Data Analysis Report Template

clinton Moshe Mwachia

8/17/2021

Table Of Contents

0.1 Introduction

This introduction of a data analysis report can include: -

1. Summary of the analysis.
2. The aim of the analysis.
3. Background of the analysis.
4. The summary of the results and its conclusions.
5. The problem statement.

0.2 Research Questions.

This section describes the questions that will guide you in the type of visualizations you can perform, statistical tests you can do, the tables and summaries you can present.

The questions are a reflection of the main aim of the analysis.

1. How is **BMI** affected by body **weight**?.
2. How does **Salary** change with **Gender**?.
3. How is **Security** affected by **Regions**?

0.3 The Libraries

1. The libraries can be loaded at each **chunk** or all in a single **chunk** at the beginning of the chunk.
2. In case the library is not installed: -

   - Load the library and it will give you the option of installing.
   - Install the library using the **Install packages** pop up window from the packages tab.
   - Install manually using **install.packages('package')**.

0.4 The Data

In this section you will tell us the source of the data and how it will help us to answer the questions. The variables of the data and their descriptions will also be described in this section including their data type and summary statistics.

To include the link to the data set, do it as follows.

The data set used was obtained from ***mydata***.

To load different data formats, you will need a given library for different data formats. Use the following guide for a given data format: -

1. **.csv** files, you can use **read__csv()/read__csv2()** from the **tidyverse** library.
2. **.tsv** files, you can use **read__tsv()** from the **tidyverse** library.
3. **.txt** files, you can use **read.table()** from the **utils** library.
4. **.xls/.xlsx** files, you can use **read__xls()/read__xlsx()** from the **readxl** library.

0.5 Data Cleaning

Data cleaning may include: -

1. Handling **NA** and **NAN** values.
2. Handling **missing** data.
3. Renaming the column names into appropriate names.
4. Changing the data types into appropriate types.
5. Formatting decimals, currencies and percentages.

0.6 Formating

1. Italicize: You can italicize text by doing this ==> *your text*.
2. Bold: You bold text by doing this ==> **your text**.
3. Underscore: You can underscore text by doing this ==> ***your text***.
4. Superscript: 2^{20}.
5. Subscript: 2_{20}.
6. Strike through: ~~your word~~
7. Equation:

$$y = mx + c$$

8. Block Quote: You can a block quote using:

   my answer. (this should be in a new line).

9. Lists

- item one.

- item two.

    - sub item two.
    - sub item two.

0.7 Graphs

Figure 1 below shows the plot and i can write my observation here. Adding **message** in the code chunk options prevents printing of any information resulting from the plot.

In case you want to adjust the **height** and **width** of the output figure, you can use the code chunk options, *out.width* and *out.height* for the width and height respectively.

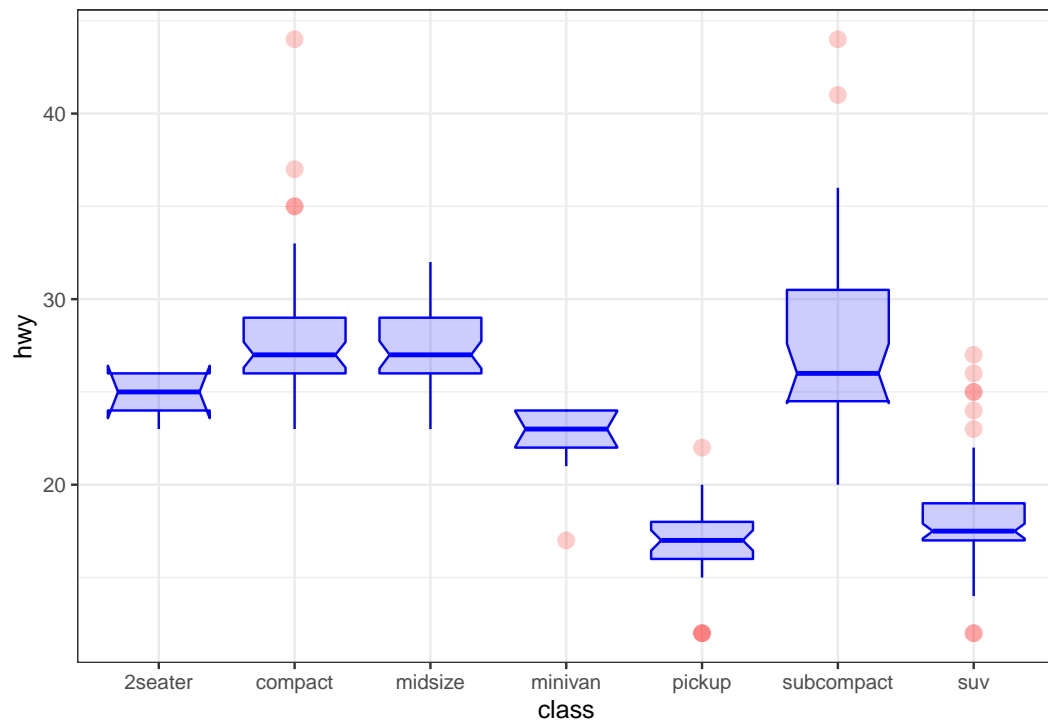NOTE: Try removing the **message** option to see the difference.

Figure 1: plot 1

0.8 Tables

A simple table.

Header	Header
Cell	Cell
Cell	Cell

3

Tables in R can be made using the **knitr** package and referenced as illustrated below. **kable** offers various formats of generating tables including *HTML*, *latex* and *simple.*

0.8.1   Simple Tables.