COMP824 Assignment
Semester 1 2023

Minh Hoang Quach
ID Number: 20107785

06 April 2023

Q1 Load & extract the data

a) Download the RMD file

No answer required

b) Download the csv file

```
library(tidyverse)
sales_all <- read_csv("COMP824_sales_data.csv")
```

c) Load data and filter

```
sales_for_myid <- sales_all %>% filter(sku_id == 320485)
print(head(sales_for_myid, n=5))
```

```
## # A tibble: 5 x 15
##   record_ID week       store_id sku_id total_p~1 base_~2 is_fe~3 is_di~4 units~5
##       <dbl> <date>        <dbl>  <dbl>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1        31 2011-01-17     8095 320485      206.    206.       0       0      28
## 2        56 2011-01-17     8094 320485      177.    177.       0       0      27
## 3        80 2011-01-17     8063 320485      177.    177.       0       0       9
## 4       107 2011-01-17     8023 320485      206.    206.       0       0      60
## 5       132 2011-01-17     8058 320485      206.    206.       0       0      22
## # ... with 6 more variables: year <dbl>, month <dbl>, day <dbl>, weekday <chr>,
## #   start_of_month <date>, end_of_month <date>, and abbreviated variable names
## #   1: total_price, 2: base_price, 3: is_featured_sku, 4: is_display_sku,
## #   5: units_sold
```

Q2 Explore the sales data

a) Date range

```
date_range = range(as.Date(sales_for_myid$week))
print(date_range)
```

```
## [1] "2011-01-17" "2013-07-09"
```

The date range of the sales data (specifically with the SKU id is 320485) is from 17/02/2011 to 09/07/2013

b) Number of stores

c) Total price

d) Discussion about total price

Q3 Analysis of monthly sales data

a) Compute monthly sales

b) Plot of monthly sales

c) Discussion about monthly sales

Q4 Analysis of store performance

a) Compute total sales per store

b) Plot of total sales per store

c) Additional performance metric (rename this section)

d) Plot of additional performance metric (rename this section)

e) Discussion of store performance

A Appendix: R Environment

```
format(Sys.time(), '%d %B %Y')
```

```
## [1] "06 April 2023"
```

```
sessionInfo()
```

```
## R version 4.2.2 (2022-10-31)
## Platform: aarch64-apple-darwin20 (64-bit)
## Running under: macOS Ventura 13.2.1
##
## Matrix products: default
## BLAS:   /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/4.2-arm64/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
##  [1] lubridate_1.9.2 forcats_1.0.0   stringr_1.5.0   dplyr_1.1.1
##  [5] purrr_1.0.1     readr_2.1.4     tidyr_1.3.0     tibble_3.2.0
##  [9] ggplot2_3.4.1   tidyverse_2.0.0 knitr_1.42
##
## loaded via a namespace (and not attached):
##  [1] pillar_1.8.1     compiler_4.2.2   tools_4.2.2      bit_4.0.5
##  [5] digest_0.6.31    timechange_0.2.0 evaluate_0.20    lifecycle_1.0.3
##  [9] gtable_0.3.3     pkgconfig_2.0.3  rlang_1.1.0      cli_3.6.0
## [13] rstudioapi_0.14  parallel_4.2.2   yaml_2.3.7       xfun_0.37
## [17] fastmap_1.1.1    withr_2.5.0      generics_0.1.3   vctrs_0.6.1
## [21] hms_1.1.2        bit64_4.0.5      grid_4.2.2       tidyselect_1.2.0
## [25] glue_1.6.2       R6_2.5.1         fansi_1.0.4      vroom_1.6.1
## [29] rmarkdown_2.20   tzdb_0.3.0       magrittr_2.0.3   scales_1.2.1
## [33] htmltools_0.5.4  ellipsis_0.3.2   colorspace_2.1-0 utf8_1.2.3
## [37] stringi_1.7.12   munsell_0.5.0    crayon_1.5.2
```