Homework #3

CS 5665, Fall 2016

Task: Visualization

- To begin, you should identify a dataset of your choice and ask a compelling question of the data. As a starting point, here a few datasets (and lists of datasets):
  - AWS Public Data Sets
  - Internet Data Sources for Social Scientists
  - Quandl, a search engine for datasets
  - http://www.kdnuggets.com/datasets/index.html
  - https://github.com/caesar0301/awesome-public-datasets

  Feel free to use your own choice of dataset, so long as the data is non-trivial.

  Based on the data of your choice, you should ask and answer a compelling question via a data visualization of your choice. You may clean, manipulate, and transform the data as you like. You may integrate multiple datasets, or none at all.

  Ultimately, you should produce:
  - A single image (JPG, PNG, or GIF) that effectively communicates the data. You may use any data visualization toolkit you like.
  - A brief explanation describing your design. (a few paragraphs, 1/2 page or so).

  Your explanation should justify the design choices you have made. Why did you choose the particular visualization type? What motivated your choice of size, color, and scale? Why are they appropriate for what you are trying to communicate?

  Your explanation should document what you are trying to communicate, the benefits of your choice of visualization, as well as any downsides to your viz (e.g., are certain connections downplayed due to your choices?).

  Answer: For this assignment, I have selected a dataset from the Cornell university website "Internet Data Sources for Social Scientists". On this website there were lots of datasets available out of which I picked "Foreign Currency Exchange Rates". This data is a historical series from the Federal Reserve Board which is a daily spot exchange rate of 24 countries for the duration of January 4, 2000 to October 7, 2016.

  On this data set there was no requirement for data cleaning, manipulation or transformation as the data is already filtered and clean and clear. I have integrated the data for all countries into a single sheet along with the "Country" column for the

purpose of categorizing the data. For implementing the visualization, I used “Tableau”. Tableau selects the size and color which is more important to compare various countries in the chart.

The dataset consists of 3 columns:

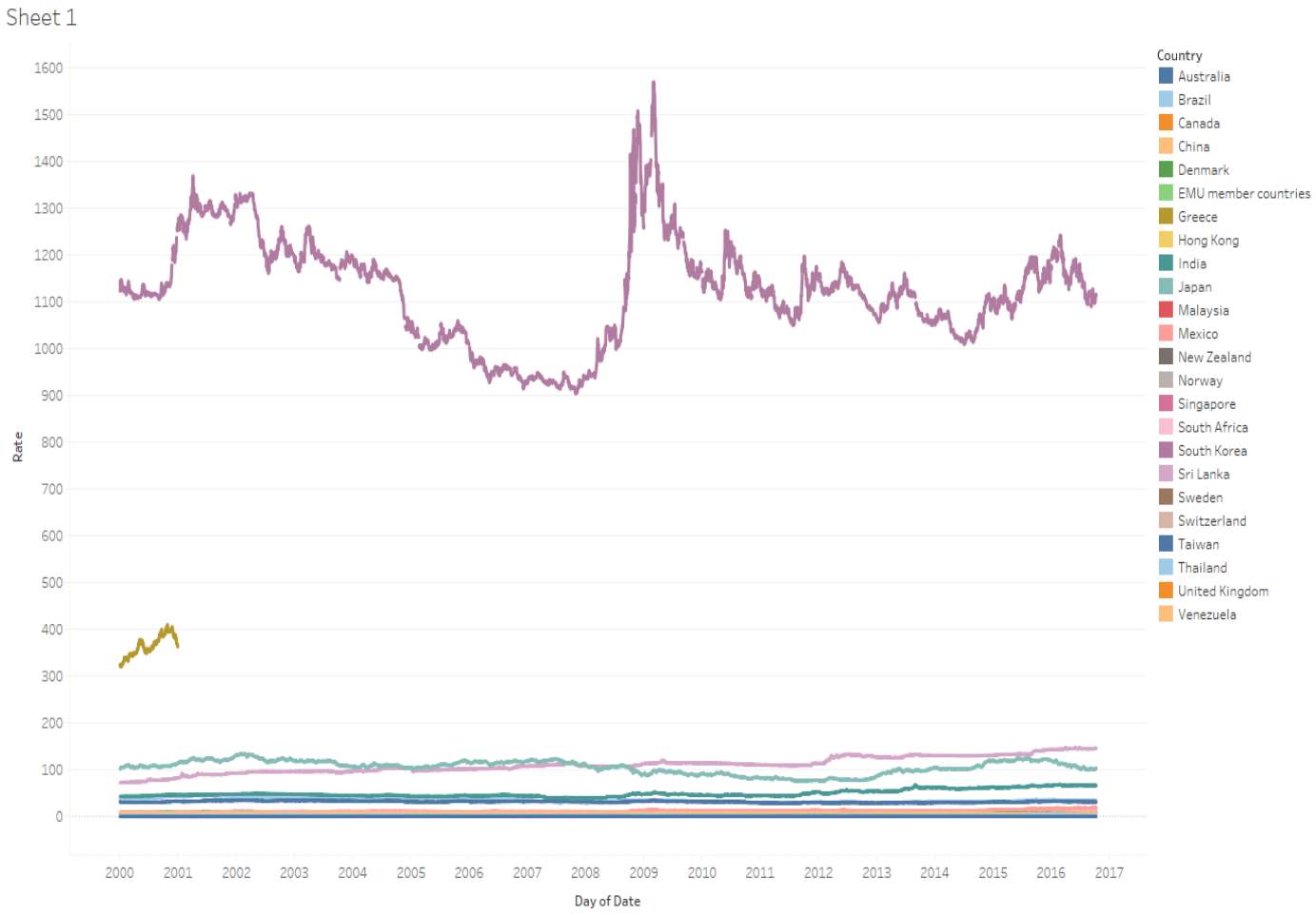
- Date
- Rate
- Country

For this data we need to compare the rate change over the time for different countries. Data of comparison type can be represented in following types of charts:

- Column charts
- Bar charts
- Circular Area chart
- Line Chart

As the rate change was to be observed over the time or duration starting from year 2000 to 2016. The time duration was of very small interval I preferred picking the “Many categories Line Chart”

In this chart different colored lines represent the change in different countries currency exchange rate value over the time from yea 2000 to October, 2016. This comparison could be very useful for understanding the economy variation of countries over the time.



In this chart I am trying to demonstrate the comparison of currency exchange rates of various companies. From the above chart we can infer following points:

1. South Korea has the highest exchange rate for the complete duration.
2. In 2008 & 2009 The value of exchange rate in South Korea increased a lot while the similar behavior cannot be seen for other countries.
3. Greece had foreign exchange rate only till 2001. After that they stopped exchange rate.
4. For Japan the exchange rate value reduced in the year 2008-09.
5. We can infer that the increase in the exchange rate value for Sri Lanka is almost linear and straight line.