Visualization Analysis About Women's Rights and Interests

Introduction

In this blog post, I will introduce how I perform data visulization and analysis about women's rights and interests by writing a website. The whole process of my analysis can be divided into three phases, getting datasets, processsing data and performing analysis. But before all the above contents, I would like to talk about why I choose this topic to do my visualization analysis.

About the Topic

- **Something Interesting**

  For the first time I learned about this assignment, I decided to collect some interesting datasets and focus on a meaningful topic, rather than just cope with the homework. Recently, no matter on the Internet or in real life, we hear more and more about the word "feminism", which has also caused various discussions on the Internet. I began to wonder, in this era when people are more and more concerned about equal rights between men and women, what is the real situation of women's rights and women's living circumstances, especially in those undeveloped regions?

- **The Indicators**

  Women's rights and interests is a big topic, I must be specific to some measurable standards, and the choice of data dimensions (or indicators) also determines what specific issues I want to explore through visualization.

  - Child Marriage

    On the data website of the United Nations Children's Fund (UNICEF), I found a group of data, whose indicator definition is as follows:

    Percentage of women aged 20 to 24 years who were first married or in union before ages 15 and 18.

  - Maternal Care

    Now that we talk about the issue of child marriage, it inevitably involves the issue of early childbirth. So is the prenatal care and postpartum care for these early childbearing mothers in place? Will there be many undeveloped areas still unable to provide basic reproductive health conditions?

The Adolescent Health Coverage Database of UNICEF provides many indicators. I simply chose three of those indicators:

  a. *Percentage of women aged 15–49 years attended at least four times during pregnancy by any provider. (ANC_4)*

  b. *Percentage of women aged 15-49 years who had a blood sample taken during any antenatal care visit. (ANC_BD)*

  c. *Percentage of women aged 15–49 years who received postnatal care within 2 days after birth. (PNC_WM)*

Note that the first two indicators refer to women who had a live birth in the last three or five years and the last one refers to women who had a live birth in a recent time period, generally two years.

By the formula below, I got a Maternal Care Index *(MNC)*, which I thought can reflect the health status of those early childbearing mothers.

$$MNC = ANC_4 * 0.3 + ANC_B * 0.2 + PNC_W * 0.5$$

○ Neonatal and Maternal Mortality

If there is a tradition of child marriage in a region, and it is unable to provide basic reproductive health conditions for early childbearing mothers, I naturally thought that this will certainly affect the local neonatal and maternal mortality.

At first I got two indicators to show the neonatal and maternal mortality:

  a. *Maternal Mortality Ratio (MMR; maternal deaths per 100,000 live births)*

  b. *Neonatal Mortality Rate (NMR)*

Then I thought the *NMR* may not be accurate enough to reflect the local reproductive health environment. So I found a dataset about the causes of child deaths. Among all the causes, I select two causes, *Intrapartum Death* and *Preterm Birth*, which I thought are main factors reflecting the local reproductive health environment.

By the formula below, I got a new index *NMR_N* to evaluate the local maternity health environment.

$$NMR_N = NMR * (P(Intrapartum Death) + P(Preterm Birth))$$

○ Other Relevant Indicators

In order to find the relationship between economic development, geographical location and gender equality culture, I introduced three more indexes, GDP per

capita, Geographical Region and gender inequality index (GII). All the above indicators come from the website of the United Nations Human Development Data.

Data Processing

After choosing the datasets, I got a series of file in *xls*, *xlsx* and *csv* format. In fact, the Microsoft Excel Software was a good data processing tool for me. I deleted useless comments and pictures in the data file and only kept useful statistics in the table. I could use *formula* in the Excel to calculate new data from the old one automatically. For example, in the dataset about Maternal Mortality Ratio, I could use the formula below to get a new column, which contains maternal deaths per 1000 live births.

```
=H:H/1000
```

Besides, I used the *function* VLOOKUP in the Excel to merge the datasets into a new one.

> VLOOKUP is an Excel function to get data from a table organized vertically. Lookup values must appear in the *first* column of the table passed into VLOOKUP. VLOOKUP supports approximate and exact matching, and wildcards (* ?) for partial matches.

