The evolution of Germany in the second half of the 20th century

Subject: "Web technologies to improve historical research"
Professor: Jorge Luis Morato Lara
Integrants: Juan Ruiz Vindel 100365483
Gabriela Fernanda Reyna Vinazza 100347854
María González da Silva 100332028
Francisco Pascual García Valcárcel 100347886
1. Summary

When analyzing the 20th century, Germany is probably the most interesting country to consider. Its role in the World War I (WWI), the Nazism, the war aftermath and its incredible recovery from the 50's onwards make this country worth studying.

The aim of this paper is to study the evolution of Germany from the 50’s to the end of the 20th century. Namely, a time span that will start with the recovery of the split Germany from the World War II (WWII), the evolution of both allies and soviet sides (the effects of the Iron Curtain) and the consequences of the URSS collapse in the last part of the century.

To do this, we will run a search using the proposed databases that will provide us an assortment of both quantitative and qualitative data. Special care will be needed to disregard biased information. Once we have this we will start the cleansing process and then the enrichment phase merging the data and building a unique database. This data will be interpreted to reach data-founded conclusions that will be written and described. Finally, both the database and the conclusions drawn from it will be uploaded to a webpage and published to let everyone have access to this information.

2. Research topic

Germany is one of the countries that has experienced more significant changes during the 20th century, especially during the second half of the century. After the World War II, the territory of Germany was divided into the Federal Republic of Germany (FRG) and the German Democratic Republic (GDR); in a less formal way, West Germany and East Germany.

This separation did not only affect the geographic field; other main aspects of the country were affected by this historical event. Furthermore, the two sides of Germany were based on different types of governments with two different types of policies. Because of this, the economic policies taken to recover from the World War II were completely different; the FRG was included in the Marshall Plan whereas the GDR focused more on its industrial and agricultural production. Moreover, the international relationships were different. This political differentiation was established until the reunification of Germany, process that started with the fall of the Berlin Wall. After the process of reunification, the role of Germany in the international framework became more relevant, especially in the European Union.
The data bases show us that the German economy suffered fluctuations, upwards and downwards, during the last decades of the 20\textsuperscript{th} century; but the overall effect is an impressive growth over this time span. We will be focused on the effect of the big shocks caused by the historical events, national and international, that affected the German economy and its population.

3. Data acquisition and cleaning tools

The first step in order to carry out our research consists of collecting information. We will carefully examine the different datasets from where we will be able to extract useful data. Some of them are already well-known, some others will be found thanks to the crawlers. It is important to be aware of the possibility of importing biased data since the very beginning, hence we will try to use as much reliable sources as possible.

Below it can be seen some of the datasets that we will use to build up ours:

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>World War II Database</td>
<td>Link</td>
</tr>
<tr>
<td>Populstat</td>
<td>Link</td>
</tr>
<tr>
<td>World Bank</td>
<td>Link</td>
</tr>
<tr>
<td>Population pyramid</td>
<td>Link</td>
</tr>
<tr>
<td>Deutsche Bundesbank</td>
<td>Link</td>
</tr>
<tr>
<td>Ancestry</td>
<td>Link</td>
</tr>
<tr>
<td>Eurostat</td>
<td>Link</td>
</tr>
<tr>
<td>Statistical Data Warehouse</td>
<td>Link</td>
</tr>
<tr>
<td>FRED</td>
<td>Link</td>
</tr>
<tr>
<td>OECD.STAT</td>
<td>Link</td>
</tr>
</tbody>
</table>

Once we have determined our resources it is time to focus on the indicators we will use to show the evolution of Germany across the second half of the 20\textsuperscript{th} century. Economic and demographic variables such as price index, population growth or industrial production will be used to study whether they reflect or not the historical and economic situation that Germany faced during this time span. Above all, we will have to pay attention to specific moments in History such as, for example, the afore mentioned fall of the Berlin Wall that had an impact for sure in many economic variables. Therefore, we will be able to draw conclusions resulting from these changes.

Importing all this information in a spreadsheet will allow us to start building up our database. In addition to that, we will have to be especially careful with possible typos in the data that could yield wrong results and therefore make our conclusions inconsistent.

