Group Project

Mark Hovsepyan, Vazgen Tadevosyan, Garen Danageozian, Davit Gadyan

July 22, 2018

Homework #4: Group Assignment

### Quandl API for Scraping Quantitative Data

The purpose of quantitave data analysis of our project is to explore and analyze the economic conditions of Armenia as compared to the most similar economies of the world, such as Romania, Hungary, Ukraine, Georgia, Belarus. Also in our analysis would be included the results of some developed or develping countries as to compare their results with Armenian indicators in order to understand what we need to do to have strong and developing economy. Notably, we are not going to get deeper into art of economics but rather to visualize and explore some interesting patterns and insights from data, taking into account the context of our Homework.

The first dataset that we are going to use is UN Gloabal Indicators, kindly provised by United Nations. Information about the documentation may be accessed through the following link>>> https://www.quandl.com/data/UGID-United-Nations-Global-Indicators
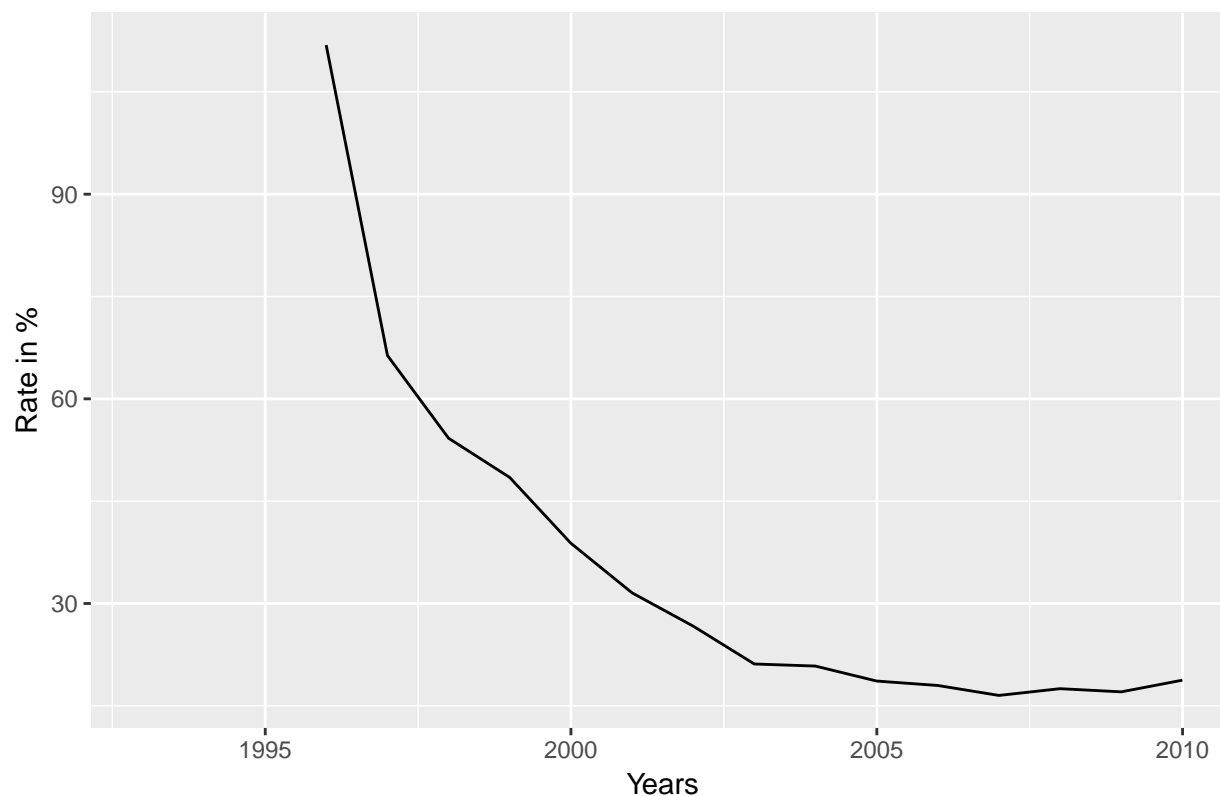
Generally we may observe that rate have been decreasing from the end of 1990s. Now let's compare with four other similar to Armenia in economic terms countries such as Romania,Hungary,Ukraine,Georgia and also Russia.
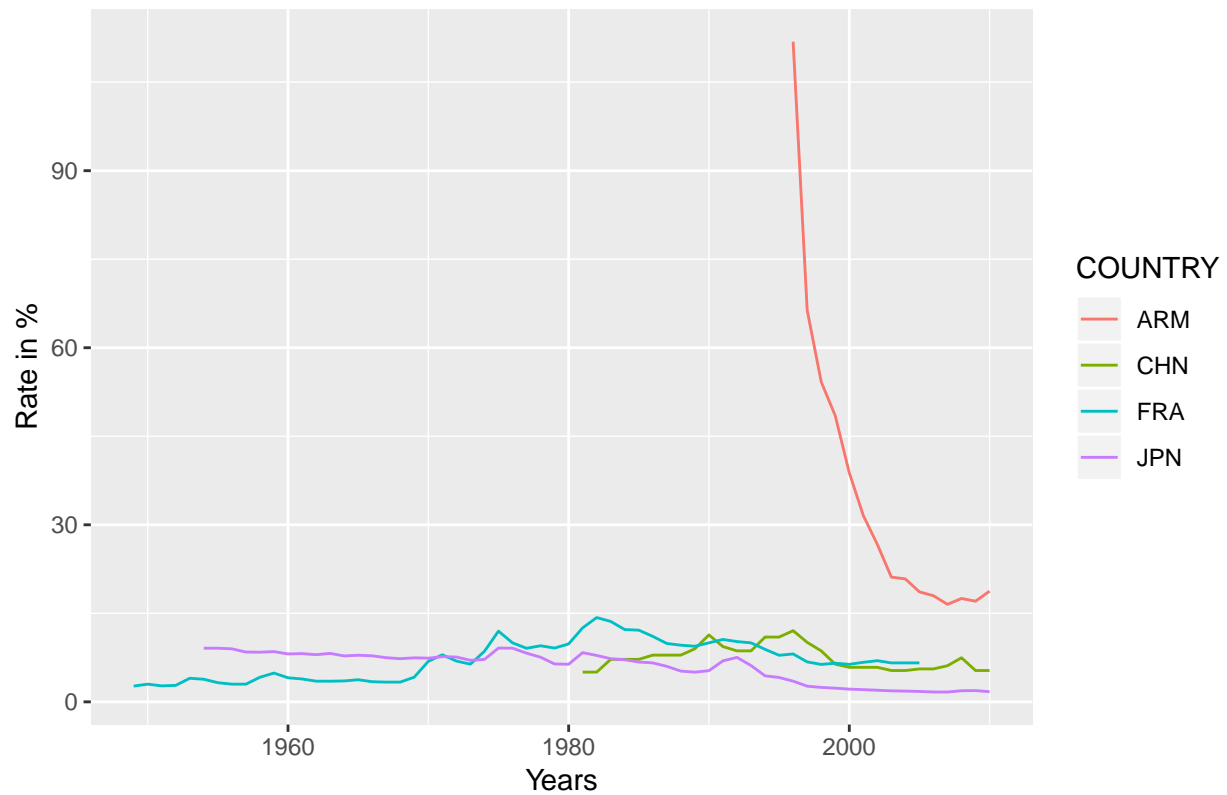
Historical Lending Rate by countries
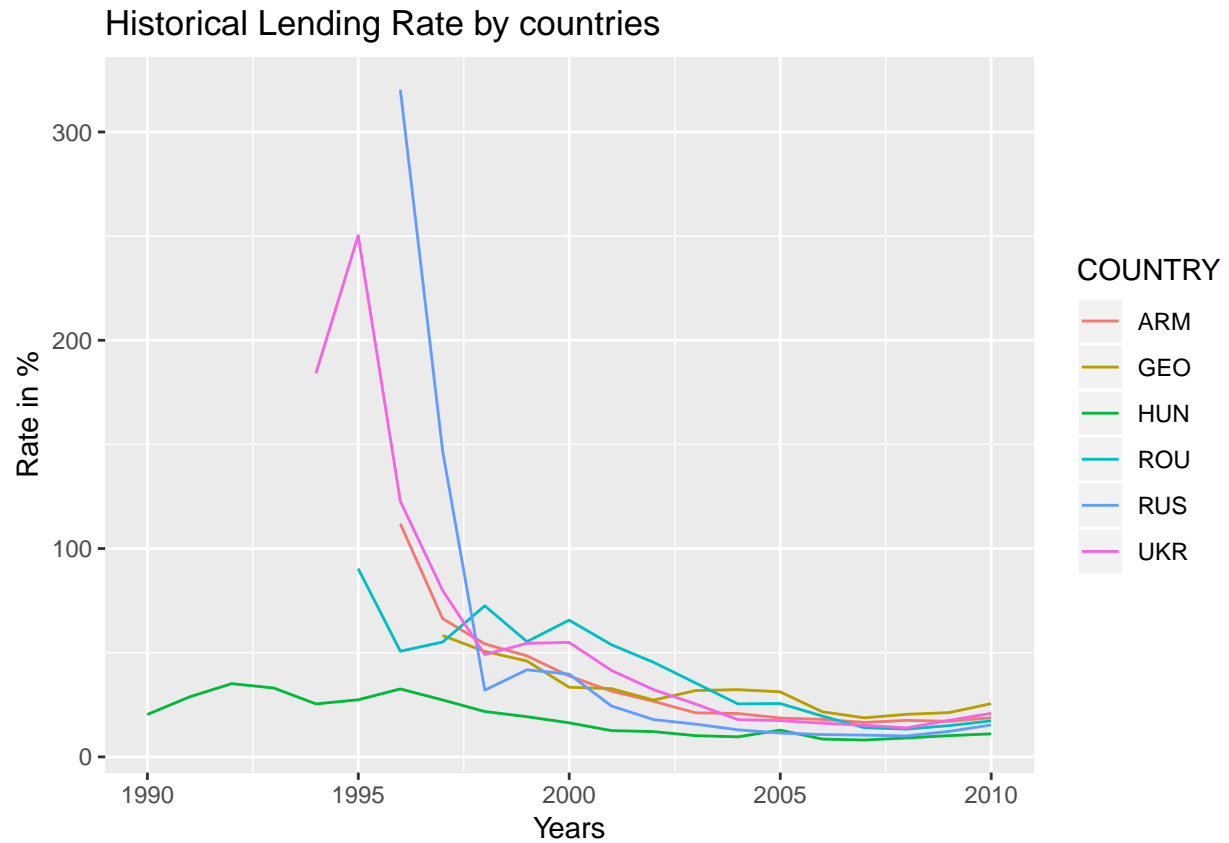
This graph depicts the situation with lending rates in Armenia and other, similar to Armenia in economic terms countries. The rates appears to be extremelly high particularly with Russia (more than 300% !!!)and Ukraine during 90th.

