

I T P 4 4 9

# Python Data Visualization

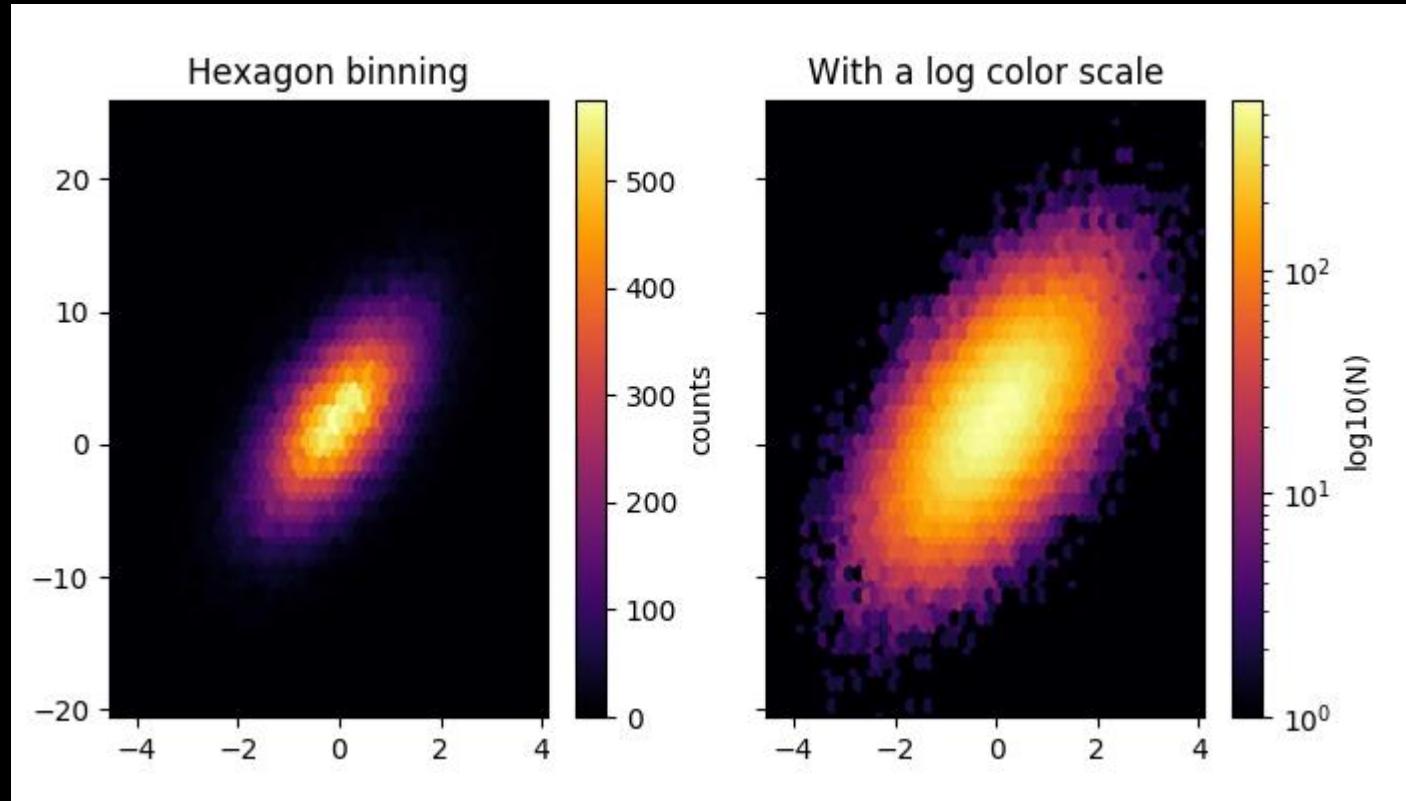
L e c t u r e 4



# Many Python packages for data visualization:

- Matplotlib
- Plotly
- Bokeh
- ggplot (also in R)
- Seaborn

# Matplotlib



[https://matplotlib.org/gallery/statistics/hexbin\\_demo.html#sphx-glr-gallery-statistics-hexbin-demo-py](https://matplotlib.org/gallery/statistics/hexbin_demo.html#sphx-glr-gallery-statistics-hexbin-demo-py)

We will use the *pyplot* sub library of matplotlib  
Matplotlib has many similarities to MATLAB  
(used extensively in engineering)

- Install *matplotlib*
- Import *matplotlib.pyplot*
- Alias *pyplot*

*import matplotlib.pyplot as plt*

- We control what and how data should be displayed using *pyplot*
- *pyplot* has many functions
  - *plt.plot(x-axis data, y-axis data) #add axes data to the plot*
  - *plt.show() #display the plot*
- You can add plots to the same chart by issuing the *plt.plot* command again
- <https://matplotlib.org/>

## 1: Project

Pro... + -

- ITP449\_Fall2020 C:\Us  
  Class  
  .idea  
  In Class Coding
  - drivers.csv
  - in\_class\_coding.py
  - salary.csv
- > venv library root

External Libraries

Scratches and Console

Run: in\_class\_coding

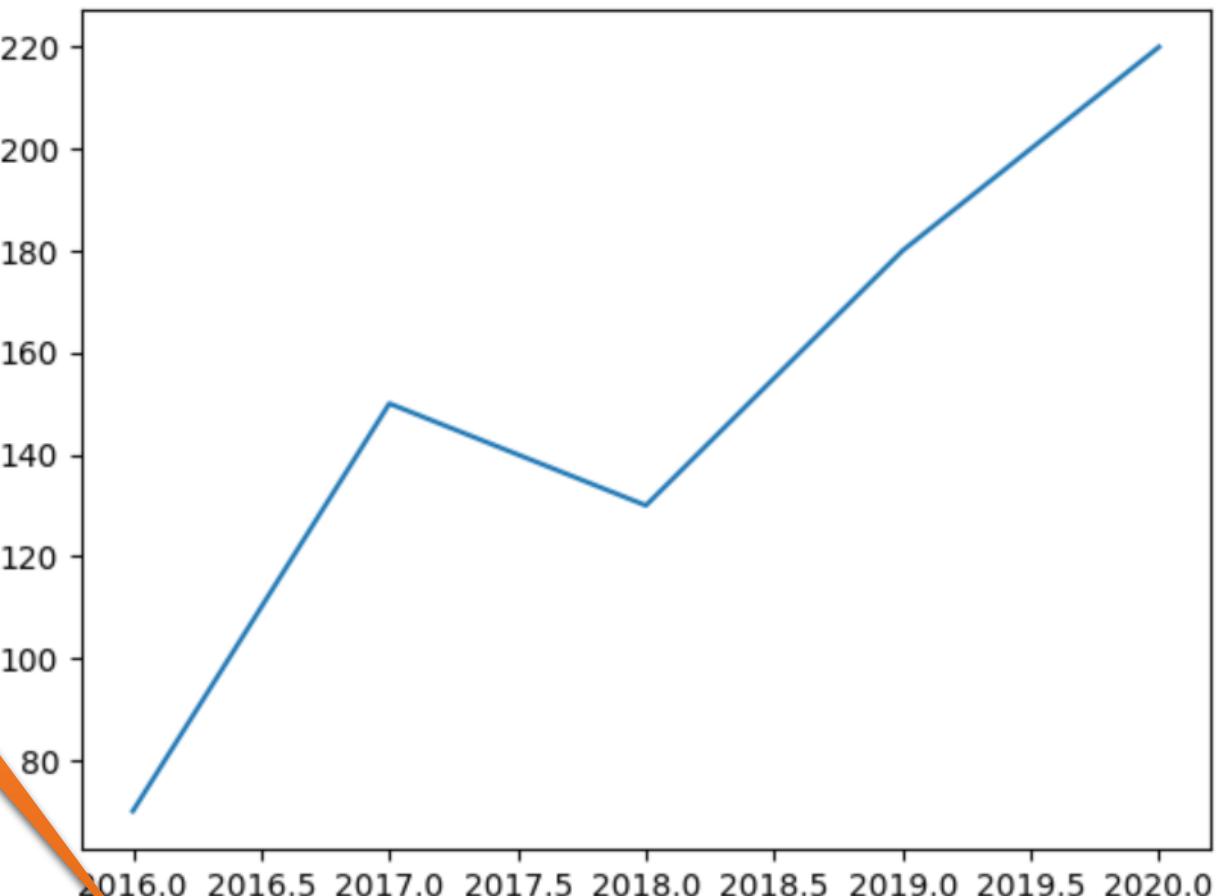
C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

## 2: Structure

## 2: Favorites

Figure 1

— □ ×



Controls



## Various options for plotting

- Change the marker
- `plt.plot(year, ITP449, marker = "o")`
- [https://matplotlib.org/api/\\_as\\_gen/matplotlib.pyplot.plot.html#matplotlib.pyplot.plot](https://matplotlib.org/api/_as_gen/matplotlib.pyplot.plot.html#matplotlib.pyplot.plot)
- Find markers here [https://matplotlib.org/2.0.2/api/markers\\_api.html](https://matplotlib.org/2.0.2/api/markers_api.html)
- Similarly you can change the line style `plt.plot(year, ITP449, marker = "o", linestyle = "- -")`
- Similarly you can set the color `plt.plot(year, ITP449, marker = "o", linestyle = "- -", color = "r")`
- [https://matplotlib.org/2.0.2/api/colors\\_api.html](https://matplotlib.org/2.0.2/api/colors_api.html)
- `plt.plot(year, ITP449, "ro--")`

# **Line Formatting, Legend, Labels**

# Color

b blue

g green

r red

c cyan

m magenta

y yellow

k black

w white

#000000 black

#ffffff white

#...

# Line Style



# Marker Style

marker	symbol
" . "	●
" , "	,
" o "	○
" v "	▼
" ^ "	▲
" < "	◀
" > "	▶
" 1 "	Y
" 2 "	L
" 3 "	L
" 4 "	Y
" 8 "	●
" s "	■
" p "	◆
" P "	+■
" * "	★
" h "	◆
" H "	◆



Version 3.3.1

Fork me on GitHub

[Installation](#) [Documentation](#) [Examples](#) [Tutorials](#) [Contributing](#)

[Search](#)

[home](#) | [contents](#) » [API Overview](#) » [matplotlib.lines](#) » [matplotlib.lines.Line2D](#)

[previous](#) | [next](#) | [modules](#) | [index](#)

## matplotlib.lines.Line2D

```
class matplotlib.lines.Line2D(xdata, ydata, linewidth=None, linestyle=None, color=None, marker=None, markersize=None, markeredgewidth=None, markeredgecolor=None, markerfacecolor=None, markerfacecoloralt='none', fillstyle=None, antialiased=None, dash_capstyle=None, solid_capstyle=None, dash_joinstyle=None, solid_joinstyle=None, pickradius=5, drawstyle=None, markevery=None, **kwargs)
```

[source]

Bases: [matplotlib.artist.Artist](#)

A line - the line can have both a solid linestyle connecting all the vertices, and a marker at each vertex. Additionally, the drawing of the solid line is influenced by the drawstyle, e.g., one can create "stepped" lines in various styles.

Create a [Line2D](#) instance with x and y data in sequences of `xdata`, `ydata`.

Additional keyword arguments are [Line2D](#) properties:

Property	Description
<code>agg_filter</code>	a filter function, which takes a (m, n, 3) float array and a dpi value, and returns a (m, n, 3) array
<code>alpha</code>	float or None
<code>animated</code>	bool
<code>antialiased</code> or <code>aa</code>	bool
<code>clip_box</code>	<a href="#">Bbox</a>
<code>clip_on</code>	bool
<code>clip_path</code>	Patch or (Path, Transform) or None
<code>color</code> or <code>c</code>	color
<code>contains</code>	unknown
<code>dash_capstyle</code>	{'butt', 'round', 'projecting'}

Table of Contents

[matplotlib.lines.Line2D](#)

- [Examples using matplotlib.lines.Line2D](#)

Related Topics

Documentation overview

- [API Overview](#)
  - [matplotlib.lines](#)
    - Previous: [matplotlib.lines](#)
    - Next: [matplotlib.lines.VertexSelector](#)

Show Page Source

## Various options for plotting

- Add *labels*
  - `plt.xlabel('Year')`
  - `plt.ylabel('Enrollment')`
- Add title
  - `plt.title('ITP Enrollment')`
- Control the axes range
  - `plt.axis([xmin, xmax, ymin, ymax])`

## 1: Project

- Pr... ▾
- + ▾
- in\_class\_coding.py ×

- ITP449\_Fall2020 ▾
  - Class
  - .idea
  - In Class C
    - drivers.
    - grade5Sc
    - in\_class
    - rainfall
    - salary.c
  - venv libra
- External Librar
- Scratches and C

```
1 import matplotlib.pyplot as plt
2
3 year = [2016, 2017, 2018, 2019, 2020]
4 ITP449 = [70, 150, 130, 180, 220]
5 plt.plot(year, ITP449, marker='o', linestyle='--', color='r')
6 # Alternative
7 # plt.plot(year, ITP449, 'ro--')
8 plt.xlabel('Year')
9 plt.ylabel('Enrollment')
10 plt.title('ITP Enrollment')
11 plt.axis([2014, 2022, 0, 300])
12 plt.show()
13
```

