As someone who wants to intersect environmentally conscious work with Data Science in order to bring positive change and practical, real world understanding from STEM to the forefront of discussions, I chose a dataset that would help facilitate this process by bringing useful information about trends and patterns to the table for analysis and easy-to-understand discussion. For my project, I am going to be exploring a New York City Environmental Data Set surrounding the Recycling Diversion and Capture Rates of each District. The recycling diversion rate is defined as the “percentage of total municipal solid waste collected by the Department of Sanitation that is disposed of by recycling,” whereas the capture rate is the “percentage of total Paper or Metal/Glass/Plastic in the waste stream that is disposed of by recycling.” Within this data set, I aim to discover the trends of the rates of recycling diversion and capture rates. I will do this by:

* Creating summations of the respective rates over each year in the data set in order to see the patterns throughout the fiscal year, which starts in July and ends in June, as well as hypothesize the reasons for the measurable differences in our data
* Create bar graphs to compare the trends of each year in order to examine the change of our RD and capture rates, and hypothesize reasons for the changes (or lack thereof) throughout the years
* Explore the trends of which districts seem to have a higher/lower rate of RD and Capture and hypothesize reasons for the differences

Doing this will require combining the districts of each borough into one [borough] in order to gain a general insight of the trends per borough. I will be using pandas and matplotlib in order to explore and discover these insights, as well as create visual representations for the data in the forms of graphs.

Additionally, if I have time to incorporate this, I will use a simple dataset from NYC that lists all the public recycling bins within each respective borough, in order to explore the possible correlations between RD and Capture rates, and the number of recycling bins per borough. I will list the links of each data set I intend to use below and will update if I find other data sets that will help me gain insight and strengthen the hypotheses that I have listed within my bullet points above.

Datasets:

<https://data.cityofnewyork.us/Environment/Recycling-Diversion-and-Capture-Rates/gaq9-z3hz>

<https://data.cityofnewyork.us/Environment/Public-Recycling-Bins/sxx4-xhzg>