## Historical Deposit Rate by countries
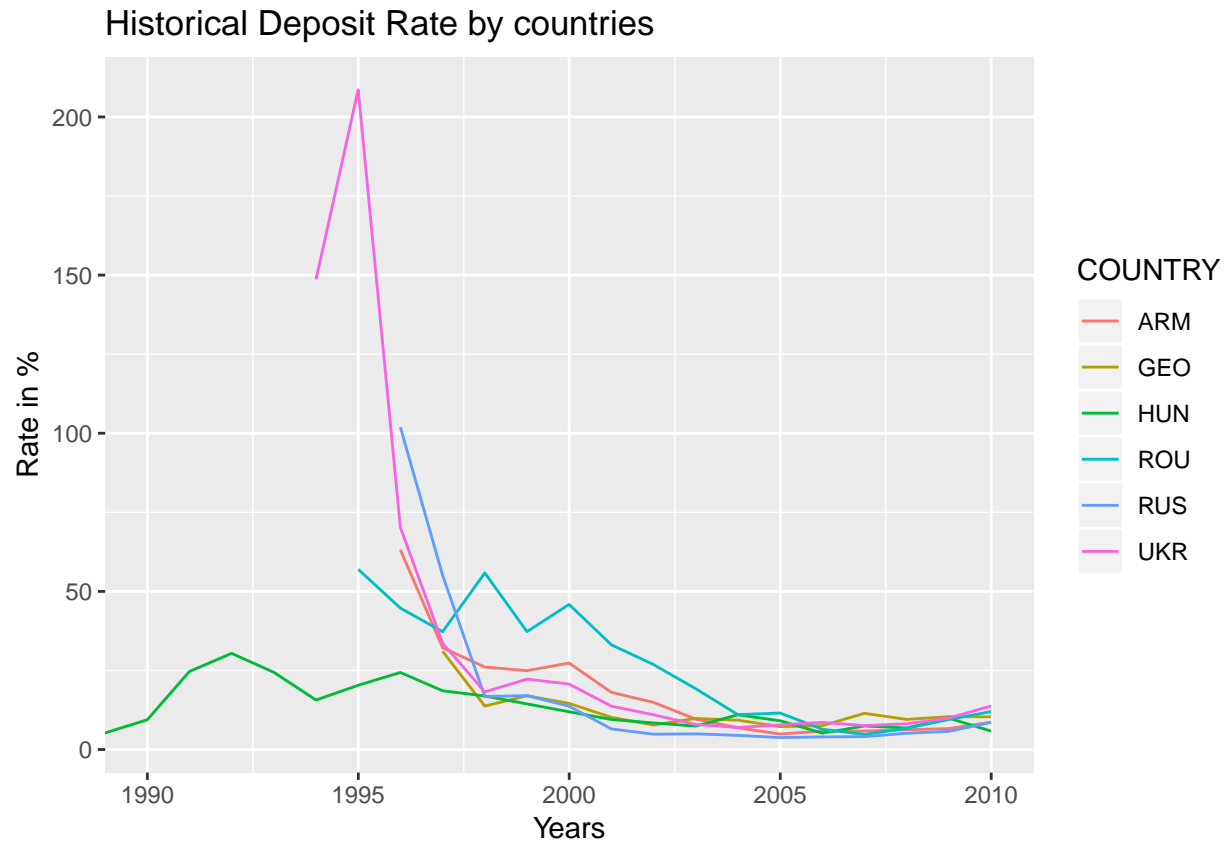
The above graph visualizes the historical deposit rates in Armenia and five other countries. We observe that again the highest rates have Ukraine more than 200% and Russia with around 100% at the beginning. However this changes with time.

Historical Lending Rate Armenia VS Developed Countries

This graphs clearly transmits the message that high lending rates during the 90s were not all over the world. Even few historical lending rates of China, France and Japan proves that high rates identified in Armenia, Russia and Ukraine were caused by some other factors.
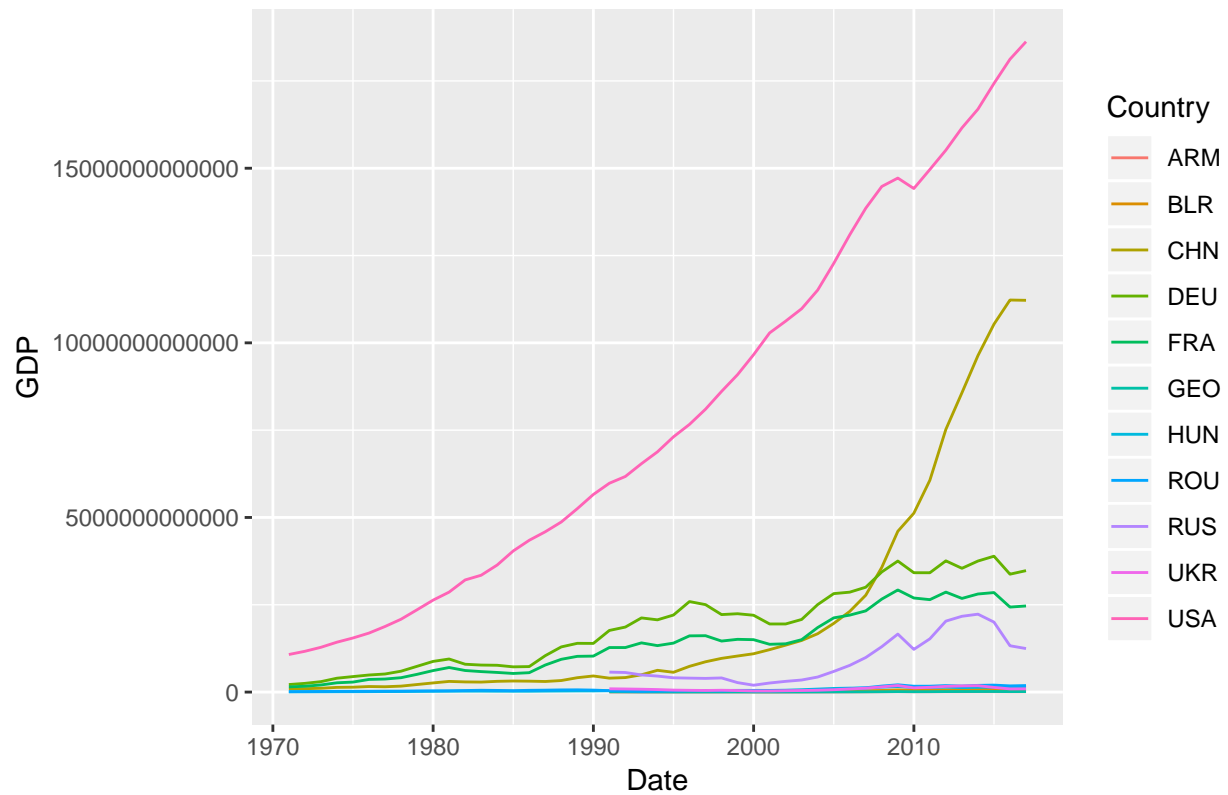
## Historical Deposit Rate Armenia VS Developed Countries
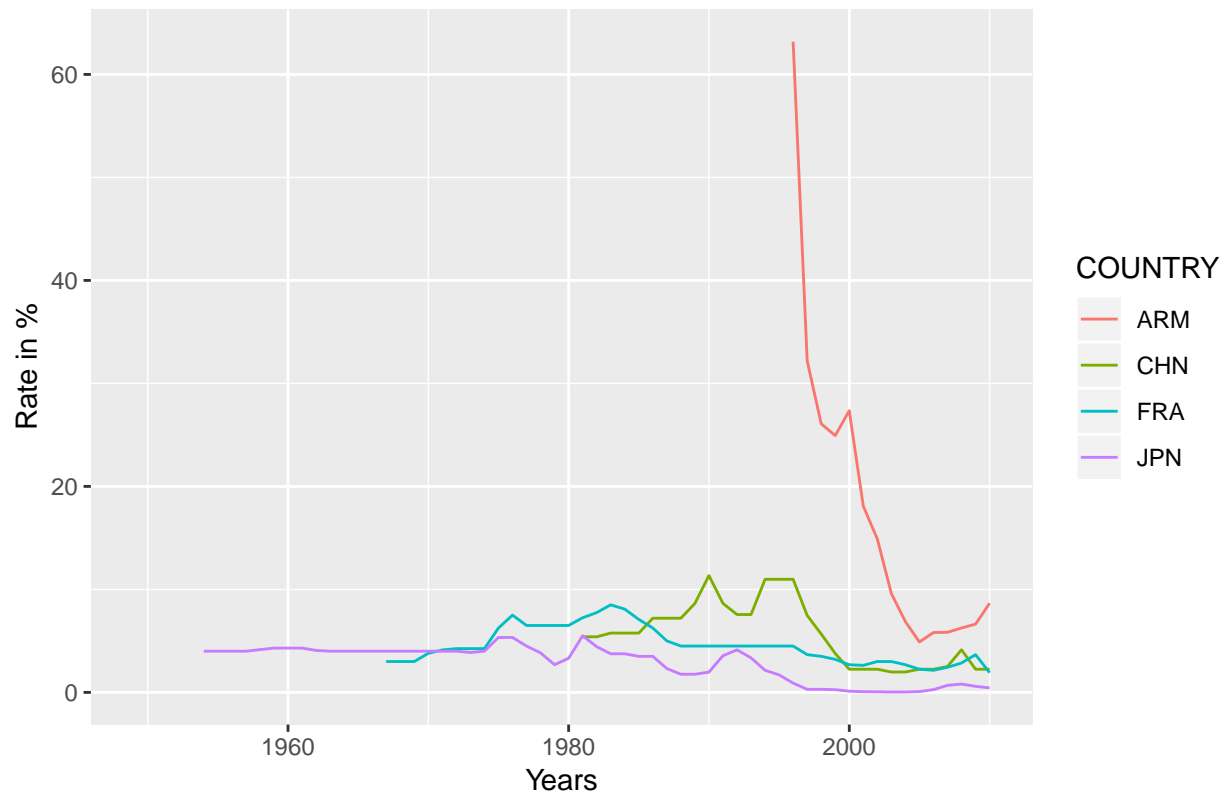
To conclude, many countries have the same fate as Armenia and even worse in terms of high Lending Rates during 90s, but gradually Lending Rates decrease and bundle around 20 percent. The ones with extremely high rates were part of USSR, which dissolved On December 25, 1991, ruining economies of countries which were once parts of it, including Armenia.

Now let's explore UN National Account Estimates dataset by using API. Documentation for this dataset follows>>> https://www.quandl.com/data/UNAE-United-Nations-National-Accounts-Estimates/documentation/data-organization

GDP of Developed and Developing countries during 1970–2016

The above graph is the first visualization which is related to UN National Accounts Estimates dataset. It shows the historical annual GDP of several Developed and Developing countries. Clearly Armenia and other developing countries are far behind USA and China; the giant drivers of Global Economy. USA shows steady increase during the whole period. Nonetheless, China has shown unprecedented growth rates from 2005 and clearly shows the patterns of catching up to USA in near future!!!

GDP of Developing countries in years 1970–2016
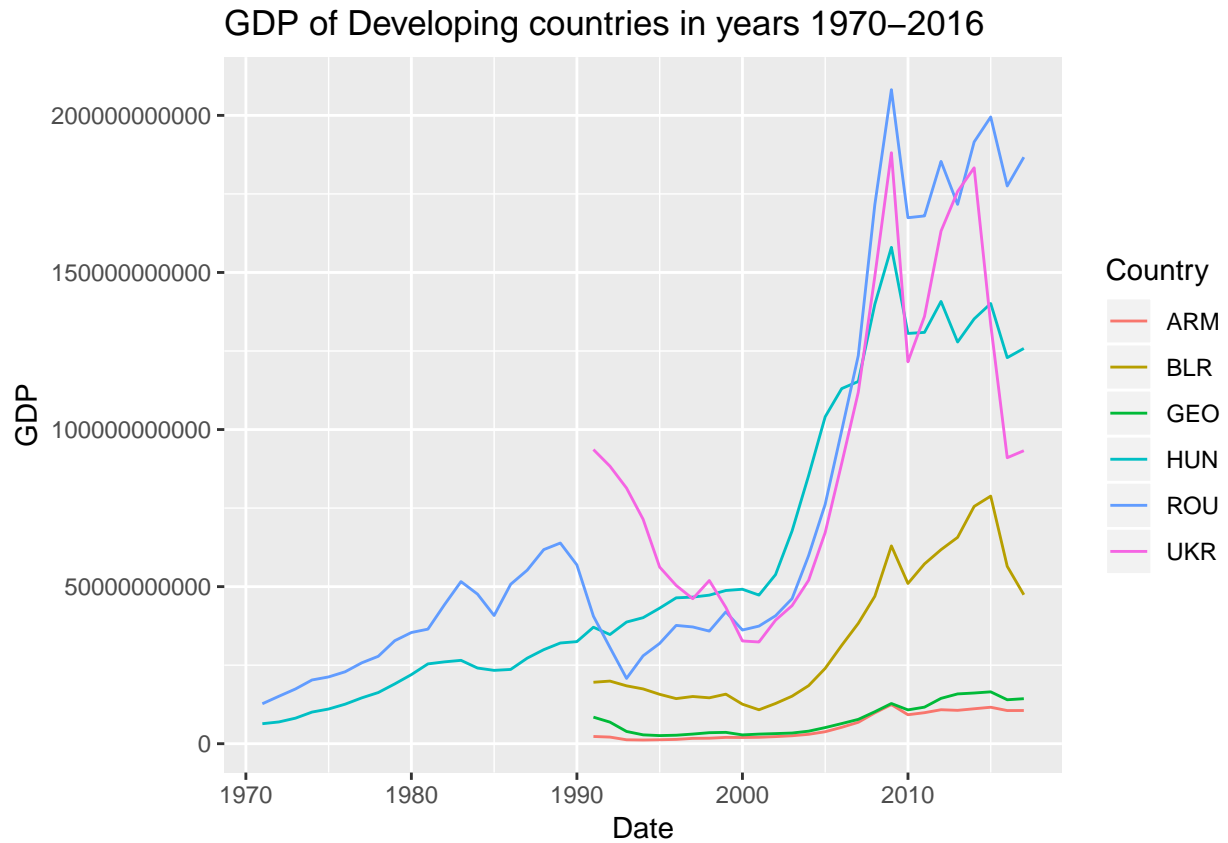
In order to see the patterns in Armenian GDP, let's turn to the subgroup in which Armenia may identified. The above graphs shows the history of Armenian GDP, which unfortunatelly is not so outstanding. Our GDP is on the last place compared to other countries of interest. Hence, we have room to develop our economy and to increase Gross Domestic Product.

