Zara EDA, DV

Andrew Bonici

2024-03-23

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(ggplot2)
library(dplyr)
library(psych)
```

```
##
## Attaching package: 'psych'
##
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```r
library(forecast)
```

```
## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo
```

1. Data Import & Structure

1. Product ID: Unique identifier for each product.
2. Product Position: The position of the product in the catalog or store layout.
3. Promotion: Indicator of whether the product is currently on promotion or not.
4. Product Category: The category of the product, such as clothing, accessories, shoes, etc.

5. Seasonal: Indicator of whether the product is part of a specific seasonal collection.
6. Sales Volume: The quantity of products sold.
7. Brand: Brand of the product.
8. URL: Product URL (e.g., if the product is sold online).
9. SKU: Stock Keeping Unit, a unique code used to identify items available for sale.
10. Name: Name of the product.
11. Description: Description of the product.
12. Price: Price of the product.
13. Currency: Currency of the product price.
14. Scraped_at: The time when the data was scraped (e.g., in web scraping process).
15. Terms: Terms or conditions of the product.
16. Section: Section or category where the product is sold in the store (e.g., women's clothing, men's clothing, children's clothing, etc.).

```r
data <- read.csv('C:\\Datasets\\Zara Sales\\zara.csv', sep = ";", quote = "\"" ,stringsAsFactors = FALS

head(data)
```

```
##   Product.ID Product.Position Promotion Product.Category Seasonal Sales.Volume
## 1     185102            Aisle        No         Clothing       No         2823
## 2     188771            Aisle        No         Clothing       No          654
## 3     180176          End-cap       Yes         Clothing      Yes         2220
## 4     112917            Aisle       Yes         Clothing      Yes         1568
## 5     192936          End-cap        No         Clothing      Yes         2942
## 6     117590          End-cap        No         Clothing       No         2968
##   brand                                                                 url
## 1  Zara       https://www.zara.com/us/en/basic-puffer-jacket-p06985450.html
## 2  Zara             https://www.zara.com/us/en/tuxedo-jacket-p08896675.html
## 3  Zara      https://www.zara.com/us/en/slim-fit-suit-jacket-p01564520.html
## 4  Zara       https://www.zara.com/us/en/stretch-suit-jacket-p01564300.html
## 5  Zara       https://www.zara.com/us/en/double-faced-jacket-p08281477.html
## 6  Zara https://www.zara.com/us/en/contrasting-collar-jacket-p06987331.html
##                sku                      name
## 1  272145190-250-2       BASIC PUFFER JACKET
## 2 324052738-800-46             TUXEDO JACKET
## 3 335342680-800-44      SLIM FIT SUIT JACKET
## 4 328303236-420-44       STRETCH SUIT JACKET
## 5  312368260-800-2       DOUBLE FACED JACKET
## 6  320298385-807-2 CONTRASTING COLLAR JACKET
##
## 1          Puffer jacket made of tear-resistant ripstop fabric. High collar and adjustable long slee
## 2                               Straight fit blazer. Pointed lapel collar and long sleeves with butto
## 3                              Slim fit jacket. Notched lapel collar. Long sleeves with buttoned cuf
## 4 Slim fit jacket made of viscose blend fabric. Notched lapel collar. Long sleeves with buttoned cuf
## 5                                                              Jacket made of faux leather faux shearl
## 6                                             Relaxed fit jacket. Contrasting lapel collar and long s
##    price currency                 scraped_at   terms section
## 1  19.99      USD 2024-02-19T08:50:05.654618 jackets     MAN
## 2 169.00      USD 2024-02-19T08:50:06.590930 jackets     MAN
## 3 129.00      USD 2024-02-19T08:50:07.301419 jackets     MAN
## 4 129.00      USD 2024-02-19T08:50:07.882922 jackets     MAN
## 5 139.00      USD 2024-02-19T08:50:08.453847 jackets     MAN
## 6  79.90      USD 2024-02-19T08:50:09.140497 jackets     MAN
```

```
str(data)
```

```
## ’data.frame':    252 obs. of  16 variables:
##  $ Product.ID      : int  185102 188771 180176 112917 192936 117590 189118 182157 141861 137121 ...
##  $ Product.Position: chr  "Aisle" "Aisle" "End-cap" "Aisle" ...
##  $ Promotion       : chr  "No" "No" "Yes" "Yes" ...
##  $ Product.Category: chr  "Clothing" "Clothing" "Clothing" "Clothing" ...
##  $ Seasonal        : chr  "No" "No" "Yes" "Yes" ...
##  $ Sales.Volume    : int  2823 654 2220 1568 2942 2968 952 2421 1916 656 ...
##  $ brand           : chr  "Zara" "Zara" "Zara" "Zara" ...
##  $ url             : chr  "https://www.zara.com/us/en/basic-puffer-jacket-p06985450.html" "https://ww
##  $ sku             : chr  "272145190-250-2" "324052738-800-46" "335342680-800-44" "328303236-420-44"
##  $ name            : chr  "BASIC PUFFER JACKET" "TUXEDO JACKET" "SLIM FIT SUIT JACKET" "STRETCH SUIT
##  $ description     : chr  "Puffer jacket made of tear-resistant ripstop fabric. High collar and adjus
##  $ price           : num  20 169 129 129 139 ...
##  $ currency        : chr  "USD" "USD" "USD" "USD" ...
##  $ scraped_at      : chr  "2024-02-19T08:50:05.654618" "2024-02-19T08:50:06.590930" "2024-02-19T08:50
##  $ terms           : chr  "jackets" "jackets" "jackets" "jackets" ...
##  $ section         : chr  "MAN" "MAN" "MAN" "MAN" ...
```

Missing value check

```
colSums(is.na(data))
```

```
##       Product.ID Product.Position        Promotion Product.Category
##                0                0                0                0
##         Seasonal     Sales.Volume            brand              url
##                0                0                0                0
##              sku             name      description            price
##                0                0                0                0
##         currency       scraped_at            terms          section
##                0                0                0                0
```

```
dim(data)
```

```
## [1] 252  16
```

```
describe(data)
```

```
##                  vars   n      mean       sd   median   trimmed      mad
## Product.ID          1 252 153370.50 26160.44 151681.5 152999.16 33372.58
## Product.Position*   2 252      1.89     0.81      2.0      1.86     1.48
## Promotion*          3 252      1.48     0.50      1.0      1.47     0.00
## Product.Category*    4 252      1.00     0.00      1.0      1.00     0.00
## Seasonal*            5 252      1.51     0.50      2.0      1.51     0.00
## Sales.Volume         6 252   1823.70   697.70   1839.5   1835.50   868.80
## brand*               7 252      1.00     0.00      1.0      1.00     0.00
## url*                 8 252    112.52    65.50    108.5    112.01    83.77
## sku*                 9 252    115.14    65.91    115.5    115.26    83.03
## name*               10 252     97.33    53.97     98.0     97.24    65.98
## description*        11 252    112.06    64.86    110.5    112.05    82.28
## price               12 252     86.25    52.08     79.9     80.92    43.14
## currency*           13 252      1.00     0.00      1.0      1.00     0.00
## scraped_at*         14 252    123.39    68.52    126.5    125.23    93.40
## terms*              15 252      2.27     1.55      1.0      2.09     0.00
## section*            16 252      1.13     0.34      1.0      1.04     0.00
##                          min    max    range  skew kurtosis      se
## Product.ID         110075.00 199631 89556.00  0.11    -1.23 1647.95
## Product.Position*       1.00      3     2.00  0.20    -1.44    0.05
## Promotion*              1.00      2     1.00  0.09    -2.00    0.03
## Product.Category*       1.00      1     0.00   NaN      NaN    0.00
## Seasonal*               1.00      2     1.00 -0.03    -2.01    0.03
## Sales.Volume          529.00   2989  2460.00 -0.11    -1.13   43.95
## brand*                  1.00      1     0.00   NaN      NaN    0.00
## url*                    1.00    228   227.00  0.05    -1.18    4.13
## sku*                    1.00    228   227.00 -0.01    -1.20    4.15
## name*                   1.00    195   194.00  0.01    -1.13    3.40
## description*            1.00    222   221.00  0.01    -1.19    4.09
## price                   7.99    439   431.01  2.36    10.99    3.28
## currency*               1.00      1     0.00   NaN      NaN    0.00
## scraped_at*             1.00    229   228.00 -0.14    -1.30    4.32
## terms*                  1.00      5     4.00  0.62    -1.28    0.10
## section*                1.00      2     1.00  2.12     2.52    0.02
```

Removing unnecessary variables

```
data <- data[,-c(8:9, 11, 14)]
data2 <- data
```

Variable type transformation

```
data$Product.Position <- as.factor(data$Product.Position)
data$Product.Category <- as.factor(data$Product.Category)
data$brand <- as.factor(data$brand)
data$terms <- as.factor(data$terms)
data$section <- as.factor(data$section)
data$name <- as.factor(data$name)
data$currency <- as.factor(data$currency)
```

Data transformation

```
# Promotion : No - \> 0 , Yes -\> 1
data$Promotion <- ifelse(data$Promotion == 'No', 0 ,
                        ifelse(data$Promotion == 'Yes',1,2))

# Seasonal : No - > 0 , Yes -> 1
```

```
data$Seasonal <- ifelse(data$Seasonal == 'No',0,
                        ifelse(data$Seasonal=='Yes',1,2))

# section : MAN - > 0 , WOMEN -> 1 ,
data$section <- ifelse(data$section == 'MAN',0,
                       ifelse(data$section=='WOMEN',1,2))

head(data)
```

```
##   Product.ID Product.Position Promotion Product.Category Seasonal Sales.Volume
## 1     185102            Aisle         0         Clothing        0         2823
## 2     188771            Aisle         0         Clothing        0          654
## 3     180176          End-cap         1         Clothing        1         2220
## 4     112917            Aisle         1         Clothing        1         1568
## 5     192936          End-cap         0         Clothing        1         2942
## 6     117590          End-cap         0         Clothing        0         2968
##   brand                      name  price currency   terms section
## 1  Zara       BASIC PUFFER JACKET  19.99      USD jackets       0
## 2  Zara             TUXEDO JACKET 169.00      USD jackets       0
## 3  Zara      SLIM FIT SUIT JACKET 129.00      USD jackets       0
## 4  Zara       STRETCH SUIT JACKET 129.00      USD jackets       0
## 5  Zara       DOUBLE FACED JACKET 139.00      USD jackets       0
## 6  Zara CONTRASTING COLLAR JACKET  79.90      USD jackets       0
```

