

QTM 150 - Final Project

Jessica Kang

December 10th, 2018

**Does more expensive food mean better food?
Does a price increase from lunch to dinner lead to
a ratings increase?**

Why these questions?

Since the semester started, I've gone out to eat multiple times. As college students on a budget, my friends and I try to find restaurants that have a good balance between quality of food and price each time we eat out. Thus, I was curious whether restaurants with more expensive food and overall higher price ranges have better ratings. Should we choose a more expensive place to eat at in order to get better food or can we choose a cheaper place without sacrificing food quality? In addition, my friends and I usually go out for dinner, but dinner tends to be more expensive than lunch. I wanted to find out whether this increase in the price of food from lunch to dinner meant higher dinner than lunch ratings. Is it worth it to go to a restaurant for dinner or should I just go for lunch?

Section 1: Original Data - Restaurant Survey Responses: Average Price vs Rating

For the first section, I sent out a survey to my friends asking them to list at least two of their favorite restaurants (excluding fast food places) and what they would rate each restaurant out of 5. I received 13 responses with 27 unique restaurants. For restaurants that were repeated, I took the average of the ratings. I then looked up the menu for each of these 27 restaurants and found the average of their cheapest and most expensive entrees to determine the average price for each restaurant. I saved this information as the `friendsrest` data frame. I then created a scatterplot to visualize the data and calculated R and R^2 to determine the strength of the relationship between average price and rating.



Food from the *Rustic Inn Restaurant*, the restaurant with the highest average price out of my friends' responses

Data Frame

The data frame was created on Excel using my friends' survey responses. The data frame includes the following: restaurant name, rating, lowest priced entree, highest priced entree, and average price.

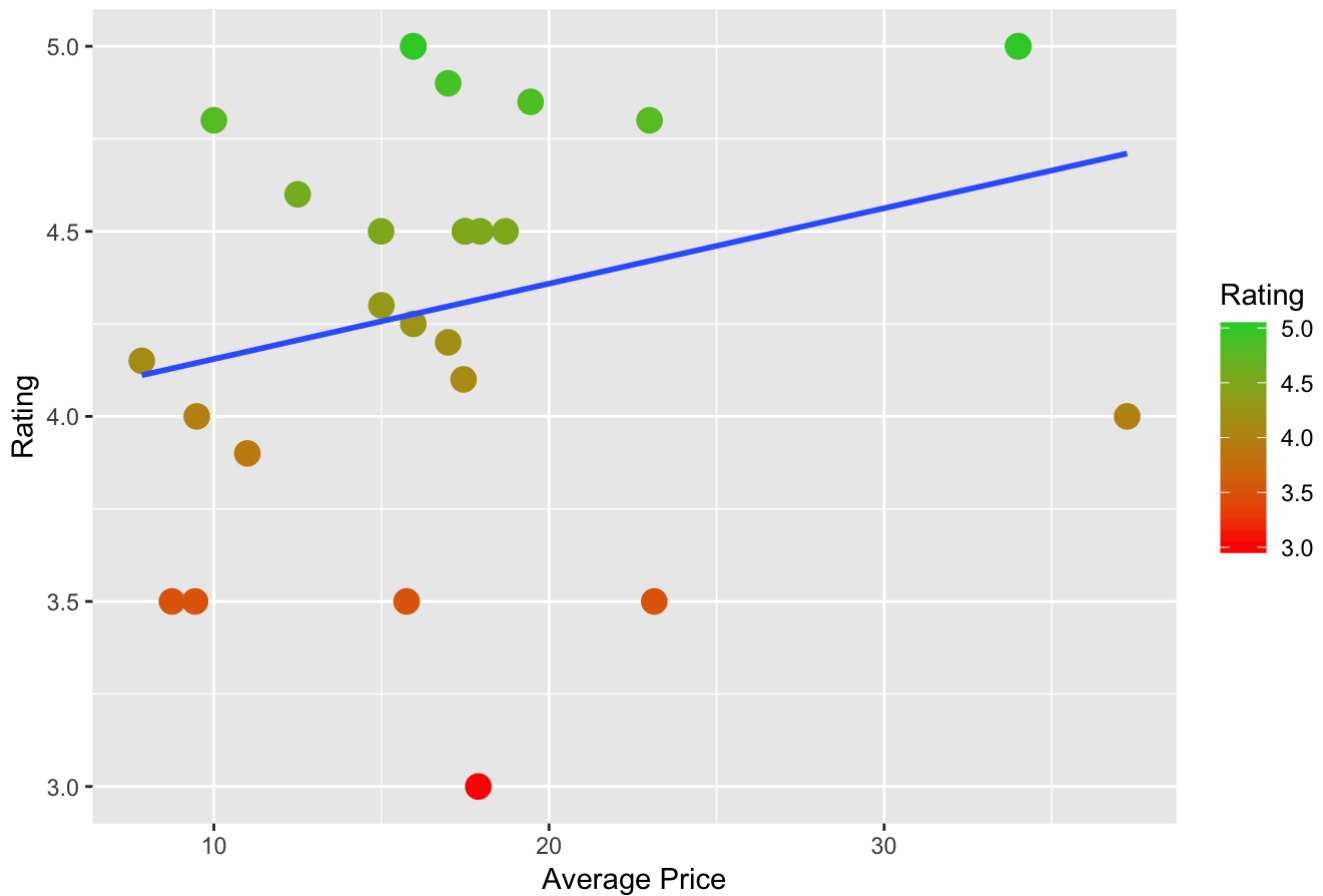
```
friendsrest <- read.csv("https://github.com/jkang4/QT-150-Final-Project/blob/master/Favorite%20Restaurants%20-%20Responses.csv?raw=true")
```