A 12 X 58 ^ v

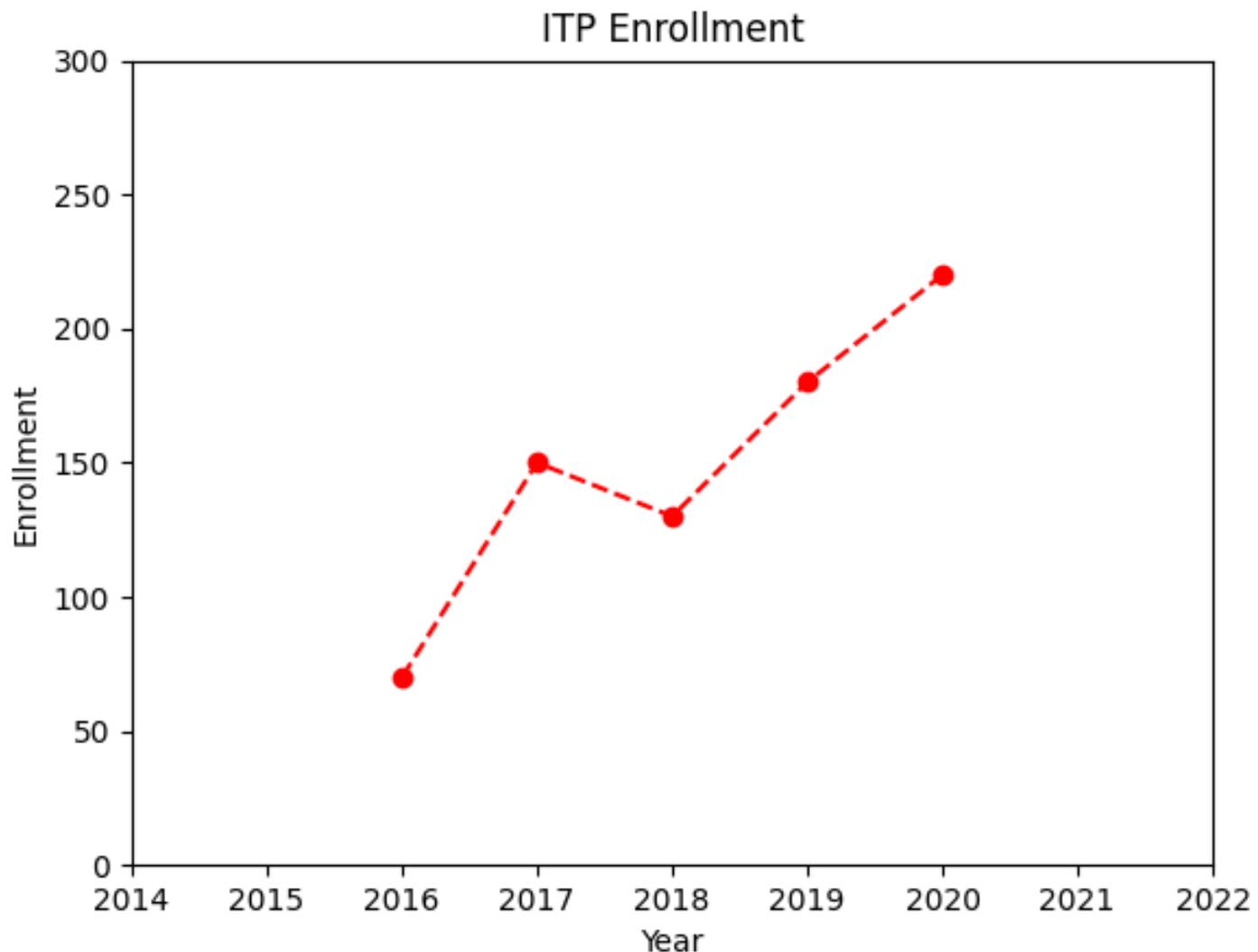
## 2: Structure

## 3: Favorites

Run: in\_class\_coding ×

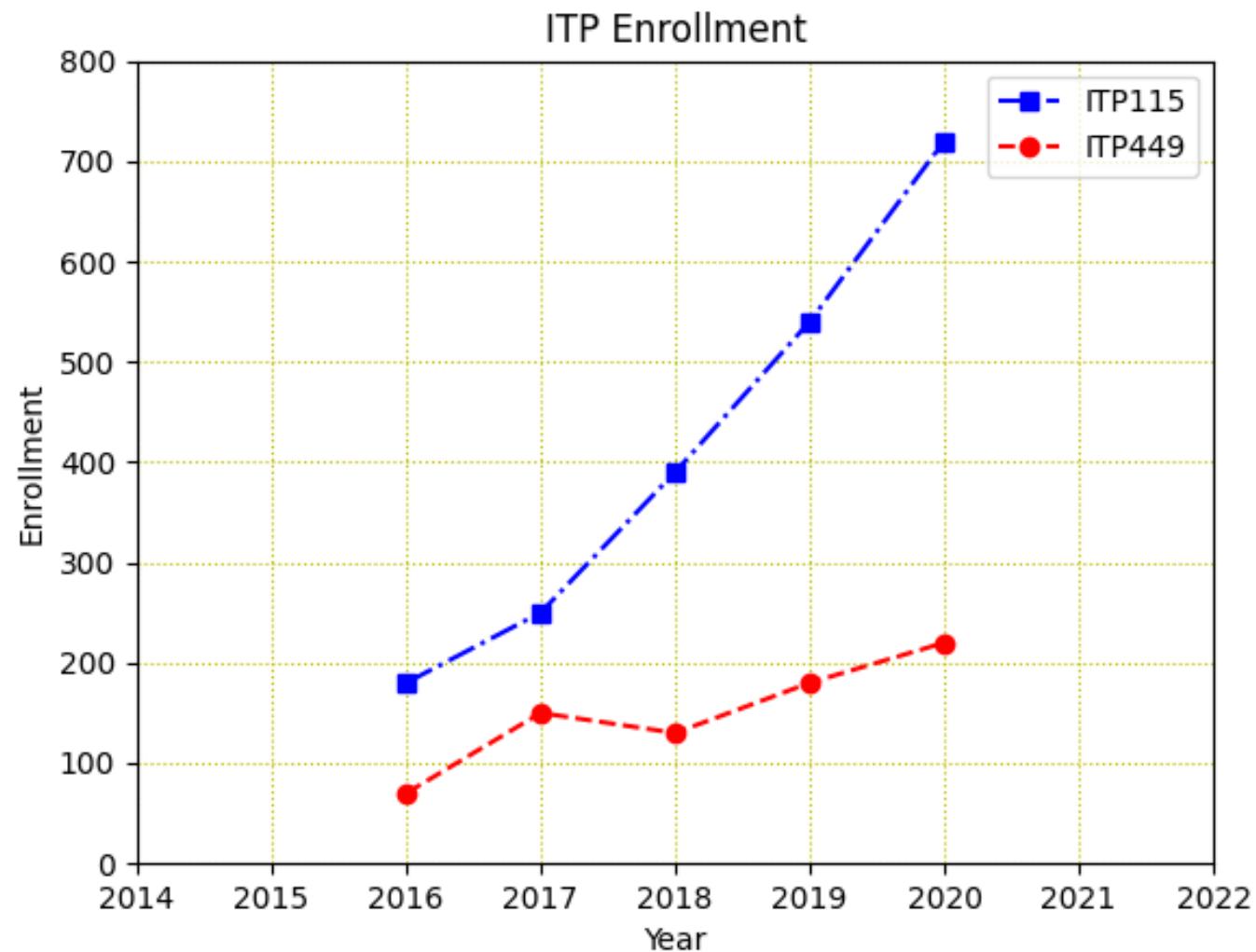
C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code 0



1: Project 1: Structure 1: Favorites

1 import matplotlib.pyplot as plt  
2  
3 year = [2016, 2017, 2018, 2019, 2020]  
4 ITP115 = [180, 250, 390, 540, 720]  
5 ITP449 = [70, 150, 130, 180, 220]  
6 plt.plot(year, ITP115, marker='s', linestyle='-.', color='b', label='ITP115')  
7 plt.plot(year, ITP449, marker='o', linestyle='--', color='r', label='ITP449')  
8 plt.xlabel('Year')  
9 plt.ylabel('Enrollment')  
10 plt.title('ITP Enrollment')  
11 plt.axis([2014, 2022, 0, 800])  
12 plt.legend()  
13 plt.grid(True, color='y', linestyle=':')14 plt.show()



Project

- ITP449\_Fall2020 C:\Us...
- Class
- .idea
- In Class Coding
  - drivers.csv
  - in\_class\_coding.py
  - salary.csv
- venv library root

External Libraries

Scratches and Console

Run: in\_class\_coding

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl...

Run

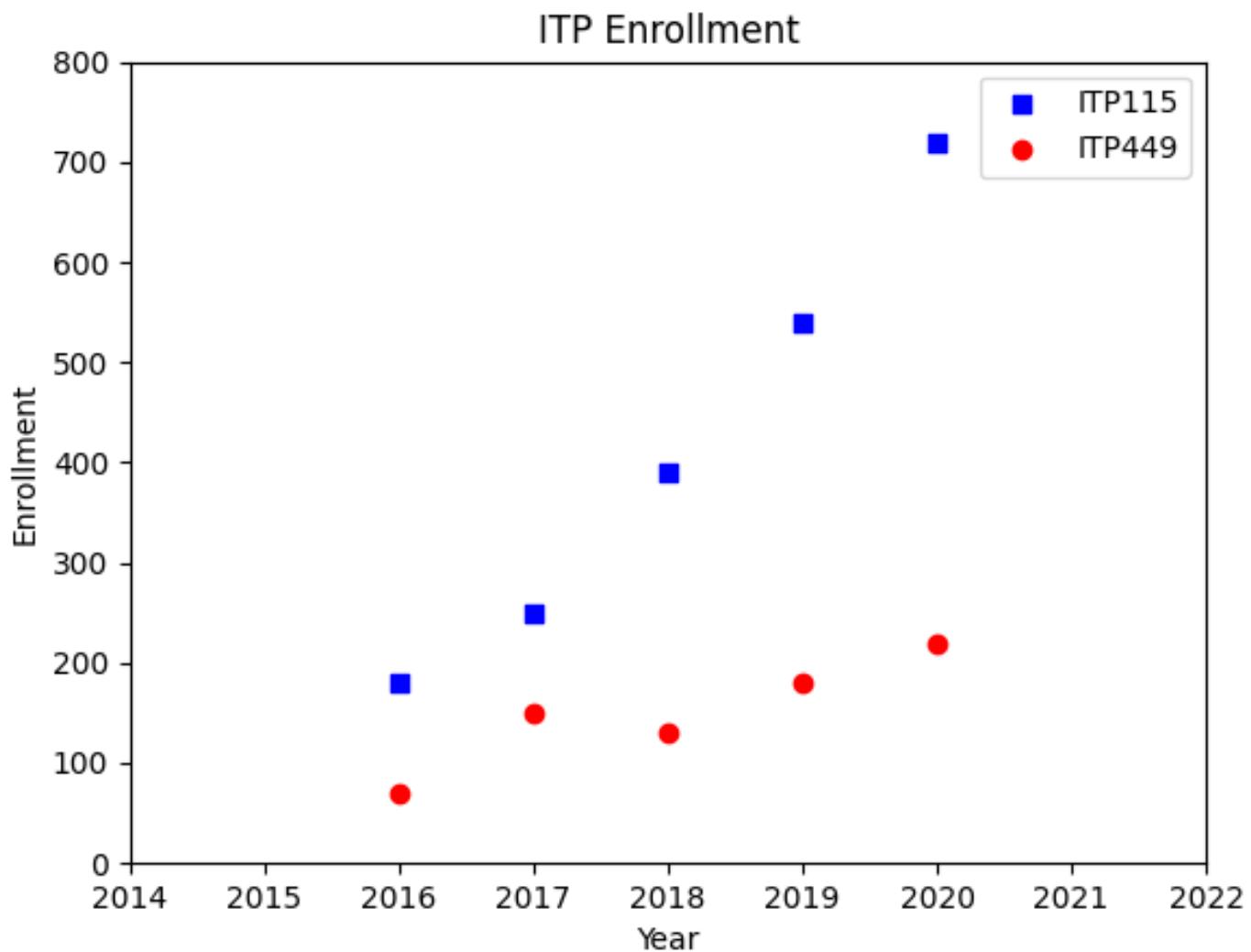
TODO

Problems

Terminal

Python Console

Event Log

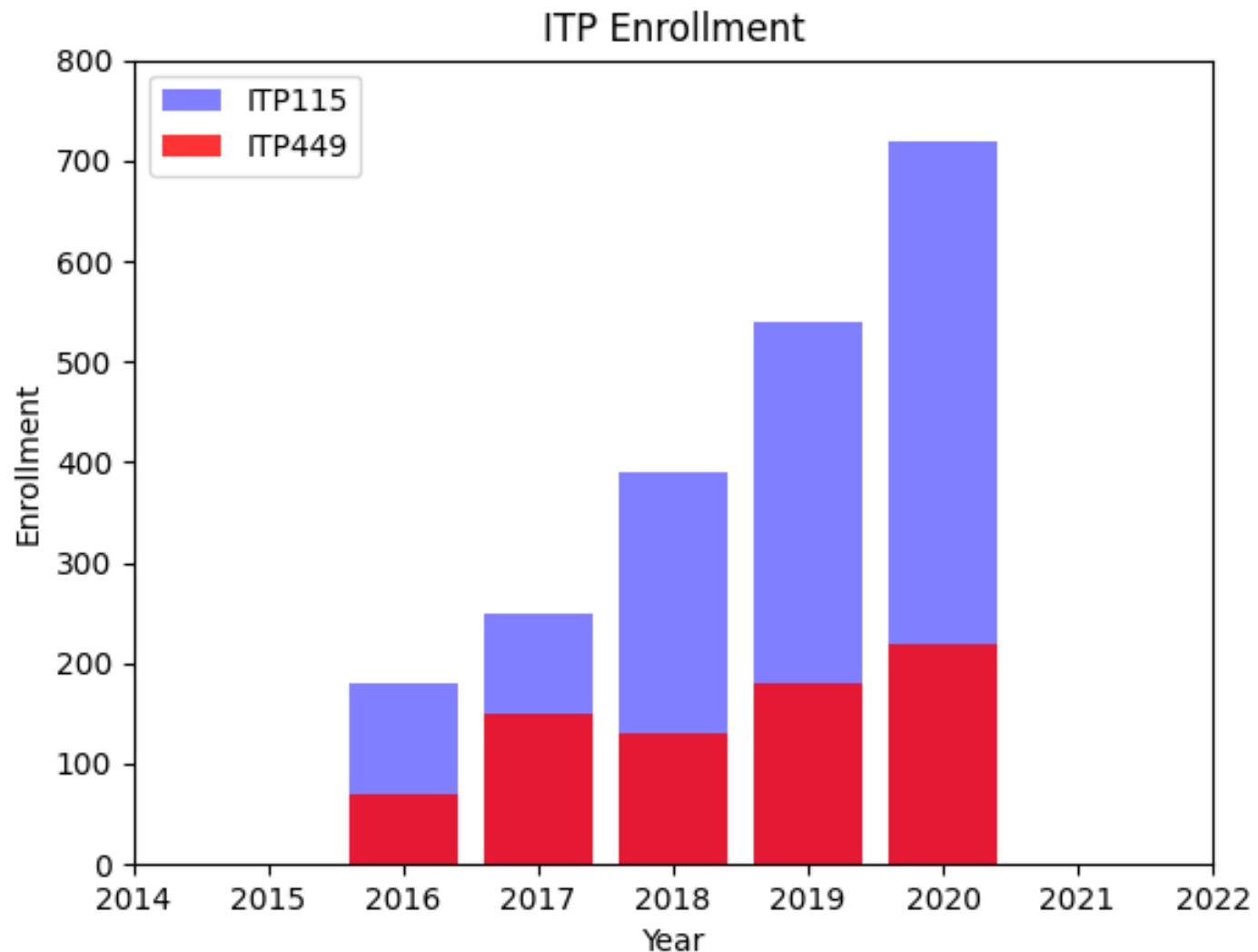


## 1: Project

 Pro...  
ITP449\_Fall2020 C:\Us  
  Class  
  .idea  
  In Class Coding  
    drivers.csv  
    in\_class\_coding.py  
    salary.csv  
  venv library root  
External Libraries  
Scratches and Console

