Group 7: Visualizing Bacterial and Chemical Data in Boston Tap Water

Meera Ravichandran, Fangyi (Chloe) Zhao, Emily Chen

DS 4200 Fall 2020 — Prof. Cody Dunne, Northeastern University

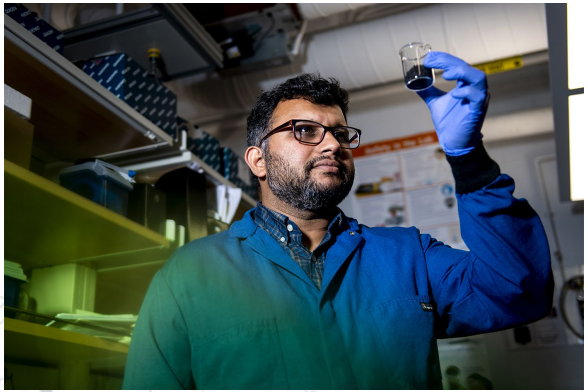
Introduction

Understand the behavior of microbial communities in relation to tap water characteristics → Improve microbial management strategies → Safer drinking water

- Over 200 bacterial communities investigated
- Tap water characteristics such as:
  - Nitrate
  - Nitrite
  - Chlorine
- Use optimal levels of chemical characteristics to store tap water
- Better water filtration devices and techniques
- Improve water distribution strategies

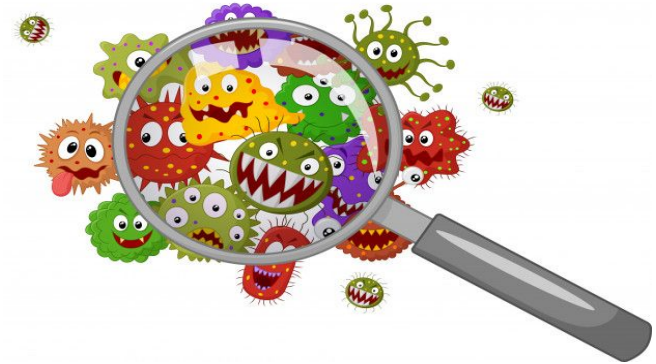
Partner → Goal

- Dr. Ameet Pinto & PhD candidate Solize Vosloo @ Northeastern University College of Engineering
  - Lab focus: characterizing bacterial communities in water to improve water filtration and distribution methods



<- Dr. Pinto!

- Visualize chemical & genetic data
  - Discover patterns to help scientists and researchers model the behavior of bacterial communities
  - Help research community improve water filtration and distribution systems



Data

Run	Date	Season	Temperature	pH	Chlorine(mg/	Ammonium(	Nitrate(micr	Nitrite(micrc	Non_ICC	ICC
Run#001_03	3/1/19	Spring	22.55	9.11	1.43	29.65	6.45	0.17	15	5
Run#005_03	3/15/19	Spring	22.85	9.24	1.44	16.15	8.87	0.11	22	9
Run#010_04	4/2/19	Spring	18.9	9.29	1.71	29.36	11.29	0.13	31	15
Run#015_04	4/19/19	Spring	22.8	9.3	1.22	28.58	8.06	0.22	30	14
Run#020_05	5/7/19	Spring	22.85	9.3	0.93	26.62	8.06	0.51	40	19
Run#025_05	5/24/19	Spring	21.4	9.6	1.42	29.36	9.68	0.23	33	9
Run#030_06	6/11/19	Summer	22.9	9.5	1.37	29.95	17.74	0.32	39	12
Run#035_06	6/28/19	Summer	22.25	9.57	1.38	28.77	14.51	0.64	32	10
Run#040_07	7/16/19	Summer	22	9.43	1.71	31.32	16.13	0.62	27	8
Run#045_08	8/2/19	Summer	22.7	9.59	1.94	32	13.71	0.57	23	4
Run#050_08	8/20/19	Summer	22.7	9.58	1.7	29.95	8.87	0.55	35	7
Run#055_09	9/6/19	Fall	22.35	9.65	1.67	29.55	11.29	0.51	52	12

Formatted data used for the bar chart and table

Our data was provided to us by Dr. Pinto and PhD candidate Solize Vosloo. The raw data corresponds to tap water samples taken twice a week (before March 2020) over a year's span. The tap water was taken from the Mugar Life Sciences building at Northeastern.

