Red Line Linear Regression Plot

Miguel Angel Bustos

06-01-2022

INTRODUCTION

Red Line Linear Regression Plot, it´s a plot that shows the regression line of a cafe, at the end of this project, it´s a conclusion about how the total will increments if the price of products goes up. "Item price" it´s the price of any beverage and "total" it´s the price of that bill or sales. And colour points are the different kind of coffe, juice or beverages

DATA

Different coffe in your cafe

```
df <- data.frame(cafe_id = c(1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2,
                             2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 3),
                 timestamp = structure(
                   c(1596283511, 1596283511, 1596287730, 1596287730, 1596287730,
                     1596370505, 1596378076, 1596386672, 1596386672, 1596283999,
                     1596283999, 1596283999, 1596291727, 1596368783, 1596368783,
                     1596369685, 1596285196, 1596285327, 1596285327, 1596286869,
                     1596291836, 1596293731, 1596293731, 1596370648, 1596370648),
                   class = c("POSIXct", "POSIXt"), tzone = "UTC"),
                 invoice_id = c(1, 1, 2, 2, 2, 3, 4, 5, 5, 1, 1, 1, 2,
                                3, 3, 4, 1, 2, 2, 3, 4, 5, 5, 6, 6),
                 item = c("coffee", "latte macchiato", "cake", "cappuchino",
                          "tea", "coffee", "cake", "espresso", "orange juice",
                          "tea", "bagel", "espresso", "coffee", "ice cream",
                          "espresso", "cake", "latte macchiato", "cake",
                          "cappuchino", "bagel", "espresso", "tea", "cake",
                          "coffee", "bagel"),
                 quantity= c(1, 1, 2, 2, 1, 2, 1, 2, 1, 2, 3, 1, 1,
                             2, 3, 1, 1, 1, 1, 1, 2, 1, 1, 2, 2),
                 item_price = c(2, 3, 3, 2.5, 2.5, 2, 3, 1.5, 2.6,
                                2.5, 3, 1.5, 2, 2.3, 1.5, 3, 3, 3,
                                2.5, 3, 1.5, 2.5, 3, 2, 3),
                 total = c(2, 3, 6, 5, 2.5, 4, 3, 3, 2.6, 5, 9, 1.5, 2,
                           4.6, 4.5, 3, 3, 3, 2.5, 3, 3, 2.5, 3, 4, 6))
```

The first lines of the data frame

```
head(df)
```

```
  cafe_id           timestamp invoice_id            item quantity item_price
1       1 2020-08-01 12:05:11          1          coffee        1        2.0
```

```
2       1 2020-08-01 12:05:11          1 latte macchiato        1        3.0
3       1 2020-08-01 13:15:30          2            cake        2        3.0
4       1 2020-08-01 13:15:30          2      cappuchino        2        2.5
5       1 2020-08-01 13:15:30          2             tea        1        2.5
6       1 2020-08-02 12:15:05          3          coffee        2        2.0
  total
1   2.0
2   3.0
3   6.0
4   5.0
5   2.5
6   4.0
```

LINEAR MODEL

Let´s see pearson correlation

```
cor(df$total, df$item_price, method = c("pearson"))
```

```
[1] 0.2886805
```

Linear model, linear model of total according to bevarage.

```
li <- lm(total ~ item_price, data=df)
li
```

```
Call:
lm(formula = total ~ item_price, data = df)

Coefficients:
(Intercept)   item_price
     1.5553       0.8439
```

```
summary(li)
```

```
Call:
lm(formula = total ~ item_price, data = df)

Residuals:
   Min     1Q Median     3Q    Max
-1.321 -1.149 -1.087  1.104  4.913

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.5553     1.4686   1.059    0.301
item_price    0.8439     0.5836   1.446    0.162

Residual standard error: 1.598 on 23 degrees of freedom
Multiple R-squared:  0.08334,   Adjusted R-squared:  0.04348
F-statistic: 2.091 on 1 and 23 DF,  p-value: 0.1617
```

