

# **Is the Relationship between Price of Airbnbs in Chicago and Accommodates depending on Property Type?**

**Group II**

**Jaeyeon Won, Charles Krebs, Christopher Wesseln**

# Background

- Dataset: Chicago Airbnb
- Sample: 500 AirBnBs in Chicago, August, 2008 to May, 2017
- Response Variable: Airbnb rental Price
- Explanatory Variables: Property Type (**Apartment\***, Condominium, House, Other), Bed Type (**Real Bed\***, Not Real Bed), Accommodates
- Purpose of investigating the interaction between Accommodates and Property Type

\*Bolded one is the reference group

## Model Selection

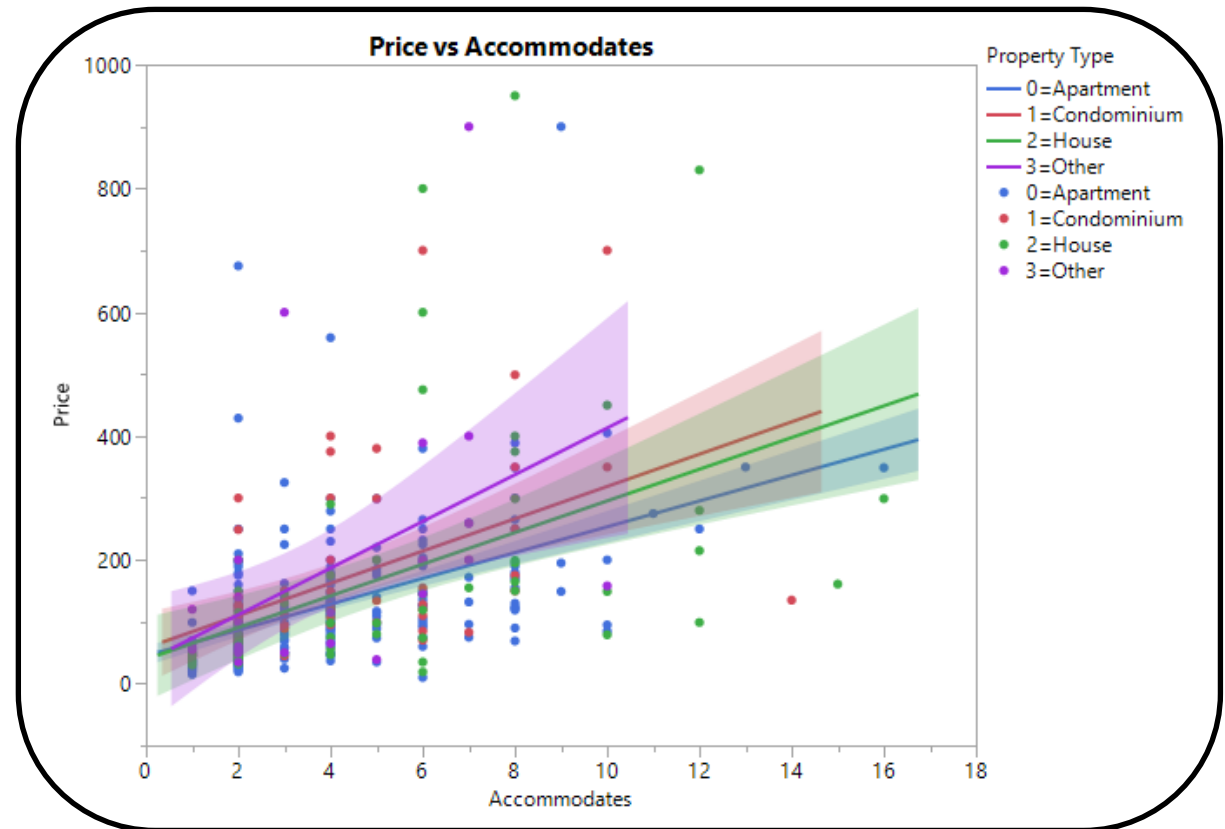
- Overall significance
- Highest adjusted  $R^2$  value
- Interaction in scatterplot
- More significant interaction terms

Analysis of Variance				
Source	DF	Sum of Squares	Mean Square	F Ratio
Model	8	2106301.7	263288	22.2691
Error	491	5805093.7	11823	Prob > F
C. Total	499	7911395.5		<.0001*

# Data Analysis

## 1. Scatterplot

The relationship between Price and Accommodates appears to be slightly different depending on different Property Types, since the slopes are slightly different. The biggest difference is between Apartment and Other types of property while Condominium and House have almost the same slopes.



