Menu Pricing for a New Italian Restaurant in New York City (IDMRAD version with regression analysis and technical appendix)

Brian W. Junker *
Department of Statistics & Data Science, Carnegie Mellon University

September 24, 2020

Abstract

We address the question of how to set menu prices for a new high-end Italian restaurant in New York City. We examine data on Italian restaurants in Manhattan collected by Zagat (2011), using exploratory data analyses presented by Sheather (2009). From exploratory data analysis, it appears that the average price of dinner is highly influenced by customer ratings of Food and Service, and somewhat less by Decor. However, these rating variables are highly correlated, and a regression analysis shows that Service matters much less after accounting for Food and Decor. The effect of location (east or west of Fifth Avenue) is more ambiguous. The maximal rating on each category is 24 or 25 out of 30, so there is room for a restaurant even more highly rated on all these dimensions. The restuaurant would be competitively priced if the average dinner price was in the \$60–\$80 or so range, although it should be noted that no restaurants currently have prices above \$65.

Keywords: Zagat, Sheather (2009), Italian Food, Food, Service, Decor

*brian@stat.cmu.edu

1 Introduction

The restaurant market in any city is competitive and difficult to thrive in, and driven by customer perceptions as much as any other quality. How should menu prices for a new restaurant be set?

This question is especially critical in the highly competitive New York City market, where we have been asked to suggest prices for a new Italian restaurant's dinner menu that are consistent and competitive with other high-end Italian restaurants in Manhattan, in or north of the Flatiron district. In particular, "The stated aims of the restaurant are to provide the highest quality Italian food utilizing state-of-the-art décor while setting a new standard for high-quality service in Manhattan. The creation and the initial operation of the restaurant will be the basis of a reality TV show for the US and international markets (including Australia)" (Sheather 2009, pp. 5–6).

In addition to answering the main question posed above, we will address the following questions:

- Which customer perception has the largest effect on pricing?
- Should the restaurant be located east or west of Fifth Avenue, to maximize menu prices?
- Can we set a "price premium" (i.e., higher price) for "setting a new standard for high quality service in Manhattan" for Italian restaurants?
- Are there any restaurants that seem unusually high- or low-priced[1], given customer perceptions?

2 Data

The data for this study come from several surveys that were conducted of Italian restaurants in or north of the Flatiron district in Manhattan in New York City. The reader should refer to Zagat (2011) for definitions, eligibility, inclusion/exclusion criteria, and so forth, for these surveys.The data, and the entire problem as stated in Section 1, are presented in Example 1.2.3, pp. 5–7, in Sheather (2009).

[1] These are outliers, in common Statistical parlance.

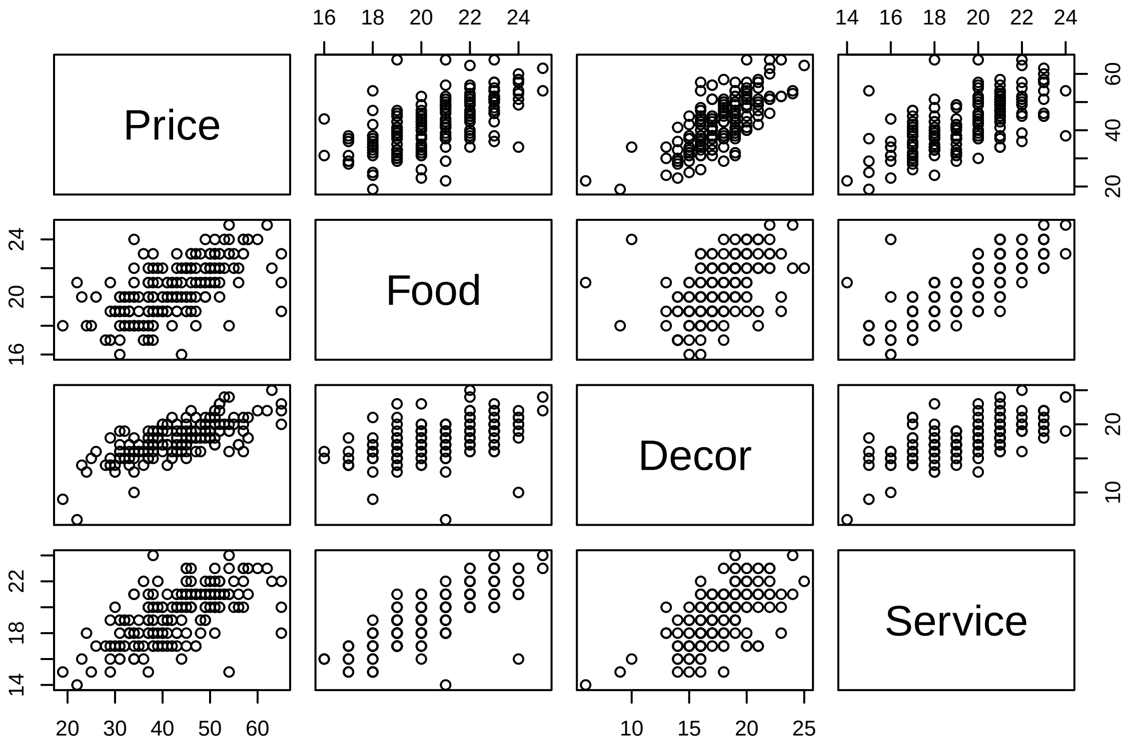


Figure 1: Scatterplot matrix of Price, Food, Décor, & Service ratings. From Sheather (2009, p. 7).

In all, 168 restaurants are represented in the data available to us, and the following variables were measured on each:

Table 1: Description of dataset of Restaurant datset.

Y	Price = the price (in $US) of dinner (including one drink & tip)
x_1	Food = customer rating of the food (out of 30)
x_2	Decor = customer rating of the décor (out of 30)
x_3	Service = customer rating of the service (out of 30)
x_4	East = dummy variable = 1 (0) if the restaurant is east (west) of Fifth Avenue

The data are available in the file `nyc.csv`, in the online supplement accompanying Sheather (2009).

In Figure 1 we show the relationships between all of the quantitative variables x_1 through x_4 in the data; each appears to be positively correlated with the response variable Price, and they are all positively correlated with each other. In Figure 2 we show the relationship of Price to location (east or west of Fifth Avenue), showing relatively symmetric price distributions with no outliers and a weak relationship with location. More details from an exploratory data analysis (EDA) can be found in Appendix A.1.

3 Methods

Our analysis[2] consists of two parts. First, we relied on visual comparison of exploratory scatter plots and box plots (Figures 1 and 2 in this report) provided by (Sheather 2009, p. 7), using the R language and environment for statistical computing (R Core Team 2020). In order to identify restaurants with unusual/outlying pricing, we have also examined the raw data in `nyc.csv`. This analysis can tell us about the individual effects of the variables, but not how they work in combination to affect pricing. Detailed R analyses can be found in Appendix A.1 and Appendix A.2.

[2]This is not the most complete possible analysis; however I hope that it is useful in demonstrating what an IDMRAD paper might look like.

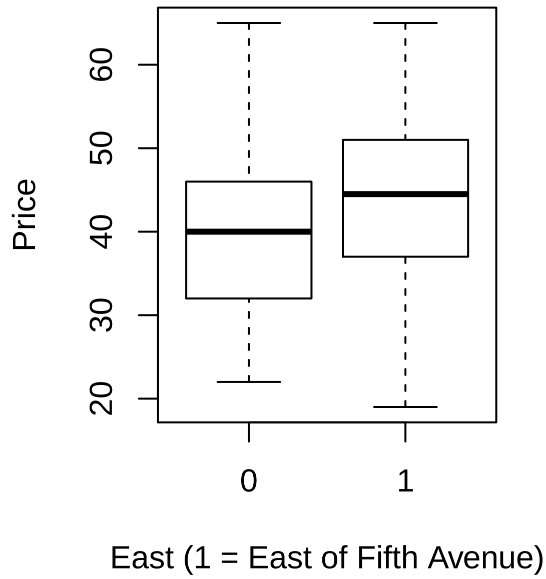


Figure 2: Box plots of Price for two levels of the dummy variable East. From Sheather (2009, p. 7).

Second, we considered two multiple regression models, also in R, predicting Price from each of the variables Food, Decor, Service, and East, since multiple regression can tell us about the effect of each individual predictor variable, after controlling for all the other predictor variables. We examined case-wise residual plots and a likelihood ratio test to select a good model, and we used that model to interpret the effects of the predictor variables and to estimate a competitive price for a "premium" restaurant. Details of these analyses in R can be found in Appendices A.3, A.4 and A.5.}

Analyses were carried out in R and RStudio (RStudio Team 2020).

4 Results

4.1 Visual Comparison of Exploratory Plots

We can crudely estimate the slope of a (univariate) regression of Price on each rating variable, by dividing the plotting range of Price by the plotting range of each rating variable, estimated by eye from the scatterplots in Figure 1. The estimated slopes are 5.0 (Food), 2.67 (Décor), and 4.0 (Service). Thus, apparently, the Food rating has the biggest influence on

Price (with a \$5 increase in price associated with each 1-point increase in rating), followed closely by Service, and somewhat more distantly by Décor.

We can examine the influence of Service rating on Price more carefully by examining the scatter plot in the upper right in Figure 1. As noted in the previous paragraph, there appears to be a strong increasing relationship between Service rating and Price. Moreover, there is scope for creating a restaurant with higher Service ratings than any in the the Zagat survey: the highest Service rating appears to be 24 or so, out of 30. On the other hand, there is cause for concern about a higher price point for restaurants with excellent service: the two restaurants with the highest service ratings (about 24) do not have the highest average dinner prices: one is priced around \$54 with ten or so restaurants charging more for dinner, with lesser service ratings; and the other prices dinner quite modestly at under \$40. Indeed, restaurants with more modest service ratings of 18 and 20 are among the most costly in the survey. Based on this analysis, it does not seem worthwhile to rest higher Prices exclusively on high Service ratings.

Interestingly, the two restaurants with modest Service ratings and maximal dinner Prices also have quite modest Food ratings, as can be seen from top row of scatter plots in Figure 1. Examining the raw data[3] in R, we find is located East of Fifth, and the other West of Fifth, so location does not appear to be a direct influence. Instead, these restaurants seem to be getting a boost from relatively high Décor ratings.

Turning to the boxplots in Figure 2, it appears that there is some influence of location (East of Fifth Avenue, vs West); the median price goes from about \$40 to about \$45, with a similar shift in the lower and upper quartiles of price. However, the interquartile range is about \$10 or so, and the entire range of prices for restaurants East of Fifth Avenue completely contains the range of prices for those West of Fifth. Since the typical effect of location is similar to the effect of a one-point increase in Food or Service rating—about \$5—it may not be worth it to choose a location East of Fifth, especially if the price of rent and other operating costs are substantially higher East of Fifth.

There do not appear to be any strong outliers in Price in this data set. The boxplots in Figure 2 do not show any outliers by the usual $1.5\times$IQR rule, and the scatterplots in the

[3]We used the R command `nyc[nyc$Price==max(nyc$Price),]`. See Appendix A.2 for details

top row of Figure 1 generally confirm this. One restaurant seems to have a relatively high dinner Price, about $54, with rather low Service and Food ratings (15 and 18, respectively). Examining the raw data[4] in R, the restaurant appears to be Nello. From the raw data record shown in Table 2,

Table 2: Raw data record for the restaurant Nello.

Case	Restaurant	Price	Food	Decor	Service	East
56	Nello	54	18	16	15	1

4.2 Regression Analysis

For regression analysis we considered the two multiple regression models

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \varepsilon \tag{1}$$

and

$$\begin{aligned} Y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 \\ &\quad + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 \\ &\quad + \beta 23 x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{34} x_3 x_4 + \varepsilon \end{aligned} \tag{2}$$

with Y = Price, x_1 = Food, x_2 = Décor, x_3 = Service, and x_4 = East, as defined in Section 2. Casewise plots of residuals (Appendices A.2 and A.3) did not reveal any substantial misfit, nor any worrying outliers or influential observations, for either model. A likelihood ratio test comparing model (1) with model (2) does not strongly favor the model with interactions, and so for the remainder of the analysis we will use the main-effects-only model in equation (1).

[4]We used the R command `nyc[nyc$Service==15,]`, since there appear to be only five restaurants with this Service rating; from there it is easy to identify Nello from the R output... See Appendix A.2 for details.

Table 3: Estimated coefficients for model 1.

	Estimate	Std. Error	t value	$P(T > \|t\|)$
$\hat{\beta}_0$ (Intercept)	-24.024	4.708	-5.102	0.000
$\hat{\beta}_1$ Food	1.538	0.369	4.169	0.000
$\hat{\beta}_2$ Decor	1.910	0.217	8.802	0.000
$\hat{\beta}_3$ Service	-0.003	0.396	-0.007	0.995
$\hat{\beta}_4$ East	2.068	0.947	2.184	0.030

Table 3 gives the estimated coefficients for model (1) along with standard errors and the usual t-test for testing whether each coefficient is significantly different from zero. We can see that Food and Decor each have a strong effect on price, worth about \$1.54 per rating point for Food, and \$1.91 per rating point for Decor. Service does not have much of a unique effect after accounting for other variables in the model. This could either be because Service is not really important in the model (which seems to contradict Figure 1, for example), or it could be that Service is highly correlated with other variables. For example in Figure 1 it can be seen in the bottom row, second plot, that as the Food quality goes up, so does the Service quality, so that once Food quality is accounted for, perhaps there is little more variation for Service quality to account for. This is an "ensemble" effect that could not be seen in the one-variable-at-a-time analyses in Section 4.1.

On the other hand, restaurant location, encoded by the variable East, has a significant but small effect on the mean price: it's worth about \$2 more in average price for a restaurant to be located East of Fifth Avenue. This seems different from the result using boxplots in Figure 2. The difference is that boxplots compare the whole distribution, so differences have to be fairly widespread across the price distribution for us to see a difference, whereas the regression analysis focuses on mean price, adjusted for the other variables in the model[5].

We can see from the exploratory analysis (Appendix A.1) that the maximum Food, Decor and Service ratings are all 24 or 25. To examine pricing for a "Premium" restaurant,

[5]Generally when you concentrate inference on the means, you get more dramatic results (because, roughly speaking, $SE_{mean} = SD_{population}/\sqrt{samplesize}$).

we estimated the average price for two fictional restaurants (Appendix A.5):

- A restaurant located West of Fifth Avenuem with Food, Decor and Service ratings of 25 each: estimated price \$62.11 with a prediction interval of $(50.36, 73.87)$.
- A restaurant located East of Fifth Avenue with Food, Decor and Service ratings of 30 each: estimated price \$81.41 with a prediction interval of $(69.08, 93.74)$.

Thus it seems that a "Premium" restaurant could charge on average between about \$60 and \$80. It should be noted that the current highest average price is \$65, so such a restaurant might wish to charge in the lower part of this range, to stay competitive with other restaurants.

Finally, in examining the residual plots for model (1) more closely in Figure 3 (see also Appendix A.3), we can see that Case 56 is the most extreme outlier (with a standardized residual of at least 3). Referring back to Table 2, we see once again that this is the restaurant Nello, which we also identified from the visual EDA. (More generally, higher prices seem to be under-predicted by the model.)

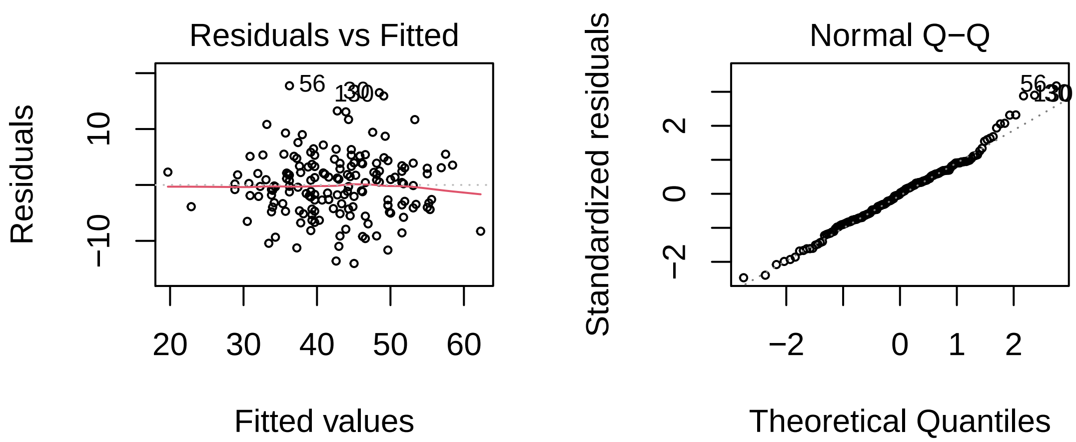


Figure 3: Residuals (raw and standardized) from model (1).

5 Discussion

The many Italian restaurants north of Flatiron in Manhattan surveyed by Zagat (2011) exhibit a wide range of Price as well as Food, Décor and Service ratings. In ourone-variable-at-a-time exploratory analysis, we found that Food and Service ratings exert high and nearly identical influence on Price, with an increase of one rating point associated with about a \$4–\$5 increase in average dinner price, while Decor has an effect about half as large. When we examined the effects of all the variables together in a multivariate regression, we fould that Food and Décor had an effect of about \$1.54 and \$1.91 per rating point, respectively, whereas Service had a negligible effect after accounting for the other two variables; this may be because Service is so highly correlated with the other variables (Food, especially). Overall, location (east or west of Fifth Avenue) does not appear to move the price distribution very much (as seen in the boxplots in Figure 2), though there is a mean price increase of about \$2 for locating east of Fifth Avenue after controlling for the other variables (Table 3). We did find one restaurant in both the exploratory analysis and the regression analysis—Nello—whose price seemed unusually high for its otherwise modest customer ratings, located east of Fifth Avenue.

If, say, you are going to open several restaurants, you may care more about the fact that the mean price can be higher East of Fifth Avenue, since on average you can charge a bit more in all your restaurants, and perhaps make substantially more money from their combined income. On the other hand if you are considering opening just one restaurant, the boxplots may be more important: the price distributions for East vs West restaurants greatly overlap, so there's little reason, in terms of competitiveness or profits, to make a location East of Fifth Avenue a primary concern.

There is scope for establishing a restaurant with higher Service ratings than any others in the Zagat surveys; indeed, ratings on all three scales top out at about 24 or 25 out of 30. From our regression analysis we concluded that a restaurant with ratings in the 25–30 range might charge between roughly \$60 and \$80, although to stay competitive such a restaurant might want to stay close to the current high average price of \$65.

The visual exploration of EDA graphs allowed us to look at univariate and bivariate relationships, and as such could not readily consider the interaction of two or more variables influencing price. Although much can be learned from visual inspection like this, the multivariate regression analysis allowed us to look at the ensemble effect of all of the variables on price. Future analyses might expand on this approach.

Our study was also limited by the use of Zagat (2011) data as reported by Sheather

(2009). It is not known how representative this data is of Italian restaurants generally in Manhattan; the Zagat data were collected to give Zagat readers useful information about individual restaurants they might be interested in, rather than to provide an unbiased characterization of Italian restaurants in the target area. In addition, the data is somewhat old; the population of Italian restaurants has undoubtedly changed in the past 17–18 years, and prices and ratings have undoubtedly changed as well.

In summary, keeping the caveats of the last two paragraphs in mind, there is scope to establish a restaurant providing “the highest quality Italian food utilizing state-of-the-art décor while setting a new standard for high-quality service in Manhattan” (Sheather 2009, pp. 5–6). Location (East or West of Fifth Avenue) does not seem to be a big price driver, on average; and there doesn’t seem to be much hope of establishing a price premium for outstanding service alone. However, a restaurant with ratings in the 25–30 range on Food, Décor and Service could have an average dinner Price in the \$60–\$80 or so range, and be competitive within the set of restaurants surveyed by Zagat (2011), for which the highest average dinner price was \$65.

A Appendix

A.1 Appendix: Initial Data Import & Exploration

Read the data in, get a general sense of the variables, and make a "pairs" plot (scatterplot matrix) of the numerical variables. Note that "Price" is the response variable.

```
nyc <- read.csv("nyc.csv")


str(nyc)
```

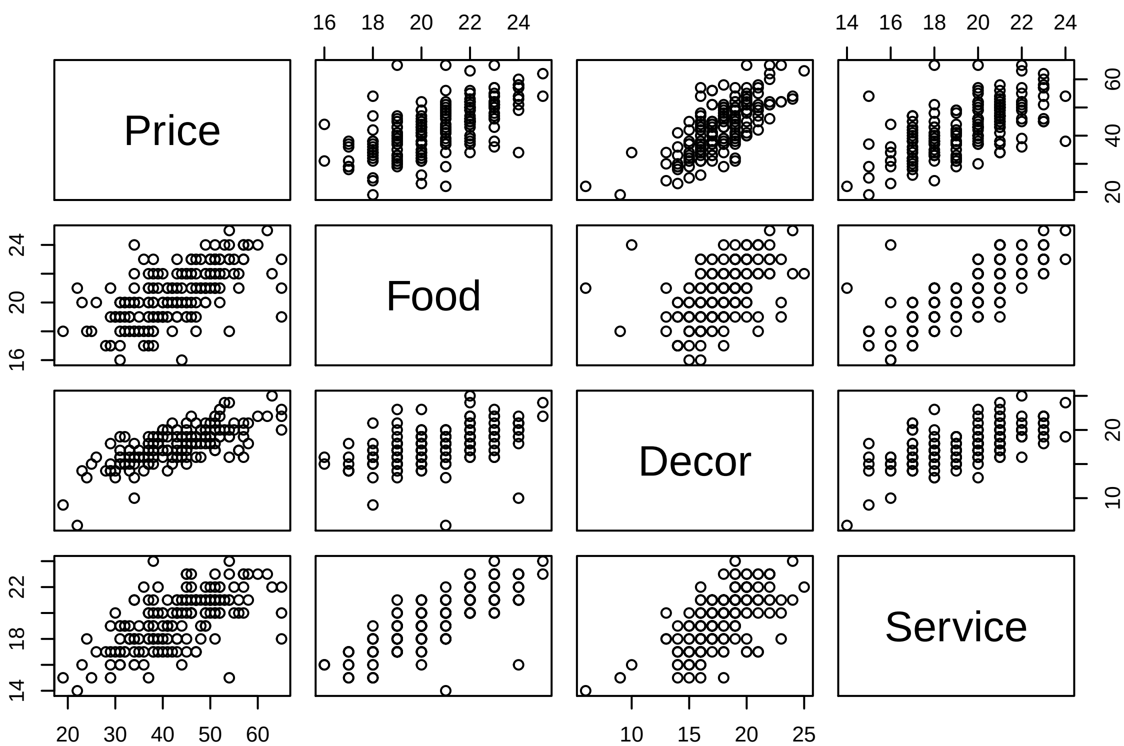
```
## 'data.frame':    168 obs. of  7 variables:
##  $ Case      : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ Restaurant: chr  "Daniella Ristorante" "Tello's Ristorante" "Biricchino" "Bottino'
##  $ Price     : int  43 32 34 41 54 52 34 34 39 44 ...
##  $ Food      : int  22 20 21 20 24 22 22 20 22 21 ...
##  $ Decor     : int  18 19 13 20 19 22 16 18 19 17 ...
##  $ Service   : int  20 19 18 17 21 21 21 21 22 19 ...
##  $ East      : int  0 0 0 0 0 0 0 1 1 1 ...
```

```
summary(nyc)
```

```
##       Case         Restaurant            Price            Food     
##  Min.   :  1.00   Length:168         Min.   :19.0   Min.   :16.0  
##  1st Qu.: 42.75   Class :character   1st Qu.:36.0   1st Qu.:19.0  
##  Median : 84.50   Mode  :character   Median :43.0   Median :20.5  
##  Mean   : 84.50                      Mean   :42.7   Mean   :20.6  
##  3rd Qu.:126.25                      3rd Qu.:50.0   3rd Qu.:22.0  
##  Max.   :168.00                      Max.   :65.0   Max.   :25.0  
##      Decor          Service          East      
##  Min.   : 6.00   Min.   :14.0   Min.   :0.000  
##  1st Qu.:16.00   1st Qu.:18.0   1st Qu.:0.000  
```

```
##  Median :18.00   Median :20.0   Median :1.000
##  Mean   :17.69   Mean   :19.4   Mean   :0.631
##  3rd Qu.:19.00   3rd Qu.:21.0   3rd Qu.:1.000
##  Max.   :25.00   Max.   :24.0   Max.   :1.000
```

```
pairs(nyc[,-c(1:2,7)])
```

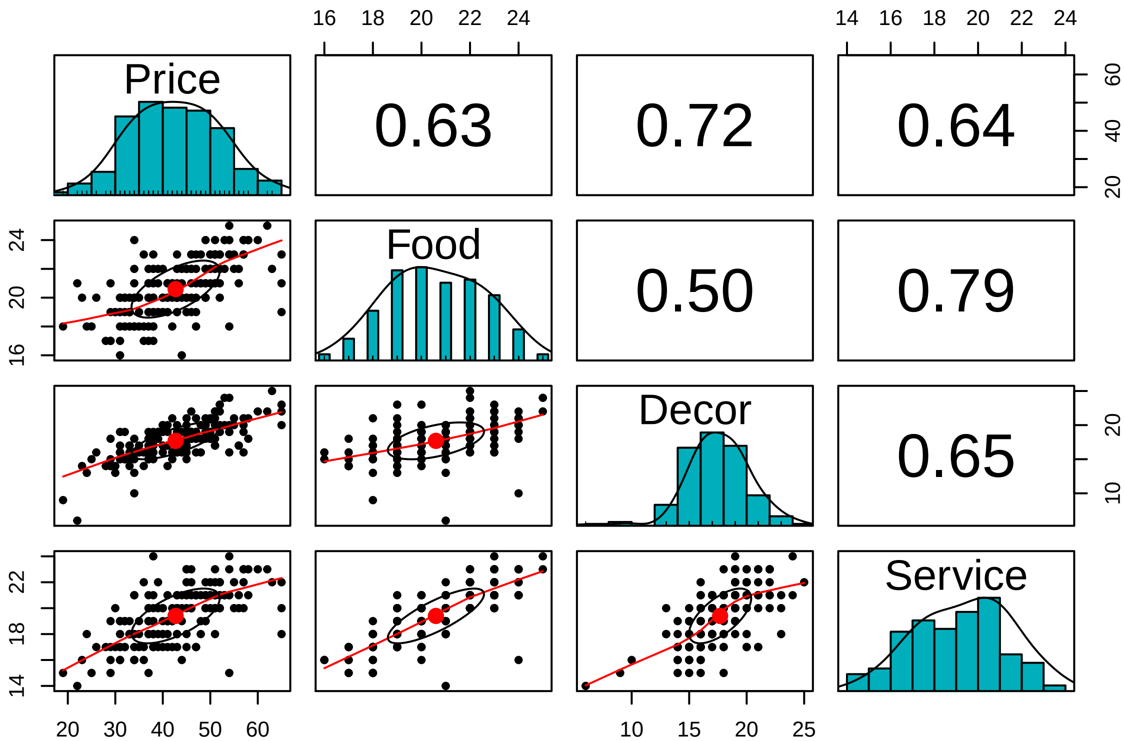


We can get a more refined look at the variables with a scatterplot matrix that also includes histograms for each variable. If the histograms revealed especially long tails or wierd outliers, we might want to transform the data, recode or delete outliers, etc.

```
library(psych)
## you would need to install the "psych" package
## one time before using this library() command...

pairs.panels(nyc[,-c(1:2,7)],
             method = "pearson",    ## correlation method
             hist.col = "#00AFBB",  ## a pretty color for histogram bars...
```

```
                 density = TRUE,         ## show density plots
                 ellipses = TRUE         ## show correlation ellipses
                 )
```

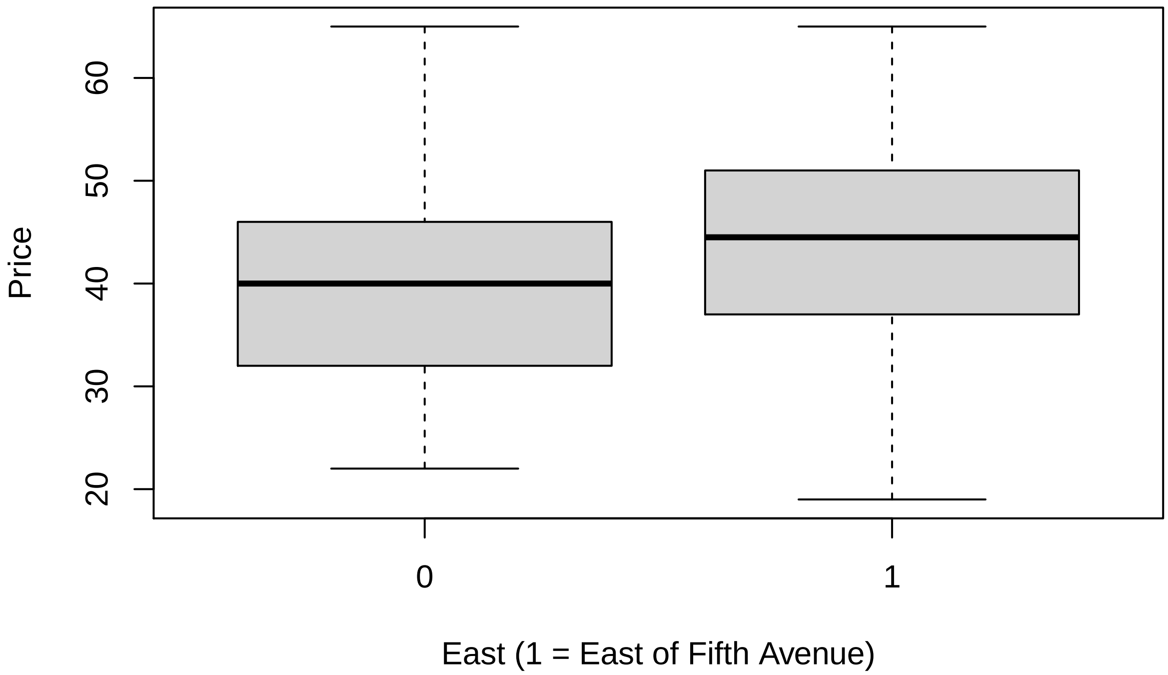


(There are lots of other packages, including ggplot, that can produce similar plots. This is really just a convenient illustration.)

The histograms don't suggest any special processing (transformations, etc.) will be needed for the variables, so we can proceed. Note that all the variables seem fairly highly correlated with one another, which makes sense, but also can affect regression results, as we'll learn later in the semester.

One of the main questions for this study is whether restaurants shoudl locate east or west of Fifth Avenue. A pair of boxplots give us a first look at this question:

```
with(nyc,boxplot(Price ~ East, xlab="East (1 = East of Fifth Avenue)"))
```



A.2 Appendix: Analysis of Apparent Outliers in the EDA Plots

To find the two restaurants with modest Service ratings and maximal dinner Prices…

```
nyc[nyc$Price==max(nyc$Price),]
```

```
##     Case     Restaurant Price Food Decor Service East
## 30    30 Harry Cipriani    65   21    20      20    1
## 130  130  Rainbow Grill    65   19    23      18    0
## 132  132   San Domenico    65   23    22      22    0
```

To find the restaurant with Service = 15….

```
nyc[nyc$Service==15,]
```

```
##     Case           Restaurant Price Food Decor Service East
## 56    56                Nello    54   18    16      15    1
## 68    68 Zucchero e Pomodori    29   17    14      15    1
## 69    69             Baraonda    37   17    18      15    1
## 100  100              Ecco-la    25   18    15      15    1
## 115  115              Lamarca    19   18     9      15    1
```

A.3 Appendix: Regression Analysis – Main Effects Only

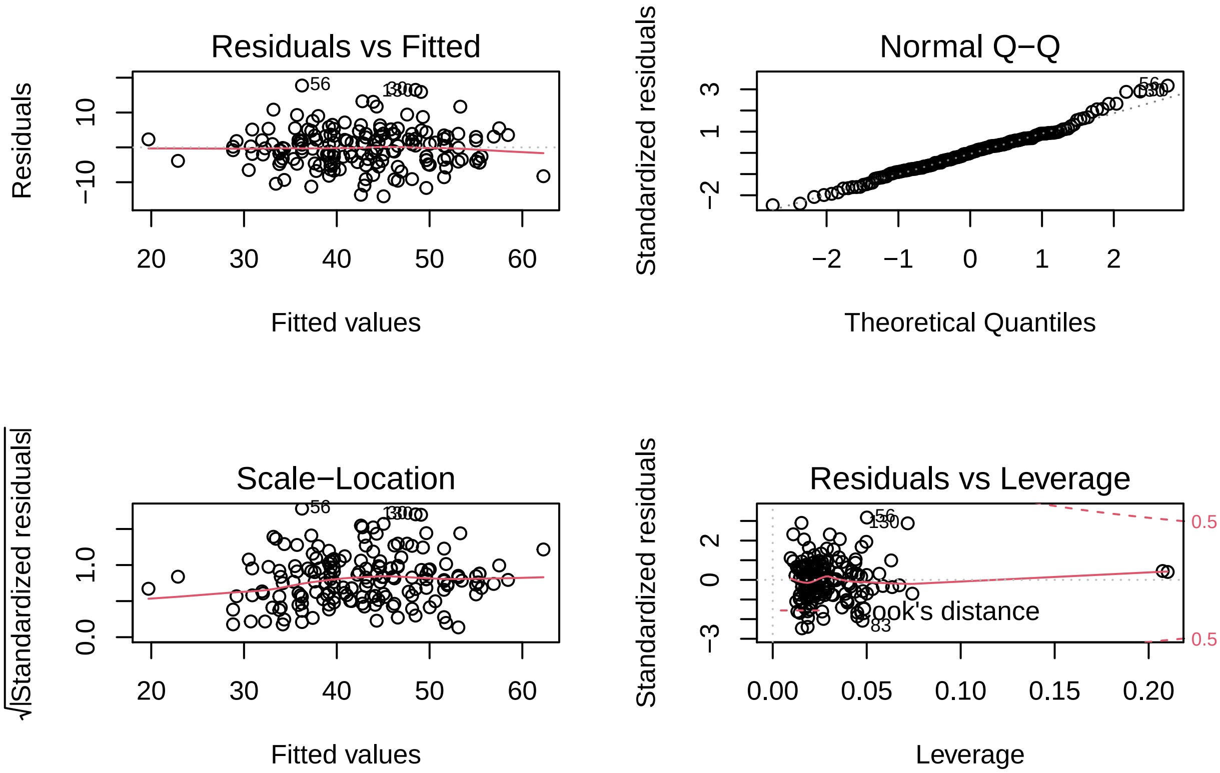
Here's a very light regression analysis to see how the variables work with one another.

```
summary(lm.0 <- lm(Price ~ . , data=nyc[,-c(1,2)]))
```

```
##
## Call:
## lm(formula = Price ~ ., data = nyc[, -c(1, 2)])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -14.0465  -3.8837   0.0373   3.3942  17.7491
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -24.023800   4.708359  -5.102 9.24e-07 ***
## Food          1.538120   0.368951   4.169 4.96e-05 ***
## Decor         1.910087   0.217005   8.802 1.87e-15 ***
## Service      -0.002727   0.396232  -0.007   0.9945
## East          2.068050   0.946739   2.184   0.0304 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.738 on 163 degrees of freedom
## Multiple R-squared:  0.6279, Adjusted R-squared:  0.6187
## F-statistic: 68.76 on 4 and 163 DF,  p-value: < 2.2e-16
```

```
par(mfrow=c(2,2))

plot(lm.0)
```



This model, which just has main effects for each of the quantitative predictor variables, suggests some interesting effects on menu prices, and the residual (casewise) diagnostic plots don't show any dramatic misfit, outliers, influential observations, etc.

From the table of coefficients, it looks like Food and Decor matter a lot for Price, but Service does not. This may be because Service is highly correlated with Food (and with Decor for that matter…).

There is also an effect for being East of Fifth Avenue; this is different from the result we got using only boxplots, because boxplots compare the whole distribution (and so differences have to be true across the distribution of prices) whereas regression analysis basically just looks at means, adjusted for the other variables in the model. Generally when you concentrate inference on the means, you get more dramatic results (because, roughly speaking, $SE_{mean} = SD_{population}/\sqrt{sample\ size}$).

If you are a policy maker (say, you have a lot of money and you are going to open several restaurants), you may care more about the fact that the mean price can be higher East of Fifth Avenue, since on average you can charge a bit more in your restaurants.

On the other hand if you are considering opening just one restaurant, the story of the boxplots may be more important: the price distributions for East vs West restaurants greatly overlap, there's little reason to make a location East of Fifth Avenue a primary concern.

A.4 Appendix: Regression analysis – Two-Way Interactions

Just for fun, we'll also try the model that has all main effects and two-way interactions, and we'll compare the two models with likelihood ratio test.

