

CA1: Task 2

Amirarsalan Shahbazi 810101451

Farjad Fallah 810101483

Farnoush Fallah 810101484

Tableau public link (Story):

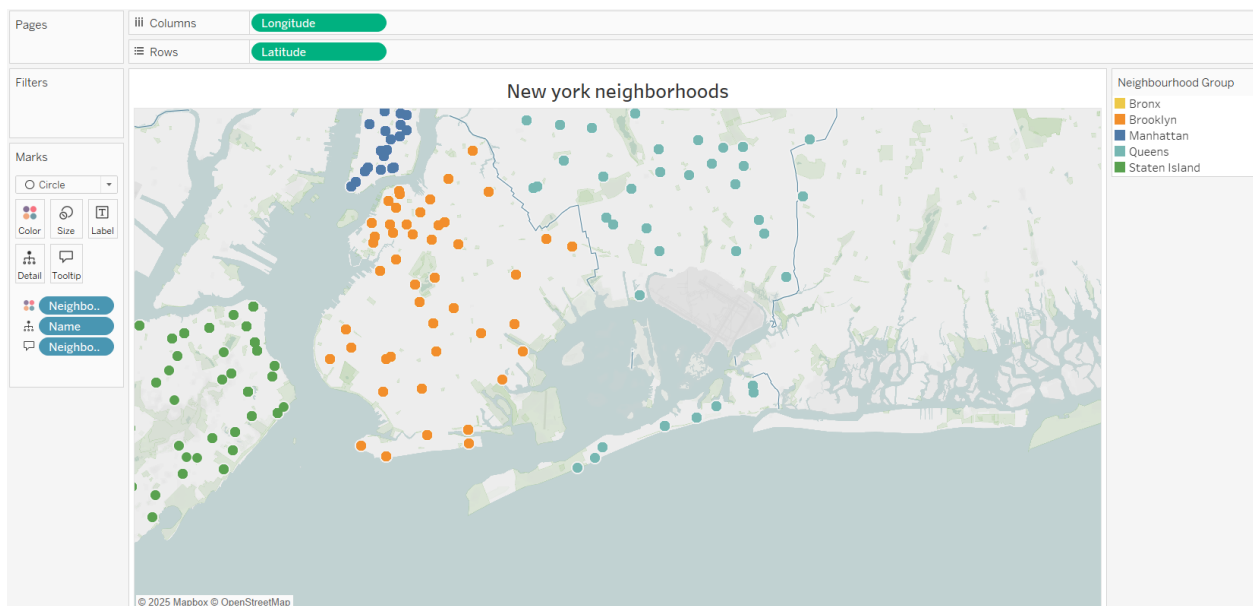
https://public.tableau.com/app/profile/amirarsalan.shahbazi/viz/CA1_17439581065030/WheretostayinNewyork?publish=yes

Our story starts with the fact that during recent years that Airbnb becomes popular among people, we can extract great facts from people using Airbnb. For example, we can check which cities are favored by people the most.

We have data for New York city and five neighborhoods of it.

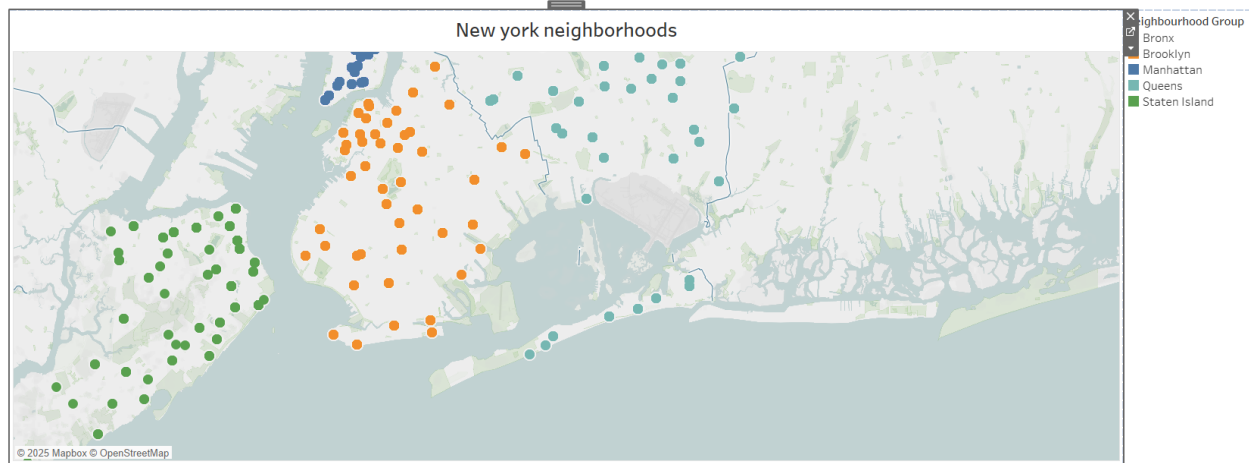
First of all, we join the two databases by common field of neighborhood group and we result in one joined data.

