Stay in the Loop

Julia Greenberger
Kate Larsen
Kristin Meier
Mari Pastor

airbnb

Agenda

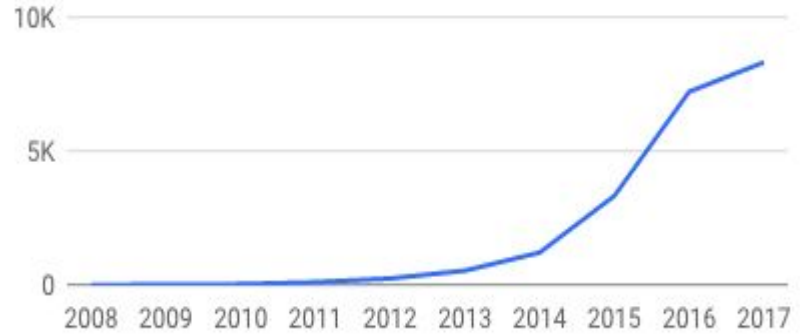
- Problem
- Audience
- Approach
- Design Choices
- Actionable Insight
- Demo

Problem

No comprehensive tool exists for prospective and current hosts to competitively price their listings and gain an understanding of the market at both the city and neighborhood level

Chicago Cumulative Listed Properties[2]

Of currently active listings



Audience

User	**Airbnb Host** *Primary User*
Goal	- Understand competition and market - Post competitive rates relative to area - Generate positive reviews
Questions	- What does the rental market look like in my neighborhood? - What rates should I set? - What features should I offer? - How do host ratings change with listing description?

Approach

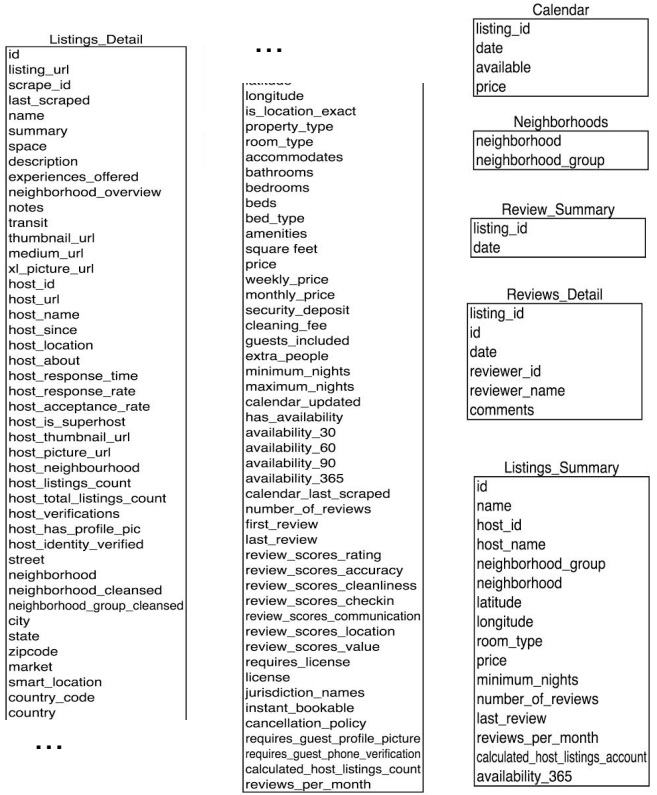
- Interactive dashboard enabling Airbnb hosts to
  - Scope out competition
  - Optimize rental property performance
  - Understand trends and new opportunities

Dashboard Breakdown

Overview	Neighborhood	Text Analytics	Amenities
City level	Neighborhood and listing level	Neighborhood level	City and neighborhood level
Provide comprehensive understanding of current airbnb market in Chicago	Breakdown listings in a particular area and allow hosts to estimate earnings	Analyze listing descriptions and other text and how it relates to ratings	Provide insights into which amenities are offered and how they differ by location

Approach - Technical Notes

- Built in Tableau, file: “StayInLoopFinal.twbx”
  - Dashboard Names:
    - “Title”
    - “Overview”
    - “Neighborhood”
    - “Text Analytics”
    - “Amenities”
  - Main Story: “Story1”
- Data set downloaded from http://insideairbnb.com/get-the-data.html
- Tables on the right outline the available attributes in the dataset



System from End to End

Data Extraction and Preprocessing

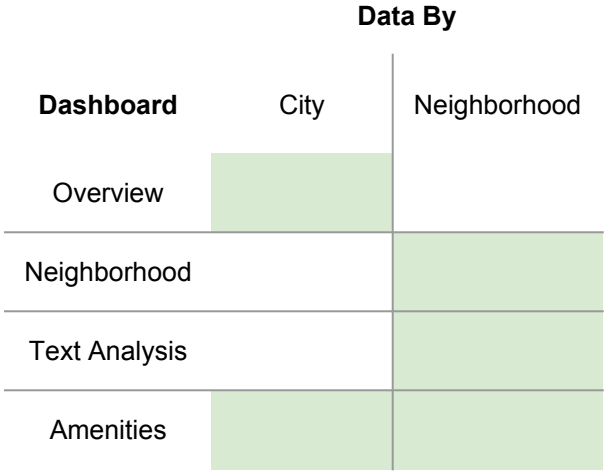
AirBnB

- **Calendar heatmap:** created aggregated file with average price and number of listings for overview and by neighborhood (performance improvement on dashboard queries)
- **Text Analytics:** developed metrics based upon semantic scoring library