```
str(data)
```

```
## 'data.frame':    252 obs. of  12 variables:
##  $ Product.ID      : int  185102 188771 180176 112917 192936 117590 189118 182157 141861 137121 ...
##  $ Product.Position: Factor w/ 3 levels "Aisle","End-cap",..: 1 1 2 1 2 2 3 1 1 1 ...
##  $ Promotion       : num  0 0 1 1 0 0 1 0 1 0 ...
##  $ Product.Category: Factor w/ 1 level "Clothing": 1 1 1 1 1 1 1 1 1 1 ...
##  $ Seasonal        : num  0 0 1 1 1 0 1 0 1 1 ...
##  $ Sales.Volume    : int  2823 654 2220 1568 2942 2968 952 2421 1916 656 ...
##  $ brand           : Factor w/ 1 level "Zara": 1 1 1 1 1 1 1 1 1 1 ...
##  $ name            : Factor w/ 195 levels "","100% FEATHER FILL PUFFER JACKET",..: 22 178 143 148 54
##  $ price           : num  20 169 129 129 139 ...
##  $ currency        : Factor w/ 1 level "USD": 1 1 1 1 1 1 1 1 1 1 ...
##  $ terms           : Factor w/ 5 levels "jackets","jeans",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ section         : num  0 0 0 0 0 0 0 0 0 0 ...
```

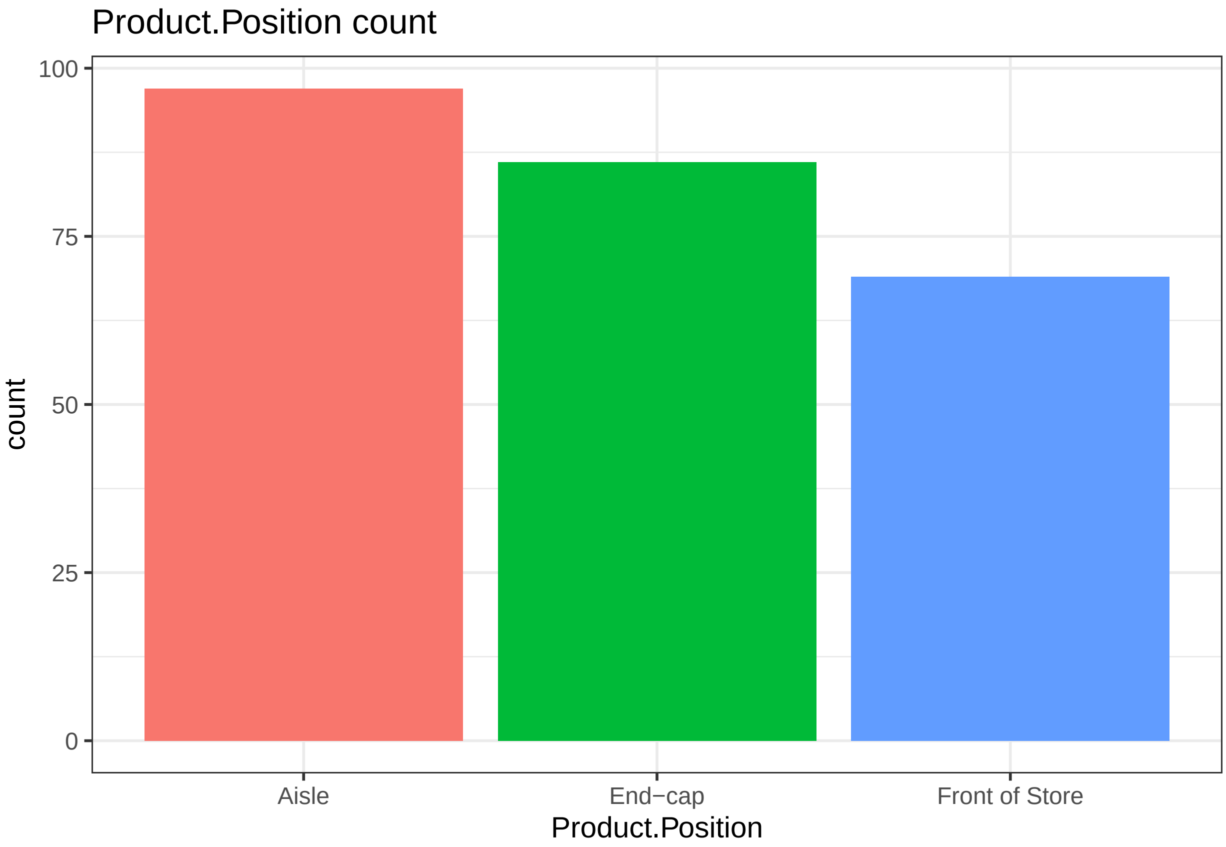
```
summary(data)
```

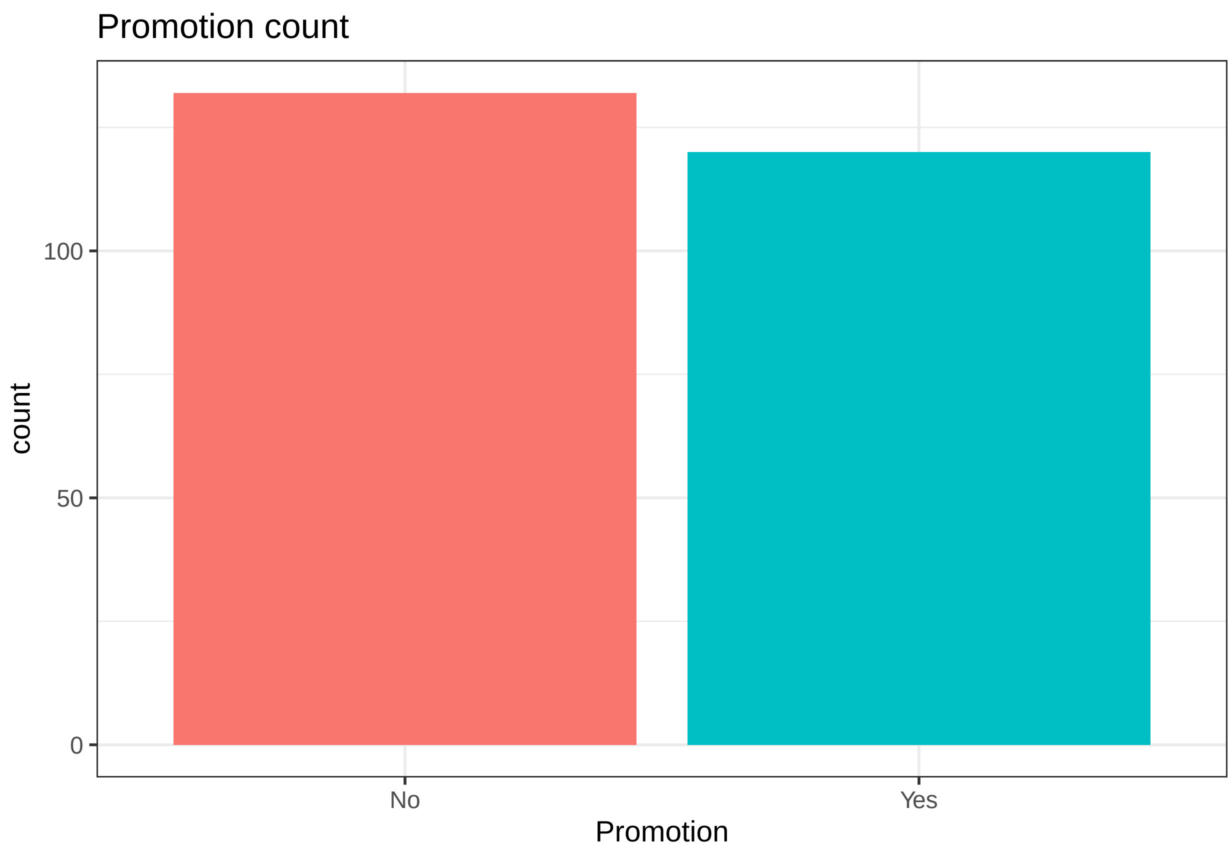
```
##    Product.ID           Product.Position   Promotion      Product.Category
##  Min.   :110075   Aisle         :97      Min.   :0.0000   Clothing:252
##  1st Qu.:131054   End-cap       :86      1st Qu.:0.0000
##  Median :151682   Front of Store:69      Median :0.0000
##  Mean   :153371                          Mean   :0.4762
##  3rd Qu.:175670                          3rd Qu.:1.0000
##  Max.   :199631                          Max.   :1.0000
##
##     Seasonal       Sales.Volume   brand
##  Min.   :0.0000   Min.   : 529   Zara:252
##  1st Qu.:0.0000   1st Qu.:1243
##  Median :1.0000   Median :1840
##  Mean   :0.5079   Mean   :1824
##  3rd Qu.:1.0000   3rd Qu.:2399
##  Max.   :1.0000   Max.   :2989
##
##                                name         price         currency
##  PLAID OVERSHIRT                 :  6   Min.   :  7.99   USD:252
##  PATCH BOMBER JACKET             :  4   1st Qu.: 49.90
##  POCKET OVERSHIRT                :  4   Median : 79.90
##  BOMBER JACKET                   :  3   Mean   : 86.25
##  CONTRASTING PATCHES BOMBER JACKET:  3   3rd Qu.:109.00
##  FAUX LEATHER BOMBER JACKET      :  3   Max.   :439.00
##  (Other)                         :229
##       terms        section
##  jackets :140   Min.   :0.0000
##  jeans   :  8   1st Qu.:0.0000
##  shoes   : 31   Median :0.0000
##  sweaters: 41   Mean   :0.2698
##  t-shirts: 32   3rd Qu.:0.0000
##                 Max.   :2.0000
##
```

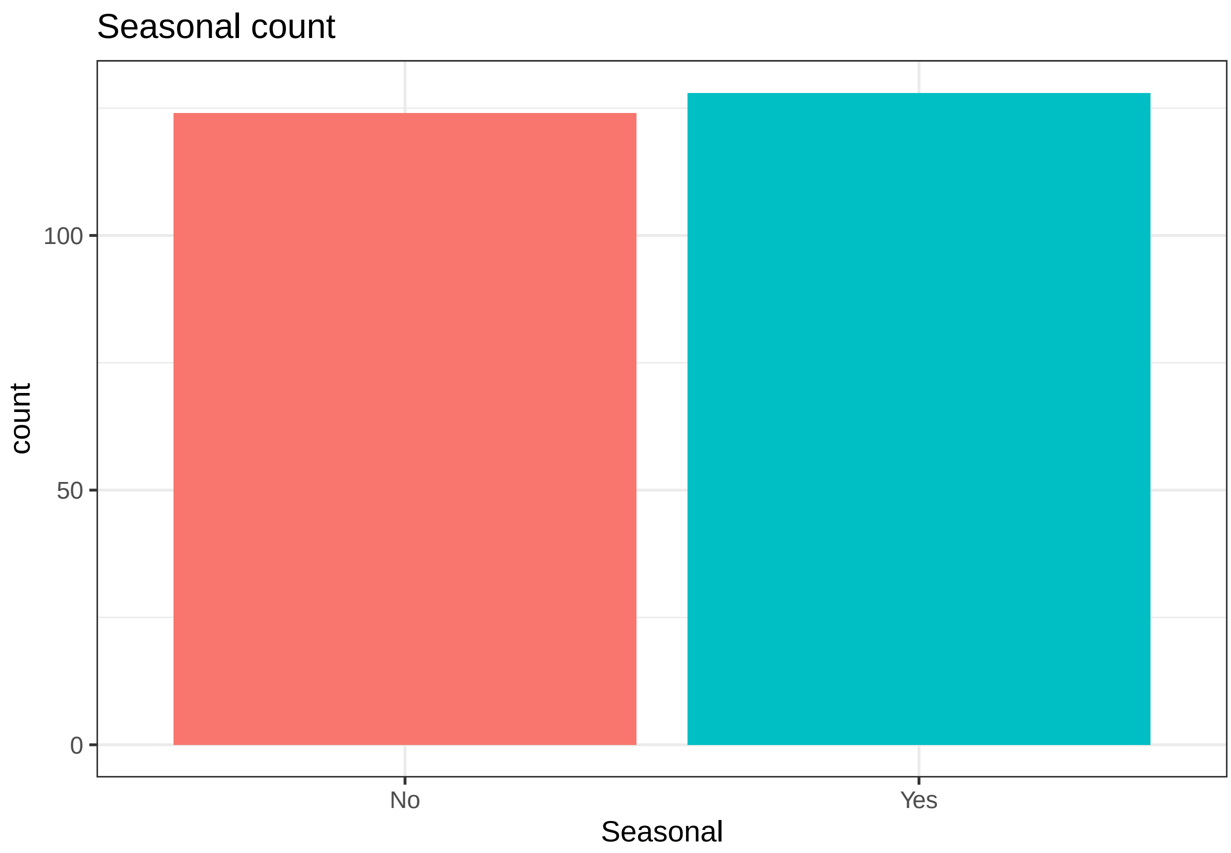
2 Data visualisation

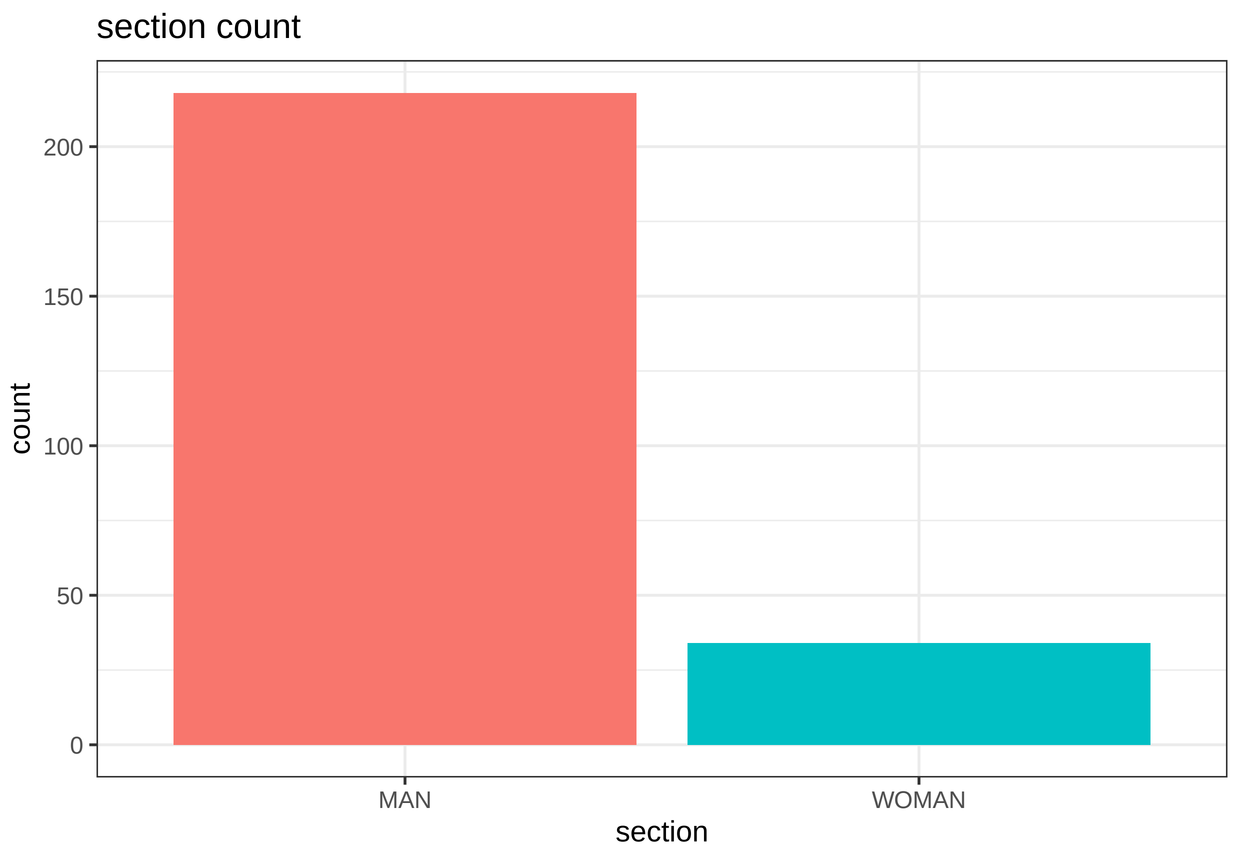
```
cols = c('Product.Position','Promotion','Seasonal','section')

