Assignment 2.2

Sathish Manthani

6/27/2020

### Packages

Import the required packages/libraries to anlayze the dataset

library(ggplot2) # plots for visualization  
library(dplyr) # for data manipulation  
library(corrplot) # for display of correlation matrix  
library(GGally) # for gallary display of charts  
library(gridExtra) # to view plots in grid  
library(purrr) # for map\_dbl function

### Read the data file

df <- read.csv("expenditures.txt", sep='\t', stringsAsFactors = FALSE)

library(sqldf)

## Loading required package: gsubfn

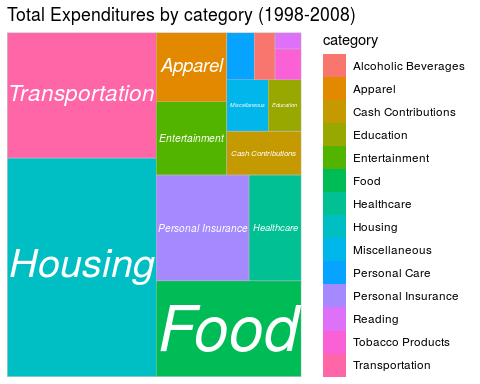
## Loading required package: proto

## Warning in fun(libname, pkgname): couldn't connect to display ":0"

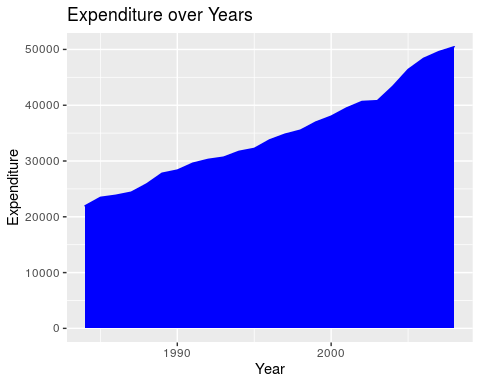
## Loading required package: RSQLite

cat.df <- sqldf("select   
 category  
 ,sum(expenditure) as expenditure   
 from df   
 group by category")

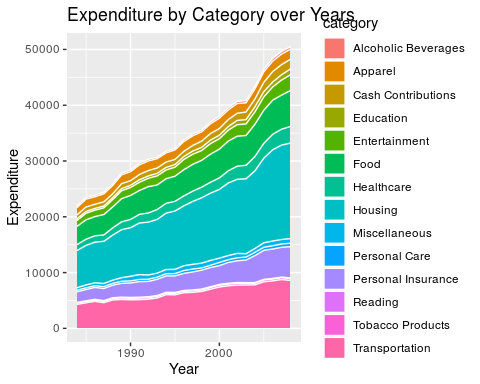
library(treemapify)  
  
ggplot2::ggplot(cat.df, ggplot2::aes(area=expenditure, fill=category, label = category)) + geom\_treemap() +  
 geom\_treemap\_text(fontface='italic',color='white',place='centre',grow=TRUE) + ggplot2::labs(title = "Total Expenditures by category (1998-2008)")



yearexp.df <- sqldf("select   
 year  
 ,sum(expenditure) as expenditure   
 from df   
 group by year")  
  
ggplot(yearexp.df, aes(x=year, y=expenditure)) +  
 geom\_area(color='blue',fill='blue') + labs(title='Expenditure over Years', x='Year', y='Expenditure')



ggplot(df, aes(x=year, y=expenditure, fill=category)) +  
 geom\_area(color='white') + labs(title='Expenditure by Category over Years', x='Year', y='Expenditure')



ggplot(yearexp.df, aes(x=year, y=expenditure)) +  
 geom\_step() + labs(title='Expenditure over Years', x='Year', y='Expenditure')

