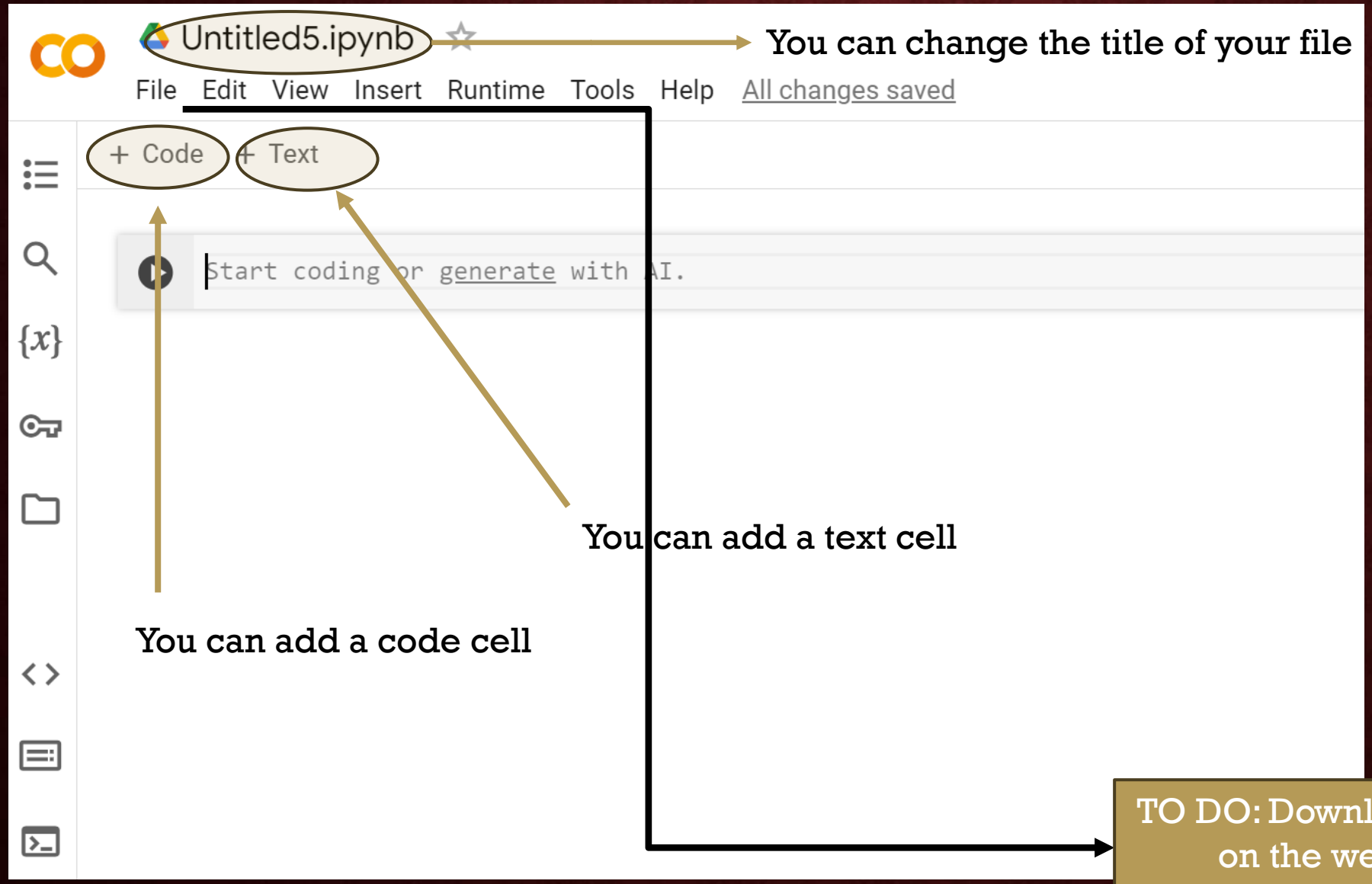
INTRO TO PYTHON 1

Be sure to get an account at

