Data Exericse CH7 Q3

Fasih Atif

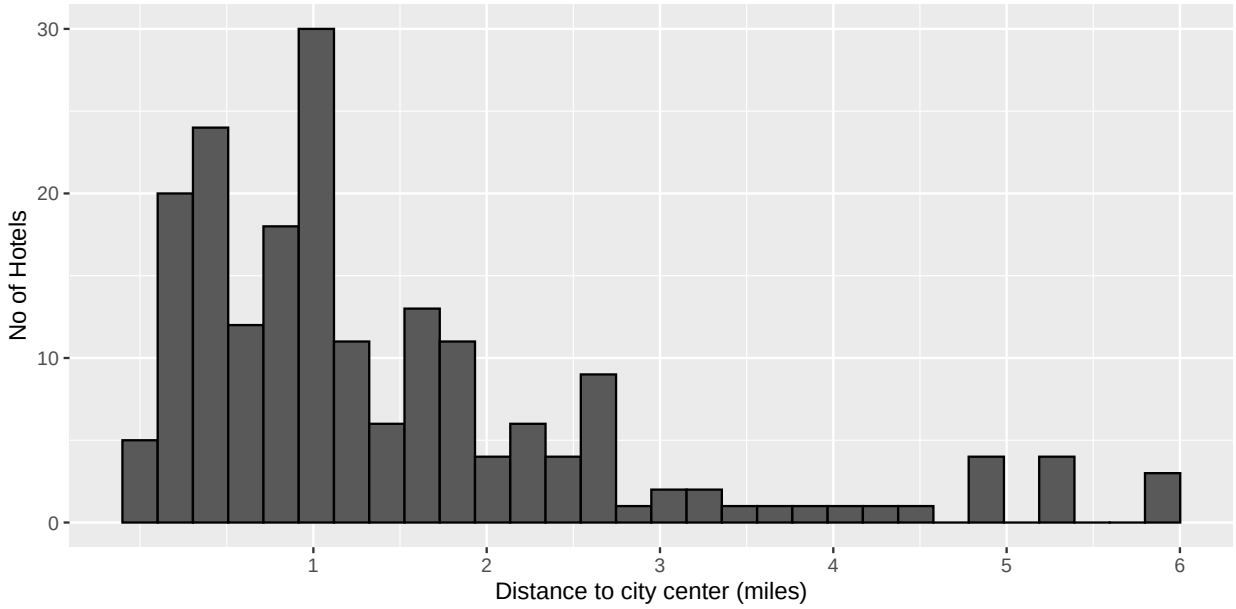
12/2/2020

Introduction

Our task is to analyze the hotel price-distance to the center pattern for another city and compare it to hotels in Austria. We will keep hotels with 3 to 4 stars for a November 2017 weekday. The city that i chose to compare with Austria is Amsterdam, Netherlands.

