ps3_solution.R

lachenar

2025-02-18

```
setwd("C:/Users/lachenar/OneDrive - Colostate/Documents/GitProjectsWithR/csu-arec-330.github.io/material

### Load necessary libraries
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.4.2
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(readr)
```

```
## Warning: package 'readr' was built under R version 4.4.2
```

```
# 1. Read in the dataset
supermarket_raw <- read_csv("https://csu-arec-330.github.io/materials/unit_00/inputs/supermarket_sales.c
```

```
## Rows: 1000 Columns: 17
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (8): invoice_id, branch, city, customer_type, gender, product_line, dat...
## dbl  (8): unit_price, quantity, tax, total, cogs, gross_margin, gross_income...
## time (1): time
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
# 2. Calculate total value of sale and verify tax
super_sales <- supermarket_raw %>%
  mutate(
    subtotal = unit_price * quantity,
    tax_verify = subtotal * 0.05
  )

# 3. Filter dataset for "Food and beverages"
food_sales <- super_sales %>%
  filter(product_line == "Food and beverages")
```

```r
# 4. Select relevant columns
food_sales_selected <- food_sales %>%
  select(city, product_line, unit_price, quantity, total, rating)

# 5. Sort by quantity in descending order
food_sales_sorted <- food_sales_selected %>%
  arrange(desc(quantity))

# 6. Calculate median sales by payment type
median_sales <- super_sales %>%
  group_by(payment_type) %>%
  summarise(median_total = median(total, na.rm = TRUE))

# 7. Compute rating per unit price (rup) and explain its insights
super_sales_rup <- super_sales %>%
  mutate(rup = rating / unit_price)

# 8. Summarize rup and unit_price by product line
summary_stats <- super_sales_rup %>%
  group_by(product_line) %>%
  summarise(
    mean_rup = mean(rup, na.rm = TRUE),
    mean_unit_price = mean(unit_price, na.rm = TRUE)
  )

# 9. Print summarized dataframe
print(summary_stats)
```

```
## # A tibble: 6 x 3
##   product_line           mean_rup mean_unit_price
##   <chr>                     <dbl>           <dbl>
## 1 Electronic accessories    0.188            53.6
## 2 Fashion accessories       0.170            57.2
## 3 Food and beverages        0.176            56.0
## 4 Health and beauty         0.182            54.9
## 5 Home and lifestyle        0.175            55.3
## 6 Sports and travel         0.170            57.0
```

```r
# 10. Display version and loaded packages
version
```

```
##                _
## platform       x86_64-w64-mingw32
## arch           x86_64
## os             mingw32
## crt            ucrt
## system         x86_64, mingw32
## status
## major          4
## minor          4.0
## year           2024
## month          04
## day            24
## svn rev        86474
## language       R
```

```
## version.string R version 4.4.0 (2024-04-24 ucrt)
## nickname       Puppy Cup
```

```
print(.packages())
```

```
## [1] "readr"     "dplyr"     "stats"     "graphics"  "grDevices" "utils"
## [7] "datasets"  "methods"   "base"
```