Downloading 311 Data

Data Visualization and Design | CUNY Graduate Center | Summer 2019

This tutorial will show you how to download an original dataset, decide what to download using the codebook and filters, and organize the data you've downloaded into folders in Tableau. If you are not familiar with 311 Service Requests,

Goals

By the end of this tutorial, you will be able to: * Download 311 data * Decide what to download using codebooks * Organize datasets with many variables

Data

NYC Open Data 311 Service Requests from 2010 to Present freely available via NYC Open Data

Steps

Exploring the dataset

1. Navigate to NYC Open Data 311 Service Requests from 2010 to Present
2. Inspect this data: pay attention to how it is collected, where it comes from. Notice that NYC Open Data actually offers an interface for visualizing the dataset, though we aren't going to use it (Tableau offers FAR more features).
3. Review the codebook. This dataset is about 311 Service Requests. The pivotal variable is the Category. This tells us which Service Request is being made. All other information relates to this variable (i.e., the date, location, etc.)
    1. Notice that most of the information here is geographic (which does NOT mean you need to make a map).
    2. Pay attention to the other variables, what do they communicate?
    3. If the geographic information is in different formats, is there any format that unifies them?

Formulating a Topic 1. Given these variables, what kinds of questions can you ask? 1. Location 2. Duration 3. Time of day/month/year 4. Submission method 5. Resolution Status 6. Description 2. Given what you know about New York City: the people, problems, areas, what variables are you most interested in? What kinds of questions can you ask? What are you curious about? 3. For

example, I see that this dataset has information about Noise Complaints. I've been listening to 99% Invisible and they are doing a series on Sound and Health. This made me think about my own sonic landscape and realize that it's actually quite noisy. I wonder if others have a similar experience in NYC. This is my premise. I would flush this story out more fully, highlighting specifics from the podcast, other research, my own experience, that make this topic worthwhile and lead me to asking these questions about noise in NYC.

Formulating a Question 1. Now that you have your topic, you'll turn it into a question. The question should address your topic AND be able to be addressed with a visualization. You will likely need to ask multiple questions in order to fully address your topic. 2. For now, we will ask: **What months have the most noise complaints?** This is intentionally a very simple example. 1. This is one step towards addressing my topic. 2. This can be answered with a bar chart showing the months. 3. You may already be thinking that we lost some potentially interesting information: What is all the months are about the same, but there are far more complaints on Wednesdays than Saturdays? That would be lost in our visualization. In this case, we would want to ask **Is there any trend in when noise complaints are made¿'**, I could use a line chart to answer this question. (I could also use a radial chart, but that is beyond our current toolset). You could also ask about time of day, the type of noise complaint at different times, etc. For our purposes here, we will stick with our question about months, but know that we have obscured some data to do this, and you would likely want to ask BOTH questions in your project.

Sketching a Visualization 1. I expect my visualization to look something like this:

Downloading the Data 1. Set your filters. I'm going to select Complaint contains Noise and the years 2018-present (because it's a lot of data). 2. Click on 'Export' and Save it as a 'csv for Excel'. If you have multiple languages on your computer, you may need to go back and save as regular 'CSV' (depending on the language settings)

Upload to Tableau" 1. Upload as Text document. If only an error code appears, return to the download page. 2. If Tableau cannot read your file, open it with Excel and save as an Excel document (we'll talk later about the dangers of this).

Organize Variables 3. Move over to Sheet1. There are a lot of variables here. I find it hard to get my bearings, I'm going to group them into folders. 4. Click on one of the many spatial variables. There is a dropdown menu. Select Folder >> Create New folder. I'm going to name mine 'Spatial Variables' 5. You can Control+Select multiple variables at once to add them all to the same folder.

You will notice that there is the option to make folders based on data selection (the more common function for folders), which helps with multiple datasets.

From here you are ready to use the NYC 311 dataset to ask and answer questions

about Complaints in NYC. You do not need to submit anything for this tutorial as it will feed directly into your blog one.

Tutorial written by Michelle McSweeney, PhD for *Introduction to Data Visualization,* a course in the M.A. in Digital Humanities at the Graduate Center at CUNY. More information about the program is available here.