

Feedback Exercise

Group we met with: *Squirrel census - NYC - Squirrels in My Park (SIMP)*

Feedback we gave/questions we had:

- For the census recorder - is there bias when recording data?
- Potentially look at different years (2018 vs post-pandemic) to compare behavior and add project complexity.
- Look into Python packages for GIS - mapping squirrels to a map
- Are there species classifications? Can you classify them into species with available data?
- Do you have access to the weather data? Fewer squirrels on a rainy day? When there's bad AQI?
- For exploratory analysis - Using geographical data, are there squirrel hotspots?

Feedback we received:

Additional data we could include:

- Do we have access to diet prior to admission?
- Were each of the patients already diagnosed with diabetes prior to admission?
- How many variables will be included? Is it robust enough?

Ethical considerations:

- Something to add to our current ethical considerations:
 - If parents could see this would it cause undue stress if they thought their child was going to go into a more severe state. There are a lot of ethical considerations when delivering really intense news to families.

Data Cleaning:

- Create a dictionary of all possible iterations of words and then create a function that will match these descriptions to classify to simplify cleaning steps.
 - Like the example of "tachy", "tachycardia", "tach"

Analysis:

- Should we do a linear regression for the Glasgow score? And logistic regression for binary outcomes? (Rather than only looking at logistic regressions)