Examining the distribution of the distance variable

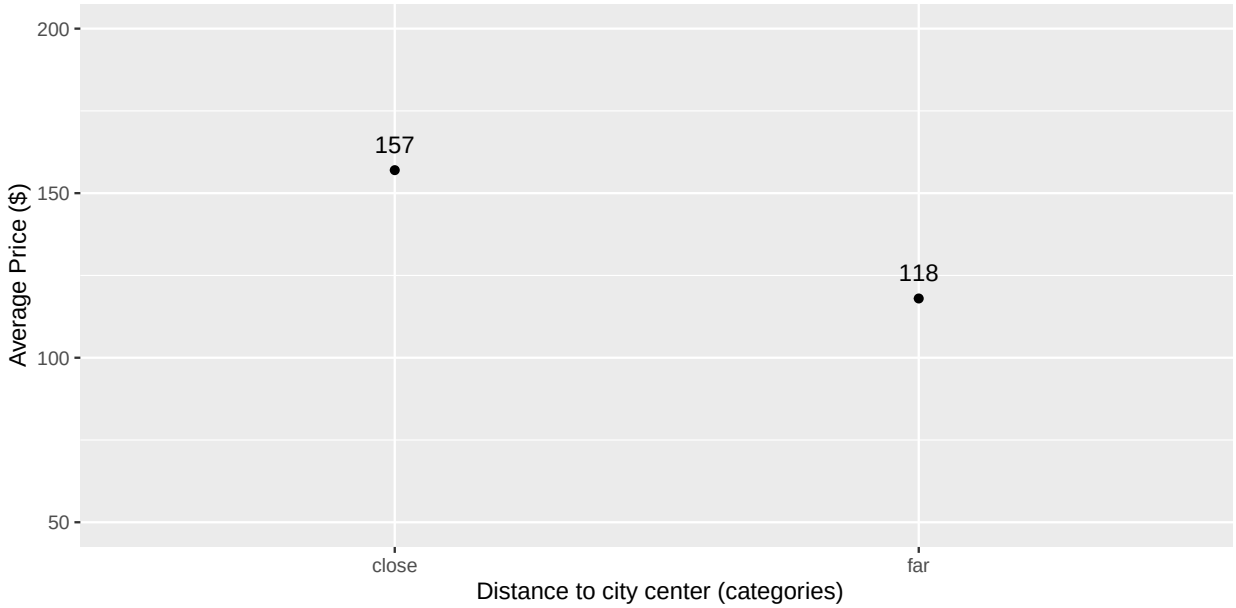
We visualize the distribution of distance of hotels in Amsterdam and observe that majority of the hotels for the criteria we defined are located within a 2 mile radius of the city center. We dont see anything out of place so we wont be dropping any observations.



Non-Parametric Regression

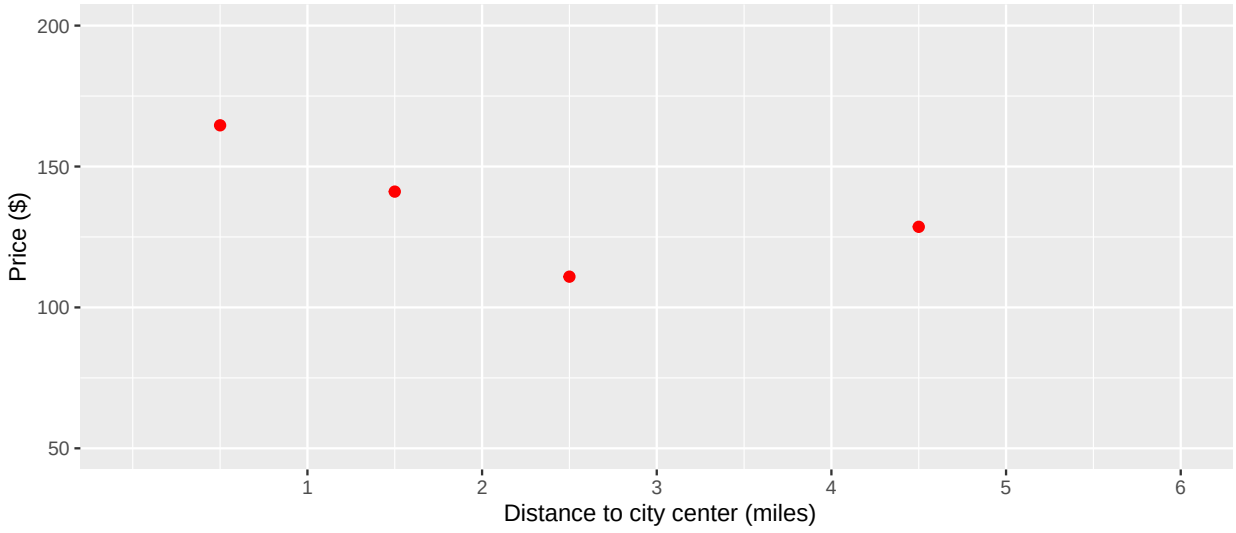
We will start off with a simple comparison and split the Amsterdam hotel data into two categories "close" and "far". Close category hotels consist of those hotels which are less than 2.0 miles away from the city center while any hotel above 2.0 miles is categorized as far hotel. These two categories will act as two bins. Differentiation has been based on the numeric quantity '2.0' miles since this is the distance that has been assumed by the author of the book for hotels in Austria. Author has taken the assumption that a couple is fine walking 2 miles for 'close' hotels.

