**CS 306**

**PROJECT STEP 1:** GROUP HAVUCLAR

<https://github.com/tansylu/Havuclar-CS306.git>

CONTRIBUTORS:

* Ada Yılmaz 29364
* Alım Gürbüz 28954
* Fatih Mutlu 27997
* Tansylu Akhmetova 30517
* Yusuf Erkam Köksal 29512

PROJECT TITLE:

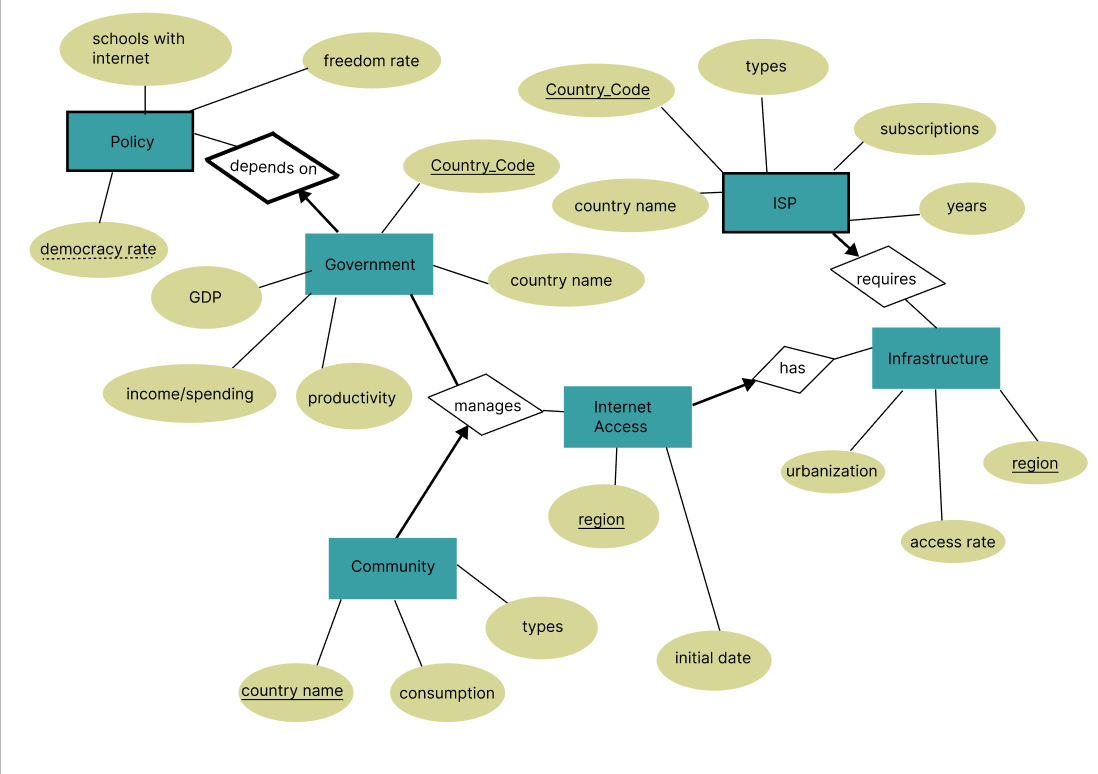
**“Analyzing the Relationship between Government, Freedom and Internet Accessibility Worldwide”**

CHOSEN WORLD PROBLEM:

The chosen World’s problem is the internet access problem all over the world and its relation with the governments’ policies of countries.

ER DIAGRAM:

In order to achieve an investigation on this topic the ER diagram illustrated is like this;



This ER diagram that examines the chosen topic is focused on the relations with internet access around the world. First of all, in order to maintain ISP( internet service providers) the weak entity and infrastructure that it is connected to are needed so the ‘has’ relation connects them. Thebn, another relation that is maintained is the ‘manages’. That connects the government and community to it. Government also has a weak entity ‘policy’ that it depends on.

In the diagram the special keys and partial keys are stated with their respective notations.

DATASETS AND STEPS :

The datasets that are used in this project are called:

* **INTERNET ACCESS**

-the year the internet was started to be used in the country

* **COMMUNITY**

-Number of Internet users

* **INFRASTRUCTURE**

-Access to electricity urban vs. rural

-Share of individuals using internet

-Share of urban population

* + **ISP**

-Fixed broadband subscriptions

-Fixed telephone subscriptions

-Mobile cellular subscriptions

* **GOVERNMENT**

-GDP

-Government spendings

-Productivity

* + **POLICY**

- Proportion of schools with access to the internet

-Primary

-lower secondary

-upper secondary

-Electoral Democracy

-Freedom Expression

The steps that are used to operate those can be summed up as;

1. The datasets are downloaded from the [www.ourworlddata.org](http://www.ourworlddata.org) website as .csv files.
2. The desired columns from datasets are chosen and put into separate .csv files.
3. The .csv files that will be used for one entity are merged together.
4. The merged files are examined and the parts with no data are cleared.
5. The data that are repeating and the ones that are not countries such as continents, and the ones that do not have Code data, are deleted.
6. Depending on the main entity, the Code or Date columns are filtered according to the required data.
7. After the csv files are in the desired format, they are imported in the MySQL Workbench in the table format.( they are in out git repository)