R Notebook

Nelson Brown and Bryce Smith

Introduction

Summary Statistics and Graphics

Quantitative Values

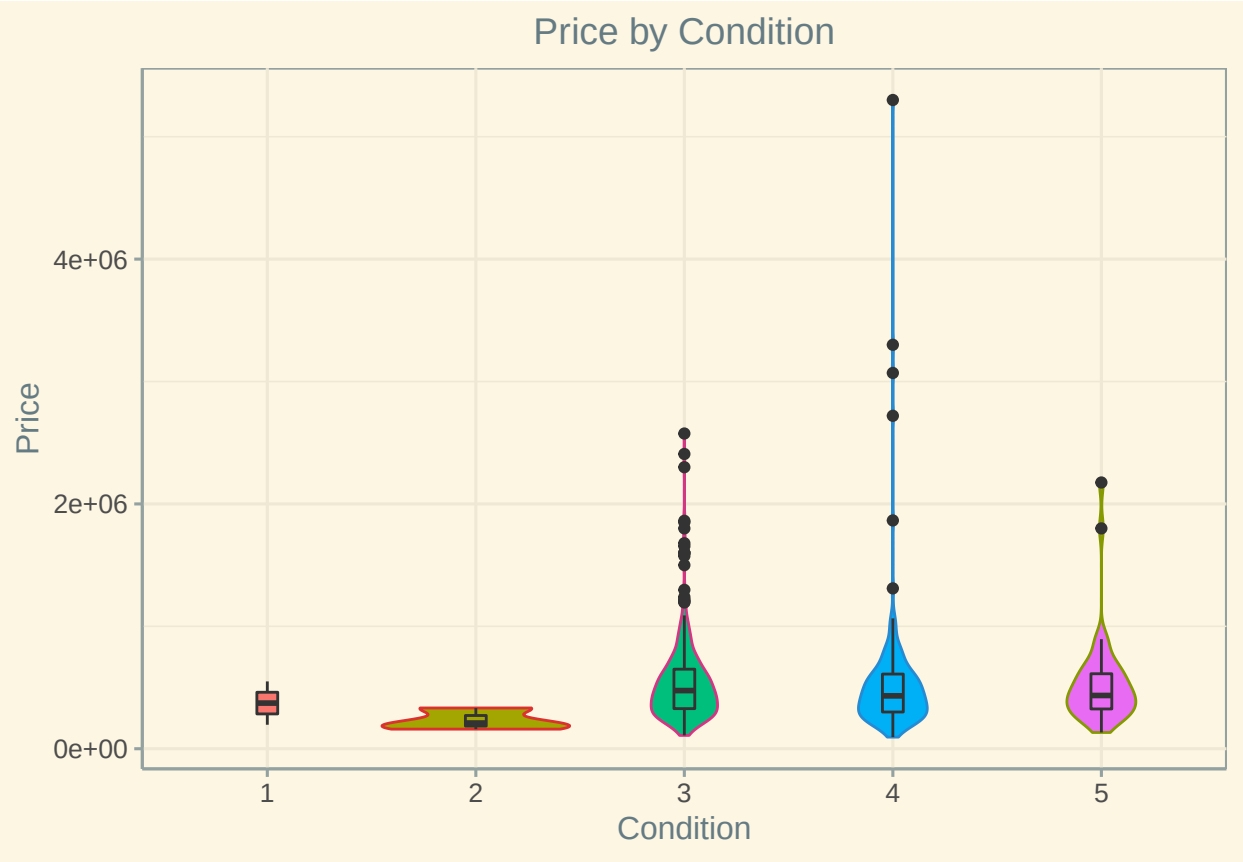
Table 1: First Four Rows for Quantitative Values on Seattlle Housing Dataframe

