Airbnb Rent Price in Toronto, Canada

By Thomas Kwok

Studying what factors influence Airbnb rent price, using exploratory analysis and geographically weighted regression.

Section 1: Introduction:

Whenever someone is going on vacation, they often choose between living in a hotel or an Airbnb. The benefit of an Airbnb over a hotel is that it can provide the person with an entire house instead of just one room. But what exactly factors into the cost of an Airbnb? Since these are property owned by individuals, the prices are set by owners instead of a corporation (like hotels), so they don't have to worry about competing against others.

In this research project, we will look at 3501 observations of Airbnb in Toronto from the July 2017 listings, in order to figure out what factors into the price of an Airbnb apartment, focusing specifically on the role spatial information plays in determining the price. The data comes from the Tomslee website, a data scientist who scraped the files from Airbnb. This data contains fifteen variables as shown in figure 1.

In this data set, we focused our study on entire home as we wanted to distinguish Airbnb options from hotel options that provide a private or shared room. We also filtered our data set to apartments that have ten or more reviews, to find apartments that were actually rented, occupied, and reviewed. This is important as anyone can post their apartment on Airbnb and set their price, but if no one rents the apartment then it would not be beneficial to our research.

Prior research shows that studies of Airbnb prices have been done in New York City and Seattle. These studies use regression analysis with little spatial analysis. The main geographical analysis of those projects look into neighborhoods or boroughs, instead of actual longitude and latitude.

Figure 1: Variables from Toronto Airbnb data set. The variables that we studied deeper are bolded.

Variable	Description
Room_ID	Integer: ID given by airbnb per room
Survey_ID	Integer: ID given by airbnb for when data is scraped. Same integer since scraped on same day.
Host_ID	Integer: ID given by airbnb per owner of room
Room_Type	Factor with 1 level: Entire home/apartment
City	Factor with 1 level: Toronto
Neighborhood	Factor with 124 levels: (ex: Waterfront)
Reviews	Integer: 10 or more reviews
Overall_Satisfaction	Number: Reviewer rating of 0-5 with 5 being best.
Accommodates	Integer: Amount of people owner states can live in house.
Bedrooms	Integer: The number of bedrooms in the house.
Price	Integer: The cost per night to rent this airbnb
Last Modified	Factor with 3501 levels: All modified 7-10-2017 but the time scrapped differ.
Latitude	Number: Latitude of the house/apartment
Longitude	Number: Longitude of the house/apartment
Location	Factor with 3501 levels: Airbnb code for location

Section 2 Methods:

In this project, we will look into the spatial data for the Toronto Airbnb locations and see the spatial impact of the locations to the prices. We will start by looking at the geographic breakdown of Airbnb price in Toronto, look at the number of apartments listed and average price for each neighborhood, and derive a geographically weighted regression analysis to predict the cost of an Airbnb price. Our variables of interest are price, neighborhood, reviews, overall satisfaction, accommodates, bedrooms, longitude, and latitude. We will use global regression (linear regression) and local regression (geographically weighted regression) to analyze the impact that geographic location plays on price.

Section 3 Exploratory Analysis:

To begin, we created a histogram of the Airbnb prices from the data set to see where most renting prices fall under. From the histogram (figure 2), we found out that the prices are heavily skewed and a quantile test (figure 4) showed that approximately 90% of the prices fall below $180 per night, because of that we created another histogram (figure 3) to see the frequency of prices under $180. The skewness here also means that we may need to take the logarithm of price to normalize the results.

Figure 2: Histogram showing frequency of Airbnb listing per price range in Toronto.

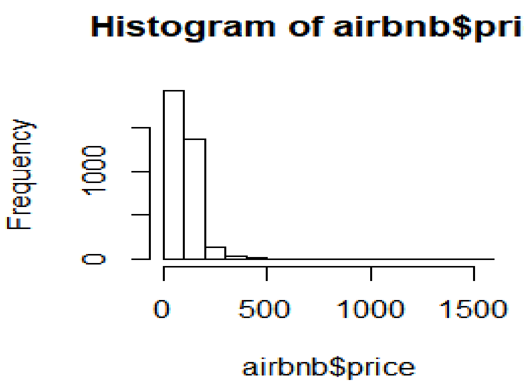


Figure 3: Histogram showing frequency of Airbnb listings under $180 in Toronto.

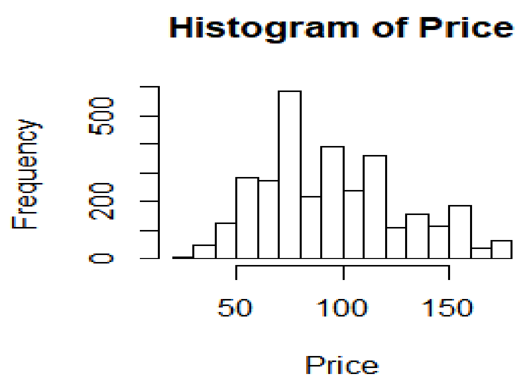


Figure 4: Airbnb Price by Quantile in Toronto

Quantile	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
Price	22	58	70	78	88	96	108	119	143	180	1599

Next, we looked at the Airbnb location by neighborhood to see where the most listings are located. We filtered the listings to the nine most popular neighborhoods, all with at least one hundred Airbnb listings. We discovered that the majority of Airbnb are listed in the Waterfront Communities (figure 5), which is located in downtown Toronto: near many of Toronto's tourist attractions. Next, we looked at the average cost per apartment in these neighborhoods, which also showed that the Waterfront Communities is the most expensive (figure 6) but the average cost per Airbnb isn't too different between neighborhoods.

