



Penn Medicine

Penn – CMU Digital Health Privacy Initiative

# An Analysis of 3rd Party Tracking on Hospital Homepages

Karim Farhat & Amey Maley

August 17, 2021



# CONTENTS

Overview

Aims

Methods

Findings

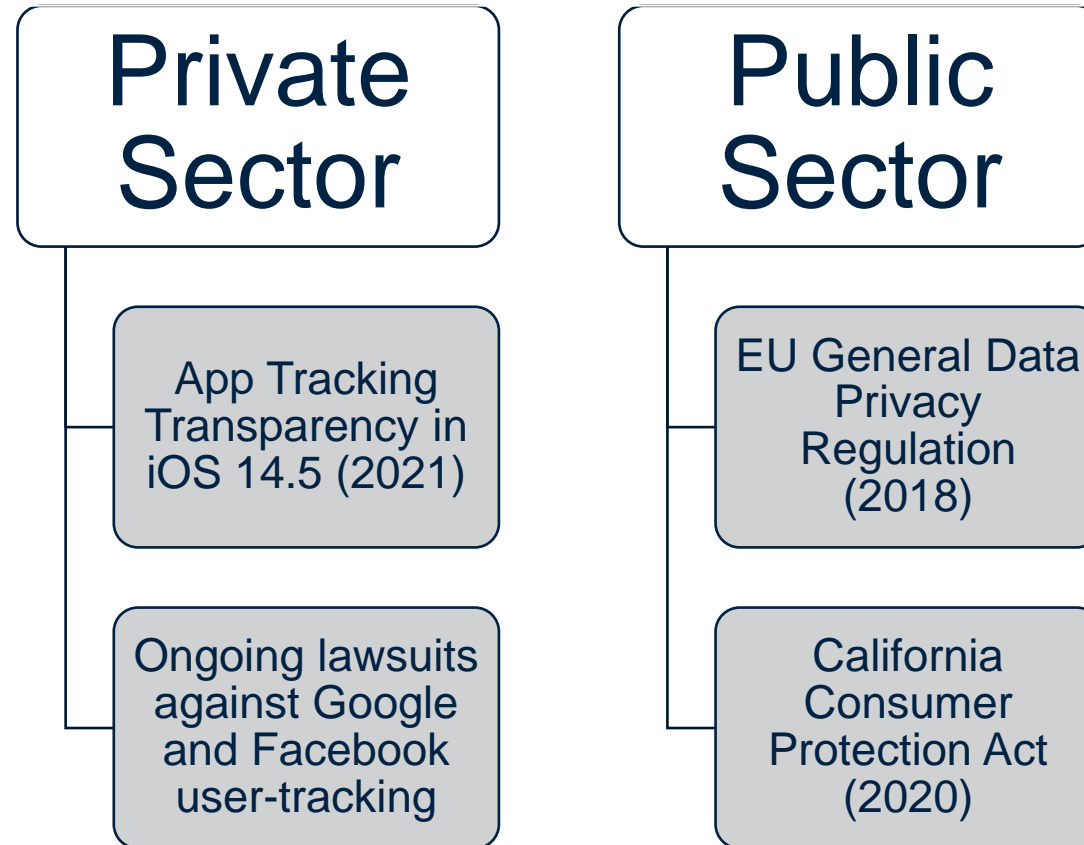
Takeaways

Lessons

