Quantium Virtual Internship - Retail Strategy and Analytics - Task 1

Solution template for Task 1

This file is a solution template for the Task 1 of the Quantium Virtual Internship. It will walk you through the analysis, providing the scaffolding for your solution with gaps left for you to fill in yourself.

Look for comments that say "over to you" for places where you need to add your own code! Often, there will be hints about what to do or what function to use in the text leading up to a code block - if you need a bit of extra help on how to use a function, the internet has many excellent resources on R coding, which you can find using your favourite search engine. ## Load required libraries and datasets Note that you will need to install these libraries if you have never used these before.

```
#### Example code to install packages
#install.packages("data.table")
#### Load required libraries
library("data.table")
library(ggplot2)
library(ggmosaic)
library(readr)
#### Point the filePath to where you have downloaded the datasets to and
#### assign the data files to data.tables
# over to you! fill in the path to your working directory. If you are on a
 ↪  Windows machine, you will need to use forward slashes (/) instead of
 ↪  backshashes (\)
filePath <- ""
transactionData <- fread(paste0(filePath,"C:/Users/Luke
 ↪  Villanueva/source/repos/LukVill/Quantium
 ↪  Program/QVI_transaction_data.csv"))
```

```
## Taking input= as a system command ('C:/Users/Luke Villanueva/source/repos/LukVill/Quantium P

## Warning in (if (.Platform$OS.type == "unix")
## system else shell)(paste0("(", : '(C:/Users/Luke
## Villanueva/source/repos/LukVill/Quantium Program/QVI_transaction_data.csv)
## > C:\Users\LUKEVI~1\AppData\Local\Temp\Rtmpw5AwMV\file531c6841aac' execution
## failed with error code 1

## Warning in fread(paste0(filePath, "C:/Users/Luke
## Villanueva/source/repos/LukVill/Quantium Program/QVI_transaction_data.csv")):
## File 'C:\Users\LUKEVI~1\AppData\Local\Temp\Rtmpw5AwMV\file531c6841aac' has size
## 0. Returning a NULL data.table.
```

```
customerData <- fread(paste0(filePath,"C:/Users/Luke
 ↪  Villanueva/source/repos/LukVill/Quantium
 ↪  Program/QVI_purchase_behaviour.csv"))
```

Exploratory data analysis

The first step in any analysis is to first understand the data. Let's take a look at each of the datasets provided. ### Examining transaction data We can use `str()` to look at the format of each column and see a sample of the data. As we have read in the dataset as a `data.table` object, we can also run `transactionData` in the console to see a sample of the data or use `head(transactionData)` to look at the first 10 rows. Let's check if columns we would expect to be numeric are in numeric form and date columns are in date format.

```
#### Examine transaction data
```