ACADGILD

SESSION 2: Introduction to working with R

Assignment 3

Submitted by: Munmun Ghosal
Login Id: munmun55@gmail.com
(M):+91-8007178659

Table of Contents

1. Problem Statement

1. How to Import SAS XPORT Files into R With The foreign package
2. How to Import SAS Files into R With The haven Package?
3. How to read Weka Attribute-Relation File Format (ARFF) files in R?
4. How to read a heavy csv/tsv file using readr package?

2. Solution

1. Import SAS XPORT Files into R With The foreign package

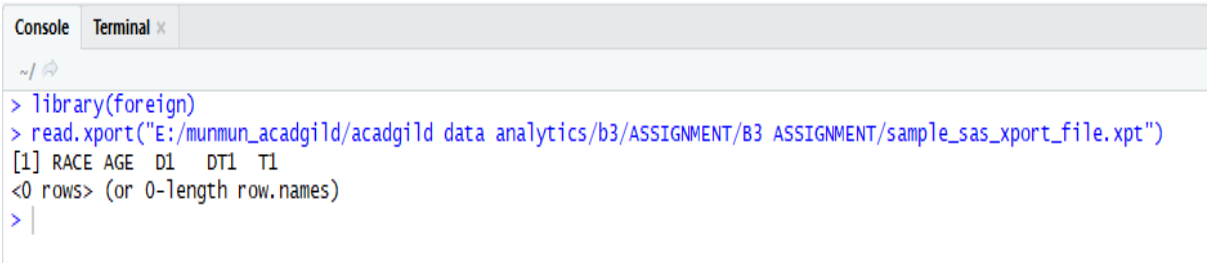
The foreign package is installed using the command: install.packages("foreign")

read.xport (file) command reads a file as a SAS XPORT format library and returns a list of data.frames ;where,file is character variable with the name of the file to read. The file must be in SAS XPORT format.

The R-script for importing SAS XPORT Files: sample_sas_xport_file.xpt into R is as follows:

```
library (foreign)
read.xport("E:/munmun_acadgild/acadgild data analytics/b3/ASSIGNMENT/B3
ASSIGNMENT/sample_sas_xport_file.xpt")
```

The output of the R-script is:



2. Import SAS Files into R With The haven Package

The haven package is installed using the command: install.packages("haven")
The haven package is used to `Import and Export 'SPSS', 'Stata' and 'SAS' Files`

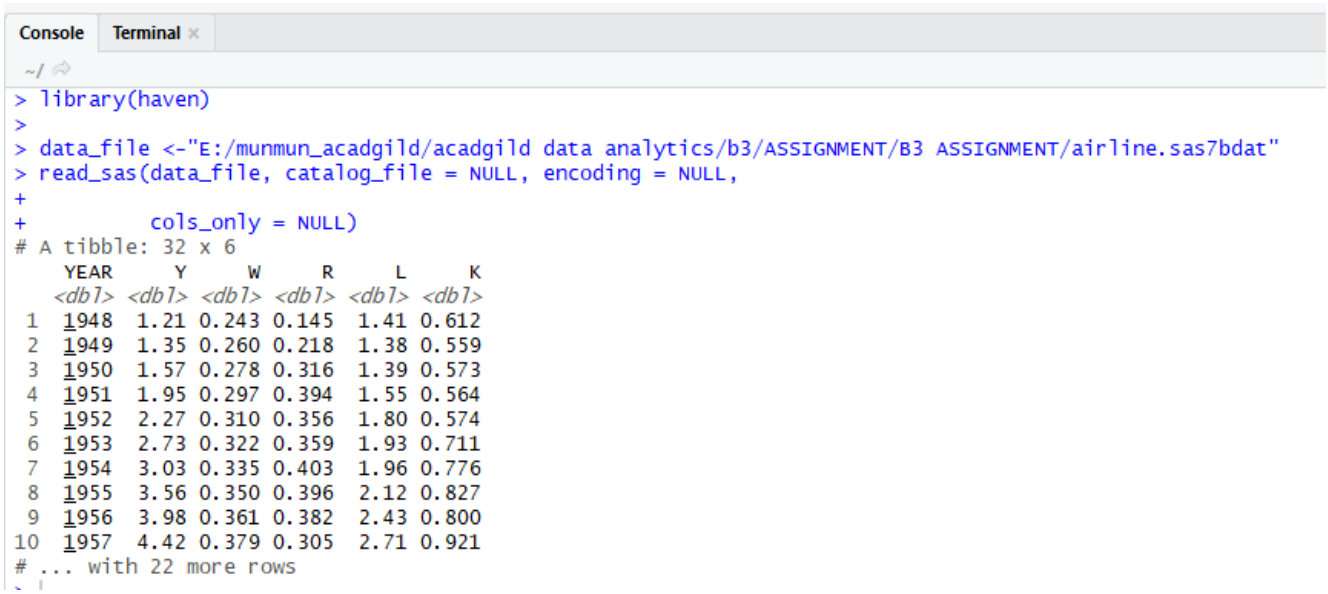
read_sas (file) function supports both sas7bdat files and the accompanying sas7bdat files that SAS uses to record value labels. read_sas (file) command reads a file airline.sas7bdat file.

