



A large, semi-transparent dark red rectangular box covers the middle-left portion of the slide, containing the project title and subtitle.

# Project – 1:

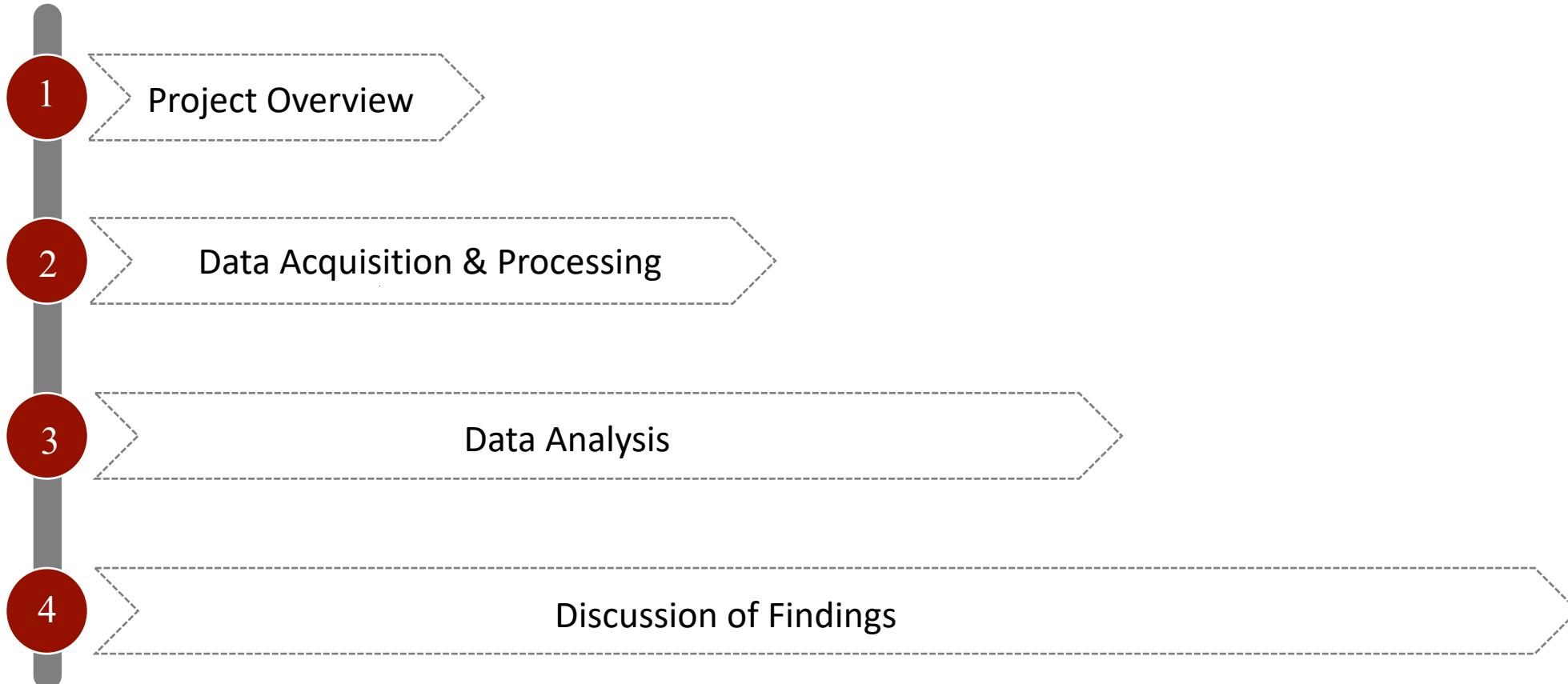
## Airbnb Rental Market Analysis

### 2017-2020

04/18/2020

Fabienne Zumbühl | Lizeth Orozco | Jairo Humberto Leon Cardoso | Stefan Blagojevic

# Agenda



# Executive Summary

## Research Question

What?

Do hosts that change their status to Superhosts charge more per night?

Why?

Superhosts have greater duties and could charge more

## Research Scope

On Location:  
New York City

On Time:  
March 2017 – March 2020

## Hypothesis

$H_1$  = Hosts increase their price when becoming a Superhost

$H_0$  = Hosts don't increase their price when becoming a Superhost

## Main Finding

Converting from Host to Superhost, the price does not increase

~~$H_1$  = Hosts increase their price when becoming a Superhost~~

## Finding

Hosts of **943 listings changed** their status from host to Superhost between 2017 and 2020

## Finding

The median Superhost charges the most in Manhattan and Brooklyn



# 1. Project Overview

```
mirror_mod.mirror_object  
operation == "MIRROR_X":  
    mirror_mod.use_x = True  
    mirror_mod.use_y = False  
    mirror_mod.use_z = False  
operation == "MIRROR_Y":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = True  
    mirror_mod.use_z = False  
operation == "MIRROR_Z":  
    mirror_mod.use_x = False  
    mirror_mod.use_y = False  
    mirror_mod.use_z = True
```

```
selection at the end -add  
    mirr_ob.select= 1  
    mirr_ob.select=1  
    context.scene.objects.active  
    "Selected" + str(modifier)  
    mirr_ob.select = 0  
    bpy.context.selected_obj  
    data.objects[one.name].sel
```

```
int("please select exactly  
--- OPERATOR CLASSES ---
```

```
types.Operator):  
    X mirror to the selected  
    object.mirror_mirror_x"  
    mirror_X"
```

# 1. Project Overview

## What was our Motivation?

- **Interest because of exposure** to Airbnb
- **Interest** in the Superhost status
- **New York** is the **biggest market** for Airbnb

## What were our main Questions?

- From 2017 to 2020, **did Airbnb Superhosts charge more** per night than a regular Airbnb hosts in the NYC ?
- **Which borough** in NYC would be **ideal to own** an Airbnb rental?
- What is the **price discrepancy** per night between **hosts & Superhosts**?



