**Final Report - Europe Database**

**INST 327 Section 0202 Team 1**

**Samir Abdalla, Iskander Lou, Asad Raheem,**

**Michael Chen, Nick Thomson**

**Introduction -**

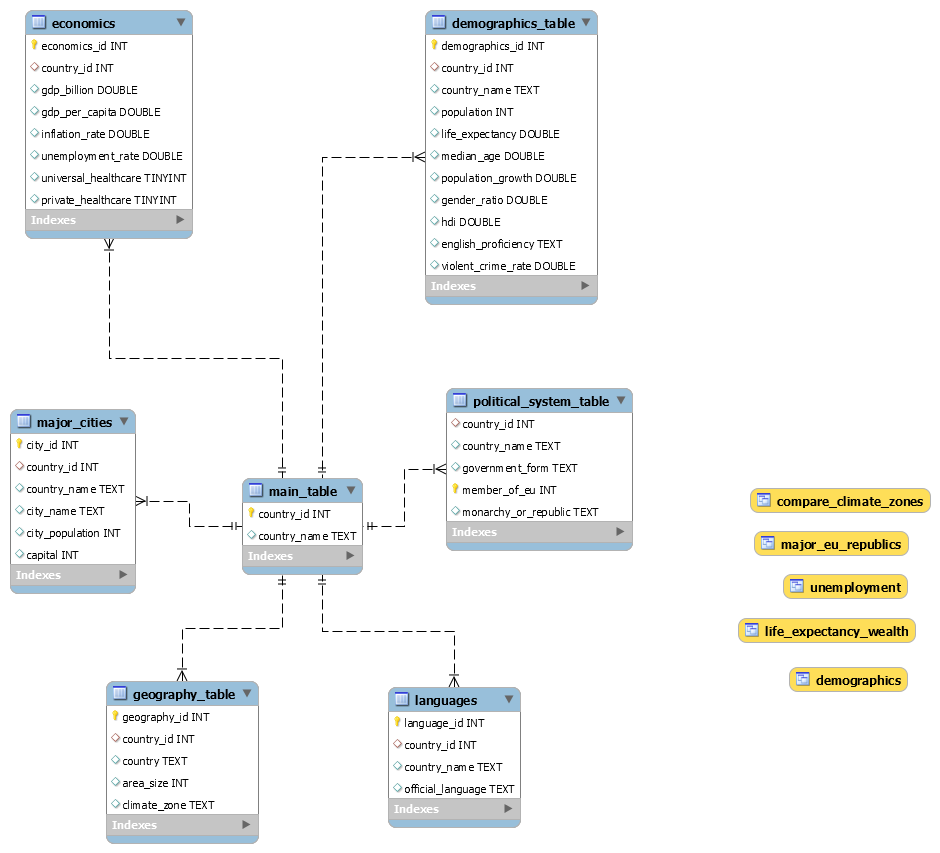
Europe is a diverse continent of numerous countries, all of which are very different from each other, and all of which have huge impacts on a global scale. The great amount of diversity, culturally, socially, and economically found in the great number of countries in Europe makes it extremely difficult to distinguish and establish ranking on any number of factors between these countries. That is why a database solely focused on outlining and highlighting different aspects of the countries within Europe will serve as a great information source for a variety of purposes.

**Database Description-**

Our database will help describe certain aspects of European nations, including statistics on their demographics, politics, economics, and geography. The main goal of this database will be providing a statistical breakdown on each individual nation, aiding in providing an objective overview of them. The time frame of this database will be 2016, as the census data on each nation from this year is quite populated. Basic statistics that will be tracked in our database will include, country name, size, population, language, gdp, and government form, which will make up the primary statistic of this database. Secondary statistics that will be tracked will include inflation rate, average temperatures, gender ratios, and violent crime rate.

* **Logical Design**

This database will highlight six key features of countries in Europe that are diverse enough that highlight different aspects of each country. As such each feature is broken into its own table, major cities, languages, demographics, political systems, economics, and geography. Along with these six tables there is a countries table that holds all the countries in Europe.



* **Physical Database**

Our Europe database that resides on SQL was successfully created in a new schema that we designed to contain all the necessary tables and columns to hold all the sample data correctly. We made sure to establish the correct relationships between all tables with appropriate primary/foreign keys to correctly establish our physical database.

* **Sample Data**

We ensured that our sample data was as accurate as possible by utilizing various reputable data sources, such as existing information regarding countries in Europe found on the CIA data website, Kaggle, and data.europa.eu. The sample data found in all of the tables are as accurate as possible, in terms of real life accuracy and in terms of accurate data type.

* **Views / Queries**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **View name** | **Req. A**  **(4 with JOIN)** | **Req. B**  **(3 with filtering)** | **Req. C**  **(2 with aggregation)** | **Req. D**  **(1 with linking table)** | **Req. E**  **(1 with subquery)** |
| **unemployment** | **✔** | **✔** | **✔** |  |  |
| **demographics** |  | **✔** | **✔** |  | **✔** |
| **compare\_climate\_zones** | **✔** | **✔** | **✔** |  |  |
| **major\_EU\_republics** | **✔** | **✔** |  | **✔** |  |
| **life\_expectancy** | **✔** |  |  |  |  |

**Changes from original design -**

There has been a huge change from the original design. In light of the current COVID-19 pandemic we had planned on developing a database focused on infectious disease outbreaks over time, but once dove deeper into this we found that it would prove extremely difficult to find sufficient and appropriate data sources. Given that we’d much rather work with actual data instead of just creating hypothetical data and the complexity of our original plan we came to the decision to focus on countries in Europe. We decided on this topic because we found ample data to work with, identified numerous aspects we could focus on, and acknowledged how useful a database such as this could be.

**Lessons Learned -**

After gaining an understanding of database ethics we made it a point to consider all aspects of our project and see if anything was unethical. We quickly realized database ethics would not be something we had to account for due to the fact that we were working with public data that neither compromised an individual’s or state’s privacy/security and because this database solely organizes existing information and is not seeking to mislead or misrepresent anything.

A lesson we learned, that is overarching, is that things do not go according to plan. Once we faced a roadblock with our original plan we had to learn that it is okay to go back to the drawing board and come up with something better. Once we were all able to come to this consensus and acknowledge that we had failed we were able to successfully pivot and ultimately create something better and more cohesive. Another major lesson came with the actual database development.

We quickly realized that the entire process will require a lot of time and trial. We thought it would be easy once we were able to identify several potential tables, data sources, and build our ERD model, but once it was time to import our data, write our queries, and build the database we soon realized that it would take multiple attempts, working together, and seeking help to accomplish what we had set out to do.

Lastly, a huge lesson learned was the entire process of building a database. We all learned so much about seeing a project from proposal to a working database and were able to implement SQL concepts we had learned into something real and useful. This will be extremely beneficial for all of us in our professional careers.

**Potential Future Work-**

The possibilities with this database are endless. We only focused on a specific number of variables associated with countries in Europe, but there is so much room to add tables and data that highlights other important factors of these countries. This database could also serve as a model for building similar databases focused on other countries, perhaps in other continents, and maybe even to compare countries of different continents. A great thing about this database is that it organizes and highlights aspects of these countries, which is useful for ranking and comparison purposes, but also that it can be utilized on a global scale for a wide array of purposes.