At first, we show on map geographical location of Airbnb places.



This result to our first dashboard that is our introduction dashboard.

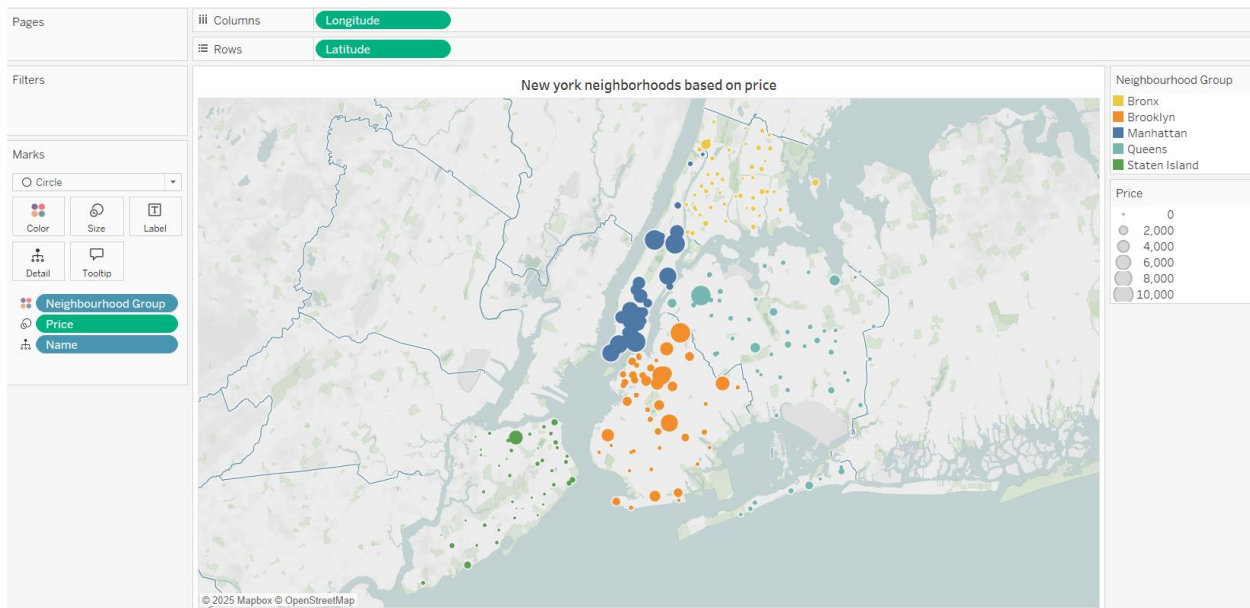
New York is a dream city for so many ones and people often try to reach their goal to visit this city at least once. In this review we aim to see which neighborhoods of this city are best to visit and then which of Airbnb places are most suited for you to stay in.



Here we have five different neighborhoods: Bronx, Brooklyn, Manhattan, Queens and Staten Island

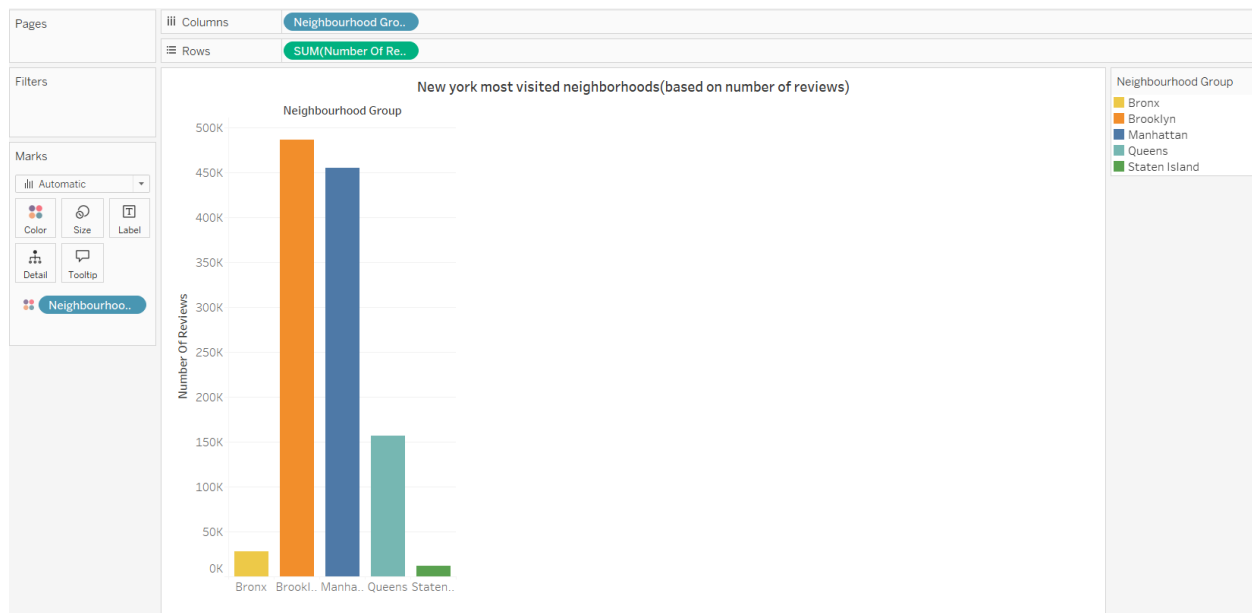
So, in this story we want to check if someone wanted to visit New York city which neighborhoods are best for them to stay in. And we aim to come up with some Airbnb places recommendation for them based on which attraction are they going to visit.

In next sheet we show price distribution in different areas.

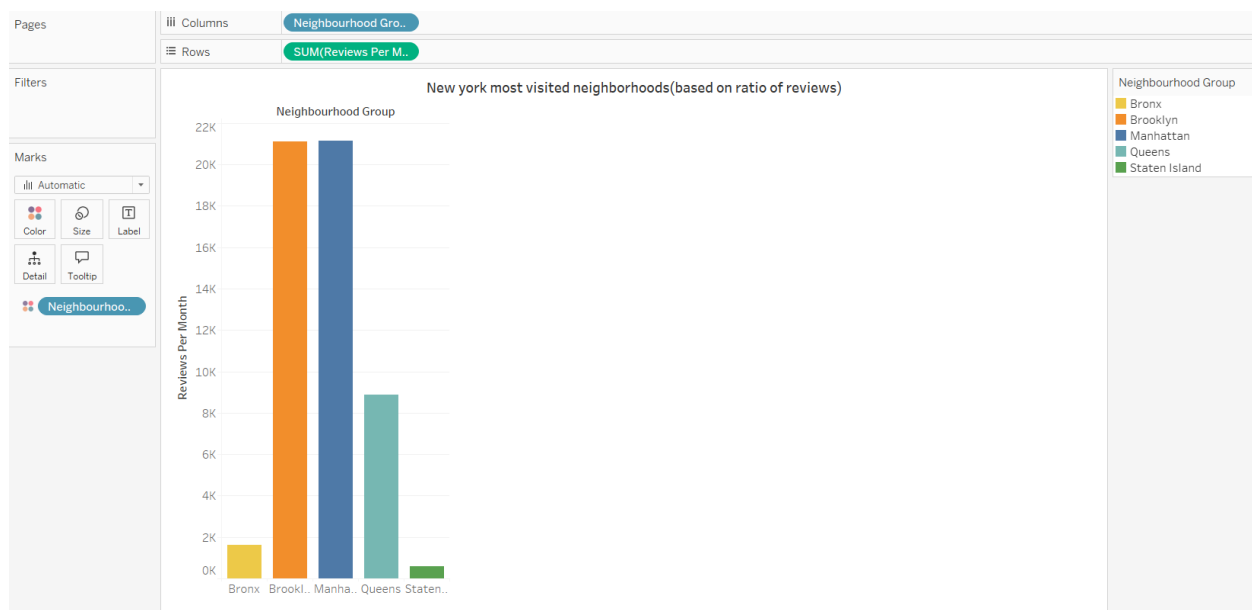


Our attention is absorbed to Manhattan and Brooklyn and we think because of their high price people may go other neighborhoods.

