Applied Statistics MATH 661 Assignment #3

September 24, 2019

Andres Castellano

---

1 Excercise 2.8 Protein and Fat

A data set containing the dietary infomation of meals from a sample or the population of first year college students can be used to study the relationship between protein and fat intake of those students. For example, the protein and fat amounts of every meal in the data set can be plotted on different axes of a scatter diagram. In this way, it may be determined whether there is a relationship between this variables.

---

2 Excercise 2.24, 2.25, and 2.26 Bone Strenght

2.1 IPC 2.24

[1]:
```
import pandas as pd
import numpy as np

data = {'Nondominant':[15.7,25.2,17.9,19.1,12.0,20.0,12.3,14.4,
                       15.9,13.7,17.7,15.5,14.4,14.1,12.3],
        'Dominant':[16.3,26.9,18.7,22.0,14.8,19.8,13.1,17.5,20.1,
                    18.7,18.7,15.2,16.2,15.0,12.9]}
frata = pd.DataFrame(data,
                             index=[np.arange(1,16)],
                             columns=pd.Index(['Nondominant','Dominant'],
                                              name='ID'))
%matplotlib inline
frata
```

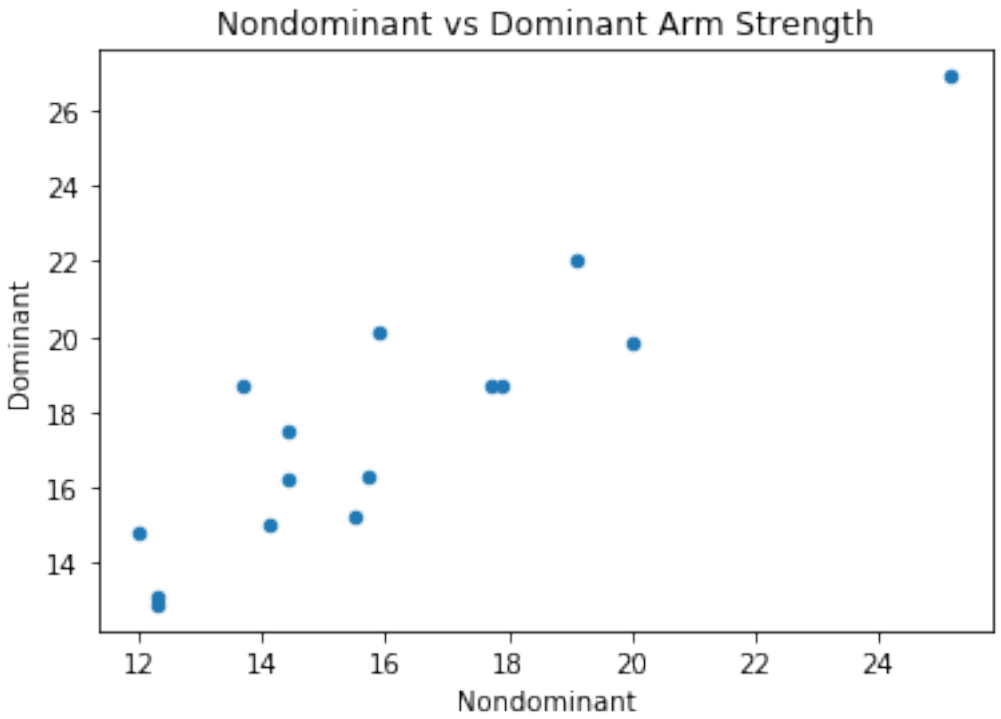
[1]:
```
ID  Nondominant  Dominant
1          15.7      16.3
2          25.2      26.9
3          17.9      18.7
4          19.1      22.0
```

5	12.0	14.8
6	20.0	19.8
7	12.3	13.1
8	14.4	17.5
9	15.9	20.1
10	13.7	18.7
11	17.7	18.7
12	15.5	15.2
13	14.4	16.2
14	14.1	15.0
15	12.3	12.9

2.1.1 a) Create a Scatterplot of The Data with the Nondominant and Dominant arm strengths on the x and y axis respectively.