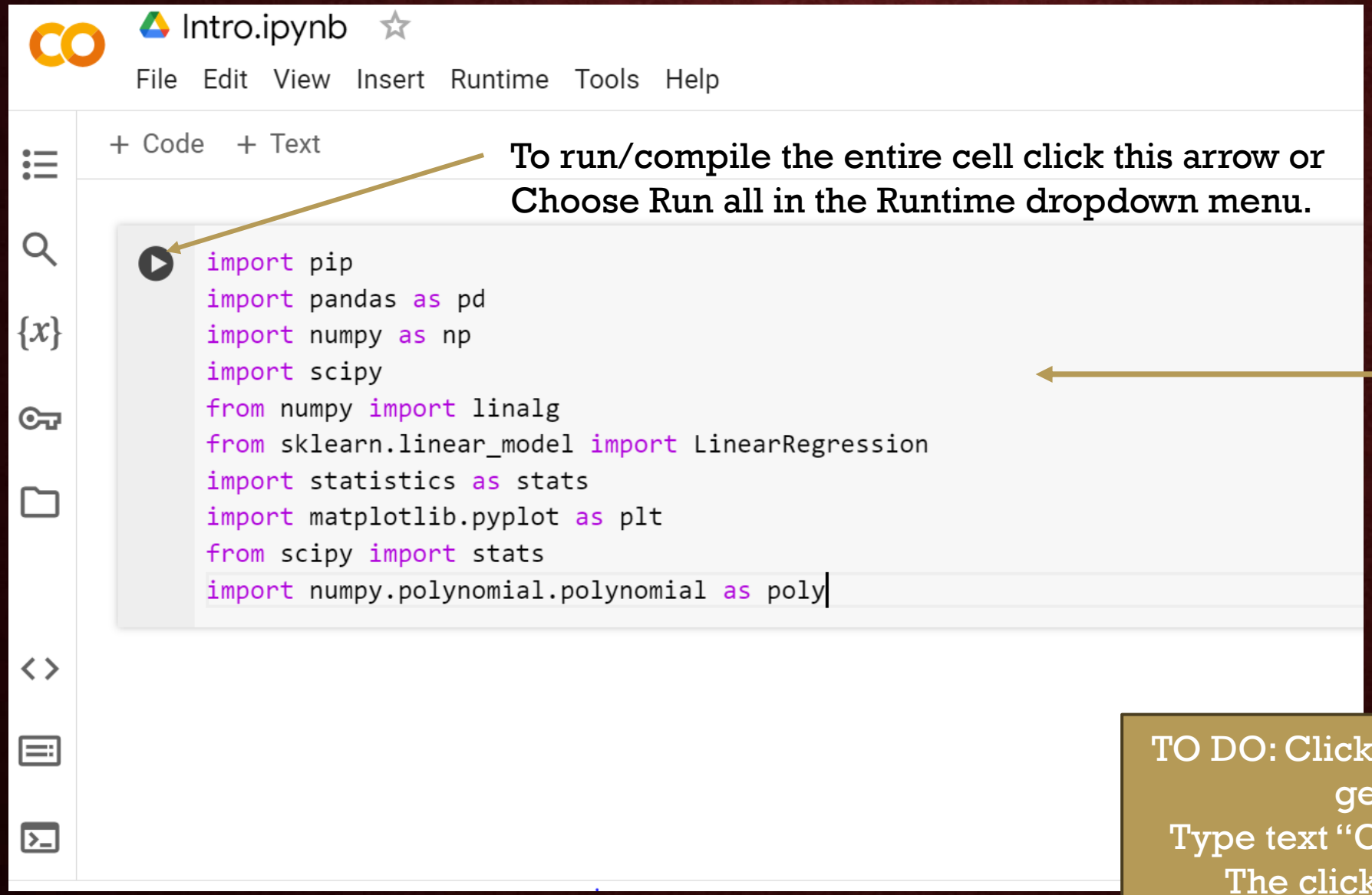
https://research.google.com/colaboratory/