for (i in cols){
    print(ggplot(data2, aes(x=data2[,i],fill = data2[,i])) + geom_bar() + ggtitle(paste(i,'count'))+ xla
}
```

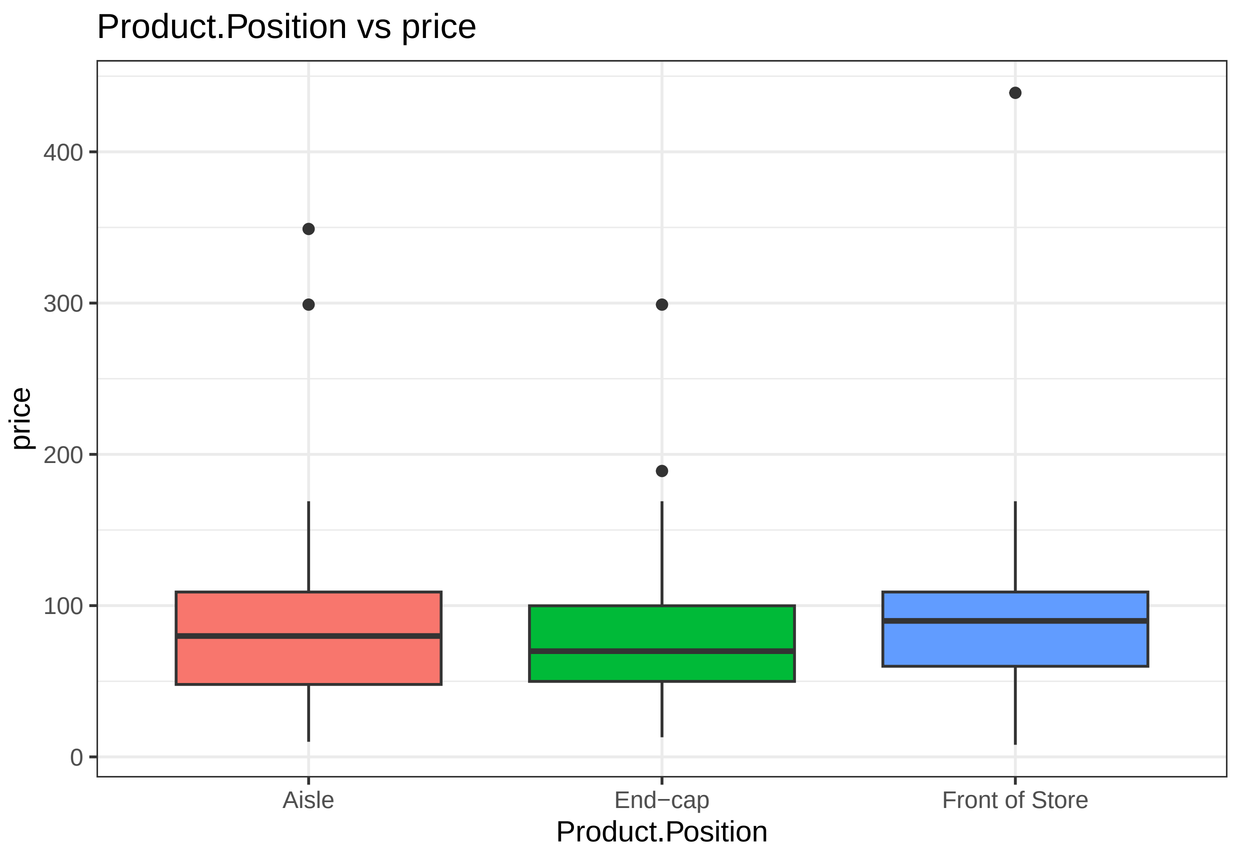

Product.Position count
100
75
50
25
0
count
Aisle
End−cap
Front of Store
Product.Position

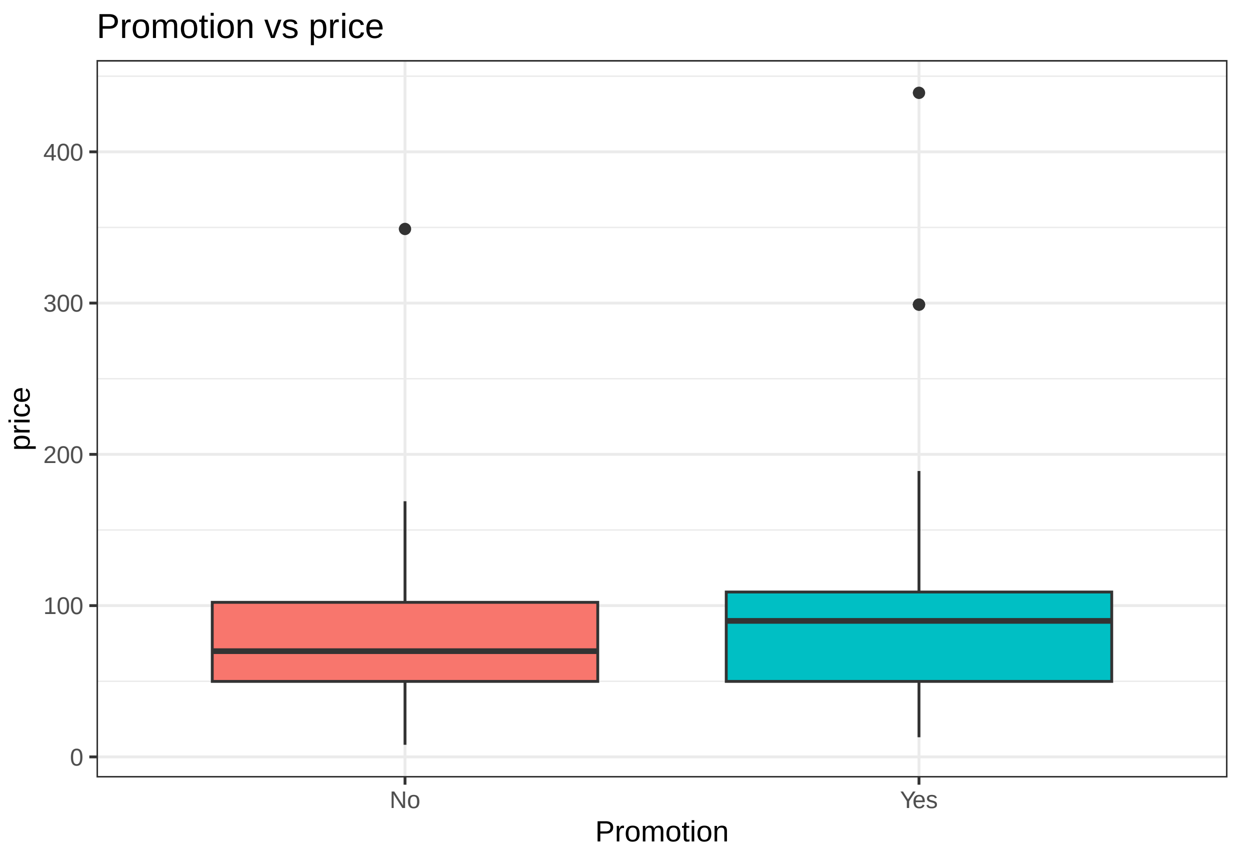

Promotion count
count
100
50
0
No
Yes
Promotion

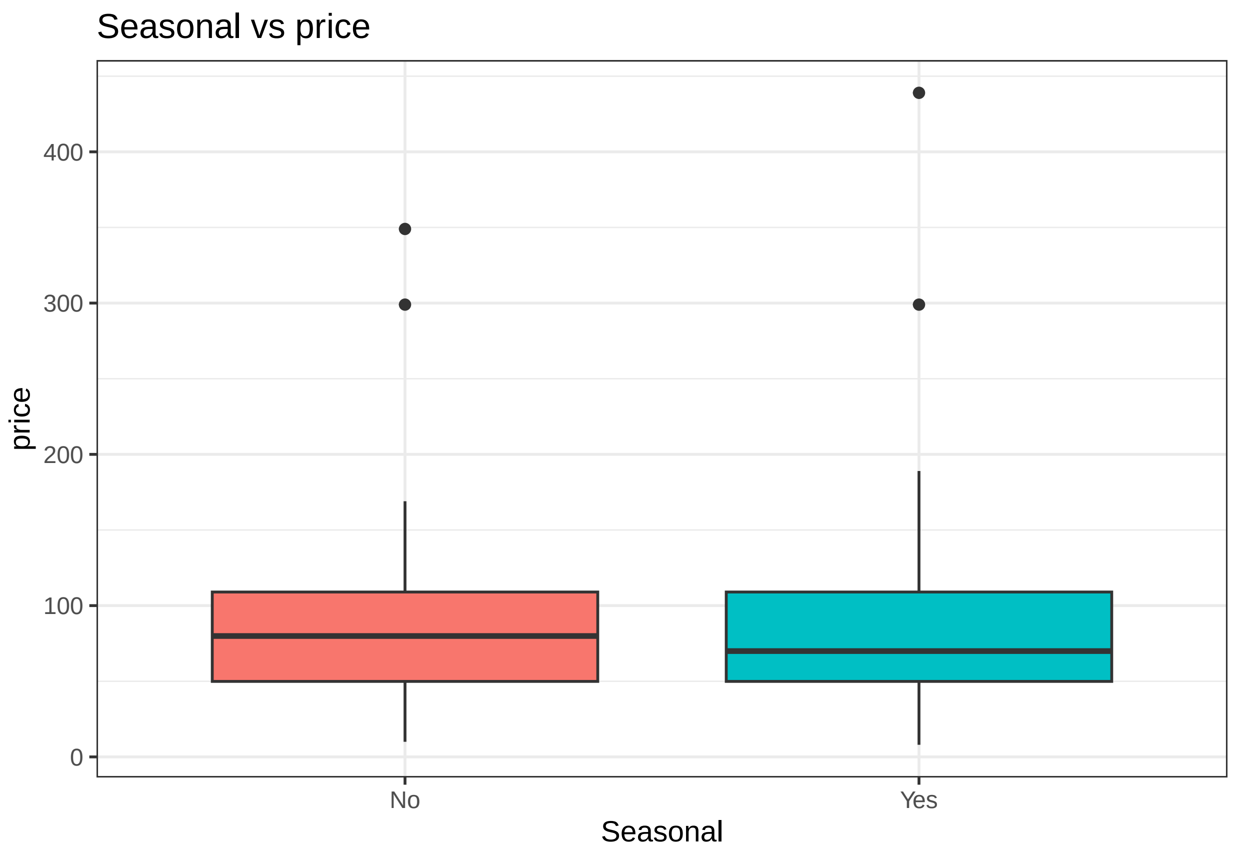

Seasonal count
count
100
50
0
No
Yes
Seasonal

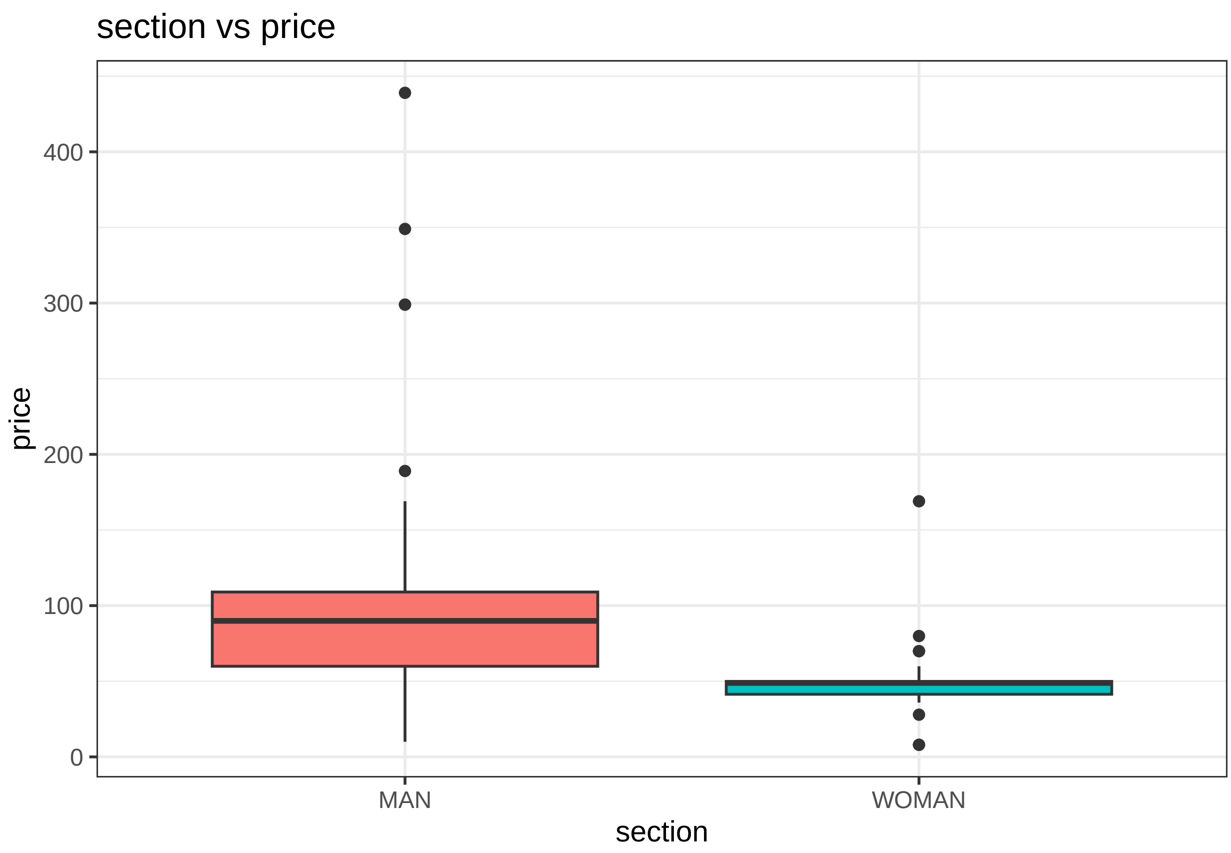


