

Global Economic Data: Comprehensive Analysis and Insights
By Jumana Jouhar, Jyothsna Sara Abey, Muhamed Adil, Sam Thomas
Data Analytics Using IBM Cognos
Team Cognova

1 INTRODUCTION

1.1 OVERVIEW

Analysing and deriving actionable insights from global economic data is important for businesses, policymakers, and researchers to make informed decisions and navigate the complex global landscape. The project aims to address the need for leveraging IBM Cognos to achieve comprehensive analysis and actionable insights from global economic data. IBM Cognos plays an important role in analysing global economic data as it enables users to create visualisations, interactive dashboards, reports, and stories. They allow various stakeholders to understand economic trends and gain useful insights, helping businesses in the long run. The tool supports trend anticipation for businesses, aids policymakers in formulating effective strategies, and helps researchers uncover significant patterns.

1.2 PURPOSE

The purpose of the project is to meet the needs of businesses, policymakers, and researchers by providing a comprehensive understanding of economic trends, enabling them to identify risks and opportunities, and supporting data-driven decisions for business growth and navigating the global economic landscape. It can address business challenges by providing practical insights, thorough analysis, easy accessibility, informed decision-making, and effective sharing of economic insights. Businesses can predict trends, adjust strategies, and make informed choices. Risk management and resource allocation is improved. Advanced analytics aid in forecasting and sustainable strategic planning. Comprehensive economic insights influence policy decisions, promoting prosperity, reducing inequalities, enhancing living standards, and fostering inclusive growth and societal unity.

2 LITERATURE SURVEY

2.1 EXISTING SOLUTIONS

Existing solutions include Bloomberg Terminal and Datastream. Bloomberg Terminal is a software used in finance, investment, and trading. It provides a wide range of financial information, tools, and analytics that help with decision-making. It thoroughly covers financial data, news, and analysis. However, it is costly to access, so it is mainly used by institutions, not individuals. Datastream is a global financial database. It offers data on stocks, currencies, economic indicators, and more for 175 countries and 60 markets. The issues with the existing solutions are that they are costly and mainly focus on financial data, often leaving out other important factors that impact the economy.

2.2 PROPOSED SOLUTION

The proposed solution is to use IBM Cognos, Python, and web frameworks like Flask to solve the

challenge of deriving comprehensive analysis and actionable insights from complex global economic data. They can gain a proper understanding of economic trends, identify risks and opportunities, and make data-driven decisions to drive business growth and navigate the global economic landscape successfully. The solution's uniqueness lies in its holistic analysis, incorporating financial and non-financial factors like corruption. Integration of IBM Cognos and Python enables thorough analysis with customizable visualisations, bridging complex data to useful insights. Stakeholders benefit from enhanced strategies and investment opportunities.

The different phases of the solution include data collection and cleaning, data exploration, creation of dashboards, reports, and story, web development and web integration using Flask. Data is collected and preprocessed for data accuracy. IBM Cognos and Python are used to explore diverse economic indicators to uncover trends and patterns. Visualisations related to the problem statement are created. These insights and visualisations are converted into interactive web dashboards which help to communicate the insights effectively. HTML and CSS contribute to the frontend of the web platform while Flask is utilised for backend functionality. The web integration allows user-friendly web platforms to be developed. It allows for the seamless presentation of visualisations and findings, enhancing accessibility and interaction. When everything is put together, a holistic solution that effectively bridges the gap between complex economic data and well-informed decision-making and caters to the needs of various stakeholders, is obtained.

