AIRBNB VISUAL ANALYTICS SYSTEM

By: Group 9
Sagar Jain (sj735)
Aayush Mandhyan (am2447)
Akshat Kumar (ak1648)

Project Description

AirBnb Analytics!!

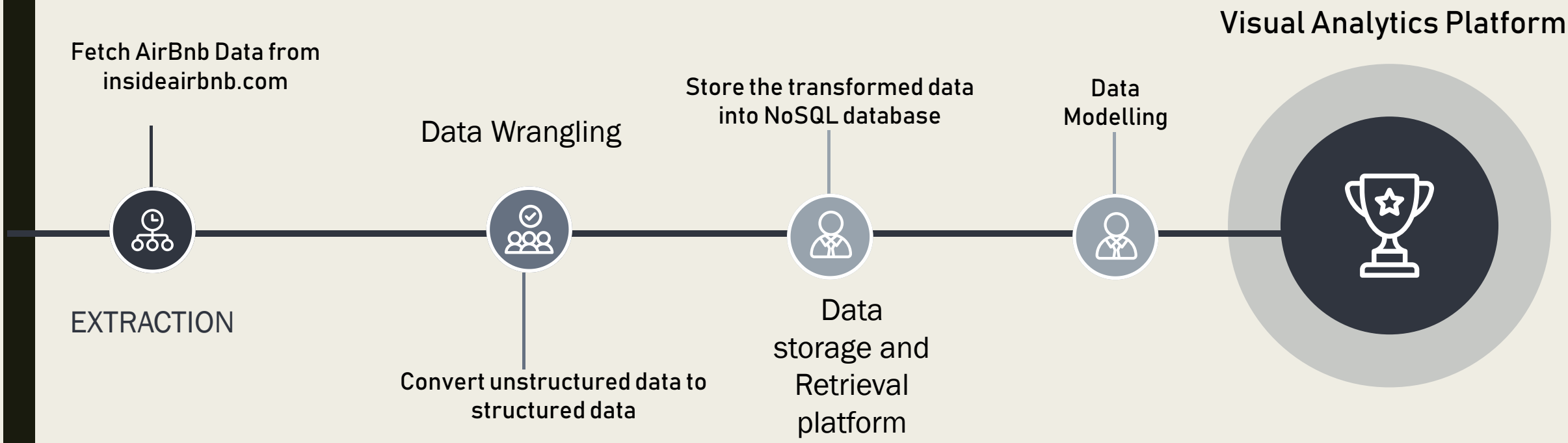
Before AirBnB, it would have been a nerve-wracking prospect to let strangers stay in your home. AirBnB has changed this—with a mission that "connects people with places to stay and things to do around the world." The company has transformed the relationship between the homeowner and the renter. Most of us are familiar with the experience as guests, renting homes to stay in on AirBnB.

Data visualization is effective because it shifts the balance between perception and cognition to take fuller advantage of the brain's abilities. Keeping this in mind we built, an Interactive Visual Analytics platform. Which caters to the user needs of finding perfect destination for travelers and helping the business users by mining Airbnb listings and finding the customer pain points and opportunities.

Project Purpose:

- Explore every listing geographically in an interactive clustered fashion.
- Filter down listing for users based on different parameters.
- Presenting analytical insights using various charts.
- View Demand fluctuation pattern across the year.
- Highlight popular or trending terms based on frequency of use and prominence.
- Finding customer pain points and opportunities for business users to connect.

Project Implementation



Data Collection

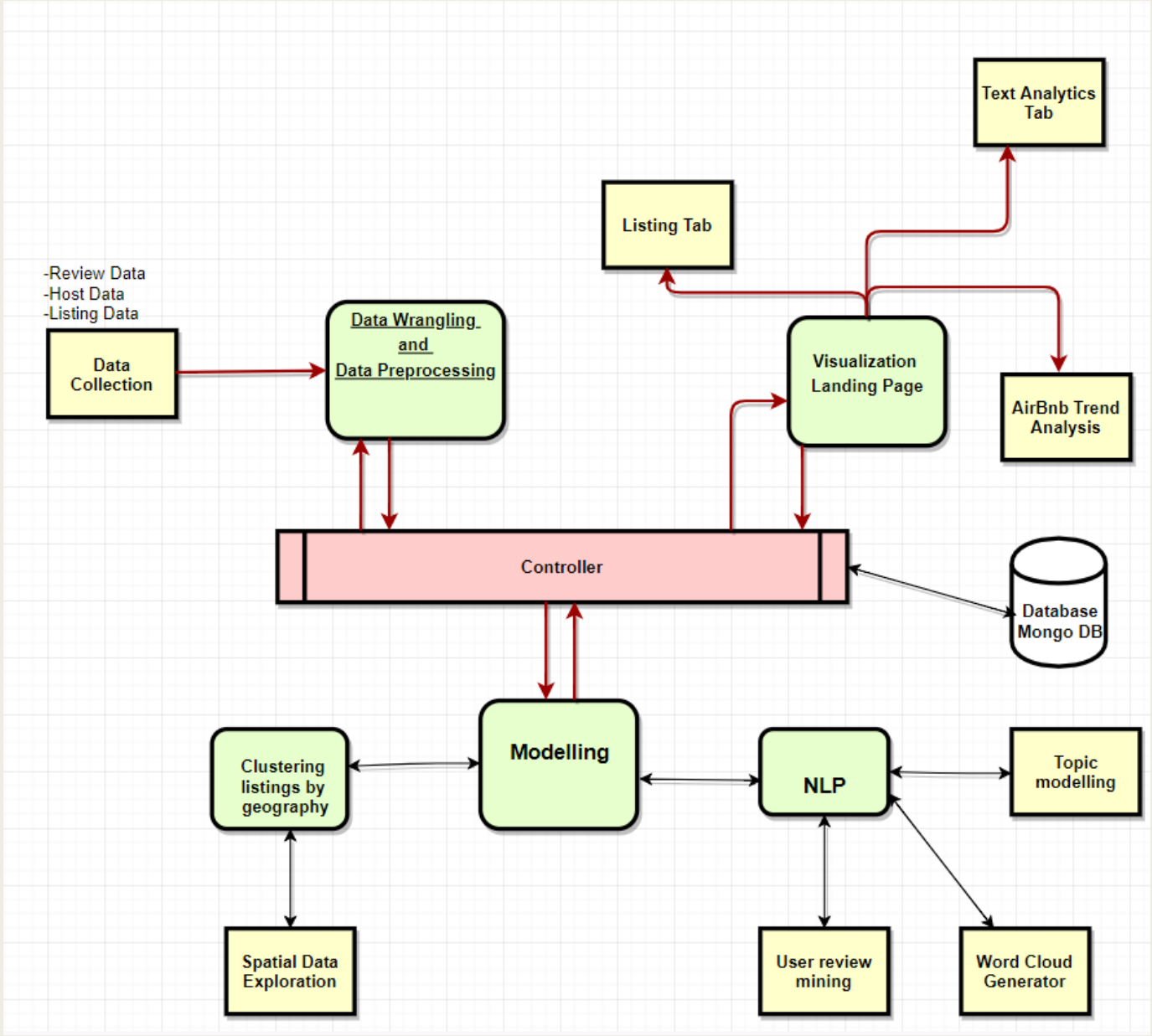
- Source: InsideAirBnb.com
- DataSet: AirBnb listings
- Format Type: CSV format
- Number of attributes: 89
- Number of sample data: 494,954
- Each set of files below for each city;
  example NYC, Asheville, etc.
- Categories of data files:
  - Listings – detailed data about each Airbnb accommodation.
  - Host – detailed description about the hosts.
  - Review – review data set for listings.

