restaurants

Chris Parrish

December 8, 2016

Italian restaurants in Manhattan

reference:
- Sheather, A Modern Approach to Regression with R, pp.5-7

Load packages.

```
library(ggplot2)
library(dplyr)
```

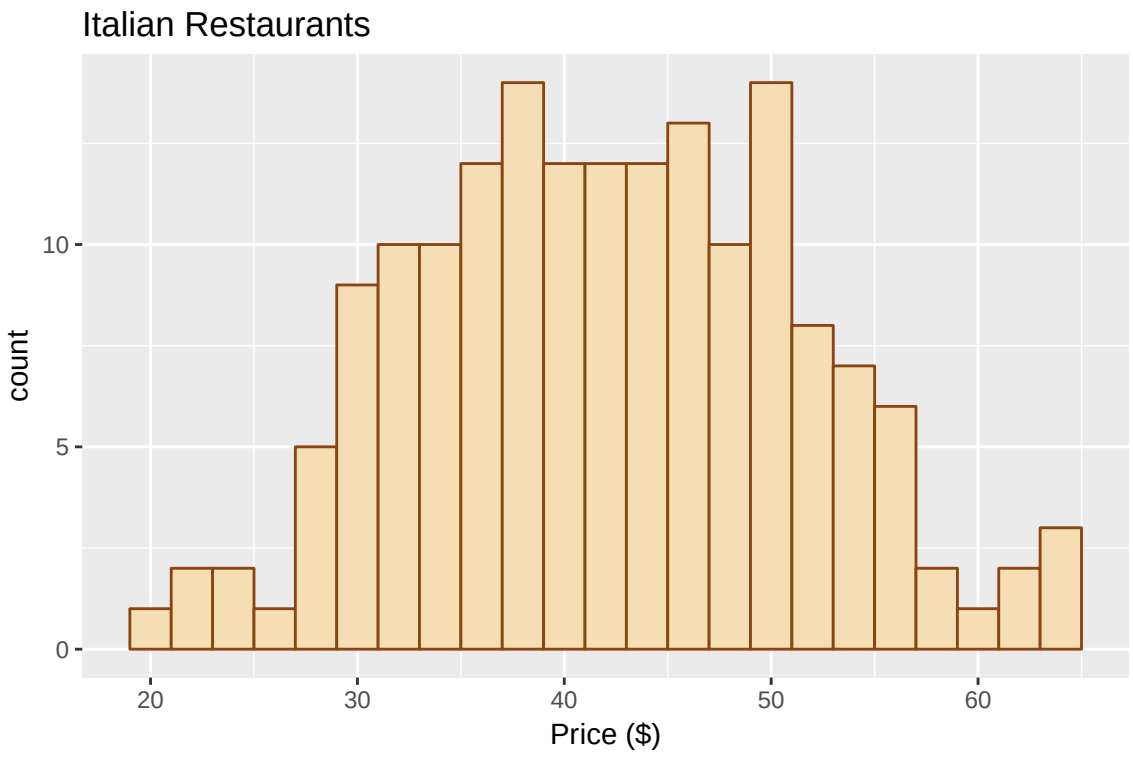
Import the data.

```
data <- read.csv("nyc.csv", header=TRUE)
head(data, 4)
```

```
##   Case          Restaurant Price Food Decor Service East
## 1    1 Daniella Ristorante    43   22    18      20    0
## 2    2  Tello's Ristorante    32   20    19      19    0
## 3    3          Biricchino    34   21    13      18    0
## 4    4             Bottino    41   20    20      17    0
```

Visualize the data.

```
ggplot(data, aes(Price)) +
  geom_histogram(binwidth = 2, color = "saddlebrown", fill = "wheat") +
  labs(x = "Price ($)", title = "Italian Restaurants")
```



Distinguish between Italian restaurants to the West and East of Fifth Avenue.

```
data$Location <- factor(data$East, labels = c("West", "East"))
ggplot(data, aes(Location, Price, color = Location)) +
  geom_boxplot() +
  scale_colour_manual(values = c("steelblue", "turquoise")) +
  labs(x = "Location", title = "Italian Restaurants")
```

Numerical summaries of prices.

reference:
- data wrangling with `dplyr` and `tidyr`

```
data %>%
  group_by(Location) %>%
  summarise(mean = mean(Price),
            sd = sd(Price),
            n = n())
```

```
## # A tibble: 2 × 4
##   Location     mean       sd     n
##     <fctr>    <dbl>    <dbl> <int>
## 1     West 40.43548 9.647256    62
## 2     East 44.01887 8.859734   106
```