Zumper

- Integrated monthly data

Dashboard Creation



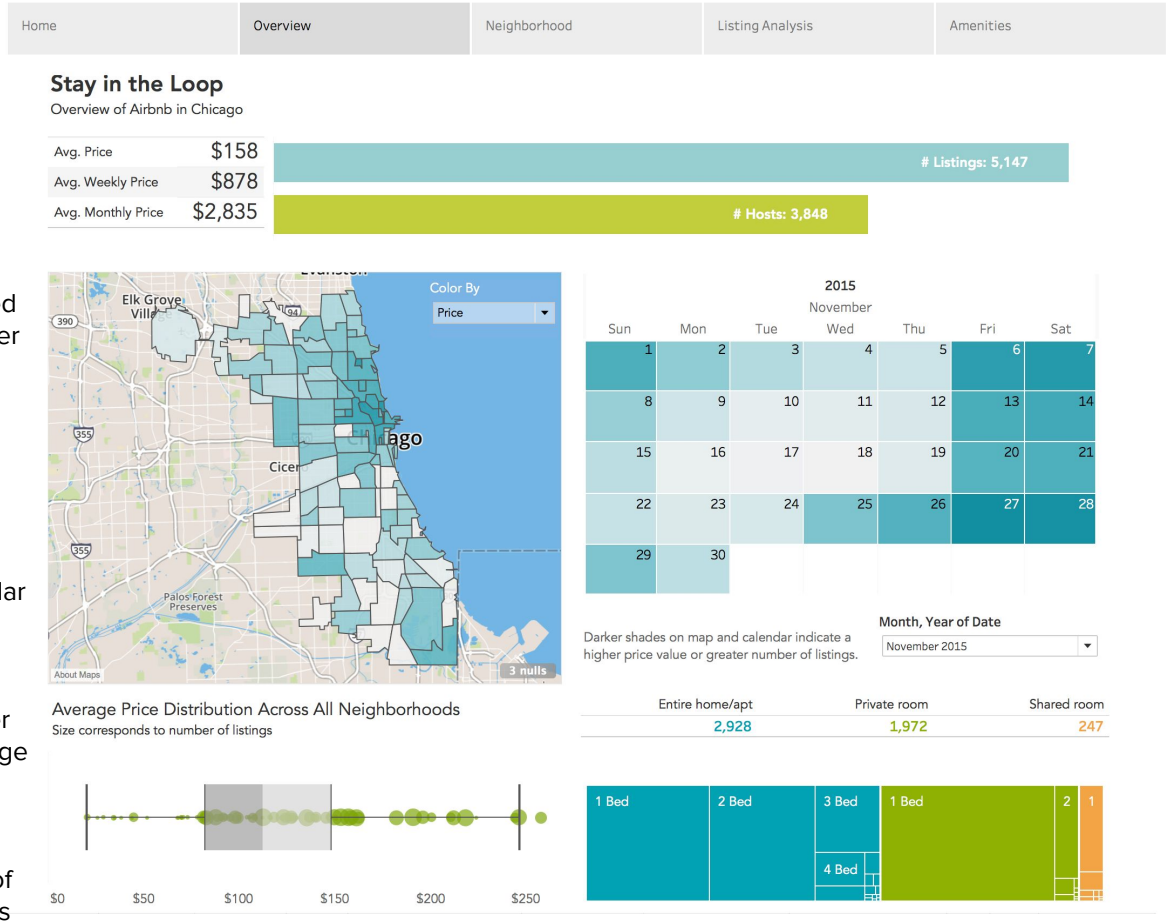
Story Creation

- Add each individual dashboard to story
- Unified style across tabs
- Add title screen explaining each dashboard