Then I decided to use Numpy and Pandas to convert my Excel Workbook file into a JSON file.

```python
def excel_to_json(self,excelfile,jsonfile):
    data=pd.read_excel(excelfile)

    data.to_json(jsonfile)
```

After all the data processing, I got a JSON file which could be used in the visulization analysis. A part of this JSON file is shown below.

```
[{"Countries and areas":"Afghanistan","GDP per capita ($)":594.5,"Gender Inequality
Index":0.67,"Married by 15":8.8,"Married by 18":34.8,"Source":"DHS
2015","Region":"South Asia","Antenatal care (four or more visits - live birth in
the last 3 or 5 years)":16.2,"Postnatal care for the mother (received postnatal
care within 2 days after birth)":37.4,"Maternal care for the mother after
filtration":29.3,"Blood sample taken during any ANC visit (for under 3 or 5
years)":28.6,"Neonatal mortality rate (during the first 28 days of life - per 1,000
live births)":33,"Percentage":56.7,"Neonatal mortality rate after
filtration":18.7,"Maternal mortality rate (maternal deaths per 1,000 live
births)":4.0},...]
```

Visualization

In this part, I decided to write a website to perform the visualization analysis. A static website was not attractive enough to me, so I wanted to implement some interesting interactive functions. This time I chose the **TauCharts**, a data focused javascript charting library, to help me with this task.

- **Scatter Diagram**

  I didn't plan to make some fancy charts.

  Scatter diagram is the most basic graph in visualization, and it is very easy to use. In the process of collecting data, I gradually realized that to reflect the relationship between various data dimensions, the use of scatter diagram was the most intuitive, and did not need redundant data processing.

- **How to be Interactive**

  It's not difficult to realize the interactive function. I used the popular Taucharts library on GitHub. The scatterplot component of the Taucharts library has interactive functions: *tool tip*, *legend*, *export to* and *trendline*. In addition, I made the label of the *XY* axis changeable. I can change the label of the *XY* axis to visualize the two chosen indicators. The radius of a plot can also represent one data dimension. Of course, this is also changeable.

  The minisize and maxsize of a plot can be dynamically modified according to the user's visual preference (through a slider component). The color of a plot is classified according to the region of the country. Then we can also filter the regions to make only the data points of the region we want appear in the chart.


Insights

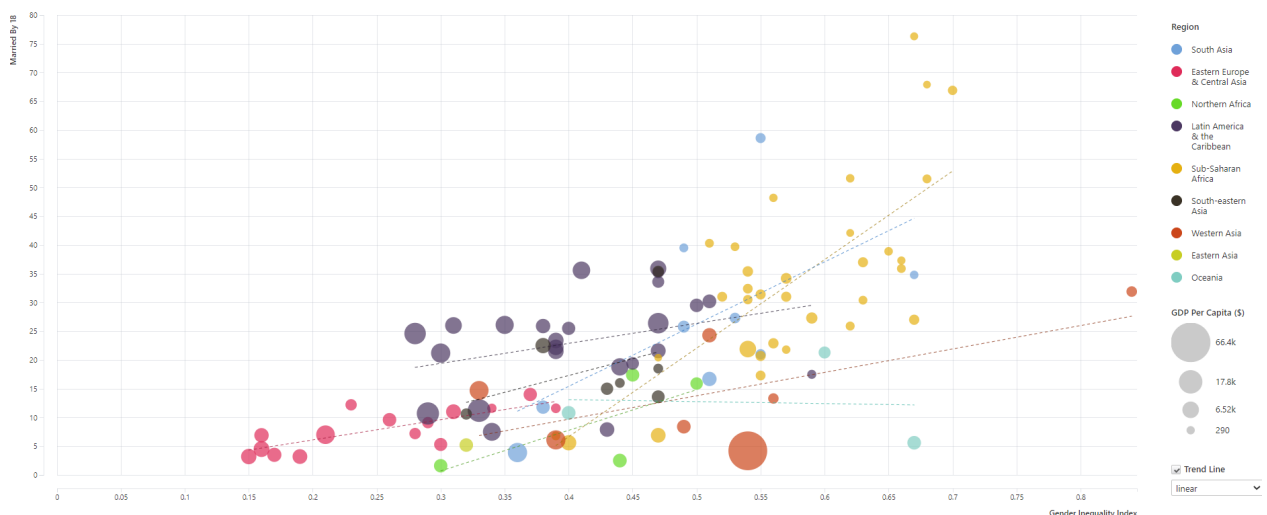- **Child Marriage and GDP Per Capita**
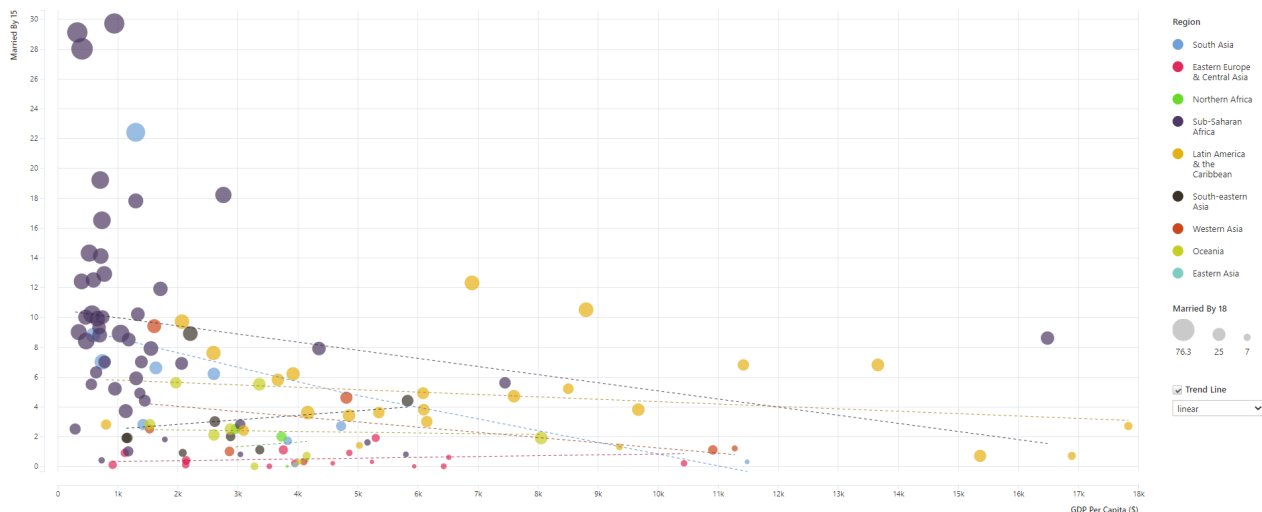
Most of the countries where child marriage is common (more than 10% married before the age of 15) have GDP per capita less than 2K, and most of them are sub-Saharan African countries. But it is not that the higher the GDP per capita, the less child marriage, as in some South American countries. It also doesn't mean that the child marriage is common when the GDP per capita is low, like most countries in Eastern Europe and Central Asia.

- **Child Marriage and GDP Per Capita**
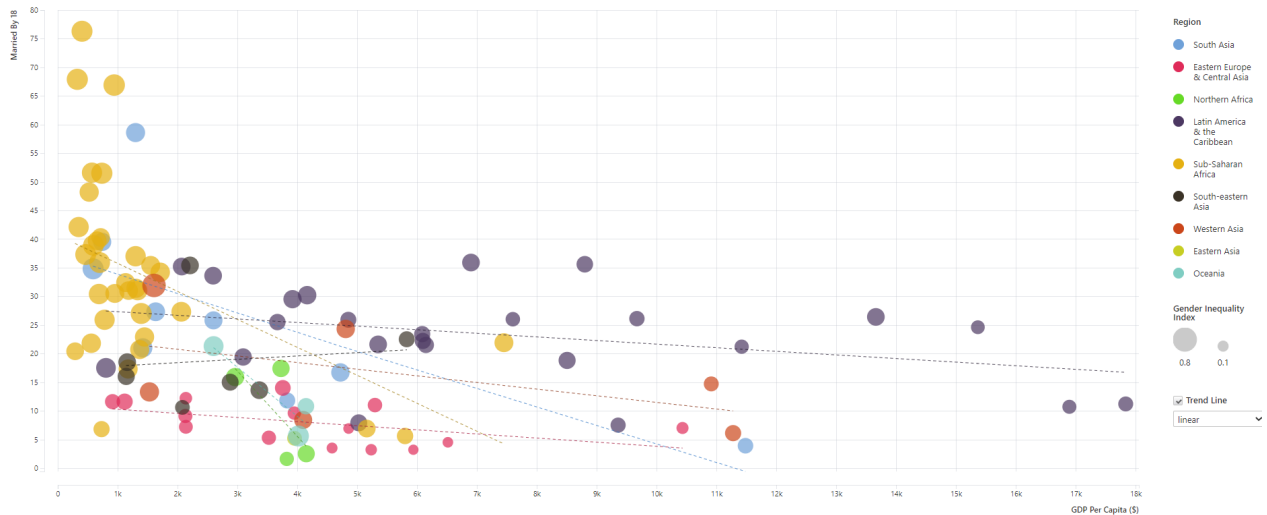
Whether child marriage is common depends not only on the development of the national economy, but also on the local polarization between the rich and the poor, the local culture and religious beliefs. For example, countries under the same GDP per capita but with low child marriage rates are basically countries with low GII.

- **Child Marriage, GDP Per Capita and GII**

Only from the data point of view, the impact of gender equality on the phenomenon of child marriage is more obvious than GDP per capita. If the GII is lower than 0.3, the child marriage rate will not exceed 25%. If the GII is higher than 0.6, the child marriage rate will show an obvious upward trend.