DataNexus
ANALYZE • INTERPRET



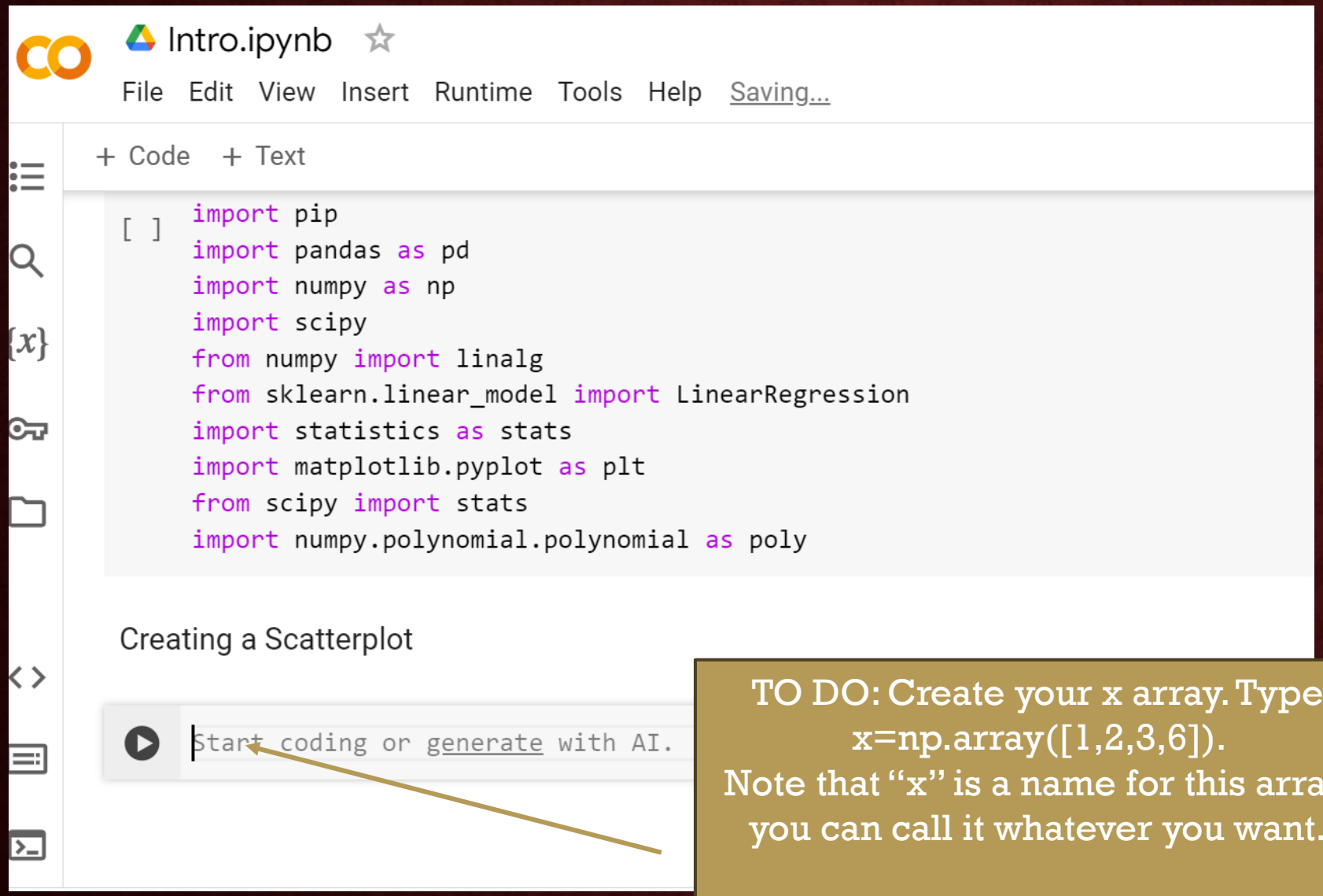
All files will be saved in your Google Drive.

TO DO: Download the file intro.ipynb on the website. Use the File dropdown to upload the file



This is a code cell. It contains a variety of packages that are useful in modeling of data and data visualization.

TO DO: Click on the + Text button to get a Text box.
Type text "Creating a Scatterplot"
The click the +Code button



There are lots of ways to input data, we will be using arrays.

Our practice data set will be {(1,0),(2,5),(3,4),(6,9)}

TO DO: Create your x array. Type x=np.array([1,2,3,6]).
Note that "x" is a name for this array, you can call it whatever you want.

Hit enter and then create an array for your y values.

Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

```
[ ]  import pip
     import pandas as pd
     import numpy as np
     import scipy
     from numpy import linalg
     from sklearn.linear_model import LinearRegression
     import statistics as stats
     import matplotlib.pyplot as plt
     from scipy import stats
     import numpy.polynomial.polynomial as poly
```

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
```

These commands store the data in the variables x and y.

Now let's try a scatterplot.

Type:plt.scatter(x,y)

in the same cell and click the arrow.

Intro.ipynb

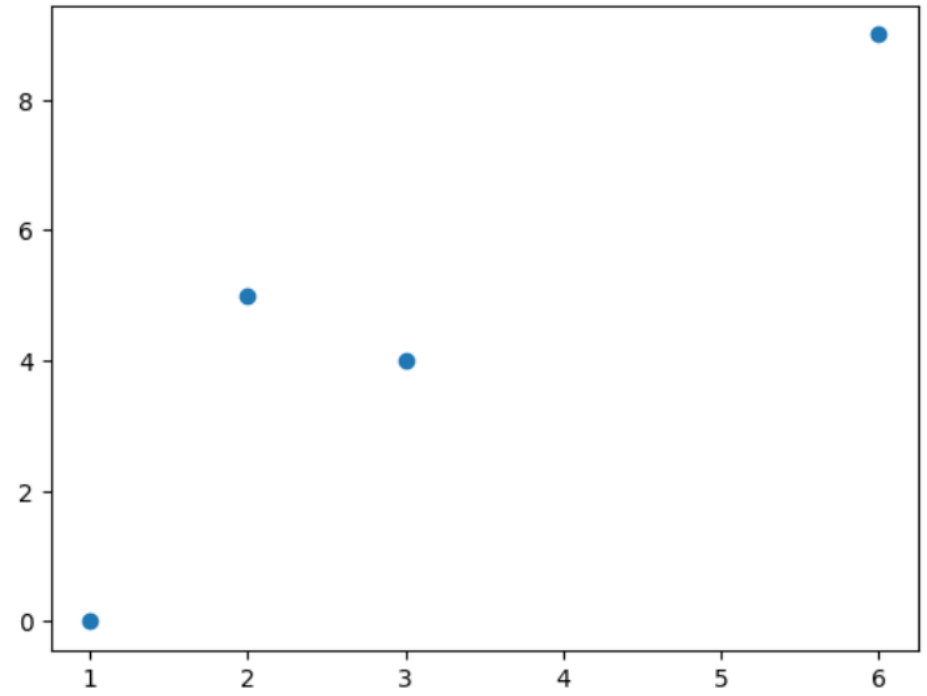
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y)
```

```
<matplotlib.collections.PathCollection at 0x7bb52624f3a0>
```



Didn't get a graph? If so, too much time may have passed.

Colab is reliant on your internet connection and if it timed out, or if you left and came back, run all of your cells by using the Runtime dropdown – Run all

Intro.ipynb ☆

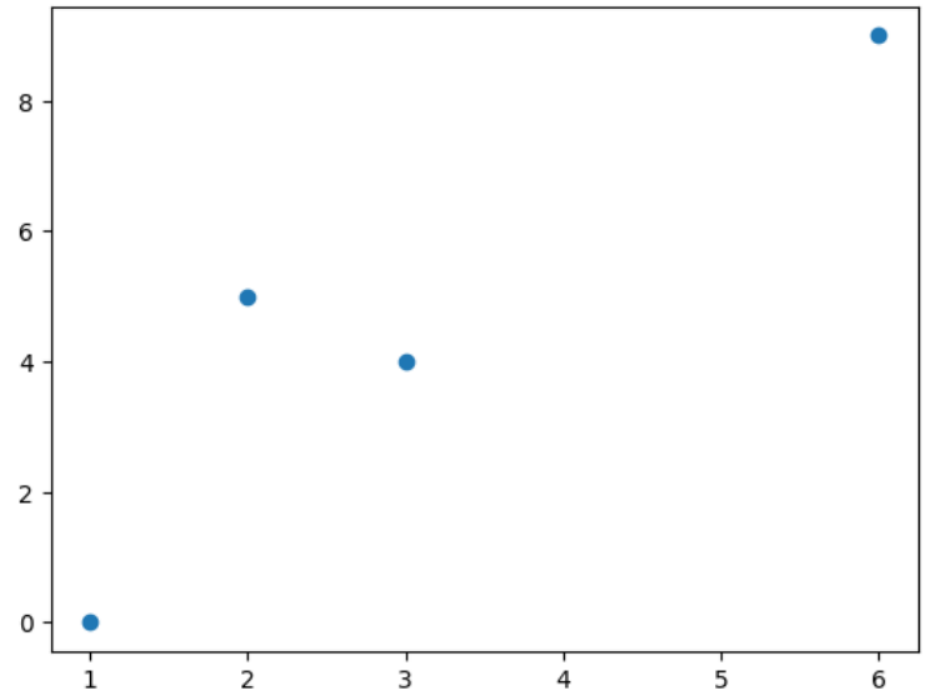
File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y)
```

