**Master Economic Data**

Team members: Lina Xiang, Lei Zhang, Gong Zhang, Nengjun Zhao

1. **Introduction of Project**

* **Introduce economic statistics**

After economic crisis in 2008, all countries that had been impacted are trying their best to recover their economy. If we look into five years economic data of these countries we can have a clear picture of each country’s economic development situation. From the economic statistics, we can easily discover the impact of the crisis to each country. If the country’s economic system got hurt seriously, or even destroyed by the economic crisis, we may consider this country’s economic system was not strong enough, without any development or improvement it may be destroyed again by another economic crisis. Taking this into consideration, I prefer not to choose these countries to start my career or business.

* **Interpret economic data properly**

Some economic metrics have been repeated thousands times, but they are often misunderstood by common people. Interpreting economic data properly is the very first step before retrieve useful information from it. Our application targets to help our users to have a clear and accuracy understanding with data analysis done in our system.

* **Extract information from economic data**

We first give consumers an overall view of worldwide economic change after economic crisis 2008. Then we present detailed information of the US economic data, which includes unemployment rate, detailed information about job market and average wages.

* **Use economic data to benefit your future**

We will not only show the economic data of high view, but also present the employment and wages of different industries and positions. These data is more meaningful for graduate students for reference in their career development decision. Users can find out the opening hire number in different industries for each month of a year and the average annual income of different positions. Looking into these charts, users will find out the answers for many questions, such as when should they graduate, which industry should be focus on.

1. **Data Collection**

* **International Monetary Fund (IMF):**

<http://www.imf.org/external/index.htm>

“The IMF promotes international monetary cooperation and exchange rate stability, facilitates the balanced growth of international trade, and provides resources to help members in balance of payments difficulties or to assist with poverty reduction.” (overview of IMF) As an international economic organization IMF have an authoritative and full database.

* **BEA and BLS**

U.S. Department of commerce, Bureau of Economic analysis,

<http://www.bea.gov/>

United States Department of Labor, Bureau of Labor Statistics:

<http://www.bls.gov/>

These two data analysis websites are running by US Department of Commerce and US Department of Labor. We got unemployment and labor statistics from them. We are sure we got the most accurate data.

* **Organization for Economic Co-operation and Development, OECD.StateExtracts:**

<http://stats.oecd.org/Index.aspx>

“OECD iLibrary is the online library of the Organization for Economic Cooperation and Development (OECD) featuring its books, papers and statistics and is the gateway to OECD’s analysis and data.”(overview of OECD) It is not a main data resource database for us, but a supplement one.

作为补充

blng

0

1. **Data Selection**

We focus on economic and labor statistics, and selected the following data for our main research objective.

* Gross Domestic Product (GDP): the market value of all officially recognized final goods and services produced within a country in a given period of time.
* Consumer Price Index (CPI): a measure of the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services.
* Population
* Unemployment Rate: a measure of the prevalence of unemployment (people are without work and actively seeking work); it is calculated as a percentage by dividing the number of unemployed individuals by all individuals currently in the labor force
* Average Wages in different positions
* Job openings and hires by month

1. **Data Processing**

* **Data Selection**

In the design of our system, we put our target users in first stage as students and newly graduated ones. All different websites providing economic statistics have a huge amount of data measuring many parameters of economy. We selected those data related to overall economic situation, job market and wages. We have GDP, CPI and population as measurement of macro economy. In the job market scale, we focused on unemployment rate, job openings and hires of major industries. Considering wages, we looked into average wages of different positions, also we had data bout averages wages and taxes of different countries. Some data is only available in our Google Datastore but not accessible from our application, because we are running out of time to finish all data visualization which we think uses will be interested in.

* **Data Cleaning**

When we extracted data from different sources, we found that some data includes incomplete values which are not fit for our visualization. We did some cleanup to the data set. Moreover because of the limitation on Google Datastore, we kept only those data mostly interested by normal users, and removed some industries to ensure data could be loaded in the Datastore properly