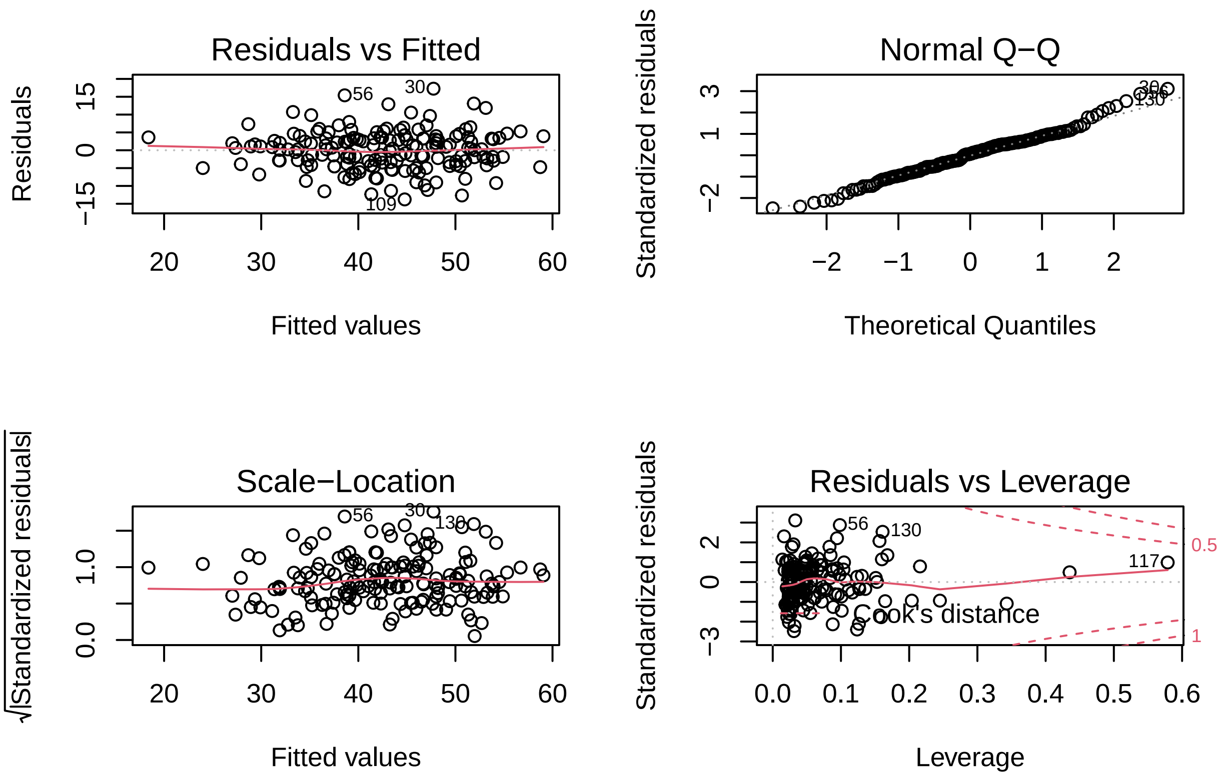
```
summary(lm.1 <- lm(Price ~ .^2 , data=nyc[,-c(1,2)]))
```

```
##
## Call:
## lm(formula = Price ~ .^2, data = nyc[, -c(1, 2)])
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -13.7758  -3.5519   0.3466   3.3383  17.2584
##
## Coefficients:
##                Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -39.32976   42.34684  -0.929  0.35444
## Food            2.61252    2.34496   1.114  0.26694
## Decor           7.26725    2.43591   2.983  0.00331 **
## Service        -4.68620    3.27542  -1.431  0.15450
## East            6.69634   10.55070   0.635  0.52656
## Food:Decor     -0.35758    0.13716  -2.607  0.01001 *
## Food:Service    0.20733    0.15317   1.354  0.17782
## Food:East       1.87559    0.89562   2.094  0.03785 *
## Decor:Service   0.10665    0.09193   1.160  0.24777
## Decor:East     -0.34309    0.46090  -0.744  0.45775
## Service:East   -1.90937    0.87262  -2.188  0.03014 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.645 on 157 degrees of freedom
## Multiple R-squared:  0.6531, Adjusted R-squared:  0.631
## F-statistic: 29.55 on 10 and 157 DF,  p-value: < 2.2e-16
```

```r
par(mfrow=c(2,2))


plot(lm.1)
```



This model is intersting in that several interactions seem to have coefficients significantly different from zero, and some of the main effects no longer do. The residual plots do not look much better (or worse) than the plots for the main-effects-only model.

As a rule, unless you have a VERY VERY VERY VERY good reason for doing otherwise, when you want to keep an interaction in a model you should also keep the main effects. Thus, if we wanted to keep the `Service:East` interaction, we should also keep the main variables Service and East, even though neither main effect is significantly different from zero.

However, in this analysis we do not need to worry so much about that, since the likelihood ratio test does not strongly favor the model with interactions; it appears we can "get away" with just the main effects models.

```
anova(lm.0,lm.1,test="LRT")
```

```
## Analysis of Variance Table
##
## Model 1: Price ~ Food + Decor + Service + East
## Model 2: Price ~ (Food + Decor + Service + East)^2
##   Res.Df    RSS Df Sum of Sq Pr(>Chi)
## 1    163 5366.5
## 2    157 5003.4  6    363.11  0.07694 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

So, we can just stick with the simpler model, lm.0.

A.5 Appendix: Predicting the price of a restaurant that has very high scores on food, Decor and Service...

```
low.premium <- data.frame(Case=1000,Restaurant="The Ritz!",Price=NA,
                          Food=25,Decor=25,Service=25,East=0)
predict(lm.0,low.premium,interval="prediction")
```

```
##        fit      lwr     upr
## 1 62.11319 50.35648 73.8699
```

```r
high.premium <- data.frame(Case=1000,Restaurant="The Ritz!",Price=NA,
                           Food=30,Decor=30,Service=30,East=1)
predict(lm.0,high.premium,interval="prediction")
```

```
##        fit      lwr      upr
## 1 81.40864 69.07858 93.73869
```

A.6 Appendix: A Table Suitable for Including in a Report

```r
round(summary(lm.0)$coefficients,3)
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -24.024      4.708  -5.102    0.000
## Food           1.538      0.369   4.169    0.000
## Decor          1.910      0.217   8.802    0.000
## Service       -0.003      0.396  -0.007    0.995
## East           2.068      0.947   2.184    0.030
```

References

R Core Team (2020), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria.
URL: *https://www.R-project.org/*

RStudio Team (2020), *RStudio: Integrated Development Environment for R*, RStudio, PBC, Boston, MA.
URL: *http://www.rstudio.com/*

Sheather, S. J. (2009), *A modern approach to regression with R*, Springer Science + Business Media LLC.

Zagat (2011), *Zagatsurvey 2001 New York City Restaurants.*