[2]:
```
%matplotlib inline
frata.plot.scatter('Nondominant','Dominant',
                   title='Nondominant vs Dominant Arm Strength')
```

[2]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x254cc3175f8>
```



2.1.2 b,c,d, and e) Describe the overall pattern in the scatterplot and any striking deviations from the pattern

The Data appears to follow a linear behavior densely distributed between 12,000 and 20,000 Newton-Meters. One notable instance occurs approximately at 26,000 Newton Meters. Although the instance seems to behave according with the overall trend of the data, it is outside of the densely populated area mentioned above. However, it is not likely that the instance is an outlier since it appears to follow the overall linear trend of the rest of the data. It is possible that the sample did not include enought data between 22000 and 24000 Newton Meters.

2.2 IPC 2.25

[3]:
```
data2 = {'Nondominant':[17.0,16.9,17.7,21.2,21.0,14.6,31.5,14.9,
                        15.1,13.5,13.6,20.3,17.3,14.6,22.6],
         'Dominant':[19.3,19.0,25.2,37.7,40.3,20.8,36.9,21.2,
                     19.4,20.4,17.1,26.5,30.3,17.4,35.0]}
frata2 = pd.DataFrame(data2,
                      index = [np.arange(16,31)],
                      columns = pd.Index(['Nondominant','Dominant'],name='ID'))
frata2
```

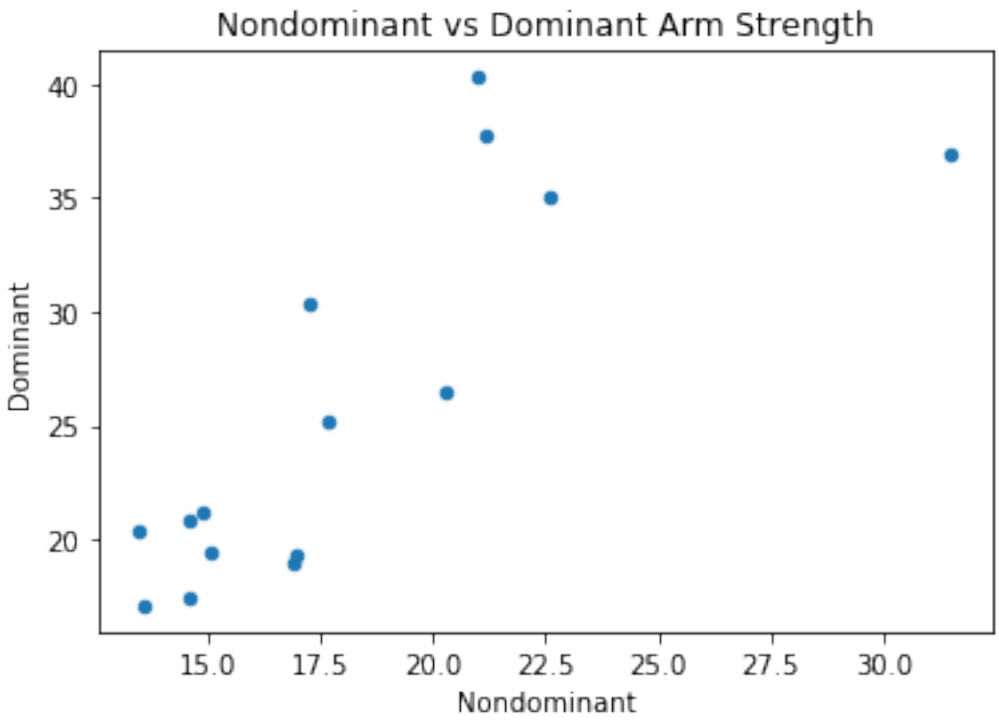
[3]:

ID	Nondominant	Dominant
16	17.0	19.3
17	16.9	19.0
18	17.7	25.2
19	21.2	37.7
20	21.0	40.3
21	14.6	20.8
22	31.5	36.9
23	14.9	21.2
24	15.1	19.4
25	13.5	20.4
26	13.6	17.1
27	20.3	26.5
28	17.3	30.3
29	14.6	17.4
30	22.6	35.0

2.2.1 a) Make a scatter plot of the data.

[5]:
```
%matplotlib inline
frata2.plot.scatter('Nondominant','Dominant',
                    title='Nondominant vs Dominant Arm Strength')
```

[5]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x254cc60b588>
```

