Announcements
Covid Pandemic
Global Cases

New reported cases

- **All time**
- **Last 90 days**

<table>
<thead>
<tr>
<th></th>
<th>Cases</th>
<th>14-DAY CHANGE</th>
<th>TOTAL REPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases</td>
<td>2,615,566</td>
<td>+180%</td>
<td>313,620,083</td>
</tr>
<tr>
<td>Deaths</td>
<td>6,647</td>
<td>+3%</td>
<td>5,502,949</td>
</tr>
</tbody>
</table>

Dates: Daily avg. on Jan. 11
Disease Modelling

• Naïve Assumptions
  • Uniform Population
  • Three types of people: Susceptible, Infected, Recovered
  • Disease lasts one unit of time
  • Each Infected exposes $r$ other people
  • A person Susceptible person that is exposed becomes infected

Susceptible \rightarrow Interected \rightarrow Recovered
Model predictions

• Exponential growth for $r > 1$
• Decline in cases when $r < 1$
• New cases: $r \times |I| \times \left( \frac{|S|}{n} \right)$
• Effective $r$ value decreases as susceptible population decreases
• Decline starts when $r = \left( \frac{n}{|S|} \right)$
Making the model continuous

• Shrink the time interval and allow multiple time periods in I
• Parameters $\beta$ and $\gamma$ for transition from S to I and from I to R
• Time step to zero to make a differential equation

\[ \begin{align*}
\frac{dS}{dt} &= -\beta SI \\
\frac{dI}{dt} &= \beta SI - \gamma I \\
\frac{dR}{dt} &= \gamma I
\end{align*} \]
Expanding the models

• SIRD – Susceptible, Infectious, Recovered, Dead
• SEIRRD - Susceptible, Exposed, Infectious, Recovered, and Dead
• Adding Vaccines
• Adding Reinfection
Covid Epidemic: South Africa
Covid Epidemic: India
Covid Epidemic – CA, FL, WA
Where does this data come from

- Different country reporting strategies and aggregation
  - Tremendous variation on levels of accuracy and sources of case data
- Reported to centralized authorities
- Aggregators, such as JHU curate data sets
  - Labor intensive – group of research assistants collect daily data
- Data made available for download
  - Github or other sources
Data repositories

• Johns Hopkins University
  • Source for majority of dashboards
• Our World in Data (Our world in data)
• US Specific sources  CDC
• Other countries
  • https://coronavirus.data.gov.uk/
  • https://sacoronavirus.co.za/

• Global data: WHO
  • https://covid19.who.int/
Data types

• Case counts and death counts
• Vaccine delivery
  • Quantity of vaccines delivered to countries is known
  • Reporting of number of immunizations is fairly good
• Covid variants – percent of different variants detected around the world
• Covid restrictions – time scale of restriction by geography
• Excess deaths
• Country demographics and maps
Many dashboards and studies are already available

- What is left to be done?
- How to compete against New York Times or Johns Hopkins University or Institute of Health Metrics and Evaluation
- Sources of data exist
  - Possible to build on top of existing data sets
  - Infrastructure exists to work with large data sets
- Identify specific directions that are not components of these existing tools
- Tools give very good overviews and summaries
  - Opportunity is doing deeper analysis: combining data sources and refining geographical analysis
Applications of modeling

• Obviously, predicting the future
  • When will Omicron peak in Seattle?
  • This quarter Omicron will sweep round the world
  • So tools would need to be designed for Omicron – but ready for Pi, Rho, and Sigma

• Matching SIR model against previous waves
  • Picking out previous waves is a start!

• Tying modeling to other data sets
  • Vaccination, Public Health Restrictions
Geographic refinement

• Predict and understand the epidemic across geographic areas using data from sub regions (e.g., county level data in the US)

• This is missing from the aggregation sources – which give good summaries but tend to be “one dimensional”

• Sub-national data is often not on aggregation sources (which was part of the motivation for one of the project areas)
  • There will be technical challenges in building appropriate data tools

• Reasons for paying attention to subnational data
  • Identify geographic structure in events
  • Correlations between different data can be stronger at the subnational level
Exploring the interaction of data sources

• Significant opportunities to investigate correlations between data sources
  • Rural-Urban vs impact
  • Cases vs death rate vs variant
  • Variant vs wave vs impact
  • Vaccine status vs public health intervention vs impact
  • Season vs climate vs wave

• Recommendation
  • Pick a subset of factors with plausible relations and build tools around good data sets
  • Option of emphasizing either tool building or data exploration
  • Map based tools or other visualizations could be included
Existing tools (from fall project)

- Available on local github, as csv files
  - Time series processing of JHU case count data
  - Decomposition of time series into waves with statistics
  - USA county adjacency map
  - CDC vaccination data (by county)
  - Co-variant data by country