The 'close' bin contains 150 hotels and the 'far' bin contains 45 hotels. Price comparison for the two bins is shown below:



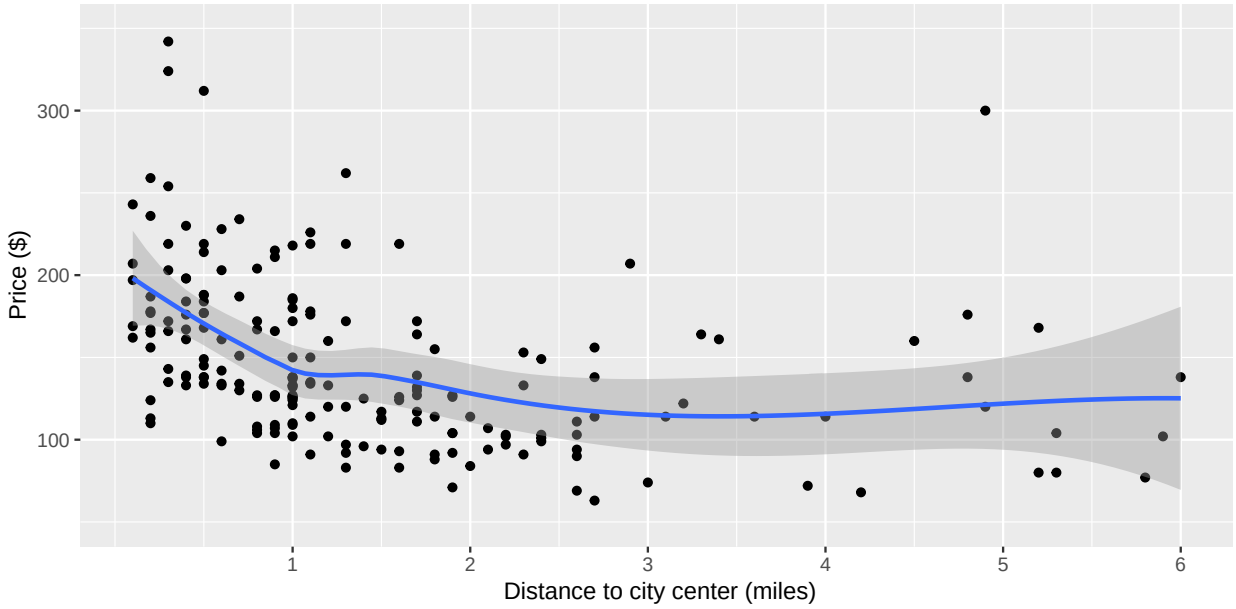
The hotels less than 2.0 miles away cost an average of $157 compared to Austria's $116 while those more than 2.0 miles away cost an average of $118 compared to Austria's $90. Austria's hotels are on average much cheaper than Amsterdam.

Now we will create more refined bins splitting the 2 bins into 4 bins. These 4 bins have the ranges of [0,1],[1,2],[2,3],a nd [3,7]. We will visualize a bin scatter to show the average prices for each bin:



This regression shows that the hotels close to city center are more expensive on average. Beyond 4 miles, the average price in the 4-7 miles bin rises. The relationship appears to be monotonic but non-linear. We see a larger negative difference between the [0,1] miles and the [1,2] miles than between adjacent bins at higher distances.

After bin scatters, we look at a smoothed non parametric loess regression:



We can see a general negative trend in price as we move away from the city center. The curve follows a nearly identical trend to that of shown for Austrian hotels. In both loess regression graphs, There is a steep decline in prices till 1 mile after which the curve flattens out till 4 miles after which it it rises slightly a bit. This smooth non parametric regression is a corresponding version of the binwidth scatter plot.