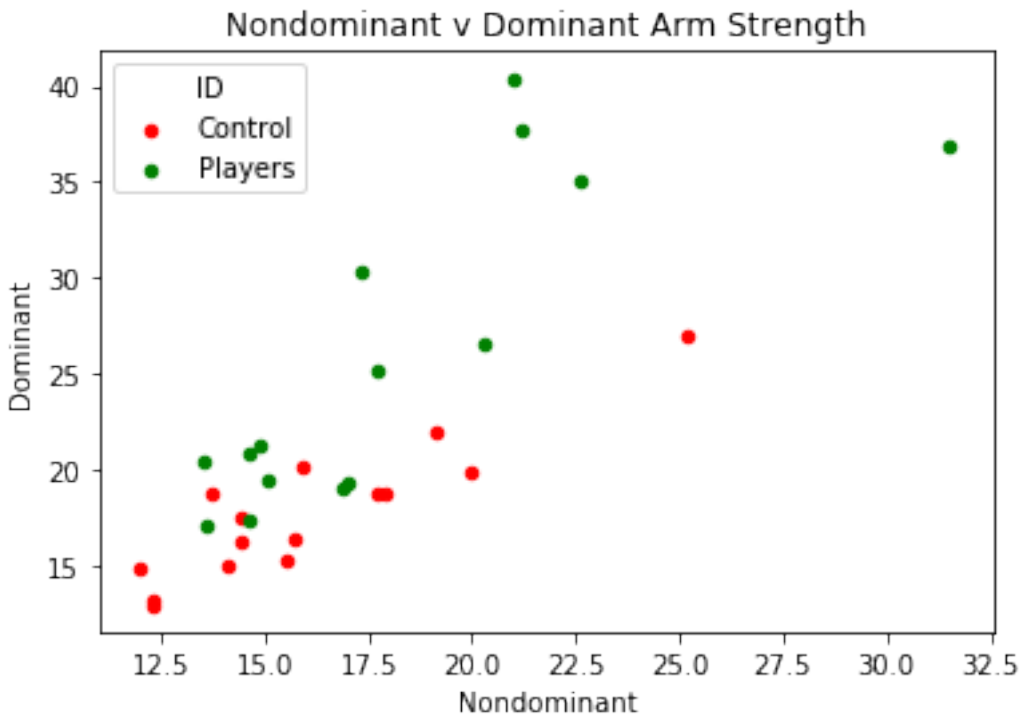


2.2.2 b,c,d,e) Describe the overall pattern in the scatterplot and any striking deviations from the pattern.

The data is clustered between 15,000 and 22,500 Newton-Meters. However, the data does not seem to be linearly related. In fact, at first glance, the data appears to follow a positive exponential distribution except for one possible outlier found at approximately 34,000 Newton-Meters.

[22]:
```
#ax1 = frata.plot.scatter('Nondominant','Dominant',
 #                  title='Nondominant vs Dominant Arm Strength');
#ax2 = frata2.plot.scatter('Nondominant','Dominant',
  #                  title='Nondominant vs Dominant Arm Strength',
   #                       color='r')
ax1 = frata.plot(kind='scatter',x='Nondominant',y='Dominant',
                 color='r',label='Control')
ax2 =frata2.plot(kind='scatter',x='Nondominant',y='Dominant',
                 color='g',ax=ax1,
                  title='Nondominant v Dominant Arm Strength',label='Players')
print(ax1==ax2)
```

```
True
```


Nondominant v Dominant Arm Strength
ID
Control
Players
Dominant
Nondominant
40
35
30
25
20
15
12.5
15.0
17.5
20.0
22.5
25.0
27.5
30.0
32.5

[ ]: