Department of Mathematical Sciences
Faculty of Applied Sciences
Wayamba University of Sri Lanka
B.Sc. (General/Joint Major/Special) Degree Program
Academic Year 2021/2022 – Semester I

STAT 3232– Data Analysis & Preparation of Statistical Reports
Tutorial #03

---

01. Consider the dataset "student" contains the variables student id, age, gender, test_score and favorite_subject in a sample of 100 people.
    a) Import the data set to R and obtain an overview of the variables.
    b) Find the average and the standard deviation of the variable "age" in the dataset?
    c) Obtain the median and range of the variable "test_score". What does the range indicate about the spread of the variable?
    d) Find the Inter Quantile Range of the "test_score".
    e) Obtain number of students for each "favorite_subject".
    f) Create a histogram to visualize the distribution of the variable "test_score" and sort the variable in descending order

02. a) Generate 150 observations from a normal distribution with $\mu=2$ and $\sigma=10$.
    b) Generate 30 random data points from binomial distribution with the sample size=50 and the probability of success=0.25.

03. Consider the following data set.

ID	age	gender	income	education	happiness_score
1	34	Female	54512.73	Bachelor's	8.12
2	45	Male	48790.25	Master's	6.99
3	27	Female	51236.49	PhD	9.35
4	56	Male	50298.81	High School	5.88
5	40	Female	45981.22	Bachelor's	7.23
6	22	Female	52203.99	Master's	8.35
7	49	Male	47490.18	PhD	6.46
8	33	Male	51473.47	Bachelor's	8.77
9	65	Female	48123.88	High School	6.57
10	30	Female	49811.33	PhD	7.68
11	52	Male	53021.76	Bachelor's	8.22
12	29	Female	49568.94	Master's	7.1
13	39	Male	50679.01	High School	7.89

14	48	Male	48299.87	PhD	6.81
15	37	Female	51409.6	Bachelor's	8.01
16	43	Female	50213.42	Master's	7.55
17	32	Male	52134.77	High School	6.92
18	58	Female	49918.65	PhD	7.29
19	31	Male	47567.92	Bachelor's	8.87
20	60	Male	48750.3	Master's	7.77
21	25	Female	50897.12	High School	6.99
22	42	Female	52601.89	Bachelor's	8.12
23	51	Male	49020.55	PhD	7.88
24	36	Male	50430.14	Master's	7.36
25	44	Female	51984.26	High School	6.68
26	38	Female	50345.78	PhD	7.92
27	53	Male	51123.9	Bachelor's	8.45
28	28	Female	48890.32	Master's	7.77
29	47	Male	49765.13	High School	6.78
30	59	Male	48321.78	PhD	7.2

a) Import the data set in to R.
b) Create Histograms for “age”, “income”, and “happiness_score” and colour the histogram as red, green and blue respectively.
c) Create a box plot to compare “happiness_score” for two separate genders and interpret.
d) Create a scatter plot for “age” over “happiness_score” and comment on the relationship.
e) Create bar plot and pie chart for education and colour it using different colours. Interpret the plots.

Submit on or before 3rd of April 2024 at 4.00 p.m.

Note that your commands should be written in R editor. Both commands and outputs should be copied into a word file and upload to the LMS as a pdf document.