Sri Lanka Institute of Information

Technology



Lab Submission

<Lab sheet No 04>

**<IT24101668>**

**<Nanayakkara N N D A S>**

**Probability and Statistics | IT2120**

Exercise

1. Import the dataset (’Exercise.txt’) into R and store it in a data frame called “branch data”.

A screenshot of a computer program

Description automatically generated

2. Identify the variable type and scale of measurement for each variable.

* Sales\_X1: Continuous, Ratio scale
* Advertising\_X2: Continuous, Ratio scale
* Years\_X3: Discrete, Ratio scale
* Branch: Categorical, Nominal

3. Obtain boxplot for sales and interpret the shape of the sales distribution.

A screenshot of a computer

Description automatically generated

4. Calculate the five number summary and IQR for advertising variable.

A screenshot of a computer

Description automatically generated

5. Write an R function to find the outliers in a numeric vector and check for outliers in years variables.

A screenshot of a computer

Description automatically generated

# Lab Exercise 4 (Descriptive Statistics)

Major League Baseball is known as ”America’s pastime.” The role of Major League Baseball has been ingrained into American culture. The heroic figures and memorable moments of Major League Baseball reflect the type of attitude that American culture is built on. Given below are some measurements observed in this significant sport during the 1998 league.

X1 = Team Attendance

(Average number of spectators for a match that the team play)

X2 = Team Salary

(Earning of the team)

X3 = Years

(Years since the team has owned a stadium)

Before starting the lab sheet, you need to create a folder in your desktop and save all your working inside the folder. Set the working directory to that folder using the following command:

setwd("paste the path of the folder")

**Eg:-** setwd("D:\\2025 - Sem 2\\IT2120\\Lab Sessions\\Lab 04")

1. Identify the variables and enter the given data set into R.

A screenshot of a computer code

Description automatically generated

1. Obtain the following for each variable
   1. Box-Plot, Histogram and Stem-Leaf Plot.

A computer screen shot of text

Description automatically generated

* 1. Mean, Median and Standard Deviation.

A screenshot of a computer code

Description automatically generated

* 1. First and Third Quartile.

A screenshot of a computer

Description automatically generated

* 1. Interquartile Range.

A close up of a logo

Description automatically generated

1. Write a function to find the modes of a given set of values. Check the function by finding the mode of the variable ”Years”.

A computer code on a white background

Description automatically generated

1. Write a function that would produce the outliers when the values are given. Check the function with the 3 variables in the dataset.

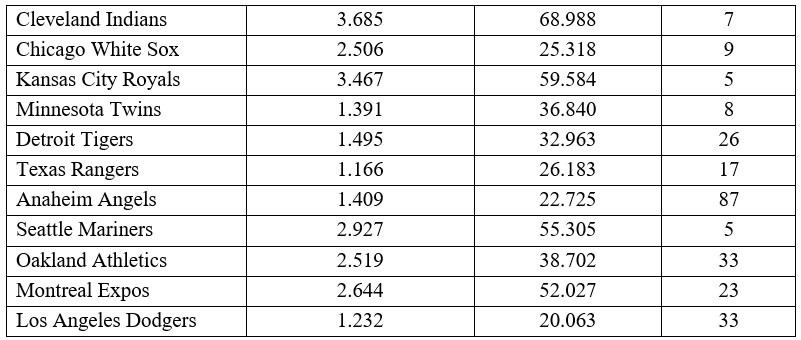
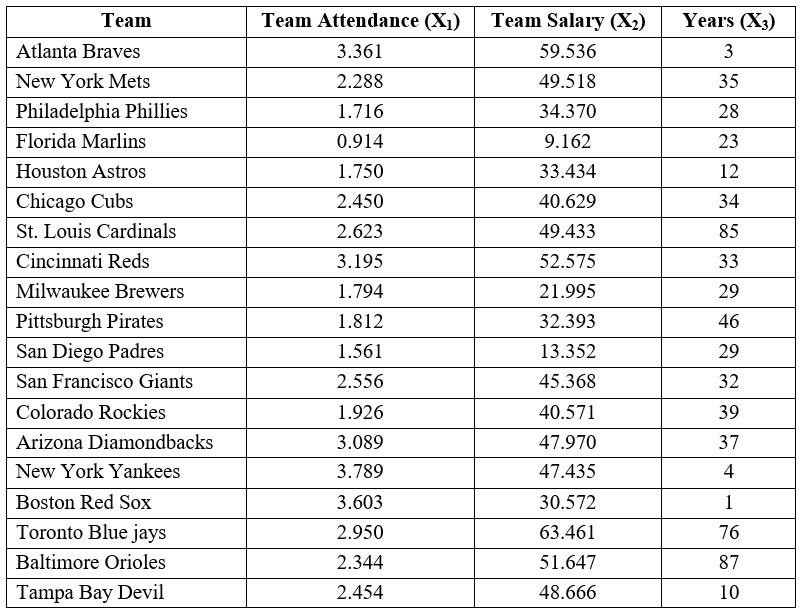
A computer screen shot of a program

Description automatically generated

A computer screen shot of a computer program

Description automatically generated

**DataSet**



Part 1

A screenshot of a computer

Description automatically generated

Part 2 - A

1. A screenshot of a computer

   Description automatically generated

A screenshot of a computer

Description automatically generated

1. A screenshot of a computer

   Description automatically generated

A screenshot of a computer

Description automatically generated

1. A screenshot of a computer

   Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer program

Description automatically generated

Part 2 – B

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Part 2 – C

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Part 2 – D

A screenshot of a computer

Description automatically generated

A screenshot of a data

Description automatically generated

Part 3

A screenshot of a computer

Description automatically generated

Part – 4

A screenshot of a computer

Description automatically generated