Linear Regression

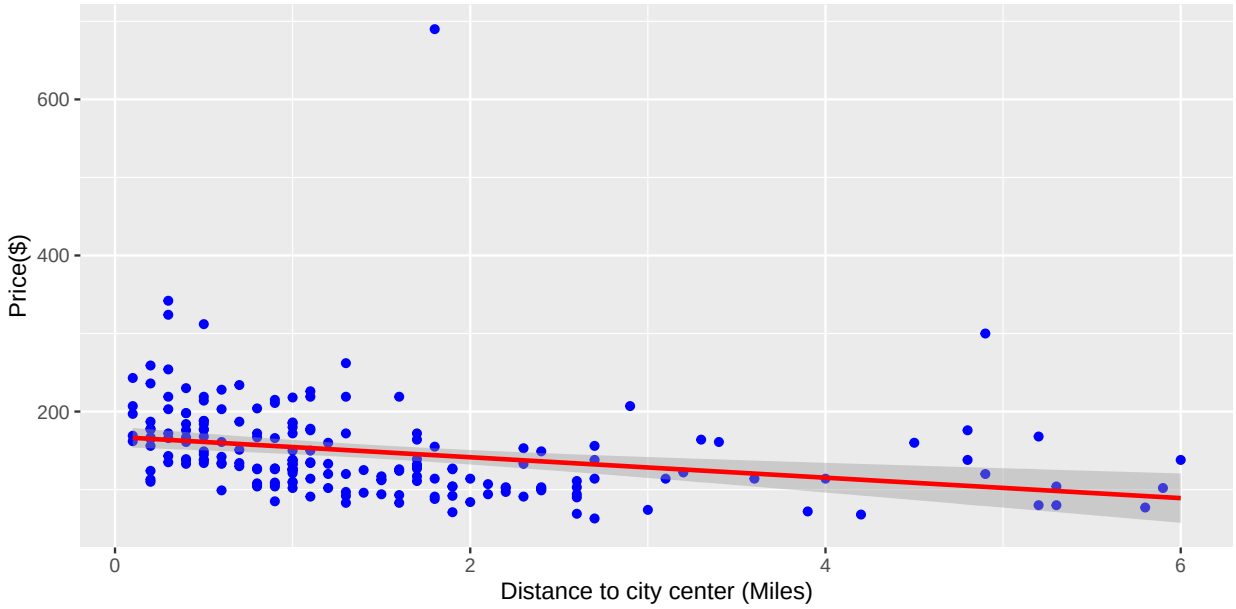
Non parametric regressions dont produce readily interpretable quantitative answers. Hence, we carry out a parametric linear regression:

```
##
## Call:
## lm(formula = price ~ distance, data = df_ams)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -71.74 -33.36 -16.52  22.40 545.95
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  167.614      6.656  25.183  < 2e-16 ***
## distance     -13.092      3.406  -3.844 0.000164 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 61.02 on 193 degrees of freedom
## Multiple R-squared:  0.07111,    Adjusted R-squared:  0.0663
## F-statistic: 14.77 on 1 and 193 DF,  p-value: 0.0001644
```

Interpretation

Hotels that are 1 mile further away from the city center are on average $13 cheaper in Amsterdam, considering all other factors are held constant. The intercept is $167.6 which tells us that the average price of a hotel situated exactly in the city center would be $167.6. It is to be kept in mind that in reality a hotel might not be present exactly at the city center but can be present a few steps away.

In comparison, Austrian hotels are $14 cheaper on average for every mile away from city center.



Conclusion

We wanted to study the relationship between weekday price and distance for hotels that are rated as 3.0 to 4.0 stars. came to the conclusion that although hotels in both cities follow a similar trend in price as we move away from the city center, hotels in Amsterdam are on average more expensive than the hotels in Austria.