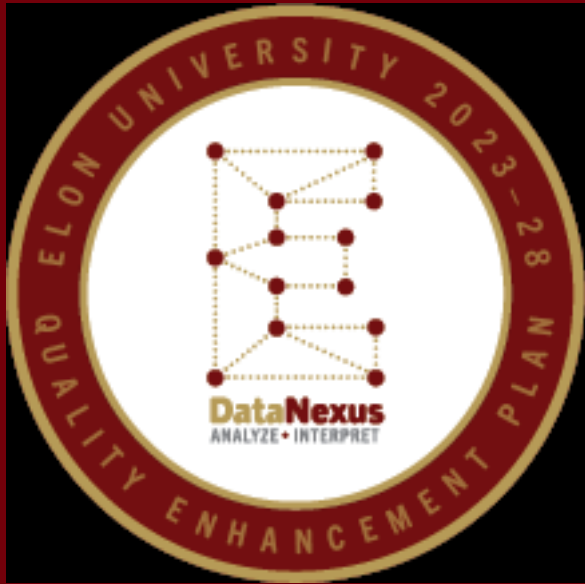


INTRO TO PYTHON 2

Be sure to get an account at

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





DOWNLOAD THE DATA FOR THIS SESSION

Before we begin download
[gunlaws.csv](#) from our website





Intro.ipynb ☆

File Edit View Insert Runtime Tools Help

+ 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
```

To run/compile the entire cell click this arrow or Choose Run all in the Runtime dropdown menu.

Click on the Folder,
Upload, Find the
Gunlaw.csv
File

TO DO: Add a Text cell and Type
“Uploading a Dataset for Creating a
Scatterplot”



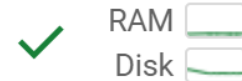
Files



sample_data

Click on this button to Upload a File
Find the gunlaws.csv file in your downloads and choose to upload it.

+ Code + Text



Colab AI



3s

```
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
```

```
/usr/local/lib/python3.10/dist-packages/_distutils_hack/
warnings.warn("Setuptools is replacing distutils.")
```

Uploading a File for Creating a Scatterplot



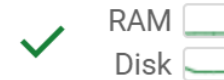
Files



- ..
- sample_data
- gunlaws.csv



+ Code + Text



Colab AI


```
[2] from numpy import linalg
from sklearn.linear_model import LogisticRegression
import statistics as st
import matplotlib.pyplot as plt
from scipy import stats
import numpy.polynomial
```

We will next Create a dataframe
Create a Code Cell
Type `df=pd.read_csv(" ")`

Uploading a File for Creating a Scatterplot

```
df=pd.read_csv("/content/gunlaws.csv")
```

Copy your file path here
Don't forget to run the cell



Click on the 3 dots next to the gunlaws.csv file to Copy the Path

The screenshot shows a Google Colab notebook titled "Intro_to_Python Part 2". The code cell contains the following Python code:

```
df=pd.read_csv("/content/gunlaws.csv")
print(df)
```

The output is a pandas DataFrame with the following columns: State, 2019 Grade, Gun Death Rate, and Gun Death Rate Per 100K. The data is as follows:

	State	2019 Grade	Gun Death Rate	Gun Death Rate Per 100K
0	Alabama	F	2	21.70
1	Alaska	F	7	20.74
2	Arizona	F	18	15.29
3	Arkansas	F	8	18.96
4	California	A	44	7.45
5	Colorado	C+	19	15.14
			45	4.91
			34	11.55
			27	12.81
			17	15.72
			48	4.03
			16	16.61
			36	10.78
			22	11.71

We will be working with arrays of data in this video.

An array is a mathematical table with rows and columns.

If **data** is the name of our array then we refer to the entry in the 0th row and 3rd column as `data[0,3]`



+ Code + Text