Listings & Review Data

id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_ni	number_of_	last_review	reviews_per_	calculated_h	availability_365
2818	Quiet Garder	3159	Daniel		Oostelijk Havengebied - In	52.36575	4.94142	Private room	59	3	250	02/01/19	2.07	1	15
20168	100%Centre-	59484	Alexander		Centrum-Oost	52.36509	4.89354	Entire home/apt	80	1	251	24/02/19	2.29	2	191
25428	Lovely apt in	56142	Joan		Centrum-West	52.37311	4.88367	Entire home/apt	125	14	2	16/02/19	0.15	2	157
27886	Romantic, st	97647	Flip		Centrum-West	52.38673	4.89208	Private room	150	2	180	04/03/19	2.07	1	125
28658	Cosy guest rc	123414	Michele		Bos en Lommer	52.37534	4.85729	Private room	65	3	444	17/02/19	4.14	2	231
28871	Comfortable	124245	Edwin		Centrum-West	52.36719	4.89092	Private room	75	2	240	02/03/19	2.31	3	129
29051	Comfortable	124245	Edwin		Centrum-West	52.36773	4.89151	Private room	55	2	397	17/02/19	4.09	3	157
31080	2-story apart	133488	Nienke		Zuid	52.35132	4.84838	Entire home/apt	219	3	32	16/10/17	0.35	1	0
38266	Nice and qui	164204	Julian		Centrum-West	52.37565	4.87969	Entire home/apt	145	3	198	05/08/15	1.89	1	13
41125	Amsterdam (	178515	Fatih		Centrum-West	52.37891	4.88321	Entire home/apt	180	5	77	04/01/19	0.76	1	29
42970	Comfortable	187580	Micha		Centrum-West	52.36781	4.89001	Private room	159	3	437	24/02/19	4.22	2	158
43109	Oasis in the r	188098	Aukje		Centrum-West	52.37537	4.88932	Entire home/apt	210	13	56	12/06/18	0.54	1	351
43980	View into pa	65041	Ym		Zuid	52.35746	4.86124	Entire home/apt	100	28	61	18/02/18	0.6	2	65
44129	Luxury desig	187728	Tanya		Centrum-West	52.38071	4.8861	Entire home/apt	250	2	176	01/10/18	1.69	5	10
44391	Amsterdam (	194779	Jan		Centrum-Oost	52.37017	4.91438	Entire home/apt	200	5	31	18/11/18	0.3	1	0
46386	Cozy loft in c	207342	Joost		De Pijp - Rivierenbuurt	52.35247	4.90825	Entire home/apt	150	3	3	03/01/18	0.03	1	0
47061	Charming ap	211696	Ivar		De Baarsjes - Oud-West	52.36799	4.87447	Entire home/apt	140	2	169	29/12/18	1.64	1	0
48076	Amsterdam (	219080	Franklin		Centrum-West	52.38042	4.89453	Entire home/apt	350	5	164	11/02/19	1.81	2	232
50518	Perfect centr	231806	Nikki		Westerpark	52.38201	4.87865	Entire home/apt	125	1	92	30/12/18	1.19	1	12
50523	B & B de 9 St	231946	Raymond		Centrum-West	52.36841	4.88413	Private room	115	2	225	06/02/19	2.26	1	180
50570	Bright Apartr	232321	Evert		Bos en Lommer	52.37774	4.84891	Entire home/apt	90	4	155	19/02/19	1.6	2	1
53067	Green studio	246493	Do & Jo		De Pijp - Rivierenbuurt	52.35339	4.90064	Private room	87	1	342	03/02/19	3.39	5	0
53671	Nice room ne	247822	Georg		Westerpark	52.38905	4.88559	Private room	75	3	268	05/03/19	2.78	1	159

listing_id	id	date	reviewer_id	reviewer_na	comments
116134	1344011	25/05/12	231288	Gail	This is a well equipped, very comfortable apartment --- the host has thought of everything.... all the kitchen equipment ima
116134	12771833	11/05/14	10598402	Michel	Wij hebben met ons drieÃ«n een volle week genoten van ons verblijf in het appartement. Deze is zeer ruim, schoon en me
116134	1344011	25/05/12	231288	Gail	This is a well equipped, very comfortable apartment --- the host has thought of everything.... all the kitchen equipment ima

