

# Data Visualization

## Final Project Proposal

Ilya Ilyankou

All three parts will be based on the same dataset, Mapping the World Prices 2017 by Deutsche Bank Markets Research.

### 1 Plot Cities on a 2D Plane

Considering each city as a point with 30+ parameters, use one of the dimensionality reduction techniques to plot the cities on a 2D plane. It will be interesting to see whether cities from the same region are clustered together, whether there are any outliers, or if all cities are scattered randomly.

In addition, implement both PCA and SVD and see if two plots are similar.

### 2 Display City Data on a Map

Represent each city as a circle on the map. The diameter of the circles will change with time. For each time  $t$ , redraw all circles so that their diameters are proportional to the value of the *parameter* $[t]$ .

This is like visualizing time series, but instead of displaying values of the same parameter over time, we display values of different parameters for the same year (2017). In my opinion, this is a great way to spot outliers in some categories: for instance, London will be represented by a big circle most of the time (its GDP per capita, average salary, transportation prices are high), but the human eye will quickly notice if the circle suddenly gets smaller. Users will be able to pause and see which parameter it is.

### 3 Compare Two Arbitrary City Variables on a Map

Although correlation does not imply causality, being able to compare any two variables (such as salary and level of happiness) may give some insight and food for thought for the researchers. By displaying one of the variables by the size of the circle, and another one by the intensity of a particular color, one can spot if two different parameters may be related.