3 THEORETICAL ANALYSIS

3.1 BLOCK DIAGRAM

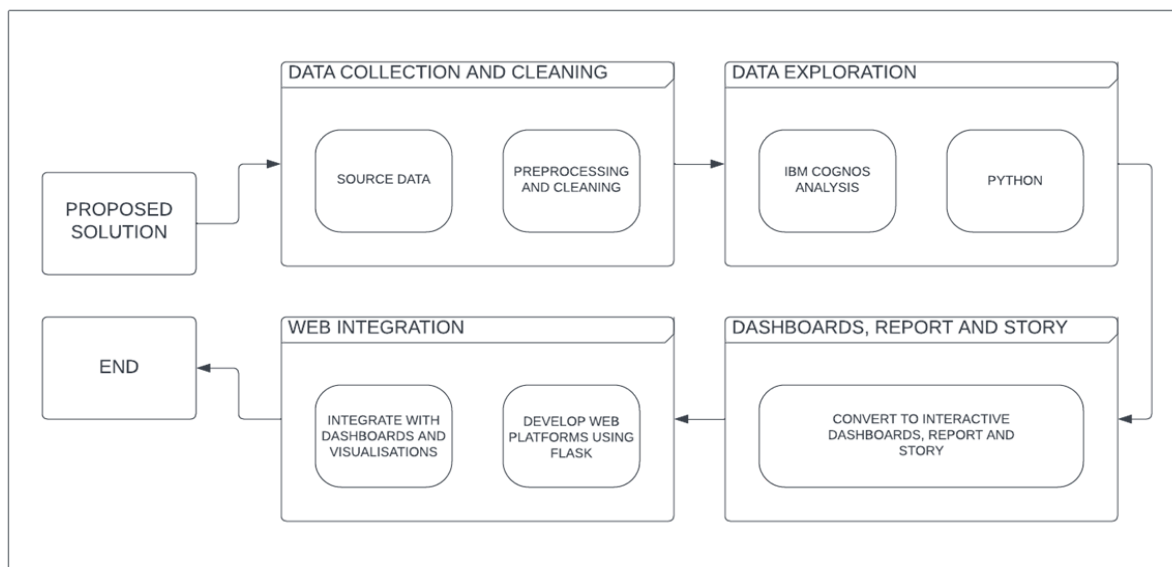


Fig 1. Block diagram showing overview of project

3.2 HARDWARE/SOFTWARE DESIGN

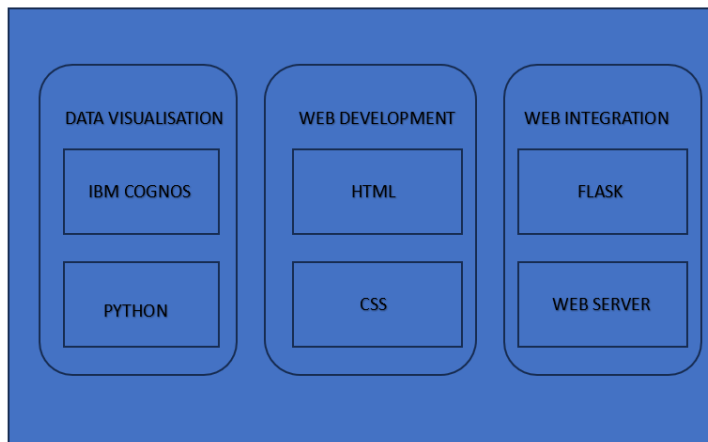


Fig 2. Software design of the project

The project is software-centred and utilises technologies such as IBM Cognos, Python, HTML, CSS, and Flask for its development. Data analysis, visualisation, and reporting tasks are carried out using IBM Cognos, which can generate insights and analytics from the created visualisations. Python, known for its versatility, is mainly employed to merge multiple CSV files. Flask, chosen as the web framework, enhances user-friendliness, accessibility, and allows seamless interaction, providing remote access to insights. The global economic data, serving as integral data sources, offer a comprehensive and accurate repository of information.

4 EXPERIMENTAL INVESTIGATIONS

There are 5 CSV files in the global economic data dataset — corruption.csv, cost_of_living.csv, richest_countries.csv, tourism.csv and unemployment.csv. Exploratory data analysis is done to gain useful insights and identify trends and patterns. Exploratory Data Analysis includes the following steps:

(i) Variable Identification

In corruption.csv, the variables are country, annual_income and corruption_index.

In cost_of_living.csv, the variables are country, cost_index, monthly_index and purchasing_power_index.

In richest_countries.csv, the variables are country and gdp_per_capita.

In tourism.csv, the variables are country, tourists_in_millions, receipts_in_billions, receipts_per_tourist and percentage_of_gdp.

In unemployment.csv, the variables are country and unemployment_rate.

There are 2 types of variables — predictor variables and target variables. There are 2 data types — character and numerical. There are 2 variable categories — categorical and continuous. country, present in all the csv files, is a categorical predictor variable of character data type.

In corruption.csv, annual income is a predictor variable and corruption_index is a target variable. annual_income and corruption_index are continuous and of numeric data type.

In cost_of_living.csv, predictor variables include monthly_income, purchasing_power and target variable is cost_index. cost_index, monthly_income, purchasing_power are continuous and are of numerical data type.

In richest_countries.csv, gdp_per_capita is a continuous target variable of numeric data type.

In tourism.csv, predictor variables include tourists_in_million, receipts_in_billion, receipts_per_tourist and target variable is percentage_of_gdp. tourists_in_million, receipts_in_billion, receipts_per_tourist and percentage_of_gdp are continuous and of numeric data type.

In unemployment.csv, unemployment_rate is a continuous target variable of numeric data type.

(ii) Univariate Analysis

Univariate analysis means analysis to understand characteristics of 1 variable. For continuous variables, central tendency (mean, median, mode) and spread (range, standard deviation) have to be assessed using histograms or box plots. For categorical variables, the focus is on frequency and proportions, using bar charts and pie charts.

The following visualisation is an example of a univariate analysis that has been done:

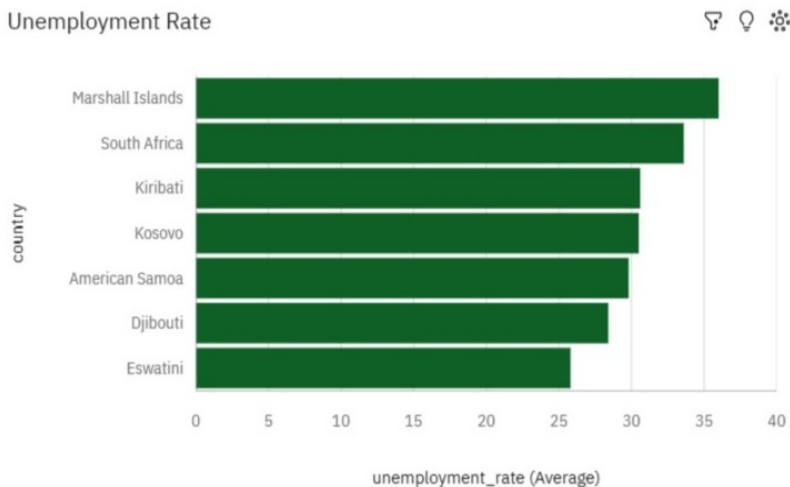


Fig 3.1 Bar graph of country vs unemployment rate

This graph shows the different unemployment rates of each country. Examples of insights that can be gained are that Marshall Islands has the highest unemployment rate and South Africa has a higher unemployment rate than Kosovo.