Let us analyze an example on how we will collect this information. The table below shows German population (source: [http://www.populstat.info/Europe/germanyc.htm](http://www.populstat.info/Europe/germanyc.htm))
We have imported it into a google spreadsheet:

![Image of spreadsheet](http://example.com/spreadsheet.png)

Afterwards, we must carry out a cleansing process of such data. To do that, we choose to use the Open Refine program. This tool will allow us not only to clean the data but also to merge it. By combining the different data-sources that we have used our figures will be much more reliable.

In order to continue with our previous example we have uploaded the data to Open Refine.
and as it can be seen, the years are not properly written. That would be for instance, a first task to carry out through this tool.

We will do the same with the remaining data and once we have cleansed and merged it, would be the moment to start building up our database. With that purpose, and taking advantage of importing all the combined data into a google spreadsheet, we will do it through excel. Once we have compiled all the information, plot it and show it graphically will help us to evaluate and understand the evolution of the chosen indicators. Hence, we will be able to explain the changes that Germany suffered during that period.

Finally, we will upload our results to a webpage to make it available for everybody. To do it we will use WordPress (https://es.wordpress.com/create/) an online tool very useful to carry out this particular task. Below this lines we can find a first trial of what would be our webpage.
4. Process

Selecting the resources

Following the aforementioned steps, we have extracted useful data from these datasets:

1. Populstat
2. World Bank
3. FRED
4. OECD.STAT

Even though there were more datasets, some of them are not actually suitable. For example, World War II Database has a wide material, but it does not reflect the situation of Germany for the period of our interest; as well as Ancestry. Deutsche Bundesbank has a similar data to the World Bank, but we have preferred to work with the last one since it is much easier. Statistical Data Warehouse has a really useful data, but the steep obstacle is to find information for our target time span. Eurostat has information for almost everything, but is also a disadvantage because it is tough to extract the data.

On the other hand, Populstat has information of the population of Germany since 1700. This allow us to observe how the different historical periods have affected German population. World Bank has certainly convenient data for our task. It features because it is undoubtedly easier to work with it. OECD.STAT is practical because it is filled with relevant information. FRED is honestly the most useful dataset which we have used. Unquestionably, the interface is neat and you can determine the target period. Moreover, the interface allows to create your own dataset. As a detail, FRED has specific data from the Federal Republic of Germany (West Germany) during the epoch in which Germany was divided into two.

Nevertheless, we have faced an unexpected measurement of the employment rate by FRED which is called ‘discontinued’. That may be referred to some specific type of contract, thus, we decided to reject such information since we were not sure about its interpretation.

Once we have established our confident resources, we have to pay attention to the indicators we will need to recreate the evolution of Germany during the second half of the 20th century. According to their importance we have chosen: unemployment rate, manufacturing, Gross domestic product, life expectancy, consumer price index, exports, population. These indicators have all in common that they represent different aspects of Germany in the second half of 20th century, in addition of this, they are correlated. In other words, large GDP is linked to a larger life expectancy. Causality is one of the facts that we are trying to prove, how historical facts affected to Germany. And they encompass generally all the interesting characteristics that they help us with our objective.

Besides, we decided to compare with other countries looking for a suitable way to understand all this data. Japan, Canada, France and Great Britain are the chosen. We include GDP information for these nations.

Firstly, we have imported all this information into a spreadsheet in order to carry out the following processes such as cleansing and merging through GoogleRefine.
Cleaning and Merging the data

Once we have compiled the data regarding the above mentioned indicators we end up with a google spreadsheet that contains all the information (PHASE II file). Then, it is time to start the cleansing process by removing the columns that we do not need and bringing together those which are repeated. Such process is shown through screenshot1, screenshot2 and screenshot3.

![Screenshot 1](image1)

![Screenshot 2](image2)
Concurrently, we are working with other data sources and one of them looked very disordered and filled with information we do not need. Hence, it is necessary to carry out a cleaning process picking up the relevant data. In this case, as it can be seen in screenshot 4, we are dealing with unemployment rates and we want to focus on German data (DEU). Therefore, through the text filter we are able to remain with the preferred information (39 rows out of 858). As it can also be observed, we need to split the column into several ones to get the correct values (see screenshot 5).
Afterwards, we transform the common *year* column by removing the inverted commas (see *screenshot 6*), which will be used to do the merge, so that we will have the same format as in the PHASE II document.