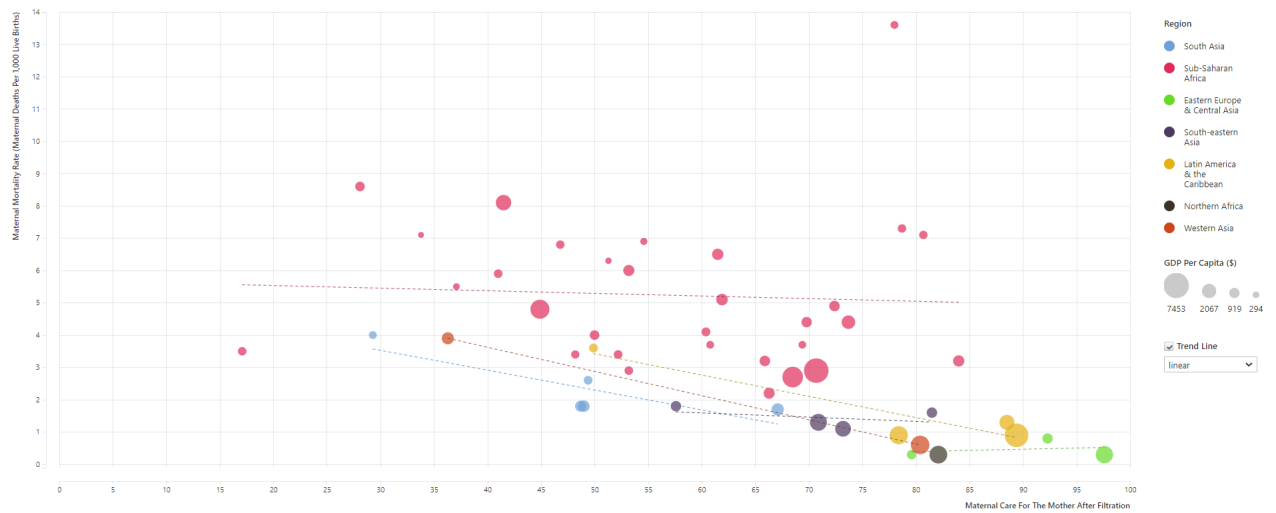- **Maternal Care and GDP Per Capita**
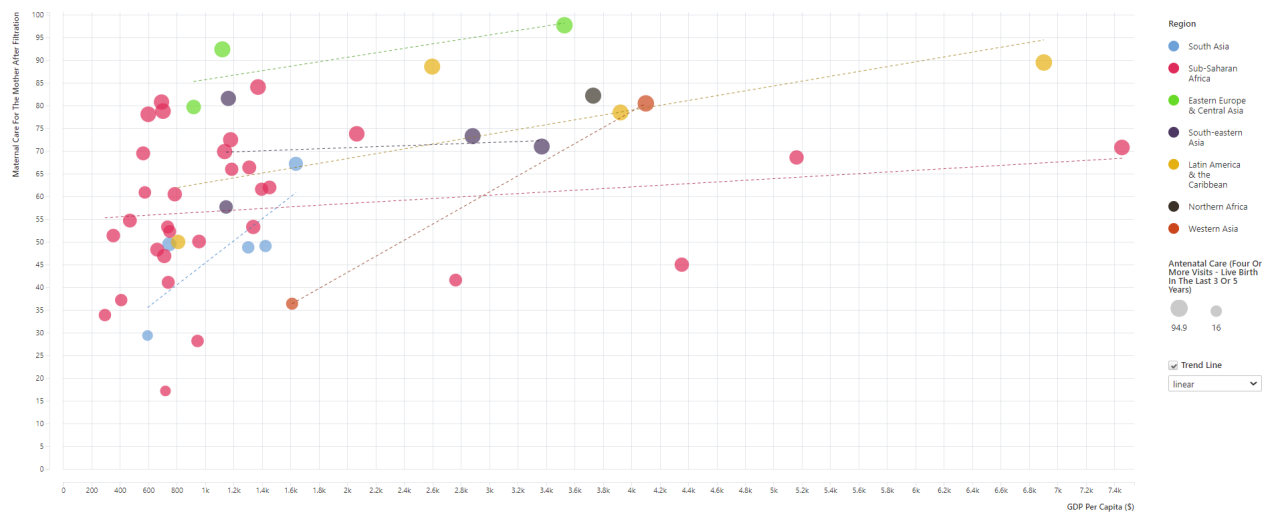
The indicators of maternal care did not fluctuate too much with GDP per capita, and basically maintained at a good level (MNC was more than 50%). Those countries with very poor maternal care are mostly sub-Saharan African countries such as Ethiopia, or the countries at war such as Afghanistan.
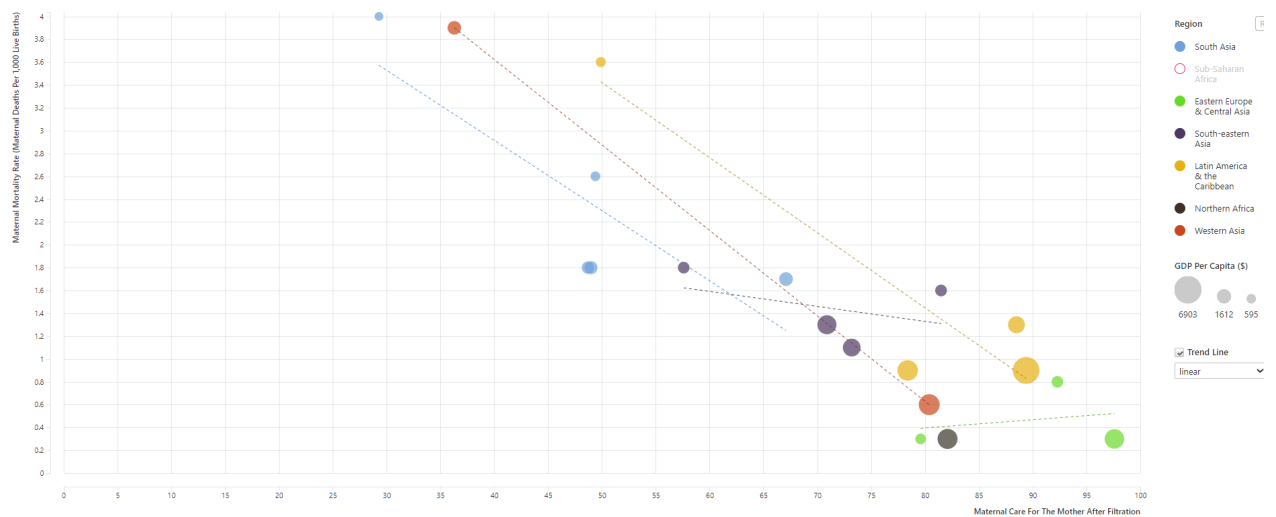
- **Maternal Care, GII and GDP Per Capita**
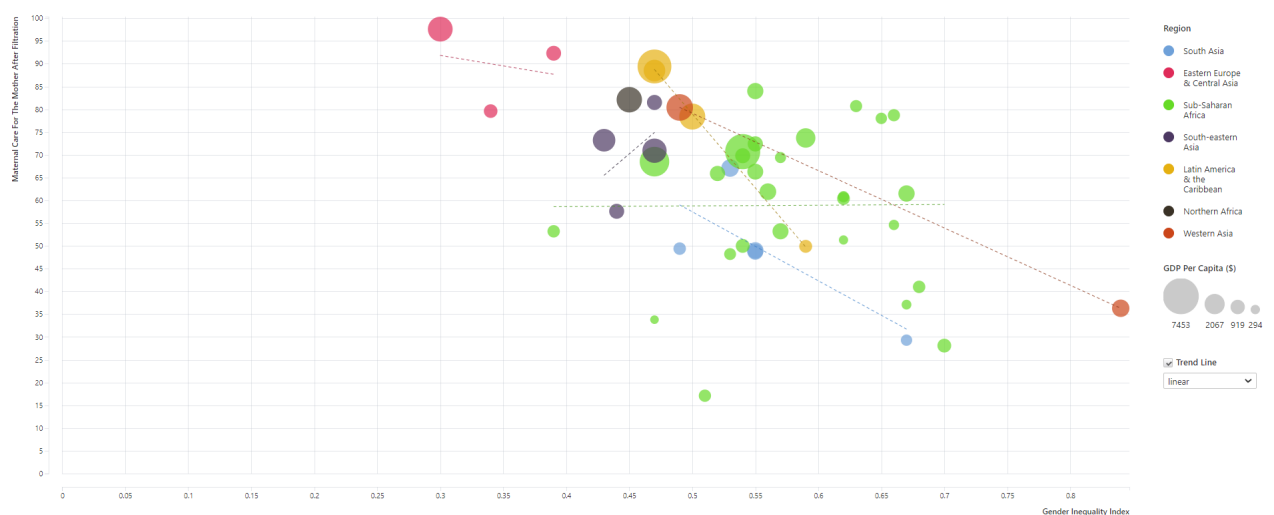
Those countries with high GII are not meant to have poor maternal care, but those with low GII usually have good maternal care. Those sub-Saharan African countries usually have many complex social problems such as corruption and polarization of social wealth, which make neither GDP nor GII able to explain the maternal care level very well in those countries.

