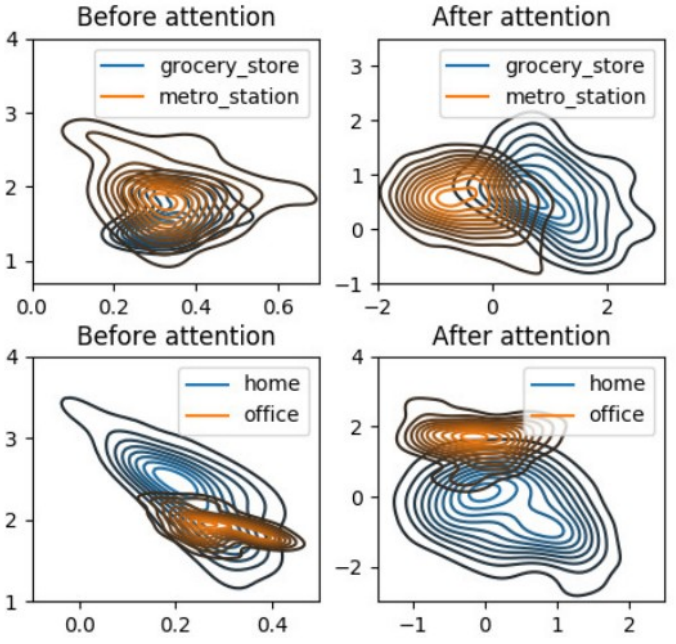
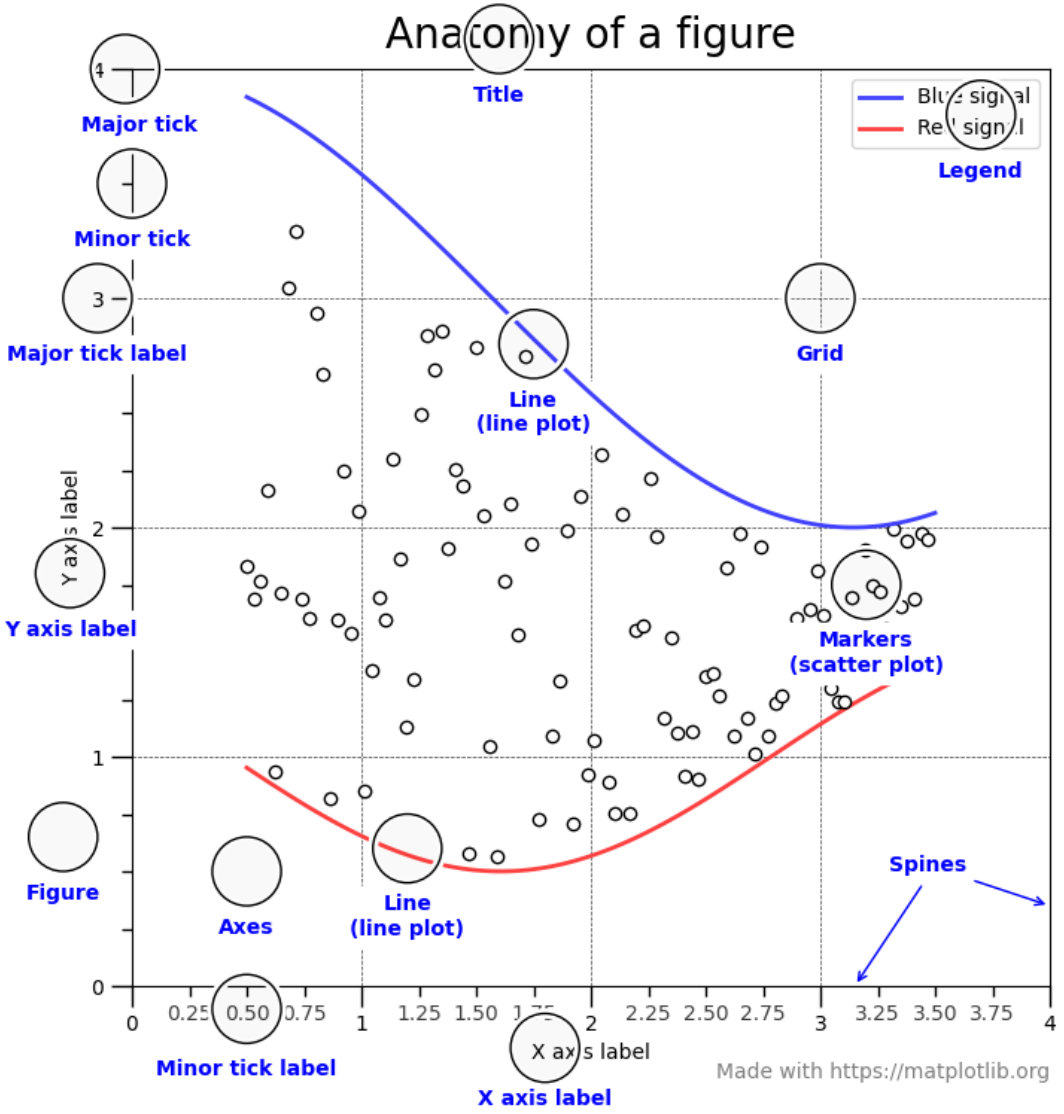