Data Visualization

I created a scatterplot of rating vs average price with a best fit linear regression model.

```
ggplot(friendsrest, aes(x=Average_Price, y=Rating))+  
  geom_point(size=4, aes(color=Rating))+  
  scale_color_gradient(name="Rating", low="red", high="#32cd32")+  
  labs(x="Average Price", y="Rating", title="Friends' Restaurant Ratings vs Average Food  
Price")+  
  theme(plot.title = element_text(hjust = 0.5))+  
  geom_smooth(method="lm", se = FALSE)
```

Friends' Restaurant Ratings vs Average Food Price



```
lm(friendsrest$Average_Price ~ friendsrest$Rating)
```

```
##
## Call:
## lm(formula = friendsrest$Average_Price ~ friendsrest$Rating)
##
## Coefficients:
##      (Intercept)  friendsrest$Rating
##           1.830           3.653
```

Correlation (R):

The correlation between average price and rating of the restaurants from the survey is 0.2728, which indicates a weak, positive linear relationship.

```
friendsrest %>%
  summarize(r_1 = cor(Average_Price, Rating))
```

```
##           r_1
## 1 0.2728147
```

Coefficient of Determination (R^2):

The percentage in variation of the data that is explained by the linear model is only %7.44, which indicates that the linear model is not a good fit.

```
lm_fr <- lm(Average_Price ~ Rating, data = friendsrest)
summary(lm_fr)
```

```
##
## Call:
## lm(formula = Average_Price ~ Rating, data = friendsrest)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.3637 -4.6102 -0.7779  0.8892 20.8086
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.830      11.191   0.164   0.871
## Rating         3.653       2.576   1.418   0.169
##
## Residual standard error: 7.321 on 25 degrees of freedom
## Multiple R-squared:  0.07443,    Adjusted R-squared:  0.0374
## F-statistic:  2.01 on 1 and 25 DF,  p-value: 0.1686
```

```
summary(lm_fr)$r.squared
```

```
## [1] 0.07442787
```

Section 2: Experimental Procedures and Data Collection

For this section, I used an external dataset I found of Kyoto Restaurant Reviews (<https://www.kaggle.com/koki25ando/tabelog-restaurant-review-dataset>) on Kaggle, which included information for 895 restaurants in Kyoto. I then generated a random sample of 80 restaurants that have numerical (non-NA) values for dinner and lunch prices as well as dinner and lunch ratings (results are hidden).

```
Kyoto_Restaurant_Info <- read.csv("https://github.com/jkang4/QTm-150-Final-Project/blob/master/Kyoto_Restaurant_Info.csv?raw=true")
```

```
rest <- Kyoto_Restaurant_Info[complete.cases(Kyoto_Restaurant_Info$LunchPrice, Kyoto_Restaurant_Info$DinnerPrice, Kyoto_Restaurant_Info$DinnerRating, Kyoto_Restaurant_Info$LunchRating),]  
str(rest)  
  
set.seed(12)  
rest1<-rest[sample(1:nrow(rest),80, replace=FALSE),]  
rest1  
str(rest1)
```

Part 1: Does More Expensive Food Mean Better Food? (Average Prices vs Overall Rating)

Part 1 of this section looks at the relationship between average prices and ratings for the sample of 80 Japanese restaurants. After changing the prices from factor to numeric vectors, I replaced the price intervals with the interval midpoints in the new `dinnerprice` and `lunchprice` numeric vectors. The average price for each of the 80 restaurants was calculated by taking the mean of each restaurant's respective lunch and dinner prices. I then created a data frame that included average price and total rating.



Food from *Hiroya* ひろや, the most expensive restaurant in the sample



Orto オルト, the top rated restaurant in the sample

Data Frame Creation

Converting dinner price and lunch price to factor vectors:

```
dinnerprice_factor<-as.factor(rest1$DinnerPrice)
dinnerprice_factor <- factor(rest1$DinnerPrice)
levels(dinnerprice_factor)
```

```
## [1] " ~¥999" " ¥1000~¥1999" " ¥10000~¥14999"
## [4] " ¥15000~¥19999" " ¥2000~¥2999" " ¥20000~¥29999"
## [7] " ¥3000~¥3999" " ¥4000~¥4999" " ¥5000~¥5999"
## [10] " ¥6000~¥7999" " ¥8000~¥9999"
```

```
lunchprice_factor<-as.factor(rest1$LunchPrice)
lunchprice_factor <- factor(rest1$LunchPrice)
levels(lunchprice_factor)
```

```
## [1] " ~¥999" " ¥1000~¥1999" " ¥15000~¥19999"
## [4] " ¥2000~¥2999" " ¥3000~¥3999" " ¥4000~¥4999"
## [7] " ¥5000~¥5999" " ¥6000~¥7999"
```

Converting the dinnerprice_factor and lunchprice_factor to numeric vectors:

```
dinnerprice<-factor(dinnerprice_factor,levels=c("500", "1500", "2500", "3500", "4500",
"5500", "7000", "9000", "12500", "17500", "25000"))
```

```
dinnerprice[dinnerprice_factor==" ~¥999"]<-"500"
dinnerprice[dinnerprice_factor==" ¥1000~¥1999"]<-"1500"
dinnerprice[dinnerprice_factor==" ¥2000~¥2999"]<-"2500"
dinnerprice[dinnerprice_factor==" ¥3000~¥3999"]<-"3500"
dinnerprice[dinnerprice_factor==" ¥4000~¥4999"]<-"4500"
dinnerprice[dinnerprice_factor==" ¥5000~¥5999"]<-"5500"
dinnerprice[dinnerprice_factor==" ¥6000~¥7999"]<-"7000"
dinnerprice[dinnerprice_factor==" ¥8000~¥9999"]<-"9000"
dinnerprice[dinnerprice_factor==" ¥10000~¥14999"]<-"12500"
dinnerprice[dinnerprice_factor==" ¥15000~¥19999"]<-"17500"
dinnerprice[dinnerprice_factor==" ¥20000~¥29999"]<-"25000"
```

```
summary(dinnerprice)
```

```
## 500 1500 2500 3500 4500 5500 7000 9000 12500 17500 25000
## 4 6 9 22 12 6 8 4 7 1 1
```

```
dinnerprice<-as.numeric(as.character(dinnerprice))
is.numeric(dinnerprice)
```

```
## [1] TRUE
```

```
dinnerprice
```

```
## [1] 3500 3500 3500 9000 2500 3500 3500 4500 9000 4500 3500
## [12] 5500 3500 9000 3500 3500 3500 12500 3500 5500 3500 12500
## [23] 3500 5500 17500 1500 7000 12500 2500 3500 4500 3500 7000
## [34] 2500 7000 2500 500 12500 12500 3500 4500 9000 5500 500
## [45] 5500 7000 1500 12500 2500 500 4500 2500 1500 3500 4500
## [56] 2500 3500 1500 4500 1500 12500 2500 4500 1500 4500 7000
## [67] 4500 2500 7000 5500 3500 3500 7000 3500 4500 25000 500
## [78] 7000 4500 3500
```

```
lunchprice<-factor(lunchprice_factor,levels=c("500", "1500", "2500", "3500", "4500", "5500", "7000", "17500"))
```

```
lunchprice[lunchprice_factor==" ~¥999"]<-"500"
lunchprice[lunchprice_factor==" ¥1000~¥1999"]<-"1500"
lunchprice[lunchprice_factor==" ¥2000~¥2999"]<-"2500"
lunchprice[lunchprice_factor==" ¥3000~¥3999"]<-"3500"
lunchprice[lunchprice_factor==" ¥4000~¥4999"]<-"4500"
lunchprice[lunchprice_factor==" ¥5000~¥5999"]<-"5500"
lunchprice[lunchprice_factor==" ¥6000~¥7999"]<-"7000"
lunchprice[lunchprice_factor==" ¥15000~¥19999"]<-"17500"
```

```
summary(lunchprice)
```

```
## 500 1500 2500 3500 4500 5500 7000 17500
## 17 34 12 8 2 4 2 1
```

```
lunchprice<-as.numeric(as.character(lunchprice))
is.numeric(lunchprice)
```

```
## [1] TRUE
```

```
lunchprice
```

```
## [1] 500 1500 2500 2500 500 2500 500 1500 2500 1500 1500
## [12] 1500 1500 4500 3500 1500 3500 4500 1500 1500 1500 5500
## [23] 500 2500 17500 1500 1500 7000 500 500 1500 1500 2500
## [34] 1500 3500 1500 500 5500 3500 1500 2500 1500 1500 500
## [45] 2500 1500 1500 5500 1500 500 2500 500 3500 500 1500
## [56] 500 2500 500 1500 1500 5500 1500 2500 500 1500 1500
## [67] 1500 1500 3500 2500 1500 500 3500 500 1500 7000 500
## [78] 3500 1500 1500
```

Calculating average price for each restaurant:

```
Avg_Price <- (dinnerprice + lunchprice)/2
Avg_Price
```

```
## [1] 2000 2500 3000 5750 1500 3000 2000 3000 5750 3000 2500
## [12] 3500 2500 6750 3500 2500 3500 8500 2500 3500 2500 9000
## [23] 2000 4000 17500 1500 4250 9750 1500 2000 3000 2500 4750
## [34] 2000 5250 2000 500 9000 8000 2500 3500 5250 3500 500
## [45] 4000 4250 1500 9000 2000 500 3500 1500 2500 2000 3000
## [56] 1500 3000 1000 3000 1500 9000 2000 3500 1000 3000 4250
## [67] 3000 2000 5250 4000 2500 2000 5250 2000 3000 16000 500
## [78] 5250 3000 2500
```

Creating data frame:

Data frame with the 80 restaurants' average price and rating.

```
Rating <- rest1$TotalRating

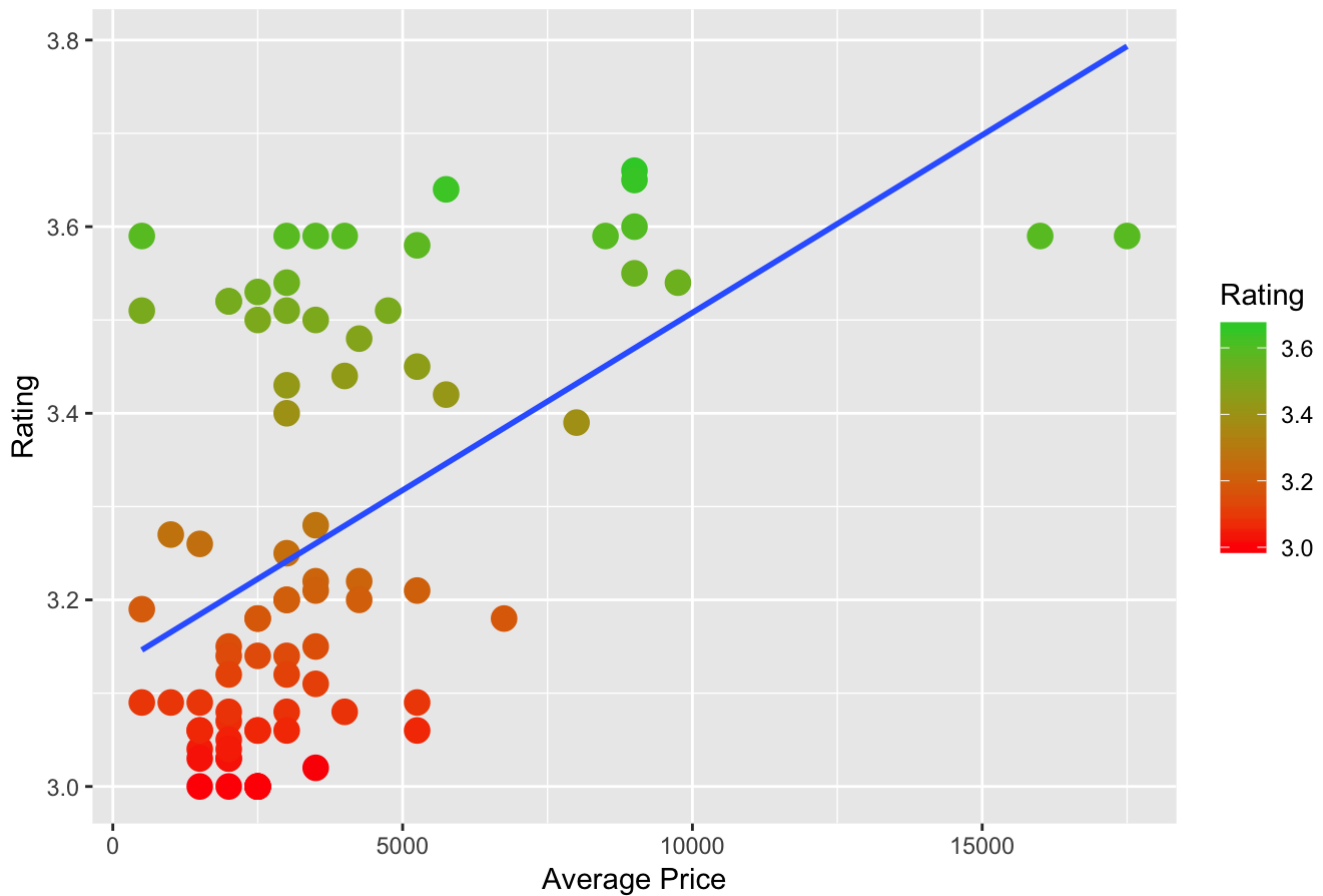
pricevsrating_df <- data.frame(Avg_Price, Rating)
pricevsrating_df
```

Data Visualization

I created a scatterplot of total rating vs average price with a best fit linear regression model. The correlation and coefficient of determination were calculated to be $R=0.5194$ and $R^2=0.2698$.

```
ggplot(pricevsrating_df, aes(x=Avg_Price, y=Rating))+
  geom_point(size=4, aes(color=Rating))+
  scale_color_gradient(name="Rating", low="red", high="#32cd32")+
  labs(x="Average Price", y="Rating", title="Restaurant Ratings vs Average Food Price")+
  theme(plot.title = element_text(hjust = 0.5))+
  geom_smooth(method="lm", se = FALSE)
```


Restaurant Ratings vs Average Food Price



```
lm(Avg_Price ~ Rating)
```

```
##
## Call:
## lm(formula = Avg_Price ~ Rating)
##
## Coefficients:
## (Intercept)      Rating
##      -19432         7089
```

Correlation (R):

```
pricevsrating_df %>%
  summarize(r_2 = cor(Avg_Price, Rating))
```

```
##           r_2
## 1 0.5194219
```

Coefficient of Determination (R^2):

```
lm_pvr <- lm(Avg_Price ~ Rating, data = pricevsrating_df)
summary(lm_pvr)
```

```
##
## Call:
## lm(formula = Avg_Price ~ Rating, data = pricevsrating_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5516.1  -975.4  -285.1   684.9 11483.9
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -19432      4327   -4.491 2.42e-05 ***
## Rating         7089      1320    5.368 7.94e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2578 on 78 degrees of freedom
## Multiple R-squared:  0.2698, Adjusted R-squared:  0.2604
## F-statistic: 28.82 on 1 and 78 DF,  p-value: 7.936e-07
```

```
summary(lm_pvr)$r.squared
```

```
## [1] 0.2697991
```

Analysis Without Average Price Outliers

In this section, I used the 1.5*IQR rule to take out outliers in terms of average price and created a new scatterplot using the new data frame that did not contain these outliers. The recalculated R and R² values are 0.2827 and 0.0799, respectively.