Figure 5: Neighborhoods with at least 100 Airbnb listings

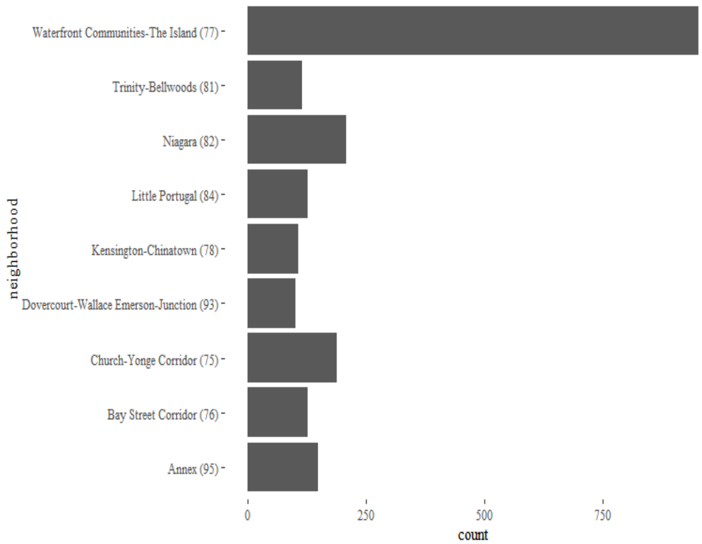
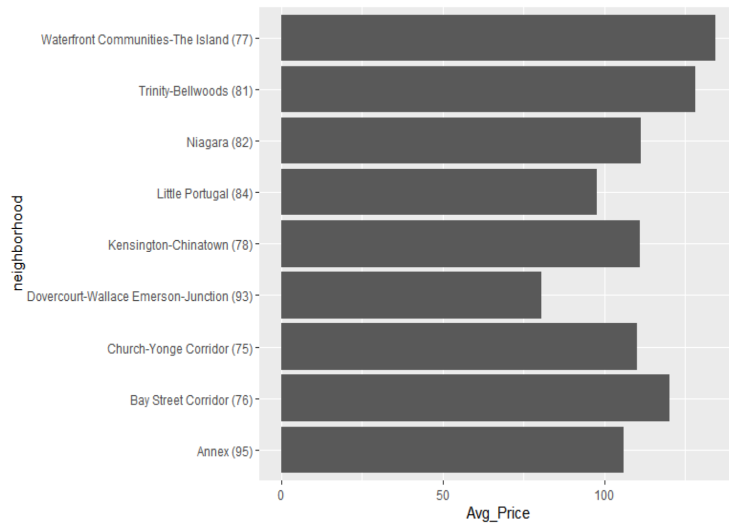


Figure 6: Average Price each these same neighborhoods



With this information, we looked at a heat map of Toronto for number of bedrooms and accommodates to see if there are any connections we can make from all three. This showed that most apartments in Toronto have approximately 2-4 bedrooms, and able to host approximately 4-6 people. There appears to be a cluster of approximately houses that is said to be able to accommodate 10-16 people near the bottom of the map, where downtown Toronto is located close to the Waterfront, and many of these seem to be between 2-6 bedrooms.

Figure 7: Heat map of number of bedrooms in Toronto.

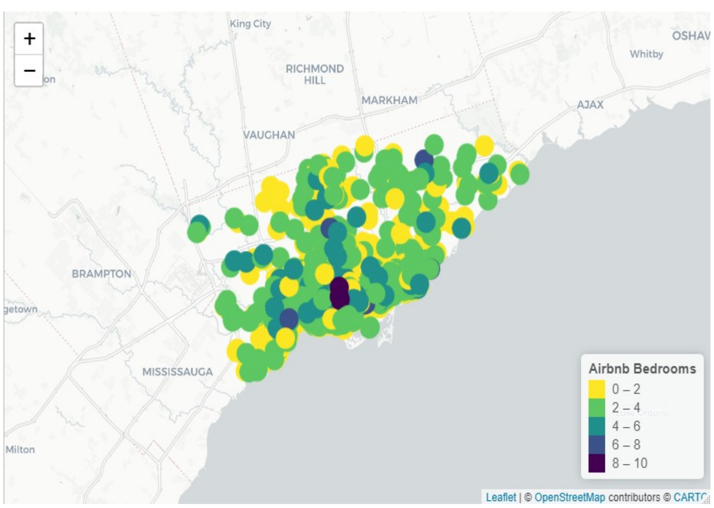
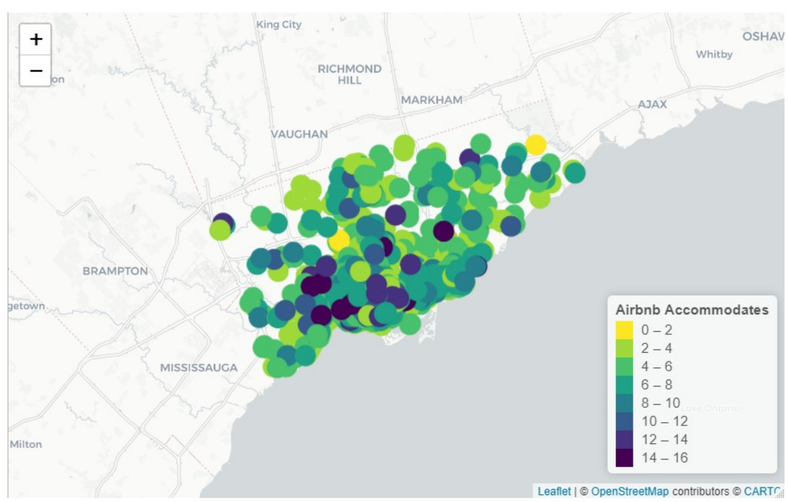


Figure 8: Heat map of number of people in Toronto the houses can accommodate.