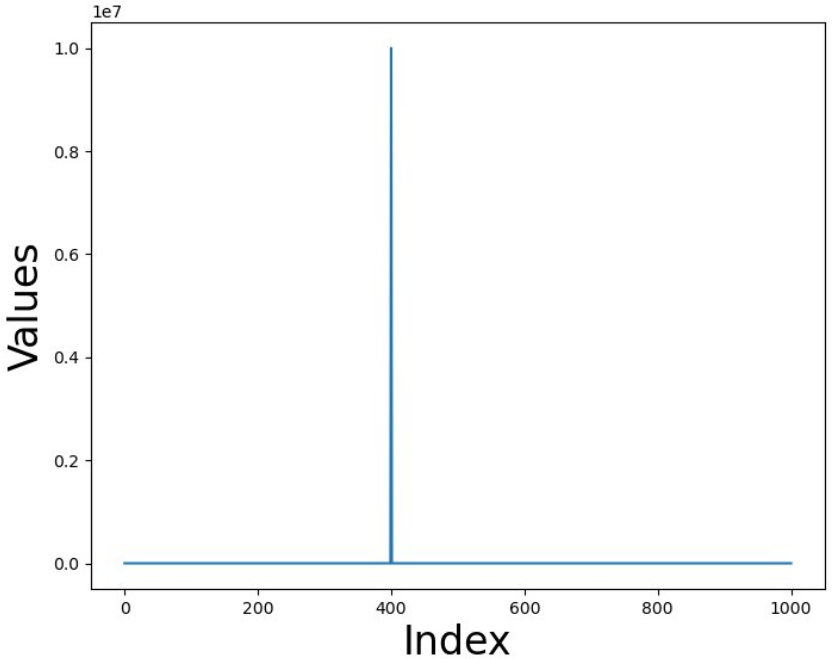
Visualization in Python