Other examples of univariate analysis include bar graphs of GDP per capita vs country, tourists in millions vs country and receipts per tourist vs country.

(iii) Bivariate Analysis

Bivariate analysis means understanding the relationship between 2 variables and how change in 1 variable affects another. If the relationship between 2 categorical variables are studied, two way tables and stacked column charts are used. For categorical and continuous, box plots can be used.

To explore the relationship between 2 continuous variables, scatter plots are used.

The following visualisation is an example of a bivariate analysis that has been done:

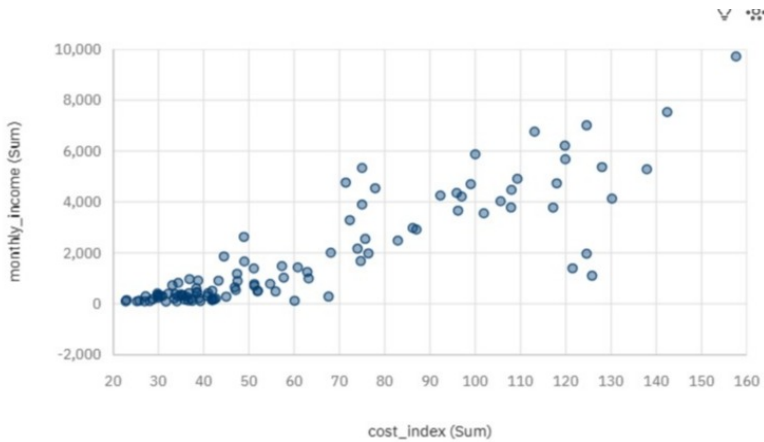


Fig 3.2 Scatter plot of monthly_income vs cost_index

A general trend that can be observed from this visualisation is that a country with a higher cost_index is likely to have a higher monthly_income. Another example of bivariate analysis is a scatter plot of receipts in billions vs tourists in millions.

Multivariate analysis typically involves analysing three or more variables simultaneously to understand complex relationships and patterns among them. Examples of multivariate analysis include a heatmap of tourists_in_millions, receipts_in_billions, receipts_per_tourist, percentage_of_gdp, receipts_in_billions vs country.

Data exploration involves various other steps, including missing value and outlier treatment, as well as variable transformation and creation. For missing values, options include removal of observations with minimal missing values or imputation using measures like mean, median (numerical), or mode (categorical). Outliers, which are unusual values, can be identified using scatter plots and box plots. Variable transformation adjusts relationships; options include logarithmic, square root, or Box-Cox transformations to normalise skewed distributions. Lastly, variable creation involves generating new information through transformations or calculations applied to existing variables.

In the project, Python code was used to merge more than 2 different csv files at a time and to avoid missing values that would cause data inaccuracies. To make visualisations, both IBM Cognos and Python could be used but IBM Cognos was mainly used as it reduces the need for more code. Scatter plots were created as well.

5 FLOWCHART

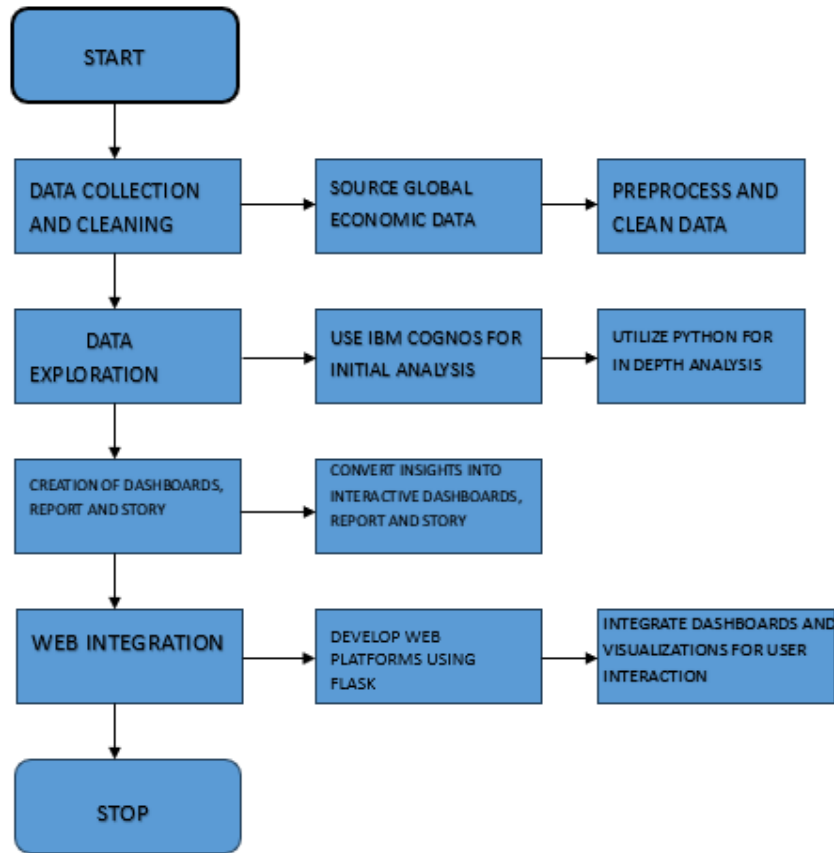


Fig 4. Control Flow of the Solution

6 RESULT

The results of the project are presented in the form of dashboards, a report and a story.

Note: The following are screenshots and thus, not interactive. In the website, hover over the different visualisations to see what each represents and use fullscreen mode for maximum functionality.

6.1 DASHBOARD

A dashboard is an interactive visual interface that presents real-time data in a concise manner. It offers a holistic view of key performance indicators for quick insights.

The dashboard has 3 tabs.

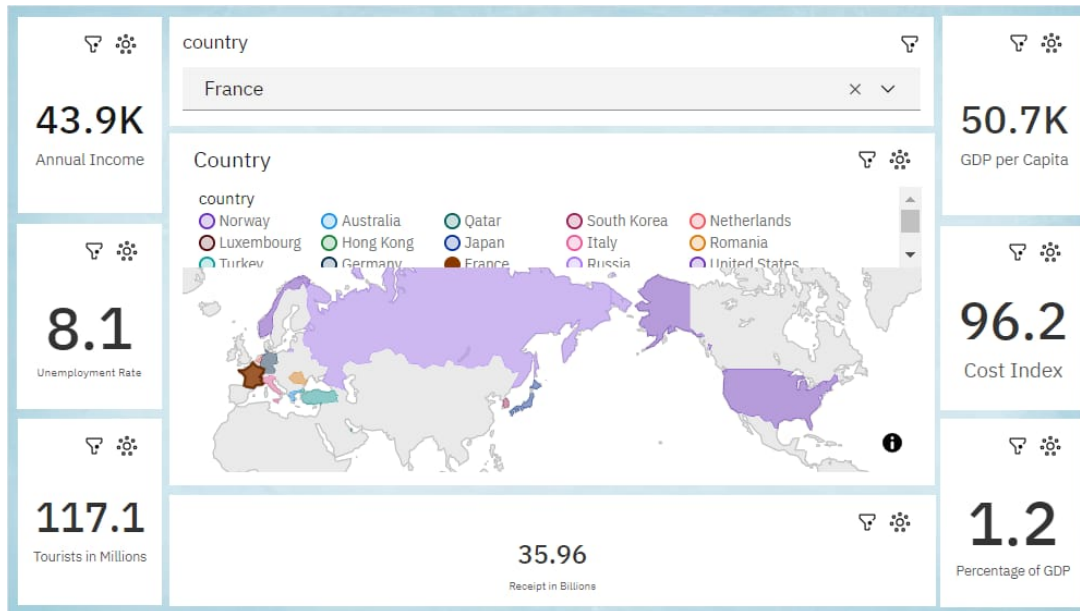


Fig 5.1. Tab 1 of the dashboard

It has a drop-down menu for users to select the country whose details they want to view. The annual income, unemployment rate, tourists in millions, GDP per capita, cost index and GDP per capita are present. Each selected country can also be viewed on the map.

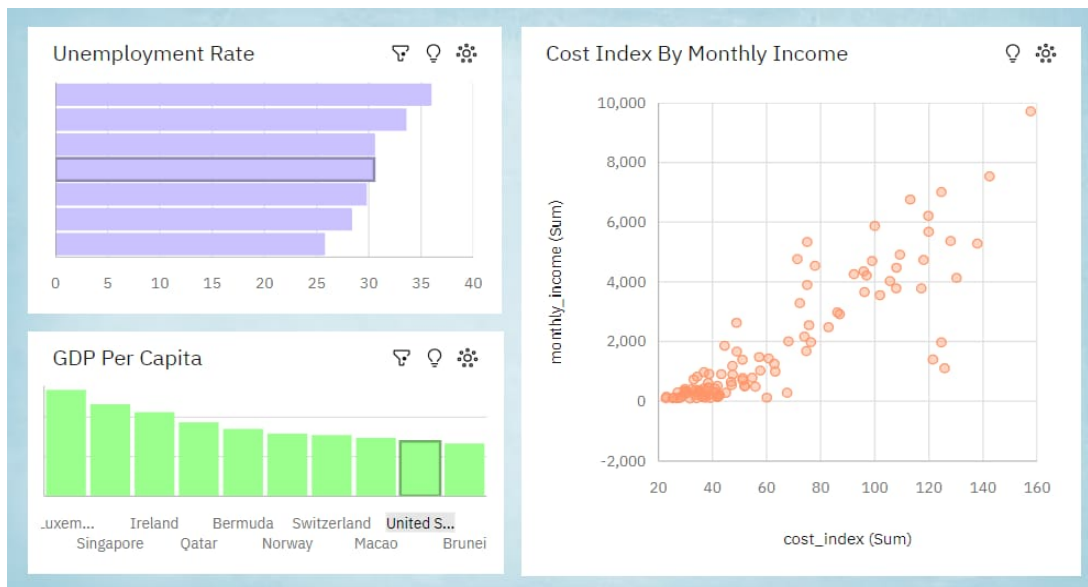


Fig 5.2 Tab 2 of the dashboard

It consists of 3 visualisations – bar graphs of unemployment rate of each country and GDP per capita of each country and a scatter plot of monthly_income vs cost_index which shows that countries with a higher monthly income are likely to have a higher cost index.