With this result, we wanted to create a heat map of the prices of Airbnb in Toronto and see if they support the same findings as the number of bedrooms and accommodates. First we created a heat map that looked at the prices of all of the airbnb listings in Toronto and then we split the prices into four groups, $0-$200, $200-$400, $400-600, and $600-$1600 and found that 3252 of the 3501 observations fall below $200 and 212 observations fall between $200 to $400. This means that less than fifty listings are above $400 only.

Figure 8: Airbnb Price in Toronto

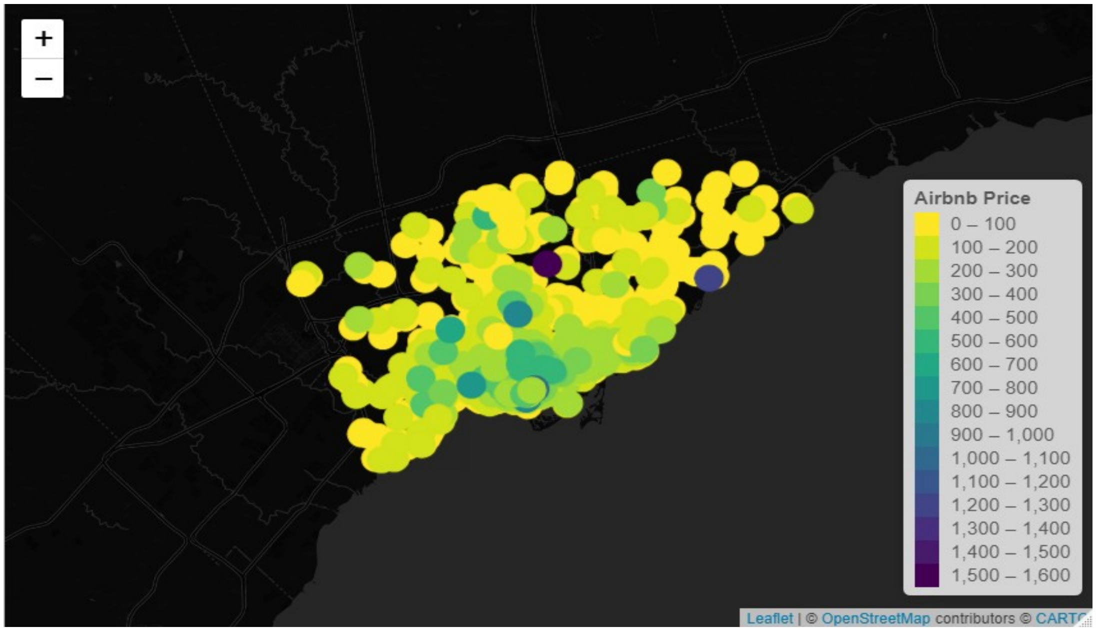


Figure 9: Toronto Airbnb between $0-$200 per night

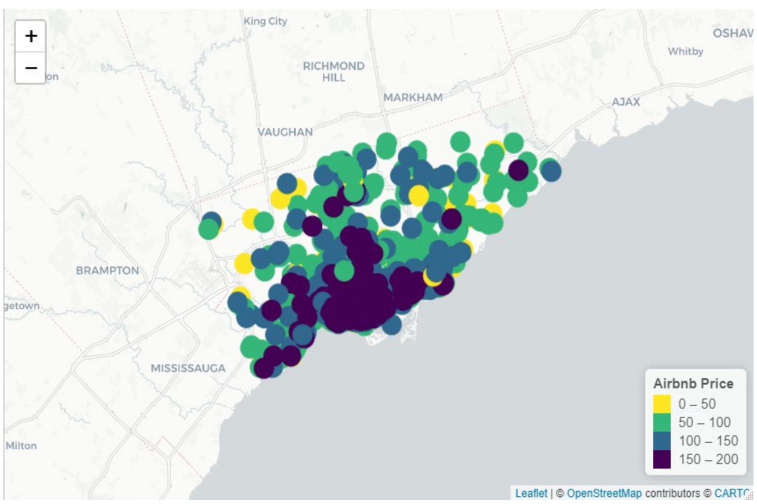
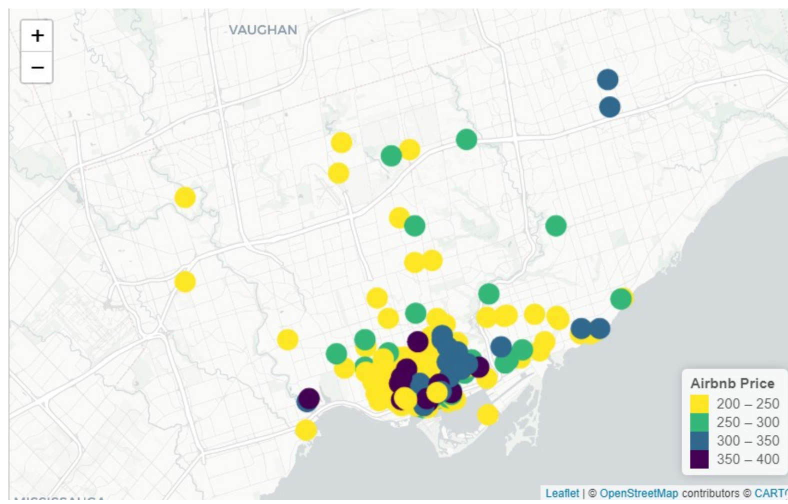


Figure 10: Toronto Airbnb between $200-$400 per night



Section 4 Results:

The Geographically Weighted (GW) model of choice used was GW regression to see if there is a spatial relationship between our dependent variable Price to our independent variables. The basic form of the GW regression model is:

$$y_i = \beta_{i0} + \sum_{k=1}^{m} \beta_{ik} x_{ik} + \epsilon_i$$

In this model, the dependent variable is dependent at location i and xik is the independent variables on location i. For this geographically weighted model, nearer observations are given more weight to price than further observations.