## 2. Multiple Regression Model

$$\widehat{Price} = 46.923 + 20.607 * Accommodates - 18.075 * Not Real Bed + 11.553 * Condominium - 7.423 * House - 10.077 * Other + 5.383 * Condominium * Accommodates + 4.974 * House * Accommodates + 16.987 * Other * Accommodates$$

### (a) Interpretations:

- **Condominium\*Accommodates:** For every additional person the AirBnB accommodates, the predicted Price of AirBnB increases by \$26.0 for Condominium type property, holding Bed Type constant.
- **House \* Accommodates:** For every additional person the AirBnB accommodates, we predict an increase Price of AirBnB of \$25.9 for House type property, holding Bed Type constant.

- **Other \* Accommodates:** For every additional person the AirBnB accommodates, the predicted Price of AirBnB increases by \$37.6 for other types of property, holding Bed Type constant.

### (b) Adjusted $R^2$

- 25.4% of the variability in Price can be explained using a linear model including Accommodates, Bed Type, and Property Type after accounting for the complexity of the model.

Summary of Fit	
RSquare	0.266236
RSquare Adj	0.254281
Root Mean Square Error	108.7336
Mean of Response	135.416
Observations (or Sum Wgts)	500

### 3. Conditions

#### Independence (Met)

- Simple Random Sample
- Data from an Airbnb would not affect the data from another Airbnb

#### Normality (Violated)

- A curved pattern in QQ plot
- Points out of the bounds
- Not much of a concern because of a large  $n = 500$

#### Conditions

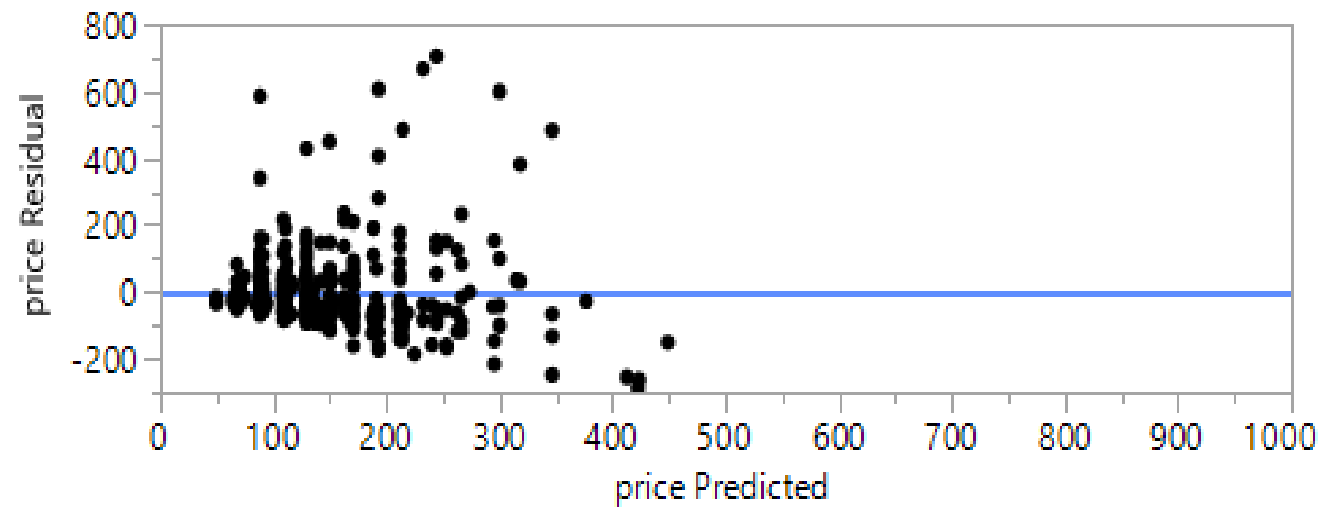
#### Constant Variance (Violated)

- Fan shape in residual plot

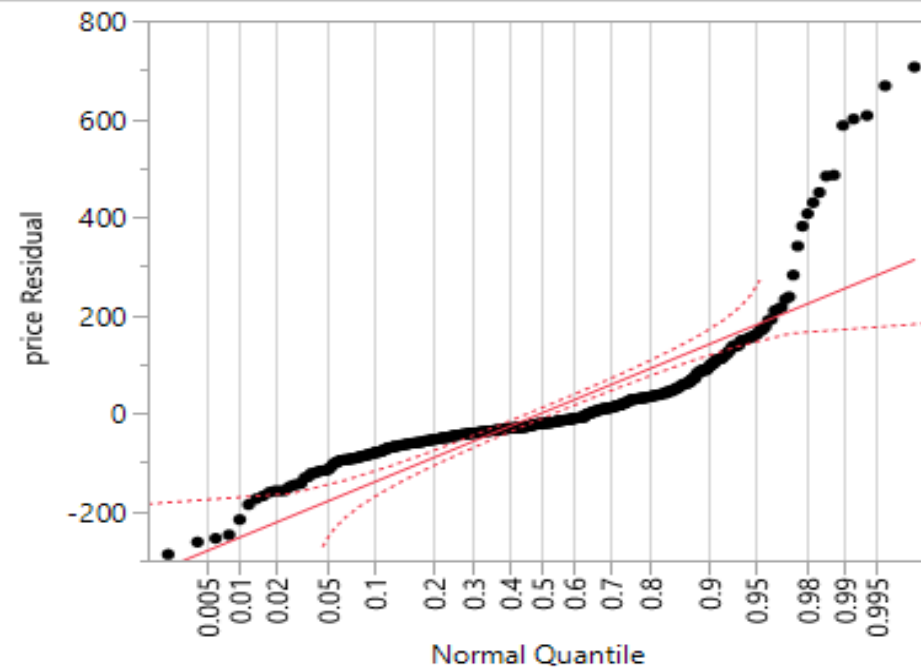
#### Linearity (Reasonably Met)

- More residual points below 0 with multiple points with high values even out
- Mean residual is about 0

**Residual by Predicted Plot**



**Residual Normal Quantile Plot**



## 4. Outliers

### (a) Standardized Residuals

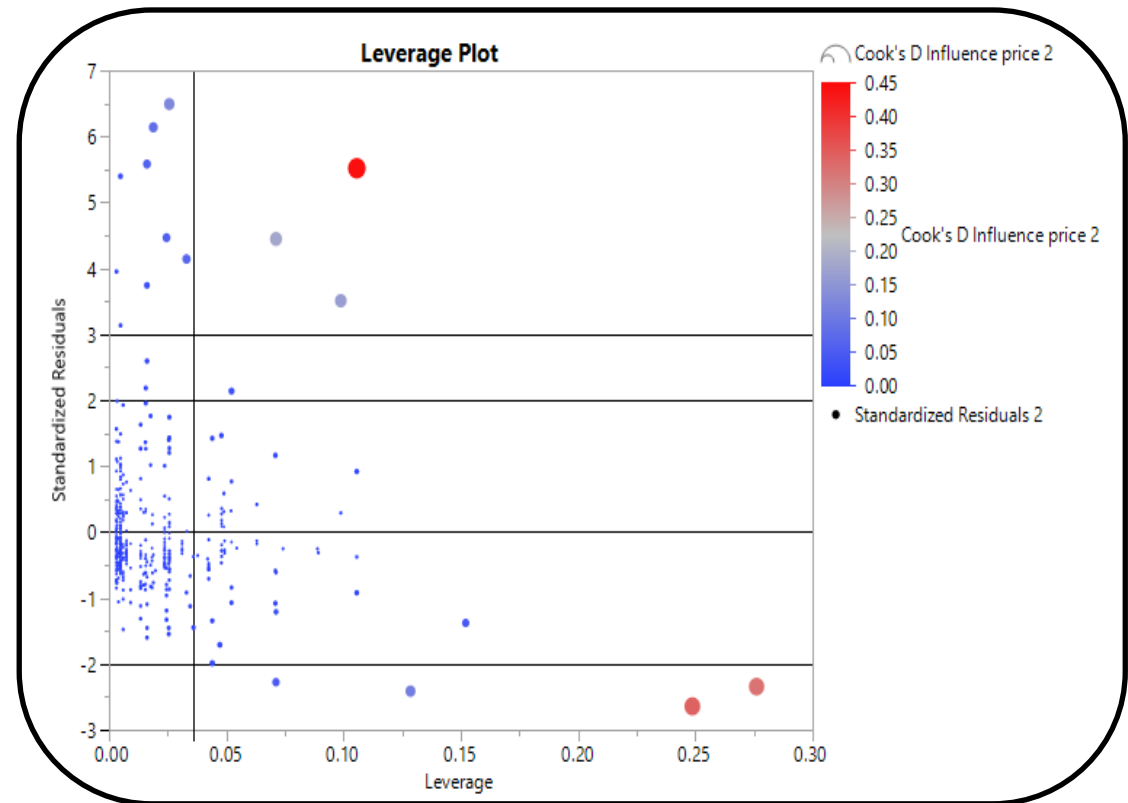
- 12 extremely unusual points with standardized residual greater than 3
- About 8 somewhat unusual points with the standardized residuals of between 2 and 3 or -2 and -3

### (b) Cook's D

- No concern (highest D  $0.45 < 1$ )

### (c) Leverage

- Cutoff Value:  $\frac{2(8+1)}{500} = 0.036$
- Multiple high leverage points





## 5. Hypothesis Tests

### (a) First Order Term

$$H_0: \beta_2 = 0, H_a: \beta_2 \neq 0$$

**Test Statistic:**  $t = -0.74$

**P-value:**  $p = 0.4625$

**Decision:** Fail to reject the null hypothesis since  $p = 0.4625 > \alpha = 0.05$

**Conclusion:** We do not have evidence that the population price of Airbnb is different for Not Real bed and Real bed, controlling for Property type and Accommodates constant.

Parameter Estimates				
Term	Estimate	Std Error	t Ratio	Prob> t
Intercept	46.923387	11.49021	4.08	<.0001*
accommodates	20.607285	2.599254	7.93	<.0001*
Not Real Bed	-18.07524	24.58404	-0.74	0.4625
Condominium	11.552727	27.81248	0.42	0.6780
House	-7.423192	24.77348	-0.30	0.7646
Other	-10.07689	35.25571	-0.29	0.7751
Condominium*accommodates	5.3834992	5.881759	0.92	0.3605
House*accommodates	4.9744929	4.452753	1.12	0.2645
Other*accommodates	16.987271	8.482824	2.00	0.0458*

## (b) Interaction Terms

$$H_0: \beta_6 = \beta_7 = \beta_8 = 0$$

$H_a$ : At least one  $\beta_i \neq 0$  ( $i = 6, 7, 8$ )

**Test Statistic:**  $F = 1.614$ , **P-value:**  $p = 0.185$

Sum of Squares	57250.290805
Numerator DF	3
F Ratio	1.6140935361
Prob > F	0.1851903783

**Decision:** Fail to reject the null hypothesis since  $p = 0.185 > \alpha = 0.05$

**Conclusion:** We don't have evidence that the interaction between property types and accommodates when predicting the price of Airbnb is significant, holding Bed type constant.

## Conclusion

We don't have evidence that the relationship between AirBnB price and accommodates is significantly different for different property types based on the hypothesis test of interaction term. However, looking at the scatterplot and the p-value of 0.185, which is fairly small, we can conclude there is a slight interaction between accommodates and property type.