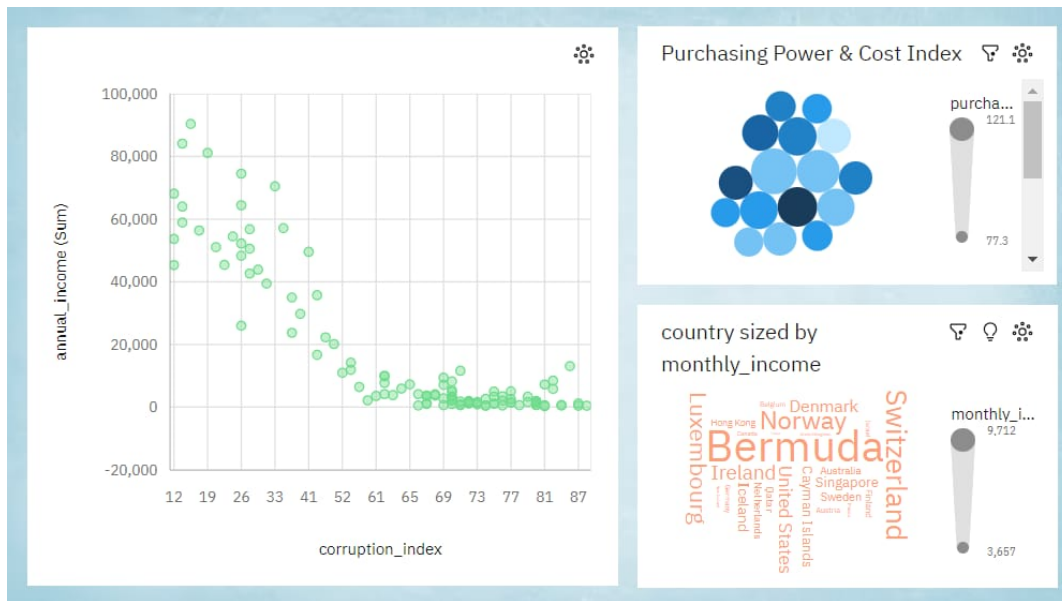


Fig 5.3 Tab 3 of the dashboard

It consists of 3 visualisations – a scatter plot of annual income vs corruption index, a packed bubble chart of purchasing power and cost index and a word cloud of countries where monthly income is represented by size. The scatter plot shows that a country with a higher corruption index is likely to have a lower annual income and the word cloud shows that Bermuda is the country with the highest monthly income. Lower cost index and higher purchasing power indicates that people can buy more goods for the same amount of money. The packed bubble chart helps to identify such countries.

6.2 REPORT

A report is a structured presentation of data, often static, offering in-depth analysis and insights to support decision-making.