Outliers in Average Price:

The third quartile is 4250, which was added to the value of 1.5*IQR. Outliers for average price are the values where Avg_Price > 7625.

```
quantile(Avg_Price)
```

```
##      0%      25%      50%      75%     100%
##      500     2000     3000     4250    17500
```

```
IQR_AP = IQR(Avg_Price)
IQR_AP
```

```
## [1] 2250
```

```
1.5*IQR_AP
```

```
## [1] 3375
```

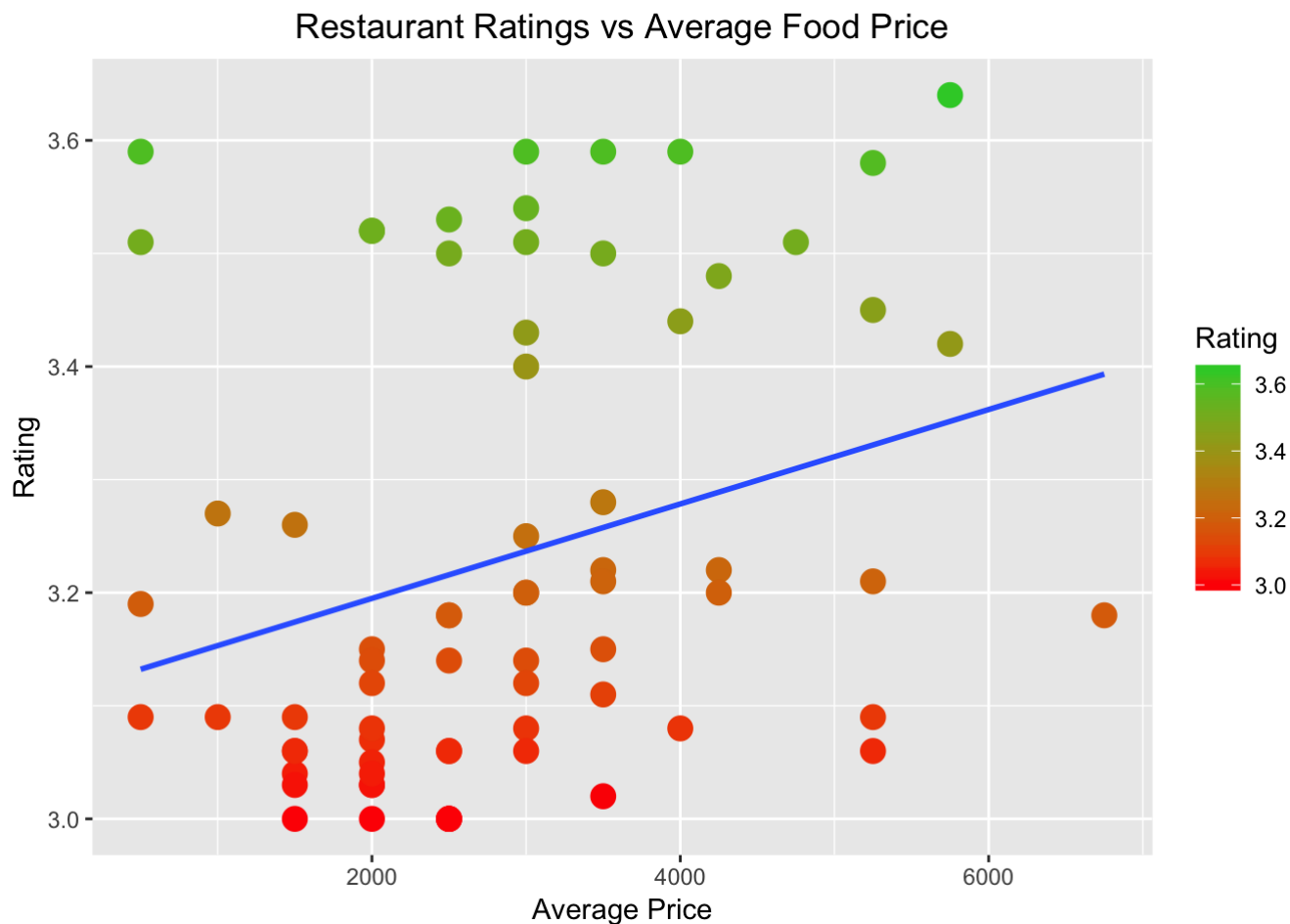
```
4250+1.5*IQR_AP
```

```
## [1] 7625
```

New Data Frame and Visualization:

This new data frame contains all the data points with average price values less than 7625. The new scatterplot shows the values of the `pricevsrating` data frame without outliers.

```
pricevsrating_df2 <- pricevsrating_df %>%  
  filter(Avg_Price <= 7625)  
  
ggplot(pricevsrating_df2, aes(x=Avg_Price, y=Rating))+  
  geom_point(size=4, aes(color=Rating))+  
  scale_color_gradient(name="Rating", low="red", high="#32cd32")+  
  labs(x="Average Price", y="Rating", title="Restaurant Ratings vs Average Food Price")+  
  theme(plot.title = element_text(hjust = 0.5))+  
  geom_smooth(method="lm", se = FALSE)
```



New Correlation (R) and Coefficient of Determination (R^2)

```
pricevsrating_df2 %>%  
  summarize(r_3 = cor(Avg_Price, Rating))
```

```
##           r_3  
## 1 0.2826657
```

```
lm_pvr2 <- lm(Avg_Price ~ Rating, data = pricevsrating_df2)  
summary(lm_pvr2)
```

```
##  
## Call:  
## lm(formula = Avg_Price ~ Rating, data = pricevsrating_df2)  
##  
## Residuals:  
##      Min       1Q   Median       3Q      Max   
## -3059.3  -707.9   -59.3   658.2  3975.1   
##  
## Coefficients:  
##              Estimate Std. Error t value Pr(>|t|)      
## (Intercept)  -3309.2     2530.5  -1.308   0.1953      
## Rating        1913.2       781.6   2.448   0.0169 *     
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 1315 on 69 degrees of freedom  
## Multiple R-squared:  0.0799, Adjusted R-squared:  0.06657   
## F-statistic: 5.992 on 1 and 69 DF,  p-value: 0.01692
```

```
summary(lm_pvr2)$r.squared
```

```
## [1] 0.07989992
```

Part 2: Does a Lunch to Dinner Price Increase Mean Better Food Quality for Dinner? (Difference in Lunch and Dinner Prices vs Difference in Lunch and Dinner Ratings)

Part 2 deals with the relationship between the difference in food quality and ratings for lunch vs dinner. The difference between dinner and lunch prices and the difference in dinner and lunch ratings were calculated for each restaurant. A data frame with the dinner and lunch price and rating differences was created.

