Transformation Exercise Key

Carolina Alves de Lima Salge

2/15/2021

Transformation

Below is the code from the slides

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.6     v dplyr   1.0.4
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
orderList <- read_csv("List of Orders.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   `Order ID` = col_character(),
##   `Order Date` = col_character(),
##   CustomerName = col_character(),
##   State = col_character(),
##   City = col_character()
## )
```

```
orderDetails <- read_csv("Order Details.csv")
```

```
##
## -- Column specification --------------------------------------------------------
## cols(
##   `Order ID` = col_character(),
##   Amount = col_double(),
##   Profit = col_double(),
##   Quantity = col_double(),
##   Category = col_character(),
##   `Sub-Category` = col_character()
## )
```

```
salesTarget <- read_csv("Sales target.csv")
```

```
##
## -- Column specification --------------------------------------------------
## cols(
##   `Month of Order Date` = col_character(),
##   Category = col_character(),
##   Target = col_double()
## )
```

```
orderDetails %>%
    filter(., Category == "Furniture")
```

```
## # A tibble: 243 x 6
##    `Order ID` Amount Profit Quantity Category  `Sub-Category`
##    <chr>       <dbl>  <dbl>    <dbl> <chr>     <chr>
##  1 B-25601      1275  -1148        7 Furniture Bookcases
##  2 B-25603        24    -30        1 Furniture Chairs
##  3 B-25608      1364  -1864        5 Furniture Tables
##  4 B-25608       476      0        3 Furniture Chairs
##  5 B-25610        30     -5        2 Furniture Furnishings
##  6 B-25612       259    -55        2 Furniture Chairs
##  7 B-25614       494     54        4 Furniture Bookcases
##  8 B-25618       362    127        1 Furniture Bookcases
##  9 B-25626      1103   -276        3 Furniture Chairs
## 10 B-25628        35     -8        2 Furniture Furnishings
## # ... with 233 more rows
```

```
orderDetails %>%
    filter(., Category == "Furniture", Quantity > 1)
```

```
## # A tibble: 223 x 6
##    `Order ID` Amount Profit Quantity Category  `Sub-Category`
##    <chr>       <dbl>  <dbl>    <dbl> <chr>     <chr>
##  1 B-25601      1275  -1148        7 Furniture Bookcases
##  2 B-25608      1364  -1864        5 Furniture Tables
##  3 B-25608       476      0        3 Furniture Chairs
##  4 B-25610        30     -5        2 Furniture Furnishings
##  5 B-25612       259    -55        2 Furniture Chairs
##  6 B-25614       494     54        4 Furniture Bookcases
##  7 B-25626      1103   -276        3 Furniture Chairs
##  8 B-25628        35     -8        2 Furniture Furnishings
##  9 B-25631        89    -89        2 Furniture Furnishings
## 10 B-25634       389    -83        3 Furniture Chairs
## # ... with 213 more rows
```

```
orderDetails %>%
    filter(., Category == "Furniture" & Quantity > 1)
```

```
## # A tibble: 223 x 6
##    `Order ID` Amount Profit Quantity Category  `Sub-Category`
```

```
##      <chr>         <dbl>  <dbl>    <dbl> <chr>      <chr>
##  1 B-25601        1275   -1148       7 Furniture Bookcases
##  2 B-25608        1364   -1864       5 Furniture Tables
##  3 B-25608         476      0       3 Furniture Chairs
##  4 B-25610          30     -5       2 Furniture Furnishings
##  5 B-25612         259    -55       2 Furniture Chairs
##  6 B-25614         494     54       4 Furniture Bookcases
##  7 B-25626        1103   -276       3 Furniture Chairs
##  8 B-25628          35     -8       2 Furniture Furnishings
##  9 B-25631          89    -89       2 Furniture Furnishings
## 10 B-25634         389    -83       3 Furniture Chairs
## # ... with 213 more rows
```

```r
orderDetails %>%
    filter(., Category == "Furniture" | Quantity > 1)
```

```
## # A tibble: 1,388 x 6
##     `Order ID` Amount Profit Quantity Category    `Sub-Category`
##     <chr>       <dbl>  <dbl>    <dbl> <chr>       <chr>
##  1 B-25601      1275  -1148       7 Furniture   Bookcases
##  2 B-25601        66    -12       5 Clothing    Stole
##  3 B-25601         8     -2       3 Clothing    Hankerchief
##  4 B-25601        80    -56       4 Electronics Electronic Games
##  5 B-25602       168   -111       2 Electronics Phones
##  6 B-25602       424   -272       5 Electronics Phones
##  7 B-25602      2617   1151       4 Electronics Phones
##  8 B-25602       561    212       3 Clothing    Saree
##  9 B-25602       119     -5       8 Clothing    Saree
## 10 B-25603      1355    -60       5 Clothing    Trousers
## # ... with 1,378 more rows
```