GDP by countries
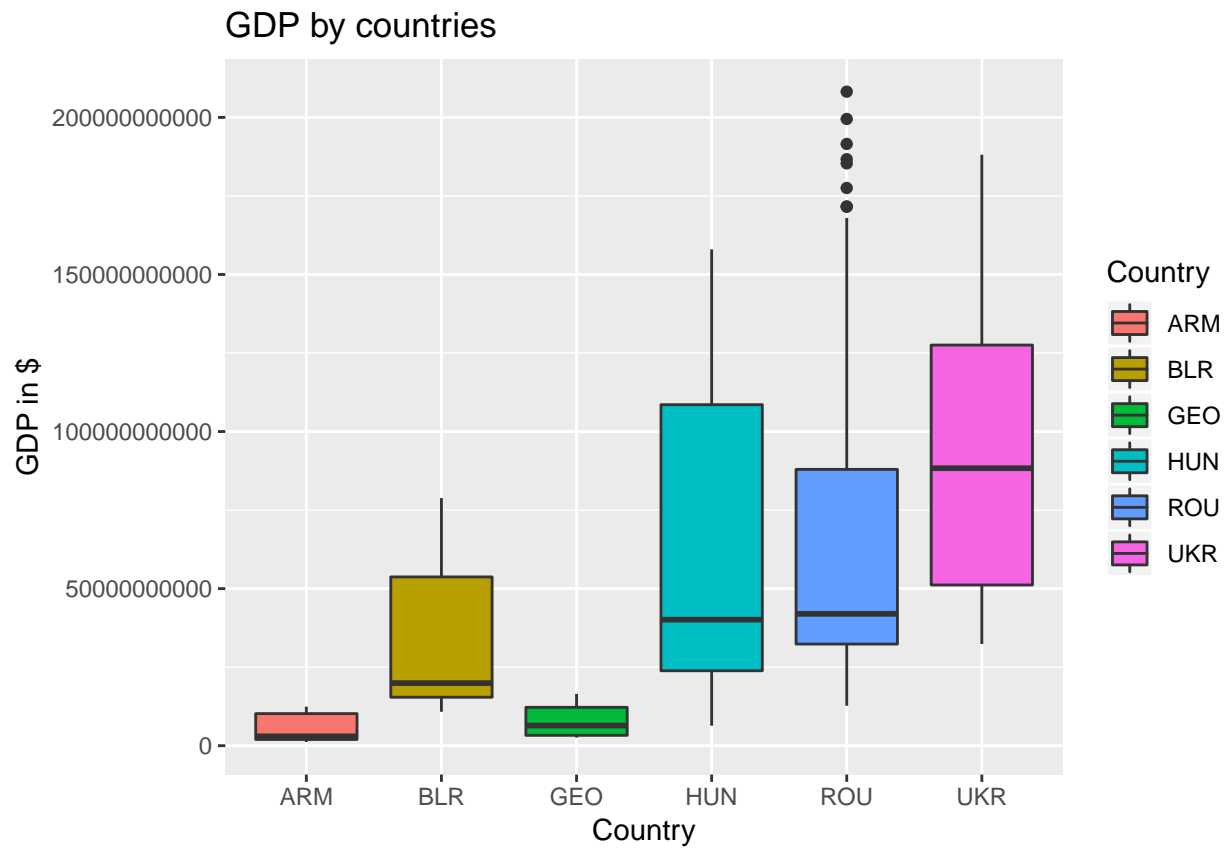
The above visualization also shows GDP of Developing countries in the format of boxplot.

GDP by developed countries VS Armenia
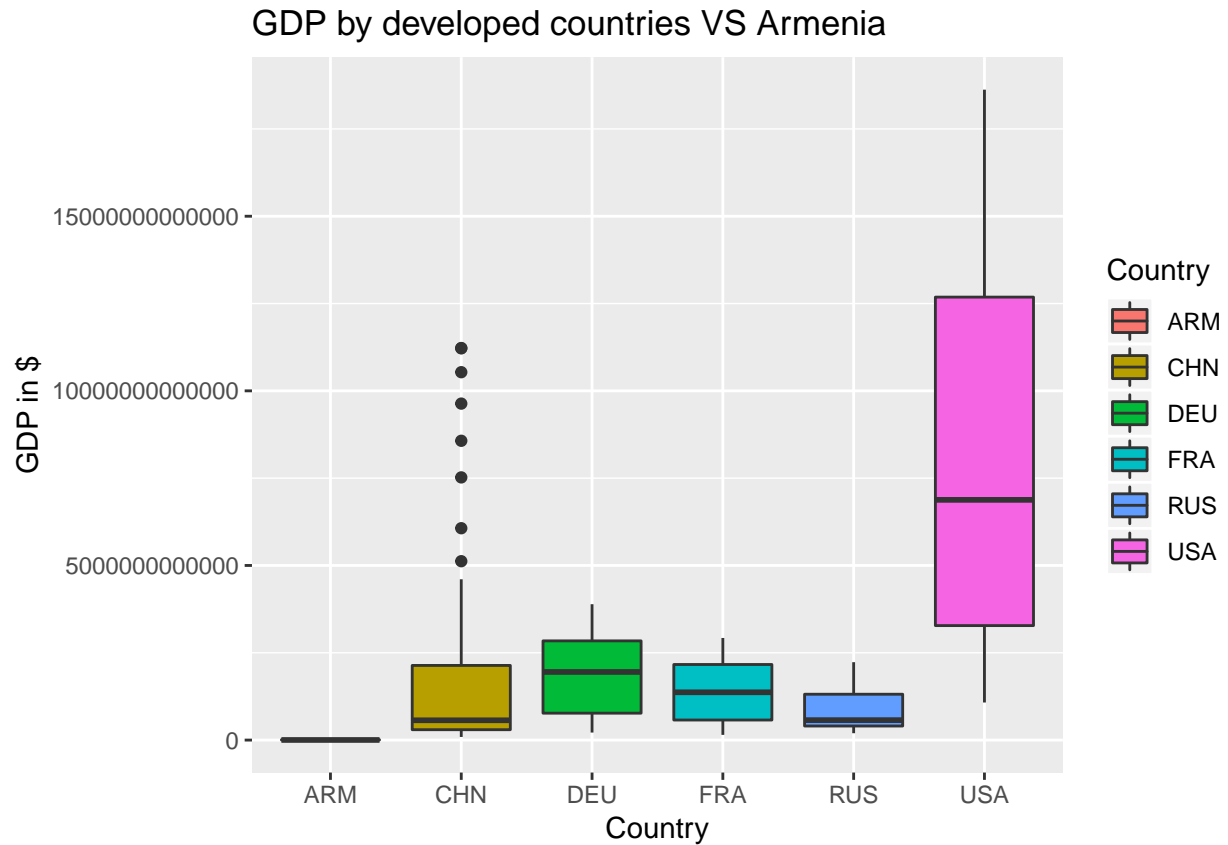
This graph may seem to be not so useful to identify new insights for Armenia, but let's turn our look to China. We may observe seven outliers and all of them lie on the upper side of boxplot which in reality is great as economy is developing and increasing with unprecendented in its history temps. This brings the question which actions China takes to achieve this and what small Armenia with its current economy and political situation may employ to achieve similar results.

Let's look what differentiates developed countries from Armenia by looking at export and import information.

Imports of goods and services
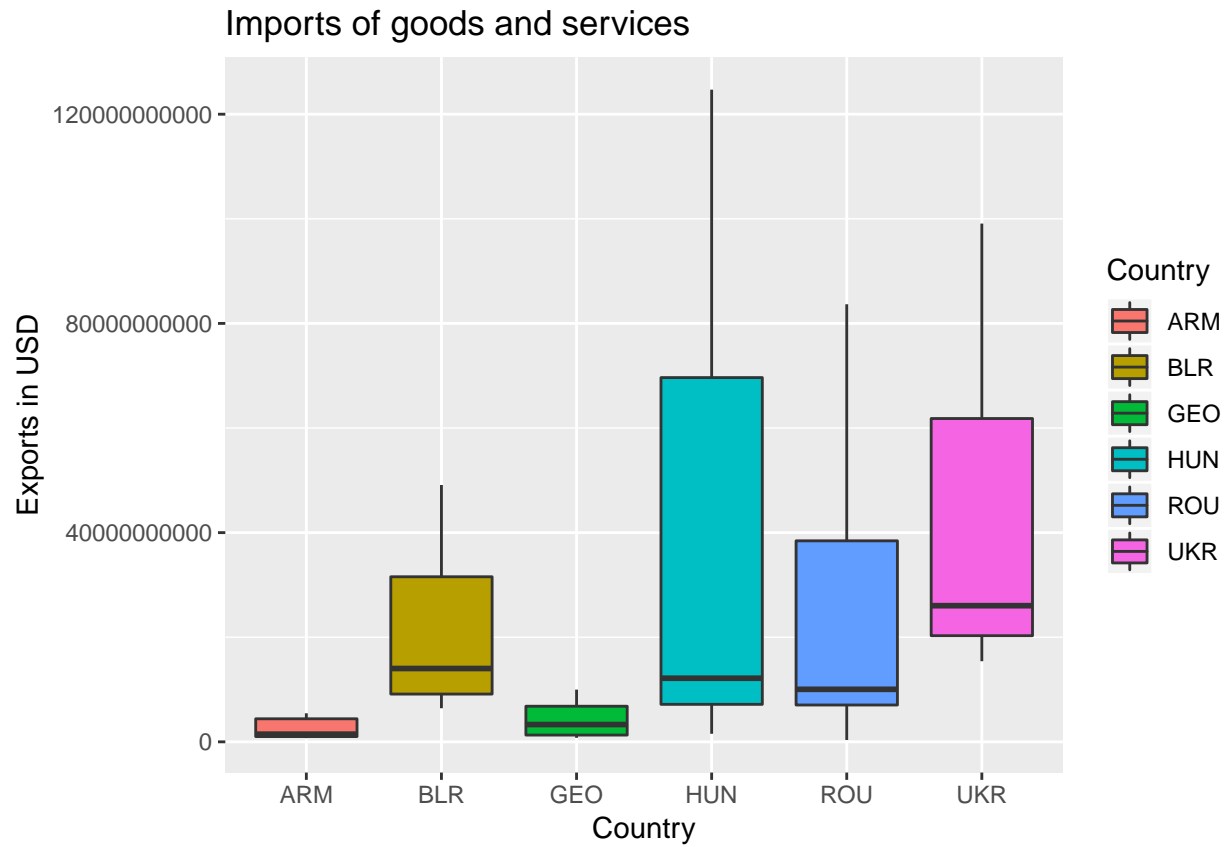
The following graph also confirms that Armenia does not import as much as other comparative countries. But by our volumes we are very close to Georgia.