We first start with a global correlation analysis, a simple linear model where there is no geographic weight plays zero impact on Airbnb price. Our linear model equation is listed below:

Airbnb_Price ~ Bedrooms + Accommodates + Reviews + Overall Satisfaction + Longitude + Latitude

The model output is listed below:

```
--------------------------------------------------------------------------------
lm(formula = price ~ ., data = airbnb_df)

Residuals:
     Min      1Q  Median      3Q     Max 
-225.34  -27.62   -5.67   18.85 1309.57 

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)    
(Intercept)          2.489e+04  2.705e+03   9.203  < 2e-16 ***
bedrooms             2.598e+01  1.646e+00  15.785  < 2e-16 ***
accommodates         1.172e+01  7.477e-01  15.674  < 2e-16 ***
reviews             -1.383e-01  2.624e-02  -5.271 1.44e-07 ***
overall_satisfaction  2.717e+01  3.629e+00   7.487 8.87e-14 ***
longitude            1.255e+02  2.288e+01   5.484 4.46e-08 ***
latitude            -3.440e+02  3.452e+01  -9.966  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 59.76 on 3494 degrees of freedom
Multiple R-squared:  0.3651,	Adjusted R-squared:  0.364 
F-statistic: 334.9 on 6 and 3494 DF,  p-value: < 2.2e-16
--------------------------------------------------------------------------------
```

The data shows that all variables are significant due to their p-value and suggests that Airbnb price is positively associated with the number of bedrooms, accommodation, satisfaction score, and longitude while negatively associated with number of reviews and latitude. The multiple R-square is 0.365 and the adjusted R-square, which penalizes using more variables is 0.364.

Next we ran a model selection, with a bisquare kernel and bandwidth of 41, comparing a combination of variables to price through their Akaike Information Criterion corrected (AICc) scores to find the smallest amount of predictors for Airbnb price. This selection runs 21 regression models with various combinations of the six variables and provides a graph comparing each model by their AICc score.

Figure 9: Model view of the stepwise specification procedure

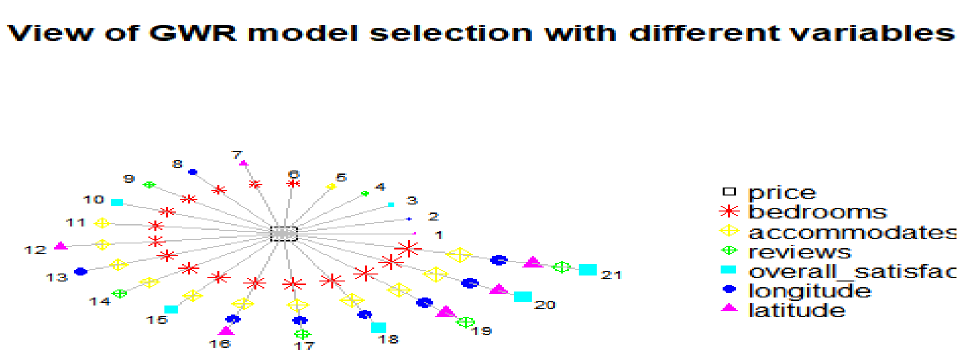
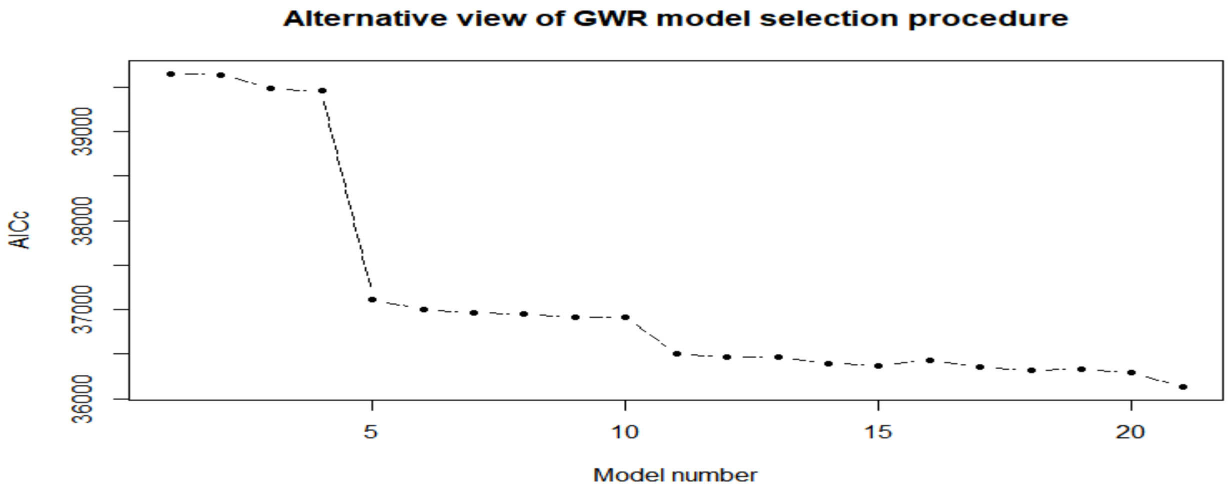
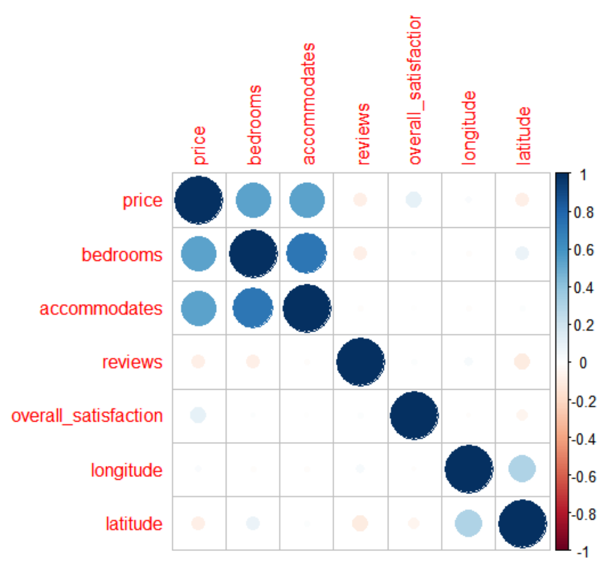