# 1. Project Overview

## Hypothesis

$H_1$  = *Hosts increase their price per night when becoming a Superhost*

$H_0$  = *Hosts don't increase their price per night when becoming a Superhost*

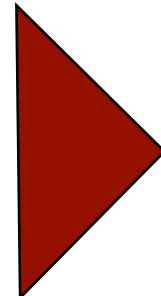
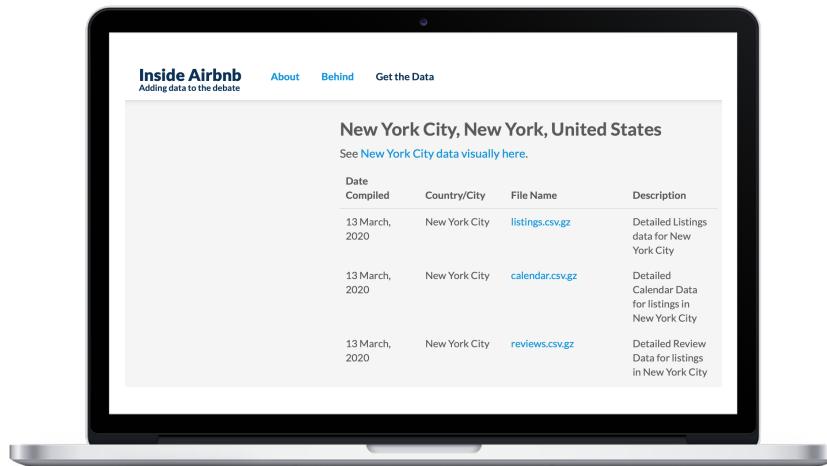
The background of the slide features a repeating pattern of binary digits (0s and 1s) in a light gray color, creating a digital or technical atmosphere.

## 2. Data Acquisition & Processing

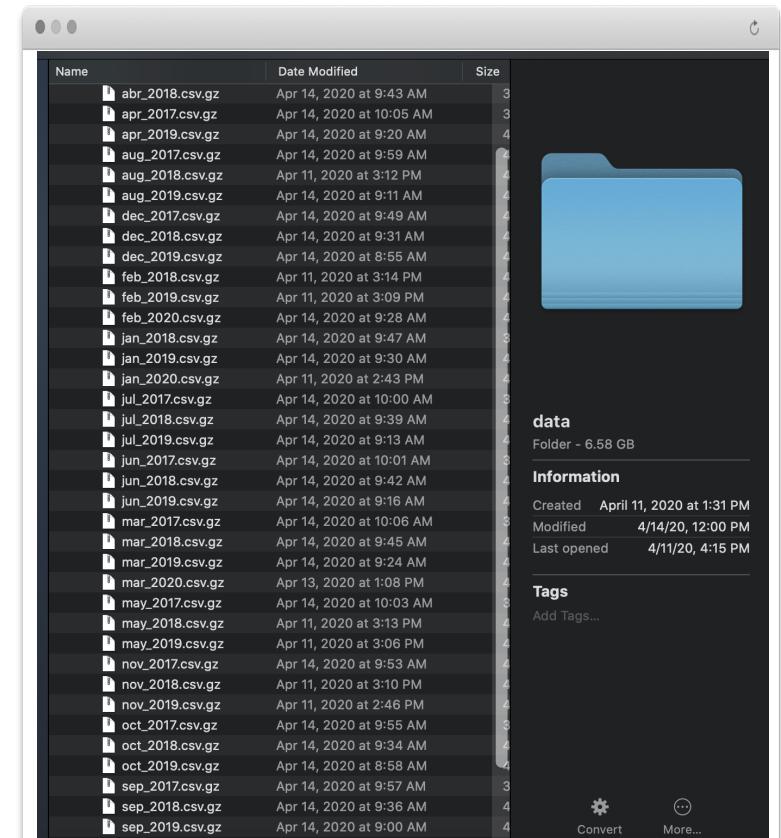
## 2. Data Acquisition & Processing

### Down/Uploading Files in the Time of the Coronavirus ... Duboce Triangle and AT&T, A Slow Story

Retrieval of 37 Zipped File Months from Inside Airbnb



**Inside Airbnb**  
(<http://insideairbnb.com/get-the-data.html>)



## 2. Data Acquisition & Processing

### Downsizing to ‘../Uploadable Files.csv’

Super Duper Work Flow!

```
#reduce original df with list of columns
mar_2020_red = original_df[col_list_work]

#filter host_listings_count != 0
mar_2020_red2 = mar_2020_red[mar_2020_red.host_listings_count != 0]
#filter host_identity_verified = t
mar_2020_red3 = mar_2020_red2[mar_2020_red2.host_identity_verified == 't']
#filter bed_type = real bed
mar_2020_red4 = mar_2020_red3[mar_2020_red3.bed_type == 'Real Bed']
#filter location = exact
mar_2020_red5 = mar_2020_red4[mar_2020_red4.is_location_exact == 't']
#filter bedrooms drop nan
mar_2020_red6= mar_2020_red5.dropna()
#replace bedrooms 0 with 1
mar_2020_red6.loc[(mar_2020_red6.bedrooms == 0 ), 'bedrooms']= 1
mar_2020_work = mar_2020_red6
```