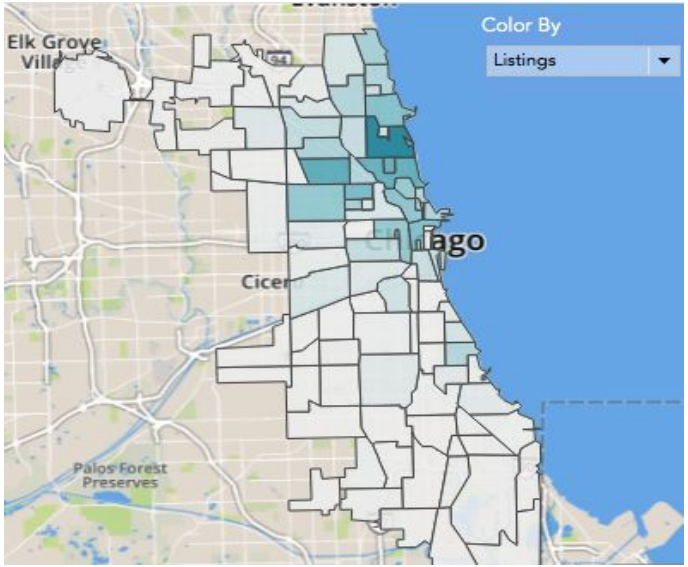
Overview

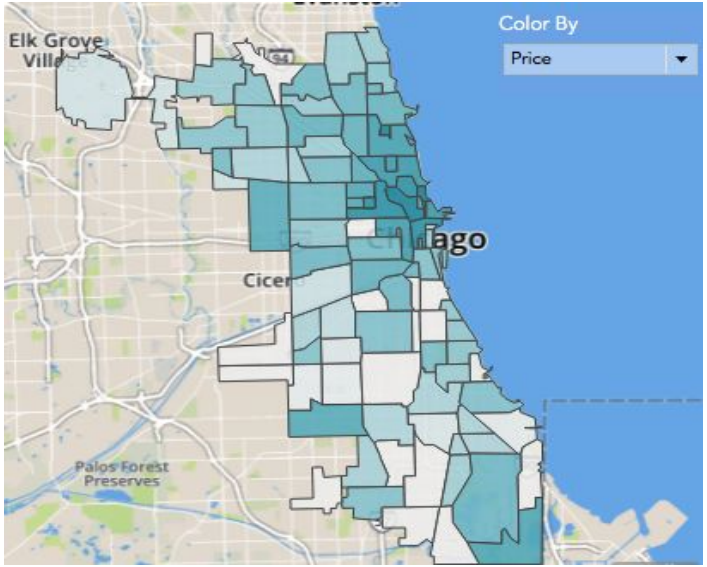
Elements of the Dashboard

- Summary statistics of price, listing count, and host count are displayed across the top for quick reference
- Both the map and calendar can be filtered using the dropdown in the top right corner of the map to display color by price or number of listings
- Map displays aggregated data at the neighborhood level and utilizes official Chicago neighborhood boundaries
- Calendar is made to mimic typical calendar format for easy interpretation. User can choose a month to display
- Average price distribution shown in lower left panel gives user a sense of price range throughout the city and allows them to identify where each neighborhood falls
- Bottom right panel displays breakdown of listings by room type and number of beds



Overview

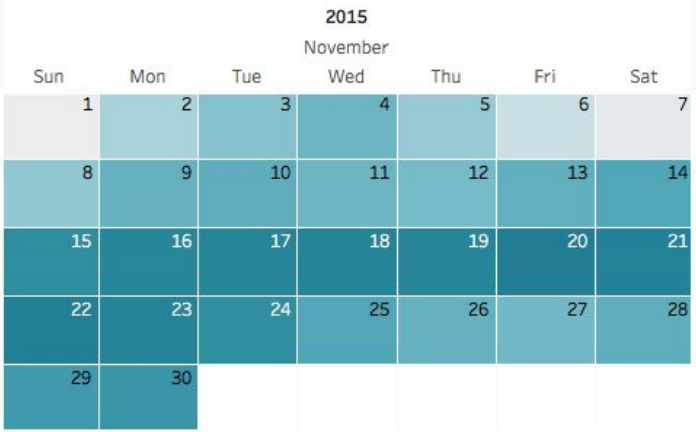




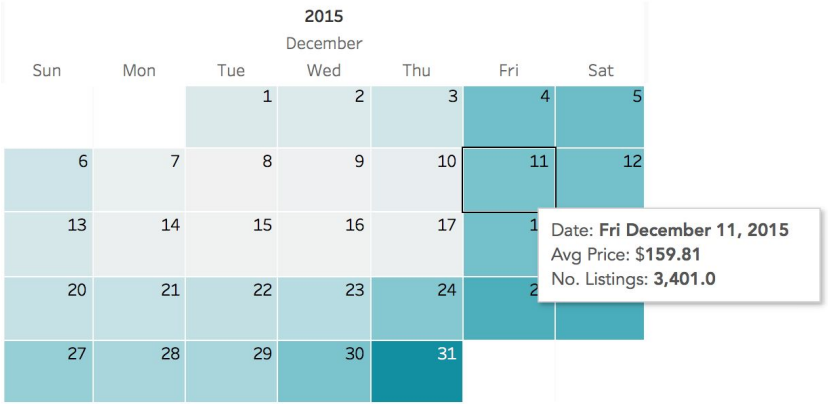
- A simple filter allows users to see the difference in distribution of listings and price

Overview

Color by Listings, November 2015



Color by Price, December 2015



- Calendar heatmap also utilizes the same filter as the choropleth map, for consistency
- Patterns are easily identified using color saturation
  - Weekend prices are higher than weekdays, holidays are more expensive
  - More listings are offered around the holidays
- The calendar can be displayed by Month or by Quarter. We chose Month as more user friendly.

Design Choices - Overview

Goals

- Display listings at the city level
- Provide key summary statistics regarding Chicago airbnb market
- Use Chicago neighborhood boundaries and aggregated data

Insights

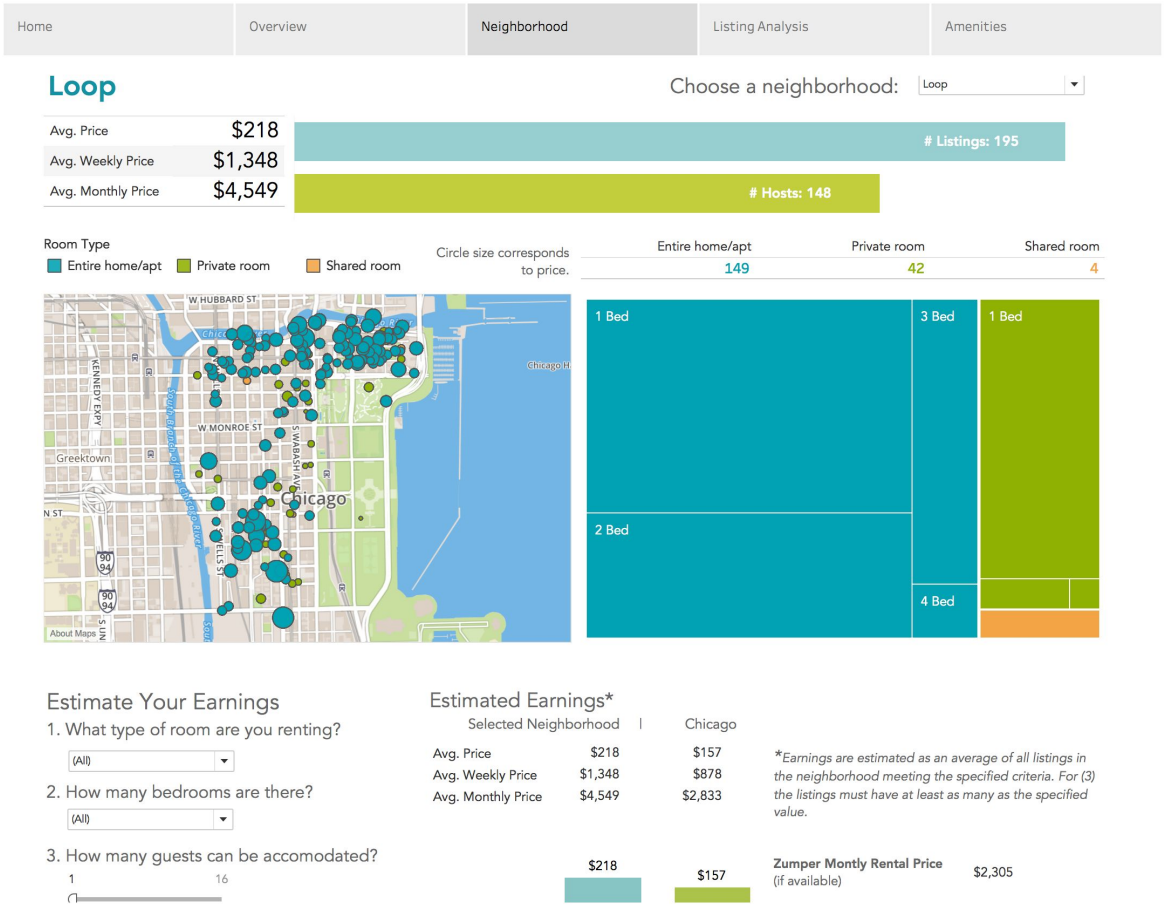
- Most expensive neighborhoods are River North and the Loop
- Most listings are in Lakeview, Lincoln Park and Logan Square
- Weekends are more expensive
- Hosts have more than one listing on average

Visualization	**Choropleth Map**	**Calendar Heatmap**	**Tree Map**	**Bar Chart**
Data	Geographic	Temporal	Hierarchical	Numeric
Marks & Channels	Areas Spatial region Color saturation	Areas Color saturation	Areas Color hue Area	Lines Length Color hue

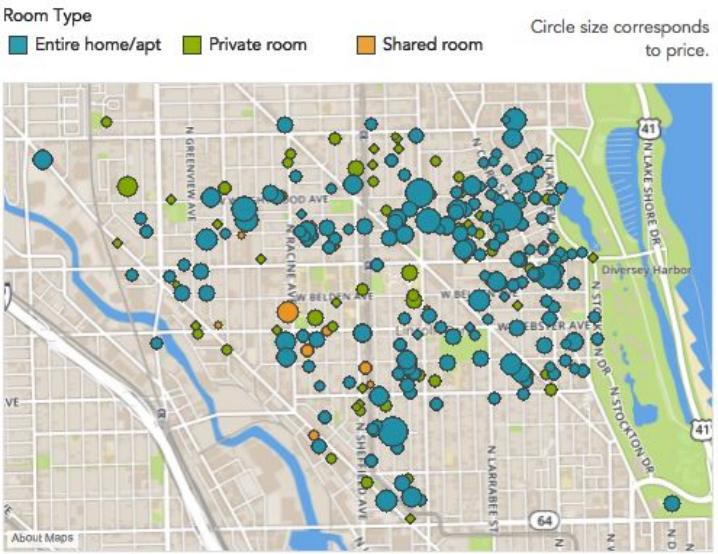
Neighborhood

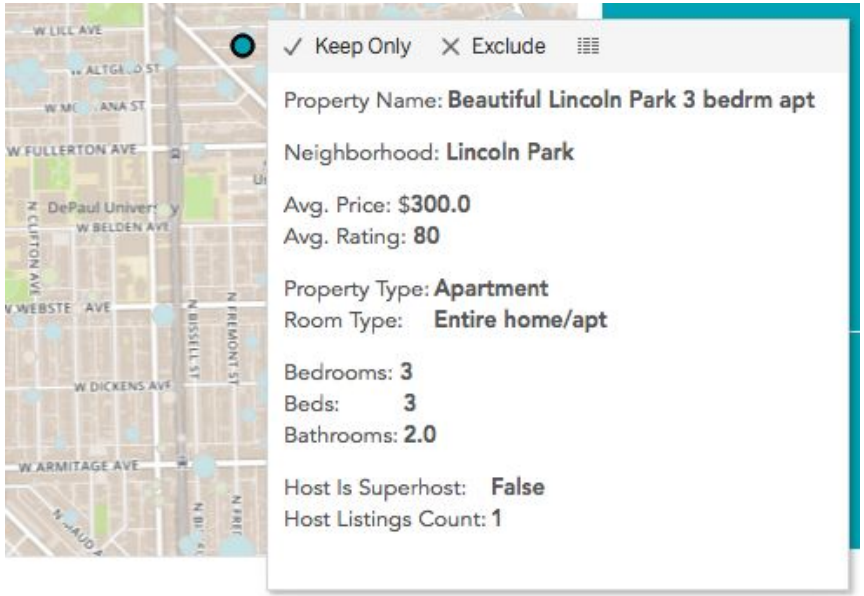
Elements of the Dashboard

- Filter on top right allows user to choose neighborhood to display
- Summary statistics of price, listing count, and host count are displayed across the top for quick reference
- Map displays distribution of listings throughout the neighborhood
- Imported map from Mapbox to customize display. By choosing a background resembling Google Maps we hope to provide a sense of familiarity for the user
- Middle right panel displays breakdown of listings by room type and number of beds
- Estimated earnings section is separated from graphs above
- Easy to use filters and clean text does not overwhelm user



Neighborhood





- Dot density map clearly displays all listing in a given neighborhood, utilizing color hue and area to differentiate by room type and price, respectively
- The zoom and tooltip function allow users to get a closer look at particular listings

Neighborhood

- Easy to use dropdown filter in top right corner of dashboard allows users to choose to a neighborhood
- Estimated earnings section is compact and follows a natural flow from top down for changing filters
- Bar charts examine comparison of estimated rental price for the user's listing compared to the Chicago average of similar listings
- Incorporating Zumper price provides context for making rental decisions
    - Zumper prices were not available for all neighborhoods and only included 1-2 bedroom apartments. Due to the caveats we decided not to highlight this value as a comparison to Avg. Monthly Price and simply provide the number as a reference.