The R-script for importing SAS Files into R With The haven Package is as follows:

```
library(haven)
data_file <-"E:/munmun_acadgild/acadgild data analytics/b3/ASSIGNMENT/B3
ASSIGNMENT/airline.sas7bdat"
read_sas(data_file, catalog_file = NULL, encoding = NULL,

      cols_only = NULL)
```

The output of the R-script is:



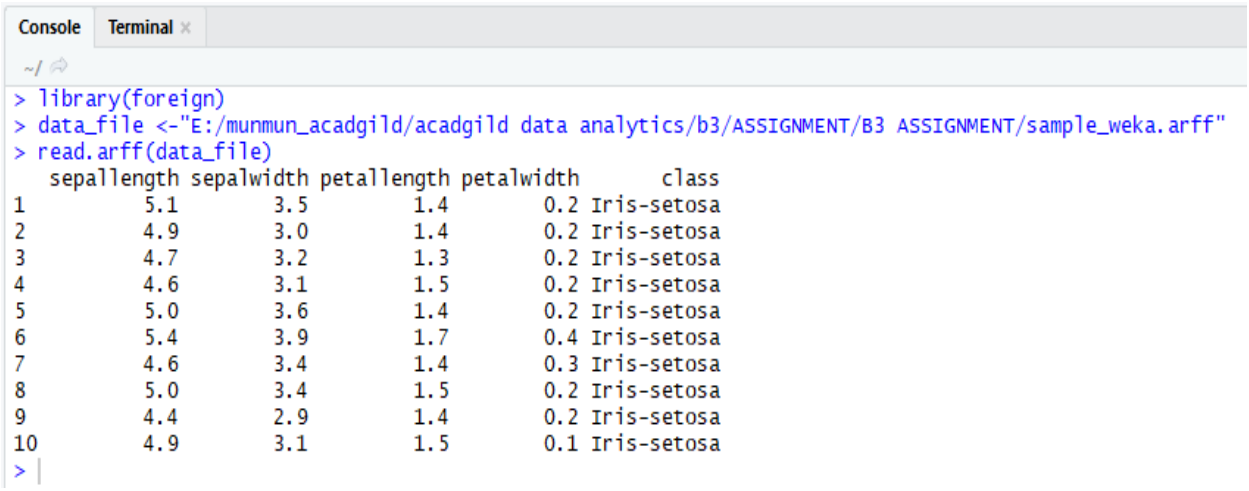
3. Read Weka Attribute-Relation File Format (ARFF) files in R

read.arff(file) reads data from Weka Attribute-Relation File Format (ARFF) files**:** sample_weka.arff file**.**

The script for reading Attribute-Relation File Format (ARFF) files in R is as follows:

```
library(foreign)
data_file<-"E:/munmun_acadgild/acadgild data analytics/b3/ASSIGNMENT/B3
ASSIGNMENT/sample_weka.arff"
read.arff(data_file)
```

The output of the R-script is:



4. Read a heavy csv/tsv file using readr package

The 'readr' package provides a fast way to read rectangular data (like 'csv', 'tsv', and 'fwf'). The `read_csv()` and `read_tsv()` are special cases of the general `read_delim()`. They're useful for reading the most common types of flat file data, comma separated values and tab separated values, respectively.

a. R-script to read a heavy csv file using readr package:

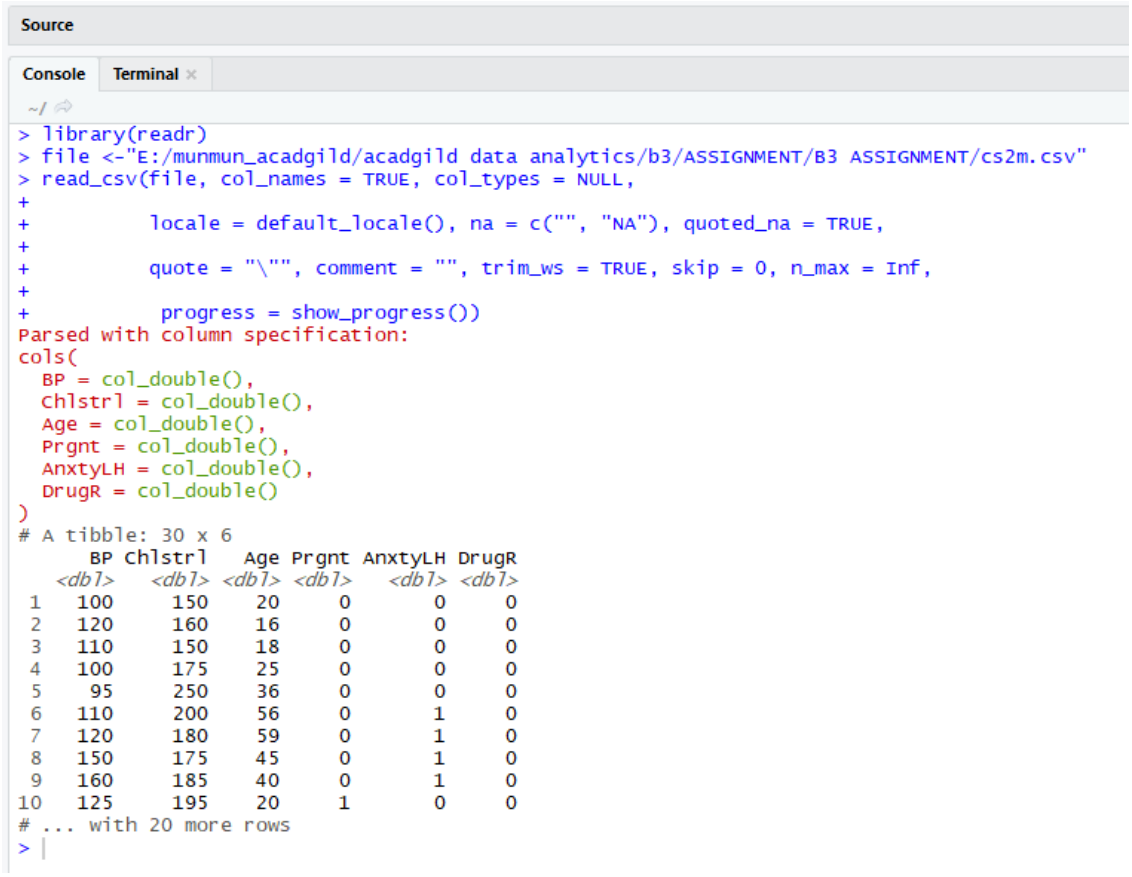
```
library(readr)
file <-"E:/munmun_acadgild/acadgild data analytics/b3/ASSIGNMENT/B3
ASSIGNMENT/cs2m.csv"
read_csv(file, col_names = TRUE, col_types = NULL,

        locale = default_locale(), na = c("", "NA"), quoted_na = TRUE,

        quote = "\"", comment = "", trim_ws = TRUE, skip = 0, n_max = Inf,

         progress = show_progress())
```

The output of the R-script is:



b. R-script to read a heavy tsv file using readr package

```
library(readr)
file <-"E:/munmun_acadgild/acadgild data analytics/b3/ASSIGNMENT/B3 ASSIGNMENT/sample_tsv_file.tsv"
read_tsv(file, col_names = TRUE, col_types = NULL,

      locale = default_locale(), na = c("", "NA"), quoted_na = TRUE,

      quote = "\"", comment = "", trim_ws = TRUE, skip = 0, n_max = Inf,

     , progress = show_progress())
```

The output of the R-script is: