

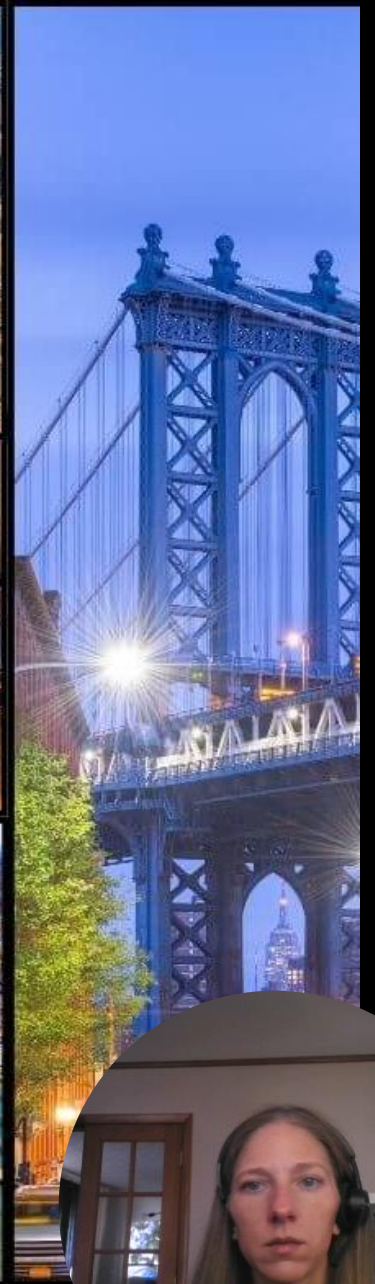
---

# NYC AIRBNB DATA ANALYSIS

LISA SIEFKER

PROJECT 1  
AU24 GTDA 5401

---



---

# DATA SUMMARY

**DATASET:** Airbnb in NYC\*

**DIMENSIONS:** 20,758 rows x 22 variables

**KEY VARIABLES:**

- Neighborhood Group (5 unique groups)
- Neighborhood (219 unique neighborhoods)
- Type of Listing (e.g., private/shared room)
- Price per night
- Minimum Nights (minimum nights required for a reservation)
- Number of Reviews (total number of reviews)
- Availability\_365 (number of days listing is available each year)
- Rating (average total rating)
- Number of Bedrooms
- Number of Beds
- Number of Bathrooms

\* <https://www.kaggle.com/datasets/vrindakallu/new-york-da>





---

# RESEARCH QUESTION

Which Airbnb is the **best deal** (i.e., the rating is above average and lowest price) in each neighborhood? The analysis will only consider Airbnbs that meet my family's criteria for an upcoming trip:

- entire home/apt
  - at least 2 beds
  - at least 1 bathroom
  - minimum nights < 4 less than 4
- 







# METHODOLOGY

## 1. Clean data – tidy format:

- Convert to tibble
- Review data structure
- Convert data types as needed
- Remove variables not relevant to analysis
- Remove rows that do not meet criteria

## 2. Analyze filtered data:|

- Tables
- Histograms
- Boxplots
- Scatterplots
- Statistical summary data



---

## RESULTS & INTERPRETATION: REVIEW, CLEAN & FILTER

```
dim(data4)
```

```
## [1] 574  12
```

```
length(unique(data4$neighborhood))
```

```
## [1] 74
```

```
length(unique(data4$neighborhood_group))
```

```
## [1] 5
```

```
print(result_original)
```

##	Category	Count	Percentage
## 1	Bronx	949	4.6
## 2	Brooklyn	7719	37.2
## 3	Manhattan	8038	38.7
## 4	Queens	3761	18.1
## 5	Staten Island	291	1.4