```
df=pd.read_csv("/content/gunlaws.csv")  
print(df)
```

	State	2019 Grade	Gun Death Rate	Gun Death Rate Per
0	Alabama	F	2	
1	Alaska	F	7	
2	Arizona	F	18	
3	Arkansas	F	8	
4	California	A	44	
5	Colorado	C+	19	
6	Connecticut	A-	45	
7	Delaware	B	34	11.55
8	Florida	C-	27	12.81
9	Georgia	F	17	15.72
10	Hawaii	A-	48	4.03
11	Idaho	F	16	16.61
12	Illinois	A-	36	10.78
13	Indiana	D	28	11.71

We are going to choose the State Gun Law Strength (which are in column 6 of the data) as the xvalues.

Type x=data[:,6]



To Do: Define y values at Gun Deaths Per 100k, which is the 3rd columns of data, and the annotations as the States, which are in the 0th column of data

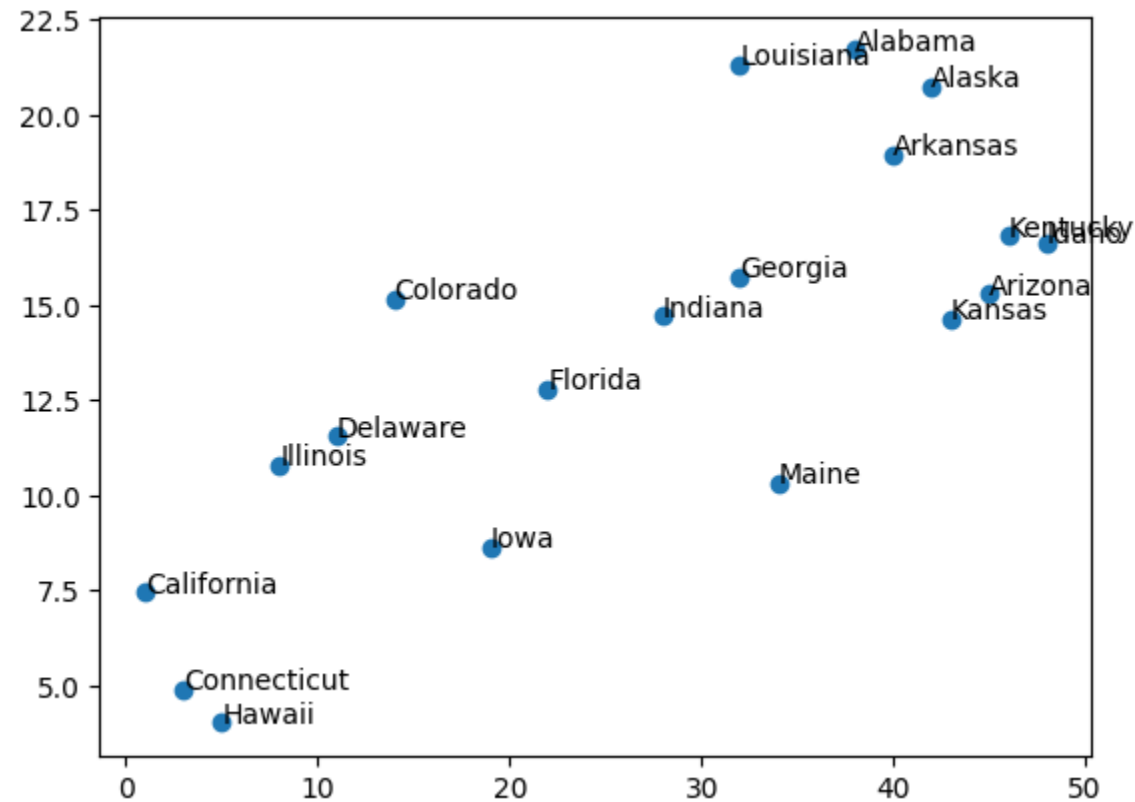


To Do: Use what you learned in Part 1 of the series to create a scatterplot using annotations as the labels.



Uploading a File for Creating a Scatterplot

```
[58] df=pd.read_csv("/content/gunlaws.csv")
data=np.array(df)
nrows=data.shape[0]
ncols=data.shape[1]
x=data[:,6]
y=data[:,3]
annotations=data[:,0]
plt.scatter(x,y)
for i, label in enumerate(annotations):
    plt.text(x[i],y[i],label)
plt.show()
```





```
df=pd.read_csv("/content/gunlaws.csv")
data=np.array(df)
nrows=data.shape[0]
ncols=data.shape[1]
x=data[:,6]
y=data[:,3]
annotations=data[:,0]
plt.scatter(x,y)
for i, label in enumerate(annotations):
    plt.text(x[i],y[i],label)
plt.show()
x=x.astype(float)
y=y.astype(float)
```

Some of your data may have come in as strings, use these commands to make sure that they are float variables.

