Title: Boston Unveiled

Abstract: Our project, "Boston Unveiled," stems from our connection to the city of Boston, with two-thirds of our team proudly calling it home. As residents, we are motivated to create a data-driven guide that captures the essence of this city. Our objective is to develop an interactive visualization platform that offers insights into Boston's key aspects, including sports, transportation, politics, entertainment, and culinary diversity. We aim to explore and present Boston's diversity through a lens of data analytics and visual storytelling, believing that by combining our personal experiences with data analysis, we can offer both residents and visitors a unique perspective on what makes Boston truly special. To achieve this, we will leverage many public datasets. Our analysis will use recent sports milestones such as the Celtics' 2024 NBA championship, voting patterns and city budget allocations from the most recent fiscal year (FY 2023), transportation data from the MBTA and reviews of top entertainment venues, and demographics of diverse culinary offerings. We will primarily use data sources including, but not limited to the City of Boston's open data portal, the MBTA database, sports analytics websites, user-generated content from food and entertainment review platforms, census data and more. Ultimately, we hope to provide a clear, data-driven understanding of Boston's city life and offer valuable insights for both long-time residents and newcomers.

<u>Team Agreement & Detailed Project Plan (28th October, 2024)</u>

<u>Group Name:</u> The Boston Trio

<u>Team Members:</u>

Zoe Price- zprice@college.harvard.edu

Cameron Hosein- cameronhosein@college.harvard.edu

Jada Pierre- jadapierre@college.harvard.edu

<u>Communication Protocol</u>

Team Meetings: Zoom for virtual meetings, in-person meetings at Lowell Study Room

Document Sharing: Google Drive shared folder

Primary Means of Communication: Text Message

<u>Regular Team Meetings:</u> Tuesdays and Thursdays, 4:00-9:00 PM EST

<u>Work Distribution:</u>

<u>Cameron (Sports & Entertainment Focus)</u>

- Sports data collection and analysis

- Celtics' 2024 NBA championship data

- Historical sports milestones

- Team performance metrics

- Entertainment venue analysis

- Venue locations and ratings

- Event attendance data

- Popular entertainment districts

- Tasks- sports visualizations, entertainment venue map, interactive timeline of sports events, event venue dashboard

Zoe (Transportation & Infrastructure Focus)

- MBTA transportation data

- Transit routes and schedules

- Ridership patterns

- Service reliability metrics

- City infrastructure data

- Bike lanes and pedestrian areas

- Traffic patterns

- Tasks- visualization on the T, interactive maps, transportation dashboard, infrastructure vis.

Jada (Politics, Culture & Culinary Focus)

- Political and budget data

- FY 2023 city budget allocations

- Voting patterns

- Neighborhood demographics

- Culinary scene analysis

- Restaurant demographics

- Cuisine diversity metrics

- Food district mapping

- Tasks- budget allocation vis, demographic distribution maps, restaurant diversity dashboard, cultural hub visualization

Shared Tasks- integration of individual components, Overall design consistency, documentation, testing and final presentation preparation

Signed:

Cameron Hosein

Zoe Price

Jada Pierre

Background and Motivation:

Our project emerges from our unique connection to Boston, with both Zoe and Jada being long-term residents and Cameron discovering the city as a newcomer. This diverse perspective—combining insider knowledge with fresh eyes—sparked our interest in creating a comprehensive, data-driven exploration of Boston's character.

Living in Boston has shown us how the city's various facets—sports culture, transit systems, political engagement, entertainment scenes, and culinary diversity—interweave to create its unique identity. However, we noticed that while there are many resources about individual aspects of Boston, there isn't a unified, interactive platform that brings these elements together through data visualization.

Related Work:

Our project draws inspiration from several sources:

Visualizing MBTA Data: http://mbtaviz.github.io/
Citymapper Boston: https://citymapper.com/boston

Living Portrait of NYC:

https://www.behance.net/gallery/164782513/Citizens-The-Living-Portrait-of-NYC?locale=en_US

NBA Data Vis: https://buckets.peterbeshai.com/app/#/playerView/201935_2015

Our data will come from multiple sources:

Sports- nba game statistics and performance metrics from NBA API, sports venue attendance from team official websites

Transportation- MBTA ridership data, blue bikes systems data

Politics- City of Boston Open data, FY 2023 budget allocations, Voting patterns from recent elections, Neighborhood demographic data from Census API

Entertainment and Culinary Data- Yelp API for Restaurant and Venue data, cultural venue info, local festival and event data

Data Cleanup

We anticipate several data cleanup challenges and have planned our approach to do this all by next week once our data sets are sourced:

We plan to calculate several key metrics from our raw data:

Sports engagement index: combining attendance rates, social media mentions, and game viewership

Transit accessibility score: measuring average commute times and service reliability across neighborhoods

Culinary diversity metric: ratio of cuisine types weighted by restaurant ratings

Neighborhood vibrancy index: combining entertainment venues, foot traffic, and event frequencies

To minimize cleanup, we are primarily using:

MBTA's preprocessed API data for transit information

NBA's official statistics API for sports data

Census Bureau's cleaned demographic datasets

Yelp's structured API for restaurant and entertainment data

MAP

Potential Audience:

1) Boston newcomers (i.e. students, tourists, young professionals who moved to Boston)

2) Long term Boston residents looking to explore the city even further

3) City planners and policymakers

4) Sports enthusiasts

5) Cultural explorers

Target Audience Selected: Boston newcomers

Why? Our team consists of both Boston locals as well as a newcomer so we would be able to combine our thoughts together into a visualization. The locals can suggest what someone must do and see and Boston while the newcomer can also state what an outsider heard about and would be interested in pursuing. The comprehensive nature of the data also suits those needing a complete city overview. Lastly, Boston is a college town so there is a constant influx of students and young professionals so this would be a perfect visualization for them.

We assume that our target audience has basic understanding of city concepts and also limited familiarity with Boston's geography and transport systems. They should also be comfortable with standard data visualizations such as maps and charts, comfortable with interactive interfaces, familiar with map-based navigation and be able to decipher more complex charts with proper annotations and instructions (they are also going to college so they should be educated,

we hope). They may not be aware of Boston-specific terminology such as the "T" but should be familiar once they view our visualization.

The interests of our target audience would range from practical information for daily life such as transport and food to entertainment and social scenes. They would also be interested in cost of living and neighborhood characteristics as well as the cultural experiences.

The detail of the presentation would be more than basic tourist information that can be easily googled but less than resident expertise. It would include clear explanations of Boston-specific terms, interactive elements and content with the use of tooltips and annotations.

Questions we considered:

1. Which neighborhoods offer the best MBTA access for daily commuting?

2. What are the most reliable T lines during peak hours?

3. How do commute times vary between different neighborhoods?

4. Where are the best-connected areas for someone without a car?

5. Which neighborhoods have the highest concentration of entertainment venues?

6. How do ticket prices vary across different sports venues?

7. What are the most popular entertainment districts for young professionals?

8. How does event attendance correlate with MBTA ridership?

9. Which areas offer the most diverse dining options?

10. How does the price range of restaurants vary by neighborhood?

11. What neighborhoods have the highest concentration of student-friendly dining options?

12. Where are the cultural hubs relative to public transportation?

13. Which neighborhoods offer the best balance of entertainment options and affordability?

14. How do entertainment options correlate with neighborhood demographics?

15. What areas have the highest concentration of young professionals?

Datasets:

-[player data](#)

//-[general nba data](#)

-[entertainment venues data](#)

-[Special Event License Applications](#)

-[Women-Owned Businesses](#)

-[MBTA Bus Ridership by Trip, Season, Route/Line, and Stop](#)

-[MBTA Rail Ridership by Time Period, Season, Route/Line, and Stop](#)

-[MBTA Bus, Commuter Rail, & Rapid Transit Reliability](#)
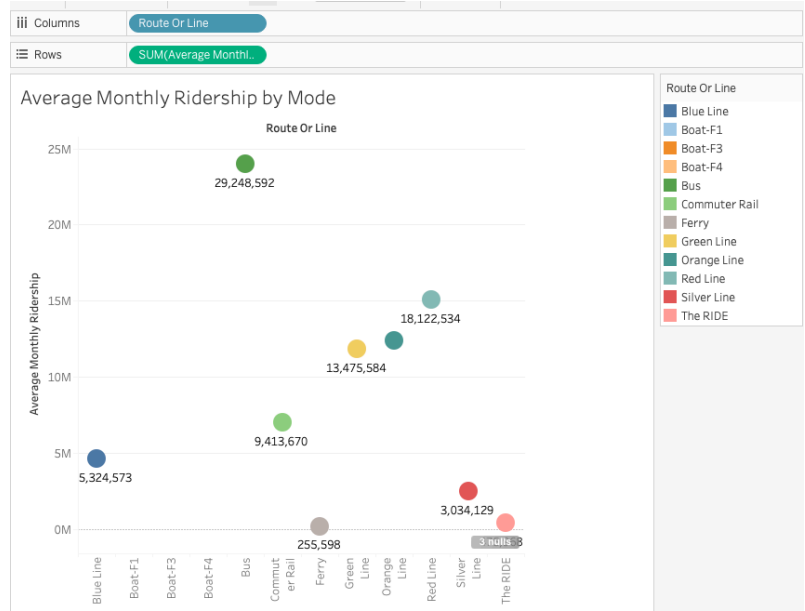
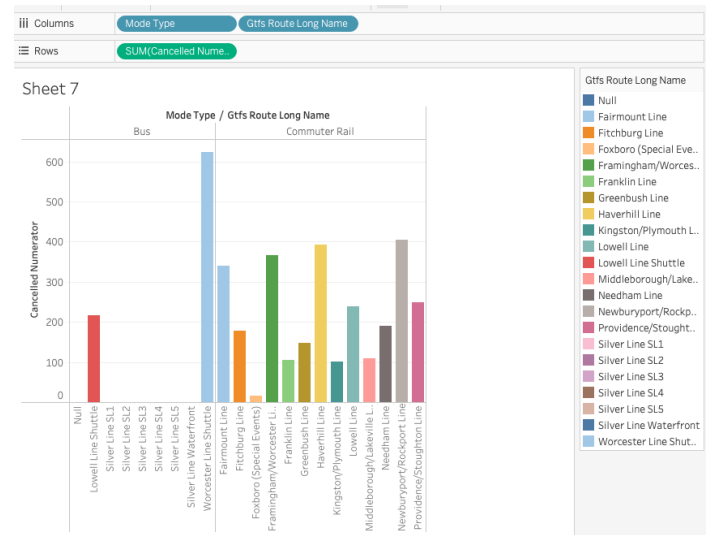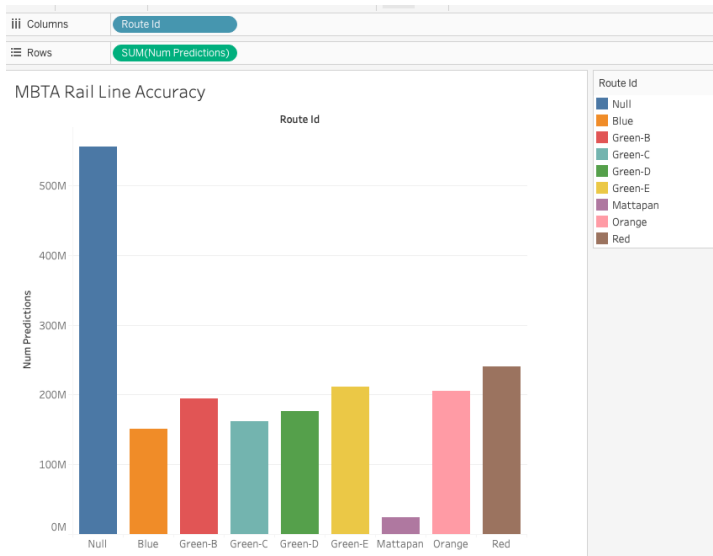-[MBTA: Fuel and Energy](#)

-[MBTA: Gated Station Entries](#)

-[Rapid Transit and Bus Prediction Accuracy Data](#)

Potential additional data:

- Political and budget data

- FY 2023 city budget allocations

- Voting patterns

- Neighborhood demographics

- Culinary scene analysis

- Restaurant demographics

- Cuisine diversity metrics

- Food district mapping

Tableau Visualizations
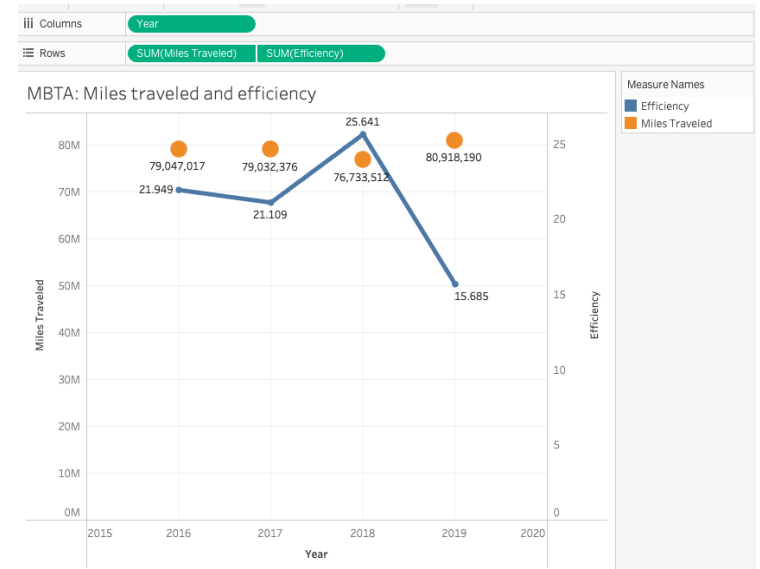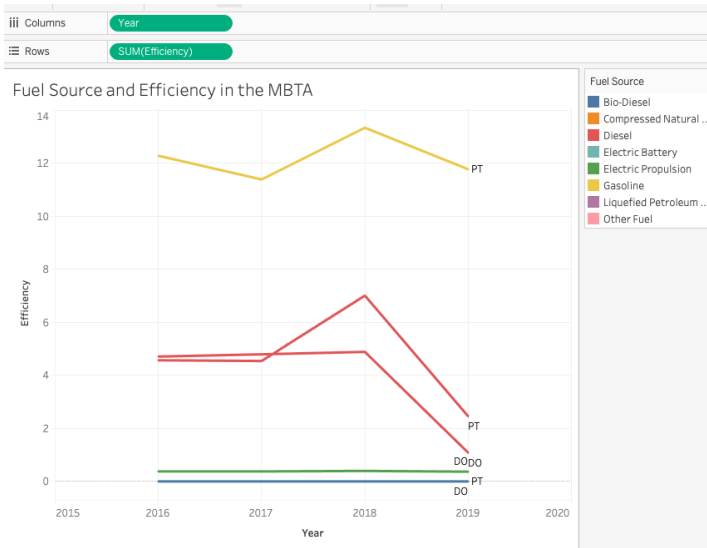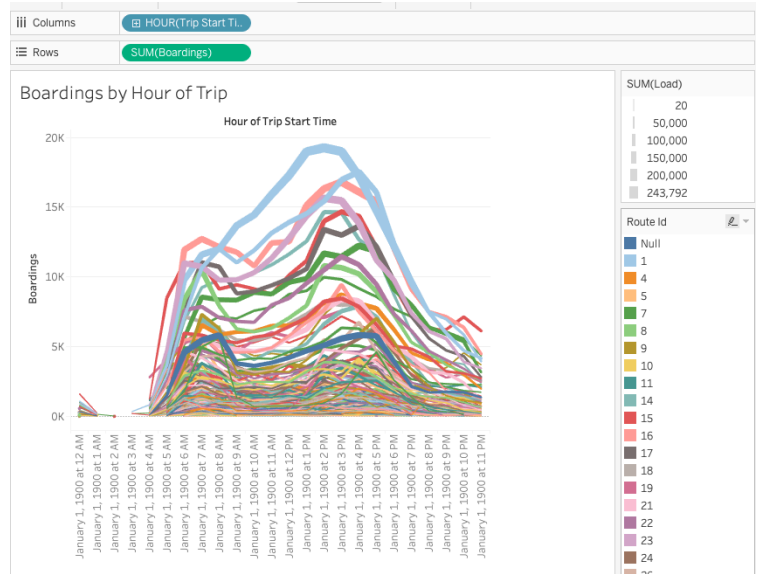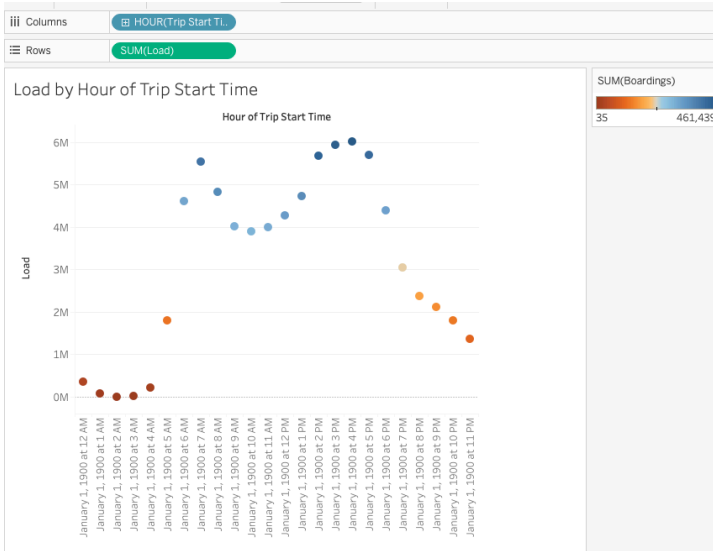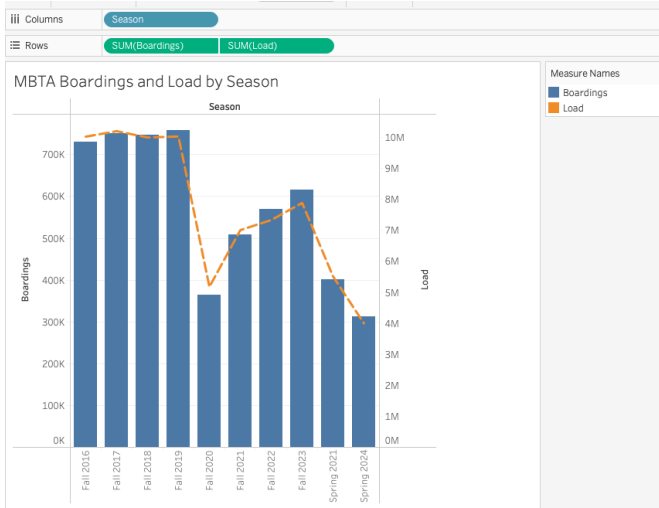
The questions we answered in Tableau differed slightly from the original team questions due to the constraints of the data we only initially found (we've found more as of week 10) and the Tableau platform. For example, while our team initially brainstormed exploratory questions like "Which neighborhoods offer the best MBTA access for daily commuting?" and "What are the most reliable T lines during peak hours?", we found that some questions were more actionable or visualizable in Tableau, such as "How do commute times vary between neighborhoods?" and "Where are the best-connected areas for someone without a car?". In our Tableau visualizations,

it is easier to see the peak MBTA times and patterns around the city regarding ridership, such as which lines have the most reliability (Red and Green-E lines).
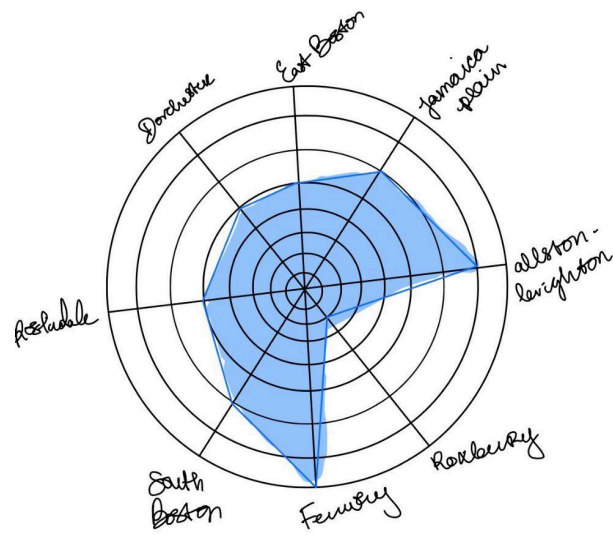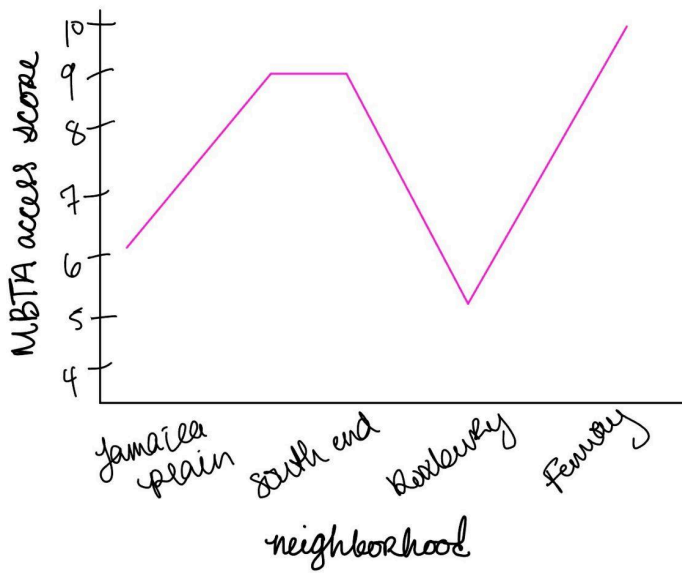

Sketches

Jada's Sketches
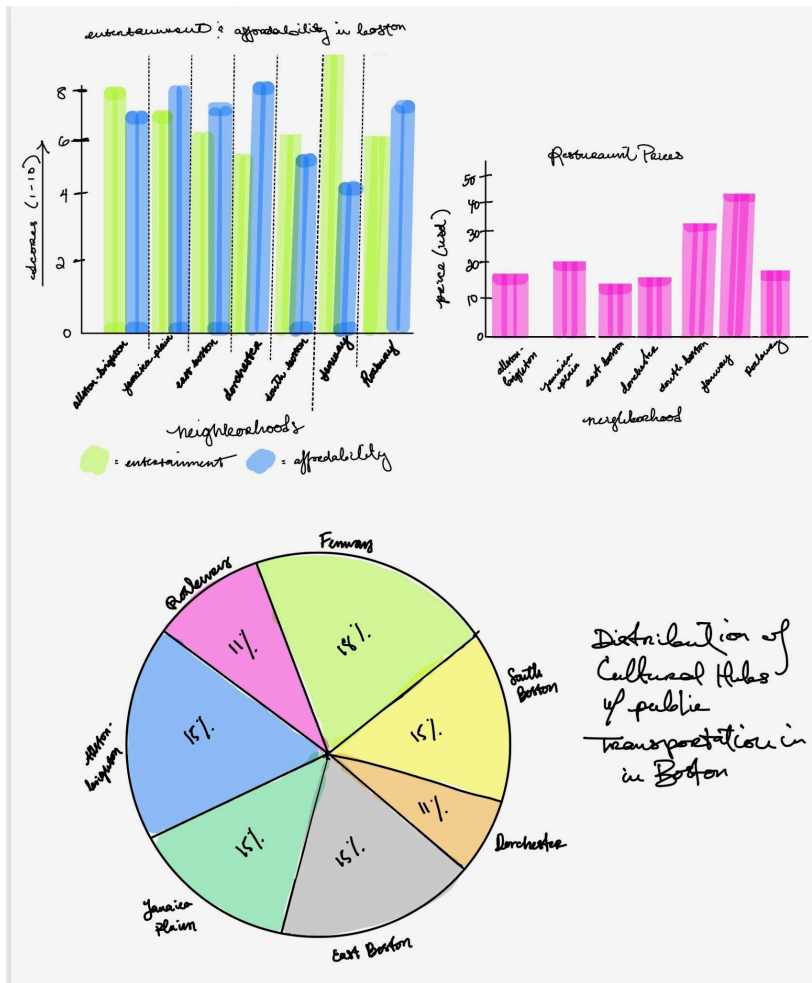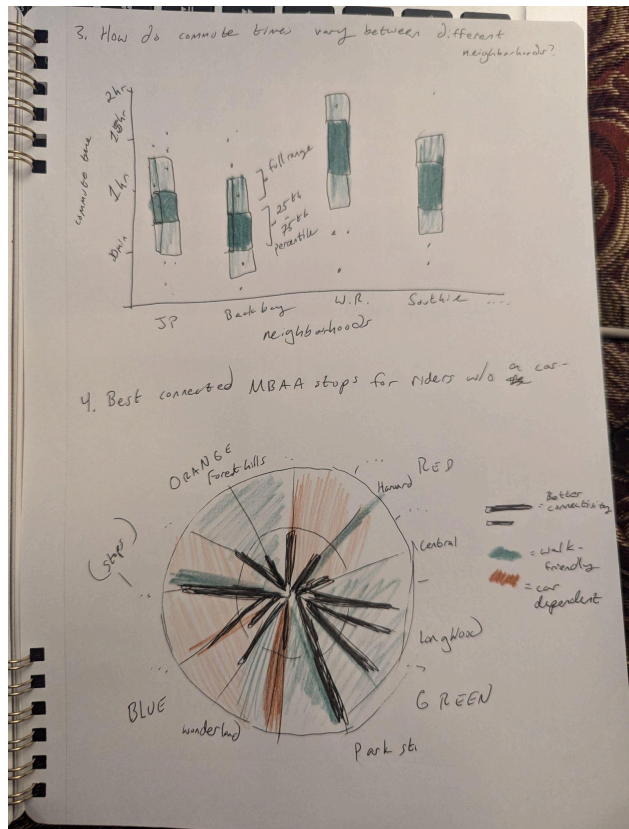

Sketch 1D: JP 1, JP2 and JP3 respectively
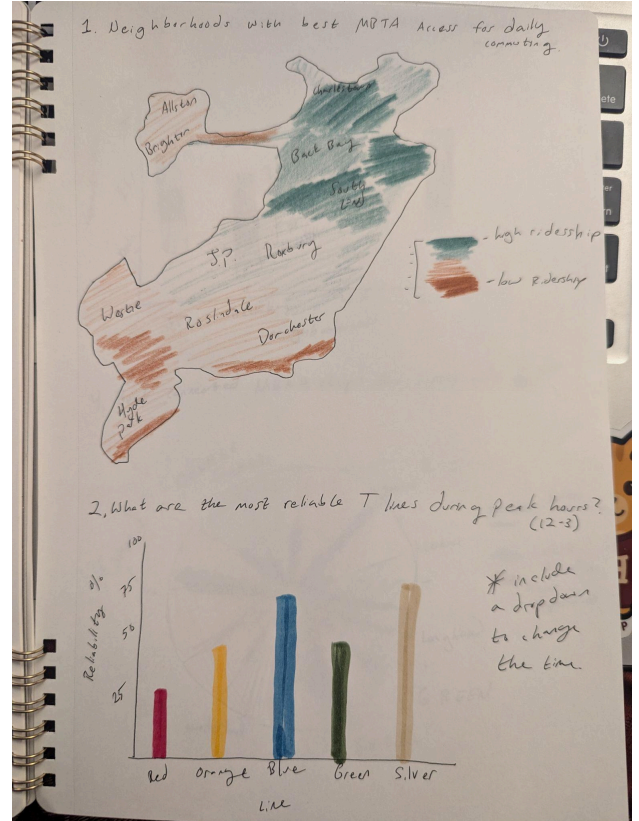
concentration of young professionals

mbta access for daily commuting

Sketch ID: JP4 and JP5 respectively

Zoe's Sketches

Sketch ID: ZP1, ZP2 respectively

Sketch ID: ZP3 and ZP4 respectively

Sketch ID: ZP5

Cameron's Sketches

1) Dining Options & Price Range. (Question # 10)

Cluster Showing Boston Restaurants Avg Price Range

Tooltip-
shows name &
category + price
Sort - sorts by type of
Restaurant - y Italian etc.

can encode
color - expense
size - popularity

☐ — cheap
☐ — middle tier
☐ - very expensive
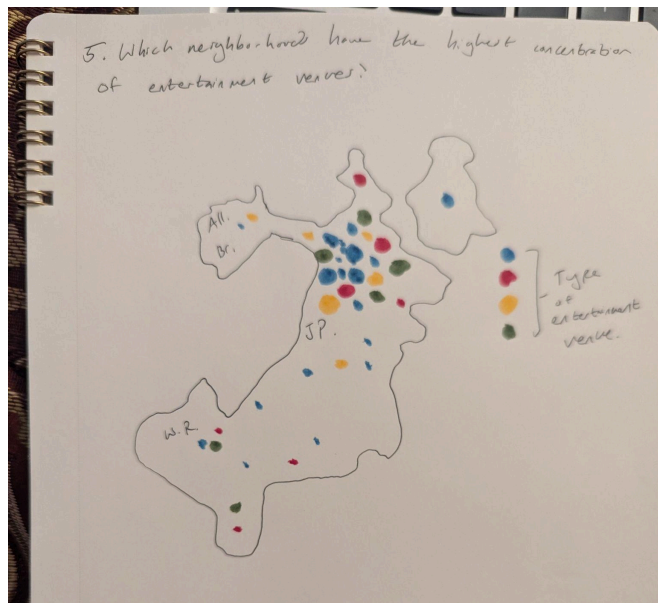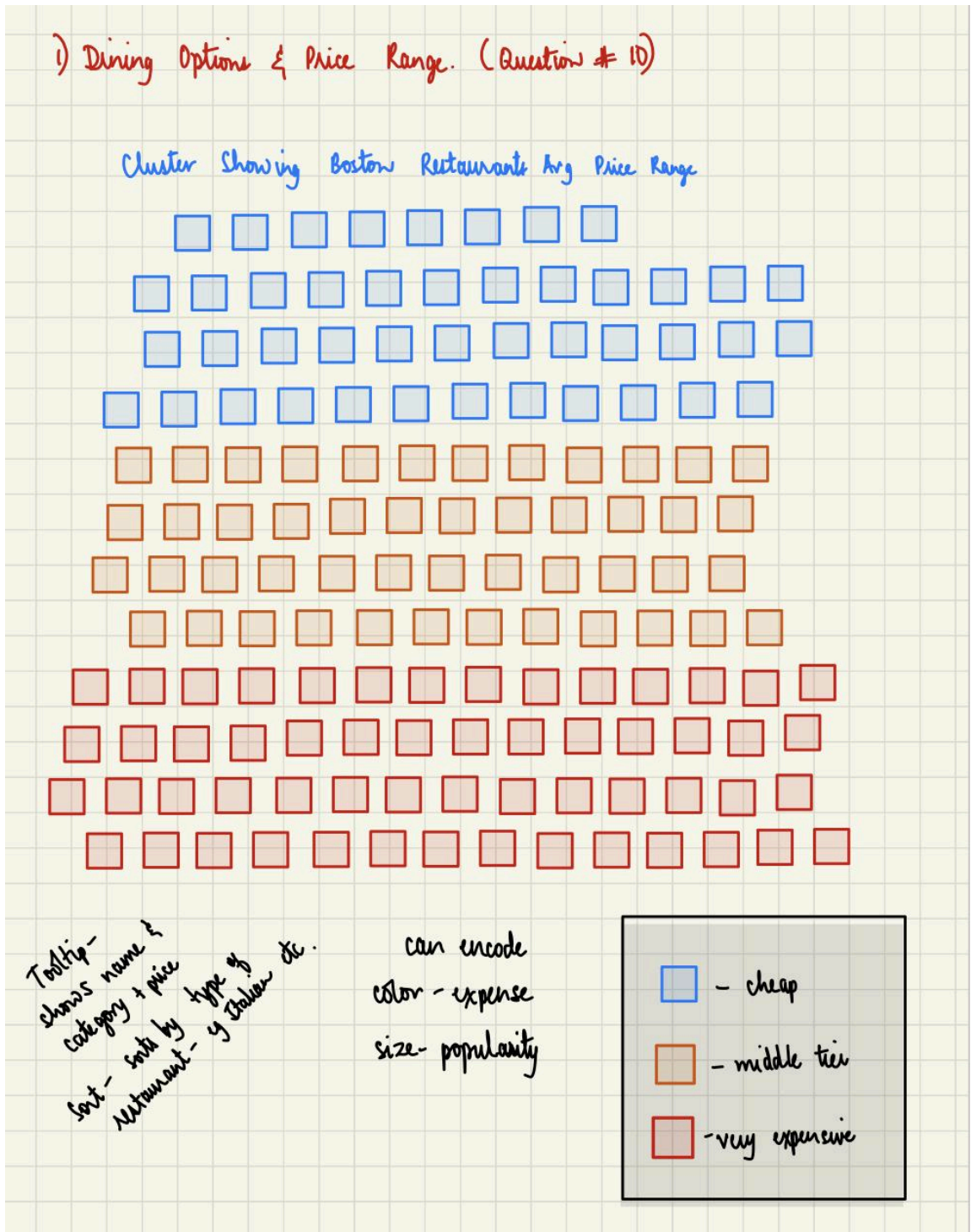
Sketch ID: CH1

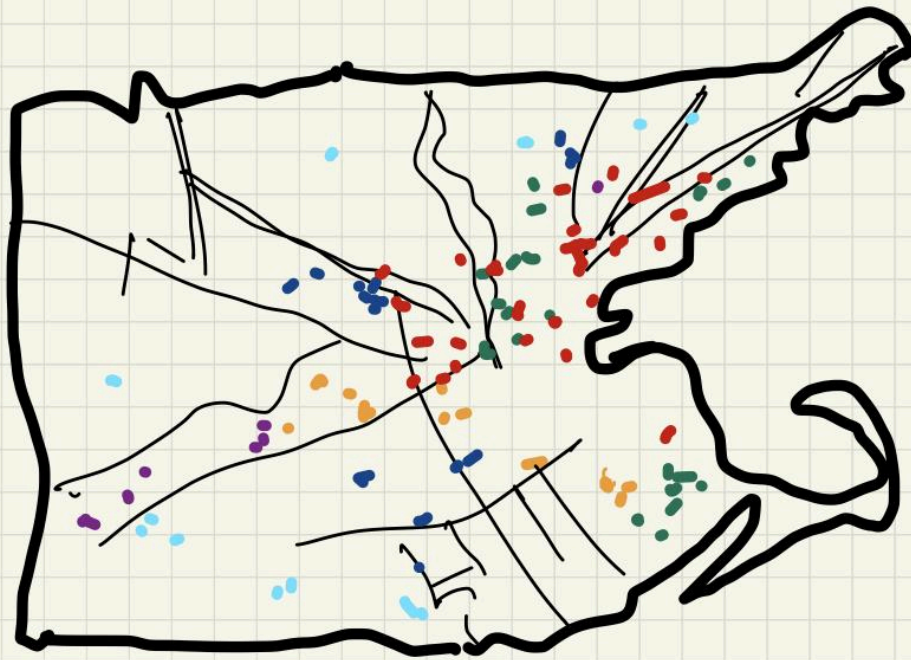2) Best Connected Areas for people without a car – (Question # 4)

Each colored dot represents different encodings

🔴 red line   🔵 blue line   🟢 green line

🟣 commuter rail   🔵 walking   🟠 bike
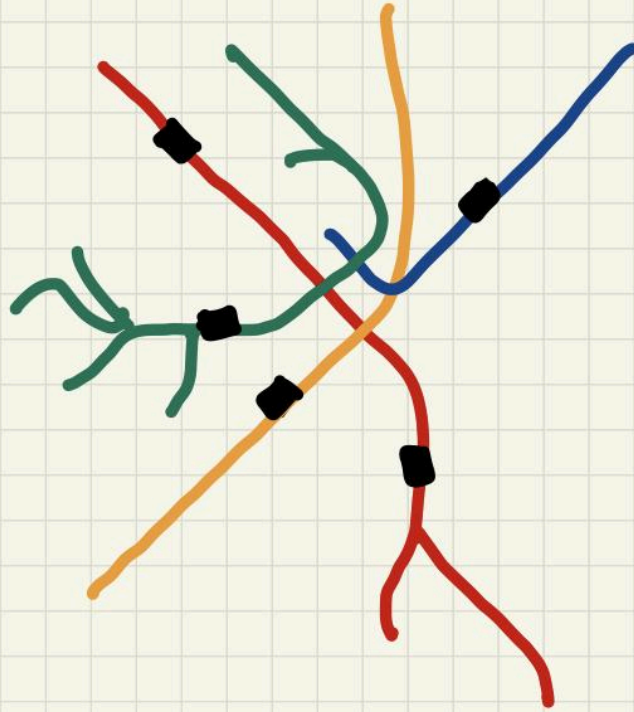
higher density – best connectivity

lower density – poor connectivity

Trend: Area closest to city has more connectivity.

Sketch ID: CH2

3) mini Map of Boston T
(Questions 1-3)

we can have mini moving trains that show peak ridership at certain hours

Tooltip to show how many people are there waiting

filter by line on T

show wait times w/ tooltip

Sketch ID: CH3

4) Ticket Prices Sports Venues (# 6)

Price

Time period

each color is
a different venue

Sketch ID: CH4

5) most Reliable MBTA (based on delays) # 2

delays

Red   Green   Blue   Orange

T line

Sketch ID: CH5

Decide

Sketch ID	Question ID	Author
JP1	5	JP
JP2	10	JP
JP3	12	JP
JP4	15	JP
JP5	1	JP
ZP1	3	ZP
ZP2	4	ZP
ZP4, CH5 (identical)	1,2	ZP, CH
ZP3	2	ZP
ZP5	5	ZP
CH1	10	CH
CH2	4	CH
CH3	1,2,3	CH
CH4	6	CH

Screenshots of Chosen Visualizations

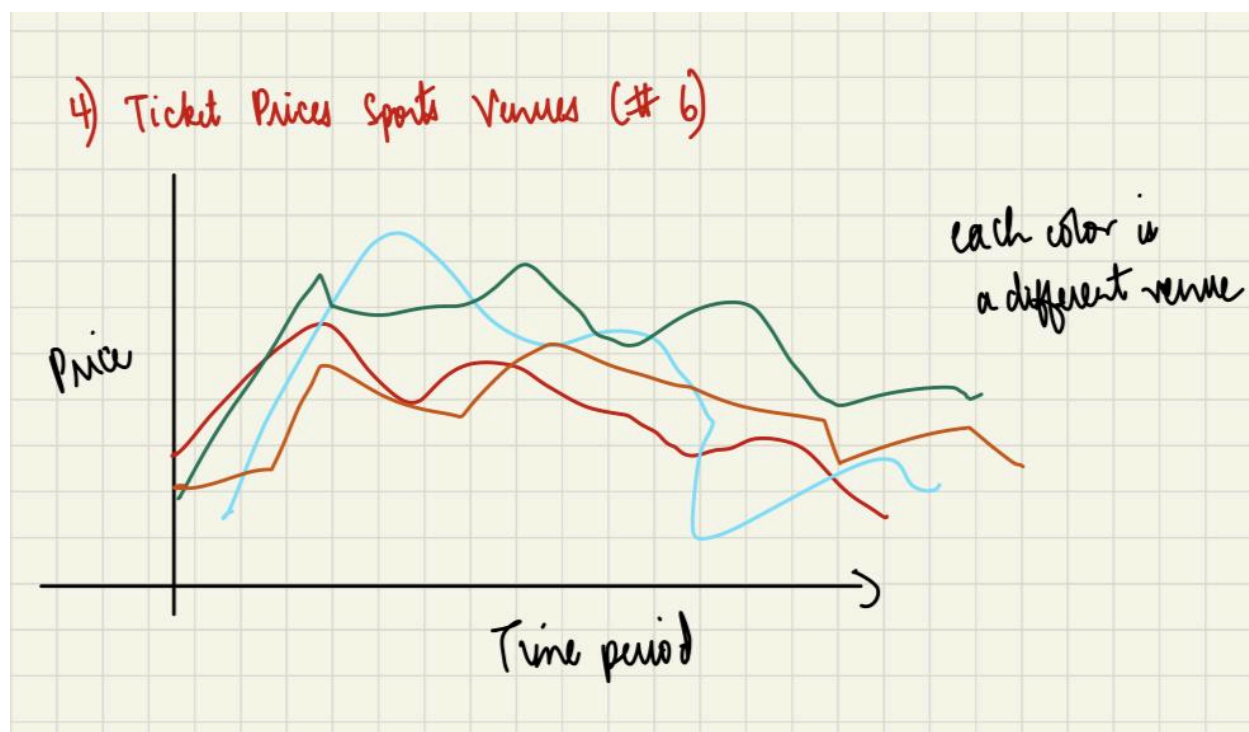3) Mini Map of Boston T
(Question 1-3)

we can have mini moving trains that show peak ridership at certain hours

Tooltip to show how many people are there waiting
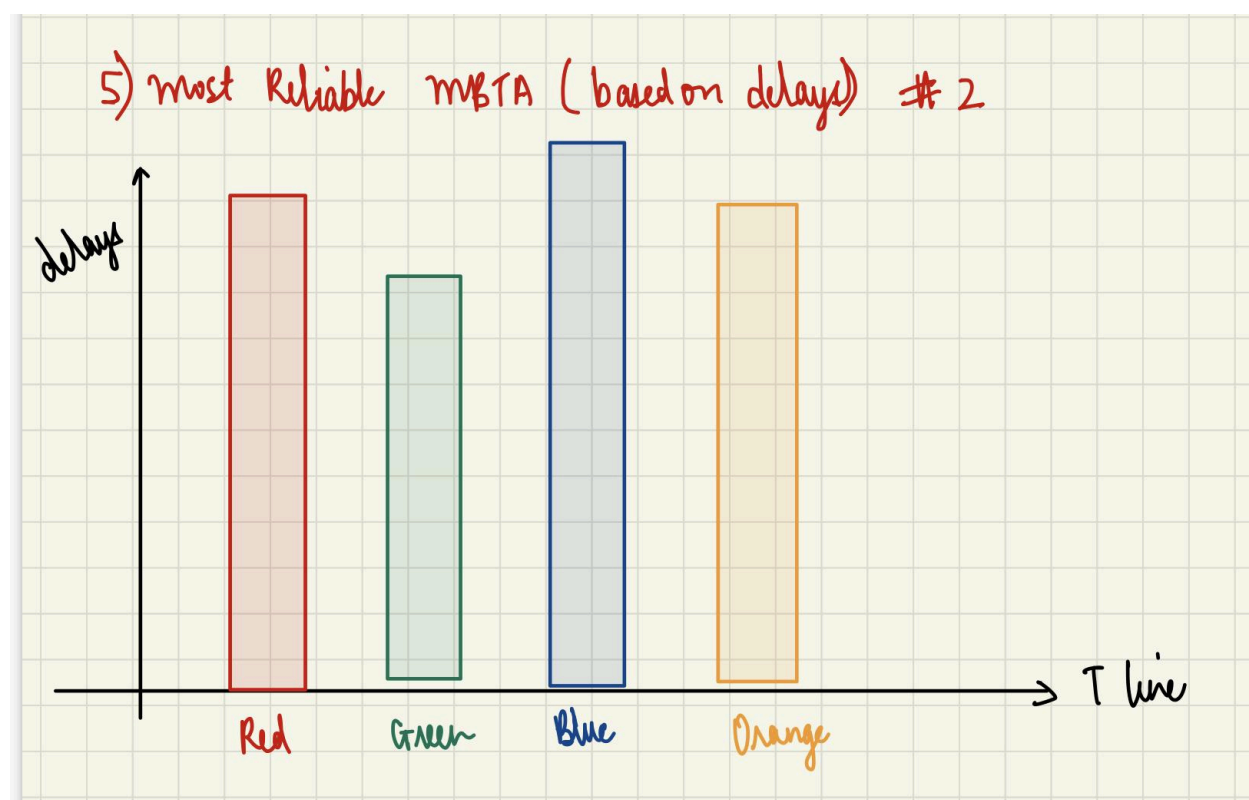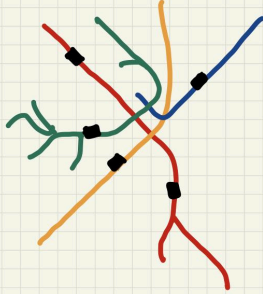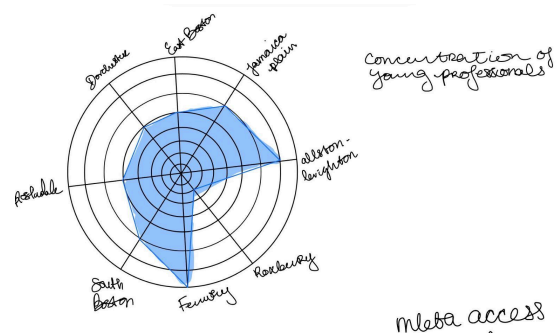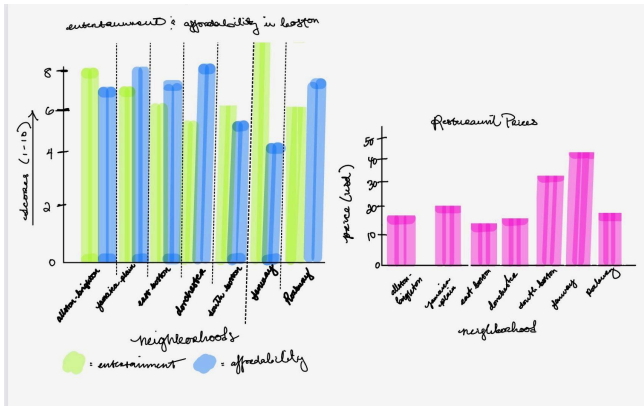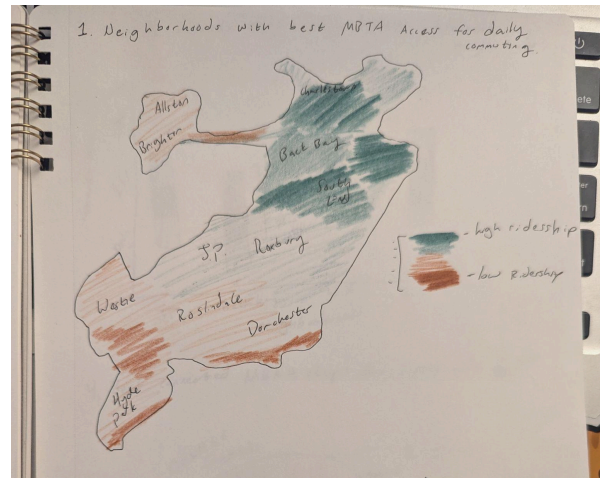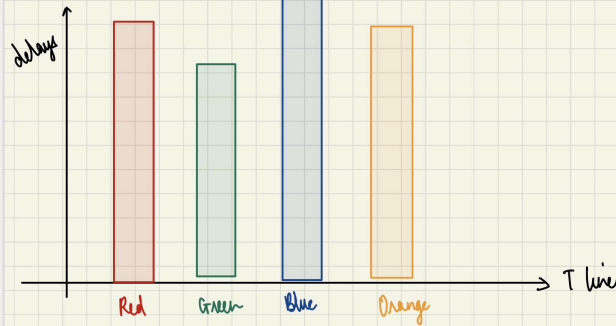
filter by line on T

show wait times w/ tooltip

5) Most Reliable MBTA (based on delays) #2

delays

Red   Green   Blue   Orange

→ T line

entertainment & affordability in boston

evenness (1-10)

neighborhoods

: entertainment   : affordability

Restaurant Prices

price (usd)

neighborhood

1. Neighborhoods with best MBTA Access for daily commuting

Allston Brighton   Charleston   Back Bay   Darby Line   J.P   Roxbury   Westie   Roslindale   Dorchester   Hyde Park

- high ridership
- low Ridership

Dorchester   East Boston   Jamaica plain   allston-brighton   Roxbury   Fenway   South Boston   Roslindale

concentration of young professionals

mbta access

1) Dining Options & Price Range. (Question # 10)

Cluster Showing Boston Restaurants Avg Price Range

Tooltip shows name & category & price

Sort - each by type of restaurant - by cuisine etc.

can encode
color - expense
size - popularity

- cheap
- middle tier
- very expensive
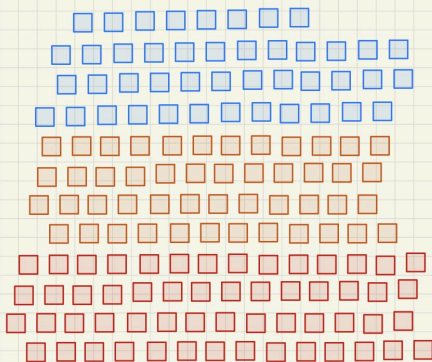
Voting Explanation

We first chose the top three MBTA-related visualizations, because they answer our essential questions: how reliable is T ridership in different neighborhoods, how accessible is the T throughout Boston neighborhoods, and how do delays impact ridership. We will then have a section on the Boston professional social environment of different neighborhoods, in our visualizations depicting the distribution of young professionals in the city and the visualization of affordability in spaces in different neighborhoods. Finally, in the entertainment section of our project, we will include the visualization on food affordability throughout the city.

StoryBoard

Hook: Our main hook is gonna be an interactive city map on the first page with multiple layers (transit, entertainment, sports and food) which allows the viewer to pick their journey.

The narrative for this project would be: Explore Boston, a historical yet innovative city where championship Celtics parades meet student influx and every neighborhood tells a unique story.

Insights

Transportation: MBTA system reliability across all lines, peak usage patterns visualization, neighborhood connectivity scores, bike lanes. (Zoe)

Culture: young professionals distribution map, neighborhood affordability, cultural locations and community demographics. (Cameron)

Entertainment/ Sports: venue locations, Celtics 2024 championships, entertainment district analysis (Jada)

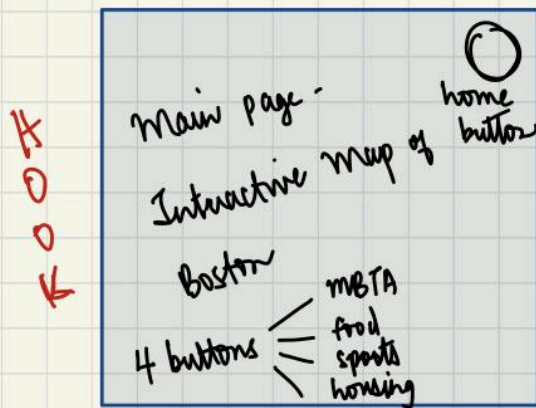**Civic Engagement:** budget patterns, voting engagement levels (Jada)

Main Insight: We could not pick one particular aspect so we decided to combine them. Boston's diverse city offers a unique combination of accessibility, culture and lifestyle– from sports near Fenway to mouthwatering culinary experiences at North End, and the rapid transit system (T) makes it quite easy to move between them.

We chose to combine Boston's diverse elements of transportation, culture, entertainment, and civic engagement as our main insight because it authentically captures the interconnected nature of city life that newcomers need to understand. Our unique team composition - with both Boston locals and a newcomer perspective - helped us recognize that focusing on just one aspect would miss the crucial interplay between what makes Boston special. Through our data analysis, we discovered that the most valuable insights emerge not from isolated aspects but from their intersection: how the T connects sports fans to Celtics games, how young professionals cluster around areas with both strong transit access and diverse entertainment options, and how neighborhood character is shaped by the mix of cultural venues, dining options, and community engagement. This comprehensive view aligns perfectly with our target audience of newcomers, who need to understand not just individual aspects of Boston but how they work together to create livable, exciting neighborhoods. The data supports this approach, showing strong correlations between transit accessibility, entertainment options, and neighborhood vitality, making it clear that Boston's true character emerges from the synthesis of these elements rather than any single factor.

Storyboard.

HOOK

Main page -
Interactive Map of
Boston
home button

4 buttons — MBTA
food
sports
housing

Narrative:
emphasis on
appeal of city to
newcomers

click on each to show
scrollytelling vis for
each of the 4

rising

rising

rising

rising

MBTA

interactive map

scroll to see
peak-times, delays,
cost etc

Sports / culture

Celtics vis

ticket prices etc

food

Restaurant vis

hover for places

Housing/neighborhood

voter patterns.

— neighborhood patterns

Solutions: **Call to action**

— encourage exploration,
sieze the opportunity
— try new places
— use our website

Implementation Week 1

So for the interactive parts of our project, so far we planned out the transportation aspect.

The first visualization would be a working and moving map of the subway system that just gives

basic information about the subway system in Boston. There would be moving trains and also a

hover effect on various stops along the 4 main lines of the subway system. After you scroll

down, we will have the graph of the delays that draws itself on the page, similar to the FIFA

homework. The message here will be along the lines of, "While Boston has an extensive subway

system, look out for delays in the train!" This will switch back and forth between a bar chart of

the most reliable. Then we will have a visualization of the outline of the trains filled as a

proportion with peak times with a hover that shows how many people are in the train at that time.

Lastly, we will have another option of bikes (Similar to the lab) and walking.

<u>TO DO:</u>

-clean up mbta

-change bar chart to lollipop

-sports vis

-heatmap