* **Data Categorization**

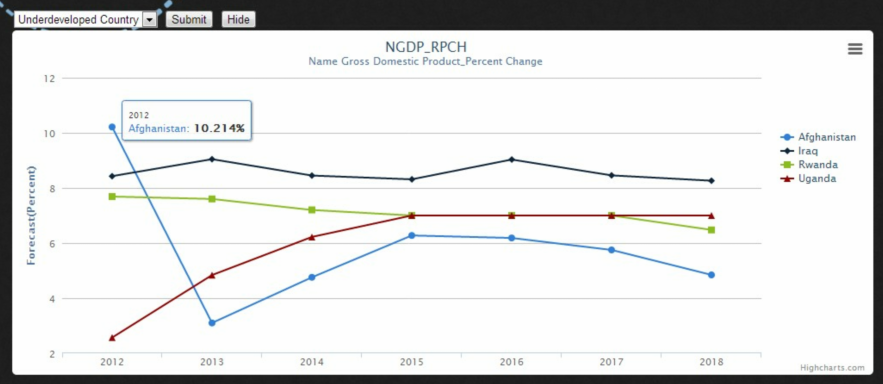
To support our goal of analyzing data in a designed context, we categorize data from the total dataset. This is a strategy to help our users to have a clear view of the meaning of different data, and be able to link the data with some of scenarios users are involved. For example we categorized the GDP data with different groups. We separated data display for job market into different industries. In showing wages of different positions, we also categorized them by industry. However, from the final visualization result we think we should come out a better way to make it less overwhelming.

1. **Technology used in Application**

For our project, we use the Google web app engine as our platform, with a back end server which is written in python. A front end website written in HTML, JQuery, JS has been created. We use the .js files from highchart website to create charts on our application. The files help us to build several beautiful charts in a short time. For our application, the most important part is the data. Most of the data source we extracted from different data sources is.csv file. We transferred it into Google data store. Unfortunately, the free version of Google application offers very limited space. Our data is too big to get through. A timeout error always shows. This is why our application could not commit into Internet to be publicly accessible.

1. **Charts**

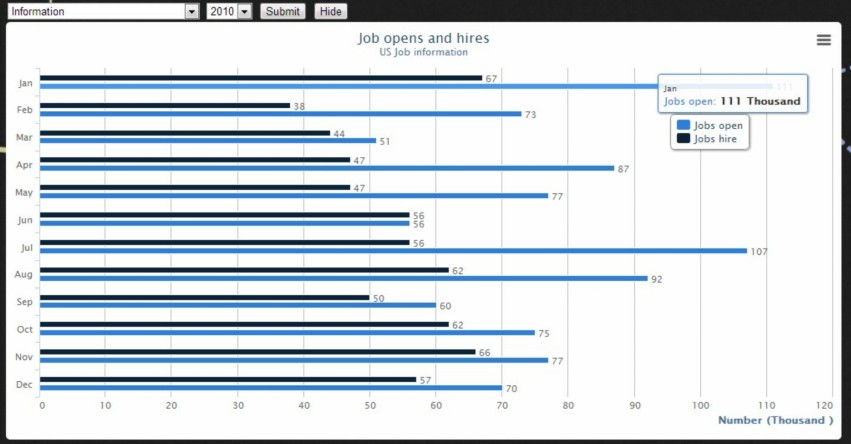
* **GDP**



It is not possible to show GDP data of all the countries in the world in one chart. In order to give users an easy and clear view of the data, we selected several countries and grouped them into 3 categories: developed country, developing country, and undeveloped country, with 4 countries for each group. Germany, US, UK, and Japan represent developed country. It shows all of these countries’ GDP will keep a stable, but slow growth in the following 6 years. The US stands out in this group obviously. Students can choose the US to start their career since its economy is booming in future years. In the group of developing countries, we selected countries in BRIC (In economics, BRIC is a grouping [acronym](http://en.wikipedia.org/wiki/Acronym) that refers to the countries of Brazil, Russia, [India](http://en.wikipedia.org/wiki/India) and [China](http://en.wikipedia.org/wiki/China), which are all deemed to be at a similar stage of newly advanced economic development. wikipedia). All these four countries have a high speed growth period in the following six years. China keeps an 8% growth rate of GDP and India keeps a 7% growth rate. Undeveloped countries group is represented by Afghanistan, Iraq, Rwanda, and Uganda. Although the GDP growth rate are not as stable as other two groups, this group of countries still have a high growth rate, between 5% to 9%.

Concluding these three GDP charts, we can say the globe economy has a bright outlook in six years. It is a good period for investment.