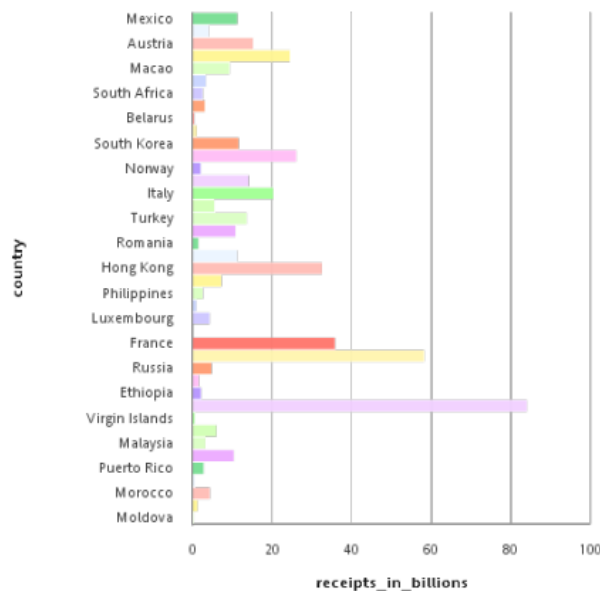


Fig 6.1. Bar graph of country vs receipts_in_billions

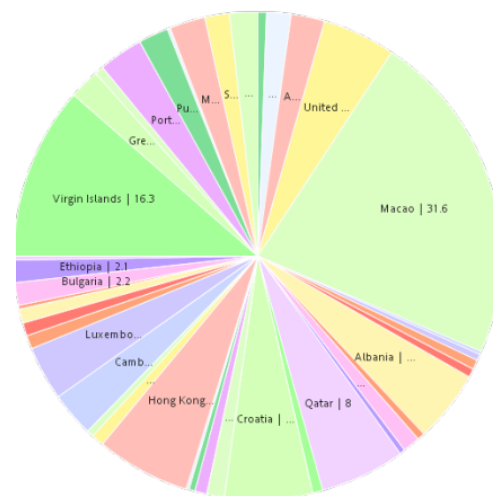


Fig 6.2 Pie chart of percentage_of_gdp of each country

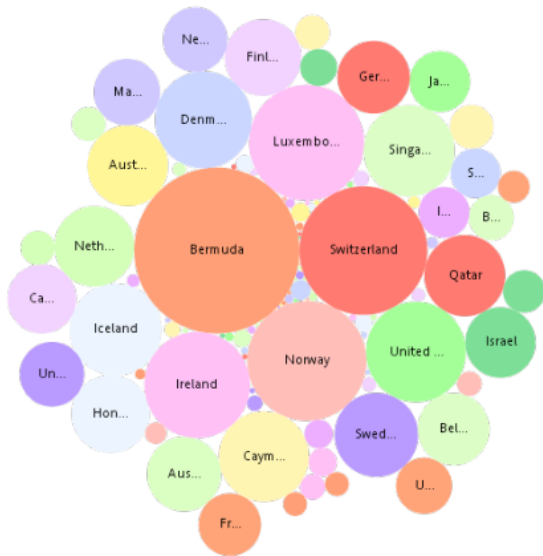


Fig 6.3 Packed bubble chart of monthly_income of different countries Fig 6.4 Word cloud of country sized by cost_index

From the report, the following insights are obtained:

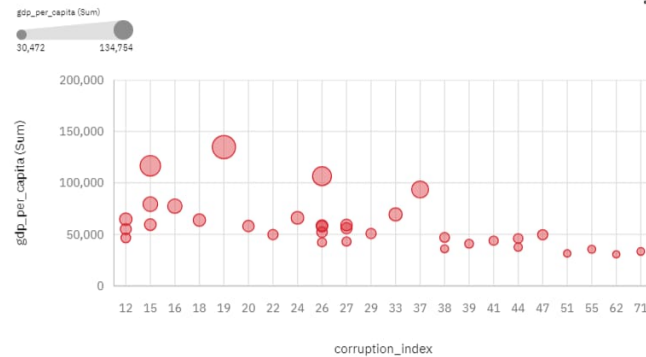
- (i) **Tourist Attraction:** Countries with a high number of tourists can be attractive markets for businesses, especially those in the hospitality, entertainment, and retail sectors. A significant tourist influx implies a potential customer base looking for various services and products.
- (ii) **Economic Impact:** Tourism is not only about attracting visitors but also about its economic contribution. Visualising tourism receipts as a percentage of GDP can provide insights into how much a country's economy relies on tourism. This information can help businesses understand the economic significance of the tourism sector in a given country.
- (iii) **Market Segment:** Tourists come from diverse backgrounds and have different preferences and spending capacities. It is important for businesses to analyse the types of tourists a country attracts. High-income tourists may seek luxury experiences, while budget travellers may want affordability. This understanding can guide businesses in customising their products or services to specific market segments.

6.3 STORY

A story guides users through insights and outcomes by sequentially presenting visualisations and explanations that focus on data-driven storytelling. The explanations have been given beside each visualisation.

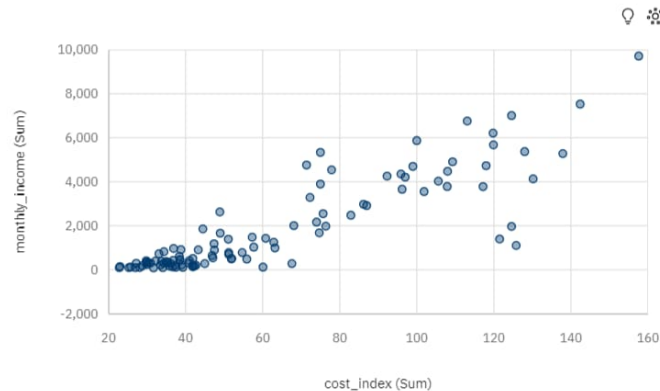
CORRUPTION AND ECONOMIC PERFORMANCE

- There's a negative correlation between corruption levels and GDP per capita. Countries with lower corruption tend to have higher GDP per capita.
- Businesses may find it advantageous to invest in countries with lower corruption levels as they often offer a more stable and favorable economic environment.



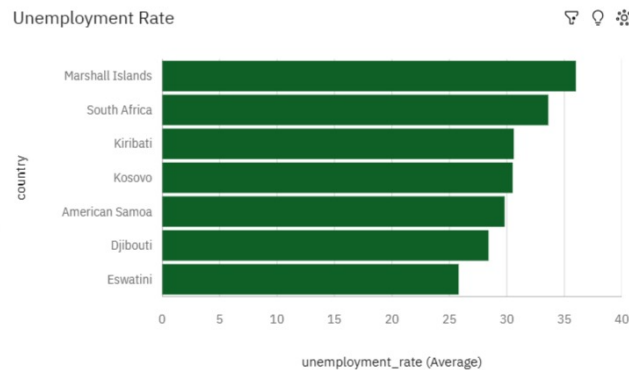
COST OF LIVING AND INCOME

- There's a positive correlation between the cost of living and annual income. As the average annual income increases, so does the cost of living in a country.
- Businesses targeting high-income consumers may need to consider the cost of living in their pricing and marketing strategies.