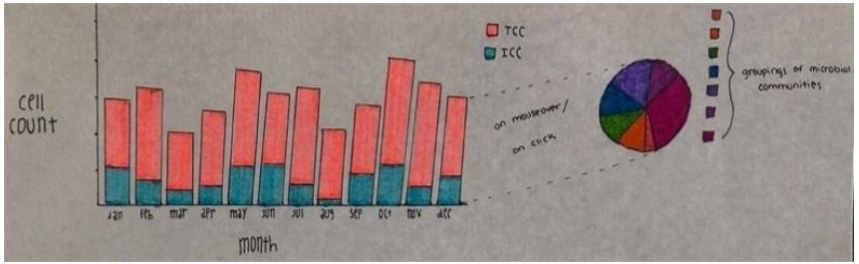
Date	Type	Percentage
3/1/2019	Betaproteobacte	5.93
3/1/2019	Nitrospira	5.37
3/1/2019	Gammaproteoba	0.987
3/1/2019	Planctomycetia	0.711
3/1/2019	Actinobacteria	0.694
3/1/2019	Oligoflexia	0.16
3/1/2019	Gemmatimonad	0.138
3/1/2019	Chlamydiia	0.121
3/1/2019	Flavobacteriia	0.111
3/1/2019	Deltaproteobact	0.0866

Formatted data used for the heatmap

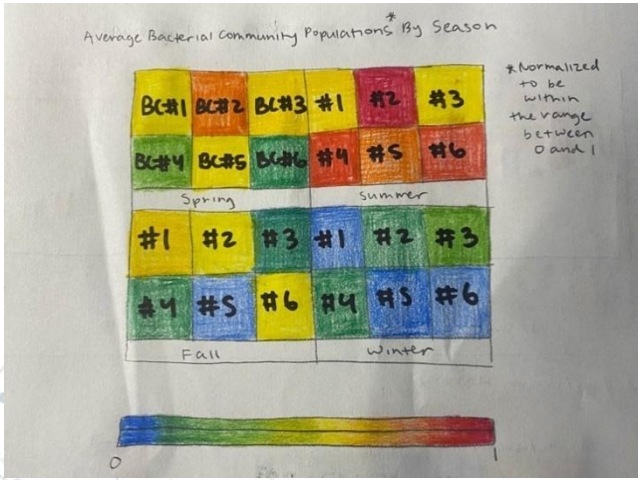
Task Analysis

Task ID	Domain Task	Analytic Task	Search Task	Analyze Task
1	How do the cell counts fluctuate over time?	Summarize	Explore	Present
2	What is the proportion of bacterial communities at a certain time?	Compare	Locate	Derive
3	Is there a relationship between specific chemical levels and bacterial communities?	Compare	Explore	Discover
4	How is pH and temperature affected by the change in time?	Summarize	Explore	Discover
5	How are the chemical levels related to the regulatory standards imposed by the MWRA?	Compare	Explore	Derive

Preliminary Sketches

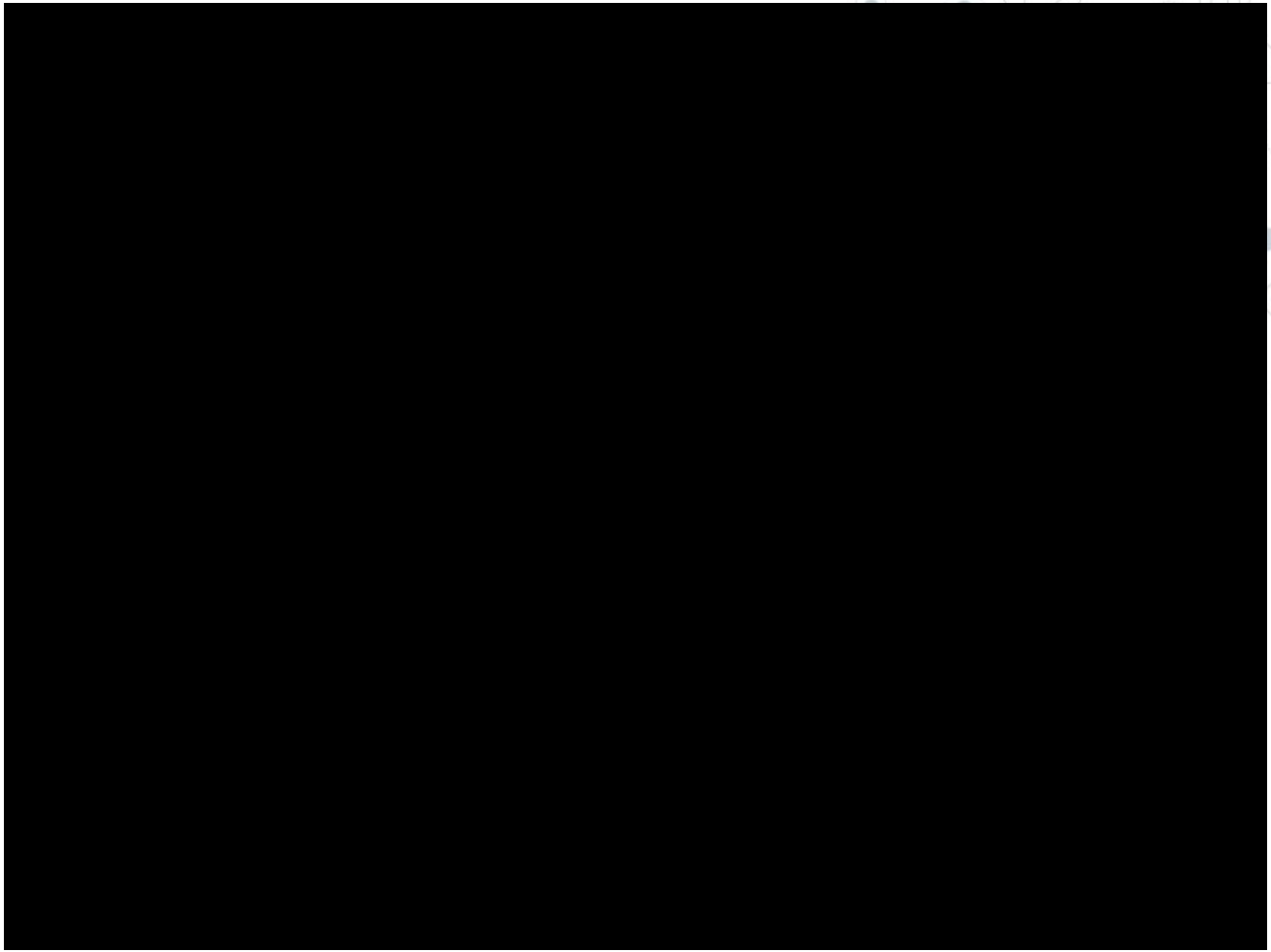


A stacked bar graph to measure bacterial populations over time. Each bar is the average cell count per month and an added dimension, color, depicts the total cell count vs the intact cell count. When the user clicks on a bar, it shows the proportions of which bacterial communities are apparent at that time in a pie chart.



A heatmap sectioned into the different bacterial communities where each square area mark represents the normalized average population for a specific bacterial community. A sequential colormap was utilized to create ordering between the range of populations, allowing for comparison between seasons as well as between different communities.

Visualization Demo



Visualization Explanation

Task #1 | Task #2 | Task #3 | Task #4

Stacked Bar Graph

- ★ Explore changes in bar height
- ★ Summarize how cell counts change over time

Details on Demand Heatmap

- ★ Compare relative percentages of different bacterial communities at certain dates

Brushing & Linking

- ★ Compare chemical levels in table with bacterial abundance in heatmap and barchart
- ★ Explore relationship between genetic and chemical data

Table

- ★ Discover relationship between pH and temperature over time

Conclusion/Acknowledgements

- Visualized chemical & genetic data
- Fulfilled our four tasks through a barchart, heatmap, and table using d3.js, CSS, and HTML techniques
- Will be demoing our webpage for Dr. Pinto's lab!

- We would like to thank all of the individuals who contributed to our project:
    - Ameet Pinto and Solize Vosloo
    - Professor Cody Dunne
    - David Saffo
    - Bl.ocks.org
    - Stack Overflow

Thank You for Listening!

Any Questions?