A story behind GDP

Milestone 2

(revised Feburary 2014)

Yezhuo Zhu, Ersi Zha, and Jiaqi He

xxxx-xxxx/0x/$xx.00 © 200x IEEE Published by the IEEE Computer Society

—————————— ◆ ——————————

# 1 Introduction

A

S economic globalization rapidly develops, Gross Domestic Products (GDP) has become one of hottest topics in recent years. But sometimes GDP cannot comprehensively and accurately reflect the living standard of a country because the living standard of country depends on many aspects such as rural and urban development, health status, poverty situation, infrastructure and so on. Today many people are still confused about the relationship between GDP and these aspects. The main purpose of our project is to explore how GDP impacts the living standard of a country and make a comparison of some major countries in the world. We plan to use various visualization approaches such as Choropleth map, histograms and bubble plot to represent our story.

In this paper, we will mainly explain the related work, data selection and visualization design ideas. After watching Hans Rosling’s TED lecture and studying the economic knowledge, we decided to select five aspects to make connection with GDP: Rural Development, Urban Development, Poverty, Health and Infrastructure. In conclusion, we will discuss the breakdown and eleven reference papers.

# 2 Related Work

In recent years, many works have been done in this field. Swedish medical doctor, academic and statistician Hans Rosling used software Trendalyzer to change tedious time series of development data into fascinating moving graphics. Bubble-shaped country icons were applied in his World Health Chart. The size and the color represented the population and the continent respectively. The user can enjoy an animated change to a new position when he changes settings. Also, he applied Dollar Street graph accompanied with some video documentation to illustrate daily income per person in the world. An interactive graph was used to show the world income distribution and he selected several countries’ statistics between 1970-1988 to display. We will list developed 11 papers in the references section.

# 3 Data Selection

|  |  |
| --- | --- |
| GDP | GDP  GDP annual growth  GDP per capita |
| Rural and Urban Development | Rural population  Urban population |
| Poverty | Income share held by highest 10%  Income share held by lowest 10% |
| Health | Population ages 65 and above (% of total)  Population ages 15-64 (% of total)  Population ages 15 and below (% of total)  Life Expectancy (years)  Total health expenditure  Infant mortality |
| Infrastructure | Internet users (per 100 people)  Telephone lines (per 100 people) |

Table 3.1

Table 3.1 shows the data we are going to visualize in our project. Since the purpose of our project is to explore and compare the living standard of different countries in the world, for this project, we will focus on 10 major typical countries in the world to make comparisons. It is obvious that the quality of life in a country is closely relevant to its economic development level and GDP is one of the most important factors to measure the economic development of a country. So our story will begin with analyzing the GDP of worldwide countries. We will give people a comprehensive overview and summary of worldwide GDP.

However, GDP cannot comprehensively evaluate the development of a country. For example, high GDP does not mean high-level living standard. Also, there are a lot of factors that can influence GDP. Therefore, we will use our selected data to explore the story behind GDP.

Firstly, we will examine the population composition of different countries, which is assumed to have a close relationship with GDP. It will be very interesting to analyze how different population compositions contribute to different GDP. We will figure out the percentage of rural population and urban population of 10 major countries. Then we will explore the population proportion of different age range of those major countries. The proportion of population ages 15-64 of a country is a very important factor to decide the country’s GDP level for the reason that population aged 15-64 would definitely make more contributions to GDP than other age groups population.

Secondly, we will examine the poverty situation of above countries. We want to compare the income gap of these major countries. Even with a high GDP, if the income gap of countries is high, the economic development seems not to be well.

Thirdly, health status of a country will also be considered. The health care industry is one of the world’s largest and fastest-growing industries. In many countries, health care industry consumes nearly 10 percent of gross domestic product (GDP). In this part, we are going to make a comparison of health status of those major countries.

Last, we will give statistics of infrastructure distribution, such as Internet users (per 100 people) and telephone lines (per 100 people), since infrastructure is as important as health to decide the living standard of a country.

# 4 Visualization Design ideas

## 4.1 Choropleth Maps of GDP

At the start of our project, we will use appropriate classification and color scheme to create a choropleth map, which illustrates global Gross Domestic Products (GDP). Countries in the choropleth map will be marked with several sequential colors. Light values are low data values, dark are high. So darker color areas in the choropleth map have larger GDP than lighter color areas. A year slider will be placed below the map. The position of the slider indicates different years from 1961 to 2013. The map will change according to the change of slider. Thus, we can have a clear understanding of the world GDP change from 1960 to 2013. Also a basic information description of a country will be displayed if we move our mouse to the corresponding area in the map. The description includes GDP annual growth and GDP per capita.

## 4.2 Relationship between Population and GDP

Population of a country is a key factor to influence the GDP level. After an overview of global GDP change, in this part, we are going to explore the population composition of some major countries and illustrate the relationship between population and GDP level. Our population data contain three aspects: total population, rural population and urban population. The method that we will use to visualize our data is Bubble Plot. In the Bubble Plot, the X-axis presents the number of total population and Y-axis presents the ratio of urban population to rural population. Each country will be displayed as a colored circle and the size of circle reflects the country’s GDP level. After clicking the circle (countries), it will transform to three plots, representing the age range of less than 14 years old, 14 years old to 65 years old, older than 65 years old respectively.

With this bubble plot module, we can analyze the relationship among the population ratio of the urban and the rural, the age distribution, and GDP.

## 4.3 Stream River of Poverty

Theme River is going to be applied to reveal the story behind the data about poverty. The X-axis represents the year (1960-2013) while Y-axis depicts the percentage of income share. Above X-axis represents highest 10% of the population who contribute to GDP. Blow X-axis represents lowest 10% of the population who contribute to GDP. We plan to select these major countries to make a series of theme river graphs for analysis.

With theme river graph, user can easily figure out how income shares are changing from year to year. Comparing the first 10% and the last 10%, we could also see how the gap between rich and poor changes with time passing.

## 4.4 Histograms of Health

Obviously health care plays a crucial role in the GDP, therefore, in this part we will discuss how GDP impacts a country’s health status. We chose three aspects to explore the relationship between health status and GDP: total health expenditure, hospital beds (per 1,000 people) and infant mortality. And we will also use these data to make a comparison of health status among several different major countries in the world. We will create three histograms to visualize the data. The X-axis presents different countries and the Y-axis presents total health expenditure, hospital beds (per 1,000 people) and infants mortality respectively. Our histograms will also show the distribution of these three aspects in the year 1960, 1980, 2000 and 2013 respectively. So we can compare different countres’ data in different years.

## Line Chart of Infrastructure

In this part, we will discuss the relationship between GDP and infrastructure development. We decided to analyze and visualize two kinds of data: Internet users (per 100 people) and telephone lines (per 100 people). The reason why we chose these two aspects is that they can present the infrastructure level of a country to some extent. The approach we use to show the data is a line chart. The X-axis presents different years (from 1960 to 2013). The Y-axis presents the number of Internet users or the number of telephone lines. In our line chart, each country has two particular colored lines: solid line and dotted line. Solid lines represent the number of Internet users and dotted lines represent the number of telephone lines.

# 5. Breakdown

We have collected several categaries of data from the World Bank. Then we selected GDP, Rural and Urban De-velopment, Poverty, Health and Infrastructure as the in-terested datasets. After that, we had a discussion about the best method to virtualize these datasets. At last we decided to use Choropleth Map, Stream River, Buble Plot, Scatter Plot, Stacked Bar Chart and Line Chart.

In the next stage, we will focus on the implementation of these visualizations.

|  |  |
| --- | --- |
| Jiaqi He 33% | 1. Choropleth Maps of GDP  2. Stream River of Poverty |
| Yezhuo Zhu 33% | 1. Bubble Plot of Population distribution and GDP  2. Scatter Plot of age distribution |
| Ersi Cha 33% | 1. Stacked Bar of Health  2. Line Chart of Science and Infrastructure |

**6. References**

1. Rosling H, Rosling R A, Rosling O. New software brings statistics beyond the eye[J]. Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making. Paris, France: OECD Publishing, 2005: 522-530
2. Havre S, Hetzler B, Nowell L. ThemeRiver: Visualizing theme changes over time[C]//Information Visualization, 2000. InfoVis 2000. IEEE Symposium on. IEEE, 2000: 115-123.
3. Chen M, Ebert D, Hagen H, et al. Data, information, and knowledge in visualization[J]. Computer Graphics and Applications, IEEE, 2009, 29(1): 12-19.
4. Bertini E, Tatu A, Keim D. Quality metrics in high-dimensional data visualization: an overview and systematization[J]. Visualization and Computer Graphics, IEEE Transactions on, 2011, 17(12): 2203-2212.
5. Mullër W, Schumann H. Visualization methods for time-dependent data-an overview[C]//Simulation Conference, 2003. Proceedings of the 2003 Winter. IEEE, 2003, 1: 737-745.
6. Havre S, Hetzler B, Nowell L. ThemeRiver: Visualizing theme changes over time[C]//Information Visualization, 2000. InfoVis 2000. IEEE Symposium on. IEEE, 2000: 115-123.
7. Havre S, Hetzler E, Whitney P, et al. Themeriver: Visualizing thematic changes in large document collections[J]. Visualization and Computer Graphics, IEEE Transactions on, 2002, 8(1): 9-20.
8. Andrienko G L, Andrienko N V. Interactive maps for visual data exploration[J]. International Journal of Geographical Information Science, 1999, 13(4): 355-374.
9. Derthick M, Kolojejchick J, Roth S F. An Interactive Visualization Environment for Data Exploration[C]//KDD. 1997: 2-9.
10. Zhu N Q. Data Visualization with D3. js Cookbook[M]. Packt Publishing Ltd, 2013.
11. Bostock M, Ogievetsky V, Heer J. D³ data-driven documents[J]. Visualization and Computer Graphics, IEEE Transactions on, 2011, 17(12): 2301-2309.