BIOINFORMATICS ANALYSIS

VISUALIZING TRANSMISSION OF CLOSTRIDIUM DIFFICILE

George Sivulka

Mentor: Theodore Pak   Teacher: Jefferey Marcucio

Contents

- What is *Clostridium Difficile*?
  - Visualization Approach
- ***The Data Science***
  - Dealing with Raw Data
  - d3.js Visualization
  - Optimization
  - Interpretation
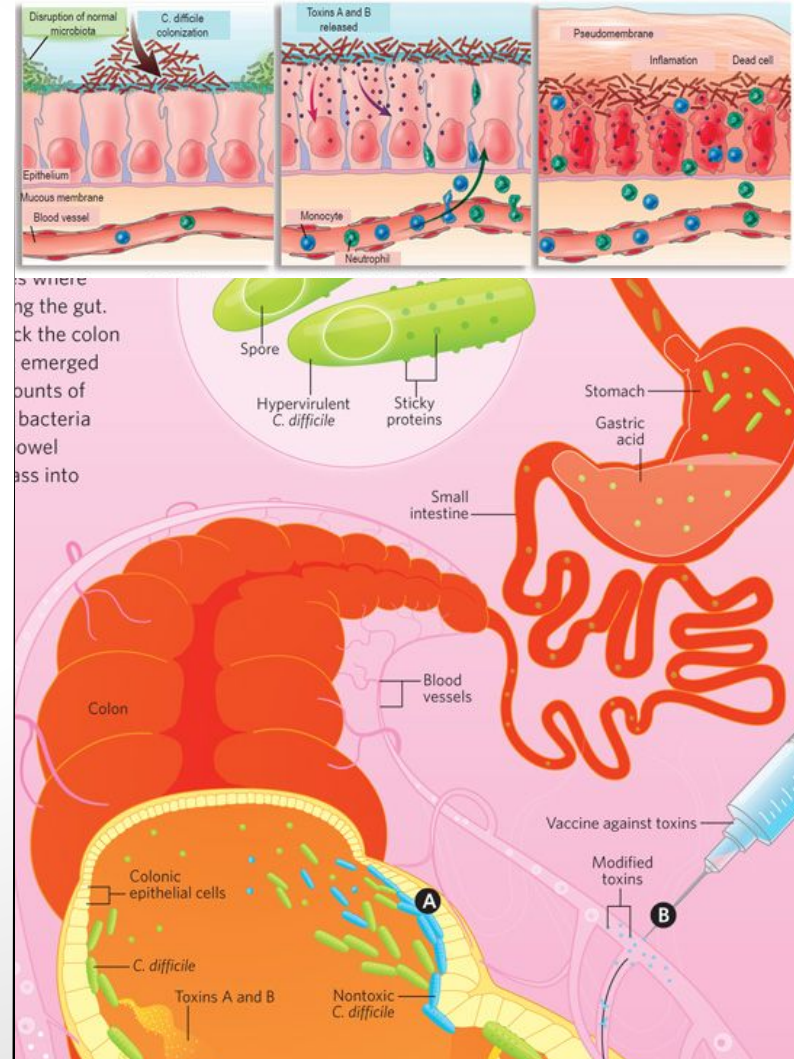- Conclusions
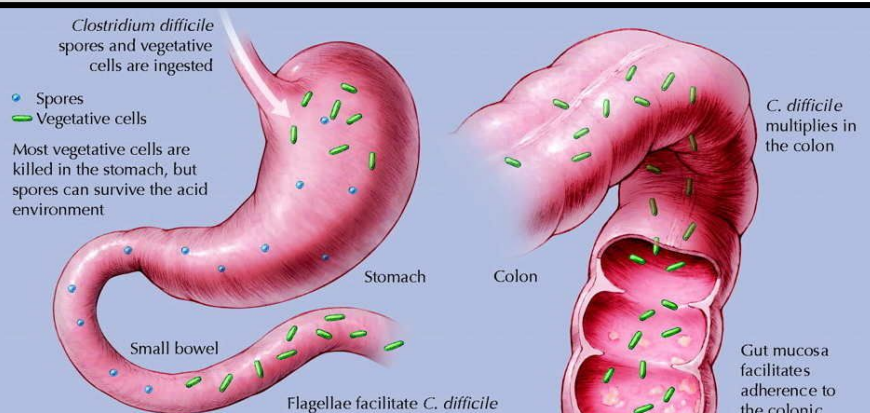 - Other Applications

Clostridium Difficile...