## 2: Structure

## 3: Favorites





ITP449\_Fall2020 > Class > In Class Coding >  in\_class\_coding.py

in class coding ▾



1: Project

7: Structure

2: Favorites

```
import matplotlib.pyplot as plt

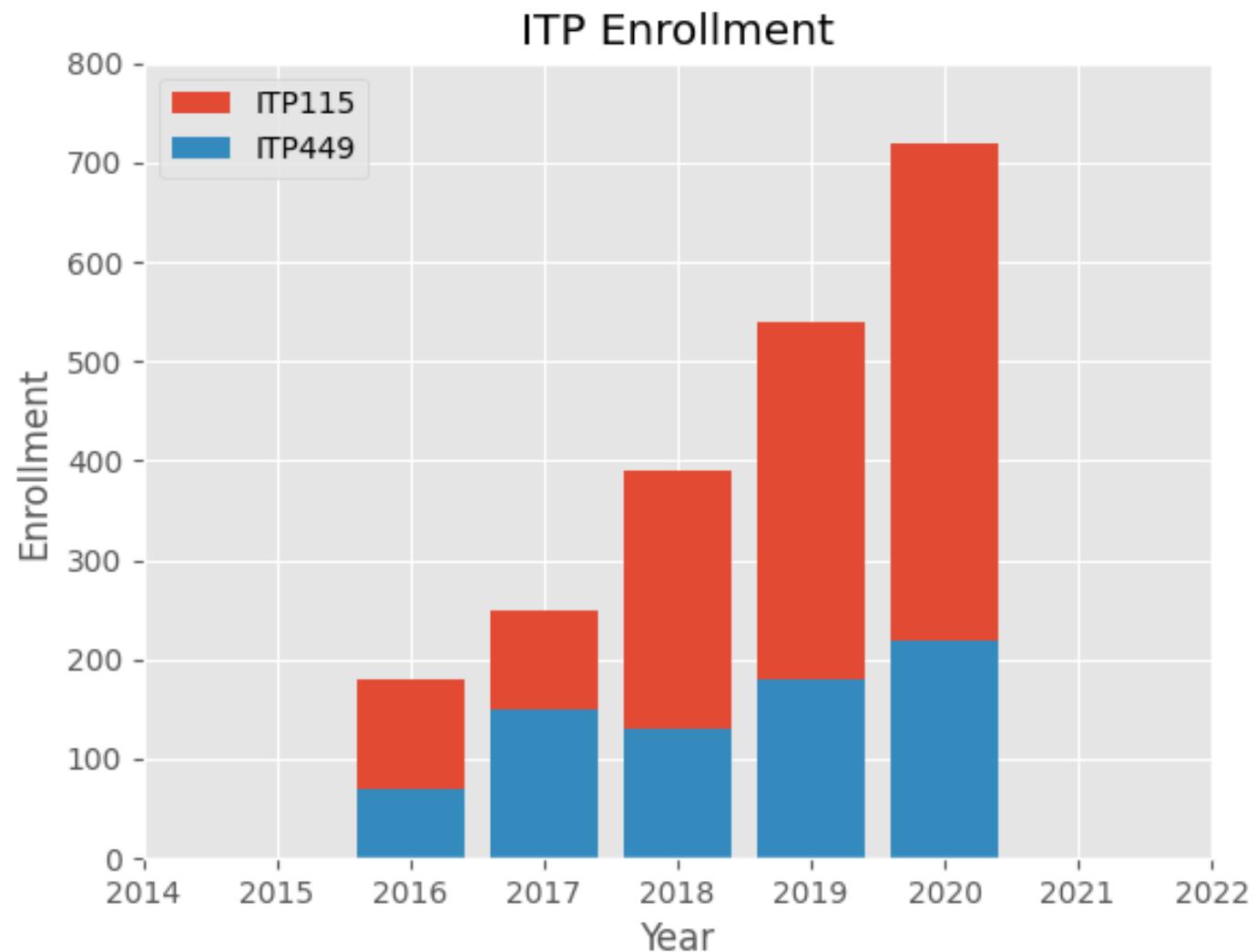
plt.style.use('ggplot')
year = [2016, 2017, 2018, 2019, 2020]
ITP115 = [180, 250, 390, 540, 720]
ITP449 = [70, 150, 130, 180, 220]
plt.bar(year, ITP115, label='ITP115')
plt.bar(year, ITP449, label='ITP449')
plt.xlabel('Year')
plt.ylabel('Enrollment')
plt.title('ITP Enrollment')
plt.axis([2014, 2022, 0, 800])
plt.legend()
plt.show()
```

Run:  in\_class\_coding >

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Packages installed successfully: Installed packages: 'seaborn' (today 6:59 PM)

15:1 CRLF UTF-8 4 spaces Python 3.8 (Class) 



Pr... in\_class\_coding.py

```
1 import matplotlib.pyplot as plt
2
3 plt.style.use('ggplot')
4 rainfall = [15, 19, 18, 13, 11, 7, 8, 7, 10, 12, 14, 15]
5 months = ['Jan', 'Feb', 'Mar', 'Apr', 'May', 'Jun', 'Jul', 'Aug', 'Sep', 'Oct', 'Nov', 'Dec']
6 plt.bar(months, rainfall)
7 plt.xticks(rotation=45)
8 plt.xlabel('Month')
9 plt.ylabel('Rainfall (cm)')
10 plt.title('Rainfall')
11 plt.savefig('rainfall.png', bbox_inches='tight')
12 plt.show()
13
```

A 11 X 57

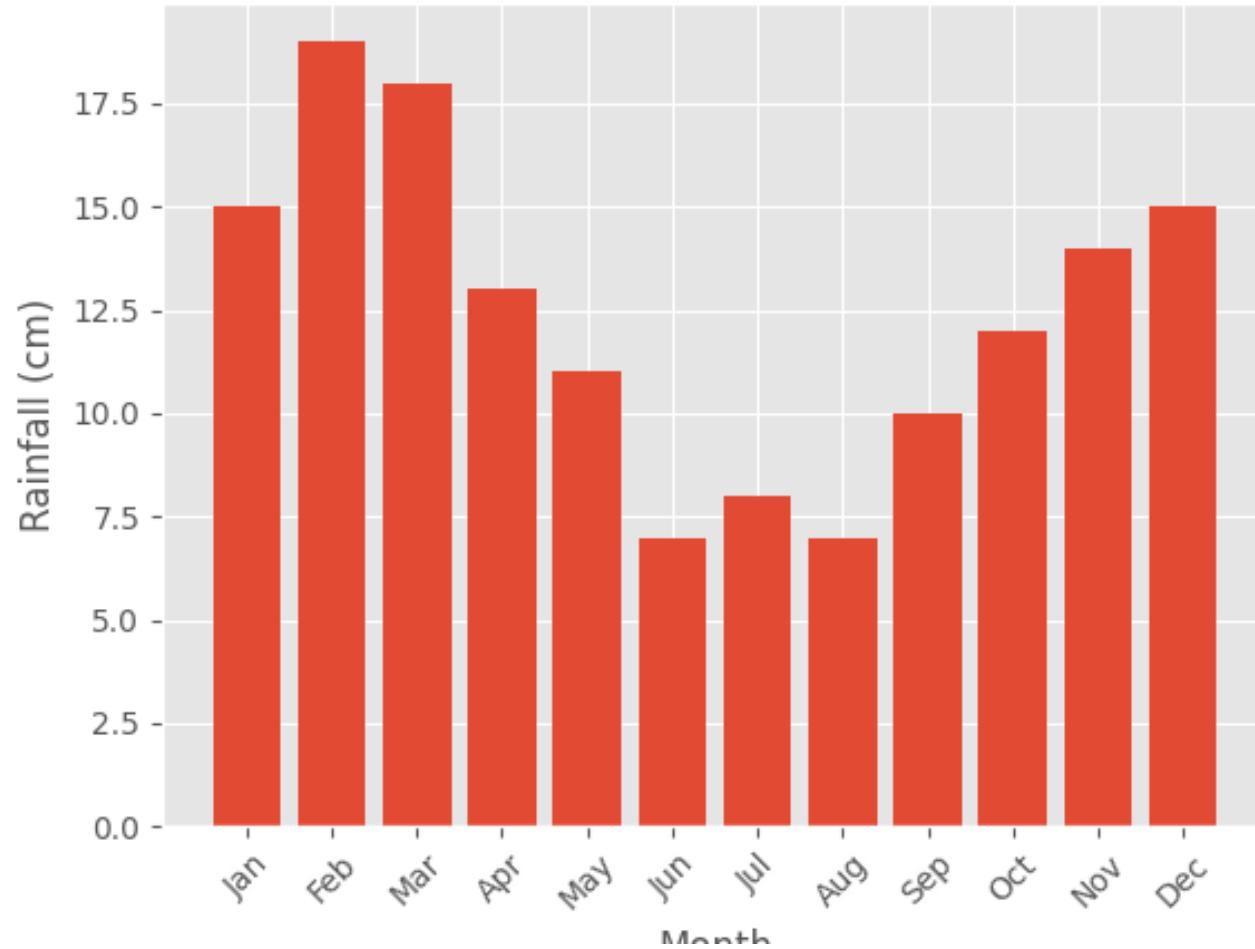
External L Scratches

Run: in\_class\_coding

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl"

Process finished with exit code 0

## Rainfall



## 1: Project

## 2: Structure

## 2: Favorites

Pr... + | in\_class\_coding.py x

```
1 import matplotlib.pyplot as plt
2
3 plt.style.use('ggplot')
4 labels = ['Chrome', 'Firefox', 'Internet Explorer', 'Safari', 'Others']
5 marketShare = [61.64, 18.98, 11.02, 4.23, 4.13]
6 plt.pie(marketShare, labels=labels, autopct='%.1f%%', startangle=45)
7 plt.title('Browser Market Share')
8 plt.axis('equal')
9 plt.show()
10
11
12
```

A 11 X 58 ^ v

Run: in\_class\_coding x



C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code -1



Run

TODO

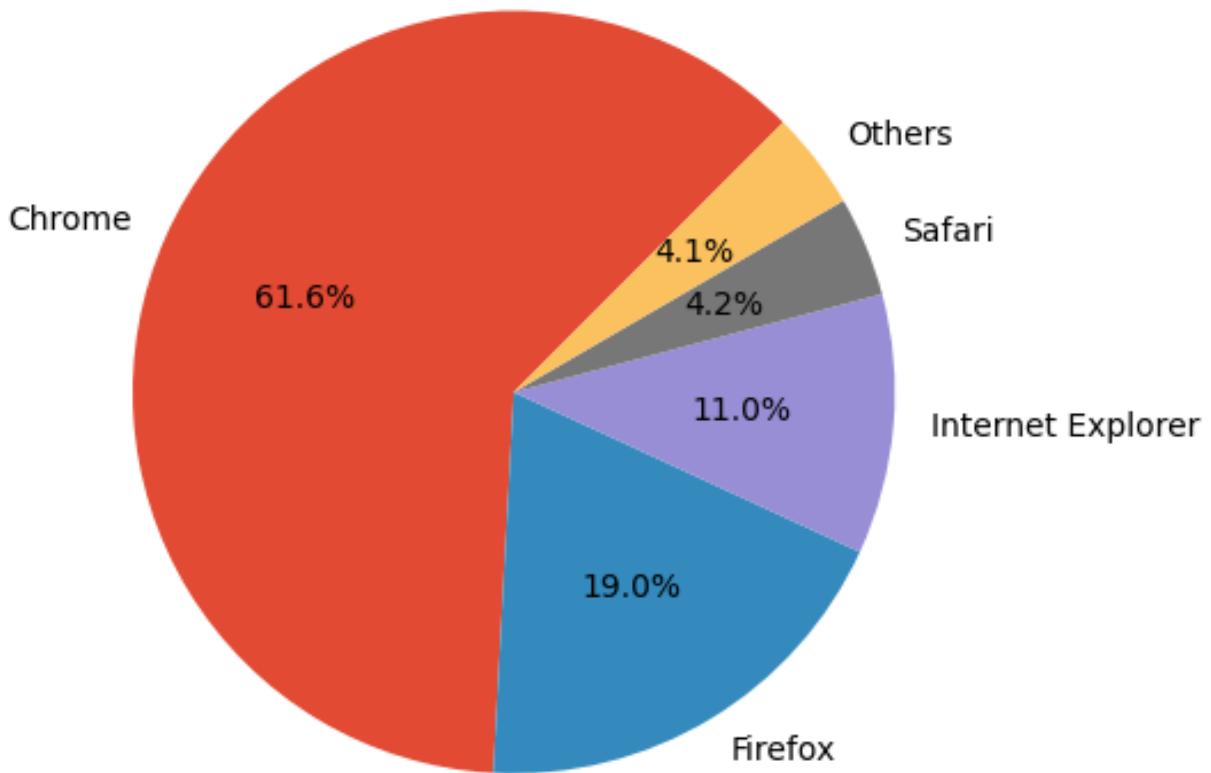
Problems

Terminal

Python Console

Event Log

### Browser Market Share



## 1: Project

 Pr... in\_class\_coding.py x  
ITP449\_Fall2020  
  Class  
    .idea  
  In Class Coding  
    drivers  
    grade5S  
    in\_class\_coding.py  
    ITPEnr.  
    rainfall  
    salary.  
  venv libra...  
External Libraries  
Scratches and C...