Import them into Jupyter Notebook

Set Columns to Work to a Variable

Filtered and cleaned

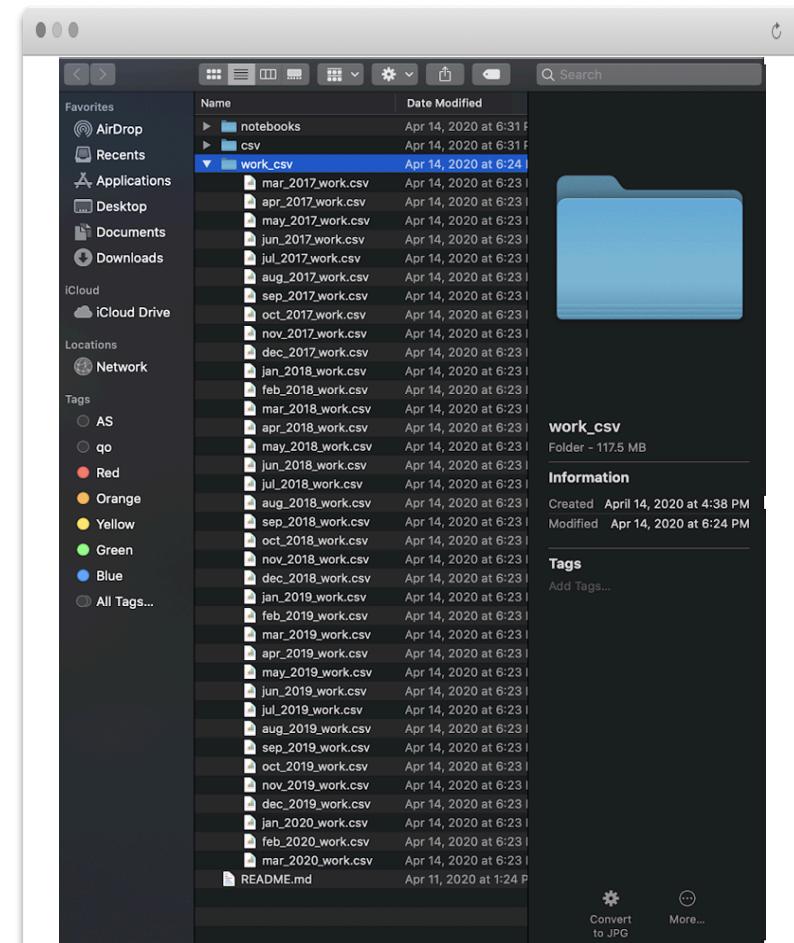
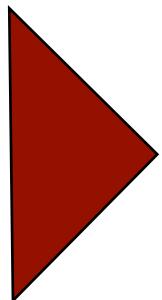
Output 37 files

## 2. Data Acquisition & Processing

Booya!!! All Up There in the GitHub Universe!

A Present  
to My Group Mates  
for Their ENTERTAINMENT!

I am done with typing! Siri,  
Activate Voice Recognition!!



## 2. Data Acquisition & Processing

### Narrowing down the Scope

- **Strategy changed due to time constraints** and reduced from monthly to yearly trending.
- Use **data for same month for the past 4 years**.
- The narrowing down **removes the variable of seasonal changes** in prices



**Creating  
a dataframe  
with 37 csv's**

**Creating  
a dataframe  
with 4 csv's**

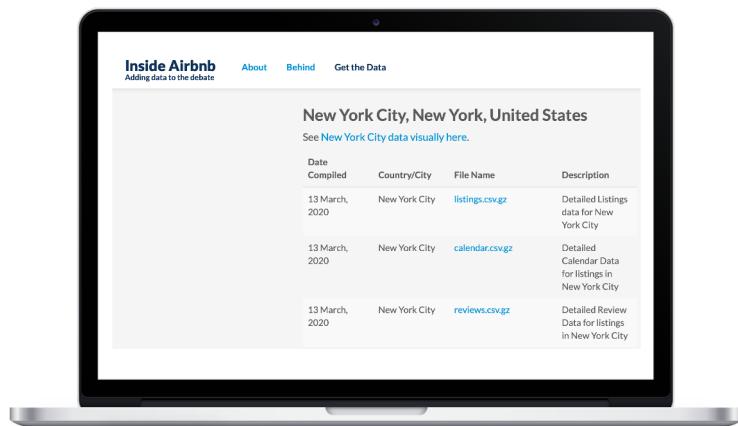
## 2. Data Acquisition & Processing

### Data Needed

- Data on **superhosts** in NYC
- Data on **prices per night per borough**
- Data on the **change of price host vs. superhost**
- Data on **object types**
- Data from **2017 – 2020**

### Data Source

- **Inside Airbnb**  
(<http://insideairbnb.com/get-the-data.html>)



### Share of market activity

Changes in the activity of the active and passive market is uncertain. Established positive trends in various market segments.



### 3. Data Analysis



### 3. Data Analysis

#### Identifying Inflection Points – The change from Host to Superhost

How **do we identify** at which point a host became a **Superhost**?

- Using the Dataframe combining **4 years of listings**, choose **2 starting points**
- How many listings in NYC were **not Superhosts in March of 2017 AND...**
- How many of those listings **became Superhosts in March 2020?**

```
change_superhost = merged_data3.loc[(merged_data3["2017"] == "f")  
& (merged_data3["2020"] == "t")]
```

### 3. Data Analysis

#### Our starting Dataframe for our data analysis

We start to see the status and price changes over the years. Progress!!!

943 rows x 17 columns

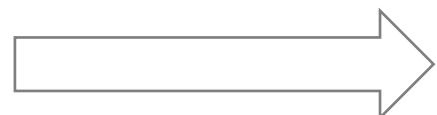
	d	2017	Host_id_17	Lat_17	Lon_17	Prop_17	Type_17	Price_17	2018	Price_18	2019	Price_19	2020	Price_20	Price_Diff_17_18	Price_Diff_18_19	Price_Diff_19_20
2	446208	f	Bronx	40.868144	-73.858745	House	Private room	48.0	t	40.0	t	44.0	t	45.0	-8.0	4.0	1.0
6	340039	f	Bronx	40.866770	-73.859378	House	Private room	48.0	t	43.0	t	45.0	t	45.0	-5.0	2.0	0.0
8	342975	f	Bronx	40.866889	-73.857756	House	Private room	45.0	t	43.0	t	43.0	t	45.0	-2.0	0.0	2.0
9	1478809	f	Queens	40.780306	-73.910030	Apartment	Entire home/apt	85.0	t	95.0	t	95.0	t	95.0	10.0	0.0	0.0
11	157876	f	Queens	40.774350	-73.912175	Apartment	Private room	59.0	t	57.0	f	51.0	t	51.0	-8.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
7579	483216	f	Brooklyn	40.718949	-73.956041	Apartment	Private room	65.0	f	69.0	f	73.0	t	85.0	0.0	8.0	12.0
7584	618546	f	Brooklyn	40.711783	-73.956843	Apartment	Private room	110.0	f	115.0	f	90.0	t	90.0	5.0	-25.0	0.0
7593	1024606	f	Brooklyn	40.712601	-73.946683	Apartment	Entire home/apt	255.0	t	255.0	t	255.0	t	255.0	0.0	0.0	0.0
7601	369575	f	Brooklyn	40.716209	-73.956416	Apartment	Entire home/apt	325.0	t	325.0	f	375.0	t	375.0	0.0	50.0	0.0
7606	938022	f	Brooklyn	40.707799	-73.948273	Apartment	Private room	98.0	t	98.0	t	99.0	t	99.0	0.0	1.0	0.0

The hosts of 943 listings have upgraded their status from a regular host to a Superhost from 2017-2020

### 3. Data Analysis

#### Change of Host to Superhost Listings throughout the four years

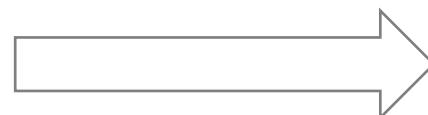
	2017	2018	% total '18	2019	% total '19	2020
Superhosts	0	468	48.7%	596	63.2%	943
Hosts	943	475	49.4%	347	36.8%	0



Hosts of **468** listings moved up as  
Superhosts



Hosts of **380** listings stayed Superhost  
Superhosts of **88** listings resigned as Superhosts

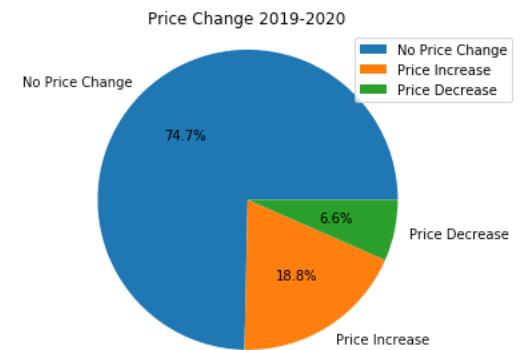
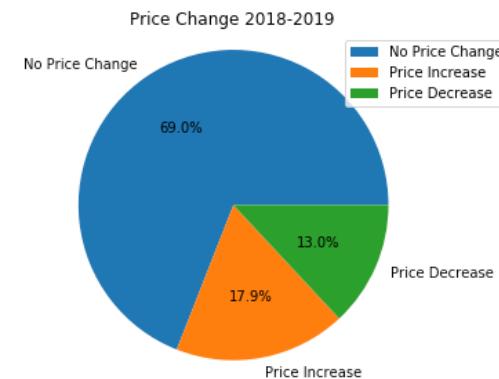
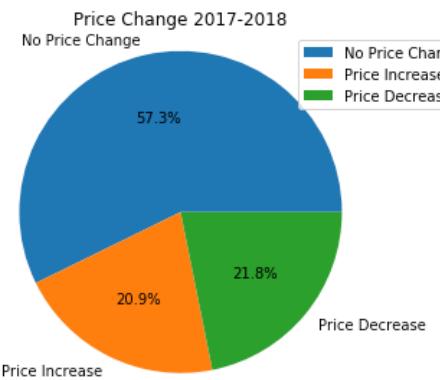


Hosts of **216** listings moved up as Superhosts



### 3. Data Analysis

#### General Price Changes in the Market



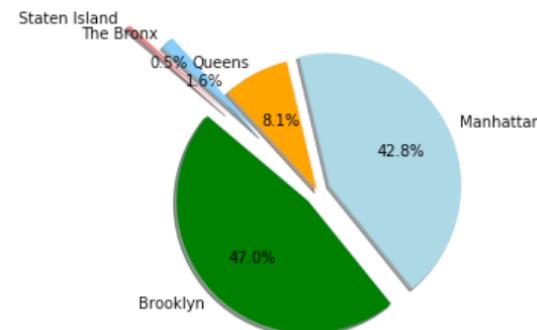
	No Price Change	Price Increase	Price Decrease
	0	540	197
Price change 2018 to 2019	No Price Change	Price Increase	Price Decrease
	0	651	169
Price change 2019 to 2020	No Price Change	Price Increase	Price Decrease
	0	704	177
			62

### 3. Data Analysis

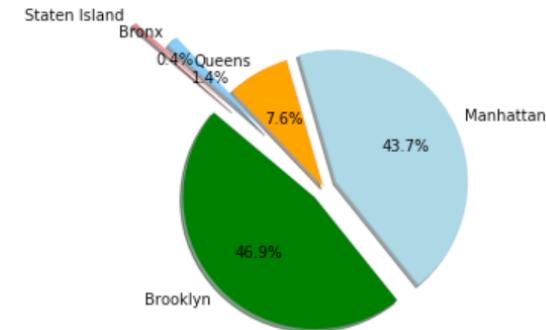
#### Market Share of Listings by Borough (2017-2020)

Host to Superhost

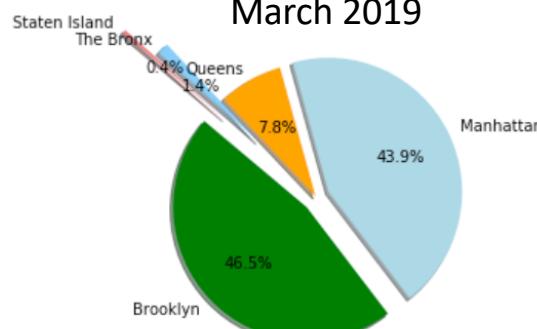
March 2017



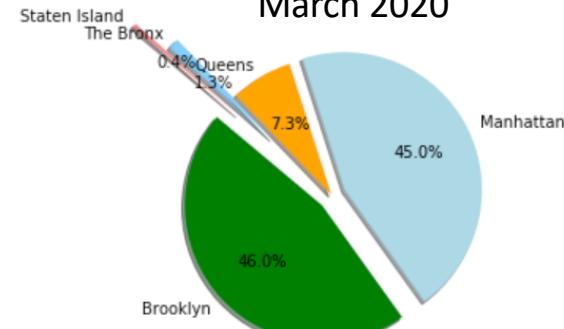
March 2018



March 2019

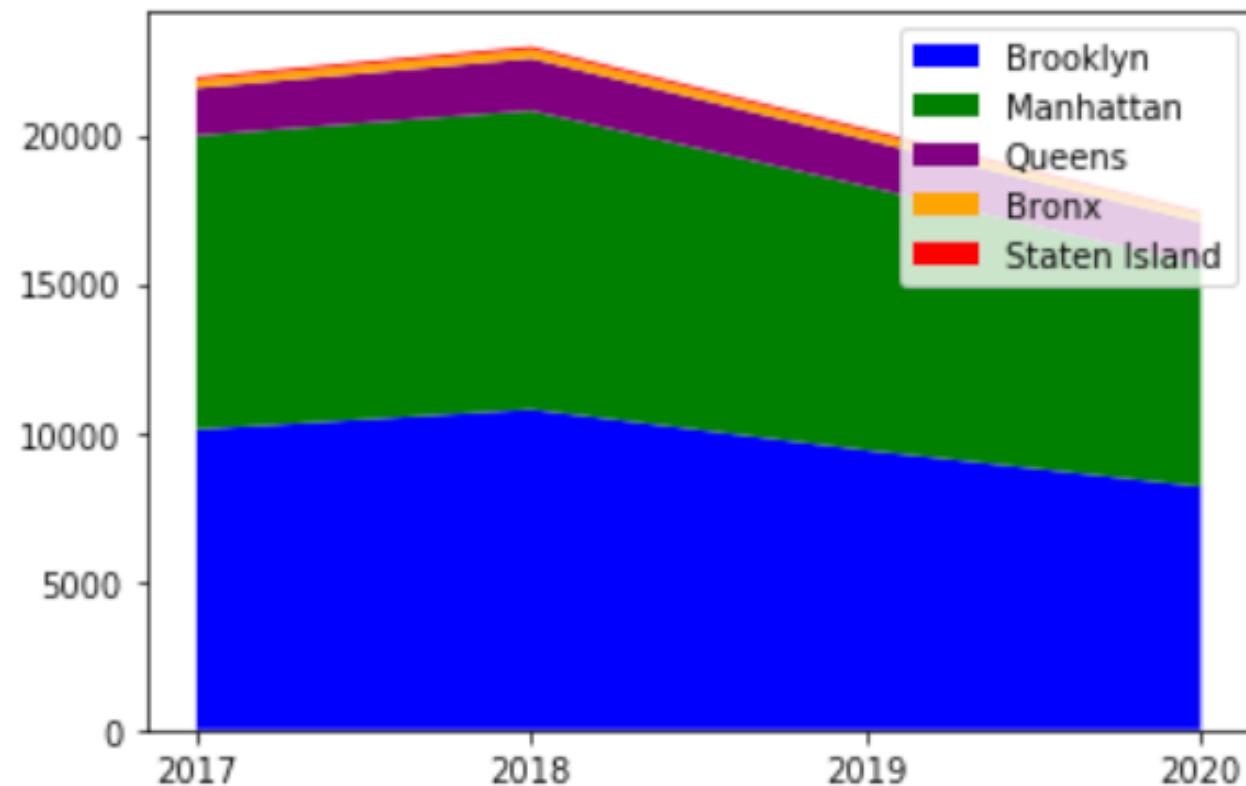


March 2020



### 3. Data Analysis

#### Listings by Borough (2017-2020)



### 3. Data Analysis

#### Paired sample t-test

$H_0$  = Hosts don't increase their listing price per night when becoming a superhost

$H_1$  = Hosts do increase their listing price per night when becoming a superhost

```
In [704]: # t-test comparing price change as a proxy for hosts vs superhosts
%matplotlib inline
from matplotlib import pyplot as plt
import numpy as np
import scipy.stats as stats

superhosts=dropped_col['Price_20']
hosts=dropped_col['Price_17']

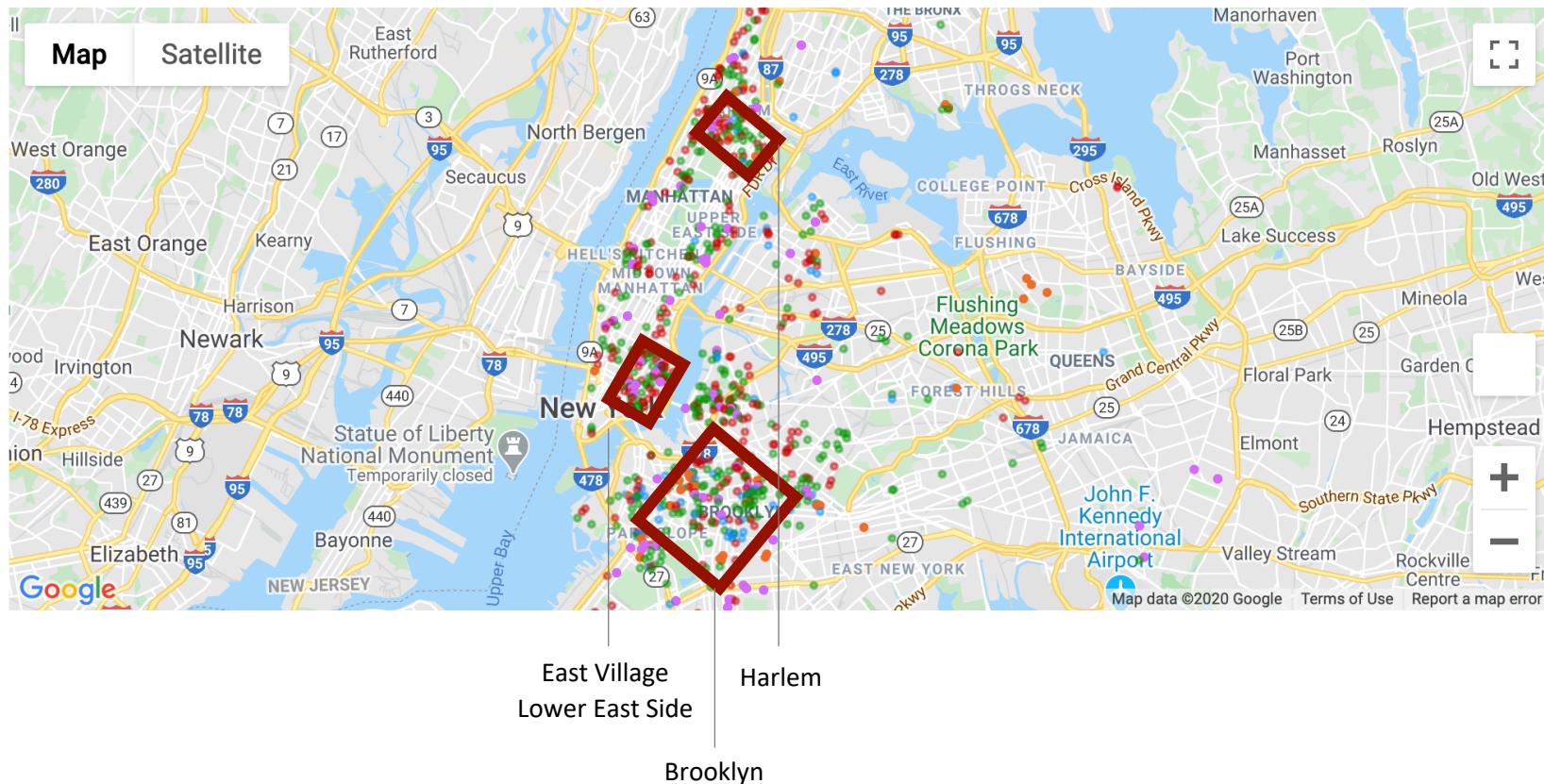
stats.ttest_rel(hosts, superhosts)
```

```
Out[704]: Ttest_relResult(statistic=-2.1513613547545436, pvalue=0.03170137822992948)
```

### 3. Data Analysis

Where exactly do the newly upgraded Superhosts live?

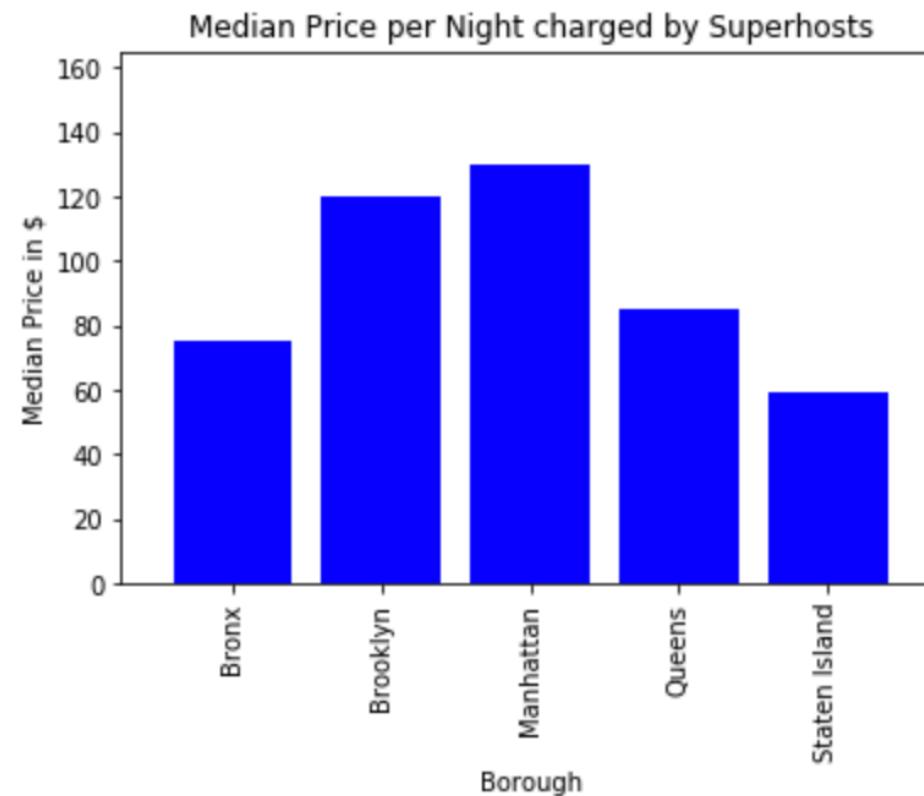
Consolidated View (2017-2020)



- Entire Apartment
- Apartment with Private Room
- Entire House
- House with Private Room
- Drop-outs 2018/2019 and new Superhosts 2019/2020
- Guesthouses

### 3. Data Analysis

The median Superhost charges the most in Manhattan and Brooklyn



### 3. Data Analysis

Let's get back to our main question: Do the Superhosts charge more when attaining their status?

468 new Superhosts in 2018

	<b>id</b>	<b>Hood_17</b>	<b>Lat_17</b>	<b>Lon_17</b>	<b>Prop_17</b>	<b>Type_17</b>	<b>2017</b>	<b>Price_17</b>	<b>2018</b>	<b>Price_18</b>	<b>Price_19</b>	<b>Price_Diff_17_18</b>
2	4462008	Bronx	40.868144	-73.858745	House	Private room	f	48.0	t	40.0		-8
6	3400359	Bronx	40.866770	-73.859378	House	Private room	f	48.0	t	43.0		-5
8	3429765	Bronx	40.866889	-73.857756	House	Private room	f	45.0	t	43.0		-2
9	14788009	Queens	40.780306	-73.910030	Apartment	Entire home/apt	f	85.0	t	95.0		10
11	1578776	Queens	40.774350	-73.912175	Apartment	Private room	f	59.0	t	51.0		-8
...	...	...	...	...	...	...	...	...	...	...	...	...
7537	1681023	Brooklyn	40.719421	-73.944668	Apartment	Private room	f	78.0	t	65.0		-13
7576	12869662	Brooklyn	40.717009	-73.956847	Apartment	Entire home/apt	f	140.0	t	250.0		110
7593	10246056	Brooklyn	40.712601	-73.946683	Apartment	Entire home/apt	f	255.0	t	255.0		0
7601	3695725	Brooklyn	40.716206	-73.956416	Apartment	Entire home/apt	f	325.0	t	325.0		0
7606	938022	Brooklyn	40.707799	-73.948273	Apartment	Private room	f	98.0	t	98.0		0

468 rows x 11 columns

216 new Superhosts in 2019

	<b>id</b>	<b>Hood_17</b>	<b>Lat_17</b>	<b>Lon_17</b>	<b>Prop_17</b>	<b>Type_17</b>	<b>2018</b>	<b>Price_18</b>	<b>2019</b>	<b>Price_19</b>	<b>Price_Diff_18_19</b>	
37	1981410	Queens	40.779065	-73.912648	Apartment	Entire home/apt	f	135.0	t	135.0		0
42	16545876	Queens	40.775599	-73.908006	Apartment	Private room	f	70.0	t	45.0		-25
55	11496210	Queens	40.680839	-73.847654	Apartment	Entire home/apt	f	80.0	t	87.0		7
56	11474664	Queens	40.681650	-73.845577	Apartment	Entire home/apt	f	105.0	t	120.0		15
144	11010008	Queens	40.761009	-73.915729	Apartment	Private room	f	70.0	t	70.0		0
...	...	...	...	...	...	...	...	...	...	...	...	
7413	3090922	Brooklyn	40.715202	-73.954450	Apartment	Private room	f	75.0	t	75.0		0
7483	3414885	Brooklyn	40.711220	-73.948928	Apartment	Entire home/apt	f	210.0	t	240.0		30
7523	6192999	Brooklyn	40.707762	-73.958864	Apartment	Entire home/apt	f	223.0	t	133.0		-90
7542	17331192	Brooklyn	40.713718	-73.963260	Apartment	Entire home/apt	f	190.0	t	179.0		-11
7562	158176	Brooklyn	40.715339	-73.959140	Apartment	Entire home/apt	f	125.0	t	120.0		-5

216 rows x 11 columns

347 new Superhosts in 2020

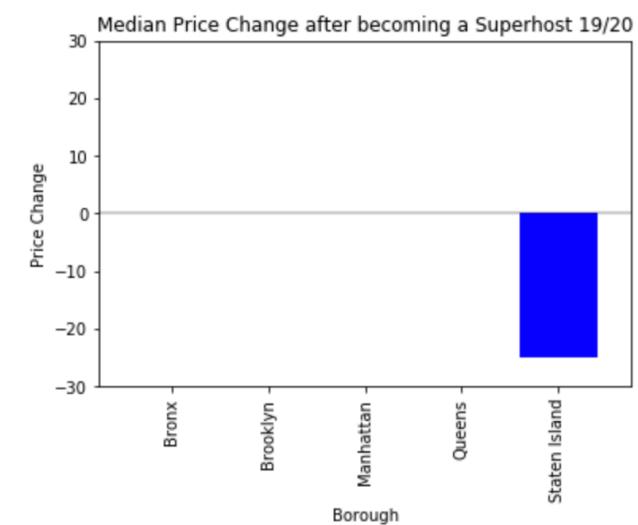
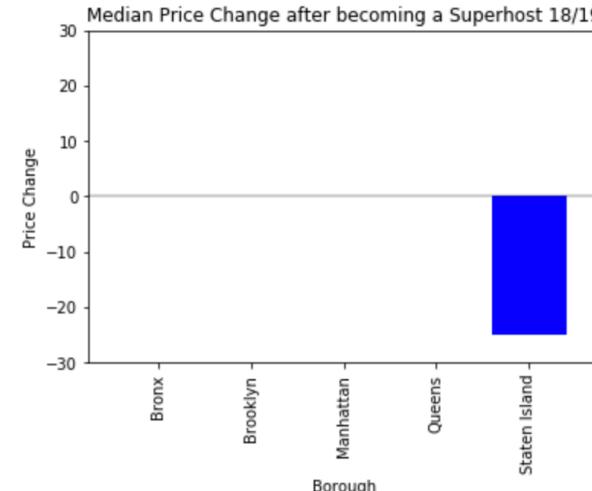
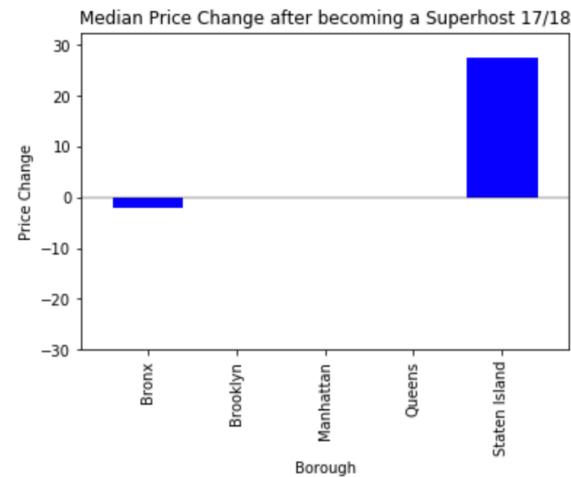
	<b>id</b>	<b>Hood_17</b>	<b>Lat_17</b>	<b>Lon_17</b>	<b>Prop_17</b>	<b>Type_17</b>	<b>2019</b>	<b>Price_19</b>	<b>2020</b>	<b>Price_20</b>	<b>Price_Diff_19_20</b>	
11	1578776	Queens	40.774350	-73.912175	Apartment	Private room	f	51.0	t	51.0		0
39	12851100	Queens	40.771414	-73.914154	Apartment	Entire home/apt	f	80.0	t	80.0		0
45	9859181	Queens	40.782744	-73.913179	House	Entire home/apt	f	88.0	t	193.0		105
74	12205604	Queens	40.587757	-73.793083	Guesthouse	Private room	f	95.0	t	95.0		0
93	74680	Queens	40.776350	-73.934261	Apartment	Entire home/apt	f	115.0	t	115.0		0
...	...	...	...	...	...	...	...	...	...	...	...	
7557	13399348	Brooklyn	40.715697	-73.959828	Apartment	Entire home/apt	f	125.0	t	165.0		40
7560	15232776	Brooklyn	40.719579	-73.961232	Apartment	Entire home/apt	f	250.0	t	350.0		100
7579	4832126	Brooklyn	40.718949	-73.956041	Apartment	Private room	f	73.0	t	85.0		12
7584	6185426	Brooklyn	40.711783	-73.956843	Apartment	Private room	f	90.0	t	90.0		0
7601	3695725	Brooklyn	40.716209	-73.956416	Apartment	Entire home/apt	f	375.0	t	375.0		0

347 rows x 11 columns



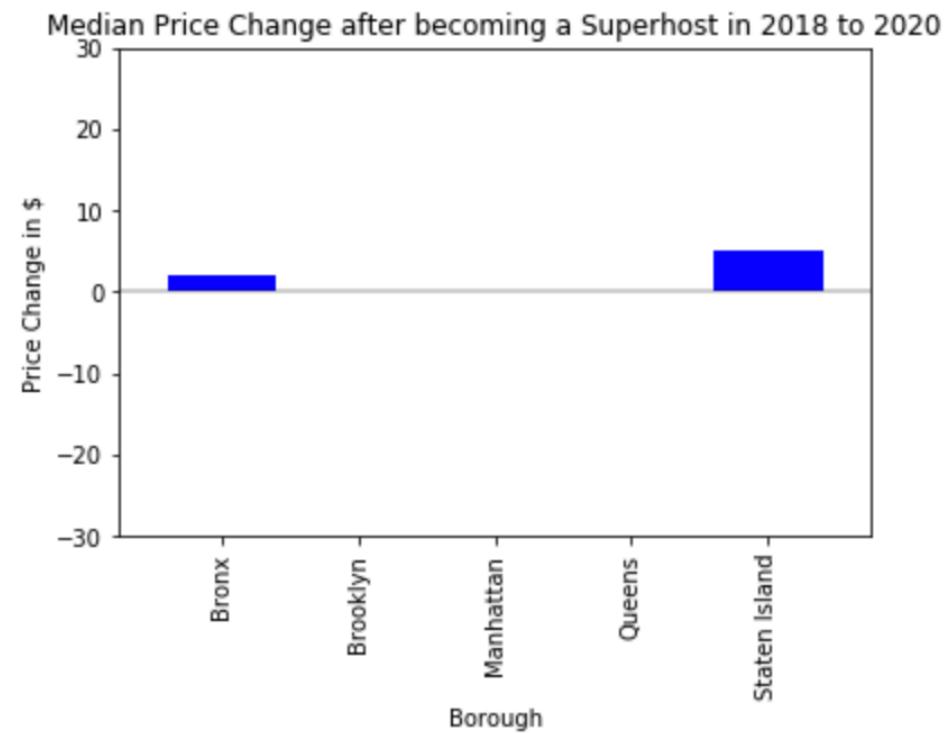
### 3. Data Analysis

The Median Price change per Borough is (almost) zero! (exception Staten Island)



### 3. Data Analysis

The same hold True for people who have been Superhosts from 3 consecutive years





## 4. Discussion of Findings

## 4. Discussion of Findings

- Based on our research approach, our finding is that **the Median Superhost Rental is not increasing its price** from year n (regular Host) to n+1 (Superhost).
- Based on the findings, the **incentives of becoming a Superhost is charging more money per customer**, but being on top of the search list, thus **attracting more customers and make more money like that** (Assumption)
- The **competitiveness in the NYC rental market** is so high that Superhost might lose customers by increasing their prices **due to price inelasticity of the demand** (Assumption)
- Limitations of Findings:
  - Dataset was **limited to 943 listings**
  - Findings are **not universally applicable for the NYC market**, but only for the years studied
  - Findings are thus also **not universally applicable for the whole Airbnb market**



## 4. Discussion of Findings

We can reject our Hypothesis

$H_1 = \text{Hosts increase their price per night when becoming a Superhost}$



$H_0 = \text{Hosts don't increase their price per night when becoming a Superhost}$

## 4. Discussion of Findings

### Lessons Learned

- **Don't plan on doing things with the data without having the data**
- **Acquiring and cleaning up the data is 50% of the job**
- **It is hard to decide which visualization to choose from** (out of the 20 charts created)
- **A research question has a lot of facets and it is important to disassemble the research question**
- **It's easy to lose track on the research question we wanted to answer**





Thank you!!!  
Questions?