

Effects of Employment on Health

A Big Data Analytic

Authors: Christian Keitel, George Dagher, Khen Price

Motivation

To determine if there exists a relationship between being unemployed and experiencing negative health symptoms. We hope to capture this correlation by analyzing hospital stay information and cross-referencing that to employment statistics from New York State.

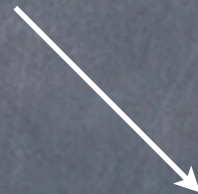
The Data

Measuring Health: Our main source for measuring health comes from a dataset that lists every patient accepted into the hospital and their length of stay along with the “severity” of their illness for all zip codes in New York.

Measuring Employment: The dataset we are using provides statistics on the percentage of people who are employed in various county’s verses those unemployed.

Data Flow

Inpatient Data Broken up by zip code.



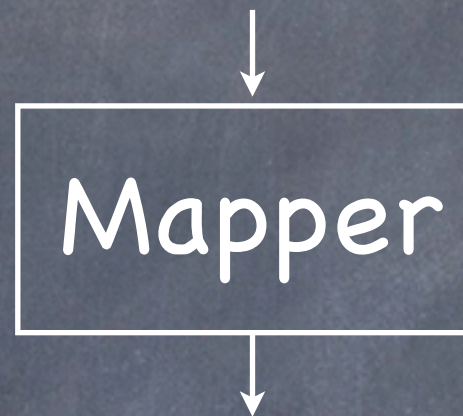
Grouping overall health and professions
by zip code will give us an indicator as
to whether a correlation exists
between the two.

Occupation Broken Up by zip Code

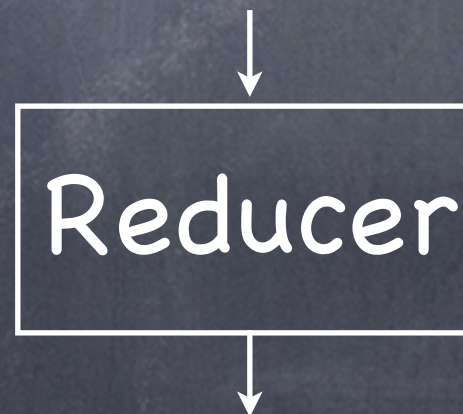


Hadoop Process

Hospital County, Zip Code Length of Stay, APR Severity of Illness Code, Total Charges



Zip, [Length of Stay, Year, Severity, Charges]



Zip, Year, Total/Avg Stays, Total/Avg Severity, Total/Avg Charges