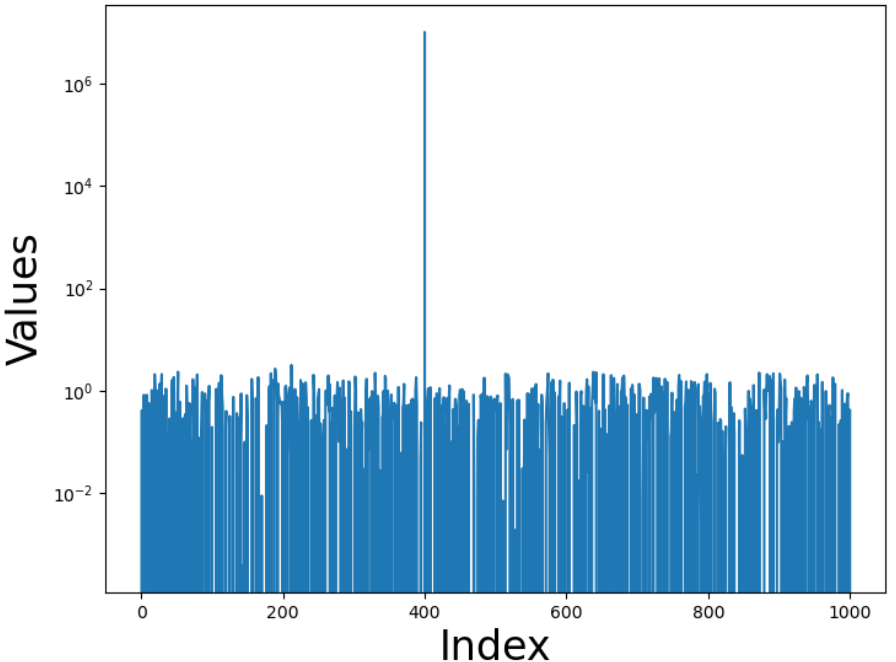

Anatomy of a figure
Title
Major tick
Minor tick
Major tick label
Legend
Grid
Line
(line plot)
Y axis label
Markers
(scatter plot)
Figure
Axes
Line
(line plot)
Spines
Minor tick label
X axis label
Made with https://matplotlib.org

Plotting dos and donts