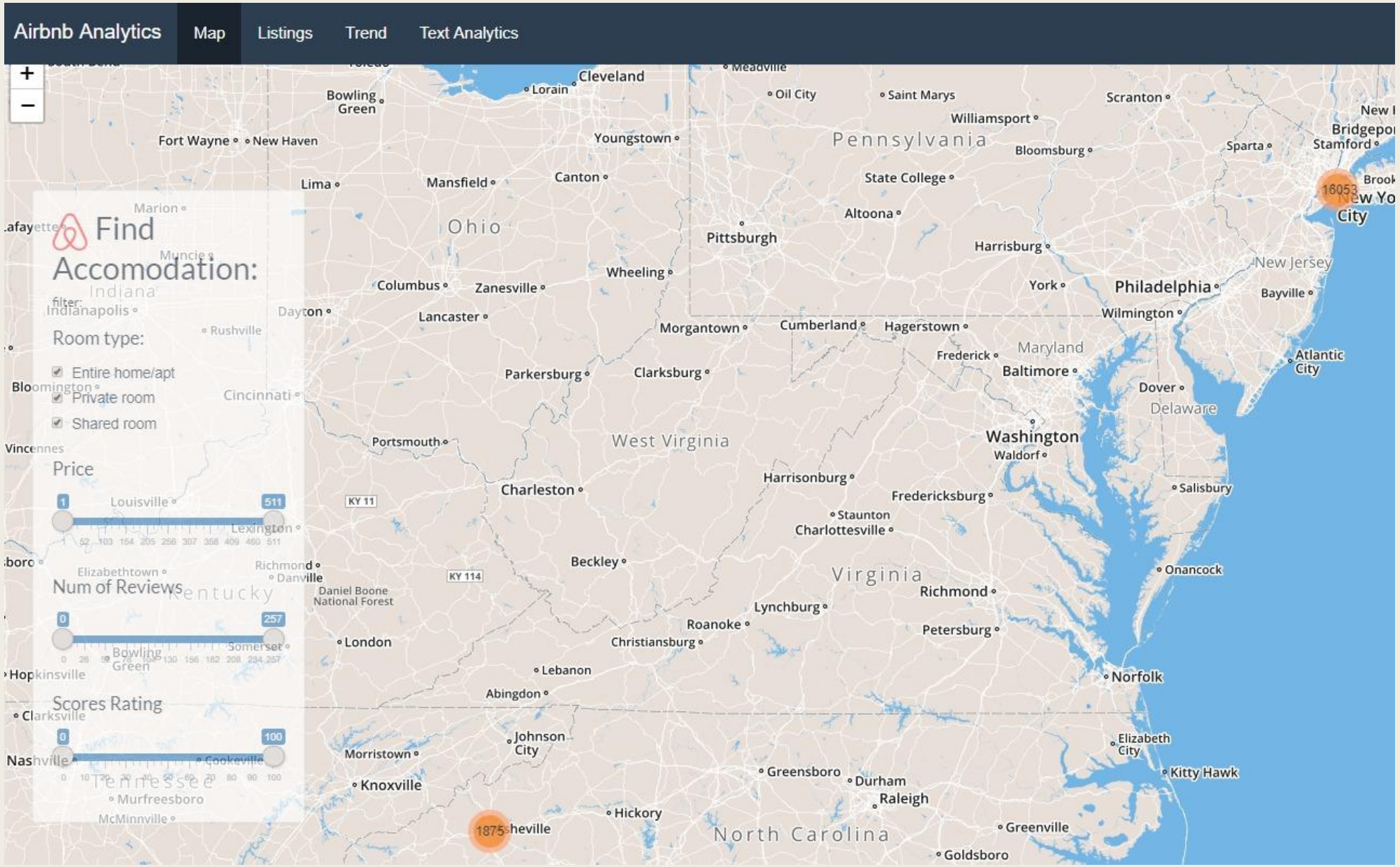
Flow Diagram:



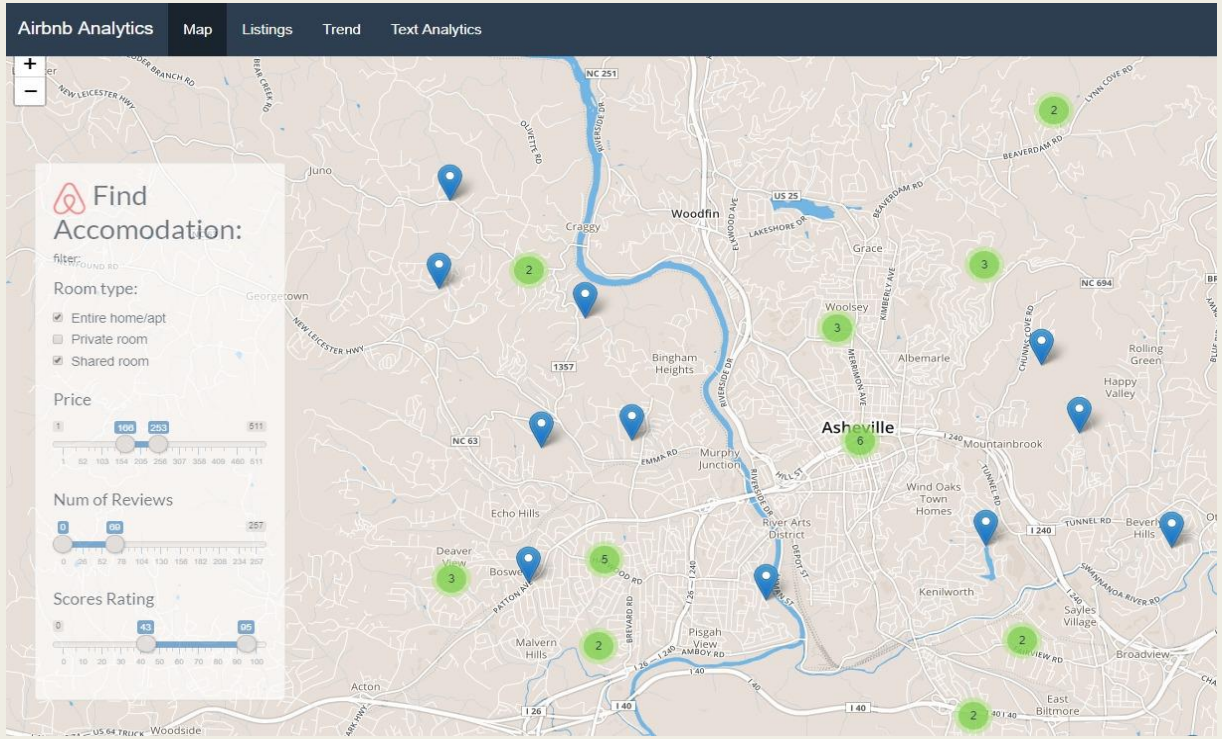
User Types & Interaction:

- Traveler:
  - *Person looking for accommodation while travelling.*
  - *Interaction: They will use the Landing Map page. Where they can drill down to the accommodation on the map. On clicking any accommodation a pop-up will appear listing the information about the listing. They can use the filter facility to filter the listings on the map. Filters provided: room type, price range, rating range, rating score.*
- Business User
  - *Person looking for business insights by mining Airbnb listings.*
  - *Example: Where, which & how many are the costlier accommodation in NYC ?*
  - *Interaction:*
    - This user can interact with the Landing Map page to view the density of accommodation clustering based on various filters.
    - Listing Tab: Query information about density of listings region wise based on different parameters. Outputs in the form of interactive plots.
    - Trend Tab: To analyze the trend and seasonality of Airbnb bookings by using an interactive time series plot.
    - Text Analytics Tab: To get insight into the trending term on the basis of frequency of use and prominence of terms, finding customer pain points by using the word cloud provided in the dashboard and topic modelling dashboard based on different filters.

Map Landing page



Map Landing page: cluster drill down & filtering



Usage:

This is a advance interactive graph with all the listing in USA appearing in a clustered fashion.
You can click on clusters to see the listing it comprises.
This gives a zoom-in view. You can further click on each listing to see details like Listing Name, Host Name, Price of the property, Property Type, Room Type.

Listing Analytics Tab

Airbnb Listings

Analytics of listings based on the neighborhood.

Using this we can observe the trend in the room type in various neighborhood of the city.

filter

Showing as:

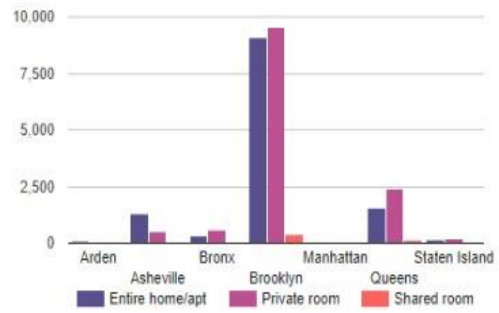
Count

Scores Rating

0 100

0 10 20 30 40 50 60 70 80 90 100

Chart Type?



filter

Showing as:

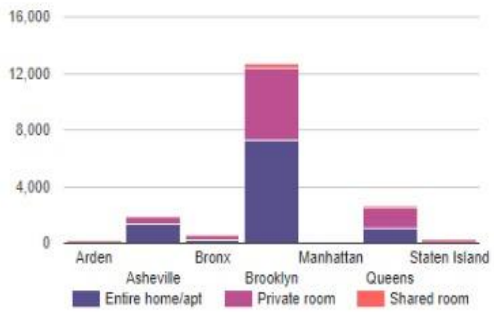
Count

Price

1 511

1 52 103 154 205 256 307 358 409 460 511

Chart Type?



filter

Showing as:

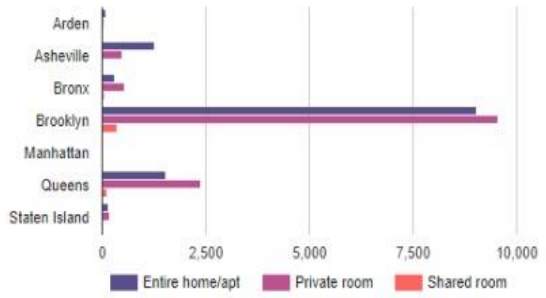
Count

Num of Reviews

0 257

0 26 52 78 104 130 156 182 208 234 257

Chart Type?



Listing Analytics Tab

- Comparison between neighbourhoods cities on the basis of scores ratings, Price and Number of Reviews.
- The user can identify the super host and target the particular group of audience.
- Charts can be configured according to the business requirement and are customized.

Trend Analytics Tab

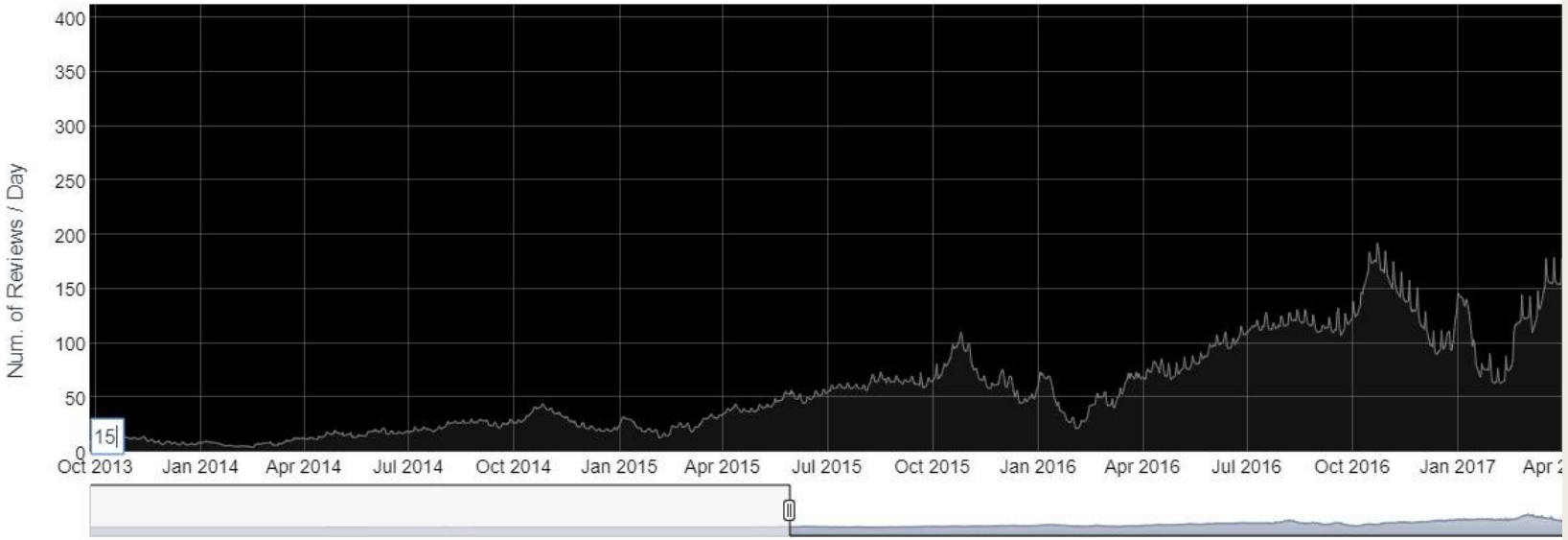
Airbnb Analytics | Map | Listings | Trend | Text Analytics

Trend

The Popularity trend of Airbnb bookings are plotted using the review count per day as our metrics.

We can zoom in the trend using the slider provided on the graph to see the magnified trend in particular period of time, to observe seasonality, peaks during weekends and so on.

Also in the box above we can specify the Moving Average Window for the chart



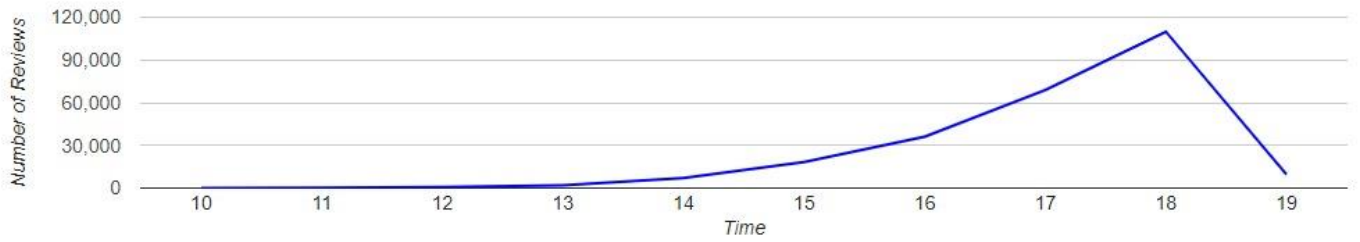
Trend Analysis

Showing data by:

Year

Select Date Range:

08/10/06 to 19/02/10



Trend Analytics Tab

- Usage:

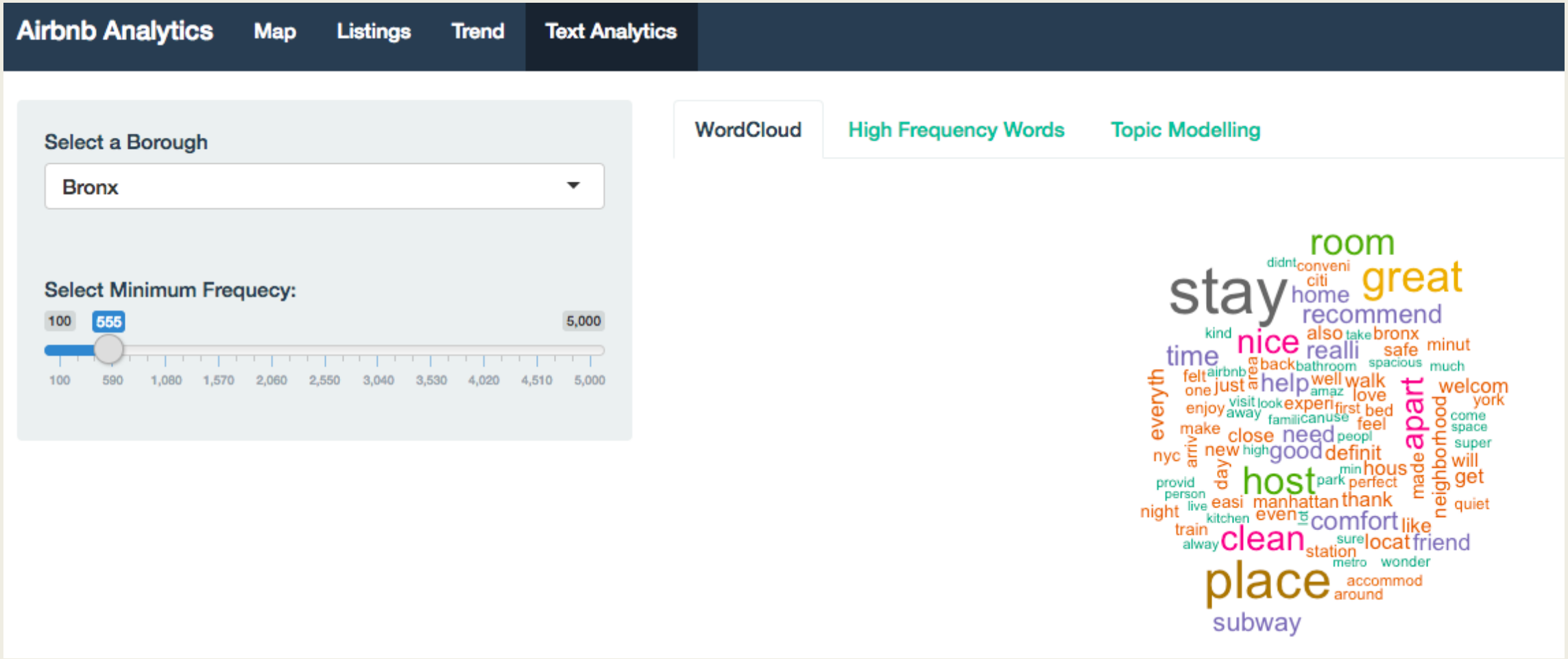
Interactive plot to observe the pattern demand of Airbnb listing through number of review per day. Drag the scale in the x axis to observe the demand for listings in a particular time period. Hover over the graph to see the total number of reviews on a particular day.