Calculating the price and rating difference between dinner and lunch:

```
pricediff <- dinnerprice - lunchprice
pricediff
```

```
## [1] 3000 2000 1000 6500 2000 1000 3000 3000 6500 3000 2000
## [12] 4000 2000 4500 0 2000 0 8000 2000 4000 2000 7000
## [23] 3000 3000 0 0 5500 5500 2000 3000 3000 2000 4500
## [34] 1000 3500 1000 0 7000 9000 2000 2000 7500 4000 0
## [45] 3000 5500 0 7000 1000 0 2000 2000 -2000 3000 3000
## [56] 2000 1000 1000 3000 0 7000 1000 2000 1000 3000 5500
## [67] 3000 1000 3500 3000 2000 3000 3500 3000 3000 18000 0
## [78] 3500 3000 2000
```

```
ratingdiff <- rest1$DinnerRating - rest1$LunchRating
ratingdiff
```

```
## [1] 0.18 0.00 -0.10 0.04 -0.12 0.21 0.09 -0.17 0.08 -0.11 -0.14
## [12] -0.04 -0.03 -0.39 -0.34 0.00 0.03 0.53 -0.08 0.16 -0.03 0.01
## [23] 0.00 -0.11 0.02 0.02 -0.35 0.04 0.00 -0.01 -0.43 -0.26 -0.21
## [34] 0.05 0.01 -0.04 -0.03 0.01 0.39 0.01 -0.03 0.20 -0.08 -0.11
## [45] 0.31 -0.15 0.05 -0.14 -0.24 -0.01 -0.33 0.00 0.35 0.02 0.13
## [56] 0.06 -0.06 -0.02 0.02 -0.01 0.11 -0.12 -0.23 -0.07 0.18 0.16
## [67] 0.06 0.05 -0.01 -0.03 0.01 0.01 0.00 0.21 -0.14 -0.10 0.04
## [78] 0.15 -0.01 -0.24
```

Creating data frame:

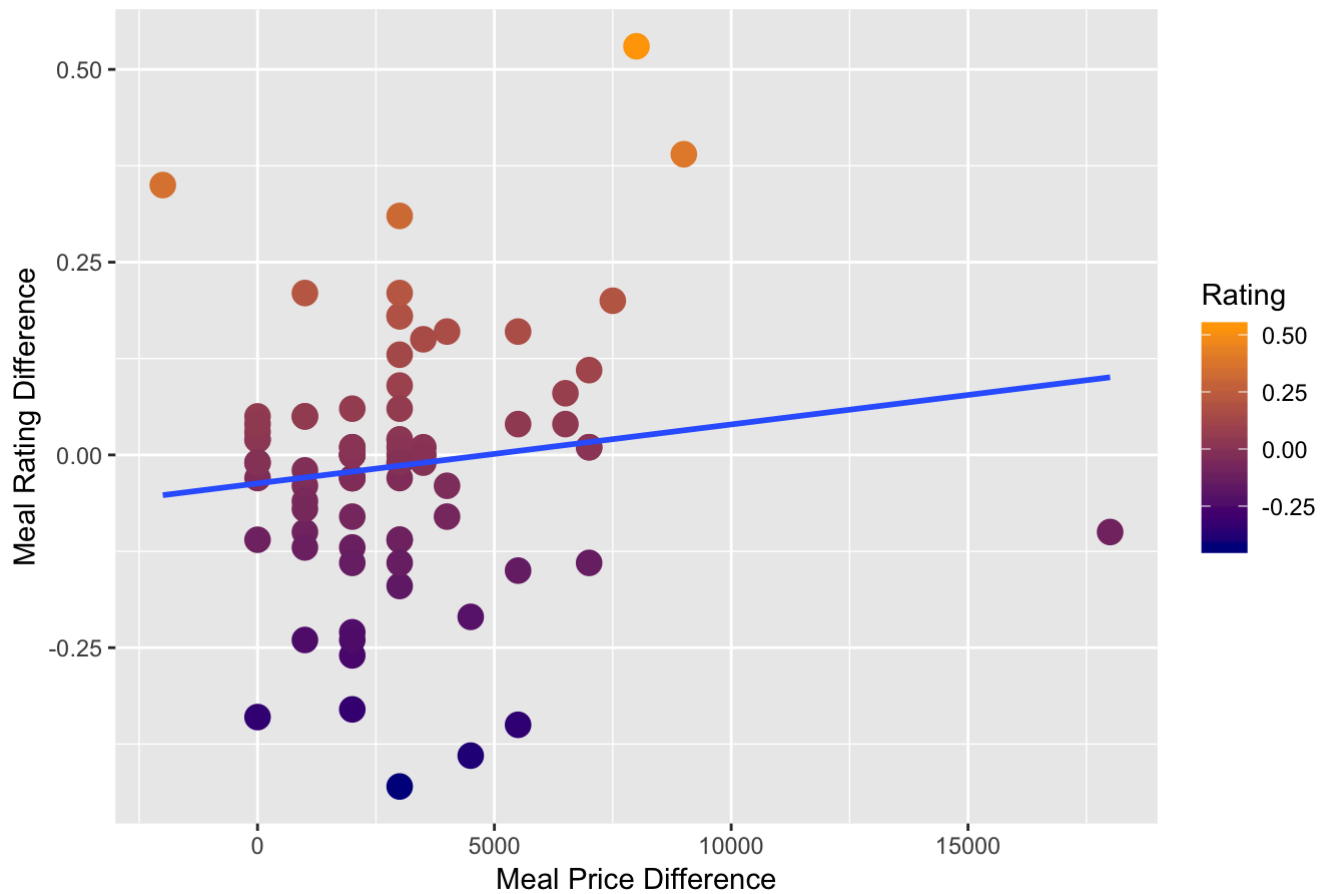
```
PvRdifference_df <- data.frame(pricediff, ratingdiff)
PvRdifference_df
```