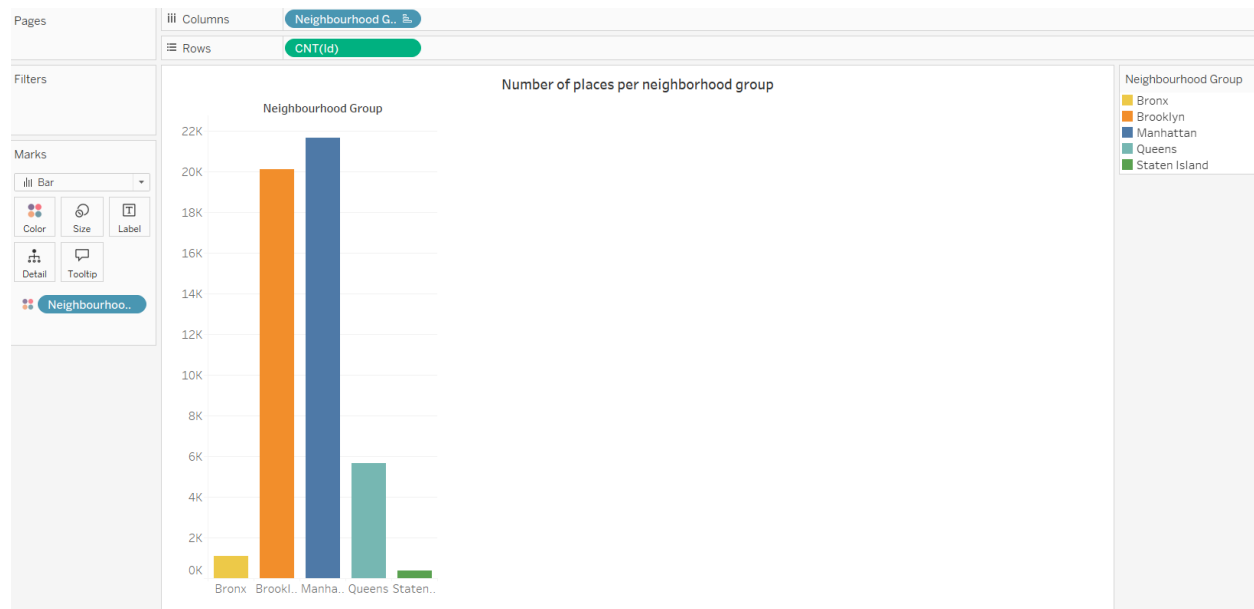
But with these three charts we see that people tend to rent a house in those two neighborhoods.



This chart shows that most reviews are recorded in Manhattan and Brooklyn.

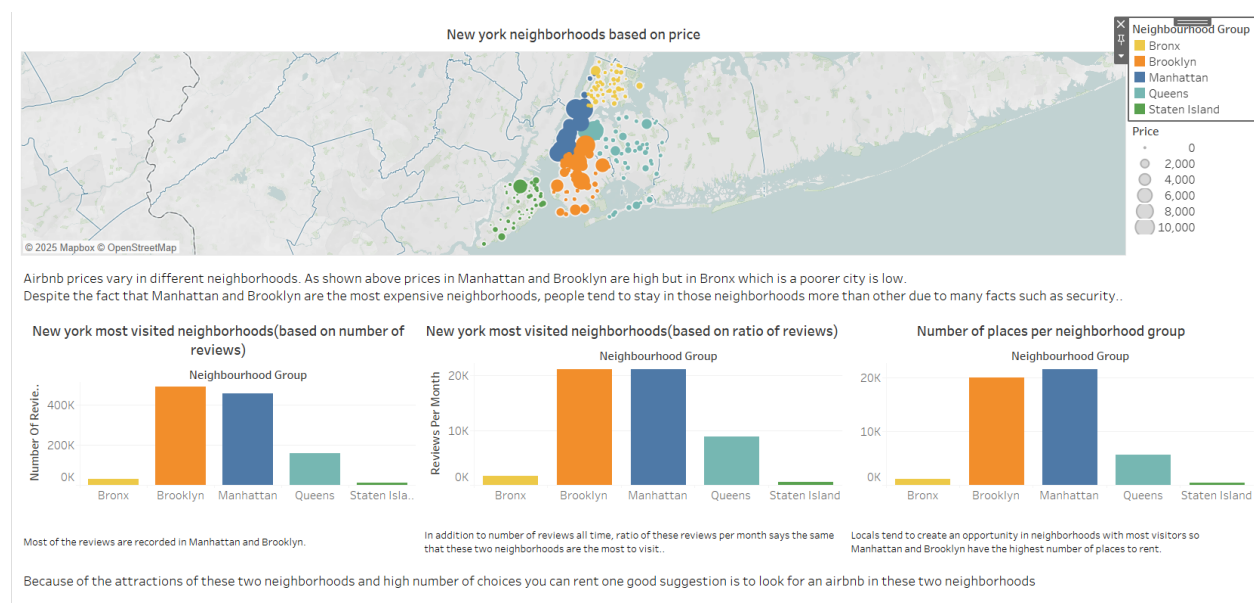


In addition to number of all reviews, ratio of recorded reviews are higher in those neighborhoods too.