Exports of goods and services
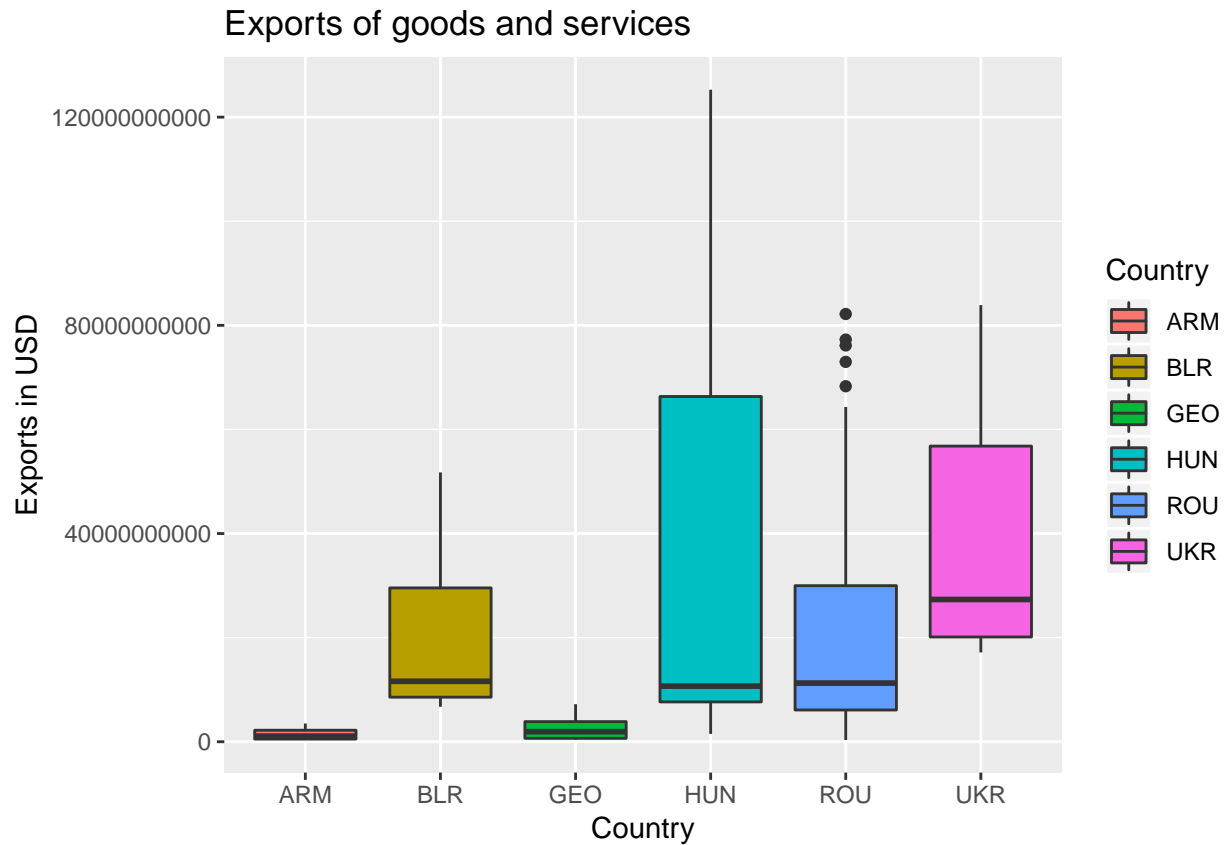
As we may clearly observe Imports and Exports of Armenia are significantly lower compared to other countries of interest. Probably the most logical reason is its isolation from two sides by our hostile neighbours. Our 'route of life' is the road which connects with Iran and Georgia. Nonetheless, we may observe the slight dominance of Imports compared to Exports in Armenia. Let's look closer to this observation by ploting the disctribution.

In addition Romania also shows very sound signs of increasing volumes in Exports as several outliers lie higher and the median is somewhat shifted towards 25-th percentile of the boxplot.

Imports of goods and services in Armenia 1990–2016

Let's explore the Exporting and Importing patters of Armenia by using density plots and compare Imports to Exports.

Exports of goods and services in Armenia 1990–2016
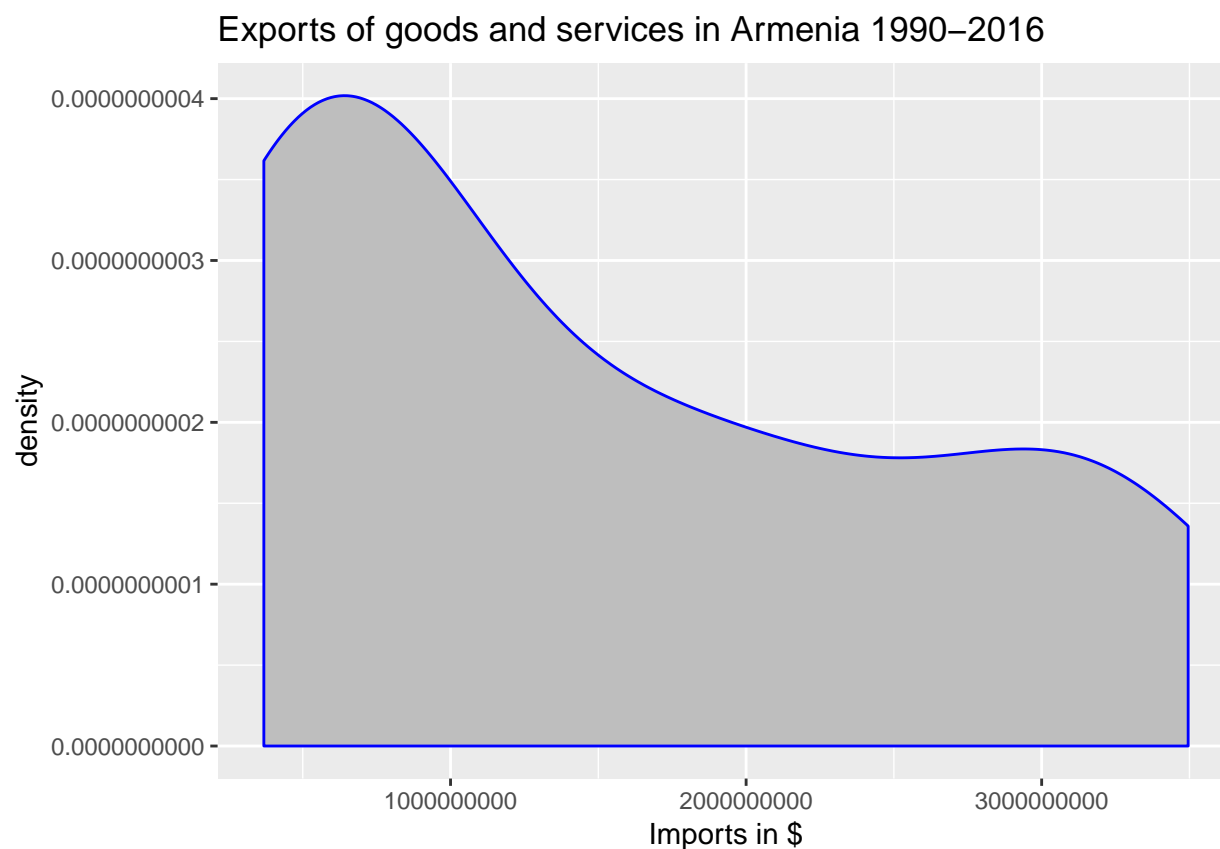
The density graphs of Exports into and Imports from Armenia clearly convey the message that Armenia is importing more than exporting. This is called 'negative trade balance'. Let's explore whether the situation is the same with other countries, particularly developed ones.