```
<matplotlib.collections.PathCollection at 0x7bb52624f3a0>
```



Let's add some axis labels and graph title.

In the same cell, type:

plt.title("Name your Graph")
plt.xlabel("Name your xaxis")
plt.ylabel("Name your yaxis")
plt.show()

Then Runtime→Runall
Or click

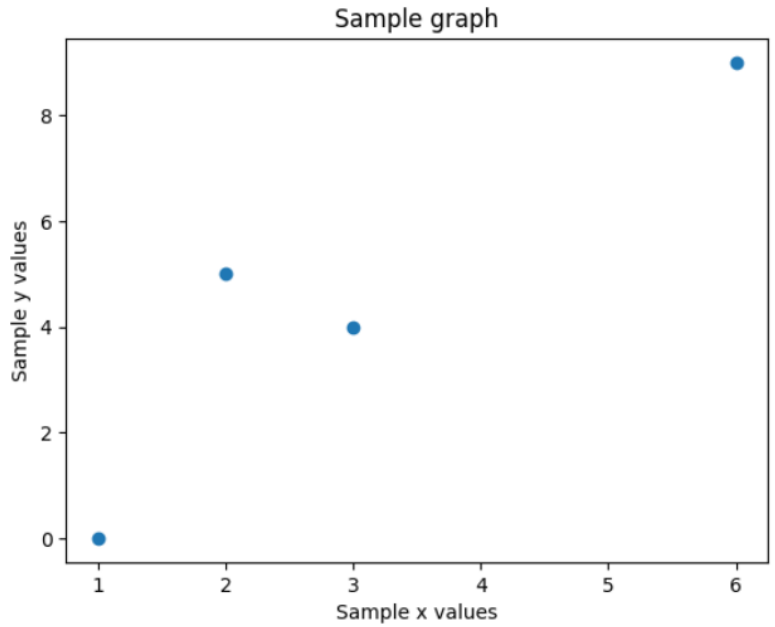
Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y)
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
plt.show()
```



Features of the scatterplot. Want to change the color or transparency of the dots

Add it here. Try it: plt.scatter(x,y,marker="s", alpha=.5,color="red")

Then Runtime→Runall
Or click

Find more markers here: https://matplotlib.org/stable/api/markers_api.html

TO DO: Make the dots more transparent by changing alpha. Change the color to green and change the marker to a Diamond.

