

SESSION 2: INTRODUCTION to working with R

Assignment 3

5. Problem Statement

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

Ans. The `read.xport` function reads SAS xport formatted files, augmenting the functionality of the `read.xport` function provided in the 'foreign' package with additional features, borrowed from `sasxport.get` in Frank Harrell's 'Hmisc' package. Namely, variables are properly coerced into the types specified by the format field. All standard numeric and string SAS formats are supported automatically, while user-defined formats are supported when the user has included the appropriate format data in the xport file via:

```
PROC FORMAT CNTLOUT=format;
```

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

Ans. Haven enables R to read and write various data formats used by other statistical packages by wrapping the fantastic ReadStatC library written by Evan Miller. Haven is part of the tidyverse. Currently it supports:

SAS: `read_sas()` reads `.sas7bdat` + `.sas7bcat` files and `read_xpt()` reads SAS transport files (version 5 and version 8). `write_sas()` writes `.sas7bdat` files.

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

Ans. `read.arff` in package `foreign` reads data from Weka Attribute-Relation File Format (ARFF) files.

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

Ans. `readr` makes it easy to read many types of rectangular data, including csv, tsv and fixed width files. Compared to base equivalents like `read.csv()`, `readr` is much faster and gives more convenient output: it never converts strings to factors, can parse date/times, and it doesn't munge the column names. Install the latest version with: