IDEAS: to explore relationship btw things visually (interactive/static)

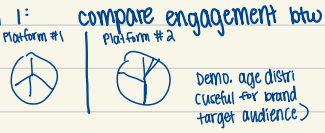
- NYC Hotel Prices: include map of hotels (hovering over each one → gives name/price/other impt info), add line graph of prices/bookings each day of week w interactive ticker that allows you to switch btw month num./location or borough?/ratings) that act like filters & another for price vs months
  - (might be easier to keep @ 1st week)
  - ↑ Bigger cities = easier to find data sets
  - City & extract Manhattan hotels/may want to cut off @ N edge of central Prk)

- Data: data scraping from Bookings.com (location/name/price/rating), City of NYC Hotel Properties Citywide (latitude/longitude for map data pnt location BUT no names only codes), dataset of Airbnb listings from insideairbnb.com

Questions to Consider:

- What're distr. of cost & rating of a std. room in diff. parts of NYC?
- How does price correlate w/ ratings?
- For each area/borough, what's the expected price for a std. room?
- Which areas are more $$$ & $?

- Social Media Dashboard: include metrics to eval. relationship btw likes/growth/comments
  - Idea 1: compare engagement btw 2 social platforms to eval. engagement

ex. 

demo. age distri (useful for brand target audience)

  - Idea 2: evaluating trends on social media w/ interactive word-cloud (to better understand the social state of a time period → relavant topics/hashtags, themes, trends)

ex. word cloud w/ features that can filter by date ranges, platforms, demographics, etc.; interaction aspect may involve clicking on spec words to see related posts; dynamic aspect can be updates from user interactions (if possible, real time data streaming)

- Data: *need to find*

- Movie Review Ratings: include diff factors of movies (genre/etc.) w/ reviewer stats (age/state/occupations) & how they influ. given movie ratings
  - Interactive Visual: interative dashboard w/ mult visualizations (bar chart - distr. of ratings, treemap - distr. of genre, bubble chart - relation. btw. ratings & num. of reviews) that users can filter through to explore diff. influences that may gain insight into movie preferences/trends for stakeholders to make informed decisions in the movie industry
  - Static Visual: heatmap of avg. ratings of movies across diff. genres (each row rep. genre & column rep ratings, color intensity rep. avg rating of movies in that genre)
- Data: Movie_Reviews.csv from GroupLens Research (Uni. of Minnesota)

Sketches below rep. potential interactive visualization ideas of idea #1

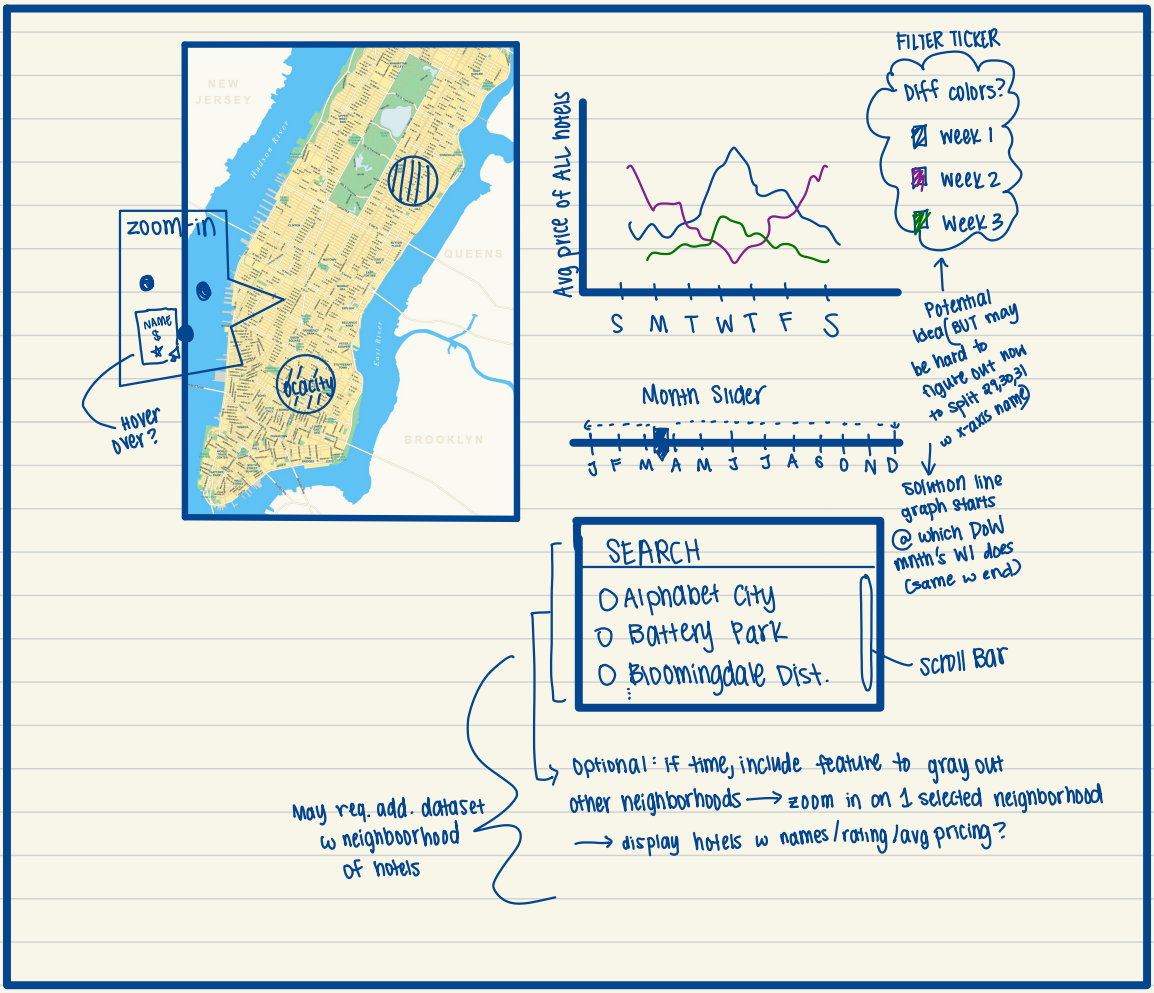
↓

SKETCH 1.

Pros: lots of ideas

Cons: too many things to focus on → next sketch: simplify

Manhattan Hotel Dashboard

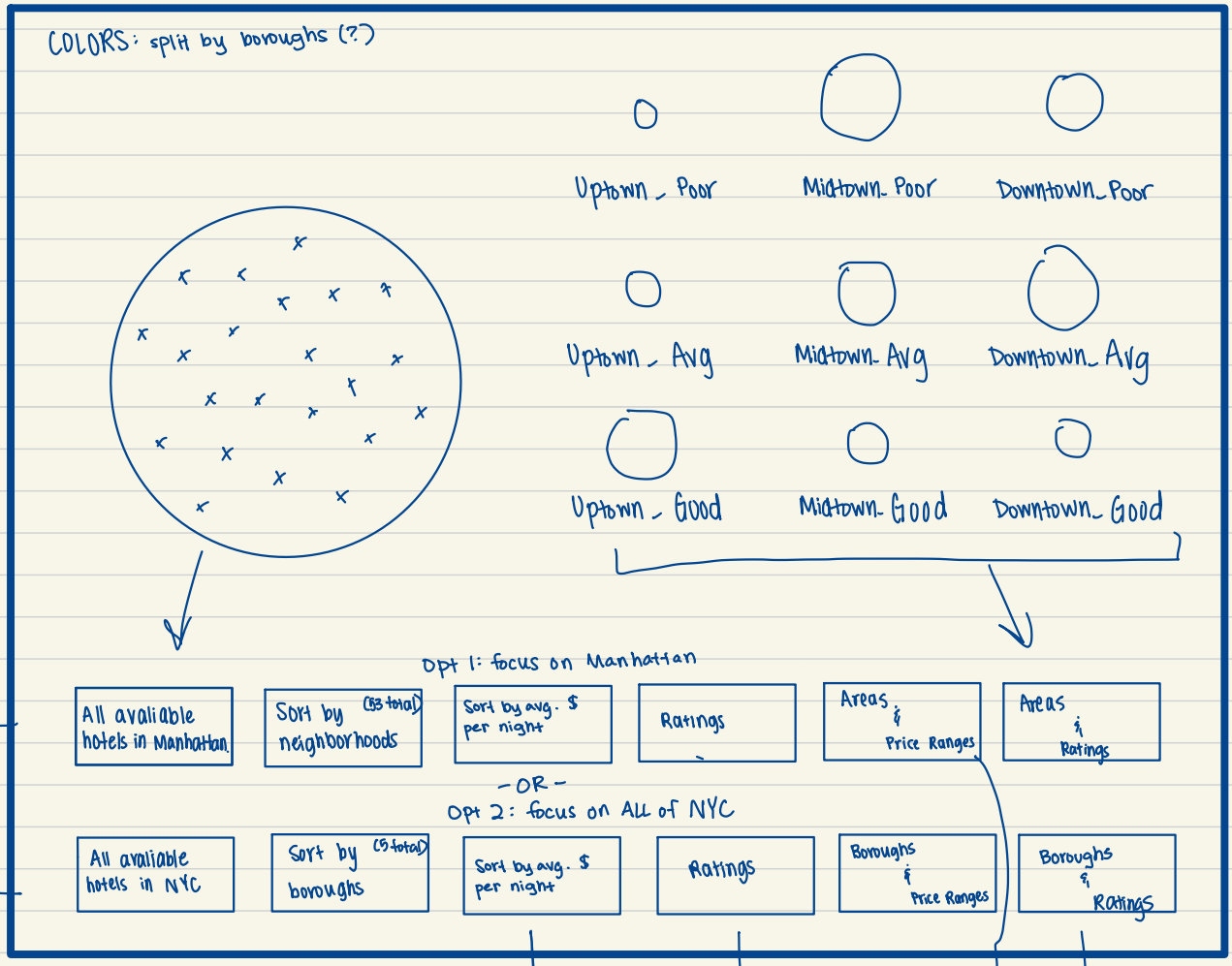


SKETCH 2.

Manhattan Hotel Availability

Pros: simple but informative

cons: splitting by areas/boroughs may seem clustered (stylistic / "aesthetics"),
finding dataset (w/o datascraping & heavy manipulations) → next sketch: Airbnbs



Need to convert $ to qual. (turn $ to ranges/intervals)

Convert to ranges/intervals
OR
round ratings to deal w/ decimals

Ratings classified into Poor, Avg., Good
(0-3.3) (3.4-6.7) (6.8-10)
or
(0-4.9) (5-7.9) (8-10)

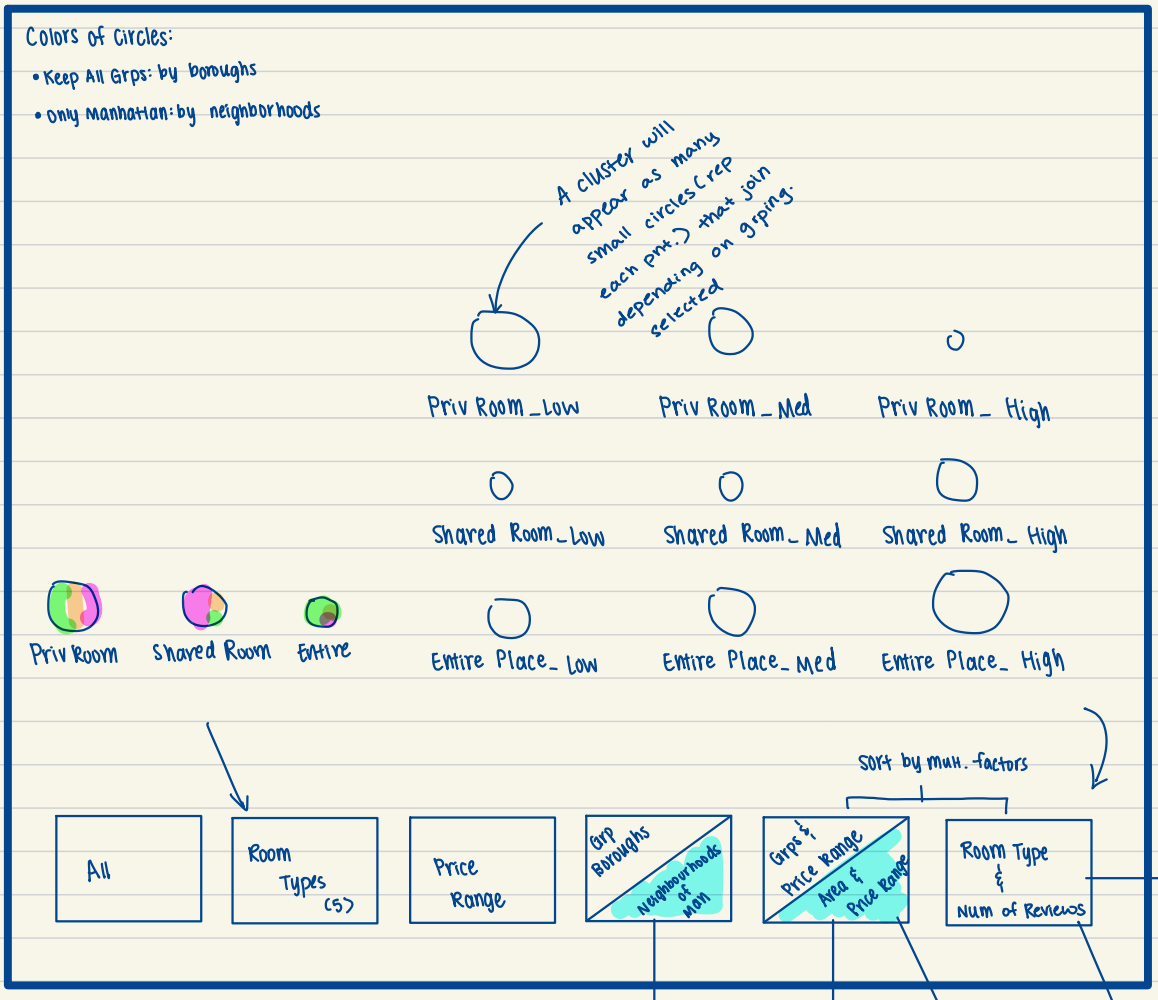
What: each individual circle rep. a single hotel

Additional features: zoom-in, display hotel name/price/rating when hovering over

Areas: downtown, midtown, uptown, (include what each borough is w/ some data manipulation)

SKETCH 3.

NYC / Manhattan Airbnb Listings



Note:

- Airbnb data includes Manhattan, Brooklyn, & Queens (w/ borough/neighborhoods)
- Some prices have min. num. of nights (thus 10K Airbnb)
- 11 observations have price of 0 bc they're temp. marked unavailable for booking/inactive (adjusting for bad reviews/etc.); solutions can be choose to remove listings from visual / change to nan → Can explore these w/ num. of total reviews
- When looking @ host_id, you notice some hosts have hundreds of listings in NYC
- Since there's no 'num. of total bookings' for a listing → use 'num. of total reviews' to indicate popularity