Figure 10: AICc values for the 21 GW regression models



As figure 10 shows, if only compared to one variable the best occurs in model 6 with Airbnb price compared to bedrooms alone, slightly outperforming model 5 which compares it to accommodates. The next sizable drop occurs in model 11 which compares Airbnb price to bedrooms and accommodates, showing us that those two variables are the best predictors of price. Running a correlation plot of all variables show that bedroom and accommodates are heavily correlated with one another, leading to potential multicollinearity. We decided for our model to only use bedroom as the number an Airbnb can accommodate is subjective to perception whereas the number of bedrooms is objective and cannot be changed.

Figure 11: Correlation matrix comparing the predictor variables to one another and to price



First, we ran a global regression model comparing Airbnb price to number of bedrooms alone, getting an R-square of 0.2856.

Model output: Global Regression: Price ~ bedroom

```
****************Results of Global Regression*****************
lm(formula = price ~ bedrooms, data = airbnb.spdf)

Coefficients:

               Estimate Std. Error t value Pr(>|t|)
(Intercept)   53.280      1.921   27.74   <2e-16 ***
bedrooms      44.152      1.180   37.40   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 63.34 on 3499 degrees of freedom
Multiple R-squared:  0.2856,    Adjusted R-squared:  0.2854
F-statistic:  1399 on 1 and 3499 DF,  p-value: < 2.2e-16
*************************************************************
```

Next we ran GW regression with four kernel methods (gaussian, bisquare, boxcar, and tri-cube), optimizing bandwidth for each method by using the number of nearest neighbors methhod and compared the results of each GW regression to one another and to the global regression method. Our analysis showed that a bisquare kernel using 41 nearest neighbors provided the best output, with an R-square of 0.646. The model output is displayed below.

```
Model output: GW Regression with bisquare kernel and bandwidth 41: Price ~ bedrooms.
****************Summary of GWR coefficient estimates:*****************
               Min.  1st Qu.   Median  3rd Qu.   Max.
   Intercept -348.624   37.812   53.421   69.648 175.09
   bedrooms   -38.483   27.327   38.514   53.536 369.43
   ***********************Diagnostic information*************************
   Number of data points: 3501
   Effective number of parameters (2trace(S) - trace(S'S)): 632.3534
   Effective degrees of freedom (n-2trace(S) + trace(S'S)): 2868.647
   AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 37623.7
   AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 36996.47
   Residual sum of squares: 6952696
   R-square value:  0.646213
   Adjusted R-square value:  0.5681984
*************************************************
```

Lastly, we ran a bisquare kernel with 131 nearest neighbors, using the log of price to account for skewness and eliminate any negative price predictions since it does not make sense for an Airbnb owner to pay their guest to stay in their location, and received an R-square of 0.518.

```
Model output: GW Regression with bisquare kernel and bandwidth 41: Log(Price) ~ bedrooms.
****************Summary of GWR coefficient estimates:*****************
               Min. 1st Qu.  Median 3rd Qu.   Max.
   Intercept 3.61619 4.02418 4.19815 4.34554 4.5685
   bedrooms  0.12342 0.27346 0.30668 0.34323 0.6051
   ***********************Diagnostic information*************************
   Number of data points: 3501
   Effective number of parameters (2trace(S) - trace(S'S)): 208.3779
   Effective degrees of freedom (n-2trace(S) + trace(S'S)): 3292.622
```

```
    AICc (GWR book, Fotheringham, et al. 2002, p. 61, eq 2.33): 2351.338
    AIC (GWR book, Fotheringham, et al. 2002,GWR p. 96, eq. 4.22): 2182.82
    Residual sum of squares: 366.0939
    R-square value:  0.5179072
    Adjusted R-square value:  0.487388
***********************************************************************
```

Section 5 Conclusion:

Comparing our global regression R-square to our GW regression R-square with bisquare kernel, it seems that geographic location may play a factor in Airbnb price as the R-square from our GW Regression was greater than the R-square for the linear regression model for all variables and also bedrooms alone. We also compared GW regression models for Airbnb price and Airbnb log price, and found that the R-square value decreased for log price compared to the R-square for price.

