ACADGILD

SESSION 2: INTRODUCTION to working with R

Assignment 3

Table of Contents

1. Introduction

This assignment will help you understand the concepts learnt in the session.

2. Objective

This assignment will test your skills on the basics of R.

3. Prerequisites

Not applicable.

4. Associated Data Files

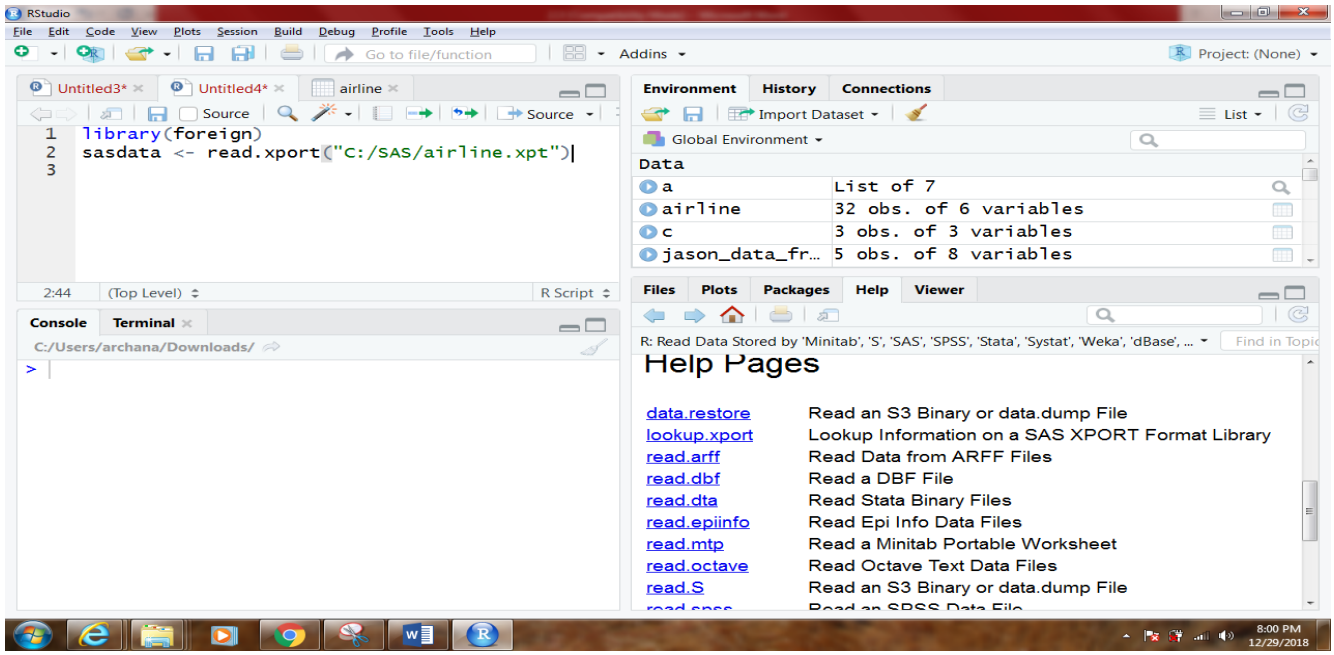
Not applicable.