Data Visualization

I created a scatterplot of rating difference vs price difference with a best fit linear regression model. The correlation and coefficient of determination were calculated to be $R=0.1275$ and $R^2=0.0163$.

```
ggplot(PvRdifference_df, aes(x=pricediff, y=ratingdiff))+
  geom_point(size=4, aes(color=ratingdiff))+
  scale_color_gradient(name="Rating", low="darkblue", high="orange")+
  labs(x="Meal Price Difference", y="Meal Rating Difference", title="Lunch and Dinner Ra
ting Difference vs Price Difference")+
  theme(plot.title = element_text(hjust = 0.5))+
  geom_smooth(method="lm", se = FALSE)
```

Lunch and Dinner Rating Difference vs Price Difference



```
lm(pricediff ~ ratingdiff)
```

```
##
## Call:
## lm(formula = pricediff ~ ratingdiff)
##
## Coefficients:
## (Intercept)    ratingdiff
##          3024          2126
```

Correlation (R):

```
PvRdifference_df %>%
  summarize(r_4 = cor(pricediff, ratingdiff))
```

```
##          r_4
## 1 0.1274855
```

Coefficient of Determination (R^2):

```
lmpvrdiff <- lm(pricediff ~ ratingdiff, data = PvRdifference_df)
summary(lmpvrdiff)
```

```
##
## Call:
## lm(formula = pricediff ~ ratingdiff, data = PvRdifference_df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5767.7 -1577.0  -406.3   532.4 15189.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   3023.5      308.7    9.793 3.15e-15 ***
## ratingdiff    2126.3     1873.1    1.135   0.26
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2751 on 78 degrees of freedom
## Multiple R-squared:  0.01625,    Adjusted R-squared:  0.00364
## F-statistic: 1.289 on 1 and 78 DF,  p-value: 0.2598
```

```
summary(lmpvrdiff)$r.squared
```

```
## [1] 0.01625256
```

Analysis Without Price Difference Outliers

Like in Part 1 of Section 2, I used the 1.5*IQR rule to take out outliers of price difference and created a new scatterplot using the new data frame that does not contain the outliers. The recalculated R and R² values are -0.0312 and 0.0010, respectively.

Outliers in Price Difference:

The third quartile is 3500, which was added to the value of 1.5*IQR. Outliers for price difference are the values where pricediff > 7250.

```
quantile(pricediff)
```

```
##      0%      25%      50%      75%     100%
## -2000    1000    3000    3500   18000
```

```
IQR_PD = IQR(pricediff)
IQR_PD
```

```
## [1] 2500
```

```
1.5*IQR_PD
```

```
## [1] 3750
```

```
3500+1.5*IQR_PD
```