- Questions Answered:

About 50% of guests review the hosts/listings, hence studying the number of review will give us a good estimation of the demand.

Consistent pattern in how demand fluctuates across the year. The demand is lowest in January and increases until October, when it begins to falls until the end of the year. Due to the holiday season kicking in, with people celebrating Thanksgiving and Christmas at home with their family.

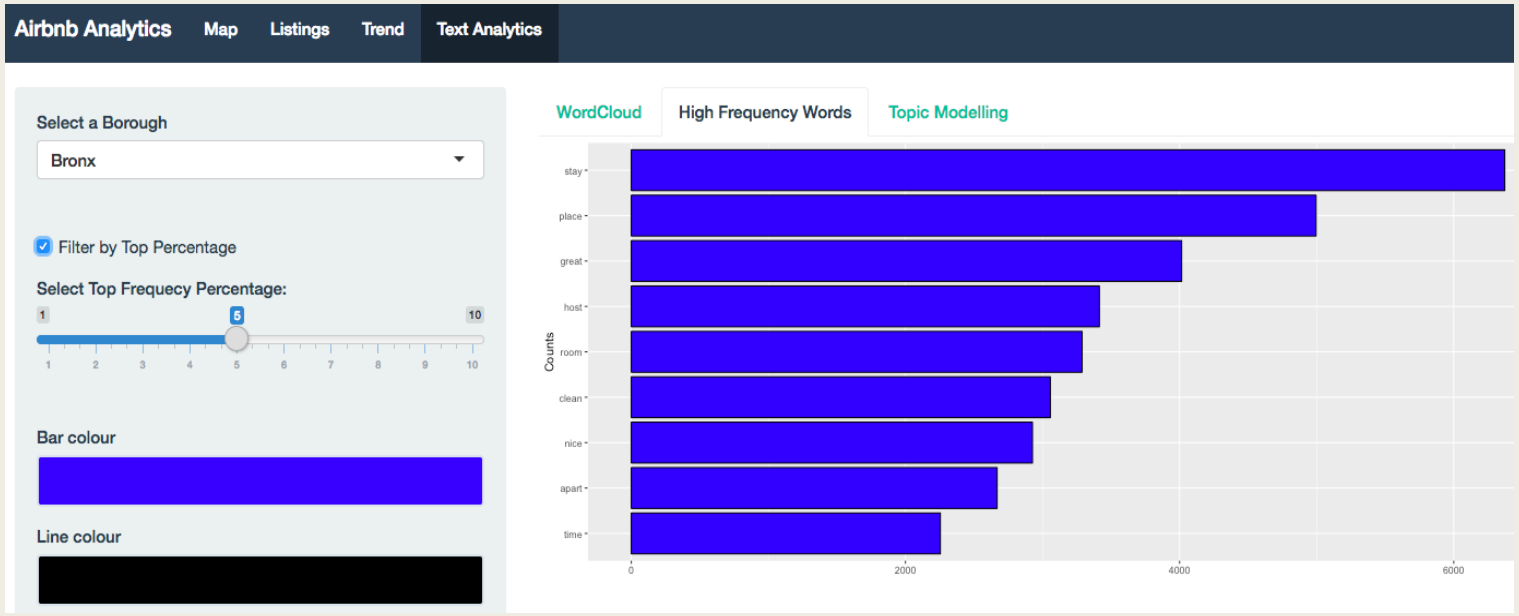
Text Analytics Tab



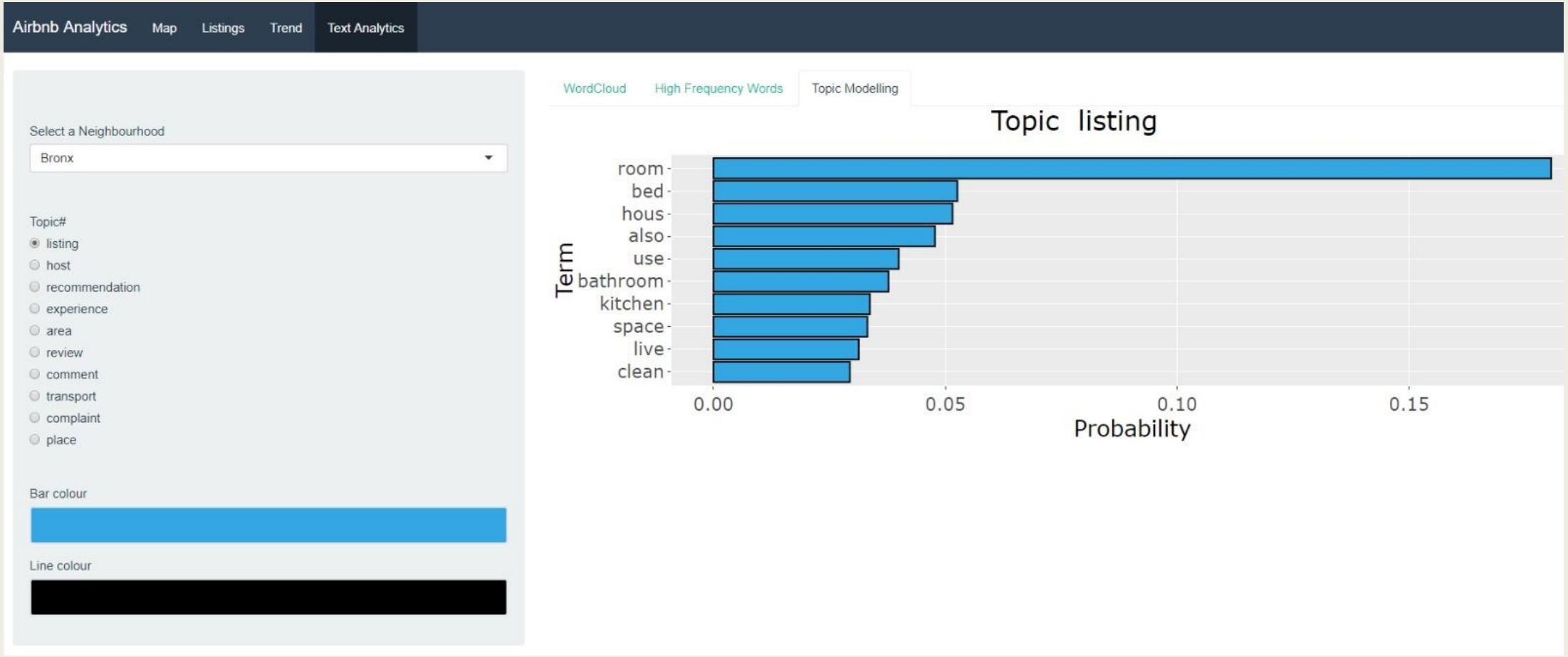
Wordcloud and High Frequency

- Highlight popular or trending terms based on frequency of use and prominence
- Finding customer pain points and opportunities to connect
- Pain points (such as “wait time,” “price,” or “convenience”) are very easy to identify with text clouds
- “Late”, “rude”, “filthy” in customer feedback should ring a warning bell.
- Potential keywords to target that your site content already uses
- Set the frequency of words coming in the reviews to show a refined output.
- Filter by top percentage

Text Analytics Tab



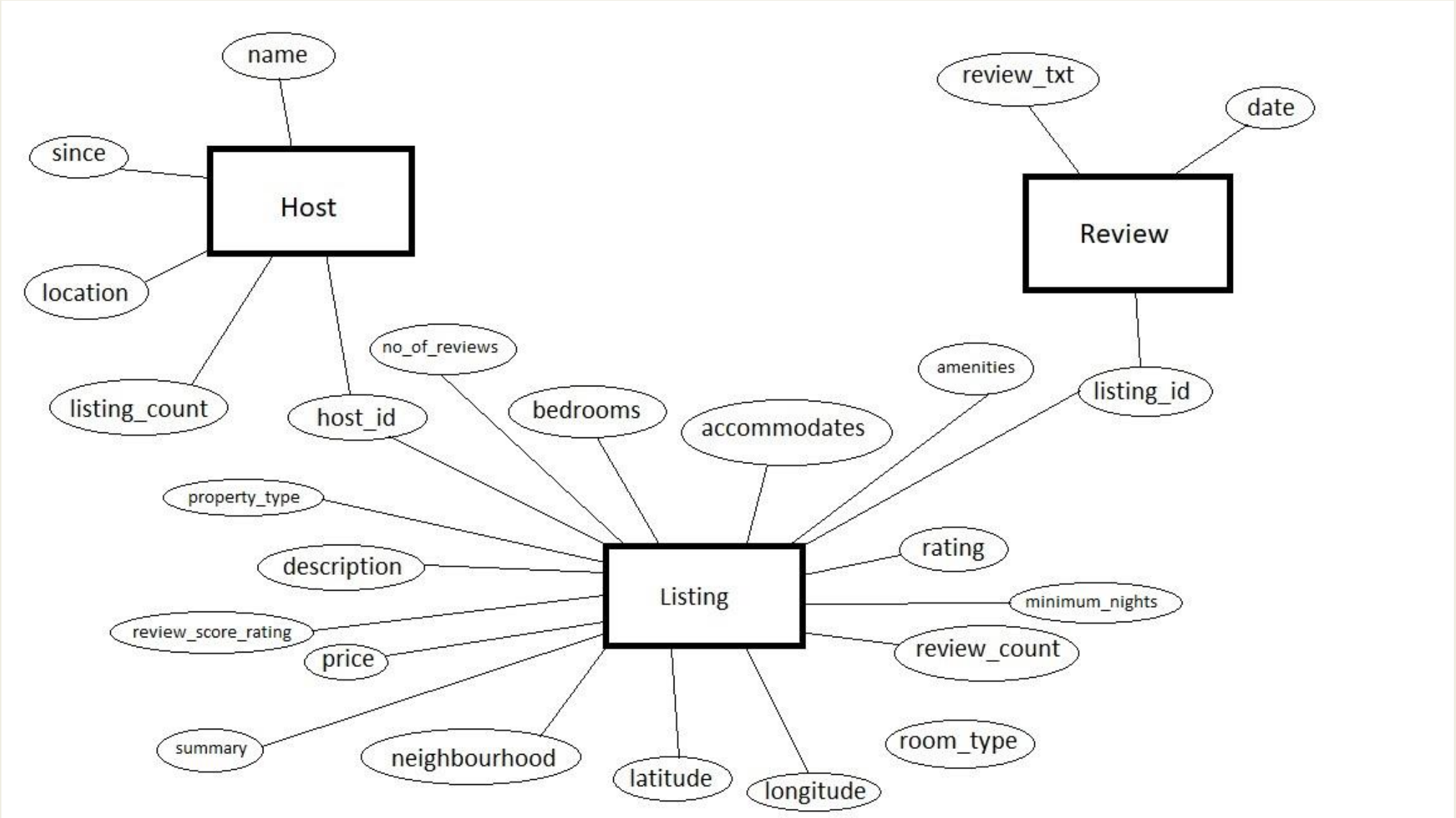
Text Analytics Tab



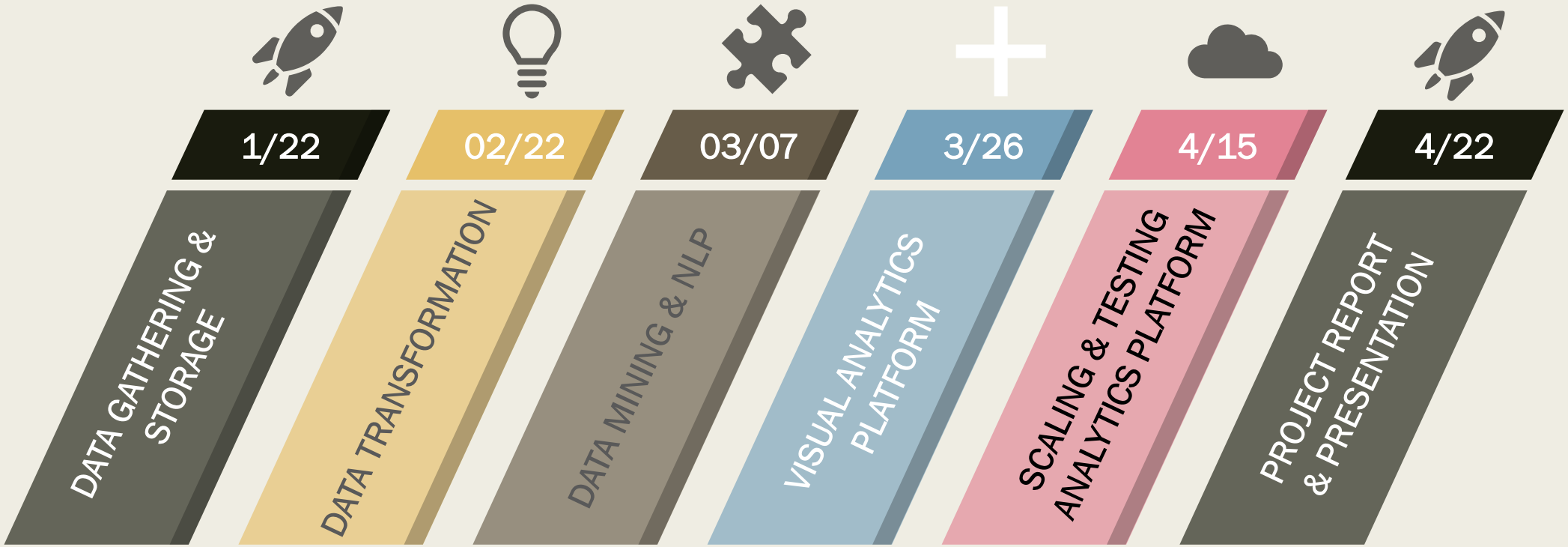
Topic Modelling

- Process of identifying topics in a set of documents.
- Wanted to quickly understand the general trends or topics that Airbnb review contains.
- Important for Airbnb to find out the most trending topics happening in listings in a particular neighborhood
- Host, complaint, listing, neighborhood, transport most popular topics in Airbnb reviews
- Choose the type of topic
- Posterior probability of the topics for each neighborhood and terms for each topic
- Select any topic for each neighborhood from the side tab
- Download, zoom in and out, reset axis options in the tab

ER Diagram:


name
since
Host
location
listing_count
host_id
review_txt
date
Review
no_of_reviews
bedrooms
accommodates
amenities
listing_id
property_type
description
review_score_rating
price
summary
Listing
rating
minimum_nights
review_count
neighbourhood
latitude
longitude
room_type

Project Timeline



Resources:

- RShiny (Web development)
- MongoDB (NoSQL Database)
- Visualization libraries R:
  - *LeafLet*
  - *GoogleVis*
  - *Ggplot.*
  - *Plotly*
  - *DyGraph*

Future Work

- Scaling the application to handle more data and at the same time be more interactive.
- Correlation between various entities in order to extract more relevant information.
- Authentication based on User type.
- Creating platform to select the features and class type to create decision trees

References:

- https://shiny.rstudio.com/gallery/
- https://cran.r-project.org/web/packages/googleVis/vignettes/googleVis_examples.html
- https://www.rdocumentation.org/packages/shiny/versions/1.2.0
- https://rdrr.io/github/jburos/GoogleVis/man/gvisBarChart.html
- http://insideairbnb.com/get-the-data.html

Acknowledgements:

- We would like to thank Professor Abello for his support and constructive criticism which really helped us in building this application.


THANK YOU