```
for(i in cols){
    print(ggplot(data2, aes(x=data2[,i], y=price, fill= data2[,i])) + geom_boxplot() + xlab(i) +ggtitle
}
```

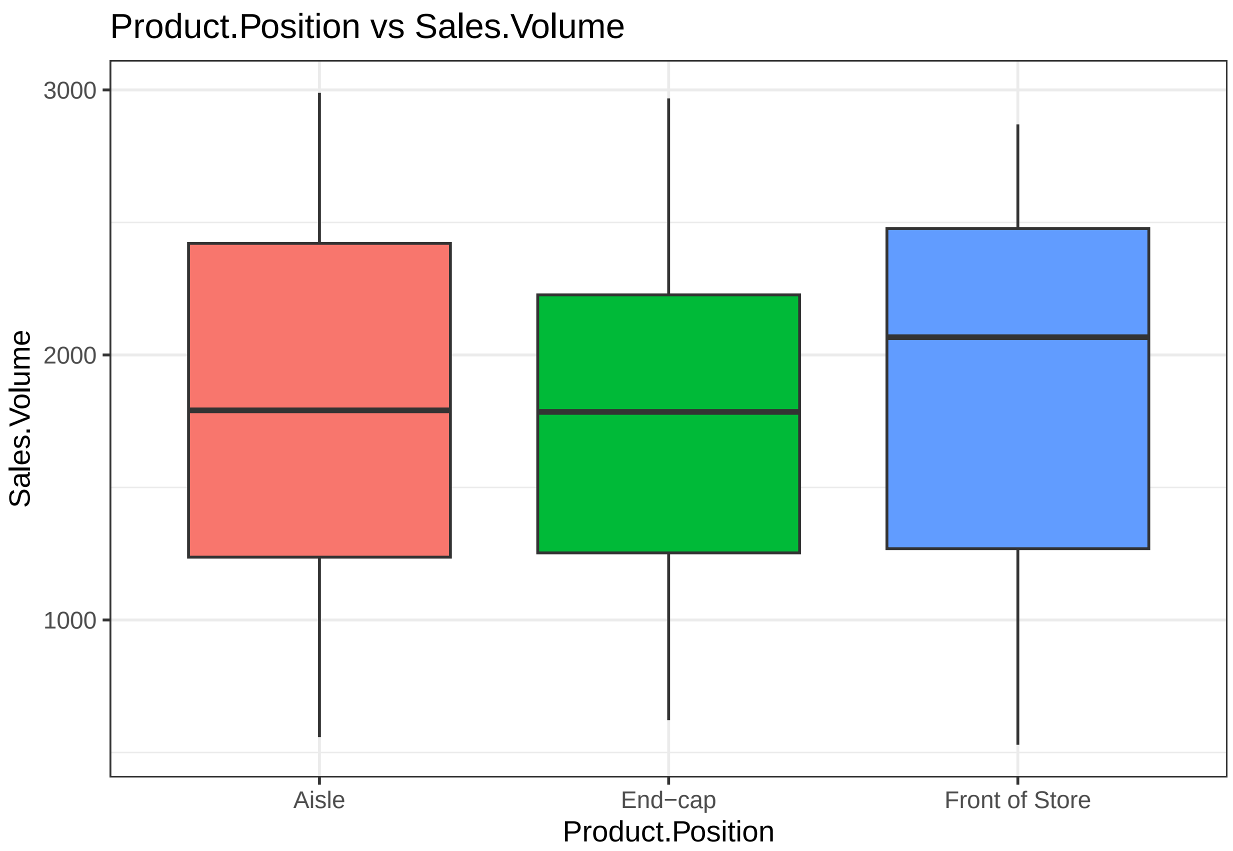

Product.Position vs price
price
400
300
200
100
0
Aisle
End−cap
Front of Store
Product.Position

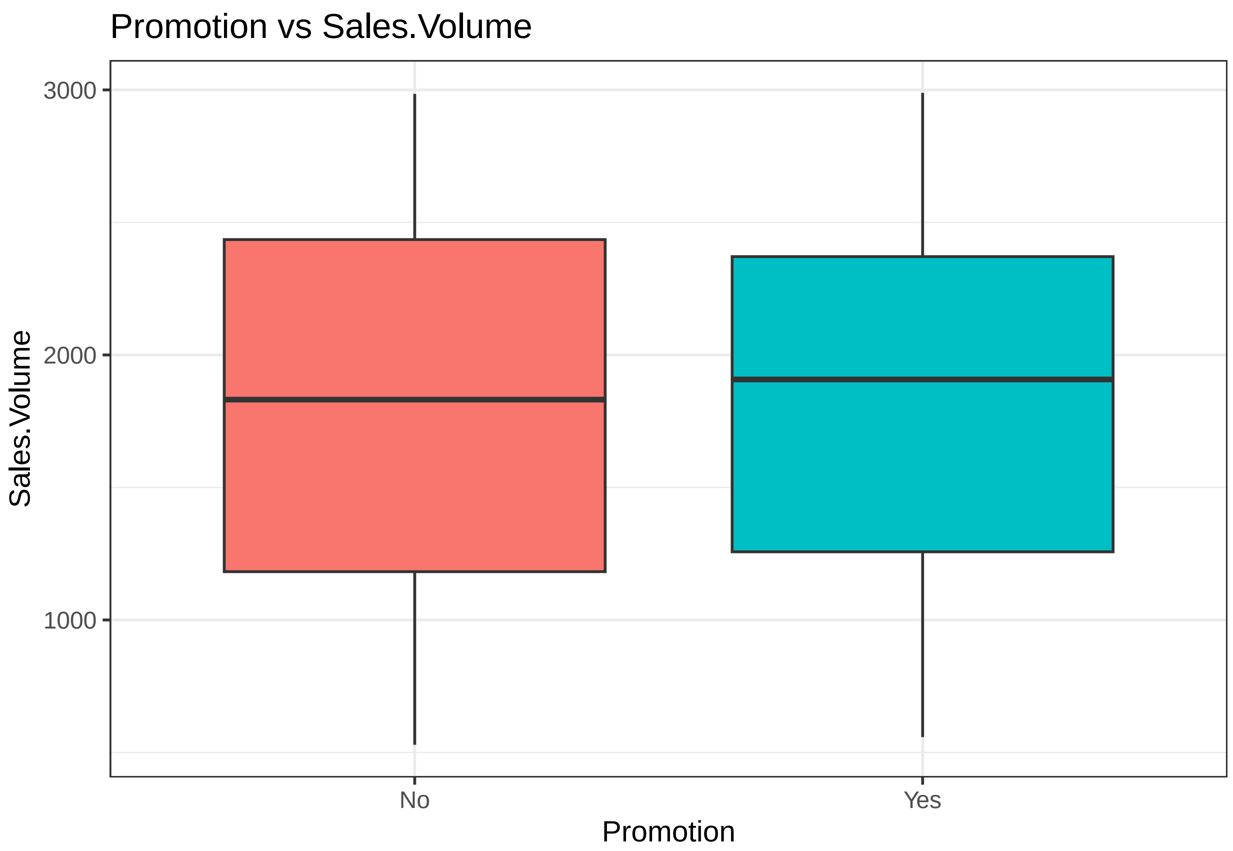

Promotion vs price
price
400
300
200
100
0
No
Yes
Promotion

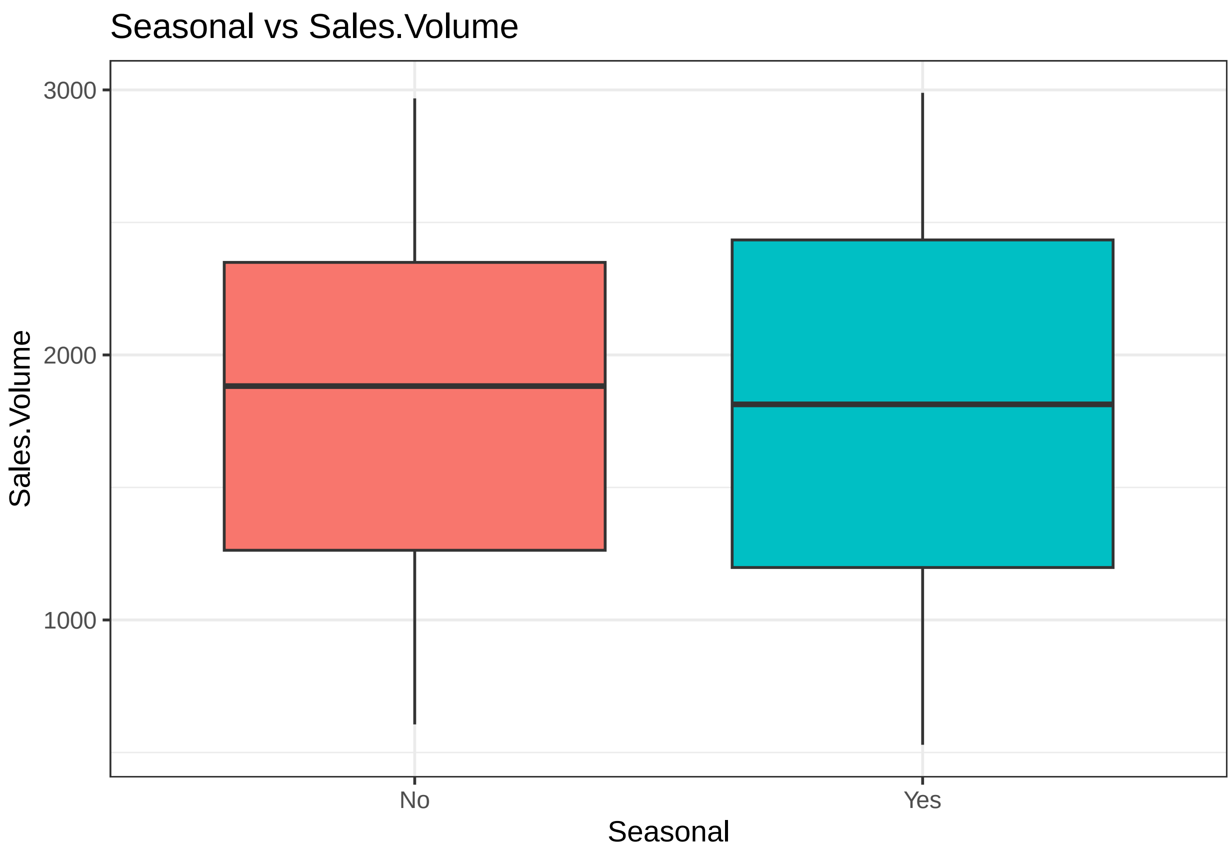

Seasonal vs price
price
400
300
200
100
0
No
Yes
Seasonal

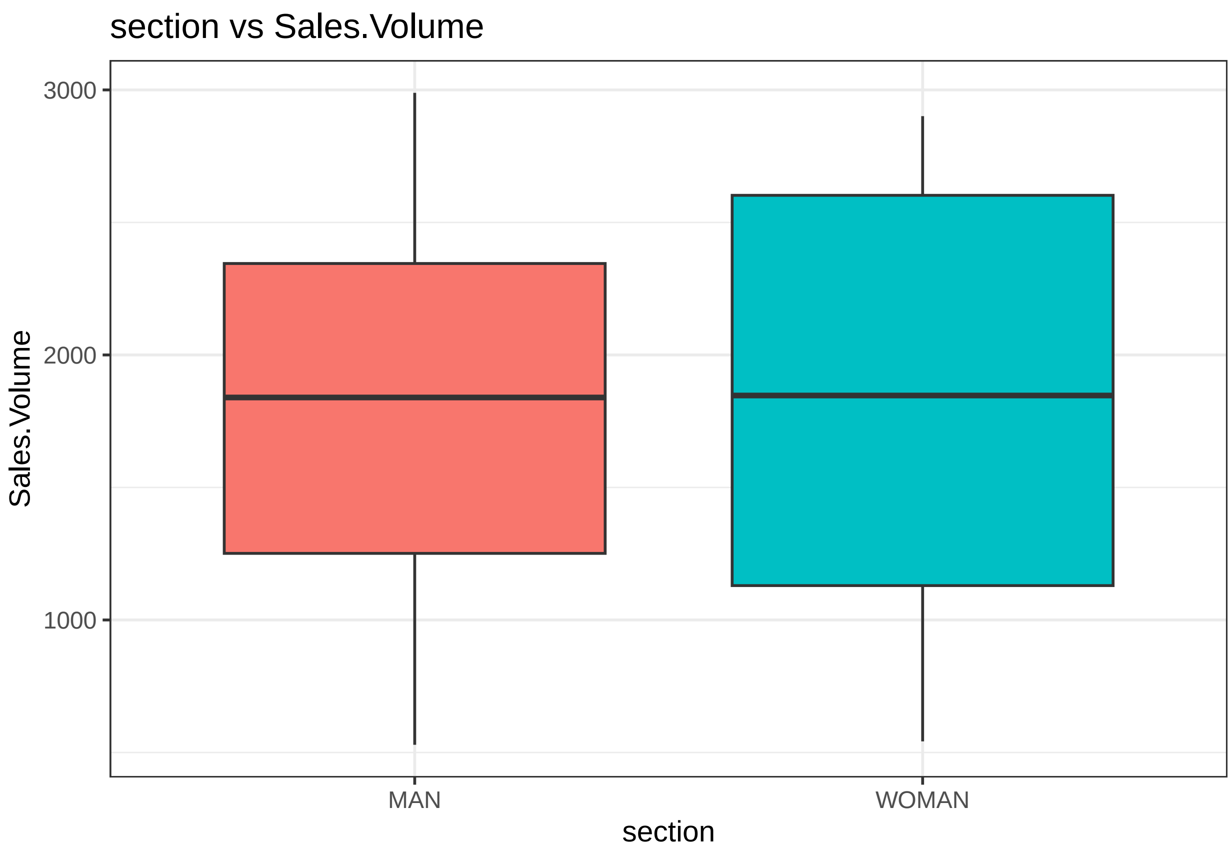


