Project Ideas

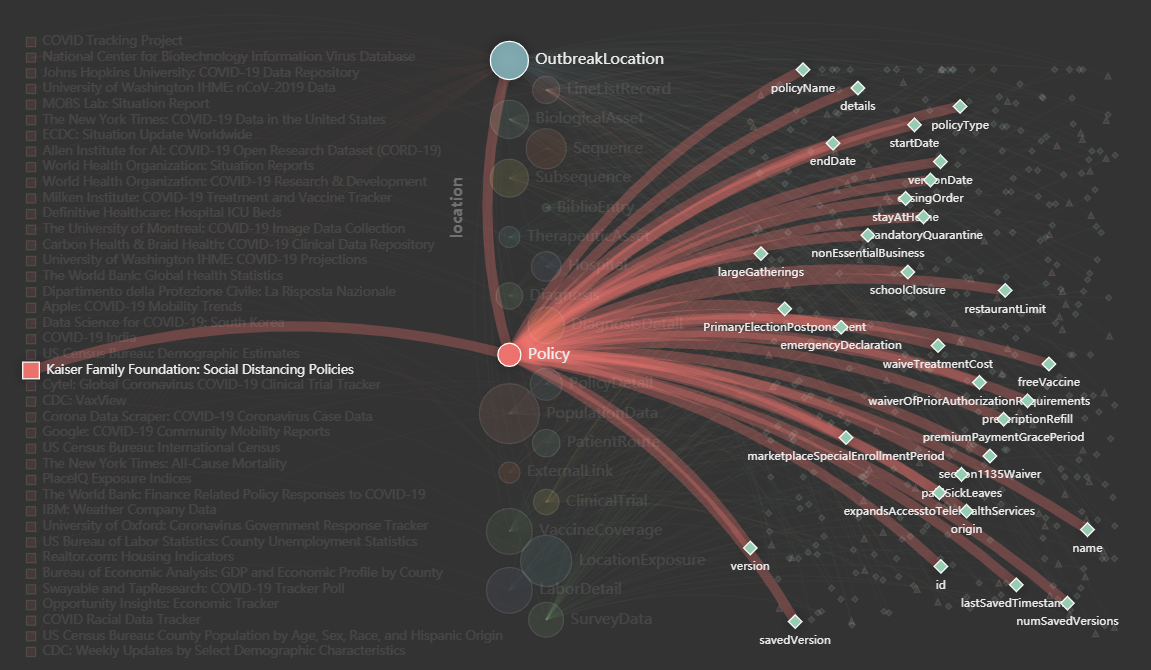
Idea #1: Baseball statistics. Datasheets and legend available on

<https://www.retrosheet.org/>

Idea #2: COVID-19 research using the C3 AI data

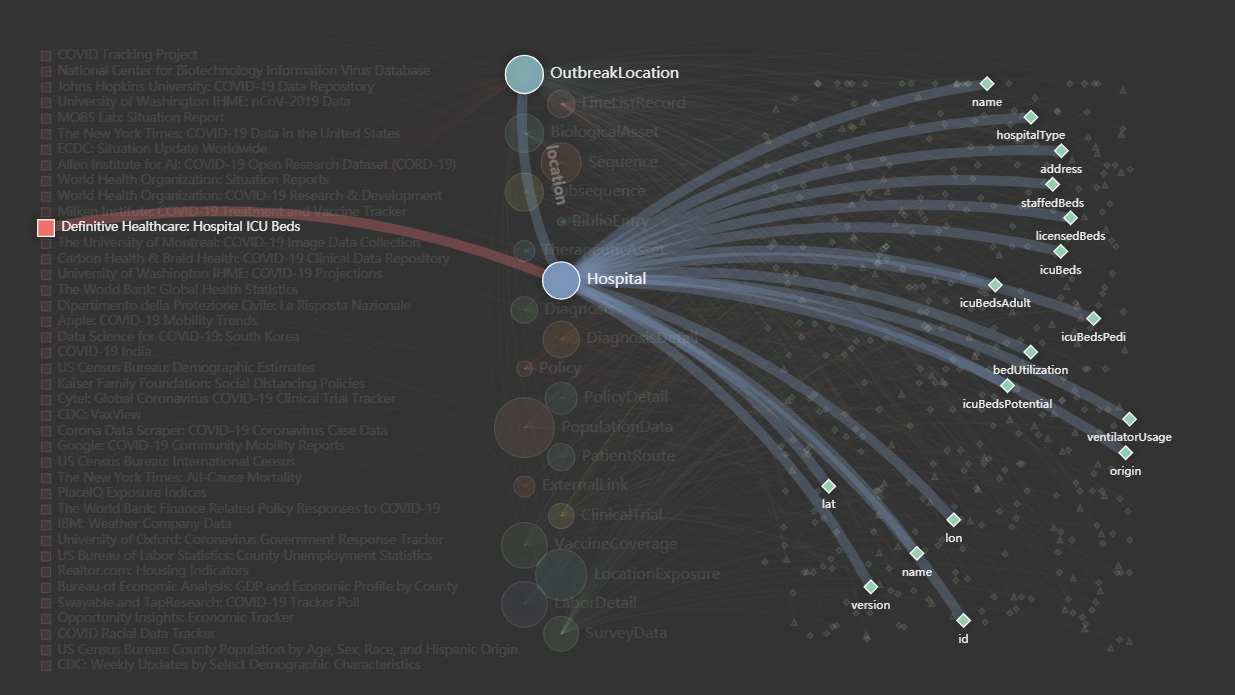
<https://c3.ai/products/c3-ai-covid-19-data-lake/> \

* Analyze public health policy decisions along with COVID death/case rate to indicate the relative effectiveness of different policies across different states in the US over time.



<https://github.com/KFFData/COVID-19-Data>

* Analyze optimum number of ICU beds per hospital depending on demographics (such as location, population, etc.)



Idea #3: Diabetes Research used by the CDC.

<https://www.kaggle.com/cdc/national-health-and-nutrition-examination-survey>

* Find the variable which effect the diabetes for people.
* Build multiple linear regression model to solve the problem about if variable effect diabetes what variable comban together will affected diabetes.
* Use different distribution to solve the diabetes problem and prove it make sense.

REQUIREMENTS:

1 Project Group

Please provide the following information regarding the project group members:

•A list of project group members.

Roman Sydorchuk

Spencer Sumner

Kacper Malysa

Liu Yurun

Mingxi Xia

•A designated project group leader.

Leader

2 Project Topic

Please provide the following information regarding the project topic area:

•Application subject area (e.g., astronomy, biology, social services, etc.)

Subject Area

•Specific data set and sources (e.g., Kaggle, Google Public Datasets, etc.)

Data Set

•Reference resources (e.g., Articles, Papers, Existing Works, etc.)

Resources

•Supplemental resources (e.g., Reference Datasets, Software, etc.)

Supplemental