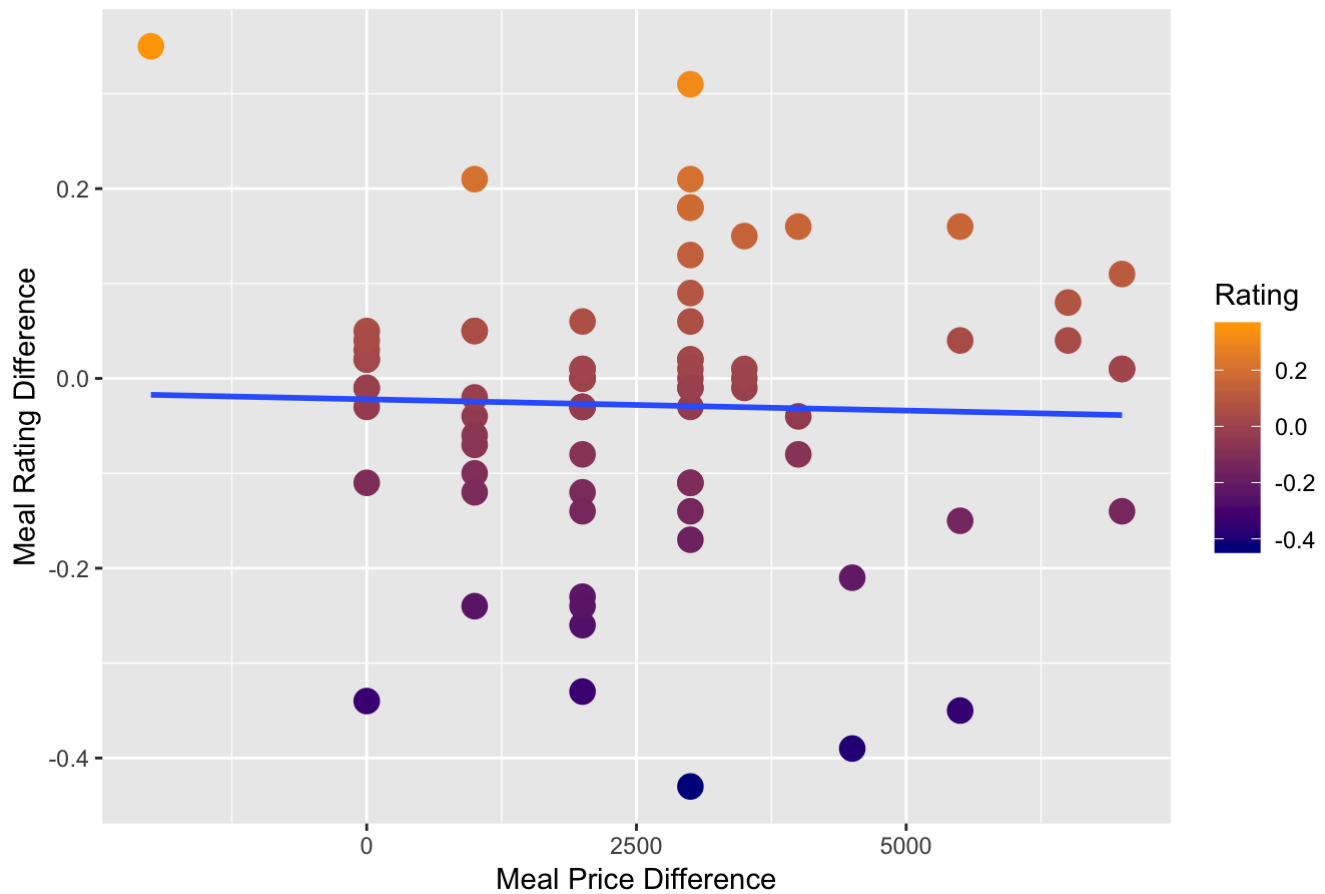
```
## [1] 7250
```

New Data Frame and Visualization:

This new data frame contains the data points with price difference values less than 7250. The new scatterplot shows the values of the `PvRdifference` data frame without outliers.

```
PvRdifference_df2 <- PvRdifference_df %>%  
  filter(pricediff <= 7250)  
  
ggplot(PvRdifference_df2, aes(x=pricediff, y=ratingdiff))+  
  geom_point(size=4, aes(color=ratingdiff))+  
  scale_color_gradient(name="Rating", low="darkblue", high="orange")+  
  labs(x="Meal Price Difference", y="Meal Rating Difference", title="Lunch and Dinner Ra  
ting Difference vs Price Difference")+  
  theme(plot.title = element_text(hjust = 0.5))+  
  geom_smooth(method="lm", se = FALSE)
```


Lunch and Dinner Rating Difference vs Price Difference



New Correlation (R) and Coefficient of Determination (R^2)

```
PvRdifference_df2 %>%
  summarize(r_5 = cor(pricediff, ratingdiff))
```

```
##           r_5
## 1 -0.03118658
```

```
lmpvrdiff2 <- lm(pricediff ~ ratingdiff, data = PvRdifference_df2)
summary(lmpvrdiff2)
```

```
##
## Call:
## lm(formula = pricediff ~ ratingdiff, data = PvRdifference_df2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4437.2 -1560.1  -156.4   638.7  4464.5
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2580.6      227.7   11.335  <2e-16 ***
## ratingdiff    -409.7     1526.5   -0.268    0.789
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1949 on 74 degrees of freedom
## Multiple R-squared:  0.0009726, Adjusted R-squared:  -0.01253
## F-statistic: 0.07204 on 1 and 74 DF,  p-value: 0.7891
```

```
summary(lmpvrdiff2)$r.squared
```

```
## [1] 0.0009726028
```

Interpretations of Data and Overall Conclusions

Section 1

In Section 1, we saw that there exists a weak linear relationship between average price and rating for the 27 restaurants from the survey I sent out to my friends.

Section 2

In Part 1, the first scatterplot, which included outliers, had a correlation of 0.5194 and $R^2=0.2698$. Thus, there is a moderately strong, positive correlation between average price and rating. However, after the outliers were taken out, both the correlation and coefficient of determination decreased to 0.2827 and 0.0799, respectively, indicating a weaker linear relationship. For both graphs, the R^2 values were extremely low, so not much variability was accounted for by the linear models.

In Part 2, the first scatterplot had a weak, positive correlation of 0.1275 and R^2 value of 0.0163. Similar to the results from Part 1, after outliers were taken out, correlation and coefficient of determination both decreased significantly to -0.0312 and 0.0010, respectively. It is interesting to note that after excluding outliers, the correlation became negative. Once again, the coefficient of determination was very small regardless of whether or not we included outliers.

Thus, from Section 1 and Section 2 Part 1, we see that there is not significant evidence to support the fact that higher ratings equate to better quality food. From Section 2 Part 2, we determined that restaurants don't necessarily have better dinner food than lunch food just because of the price increase from lunch to dinner. Overall, higher prices don't mean higher ratings, whether it is across restaurants or within an individual restaurant

(lunch and dinner). My friends and I should therefore choose to go to the cheapest option when we are deciding where to eat out. We should also try to only eat out for lunch and avoid going out for dinner since dinners tend to be more expensive but don't have better food.