Project 1: World Database

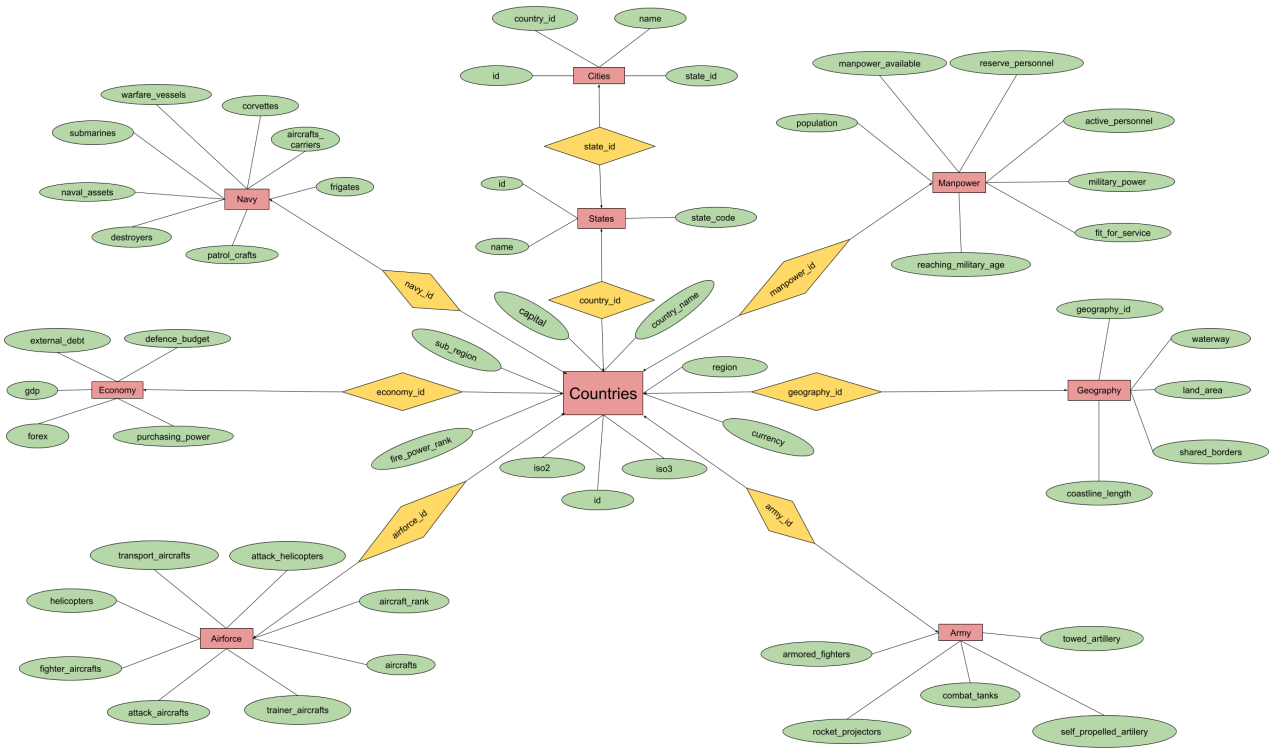
Kuriseti Ravi Sri Teja (2019CS10369) Aryan Santosh Sahoo(2019CS10335)
Gattu Karthik(2019CS10348)

March 1, 2022

1 Section 1

We made a database that consists of various characteristics of nations around the world i.e.,states,cities,military strength of triad forces (army,navy and airforce)(This new addition was motivated by Russia vs Ukraine war currently going on),geographical characteristics(Say Area,coastline etc..),economic characteristics(GDP,PPP,Forex Reserves etc.).For this project,we found motivation from sub-domains of Wikipedia such as ProofWiki,Wikitex that hosts data about a particular topic which are open source and can be accessed and edited by anyone.

1.1 ER-Diagram



1.2 Entities and Attributes

Entity: country

Attribute-Name
id
country_name
iso3
iso2
capital
currency
region
subregion
fire_power_rank
army_id
navy_id
airforce_id
economy_id
manpower_id
geography_id

Entity: army

Attribute-Name
army_id
combat_tanks
armored_fighters
self_propelled_artillery
towered_artillery
rocket_projectors
geography

Entity: navy

Attribute-Name
navy_id
naval_assets
aircraft_carriers
frigates
destroyers
corvettes
submarines
patrol_crafts
warfare_vassals

Entity: airforce

Attribute-Name
army
navy
airforce
economy
manpower
country
geography
states
cities

Entity: economy

Attribute-Name
economy_id
defence_budget
external_debt
forex
purchasing_power
gdp

Entity: manpower

Attribute-Name
manpower_id
population
man_power_available
fit_for_service
reaching_military_age
military_power
active_personnel
reserve_personnel

Entity: geography

Attribute-Name
geography_id
land_area
coastline_length
shared_border
waterway

Entity: states

Attribute-Name
id
name
country_id
state_code

Entity: cities

Attribute-Name
id
name
state_id
country_id

2 Section 2

1. We downloaded the data from Kaggle from the links
   - **https://www.kaggle.com/tanweerulhaque/countries-states-cities-dataset**
   - **https://www.kaggle.com/blitzr/gfp2017?select=GlobalFirePower_multiindex.csv**
2. We have used the readymade data that was available on Kaggle as it would take a lot of effort to scrape the web for the data.
3. Since the data was from two different sources,we had to verify whether the same attribute(country id) was different in different sources and it was indeed different.So we used pandas package to clean up data by using hashmaps to map the id's from one dataset to another.Also in one dataset the table was not in EBNF and hence we cleaned it up to convert into EBNF by creating new Id's.
4. Statistics (Include a table which lists the name of the table, number of tuples in the table, time to load, size of raw dataset, size of raw dataset *after* cleanup.)

2.1 Statistics

Table-1:Number of tuples in tables and time to load.

Table Name	Time to load in milliseconds	Number of Tuples
army	217	133
navy	276	133
airforce	137	133
economy	146	133
manpower	161	133
country	169	133
geography	149	133
states	168	3955
cities	596	143560

Table-2:Size of table before and after clean-up

Table Name	Size before cleanup	Size after cleanup
army	3 KB	3 KB
navy	3 KB	3 KB
airforce	4 KB	4 KB
economy	6 KB	6 KB
manpower	8 KB	8 KB
country	88 KB	11 KB
geography	4 KB	4 KB
states	275 KB	112 KB
cities	8560 KB	3837 KB

3 Section 3

User's view of the system (this should be an itemized list)

1. The user can view the
   - Key details of all the countries in a table and can search for anykey in the Search Menu
   - Enter the name of country to get all the states in that country in a table
   - Enter the name of state to get all the cities in that country in a table
   - View the economic,geographic details,military details,man_power and fire_power rankings in respective tables.
2. The system can view the
   - Insertion of the data city-name,city-id,state-name,state-id,country-name,country-id into tables which cause update of city table.(We will add cities based on state-id).
   - Since only there is a limited set of input there is no need to fire triggers.
3. **Table:Queries tested and runtimes**

Query Name	Input	Time taken
Get_cities	Manipur	100 ms
Get_States	India	20 ms
Fire_Power_Rankings	No-Input	10ms
Army_Details	No-Input	15 ms