## 2: Structure

```
1 import matplotlib.pyplot as plt
2
3 plt.style.use('seaborn')
4 labels = ['Chrome', 'Firefox', 'Internet Explorer', 'Safari', 'Others']
5 marketShare = [61.64, 18.98, 11.02, 4.23, 4.13]
6 plt.pie(marketShare, labels=labels, autopct='%.1f%%', startangle=45,
7         explode=[0, 0, 0, 0.5, 0.5])
8 plt.title('Browser Market Share')
9 plt.axis('equal')
10 plt.legend(labels, loc='lower left')
11 plt.show()
12
```

A 11 X 58 ^ v

## 2: Favorites

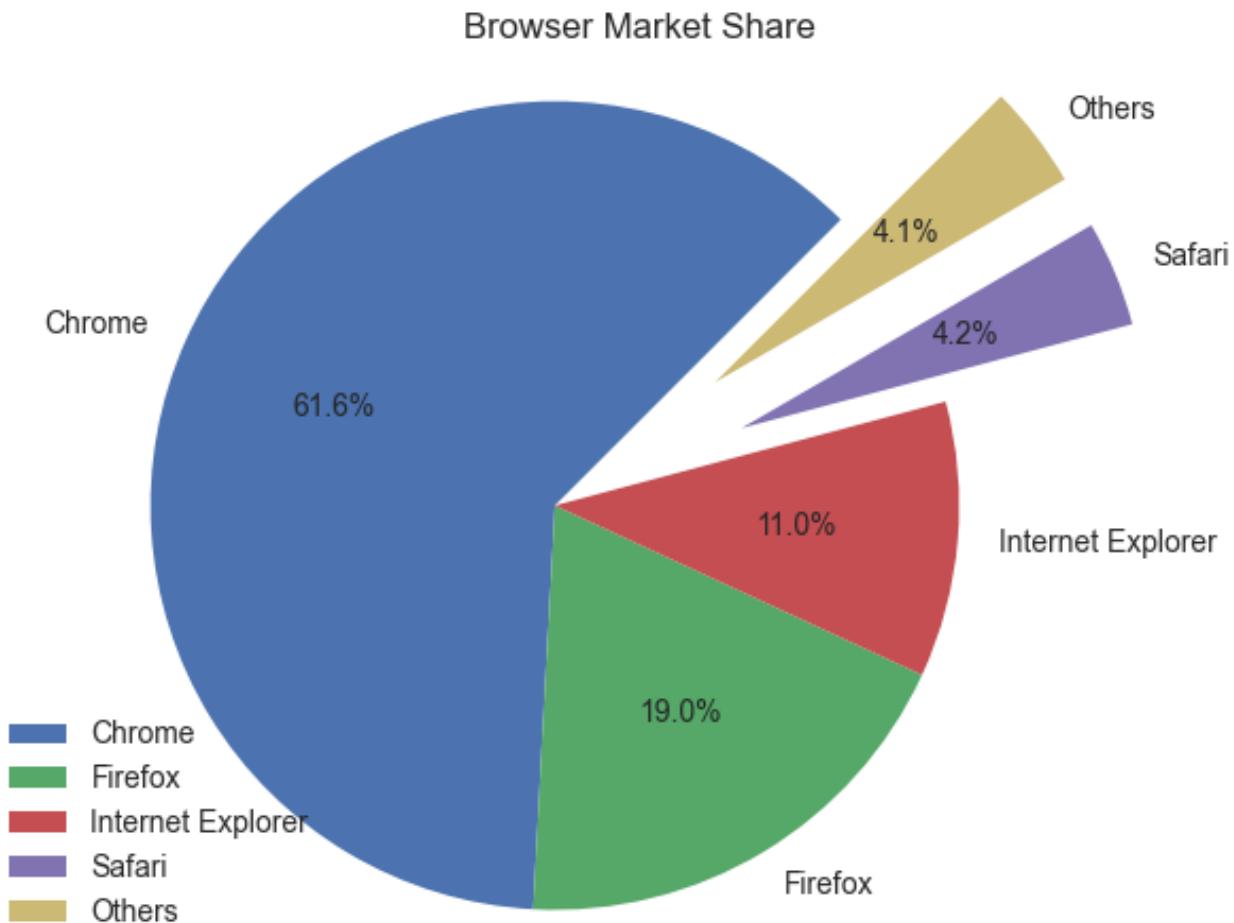
Run: in\_class\_coding x

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl...

Process finished with exit code 0

## 3:

Run TODO Problems Terminal Python Console Event Log



1: Project 1: Structure 1: Favorites

1 in\_class\_coding.py x

ITP449\_Fall2020

- Class
- .idea
- In Class C
  - drivers.
  - grade5Sc
  - in\_class
  - rainfall
  - salary.c
- venv libra

External Librar

Scratches and C

```
1 import matplotlib.pyplot as plt
2
3 plt.style.use('ggplot')
4 plt.plot([0, 1, 2, 3, 4], [0, 1, 8, 27, 64])
5 plt.axis([0, 4, 0, 64])
6 plt.xlabel('X')
7 plt.ylabel('Y')
8 plt.title('X^3')
9 plt.show()
10
```

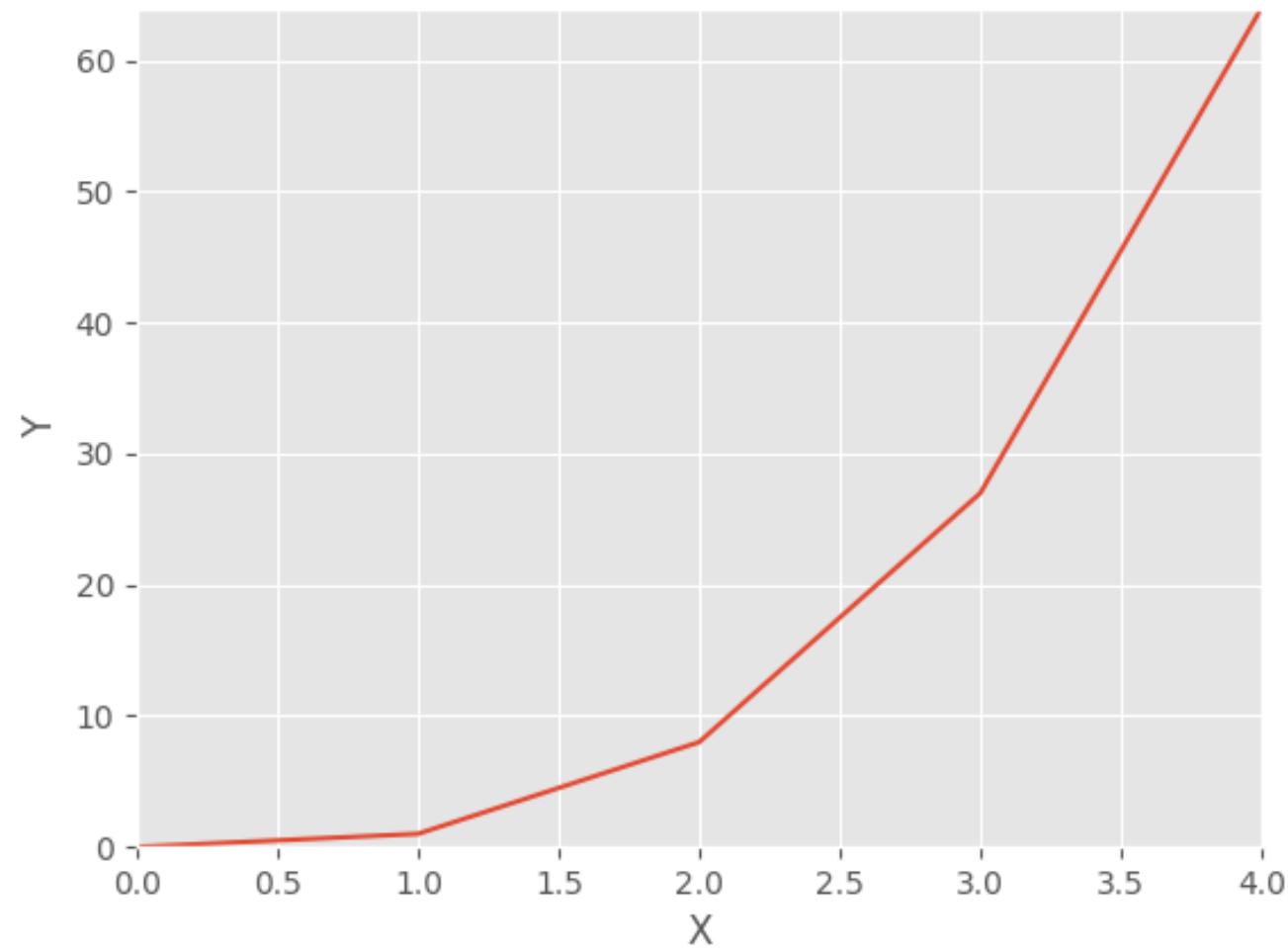
Run: in\_class\_coding x



```
C:\Users\Reza\Desktop\ITP449_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449_Fall2020/Cl"
```

```
Process finished with exit code 0
```

$X^3$



## 1: Project

ITP449\_Fall2020  
  Class  
  .idea  
  In Class Coding  
    drivers.  
    grade5Sc  
    in\_class\_coding.py  
    rainfall  
    salary.c  
  venv libra  
External Librar  
Scratches and C

## 2: Structure

ITP449\_Fall2020  
  Class  
  .idea  
  In Class Coding  
    drivers.  
    grade5Sc  
    in\_class\_coding.py  
    rainfall  
    salary.c  
  venv libra  
External Librar  
Scratches and C

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 plt.style.use('ggplot')
5 x = np.arange(0, 4.1, 0.1)
6 plt.plot(x, x**3)
7 plt.axis([0, 4, 0, 64])
8 plt.xlabel('X')
9 plt.ylabel('Y')
10 plt.title('X^3')
11 plt.show()
12
```

A 12 X 62 ^ v

## Run: in\_class\_coding

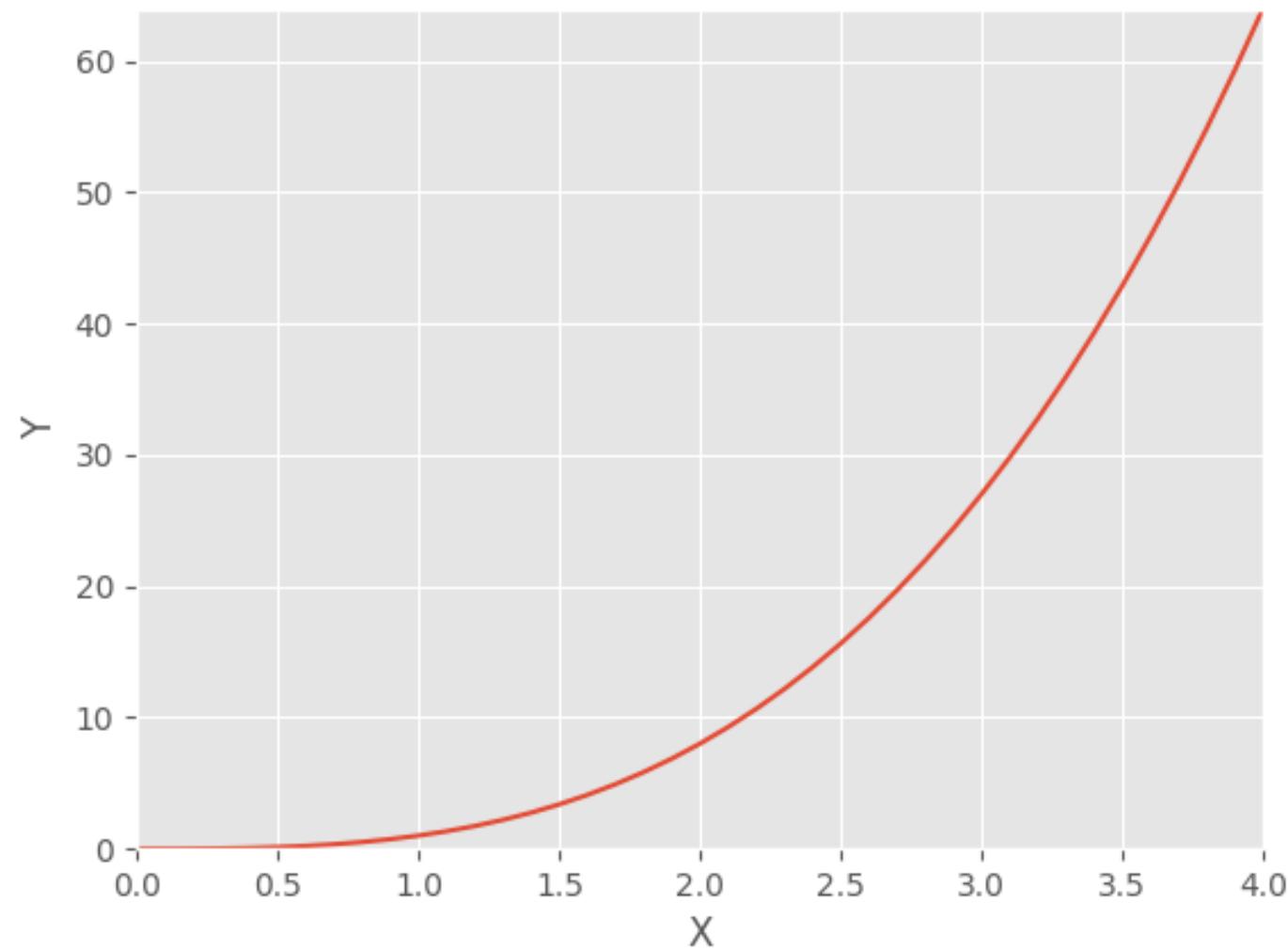


```
C:\Users\Reza\Desktop\ITP449_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449_Fall2020/Cl"
```

```
Process finished with exit code 0
```

## 2: Favorites

$X^3$



## 1: Project

## 2: Structure

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 plt.style.use('ggplot')
5 x = np.arange(-4, 4.1, 0.1)
6 plt.plot(x, x**2, x, x**3, x, x**4)
7 plt.axis([-4, 4, -70, 70])
8 plt.xlabel('X')
9 plt.ylabel('Y')
10 plt.legend(['x^2', 'x^3', 'x^4'])
11 plt.show()
12
```

A 12 X 54 ^ v

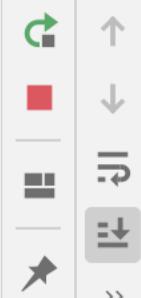
## Run:

in\_class\_coding



C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

## Favorites



## 2: Favorites



Run

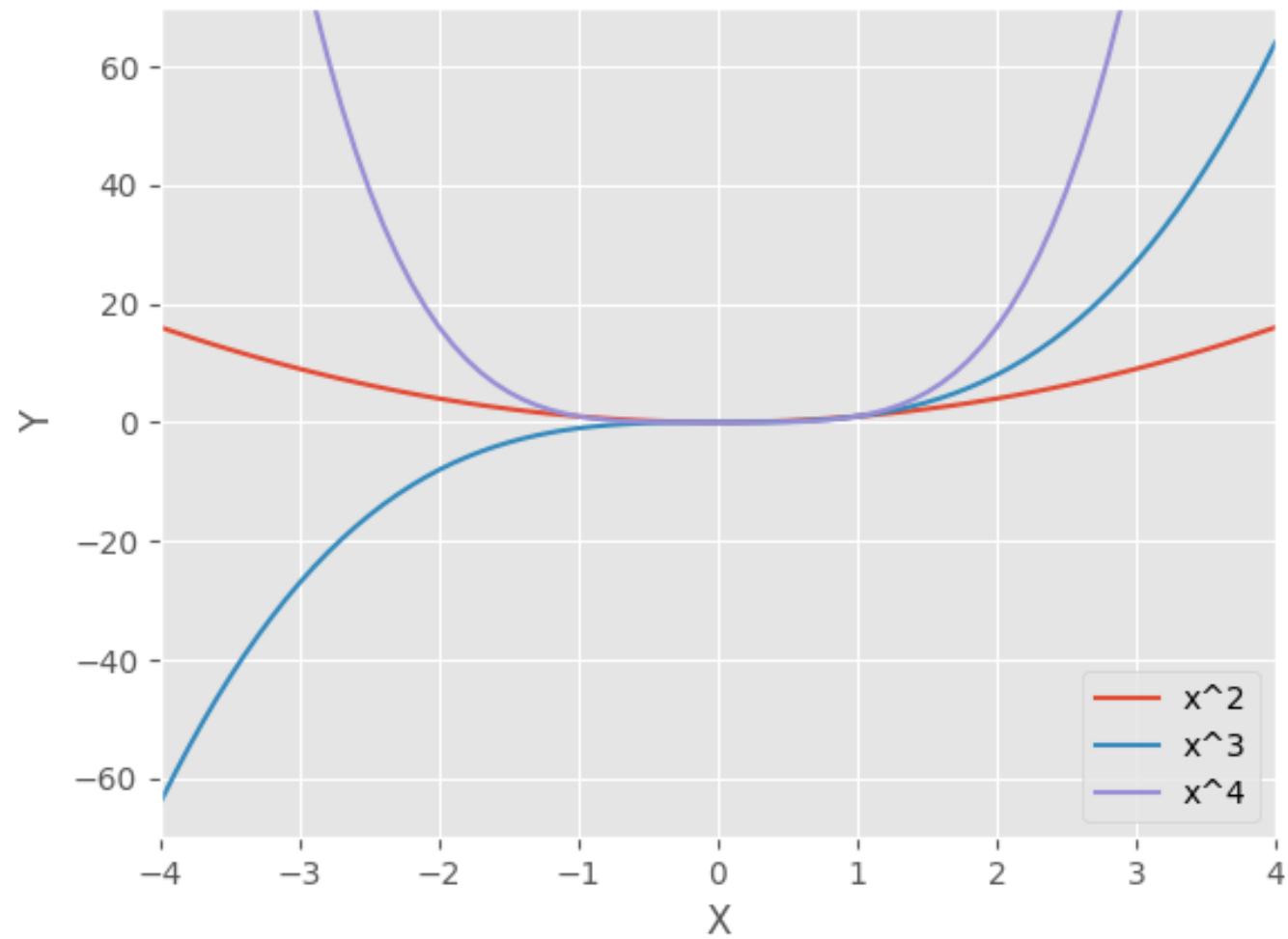
TODO

Problems

Terminal

Python Console

Event Log



# Subplots

- A set of plots. Each plot focuses on a slice of the overall data.
- An array of plots
- Each plot has an index so that each plot can be addressed
- *plt.subplot(# rows of plots, # columns of plots, chart #)*