```
for(i in cols){
    print(ggplot(data2, aes(x=data2[,i], y=Sales.Volume, fill= data2[,i])) + geom_boxplot() + xlab(i) +
}
```

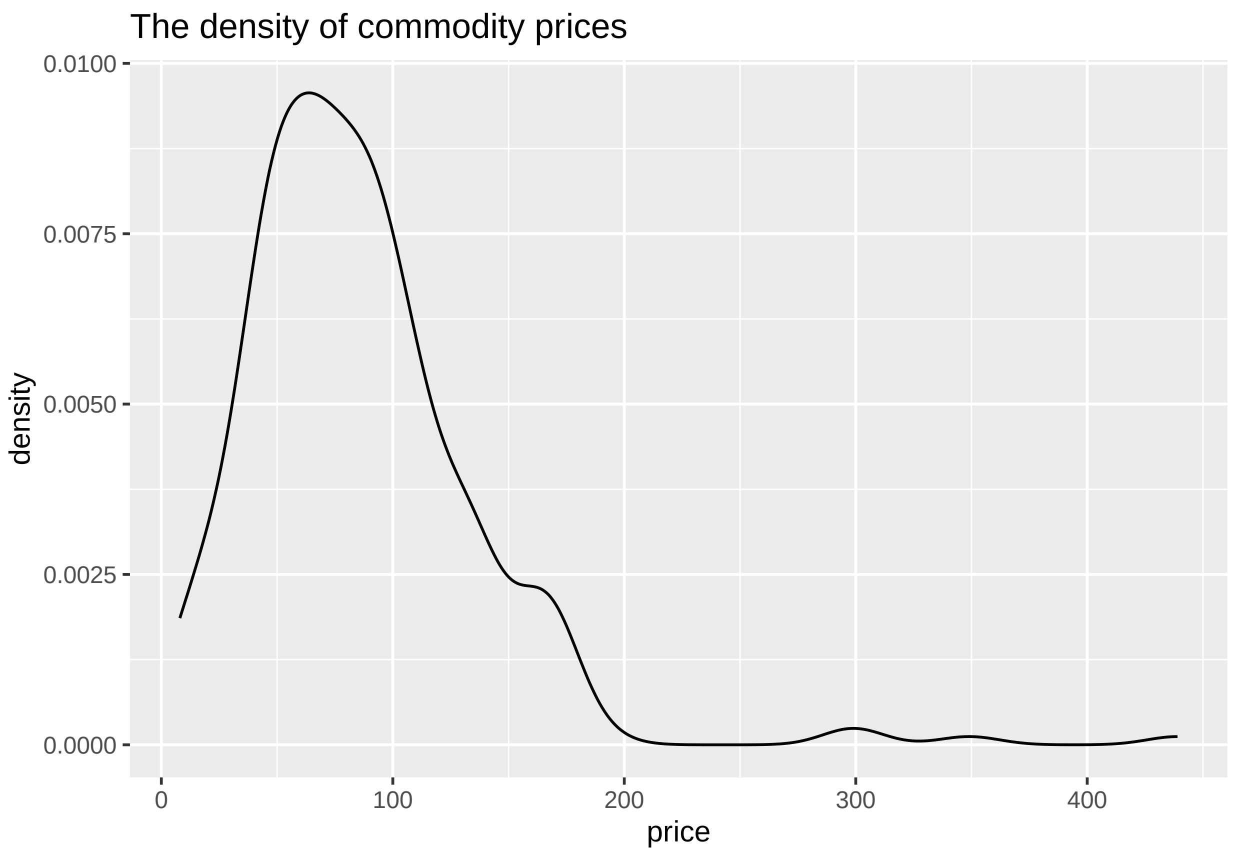

Product.Position vs Sales.Volume
Sales.Volume
3000
2000
1000
Aisle
End−cap
Front of Store
Product.Position


Promotion vs Sales.Volume
3000
2000
1000
Sales.Volume
No
Yes
Promotion


Seasonal vs Sales.Volume
3000
2000
1000
Sales.Volume
No
Yes
Seasonal



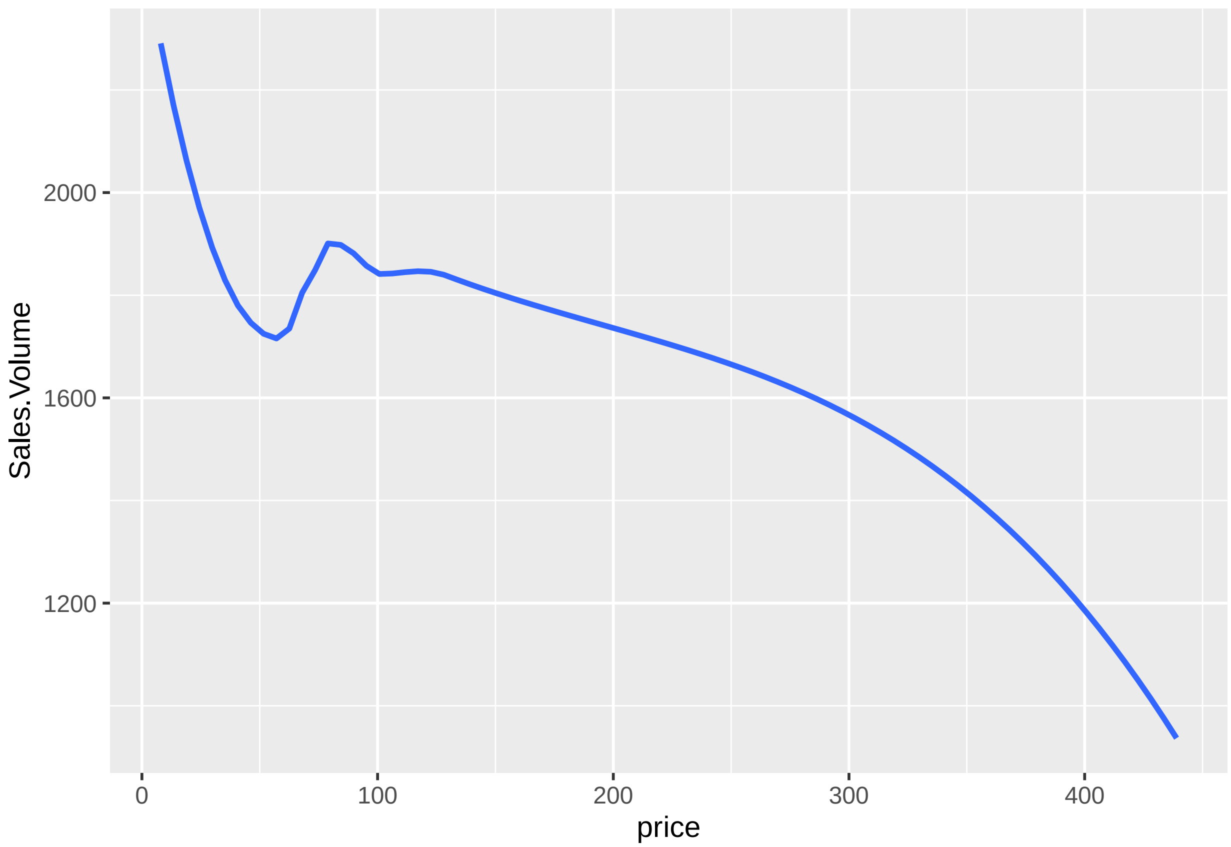
```
ggplot(data2, aes(x=price)) + geom_density() + ggtitle('The density of commodity prices') + xlab('price
```



Most items are under $100.