Next, we compared the coefficient estimates for both models, comparing them to one another to see if we should take the log of price. These models are provided below:

Figure 12: Basic GW regression coefficient for bedrooms to Airbnb Price estimate

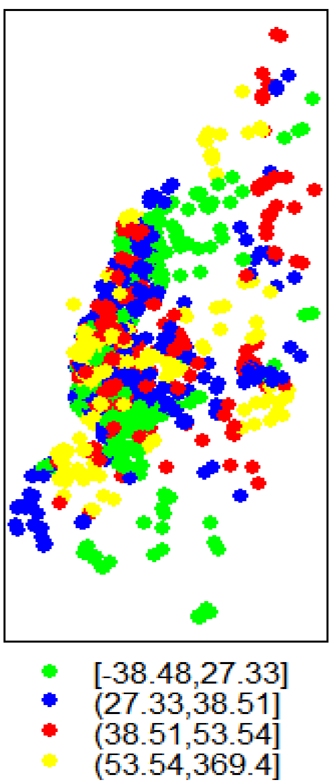
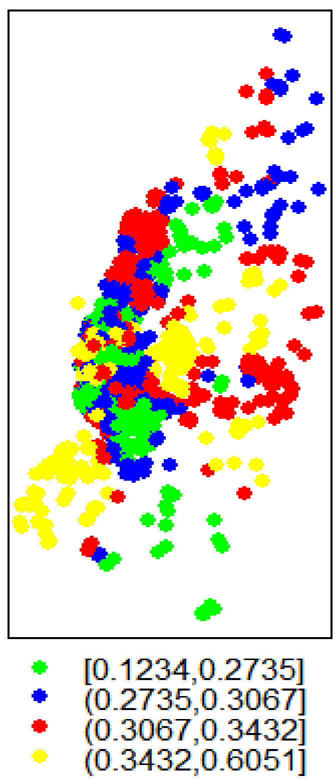


Figure 13: Basic GW regression coefficient for bedrooms to Airbnb Log(Price) estimate



Looking at the coefficient estimate graphs, we notice that the range for the coefficients is much larger for price to bedroom than log price to bedroom, with the range of the green points containing zero, which is not meaningful.

Next, we looked at a t-score statistical significance between price and log price to bedroom to see the geographic significance of bedrooms to price in both instances. In our graphs, the red circles show that the variable is not statistically significant and the green shows it may be statistically significant.

Figure 14: Statistical Significance for bedroom to Airbnb Price

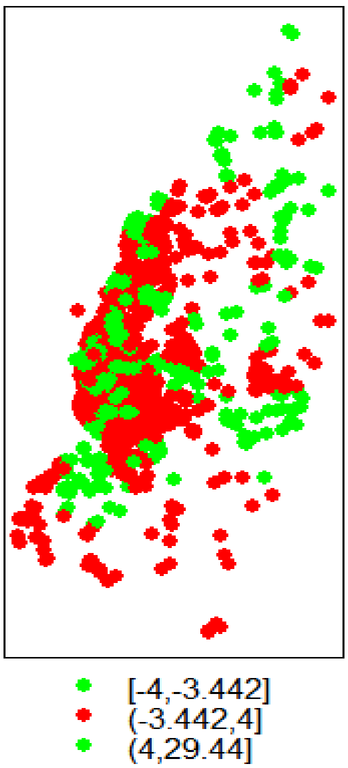
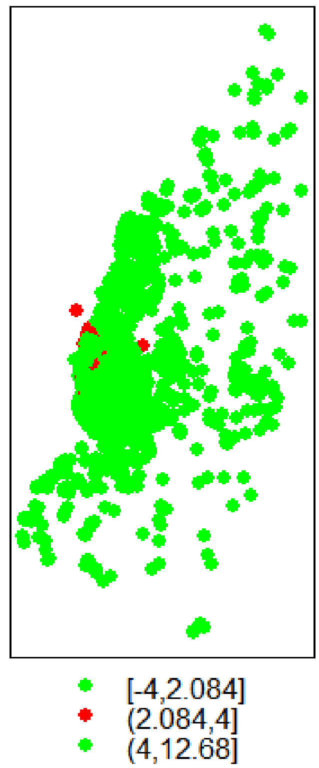


Figure 15: Statistical Significance for bedroom to Airbnb Log(Price)



The data shows that when using log price, there are less statistically insignificant points for bedroom than when using price. Using the coefficient estimate graph, the skewness of price, and the fact that the GW regression for price produces negative price values, the best model is to do a GW regression of log price to bedroom. Although this may produce a smaller R-square value, using the log price seems to produce more statistically significant results than using price.

In the end, the number of bedrooms is also the most important independent variable for determining Airbnb price. A geographic model is preferred over a global regression model, as location plays a factor in Airbnb price. The preferred weighting method is using a bisquare kernel with a bandwidth of 131 nearest neighbor Airbnb, and a regression formula that compares log price to bedroom.

Section 6 References:

1. Slee, Tom. “Airbnb Data Collection: Get the Data.” <http://tomslee.net/airbnb-data-collection-get-the-data> [2017].

2. Gollini, Isabella et al. “GWmodel: An R Package for Exploring Spatial Heterogeneity Using Geographically Weighted Models.” pp. 1-44. In *Journal of Statistical Software*. [2015].

3. RH. “Introduction to Geographically Weighted Regression.” pp. 1-9. <https://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/gwr.pdf> [2009].

4. “Advanced Spatial Analysis with R.” pp. 1-8. In *Spatial Statistics in R*. <https://www.geos.ed.ac.uk/~gisteac/gt/r/pracs/gwr_using_r/prac_4_1.pdf> [2013].

