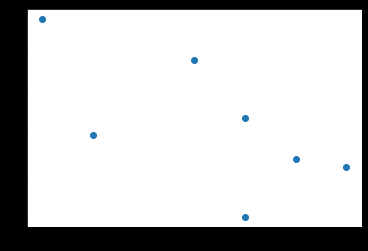
**DATA SCIENCE**

**STATISTICS**

**ASSIGNMENT-4**

**EXERCISE 1.**Use the data in the table below to generate a scatter plot, calculate Pearson’s correlation, and interpret the relationship between two variables.

| **X** | **Y** |
| --- | --- |
| 0 | 88 |
| 6 | 52 |
| 4 | 64 |
| 1 | 60 |
| 5 | 54 |
| 3 | 78 |
| 4 | 40 |



r = -0.6736717993982941

**Interpretation :**

The correlation between x and y variables is around -0,67.

As you see in scatter plot above, y tends to decrease as x increase.

For this reason the reationship between these variables is strong and slope of this relationship is negative.

**Note :** This interpretation is based on the following data.

 .00-.19 “very weak”

 .20-.39 “weak”

 .40-.59 “moderate”

 .60-.79 “strong”

 .80-1.0 “very strong”

**EXERCISE 2.**In some cases we, data analysts, make calculations using online calculators. Answer the following questions using the online calculator available at this link: [Normal Distribution Calculator](http://onlinestatbook.com/2/calculators/normal_dist.html)

1. The heights of male students in a particular town are normally distributed with a mean of 65 inches and a standard deviation of 1.7. What percentage of these students is taller than 66.7 inches?
2. A data set was created by asking 300 students about their weights. The mean is 60 kg. standard deviation is 8 kg. How many students weighed more than 52 kg and less than 68 kg?

1. % 15.87

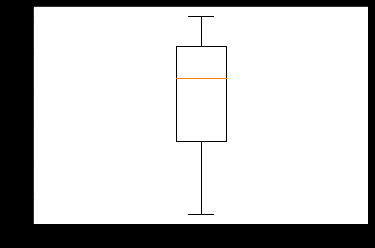
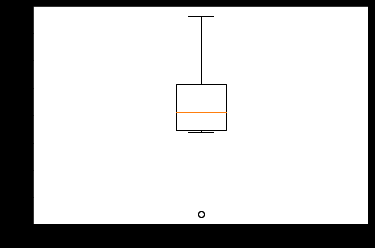
2. 204-205 students

**EXERCISE 3.** The last 10 game results between LA Lakers and LA Clippers are given in the table below. Draw two boxplots for the points scored by these 2 teams and answer the following questions.

|  |  |  |
| --- | --- | --- |
| **Date** | **LA Lakers** | **LA Clippers** |
| July 31, 2020 | 103 | 101 |
| March 8, 2020 | 112 | 103 |
| December 26, 2019 | 106 | 111 |
| October 23, 2019 | 102 | 112 |
| July 7, 2019 | 87 | 93 |
| April 6, 2019 | 122 | 117 |
| March 5, 2019 | 105 | 113 |
| February 1, 2019 | 123 | 120 |
| December 29, 2018 | 107 | 118 |
| October 7, 2018 | 87 | 103 |

1. Which team has a greater median, Lakers or Clippers?
2. Which team has greater IQR, Lakers or Clippers?
3. Which team has a higher percentage of scores above its median?

Boxplot for Lakers Boxplot for Clippers



|  |  |  |
| --- | --- | --- |
|  | LAKERS | CLİPPERS |
| Median | 105.5 | 111.5 |
| Q1 | 102.25 | 103 |
| Q2 | 105.5 | 111.5 |
| Q3 | 110.75 | 116 |
| IQR | 8.5 | 13 |

1. Clippers
2. Clippers
3. Every data set have %50 of values above and below their median. So Lakers and Clippers have same percentage of scores above their median.