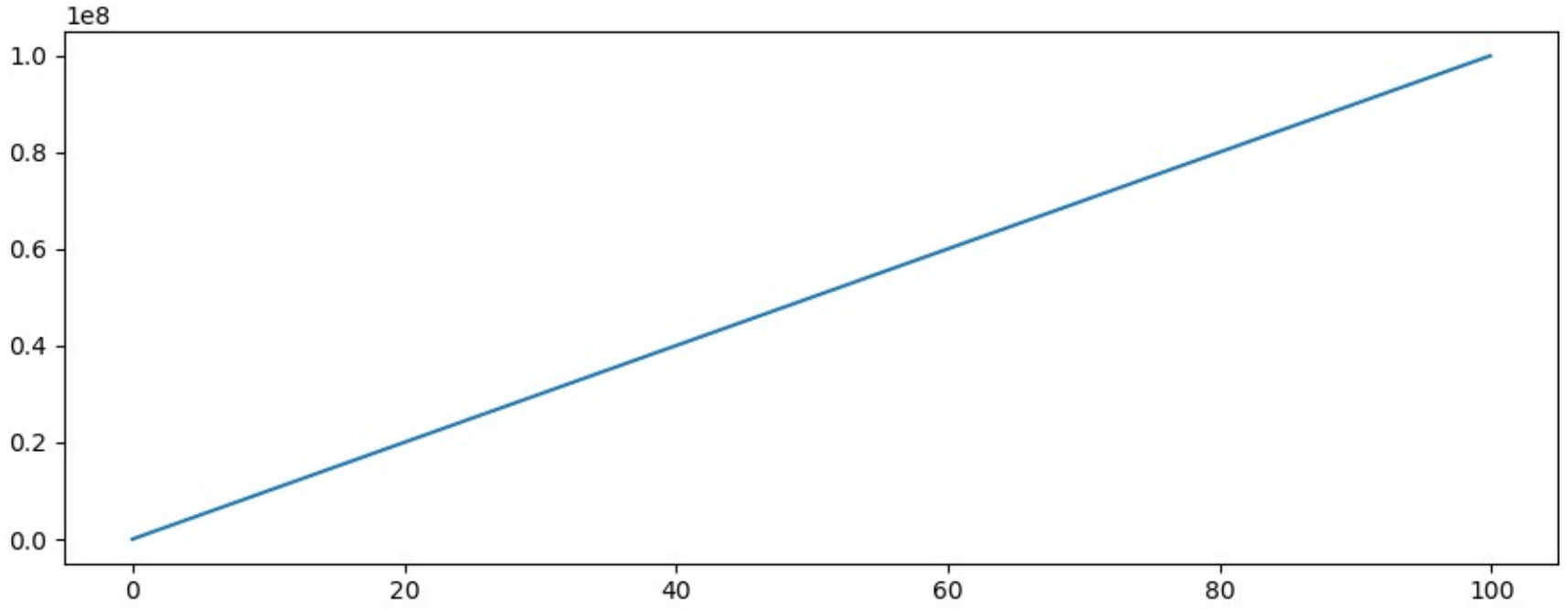
- Axis ranges: must be same if you want to compare figures
- Axis scale: linear or log, based on range of values
- Axis labels: always label the axis. Give name and units. eg. response time (sec)
- Axis tick labels: format it right eg. 0.000000034m vs 34nm



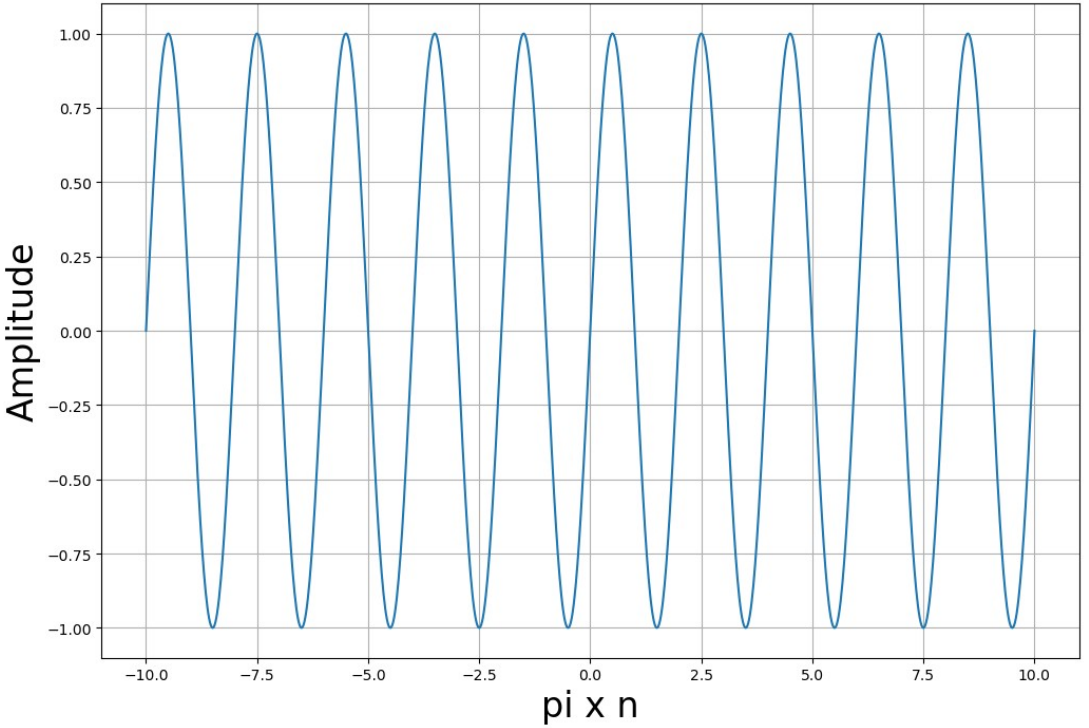
Linear scale in y axis



Log scale in y axis



Misleading axis scales



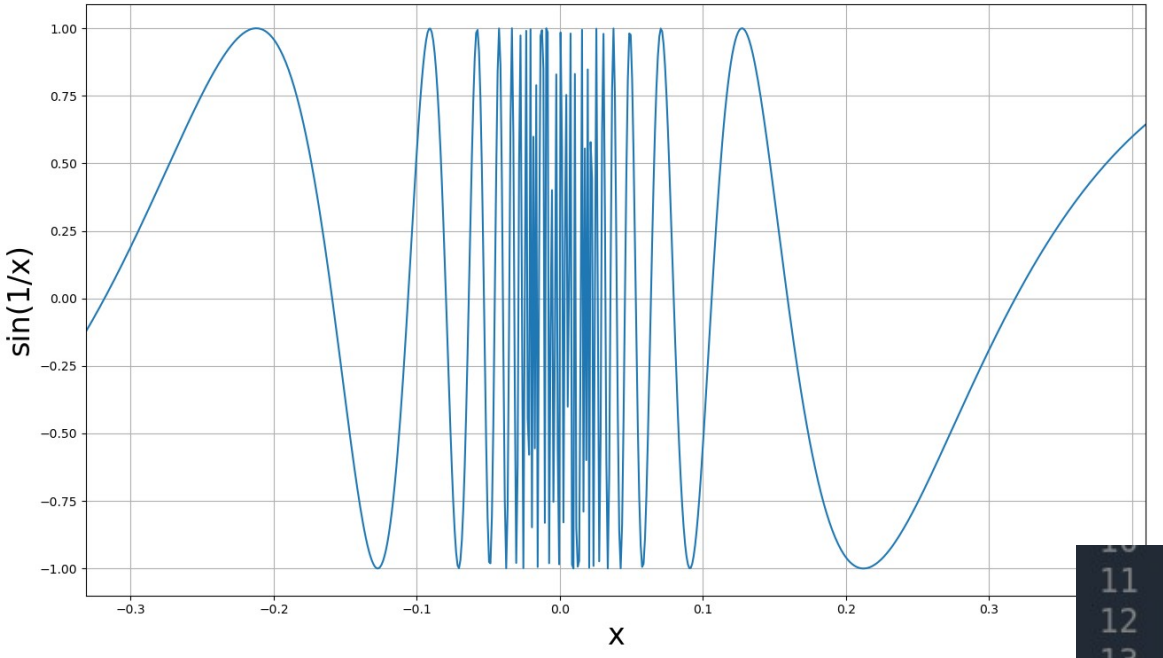
A good plot

sin(x)

```python
# plotting a sine wave

import numpy as np
import matplotlib.pyplot as plt

xticks = np.linspace(-10, 10, num=1000)

xvalues = xticks * np.pi
yvalues = np.sin(xvalues)

plt.plot(xticks,yvalues)
plt.grid(True)
plt.xlabel('pi x n', fontsize=24)
plt.ylabel('Amplitude', fontsize=24)