```
ggplot(data2, aes(x=price, y= Sales.Volume)) + geom_smooth(se=F)
```

```
## ‘geom_smooth()‘ using method = ’loess’ and formula = ’y ~ x’
```



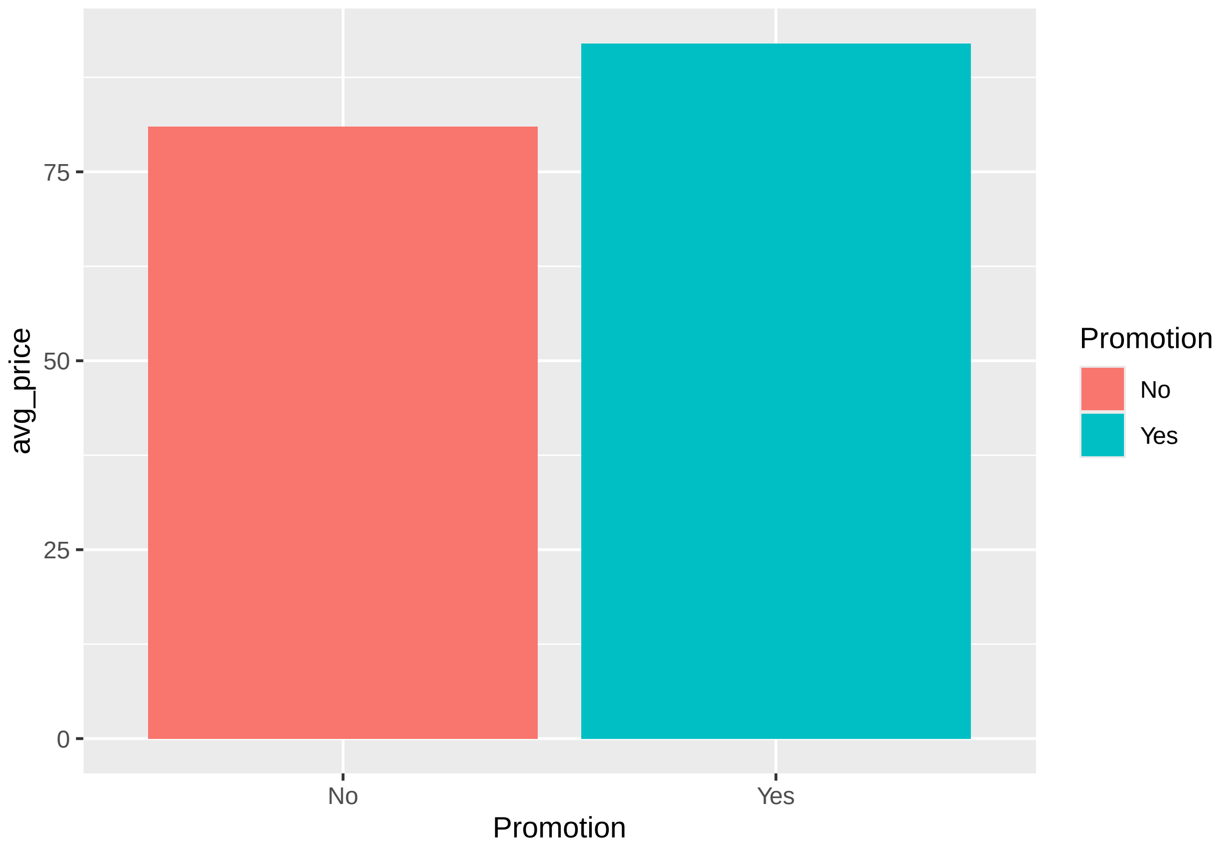
Relation between price and sales.

Checking average price by group.

```
data2_promotion <- data2 %>% group_by(Promotion) %>% summarize(N = n(), avg_price = round(mean(price,na
data2_promotion
```

```
## # A tibble: 2 x 3
##   Promotion     N avg_price
##   <chr>     <int>     <dbl>
## 1 No          132        81
## 2 Yes         120        92
```

```
ggplot(data2_promotion, aes(x=Promotion, y= avg_price,fill = Promotion)) + geom_col()
```

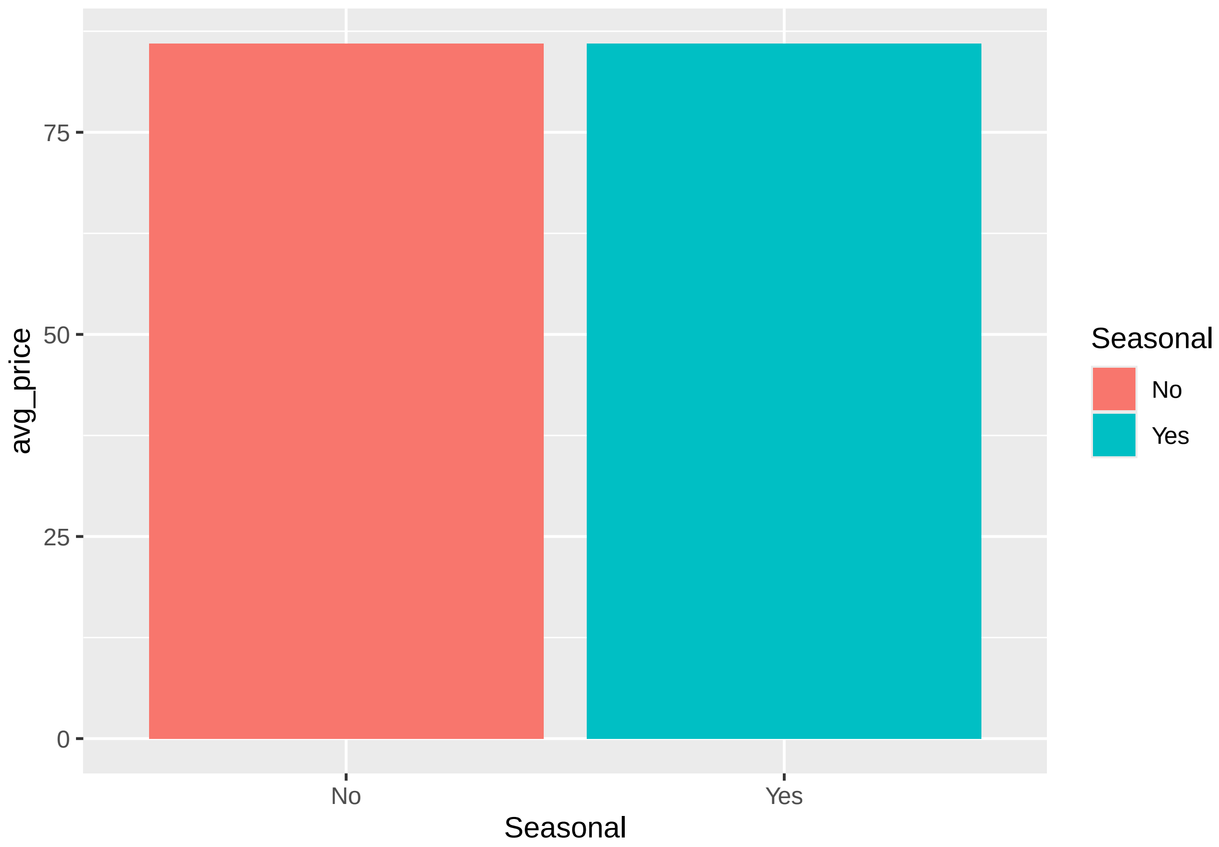


Average price of productions on promotion.

```
data2_Seasonal <- data2 %>% group_by(Seasonal) %>% summarize(N = n(), avg_price = round(mean(price,na.r
data2_Seasonal
```

```
## # A tibble: 2 x 3
##   Seasonal      N avg_price
##   <chr>     <int>     <dbl>
## 1 No          124        86
## 2 Yes         128        86
```

```
ggplot(data2_Seasonal, aes(x=Seasonal,y=avg_price, fill = Seasonal)) + geom_col()
```



Not much of a difference in average price.

```
# Average calculation according to the number of people allocated
data2_section <- data2 %>% group_by(section) %>%  summarize(N = n(), avg_price = round(mean(price,na.rm=
data2_section
```

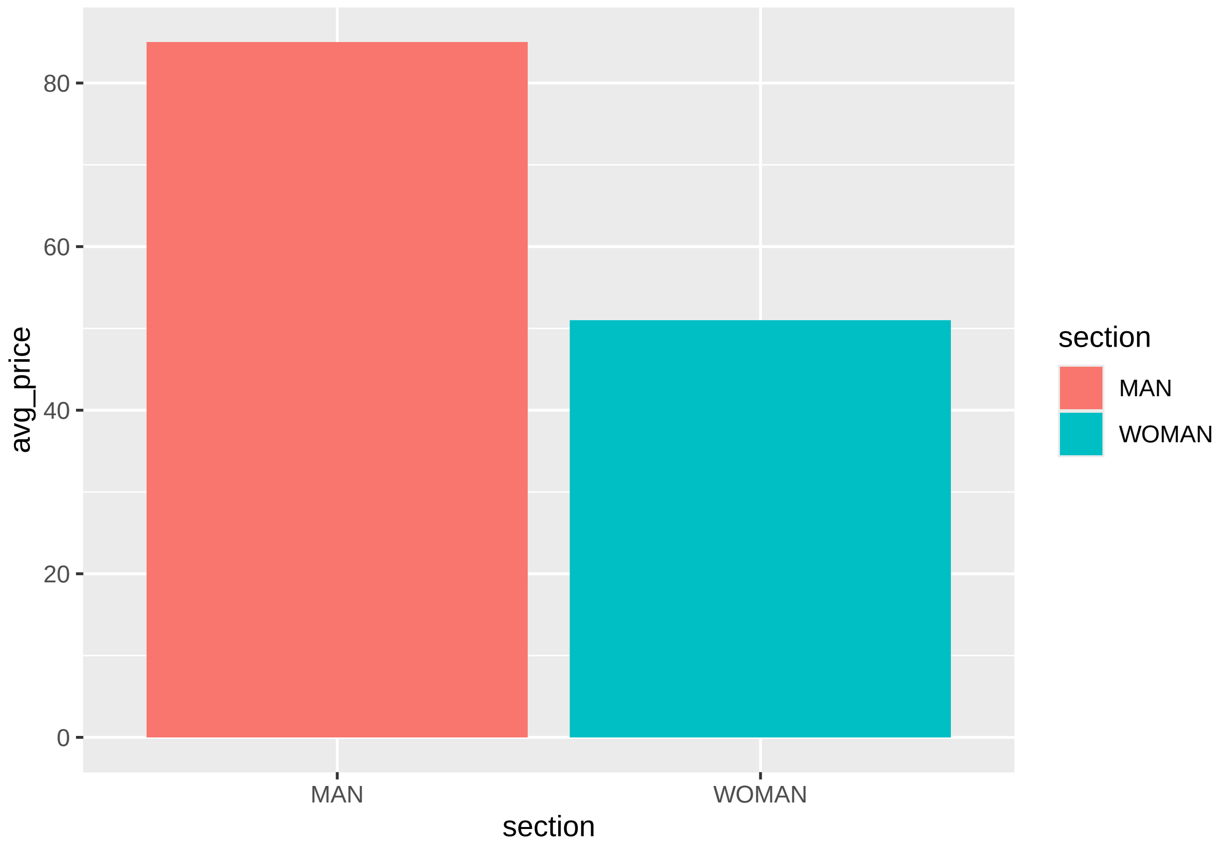
```
## # A tibble: 2 x 3
##   section     N avg_price
##   <chr>   <int>     <dbl>
## 1 MAN       218        92
## 2 WOMAN      34        51
```

```
# If pick 30 people and average them

data2_section30 <- data2 %>% group_by(section) %>% sample_n(size = 30)%>% summarize(N = n(), avg_price =
data2_section30
```

```
## # A tibble: 2 x 3
##   section     N avg_price
##   <chr>   <int>     <dbl>
## 1 MAN        30        85
## 2 WOMAN      30        51
```

```
ggplot(data2_section30, aes(x=section,y=avg_price, fill = section)) + geom_col()
```

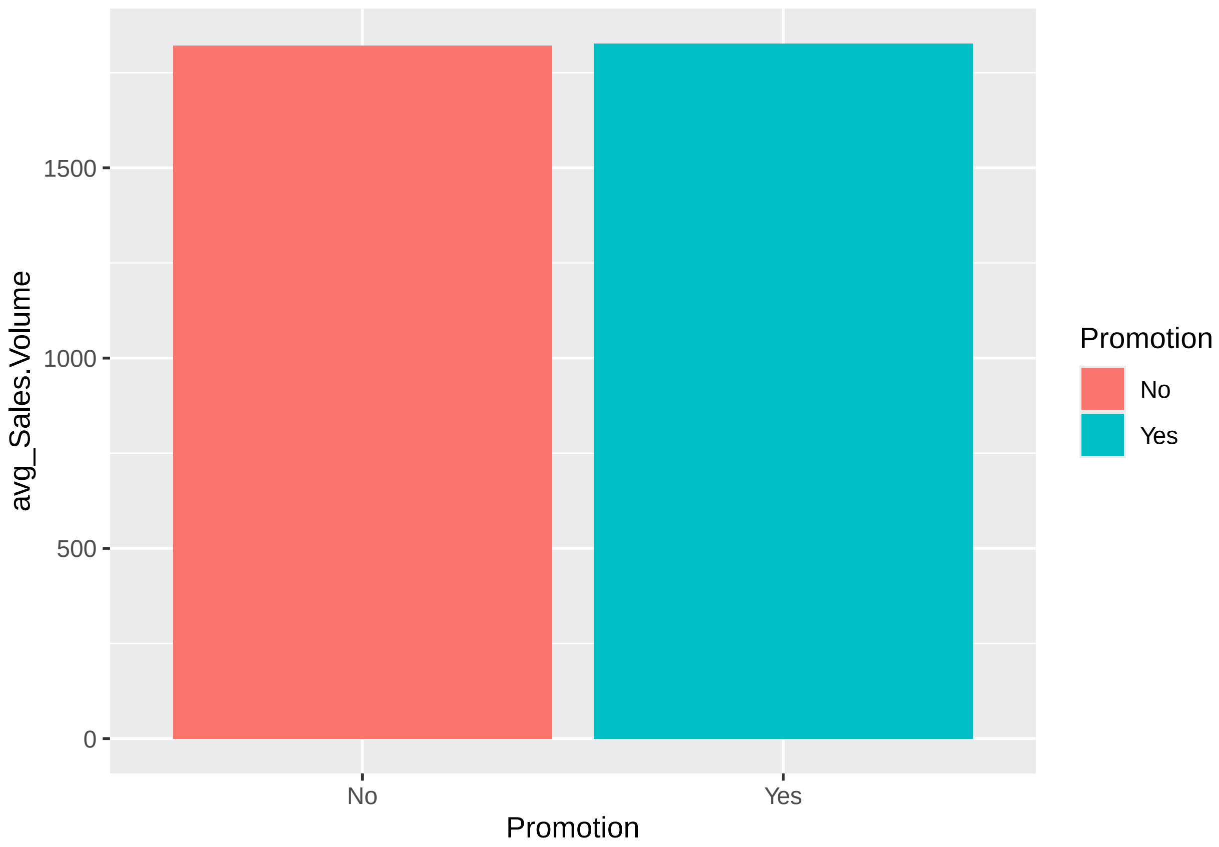


Average price of male products is considerably higher.

