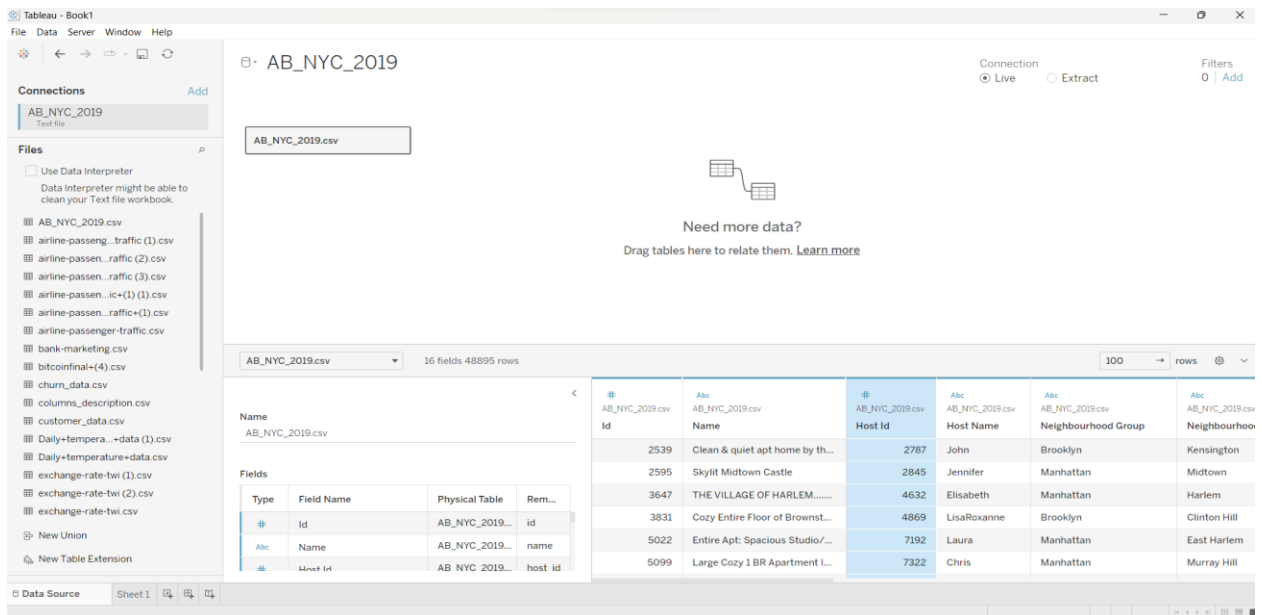


# Methodology Document for Airbnb, NY Dataset

The data cleaning, transformation, and visualization were done on the Tableau tool. All the necessary steps on the data have been performed before creating the PPT. Below are the methodology/steps followed:

## **Data Cleaning:**

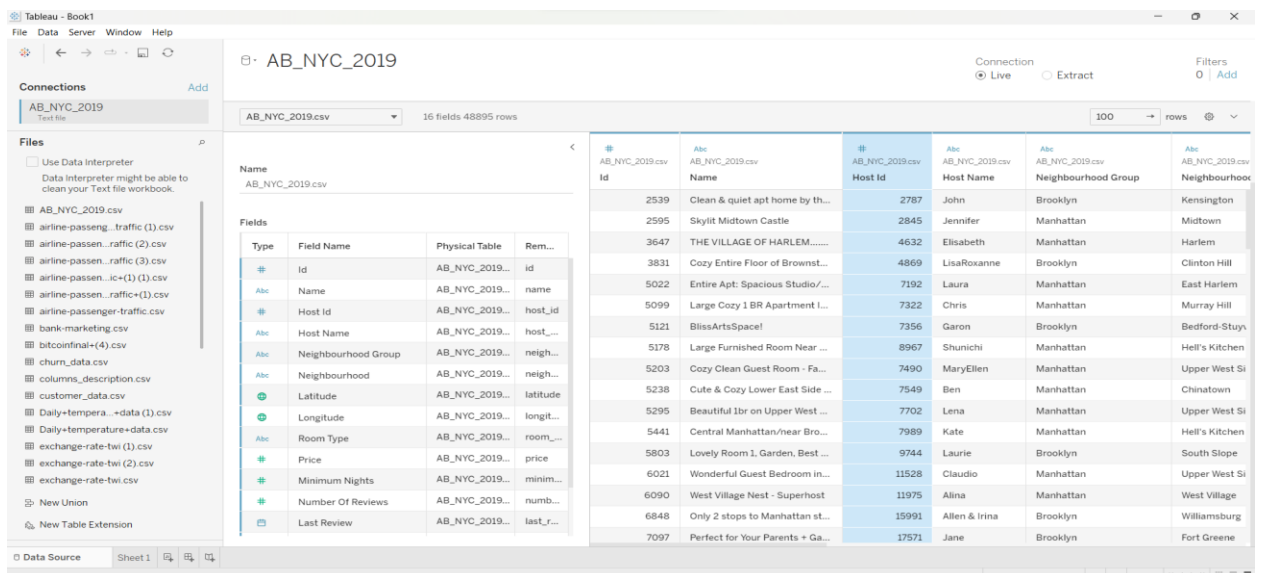
### 1. Entire Airbnb data was loaded in Tableau



The screenshot shows the Tableau interface with the 'AB\_NYC\_2019' dataset loaded. The 'Connections' pane on the left lists various data sources, including 'AB\_NYC\_2019.csv'. The 'Files' pane shows a list of files, including 'AB\_NYC\_2019.csv'. The central view displays the dataset name 'AB\_NYC\_2019' and a preview of the data. The preview table shows columns: Id, Name, Host Id, Host Name, Neighbourhood Group, and Neighbourhood. The data preview table is as follows:

Id	Name	Host Id	Host Name	Neighbourhood Group	Neighbourhood
2539	Clean & quiet apt home by th...	2787	John	Brooklyn	Kensington
2595	Skyliit Midtown Castle	2845	Jennifer	Manhattan	Midtown
3647	THE VILLAGE OF HARLEM.....	4632	Elisabeth	Manhattan	Harlem
3831	Cozy Entire Floor of Brownst...	4869	LisaRoxanne	Brooklyn	Clinton Hill
5022	Entire Apt: Spacious Studio/...	7192	Laura	Manhattan	East Harlem
5099	Large Cozy 1 BR Apartment L...	7322	Chris	Manhattan	Murray Hill

### 2. Checking Datatypes of each column:



The screenshot shows the Tableau interface with the 'AB\_NYC\_2019' dataset loaded. The 'Connections' pane on the left lists various data sources, including 'AB\_NYC\_2019.csv'. The 'Files' pane shows a list of files, including 'AB\_NYC\_2019.csv'. The central view displays the dataset name 'AB\_NYC\_2019' and a preview of the data. The preview table shows columns: Id, Name, Host Id, Host Name, Neighbourhood Group, and Neighbourhood. The data preview table is as follows:

Id	Name	Host Id	Host Name	Neighbourhood Group	Neighbourhood
2539	Clean & quiet apt home by th...	2787	John	Brooklyn	Kensington
2595	Skyliit Midtown Castle	2845	Jennifer	Manhattan	Midtown
3647	THE VILLAGE OF HARLEM.....	4632	Elisabeth	Manhattan	Harlem
3831	Cozy Entire Floor of Brownst...	4869	LisaRoxanne	Brooklyn	Clinton Hill
5022	Entire Apt: Spacious Studio/...	7192	Laura	Manhattan	East Harlem
5099	Large Cozy 1 BR Apartment L...	7322	Chris	Manhattan	Murray Hill

### 3. Remove Unwanted column:

Tableau - Book1

File Data Server Window Help

Connections: AB\_NYC\_2019 (Text file)

Files: Use Data Interpreter (Data Interpreter might be able to clean your Text file workbook.)

AB\_NYC\_2019.csv

Need more data? Drag tables here to relate them. [Learn more](#)

Connection: Live Extract Filters: 0 Add

AB\_NYC\_2019.csv 16 fields 48895 rows

Name: AB\_NYC\_2019.csv

Fields:

Type	Field Name	Physical Table	Rem...
#	Id	AB_NYC_2019...	id
Abc	Name	AB_NYC_2019...	name
#	Host Id	AB_NYC_2019...	host id

Describe...

AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv
Id	Name	Host Id	Host Name	Neighbourhood Group	Neighbourhood
2539	Clean & quiet apt home by th...	2787	John	Brooklyn	Kensington
2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown
3647	THE VILLAGE OF HARLEM.....	4632	Elisabeth	Manhattan	Harlem
3831	Cozy Entire Floor of Brownst...	4869	LisaRoxanne	Brooklyn	Clinton Hill
5022	Entire Apt: Spacious Studio/...	7192	Laura	Manhattan	East Harlem
5099	Large Cozy 1BR Apartment L...	7322	Chris	Manhattan	Murray Hill

Data Source: Sheet 1

### 4. Add filter condition many columns to remove the null values

Tableau - Book1

File Data Server Window Help

Connections: AB\_NYC\_2019 (Text file)

Files: Use Data Interpreter (Data Interpreter might be able to clean your Text file workbook.)

AB\_NYC\_2019.csv

Need more data? Drag tables here to relate them. [Learn more](#)

Connection: Live Extract Filters: 3 Edit

AB\_NYC\_2019.csv 16 fields 31268 rows

Name: AB\_NYC\_2019.csv

Fields:

Type	Field Name	Physical Table	Rem...
#	Id	AB_NYC_2019...	id
Abc	Name	AB_NYC_2019...	name
#	Host Id	AB_NYC_2019...	host id
Abc	Host Name	AB_NYC_2019...	host name
Abc	Neighbourhood Group	AB_NYC_2019...	neigh...
Abc	Neighbourhood	AB_NYC_2019...	neigh...
#	Latitude	AB_NYC_2019...	latitude
#	Longitude	AB_NYC_2019...	longit...
Abc	Room Type	AB_NYC_2019...	room...

Filter

Filter	Details
Availability 365	includes values greater than or equal to 1
Minimum Nights	includes values less than or equal to 180
Price	includes values greater than or equal to 1

Add... Edit... Remove OK Cancel

Describe...

AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv	AB_NYC_2019.csv
Longitude	Room Type	Price	Minimum Nights	Number O
-73.9723700	Private room	149	1	
-73.9837700	Entire home/apt	225	1	
-73.9419000	Private room	150	3	
-73.9597600	Entire home/apt	89	1	
-73.9750000	Entire home/apt	200	3	
-73.9849300	Private room	79	2	
-73.9903700	Entire home/apt	150	1	
-73.9654500	Entire home/apt	135	5	
-73.9886700	Private room	85	2	
-73.9877900	Private room	89	4	
-73.9611300	Private room	85	2	
-73.9535200	Entire home/apt	140	2	

Data Source: Sheet 1



## Data Preparation

1. Created a conditional columns “Days Stayed” and “Price Range” for more insightfull analysis

Tableau - Book1

File Data Server Window Help

Connections **AB\_NYC\_2019** Text file

Files

- ☐ Use Data Interpreter
- Data Interpreter might be able to clean your Text file workbook.
- AB\_NYC\_2019.csv
- airline-passeng...traffic (1).csv
- airline-passen...raffic (2).csv
- airline-passen...raffic (3).csv
- airline-passen...ic+(1) (1).csv
- airline-passen...raffic+(1).csv
- airline-passenger-traffic.csv
- bank-marketing.csv
- bitcoinfinal+(4).csv
- churn\_data.csv
- columns\_description.csv
- customer\_data.csv
- Daily+tempera...+data (1).csv
- Daily+temperature+data.csv
- exchange-rate-twi (1).csv
- exchange-rate-twi (2).csv
- exchange-rate-twi.csv
- New Union
- New Table Extension

AB\_NYC\_2019

AB\_NYC\_2019.csv

Need more data?

Drag tables here to relate them. [Learn more](#)

AB\_NYC\_2019.csv 16 fields 31288 rows

100 rows

Type	Field Name	Physical Table	Rem...
#	Id	AB_NYC_2019...	id
Abc	Name	AB_NYC_2019...	name
Abc	Neighbourhood Group	AB_NYC_2019...	neigh...
Abc	Neighbourhood	AB_NYC_2019...	neigh...
+	Latitude	AB_NYC_2019...	latitude
+	Longitude	AB_NYC_2019...	longit...
Abc	Room Type	AB_NYC_2019...	room_...
#	Price	AB_NYC_2019...	price
Abc	Price Range	Calculation	Calcul...

NYC_2019.csv	# AB_NYC_2019.csv	Abc Calculation	# AB_NYC_2019.csv	Abc Calculation	# AB_NYC_2019.csv	Abc Calculation
m Type	Price	Price Range	Minimum Nights	Day's Stayed	Number Of Reviews	Last Revi
ate room	149	Less than 500	1	Less than 5 Day's	9	10/19/2
re home/apt	225	Less than 500	1	Less than 5 Day's	45	5/21/20
ate room	150	Less than 500	3	Less than 5 Day's	0	n
re home/apt	89	Less than 500	1	Less than 5 Day's	270	7/5/201
re home/apt	200	Less than 500	3	Less than 5 Day's	74	6/22/20
ate room	79	Less than 500	2	Less than 5 Day's	430	6/24/20
re home/apt	150	Less than 500	1	Less than 5 Day's	160	6/9/20
re home/apt	135	Less than 500	1	Less than 5 Day's	53	6/22/20
ate room	85	Less than 500	2	Less than 5 Day's	188	6/23/20
ate room	89	Less than 500	4	Less than 5 Day's	167	6/24/20
ate room	85	Less than 500	2	Less than 5 Day's	113	7/5/201
re home/apt	140	Less than 500	2	Less than 5 Day's	148	6/29/20

Data Source Sheet 1