- Bacterium
  - **Spore-forming**, gram-positive
  - "Colitis"
- Infection
  1. Spores survive stomach acid
  2. Bile acids activate
  3. Latch to epithelial Cells
  4. Enterotoxin & Cytotoxin
  5. Diarrhea, Inflammation, etc.
- Fecal-Oral Route in hospitals

A Nosocomial Disease

- Spores resistant to cleaners
    - not killed by alcohol/heat
    - survive in clinical environments for long periods
- Antibiotics help!
    - Replaces normal gut flora (impaired)
    - Further antibiotic treatment difficult

Clostridium difficile spores and vegetative cells are ingested

- Spores
- Vegetative cells

Most vegetative cells are killed in the stomach, but spores can survive the acid environment

Stomach

Small bowel

Flagellae facilitate *C. difficile*

C. difficile multiplies in the colon

Colon

Gut mucosa facilitates adherence to the colonic

As a Healthcare Problem

- 29,000 deaths in US in 2011 [1]
- Acquisition
  - 13% (stays of up to two weeks)
  - 50% (stays longer than four weeks)
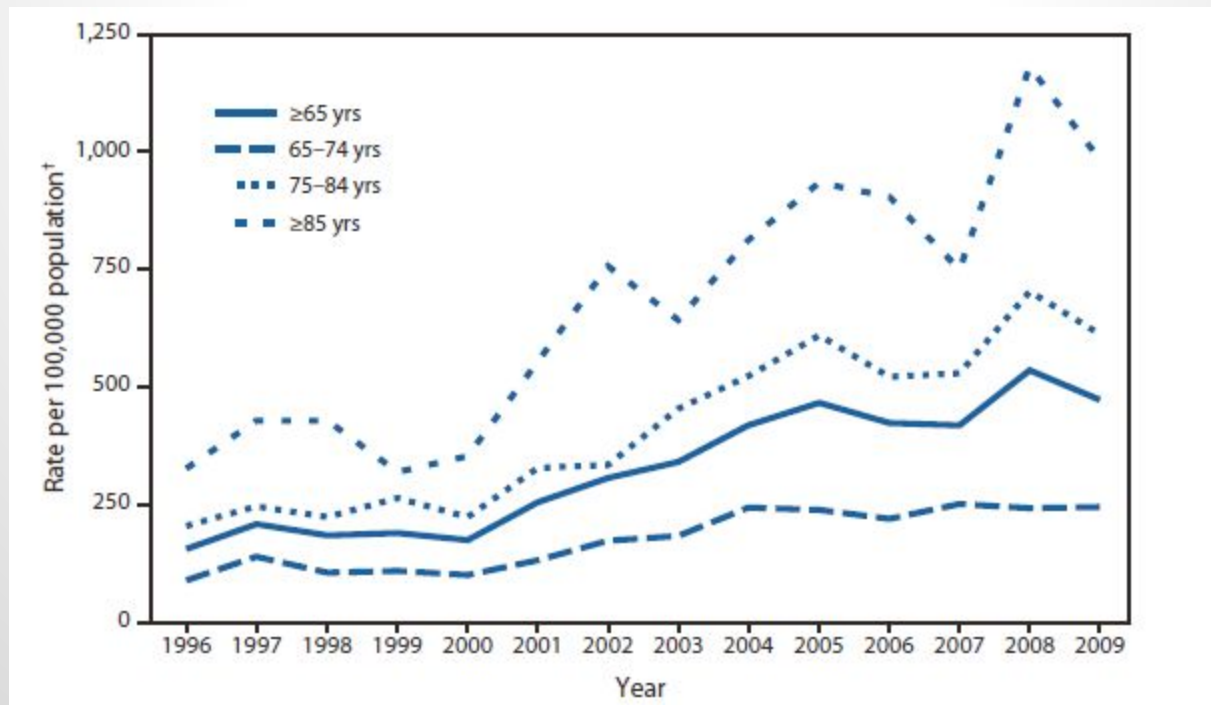
Review of Literature

1. McFarland, Lynne V., et al. "Nosocomial acquisition of Clostridium difficile infection." New England journal of medicine 320.4 (1989): 204-210.

