1. (40 pts) The file `segment.csv` has data from a survey to 300 customers and potential customers from a company offering a subscription service. It includes the age, gender, income, number of children, whether they own or rent their homes, and whether they currently subscribe to the service or not. It is of interest to know whether clients are just as likely to subscribe or not without regard to home ownership.

   a) Construct a single two-way crosstab table to count the number of individuals that are subscribers (or not) and are home owners (or not).
   b) Test if the proportion of subscribers is the same between home owners and home renters.

2. The file `brands.csv` has ratings from some perceptual attributes on a set of brands labeled as `a,b,...,j`. The data comes from a survey to 100 customers. The attributes are as follows.

Perceptual attribute	Example
perform	Brand has strong performance
leader	Brand is a leader in the field
latest	Brand has the latest products
fun	Brand is fun
serious	Brand is serious
bargain	Brand products are a bargain
value	Brand products are a good value
trendy	Brand is trendy
rebuy	I would buy from Brand again

   a) (10 pts) Find the average rating of each brand on each attribute and store it in dataframe `df1`. Let column `brand` be the rownames of `df1`

```
rownames(df1) = df1$brand
df1$brand = NULL
```

   b) (10 pts) Display a heatmap using the average ratings from `df1` using the following commands

```
library(gplots)
library(RColorBrewer)
heatmap.2(as.matrix(df1),col=brewer.pal(9, "GnBu"),
trace="none", key=FALSE, dend="none",main="")
```

   What brands are highly rated on attributes *leader*, and *serious*?
   c) (10 pts) Scale the brands data to find principal components (call `prcomp1` the resulting object). How many principal components explain at least 80% of the variation in the customer's brand ratings? Display a lineplot of the Cumulative PVEs.
   d) (10 pts) Construct a biplot from `prcomp1`
   e) (10 pts) Find principal components from `df1`. Use `prcomp2 = prcomp(df1,scale=T)`.
   f) (10 pts) Construct a biplot from `prcomp2`. This is called a perceptual map of the brands. It helps answer the question *What is the average position of the brand on each attribute?* What brands are highly rated on attributes *leader*, and *serious*? Which are highly rated on *bargain*, and *value*?

Submit your report (code and output) as a pdf file onto Blackboard (no screen captures).