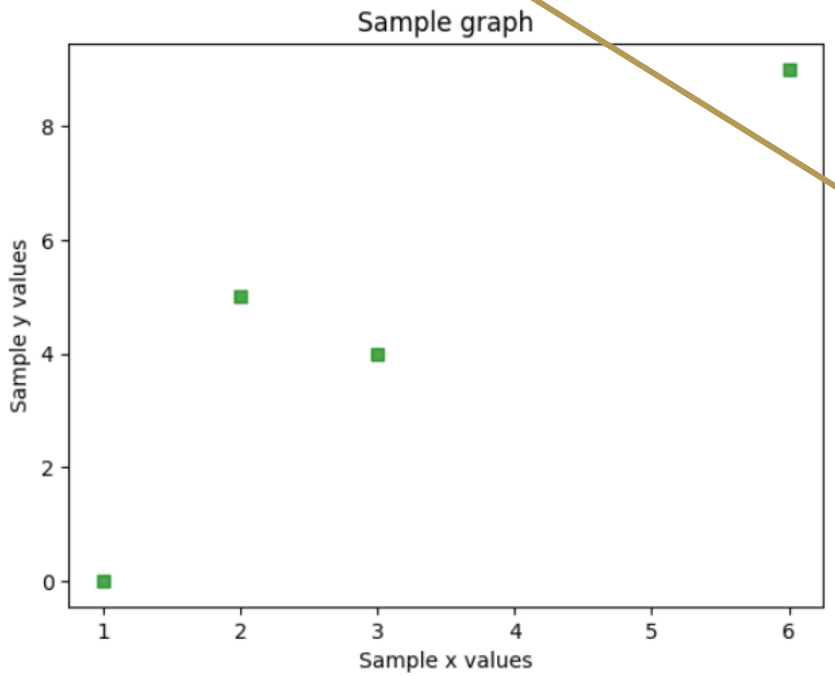
Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y,marker="s",alpha=.7,color="green")
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
plt.show()
```



Want to highlight just one data point with different features?

Python arrays start counting with the number 0.

For example x[0]=1 and y[1]=5.

So plt.scatter(x[1],y[1]) would plot the point (2,5).

TO DO: Add a plt.scatter line where the arrow is pointing to plot the point (2,5) in red as a diamond. Don't forget to run your code!

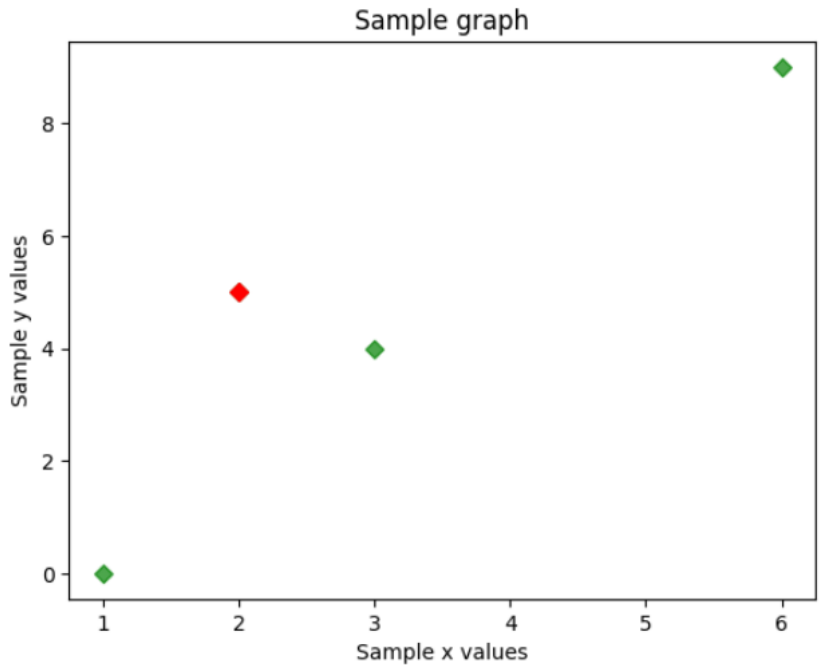
Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y,marker="D",alpha=.7,color="green")
plt.scatter(x[1],y[1],marker="D",color="red")
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
plt.show()
```



Want to label your data?

Create annotations

annotations=[“Sam”,”Qian”,”Rath”,”Kalu”]

Then you have to use a for loop to add the labels. We will talk more about for loops later.

For i, label in enumerate(annotations):

→ plt.text(x[i],y[i],label)