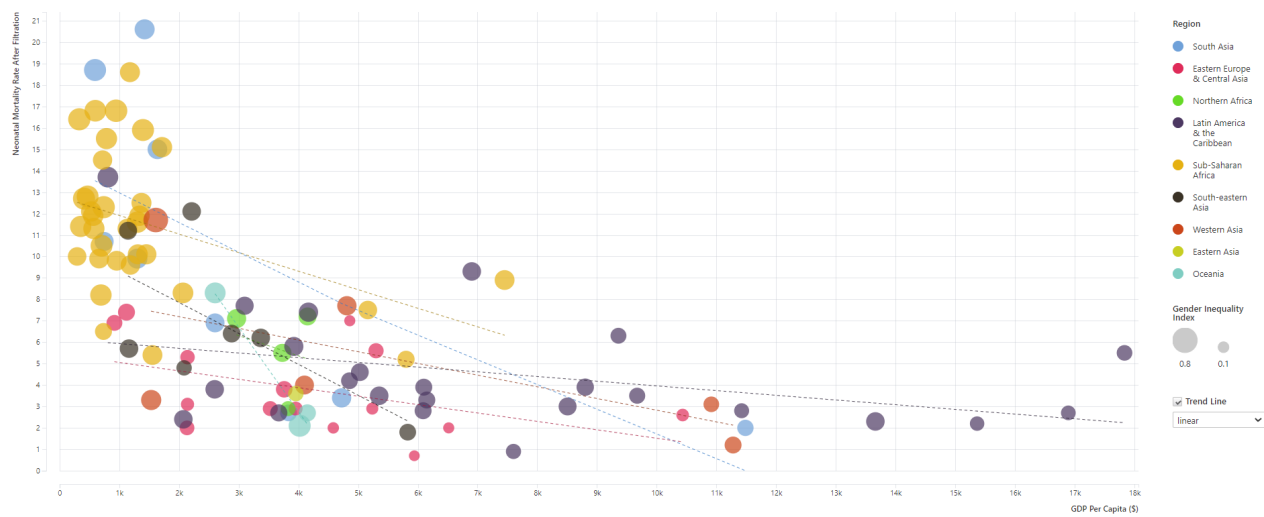- **Maternal Care and Maternal Mortality Rate**

