INFORMATION VISUALIZATION

Student work at the School of Information, Pratt Institute

FILMING LOCATIONS ACROSS NEW YORK CITY

David Bradshaw

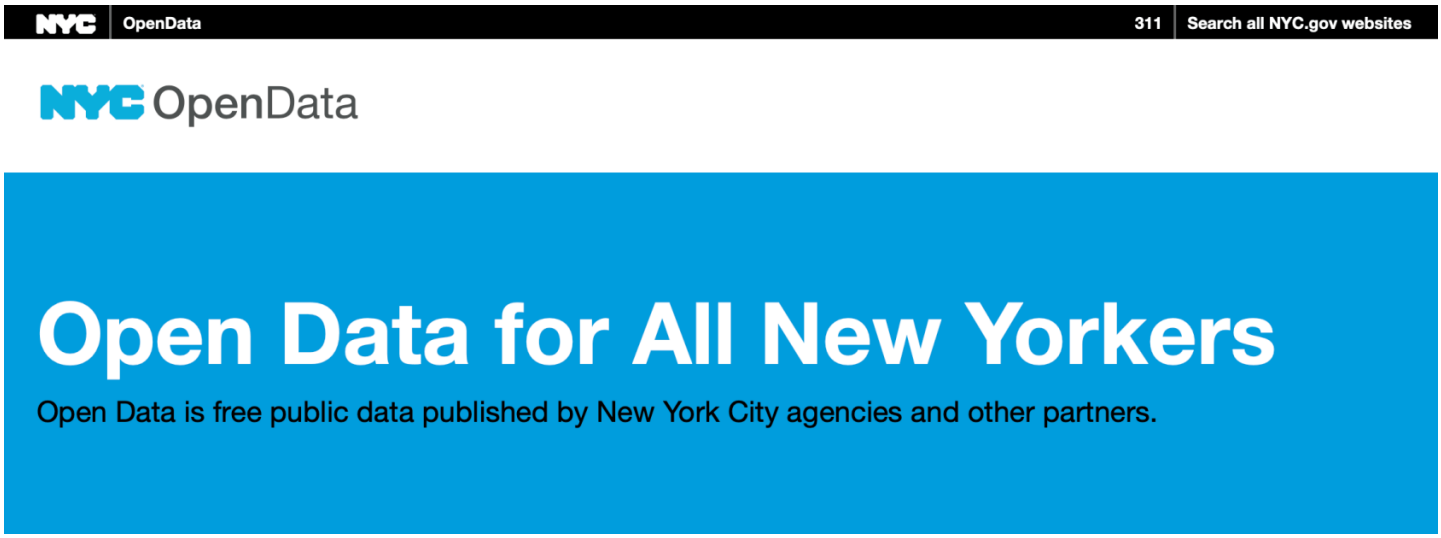
Final Projects

Made in NY

Introduction

For this project I want to create a visualization showing the most filmed locations in New York City. This data will be from the NYC Open Data project which has data from applications for file shoots and other media type items across the city. The data comes from the Mayor's Office of Media and Entertainment The data found here includes over 67,000 file permits issues going back to January 2012. This should give a good qualitive view across each year which can yield averages for the seasons.

Material



NYC Open Data

The data used for this experiment was derived from the Filming Permits issued across the city's five boroughs which are issued whenever a crew wants to film in a specific location or have access to the streets without cars and other interferences and this data came from the New York City Open Data project.

Tableau Public

For processing the data, a tool called Tableau Public which offers an intuitive drag and drop system for creating charts and data visualizations. The visualizations are then taken from the local software and stored in Tableau Public's site to be seen and shared with others.

For the UX study part of this project the tool Cisco Webex was used to connect to others for UX feedback while practicing social distance during the COVID-19 pandemic. This software allows for the participant to share their screen so that I can see their mouse movements and interactions with the prototype while also watching their facial expressions all on one

screen.

Process

After securing the data the data was loaded into Tableau. Once in Tableau I reviewed the Data Source content and noticed a few issues. One was the precise locations I intended to work off of were in human readable formats. For example, I originally intended to use the street closure data to identify precise locations within the city where filming took place. However, the data in this column was stored as street names and cross intersections in multiple patterns. After researching Google Maps and MapQuest APIs I came to the conclusion there was no clean way to retrieve the latitude and longitude coordinates, since the human readable versions would yield assumptions (best guesses) by the APIs and without a manual review it could incorrectly show the outcome. Because of this the precision was raised to be the zip codes.

Some of the data also included multiple points. For example, some entries had multiple zip codes separated by commas in the field. Because this would lead to incorrect rendering of the charts, I took these fields and split them by comma and only kept the first entry as the primary filming location.

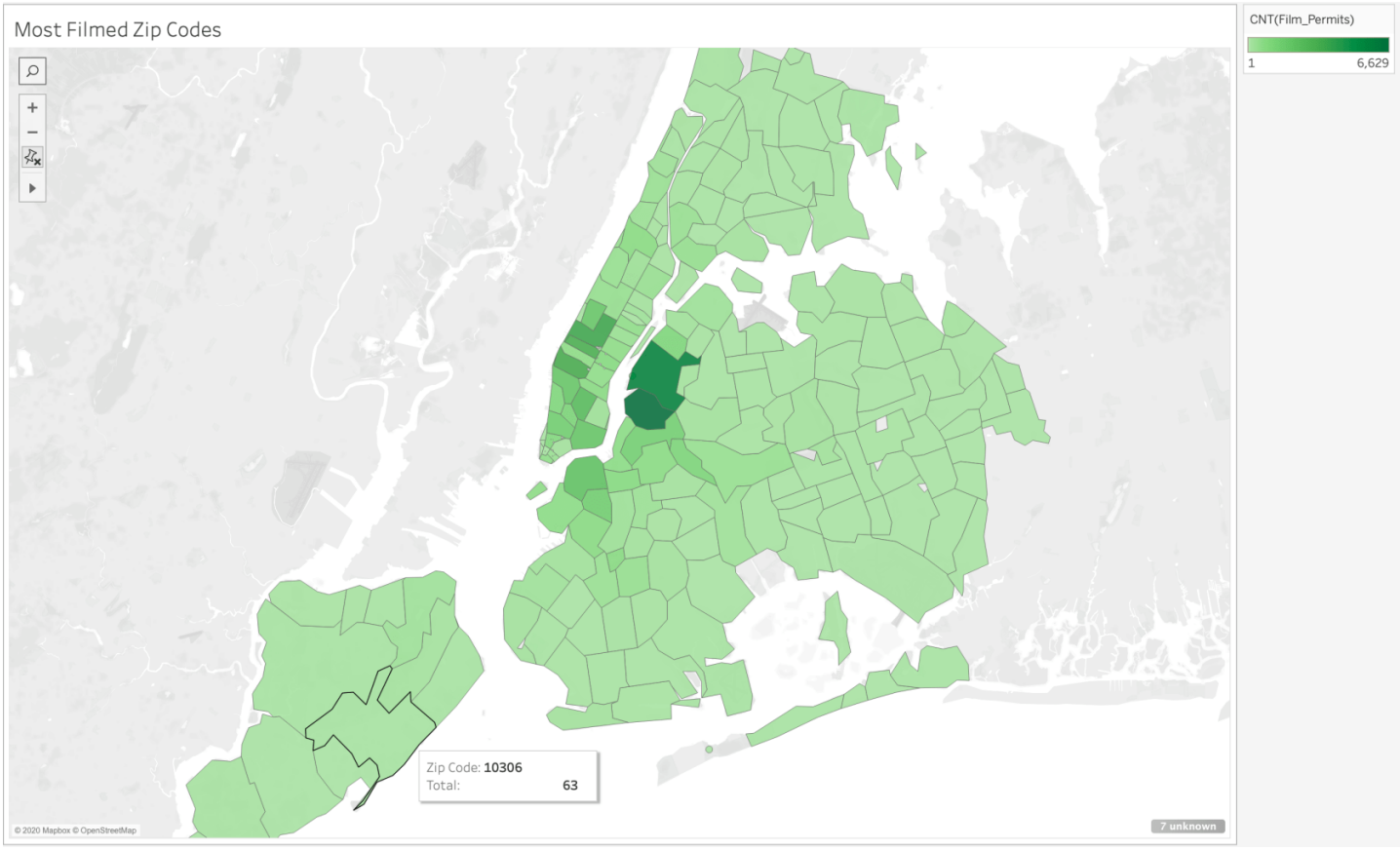
Now that the data was cleaned and ready to render, I created four worksheets to represent different aspects of the data. These will be shown in more detail in the Outcome section. But for this I created a Map which shows the heat scale of the zip codes by numbers of permits, a breakdown of when the most start filming dates were across all five boroughs in an Area Chart, a stacked Line graph showing which months were the most popular across the last 8 years and finally a Packed Bubbles chart showing the breakdown of what type of filming took place. And finally, a Dashboard that shows all in context. Each worksheet and the Dashboard have shared filters to filter all data by the borough the filming took place in and for the month breakdown an extra filter showing the months to allow for multiple months to be compared in a clean method.

