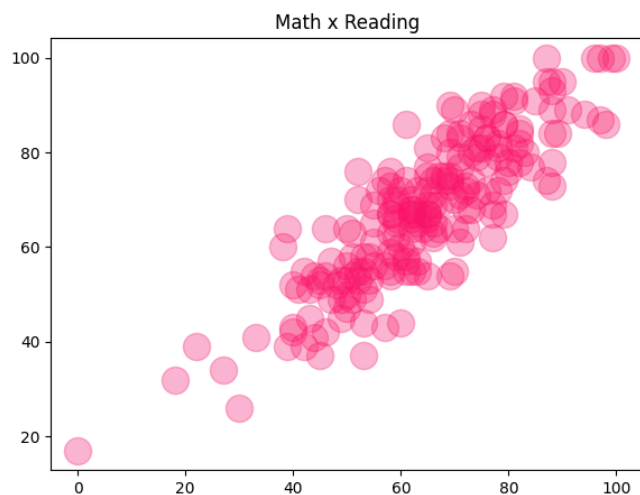
For the axis on the right, display a scatterplot of "writing_score" and "reading_score". Set the alpha to .3 and the size to 300. Choose a custom color for the plot and set its title to "Writing x Reading".

[HTML Color Codes](#)

```
fig, (left, right) = plt.subplots(1, 2, figsize = (15,5))

left.scatter(x=math_score, y=reading_score, alpha=.3, s = 300, c = "#FA166C")
right.scatter(x=writing_score, y=reading_score, alpha=.3, s = 300, c = "#16FAD4")

left.set_title("Math x Reading")
right.set_title("Writing x Reading");
```



- 4.) Create a figure with a single axis. Set the figsize to (10,5). Set the xlabel to "Reading Score", the ylabel to "Math/Writing Score", and the Title to "Test Scores". Display the two scatterplots from above within the axis. Add a legend to the upper center displaying "Math" and "Writing", respectively, in two columns.

See if you can gather an insight from the visualization.

```
fig, ax = plt.subplots(figsize = (10,5))

ax.scatter(x=math_score, y=reading_score, alpha=.3, s = 300, c = "#FA166C", label="Math")
ax.scatter(x=writing_score, y=reading_score, alpha=.3, s = 300, c = "#16FAD4", label="Writing")

ax.set_xlabel("Reading Score")
ax.set_ylabel("Math/Writing Score")
ax.set_title("Test Scores");

ax.legend(loc="upper center", ncol=2);
```

