Historical Analysis

Database Creation and Analysis

Project Proposal

Project Proposal

Economics Data

Project Title - The Rhymes of History

In this project we look into historical data and store that data to study whether they can determine any patterns across numerous economic indicators. This provides the best chance to be positioned well going into turbulent economic conditions and assist with business decisions moving forward.

We look into the different data involving Analysis of Government Debt, Inflation Rate and Foreign Direct Investment

Team Members

Maria Buenaventura

Parkavi Jayachandran

Rajiv Wick

Balvinder Rajbans

Project Description/Outline

Analysis on Government Debt, Inflation Rate and Foreign Direct

Investment.

To create python code that is able to accept csv files, transforms the data to be used and stored in an SQL database. This is to be done alongside the creation of formal specification documentation so that any new data can follow these guidelines for future import into the database and seamless work with the other tables stored.

Research Questions to Answer

1. Identify any relationships between Government Debt, Inflation Rate and Foreign Direct Investment
2. Using the data available, model any relationships visually

Datasets to Be Used

https://data.world/brianray/gapminder-inflation-annual
Inflation_annual.csv

https://data.world/brianray/gapminder-foreign-direct-inves
Foreign_direct_investment.csv

https://stats.oecd.org/index.aspx?lang=en
-Central Bank Debt

(Finance/Central Government Debt/Total central government debt (% GDP))

Government_Debt.csv

Dataset Tables - Raw

Inflation annual.csv

	Inflation, GDP deflator (annual %)	1961	1962	1963	1964	1965	1966	1967	1968	1969	...	2002	2003	2004	2005	2006
0	Abkhazia	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
1	Afghanistan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	3.845357	6.780488	8.405298	2.413906
2	Akrotiri and Dhekelia	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
3	Albania	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	3.300196	3.383486	6.007745	3.469252	1.995241
4	Algeria	3.47172	2.35128	0.549331	1.695183	1.501331	1.817815	1.312041	3.142056	1.921084	...	1.906329	8.323803	10.629329	16.459258	11.282812

5 rows × 52 columns

Foreign direct investment net inflows of GDP.csv

	Foreign direct investment, net inflows (% of GDP)	1970	1971	1972	1973	1974	1975	1976	1977	1978	...	2002	2003	2004	2005	2006
0	Afghanistan	0.013151	0.024575	0.009401	0.015577	NaN	NaN	0.001565	-0.008126	NaN	...	1.152364	1.212725	3.276532	3.976666	3.082130
1	Albania	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	3.034135	3.149791	4.572142	3.133523	3.561523
2	Algeria	1.647378	0.011817	0.613595	0.585191	2.710127	0.764883	1.054808	0.850913	0.51263	...	1.866684	0.931657	1.037359	1.056292	1.532824
3	American Samoa	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN
4	Andorra	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	NaN	NaN	NaN

5 rows × 43 columns

GOV_DEBT_27092022120932650.csv

	COU	Country	DTYP	Type	FREQ	Frequency	UNIT	Unit	DVAR	Variable	TIME	Time period	Value	Flag Codes	Flags
0	AUS	Australia	AMT	Stocks: Outstanding amounts	A	Annual	PCT	Percentage	P1	Total central government debt % of GDP	1997	1997	18.476	NaN	NaN
1	AUS	Australia	AMT	Stocks: Outstanding amounts	A	Annual	PCT	Percentage	P1	Total central government debt % of GDP	1998	1998	15.578	NaN	NaN
2	AUS	Australia	AMT	Stocks: Outstanding amounts	A	Annual	PCT	Percentage	P1	Total central government debt % of GDP	1999	1999	13.741	NaN	NaN
3	AUS	Australia	AMT	Stocks: Outstanding amounts	A	Annual	PCT	Percentage	P1	Total central government debt % of GDP	2000	2000	11.361	NaN	NaN
4	AUS	Australia	AMT	Stocks: Outstanding amounts	A	Annual	PCT	Percentage	P1	Total central government debt % of GDP	2001	2001	9.552	NaN	NaN

Applications:

Python - libraries: Pandas, SQLAlchamy
Juptyer Notebook
Database - PGAdmin (PostgresSQL)

Process:

Open python file
Import csv files into Pandas - > transform tables to formal specification spec -> connect to postgresSQL database -> load data.

Formal specification to be created that defines the tables structures importable into postgresSQL database.

Exact :

The datasets are all in csv format, using Python's Pandas library we can extract the csv and store them as Dataframes to allow for transformation.

Transform:

Filter data for countries that are being focused on for the analysis and the year range.

Case 1:

Country - Australia

Year Range - 2001 to 2010

Manipulate the tables necessary in order to allow them to be uploaded to the postgresql database inline with the tables created. Two of the tables will be straightforward and the dataset on Government Debt will need to be restructured to comply with the table structure in the database.

Data Modelling - Tables :

Table: Government_Debt
Primary Key - Country

Table: Foreign_Investment
Primary Key - Country

Table: Inflation
Primary Key - Country

Project Title	What makes the world go around?
Team Members	Maria Buenaventura Parkavi Jayachandran Rajiv Wick Balvinder Rajbans
Project Description/Outline	Analysis on Government Debt, Inflation Rate and Foreign Direct Investment
Research Questions to Answer	Identify any relationships between Government Debt, Inflation Rate and Foreign Direct Investment
Datasets to Be Used	https://data.world/brianray/gapminder-inflation-annual https://data.world/brianray/gapminder-foreign-direct-inves https://stats.oecd.org/index.aspx?lang=en -Central Bank Debt

Breakdown of Tasks:

- Project proposal (20 points)

  - ☐ Proposal cites at least two sources of data
  - ☐ Proposal includes the type of final production database to load the data into (relational or non-relational)
  - ☐ Gives relevant and succinct description of findings (2–3 sentences)

- Document Dataset - Analysis

  - Dataset 1
  - Dataset 2
  - Dataset 3

- Document Data Transforming - Table Structures/PandasDataframes

    - Transforming the data (cleaning, joining, filtering, aggregating, etc.)

      Extract: indicates the original data sources and how the data were formatted at a professional level

      Transform: explains what data clearing or transformation was required at a professional level

- Create database
    - Load

- Technical report (20 points)

    - [ ] Extract: indicates the original data sources and how the data were formatted at a professional level
    - [ ] Transform: explains what data clearing or transformation was required at a professional level
    - [ ] Load: explains the final database, tables/collections, and why the topic was chosen at a professional level

- GitHub repository (20 points)

    - [ ] Successfully uploaded to GitHub; demonstrating professional quality of presentation
    - [ ] GitHub repository is free of unnecessary files and folders and has an appropriate .gitignore in use
    - [ ] The README is customized to a professional level