```
data2_promotion2 <- data2 %>% group_by(Promotion) %>% summarize(N = n(), avg_Sales.Volume = round(mean(
data2_promotion2
```

```
## # A tibble: 2 x 3
##   Promotion     N avg_Sales.Volume
##   <chr>     <int>            <dbl>
## 1 No          132             1821
## 2 Yes         120             1827
```

```
ggplot(data2_promotion2,aes(x=Promotion, y= avg_Sales.Volume,fill = Promotion)) + geom_col()
```

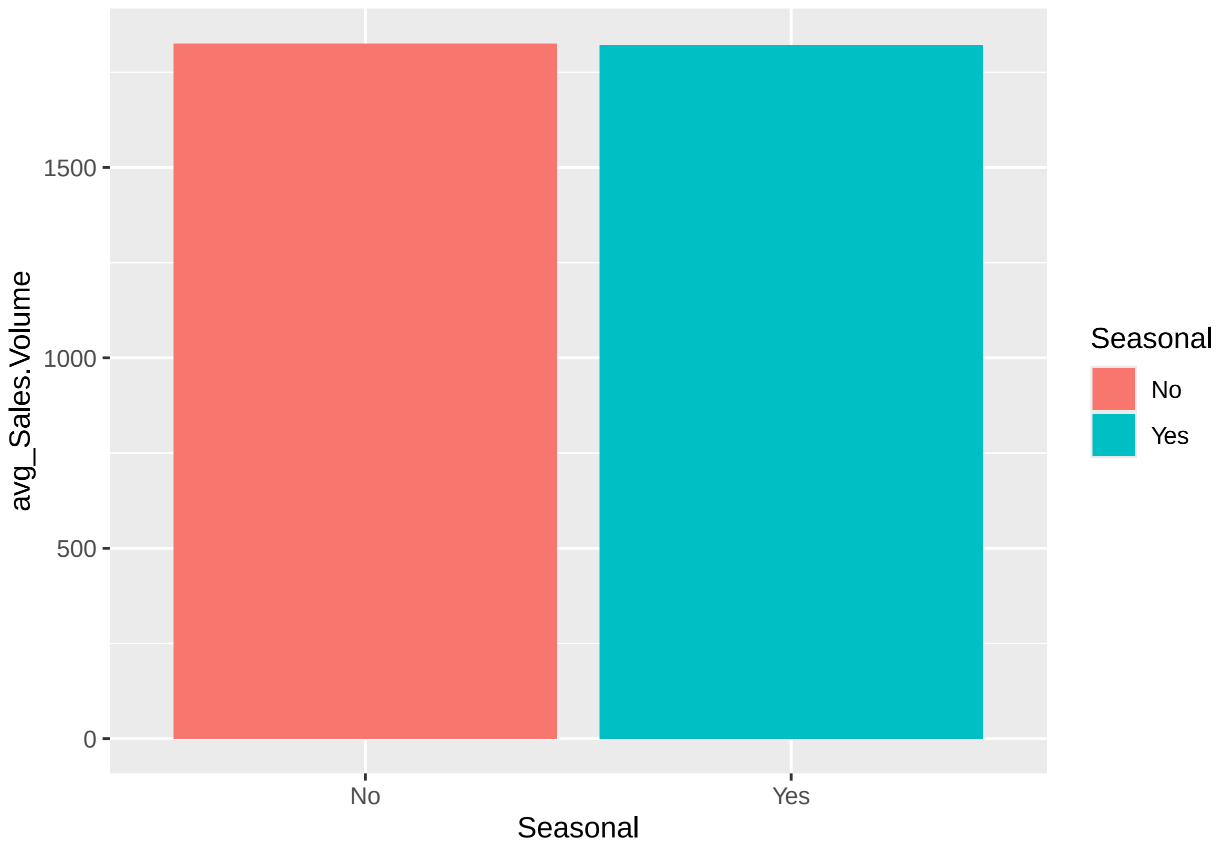


Very small difference in average sales volume.

```
data2_Seasonal2 <- data2 %>% group_by(Seasonal) %>% summarize(N = n(), avg_Sales.Volume = round(mean(Sa
data2_Seasonal2
```

```
## # A tibble: 2 x 3
##   Seasonal      N avg_Sales.Volume
##   <chr>     <int>            <dbl>
## 1 No          124             1826
## 2 Yes         128             1822
```

```
ggplot(data2_Seasonal2, aes(x=Seasonal,y=avg_Sales.Volume, fill = Seasonal)) + geom_col()
```



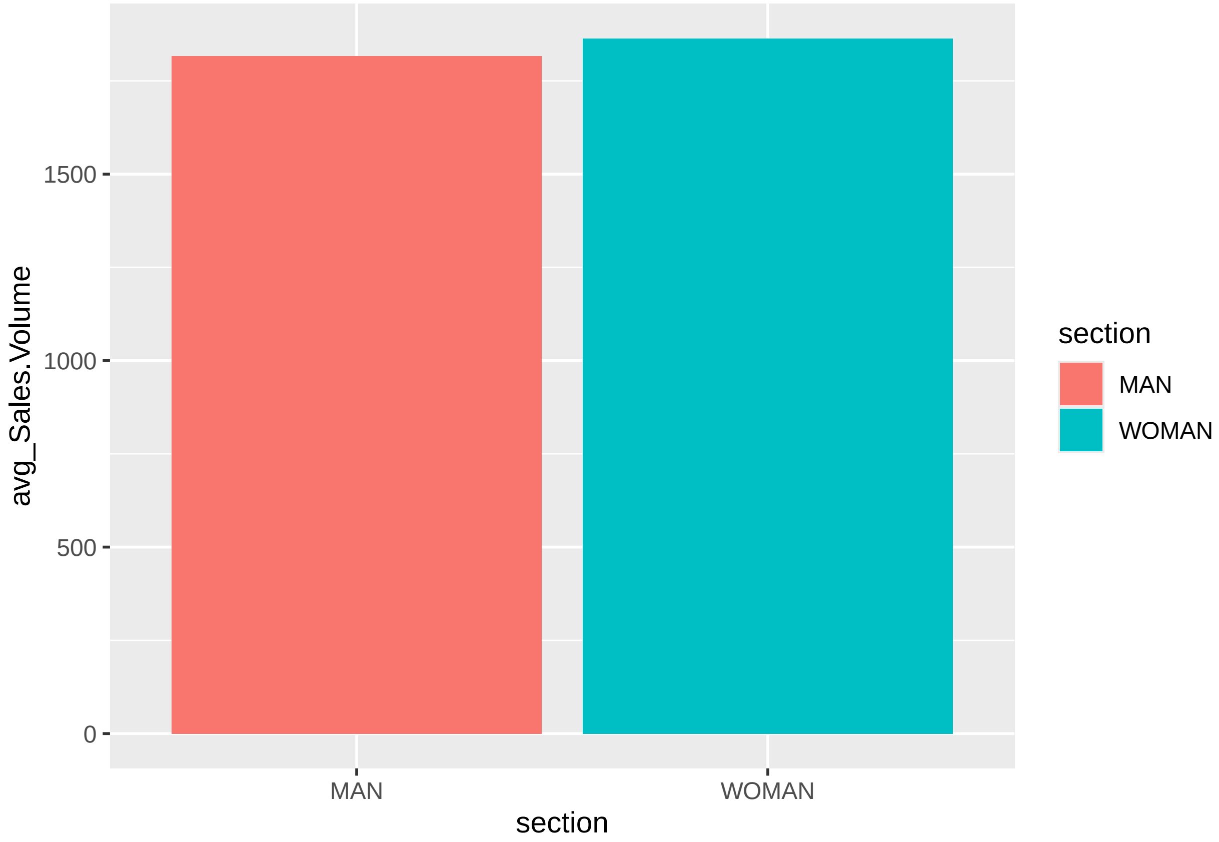
Very similar in average sales volume.

Average calculation according to the number of people allocated

```
data2_section2 <- data2 %>% group_by(section) %>%  summarize(N = n(), avg_Sales.Volume = round(mean(Sal
data2_section2
```

```
## # A tibble: 2 x 3
##   section     N avg_Sales.Volume
##   <chr>   <int>            <dbl>
## 1 MAN       218             1817
## 2 WOMAN      34             1864
```

```
ggplot(data2_section2, aes(x=section,y=avg_Sales.Volume, fill = section)) + geom_col()
```



The average sales volume for men and women are similar, but shows women buy more because the number of women is smaller.

3. Price & sales.volume Prediction

Modeling

Linear regression

```
md_lr <- lm(price ~Promotion + Seasonal+ section + Sales.Volume ,data=data)
```

```
summary(md_lr)
```

```
##
## Call:
## lm(formula = price ~ Promotion + Seasonal + section + Sales.Volume,
##     data = data)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -84.64 -28.77  -5.37  17.67 333.92
##
## Coefficients:
```

```
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)   93.542277   9.879440   9.468   <2e-16 ***
## Promotion     12.331567   6.309221   1.955   0.0518 .
## Seasonal       2.779143   6.340179   0.438   0.6615
## section      -20.928691   4.640951  -4.510    1e-05 ***
## Sales.Volume  -0.004895   0.004526  -1.081   0.2805
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 50.01 on 247 degrees of freedom
## Multiple R-squared:  0.09266,    Adjusted R-squared:  0.07796
## F-statistic: 6.306 on 4 and 247 DF,  p-value: 7.58e-05
```

```
step(md_lr,direction = "backward")
```

```
## Start:  AIC=1976.73
## price ~ Promotion + Seasonal + section + Sales.Volume
##
##                Df Sum of Sq    RSS    AIC
## - Seasonal      1       481 618270 1974.9
## - Sales.Volume  1      2925 620715 1975.9
## <none>                      617789 1976.7
## - Promotion     1      9555 627344 1978.6
## - section       1     50864 668654 1994.7
##
## Step:  AIC=1974.92
## price ~ Promotion + section + Sales.Volume
##
##                Df Sum of Sq    RSS    AIC
## - Sales.Volume  1      2938 621208 1974.1
## <none>                      618270 1974.9
## - Promotion     1      9549 627819 1976.8
## - section       1     50392 668662 1992.7
##
## Step:  AIC=1974.12
## price ~ Promotion + section
##
##             Df Sum of Sq    RSS    AIC
## <none>                   621208 1974.1
## - Promotion  1      9504 630712 1975.9
## - section    1     50973 672182 1992.0
```