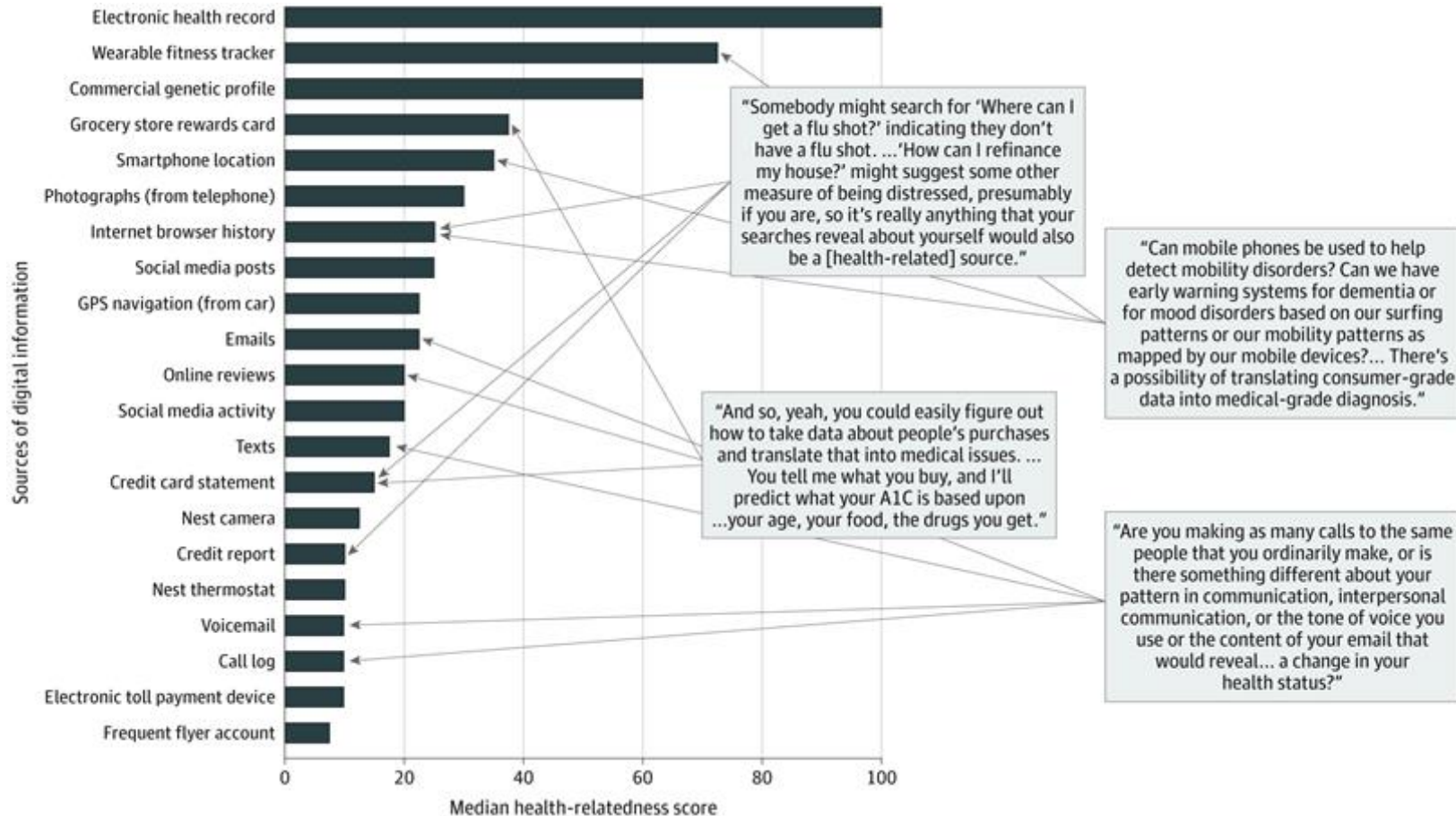
Closing Remarks

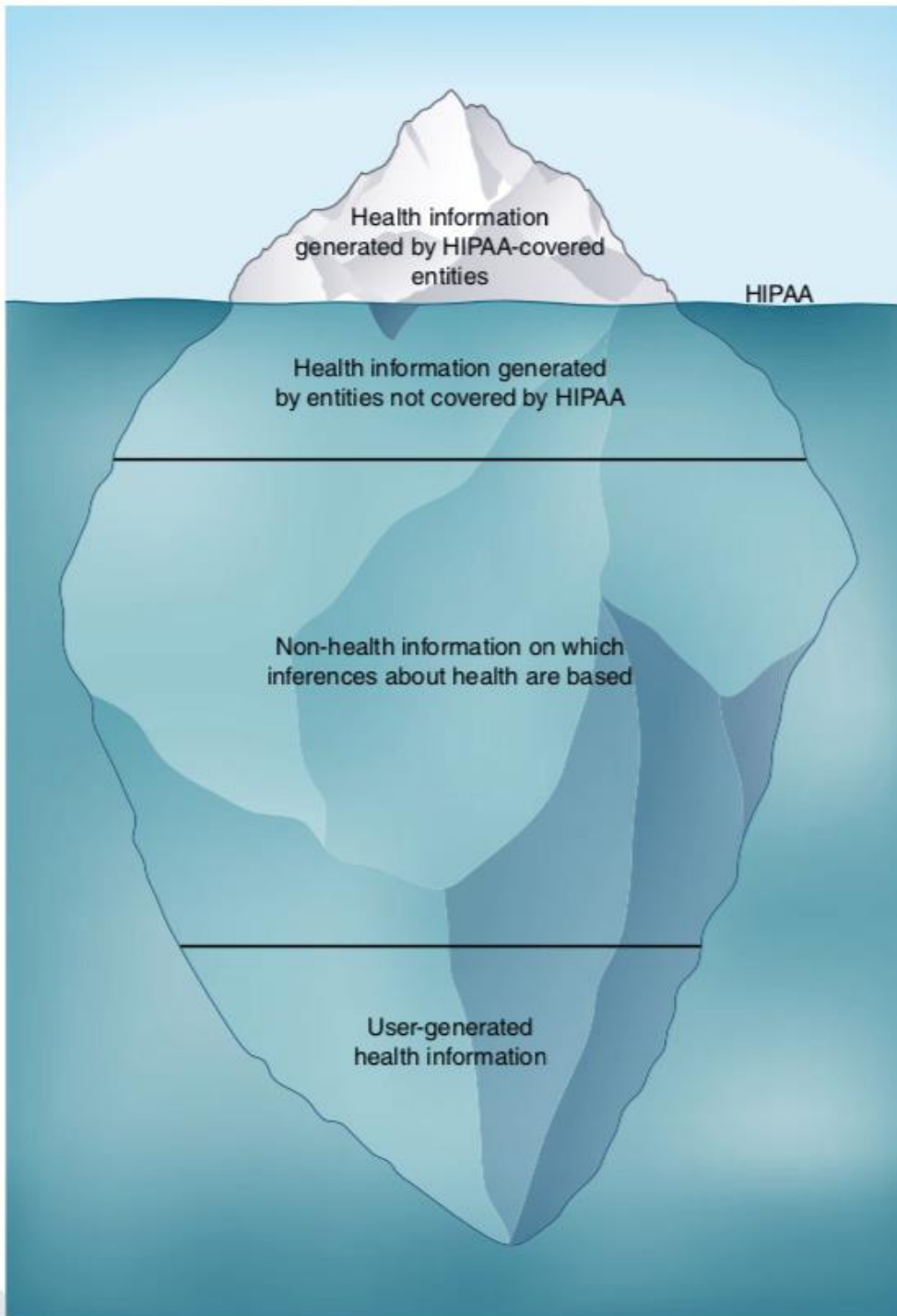


Consumer privacy rights are an increasing issue of contention.



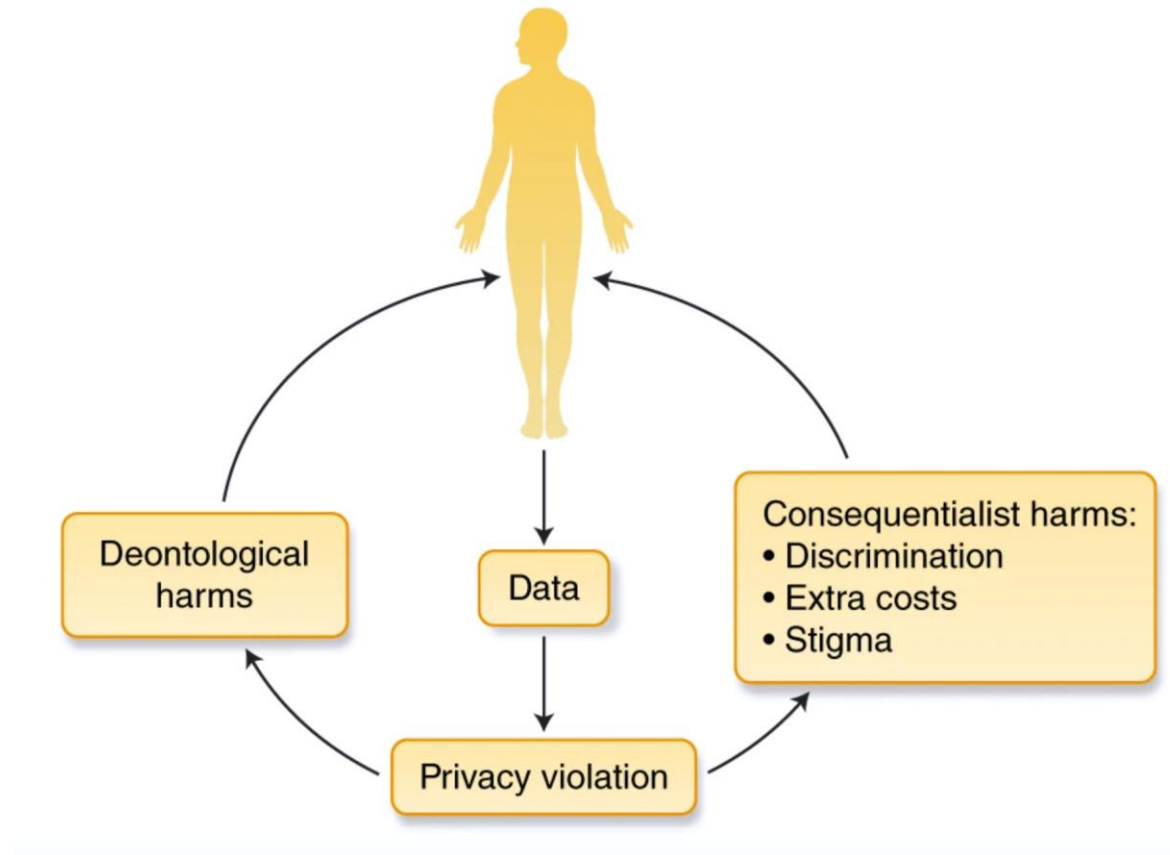
# A variety of activity contributes to our digital health footprint.