Design Choices - Neighborhood

Goals

- Consistency with Overview tab
- Provide detailed breakdown of listings at the neighborhood level
- Enable hosts to estimate income using similar listings
- Compare with Zumper rental prices, when available

Insights

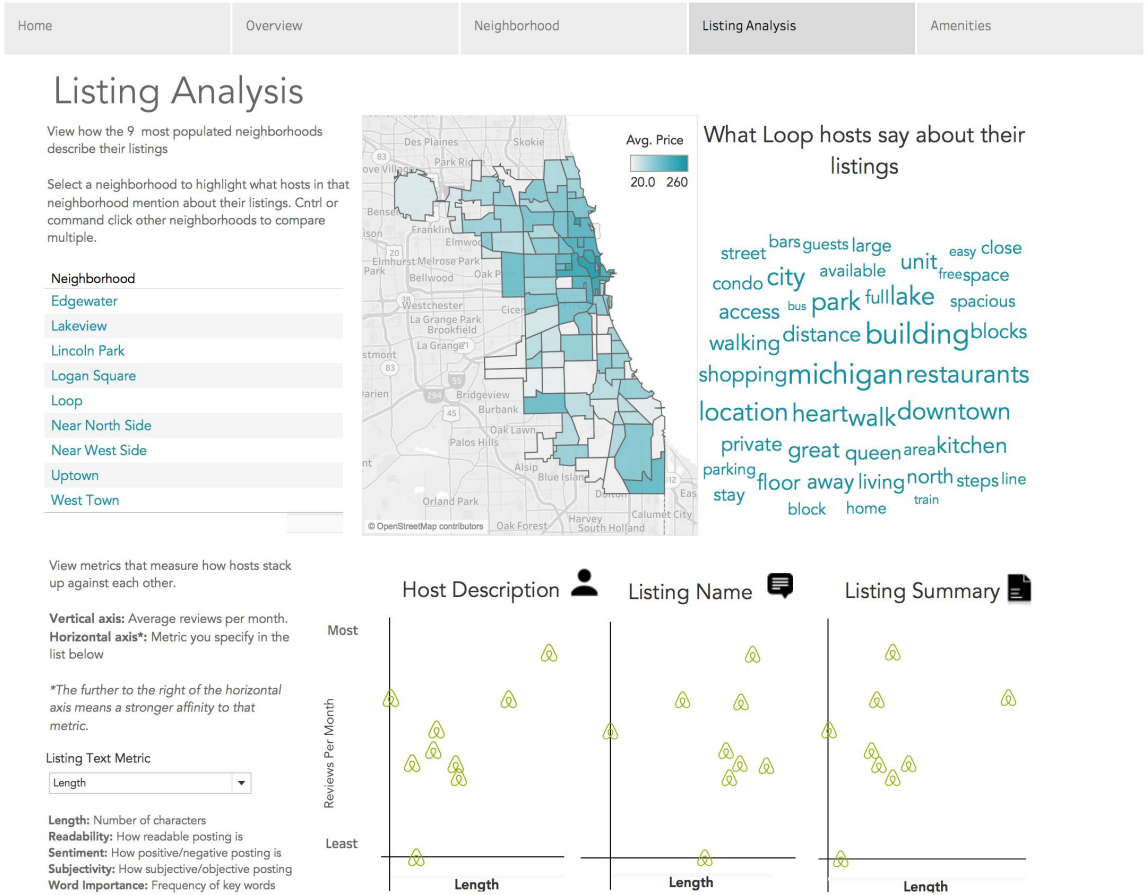
- Airbnb is more profitable than renting
- Room type breakdown is dependent on neighborhood

Visualization	Dot Density Map	Calendar Heatmap	Tree Map	Bar Chart	Table
Data	Geographic	Temporal	Hierarchical	Numeric	Numeric
Marks & Channels	Points Spatial region Color hue Area	Areas Color saturation	Areas Color hue Area	Lines Length Color hue	

Text Analysis

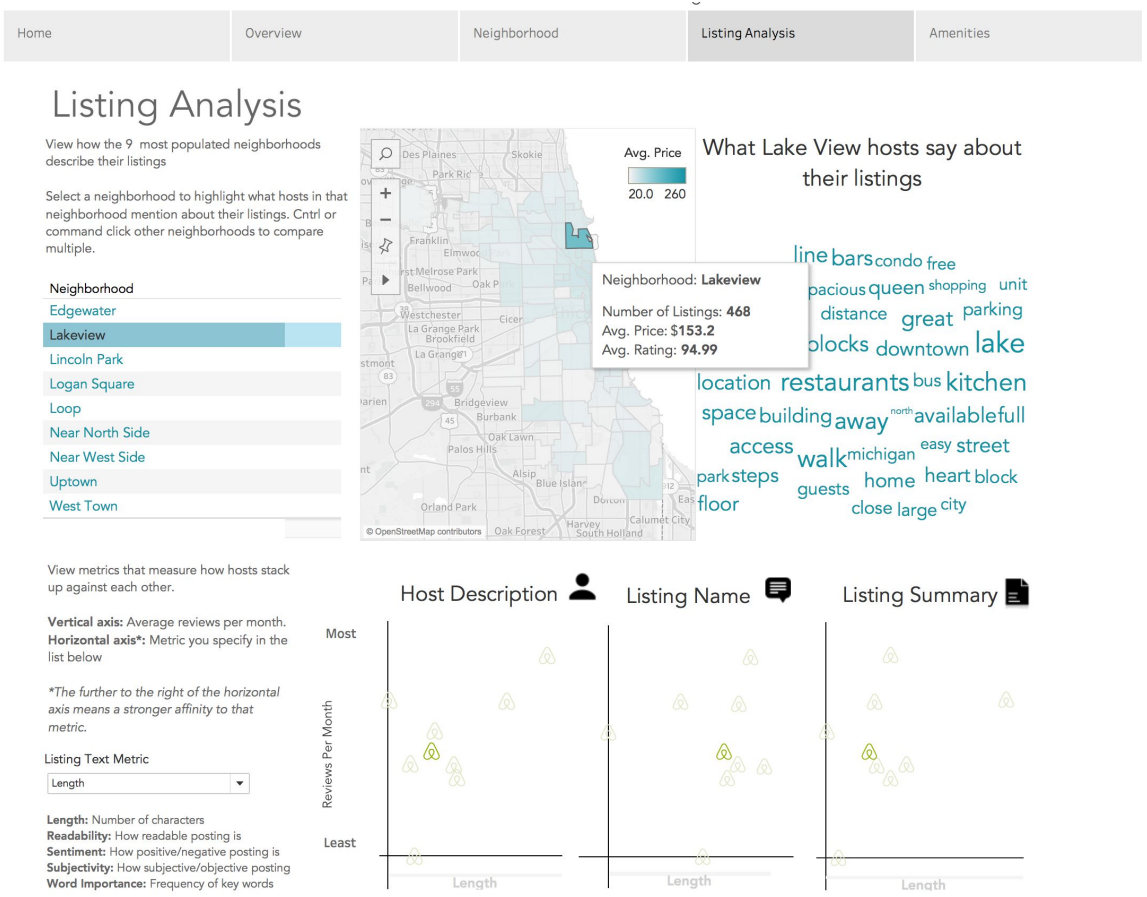
Elements of the Dashboard

- Text on the left to guide the user on how to use the dashboard.
- Menu on the left that allows the user to drill-down on a neighborhood level
- Word cloud that shows the top 50 words most frequently mentioned when describing the listing
- Filter in the bottom right allows the user to change the metric measured on the x-axis of the scatterplots
- Scatter plot shows the relationship between the metric chosen and how many reviews a host may receive monthly



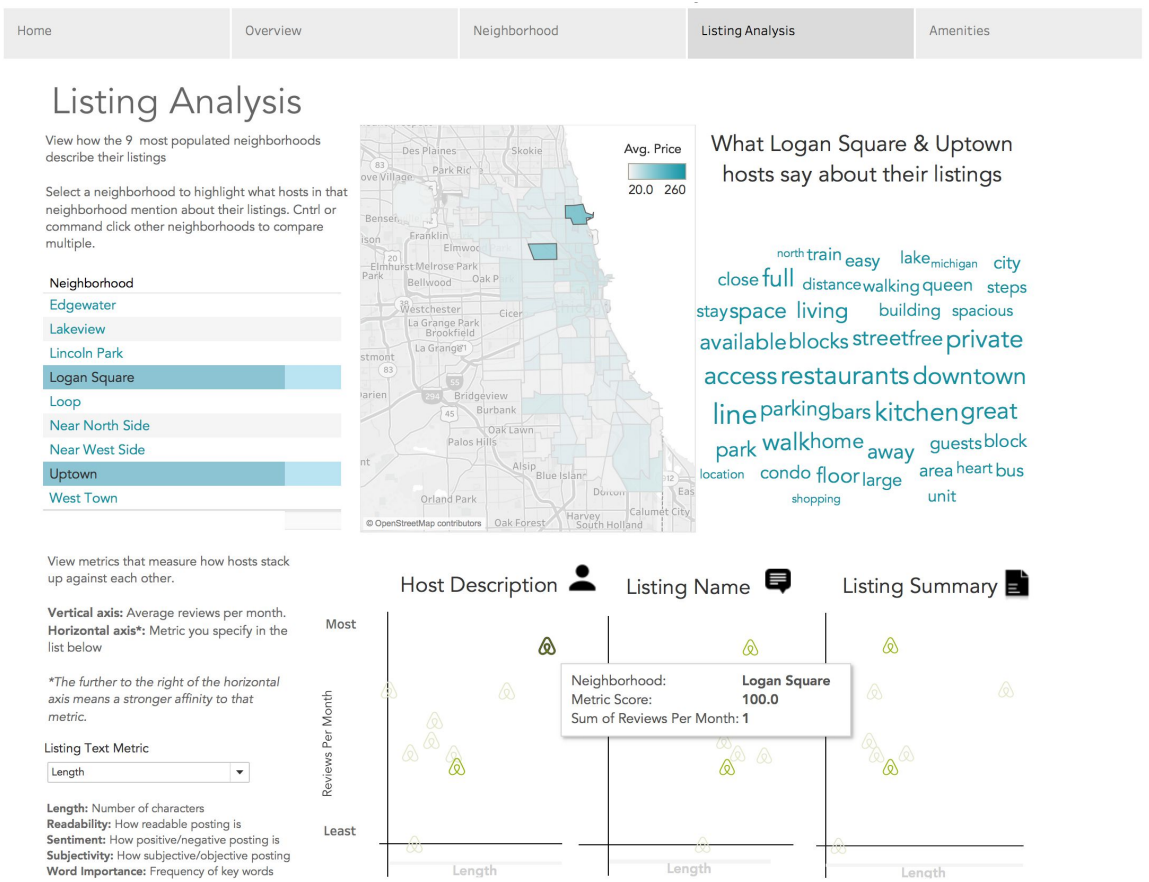
Text Analysis

- Clicking on a neighborhood highlights the observation on the map, and the scatter plots
- Tooltip gives an extra layer of information to the user
- Word Cloud updates based on your neighborhood of choice



Text Analysis

- Clicking on multiple neighborhoods performs the same functionality as it highlights the neighborhoods chosen. The tooltip helps us identify where the neighborhood we selected is in the scatter plot
- Word Cloud adds in the additional neighborhoods selected and selects the top 50 words shown in both neighborhood listings



Design Choices - Text Analysis

Goals

- Explore how hosts describe their listings
- Give a prospective host an idea on how to describe his / her listing

Insights

- Most hosts talk about restaurants nearby or having a kitchen available
- The more keywords your posting contains, the more users contact you for a booking

Visualization	**Choropleth Map**	**Word Cloud**	**Scatter Plot**
Data	Geographic	Text	Numeric
Marks & Channels	Areas Spatial region Color saturation	Words Size	Points Position Area

Visualizations - Text Analysis

Problems - Text Analysis

Design Issues

- Only 9 neighborhoods with the most listings can be selected due to the lack of listings to get a top 50 word cloud
- The separation between neighborhoods in the scatterplots was only visually apparent when we scaled the data with a 0 - 1 scale

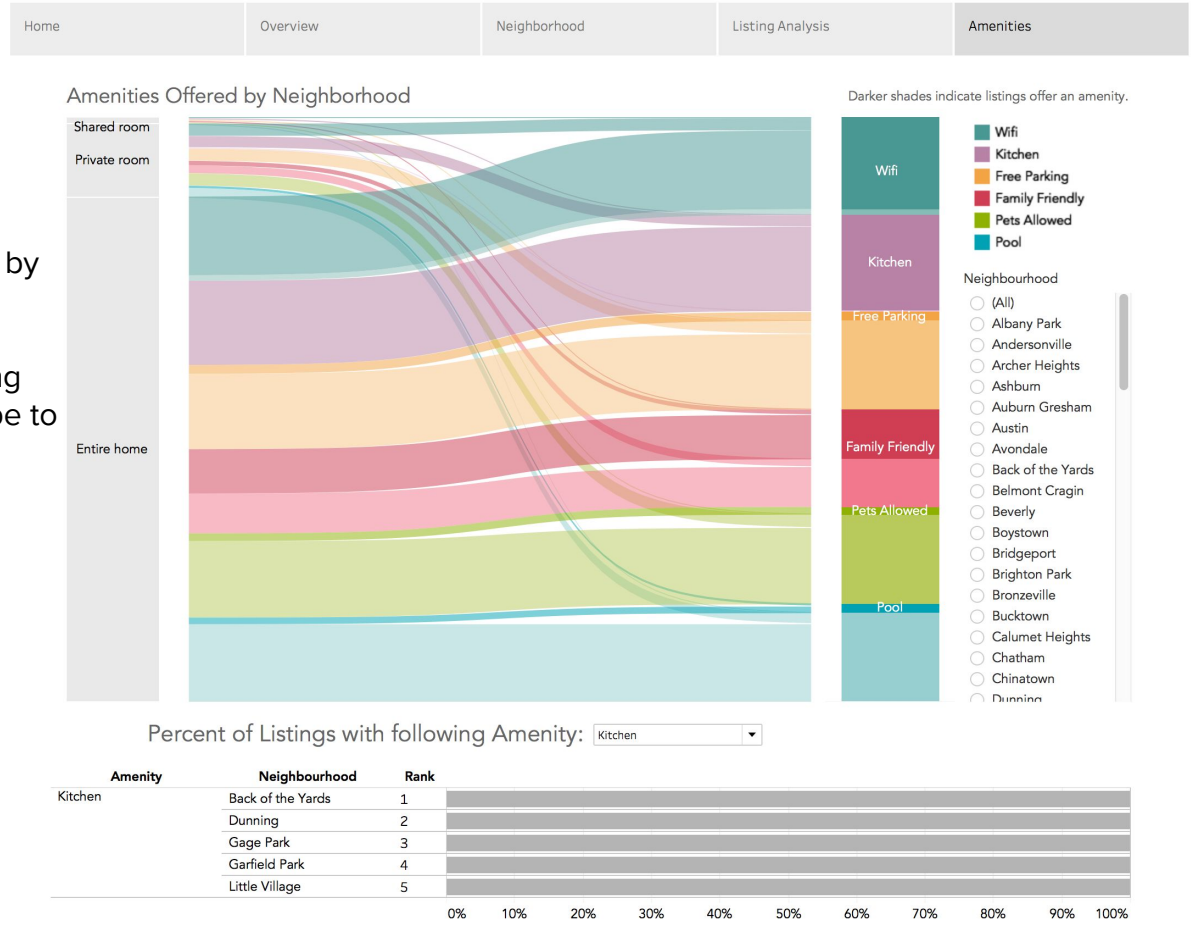
Tableau Issues

- No functionality to highlight the tooltip from the neighborhood selection menu
- Significant pre-processing done in R and Python to alter the data in the format tabeau accepts it in

Amenities

Elements of the Dashboard

- Sankey diagram shows relationship between room types and amenities by neighborhood
- Looking from left to right, each listing has 6 lines connecting the room type to each amenity category



Design Choices - Amenities

Goals

- Explore differences in amenities offered by neighborhood
- Convey relationships between categorical features

Insights

- Room type breakdown varies by neighborhood
- Stay downtown for pool access, but free parking is limited

