Hillel Krief February 16, 2022

Task 1.4

Part 1- Population Data

1

# Data Source

This is an external data source including location by county and population count by gender and age group. This data is owned by the US Census Bureau which is a government organization and is therefore trustworthy.

# Collection Method

Census data is a mixture between administrative and survey data. The data is collected manually from surveys but also taken from other government sources. There is no time-lag

# Data Contents

County, year, total population, male population, female population, and population by age.

2

The Data isn’t biased as its factual. It’s collected once per year. And because of its collection method, it can contain manual errors.

3

This data is super relevant since it has population by location and age range. This helps us identify the populations that are more vulnerable

Part 2- Laboratory and Patient Visits

1

# Data Source

These are an external data as the source is the CDC and therefore it is trustworthy. (The CDC is a government organization).

# Collection Method

This is administrative data collected by the CDC from places like: hospital, Public Health providers, labs, and other healthcare offices. This data is manually collected. Data is updated once per week.

# Data Contents

Lab tests has state, year, positive test, number of people, percent positive and which type of virus.

Influenza visits has state, year, week, total patients, age range, number of providers, and more

2

This is survey data and therefore it might not be complete as well as biased.

3

This data is not relevant since it is largely incomplete and outdated

Part 3 Children Flu Shot

1

# Data Source

These are an external data as the source is the CDC and therefore it is trustworthy. (The CDC is a government organization).

# Collection Method

This is administrative and survey data collected by the CDC from places like: National Immunization Surveys and the University of Chicago. The data is manually collected

# Data Contents

This data had 73 variables. Some important variables include flu shot data, age, state, demographic, financial status

2

This is partially survey data which the demographics not verified. However, the flu shot data is verified and accurate.

3

The data is not relevant as it pertains to children who don’t have much to do with our hypothesis