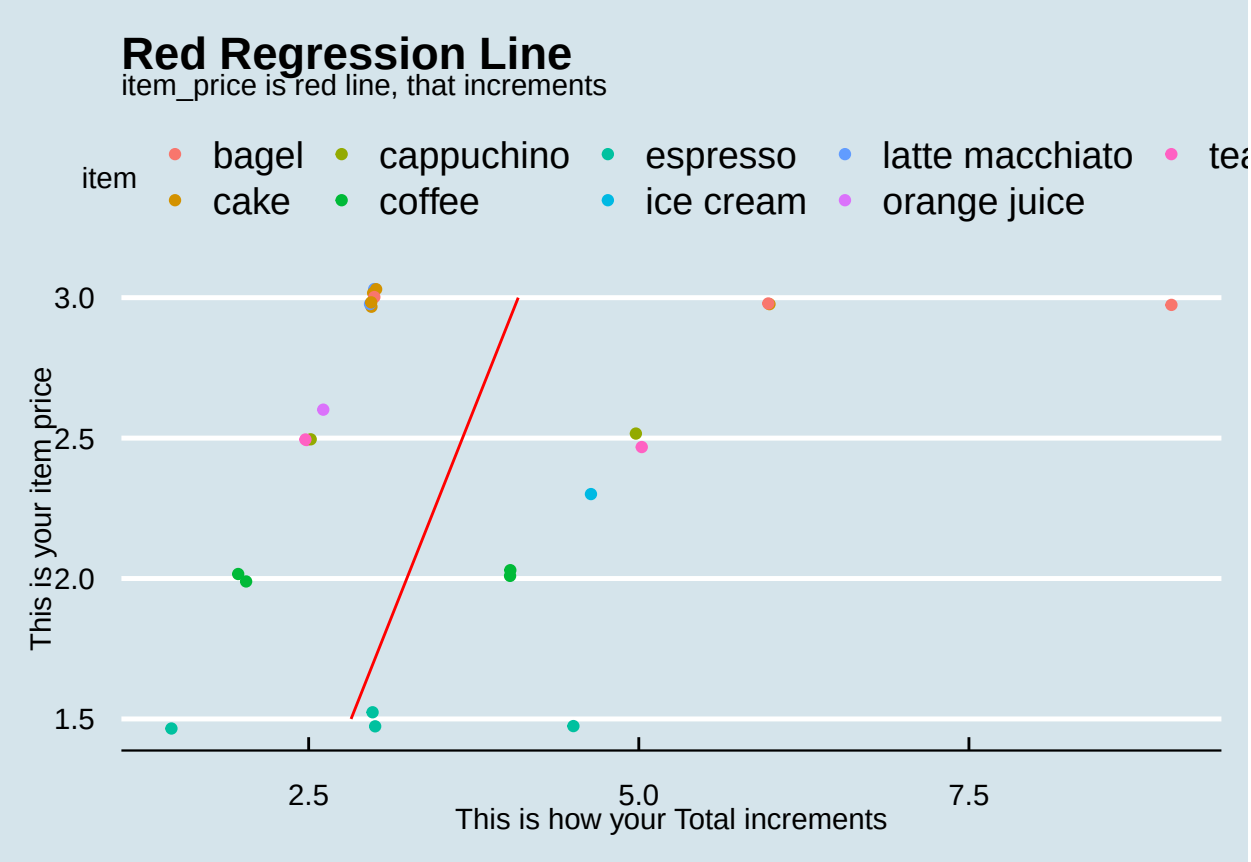
Data frame with the variable you want to plot against item_price

```
predic <- data.frame(yourprediction = predict(li, df), itemprice=df$item_price)
```

SCATTER PLOT LINEAR REGRESSION

predicted line of linear regression

```
ggplot(data = df, aes(x = total, y = item_price)) +
  geom_point(position = "jitter", aes(color = item))+
  geom_line(color='red',data = predic, aes(x=yourprediction, y=itemprice))+
  ylab("This is your item price")+
  xlab("This is how your Total increments")+
  ggtitle("Red Regression Line", subtitle = "item_price is red line, that increments")+
  theme_economist()
```



CONCLUSION

In conclusion, the total will increment, if you increment the price, but your total don´t increment too much as x_axis shows