*plt.subplot(2, 2, 1)*

*plt.subplot(2, 2, 2)*

*plt.subplot(2, 2, 3)*

*plt.subplot(2, 2, 4)*

ITP449\_Fall2020 > Class > In Class Coding >  in\_class\_coding.py

in\_class\_coding ▾



1: Project

7: Structure

Favorites

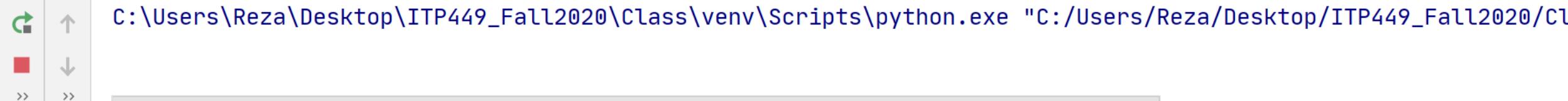
★

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 plt.style.use('ggplot')
5 x = np.arange(0, 4, 0.1)
6 plt.subplot(1, 2, 1) # 1 row, 2 cols, chart 1
7 plt.plot(x, x**0.5, 'y^', x, x**0.3, 'bo', x, x**0.25, 'r--')
8 plt.axis([0, 4, 0, 3])
9 plt.legend(['x^(1/2)', 'x^(1/3)', 'x^(1/4)'])
10 plt.subplot(1, 2, 2) # 1 row, 2 cols, chart 2
11 plt.plot(x, x**2, 'y^', x, x**3, 'bo', x, x**4, 'r--')
12 plt.axis([0, 4, 0, 70])
13 plt.legend(['x^2', 'x^3', 'x^4'])
14 plt.show()
```

Run:  in\_class\_coding >

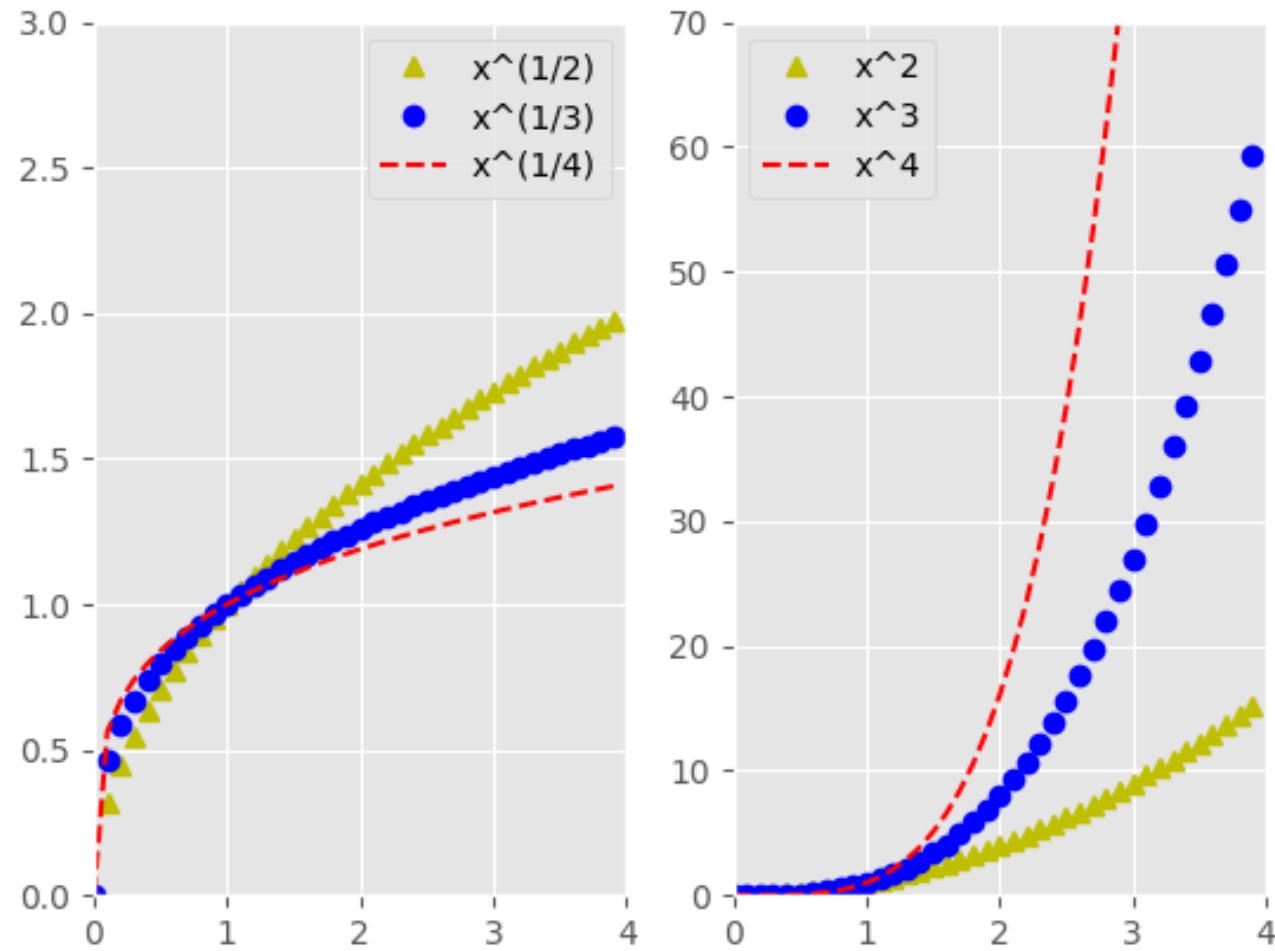


C:\Users\Reza\Desktop\ITP449\_Fall2020\class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl



Run TODO Problems Terminal Python Console Event Log

Packages installed successfully: Installed packages: 'sea... (yesterday 6:59 PM 15:1 CRLF UTF-8 4 spaces Python 3.8 (Class) ↵



# Subplots

- You can create a blank figure
- And then add subplots to it

ITP449\_Fall2020 > Class > In Class Coding >  in\_class\_coding.py

in\_class\_coding ▾



in\_class\_coding.py

```
1 import matplotlib.pyplot as plt
2
3 myFig = plt.figure()
4 ax1 = myFig.add_subplot(2, 2, 1)
5 ax2 = myFig.add_subplot(2, 2, 2)
6 ax3 = myFig.add_subplot(2, 2, 3)
7 plt.show()
```

A 9 ✓ 62 ^ <

Run:  in\_class\_coding

—

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl



٦٤

4: Run

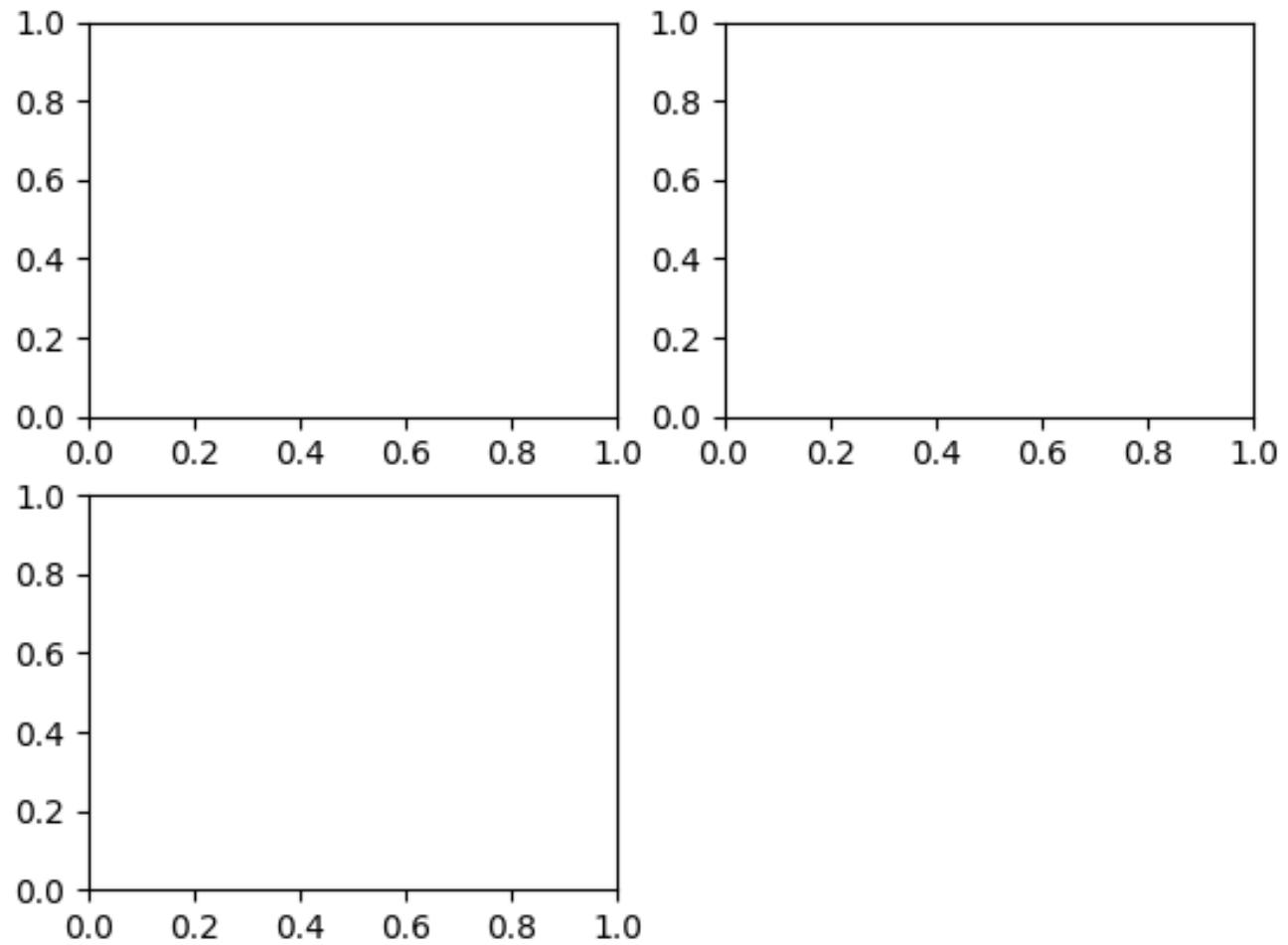
TODO ! 6:

## Problems

>\_ Terminal

Python Console

### Event Log



Pr... + in\_class\_coding.py x

ITP449\_Fall2020

- Class
- .idea
- In Class Coding
- dri
- gra
- in\_
- ITP
- rai
- sal

venv

External L

Scratches

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 myFig = plt.figure()
5 ax1 = myFig.add_subplot(2, 2, 1)
6 ax2 = myFig.add_subplot(2, 2, 2)
7 ax3 = myFig.add_subplot(2, 2, 3)
8 ax1.hist(np.random.randn(1000), bins=50, color='m')
9 ax2.scatter(np.arange(30), np.arange(30) + 2 * np.random.randn(30), color='b')
10 ax3.plot(np.random.randn(50).cumsum(), 'g--')
11 plt.show()
12
13
14
15
16
```

randn(50) generates a standard normal dist of 50 numbers.  
mean 0, variance 1

Adding some noise to data

Cumulative Sum

Run: in\_class\_coding

&gt;&gt;

-

Run

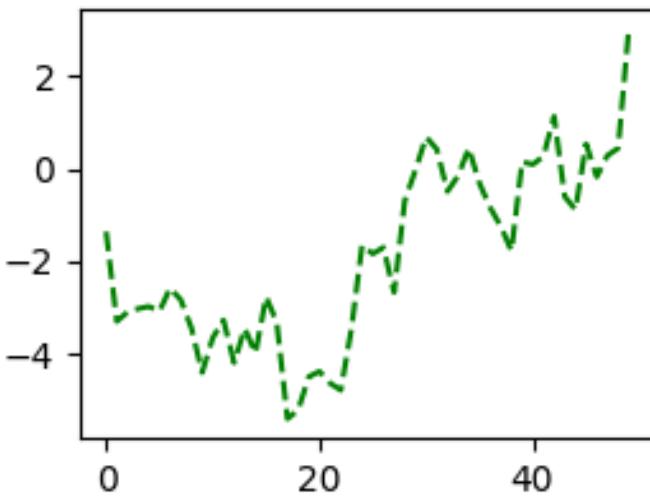
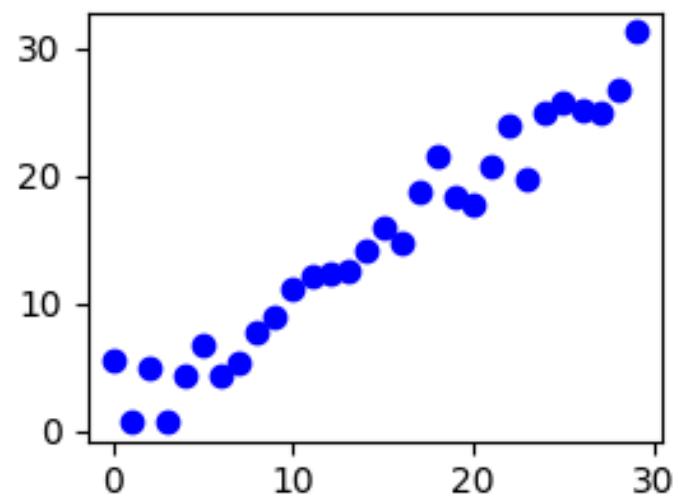
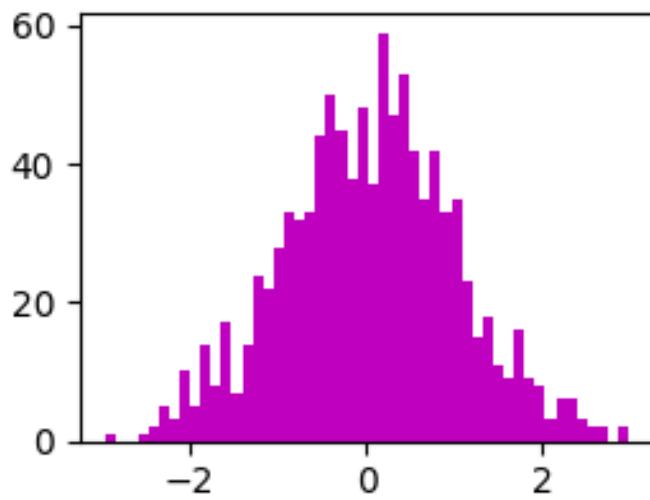
TODO