2. Clabots CR et al. (September 1992). "Acquisition of Clostridium difficile by hospitalized patients: evidence for colonized new admissions as a source of infection". The Journal of Infectious Diseases 166(3):

3. Bostock Mike et al. (May 2015). "D3.js Documentation". https://github.com/mbostock/d3

Problem Statement and Hypothesis

- The purpose of this research was to investigate whether a data-driven visualization could be created to prove a higher rate of nosocomial infection of Clostridium difficile associated with different hospital caregiver jobs, ultimately shedding light on previously faulty infection control practices in the Mount Sinai Medical System.

- It was hypothesized that using the past hospital logs sufficient trends and correlations could be proven for future reference.

Methodology - Getting Data

- Queried Mt. Sinai by Primary/Secondary Diagnosis
  ↳ `"008.45 INTESTINAL INFECTION DUE TO CLOSTRIDIUM DIFFICILE"`
  ↳ for Patient's `"LENGTH_OF_STAY"`, `"AGE_IN_YEARS"`, `"MASKED_MRN"`, `"CHECK_IN_DATE"`
  ↳ and their caregivers `"FIRST_NAME"`, `"LAST_NAME"`, `"CAREGIVER_ROLE"`
-

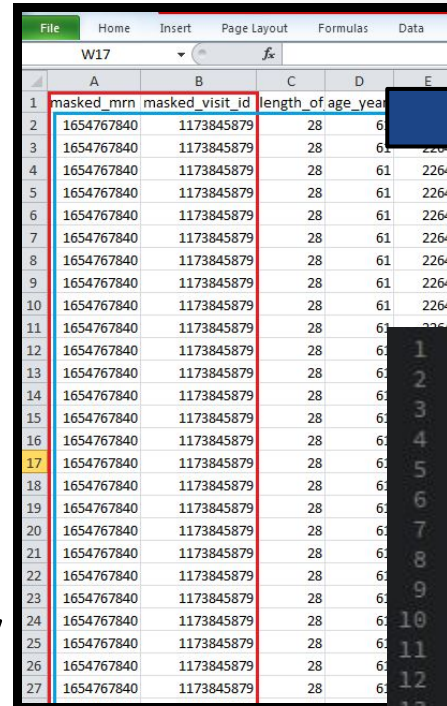|				
cdiff-caregivers-dates	5/18/2015 5:17 PM	Microsoft Excel C...	7,426 KB
cdiff-caregivers-dates	5/18/2015 5:18 PM	SQL File	1 KB
cdiff-caregivers-visits	3/28/2015 8:52 PM	Microsoft Excel C...	25,809 KB

masked_mrn	masked_visit_id	length_of	age_year	age_in_da	gender	address_z	date_of_c	principal	secondar	deceased	discharge	admission	id	caregiver_role	first_name	last_name	id
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Admitting	Brian	Markoff	224
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Admitting	Erin	Rule	224
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Admitting	Peter	Taub	224
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Admitting	Steven	Weinfeld	224
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Anesthetist	O	Guttman	225
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Assistant	David	Forsh	225
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Attending	Ageliki	Vouyouka	225
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Attending Anesthesiologist	J	Eisenkraft	225
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Case Manager	Lea	Kantelinen	225
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Dictating	David	Forsh	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Dictating	Steven	Weinfeld	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Discharge Manager	Lea	Kantelinen	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Discharging	Adam	Hedaya	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Discharging	Ageliki	Vouyouka	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Discharging	David	Forsh	226
1654767840	1173845879	28	61	22645	Male	10021		140.24 ATI	INTESTIN	No	SKILLED N	Non Healt	168115	Discharging	Marianne	Gelber	226
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Entered By	Mary Ann	Walker	228
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Entered By	Rose	Cabornero	228
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Entered By	Sabine	Jacques	228
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Entered By	Siddique	Abbasi	228
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Entered By	Susan	Alsamarai	228
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Ordering	Abraham	Vatakencherry	229
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Elizabeth	Marciak	234
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Eric	Mayott	234
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Henry	Chu	234
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Melissa	Quispe	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Rolly	Wong	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Technician	Victor	Zaretsky	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Utilization Manager	Genevieve	Annuy-Salvador	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Utilization Manager	Glory	Lemons	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Utilization Manager	Kantelinen	Lea	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Utilization Manager	Lea	Kantelinen	235
1654767840	1173845879	28	61	22645	Male	10021		140.2 ATI	INTESTIN	No	SKILLED HE	Non Healt	168115	Utilization Manager	Sandra	Leslie	235
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Admitting	Lester	Katz	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Anesthetist	E	Goren	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Assistant	Mark	Reiner	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Assistant	Richard	Crockett	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Attending	Lester	Katz	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Attending Anesthesiologist	A	Leibowitz	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Case Manager	Colleen	Kelley	2670
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Created By	Colleen	Kelley	267
743879745	153889986	19	77	28203	Female	10552		√55.2 ATT	INTESTIN	No	HOME HE/	Non Healt	2221	Dictating	Lester	Katz	267

cdiff-caregivers-visits

Methodology - Parsing

- JavaScript
- *Force Layout needs specific type of data*
- Underscore.js
  - ↳ CSV > JSON format
  - ↳ `[.each]` function
- Objects/Array Hierarchy
  - ↳ **Connections>Edges Array**
  - ↳ **People > Nodes Array**
- 63,254 lines of data
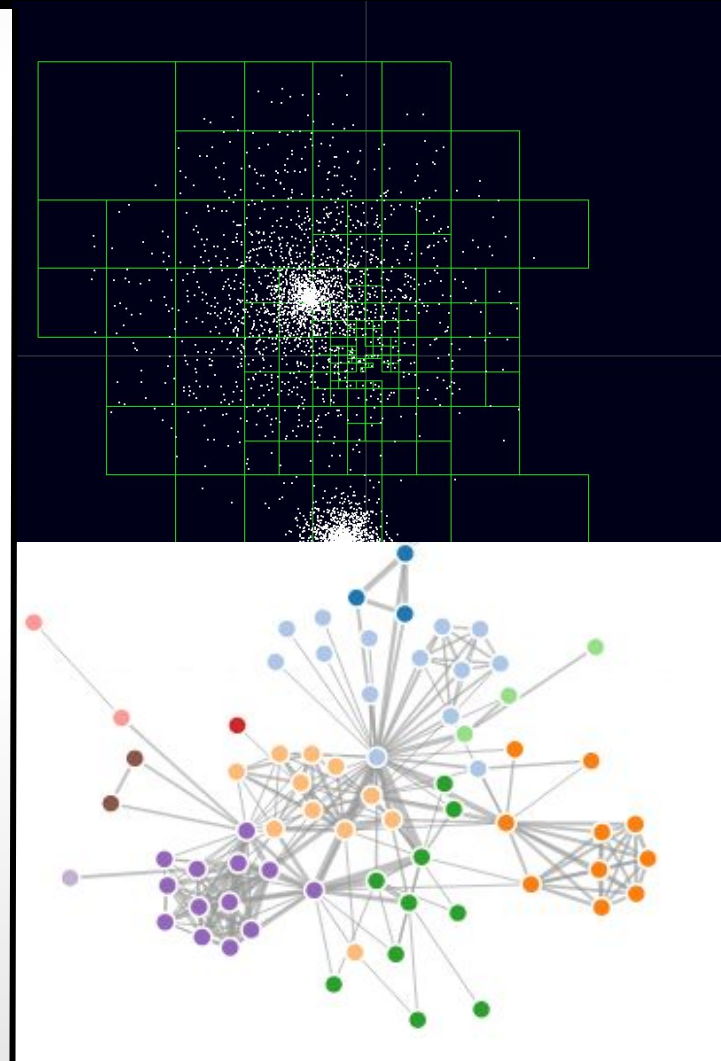
Methodology - Visualization

- Force Layout in d3.js
  ↳ ***Verlet integration*** to constraints
    - fixed-distance geometric edges
  ↳ ***Barnes–Hut approximation*** for charge interaction
 ↳ "tick" event