Outcome

Click here to view the full interactive dashboard

The final results of the project yielded four worksheets and one dashboard with the interactive visualizations. The interactive visualizations include filters to help visualize specific boroughs or specific months for finding the most filmed months.

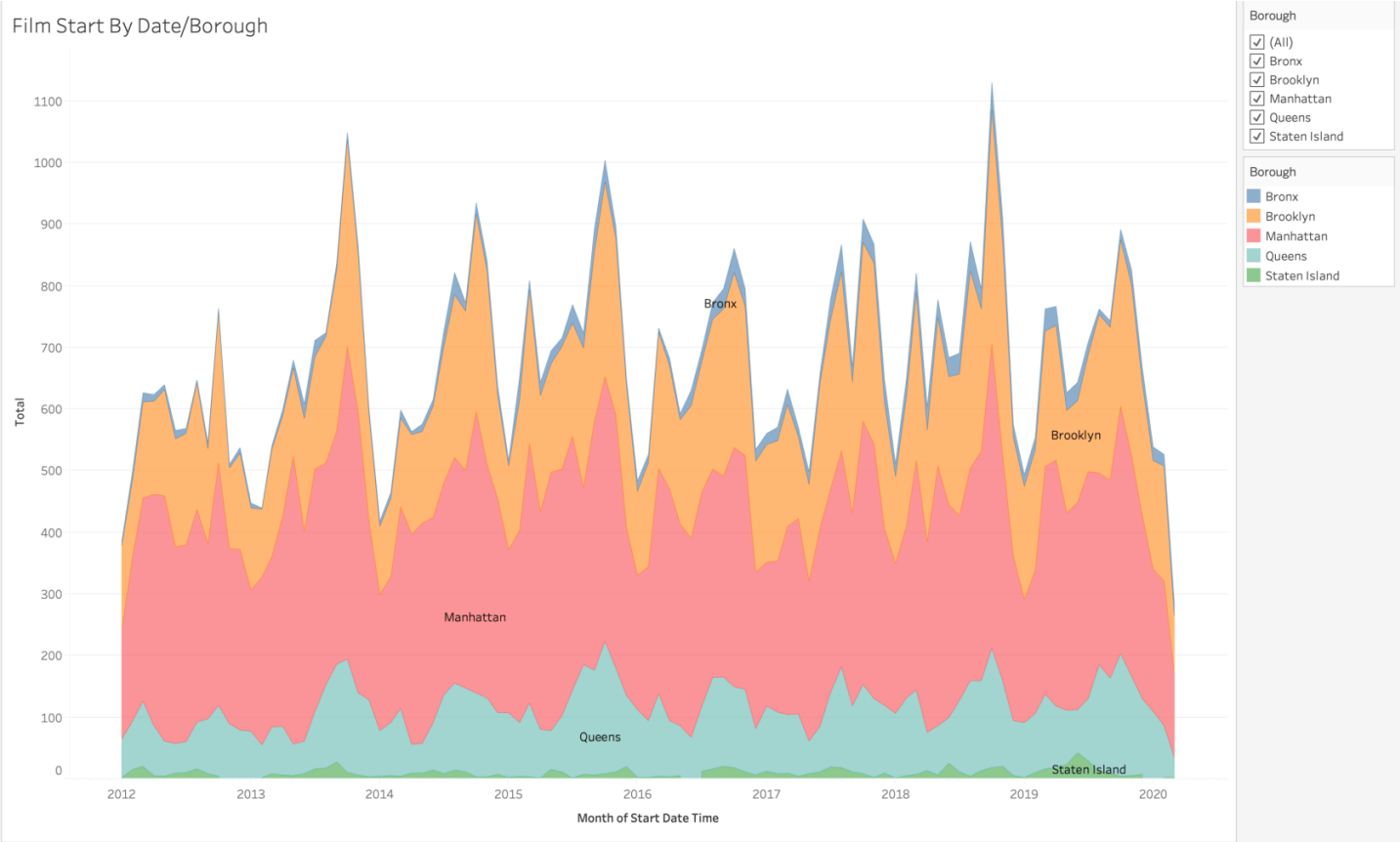
Worksheet 1: Most Filmed Zip Codes



Open Interactive

For the first worksheet we visualize a map of New York City with zip code boundaries showing a more intense green color of more popular areas for filming. This data showed that the area of Long Island City in Queens and parts of West Midtown were the most popular places. Consideration for this could be the Long Island City view provides an amazing view of Manhattan's skyline. A review of IMDb shows that over 187,000 movies marked as having filming locations of Midtown Manhattan and over 153,000 showing Long Island City. Many of the Manhattan findings from IMDb were from older movies that would have been outside of the 2012 scope of the permits.

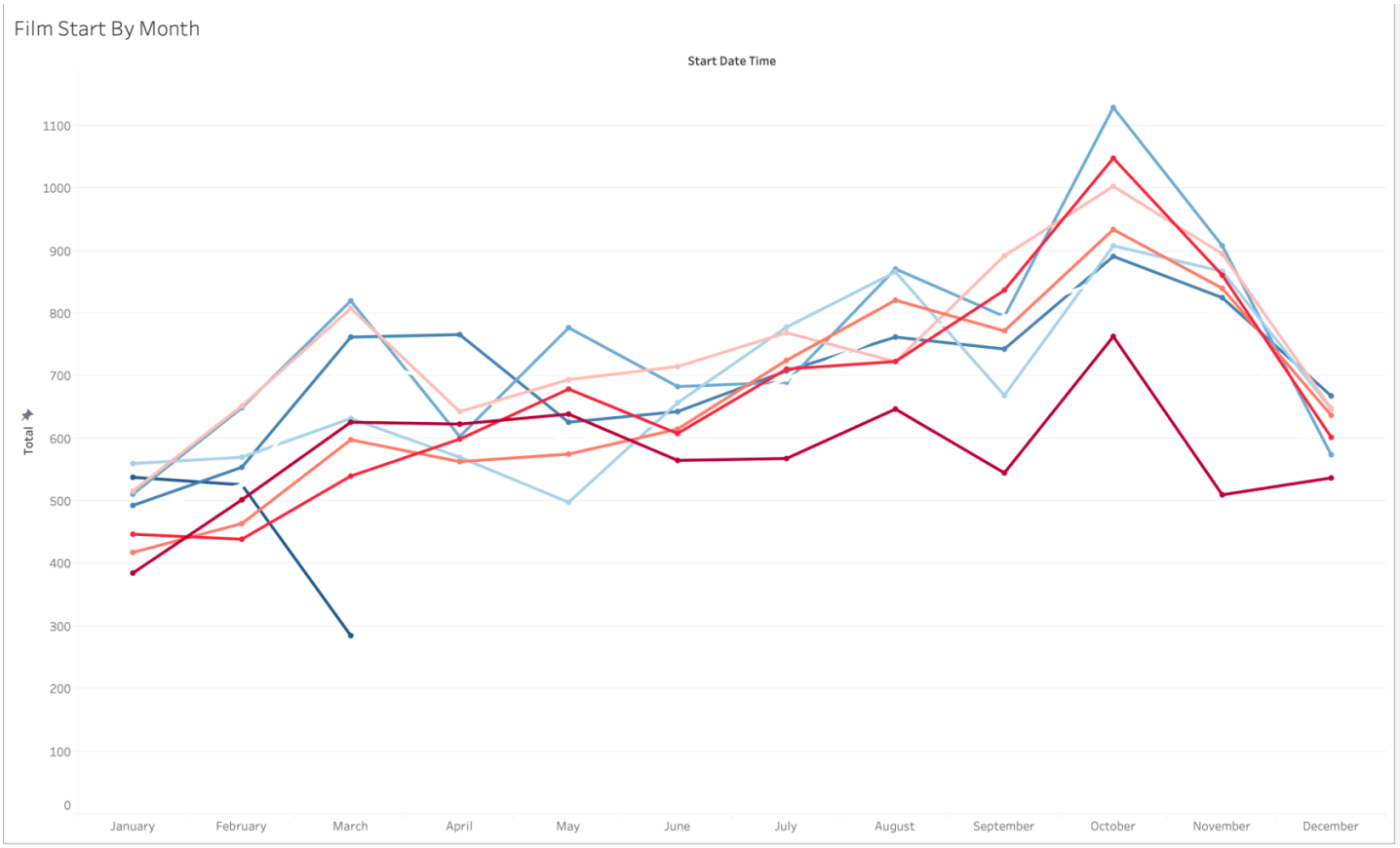
Worksheet 2: Film Start by Date & Borough



Open Interactive

For the second worksheet we visualize via an area chart where the X-axis is the ascending time by Year and month (on the minor) and the stacked charts show the breakdown across the boroughs for each month of the year. The user can hover over these and and see the breakdown of how many film shoots were in each borough on a month. Looking at this data it seems the percentage of borough usage of the month stays pretty consistent not showing any borough having more film shoots than other months and instead follows a pretty steady flow with the totals.

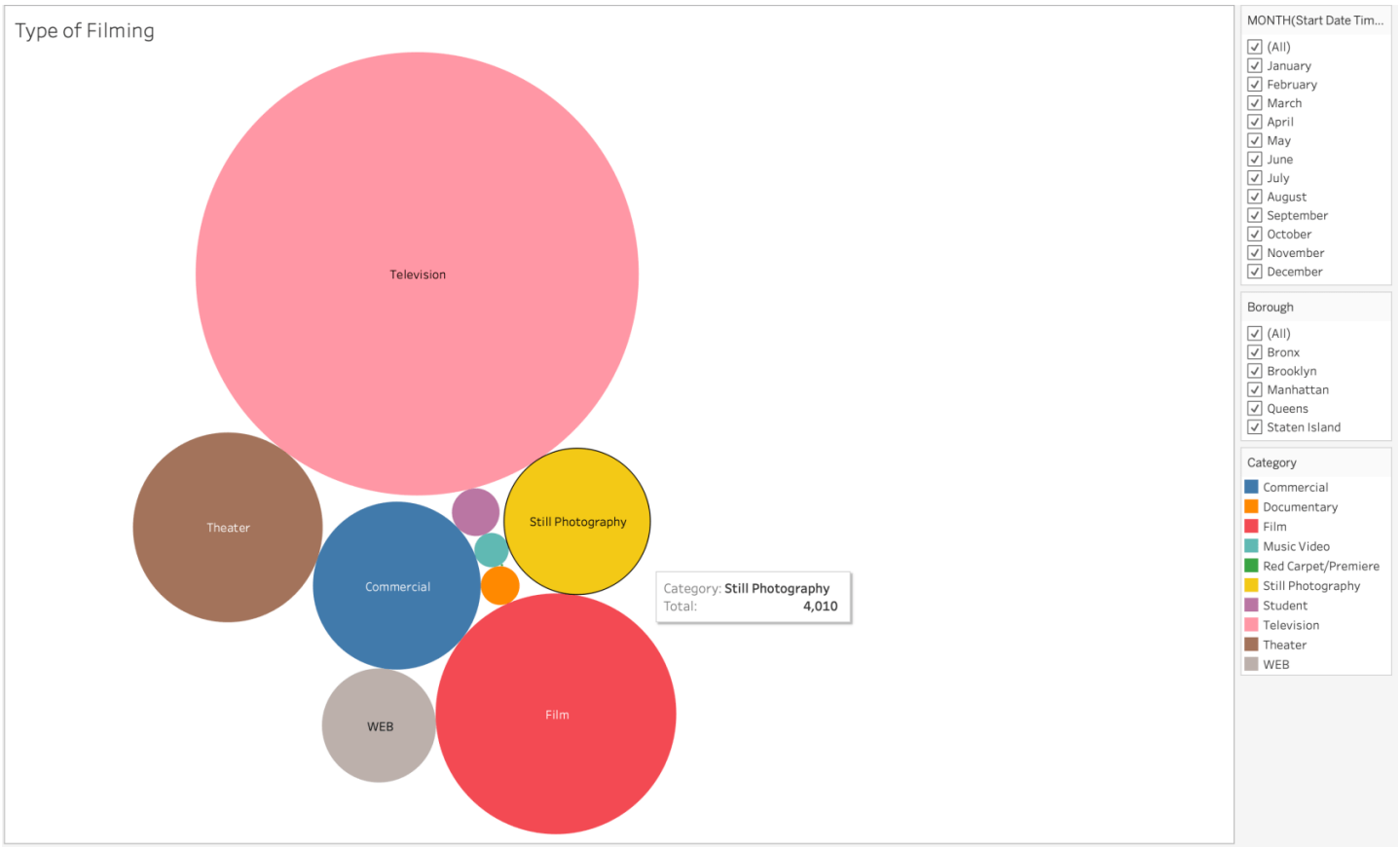
Worksheet 3: Film Start by Month



Open Interactive

Going into this study I was curious on which months were the most popular with film shoots. This stacked line chart shows this data. By showing each month and its total number of film shoots. Interestingly October is consistently the most popular. One can assume for the fall mild weather and changing of the leaves. January and February is consistently the least favorable month with May 2017 being the only exception. Most likely the extreme cold is the factor here, but I was surprised expecting more desire to film with the snowy backdrop. I went to explore if anything happened in May of 2017 but did not see any reason for this drop. The weather was on par with the average and only thing in the news was an incident of a car driving into pedestrians hurting many in Time Square LINK, but no way to know if this is related. This graph also allows the user to filter by the months to compare individual months for a cleaner representation of the data.

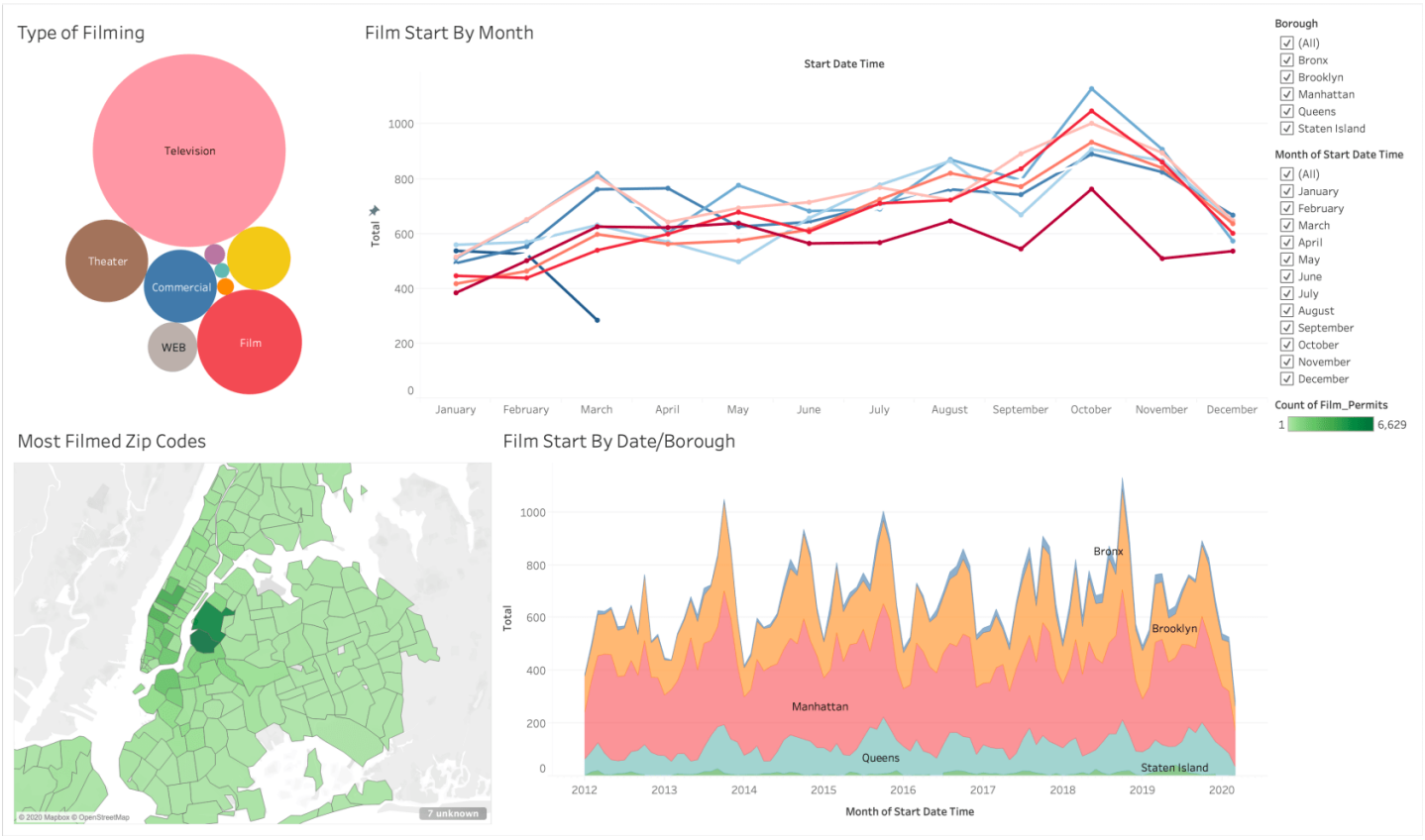
Worksheet 4: Type of Filming



Open Interactive

Finally, I was curious what type of filming happened across the boroughs. To visualize this, I used packed bubbles to quickly identify the largest types of filming. The categories are printed in the bubble when large enough and are available on the legend on the right. The most popular ones were television followed by film. Across the five boroughs television was the primary with Queens having the most percentage of television, and Manhattan having the largest grouping of theater production; which is not surprising with Broadway being in Manhattan. This provides an interesting insight into what were the most popular use cases for each borough. This can also be filtered by months to see a scope of what the most popular use cases were for each month.

Dashboard



Open Interactive

The final dashboard provides a view of all four worksheets with shared filters on the right for filtering by borough and months. This allows for an interactive way to see how the breakdown of location, months and type of media was influences across months and boroughs in one view. All items on the dashboard support useful tooltips on hover that show full specs and name to help the user identify the area and count.

UX Study Plan

The goal for this UX Study is to test the created visualizations and understand if the user can figure out how to answer questions and interpret the data to create useful theories backed by the data. While the visualizations were created for specific theories going into it I also hope to see if users can identify extra ones as well as provide feedback if the original explorations can be identified.

For this UX study I found the following participants:

- **Participant 1**: Sicheng Yu is a restaurant employee living in Manhattan, New York. He regularly enjoys seeing movies at the theater and is originally from China but moved to New York City after living in New Zealand and San Francisco. Primary language is Chinese. This UX interview was done in person.
- **Participant 2**: John Sun is a UX Designer living in San Francisco, California. Primary language is English. He regularly enjoys exercising and painting in his free time. This UX interview was done via Cisco Webex.

Each study followed the same script and pattern, even if the interview was done in a different format. To start off with we introduced the participant to the data and partial goals.

"For this research we want to show a visualization of filming locations across New York City. We hope to explore where are the most popular places to film and find out when is the most popular times for producers and directors to come to New York City. The data being presented is based on film permits issues by the City of New York. You can use the primary view which contains all four visuals and you can also enlarge any of them as you need. We ask that you think out lout and if you have any questions please feel free to ask them. To begin please review the dashboard here. Feel free to click around and when you are ready to continue let me know."

Once the user gains an initial interaction and is ready, we will ask the following to test our experiments:

1. "Can you identify which parts of the City are the most filmed locations, do you recognize these areas, or do you have any thoughts on why they are popular?"
2. "Can you identify which is the most popular season to have film shoots?"
3. "Can you see which is the most popular month to film in The Bronx?"
4. "Can you see what part of New York City is the most popular to film in during January?"
5. "What is the most popular type of media (movies, TV, commercials…) in the City?"
6. "What is the most popular type of media in Queens?"
7. "Is there any patterns or theories you can identify or can explore looking at this?"
8. "Do you have any thoughts or concerns about this visual?"
9. "Is there any extra feedback you would like to provide?"

During this time, I will also be taking notes of their time to click and answer as well as facial expressions. The raw notes will be available in the UX Study Results below.

UX Study Results

Participant 1

Spent about 3 minutes looking over the charts. Avoided the Film Start by month. Did not notice the checkbox filters at first and assumed the legend was filterable.

Questions:

1. Long Island city, that's it. Yes, he recognizes it. Thinks they are popular because you can see Manhattan.
2. He is focusing on the Date/Borough graph. Says Winter is the most popular. Mixed up the legend. Thought the legend for the months (which is near this graph) represented the colors instead of it being the Borough colors. **CRITICAL: User was confused with legends. Moving legend away from this or changing colors can help.**
3. Thinks October is most popular. Clicked the borough filter which did not do a real filter but just highlighted. Changed mind to November after seeing this. **CRITICAL: Clicking filter and highlighting and confusion of filter seems to exist here.**
4. Manhattan, used checkbox to filter for January and found answer. Tried to click January filter but it had no affect so tried checkbox. However, user used the Borough/Date chart and hovered over each item.
5. Says television. Saying it is the biggest circle.
6. Hovers over the circle and says television again. I ask how he knows and then starts to explore checkboxes for boroughs and unchecks all but Queens and then confirms again, saying it is television.
7. What's the most filmed zip codes, which filming is the most popular, and when they start.
8. Thinks it is a good way to figure out where to film if was a movie person.
9. Says the most filmed zip code was interesting. Thought it would be Manhattan, but it was queens.

Participant 2

After looking at the map went straight to trying filters and buttons that appeared, became confused by the month graph and thought colors go with the type of filming. The plus and minus actions on the Y axis lead to extra confusion. The user also thought the stacked graph meant that the Bronx had the highest number based on Y-axis and the fact that the Bronx was at the top. But when hovering realized that the total count makes it the lowest.

1. Used filmed zip code. Said the darker green is the higher. Said it was zip code 11222.
2. Went to checkboxes and started exploring. Said it looks like October, so he assumes it is the fall season.
3. Tried using the film start by date/borough. October is most popular. After filtering things mostly August.
4. Filtered for just January. Would say Manhattan
5. Television
6. Television.
7. Long Island city ever has the most and October has a lot of filming going on then. And television is the main production.
8. Took a few times to figure out that checkboxes effect the three graphs (not the most filmed part was assumed to be controlled by this).
9. He guessed with the map it was not as clear since the filters did not have any, did not realize it was being altered since it did not alter drastically or re-render.

UX Study Final Thoughts and Revisions

After the final UX reviews were completed I explored the feedback and tried to make a plan to address the issues. Some useful changes from the data above would be:

- Labels that are static of the years or a legend to indicate year. Some users were not sure what the line was for.
- More feedback on updates on the map so it shows that filters took effect on it. This is because the data never really changes over the filter.
- Stacked graph with clearer representation of each stack adding to the entire New York

City.

- Have a wider scale of color palettes for each graph to ensure user do not connect similar colors as having a connection.

Reflections

I was surprised with the data and was expecting more in Manhattan and the Bronx, based on recent movies that I have seen with filming in the Bronx. But I guess overall the Manhattan skyline is more of the desired item in films explaining the high number in Queens (Long Island City).

I wish the data for the street locations of filming was more data oriented to allow for getting the geocoordinates easier and in a reliable way. I also wish that IMDb had an export or data format I could use for free to compare these figures to number of movies and also show the most popular, based on critic or use review numbers, for each area.

But overall the data was an interesting insight into the city's "Made in NY" data and showed some insight into how producers pick areas. For this project the data past January 2020 was not considered since it all led to a sudden drop due to the COVID-19 pandemic which stopped film shoots and TV production.

Works Cited

- Film Permits – NYC Open Data
- Filed in New York City – IMDb
- NYC Media and Entertainment – City of New York
- 2017 Time Square Car Crash – Wikipedia

No related posts.

July 23, 2020 · David Bradshaw · Leave a comment