Just like I thought, maternal care level has a negative influence on the maternal mortality rate. But in those undeveloped countries, mostly sub-Saharan African countries, we can see that GDP per capita has a more obvious impact on the maternal mortality than maternal care level.

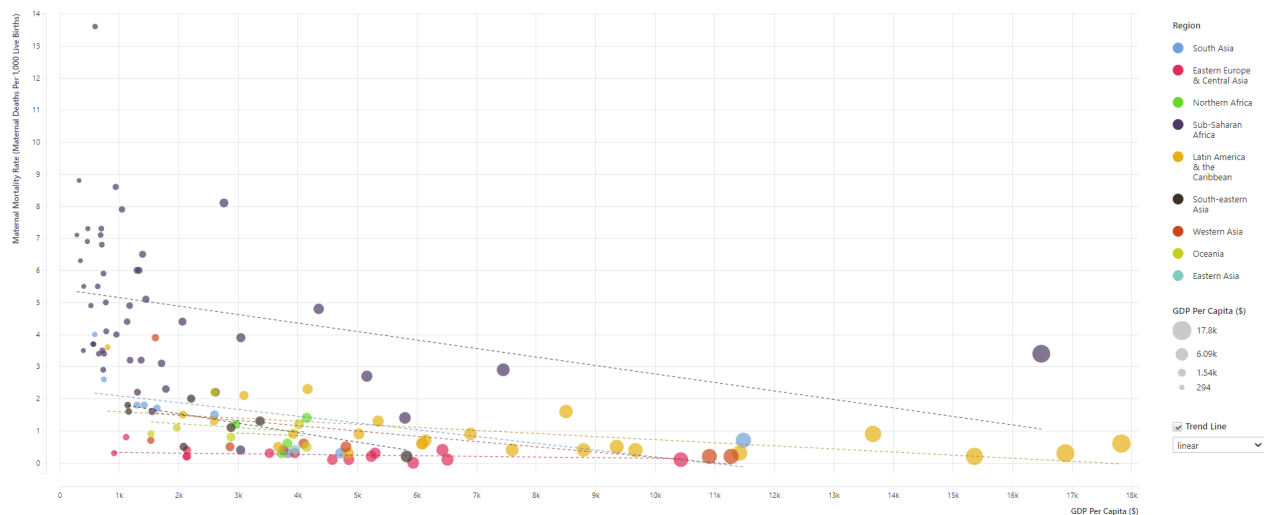- **Maternal Care and Maternal Mortality Rate**

