Exercises

(Optional) If you are using Google Colab, start again with uploading the required files 'exercise_data.csv'.

In [1]:
```
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

Use the included data matrix 'exercise_data.csv'. Solve the following exercises.

1. Load the data into a Pandas data frame.
2. Find all objects of class 1 whose height is over 25 units. Print all of their attributes. *Hint*: you can combine the results of two comparisons using the '&' operator (class is 1 *and* height is over 25). Parentheses may be needed.
3. Compute the mean of each attribute for the whole data.
4. Make a plot of weights (y axis) vs. heights (x axis).
5. Make a plot of weights (y axis) vs. "size" (=$height \cdot width^2$, x axis).
6. Like 4-5 but limited to objects of class 1.
7. Do you see anything interesting in the plots?

In [2]:
```
# 1
ex1 = pd.read_csv('exercise_data.csv')
print(ex1)
```

```
     Unnamed: 0     height     width      weight  class
0             0  18.022890  5.182596   50.527110      1
1             1  16.194651  4.383156   28.829918      1
2             2  18.150373  4.556691   38.565482      1
3             3  13.846924  3.798662   16.413877      1
4             4  27.364656  7.777790  166.997659      1
..          ...        ...       ...         ...    ...
124         124  14.394126  1.623162   20.133995      3
125         125  14.715369  1.571027   18.098519      3
126         126  14.995934  1.613972   20.634388      3
127         127  14.345058  1.604480   18.507403      3
128         128  14.879172  1.675843   20.413188      3

[129 rows x 5 columns]
```

In [3]:
```
# 2
ex2 = ex1[(ex1["height"] > 25) & (ex1["class"] == 1)]
print(ex2)
```

```
    Unnamed: 0     height     width      weight  class
4            4  27.364656  7.777790  166.997659      1
16          16  25.640493  6.852879  122.388594      1
21          21  25.343498  7.226682  135.671038      1
33          33  26.893780  7.624959  164.284870      1
44          44  32.085549  9.016029  263.542786      1
```

In [4]:
```
# 3
print(ex1.mean())
```

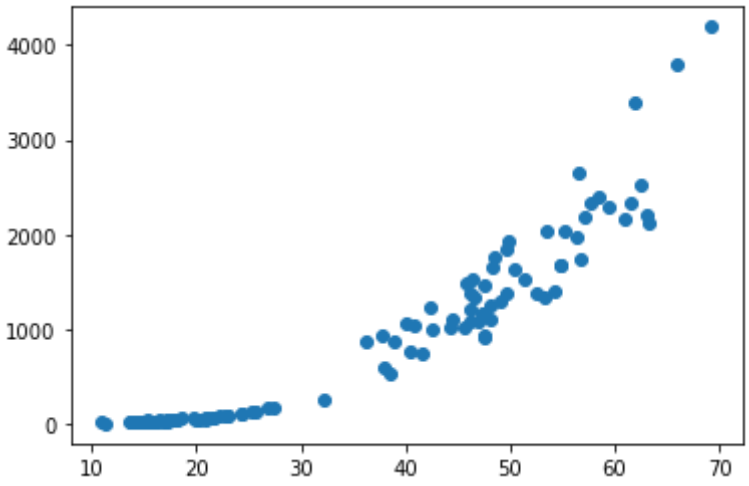
```
Unnamed: 0     64.000000
height         31.938985
```

```
width          15.431009
weight        735.292601
class           1.860465
dtype: float64
```

In [5]:
```
# 4
y = ex1["weight"]
x = ex1["height"]

plt.scatter(x,y)
```

Out[5]: <matplotlib.collections.PathCollection at 0x249ad75afa0>

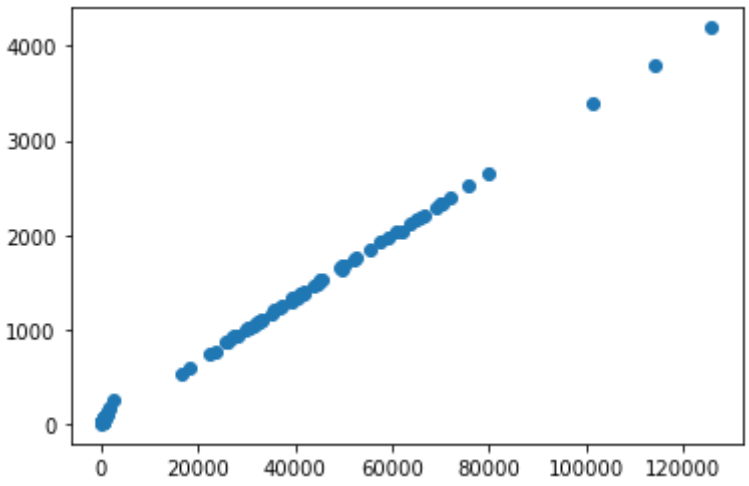


In [6]:
```
# 5
w = ex1["width"]

w2 = pow(w,2)
size = x * w2

plt.scatter(size,y)
```

Out[6]: <matplotlib.collections.PathCollection at 0x249ad856b50>



In [7]:
```
# 6
ex6 = ex1[ex1["class"] == 1]

y2 = ex6["weight"]
height = ex6["height"]

width = ex6["width"]
```

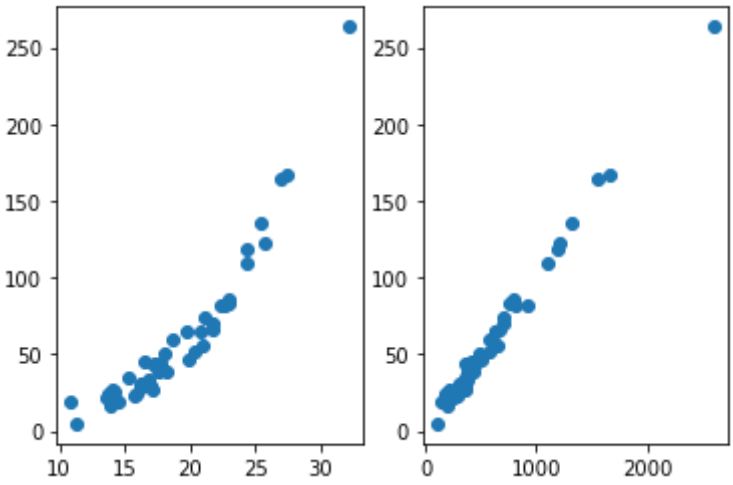
```
width2 = pow(width,2)
size2 = height * width2

plt.scatter(height, y2)

plt.subplot(1,2,1)
plt.scatter(height, y2)

plt.subplot(1,2,2)
plt.scatter(size2, y2)
```

Out[7]: <matplotlib.collections.PathCollection at 0x249ad92d700>



7

They are kind of messy with overlaps (although might be because my mistakes in some part). Also they are exponential in general. As weight increases so does height and size generally.