```
##
## Call:
## lm(formula = price ~ Promotion + section, data = data)
##
## Coefficients:
## (Intercept)    Promotion      section
##       86.01        12.30       -20.82
```

Select variables

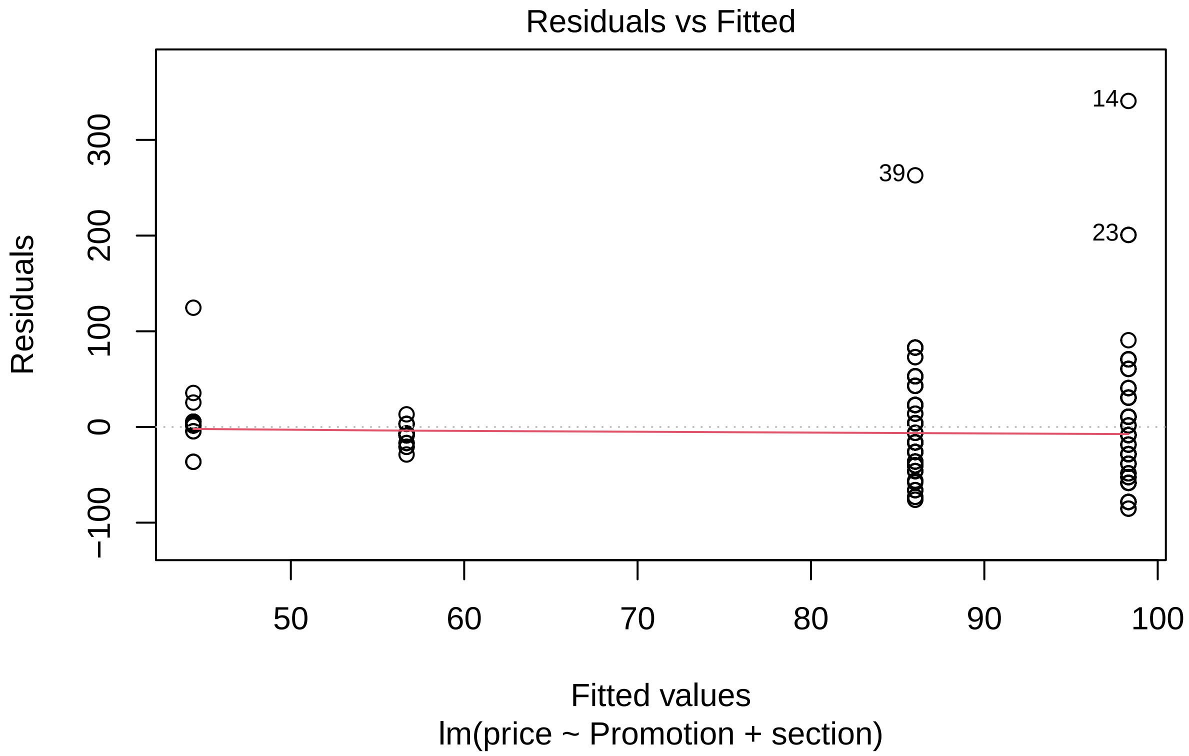
```
md_lr <- lm(price ~ Promotion + section , data = data)
```

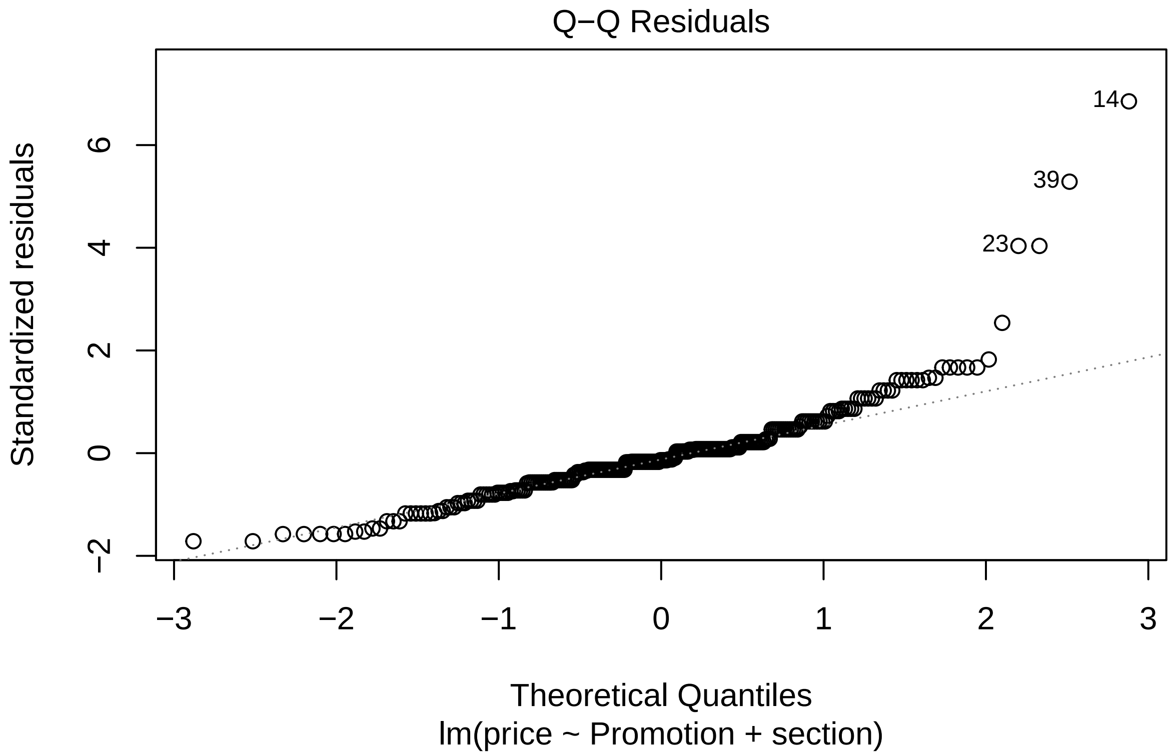
```
summary(md_lr)
```

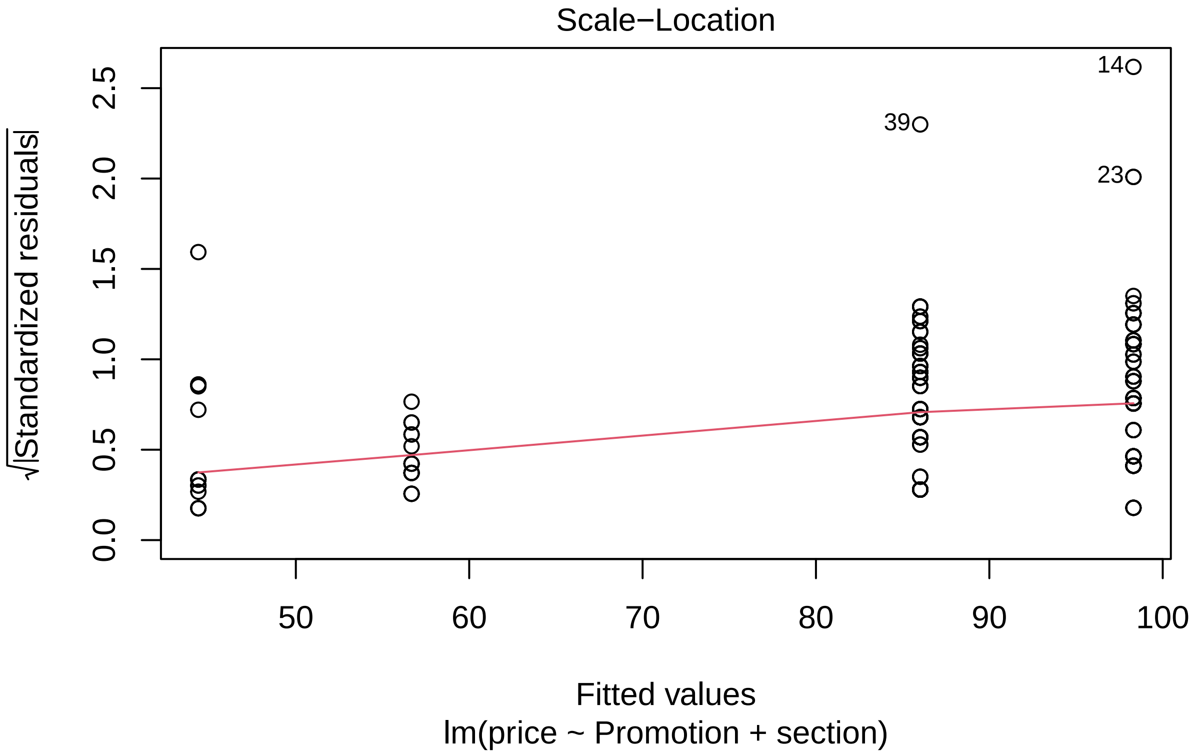
```
## 
## Call:
## lm(formula = price ~ Promotion + section, data = data)
## 
## Residuals:
##    Min     1Q Median     3Q    Max 
## -85.32 -28.34  -6.78  16.16 340.69 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   86.014      4.506  19.087  < 2e-16 ***
## Promotion     12.298      6.301   1.952   0.0521 .  
## section      -20.819      4.606  -4.520 9.56e-06 ***
## ---
## Signif. codes:  0 ’***’ 0.001 ’**’ 0.01 ’*’ 0.05 ’.’ 0.1 ’ ’ 1
## 
## Residual standard error: 49.95 on 249 degrees of freedom
## Multiple R-squared:  0.08764,    Adjusted R-squared:  0.08031 
## F-statistic: 11.96 on 2 and 249 DF,  p-value: 1.099e-05
```

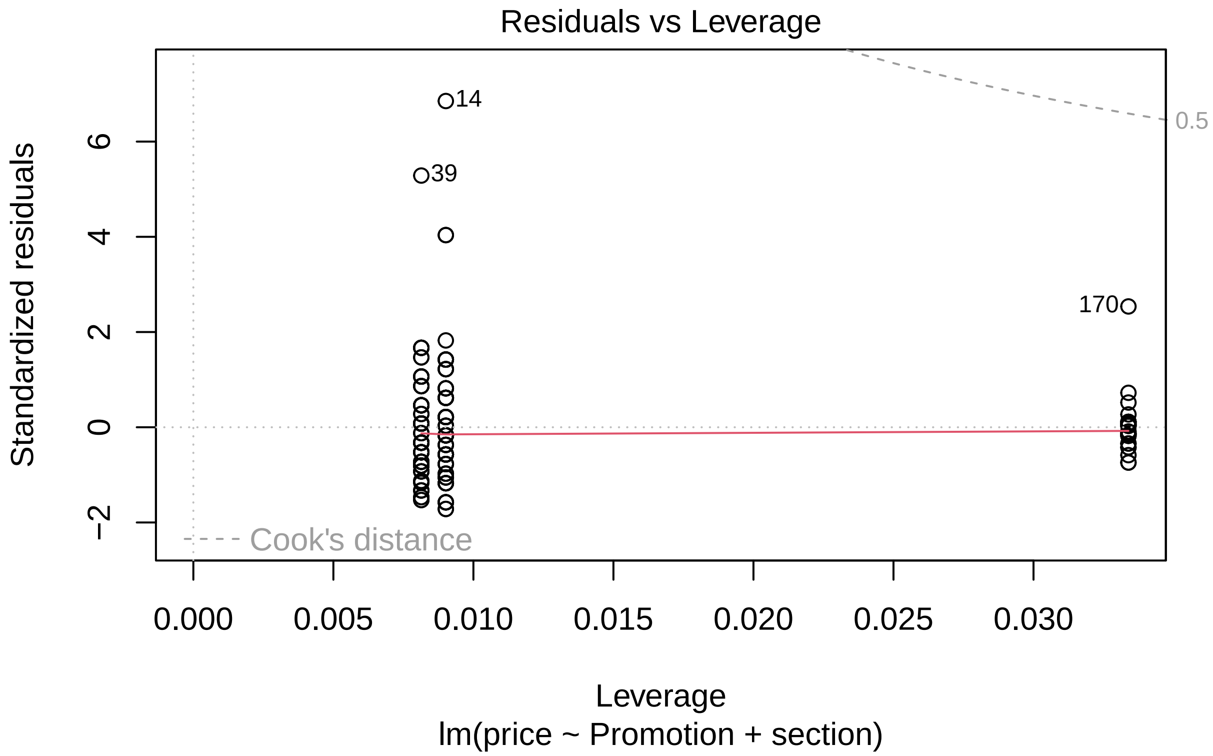
Regression equation: 86.014 + Promotion * 12.298 + section * -20.819

```
plot(md_lr)
```


Residuals vs Fitted
Residuals
300
200
100
0
−100
50
60
70
80
90
100
Fitted values
lm(price ~ Promotion + section)
14
39
23


Q−Q Residuals
Standardized residuals
−2
0
2
4
6
−3
−2
−1
0
1
2
3
14
39
23
Theoretical Quantiles
lm(price ~ Promotion + section)


Scale−Location
√|Standardized residuals|
0.0
0.5
1.0
1.5
2.0
2.5
50
60
70
80
90
100
14
39
23
Fitted values
lm(price ~ Promotion + section)



```
pred <- 86.014 + data$Promotion * 12.298 + data$section * -20.819
```

```
accuracy(data$price,pred)
```

```
##                       ME     RMSE      MAE        MPE     MAPE
## Test set -0.0001746032 49.64989 32.74035 -0.1927524 37.47007
```

```
md_lr2 <- lm(Sales.Volume ~  price + Promotion + Seasonal+ section  ,data=data)
```

```
summary(md_lr2)
```

```
##
## Call:
## lm(formula = Sales.Volume ~ price + Promotion + Seasonal + section,
##     data = data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1352.24  -587.60    21.54   557.96  1202.32
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1899.7760   107.5139  17.670   <2e-16 ***
## price         -0.9629     0.8904  -1.081    0.281
## Promotion     17.8684    89.1685   0.200    0.841
## Seasonal      -4.8806    88.9623  -0.055    0.956
## section        3.5204    67.7210   0.052    0.959
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 701.5 on 247 degrees of freedom
## Multiple R-squared:  0.005279,   Adjusted R-squared:  -0.01083
## F-statistic: 0.3277 on 4 and 247 DF,  p-value: 0.8592
```

The regression model that predicts Sales.Volume is not statistically significant.