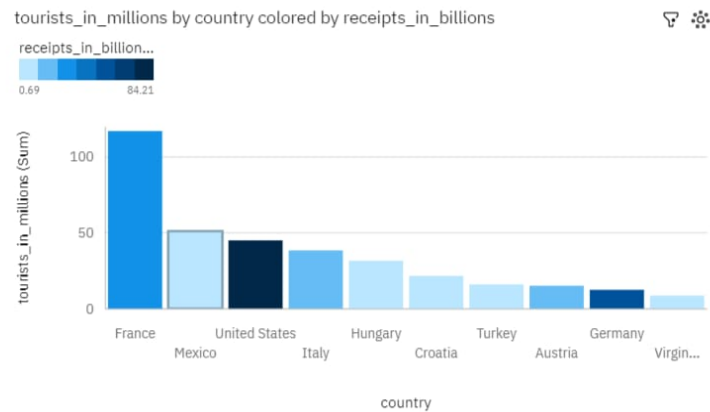
UNEMPLOYMENT RATE

- Marshall Islands, South Africa, and Kiribati have the highest unemployment rates, indicating challenges in these regions for recruiting skilled labor.
- Countries like Qatar, Luxembourg, and Ethiopia have lower unemployment rates, suggesting a potentially larger pool of available talent for businesses.



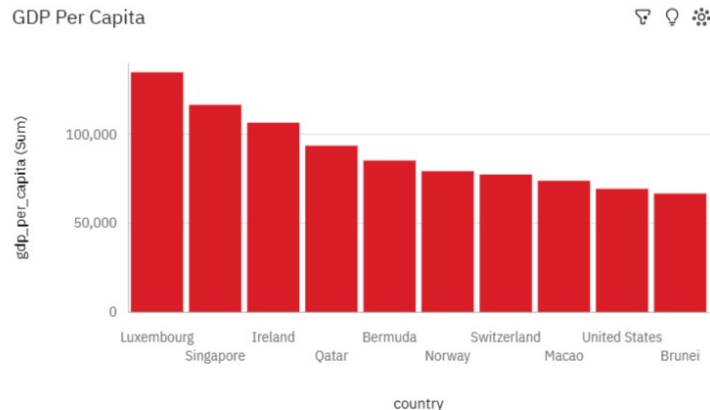
RICHEST COUNTRIES AND TOURISM

- France, Mexico, and the United States attract the highest numbers of tourists, making them lucrative markets for businesses in the tourism and hospitality sectors.
- Countries like Luxembourg and Qatar have a lower number of tourists but high receipts per tourist, suggesting a focus on high-end tourism.



RICHEST COUNTRIES AND TOURISM

- The data on the richest countries (GDP per capita) doesn't directly correlate with the number of tourists or tourism receipts. For example, some high-GDP countries like Luxembourg and Singapore have relatively fewer tourists.
- Businesses should not assume that high-GDP countries are automatically the most attractive markets for tourism-related ventures. Other factors, such as infrastructure and tourism policies, also play a crucial role.



7 ADVANTAGES & DISADVANTAGES

The proposed solution has many advantages such as the exceptional data analysis and visualisation capabilities offered by IBM Cognos. This enables businesses to create interactive dashboards and reports that help to comprehend global economic data. The integration of real-time data sources allows stakeholders to access current insights that enable them to make informed decisions. This helps organisations to optimise their operations, explore new markets, and maintain a competitive edge. This promotes better risk management, improved resource allocation, and sustainable growth strategies. The solution has positive impacts on not just businesses but on society as well. IBM Cognos provides comprehensive economic insights that contribute to shaping informed policy decisions. These promote societal prosperity, reduce inequalities, and elevate living standards. Through inclusive growth and

strategic resource allocation, it has the potential to foster societal cohesion and support equitable development.

However, IBM Cognos does have its disadvantages. To integrate the tool into existing systems and data sources would require intricate specialised expertise and potentially cause initial delays in implementation. The accuracy and completeness of input data are important for deriving meaningful insights, demanding stringent data quality assurance measures. The process of training stakeholders to utilise IBM Cognos effectively may require a lot of time. Considerations also extend to financial aspects, encompassing licensing fees, infrastructure setup costs, and ongoing maintenance expenses. While IBM Cognos provides remarkable capabilities, it is important to recognize that its capacity to predict economic shifts accurately does have limitations and harnessing its potential for interpreting economic insights effectively may demand a degree of domain expertise. Overall, even with some limitations of its own, IBM Cognos is a very useful tool that can bring myriads of benefits to many organisations.

