Technical Report: Predicting Yelp Price Ranges

Josh Dey, Emmett Powers, & Giorlando Ramirez

12/10/2019

Abstract

In this project, we leverage a large Yelp data set comprised of about 200k observations in order to build a predictive model of restaurant prices. We go about answering this question deploying three supervised learning models: linear regression, ordinal regression, and classification trees. We find that the available variables have fairly limited predictive power across our models.

Preparing Libraries

```
library(ggplot2)
library(Amelia)
```

```
## Loading required package: Rcpp
```

```
## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.7.5, built: 2018-05-07)
## ## Copyright (C) 2005-2019 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##
```

```
library(rlist)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------
```

```
## v tibble  2.1.1      v purrr   0.3.2
## v tidyr   1.0.0      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------------- tidyv
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(jsonlite)

##
## Attaching package: 'jsonlite'

## The following object is masked from 'package:purrr':
##
##     flatten

library(rcompanion)
library(MASS)

##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

Introduction

In this project we want to use a large dataset provided by Yelp to predict restaurant prices. Economic intuition would have it that there are certain things restaurants might sell and/or provide that would make them more or less expensive, a prime example being the sale of alcohol. Thus, with this project we wish to put this intuition to the test. The Yelp dataset is comprised of many individual restaurant observations with variables that are likely determinants of the restaurants prices, such as whether they provide outdoor seating or the expected attire, so with this data we go about building a variety of models which might predict a given restaurant's price range.

The Data

The data used in this project comes from the Yelp Open Dataset. This is a massive dataset Yelp releases for educational purposes, which is comprised of three main sub datasets: reviews, business, and images. In this project we use the business dataset, a dataset which contains a variety of descriptive variables for 192,609 business location. However, in this project we're focused specifically on restaurants, so when we trim the data to only include restaurants we end up with 59,371 total observations. Each observation is a given restaurant with all of the associated variables, from name of the restaurant, star rating through whether the restaurant serves alcohol. There are a total of 53 variables.

In order to perform any meaningful analysis we needed to clean the data. Thus, as mentioned previously, we began this process by cutting the data to only include restaurants. This resulted in a total of 59,371 observations. We then filtered the data set to only include predictor variables we believed would have the strongest predictive capability- such as: # of stars, whether the restaurant served alcohol, etc.- and proceeded to further cut the data to only include observations which were complete (included an observation for every variable). This resulted in a dataset with 40,584 observations, a loss of about 30% of original observations. We have no reason to believe there exist systematic reasons as to why restaurants would omit data on these variables, and thus chose this as our final dataset, which we use to perform our analysis and create our models.

Exploratory Data Analysis

Explore the structure of the data through graphics. Here you can utilize both traditional plots as well as methods from unsupervised learning. Understanding the distribution of your response is particular important, but also investigate bivariate and higher-order relationships that you expect to be particular interesting.

In this section we look over the data in order to identify trends, concerns, and other consideration we might need to take into account when creating our models.

Initial Data Prep

First, we load up the data before we begin our exploratory analysis. We also filter the data for the explanatory variables we wish to use.

```
business <- stream_in(file("business.json"))
```

```
## opening file input connection.
```

```
##
 Found 500 records...
 Found 1000 records...
 Found 1500 records...
 Found 2000 records...
 Found 2500 records...
 Found 3000 records...
 Found 3500 records...
 Found 4000 records...
 Found 4500 records...
 Found 5000 records...
 Found 5500 records...
 Found 6000 records...
 Found 6500 records...
 Found 7000 records...
 Found 7500 records...
 Found 8000 records...
 Found 8500 records...
 Found 9000 records...
 Found 9500 records...
 Found 10000 records...
 Found 10500 records...
 Found 11000 records...
 Found 11500 records...
 Found 12000 records...
 Found 12500 records...
 Found 13000 records...
 Found 13500 records...
 Found 14000 records...
 Found 14500 records...
 Found 15000 records...
 Found 15500 records...
 Found 16000 records...
 Found 16500 records...
 Found 17000 records...
 Found 17500 records...
 Found 18000 records...
```

```
Found 18500 records...
Found 19000 records...
Found 19500 records...
Found 20000 records...
Found 20500 records...
Found 21000 records...
Found 21500 records...
Found 22000 records...
Found 22500 records...
Found 23000 records...
Found 23500 records...
Found 24000 records...
Found 24500 records...
Found 25000 records...
Found 25500 records...
Found 26000 records...
Found 26500 records...
Found 27000 records...
Found 27500 records...
Found 28000 records...
Found 28500 records...
Found 29000 records...
Found 29500 records...
Found 30000 records...
Found 30500 records...
Found 31000 records...
Found 31500 records...
Found 32000 records...
Found 32500 records...
Found 33000 records...
Found 33500 records...
Found 34000 records...
Found 34500 records...
Found 35000 records...
Found 35500 records...
Found 36000 records...
Found 36500 records...
Found 37000 records...
Found 37500 records...
Found 38000 records...
Found 38500 records...
Found 39000 records...
Found 39500 records...
Found 40000 records...
Found 40500 records...
Found 41000 records...
Found 41500 records...
Found 42000 records...
Found 42500 records...
Found 43000 records...
Found 43500 records...
Found 44000 records...
Found 44500 records...
Found 45000 records...
```

```
Found 45500 records...
Found 46000 records...
Found 46500 records...
Found 47000 records...
Found 47500 records...
Found 48000 records...
Found 48500 records...
Found 49000 records...
Found 49500 records...
Found 50000 records...
Found 50500 records...
Found 51000 records...
Found 51500 records...
Found 52000 records...
Found 52500 records...
Found 53000 records...
Found 53500 records...
Found 54000 records...
Found 54500 records...
Found 55000 records...
Found 55500 records...
Found 56000 records...
Found 56500 records...
Found 57000 records...
Found 57500 records...
Found 58000 records...
Found 58500 records...
Found 59000 records...
Found 59500 records...
Found 60000 records...
Found 60500 records...
Found 61000 records...
Found 61500 records...
Found 62000 records...
Found 62500 records...
Found 63000 records...
Found 63500 records...
Found 64000 records...
Found 64500 records...
Found 65000 records...
Found 65500 records...
Found 66000 records...
Found 66500 records...
Found 67000 records...
Found 67500 records...
Found 68000 records...
Found 68500 records...
Found 69000 records...
Found 69500 records...
Found 70000 records...
Found 70500 records...
Found 71000 records...
Found 71500 records...
Found 72000 records...
```

```
Found 72500 records...
Found 73000 records...
Found 73500 records...
Found 74000 records...
Found 74500 records...
Found 75000 records...
Found 75500 records...
Found 76000 records...
Found 76500 records...
Found 77000 records...
Found 77500 records...
Found 78000 records...
Found 78500 records...
Found 79000 records...
Found 79500 records...
Found 80000 records...
Found 80500 records...
Found 81000 records...
Found 81500 records...
Found 82000 records...
Found 82500 records...
Found 83000 records...
Found 83500 records...
Found 84000 records...
Found 84500 records...
Found 85000 records...
Found 85500 records...
Found 86000 records...
Found 86500 records...
Found 87000 records...
Found 87500 records...
Found 88000 records...
Found 88500 records...
Found 89000 records...
Found 89500 records...
Found 90000 records...
Found 90500 records...
Found 91000 records...
Found 91500 records...
Found 92000 records...
Found 92500 records...
Found 93000 records...
Found 93500 records...
Found 94000 records...
Found 94500 records...
Found 95000 records...
Found 95500 records...
Found 96000 records...
Found 96500 records...
Found 97000 records...
Found 97500 records...
Found 98000 records...
Found 98500 records...
Found 99000 records...
```

```
Found 99500 records...
Found 1e+05 records...
Found 100500 records...
Found 101000 records...
Found 101500 records...
Found 102000 records...
Found 102500 records...
Found 103000 records...
Found 103500 records...
Found 104000 records...
Found 104500 records...
Found 105000 records...
Found 105500 records...
Found 106000 records...
Found 106500 records...
Found 107000 records...
Found 107500 records...
Found 108000 records...
Found 108500 records...
Found 109000 records...
Found 109500 records...
Found 110000 records...
Found 110500 records...
Found 111000 records...
Found 111500 records...
Found 112000 records...
Found 112500 records...
Found 113000 records...
Found 113500 records...
Found 114000 records...
Found 114500 records...
Found 115000 records...
Found 115500 records...
Found 116000 records...
Found 116500 records...
Found 117000 records...
Found 117500 records...
Found 118000 records...
Found 118500 records...
Found 119000 records...
Found 119500 records...
Found 120000 records...
Found 120500 records...
Found 121000 records...
Found 121500 records...
Found 122000 records...
Found 122500 records...
Found 123000 records...
Found 123500 records...
Found 124000 records...
Found 124500 records...
Found 125000 records...
Found 125500 records...
Found 126000 records...
```

```
Found 126500 records...
Found 127000 records...
Found 127500 records...
Found 128000 records...
Found 128500 records...
Found 129000 records...
Found 129500 records...
Found 130000 records...
Found 130500 records...
Found 131000 records...
Found 131500 records...
Found 132000 records...
Found 132500 records...
Found 133000 records...
Found 133500 records...
Found 134000 records...
Found 134500 records...
Found 135000 records...
Found 135500 records...
Found 136000 records...
Found 136500 records...
Found 137000 records...
Found 137500 records...
Found 138000 records...
Found 138500 records...
Found 139000 records...
Found 139500 records...
Found 140000 records...
Found 140500 records...
Found 141000 records...
Found 141500 records...
Found 142000 records...
Found 142500 records...
Found 143000 records...
Found 143500 records...
Found 144000 records...
Found 144500 records...
Found 145000 records...
Found 145500 records...
Found 146000 records...
Found 146500 records...
Found 147000 records...
Found 147500 records...
Found 148000 records...
Found 148500 records...
Found 149000 records...
Found 149500 records...
Found 150000 records...
Found 150500 records...
Found 151000 records...
Found 151500 records...
Found 152000 records...
Found 152500 records...
Found 153000 records...
```

```
Found 153500 records...
Found 154000 records...
Found 154500 records...
Found 155000 records...
Found 155500 records...
Found 156000 records...
Found 156500 records...
Found 157000 records...
Found 157500 records...
Found 158000 records...
Found 158500 records...
Found 159000 records...
Found 159500 records...
Found 160000 records...
Found 160500 records...
Found 161000 records...
Found 161500 records...
Found 162000 records...
Found 162500 records...
Found 163000 records...
Found 163500 records...
Found 164000 records...
Found 164500 records...
Found 165000 records...
Found 165500 records...
Found 166000 records...
Found 166500 records...
Found 167000 records...
Found 167500 records...
Found 168000 records...
Found 168500 records...
Found 169000 records...
Found 169500 records...
Found 170000 records...
Found 170500 records...
Found 171000 records...
Found 171500 records...
Found 172000 records...
Found 172500 records...
Found 173000 records...
Found 173500 records...
Found 174000 records...
Found 174500 records...
Found 175000 records...
Found 175500 records...
Found 176000 records...
Found 176500 records...
Found 177000 records...
Found 177500 records...
Found 178000 records...
Found 178500 records...
Found 179000 records...
Found 179500 records...
Found 180000 records...
```

```
 Found 180500 records...
 Found 181000 records...
 Found 181500 records...
 Found 182000 records...
 Found 182500 records...
 Found 183000 records...
 Found 183500 records...
 Found 184000 records...
 Found 184500 records...
 Found 185000 records...
 Found 185500 records...
 Found 186000 records...
 Found 186500 records...
 Found 187000 records...
 Found 187500 records...
 Found 188000 records...
 Found 188500 records...
 Found 189000 records...
 Found 189500 records...
 Found 190000 records...
 Found 190500 records...
 Found 191000 records...
 Found 191500 records...
 Found 192000 records...
 Found 192500 records...
 Found 192609 records...
 Imported 192609 records. Simplifying...
```

closing file input connection.

```
clean.business <- business

## Filtering for restaurants
clean.business$Restaurant <- 0
new.business <- mutate(clean.business, Restaurant=grepl("Restaurants", clean.business$categories))
restaurants <- filter(new.business, Restaurant == "TRUE")

## Isolating the important variables
attribute <- restaurants$attributes
restaurant_data <- restaurants
restaurant_data$attributes <- NULL
restaurant_data$Restaurant <- NULL
restaurant_data <- cbind(restaurant_data, attribute)
voi <- c("name", "city", "state", "latitude", "longitude", "stars", "review_count", "RestaurantsTakeOut
         "RestaurantsPriceRange2", "OutdoorSeating", "Alcohol", "categories", "RestaurantsAttire")
rdata <- subset(restaurant_data, select=voi)
```
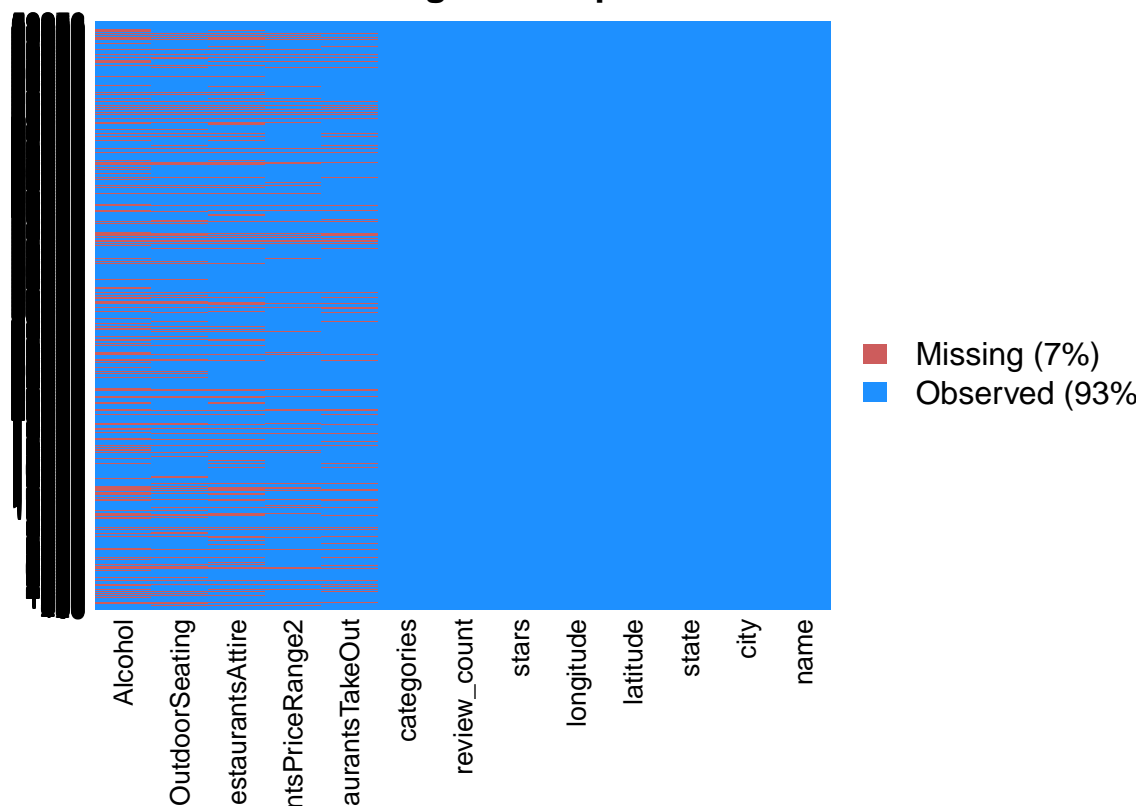
Missingness

Now we look at the missingness in the data, and we decouple it into response and some predictor variables.

```
## Missingness when including all of the restaurant observations
missmap(rdata)
```
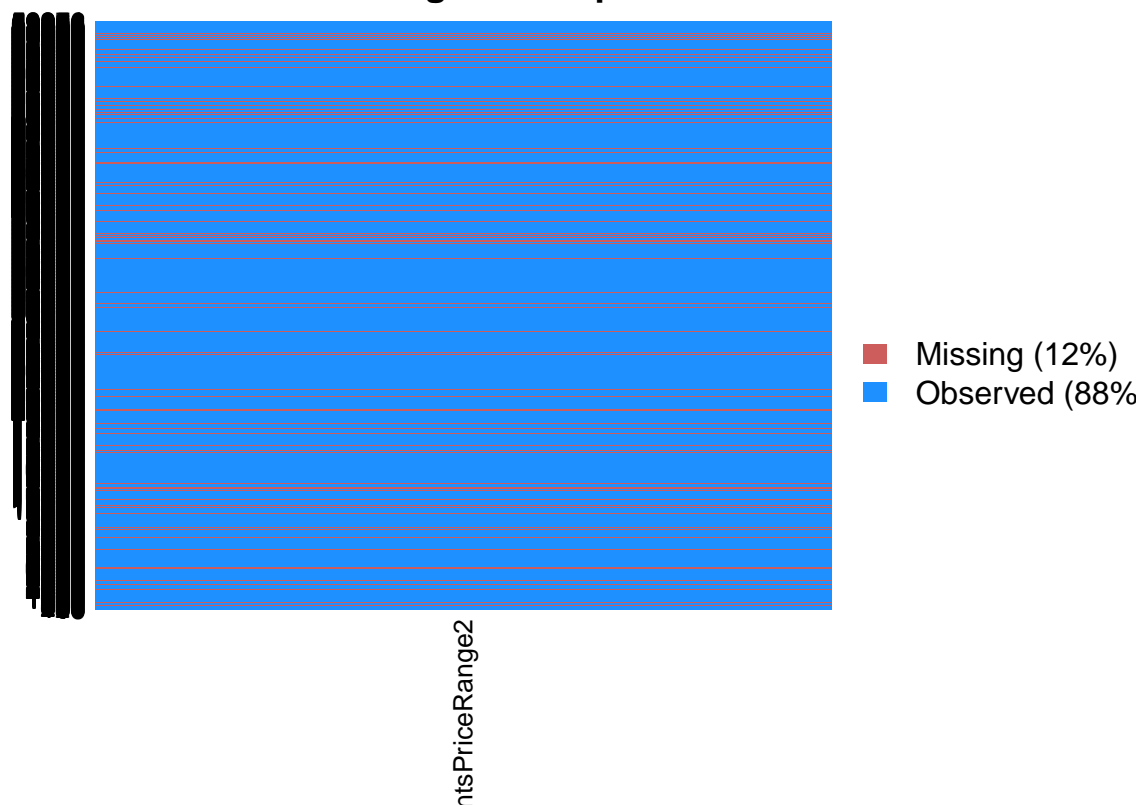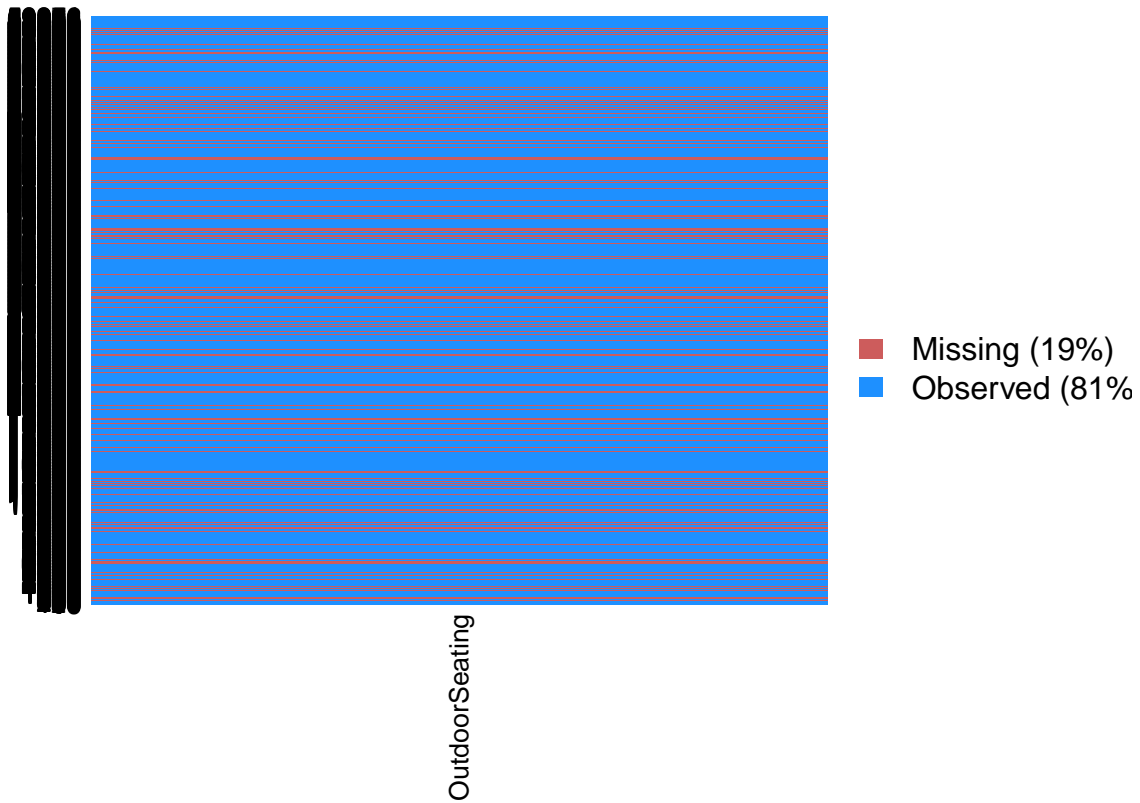
Missingness Map

Legend:
- Missing (7%)
- Observed (93%

Column labels (left to right): Alcohol, OutdoorSeating, estaurantsAttire, ntsPriceRange2, aurantsTakeOut, categories, review_count, stars, longitude, latitude, state, city, name

```
## Missingness in the response
prices <- data.frame(rdata$RestaurantsPriceRange2)
missmap(prices)
```

Missingness Map

Missing (12%)
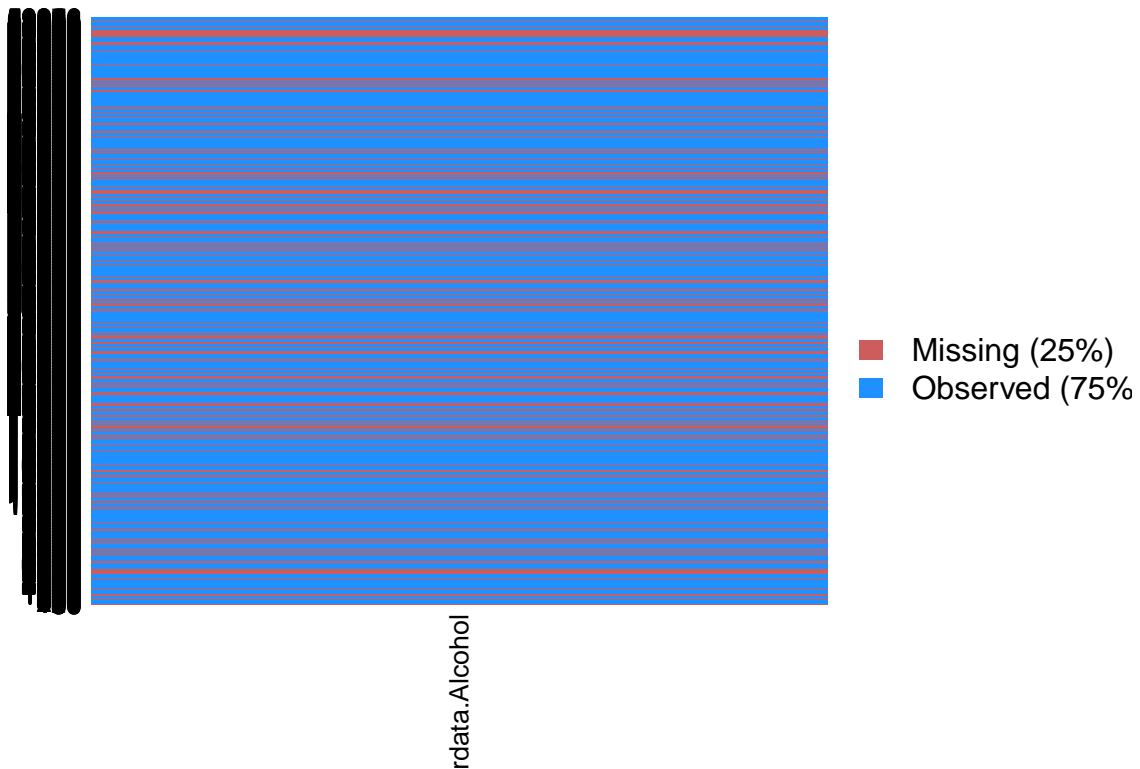Observed (88%

ntsPriceRange2

```
## Missingness in the explanatory variables
alcohol <- data.frame(rdata$Alcohol)
outdoorseating <- data.frame(rdata$OutdoorSeating)
dresscode <- data.frame(rdata$RestaurantsAttire)
missmap(outdoorseating)
```
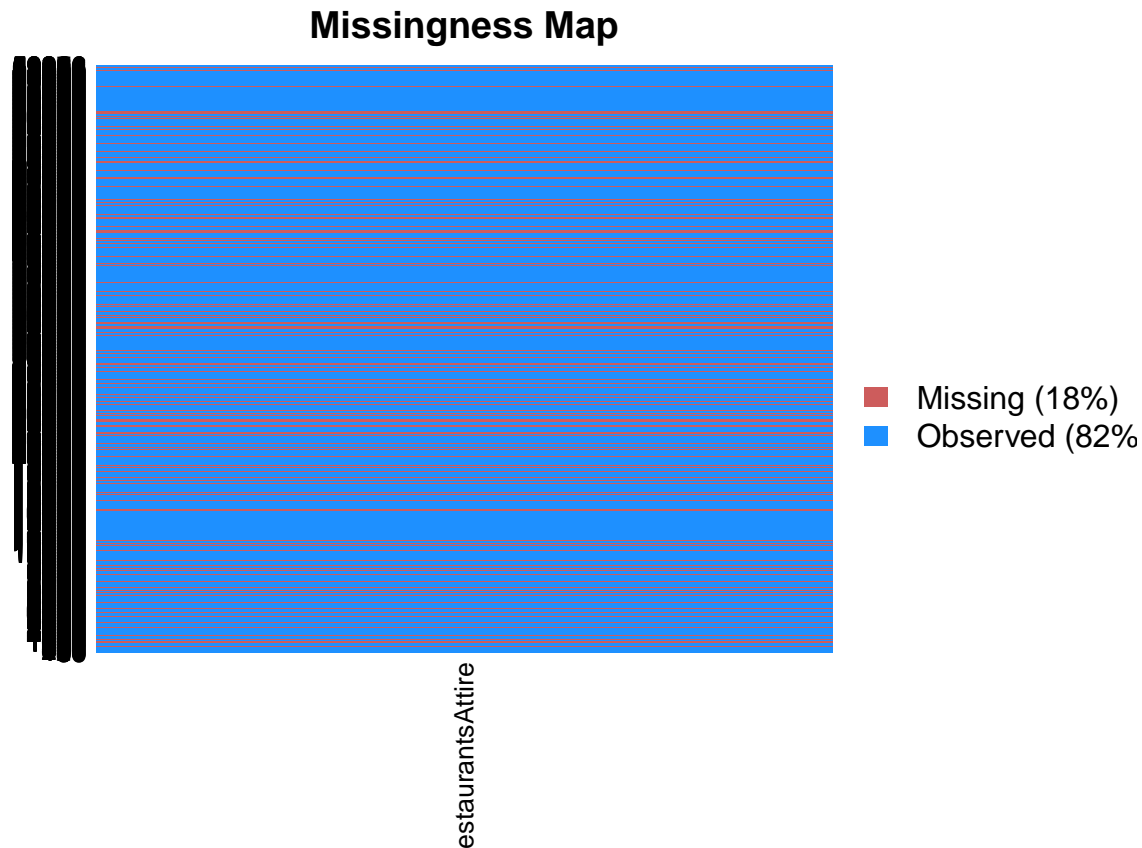
Missingness Map

Missing (19%)
Observed (81%

OutdoorSeating

```
missmap(alcohol)
```

Missingness Map

Missing (25%)
Observed (75%

rdata.Alcohol

```
missmap(dresscode)
```

Missingness Map

The missingness maps reveal that though our overall missingness is minimal, there are a couple things worth noting. In terms of the response variable, it is only 12% which is not too concerning since we still have a value for the major of observations. Further, we notice that the missingness is concentrated around the more niche variables, which are the determinants of the price range. This would be an issue if we considered there to exist a systematic issue in the missingness occurring, however we do not think this is the case. We believe the missingness is randomly distributed, so we assume it is not problematic to simply get rid of observations which do not have values for all of our variables. Thus, the final data we use will contain only complete observations. Now that we have a cursory look at the overall data and know that we will get rid of incomplete observations, we turn to preparing it for a survey of the response and predictive variables.

Preparing Data

```
##Cleaning Data to only include complete observations
rdata.clean <- rdata[complete.cases(rdata),]

## Fixing Variables
rdata.clean$OutdoorSeating <- as.logical(rdata.clean$OutdoorSeating)
rdata.clean$OutdoorSeating <- as.numeric(rdata.clean$OutdoorSeating)

rdata.clean$RestaurantsTakeOut <- as.logical(rdata.clean$RestaurantsTakeOut)
rdata.clean$RestaurantsTakeOut <- as.numeric(rdata.clean$RestaurantsTakeOut)

#Assigning categorical levels 1, 2, or 3 for alcohol
```

```
rdata.clean$Alc <- 0
lvlone <- c("'none'", "u'none'")
lvltwo <- c("'beer_and_wine'", "u'beer_and_wine'")
lvlthree <- c( "'full_bar'", "u'full_bar'")
rdata.clean$Alc <- ordered(rdata.clean$Alcohol, levels = c(lvlone, lvltwo, lvlthree))
rdata.clean$Alc <- as.numeric(rdata.clean$Alc)
rdata.clean$Alc[rdata.clean$Alc == 2] <- 1
rdata.clean$Alc[rdata.clean$Alc == 3] <- 2
rdata.clean$Alc[rdata.clean$Alc == 4] <- 2
rdata.clean$Alc[rdata.clean$Alc == 5] <- 3
rdata.clean$Alc[rdata.clean$Alc == 6] <- 3
# rdata.clean$Alc <- as.factor(rdata.clean$Alc)

# Doing the same for Attire; had to do it differently for some reason
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "'casual'"] <- 1
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "u'casual'"] <- 1
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "'dressy'"] <- 2
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "u'dressy'"] <- 2
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "'formal'"] <- 3
rdata.clean$RestaurantsAttire[rdata.clean$RestaurantsAttire == "u'formal'"] <- 3
# rdata.clean$RestaurantsAttire <- as.factor(rdata.clean$RestaurantsAttire)

rdata.cleaner <- na.omit(rdata.clean)
d1 <- rdata.cleaner[c(-1 , -2, -3, -11, -12 )]
d1 <- d1 %>% rename(Takeout = RestaurantsTakeOut, PriceRange = RestaurantsPriceRange2, Attire = Restaura
d1$Attire[d1$Attire == "None"] <- 1
d1$PriceRange[d1$PriceRange == "None"] <- NA
d1 <- na.omit(d1)
```

Now we have d1, our final data set which we use to analyze the response & predictor variables, as well as for the construction of our models.
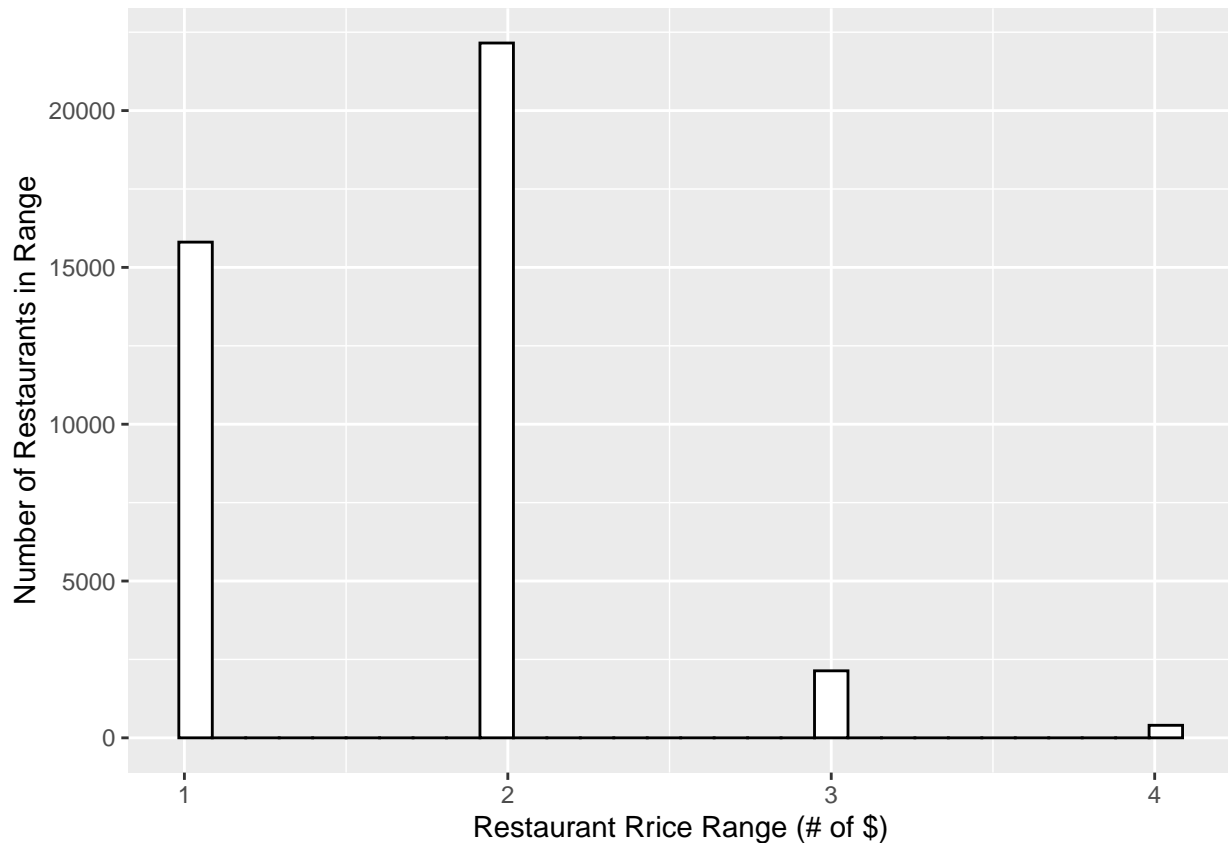

Response Variable Analysis

```
d1$PriceRange <- as.numeric(d1$PriceRange)
ggplot(d1, aes(x=PriceRange)) +
  geom_histogram(color="black", fill="white") +
  xlab("Restaurant Rrice Range (# of $)") +
  ylab("Number of Restaurants in Range")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```
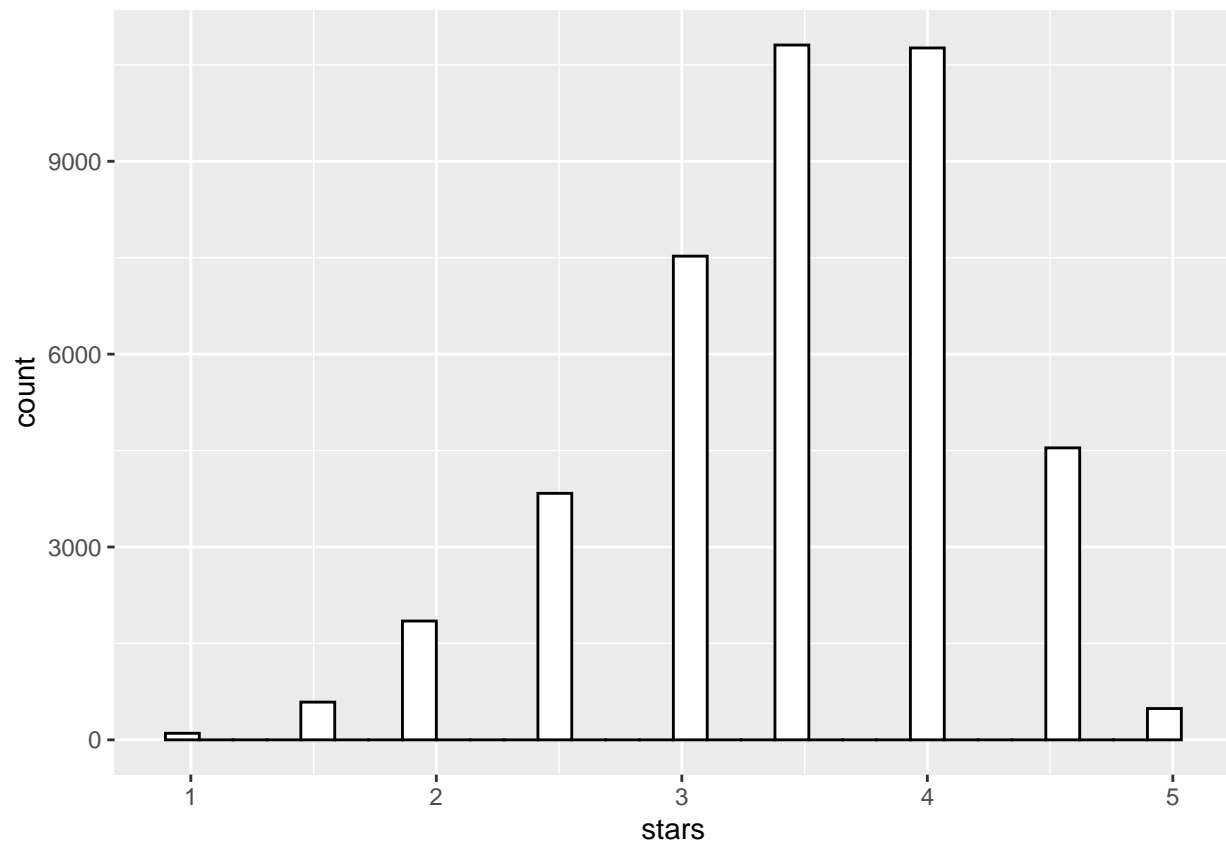
Here we see that our response variable (the price range) is heavily right skewed, with a large concentration of observations around 2 (about 53%). This is worth noting, as we might need to address it in our model construction.
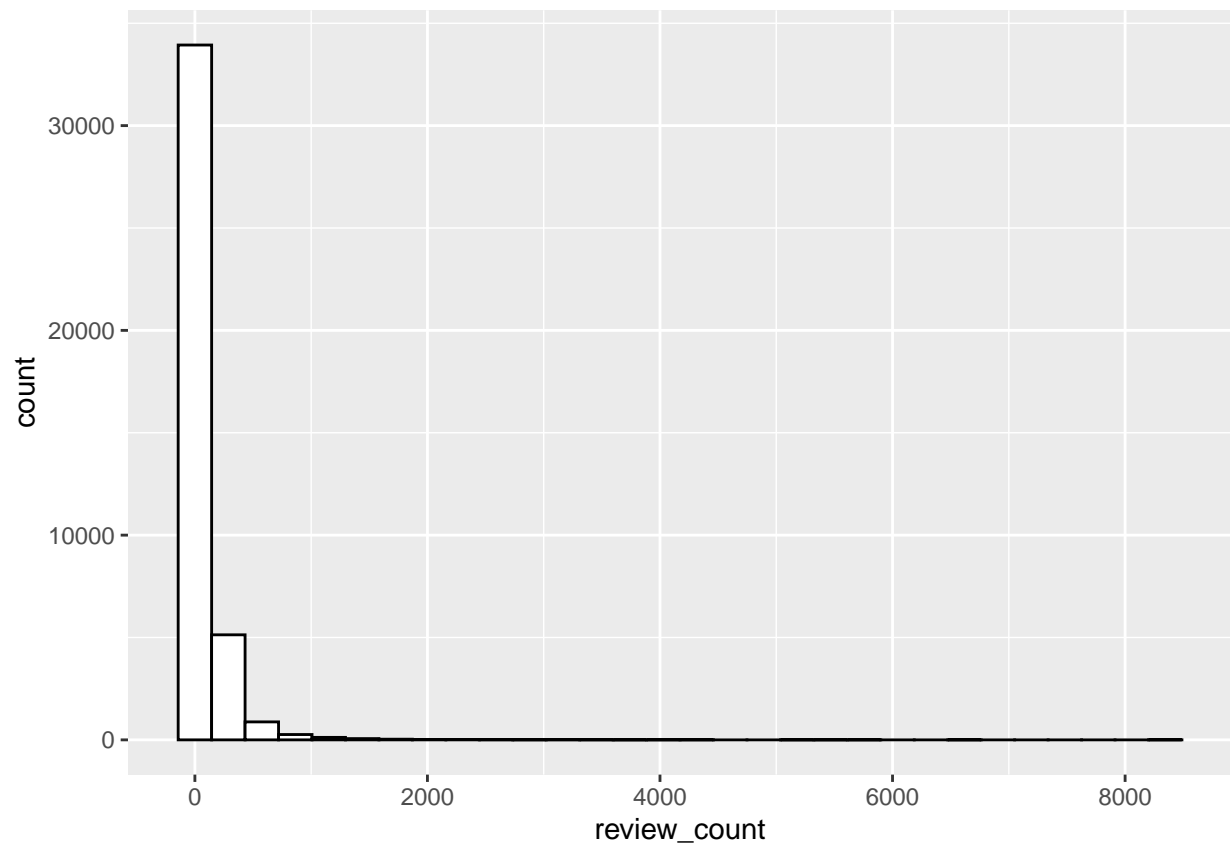
Predictors

```
## Numerical
ggplot(d1, aes(x=stars)) +
  geom_histogram(color="black", fill="white")

## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```
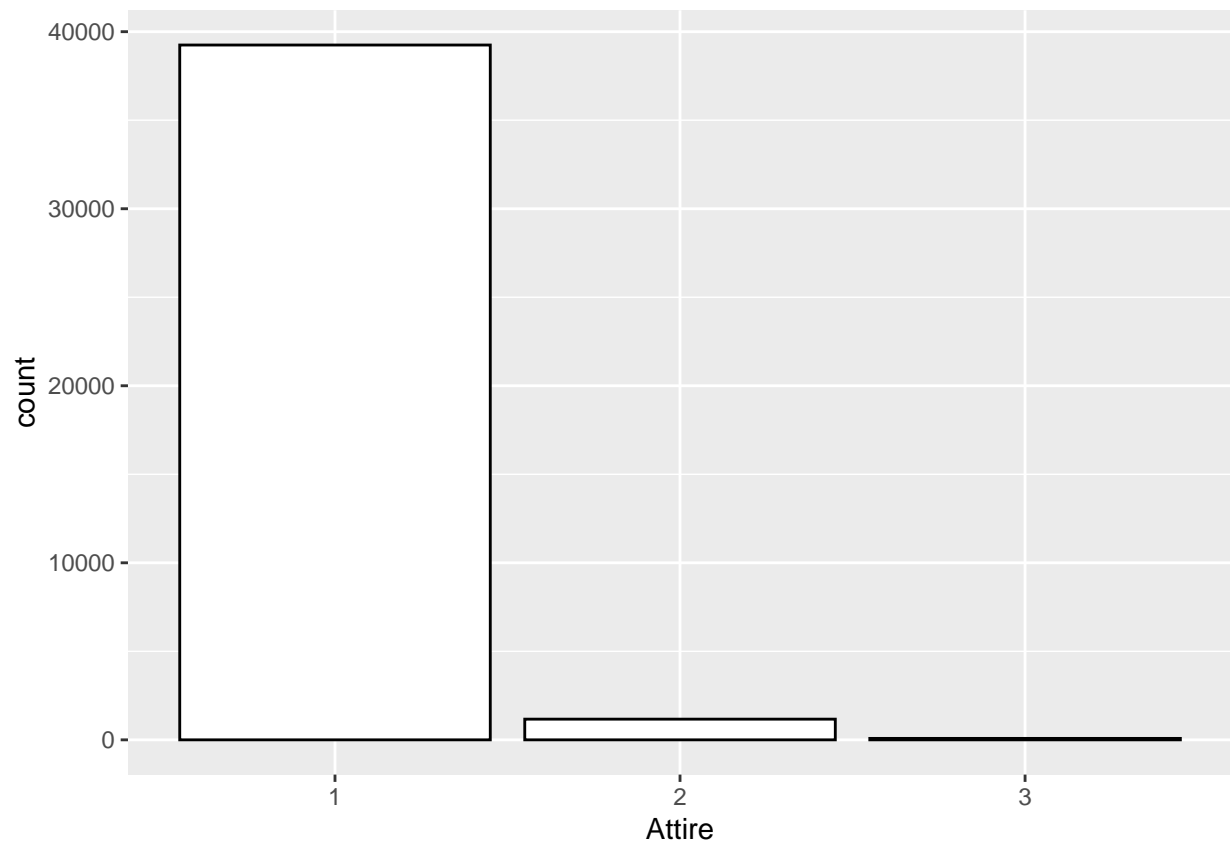
```r
ggplot(d1, aes(x=review_count)) +
  geom_histogram(color="black", fill="white")
```
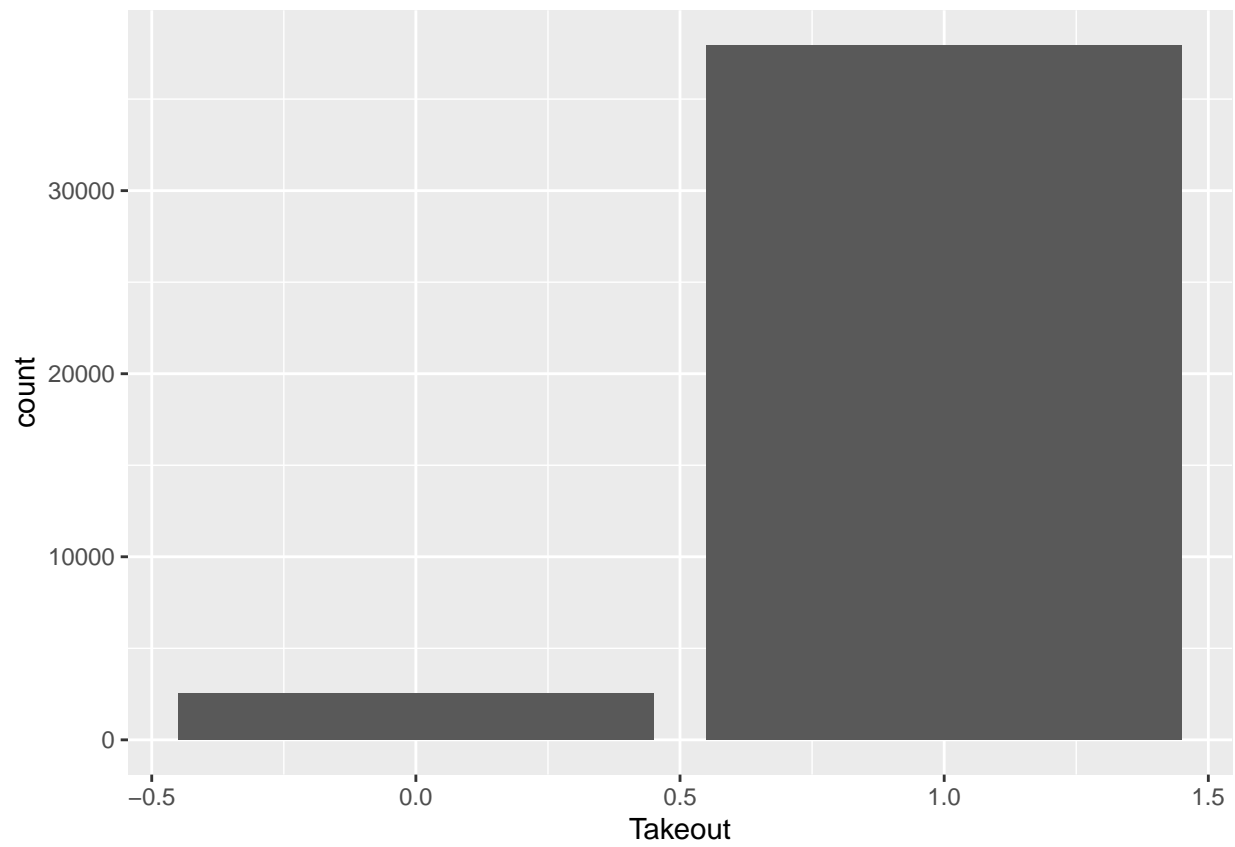
`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
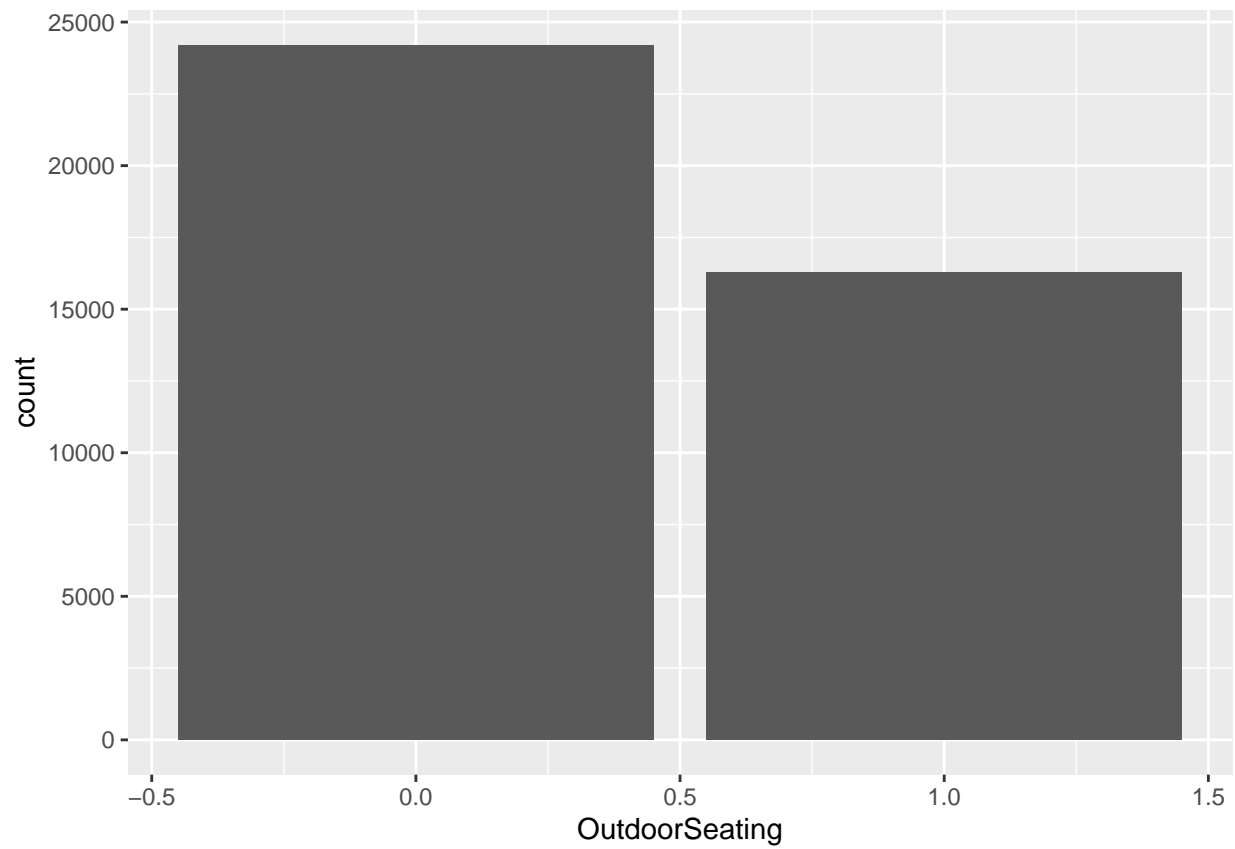
```
## Categorical
ggplot(d1, aes(x=Attire)) +
  geom_bar(color="black", fill="white")
```
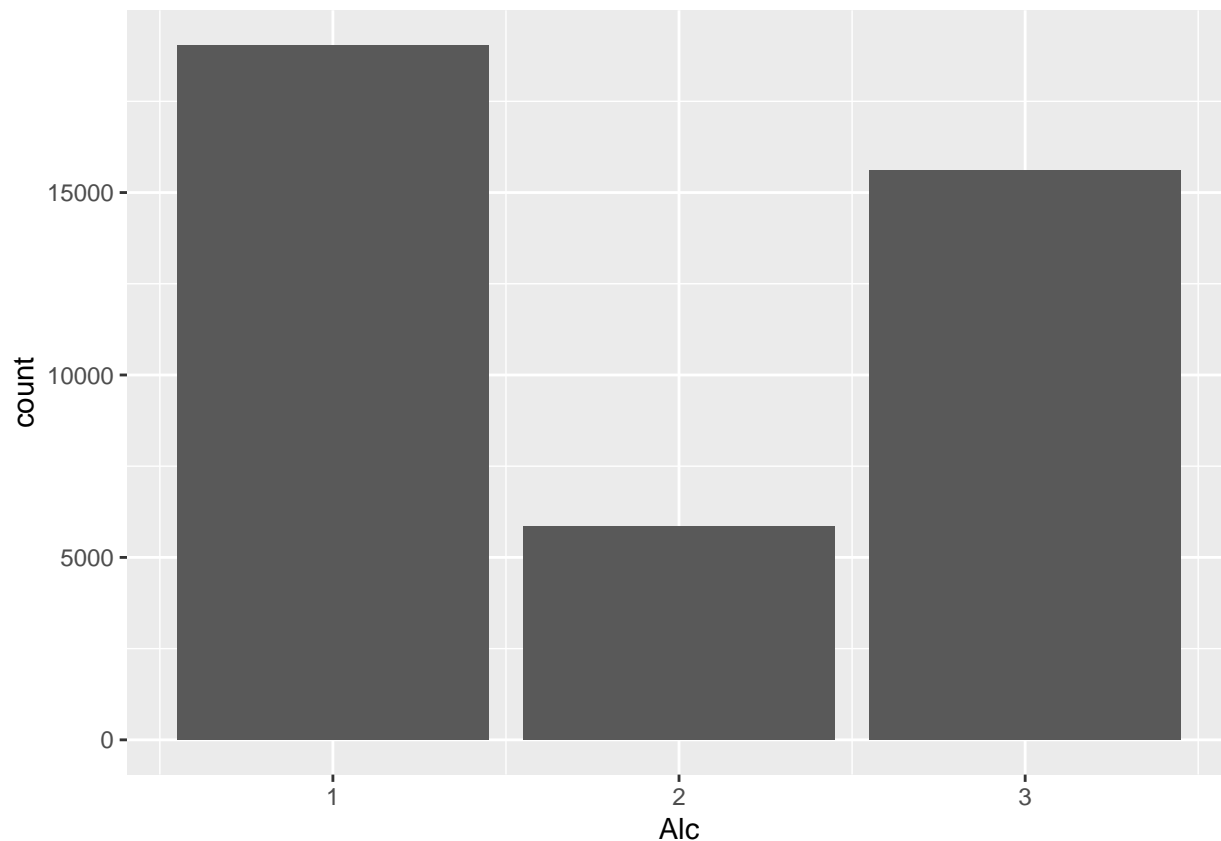
```
ggplot(d1, aes(x=Takeout)) +
  geom_bar()
```

```
ggplot(d1, aes(x=OutdoorSeating)) +
  geom_bar()
```

```
ggplot(d1, aes(x=Alc)) +
  geom_bar()
```

Numerical

There are a couple of interesting things about our predictive variables. The stars variable has a slightly left skewed distribution, with most restaurants falling under the 3.5 - 4 star rating. Review count appears to be centered around a couple of reviews, with very major outliers.

Categorical

It is difficult to make any meaningful assertions from a cursory look at the categorical, however there are a few things worth noting. Attire is heavily concetrated around the first category, "casual", which is intuitive since most restaurants are casual walk-in locations. Take out is also very heavily concentrated around 1 (which means they do Take Out), which is again intuitive since we'd expect most places to do take out. Outdoor seating has a fairly even split, though there are a larger number of restaurants that do not have it available. And alcohol has a similar amount under the categories of no alcohol & full bar, with a fairly smaller number under "beer and wine".

Modeling

Descriptive Classification Tree

```
library(rpart)
library(rpart.plot)
d1$PriceRange <- as.factor(d1$PriceRange)
```

```
m1 <- rpart(PriceRange ~. , data = d1,
  control = rpart.control(minsplit = 2))
plot1<-prp(m1)
```
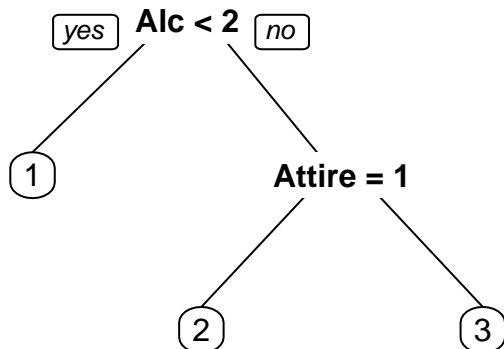

Preparing the Data: Training and Test

Before building our supervised learning models, we must first create our training and test sub datasets.

```
s.size <- floor(0.75 * nrow(d1))
set.seed(10)
train.data <- sample(seq_len(nrow(d1)), size = s.size)
train <- d1[train.data, ]
test <- d1[-train.data, ]
```

Linear Model

Though this is a classification exercise, the first model we developed was a linear model with rounding, as follows:

```
d1$PriceRange <- as.numeric(d1$PriceRange)
m2 <- lm(PriceRange ~.,  data = train)

## Warning in model.response(mf, "numeric"): using type = "numeric" with a
## factor response will be ignored

## Warning in Ops.factor(y, z$residuals): '-' not meaningful for factors
## Predicting Price
predicted.price <- predict(m2, newdata = test)
predicted.price.rounded <- round(predicted.price, digits = 0)
yp <- predicted.price.rounded

## Misclassification rate:
yt <- test$PriceRange

mcr2 <- table(yp,yt)

1-sum(diag(mcr2))/sum(mcr2)

## [1] 0.2852346
```

A couple of things are worth noting about our process in building this model. First, we aimed to address the issue around the right skewedness of the data by performing some transformations. However, none of

these changes significantly improved the predictive strength of the model, thus we chose to not keep any of the transformations. Regarding the explanatory variables we chose, we also tried a variety of models with different variables and opted to choose this one.

With this model we find a misclassification rate of 27.52%. While this modeling method is not necessarily made for this exercise, it is quite robust since it identifies just over 70% of restaurant price ranges correctly. However, we wished to further develop our understanding of the relationship between restaurant prices and the variables, so we continue our explorations with an Ordinal Regression, which should more aptly tackle the classification nature of this problem.

Ordinal Regression

```
class <- ordered(d1$PriceRange, levels= c("1", "2", "3", "4"))
d1$class <- class
set.seed(20)
olr.train.data <- sample(seq_len(nrow(d1)), size = s.size)
olr.train <- d1[olr.train.data, ]
olr.test <- d1[-olr.train.data, ]
olr.train$PriceRange = NULL
olr.test$PriceRange = NULL

m3 <- polr(class ~. ,  data = olr.train,  Hess = TRUE)
m3
```

```
## Call:
## polr(formula = class ~ ., data = olr.train, Hess = TRUE)
##
## Coefficients:
##         latitude        longitude            stars   review_count         Takeout
##      0.095364422      0.001901830      0.115547402      0.001083351     -1.561996117
## OutdoorSeating           Attire2          Attire3              Alc
##     -0.058069205      3.139967674      1.163672626      1.411391614
##
## Intercepts:
##        1|2        2|3        3|4
##   4.599408   9.417342 12.236270
##
## Residual Deviance: 39072.24
## AIC: 39096.24
```

```
predicted.price3 <- predict(m3, newdata = olr.test)
predicted.price3
```

```
##     [1] 2 2 2 1 2 1 2 2 1 1 2 2 1 2 2 1 1 2 1 1 2 2 2 2 1 2 1 2 2 2 1 2 2 2
##    [35] 1 2 2 2 2 2 2 2 2 2 2 2 2 1 1 2 1 1 2 2 2 1 1 1 2 2 1 2 2 2 1 1 1 1
##    [69] 1 2 1 1 2 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 1 2 2 1 2 1 2 1 2 2
##   [103] 2 1 1 2 3 1 1 1 1 2 2 1 1 1 1 2 1 2 1 2 2 1 1 1 3 2 1 2 2 2 1 1 2 2
##   [137] 2 2 1 2 2 1 2 2 2 1 2 2 2 2 1 1 2 1 1 1 1 2 2 2 2 1 2 2 1 2 1 2 2 1
##   [171] 2 1 2 1 1 2 2 2 2 2 2 1 2 1 2 2 1 2 2 2 2 1 1 2 1 2 2 1 2 1 1 1 1 2
##   [205] 1 2 2 1 1 1 1 3 2 1 2 2 1 2 2 1 1 2 2 1 1 2 2 2 2 1 3 2 1 2 1 1 1 2
##   [239] 2 2 1 2 2 1 2 2 1 2 2 1 1 1 1 2 2 2 2 1 1 2 2 2 2 1 1 2 1 2 2 1 1 2
##   [273] 2 1 2 2 2 2 2 1 2 1 1 2 2 2 1 1 2 1 1 2 1 1 2 1 2 2 1 2 2 2 1 2 2 2
##   [307] 1 2 2 1 2 1 2 2 2 1 2 1 3 1 1 1 1 2 1 2 2 1 1 2 2 1 1 1 1 2 2 2 1 1
##   [341] 2 1 2 3 1 1 1 1 2 2 1 2 2 2 1 1 2 2 2 3 1 1 1 2 2 2 2 1 2 2 2 1 1 3
```

```
##    [375] 2 2 1 4 2 2 1 1 2 1 2 1 2 1 2 1 1 3 1 2 2 1 2 2 2 2 2 1 2 1 2 1 1 2
##    [409] 2 2 1 2 2 1 1 1 2 2 2 1 2 2 1 1 2 1 1 2 1 1 1 2 1 2 2 2 1 1 1 1 1 2
##    [443] 2 1 2 2 2 1 2 2 2 1 2 1 1 2 2 1 1 1 1 2 1 2 2 1 2 1 2 1 2 2 1 1 2 1
##    [477] 2 2 1 1 1 1 1 2 2 2 2 1 2 2 2 1 2 2 2 2 2 2 2 1 2 1 1 2 1 2 1 2 2 1
##    [511] 2 3 1 1 2 2 2 2 1 1 2 1 1 1 2 2 2 1 2 2 2 1 1 2 1 2 2 1 1 1 1 1 1 1
##    [545] 1 1 2 1 2 2 1 1 1 2 2 1 1 1 2 2 1 1 1 1 1 2 2 1 1 1 2 1 1 1 2 2 2 1
##    [579] 2 2 1 2 1 1 1 1 2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2 2 1
##    [613] 2 1 1 2 2 1 1 2 3 1 1 1 1 2 1 2 1 2 2 2 1 1 1 2 1 1 2 1 1 1 1 2 1 2
##    [647] 1 2 1 2 1 2 2 1 2 2 1 1 1 1 1 2 1 1 1 2 2 2 2 1 1 2 3 1 2 2 2 1 1 1
##    [681] 4 2 1 1 2 2 1 1 2 1 2 2 1 2 2 2 1 1 2 2 2 2 1 1 2 1 1 2 2 1 1 1 1 1
##    [715] 1 1 2 2 2 3 1 2 2 2 1 2 2 1 2 2 2 2 2 1 2 1 1 2 2 2 2 2 2 1 2 1 1 1
##    [749] 2 1 1 2 2 1 2 2 2 1 2 1 2 1 1 2 2 1 2 1 2 2 2 2 2 1 1 2 2 2 1 1 2 1
##    [783] 1 2 2 2 1 2 1 2 2 2 2 2 1 1 2 1 2 2 2 1 1 1 1 2 2 2 1 2 1 2 2 2 2 2
##    [817] 2 2 3 2 1 2 2 2 1 2 3 2 2 2 2 1 2 2 1 1 2 2 2 1 1 1 1 2 2 2 2 2 1 2
##    [851] 1 1 1 2 2 1 1 2 1 1 1 2 2 2 1 2 3 2 2 1 2 1 2 2 1 2 2 1 1 1 1 2 2 2
##    [885] 2 1 2 1 2 1 2 1 1 3 2 2 2 2 2 1 1 2 1 1 2 2 1 1 2 1 2 2 1 3 2 2 1
##    [919] 2 1 2 1 2 2 2 2 2 1 2 2 2 1 1 2 1 1 1 2 2 1 2 3 2 2 2 1 1 2 1 1 1 2
##    [953] 2 2 2 2 2 1 2 2 2 2 2 1 1 2 2 1 1 2 2 1 1 2 2 1 2 1 1 2 1 2 2 1 2 2
##    [987] 1 2 3 1 1 2 2 2 1 1 1 2 2 1 1 2 2 2 1 2 2 1 2 1 2 1 1 2 1 2 2 2 2 2
##   [1021] 1 1 2 1 1 2 2 3 1 2 2 1 2 2 1 2 1 2 1 2 2 2 2 2 2 1 1 1 1 1 2 2 1 2
##   [1055] 2 1 2 2 1 1 2 2 2 2 2 1 2 1 1 1 1 1 2 1 1 1 1 2 2 2 1 1 1 2 1 1 1 1
##   [1089] 2 1 1 2 1 2 2 2 2 2 2 1 1 2 1 2 2 2 1 2 2 1 1 2 1 2 1 1 2 1 2 1 2 1
##   [1123] 2 2 2 2 2 2 1 1 1 2 1 2 2 2 2 1 1 2 2 1 2 1 2 2 1 2 1 1 1 1 2 3 1 1
##   [1157] 1 2 2 2 1 1 3 1 2 2 2 2 2 2 1 2 2 2 1 1 2 2 1 2 1 1 2 1 2 2 1 2 2
##   [1191] 2 1 2 1 1 1 3 2 1 1 2 1 1 2 2 3 2 2 2 1 1 2 2 1 2 1 2 2 1 2 1 1 2 2
##   [1225] 2 2 1 2 2 1 1 2 1 2 1 2 2 1 2 1 1 2 1 1 1 1 2 2 2 2 2 2 2 2 3 2 2
##   [1259] 1 1 2 2 1 2 2 1 2 1 1 1 1 2 2 1 2 2 3 1 2 2 1 2 2 2 1 1 1 2 1 1 2 2
##   [1293] 1 1 1 2 1 2 1 1 1 1 4 2 2 1 2 2 1 1 1 2 1 2 1 1 2 2 1 1 2 2 1 2 1 1
##   [1327] 2 1 2 1 1 1 2 2 1 2 2 1 1 1 2 2 2 2 2 4 2 1 2 1 1 1 2 2 2 1 2 2 1 2
##   [1361] 2 1 1 2 1 1 2 1 3 1 1 2 3 2 1 1 1 1 1 2 2 3 2 2 1 1 1 1 2 2 1 1 2 2
##   [1395] 2 2 1 2 1 2 2 1 2 2 3 3 2 1 2 1 2 1 2 1 2 2 1 2 1 1 1 1 1 2 1 2 2 2
##   [1429] 1 1 2 2 2 2 2 2 1 2 2 1 2 1 2 1 1 1 1 1 2 2 2 2 2 1 1 2 2 1 1 2 2 2
##   [1463] 2 1 2 2 1 2 2 2 2 2 2 2 2 2 1 2 2 1 2 2 1 2 1 2 1 2 1 1 3 2 2 2 2 3
##   [1497] 2 2 1 1 1 2 1 2 2 2 2 1 2 2 1 1 2 1 2 1 2 2 3 1 2 2 2 1 2 1 1 1 1 1
##   [1531] 2 2 2 2 2 2 2 3 2 2 2 2 2 2 1 1 1 2 2 1 2 1 2 1 2 2 2 1 1 2 1 1 1 1
##   [1565] 1 1 2 1 2 2 1 1 2 1 1 2 1 1 2 1 2 2 1 2 1 2 1 2 2 2 2 2 1 1 1 2 1 2 1 2
##   [1599] 1 1 1 1 2 1 2 1 1 2 2 1 2 1 2 1 2 2 2 2 1 2 1 2 1 1 1 2 1 2 2 2 2 2 2 2
##   [1633] 2 2 2 1 2 1 2 2 1 1 1 1 1 2 1 2 2 1 2 1 1 3 2 1 2 2 2 1 2 2 1 2 2 1
##   [1667] 1 2 2 2 2 2 2 2 1 1 2 1 2 1 2 2 2 2 1 1 2 1 1 2 1 1 2 2 1 2 2 2 2 1 2
##   [1701] 1 3 1 2 2 1 1 2 1 2 1 2 1 2 1 2 2 2 1 2 2 1 1 1 2 2 2 3 2 2 1 2 2 2 1 1
##   [1735] 2 2 2 1 2 2 2 1 3 2 1 2 2 2 2 2 2 2 3 2 2 2 2 1 2 2 1 1 2 2 1 2 2 2
##   [1769] 2 1 1 1 1 2 2 1 2 2 2 2 2 1 2 2 2 2 1 1 2 2 2 1 1 2 2 1 2 1 1 1 1 1
##   [1803] 3 2 2 2 1 2 1 2 2 3 2 1 2 2 2 1 2 2 1 1 2 2 1 2 1 2 2 2 2 1 1 1 2 2
##   [1837] 2 2 1 1 1 1 2 1 1 2 2 2 2 1 1 2 1 1 1 1 2 1 1 2 2 2 2 2 2 2 1 1 2 2 1
##   [1871] 1 1 1 1 1 1 2 1 1 1 2 2 2 1 1 2 2 2 1 2 2 1 1 1 1 2 2 1 1 1 2 2 2
##   [1905] 1 3 2 2 2 1 2 1 2 2 2 2 1 1 1 2 2 2 1 1 2 1 1 2 3 1 2 1 2 2 2 2 2 2
##   [1939] 1 2 1 1 2 2 1 2 1 1 2 2 1 1 2 2 2 1 1 2 1 2 1 2 1 2 1 2 2 2 1 2 2 2 2 2
##   [1973] 1 1 1 2 2 1 2 1 2 2 1 2 2 2 2 1 2 2 1 1 2 2 1 2 1 2 2 2 2 1 2 1 1 2 1
##   [2007] 1 2 1 2 2 2 3 2 2 2 2 2 1 1 2 1 2 2 2 2 2 2 2 2 2 1 1 1 2 1 2 2 2
##   [2041] 1 1 1 2 1 2 1 1 1 1 1 2 1 2 1 1 2 2 1 1 2 2 2 2 2 2 2 2 1 1 1 1 1 1
##   [2075] 1 2 2 2 1 1 2 2 2 1 2 1 2 1 1 1 2 2 2 2 2 1 2 2 2 1 2 2 2 1 2 1 2 1
##   [2109] 2 2 1 2 1 1 1 2 1 1 1 2 2 1 2 1 1 2 1 1 2 2 2 1 1 1 2 2 1 2 2 1 1 1
##   [2143] 2 2 2 2 2 2 2 2 1 1 1 3 1 1 1 1 2 1 2 2 2 1 1 2 2 1 1 1 2 1 2 1 2 1
##   [2177] 1 1 2 1 2 1 1 2 2 1 1 1 1 1 1 1 1 2 1 1 2 1 1 2 2 1 2 2 1 2 2 2 1 1 1
```

25

```
## [2211] 2 1 1 2 1 1 2 1 1 2 1 2 2 2 1 1 2 2 2 2 1 1 1 1 1 1 2 2 1 1 1 1 2 2
## [2245] 1 2 2 2 2 1 2 2 3 2 2 1 1 2 2 1 1 1 2 2 1 1 2 2 1 2 1 1 2 2 1 2 2 2
## [2279] 2 1 2 3 2 2 1 2 1 1 2 1 1 1 2 1 1 2 1 2 2 2 2 2 2 2 2 1 1 1 1 2 2 1
## [2313] 2 1 2 2 2 2 1 1 1 1 2 2 1 1 1 1 1 2 1 2 2 1 2 1 2 1 1 2 2 1 1 1 1 2
## [2347] 2 2 1 1 2 2 1 1 1 1 1 2 2 1 1 2 2 2 1 1 1 1 1 1 2 2 2 2 1 2 1 2 1
## [2381] 1 2 2 1 2 4 1 2 2 1 2 2 3 2 2 1 1 2 2 1 1 1 1 2 2 2 1 1 1 2 1 1 2 1
## [2415] 2 2 1 2 1 1 2 2 1 2 1 1 1 1 2 1 1 1 1 2 2 2 1 1 1 2 1 2 2 1 2 2 1 2
## [2449] 2 2 2 1 1 2 2 2 1 2 1 2 1 1 1 2 1 3 1 1 2 2 2 1 2 1 1 2 2 2 2 2 3 2
## [2483] 2 4 1 1 3 2 1 1 2 2 2 2 1 1 2 2 2 2 2 1 2 2 2 2 1 1 2 2 2 2 1 1 2
## [2517] 1 2 1 1 1 1 2 1 1 2 2 2 2 2 2 1 2 1 2 2 2 3 2 2 3 1 2 2 2 1 1 1 2
## [2551] 2 2 1 2 1 1 1 2 2 2 2 2 1 1 1 1 1 2 1 1 2 2 2 1 1 1 2 1 1 2 2 1 2
## [2585] 2 1 2 1 1 1 2 1 1 2 1 1 2 2 2 1 2 2 2 1 1 1 2 2 2 1 1 1 2 3 1 3 2 2
## [2619] 1 2 2 1 2 2 2 2 1 1 2 2 1 2 2 1 2 2 1 2 2 2 1 2 2 2 2 1 1 1 1 1 2 2
## [2653] 2 2 1 2 2 1 1 2 2 2 2 1 3 1 1 2 2 1 2 2 2 2 1 1 1 1 1 1 3 2 3 2 2 1
## [2687] 1 2 2 1 2 1 2 2 1 2 2 2 2 1 2 1 1 2 1 2 2 2 1 2 2 2 1 2 1 2 1 2 1 2
## [2721] 1 1 2 2 2 2 1 2 2 2 2 2 2 2 2 2 1 1 1 2 2 1 1 2 2 1 2 2 1 1 2 2 2
## [2755] 1 2 1 1 1 2 1 2 1 2 2 2 2 1 2 2 1 2 1 1 1 1 1 2 1 1 2 2 2 1 2 1 1 1
## [2789] 2 1 2 2 2 1 2 2 1 2 1 1 2 1 2 2 2 2 1 1 1 1 2 2 2 2 4 2 2 1 1 1 3 3
## [2823] 2 1 2 2 1 2 1 2 2 1 1 1 1 1 1 2 2 3 2 1 1 1 2 1 1 2 2 1 3 2 2 1 1 2
## [2857] 2 2 1 2 2 2 1 2 2 1 2 1 1 1 2 1 1 1 2 1 1 2 1 2 2 2 2 2 2 2 1 2 2 1
## [2891] 2 1 1 2 1 1 1 2 2 2 1 2 1 2 2 2 2 1 1 1 2 1 1 1 1 2 1 1 1 1 1 2 2 2
## [2925] 2 2 2 2 1 2 2 1 2 2 2 1 2 2 2 2 2 1 1 1 1 1 2 3 1 2 1 1 1 1 2 2 2 2
## [2959] 2 2 2 2 1 1 2 2 1 1 1 1 2 1 1 1 2 1 2 1 1 1 2 1 1 2 2 2 1 1 1 1 2 1
## [2993] 2 1 2 1 2 2 1 1 2 2 2 1 3 2 2 2 1 2 2 2 2 2 1 2 1 1 1 2 2 1 2 1 1 1
## [3027] 2 1 1 2 1 3 1 2 1 1 1 2 2 2 1 1 2 2 2 2 1 1 1 1 1 2 1 2 2 1 2 2 1 1
## [3061] 2 1 2 1 2 1 1 1 3 2 2 1 2 1 1 1 2 1 2 2 2 1 2 1 2 1 1 2 1 1 2 1 1 2
## [3095] 1 2 1 1 2 1 1 2 2 1 2 2 2 2 2 1 1 2 1 1 2 1 2 2 1 1 2 2 2 1 1 2 2 1
## [3129] 2 2 1 1 1 2 2 1 2 1 2 1 2 1 1 1 3 2 1 1 1 1 2 2 1 1 2 1 2 2 2 2 2 1
## [3163] 4 1 3 2 3 2 2 1 2 1 1 1 1 2 1 2 1 2 2 2 2 1 2 2 1 2 2 1 1 2 2 2 2 1
## [3197] 1 1 2 2 1 2 1 2 2 1 1 2 2 1 2 1 1 2 1 2 2 2 1 1 2 2 2 1 2 2 2 2 2 2
## [3231] 1 2 2 1 2 2 2 2 2 1 2 2 2 1 2 1 2 2 1 2 1 2 1 2 2 1 1 2 1 2 1 1 2 1
## [3265] 2 2 2 1 1 1 2 2 2 2 1 1 3 2 2 2 2 1 2 1 1 2 3 2 2 1 1 1 1 1 1 2 2 2
## [3299] 2 1 1 2 2 2 1 2 1 1 1 2 2 2 1 1 2 1 1 2 3 1 2 1 2 1 1 2 1 1 2 1 2 1
## [3333] 2 2 2 2 1 1 2 1 2 1 1 1 2 2 1 2 2 1 2 2 1 2 1 1 2 1 1 2 1 2 2 2 1 3
## [3367] 1 2 1 2 1 2 1 2 2 2 2 2 2 1 1 2 1 1 2 1 2 2 2 2 1 1 2 2 2 1 1 1 2 2
## [3401] 1 2 1 2 1 2 2 1 2 2 2 2 1 1 2 1 1 2 1 2 1 2 1 2 2 2 2 2 1 2 1 2 2 2
## [3435] 2 2 1 2 2 1 1 2 2 2 1 2 2 1 1 2 2 2 1 3 1 1 2 2 2 1 2 2 1 3 1 1 2 2
## [3469] 1 1 2 2 1 2 1 1 1 2 2 2 2 1 2 2 2 1 1 2 2 2 2 1 1 1 1 2 1 1 1 2 1
## [3503] 2 2 2 1 1 2 2 2 2 2 2 2 1 2 2 1 2 2 1 3 2 2 1 2 1 1 3 2 1 2 1 2 2 1
## [3537] 1 2 2 2 1 2 1 2 1 2 1 2 1 1 1 2 2 2 2 3 1 2 1 2 2 1 2 1 2 2 1 2 1 1
## [3571] 1 2 2 2 1 2 1 2 2 1 2 1 1 2 1 1 2 1 1 2 2 1 2 2 2 3 1 2 2 1 1 1 2 1
## [3605] 1 1 2 1 1 1 1 2 2 1 2 1 1 2 1 1 2 2 2 2 2 2 2 2 1 1 2 1 1 2 2 1 1 2
## [3639] 2 2 2 1 1 1 2 1 2 1 2 1 1 1 1 2 1 1 2 1 1 1 2 2 2 2 1 1 1 2 2 1 1 2
## [3673] 2 2 1 1 1 2 1 1 2 2 2 1 2 1 2 1 1 1 2 2 2 3 1 1 2 1 1 2 2 3 2 1 2 2
## [3707] 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 1 2 1 1 1 2 1 1 1 2 2 2 2 2 2 2 2 1 1
## [3741] 1 3 2 2 2 1 1 2 2 2 1 1 1 1 1 1 2 2 2 1 2 2 2 2 1 2 2 1 2 1 1 2 1 3
## [3775] 2 2 2 1 2 1 2 1 1 1 1 2 2 2 2 3 2 1 1 1 1 2 2 2 2 1 1 2 1 1 2 1 2 1
## [3809] 1 2 2 1 2 1 2 2 2 1 1 2 2 2 1 2 2 1 1 2 2 2 2 1 1 2 1 2 1 2 1 1 2 2 1
## [3843] 1 1 1 2 1 2 2 2 1 2 2 2 1 2 2 1 3 1 2 2 1 1 1 2 1 2 2 3 2 1 2 2 1 1
## [3877] 1 2 3 1 2 1 1 2 2 2 2 1 2 1 2 2 2 2 2 1 2 1 2 2 1 2 2 2 2 2 1 1 1 1
## [3911] 2 1 1 2 2 2 2 2 1 2 1 1 1 2 2 2 2 2 2 2 2 3 1 1 1 2 2 2 2 2 1 1 1 2
## [3945] 1 2 1 1 1 1 1 2 3 1 2 1 2 1 2 2 2 2 2 1 1 2 2 2 2 2 2 1 2 2 2 1 1 1
## [3979] 2 2 2 2 3 2 2 2 1 2 3 2 2 2 2 2 1 2 2 1 2 1 1 2 2 2 2 1 2 2 2 2 1 2
## [4013] 1 2 1 2 2 2 2 1 2 2 1 1 1 1 1 1 1 2 2 3 1 2 1 1 1 2 2 2 2 1 1 3 1 1
```

26

```
## [4047] 1 2 1 1 1 2 2 1 1 1 3 2 1 1 1 1 1 3 2 1 2 2 2 2 1 2 2 2 2 1 1 1 1 1
## [4081] 1 2 2 2 1 1 1 2 2 2 1 1 1 1 3 2 2 2 1 1 2 2 2 1 2 1 1 1 2 2 1 1 2 1
## [4115] 1 1 2 2 1 1 2 1 2 2 1 2 2 1 2 2 2 2 2 1 2 2 3 1 2 1 2 2 1 1 1 1 2 2
## [4149] 2 1 1 2 2 3 1 1 1 2 1 1 2 1 1 2 2 2 2 2 1 1 1 2 2 1 1 1 1 2 1 2 1
## [4183] 1 2 1 1 1 1 2 2 1 1 2 2 1 2 1 1 2 2 2 2 2 1 2 1 1 2 2 2 1 2 2 2 2
## [4217] 2 2 2 2 1 1 1 2 1 2 2 2 2 2 1 2 2 2 1 1 1 2 2 2 2 2 1 2 1 2 2 1 1 2
## [4251] 2 1 2 1 1 2 2 1 2 1 1 2 1 2 2 1 1 2 2 2 2 2 1 3 2 2 2 1 1 2 1 1 2 2
## [4285] 2 2 1 1 2 2 2 1 2 2 1 2 1 1 1 3 2 1 1 2 2 2 1 1 2 2 1 2 2 1 2 2 2 2
## [4319] 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 1 2 1 1 2 2 3 2 1 2 2 1 1 2 2 2 2 1 2
## [4353] 2 1 2 1 1 2 1 1 1 2 1 1 2 1 2 1 2 1 2 3 1 1 2 2 2 2 1 1 1 1 1 2 2 2
## [4387] 2 3 2 1 2 1 1 2 2 2 1 2 1 2 1 1 1 1 1 2 1 2 2 1 1 2 2 1 2 2 1 1 1 2
## [4421] 2 1 2 1 1 2 1 1 4 2 1 2 1 1 1 2 2 1 2 1 2 1 2 2 1 1 2 1 2 1 2 2 2 1
## [4455] 2 2 2 2 1 2 2 2 2 3 2 1 1 1 2 2 1 2 2 2 2 1 1 2 2 1 1 2 2 2 1 2 2 1
## [4489] 2 1 2 1 2 2 1 1 2 2 1 2 1 1 1 2 2 2 1 2 1 2 2 1 2 1 2 2 2 2 1 2 2 1
## [4523] 2 2 1 1 1 1 2 2 2 2 1 2 2 1 2 2 2 2 2 1 1 1 1 1 2 3 2 1 2 2 2 2 2 1
## [4557] 1 1 1 1 2 2 2 1 1 2 1 1 2 1 1 1 1 2 2 1 1 1 1 2 1 1 2 1 1 1 2 1 2 2
## [4591] 2 1 2 1 1 2 1 1 2 1 1 2 1 2 2 1 2 2 2 2 1 2 2 2 2 2 2 1 3 2 2 3 1 1
## [4625] 1 2 1 1 1 2 1 2 1 1 2 2 1 1 2 1 2 1 2 1 2 2 1 1 2 1 2 2 2 1 2 2 2 2
## [4659] 2 1 1 1 2 1 2 2 2 2 2 2 1 2 2 1 1 1 1 1 2 2 1 2 1 2 2 1 1 2 2 2 1 2
## [4693] 2 3 2 2 1 2 1 1 1 2 2 1 1 1 2 1 2 1 1 1 2 1 2 1 2 2 1 2 1 2 2 2 1 2
## [4727] 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1 1 1 2 1 2 2 2 2 2 2 1 1 1 2 1 2 1
## [4761] 2 2 2 2 2 1 2 1 1 2 1 2 2 1 2 2 1 1 1 2 2 1 2 2 1 1 2 1 2 1 2 2 1 1 2 1
## [4795] 2 2 2 2 2 1 1 2 1 1 2 2 2 1 1 1 2 2 1 2 1 1 2 1 2 2 1 1 2 2 1 2 2 2
## [4829] 2 1 1 1 2 2 2 2 2 2 1 2 2 2 1 2 2 2 1 1 1 2 1 2 1 2 2 2 2 1 2 2 2 1
## [4863] 1 1 1 2 2 2 2 3 2 1 2 1 1 1 1 2 2 2 2 2 2 2 2 2 1 2 3 2 2 1 1 2 2 1
## [4897] 1 1 1 1 2 1 2 2 2 1 1 1 1 2 2 1 1 1 2 2 1 2 2 2 2 1 2 1 2 1 2 1 2 1
## [4931] 2 2 2 1 2 1 2 2 1 1 2 1 2 1 2 1 1 1 2 2 1 2 2 2 1 1 2 2 2 2 1 2 1 1
## [4965] 2 2 1 2 2 2 1 2 2 1 1 2 2 2 2 2 1 2 1 2 2 2 2 1 1 1 3 2 2 1 2 1 2 2
## [4999] 2 2 1 1 1 2 2 2 1 1 2 2 2 1 1 1 1 1 2 2 2 1 2 1 1 2 2 1 2 1 2 2 1 2
## [5033] 1 1 1 2 2 1 1 1 1 1 2 1 2 2 1 2 2 2 1 2 1 2 3 1 2 2 2 1 1 1 2 1 1 1
## [5067] 1 1 2 2 2 1 1 2 2 2 2 1 1 2 2 2 2 2 2 1 2 3 1 1 2 3 1 3 2 3 1 1 1 2
## [5101] 1 1 2 2 2 1 1 2 2 1 1 2 1 2 2 2 2 1 2 1 2 2 2 2 2 2 1 1 2 2 1 1 2 2
## [5135] 2 2 2 2 2 2 1 1 2 2 2 1 1 1 1 1 1 1 1 2 2 2 2 1 1 2 1 1 2 2 2 1 3
## [5169] 2 2 2 1 1 1 2 1 1 2 1 2 2 2 2 2 2 2 1 2 2 2 1 1 2 1 1 2 1 1 2 1 1 1
## [5203] 1 2 3 1 1 2 1 2 2 1 1 1 1 2 1 1 2 2 2 1 1 1 2 1 1 2 2 2 2 1 1 2 1 3
## [5237] 2 2 1 1 1 2 2 1 1 1 2 2 1 2 2 1 2 2 1 2 2 1 2 2 1 2 1 2 1 2 1 1 1 1 2
## [5271] 2 2 2 2 2 2 2 1 1 1 1 2 2 2 1 1 2 1 1 2 1 2 2 2 2 1 2 2 2 2 2 2 2 2
## [5305] 1 1 1 1 1 2 2 2 2 3 2 2 2 2 1 2 1 1 1 2 2 2 2 2 2 1 2 2 2 2 2 1 2 1
## [5339] 2 1 1 2 2 1 1 1 1 1 2 1 1 1 2 1 1 2 2 1 2 2 2 1 2 1 2 3 1 2 1 1 1 2
## [5373] 1 1 2 1 2 1 2 1 1 3 2 1 1 2 2 2 1 1 2 1 2 1 2 2 1 1 1 2 1 1 2 1 2 2 2 2
## [5407] 1 1 1 2 1 2 1 1 1 1 1 2 1 1 2 2 2 1 3 2 2 2 2 1 2 1 2 1 2 1 1 1 2 1
## [5441] 2 1 1 2 2 2 1 1 2 1 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 1 1 1 2 1 2 1 2 2
## [5475] 1 1 2 2 2 1 1 2 1 1 2 2 1 1 2 1 2 2 2 2 2 1 2 1 2 1 1 2 2 1 2 1 1 1
## [5509] 2 2 2 2 1 2 2 1 2 2 2 1 2 1 2 1 2 2 1 2 1 2 2 1 2 1 1 1 2 1 2 2 2 2
## [5543] 2 2 2 1 2 3 2 2 2 2 2 2 2 2 2 1 1 2 2 1 2 1 2 1 2 2 1 1 1 2 1 2
## [5577] 2 1 2 2 2 1 3 2 1 1 2 2 1 2 1 2 1 2 2 2 1 1 1 1 2 1 1 2 1 2 2 2 2 2
## [5611] 1 1 2 2 1 1 1 2 1 1 2 2 2 1 2 2 1 2 1 2 1 2 2 1 2 2 1 1 2 2 1 2 2 1
## [5645] 2 2 2 2 2 1 1 3 1 2 2 1 1 1 1 2 1 1 2 1 1 2 2 1 2 2 2 2 2 1 2 2 1 2
## [5679] 2 2 2 2 1 1 2 2 1 2 1 2 2 1 1 2 2 2 2 1 1 2 2 1 2 2 2 2 1 2 1 1 2 1
## [5713] 3 2 1 2 2 1 2 2 1 2 2 1 2 2 1 1 1 2 1 4 1 1 2 2 1 2 1 1 1 1 2 2 1 1
## [5747] 2 2 3 1 1 2 2 1 2 3 1 1 1 1 2 1 1 1 1 1 2 1 2 2 1 1 2 2 2 2 2 2 1 2
## [5781] 1 2 1 2 2 2 1 2 1 2 2 2 3 1 2 1 2 2 1 1 2 2 2 1 2 2 2 1 1 2 2 1 1 1
## [5815] 2 2 2 2 1 2 2 1 1 1 1 1 2 1 1 2 1 2 1 2 1 1 1 1 2 2 2 1 2 1 2 2 1 1
## [5849] 2 2 1 2 3 2 1 2 2 1 2 1 2 1 1 2 2 2 2 1 2 1 3 2 1 2 2 2 2 2 2 2 2 1
```

```
## [5883] 1 2 2 2 2 1 2 2 1 2 1 2 2 1 1 2 4 1 1 1 2 2 1 1 1 1 1 2 2 1 1 1 2 2
## [5917] 2 2 2 2 2 1 1 1 1 1 1 2 1 1 2 2 2 1 2 2 1 1 1 2 2 1 1 1 2 2 2 1 2 2
## [5951] 1 2 1 1 1 2 2 1 2 2 2 2 2 1 2 1 2 1 1 2 1 2 2 1 1 2 1 1 2 1 2 1 2 2
## [5985] 1 1 1 1 2 3 1 1 2 1 1 1 2 2 1 2 2 2 2 2 1 1 2 2 1 2 2 1 2 1 1 1 1 2
## [6019] 2 1 2 1 1 2 1 1 2 1 2 2 2 2 2 1 2 2 1 2 2 2 1 1 1 1 1 3 2 2 2 3 1 2
## [6053] 1 2 2 4 2 2 2 1 1 2 1 2 1 1 1 1 2 1 2 2 2 1 2 2 2 2 4 1 2 1 2 1 2 2
## [6087] 1 1 1 2 2 2 1 1 1 1 1 2 1 2 2 1 2 1 2 2 1 2 1 1 2 1 3 1 2 2 1 1 3 2
## [6121] 2 1 1 1 4 1 2 1 2 2 2 2 2 2 2 2 1 1 1 2 2 2 2 2 2 1 2 2 1 1 2 1 2
## [6155] 1 2 3 2 1 1 1 2 2 2 1 1 2 2 1 1 1 1 2 3 1 2 1 1 1 1 1 2 2 2 1 2 2 2
## [6189] 1 2 2 1 2 2 2 1 2 1 2 1 1 2 2 2 2 2 1 1 1 2 1 1 1 2 2 1 2 2 2 1 1 2
## [6223] 1 1 2 1 1 2 2 2 2 1 1 1 2 2 2 1 2 2 2 3 1 2 3 1 2 2 2 2 1 2 1 2 2 2
## [6257] 2 2 1 2 2 2 1 1 1 2 1 1 2 2 2 2 1 1 1 1 1 1 1 2 1 2 2 1 1 1 3 1 1 1
## [6291] 2 1 2 1 2 1 2 2 2 3 1 1 1 3 2 2 2 1 2 2 1 1 1 1 2 1 2 1 1 1 2 2 1 1
## [6325] 1 1 2 1 1 1 2 2 1 2 2 1 1 1 2 2 2 1 2 1 1 1 2 2 2 1 1 1 1 1 1 2 2 2
## [6359] 2 2 2 1 1 2 1 3 2 2 1 2 2 1 2 1 2 2 2 2 1 1 2 2 2 1 1 2 1 2 2 2 1 2
## [6393] 2 2 1 1 2 1 2 1 1 2 2 1 2 1 2 2 2 1 2 2 2 2 1 1 2 2 2 2 2 1 2 1 2 2
## [6427] 2 1 2 2 1 2 1 2 2 1 1 1 2 1 1 2 1 1 2 2 1 1 1 2 2 1 2 2 2 1 2 2 1 1
## [6461] 2 2 2 2 1 2 1 2 2 1 2 2 2 1 1 1 2 2 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1
## [6495] 1 2 2 2 2 2 2 1 2 2 2 2 2 1 2 2 1 1 1 3 1 1 2 1 2 2 1 1 2 2 1 1 2 2
## [6529] 2 2 2 2 2 1 2 1 1 1 2 1 2 2 1 2 1 2 1 2 1 1 2 2 2 2 1 2 2 2 1 2 1 2
## [6563] 1 2 2 1 1 2 1 2 2 1 1 1 2 1 1 2 2 2 2 2 2 2 1 1 2 3 1 2 2 1 1 1 1
## [6597] 1 2 1 2 2 1 1 2 2 2 2 2 2 2 1 2 1 1 1 2 2 1 1 1 1 2 2 2 1 2 2 1 2
## [6631] 3 1 2 2 1 1 1 1 2 3 2 1 2 2 1 1 2 2 2 2 2 2 1 1 2 2 1 1 1 2 2 2 2 1
## [6665] 1 1 1 2 2 1 2 1 1 2 2 2 2 2 1 1 1 2 2 1 2 2 2 2 2 2 1 2 1 1 2 1 1 1
## [6699] 2 2 1 2 2 1 1 1 2 1 1 1 2 2 1 1 2 2 2 1 2 1 2 3 1 2 2 1 1 2 2 2 1 2
## [6733] 1 2 1 2 2 1 1 3 2 1 1 1 1 2 1 1 1 1 2 1 2 1 2 2 1 1 2 1 1 2 1 2 1 1
## [6767] 2 1 1 2 2 2 2 2 1 1 2 1 1 2 2 1 1 1 1 1 1 1 1 1 3 2 1 1 2 2 2 1 1
## [6801] 1 2 1 4 2 2 1 2 2 2 1 2 1 2 1 2 1 2 1 1 1 2 2 2 1 2 1 1 1 1 4 2 2 2
## [6835] 2 1 2 2 2 2 1 2 1 2 2 2 2 1 2 2 1 1 2 2 2 1 2 2 2 1 1 1 2 2 2 1 2 2
## [6869] 1 1 2 2 2 1 2 2 1 2 1 1 2 2 2 1 1 2 1 1 2 2 2 2 1 1 1 1 2 1 2 1 1 2
## [6903] 1 1 1 1 1 1 2 1 2 2 2 2 2 1 2 2 2 2 3 1 2 1 1 2 2 2 1 2 2 2 1 1 1 3
## [6937] 1 2 2 1 1 2 2 2 1 2 2 1 1 1 1 2 2 2 1 1 1 2 2 2 2 2 1 2 1 1 1 2 1 1
## [6971] 2 2 2 2 2 2 2 1 1 2 2 3 2 2 1 1 1 2 1 2 2 2 2 1 3 2 1 1 2 2 2 2 2 2
## [7005] 2 1 2 2 2 1 2 1 1 2 1 2 1 1 1 2 2 1 2 2 1 1 2 1 1 1 1 2 1 1 2 1 2 1
## [7039] 1 1 1 2 1 1 1 2 2 2 1 2 2 2 1 2 1 2 2 1 2 2 1 1 2 1 1 2 2 1 1 2 1 1
## [7073] 2 1 2 2 1 2 1 2 1 2 2 3 2 2 1 1 1 2 1 2 1 1 2 2 1 1 2 1 2 2 2 2 1 1
## [7107] 1 2 2 2 2 1 1 2 2 2 2 1 2 2 1 1 2 1 2 1 2 1 2 1 2 2 1 2 2 2 1 2 2 1
## [7141] 1 2 2 2 2 1 1 2 2 1 1 2 2 2 2 2 2 2 2 1 2 2 1 1 2 2 1 1 1 3 1 1 1
## [7175] 1 2 2 1 2 2 2 1 2 1 2 1 3 2 2 1 2 1 2 1 1 2 2 2 1 1 2 2 1 1 1 1 1 2
## [7209] 2 1 2 3 2 1 1 1 2 1 2 1 1 1 1 2 1 2 2 1 1 2 1 1 2 1 1 1 1 1 3 1 2 1
## [7243] 1 1 2 2 2 2 1 2 2 1 1 1 1 2 1 1 2 2 1 2 2 2 2 1 2 1 2 1 1 1 2 2 2 1
## [7277] 2 1 2 2 2 2 1 2 1 2 2 2 2 1 1 2 1 2 2 1 2 2 2 2 1 2 1 1 1 2 2 1 2 1
## [7311] 2 2 2 2 1 1 2 2 1 2 2 1 2 2 2 1 2 1 1 1 2 2 1 1 1 1 1 1 2 2 1 2 1
## [7345] 1 1 2 2 1 1 2 2 1 2 2 1 1 2 1 1 1 2 1 2 2 2 2 1 2 1 1 1 1 2 2 2 2 2
## [7379] 1 2 1 1 2 3 1 2 1 2 1 2 2 1 2 2 1 2 1 1 1 2 2 2 2 2 2 1 2 2 1 2 1 1
## [7413] 1 2 2 1 1 1 1 2 2 2 1 2 1 2 2 2 2 2 2 1 1 2 2 1 2 1 4 1 1 1 2 2 1 1
## [7447] 1 2 1 1 2 2 1 1 2 1 1 1 2 2 2 2 1 1 1 1 2 2 2 2 1 2 2 1 2 2 2 2 2 2
## [7481] 1 2 2 1 2 1 1 2 3 2 2 2 1 2 1 2 2 1 1 2 1 1 2 1 2 1 1 1 2 2 2 1 1 2
## [7515] 1 2 1 1 2 1 2 2 1 2 2 1 2 1 1 2 1 2 1 1 2 1 2 1 2 2 1 2 2 2 3 2 1 2
## [7549] 1 2 1 2 2 2 1 3 2 2 1 2 2 2 2 2 2 1 2 2 2 1 1 1 2 2 2 2 2 2 2 1 2 2
## [7583] 2 2 1 1 2 1 2 2 2 1 2 1 1 2 2 2 1 1 2 2 1 1 1 1 1 2 1 1 1 1 2 1 2 1
## [7617] 1 1 1 2 2 2 2 1 2 2 1 1 1 2 1 2 2 1 1 2 2 2 1 2 1 2 2 1 1 2 2 2 1 1
## [7651] 2 2 1 1 2 2 2 3 1 1 2 1 2 1 2 2 2 2 1 1 1 1 1 1 2 2 1 1 1 1 1 1 1 2
## [7685] 2 2 1 1 2 2 2 1 2 1 1 1 1 2 2 1 2 3 2 2 2 2 2 2 2 2 1 2 1 2 2 1 2 2
```

28

```
## [7719] 1 2 1 1 1 1 2 2 1 2 2 1 2 2 1 2 1 1 1 1 2 2 1 1 2 2 1 3 1 1 2 1 2 2
## [7753] 1 1 2 1 2 2 2 2 1 1 1 1 2 2 1 2 1 1 2 2 2 2 1 2 2 1 2 1 1 1 2 1 2 2
## [7787] 1 1 1 2 2 1 1 2 1 2 2 1 2 2 1 1 3 2 3 2 1 1 1 1 1 1 1 4 2 1 2 2 2 2
## [7821] 1 2 2 2 1 2 1 2 2 1 1 1 2 2 2 2 1 2 2 2 2 1 1 2 1 1 1 1 2 2 1 2 2 1
## [7855] 2 2 1 2 3 1 1 2 2 1 1 1 2 2 1 1 2 2 2 1 2 2 1 2 2 1 1 1 2 2 2 2 3 2
## [7889] 1 2 1 2 2 2 1 1 2 2 1 1 2 2 2 1 2 1 2 1 2 2 2 2 3 1 2 1 1 1 2 2 2 2
## [7923] 2 1 2 2 1 1 1 2 2 2 2 2 1 2 1 1 2 2 2 2 1 2 1 1 1 2 2 2 1 2 1 2 2 1
## [7957] 1 2 1 1 1 1 3 2 2 1 2 1 2 2 2 2 1 1 2 2 1 1 2 2 1 1 2 2 2 2 1 1 2 2
## [7991] 1 2 1 1 2 1 2 1 2 2 1 1 1 1 2 2 2 1 1 1 2 1 1 1 2 1 2 1 2 2 2 2 2 2
## [8025] 2 3 2 2 1 1 2 2 1 3 2 1 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 1 1 2 1 2
## [8059] 2 2 1 2 1 2 1 1 1 2 2 2 2 1 1 1 1 1 2 1 2 2 2 2 1 1 1 1 1 2 1 2 2 2
## [8093] 2 2 2 1 1 2 1 1 2 2 2 2 2 1 2 2 2 2 2 1 1 2 2 2 1 1 2 2 2 3 2 2 1 1
## [8127] 2 2 1 1 2 1 1 1 1 2 1 2 1 2 1 2 1 1 2 2 1 1 1 2 1 2 2 2 2 1 1 2 2 2
## [8161] 1 1 2 2 2 2 1 1 2 1 2 1 2 2 1 2 2 2 1 2 2 2 2 2 2 2 1 2 2 1 1 2 1 2
## [8195] 1 1 2 2 2 2 2 2 1 2 1 1 1 2 2 2 4 2 2 1 1 1 2 2 2 2 1 2 2 2 1 1 2 1
## [8229] 1 1 1 2 2 2 2 1 1 1 1 1 1 2 1 1 2 1 1 1 1 2 2 2 2 1 1 1 1 1 2 1 2 1
## [8263] 2 3 2 2 1 1 1 1 2 2 2 1 1 1 2 2 3 1 2 2 2 1 2 2 1 2 1 1 1 1 2 1 2 2
## [8297] 1 2 2 2 1 1 2 2 2 2 2 3 2 2 1 2 1 1 1 2 2 2 1 1 1 2 1 1 1 1 2 1 1 1
## [8331] 2 1 1 1 2 1 2 1 2 2 2 2 1 1 1 2 2 2 2 1 2 2 2 1 1 1 2 2 1 1 2 2 1 2
## [8365] 1 1 3 3 1 1 1 2 2 1 2 2 1 1 1 2 2 1 1 2 1 2 2 1 1 2 1 1 2 2 2 2 2 2
## [8399] 2 1 2 1 2 2 2 2 2 1 1 2 1 2 2 1 2 2 2 2 1 1 2 1 2 2 1 1 1 1 2 1 1 1
## [8433] 1 2 2 2 1 2 2 1 2 1 2 2 2 2 1 1 2 2 1 2 2 1 2 1 1 2 2 2 2 2 1 1 1 2
## [8467] 1 2 1 1 1 2 2 1 3 1 1 2 1 2 2 2 1 2 2 2 1 2 1 1 1 2 2 2 2 2 2 1 2 1
## [8501] 1 2 1 2 2 2 2 2 2 2 1 1 1 2 2 1 1 1 1 1 2 1 1 2 2 2 1 2 2 2 1 2 2
## [8535] 1 2 1 1 2 1 1 1 2 2 1 2 3 1 2 1 2 2 2 1 2 2 1 2 2 1 2 2 1 2 2 2 1 1
## [8569] 3 2 2 2 2 1 1 1 1 1 2 2 2 2 2 2 2 3 1 2 1 1 1 2 1 1 2 1 2 1 1 2 1 1
## [8603] 3 2 1 1 1 1 1 2 2 1 2 2 2 1 1 1 2 1 1 2 1 1 1 2 2 2 2 2 2 1 1 2 3 2
## [8637] 1 1 2 1 2 2 1 4 1 2 2 1 1 2 2 1 1 1 2 1 2 2 2 1 1 2 1 2 1 2 4 1 2 1 2 2
## [8671] 2 3 1 1 2 2 2 1 2 1 2 2 1 2 1 1 2 2 1 1 1 1 1 2 2 1 2 1 2 1 2 2 1 1 2 1
## [8705] 1 2 3 2 2 1 2 1 1 2 1 2 2 3 1 2 3 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 2 1
## [8739] 1 1 1 2 3 2 2 2 1 1 2 1 1 1 1 3 2 1 1 2 1 2 1 2 1 1 2 1 2 2 2 2 2 1
## [8773] 2 2 1 2 1 1 1 1 1 1 2 2 1 1 1 2 2 1 1 1 2 2 1 2 2 2 2 2 2 2 1 2 1 2
## [8807] 1 2 2 1 1 2 1 2 2 1 1 2 2 2 2 1 1 1 1 1 2 2 2 1 2 2 2 2 1 1 1 1 1 1
## [8841] 1 1 2 2 2 2 2 2 1 2 2 2 2 1 2 2 1 1 2 2 1 2 2 2 1 2 2 2 1 2 1 2 1 1
## [8875] 2 1 2 2 1 1 1 1 2 1 2 2 1 1 2 1 2 1 1 1 2 1 2 1 1 1 2 2 1 2 3 1 2 2
## [8909] 1 1 2 2 2 2 2 4 2 1 2 2 2 2 2 2 2 1 1 2 1 1 1 2 2 1 3 2 2 1 1 1 1 2
## [8943] 1 2 1 1 1 2 2 1 2 2 2 2 2 1 2 2 2 2 2 2 2 1 2 2 1 1 1 2 2 2 2 2 2 1
## [8977] 3 2 2 2 1 1 2 2 2 1 2 2 2 2 2 2 1 2 2 1 2 1 1 2 1 2 1 1 1 2 2 2 1 1
## [9011] 1 2 2 2 2 1 2 1 1 2 2 1 2 1 2 1 2 2 2 1 2 2 2 2 1 1 1 1 2 1 3 1 2 2
## [9045] 1 1 1 1 1 2 1 2 2 2 3 1 2 1 2 2 2 2 1 1 1 1 1 2 2 1 2 1 2 1 1 2 2
## [9079] 2 2 1 2 2 1 1 1 1 2 2 2 1 2 1 2 1 3 2 2 1 3 2 2 2 2 2 1 2 2 2 1 2 2
## [9113] 1 2 1 2 2 2 1 3 2 1 2 1 2 2 1 2 2 1 1 1 1 1 2 2 1 1 2 2 2 1 1 1 1 2
## [9147] 1 2 1 1 2 2 2 1 1 2 1 2 2 2 1 1 1 2 1 1 2 2 1 1 1 2 1 2 2 1 2 2 2 3
## [9181] 2 1 1 2 2 2 2 2 1 2 2 2 1 2 2 1 2 1 2 2 1 1 2 1 1 1 3 1 1 1 1 1 1 2
## [9215] 2 1 2 1 1 2 1 1 1 1 1 1 1 2 1 2 2 1 1 1 2 2 1 1 2 1 1 1 2 2 2 1 2 2
## [9249] 1 2 1 2 2 1 2 1 2 1 2 2 2 2 2 2 2 2 2 1 1 2 1 1 2 1 2 2 2 1 1 1 1 2
## [9283] 2 1 2 1 1 2 1 2 2 1 2 1 1 2 1 1 2 2 1 2 2 1 1 2 2 2 1 2 1 2 1 1 2 2
## [9317] 1 1 2 2 1 2 1 1 2 2 2 1 2 2 2 2 2 2 1 3 2 2 1 4 2 2 1 1 2 1 1 1 2 2
## [9351] 2 2 1 2 1 1 1 2 2 1 2 2 2 1 2 2 2 1 2 1 2 1 1 1 2 1 1 1 1 1 2 2 2 1
## [9385] 2 2 1 2 1 1 1 2 2 2 1 1 2 1 2 1 1 2 2 2 2 2 2 2 1 2 3 1 1 2 1 1 1 2
## [9419] 2 2 2 1 1 2 1 1 1 1 2 2 2 2 2 2 1 1 2 2 1 3 2 2 2 1 2 2 2 2 1 1 2 1
## [9453] 2 1 2 2 2 1 2 1 1 2 2 2 2 1 2 1 2 1 2 1 2 2 1 1 2 2 1 1 1 2 2 1 1 2
## [9487] 1 1 2 2 2 2 2 2 2 2 1 2 2 2 1 1 2 1 1 2 3 1 1 2 1 1 1 2 2 2 2 2 1 1
## [9521] 1 1 2 1 1 1 2 2 1 2 2 1 1 2 1 1 1 1 1 2 1 1 2 2 2 1 1 2 2 1 1 1 1 1
```

29

```
##  [9555] 2 1 3 2 2 1 2 2 2 2 1 2 2 2 2 2 1 1 2 2 2 1 1 1 1 2 1 2 2 2 2 2 1 1
##  [9589] 2 1 1 2 1 2 1 1 1 1 1 2 2 1 2 1 1 1 1 2 2 2 2 2 2 1 2 3 2 1 2 2 1 1
##  [9623] 1 2 1 1 2 1 2 2 1 3 2 1 2 2 2 2 2 2 2 1 1 3 1 2 1 2 2 2 1 1 1 1 2 1
##  [9657] 2 2 2 2 2 1 2 1 3 2 2 1 2 2 2 1 1 1 1 2 1 2 1 2 2 1 1 1 2 2 2 1 2 2
##  [9691] 2 1 2 2 2 2 1 3 2 2 1 1 2 1 1 2 2 2 1 1 1 2 1 2 1 2 1 2 1 2 1 4 1 2
##  [9725] 2 2 1 2 1 3 1 1 2 2 2 1 1 2 1 1 1 2 2 2 2 1 1 2 2 2 1 2 1 2 1 2 1 2
##  [9759] 2 2 2 2 2 2 1 1 2 1 1 2 1 1 1 2 1 1 1 2 2 1 2 2 2 1 1 1 2 2 1 2 1 2
##  [9793] 2 2 2 1 2 1 1 1 2 2 1 1 2 1 1 1 2 2 2 1 2 2 1 2 2 2 1 2 2 1 1 2 1 2
##  [9827] 2 1 1 1 1 1 1 2 1 2 2 1 1 2 1 2 2 1 1 1 1 1 2 2 2 2 2 1 2 2 1 1 1 1
##  [9861] 1 2 1 2 2 1 2 2 2 2 1 1 1 1 1 2 2 1 1 1 2 1 2 2 2 1 1 2 1 2 2 1 1 2
##  [9895] 1 1 1 2 1 1 1 1 1 1 2 2 2 2 1 1 2 1 1 1 2 1 2 1 2 2 2 2 1 2 2 2 2 1
##  [9929] 2 2 2 2 2 2 2 1 2 1 1 1 2 1 1 1 2 2 2 1 2 1 1 2 2 1 3 2 3 2 1 1 2 2
##  [9963] 2 2 2 2 2 2 2 2 2 1 1 2 1 1 2 2 1 2 3 1 2 1 1 1 1 1 2 1 2 1 1 2 3 1
##  [9997] 2 2 1 1 2 2 2 2 2 2 1 2 1 2 2 2 1 2 2 2 1 2 1 1 2 1 2 1 1 1 1 1 1 1
## [10031] 2 1 2 2 2 2 1 1 1 2 2 1 2 1 1 2 2 2 1 2 2 1 1 2 2 1 1 1 2 1 2 2 2 1
## [10065] 2 1 1 2 2 1 1 2 2 2 2 1 1 1 2 2 1 2 1 1 1 2 2 1 1 1 1 2 2 2 1 3 2 1
## [10099] 1 2 1 1 1 2 1 1 2 3 2 2 2 1 2 2 1 2 2 2 2 2 1 1 2 1 2
## Levels: 1 2 3 4
```

```r
# misclassification rate:
yp2 <- predicted.price3
yt2 <- olr.test$class
mcr2 <- table(yp2,yt2)
1-sum(diag(mcr2))/sum(mcr2)
```

```
## [1] 0.2777284
```

Discussion

The variables provided to us with the data set from Yelp included: the name of the business, city, state, latitude, longitude, stars, review count, whether they do take out or not, the price range, whether or not they offer outdoor seating, whether or not they have a full bar inside, the category that the business falls into, and the attire of the restaurant. From this data set, we decided to try and create a model that predicted the price range of the restaurant; i.e, the number of dollar signs displayed on Yelp. We attempted to model this potential relationship with a linear regression (both with and without rounding due to only integer number of dollar signs), ordinal regression, and classification trees. From this data, we decided to use the numerical variables: latitude, longitude, number of stars, and review count; as well as the qualitative variables: alcohol, attire, outdoor seating, and takeout. We found that our qualitative variables (such as stars, alcohol, attire) were fairly strong predictors of price range.

Using a linear model with rounding we were able to obtain a misclassification rate of 27.62%. Using the ordinal regression model, we were able to obtain a misclassification rate just slightly better than the linear model: a misclassification rate of 27.5%. Both these models are somewhat accurate as just guessing 2$ for each restaurant would yield roughly a 50% accuracy rate, and both these models would be 22.38% and 22.5% more accurate, respectively. However, both of these models could be improved upon with the addition of more variables, perhaps from the entirety of Yelp's dataset.

Our classification tree using only numerical variables resulted in all branches ending in a 1 or 2, not surprising since the bulk of the restaurants in the data are rated as 1 or 2 dollar signs. However, utilizing the qualitative predictors show that they are more important as the result in a single branch ending in a 3.

Future research into this model may consider utilizing more variables from the Yelp data set that were omitted from this analysis due to missingness or lack of theoretical support, and might also incorporate the actual text reviews into the prediction. Our model controls for variations in price ranges across the 10 geographical areas, so we do not believe that the addition of new metropolitan areas would make the models more robust,

since we assume that the price distribution that exists with these 10 would extend to any additional metro areas. Considering some unsupervised techniques, like PCA, may also be helpful in generating the model. It may also help to explore whether the type of restaurant (whether it be Chinese food or a steakhouse) affects the price range as it does seem like a potential indicator. The drawback of using this data set was that it characterized each restaurant as different types and not a single characteristic. It may also be worthwhile to see whether or not the variable we included effect the number of stars that a restaurant has on Yelp.

References

"Yelp Open Dataset." Yelp Dataset, 2019, www.yelp.com/dataset.