Name: Prerana.M.S Class: 3 E

SRN: PES2UG19CS296 Date: 28.08.2020

STATISTICS FOR DATA SCIENCE

ASSIGNMENT 10

1. The article “The Selection of Yeast Strains for the Production of Premium Quality South African Brandy Base Products” (C. Steger and M. Lambrechts, Journal of Industrial Microbiology and Biotechnology, 2000:431– 440) presents detailed information on the volatile compound composition of base wines made from each of 16 selected yeast strains. Following are the concentrations of total esters (in mg/L) in each of the wines.

a. Compute the mean concentration.

b. Compute the median concentration.

c. Compute the first quartile of the concentrations.

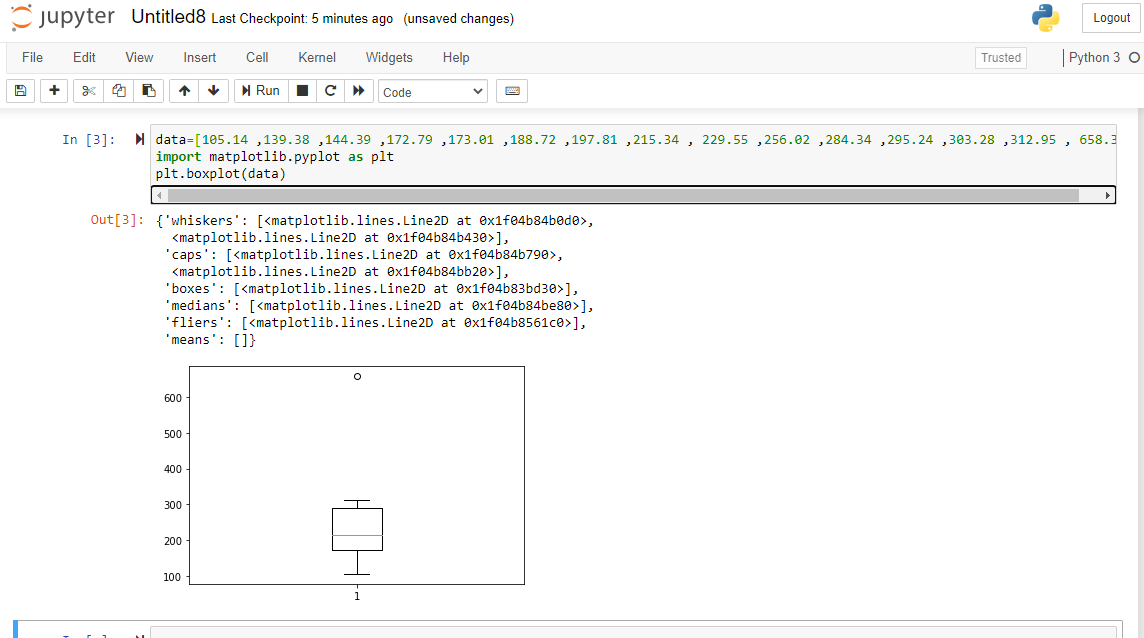
d. Compute the third quartile of the concentrations.

e. Construct a boxplot for the concentrations. What features does it reveal?

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 284.34 | 173.01 | 229.55 | 312.95 | 215.34 | 188.72 | 144.39 | 172.79 | 139.38 |
| 197.81 | 303.28 | 256.02 | 658.38 | 105.14 | 295.24 | 170.41 |  |  |

Answer.

Concentrations: 105.14 ,139.38 ,144.39 ,170.41 ,172.79 ,173.01 ,188.72 ,197.81 , 215.34 , 229.55 ,256.02 ,284.34 ,295.24 ,303.28 ,312.95 , 658.38

1. Mean concentration: 3846.75 / 16 =**240.42**
2. Median concentration: **206.58**
3. First quartile of the concentrations : 105.14 ,139.38 ,144.39 ,172.79
4. Third quartile of the concentrations : 295.24 ,303.28 ,312.95 , 658.38
5. 

This plot reveals that more data is distributed above the median. It also shows one outlier around 650 value.