5. Lu, Binbin et al. pp. 1-42. In *Science Foundation Ireland*. <https://arxiv.org/ftp/arxiv/papers/1312/1312.2753.pdf> [2013].

Section 7 R-Code:

```
setwd("C:/Users/thoma/Desktop/Spatial HW/Project/")
airbnb <- read.csv("Project-Airbnb_Toronto.csv")
library(leaflet)
library(leaflet.extras)
library(rgdal)
library(GWmodel)
library(ModelMap)
library(gstat)
library(tidyverse)
library(ggplot2)
library(ggfittext)
library(ggthemes)
library(gridExtra)
library(RColorBrewer)
library(corrplot)

summary(airbnb)

#summary of prices
summary(airbnb$price)
quantile(airbnb$price, seq(0,1,by=0.1))
quantile(airbnb$price, seq(.9,1,by=0.01))

#histogram
hist(airbnb$price)

#Prices filtered
airbnb_filter <- airbnb %>%
  filter(price <=180)
Price <- airbnb_filter$price

#Under $180 per night
hist(Price)

#Neighborhood filters
airbnb_filter3 <- airbnb %>%
  group_by(neighborhood) %>%
  filter(n() >= 100)

#ggplot of 9 most available neighborhood
ggplot(data=airbnb_filter3, aes(x=neighborhood)) +
```

```
  geom_bar() +
  theme_tufte()+
  coord_flip()

#comparing average cost of airbnb and neighborhood
price_lattice <- airbnb_filter3 %>%
  group_by(neighborhood) %>%
  summarise(Avg_Price=mean(price, na.rm = TRUE))
ggplot(data=price_lattice, aes(y=Avg_Price, x=neighborhood))+geom_bar(stat="identity")+coord_flip()

#using leaflet

#heatmap of bedrooms
pal_bed <- colorBin("viridis", domain=airbnb$bedrooms, bins = 5, pretty = TRUE,
               na.color = "#808080", alpha = FALSE, reverse = TRUE,
               right = FALSE)

leaflet(data = airbnb) %>%
  addProviderTiles("CartoDB.Positron") %>%
  addCircleMarkers(~longitude, ~latitude, color = ~pal_bed(bedrooms), weight = 1, radius=10, fillOpacity = 1, opacity = 1,
                   label = paste("Name:", airbnb$bedrooms)) %>%
  addLegend("bottomright", pal = pal_bed, values = ~bedrooms,
            title = "Airbnb Bedrooms",
            opacity = 1
  )

#heatmap accommodates
pal_accommodates <- colorBin("viridis", domain=airbnb$accommodates, bins = 10, pretty = TRUE,
                    na.color = "#808080", alpha = FALSE, reverse = TRUE,
                    right = FALSE)

leaflet(data = airbnb) %>%
  addProviderTiles("CartoDB.Positron") %>%
  addCircleMarkers(~longitude, ~latitude, color = ~pal_accommodates(accommodates), weight = 1, radius=10, fillOpacity = 1, opacity = 1,
                   label = paste("Name:", airbnb$accommodates)) %>%
  addLegend("bottomright", pal = pal_accommodates, values = ~accommodates,
            title = "Airbnb Accommodates",
            opacity = 1
  )

#heatmap price
pal <- colorBin("viridis", domain=airbnb$price, bins = 12, pretty = TRUE,
               na.color = "#808080", alpha = FALSE, reverse = TRUE,
               right = FALSE)

leaflet(data = airbnb) %>%
  addProviderTiles(providers$CartoDB.DarkMatterNoLabels) %>%
  addCircleMarkers(~longitude, ~latitude, color = ~pal(price), weight = 1, radius=10, fillOpacity = 1, opacity = 1,
                   label = paste("Name:", airbnb$price)) %>%
  addLegend("bottomright", pal = pal, values = ~price,
            title = "Airbnb Price",
            opacity = 1
  )

################## Price Categories
#mutate price
airbnb_data <- airbnb %>%
  mutate(price_categories = cut(price,
                                breaks = c(0, 200, 400, 600, 2000),
                                labels = c('$0-199', '$200-399', '$400-599', '$600-2000'),
                                include.lowest = TRUE,
                                right = FALSE))

#heatmap categories
grp1 <- airbnb_data %>%
  filter(price_categories == '$0-199')
grp2 <- airbnb_data %>%
  filter(price_categories == '$200-399')
grp3 <- airbnb_data %>%
```

```
  filter(price_categories == '$400-599')
grp4 <- airbnb_data %>%
  filter(price_categories == '$600-2000')


#grp1
pal1 <- colorBin("viridis", domain=grp1$price, bins = 4, pretty = TRUE,
                 na.color = "#808080", alpha = FALSE, reverse = TRUE,
                 right = FALSE)


leaflet(data = grp1) %>%
  addProviderTiles("CartoDB.Positron") %>%
  addCircleMarkers(~longitude, ~latitude, color = ~pal1(price), weight = 1, radius=10, fillOpacity = 1, opacity = 1,
                   label = paste("Name:", grp1$price)) %>%
  addLegend("bottomright", pal = pal1, values = ~price,
            title = "Airbnb Price",
            opacity = 1
  )


#grp2
pal2 <- colorBin("viridis", domain=grp2$price, bins = 4, pretty = TRUE,
                 na.color = "#808080", alpha = FALSE, reverse = TRUE,
                 right = FALSE)


leaflet(data = grp2) %>%
  addProviderTiles("CartoDB.Positron") %>%