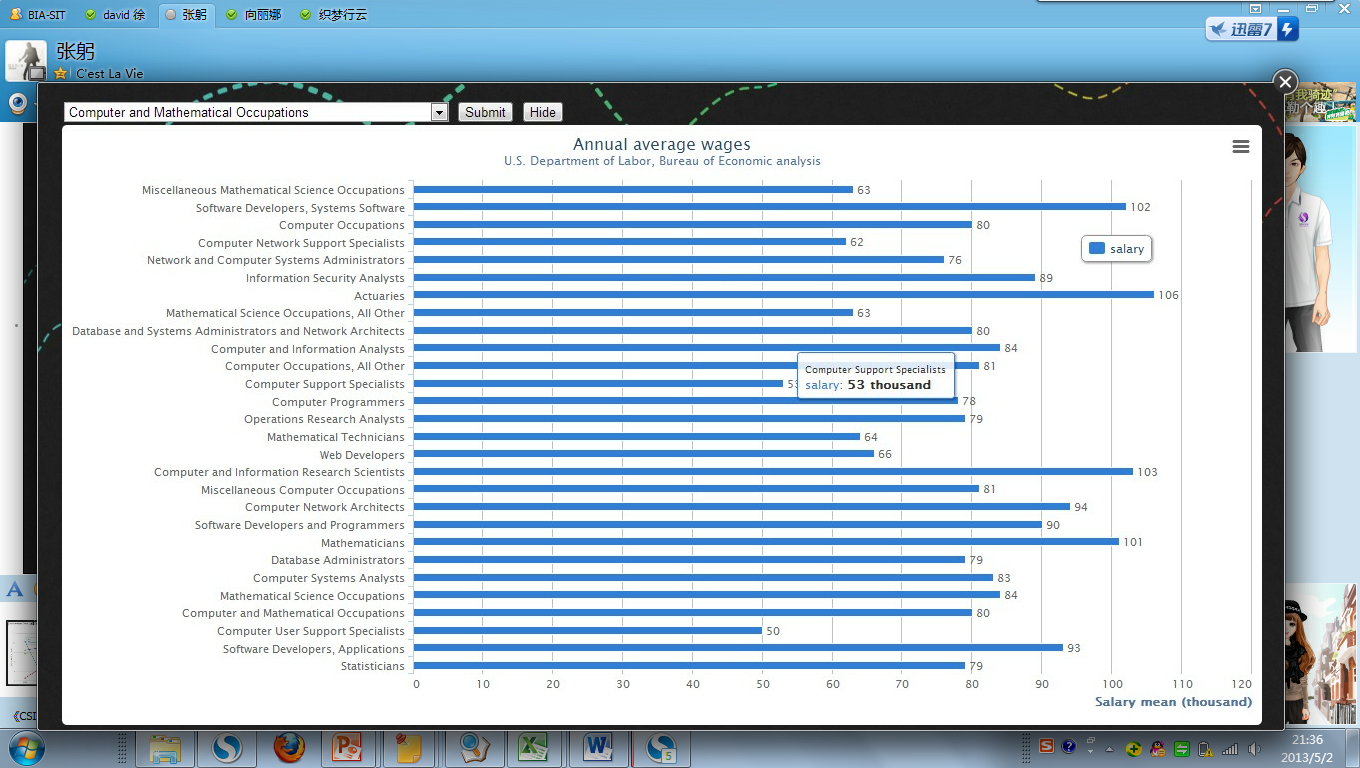
* **Job Analysis**



One of features in our system is to help users to know the job market situation. With comparison of job openings and hires in different industries, users can choose an industry easier to enter at their career start. The chart can also help users to find a better time to apply job in a year.

We built a job analysis chart with the data of job openings and job hires of each month. Users are able to choose from six industries in three years for different data view in a bar chart. For example, I am a student of Business Intelligence & Analytics, after I compared six industries. I found the hire rate (hire number divide open number) is high in information industry. I plan to focus on this industry to apply a job. Then I will compare three year (2010, 2011, and 2012). I found the Jan. Jul. and Aug. have a larger job open number than other months. I will send our more applications in these three months.

* **Annual average wages**



Wages are one of the most important factors considered when students choose a job. After users decide the industry they want to work in, they can use the data for position selection.

We chose data of the average wages of each position from different industries. Users can find an average wage of a position by submit the industry. And the chart can help them to compare different positions of salaries and find a most beneficial position matching with their skills.

For example, I can code in C++, Java, and Python. I can choose to be a software developer, web developer, or programmer. But the chart shows that to be a software developer can get the highest annual average wage. I will try to find a job in software developing, and I may get a higher wage than other job.

1. Conclusion

The current application we implemented is like an example of visualizing economic data in an easy and understandable way. We believe this could be an interesting field to be worked on, and we do see it helps the public to understand and use economic data more effectively. Our target is to enable economic data to be accessible and understandable for the public, and use the information to make decisions in their life.

In our implementation of this project, we also find that it is a really big challenge to achieve the goal. It is not just visualize the data, but how to make it meaningful for users, as well as technically practical in implementation. We run into the situation that we have too many information to show our users, but could not find those ones which help our users the most. To make the application a success, we will need more research and design for our system.