Statistics 141A

Troy lui

Due: October 30, 2018

1. Data Overview

The data set given is a data set provided by Craigslist. The data set represents apartment rentals in California on Craigslist taken on October 15th, 2018 specifically. To preface, it is good to note that the features pets, laundry, parking, place, city, county, and state were taken from the variable text and then added to the data set.

There are 21,948 rows within the data set. Each row represents a different post on Craigslist, which means that there are 21,948 different posts in this data set. Characteristics of the data to note is that it has twenty variables, or columns, that give specific information about each row's apartment. The twenty columns have variable names of title, text, latitude, longitude, city text, data posted, date updated, price, deleted, sqft, bedrooms, bathrooms, pets, laundry, parking, craiglist, place, city, state, and county. The more non intuitive column variables such as text, city text, and deleted represent apartment description, address, and whether or not the craigslist post was deleted respectively.

About the column variables, certain units that should be clarified include price, bathrooms, and bedrooms. For the variable price, it is recorded as a monthly price. As for bathrooms, the variable can take a value such as 1.5 since it represents 1 bathroom plus an extra sink. Lastly, bedrooms can take a variable of 0 if a studio apartment.

When taking a look at the date of the advertisements posted, all of the data set encompases the year 2018 except for one post that doesn't record any data except for the fact that it was deleted. This observation can be taken out of the data set since it doesn't provide useable information.

A reoccurring pattern to note about the data is that the variable text often contains information that could go under city text. For example, row 8 shows that the Terrace Apartment Homes is located in the city Santa Clarita in the text variable, but the city text variable shows missing data, which could be due to varying user inputs. In addition to this, another pattern is that some advertisements have the text variable with "Call for details", which usually means that other variables will be omitted. For example, in row 16, the apartment's text is shown to call for more details, so it doesn't provide details about sqft.

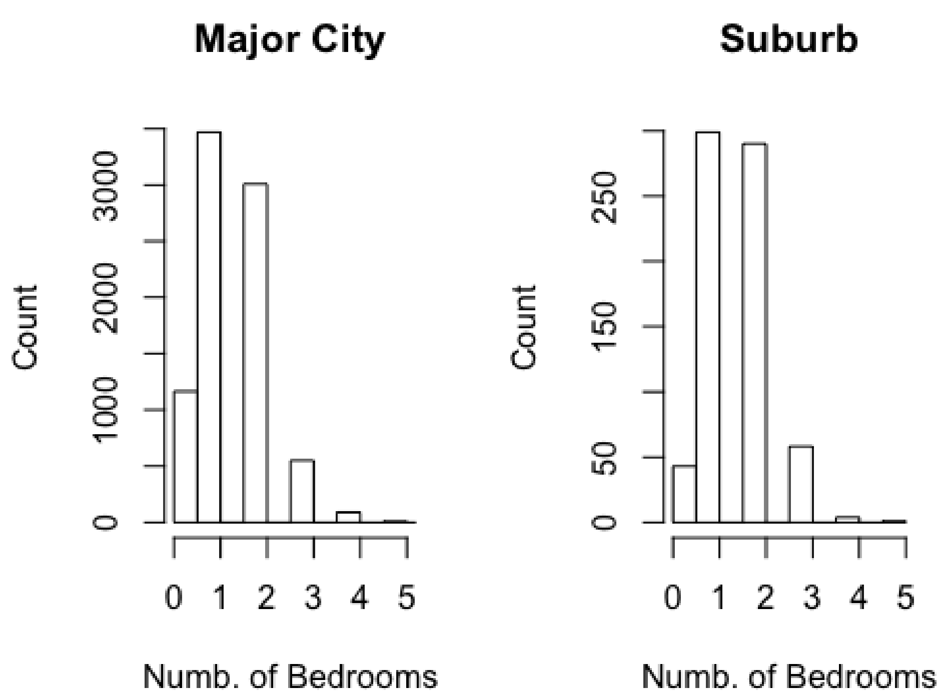
Viewing the data, there are many apartment advertisements that look as though they are duplicate advertisements; however, rows 8 to 20 show the same latitude and longitude but different data for bedrooms, baths, price, and etc. Thus, I believe that these duplicates represent an apartment complex and each advertisement represents a different apartment in that complex that is available.

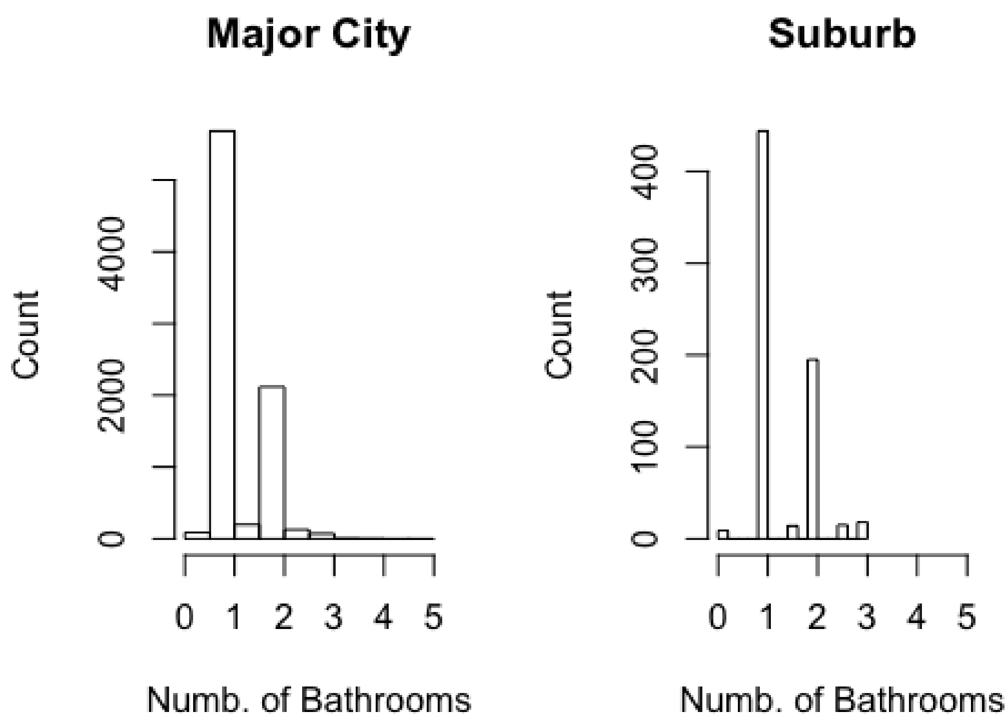
Aside from the observation that can be taken out of the data set (row 17902), all the advertisements status are active, or not deleted. When thinking back on row 17902, it could have no information available since the advertisement was deleted.

Now we'll take a look at the parameters of price, bedrooms, bathrooms, and square feet. The histgrams for the four variables show potential outliers, so median will be used. The medians of the variables are $2288 per month, 1 bathroom, 1 bedroom, and 835 square feet. When taking a look at the maximum and minimum values of these variables; however, there are values that may have been recorded wrong or an error due to user input. The three values that stand out is a 200,000 square feet apartment, 9951095 per month rent, and an even higher 34083742 per month rent. When looking at the text variable of the advertisements for the 9951095 and 34083742 per month rents, I found that the user inputted a range of 995-1095 and 3408-3742 per month rent. Therefore, to lessen the magnitude of the error, I set the data set 9951095 = 995 and 34083742 = 3408. As for the 200,000 square foot value, no information was found in the text variable; therefore, the value can't be fixed. To help this issue, I set the value to NA, or a missing value.

2.

2.1: Family-Friendly ~ Suburban vs. Major City Apartments





Major City Apartments: Partial Pets Proportions

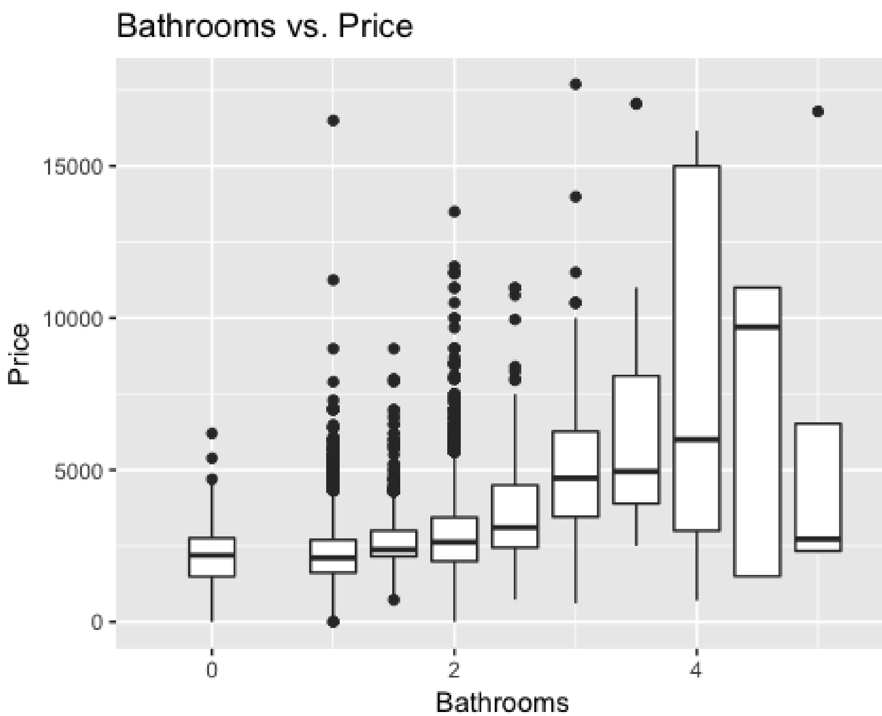
Both	Cats	Dogs	Total
0.556843203	0.095260169	0.012631823	.6647

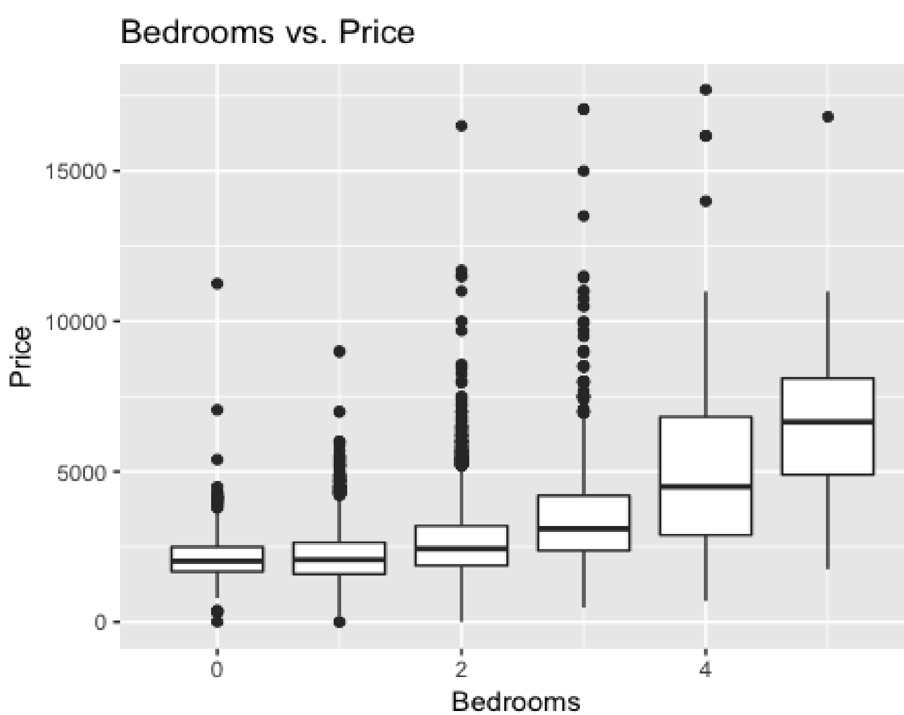
Suburban Apartments: Partial Pets Proportions

Both	Cats	Dogs	Total
0.581329562	0.053748232	0.004243281	.6393

When comparing suburbs and major city apartments, I first used San Francisco, Oakland, San Jose, Sacramento, Los Angeles, and San Diego to represent the major city apartments, and I used West Sacramento, Alhambra, Palo Alto, Folsom, Livermore, and Sunnyvale to represent the suburban apartments. To see if suburban apartments are more likely to be family-friendly, I compared the major city and suburban aparments through the variables of, bedrooms, bathrooms, and whether or not pets were allowed. The only variable that showed a slight edge for family-friendliness of suburban aparments over city apartments was number of bedrooms. The median number of bedrooms for suburban aparments from this data set equal to 2 while the median for major city apartments equaled to 1. Other than that, bedroom medians were the same and pet allowance actually favored major city apartments with a percentage of 66.47% guranteed allowance versus the suburban apartments 63.93% guranteed allowance. All in all, the only variable suburbs were more likely to be family-friendly in were the number of bedrooms.

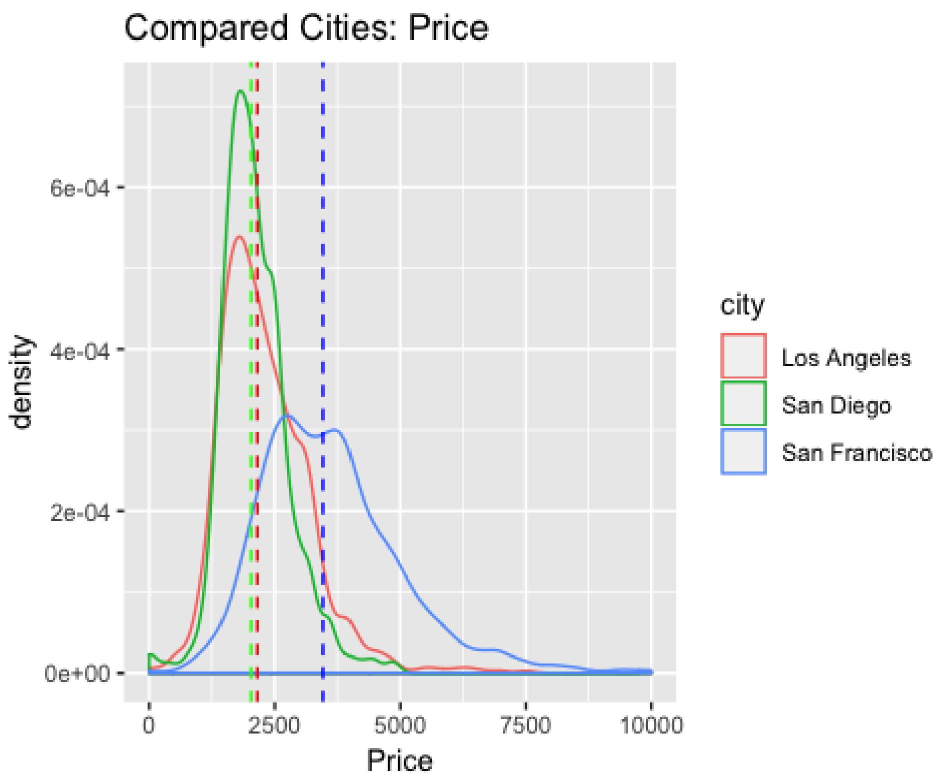
2.2: Which adds more rent ~ Bedrooms vs. Bathrooms



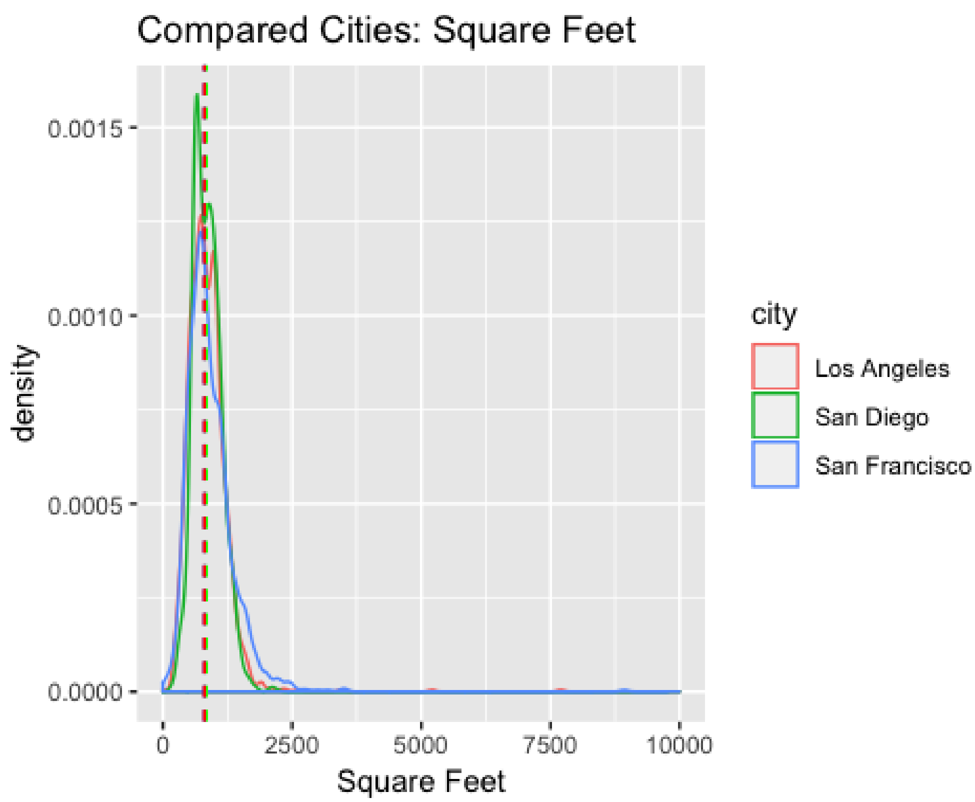


Using a visual diagnostic, bathrooms seem to raise price more than bedrooms. When comparing 1 bathroom to 1 bedroom throughout the boxplot, both 1 bathroom and 1 bedroom start off at approximately at the same median price of 2,500. Both bathroom and bedroom seem to have the same values when comparing 0:0, 1:1, and 2:2; however, at 3:3 and 4:4, bathrooms are most clearly ahead of bedrooms. With the comparison of 5:5, median bathroom price actually takes a large dip downward and median bedroom price rises significantly.

2.3: Apartments in Similar Geographical Areas



For similar geographical areas, I will relate San Francisco, San Diego, and Los Angeles since they are all similar major cities with dense populations. When comparing median prices, LA and SD have similar median prices of 2150 and 2025 respectively; however, SF is significantly different with a median price of 3462. Interestingly, I would have thought that LA would have the highest median rent price since, as of 2017, its population was approximately 4 million compared to SF's 884 thousand.



When taking a look at median square feet for SF, SD, and LA, they are shockingly similar. With medians of 805, 840.5, and 800, their median square footage doesn't vary too much at all. In retrospect, this isn't as initially shocking since big cities have limitted space for apartments.

SF, SD, and LA bathrooms tend to also have the same number of bathrooms, which makes sense since their square footing is roughly the same. When taking a look at the median values of bathrooms they all equal to 1. When accounting for bedrooms as well, the median values of the three cities all equal 1 as well.

3. Questions

One question that can be answered from this data set is given specific financial limitations, where would be an apartment available to live within a given city. This question is meaningful because it can allow people to filter out apartments that they can't live in due to location and financial situation. (1)

Another question that can be answered is whether big city apartments have a higher proportion of available parking in their complex compared to suburban areas. This question can be significant for people who can't stand having to park on the streets and could help them lean towards finding a suburban or big city apartment. (2)

Does the allowance of a dog, cat, or both tend to increase the price of rent? This question is important since it could allow college graduates moving out of their parent's house valuable infomration towards whether they should bring their pet along. It also could help people decide whether or not they should get a pet. (3)

How much does having a laundry machine inside the unit tend to increase the price of rent? This question is important since it would give people who want to save money and like to have their laundry machine inside their apartment valuable information towards whether the trade off of a laundry machine inside the apartment is worth the amount more they'll have to spend per month. (4)

For this data set, what month has the most amount of advertisements just posted? This question is important because if someone is casually looking for an apartment replacement, they could look during certain months to decide upon the most options. (5)

Which city has the best bang-for-your-buck? In other words, which city has the least price per square foot? This is important because someone could want to optimize the amount of space they get for how much they are paying monthly. (6)

Which 2-bedroom apartment is the cheapest in a given city? This information would be significant for someone who is looking to share an apartment with a significant other but wants to save as much money as possible. (7)

Out of all the advertisements, which are the five most recent posts? This information can help someone who is looking at apartments and wants to make sure that the website isn't filled with a lot of old posts that hadn't been deleted. (8)

Which advertisements are posted in the most recent month (including those that have been updated in the most recent month)? This question is important because a person may be tired of calling apartments that have already been taken. Therefore, getting the data from the most recent month helps ensure that they are still open. (9)

On average, how much does price tend to increase each .5 increase or 1 increase in bathrooms. This question is important because it can give any general person an idea of how much extra a .5 or 1 bathroom will cost per month. (10)

4. Question Answers

Question 1: Given a financial limitation of 850 per month, where would be the most optimal (sqft to price) apartment in San Jose?

<u>Hypothesis:</u> Taking a quick look at the data, I would hypothesize the best for this category would be Centerra's apartment complex since it's the lowest price with fairly high square feet.

The best apartment is Centerra's apartment complex that has a ratio of 4.3 square feet per dollar per month. Interestingly, my hypothesis was correct; however, there is potential that there was a user input error in this data since the ratio is so relatively high.

Question 2: Do big city apartments have less available parking for their residents than suburban aparmtnets?

Hypothesis: I think that since big city apartments are so dense with people, their apartments will be forced to have available parking for their residents.

Major City Apartments: Partial Parking Proportions

Covered	Garage	Paid	Valet	Total
0.194405757	0.337163417	0.007428041	0.013579387	.5526

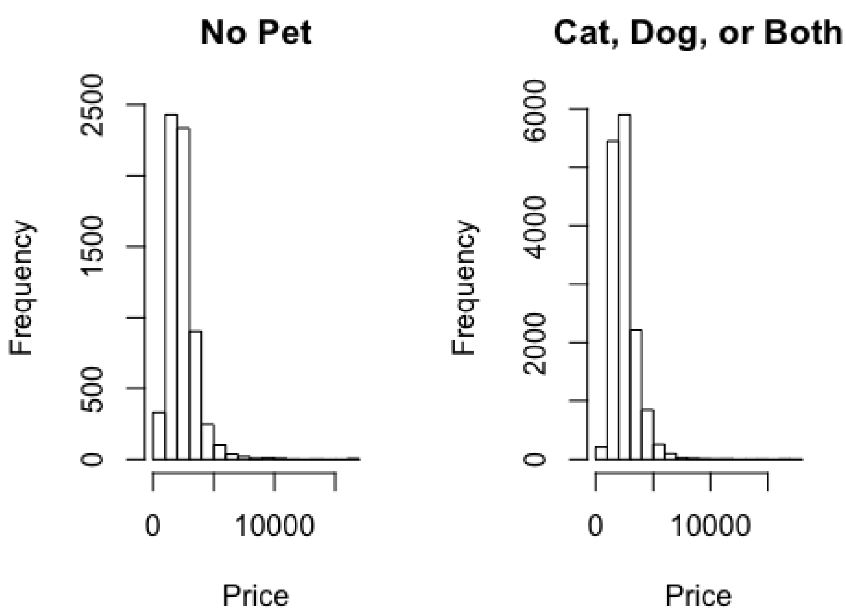
Suburban Apartments: Partial Parking Proportions

Covered	Garage	Paid	Valet	Total
0.572637518	0.194640339	0.001410437	0	.7687

Assuming that off-street, none, and street parking means that the apartments don't reserve a spot or parking structure for their residents, major cities have an available parking percentage of 55.26% while suburb cities have 76.87%. In retrospect, this seems to make more sense since suburban areas will have more available area to make a parking structure. Like West Village in Davis, there will be a lot of room to create covered parking spaces.

Question 3: Does the allowance of dogs, cats, or both tend to increase rent price?

Hypothesis: I would hypothesize that allowing dogs, cats, or both would increase rent price due to potential damages the pet could cause (scratching wood, etc.).



The histograms do look roughly the same; however, when taking the median of both, subsetted data frames, the median of no pets is equal to 2200 while the median of cats, dogs, or both is equal to 2295. Given their total numbers, a difference of 95 isn't too significant; however, over the months it has to potential to make a dent. All in all, yes the pet allowance does increase the median price of rent.

Question 4: How will a suburban and major city apartment differ even if they're geographical locations are very close?

<u>Hypothesis:</u> I hypothesize that square foot per price, number of bathrooms, and number of bedrooms will all favor the suburban aparmtment.

When taking a look at all the medians, the medians for bathrooms and bedrooms are equal to each other at 1:1 and 2:2. In addition, price to square foot has little difference being that the medians of Woodland and Sacramento are 0.598 and 0.572 respectively. In short, at least for this observation, the difference between suburban and major city apartment didn't matter. Note that no complete conclusion can be made due to it being one instance.

Question 5: On average, how much does adding 1 bathroom increase the price of rent and how does it compare to the average increase of 1 bedroom?

<u>Hypothesis:</u> Like the display of boxplots in 2.1, I think that for every 1 increase in bathrooms, the median of that will be greater than the median increase in price of bedrooms.

Interestingly, for every 1 bedroom increase, median price raised by 2270 on average and bathroom median price increased by 2088.417 on average. Although somewhat close, despite what the boxplots suggest, bedrooms increase the median price higher than bathrooms do. Note that .5 values of bathrooms weren't taken into account. The boxplots seem to show that bathrooms raise the price more; however, the values at bathrooms and bedrooms equal to 5 are so dramatic that it changes the median price raised on average to favor bedrooms costing more.

5. Limitations

Limitations of the data include the 26,316 missing data values. With the largest amount of data missing from the variable date updated. This does make sense; however, since if a profile isn't updated, it is recorded as N/A. Therefore, it would be more accurate to say that there are 13,177 missing values since the N/A values in the date updated variable gives information. Other information that is nice to note about the data is that the variable text often contains information that could go under city text. For example, row 8 shows that the Terrace Apartment Homes is located in the city Santa Clarita in the text variable, but the city text variable shows missing data, which could be due to user input. To be clear, the same can't be said about the variable parking and pets since no parking and no pets is denoted by "none" rather than left blank.

Outliers and anomalies existed in this data most likely due to user input error. In analysis, a good way to account for the anomalies is to change it to its proper value if it was an error. For example, in the price variable, anomalies included values of 9551095 and 34083742. These values were actually supposed to be a range of 955-1095 and 3408-3742 as shown by the text variable; therefore, this kind of error can be accounted for by replacing the extreme value for 955 and 3408. Also, in the square foot variable, a 200000 square foot apartment was recorded. The text variable of this advertisement didn't show any data of the real square footage; thus, the observation was just ommitted. Other ways to avoid these anomalies and outliers is to use the median rather than mean when computing parameters; however, fixing the data is the best solution.

This data was generated by users typing in information about their own apartment. Since this is the case, there is potential for the user to type in bias data. Also, errors will be present due to human error. For example, in row 17476, the price of the apartment was listed at 200 per month when in reality it is 2000 per month as shown by the text variable. In this way, user input error can happen easily by adding an extra 0 to price or square feet. Another example would be in row 1, which shows a price of 895 but the text variable lists it as 995 per month. This affects the reliability of the results because if the errors are present enough, the data set isn't representative of the true data, which makes the results of these diagnostics invalid as well.

When it comes to cities, many cities don't a significant amount of observations, which means that there's a high chance that the data collected isn't very representative of the population data.

Lastly, a limitation of this data set is that it can only apply to observations in the data set. Since factors such as economy and inflation can affect the price of rent, the observations can only be limitted to this data set.

Appendix

```
###1.
Apartment = readRDS("/Users/Troy/Desktop/cl_apartments.rds")
#View(Apartment) to get an overview of the data set
#colnames(Apartment)
#ncol(Apartment)
#nrow(Apartment)
#library(dplyr)
Year = subset(Apartment, grepl(('2018'), date_posted))
#grepl used to see the year of posts
#View(Year)
YearN = subset(Apartment, !grepl('2018', date_posted))
#all but one year was shown to be 2018 so !grepl to see last value
#Row 17902 looks like a bad observation (no information other than deletion)
#View(YearN)
Apt_fix = Apartment[-c(17902),]
#Gets rid of 17902 row in apartment data
YearFix = subset(Apt_fix, !grepl('2018', date_posted))
#View(YearFix) check to make sure deleted right row
#typeof(Apt_fix$bathrooms)
#table(Apartment$deleted)
#the one true value has no data (row 17902)
#hist(Apt_fix$price)
#some extreme max values
Price = sort(Apt_fix$price)
#table(Price)
#head(Price)
#tail(Price)
#hist(Price)
#median(Price)

Bedrooms = sort(Apt_fix$bedrooms)
#table(Bedrooms)
#head(Bedrooms)
#tail(Bedrooms)
#hist(Bedrooms)
#median(Bedrooms)

Bathrooms = sort(Apt_fix$bathrooms)
#table(Bathrooms)
#head(Bathrooms)
#tail(Bathrooms)
#hist(Bathrooms)
#median(Bathrooms)

Sqft = sort(Apt_fix$sqft)
#table(Sqft)
#head(Sqft, 20)
#tail(Sqft)
```

```r
#hist(Sqft)
#median(Sqft)
#quick fix of the data before continuing
#price and square feet adjustment
#tail(Sqft) shows odd value of 200,000 square feet
#tail(Price) shows odd values of 9551095 and 34083742
OddPrice = subset(Apartment, Apartment$price == "9951095")
OddPrice2 = subset(Apartment, Apartment$price == "34083742")
OddSqft = subset(Apartment, Apartment$sqft == 200000)
#View(OddPrice)
#Text description shows that $995-$1095 was entered but recorded as 9951095
#View(OddPrice2)
#Text description shows that $3408-$3742 was entered but recorded as 34083742
#View(OddSqft)
#This one must be a user input error

Apt_fix$sqft[Apt_fix$sqft == 200000] = NA
Apt_fix$price[Apt_fix$price == 9951095] = 995
Apt_fix$price[Apt_fix$price == 34083742] = 3408
#Patrick Discussion 4 Notes
###2.1
major_city1 = c("San Francisco", "Oakland", "San Jose",
                "Sacramento", "Los Angeles", "San Diego")
major_city = Apt_fix[Apt_fix$city %in% major_city1,]
#Patrick Discussion 4 Notes
suburb1 = c("West Sacramento", "Alhambra", "Palo Alto",
           "Folsom", "Livermore", "Sunnyvale")
suburb = Apt_fix[Apt_fix$city %in% suburb1,]
#table(Apt_fix$city) take a look at all the cities and determine 6 suburbs an
d 6 major cities
#suburbs were found on Google
#View(major_city)
#View(suburb)
#table(major_city$bedrooms)
#table(major_city$bathrooms)
#table(major_city$pets)
#table(major_city$laundry)
#table(major_city$parking)
#on second thought... won't include laundry or parking since that's just ever
yone-friendly

#table(suburb$bedrooms)
#table(suburb$bathrooms)
#table(suburb$pets)
#table(suburb$laundry)
#table(suburb$parking)

#median(major_city$bedrooms, na.rm = TRUE)
#median(suburb$bedrooms, na.rm = TRUE)
par(mfrow = c(1,2))
```

```r
hist(major_city$bedrooms, ylab = "Count", xlab = "Numb. of Bedrooms", main =
"Major City", xlim = c(0,5))
hist(suburb$bedrooms, ylab = "Count", xlab = "Numb. of Bedrooms", main = "Sub
urb")
#median bedroom comparison

#median(major_city$bathrooms, na.rm = TRUE)
#median(suburb$bathrooms, na.rm = TRUE)
par(mfrow = c(1,2))
hist(major_city$bathrooms, ylab = "Count", xlab = "Numb. of Bathrooms", main
= "Major City")
hist(suburb$bathrooms, ylab = "Count", xlab = "Numb. of Bathrooms", main = "S
uburb", xlim = c(0,5))
#median bathroom comparison

Major_pets = table(major_city$pets)
(Major_pets/sum(Major_pets))
Suburb_pets = table(suburb$pets)
(Suburb_pets/sum(Suburb_pets))

###2.2
library(ggplot2)
#plot(x = Apt_fix$bedrooms, y = Apt_fix$price)
#plot(x = Apt_fix$bathrooms, y = Apt_fix$price)
ggplot(Apt_fix, aes(group = bedrooms, y = price, x = bedrooms)) + geom_boxplo
t() + xlim(c(-.5,5.4)) + labs(y = "Price", title = "Bedrooms vs. Price", x =
"Bedrooms")
ggplot(Apt_fix, aes(group = bathrooms, y = price, x = bathrooms)) + geom_boxp
lot() + xlim(c(-.5,5.3)) + labs(y = "Price", title = "Bathrooms vs. Price", x
= "Bathrooms")

###2.3
SimCity1 = c("San Francisco", "Los Angeles", "San Diego")
SimCity = Apt_fix[Apt_fix$city %in% SimCity1,]
SF = subset(Apt_fix, Apt_fix$city == "San Francisco")
SD = subset(Apt_fix, Apt_fix$city == "San Diego")
LA = subset(Apt_fix, Apt_fix$city == "Los Angeles")
#Price
SFmed = median(SF$price, na.rm = TRUE)
SDmed = median(SD$price, na.rm = TRUE)
LAmed = median(LA$price, na.rm = TRUE)