```
+ Code + Text
0s
nrows=data.shape[0]
ncols=data.shape[1]
x=data[:,6]
y=data[:,3]
annotations=data[:,0]
plt.scatter(x,y)
for i, label in enumerate(annotations):
    plt.text(x[i],y[i],label)
plt.show()
x=x.astype(float)
y=y.astype(float)
slope, intercept, r_value, p_value, std_err=stats.linregress(x, y)
```

This command gives the

- slope
- y intercept
- r value related to linear correlation
- pvalue related to a 2-sided hypothesis test for the slope=0
- and standard error of the estimate



TO DO: To see r^2

Type `print(r_value**2)`

```
+ Code + Text
Uploading a File for Creating a Scatterplot

df=pd.read_csv("/content/gunlaws.csv")
data=np.array(df)
nrows=data.shape[0]
ncols=data.shape[1]
x=data[:,6]
y=data[:,3]
annotations=data[:,0]
plt.scatter(x,y, label='original data')
for i, label in enumerate(annotations):
    plt.text(x[i],y[i],label)

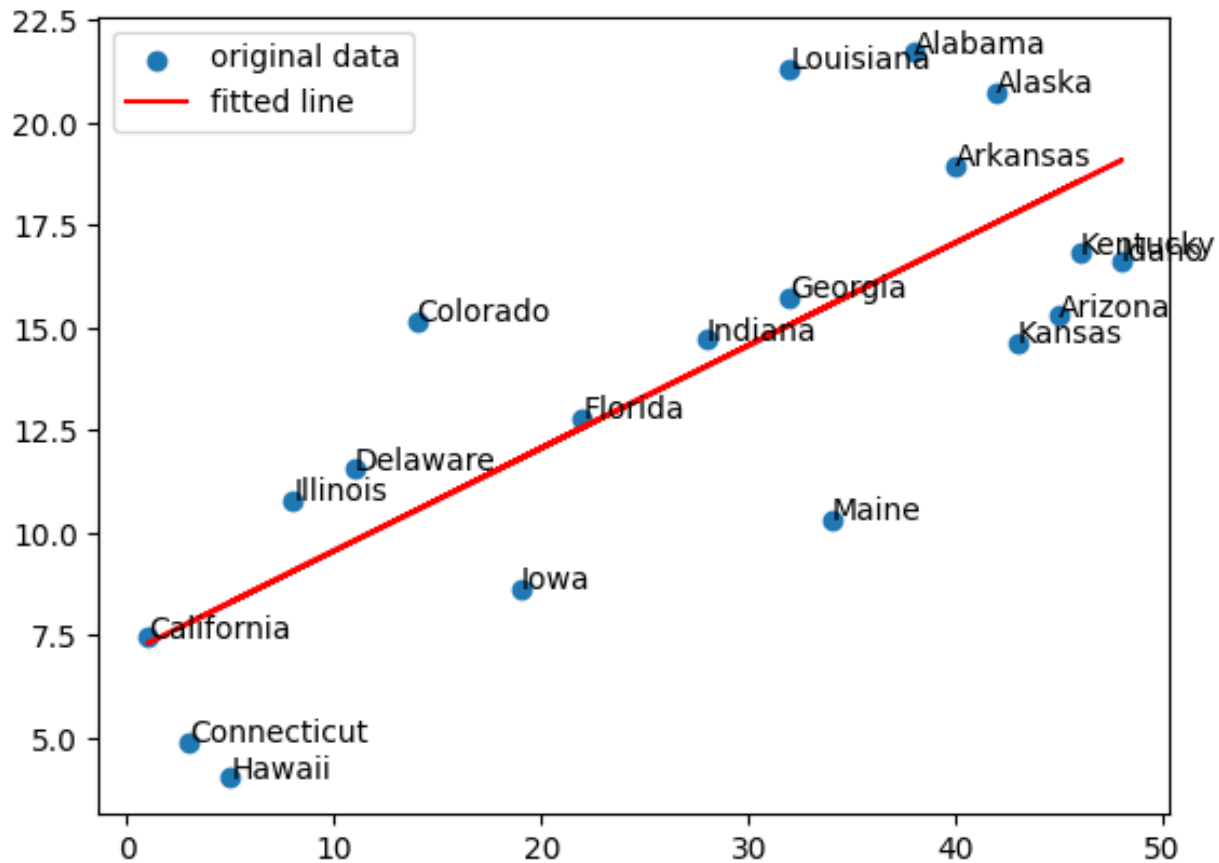
x=x.astype(float)
y=y.astype(float)
slope, intercept, r_value, p_value, std_err=stats.linregress(x, y)
print("rsquared",r_value**2)
plt.plot(x, intercept + slope*x, 'r', label="fitted line")
plt.legend()
plt.show()
```

Add a label to your original data graph

Move plt.show() to combine graphs

Including the simple regression line on your plot

Add a legend



**WATCH THE 3RD
PYTHON VIDEO
TO LEARN HOW
TO CRATE OTHER
DATA
VISUALIZATIONS**