Assignment 1
Cs4406

Submitted by-Yash Pandey
17317629
D-stream

Answer 1-

One of the most interesting data visualizations that i came across over the web is " After babylon", designed by Density Design Lab it offers easy, fast and interesting way to understand al the different languages spoken around the world and where each language is spoken, it's origin and by who is it spoken. Not only this it even shows the relation between different languages such as language families, common words,etc. It is visualised using a collection of maps and graphs using world atlas of language structures. There are lots of languages in the the world around 3000 that we know of, it's very to internalise how many exactly are there, their influence on each other, their prevalence to the world and where they are spoken and hence this creative visualisation serves up an insight that educates, inspires and provides a useful resource to aid such problems.

http://www.puffpuffproject.com/languages.html

Answer 2-

One of the most complex data sets is high frequency trading data. The microstructure data that we can say is the typical driving data behind High Frequency Trading is available directly from the exchanges that is the details of every order placed, every execution and cancelation and thereby it is possible to reconstruct full limit order book both of past and in real time. Two challenges that we will face are scaling and interpreting of this data. Scaling is a more technological challenge as everyday gigabyte of data is generated but the main challenge is of interpretation of this data, that is what features we can extract from this granular data to build a predictive model for the trading problem in hand. We can use machine learning tool to visualise this data perhaps there will always be a human in this loop but the important thing is that machine learning application is absolutely necessary given the enormous size and complexity of the microstructure data that confronts us today. Through the use of informationally and computationally efficient algos we can infer good predictive models for such large data sets. It would greatly affect all the individuals involved in trading, whether its investors, traders , etc.

References- Machine Learning for Market Microstructure and High Frequency Trading by Michael Kearns Yuriy Nevmyvaka

One of the map showing languages spoken in different parts of the world