If we exclude those data from the sub-Saharan African countries, we can see a more obvious negative correlation between maternal care level and maternal mortality rate.
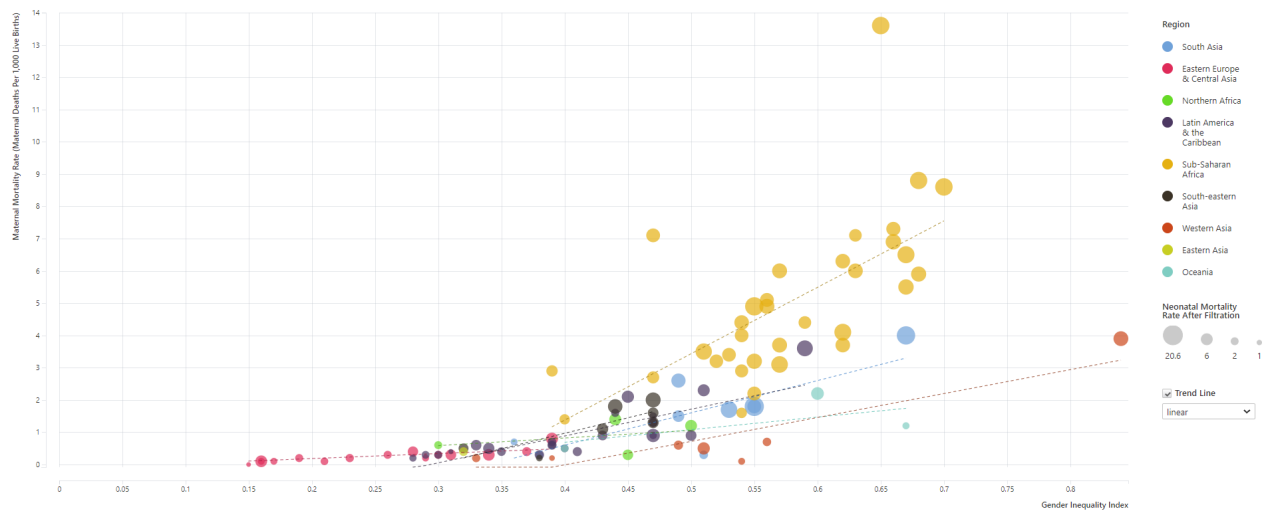
- **Neonatal Mortality Rate and GDP Per Capita**

The neonatal mortality rate is mainly limited by the medical level, which partly reflects the economic level of a country. With the increase of GDP per capita, the neonatal mortality rate showed an obvious downward trend.

- **Maternal Mortality Rate and GDP Per Capita**

Unlike neonatal mortality rate, maternal mortality rate is not so influenced by GDP per capita. As we can see from the figure, in those countries with maternal mortality below 2‰, the maternal mortality barely changes with the GDP per capita, especially in those Eastern European and Central Asian countries.

- **Maternal Mortality Rate and GII**

The maternal mortality rate has a more obvious correlation with the GII than the GDP per capita. The country with the lowest maternal mortality, the Belarus, is exactly the country with the lowest GII, which partly supports the previous point of view.