Exports VS Imports by countries
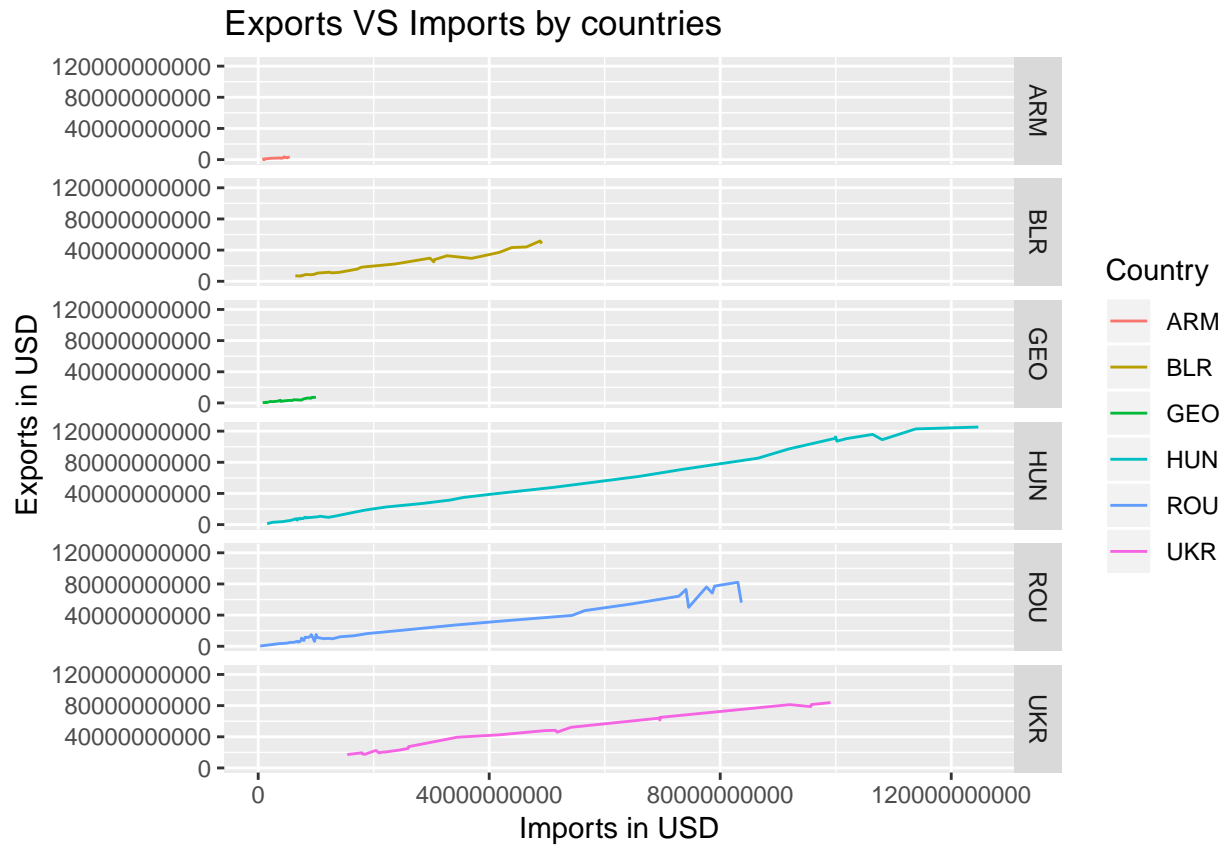
The faceted graph shows that this is not the deal and conversely all countries have steady increasing trend of Exports as compared to Imports. Whereas Armenia shows a somewhat straight line or even downward trend. Let's explore further.

Imports VS Exports of goods and services in Armenia
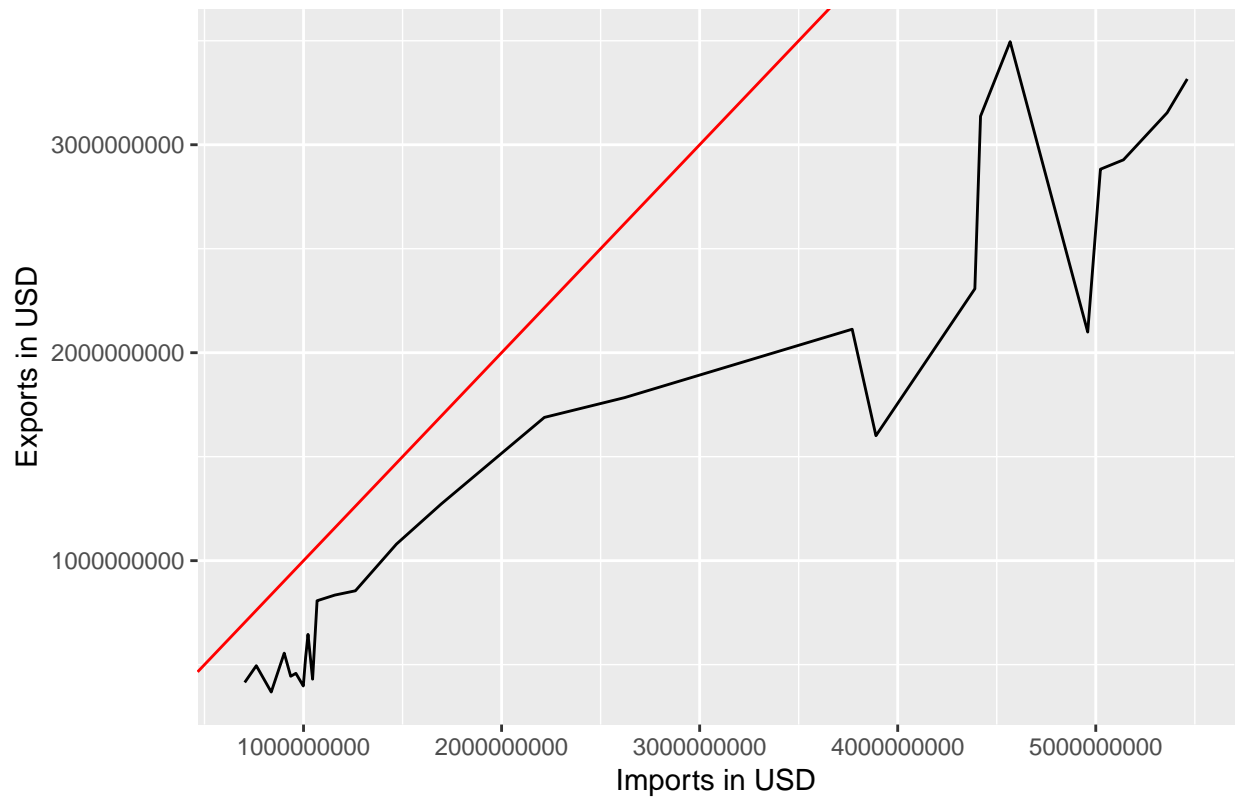
This graph clearly shows the situation with Armenian Exports and Imports. **The red line** shows the **Zero Trade Balance**, which may be observed when a country exports as much as it imports. Unfortunately, this is not the deal with Armenia, which lag behind its imports. To support our theory let's look at the Export and Imports of some developed countries, particularly China and USA.