Problems

Terminal

Python Console

Event Log



1: Project 1: Structure 2: Favorites

1 import matplotlib.pyplot as plt  
2 import numpy as np  
3  
4 myFig, myAxes = plt.subplots(2, 2)  
5 plt.show()

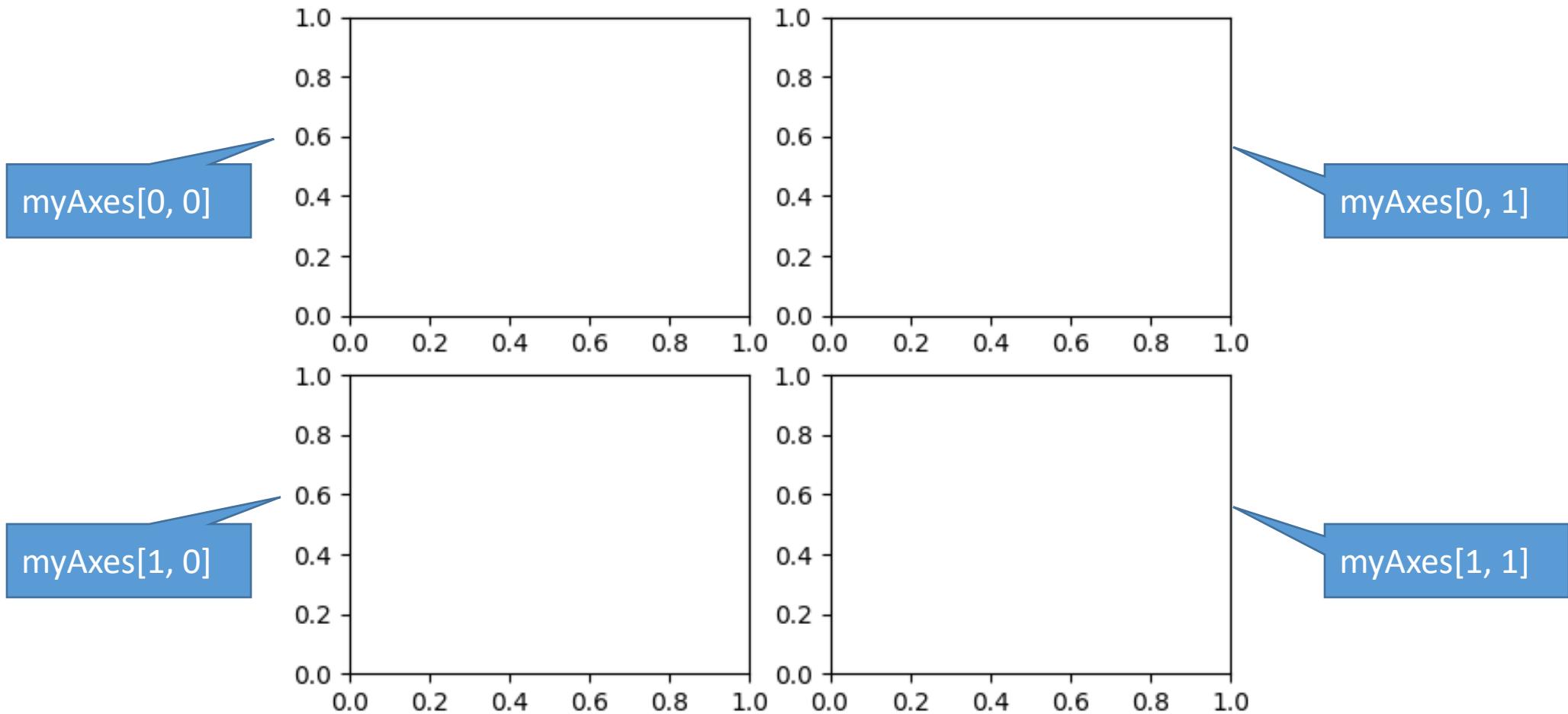
6  
7  
8  
9  
10  
11

Generates an array of subplots 2 x 2  
myFig refers to figure  
myAxes refers to axes  
myAxes[i, j] refers to a subplot

Run: in\_class\_coding

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl"

Process finished with exit code 0



## 1: Project

Pr... + in\_class\_coding.py x

- ITP449\_Fall2020
  - Class
  - .idea
  - In Class Coding
    - 7.p
    - dri
    - gra
    - in\_class\_coding.py
    - ITP449\_Fall2020
    - rai
    - sal
  - venv
- External L
- Scratches

## 2: Structure

```
1 import matplotlib.pyplot as plt
2 import numpy as np
3
4 myFig, myAxes = plt.subplots(2, 2, sharex=True, sharey=True)
5 for i in range(2):
6     for j in range(2):
7         myAxes[i, j].hist(np.random.randn(1000), bins=50, color='k', alpha=0.5)
8 plt.show()
```

Share x and y axes

## 2: Favorites

Run: in\_class\_coding



C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Class/in\_class\_coding.py"

Process finished with exit code 0

## 3:

Run

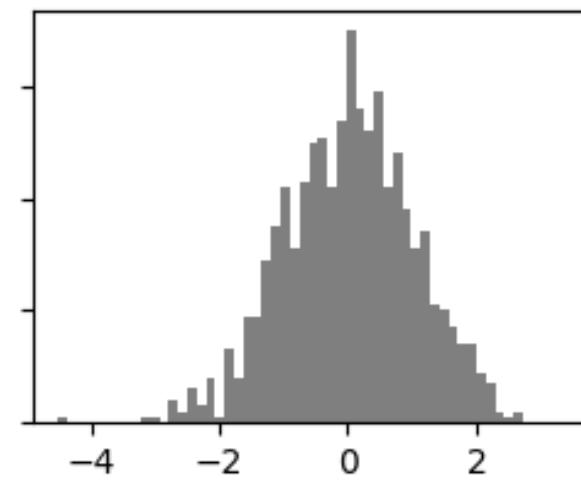
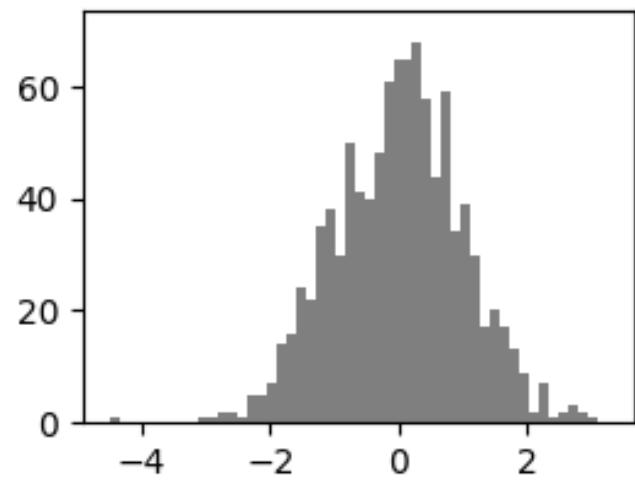
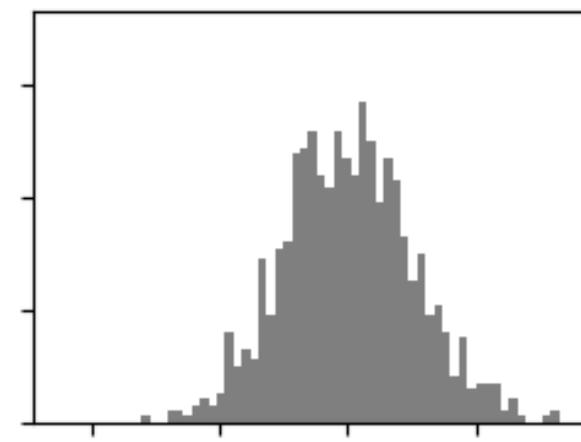
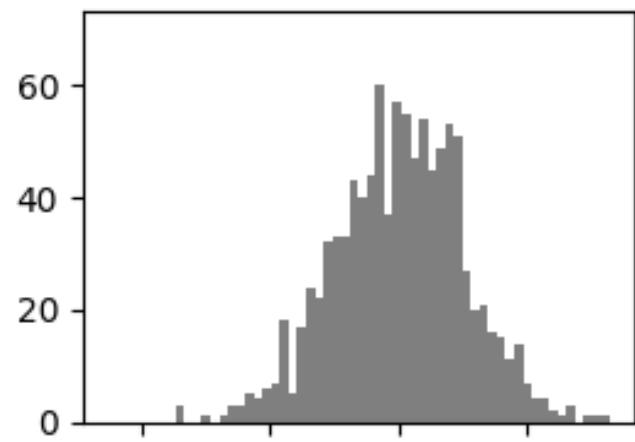
TODO

Problems

Terminal

Python Console

Event Log



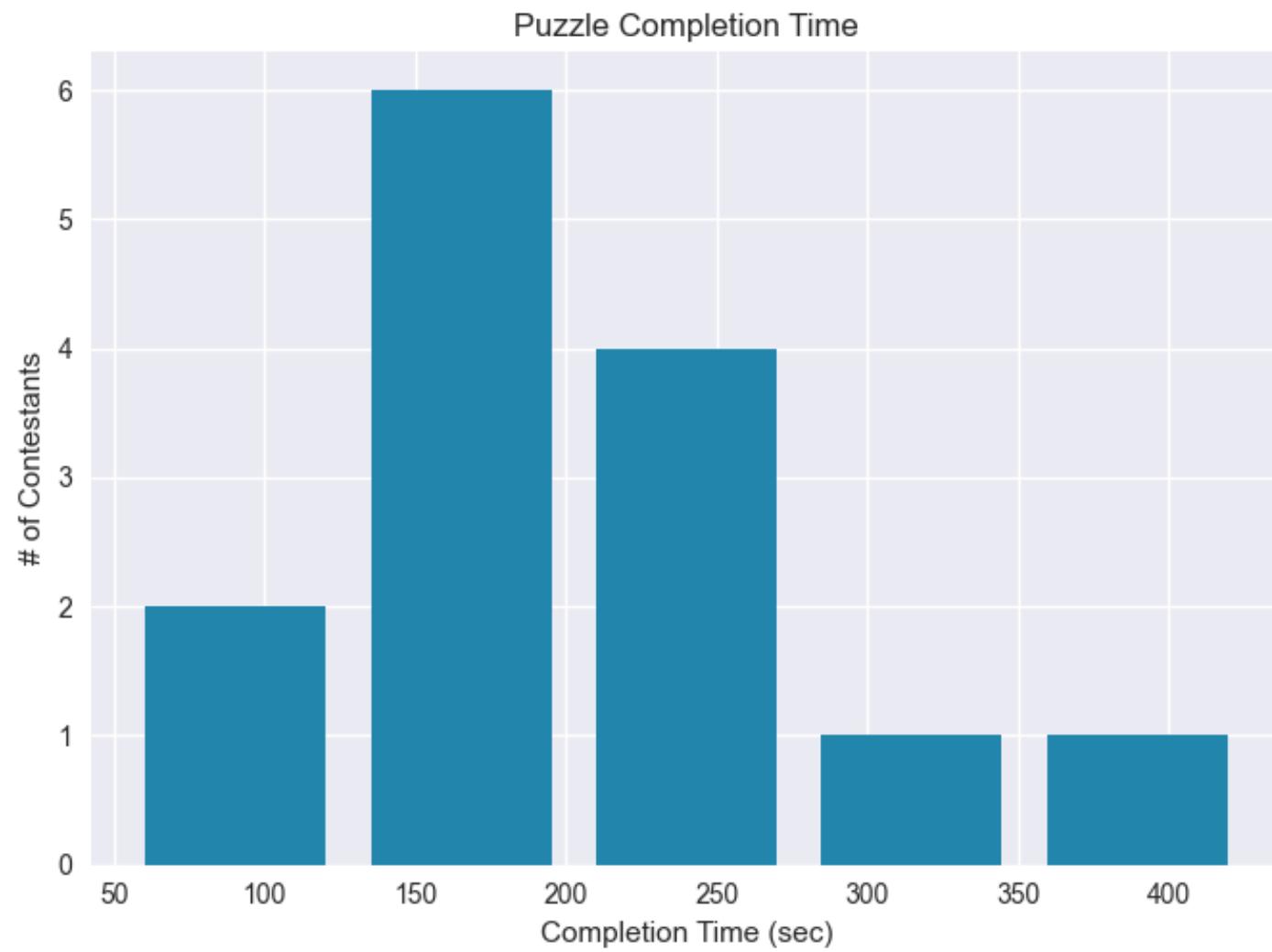
1: Project 1: Structure

```
1 import matplotlib.pyplot as plt
2
3 plt.style.use('seaborn')
4 puzzleCompletionTime = [139, 53, 96, 166, 182, 197, 211, 427, 215, 232, 297, 180, 144, 268]
5 plt.hist(puzzleCompletionTime, bins=5, color="#2185AC", rwidth=0.8)
6 plt.xlabel('Completion Time (sec)')
7 plt.ylabel('# of Contestants')
8 plt.title('Puzzle Completion Time')
9 plt.show()
10
11
```

Run: in\_class\_coding

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl"

Process finished with exit code 0



# **Loading data from a csv file in Pandas**

## 1: Project

ITP449\_Fall2020  
Class  
.idea  
In Class Coding  
7.png  
drivers.cs  
grade5Scie  
in\_class\_coding.py  
ITPEnr.png  
rainfall.p  
salaries.c  
venv library  
External Libraries  
Scratches and Cons