- **Gender Inequality Index and GDP Per Capita**
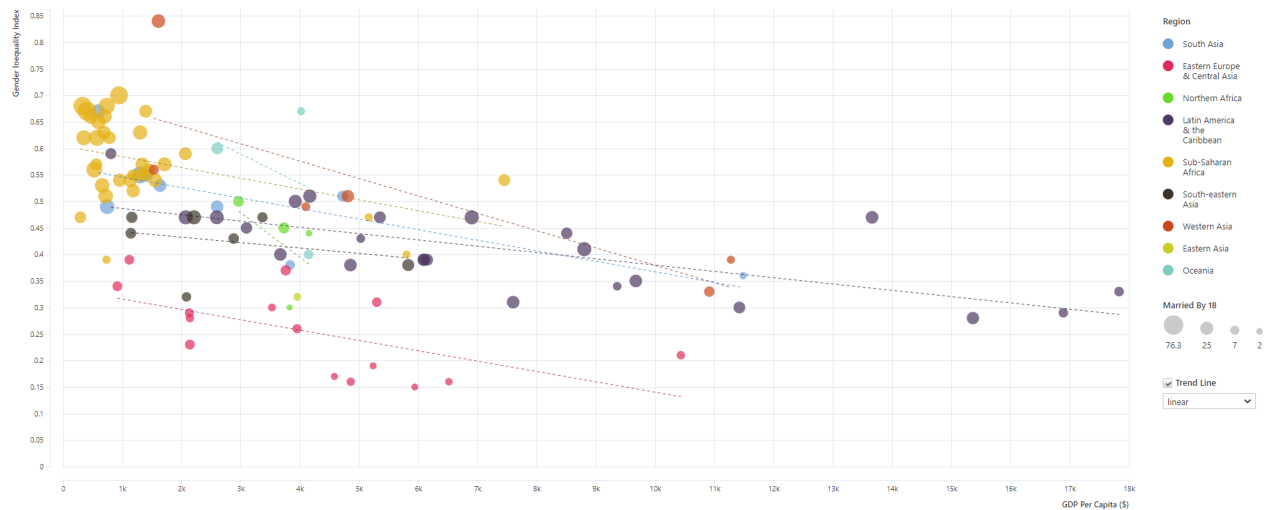
The GII is obviously affected by the economy. Basically, with the growth of GDP per capita, the GII decreases. When the GDP per capita is less than 1K, the GII is mostly above 0.5.

- **Gender Inequality Index and GDP Per Capita**
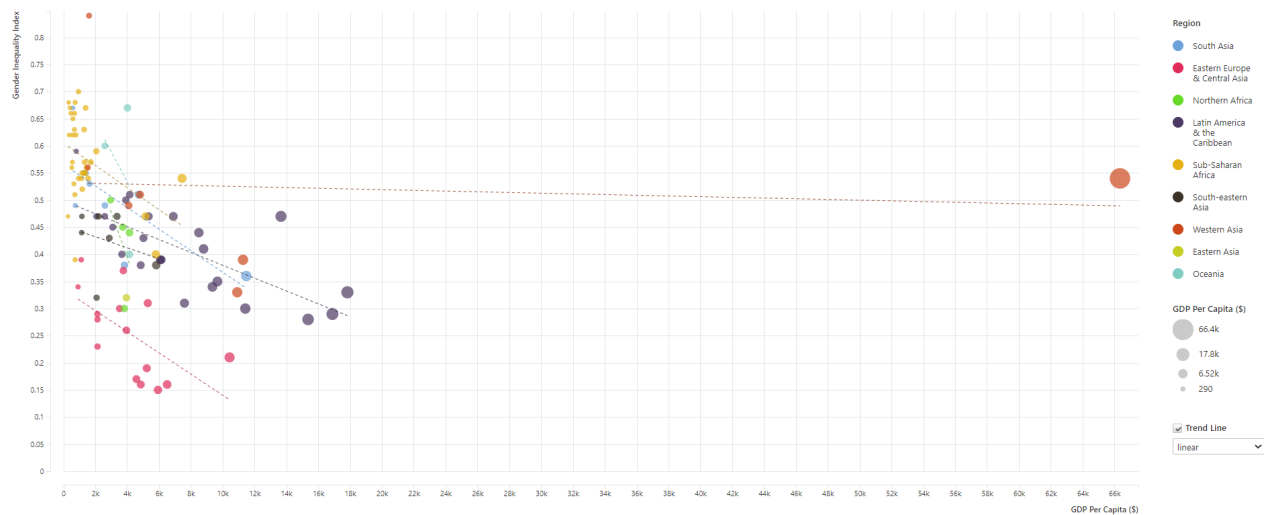
However, the GII also needs to be connected with local cultural and religious environment. For example, in the Qatar, although the GDP per capita is more than 66K, women's rights are severely restricted due to the strict enforcement of Islamic rules, which make the GII reach 0.54.

Conclusion

From the whole visualization, I have noticed that generally the more developed one country is, the more women's rights and interests can be protected in this country. But in some sub-Saharan African countries with complex social problems, such as Equatorial Guinea, Nigeria and Angola, people's living conditions are extremely poor, not to mention women, which can not be simply explained by the GDP per capita.

Compared with the GDP per capita, the GII can better reflect whether women's rights and interests are well protected, in terms of both living conditions and social status.

Data Sources

- UNICEF Data

  https://data.unicef.org/resources/resource-type/datasets/

- World Health Organization Global Health Observatory

  https://www.who.int/data/gho

- United Nations Human Development Reports

  http://hdr.undp.org/en/composite/GII

- United Nations Statistics Division

  https://unstats.un.org/unsd/gender/worldswomen.html

Tools

- TauCharts

  https://github.com/TargetProcess/tauCharts