5. Problem Statement

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

Ans-

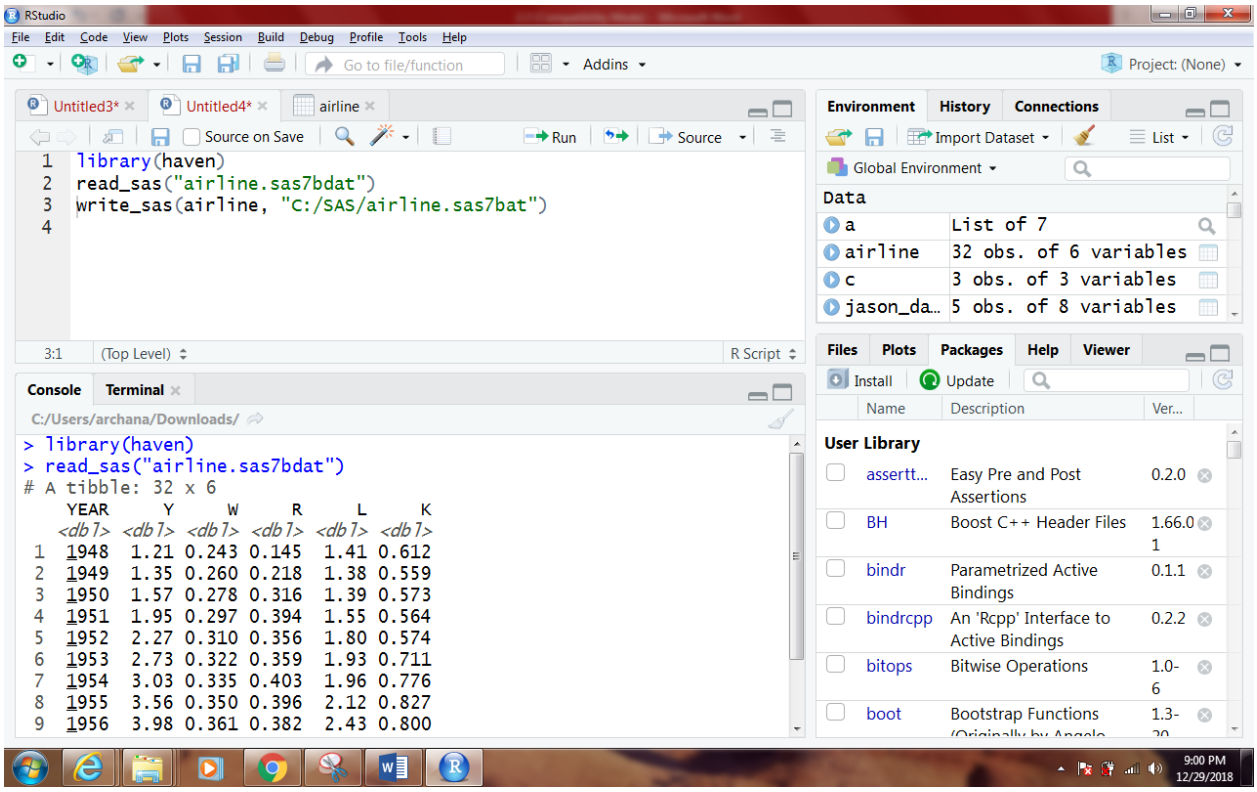
To import SAS data files in R, first install and load the package `foreign`. The functions `read.xport()` of the package `foreign` import SAS Transport data file



2. How To Import SAS Files into R With The haven Package?

- Install it,
- install.packages("haven"),
- Load it
- library(haven),
- Then pick the appropriate read function:
    - SAS: read_sas()

Output is shown in below screen shot: sample sas file is airline



1. How to read Weka Attribute-Relation File Format (ARFF) files in R?

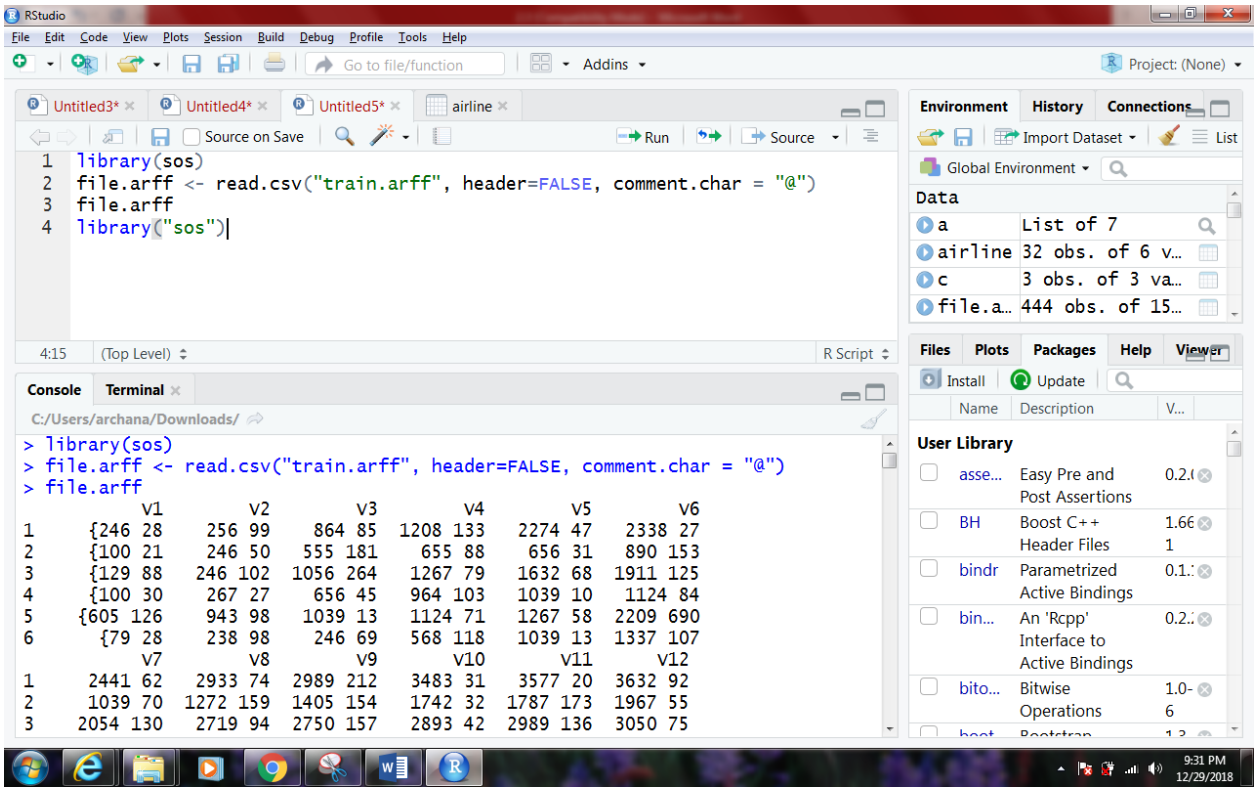
Ans:

This is done by installing package “sos”

```
install.packages("sos")

read.csv("data.arff", header=FALSE, comment.char = "@")
```

Below is screen shot. File name is train.arff

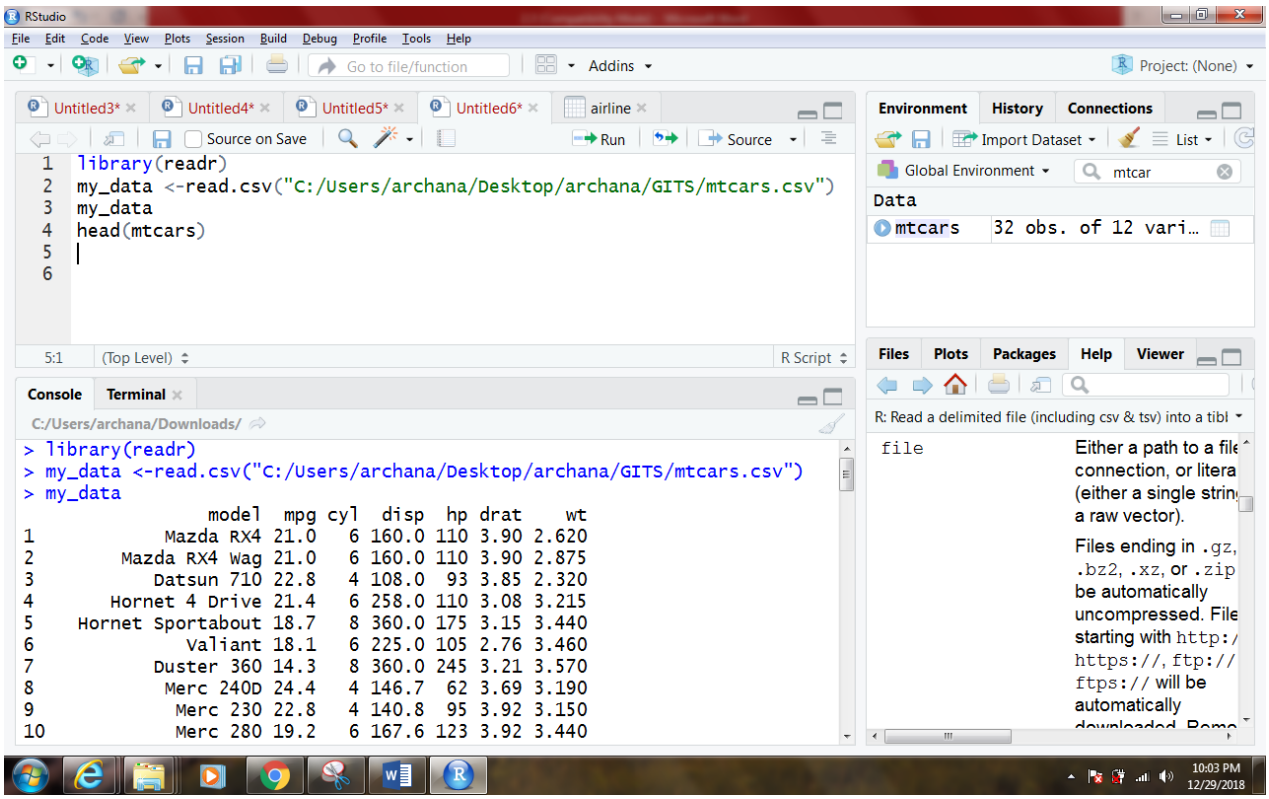


2. How to read a heavy csv/tsv file using readr package?

```
# Installing
install.packages("readr")
# Loading
library("readr")
```

```
read_delim(file, delim, col_names = TRUE)
 # Read comma (",") separated values
 read_csv(file, col_names = TRUE)
# Read semicolon (";") separated values
# (this is common in European countries)
 read_csv2(file, col_names = TRUE)
 # Read tab separated values
 read_tsv(file, col_names = TRUE)
```

```
Read a tsv file
my_data <- read_tsv(file.choose())
# Read a csv file
my_data <- read_csv(file.choose())
```



6. Expected Format

1. R file should be submitted where applicable.
2. R file should be in PDF or in .r format
3. Proper screenshots of the outputs should be submitted as well
4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

7. Approximate Time to Complete Task

20 mins.