TeamDK

Dolev Artzi
Raymond Li

Domain

- The domain of framework is numerical data that has to do with countries. We will be using OurWorldInData.org · Apiary as an API to interact with our world in data's database. Additionally, our plugins will support local CSV files extracted from https://data.un.org/Default.aspx .
- These types of data sets involve quantitative features, such as life expectancy and categorical data, such as countries within a continent (obviously).
- The framework will involve translating data into visualizations that can display comparisons between two features- one quantitative and one categorical. We will also provide the ability to focus on one country over a specified time period.
- There is a lot of room for reuse, since it is quite common to display data that compares counts of a field based on a category.
    - Bar charts
    - Pie charts
- The framework should be able to support datasets in both csv and json (through API) formats. Two different data plugins will be implemented to do this- one that is specific to the API, and another that parses a csv converted dataset from a different source.

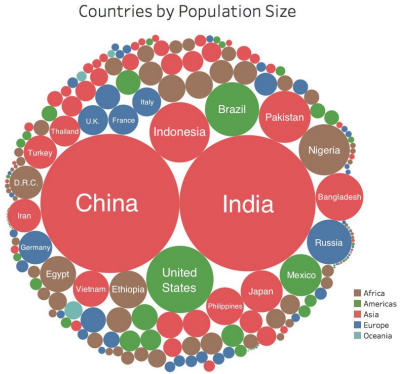
Response Body Real Diff Specification

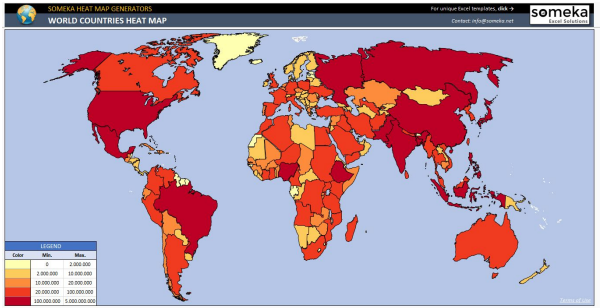

```
{
  "success": "true",
  "data": [
    {
      "id": "3905",
      "value": "391.3",
      "variable": "Household income",
      "from": "1700-01-01 00:00:00",
      "to": "1799-12-31 00:00:00",
      "entityId": "46"
    }
  ]
}
```

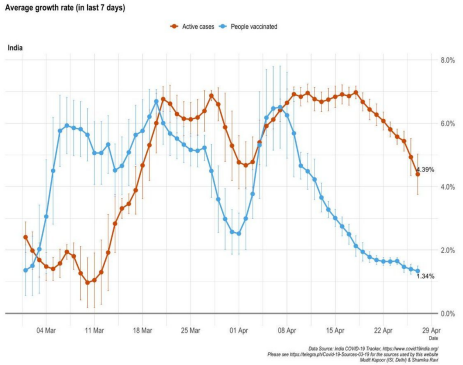
Generality

- The framework is unique in that it displays information about multiple kinds of data: categorical and numerical. This is very generalizable. Our data opportunities are expansive, but still within the scope of country-related data. For example, data could be the current population/GDP/crime rate of each country, or could be any of those statistics for a given country over a time period.

- The visualizations of choice will tentatively be a bubble graph and a heat map- as it is a unique type of graph while still being relevant to the data plugins used. We have also considered a time series plot, and the group plans on implementing two of those visualizations.







Plugin Interfaces

Split into three components- data, visual, framework

- Data extraction
  - importData()
  - getOptions(ArrayList<String> options)
  - extractData(String dataType)
  - extractData(String dataType, String startYear, String endYear)
- Visualization
  - filterByCategory()
  - chooseData(), chooseDataByField()
  - setDataType(String dataType)
  - chooseDataInRange(String dataType, String startYear, String endYear)
- Framework - integration and data cleaning
  - assignNullValues(), removeNullValues()
  - Setup -> onRegister(), onEnd()

The Idea

The data plugins are able to retrieve information about numerical values associated with countries. The data will be current unless a time range is specified. The user selects the data type that they want to see information about. Unless otherwise specified, it will be for all countries, but if a specific country is selected then it will show that data point over time for that country. Our framework will more or less be a pipeline data → visualization, with the potential processing being sorting countries into continents and computing values across categories.