ASP 460 2.0 Data Visualisation

2020 - Semester 1: Final Project

Due date: 5 October 2020

Problem description

Food industry is one of the fastest moving industrial sectors. The glamorous and glittering retail shops and supermarkets are expanding very fast all over the country. The majority of food is pre-packed and presented to the consumer in a labelled container. The data for the project is based on a study attempts to evaluate the consumers' attitude towards food labels and awareness of information printed on food labels.

The dataset is available at: https://thiyanga.netlify.app/project/datasets/

Use the following code to read the dataset into R

```
library(readr)
url <- "https://thiyanga.netlify.app/project/datasets/foodlabel.csv"
foodlabel <- read_csv(url)
head(foodlabel, 2)
```

```
# A tibble: 2 x 80
  Gender   Age Education Employment Income Housesize children marital fshopper
   <dbl> <dbl>     <dbl>      <dbl>  <dbl>     <dbl>    <dbl>   <dbl>    <dbl>
1      1    22         5          4      3         5        2       0        0
2      1    61         5          6      3         5        1       1        1
# ... with 71 more variables: mplanner <dbl>, place <dbl>, FA <dbl>,
#   Diabetes <dbl>, `Metabolic cyndrents` <dbl>, Other <dbl>, specific <dbl>,
#   job1 <dbl>, job2 <dbl>, Exercise <dbl>, Health <dbl>, taste <dbl>,
#   easy <dbl>, familiarity <dbl>, friends <dbl>, Useful <dbl>, Easiness <dbl>,
#   Sufficient <dbl>, Trusfulness <dbl>, Clear <dbl>, `attractive pack` <dbl>,
#   `hc/nutriclaims` <dbl>, graphical <dbl>, `Free/prize` <dbl>, source <dbl>,
#   netquan <dbl>, `low in fat` <dbl>, `low in cho` <dbl>, sodium <dbl>, `e
#   labels` <dbl>, place2 <dbl>, fa2 <dbl>, Health_1 <dbl>, X43 <dbl>,
#   f1 <dbl>, f2 <dbl>, f3 <dbl>, f4 <dbl>, f5 <dbl>, f6 <dbl>, f7 <dbl>,
#   f8 <dbl>, f9 <dbl>, f10 <dbl>, f11 <dbl>, f12 <dbl>, f13 <dbl>, f14 <dbl>,
#   f15 <dbl>, f16 <dbl>, f17 <dbl>, f18 <dbl>, i1 <dbl>, i2 <dbl>, i3 <dbl>,
#   i4 <dbl>, i5 <dbl>, i6 <dbl>, i7 <dbl>, i8 <dbl>, i9 <dbl>, i10 <dbl>,
#   i11 <dbl>, i12 <dbl>, i13 <dbl>, i14 <dbl>, i15 <dbl>, i16 <dbl>,
#   i17 <dbl>, i18 <dbl>, cluster <dbl>
```

The questionnaire used to collect data are given in the pdf document `food2queslabelled.pdf`. The column names are mapped with corresponding question in the questionnaire in red hand-written text (only the columns relevant to your final projects).

Perform an exploratory data analysis using R programming to help the researcher to answer the following objectives.

1. To assess the association between demographic, socio-economic characteristics and health related factors on consumers attitude towards food labels.
2. To assess the association between demographic, socio-economic characteristics and health related factors on awareness of information printed in food labels.

Create a (reproducible) report using Rmarkdown. Interpret all of the graphs in your report.

Create a flex-dashboard including the most important plots.

Email your .rmd files and final outputs to ttalagala@sjp.ac.lk

Should you need any further information, please do not hesitate to contact me.