Project Benson

Victoria M. Ng & Nicholas Thomas

"Is anyone interested in a Gala?"

Increase Match Rate

✓ High traffic

✓ High interest

Datasets

- MTA Turnstile data

- American Community Survey

| Question | Data | Cleaning | Model | Visualization |

Cleaning

Reformatted	Dropped	Identified	Notable Anomalies
☐ Columns	⃠ N/A	✓ Unique ID	? Backward counter
☐ Entries		✓ Potential Duplicate Station Names	? Non-regular audits
☐ Time + Date	⃠ Timecount < 1 row		? 3/11-3/17-3/23
			? Timecount < 1 row
	⃠ 3/11, 3/17-3/23 (inclusive)		
	⃠ Non-regular audits		

Question	Data	Cleaning	Model	Visualization

Model

- Sum across ~4 weeks of entries
- Visually represented data by entry traffic in the 90th quartile
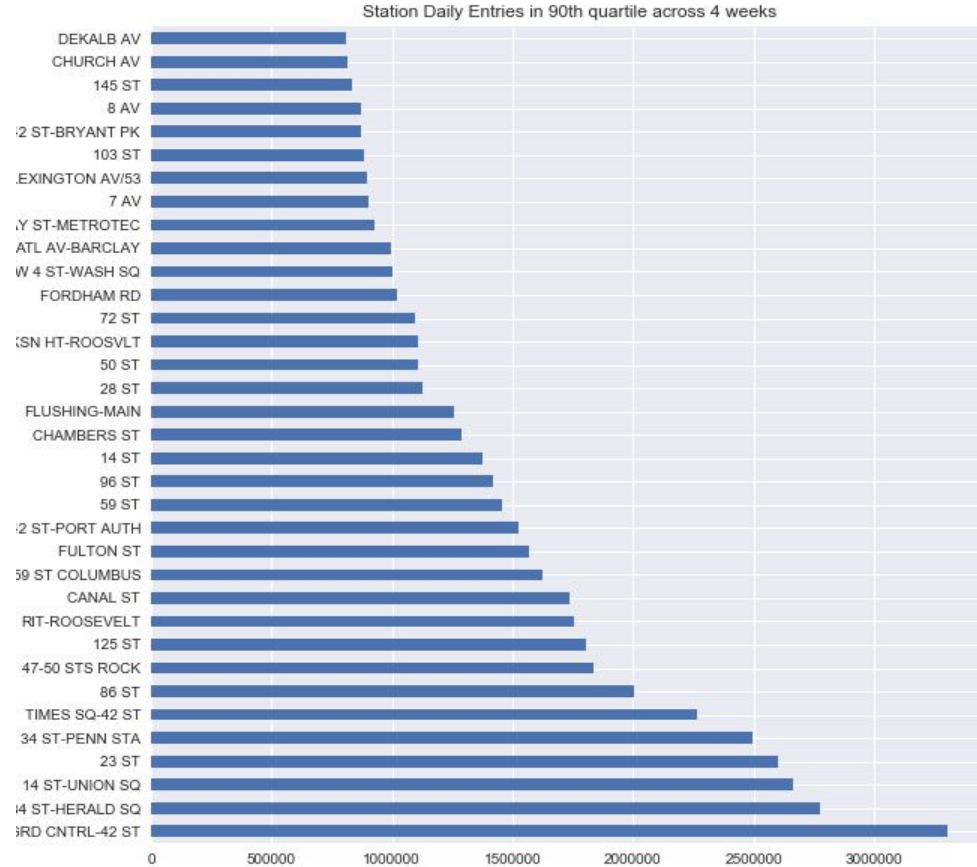- Visually represented top 35 stations with highest concentration of PhDs in zip code.

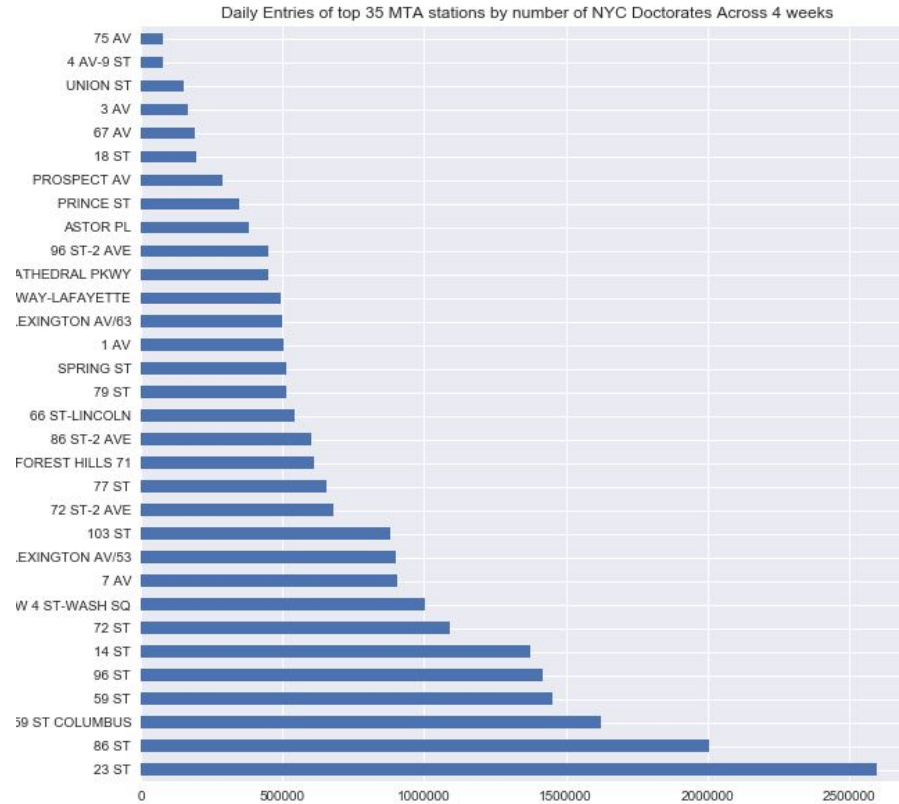| Question | Data | Cleaning | Model | Visualization |

Station Daily Entries in 90th quartile across 4 weeks

| Question | Data | Cleaning | Model | Visualization |

Daily Entries of top 35 MTA stations by number of NYC Doctorates Across 4 weeks

| Question | Data | Cleaning | Model | Visualization |

Entries by Station

Station

Daily Entries (across 4 weeks)
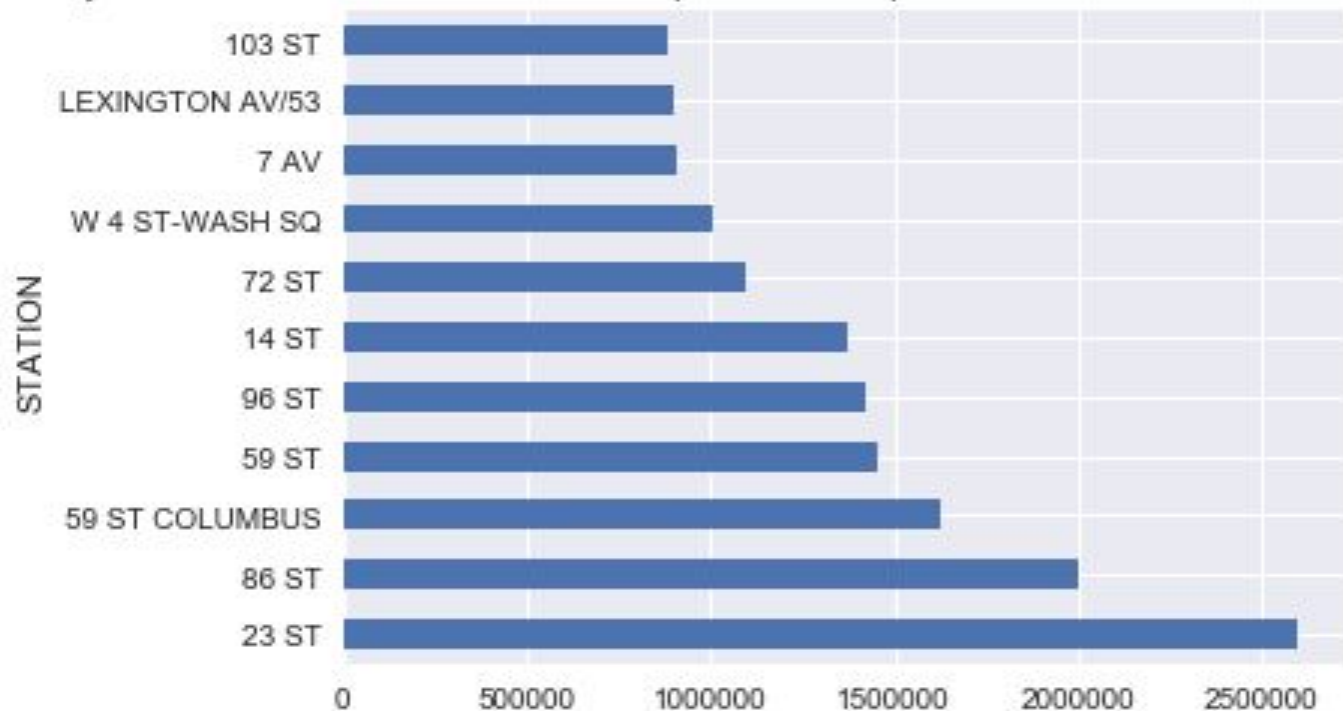
90th Quartile
Phd Top 35
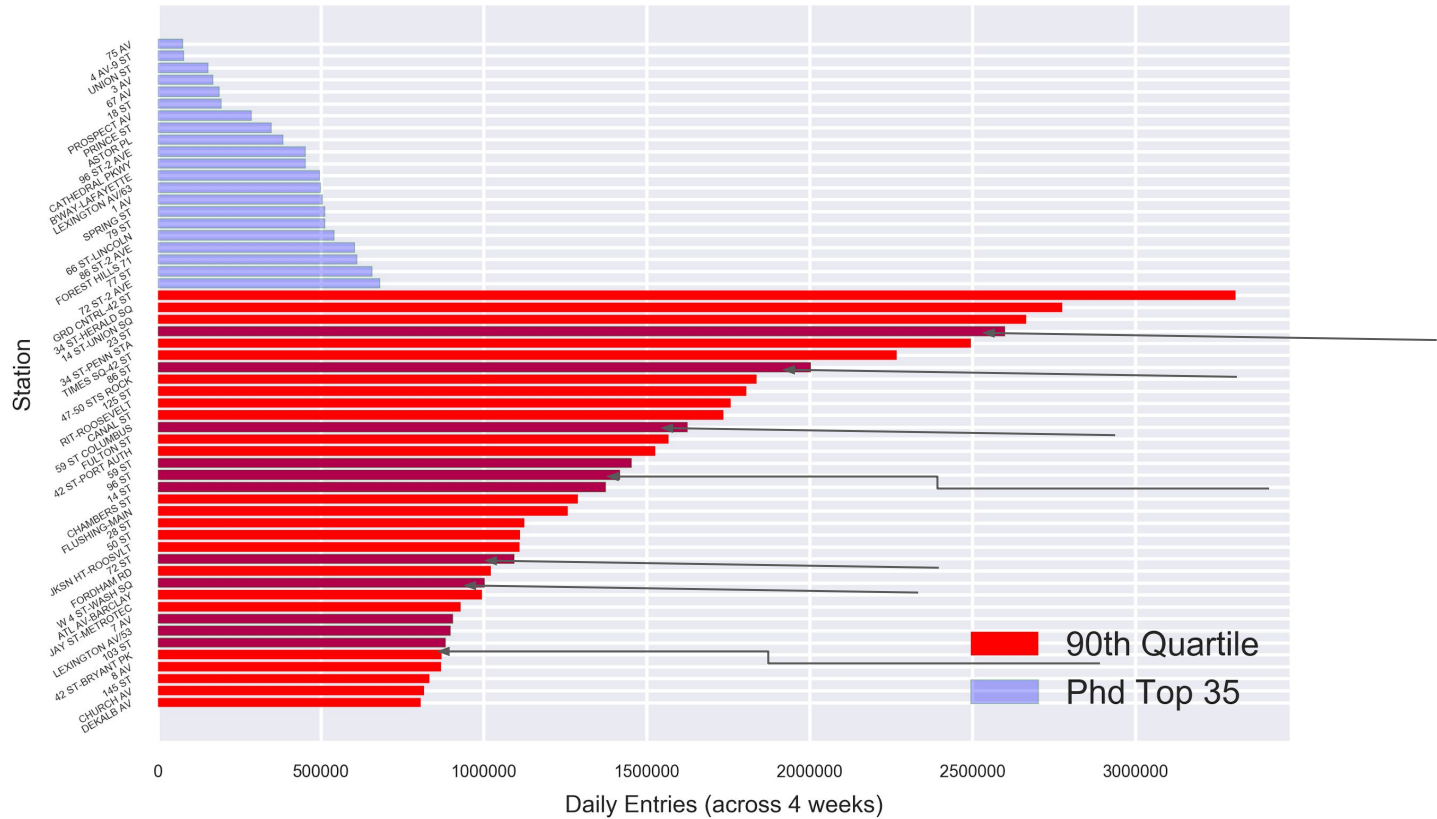
Question ＞ Data ＞ Cleaning ＞ Model ＞ Visualization

Daily Entries of 11 MTA stations in 90th quartile and top 35 number of NYC Doctorates Across 4 weeks

| Question | Data | Cleaning | Model | Visualization |

Conclusions

- WomenTechWomenYes should focus on the following stations:
    - 103rd St
    - Lexington Av/53rd
    - 7th Av
    - W. 4th St/ Washington Square
    - 72nd St
    - 14 St
    - 96 St
    - 59th St
    - 59th and Columbus
    - 86th St
    - 23rd St

Thank You

Questions?

Citations

- Steven Manson, Jonathan Schroeder, David Van Riper, and Steven Ruggles. *IPUMS National Historical Geographic Information System: Version 12.0* [Database]. Minneapolis: University of Minnesota. 2017