```

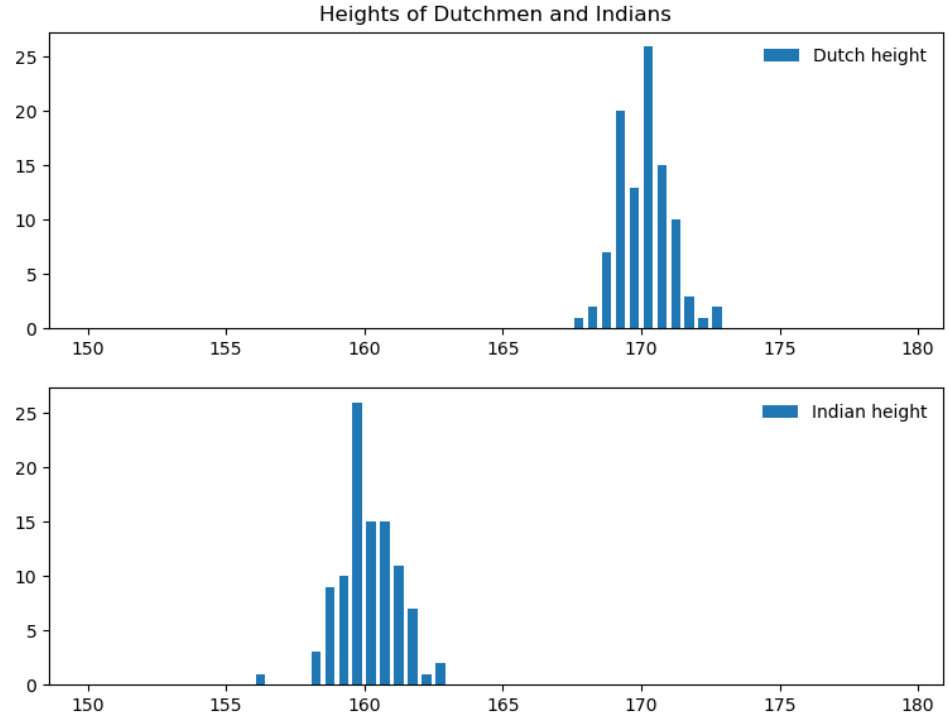
Some attributes of Numpy array a

`a.ndim`
`a.shape`
`a.size`



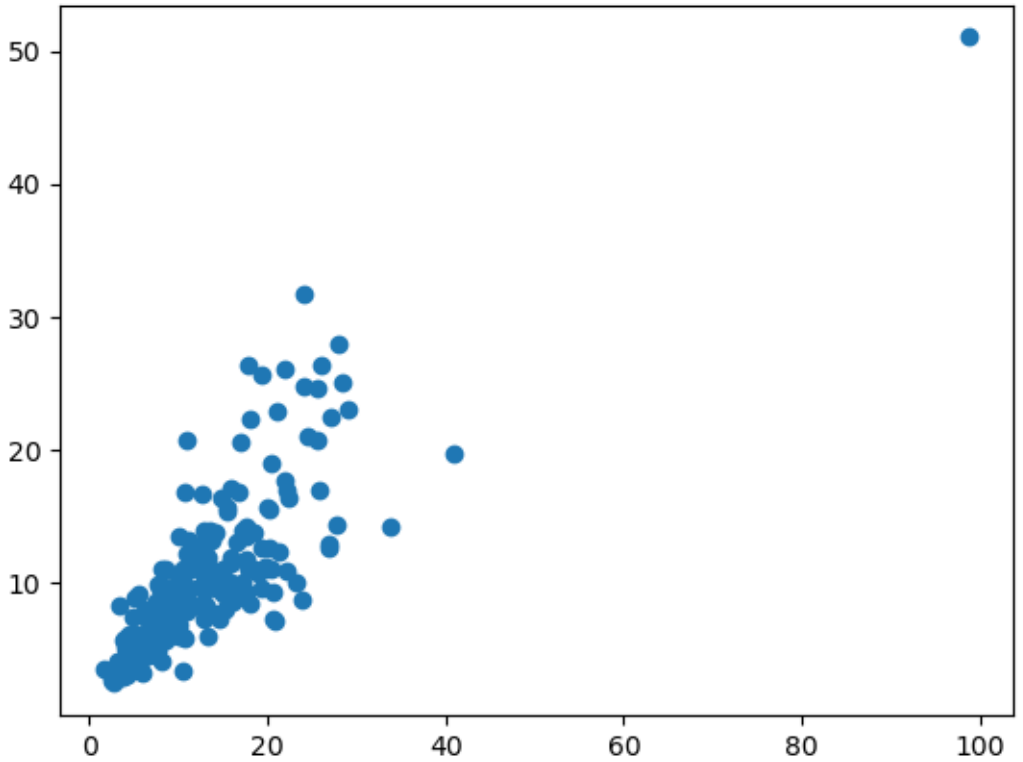

```
11  import numpy as np
12  import matplotlib.pyplot as plt
13
14  xticks = np.linspace(-5, 5, num=10000)
15
16  xvalues = 1/xticks
17  yvalues = np.sin(xvalues)
18
19  plt.plot(xticks,yvalues)
20  plt.grid(True)
21  plt.xlabel('x', fontsize=24)
22  plt.ylabel('sin(1/x)', fontsize=24)
23
```