  addCircleMarkers(~longitude, ~latitude, color = ~pal2(price), weight = 1, radius=10, fillOpacity = 1, opacity = 1,
                   label = paste("Name:", grp2$price)) %>%
  addLegend("bottomright", pal = pal2, values = ~price,
            title = "Airbnb Price",
            opacity = 1
  )

########################################################################3
#select only important info in airbnb
airbnb_df <- airbnb %>%
  dplyr::select(price, bedrooms, accommodates, reviews, overall_satisfaction, longitude, latitude)


#global regression = All dataset
global <- lm(price~., data=airbnb_df)
summary(global)


##model selection: See the importance of each variable and the model
library("GWmodel")
library(spgwr)
airbnb.spdf <- SpatialPointsDataFrame(airbnb[, 13:14], airbnb)


DeVar <- "price"
InDeVars <- c("bedrooms", "accommodates", "reviews", "overall_satisfaction", "longitude", "latitude")
model.sel <- model.selection.gwr(DeVar, InDeVars, data = airbnb.spdf,
                                 kernel = "bisquare", adaptive = TRUE, bw = 41)
sorted.models <- model.sort.gwr(model.sel, numVars = length(InDeVars),
                                ruler.vector = model.sel[[2]][,2])
model.list <- sorted.models[[1]]
model.view.gwr(DeVar, InDeVars, model.list = model.list)
plot(sorted.models[[2]][,2], col = "black", pch = 20, lty = 5,
     main = "Alternative view of GWR model selection procedure",
     ylab = "AICc", xlab = "Model number", type = "b")


#correlation between variables of interest
cplot <- cor(airbnb_df) #this tells me bedroom and accommodates is biggest factor for price. Interesting overall satisfaction isn't big factor
corrplot(cplot, method="circle")


#Create distance matrix
library("GWmodel")
airbnb.spdf <- SpatialPointsDataFrame(airbnb[, 13:14], airbnb)
DM <- gw.dist(dp.locat = coordinates(airbnb.spdf))


#global regression = No spatial data, R-square .2856
```

```
global2 <- lm(price~bedrooms, data=airbnb.spdf)
summary(global2)

#gaussian (n=24) R-square .445
bw.gwr1 <- bw.gwr(price~bedrooms, data=airbnb.spdf, approach="AICc", kernel="gaussian", adaptive=TRUE)
bgwr.res <- gwr.basic(price~bedrooms, data=airbnb.spdf, bw=bw.gwr1, kernel="gaussian", adaptive=TRUE)
print(bgwr.res)

#bisquare (n=41), R-square .646
bw.gwr2 <- bw.gwr(price~bedrooms, data=airbnb.spdf, approach="AICc", kernel="bisquare", adaptive=TRUE)
bgwr.res2 <- gwr.basic(price~bedrooms, data=airbnb.spdf, bw=bw.gwr2, kernel="bisquare", adaptive=TRUE)
print(bgwr.res2)

#boxcar (n=19) R-square = .597
bw.gwr3 <- bw.gwr(price~bedrooms, data=airbnb.spdf, approach="AICc", kernel="boxcar", adaptive=TRUE)
bgwr.res3 <- gwr.basic(price~bedrooms, data=airbnb.spdf, bw=17, kernel="boxcar", adaptive=TRUE)
print(bgwr.res3)

#tri-cube (n=41) R-square = .635
bw.gwr5 <- bw.gwr(price~bedrooms, data=airbnb.spdf, approach="AICc", kernel="tricube", adaptive=TRUE)
bgwr.res5 <- gwr.basic(price~bedrooms, data=airbnb.spdf, bw=bw.gwr5, kernel="tricube", adaptive=TRUE)
print(bgwr.res5)

#Log-Price to deal with skew and remove negative prices = Bandwidth = 131 but R-square is .518
bw.gwr6 <- bw.gwr(log(price)~bedrooms, data=airbnb.spdf, approach="AICc", kernel="bisquare", adaptive=TRUE)
bgwr.res6 <- gwr.basic(log(price)~bedrooms, data=airbnb.spdf, bw=bw.gwr6, kernel="bisquare", adaptive=TRUE)
print(bgwr.res6)

##########################################################################3
#coeff between price and bedrooms
colours = c("green", "blue", "red", "yellow")
spplot(bgwr.res2$SDF, "bedrooms", cuts=quantile(bgwr.res2$SDF$bedrooms),
      col.regions=colours,
      main = "Basic GW regression coefficient estimates for price and bedroom")

#statistical significance
t = bgwr.res2$SDF$bedrooms / bgwr.res2$SDF$bedrooms_SE
sig.map = SpatialPointsDataFrame(airbnb.spdf, data.frame(t))
colours=c("green","red","green")
breaks=c(min(t),-4,4,max(t))
spplot(sig.map, cuts=breaks, col.regions=colours, main = "Statistical Significance of bedroom to Price")

###########################################################################3
#log Price to combat negatives that don't make sense

#coeff between price and bedrooms
colours = c("green", "blue", "red", "yellow")
spplot(bgwr.res6$SDF, "bedrooms", cuts=quantile(bgwr.res6$SDF$bedrooms),
      col.regions=colours,
     main = "Basic GW regression coefficient estimates for log(price) and bedroom")

#statistical significance
t = bgwr.res6$SDF$bedrooms / bgwr.res6$SDF$bedrooms_SE
sig.map = SpatialPointsDataFrame(airbnb.spdf, data.frame(t))
colours=c("green","red","green")
breaks=c(min(t),-4
       ,4,max(t))
spplot(sig.map, cuts=breaks, col.regions=colours, main = "Statistical Significance of bedroom to log(Price)")
```