ggplot(SimCity, aes(x=price, colour = city)) + geom_density() + labs(title =
"Compared Cities: Price", x = "Price") + geom_vline(aes(xintercept = SFmed),
linetype = "dashed", color = "blue") + xlim(c(0,10000)) + geom_vline(aes(xint
ercept = SDmed), linetype = "dashed", color = "green") + geom_vline(aes(xinte
rcept = LAmed), linetype = "dashed", color = "red")
#Square Foot
SFmedft = median(SF$sqft, na.rm = TRUE)
SDmedft = median(SD$sqft, na.rm = TRUE)
LAmedft = median(LA$sqft, na.rm = TRUE)
```

```r
ggplot(SimCity, aes(x=sqft, colour = city)) + geom_density() + labs(title = "
Compared Cities: Square Feet", x = "Square Feet") + geom_vline(aes(xintercept
= SFmedft), linetype = "dashed", color = "blue") + xlim(c(0,10000)) + geom_vl
ine(aes(xintercept = SDmedft), linetype = "dashed", color = "green") + geom_v
line(aes(xintercept = LAmedft), linetype = "dashed", color = "red")
#Bathrooms
#par(mfrow = c(1,3))
#hist(SF$bathrooms, xlim = c(0,3), main = "SF Bathrooms", xlab = "Numb. Of Ba
throoms")
#hist(SD$bathrooms, xlim = c(0,3), main = "SD Bathrooms", xlab = "Numb. Of Ba
throoms")
#hist(LA$bathrooms, xlim = c(0,3), main = "LA Bathrooms", xlab = "Numb. Of Ba
throoms")
#median(SF$bathrooms, na.rm = TRUE)
#median(SD$bathrooms, na.rm = TRUE)
#median(LA$bathrooms, na.rm = TRUE)
#median(SF$bedrooms, na.rm = TRUE)
#median(SD$bedrooms, na.rm = TRUE)
#median(LA$bedrooms, na.rm = TRUE)

###4.1
BestDeal = (Apt_fix$sqft/Apt_fix$price)
#higher value indicates better deal
#Apt_fix$BestDeal = BestDeal
#makes a new column for sqft to price ratio
#View(Apt_fix)
SJ = subset(Apt_fix, Apt_fix$city == "San Jose" & Apt_fix$price <= 850)
#View(SJ)


###4.2
suburb = Apt_fix[Apt_fix$city %in% suburb1,]
park_major = table(major_city$parking)
park_suburb = table(suburb$parking)
park_major/(sum(park_major))
park_suburb/(sum(park_suburb))


###4.3
frame = Apt_fix
frame$pets[frame$pets == "cats"] = "both"
frame$pets[frame$pets == "dogs"] = "both"
#changing cats and dogs to both for one universal outcome
none = subset(frame, frame$pets == "none")
both = subset(frame, frame$pets == "both")
par(mfrow = c(1,2))
hist(none$price, xlab = "Price", main = "No Pet")
hist(both$price, xlab = "Price", main = "Cat, Dog, or Both")
#median(none$price, na.rm = TRUE)
#2200
```

```
#median(both$price, na.rm = TRUE)
#2295

###4.4
#woodland and sacramento
WL = subset(Apt_fix, Apt_fix$city == "Woodland")
SAC = subset(Apt_fix, Apt_fix$city == "Sacramento")
#square foot per price
#median(WL$BestDeal, na.rm = TRUE)
#median(SAC$BestDeal, na.rm = TRUE)
#bathrooms
#median(WL$bathrooms, na.rm = TRUE)
#median(SAC$bathrooms, na.rm = TRUE)
#bedrooms
#median(WL$bedrooms, na.rm = TRUE)
#median(SAC$bedrooms, na.rm = TRUE)

###4.5
bed0 = subset(Apt_fix, Apt_fix$bedrooms == 0)
bed1 = subset(Apt_fix, Apt_fix$bedrooms == 1)
bed2 = subset(Apt_fix, Apt_fix$bedrooms == 2)
bed3 = subset(Apt_fix, Apt_fix$bedrooms == 3)
bed4 = subset(Apt_fix, Apt_fix$bedrooms == 4)
bed5 = subset(Apt_fix, Apt_fix$bedrooms == 5)
bed0M = median(bed0$price, na.rm = TRUE)
bed1M = median(bed1$price, na.rm = TRUE)
bed2M = median(bed2$price, na.rm = TRUE)
bed3M = median(bed3$price, na.rm = TRUE)
bed4M = median(bed4$price, na.rm = TRUE)
bed5M = median(bed5$price, na.rm = TRUE)
bed_avg_raise = ((bed1M-bed0M)+(bed2M-bed1M)+(bed3M-bed2M)+(bed4M-bed3M)+(bed
5M+bed4M))/6
#For every 1 bedroom increase, median price raised by 2270 on average
bat0 = subset(Apt_fix, Apt_fix$bathrooms == 0)
bat1 = subset(Apt_fix, Apt_fix$bathrooms == 1)
bat2 = subset(Apt_fix, Apt_fix$bathrooms == 2)
bat3 = subset(Apt_fix, Apt_fix$bathrooms == 3)
bat4 = subset(Apt_fix, Apt_fix$bathrooms == 4)
bat5 = subset(Apt_fix, Apt_fix$bathrooms == 5)
bat0M = median(bat0$price, na.rm = TRUE)
bat1M = median(bat1$price, na.rm = TRUE)
bat2M = median(bat2$price, na.rm = TRUE)
bat3M = median(bat3$price, na.rm = TRUE)
bat4M = median(bat4$price, na.rm = TRUE)
bat5M = median(bat5$price, na.rm = TRUE)
bat_avg_raise = ((bat1M-bat0M)+(bat2M-bat1M)+(bat3M-bat2M)+(bat4M-bat3M)+(bat
5M+bat4M))/6
```

```
###5
#For every 1 bathroom increase, median price raised by 2088.417 on average
#NAs = sapply(Apartment,is.na)
#Frequency_Of_Missing_Values = apply(NAs, 2, sum)
#Frequency_Of_Missing_Values
#sum(Frequency_Of_Missing_Values)
```

Sources:

- https://www.google.com/search?sa=X&q=los+angeles+population&stick=H4sIAAAAAAAAAOPgUeLSz9U3MDYoTDIu0dLKTrbSz8lPTizJzM_TLy4B0sUlmcmJOfFFqelAIauC_ILSHLBsF6MFF5JWIZK0mnFxgLQaxqeYkKUxxSyngBSNALvIxYTnAAAA&ved=2ahUKEwju08KqkK7eAhUMBnwKHQccAecQth8wAXoECAUQAw (United States Census Bureau)