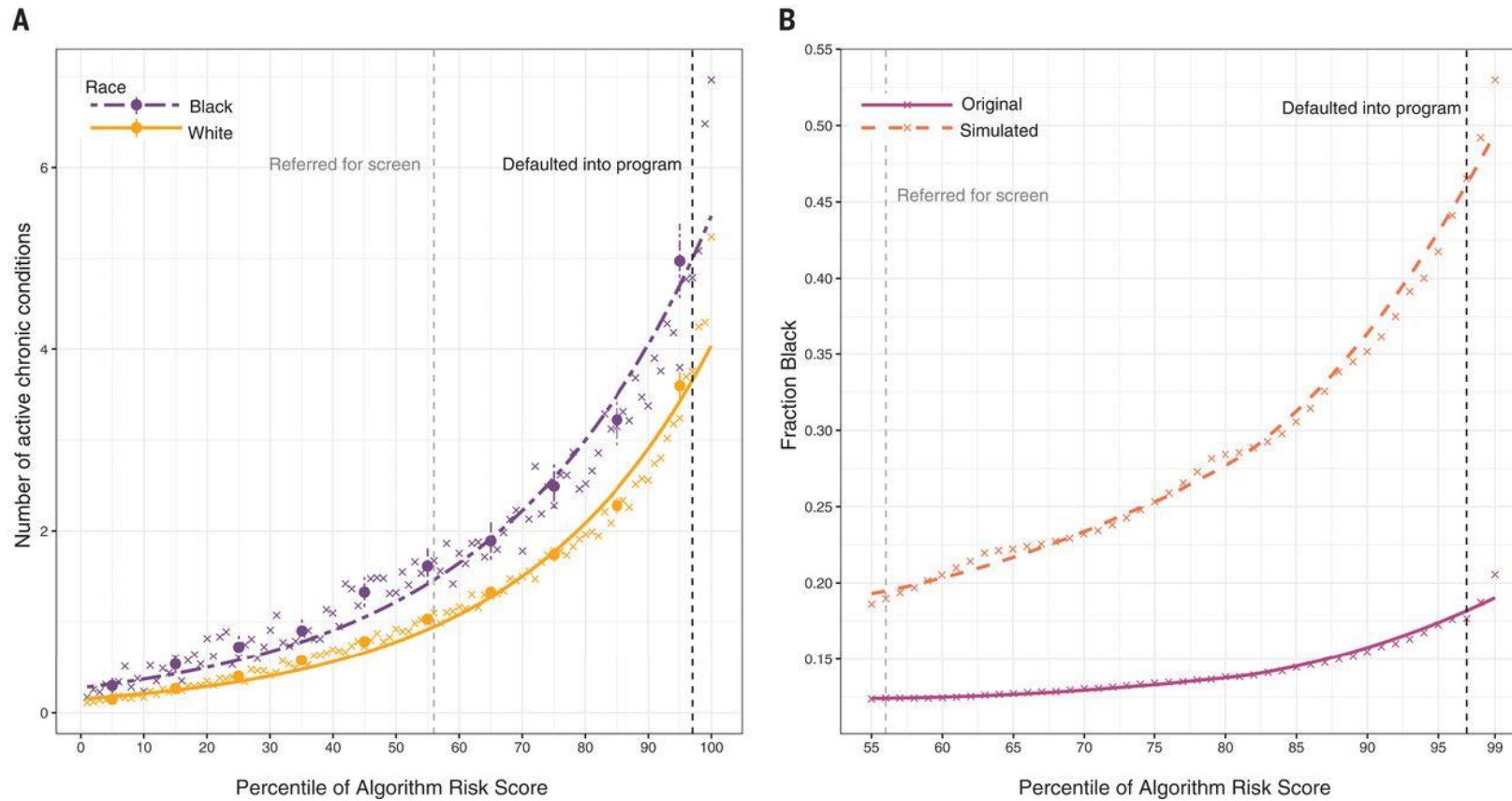


Our digital health footprint is particularly vulnerable due to algorithmic de-anonymization and limitations of HIPAA's scope.

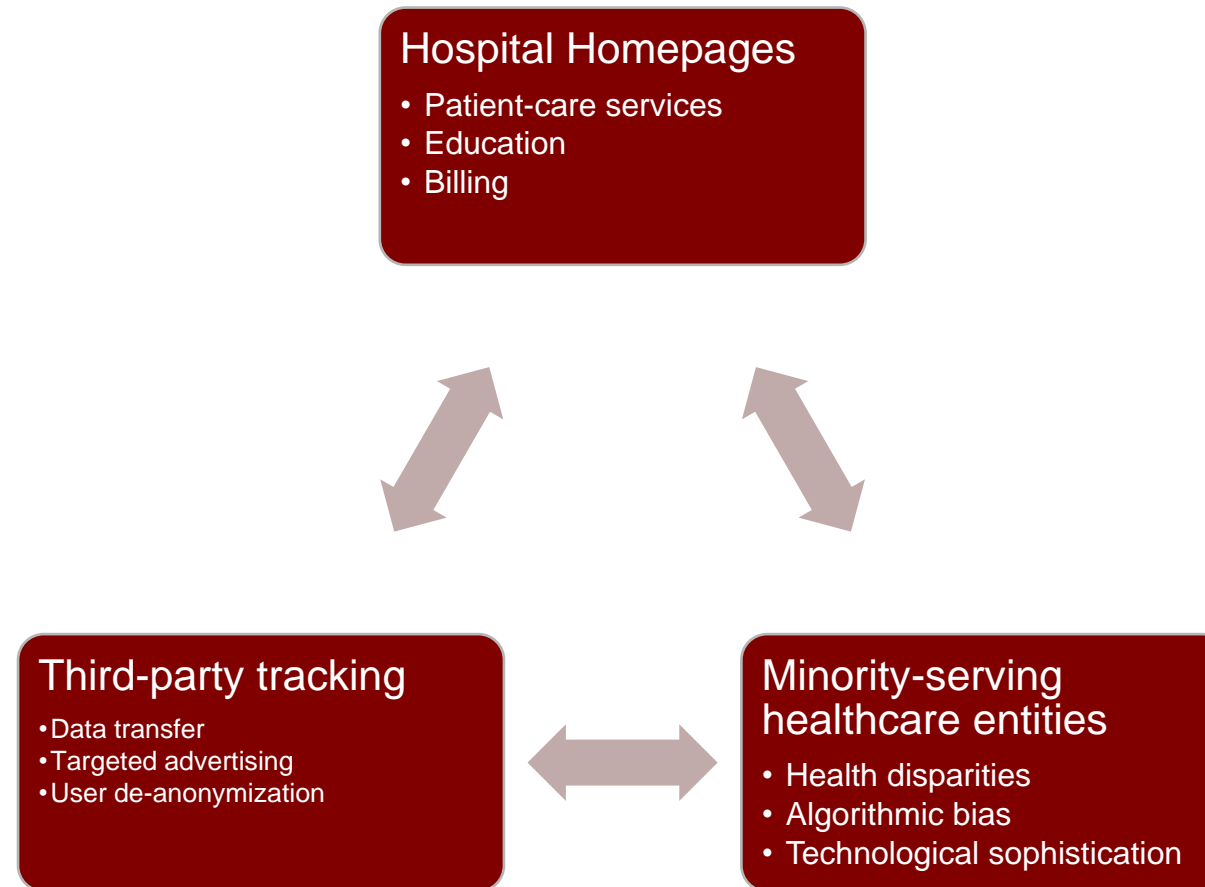
# Violations of digital health privacy pose both normative and tangible concerns.



# Algorithms informed by digital health data have been observed to perpetuate discrimination.



# The nexus of third-party tracking, hospital homepages, and minority-facing healthcare is of interest.



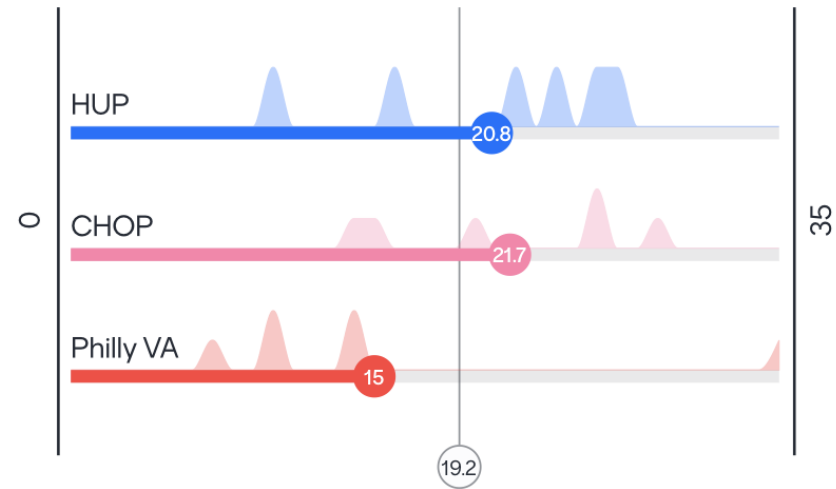


# A quick poll intermission!

Go to [www.menti.com](http://www.menti.com) and use the code 2132 0266

How many third-party data requests do you think the \_\_\_ homepage receives?

Mentimeter



# CONTENTS

Overview

Aims

Methods

Findings

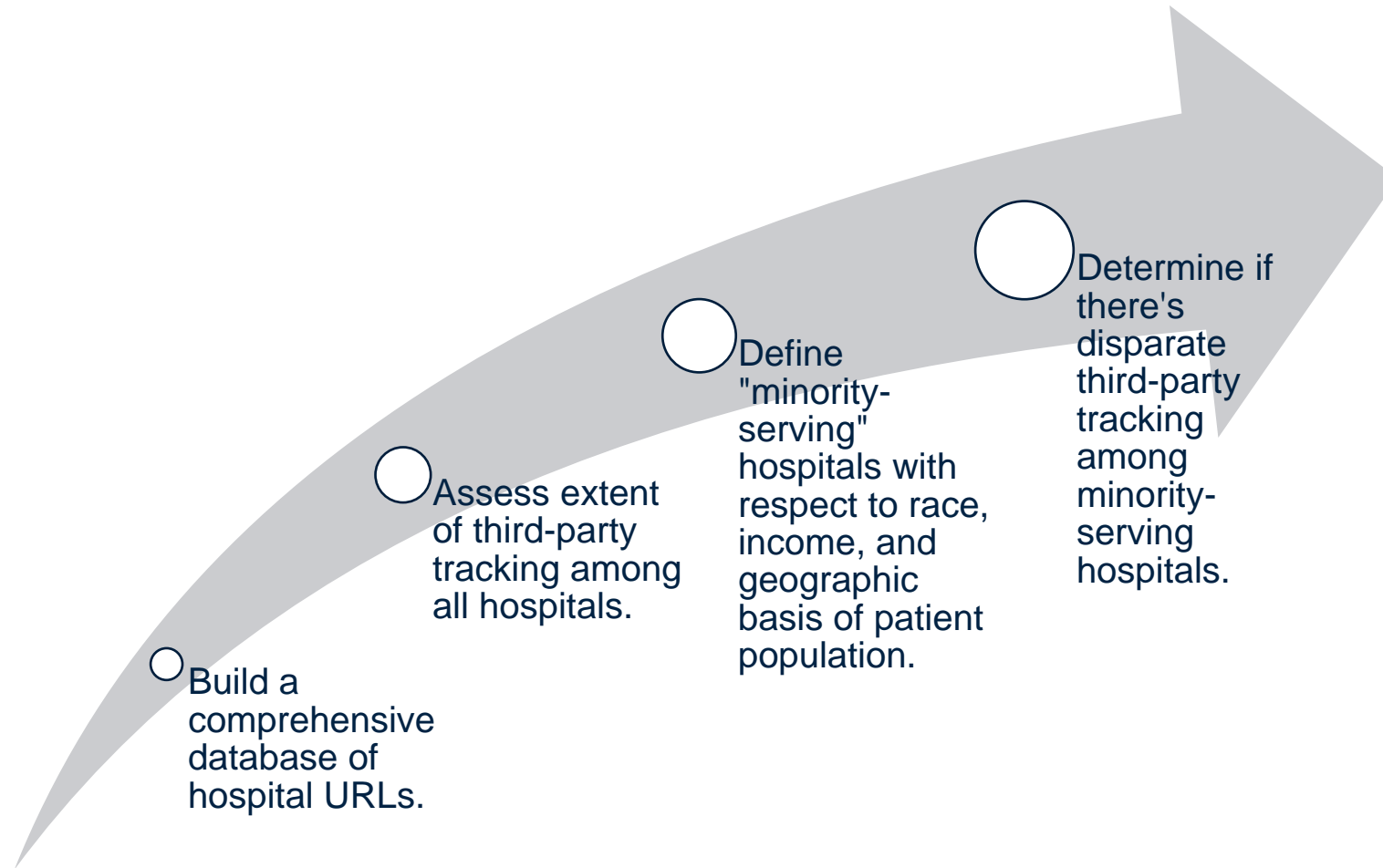
Takeaways

Lessons

Closing Remarks



# The aims for this project built off one another.



# CONTENTS

Overview

Aims

Methods

Findings

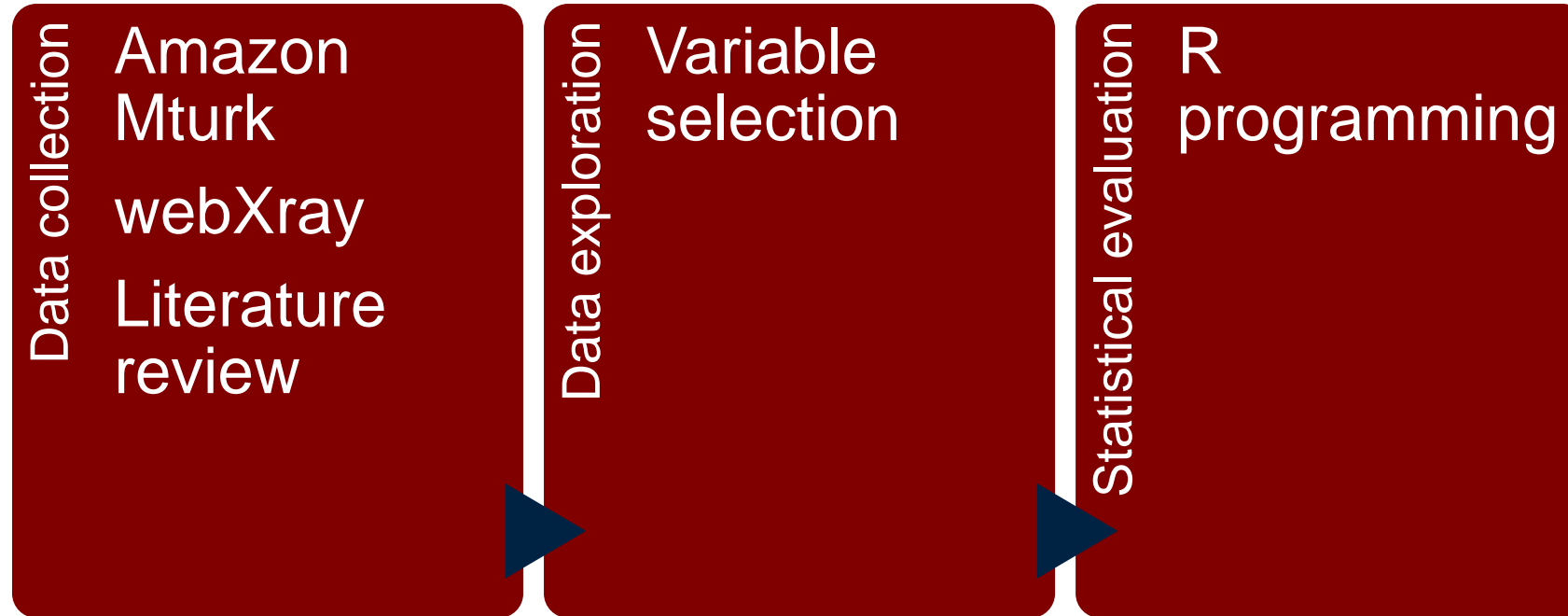
Takeaways

Lessons

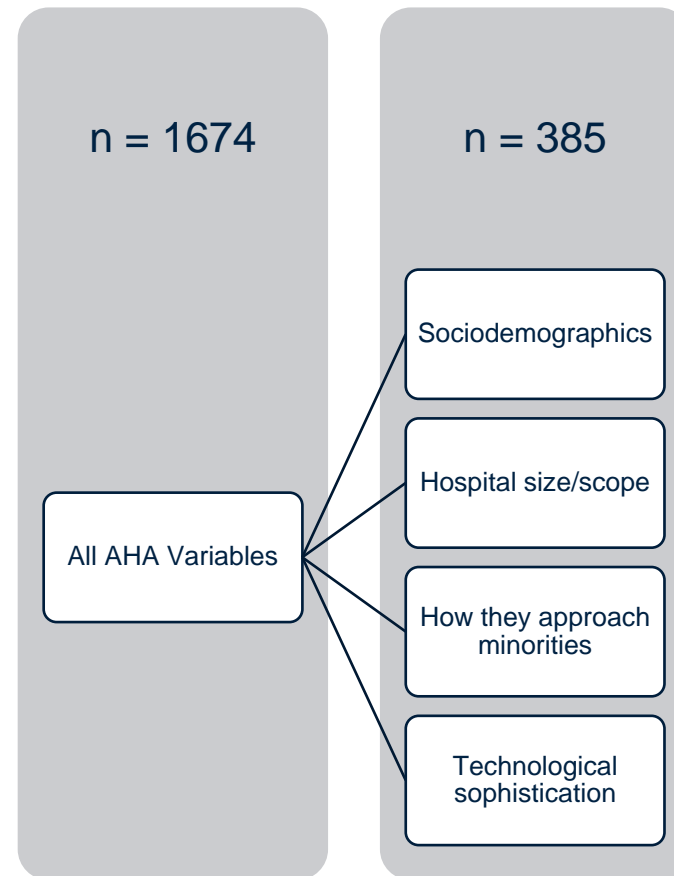
Closing Remarks



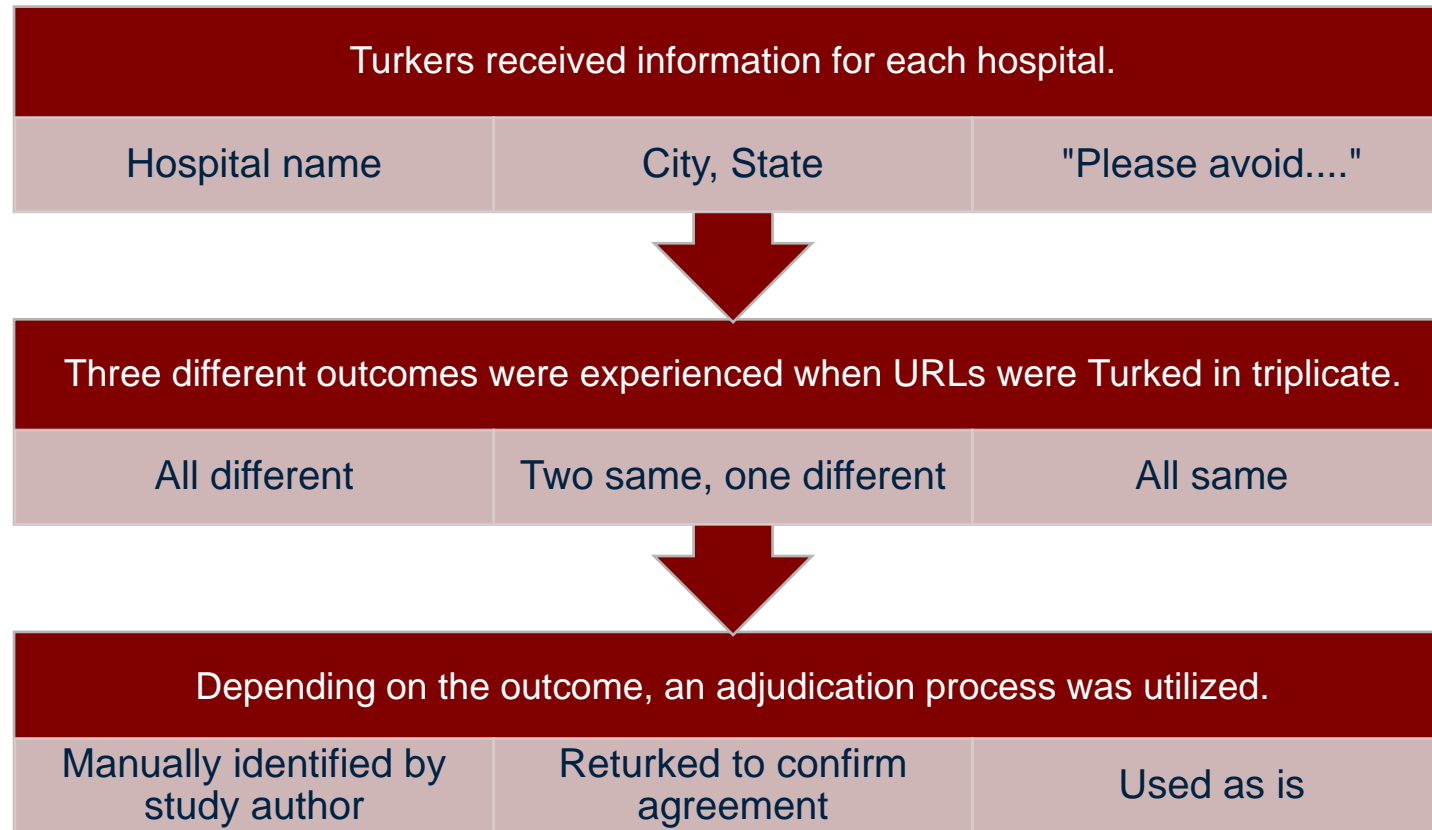
# Survey of methods utilized.



Manual variable selection was conducted to identify relevant variables in the American Hospital Association data set.



# Amazon Mechanical Turk was utilized to crowdsource hospital homepage URLs.



Webxray was run on collected URLs to obtain web-tracking info.

*“webXray detects third-party HTTP requests and matches them to the companies which receive user data”*

For each URL:

- URL of 3P request
- Request domain owner

For entire URL database:

- Aggregated tracking distribution
- Percent of URLs w/ 3P data requests & cookies



Relevant literature regarding minority-serving hospitals was examined to develop data subsetting criteria.

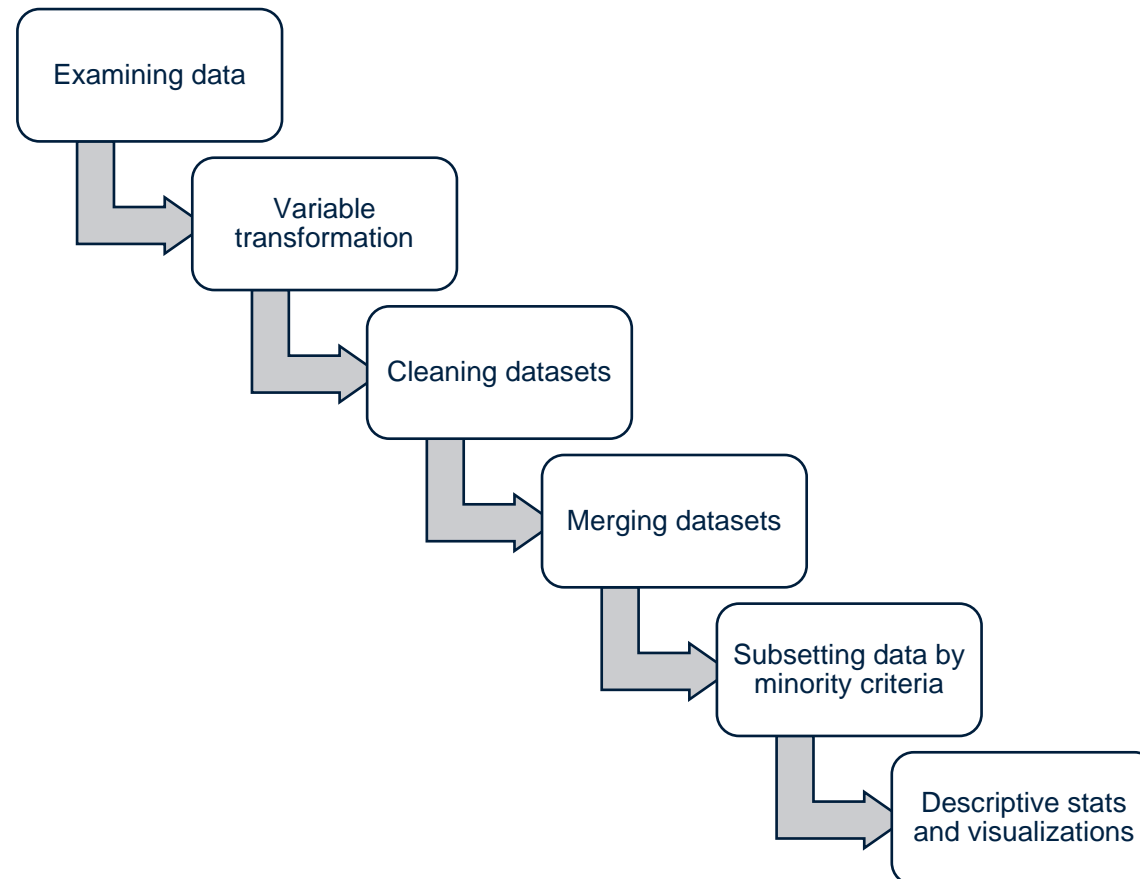
Identify papers addressing outcomes in “minority-serving” hospitals.

Isolate data sets characterizing various minority demographics.

Aggregate zip code data to obtain demographic make-up at an HSA-level.

Generate subsets of merged data using HSA population percentages for a given hospital.

Upon collecting the relevant data, various functions were executed in R.



# CONTENTS

Overview

Aims

Methods

Findings

Takeaways

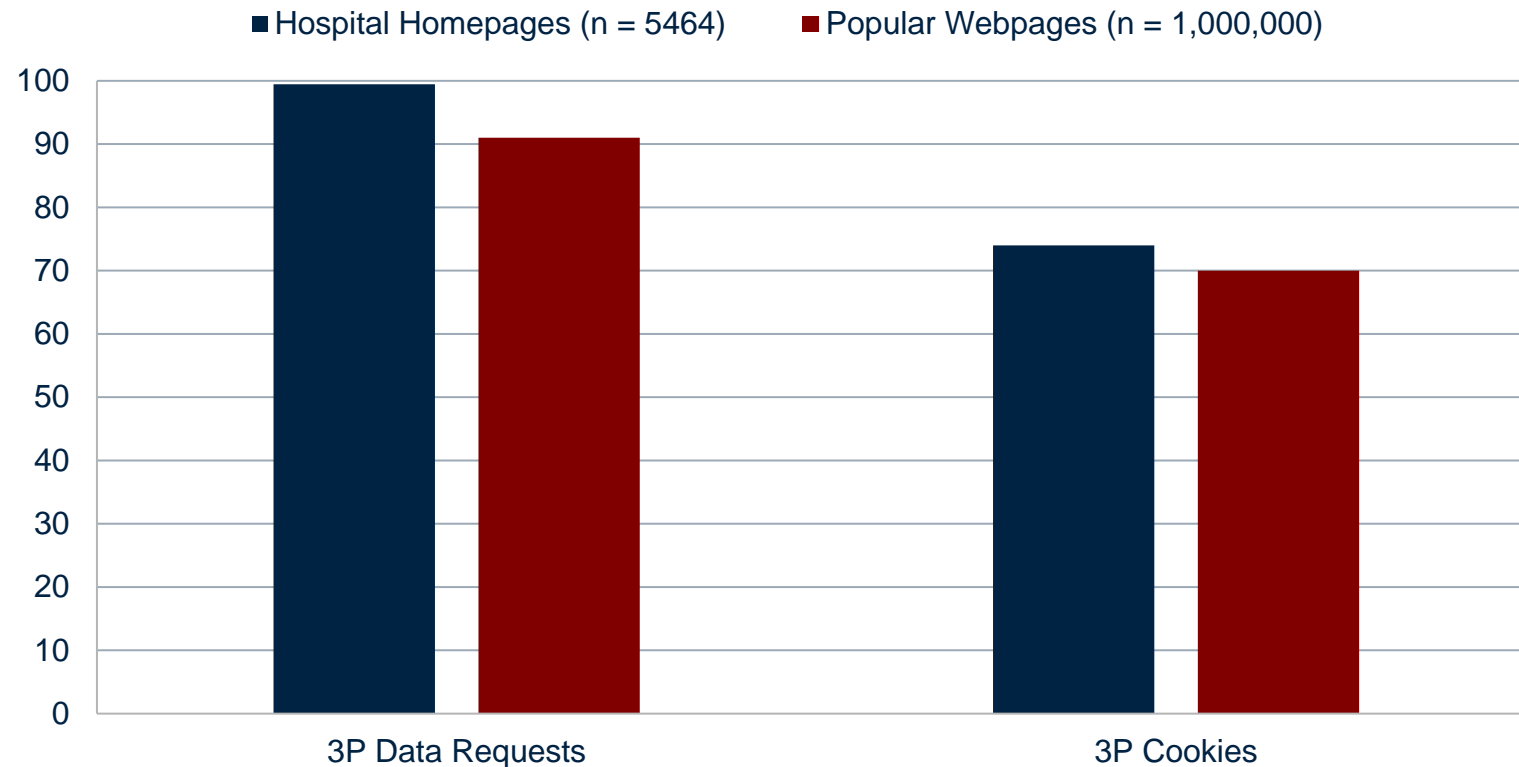
Lessons

Closing Remarks

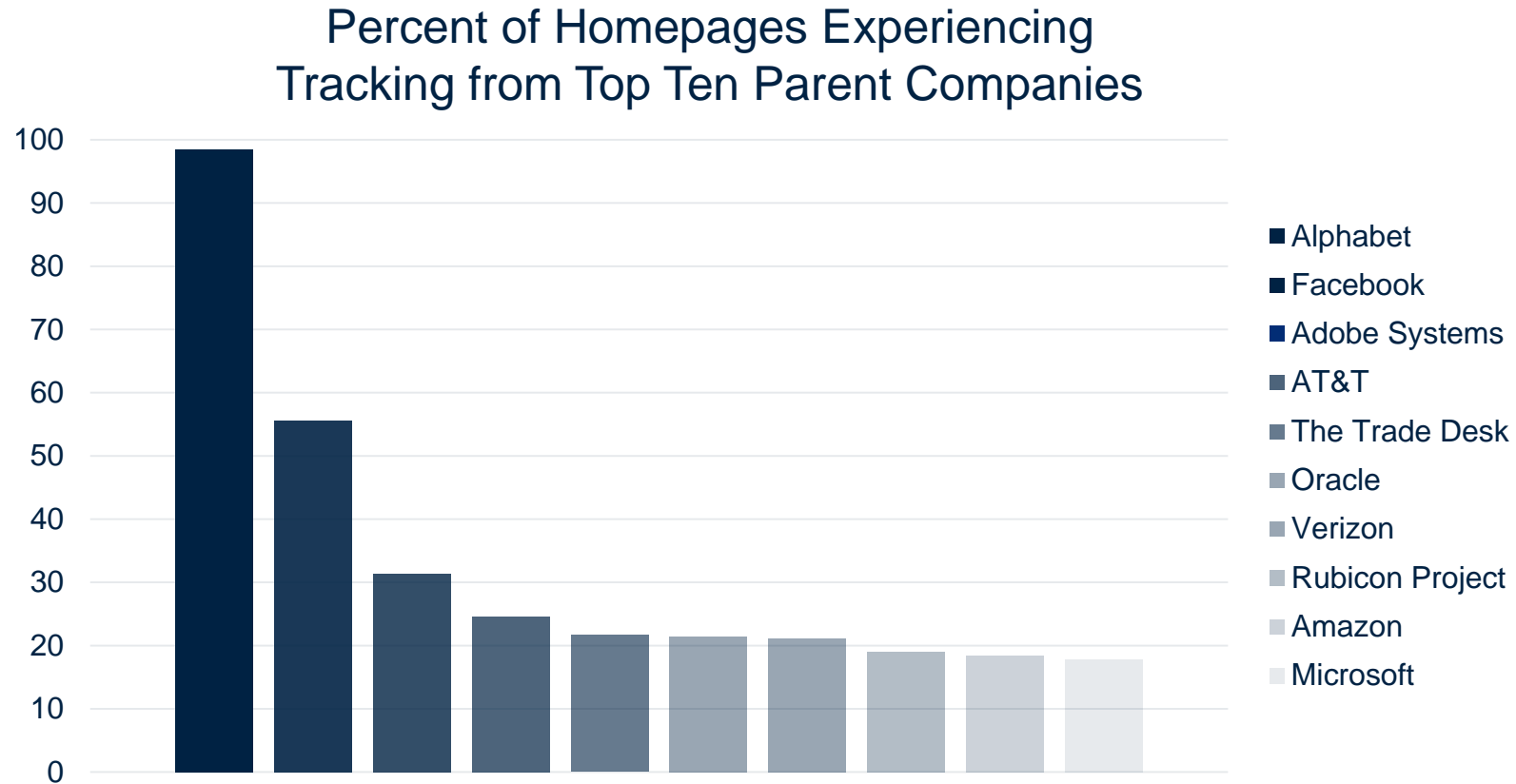


# 3<sup>rd</sup> party tracking on hospital homepages is prevalent.

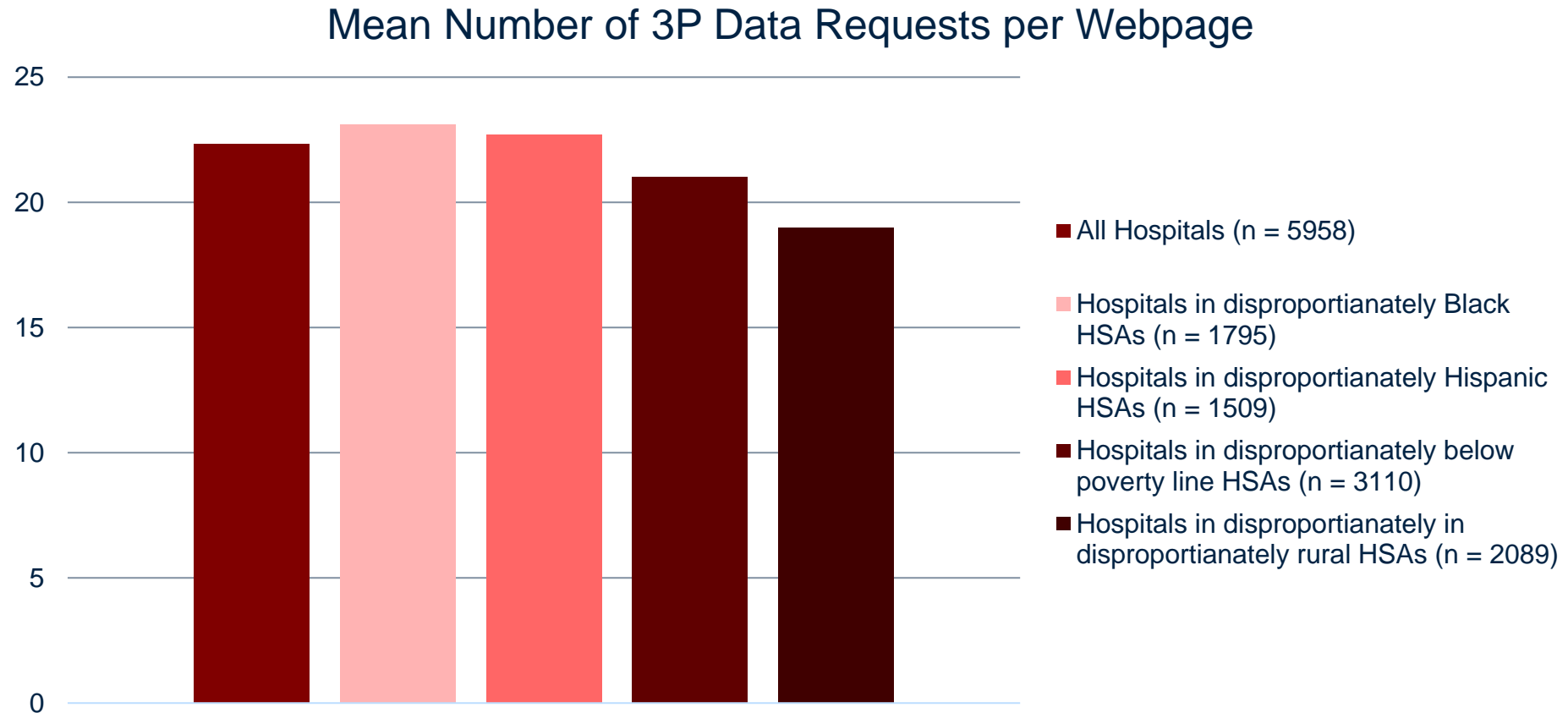