```r
orderDetails %>%
    arrange(., desc(Profit))
```

```
## # A tibble: 1,500 x 6
##     `Order ID` Amount Profit Quantity Category    `Sub-Category`
##     <chr>       <dbl>  <dbl>    <dbl> <chr>       <chr>
##  1 B-25973      4141   1698      13 Electronics Printers
##  2 B-25602      2617   1151       4 Electronics Phones
##  3 B-25761      2188   1050       5 Furniture   Bookcases
##  4 B-25923      3873    891       6 Electronics Phones
##  5 B-25830      1954    782       3 Electronics Phones
##  6 B-26073      1514    742       4 Electronics Printers
##  7 B-25853      2093    721       5 Furniture   Chairs
##  8 B-26093      2847    712       8 Electronics Printers
##  9 B-25862      2061    701       5 Furniture   Bookcases
## 10 B-25656      1389    680       7 Clothing    Saree
## # ... with 1,490 more rows
```

```r
orderDetails %>%
    rename(., profit = Profit)
```

```
## # A tibble: 1,500 x 6
##    ‘Order ID‘ Amount profit Quantity Category    ‘Sub-Category‘
##    <chr>       <dbl>  <dbl>    <dbl> <chr>       <chr>
##  1 B-25601      1275  -1148        7 Furniture   Bookcases
##  2 B-25601        66    -12        5 Clothing    Stole
##  3 B-25601         8     -2        3 Clothing    Hankerchief
##  4 B-25601        80    -56        4 Electronics Electronic Games
##  5 B-25602       168   -111        2 Electronics Phones
##  6 B-25602       424   -272        5 Electronics Phones
##  7 B-25602      2617   1151        4 Electronics Phones
##  8 B-25602       561    212        3 Clothing    Saree
##  9 B-25602       119     -5        8 Clothing    Saree
## 10 B-25603      1355    -60        5 Clothing    Trousers
## # ... with 1,490 more rows
```

```
orderDetails %>%
    mutate(., ProfitN = (Profit - min(Profit)) / (max(Profit) - min(Profit)))
```

```
## # A tibble: 1,500 x 7
##    ‘Order ID‘ Amount Profit Quantity Category    ‘Sub-Category‘   ProfitN
##    <chr>       <dbl>  <dbl>    <dbl> <chr>       <chr>              <dbl>
##  1 B-25601      1275  -1148        7 Furniture   Bookcases          0.226
##  2 B-25601        66    -12        5 Clothing    Stole              0.535
##  3 B-25601         8     -2        3 Clothing    Hankerchief        0.538
##  4 B-25601        80    -56        4 Electronics Electronic Games   0.523
##  5 B-25602       168   -111        2 Electronics Phones             0.508
##  6 B-25602       424   -272        5 Electronics Phones             0.465
##  7 B-25602      2617   1151        4 Electronics Phones             0.851
##  8 B-25602       561    212        3 Clothing    Saree              0.596
##  9 B-25602       119     -5        8 Clothing    Saree              0.537
## 10 B-25603      1355    -60        5 Clothing    Trousers           0.522
## # ... with 1,490 more rows
```

```
orderDetails %>%
  group_by(Category, `Sub-Category`) %>%
  summarize(`Average Profit` = mean(Profit, na.rm = TRUE)) %>%
  arrange(desc(`Average Profit`))
```

```
## ‘summarise()‘ has grouped output by ’Category’. You can override using the ‘.groups‘ argument.
```

```
## # A tibble: 17 x 3
## # Groups:   Category [3]
##    Category    ‘Sub-Category‘  ‘Average Profit‘
##    <chr>       <chr>                      <dbl>
##  1 Electronics Printers                    80.6
##  2 Clothing    Trousers                    73
##  3 Furniture   Bookcases                   61.9
##  4 Electronics Accessories                 49.4
##  5 Electronics Phones                      26.6
##  6 Clothing    T-shirt                     19.5
##  7 Clothing    Shirt                       16.4
##  8 Clothing    Stole                       13.3
```

```
##  9 Furniture   Furnishings              11.6
## 10 Clothing    Hankerchief              10.6
## 11 Furniture   Chairs                    7.80
## 12 Clothing    Leggings                  4.91
## 13 Clothing    Kurti                     3.85
## 14 Clothing    Skirt                     3.67
## 15 Clothing    Saree                     1.68
## 16 Electronics Electronic Games        -15.6
## 17 Furniture   Tables                  -236.
```