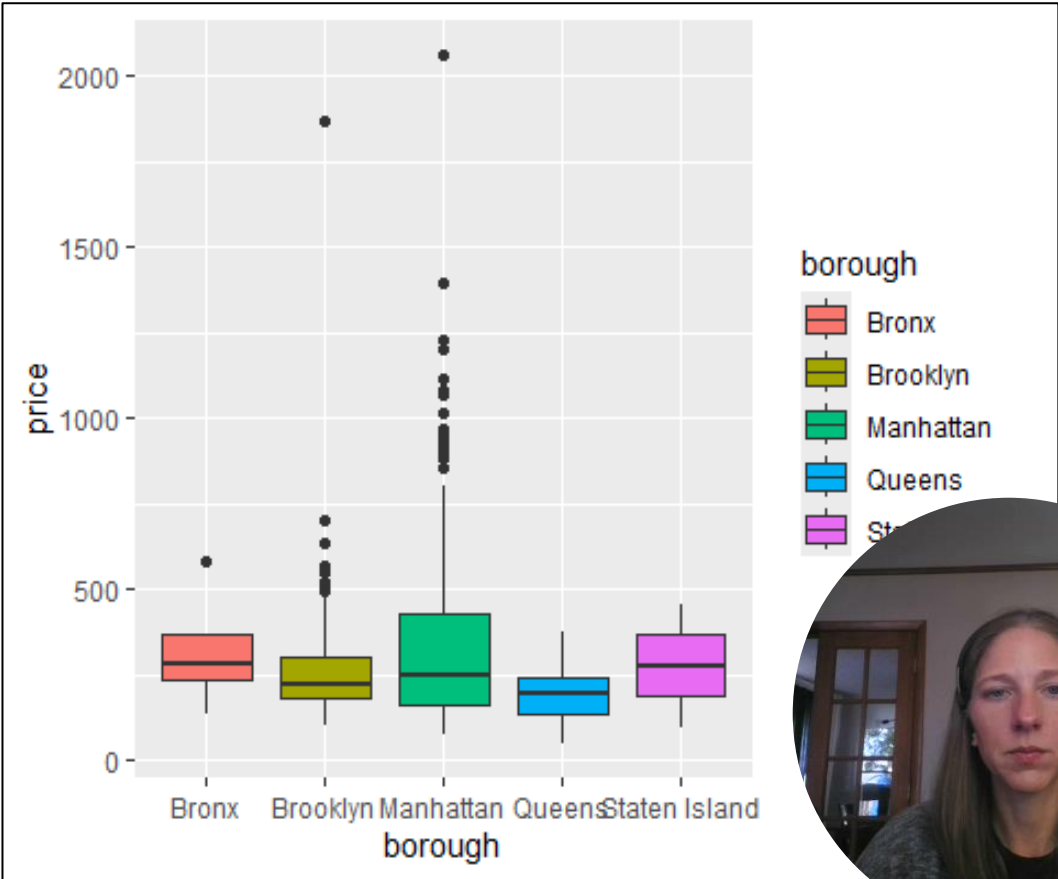
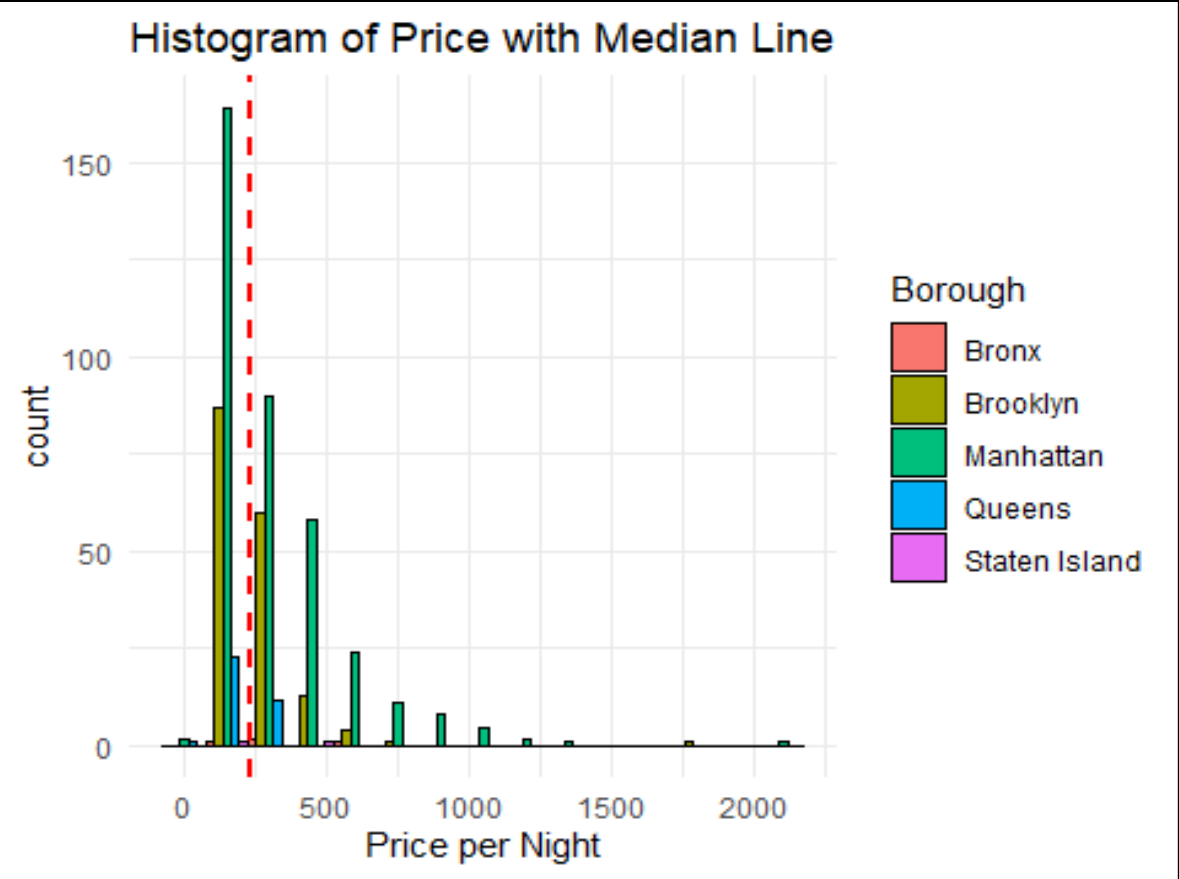
```
print(filtered_result)
```

##	Category	Count	Percentage
## 1	Bronx	4	0.7
## 2	Brooklyn	166	28
## 3	Manhattan	366	
## 4	Queens	36	
## 5	Staten Island	2	



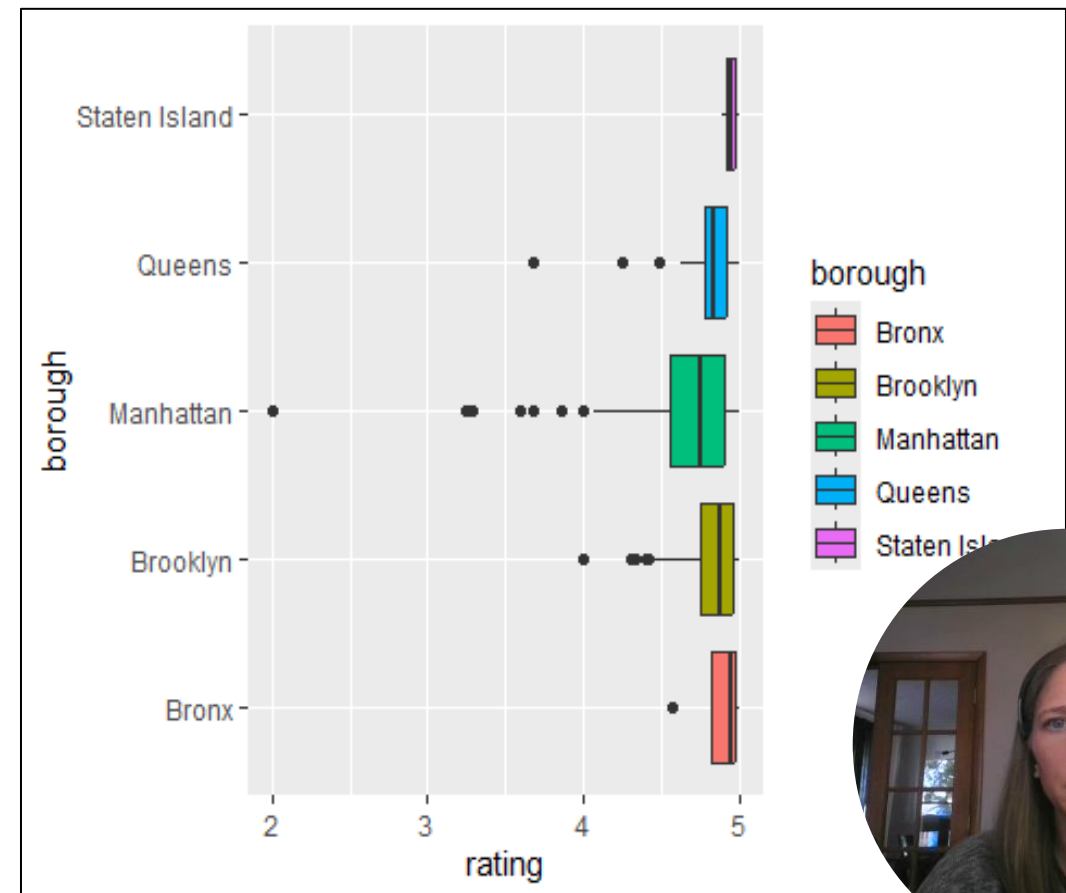
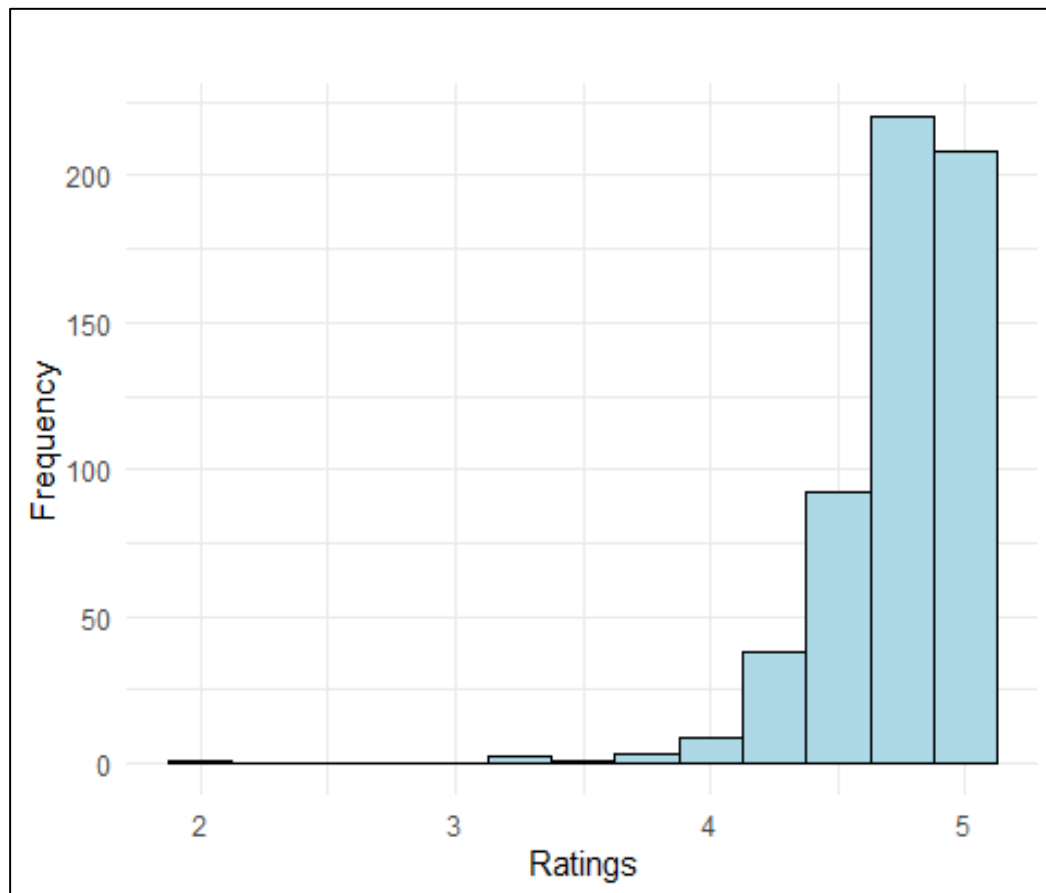
# RESULTS & INTERPRETATION: PRICE

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
49.0	164.2	235.5	299.2	356.8	2064.0

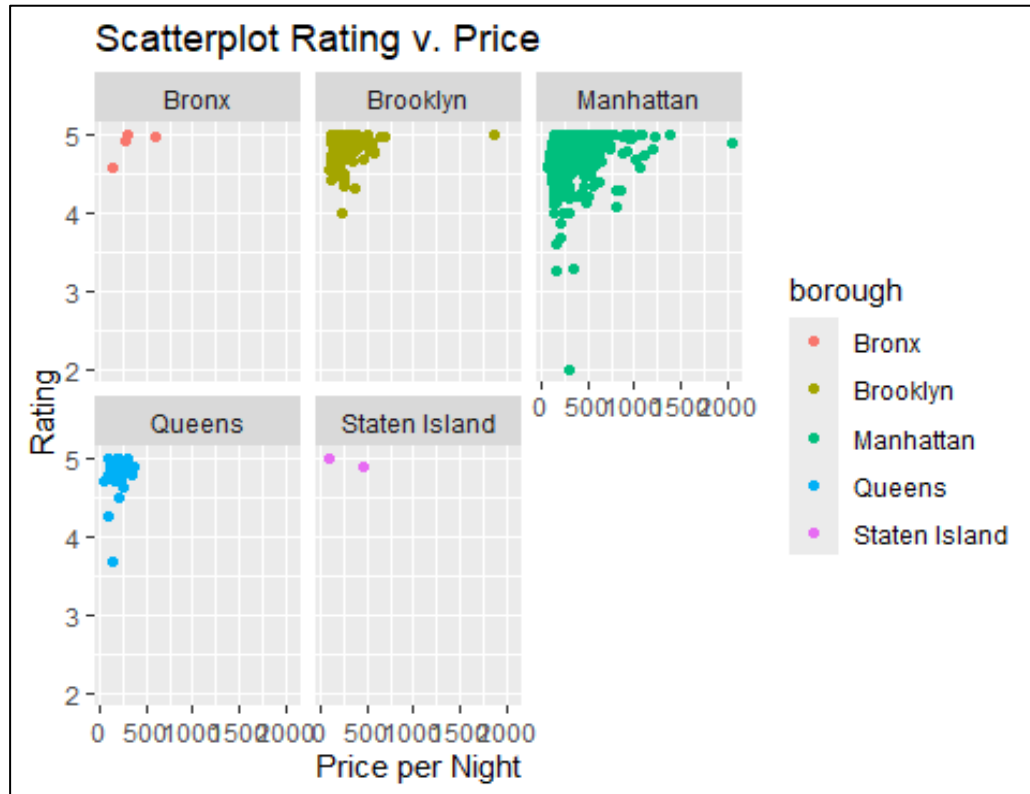


# RESULTS & INTERPRETATION: RATING

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
2.000	4.620	4.800	4.735	4.920	5.000



# RESULTS & INTERPRETATION: RATING & PRICE



*#Identify best deal for each borough*

```
## # A tibble: 4 × 6
```

```
##   borough      price rating bedrooms baths
##   <chr>      <dbl> <dbl>      <dbl> <dbl>
## 1 Brooklyn    107     5           1     1
## 2 Manhattan    97   4.76          1     1
## 3 Queens       87   4.78          1     1
## 4 Staten Island 95     5           3   1.5
```





---

# SUMMARY

- Airbnbs in Manhattan were more likely to meet my filter criteria than those in other boroughs
  - Identified Airbnbs that are the best deal (i.e., rating above average and lowest price) in each borough
  - Price did not have a relationship to rating
  - Airbnbs above and below the median price were available in all 5 boroughs
  - The most expensive Airbnbs were located in Manhattan, but otherwise price did not have a relationship to borough
- 