8 APPLICATIONS

The proposed solution offers a wide spectrum of practical applications. In business planning, it helps organisations grow and use resources better. For investments and finance, it supports smart investment choices and managing risks. It also helps governments make good policies and put them into action. In fields like international trade, it helps businesses with supply chain management and finding new markets. Healthcare planning gets better with it, making sure resources are used right and patient care improves. The same goes for supply chains and logistics, where it reduces problems and makes distribution smoother. It is also useful for sustainability and development projects, aiding in planning for resilient communities and managing resources well. This solution is useful in various fields like tourism, research, real estate, and construction, which goes onto show how flexible and far-reaching its impact is.

9 CONCLUSION

IBM Cognos is a transformative tool that converts global economic data into actionable insights, enabling organisations to make informed decisions. This capability fosters a deep comprehension of economic trends, enabling proactive identification of potential risks and opportunities. Consequently, businesses are better equipped to position themselves advantageously while mitigating potential risks. Empowering stakeholders with up-to-date data, IBM Cognos facilitates informed decisions, replacing speculation with data-driven choices. This leads to business growth, operational efficiency, and resilience in the face of economic uncertainties and helps organisations align their strategies and operations with the ever-changing global economic landscape.

The project provides insights through an interactive dashboard which provides a real-time overview of key economic indicators, a comprehensive report, and a story. The dashboard includes a dropdown menu for choosing specific countries and provides information such as annual income, unemployment rates, tourists in millions, GDP per capita, cost index, and GDP per capita. It also shows the selected country on a world map. The report uncovers useful insights like the correlation between corruption and GDP per capita, the positive relationship between cost of living and income, and the dynamics of the richest countries in terms of tourism. The report's insights extend to important aspects like understanding tourist attraction and its economic impact and exploring the segmentation of the tourist market based on

preferences and spending patterns. The story enables users to understand outcomes and insights effectively. The findings include the negative correlation between corruption and GDP per capita which implies that it is better to invest in a country with lower corruption levels. Similarly, the positive correlation between the cost of living and annual income offers important insights for businesses targeting high-income consumers. Other insights include the relationship between the richest countries and tourism dynamics, showing that GDP per capita doesn't always correlate directly with success in tourism. This insight encourages businesses to consider various factors beyond GDP when assessing market appeal. Overall, the project combines dynamic visuals, in-depth analysis, and storytelling to produce actionable insights. These insights empower stakeholders with a comprehensive understanding of economic trends, enabling better-informed decisions across diverse sectors.

10 FUTURE SCOPE

Future enhancements to our solution could include the use of AI and advanced predictive analytics. As IBM Cognos advances, adding more advanced AI capabilities can help us not only understand current market trends but also predict future changes. This predictive power helps in decision-making, enabling organisations to adapt to changing economic landscapes and take advantage of new opportunities. Real-time data integration and interactive visualisation can be used to simplify economic data analysis. This makes it easier for decision-makers to analyse trends, correlations, and outliers. This leads to better strategies, informed investment choices, and agile responses to market dynamics.

11 BIBLIOGRAPHIES

1. Kaggle. (n.d.). World Economic Data Complete Data Analysis. Retrieved from <https://www.kaggle.com/code/saisandeepjallepalli/world-economic-data-complete-data-analysis/input>
2. Bloomberg. (n.d.). Bloomberg Terminal - Professional Financial Data and Analysis Solution. Retrieved from <https://www.bloomberg.com/professional/solution/bloomberg-terminal>
3. European University Institute (EUI) Data Portal. (n.d.). DataStream. Retrieved from <https://www.eui.eu/Research/Library/ResearchGuides/Economics/Statistics/DataPortal/datastream>

APPENDIX

A. Source Code

1. For merging CSV files

```

import pandas as pd
import os

directory_path = 'C:/Users/Adil'
csv_files = [f for f in os.listdir(directory_path) if f.endswith('.csv')]

common_column = "country"
merged_data = pd.DataFrame()

for file in csv_files:
    file_path = os.path.join(directory_path, file)
    data = pd.read_csv(file_path)

    if merged_data.empty:
        merged_data = data
    else:
        merged_data = pd.merge(merged_data, data, on=common_column, how='inner')

merged_data.to_csv('C:/Users/Adil/merged_data.csv', index=False)

```

A Python library called Pandas is used to merge multiple CSV files containing data related to different countries. It begins by importing the necessary `pandas` and `os` modules and then looks through the CSV files in the directory and reads the data from each file using `pd.read_csv()`. The common column for merging is "country". If the `merged_data` DataFrame is empty, the first set of data is added to it. Similarly, the code merges the subsequent datasets with the existing `merged_data` using an inner join. The merged data is then saved as 'merged_data.csv' in the same directory. This process effectively combines data from multiple CSV files into a single unified dataset based on the shared "country" column.

2. For web development using HTML and CSS (Refer to Github)
3. For web integration using Flask

```

app.py > ...
1  from flask import Flask, render_template, url_for
2  app = Flask(__name__)
3
4
5  @app.route('/')
6  def index():
7      return render_template('index.html')
8  if __name__ == '__main__':
9      app.run()

```

A Python web framework called flask is used to set up a basic web application. The app instance is created using Flask, and a route for the root URL ("/") is defined. When a user accesses the root URL, the `index()` function is triggered. Inside this function, an HTML template named 'index.html' is rendered and displayed in the browser. The script checks if it is run directly. If so, it starts the development server using `app.run()` and a simple Flask application is created.