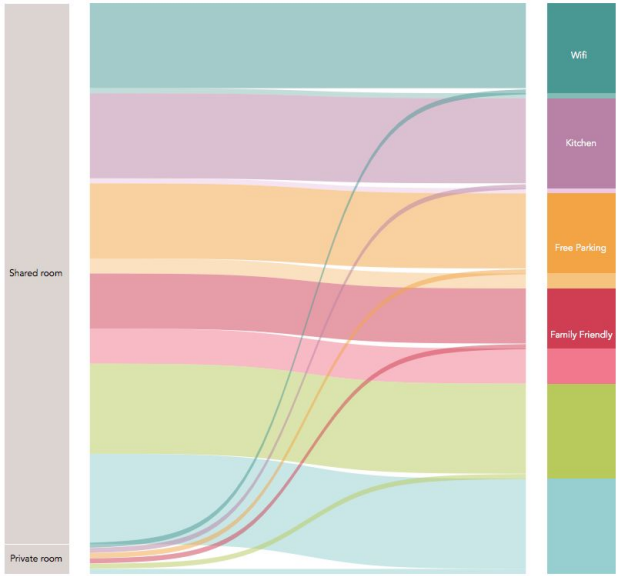
Visualization	**Sankey Diagram**
Data	Categorical
Marks & Channels	Lines Areas Color saturation Color hue Area

Visualizations - Amenities

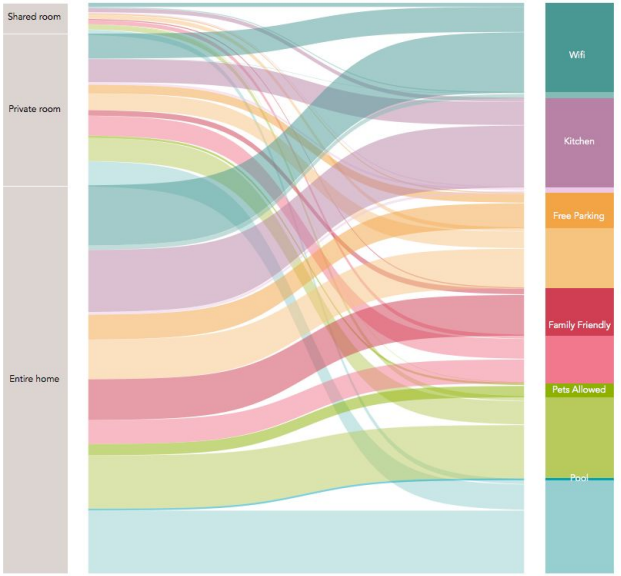
Compare room type and amenities offered in two of Chicago's poorest and wealthiest neighborhoods

Neighborhood Median Household Income

Englewood
$24,000



Lincoln Park
$83,000



Problems - Amenities

Design Issues

- Amenities category is not mutually exclusive so listings can have more than one amenity

Tableau Issues

- When no listings in a neighborhood offer an amenity, the amenity category label does not appear
- Tooltip doesn't currently show the number of listings

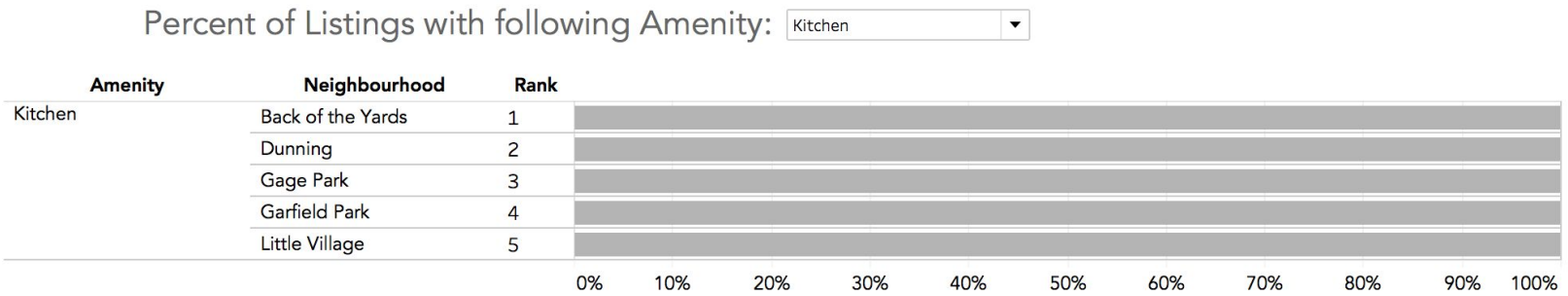
Amenities

Design Issues

- Some amenities are more common (more than 5 neighborhoods have 100% of listings offering an amenity)
- To solve this, neighborhoods must have at least 5 listings to appear in ranking

Tableau Issues

- Couldn't factor multiple amenities in ranking



Actionable Insight

- Stay downtown for pool access!
- Always mention nearby restaurants in your listings
- Airbnb is more profitable than renting*
- Offering up your entire apartment can result in significantly more earnings

* *Zumper rental prices not available for all neighborhoods*

Challenges

- Data Limitations
  - Airbnb neighborhoods did not always line up with Chicago map neighborhood labels. This required preprocessing in order to ensure that the neighborhoods would show accurately on the map we created.
  - There were limited postings for certain neighborhoods, which made it challenging to find insights across all neighborhoods.
  - We integrated with Zumper data. However, data is only available at the month level. Thus, we were able to make comparisons at the higher level, but not at the nightly average level.

Alternatives

- Alternative Designs
  - Calendar
    - One option is to view multiple months at once on the overview page to see a longer time series of information; however, we decided to limit the view to one month for easier readability on a summary screen.
- Additional Visualizations
  - Network Graph
    - One potential addition is creating a network graph among hosts and users within and across neighborhoods. This may be beneficial in understanding how neighborhoods are connected across people and listings. A host can learn from hosts with multiple listings about where to next expand.
  - Availability Data Map
    - Integrate the availability of listings from the dataset to show which neighborhoods are most in demand or where high demand is, but availability is low for potential hosts.