Incidence of 3P Tracking on Hospital Homepages Compared to Popular Webpages



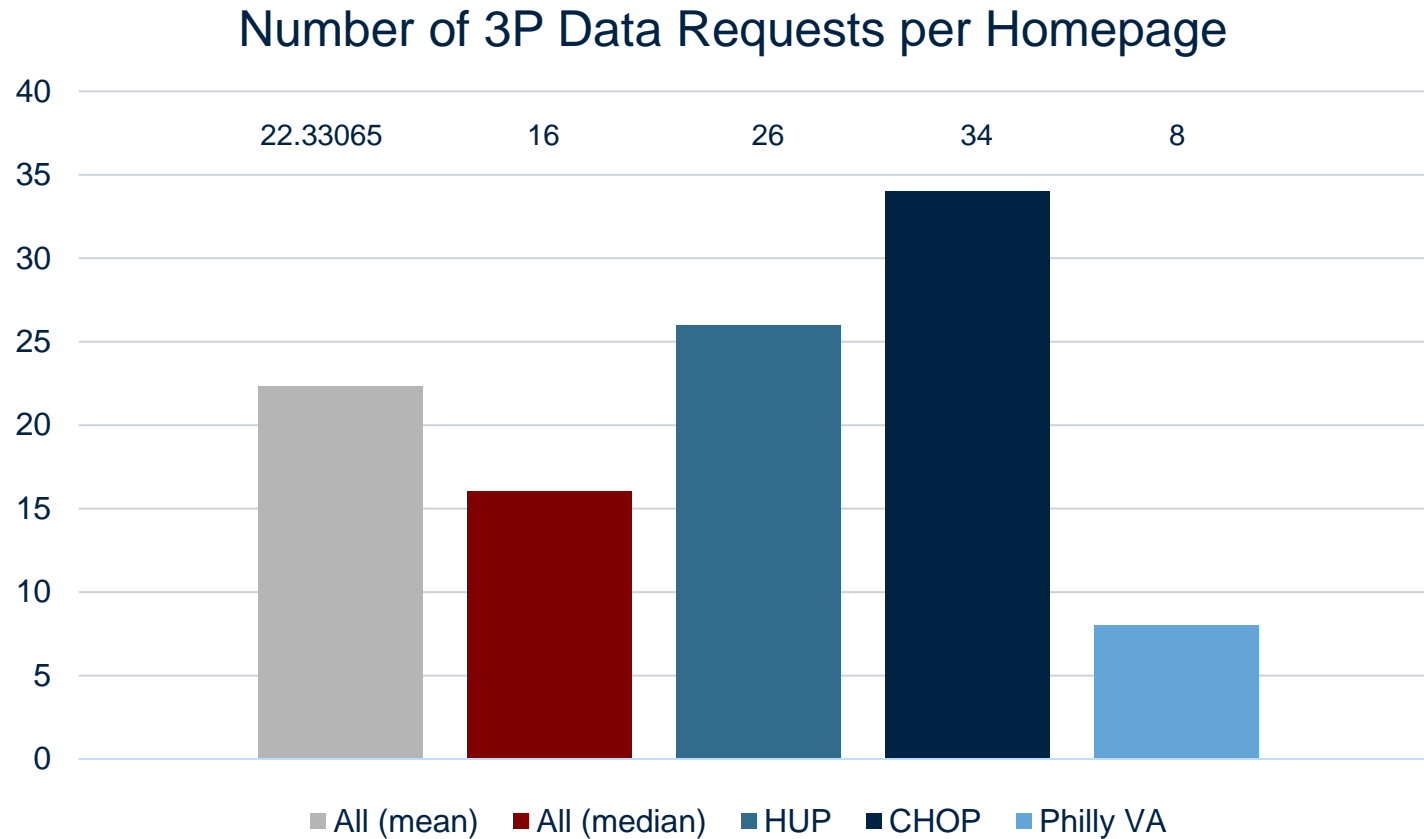
# Top top ten most prolific tracking companies.



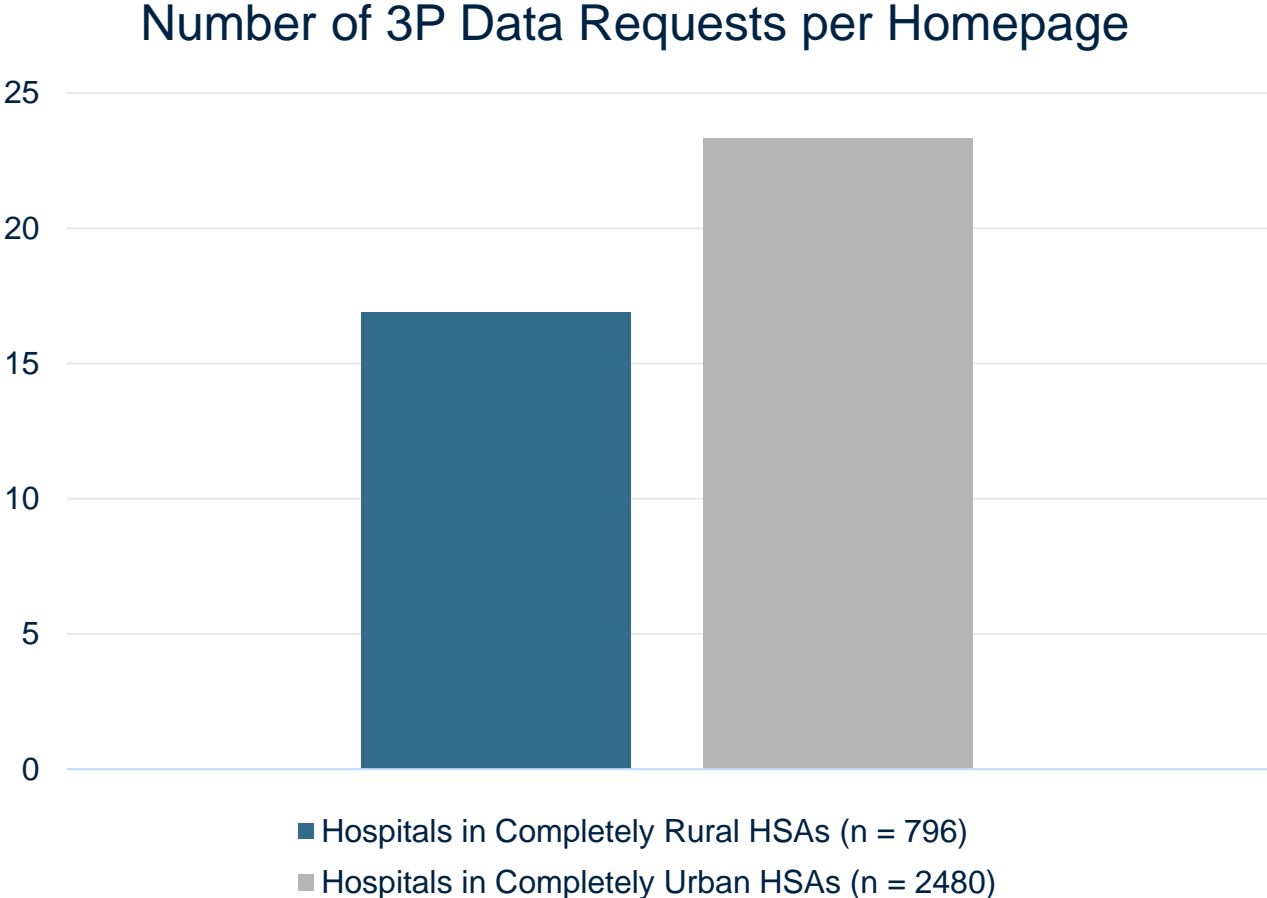
3<sup>rd</sup> Party data tracking exists on the homepages of minority-serving and non-minority serving hospitals alike.



# Where do our polled hospitals stand?



The disparity in tracking between rural and urban hospitals is of note.





# CONTENTS

Overview

Aims

Methods

Findings

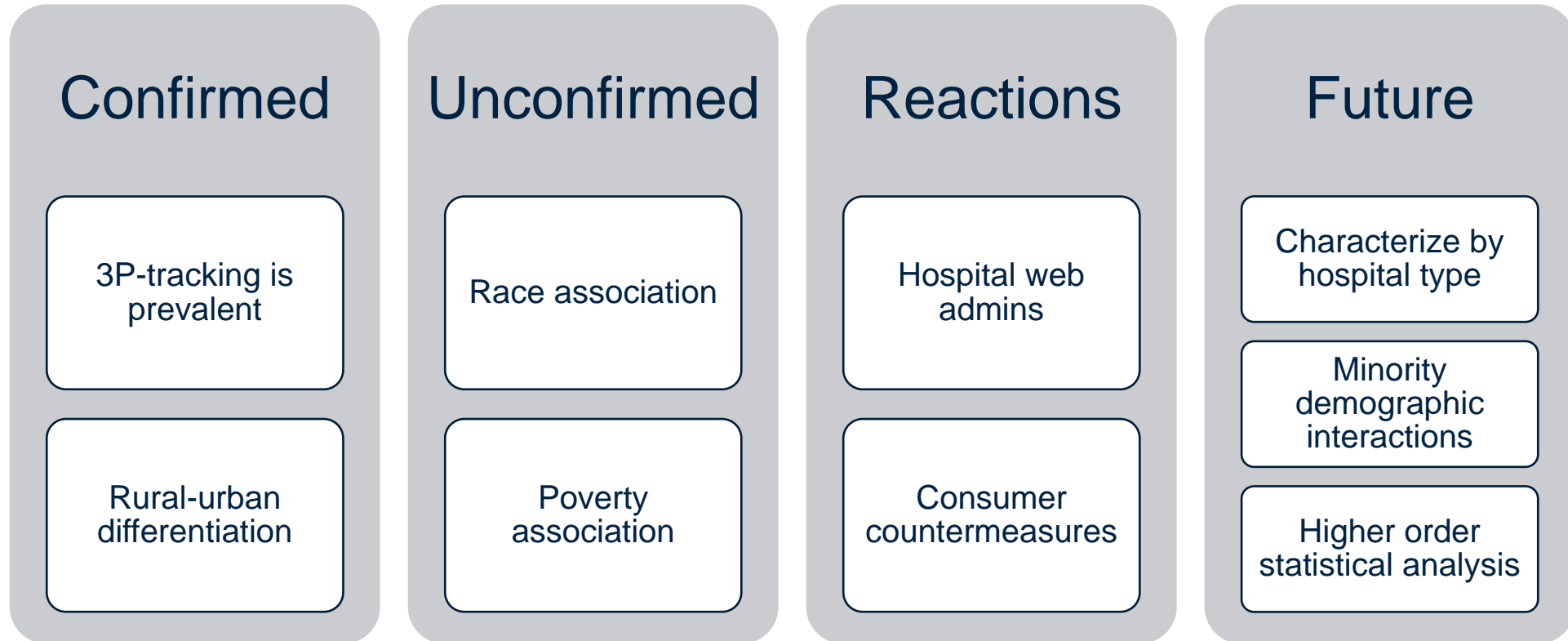
Takeaways

Lessons

Closing Remarks



These results affirm and contradict some of our initial hypotheses while opening the door for future analysis.



# CONTENTS

Overview

Aims

Methods

Findings

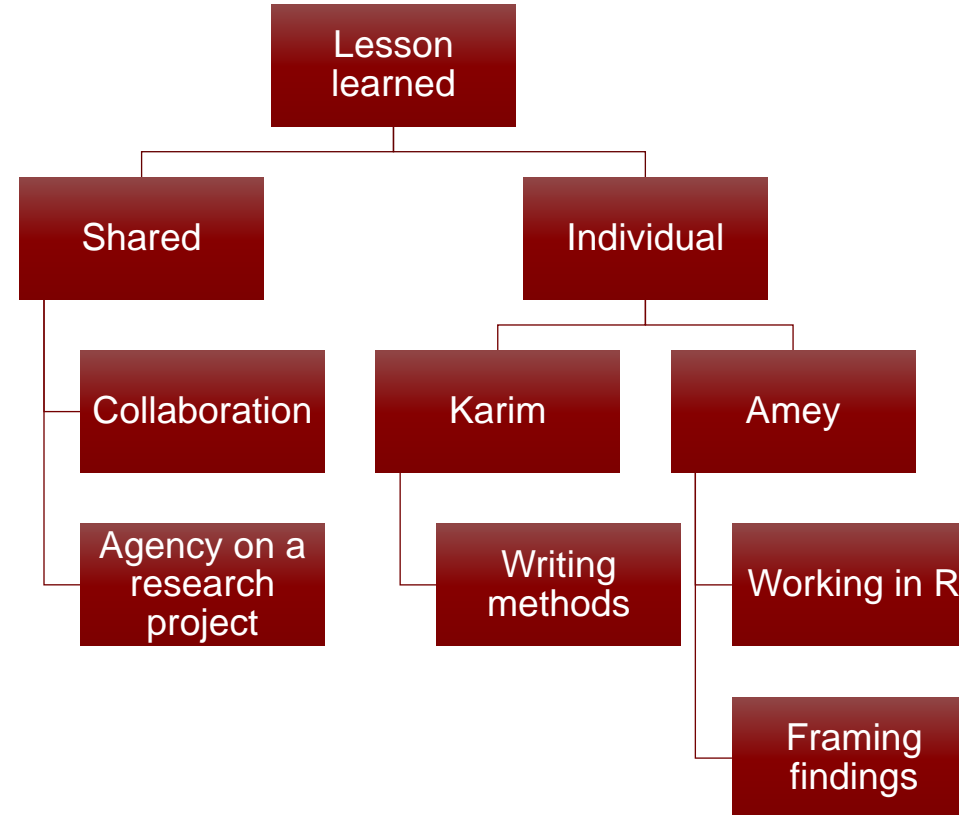
Takeaways

Lessons

Closing Remarks



This project was a challenging yet rewarding way to expand our capabilities as researchers.



# CONTENTS

Overview

Aims

Methods

Findings

Takeaways

Lessons

Closing Remarks



# Acknowledgements & Q/A

- ▶ Thank you to our amazing mentors Ari Friedman and Matthew McCoy!
- ▶ Shoutout to Joanne and Joanna putting together an amazing SUMR program.