Plotting histograms

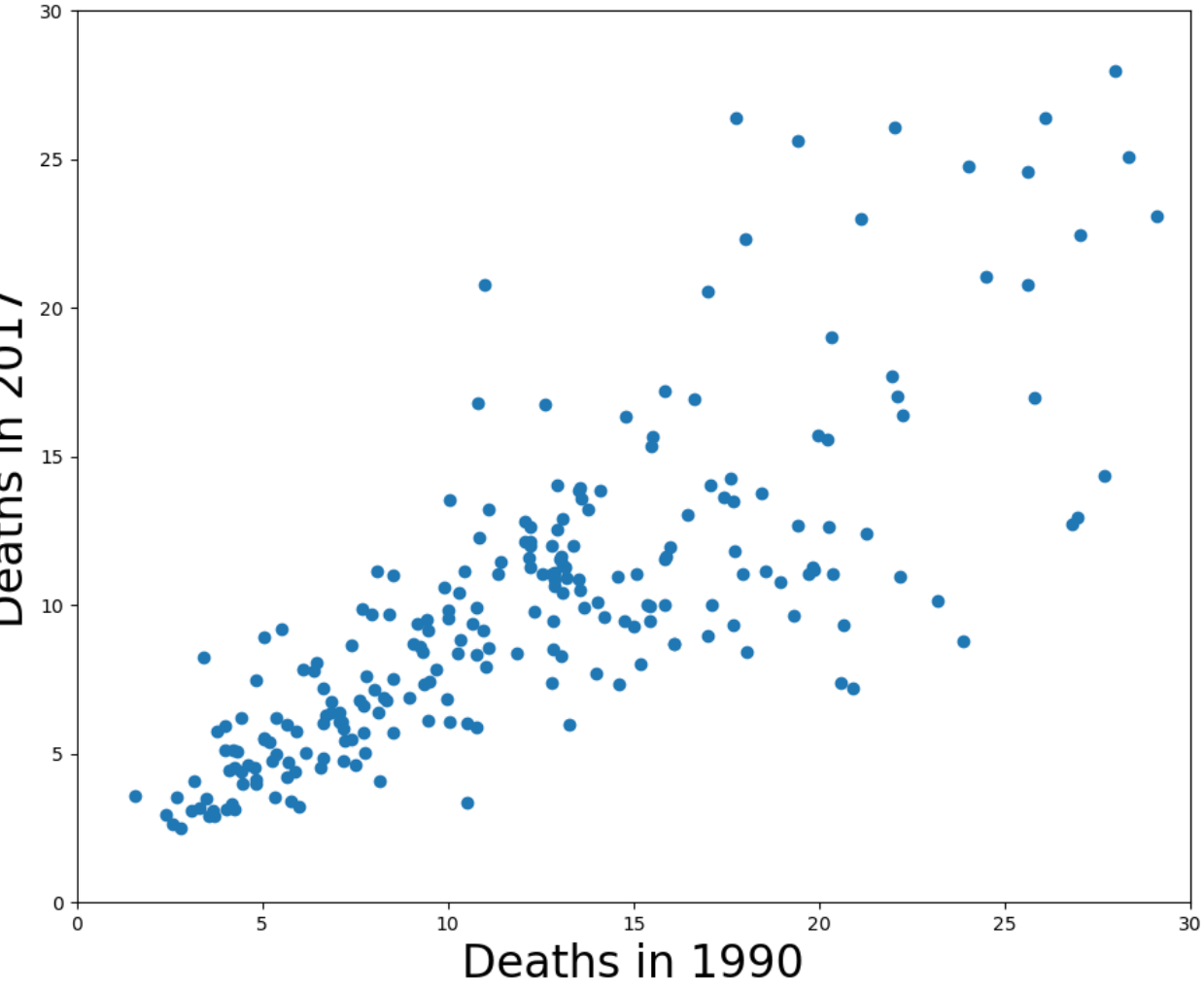



```python
import numpy as np
import matplotlib.pyplot as plt


hBins = np.arange(150,180,0.5)

plt.subplot(211)
histD = plt.hist(htD,label='Dutch height',bins=hBins,rwidth=0.7)
plt.title("Heights of Dutchmen and Indians")
leg = plt.legend(frameon=False)
plt.subplot(212)
histI = plt.hist(htI,label='Indian height',bins=hBins,rwidth=0.7)
leg = plt.legend(frameon=False)

