* An explanation of the central idea behind your final project (what is the idea? which datasets do you need to explore the idea? why is it interesting?)

As we are in a severe situation caused by the Cov-19 virus, a question is how we can know better about it. One way to tackle this problem is to compare it with other infectious diseases that we already know, so we could have a better picture of it.

The dataset includes several famous viruses like Ebola, SARS, H1N1 and so on. By comparing metrics like infected cases, death cases, the death rate, we could find out which disease Cov-19 most similar to and if it is really the worst pandemic the last 2 decades.

* A mock-up of the visualization that you wish to build. (Anything is fine here. Pen and paper, MS Paint, Inkscape, D3, anything.).
* Make sure you answer the questions
  + What genre is it? (for *Genres*, see section 4.3 of the Segel and Heer paper).

In this preliminary phase, we will classify the visualization outcome of our projects as a combination of genres. To be more specific, we will present the visualizations by using annotated graphs(genre 1) within a slide a slide show(Genre 2) format.

* + Why is that genre right for telling the story you want to communicate with the data

We aim to built an author and reade driven visualization and thus we will use annotated graphs to convey our messages but we will, also implement interactive slide show which will allow to the reader/user to explore the results according to his/her will.

* An outline of the elements you'll need to get to your goal./The implementation plan.

1. Organize our data, since we collect several datasets, a preliminary step is to combine them together and extract columns that we need for the next step.
2. A quick data-analysis on the dataset to get an idea about the data. For example, the size of the data, how many samples in each category and so on.
3. Data statistics and visualization. we would select a couple of models in Bokeh to show the comparison.

* A walk-through of your preliminary data-analysis, addressing
  + What is the total size of your data? (MB, number of rows, number of variables, etc)
  + What are other properties? (What is the date range? Is it geo-data? then a quick plot of locations, etc.)
  + Show the fundamental distributions of the data (similar to the work we did on SF crime data for lecture 3)

93389 rows of data include Date, Country, confirmed cases and death cases among several diseases.