**Student: Seif Kungulio**

**Date: 01/15/2025**

**Subject: Project 1**

**Class: DSCI 502**

**Section: 01W**

**Instructor: Sean Yang**

**File Name: Project1\_Kungulio\_Seif.docx**

1. Read the dataset in CarInsurances.xlsx into R. Call the loaded data Insurance. Make sure that you have the directory set to the correct location for the data.

A screenshot of a computer

Description automatically generated

> # 1. Read the dataset in CarInsurances.xlsx into R. Call the loaded data

> # Insurance. Make sure that you have the directory set to the correct

> # location for the data.

>

> # Set the working directory

> setwd("C:/Projects/DSCI 502/Week 1")

>

> # Import necessary libraries

> library(readxl)

>

> # Import the data set

> Insurance <- read\_excel("CarInsurances.xlsx")

>

> # Display the dimension of the data frame

> dim(Insurance)

[1] 52 4

1. How many rows in the data set?

A black background with white text

Description automatically generated

> # 2. How many rows in the data set?

>

> # Display the number of rows

> cat("There are", nrow(Insurance), "number of rows\n")

There are 52 number of rows

1. How many columns in the data set?

A black background with white text

Description automatically generated

> # 3. How many columns in the data set?

>

> # Display the number of columns

> cat("There are", ncol(Insurance), "number of columns\n")

There are 4 number of columns

1. Assign the first eight rows of the data set to a variable: first.eight.rows and print it out using print() function.

A screenshot of a computer program

Description automatically generated

> # 4. Assign the first eight rows of the data set to a variable:

> # first.eight.rows and print it out using print() function.

>

> # Assign the first eight rows to first.eight.rows variable

> first.eight.rows <- head(Insurance, 8)

>

> # Print the first eight rows

> print(first.eight.rows)

# A tibble: 8 × 4

State MRC FC AD

*<chr>* *<dbl>* *<dbl>* *<dbl>*

1 Average 699 1537 838

2 Alabama 563 1367 804

3 Alaska 419 1105 686

4 Arizona 684 1527 843

5 Arkansas 578 1449 871

6 California 629 1654 1025

7 Colorado 713 1738 1025

8 Connecticut 1029 1984 955

>

1. Assign the last five rows of the data set to a variable: five.rows and print it out using print() function.

A screenshot of a computer program

Description automatically generated

> # 5. Assign the last five rows of the data set to a variable:

> # five.rows and print it out using print() function.

>

> # Assign the last five rows to five.rows variable

> five.rows <- tail(Insurance, 5)

>

> # Print the last five rows

> print(five.rows)

# A tibble: 5 × 4

State MRC FC AD

*<chr>* *<dbl>* *<dbl>* *<dbl>*

1 Virginia 431 1039 608

2 Washington 545 1009 464

3 West Virginia 635 1501 866

4 Wisconsin 491 1084 593

5 Wyoming 329 1085 756

>

1. List all objects in the memory using two methods.

A screenshot of a computer error

Description automatically generated

> # 6. List all objects in the memory using two methods.

>

> # Use ls() method to list all the objects

> ls()

[1] "first.eight.rows" "five.rows" "Insurance"

>

> # Use objects() method to list all the objects

> objects()

[1] "first.eight.rows" "five.rows" "Insurance"

>

1. We want to summarize the data. To do it, we may use the summary function. Before asking others for help, it’s generally a good idea for you to try to help yourself either using help() function or google it. Please help yourself and summarize the data first. Then answer the following questions:

A screenshot of a computer screen

Description automatically generated

> # 7. We want to summarize the data. To do it, we may use the summary function.

> # Before asking others for help, it’s generally a good idea for you to try to

> # help yourself either using help() function or Google it. Please help yourself

> # and summarize the data first. Then answer the following questions:

>

> # Remove the first row which represents the Averages

> Insurance <- Insurance[-1, ]

>

> # Display the statistical summary of the "Insurance" data frame

> summary(Insurance)

State MRC FC AD

Length: 51 Min. : 309.0 Min. : 961 Min. : 464.0

Class : character 1st Qu.: 492.0 1st Qu.:1106 1st Qu.: 651.0

Mode : character Median : 585.0 Median :1449 Median : 795.0

Mean : 700.2 Mean :1539 Mean : 838.6

3rd Qu.: 825.0 3rd Qu.:1695 3rd Qu.: 999.5

Max. :2696.0 Max. :3986 Max. :1794.0

>

* 1. What is the mean of MRC (annual premium of Minimum Required Coverage)?

A screen shot of a computer

Description automatically generated

> ### 7.1. What is the mean of MRC (annual premium of Minimum Required Coverage)?

>

> # Display mean of MRC

> round(summary(Insurance$MRC)["Mean"], 1)

Mean

>

7.2 What is the mean of FC (annual premium of Full Coverage)?

A black screen with white text

Description automatically generated

> ### 7.2. What is the mean of FC (annual premium of Full Coverage)?

>

> # Display mean of FC

> round(summary(Insurance$FC)["Mean"], 0)

Mean

>

7.3 What is the mean of AD (annual premium differences between MRC and FC)?

A screen shot of a computer

Description automatically generated

> ### 7.3. What is the mean of AD (annual premium differences between MRC and FC)?

>

> # Display mean of AD

> round(summary(Insurance$AD)["Mean"], 1)

Mean

838.6

>