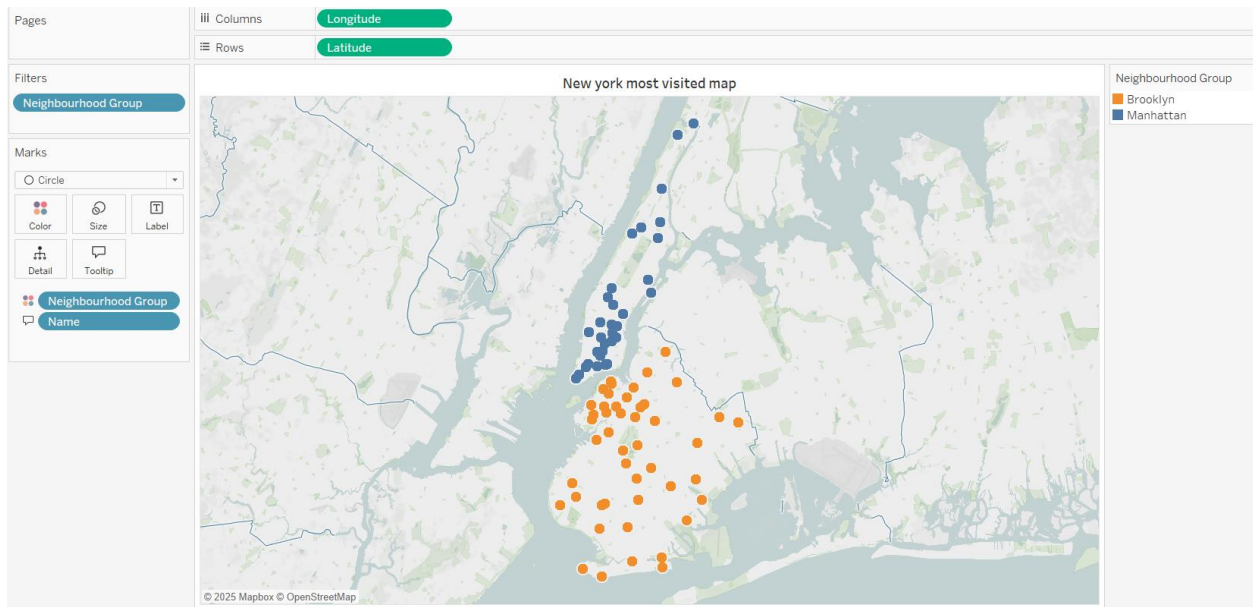


And at last, you can see number of places which are for rent is more in Brooklyn and Manhattan.

Charts above result in second dashboard.



By conclusion from above dashboard, we aim to zoom in Manhattan and Brooklyn statistics.



Now we come up with an idea of finding some famous attractions in those two neighborhoods. First of all we create a parameter called Selected Attraction. User can select the attraction that wants to visit during its presence.

Name

Selected Attraction

Properties

Data type: String

Display format: Rockefeller Center

Current value: Rockefeller Center

Value when workbook opens: Current value

Allowable values

☐ All ☒ List ☐ Range

Value	Display As
Central Park	Central Park
Empire State Building	Empire State Building
Times Square	Times Square
Rockefeller Center	Rockefeller Center
The Met Museum	The Met Museum
One World Trade Center	One World Trade Center
Brooklyn Bridge	Brooklyn Bridge

☒ Fixed ☐ When workbook opens

Add values from

Remove Selected

Cancel OK

For each attraction we create a calculated field for their longitude and latitude.

ected Attraction Latitude

×

```
IF [Selected Attraction] = "Central Park" THEN 40.785091
ELSEIF [Selected Attraction] = "Empire State Building" THEN
ELSEIF [Selected Attraction] = "Times Square" THEN 40.7588
ELSEIF [Selected Attraction] = "Rockefeller Center" THEN 40
ELSEIF [Selected Attraction] = "The Met Museum" THEN 40.77
ELSEIF [Selected Attraction] = "One World Trade Center" THEN
ELSEIF [Selected Attraction] = "Brooklyn Bridge" THEN 40.70
ELSEIF [Selected Attraction] = "Prospect Park" THEN 40.660
ELSEIF [Selected Attraction] = "Brooklyn Museum" THEN 40.67
ELSEIF [Selected Attraction] = "Coney Island" THEN 40.5749
END
```

The calculation is valid.

11 Dependencies ▾

Apply

OK

ted Attraction Longitude

×

```
IF [Selected Attraction] = "Central Park" THEN -73.968285
ELSEIF [Selected Attraction] = "Empire State Building" THEN
ELSEIF [Selected Attraction] = "Times Square" THEN -73.985
ELSEIF [Selected Attraction] = "Rockefeller Center" THEN -
ELSEIF [Selected Attraction] = "The Met Museum" THEN -73.9
ELSEIF [Selected Attraction] = "One World Trade Center" TH
ELSEIF [Selected Attraction] = "Brooklyn Bridge" THEN -73.
ELSEIF [Selected Attraction] = "Prospect Park" THEN -73.96
ELSEIF [Selected Attraction] = "Brooklyn Museum" THEN -73.
ELSEIF [Selected Attraction] = "Coney Island" THEN -73.985
END
```

The calculation is valid.

11 Dependencies ▾

Apply

OK

Also, we create another parameter to filter top hotel ranks in a feature we select.

Edit Parameter [Number of hotels to check] ×

Name

Number of hotels to check

Properties

Data type

Integer

▼

Display format

5,000

▼

Current value

5,000

Value when workbook opens

Current value

▼

Allowable values

☐ All ☐ List ☒ Range

Range of values

☐ Minimum

1

☐ Maximum

100

☐ Step size

1

☒ Fixed

☐ When workbook opens

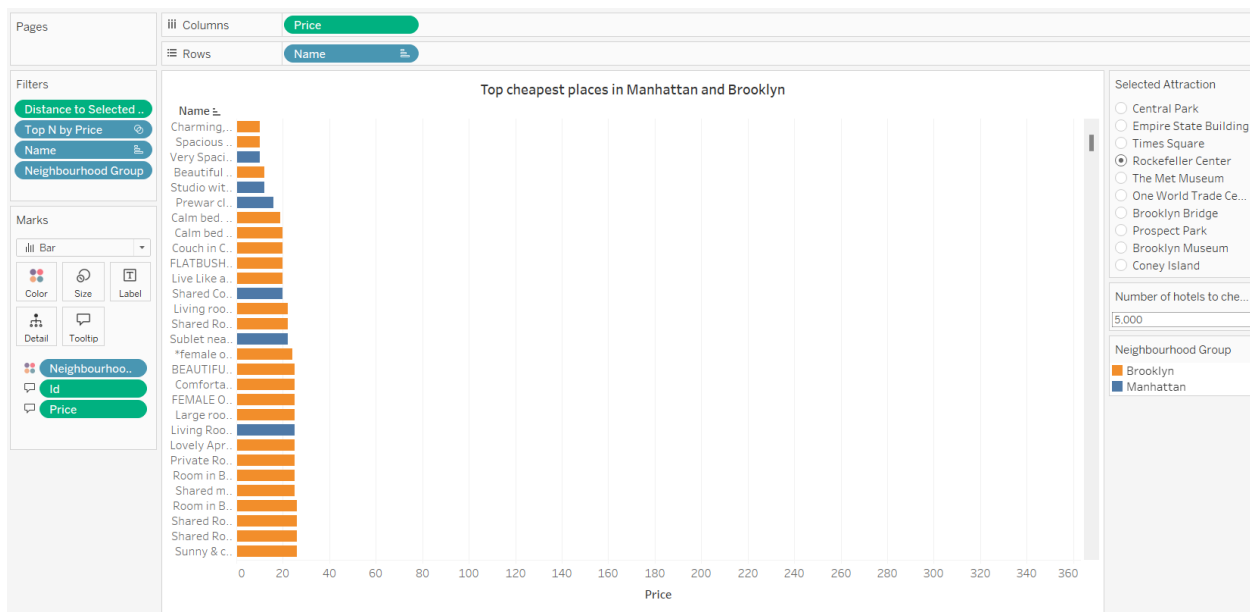
Add values from

▼

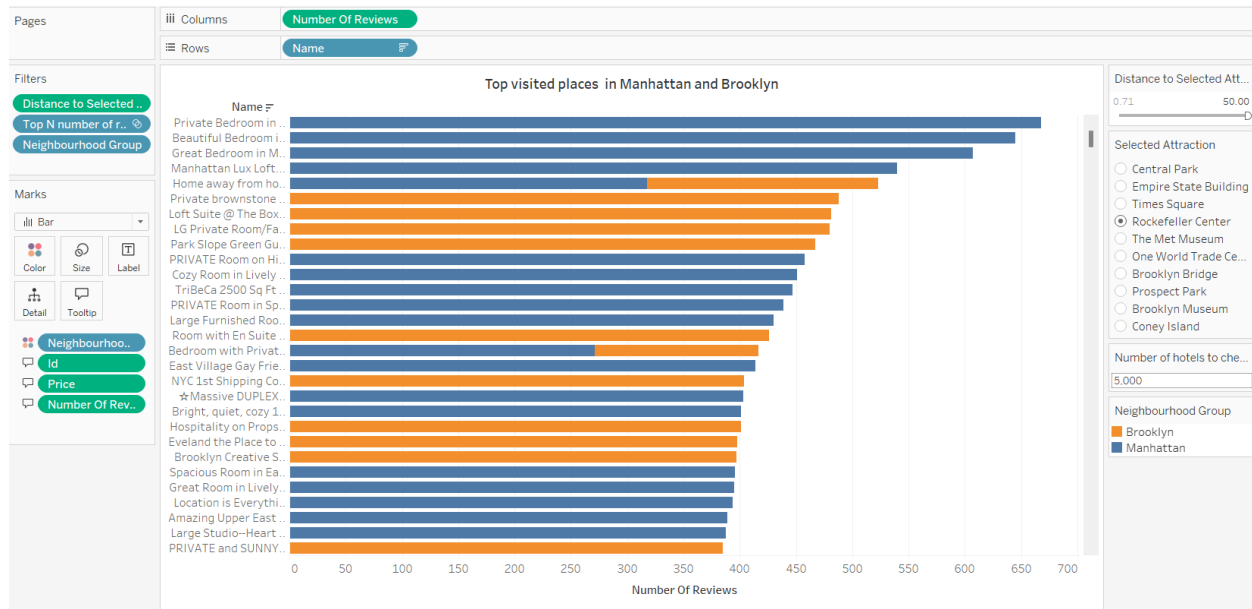
Cancel

OK

For example, in next sheet we rank top 5000 cheapest places in Manhattan and Brooklyn.



We do the same for the most reviewed places in those neighborhoods too.



Next, we create a new calculated field for calculating the distance between hotels to attraction. This field is called Distance to Attraction.

Distance to Selected Attraction

111.32 * SQRT (

([Latitude] - [Selected Attraction Latitude]) ^2 +

(([Longitude] - [Selected Attraction Longitude]) * COS (RADIANS ([Latitude]))) ^2

)

The calculation is valid.
10 Dependencies
Apply
OK

By using Latitudes and Longitudes we calculate distance between two points in kilometer.

With this field we find top nearest hotels to attraction selected by user which is shown below.

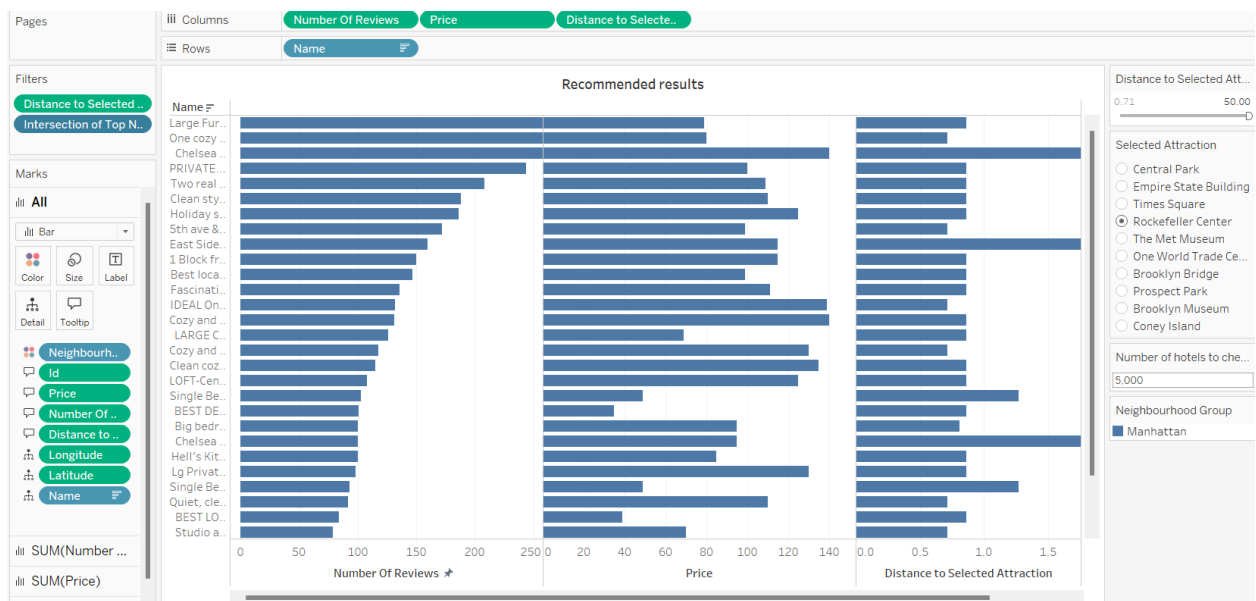
Intersection of Top Ns

[Top N nearest hotels] AND [Top N by Price] AND [Top N number of reviews]