Methodology - Differentiation

- Vis created, "Now what?"
  - just a bunch of dots
- NEED TO FIND TRENDS
↳ Differentiation
  - caretakers vs. patients
  - roles/jobs
  - time
  - other information

```javascript
function color(d) {
    // TODO: check d.group ?
    return d.group == "patients" ? "#3182bd" : "#c6dbef";
}

function outline(d) {
    if (d.role == "Attending") {
        return "pink";
    }
    if (d.role == "Technician") {
        return "grey";
    }
    if (d.role == "Assessor") {
        return "red";
    }
    if (d.role == "radiologist") {
        return "green";
    }
    if (d.role == "Primary Surgeon") {
        return "blue";
    }
    if (d.role == "Attending Anesthesiologist") {
        return "black";
    }
    if (d.role == "Rn - Scrub") {
        return "orange";
    } else {
        return "#3182bd";
    }
}
```

Results

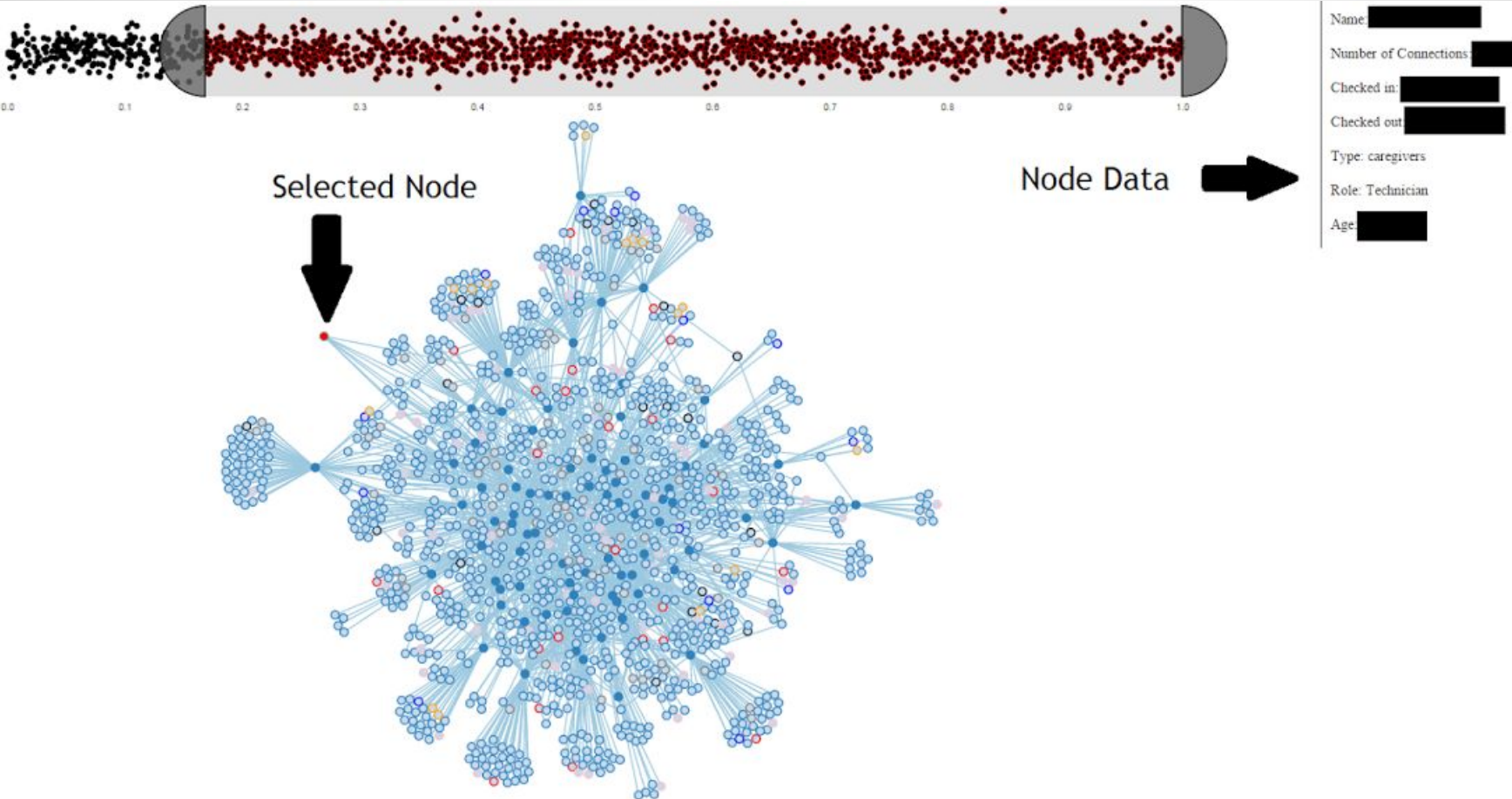127.0.0.1/index4.html

Results - Trends
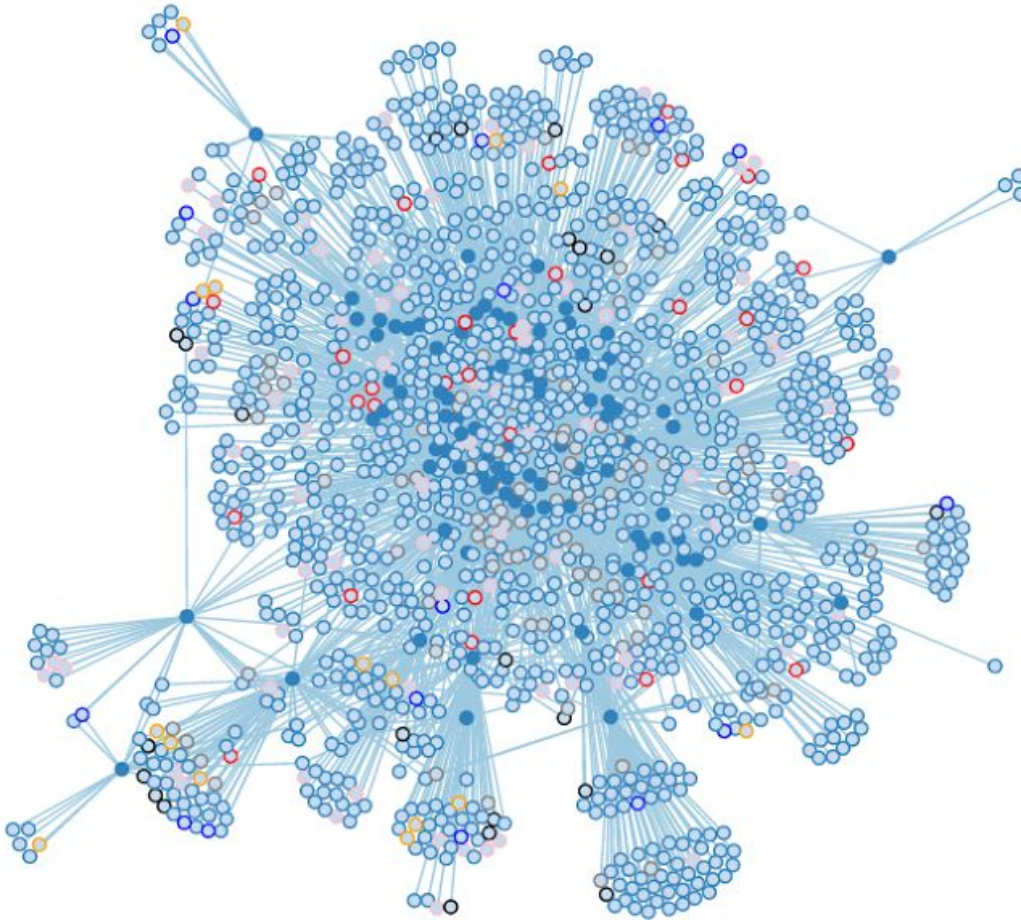
Large Amounts of Grey and Pink Outlined Nodes in the Center=

More Connectivity among Technicians and Attendants with Ill Patients

Results - Node Selection

Selected Node

Node Data ➡

Name: ▉
Number of Connections: ▉
Checked in: ▉
Checked out: ▉
Type: caregivers
Role: Technician
Age: ▉

Discussion/Conclusion

1. **High** Amount of Connectivity
2. Caretaker Roles Trends
   - ✓ Anesthesiologists/Surgeons
   - ↘ Technicians/Attending
3. Individual Trends
   - ↳ Quantitative Analysis

   (on top of visual)

- Future Work
  - ↳ DNA sequences
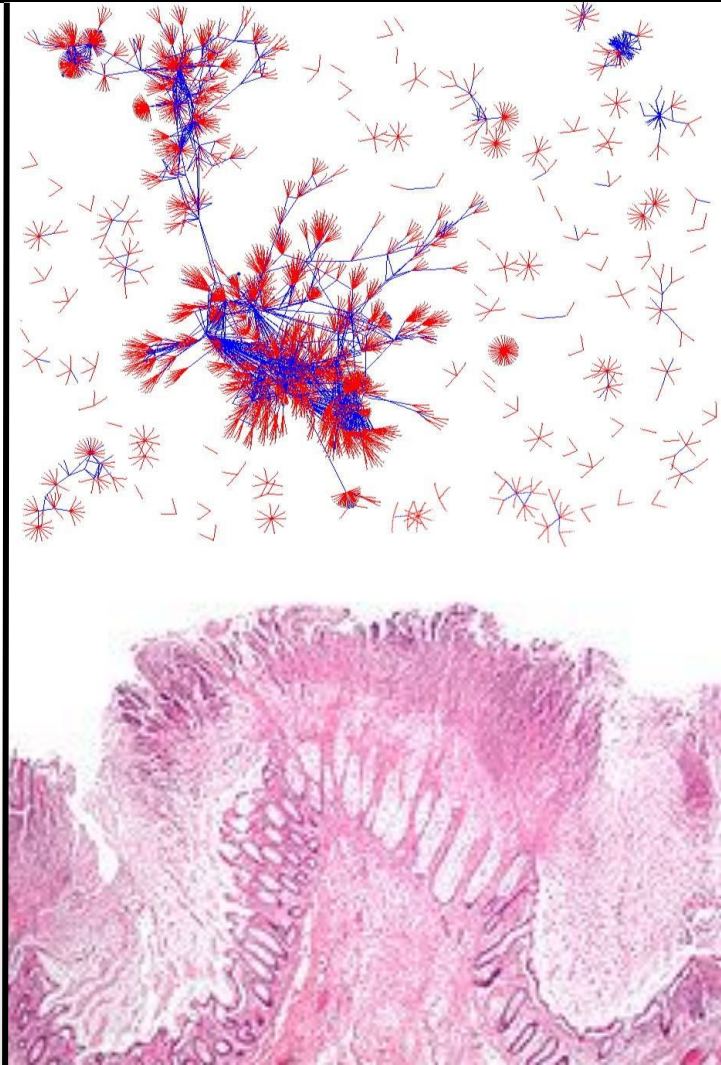
Sources

1. McFarland, Lynne V., et al. "Nosocomial acquisition of Clostridium difficile infection." New England journal of medicine 320.4 (1989): 204-210.
2. Clabots CR et al. (September 1992). "Acquisition of Clostridium difficile by hospitalized patients: evidence for colonized new admissions as a source of infection". The Journal of Infectious Diseases 166(3):
3. Bostock Mike et al. (May 2015). "D3.js Documentation". https://github.com/mbostock/d3
4. Lessa, Fernanda C et al. (26 February 2015). "Burden of Infection in the United States". New England Journal of Medicine 372
5. Halsey J (2008). "Current and future treatment modalities forClostridium difficile-associated disease". American Journal of Health-System Pharmacy : AJHP : Official Journal of the American Society of Health-System Pharmacists 65 (8): 705–15
6. Anglim, Anne M., and Barry M. Farr. "Nosocomial diarrhea due to Clostridium difficile." Current Opinion in Infectious Diseases 7.5 (1994): 602-608.
7. Kurz, Andreas, and A. Min Tjoa. "Data warehousing within intranet: prototype of a web-based executive information system." Database and Expert Systems Applications, 1997. Proceedings., Eighth International Workshop on. IEEE, 1997.
8. Theisel, Holger, and Matthias Kreuseler. "An enhanced spring model for information visualization." Computer Graphics Forum. Vol. 17. No. 3. Blackwell Publishers Ltd, 1998.
9. National Hospital Discharge Survey, Annual Files, 1996--2009. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6034a7.htm