Imports VS Exports of goods and services in United States

Imports VS Exports of goods and services in China
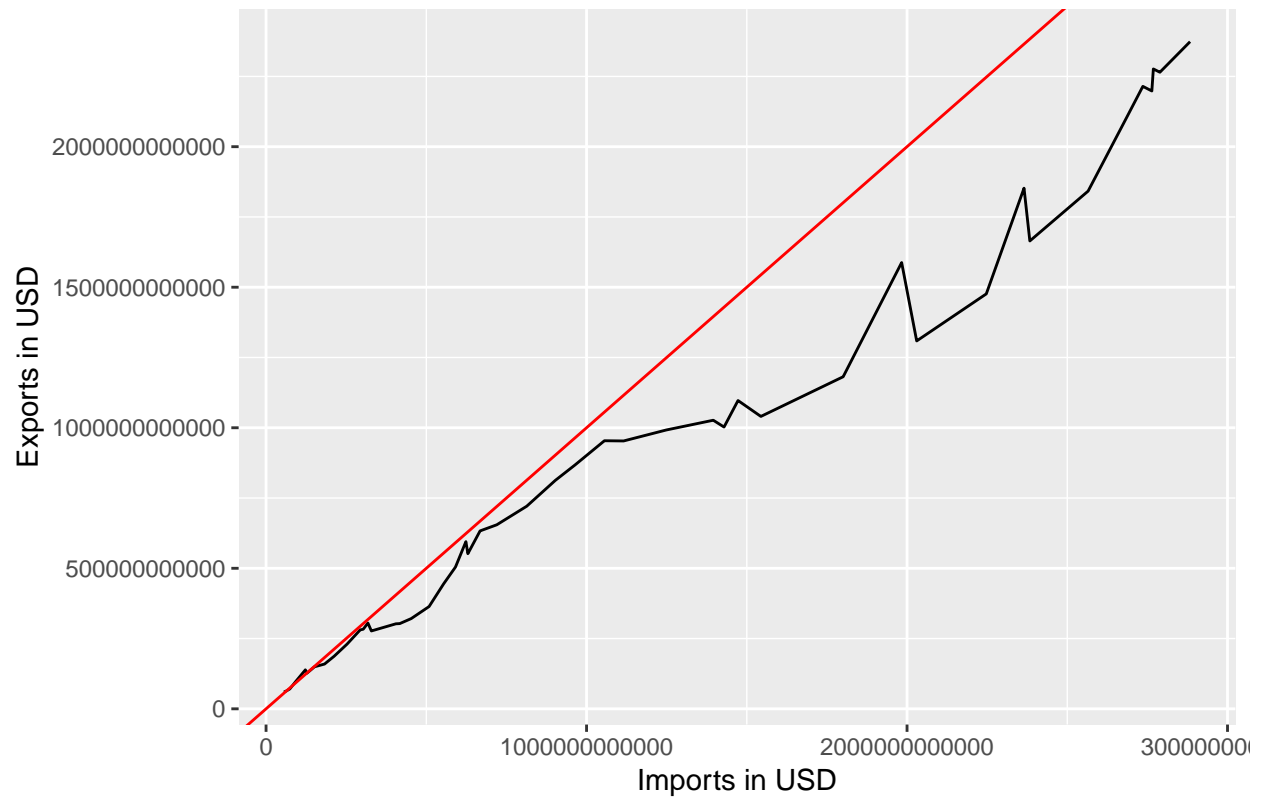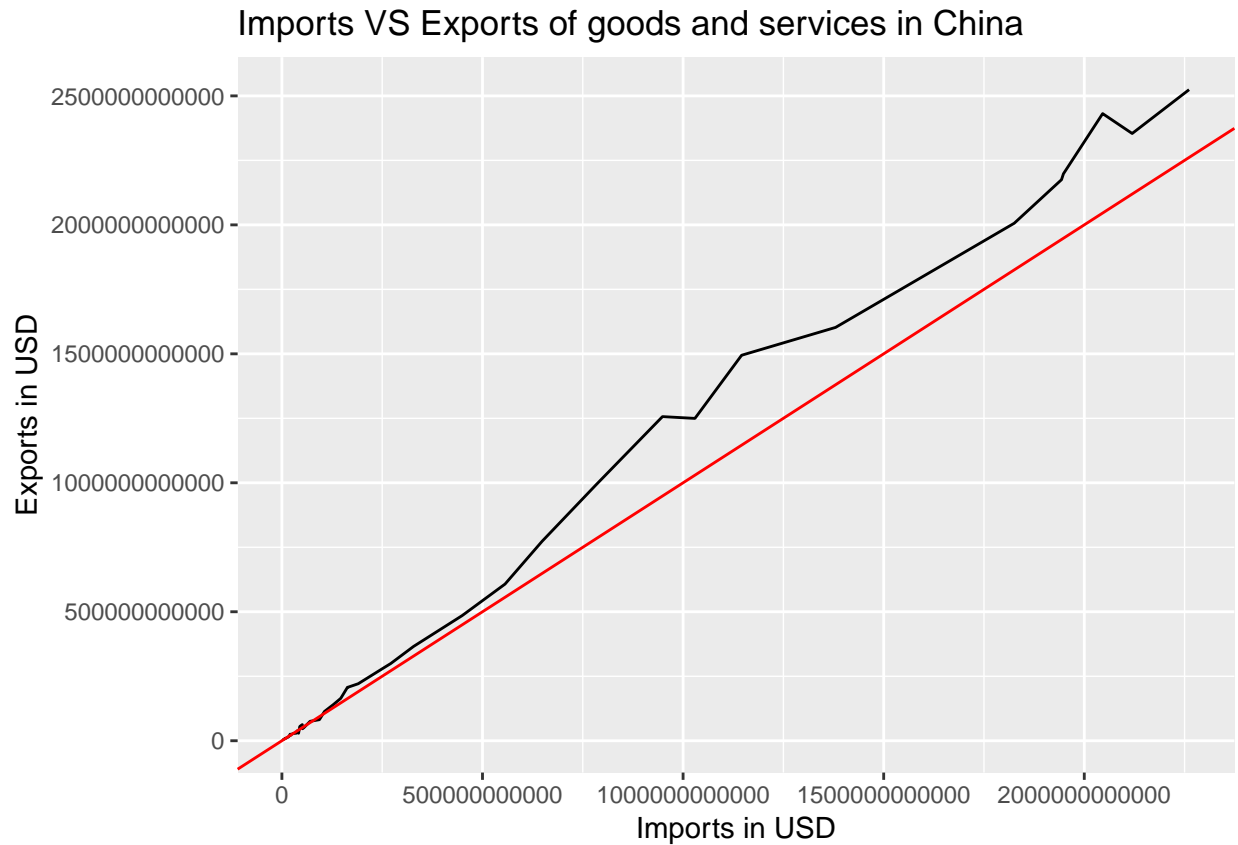
As a matter of fact, as previously mentioned **the red line** in each graph denotes the **Zero Trade Balance** for that particular country. We may clearly observe that Armenia is under the **Zero Trade Balance** line, hence it imports more than exports, whereas for developed countries such as China or USA their respective Trade Balance lines are very close or even higher than Zero Trade Balance.In case of China it is substantially increasing from year to year making China number one exporter of goods worldwide.

Historical Exports of goods and services in Armenia
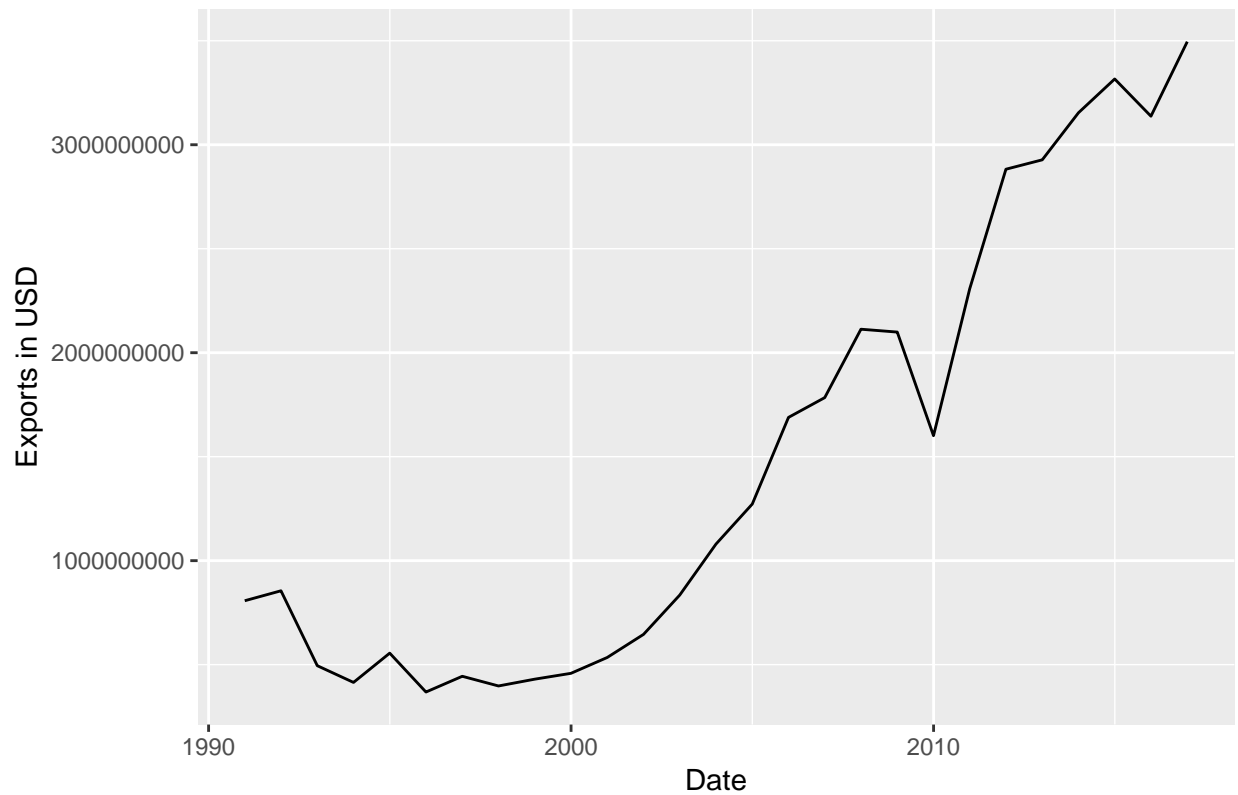
By further examining the historical patterns in exports of Armenia we may observe that the were some shock periods, particularly between 2009-2010 when the echoes of Global Financial Crisis reached Armenian economy but fortunatelly it was able to recover.