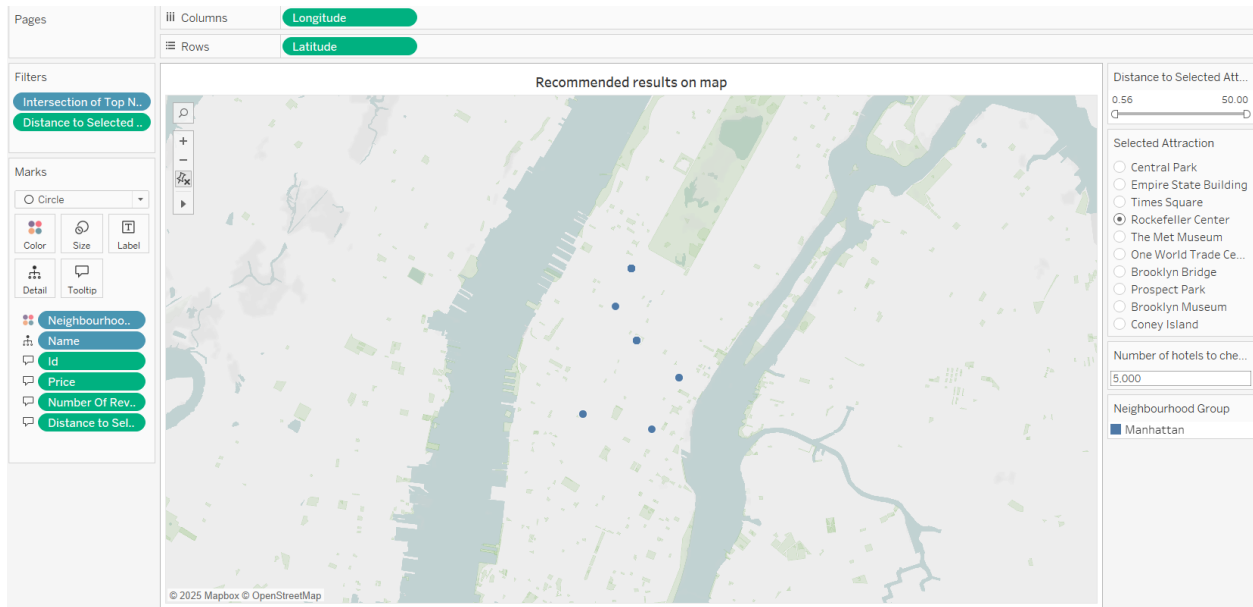
The calculation is valid.
3 Dependencies
Apply
OK

We find hotel names that are ranked top in all of these three sets.

If we choose number of hotels we check, a higher number, we will come up with more hotels to recommend.

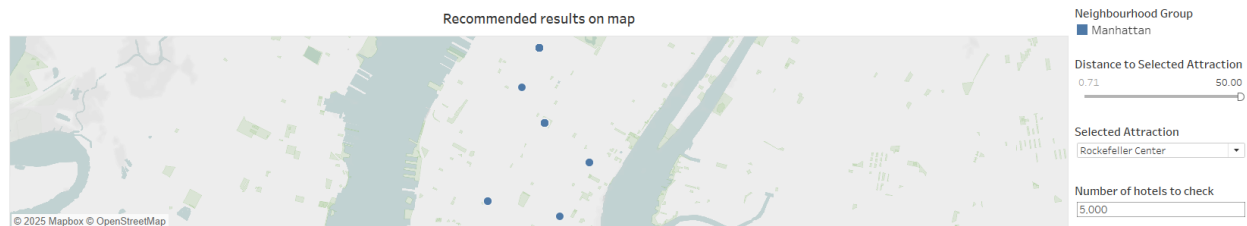


In this sheet we find out places that are ranked top in each factor. We can sort by each factor that is more important to us to find the best choice for our place.



Next, we show geographical locations of recommended places too.

These two sheets lead to our final dashboard which is used to recommend hotels based on the factors and also show the user its geographical location.



Location of recommended hotels are shown above and you can be aware of the geographical position of your place.



The best results by three factors of Number of reviews, price, distance to attractions are listed above. You can sort by the factor which is more important to you to see the best choices. You can set your favorite attraction and check which hotels are best suited for your condition

And at the end user can get different suggestions based on filters that are shown.

So, in conclusion, we start our story with New York places to rent. Then we come up with the idea to limit our journey by just choose one neighborhood between Manhattan or Brooklyn. Then based on an attraction that we want to attend, we come up with some hotel suggestion based on three factors and at last we aggregate our result in order to find a place which is good in all three factors.