Indent just one tab

TO DO: Add annotations to your graph

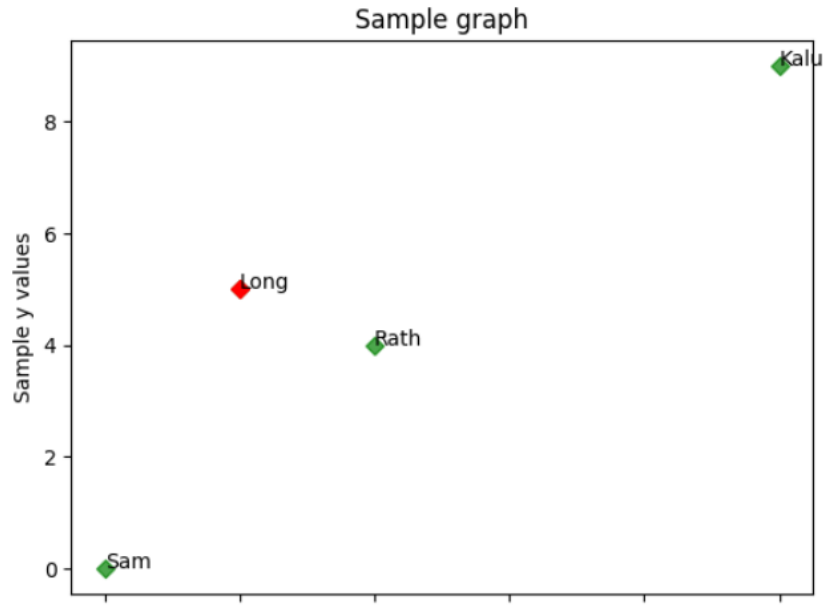
Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y,marker="D",alpha=.7,color="green")
plt.scatter(x[1],y[1],marker="D",color="red")
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
annotations=["Sam","Long","Rath","Kalu"]
for i, label in enumerate(annotations):
  plt.text(x[i],y[i],label)
plt.show()
```



You might want to move your label so that they are not on top of your points.

You can choose how much to move the labels in both the x and y directions, relative to the data.

For example, plt.text(x[i]+.1,y[i]-.02, label)
moves labels right .1 from their dot
moves labels down .02 from their dot

TO DO: Move your labels left .1 and up .2

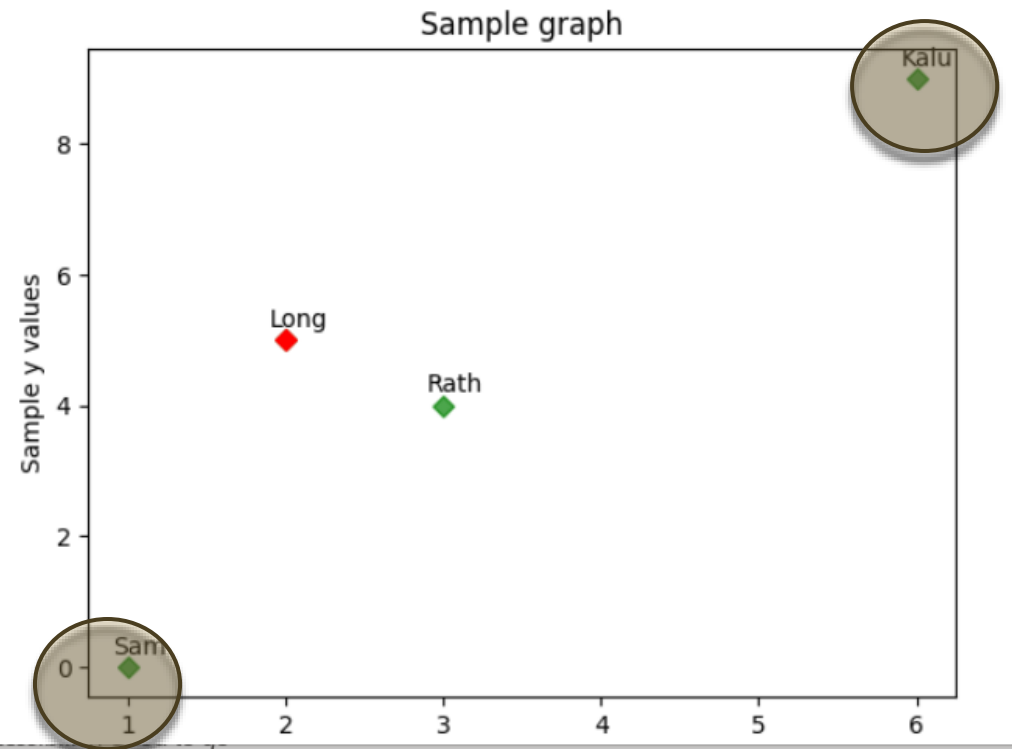
Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y,marker="D",alpha=.7,color="green")
plt.scatter(x[1],y[1],marker="D",color="red")
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
annotations=["Sam","Long","Rath","Kalu"]
for i, label in enumerate(annotations):
  plt.text(x[i]-.1,y[i]+.2,label)
plt.show()
```



You might want to change the x range and y range to include (0,0) or make sure you can see your label for Kalu.

plt.xlim([minimum x, maximum x])
plt.ylim([minimum y, maximum y])

Make sure to put this before plt.show()

TO DO: Change your x range to [0,7] and y range to [0,10]

WATCH THE 2ND PYTHON VIDEO TO LEARN HOW TO IMPORT DATA, AND FIND AND GRAPH A SIMPLE LINEAR REGRESSION MODEL

Intro.ipynb

File Edit View Insert Runtime Tools Help All changes saved

+ Code + Text

Creating a Scatterplot

```
x=np.array([1,2,3,6])
y=np.array([0,5,4,9])
plt.scatter(x,y,marker="D",alpha=.7,color="green")
plt.scatter(x[1],y[1],marker="D",color="red")
plt.title("Sample graph")
plt.xlabel("Sample x values")
plt.ylabel("Sample y values")
annotations=["Sam","Long","Rath","Kalu"]
for i, label in enumerate(annotations):
  plt.text(x[i]-.1,y[i]+.2,label)
plt.xlim([0,7])
plt.ylim([0,10])
plt.show()
```