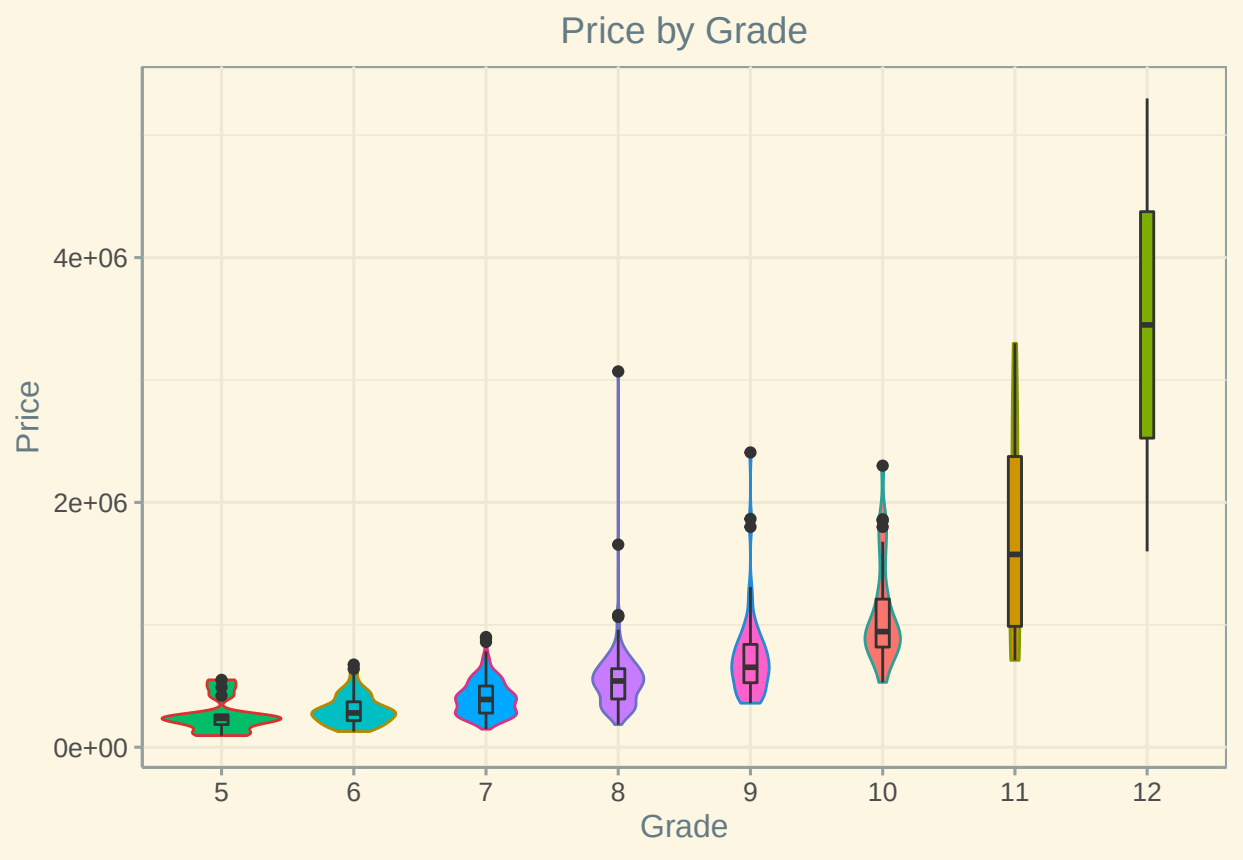
price	sqft_living	sqft_lot	sqft_above	sqft_basement
359,950	1,570	6,975	1,040	530
909,950	3,050	8,972	3,050	0
318,000	1,570	12,506	1,570	0
272,000	1,390	10,660	1,030	360

Table 2: Summary Statistics for Quantitative Values on Seattlle Housing Dataframe

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Pctl(75)	Max
price	613	545,427.700	408,545.900	95,000	315,000	631,500	5,300,000
sqft_living	613	2,073.669	963.763	380	1,370	2,550	7,390
sqft_lot	613	15,967.970	46,698.890	740	5,100	10,585	871,200
sqft_above	613	1,793.571	873.153	380	1,130	2,313	6,530
sqft_basement	613	280.098	424.835	0	0	570	2,390

Discrete and Categorical Values





Analysis

Results and Conclusions

Appendix: All Code for This Report

```
library(ggplot2)
library(ggthemes)
library(lubridate)
library(stargazer) # Used for latex tables to summarize the data and models
# Read the Data
df <- read.csv('Seattle.csv', strip.white = TRUE, stringsAsFactors = FALSE)
quant.columns <- c(3,6,7,13,14)
df$condition <- as.factor(df$condition)
df$grade <- as.factor(df$grade)

# Clean the Data
df$date <- ymd(substr(df$date,1,nchar(df$date) - 7)) # Convert string to date object
# Print head of initial dataframe
stargazer(df[1:4,quant.columns], rownames=FALSE, summary=FALSE, header=FALSE, title="First Four Rows fo
# Summarize initial dataframe
stargazer(df[,quant.columns],header=FALSE, title="Summary Statistics for Quantitative Values on Seattll
ggplot(df, aes(x=condition, y=price, fill=condition)) +
  theme_solarized(light = TRUE) +
  scale_colour_solarized("red") +
  geom_violin(aes(color=condition)) +
  geom_boxplot(width=0.1) +
  ggtitle("Price by Condition") +
  xlab("Condition") +
  ylab("Price") +
  theme(legend.position="none", plot.title = element_text(hjust = 0.5))

ggplot(df, aes(x=grade, y=price, fill=grade)) +
  theme_solarized(light = TRUE) +
  scale_colour_solarized("red") +
  geom_violin(aes(color=grade)) +
  geom_boxplot(width=0.1) +
  ggtitle("Price by Grade") +
  xlab("Grade") +
  ylab("Price") +
  theme(legend.position="none", plot.title = element_text(hjust = 0.5))
```