Erin LaFarge

Data Visualization,

Mapping, and

Storytelling

12/12/23

Data Visualization Methodology

Community Scorecard map:

1. Downloaded shapefile of Dept. of Sanitation cleaning sections <https://data.cityofnewyork.us/City-Government/DSNY-Sections/7vgu-qbur>
2. Downloaded pdf file of the Community Board Fiscal year 2022 report of street cleanliness (cleanliness scorecard) <https://www.nyc.gov/site/operations/performance/scorecard-street-sidewalk-cleanliness-ratings.page#:~:text=The%20Scorecard%20Inspection%20program%20measures,your%20Sanitation%20Department%20cleaning%20section>.
   1. Converted pdf to jpeg and used tabular ocr to copy data into excel
   2. Cleaned table in excel and then exported as csv
3. Loaded DSNY shapefile into qGIS
   1. Both datasets are divided by DSNY cleaning section and so can be joined along that column in both sets
   2. I saved that new joined file as a new layer called Percentage of clean streets
   3. I used symbology to make a choropleth map, with cleanest being the deepest green

A screenshot of a computer

Description automatically generated

* 1. I created a print map in the print layout tab

1. Street Cleaning frequency map
   1. I downloaded the csv for all city signs and regulations from open data <https://data.cityofnewyork.us/Transportation/Parking-Regulation-Locations-and-Signs/nfid-uabd>
   2. I read that file into R where I was able to clean and filter for only street cleaning signs and the geographic data that I needed for qGIS (script is asp.R)
   3. I wrote that to a new csv file (asp\_regulations.csv)
   4. Then I loaded this csv file into R
   5. I also loaded the DSNY sections shapefile into R to serve as the basemap. I chose this basemap because I wanted to be able to compare this map with the choropleth map I made of cleanliness.
   6. Since each sign has a unique geo location, I was able to map them as points on the map.
   7. I chose to remove the highest frequency and the weekend street cleaning signs because those are primarily in commercial areas
   8. I focused on mapping only the 2x and 4x a week signs in residential neighborhoods.
   9. I symbolized these as two different color points with the outlines removed and with lower opacity so that they created an overlapping area with those characteristics.
   10. As an experiment, I also overlayed the cleaning frequency on top of the cleanliness map and reduced the opacity further so you could see one through the other.

A map of a city

Description automatically generated

* 1. I ultimately decided not to use this version in my paper because I preferred the side by side comparison.

1. Line Graph of Cleanliness scores over time
   1. From the pdf files of community board cleanliness scorecard, I was able to pull the average cleanliness score by borough for 2016-2023. I manually entered that information into an excel spreadsheet that I exported to a csv.
   2. I read that csv into R and had to pivot longer to make the data graphable
   3. I then created a line chart showing the change over time for each borough
   4. The R script is called cleanliness\_scores.R
2. Parking Violations Graph
   1. This graph was the biggest challenge and I spent over a week attempting to get this data
   2. I pulled this data from the Open Data portal <https://data.cityofnewyork.us/City-Government/Open-Parking-and-Camera-Violations/nc67-uf89>
   3. First, I attempted to filter the data on the open data website so that I was only pulling violations for street cleaning but the portal kept crashing
   4. I attempted an API call but that wouldn’t allow me to call that volume of data either
   5. I downloaded the entire 28gig dataset to my computer and attempted to open in R and another software specificially for large datasets. Neither worked
   6. Finally I was able to filter the data on the open data portal on my desktop in my office (instead of my personal laptop) but then I couldn’t get the data off the computer. Eventually burning it to a DVD and stealing the CD burner from my office succeeded and I was finally able to get this data into R
   7. Once in R, I was able to clean the names and perform some checks on the data. For example, there were nonsensical years in the database like 2099 but I was able to confirm that this was part of the original data and not a coding error.
   8. I also checked for duplicates because the sheer volume of tickets was over 1 million per year and this did not match with news reports that I had read. But a check for duplicate summons numbers revealed that there were no duplicates.
   9. Finally I was able to graph this in a line graph.
   10. The script is called street\_cleaning\_violations.R
3. 311 Complaints graph
   1. This was another huge dataset that I managed by downloading the entire thing and filtering in R.
   2. I created a new csv with only complaints for parking sign violations and double parking which are both related to alternate side parking
   3. I deleted the 24gig 311 complaint csv and started a new R project for the smaller dataset
   4. I grouped these complaints by year and graphed them in a line chart over time
   5. The script for this is called “311 parking violations per year.R”
   6. Since I had made many graphs in R already, I opted to give datawrapper a try
   7. I exported a csv file of just the number of complaints per year and loaded that into datawrapper which quickly generated a line graph for me
4. Link to ASP survey I conducted results: https://docs.google.com/forms/d/1ZfleX-TVeAPKJ65Xh9V6aXyCUwfGjBWLCVI9I-W8nBg/edit?ts=657363ca#responses

General Bibliography (all sources are also sited specifically in context within the article)

1. <https://edc.nyc/article/new-yorkers-and-their-cars#:~:text=While%20almost%20half%20the%20households,of%20Manhattanites%20drive%20to%20work>.
2. <https://www.nytimes.com/2023/10/09/nyregion/nyc-parking-cars.html>
3. <https://www.thecity.nyc/2023/02/03/nyc-streets-getting-filthier-statistics-unclear/>
4. <https://www.nytimes.com/2022/07/05/nyregion/nyc-alternate-side-parking.html>
5. <https://abc7ny.com/alternate-side-parking-street-sweeper-new-york-city-traffic/12019513/#:~:text=But%20the%20Department%20of%20Sanitation,with%20a%20possibly%20weekly%20ticket>.
6. <https://nyc.streetsblog.org/2023/04/19/komanoff-dissects-new-york-citys-car-baby-boom>
7. <https://sunrisesanitation.com/waste-management-strategies-by-some-of-the-cleanest-cities-in-the-world/>
8. <https://www.worldsweeper.com/Country/China/HK_Manual.html#:~:text=All%20of%20the%20streets%20of,with%20brooms%20and%20plastic%20containers>.
9. <https://www.wsj.com/articles/SB10001424052970204394804577012304244132124>
10. <https://www.curbed.com/2022/08/alternate-side-parking-nyc-trash.html>
11. <https://www.boweryboyshistory.com/2023/06/the-new-york-parking-wars-how-cars-took-over-the-curb.html>
12. <https://www.nytimes.com/2021/01/05/nyregion/nyc-residential-parking.html>
13. <https://inrix.com/press-releases/parking-pain-us/#:~:text=With%20the%20goal%20of%20analyzing,wasted%20time%2C%20fuel%20and%20emissions>.
14. <https://nyc.streetsblog.org/2022/03/21/opinion-we-need-an-alternative-to-free-alternate-side-parking>
15. <https://www.curbed.com/2022/08/sanitation-department-dsny-excellent-tiktok-nyc.html>
16. <https://www.nytimes.com/1996/03/17/realestate/streetscapes-cars-when-streets-were-vehicles-for-traffic-not-parking.html>
17. <https://gothamist.com/news/cleanest-streets-are-in-staten-island>
18. <https://www.nytimes.com/2023/10/09/nyregion/nyc-parking-cars.html>
19. <https://www.nytimes.com/2007/03/29/opinion/29shoup.html>
20. <https://wellango.github.io/posts/2021/06/who-owns-cars-in-nyc/>
21. <https://www.curbed.com/2022/04/mary-norris-alternate-side-parking.html>