````
---
title: "Assignment 11"
author: "Sathish Prasad V T, sp702"
date: "`r Sys.Date()`"
output:
  pdf_document: default
  html_document: default
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

## Instructions

*This is an **individual** assignment.*

The objective of this assignment is to give you experience composing programs in R with an emphasis on using Tidyverse packages.

Please meet the specific requirements exactly. Do not change or remove any of the instructions; just add you name and NetID in the author metadata area above and code in the code cells as indicated (you may edit/remove the "TODO" comments).  Please make sure to test your programs before submitting your solutions on Canvas!

## Problem 1: Marketing Campaign Analysis

This problem builds off of Problem 1 from Assignment 10 and uses the same data (`marketing_campaigns.csv`).  You may want to start with reusing some of your code from that assignment!

First, filter the campaigns to only consider campaigns with a daily average CTR > 0.02.

Then group the data into monthly data using the `lubridate` method `floor_date` so that you have, for each campaign, the monthly number of clicks, number of Impressions, and CTR.

Then, use `pivot_longer` to turn the data into a "wide" format with one column for each month and three rows for each campaign (clicks, impressions, and CTR).

Finally, write out the resulting data to `campaigns_filtered.csv`; the first four lines of this file should appear as follows.

```
Campaign,Metric,2023-03-01,2023-04-01,2023-05-01,2023-06-01,2023-07-01
Campaign B,Clicks,NA,1583,1713,NA,NA
Campaign B,CTR,NA,0.0485374379101,0.0504982017569719,NA,NA
Campaign B,Impressions,NA,32614,33922,NA,NA
```

```{r problem1}
# TODO: Insert your code here.
library(readr)
library(tidyverse)
marketing_data = read_delim('marketing_campaigns.csv',show_col_types = F)
marketing_data = marketing_data %>% mutate(CTR=Clicks /Impressions)
filtered_df <- marketing_data %>%
  group_by(Campaign) %>%
  mutate(avg_CTR = mean(CTR)) %>%
  filter(avg_CTR > 0.02)
View(filtered_df)
filtered_df$Date = floor_date(filtered_df$Date, unit = "month")

monthly_summary = filtered_df %>%
````
````

```
  group_by(Campaign, Date) %>%
  summarize(
    Clicks = round(sum(Clicks)),
    Impressions = round(sum(Impressions)),
    CTR = Clicks/Impressions,
    .groups = "drop"
  )

monthly_summary = monthly_summary %>%
  pivot_longer(cols = c(Clicks,CTR,Impressions),
               names_to = "Metric",
               values_to = "Value")

monthly_summary = monthly_summary %>%
  pivot_wider(names_from = Date, values_from = Value)

write_csv(monthly_summary,'campaigns_filtered.csv',col_names = T)
```


## Problem 2: Manipulating Product Data

This problem builds off of Problem 2 from Assignment 10 and uses the same data (`products.csv`).  You may want to reuse some of your code from that assignment!

Your manager has decided that the product categories can be simplified to the following.

```
Bags (634 items)
Belts (164 items)
Shoes (696 items)
Jewelry (222 items)
Hats (181 items)
Keychains (18 items)
Wallets (190 items)
Neckties (106 items)
Outerwear (16 items)
Sunglasses (377 items)
Suspenders (20 items)
Umbrellas (25 items)
Watches (351 items)
```

Use `fct_collapse` to assign each product to its new category label. Yes, you do need to figure out which old categories belong to which of the simplified categories.  You are free to discuss this with your peers if needed, but try to figure it out on your own if you can!  The method `fct_count` may also come in handy to check your work!

With these new categories, recompute the average margins per category and write out the results to `avg_margin_by_new_category.csv`; the first five lines of this file should look as follows.

```
Simplified.Category,avg_Margin
Bags,0.165022591140384
Belts,0.165824075527352
Shoes,0.159532640824002
Jewelry,0.164746143437297
```

```{r problem2}
# TODO: Insert your code here.
library(forcats)
data = read_delim('products.csv',show_col_types = F)
```

```
data = data %>% mutate(Margin=(Retail.Price-Unit.Cost)/Retail.Price)
data = data %>% separate(Item, into =c('Item','Category'),sep ='- ')
unique_items <- unique(data$Category)

data$Simplified.Category <- fct_collapse(
  data$Category,
  Bags = c("Crossbody Bags", "Travel Duffels", "Clutches", "Kids' Backpacks", "Drawstring
Bags","Sports Duffels", "Messenger Bags", "Backpacks", "Waist Packs","Luggage","Shoulder
Bags","Casual Daypacks","Totes"),
  Belts = c("Belts","Belt Buckles"),
  Shoes = c("Platforms & Wedges", "Flats", "Pumps", "Fashion Sneakers", "Heeled Sandals",
"Oxfords", "Sandals", "Boots","Slippers","Shoelaces"),
  Jewelry = c("Necklaces", "Brooches & Pins", "Jewelry"),
  Hats = c("Visors", "Hats & Caps", "Sun Hats"),
  Keychains = c("Keyrings & Keychains"),
  Wallets = c("Wallets","Money Clips"),
  Neckties = "Neckties",
  Outerwear = "Outerwear",
  Sunglasses = "Sunglasses",
  Suspenders = "Suspenders",
  Umbrellas = "Umbrellas",
  Watches = "Wrist Watches"
)

data = data %>%
  group_by(Simplified.Category) %>%
  summarize(avg_margin = mean(Margin))

write_csv(data,"avg_margin_by_new_category.csv")
```