

Process Book
Michelle Danoff, Maria Lovett, Mirhee Kim.
City Cycle

Preproposal work:

As our preproposal discusses, our team was initially drawn to the concept of biking data due to our interest in environmentalism and sustainable living. In New York, where Mirhee and Emily are from, bike shares completely changed the nature of city living http://www.huffingtonpost.com/jared-green/bike-share-is-growing-up_b_6188212.html , and bike shares are beginning this year in Philadelphia, where Michelle is from http://articles.philly.com/2015-02-12/news/59048312_1_bike-sharing-program-bikes-and-docking-stations-independence-blue-cross.

We wanted to create a visualization that would help these bike share companies understand their own businesses and make smart business choices. Since many of these companies manually move bikes around during the day, understanding the bike traffic patterns better would help them know the most efficient way to arrange this movement. We also think a visualization of bike share growth will appeal to the general public, who often lack knowledge of the scale of bike shares in New York City.

There are various questions we hope our visualization will help illustrate. We hope that the animation will show how travel is different on regular work days and public holidays. We also hope to better understand what travel looks like during rush hour and how those rush hour peaks vary based on the day of the week. We also hope that our visualizations will illustrate differences between male and female users, as well as frequent subscription CitiBike users and one time customers. These questions are both personally interesting to us since they illustrate how individuals approach eco-friendly transport and also the most relevant to CitiBike for its understanding of its user base and how to run its business.

Sunday April 12:

Our team conceptualized in detail more what the visualization for the project would look like to have a solid vision before we start coding. We also made data requests to finalize data collection and submitted a weather data request to the National Climatic Data Center, a branch of the National Oceanic and Atmospheric Administration. We need to wait until the work week begins (tomorrow) for the organization to approve our data request. We were lucky to find all of our data (both the bike and the weather data) in CSV form so we won't have to do much data manipulation. We already have our bike data (we included that in the pre-proposal) so the only remaining question with the bike data is how long of a time period we want to visualize.

We are debating between showing several consecutive days and selecting a few particularly interesting days to display. We now have a strong idea about what aspects we want our visualization to have (a map that animates the motion of bikes around the city, a timeline

where users can select where to see data from, graphs (most likely bar graphs) that compare bike usage this day to other days, and information about the day such as weather.

Tasks Accomplished: All data requests submitted, initial concept complete

Monday April 13:

Our team met with our mentor Mohammad. The meeting made us aware of some ways that our original design was vague and our team spent several hours after conceptualizing exactly what the graphs and visualizations would look like. We then began coding. We spent the bulk of the afternoon and evening learning how to use Google maps and set up the correct Google map window with station nodes that scaled appropriately with zooming. At the end of the day, unfortunately realized that it would be impossible to do the movement animation with Google maps, so decided to approach the map animation with MapBox and Leaflet. While it is extremely frustrating that we won't be able to use the code we spent most of the day working on, we still learned about Google maps and also realized the importance of researching the limitations of software before we start implementing. properly. We also changed our time frame and resubmitted the weather data request. By 4 pm we received the correct weather data, making our entire data collection complete, and all the data is loaded into the data folder of our github page.

Tasks Accomplished: Data collection complete. Time slider working. We know now not to use Google maps.

Tuesday April 14:

Today we dug into implementing the maps in a way that will enable animation. First, we came to the in class design studio and got useful feedback from the other team we worked with. They were very enthusiastic about the animated bike movement, and helped us conceptualize what the comparison graphs would look like. We settled on having two separate comparison graphs: one comparing the displayed day to other days (within that we will have several divisions -- other days that had the same weather, other days in the same month, etc.), and one comparing the hour displayed to the other hours of the day (so it will be easy to see if it is a peak traffic time).

After class, we dove into the implementation of the map post Google maps. We got the map up and added nodes to represent stations. We debated the best algorithm to use for bike movement: either using an adaptation of the midpoint formula or using trigonometry. We decided on using the midpoint formula, see our attached photos for the preliminary algorithm.

Tasks Accomplished: Map that will allow animation up and running, station nodes in place.

Wednesday April 15:

Today we focused on the animation. We fixed the station nodes so they scale properly when the user zooms in and out. We added bike nodes and began to think about the animation. One the side, we started working on the daily informational graph, but were slightly derailed by an extremely slow computer that kept crashing.

Thursday April 16:

We realized that our current slider would not work functionality wise with how we envisioned our animation so we are going to modify the slider. Also got the home page working which shows a citibike logo before beginning the animation.

Friday April 17:

Finished the write up of the design studio. The animation is fully functional, however we have to adjust the slider to play the animation correctly. Right now, moving it forward causes the animation to move by 1 minute, as opposed to the actual difference in minutes. It also does not play backwards yet. We also are displaying current weather data and the day, which corresponds correctly to the slider. We have also noticed an error in the slider, when one clicks on a new time -ie not dragging it-, the event fires twice. This is problematic for us.

Monday April 20:

Decided to use a JQuery slider to fix the incorrect event firing.

Tuesday April 21:

Programed the slider to move backwards. Made it so a large jump in time, just places the bikes in the right place without an animation as this is more efficient. We also decided to add filters to break down bike usage by gender and subscription level. We had initially only planned to add the bar charts by gender, but realized that our design would be far more interesting viewers could see patterns based on user type.

Thursday April 23: Added station pings when a new trip starts. This added significantly to the visual appeal of the project, since it is clear when trips are beginning. Especially during rush hour, it can be hard to tell when trips are starting versus just moving past stations, and this makes movement patterns clear.

Sunday, April 26: Added the bar chart without filters. We also had a lengthy discussion about the best information to display on the bar charts. We initially thought it would be interesting to show comparisons that included this hour vs. average hour today but realized it would be more interesting to show a comparison between the days that the users can switch between, especially since a comparison between the different hours in the currently displayed day is better illustrated by the line chart that we added under the timeline.

Tuesday, April 28: Started work on the filters. Made a to do list for the remaining time:

To Do:

- change day
- line graph above the timeline
- line graph that allows you to add stations to it, we want that's axis to just be to current time and should "grow" as time passes. Show at most 3.
- bar charts -> break into smaller
- make sure that all things work for play button and click and drag timeline

- weather data/day data
- process book
- make it pretty/write the story

Thursday April 30: Today we worked to optimize the functionality of the animation. We realized that initializing the animation was taking a long time to load because the script was looping through all the data before extracting the information relevant to the days we wanted to display. We made a python script to extract 1 day's data in a CSV and then only used these 7 CSVs for data, which makes loading and running the visualization much more seamless. We also made buttons to allow switching from day to day.

Friday May 1: Made the area graph above the timeline. This makes it easy for a user to contextualize the bikes they see in the current animation based on the rest of the day (for example, is the current display typical of an hour on Monday or due to a rush hour peak) We also worked on the homepage, including working on parallax scrolling and the moving bike.

Sunday May 3: Much of today was spent reformatting data in order to make the graphs filter by gender and subscription on click. Our previous method for formatting data had wrangled the data in the main vis file and then passed it into the bar graph files already formatted -- making it impossible to further filter based on other factors and also violating proper MVC formatting. Made the station line graph.

Monday May 4: Made the design's layout more attractive. Added the ability to click stations to add it to the graph and to switch between trips starting and trips ending at that station. We also made the final decision today about what days to include in our visualization. We initially planned to include days throughout the year, but realized that this might lead to uncontrolled comparisons. For example, we noticed that CitiBike usage was far higher during some winter days than some summer days, but think that this difference is probably not due to higher biker rates, but rather the fact that CitiBike was more popular after the startup had been around for 6 more months. If the purpose of the visualization is to demonstrate bikers patterns based on the day, then these patterns are only reflective of the growth of one particular bike company, and not normal biker patterns.

As a result, we decided to include a consecutive week of bike information, and selected a week that included a holiday. That way, we are visualizing 4 normal weekdays, 2 weekends, and one federal holiday, and will not have the control issues that we experienced before. It would be extremely easy to select different days to display, we simply chose this week as an example.

Final Reflections and Evaluation:

With the final visualization, we are able to see several patterns in bike usage. Firstly, we noticed that there are two distinct rush hour peaks in the morning and early evening on non-holiday weekdays, such as June 30 and July 1. However, on Weekends, such as June 29

and July 5, we see generally high bike usage in the afternoon, presumably people using bikes for leisure activities, but no commuting peaks.

Additionally, we found that bike usage is far different on holidays. We see an interesting pattern with bike traffic on July 4 which shows bikers biking to the waterfront to watch the fireworks. Bike traffic before the fireworks (in the evening) is noticeably toward the waterfront, and we see a sudden movement of bikes away from the waterfront as soon as the fireworks end, around 10:30 p.m. Unsurprisingly, bike usage is also far lower on holidays.

We see interesting patterns with subscriber versus customer behavior. For example, one time customers (non-subscription users) constitute a very small percentage of riders on weekdays and holidays, but about a third of total users on weekends.

We also see interesting usage differences between males and females. On the average workday, male riders outnumber female riders about three to one, according to the bar charts. However, on weekends, male riders only outnumber female riders 2:1. This indicates that females are proportionately more likely than males to use bikes for recreational use as opposed to transportation to work.

Overall, we are all proud of the visualization we created and think it illustrates interesting and useful patterns. We went far beyond the scope of the project that we initially envisioned. However, one improvement could be to find a way to better compare yearly data. We made the decision to only include the week we did since CitiBike had only been in NYC for a little over a year, so the growth in customer base between the beginning of the year was substantial, and made comparisons of net riders indicative of only the growth, and not biker trends depending on the time of year. However, it would be interesting to show a comparison between different times of the year once CitiBike is not growing as rapidly, and such comparisons could be meaningful.