```

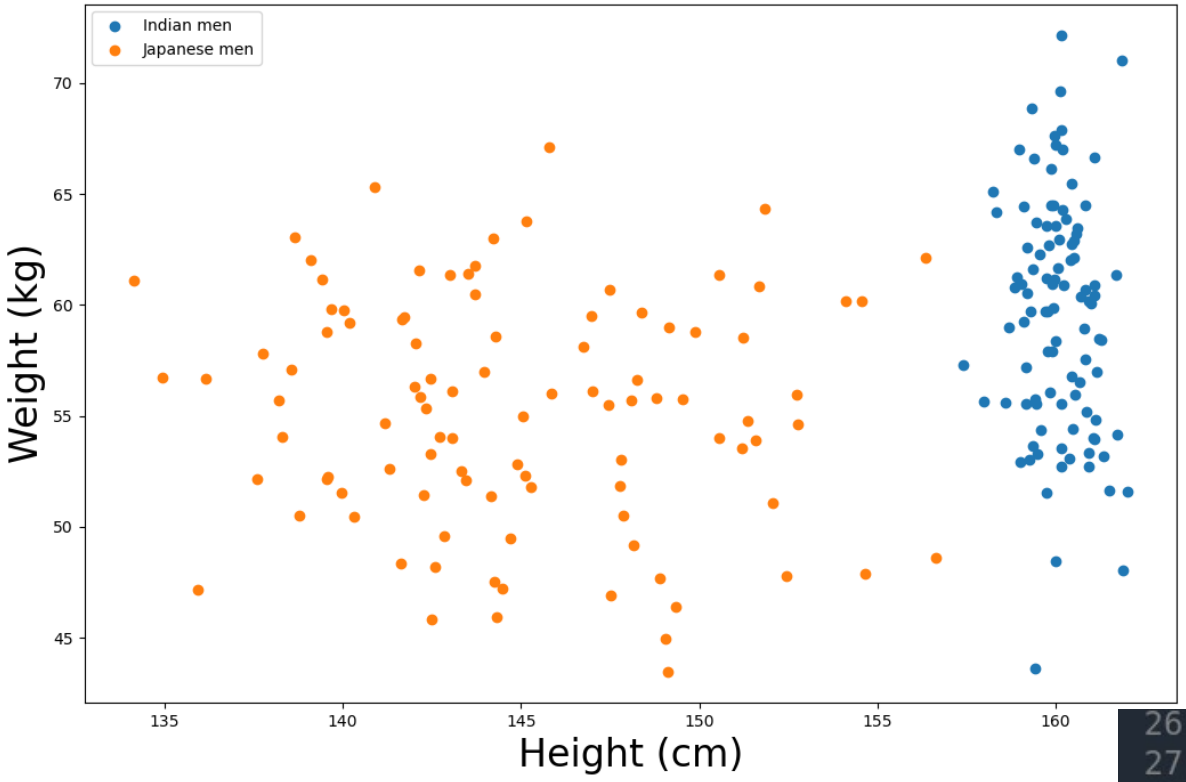
Scatter plots



Not a good plot




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

data = pd.read_csv('suicide_data.csv')
arr = data.to_numpy()
oldValues = arr[:,0]
newValues = arr[:,1]
plt.scatter(oldValues,newValues)
plt.xlim(0,30)
plt.ylim(0,30)
plt.xlabel('Deaths in 1990', fontsize=24)
plt.ylabel('Deaths in 2017', fontsize=24)
plt.show()
```




```
26
27  plt.scatter(x, y)
28  plt.scatter(x1, y1)
29  plt.legend(['Indian men', 'Japanese men'])
30  plt.xlabel('Height (cm)', fontsize=24)
31  plt.ylabel('Weight (kg)', fontsize=24)
32  plt.show()
33
```