After that, we remove the useless columns and rename them in order to facilitate the merge. As a result, we will have a spreadsheet as the one shown in the *screenshot 7*.
However, while trying to do the merge we realize we need to transform the text values of the *year* column into a number parameter (see the transformation in *screenshot 8*).

Finally, we merge the data from both files in order to include the column of the unemployment rates in our principal document (PHASE II). Using the GREL language (*screenshot 9*) we finally reach the desired result (*screenshot 10*).
Because we want to evaluate the development of the significant indicators during the second half of the 20th century in Germany, we need to compare those indicators with the ones of other countries. We focus on comparing the Life expectancy of Germany with the Britain and the French life expectancy and the German Unemployment rate with the Japanese and the Canadian rate. Moreover, we think that the most important indicator to compare is the GDP and because of this, we search the GDP evolution of the four mentioned countries during the same period of time. Now, we have to repeat the process of cleaning with the four different country data as we have done before. Next, we should merge the four spreadsheets with the principal (PHASE II). To do this, we repeat the merge process with the indicator previously mentioned. As an example here is the screenshot11.
As all the columns added are based on the Years column. We need to reorder the columns to have a more clarified idea of our results (screenshot12). Finally, we will export the project in order to carry out the visual and data analysis (screenshot13).
Visual and data analysis

Once we have done all the steps detailed above is due time to try to give an historical explanation to all of it. We have plotted all the not-disregarded data as it can be seen in the following graphs. We will go over them to give an historical and economic interpretation and we will see that all the indicators are affected by almost the same historical events.

In this two first graphs it can be seen the amount of exports of Germany and the manufacturing production both from the 60s onwards.

Obviously both have an upward sloping trend (which is, in general terms, the same in most of the developed countries). However, there are some particular fluctuations in both of them that are interesting. As for the exports, there are two big slowdowns in the trend: the first of it starts in the late 70s and lasts until 1986 more or less; and the second one is the year span between 1989 and 1993. What could be the reason for these fluctuations?

The first export shrink is likely to be caused by the OPEP crisis which was a global affair and affected almost every country in Europe in one way or the other. However, we strongly believe that the second shrink is a more endogenous thing: the fall of the Berlin Wall in November 10th 1989 due to the looming URSS collapse.
Besides, it is interesting how fast the exports post-depression growth was (from 1993 onwards). This only explains that this shrink was just because the social and political instability of the time which probably made investors and markets cautious about it; but the overall German economy was in a very good shape.

As for manufacturing it could be discussed something like a slowdown in 1973 (OPEP’s crisis), but we found more obvious the fall of the Berlin Wall issue.

When talking about Germany’s total population, it can be found a disturbance in the growth from 1973 to 1986. Although it starts in the same year than the OPEP crisis we do not feel very comfortable with stating that this entire disturbance is caused by this event. The reason for this is that the slowdown lasts too long in order to be the OPEP the unique answer; and it is common knowledge that Germany did a really good job when coping with the inflationary shock (see Consumer Price Index graph below). However, what we can see in this graph for sure is the effect of the end of the East Germany. This shock could have supposed a decrease in the birth rate from 1989 to 1991 so that the total population stayed constant these years: birth and death rates were equalized.
The life expectancy does not yield any strong evidence of Germany’s development (DEU in the graph). This is understandable since from 1960 onwards there was not any historical event which could have caused a shrink in this index. However, it can be note that although Germany’s life expectancy has been always parallel to France’s life expectancy; it does have a different growth over time when comparing it when with UK’s. In 1960 Germans died on average almost two years before their British counterparts. This gap turned upside down by the year 2000, when Germans died slightly after Britons.

When looking the Gross Domestic Product, we can see the above mentioned OPEP crisis in the 70s. Besides, there is a slowdown around 1989 which also exists, oddly enough, in both France (FRA) and United Kingdom (GBR). The biggest shrink, however, was triggered in 1995. It is way too late to consider it an effect of the East Germany end. Maybe it has something to do with the 1992 crisis which supposed a big shock for the European Union.

In a nutshell, we think it can be fairly stated that Germany’s economic development was huge. It is difficult to think, when looking these graphs starting in 1960, that Germany was incredibly destroyed two times in the first half of the twentieth century. This country’s resilience is study-worthy and whether other different country would have been able to bear the same historical events and have this fast growth in a couple of decades after is, in our opinion, unlikely.