## 2: Structure

```
1 import matplotlib.pyplot as plt
2 import pandas as pd
3
4 plt.style.use('seaborn')
5 puzzleCompletionTime = pd.read_csv('grade5ScienceFairProject.csv')
6 plt.hist(puzzleCompletionTime['Seconds'], bins=5, color='#2185AC', rwidth=0.8)
7 plt.xlabel('Completion Time (sec)')
8 plt.ylabel('# of Contestants')
9 plt.title('Puzzle Completion Time')
10 plt.show()
11
12
```

## 2: Favorites

Run: in\_class\_coding

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code 0

## 3: Bottom Bar

▶ 4: Run

TODO

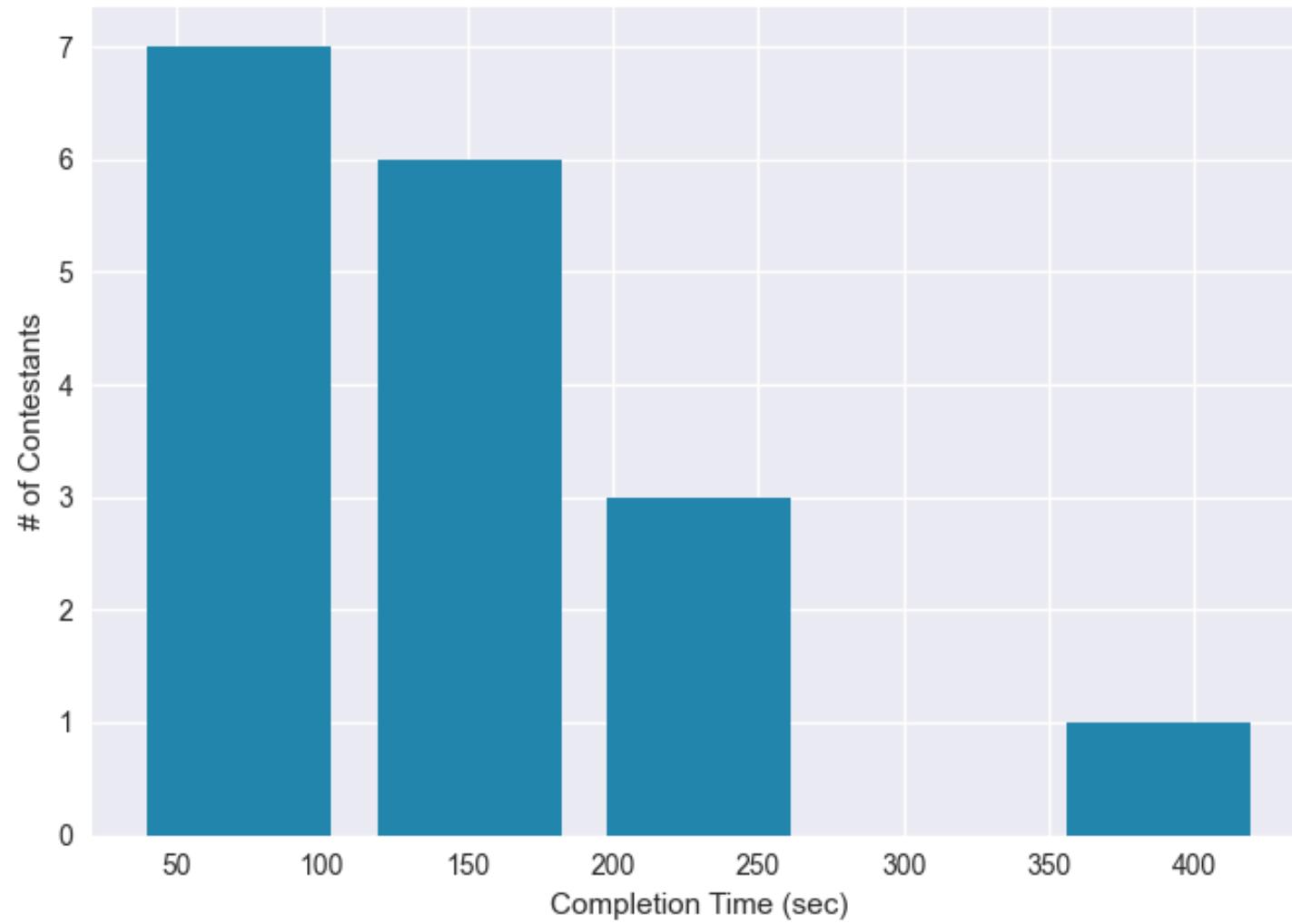
6: Problems

Terminal

Python Console

Event Log

Puzzle Completion Time



# **Statistical Plotting using Seaborn**

## 1: Project

Pro... ▾ -

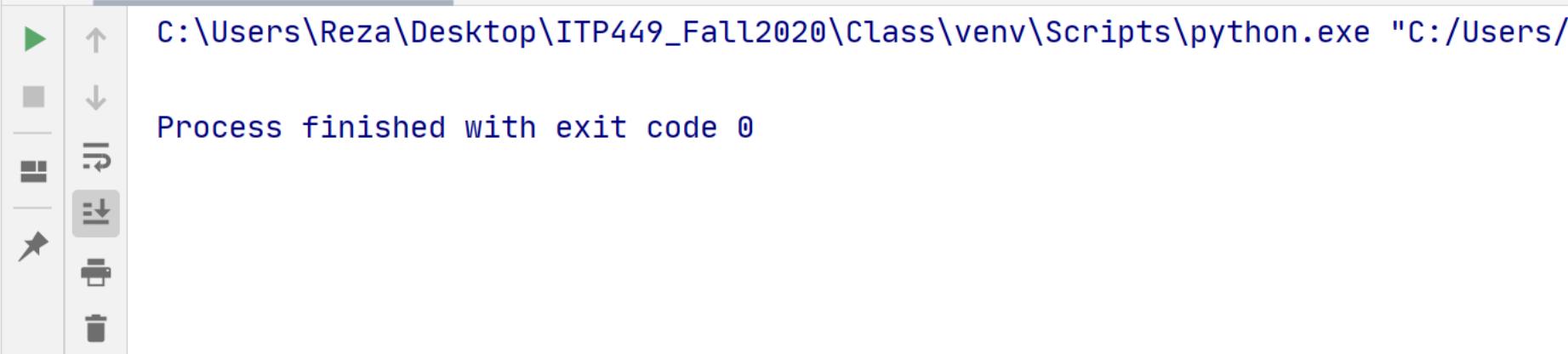
ITP449\_Fall2020 C:\Us  
  Class  
  > .idea  
  In Class Coding  
    drivers.csv  
    grade5ScienceF  
    in\_class\_coding.py  
    rainfall.png  
    salary.csv  
  > venv library roo  
External Libraries  
Scratches and Console:

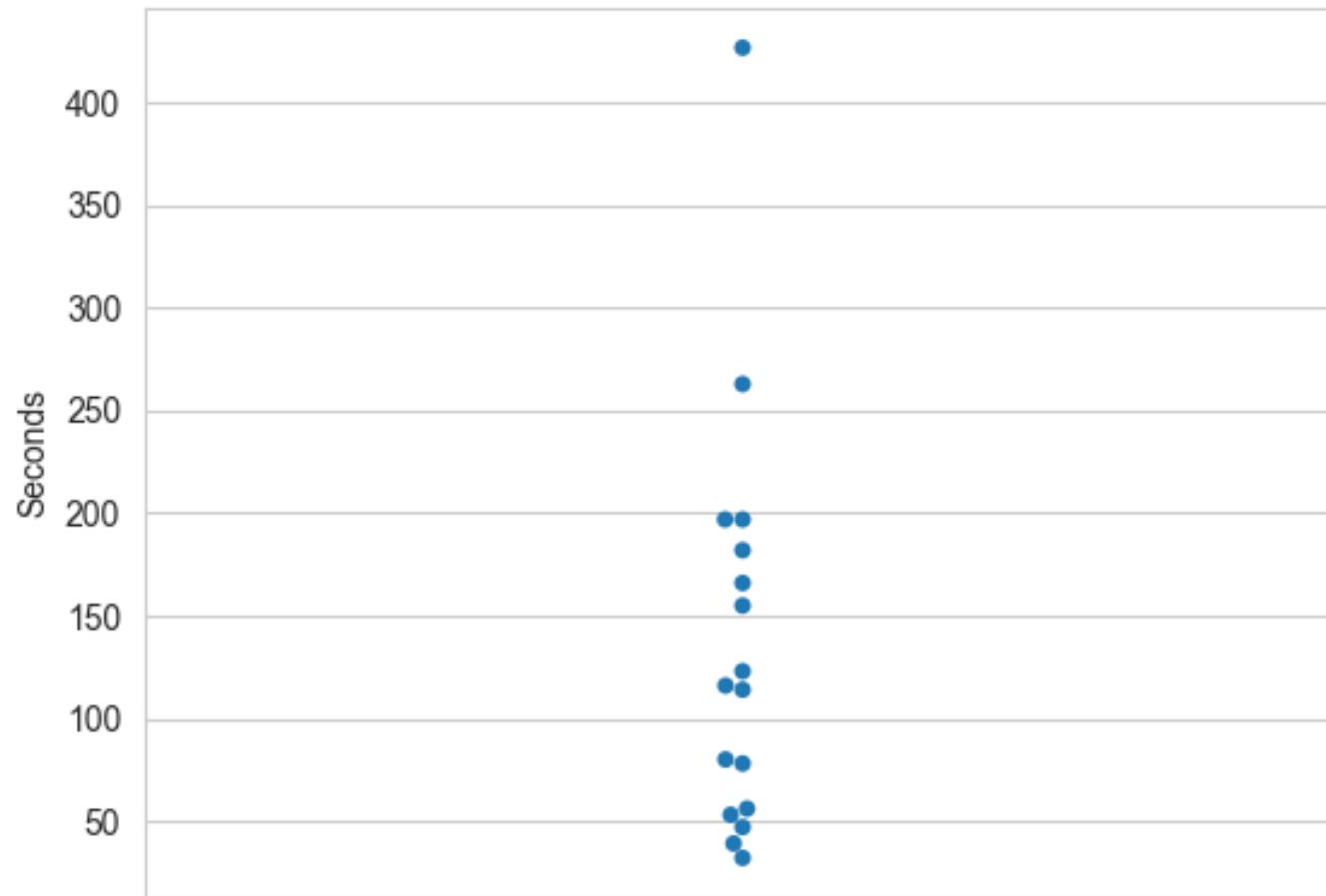
## Run: in\_class\_coding ×

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code 0

## 2: Favorites





ITP449\_Fall2020 > Class > In Class Coding >  in\_class\_coding.py

in class coding ▾



1: Project

7: Structure

★ 2: Favorites

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

sb.set_style('whitegrid')
puzzleCompletionTime = pd.read_csv('grade5ScienceFairProject.csv')
sb.swarmplot(x='Music', y='Seconds', data=puzzleCompletionTime, size=10)
plt.show()
```

Run:  in\_class\_coding ×

—

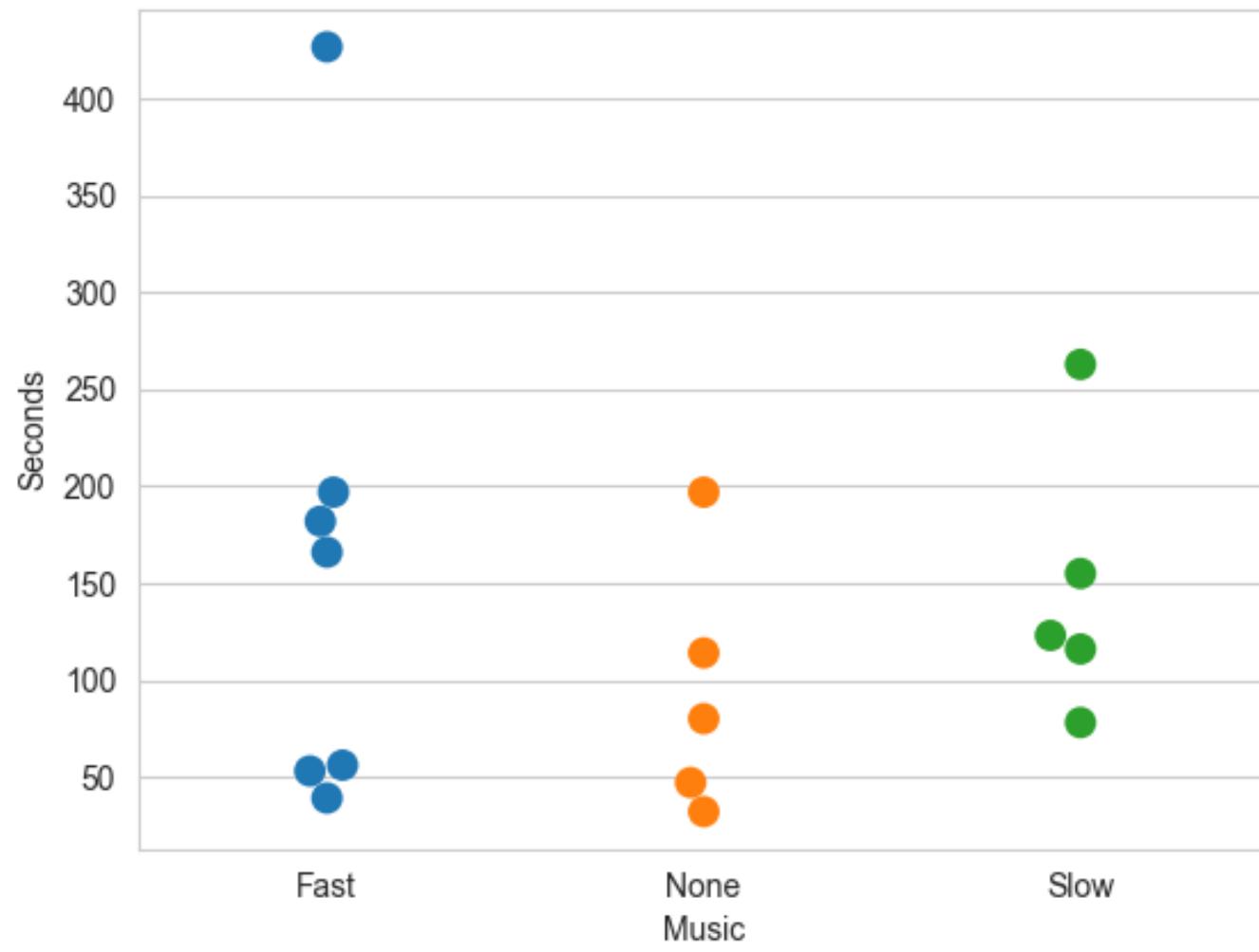
c:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code 0



▶ 4: Run    TODO    ⚡ 6: Problems    Terminal    Python Console    1 Event Log

Packages installed successfully: Installed packages: 'seab... (yesterday 6:59 PM) 4:1 CRLF UTF-8 4 spaces Python 3.8 (Class) 🔒





1: Project

7: Structure

★ 2: Favorites

The screenshot shows the PyCharm file explorer with the following structure:

- ITP449\_Fall2020 (Project)
  - Class
    - .idea
  - In Class Coding
    - drivers.csv
    - grade5Science.pdf
    - in\_class\_coding.ipynb
    - rainfall.png
    - salary.csv
  - venv
- External Libraries
- Scratches and Console

Run:  in\_class\_coding ×

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code

## Superimpose boxplots on the previous swarmplots

► 4: Run

## ☰ TODO

! 6: Pro

Problems >

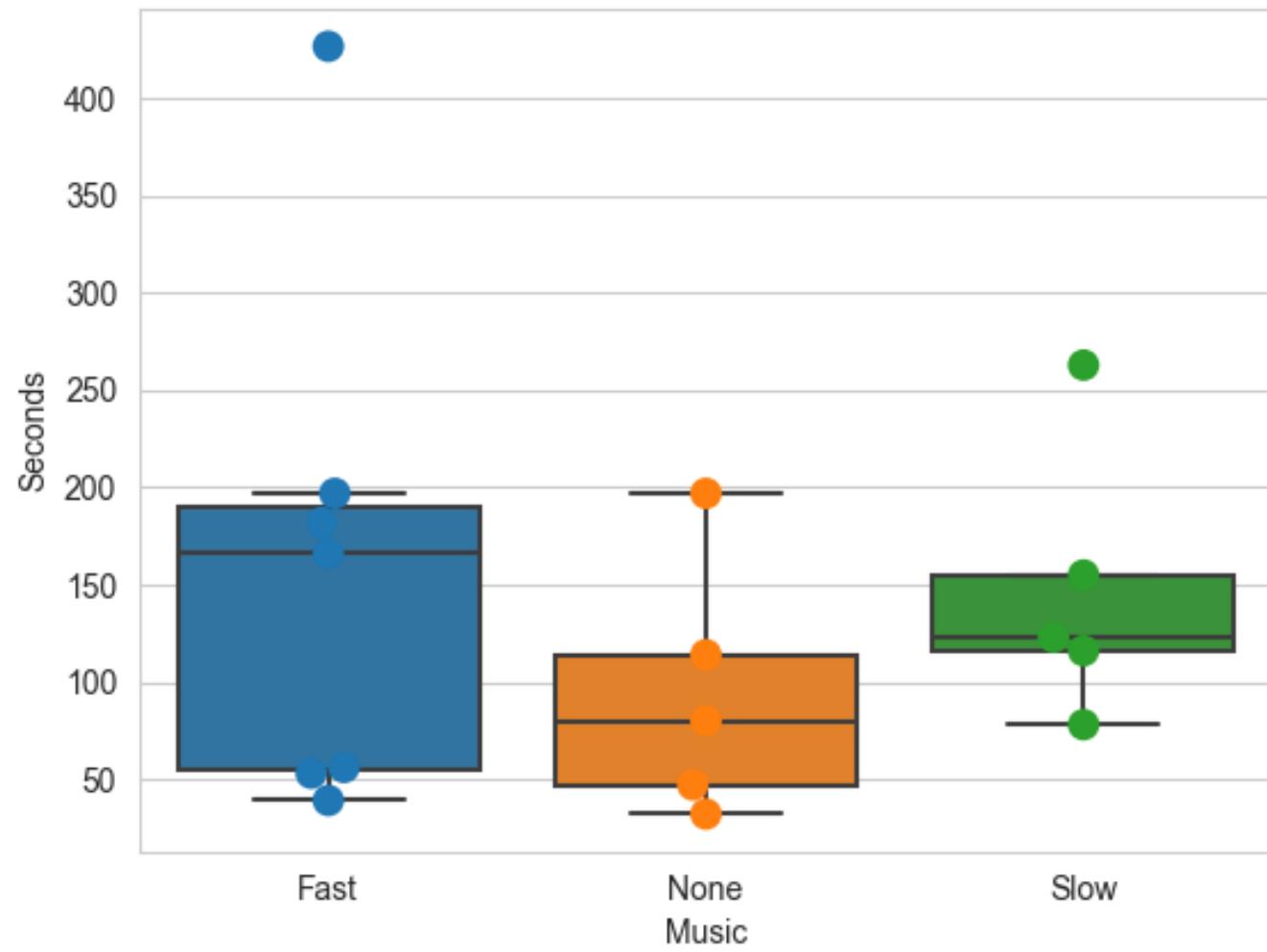
## Terminal

 Python Console

1 Event Log

Packages installed successfully: Installed packages: 'sea... (yesterday 6:59 PM)

10:1 CRLF UTF-8 4 spaces Python 3.8 (Class) 🔒



## 1: Project

Project -

- ITP449\_Fall2020 C:\Use
  - Class
  - .idea
  - In Class Coding
    - 7.png
    - drivers.csv
    - grade5ScienceFa
    - in\_class\_coding
    - ITPEnr.png
    - rainfall.png
    - salaries.csv
  - venv library root
- External Libraries
- Scratches and Consoles

## 2: Structure

Run: in\_class\_coding ×

C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/Cl

Process finished with exit code 0

## 2: Favorites



Run

TODO

Problems

Terminal

Python Console

Event Log



1: Project 1: Structure 1: Favorites

Project: ITP449\_Fall2020 C:\Users\Reza\Desktop\ITP449\_Fall2020\Class

1 in\_class\_coding.py

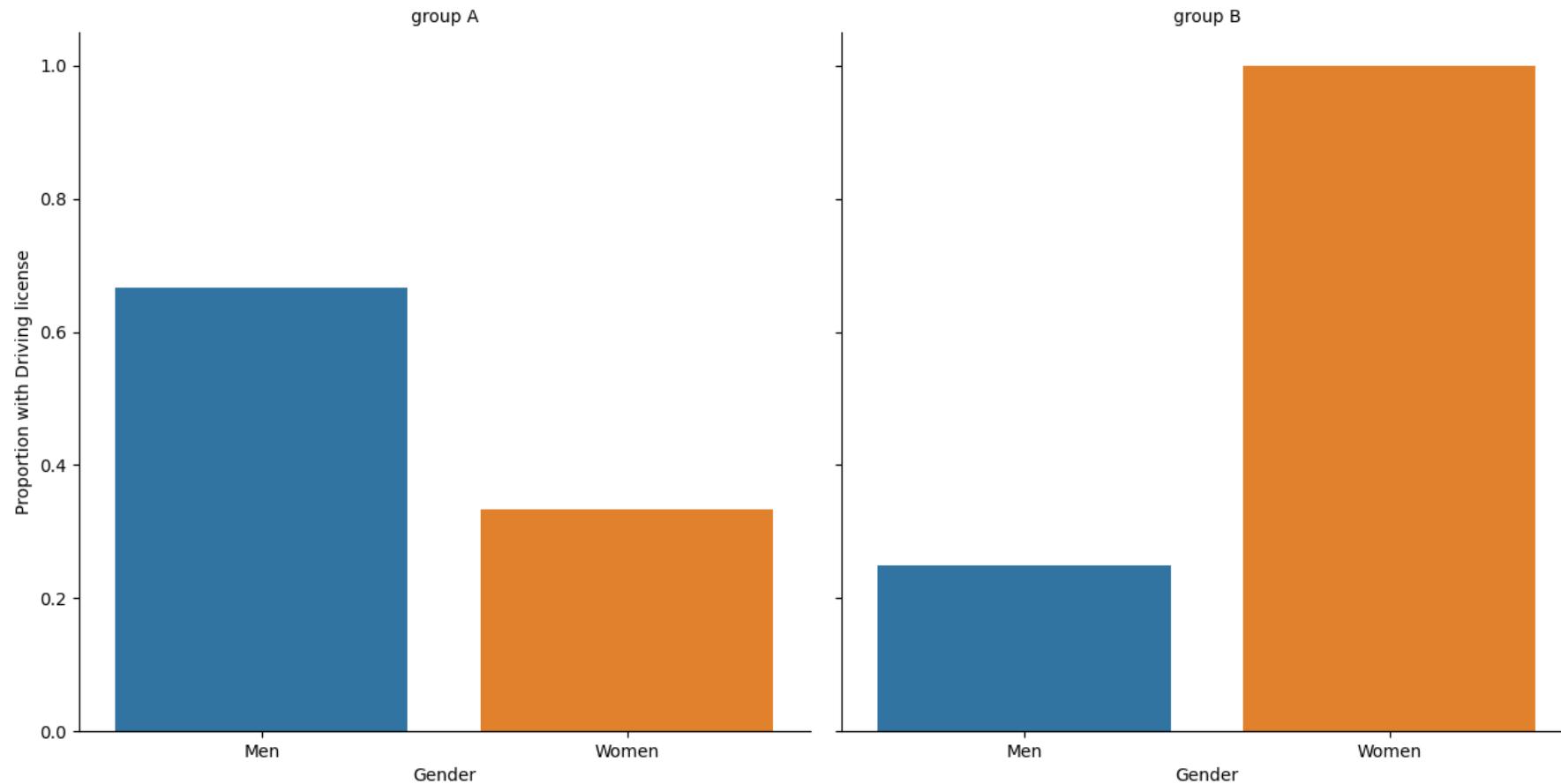
```
1 import matplotlib.pyplot as plt
2 import pandas as pd
3 import seaborn as sb
4
5 driverData = pd.read_csv('drivers.csv')
6 g = sb.catplot(x='gender', y='license', col='group',
7                 data=driverData, kind='bar', ci=None)
8 g.set_axis_labels('Gender', 'Proportion with Driving license')
9 g.set_xticklabels(['Men', 'Women'])
10 g.set_titles('{col_var} {col_name}')
11 plt.show()
12 |
```

Run: in\_class\_coding



C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/CL

Process finished with exit code 0



1: Project

ITP449\_Fall2020

- Class
- .idea
- In Class Cod:
  - 7.png
  - drivers.cs
  - grade5Scie
  - in\_class\_c
  - ITPEnr.png
  - rainfall.p
  - salaries.c
- venv library

External Libraries

Scratches and Cons

in\_class\_coding.py

```
1 import matplotlib.pyplot as plt
2 import seaborn as sb
3
4 titanic = sb.load_dataset('titanic')
5 g = sb.catplot(x='who', y='survived', col='class',
6                 data=titanic, kind='bar', ci=None, aspect=0.5)
7 g.set_axis_labels('', 'Survival Rate')
8 g.set_xlabels(['Men', 'Women', 'Children'])
9 g.set_titles('{col_name} {col_var}')
10 plt.show()
```

Sample dataset in Seaborn

Run: in\_class\_coding



C:\Users\Reza\Desktop\ITP449\_Fall2020\Class\venv\Scripts\python.exe "C:/Users/Reza/Desktop/ITP449\_Fall2020/CL

Process finished with exit code 0

2: Favorites

3:

Run

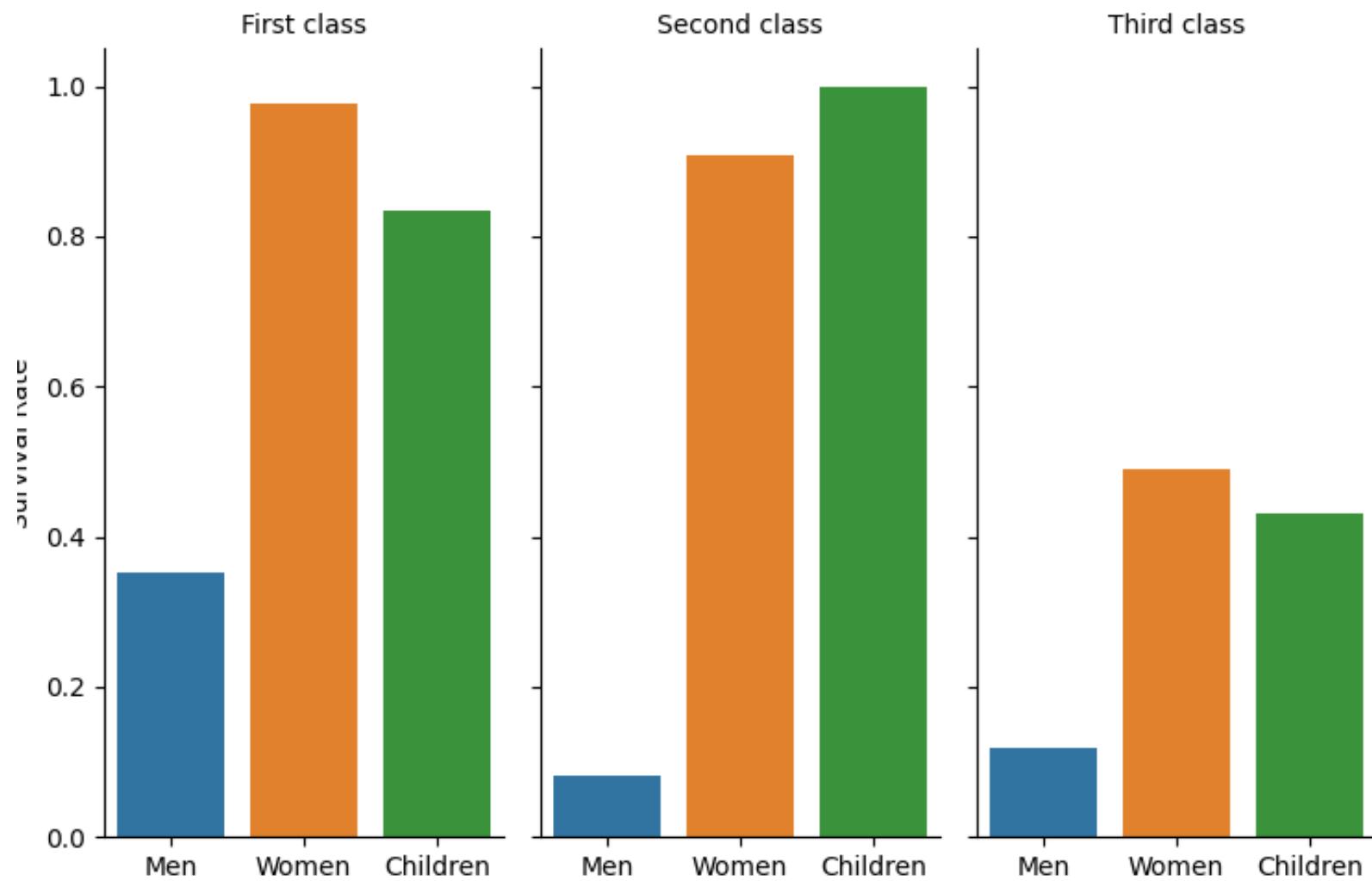
TODO

Problems

Terminal

Python Console

Event Log



# Summary

In this lecture we covered data visualization using `pyplot` and *Seaborn*

`pyplot` gives you extensive control on how data are visualized

*Seaborn* is useful in making statistical graphics in Python

We will use *matplotlib* (and other packages) for visualizing the results of machine learning models in this course