To summarize, this mini project may not cover all vital points for comprehensive analysis of Armenian economy taking into the context of homework.

Nevertheless this work may be a good starting point for future analysis, which will aim to identify useful patters to use in order to improve Armenian economy and develop the country.

Web Scraping and Text Mining

In this part we are going to use web scraping, text mining, regular expressions and data manipulation techniques in order to conduct an analysis of recent events that happened in Armenia. To further explain our goal, we are going to analyse the "Armenian Velvet Revolution" phenomenon from different perspectives. We will use International (Europe, US, etc.), Russian and Armenian news sources for conducting an analysis of attitudes towards the event and understanding the differences.

Our first step was to prepare tools and variables for retrieving and keeping information from news sources. We have created an empty data-frame with variables corresponging to Articles, Dates, Regional category and timeline indicator (before or during).

Next step was the creation of a universal function for scraing texts from web. The function operates by taking url, index, regional indicator and a text selector coupled with a date selector of webpages and putting all together into the previously initialized data-frame.

```
## [1] "Russian_Russia.1251"
```

We have 21 source webpages that are classified by 3 tags: Global, Russian and Armenian.

Here are the links for the webpages:

Global Sources:

- https://www.bbc.com/news/world-europe-43948181
- https://www.npr.org/sections/thetwo-way/2018/05/08/609364542/leader-of-armenias-velvet-revolution-takes-power-afte
- http://www.latimes.com/world/europe/la-fg-armenia-election-20180508-story.html
- https://www.theguardian.com/world/2018/may/08/hes-not-a-populist-hes-popular-nikol-pashinyan-becomes-armenian-p
- http://www.latimes.com/world/europe/la-fg-armenia-election-20180507-story.html
- https://www.aljazeera.com/indepth/opinion/armenia-velvet-revolution-masterclass-socialism-180507093619273.html
- https://www.nytimes.com/2018/04/23/world/europe/armenia-prime-minister-protests.html
- https://old.civil.ge/eng/article.php?id=31041
- https://www.washingtonpost.com/world/europe/armenias-pro-democracy-movement-defiant-in-face-of-parliament-reject 2018/05/02/dbb57c0a-4de7-11e8-84a0-458a1aa9ac0a_story.html?noredirect=on&utm_term=.016bf6ed063e
- https://www.economist.com/europe/2018/05/03/armenias-revolution-continues-as-its-opposition-leader-nears-power

Armenian Sources:

- https://news.am/eng/news/447914.html
- https://news.am/eng/news/446477.html
- http://arminfo.info/full_news.php?id=30467&lang=3
- https://www.azatutyun.am/a/29214870.html
- https://www.civilnet.am/news/2018/04/27/How-The-World-Sees-The-Velvet-Revolution/335224

Russian Sources:

- https://meduza.io/news/2018/04/22/v-moskve-u-armyanskoy-tserkvi-zaderzhali-mitinguyuschih-oni-ustroili-aktsiyu-v-p
- https://www.rbc.ru/rbcfreenews/5ae046e89a79471c04280ebe
- https://www.rbc.ru/rbcfreenews/5adc9bba9a794727c661e540
- https://www.svoboda.org/a/29174776.html
- https://www.currenttime.tv/a/29187295.html
- http://www.ntv.ru/novosti/2005924/
- https://rtvi.com/stories/pobeda-barkhatnoy-revolyutsii-v-postsovetskoy-strane/

Now the Russian sources have the articles mainly in Russian, and as acquiring the Google translate api key was not possible at the time, we manually translated these documents using Google and stored them in a csv file. We also added the missing dates in the dataset. Although loss in translation can occur, we are mostly concerned about the words and not the meaning, so this time of machine translation will suffice.

Note: We will attach the required "Ru_En_Articles.csv" file, so please have it in the directory

Before we clean the articles, we need to format the dates. However, in web scraping, dates can be found in various formats and in our case, no specific format could have been salvaged that could fit all of the dates.

However, through observation we noticed that all the articles had May 8th as the latest possible date, hence any article with a date other that the 8th would be recording the events during the revolution, otherwise it takes after.

All that remained was to extract unneeded characters such as hours, year and month dates. We used several regular expressions for this goal.

Finally, we are left with:

```
## [1] 23 15 16  8 27 22
```

Having done so, we were able to determine the timeline of the articles and begin working with wordclouds and other types of analysis.

For the Russian sources, here are the top 5 appearing words:

```
##              terms freq
## protest protest   29
## peopl     peopl   17
## opposit opposit   15
## polic     polic   15
## minist   minist   12
```

An initial guess would be that the Russian outlets may be focused on the chaos during the revolution.

It is worth noting that none of the Russian sources have a date of may 8th.
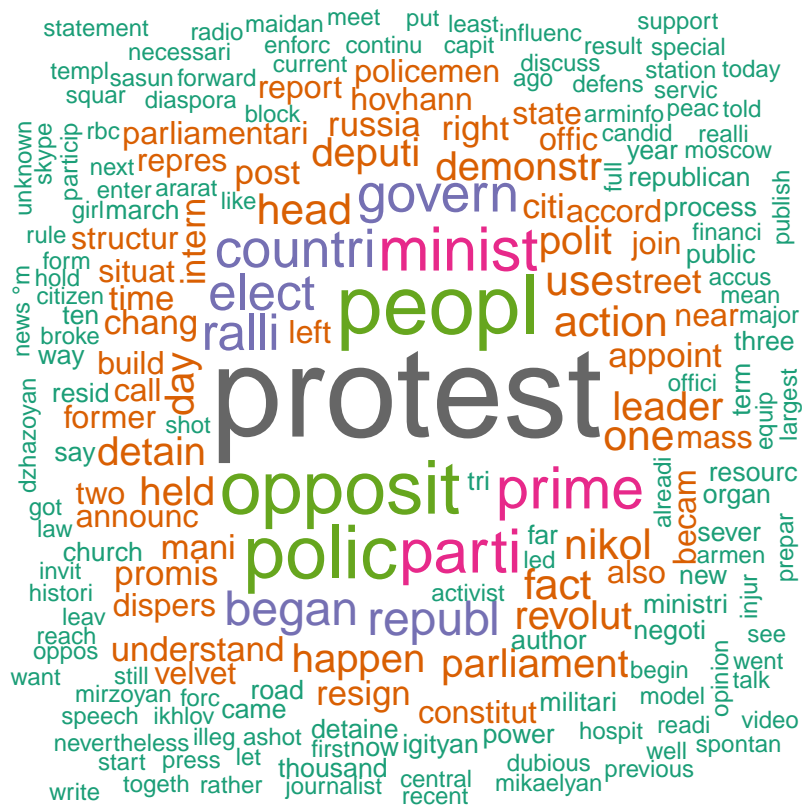
The focus is mostly on the governing body and the opposition that it is facing.

Here are the top 5 words for Armenian sources along with the word cloud:

```
##              terms freq
## protest protest   16
## elect      elect   13
## opposit opposit   13
## prime      prime   13
## minist    minist   12
```

Although the top 5 words were similar to the Russian sources, but the wordcloud displays more supportive words within the context of the revolution compared to the Russian counterpart.

Lastly, the global (international) sources:

```
##              terms freq
## protest protest   63
## prime     prime   55
## parti     parti   50
## minist   minist   49
## elect     elect   39
## peopl     peopl   38
```

The top words are once again the same. But the wording seems balanced and not skewed in any particular dimension.

We now move on to the final aspect, that is sentiment analysis.

With the use of the qdap library, we were able to extract the following analysis from the articles with the comparison point being whether they were published before or after May 8th.

```
##   Timeline total.sentences total.words ave.polarity sd.polarity stan.mean.polarity
## 1    After               4        4327       -0.230       0.069             -3.329
## 2   During              17       10477       -0.229       0.452             -0.506
```

The sentiment analysis results show that the overall polarity is negative regardless of the time of publishing. This is to be expected as it may be abnormal to see overly positive sentiments from these type of political reports.

Let's do the same by region.

```
##     Region total.sentences total.words ave.polarity sd.polarity stan.mean.polarity
## 1 Armenian               5        2219       -0.049       0.653             -0.074
## 2   Global              10       10199       -0.307       0.345             -0.889
## 3  Russian               6        2386       -0.249       0.237             -1.053
```

We See that the Armenian news outlets show a less negative polarity, this may be due to the inclusion of pro revolution websites within the sources. Other than that, the global sources (10 articles) show the most negative polarity overall.

Lastly, we form a comparison cloud between Armenian and Global sources for a better contrast.

hhk Armenia also
oppositionist call armenia'
public support the
day opposit yelk
pashinyan say
nnn russia countri
sargsyan soviet
new revolut russian
share Global capit

From the last result, we understand that the global outlets, were concerned about the revolution and about the Soviet connection throughout, while the Armenian sources mostly displayed an outlook supportive of the revolution and the change.