# Gender Equality Impact on Economic Growth

Submitted in partial fulfilment of the requirements for the award of the degree of

**Bachelor of Computer Applications**


**Submitted to:**

Dr. Cosmena Mahapatra

Assistant Professor

VIPS-TC

**Submitted by:**

Rupashi Maurya

01717702022

BCA 5A


**(Batch 2022-25)**


Vivekananda Institute of Professional Studies

(Affiliated to Guru Gobind Singh Indraprastha University)

# SELF CERTIFICATE

This is to certify that the dissertation/project report entitled "Gender Equality Impact on Economic Growth" is done by me is an authentic work carried out for the partial fulfilment of the requirements for the award of the degree of Bachelor of Computer Applications under the guidance of Dr. Cosmena Mahapatra. The matter embodied in this project work has not been submitted earlier for award of any degree or diploma to the best of my knowledge and belief.

Signature of the student

| | |
|---|---|
| Ishita Sharma | Rupashi Maurya |
| 12917702022 | 01717702022 |

# ACKNOWLEDGEMENTS

We would like to express our heartfelt gratitude to everyone who contributed to the successful completion of project titled “Gender Equality Impact on Economic Growth”. First and foremost, we extend our deepest thanks to our esteemed faculty and the institution for providing us with the opportunity and support necessary for this project.

A special note of appreciation goes to our mentor and guide, **Dr. Cosmena Mahapatra and Dr. Meenu Chopra** faculty VSIT, CSD, VIPS-TC for their invaluable guidance, encouragement, and insights throughout this journey. Her expertise and dedication were instrumental in shaping our work and ensuring its successful implementation.

Lastly, we are grateful to our peers and everyone who supported us during the development of this project. Your encouragement and feedback have been a constant source of motivation.

Project Team Members- Rupashi Maurya and Ishita Sharma

# CERTIFICATE

This is to certify that this project entitled “Gender Equality Impact on Economic Growth” submitted in partial fulfillment of the degree of Bachelor of Computer Applications to the Dr. Cosmena Mahapatra done by Ms. Rupashi Maurya, Roll-No. 01717702022 is an is an authentic work carried out by him/her at under my guidance. The matter embodied in this project work has not been submitted earlier for award of any degree to the best of my knowledge and belief.

Signature of the student

Signature of the Guide

# TABLE OF CONTENTS

**Sign**

# Project Synopsis

**Title of the Project**: Impact of Gender Inequality on Economic Growth

**Problem Statement**: Gender inequality continues to be a significant barrier to sustainable economic growth worldwide. Disparities in education, labor force participation, health outcomes, and political representation limit the full economic participation of women and hinder comprehensive development. This project seeks to investigate how gender inequality affects economic performance, using India as a case study.

**Why This Topic Was Chosen**: The persistent impact of gender inequality on economic progress makes this topic critical for analysis. By understanding these dynamics, policymakers and stakeholders can be better equipped to create strategies that promote gender equality, which, in turn, supports economic growth. The choice of this topic also aligns with the pursuit of Sustainable Development Goals (SDGs), particularly SDG 5 (Gender Equality) and SDG 8 (Decent Work and Economic Growth).

**Objective and Scope of the Project**: The primary objective of this project is to analyze socio-economic data to understand the influence of gender inequality on economic growth and to propose data-driven strategies for promoting gender equality. The scope includes:

- Collecting and analyzing data related to gender disparities and economic indicators.
- Identifying key areas where gender inequality most impacts economic outcomes.
- Understanding trends over time and across regions.
- Developing predictive models to project potential economic benefits of reducing gender inequality.
- Proposing actionable policy recommendations to promote inclusive growth.

**Methodology**: This project involves multiple phases:

1. **Data Collection**: Socio-economic data were sourced from reliable platforms such as Kaggle, including datasets on the Gender Inequality Index (GII) and Gross National Income (GNI) per capita.
2. **Data Preprocessing**: The data were inspected for null values and cleaned to ensure consistency. Rows with missing values were removed, and relevant features were selected to create a refined dataset.

3. **Descriptive Analysis**: Preliminary data analysis was conducted using statistical functions to identify mean values, distributions, and key trends.
4. **Correlation Analysis**: A heatmap was used to examine the correlation between GII and various economic indicators, including GNI per capita and maternal mortality.
5. **Machine Learning Models**: A Linear Regression model was employed to assess the relationship between GII and economic outcomes. The model's performance was evaluated using $R^2$ scores and Mean Squared Error (MSE).
6. **Data Visualization**: Interactive visuals, including scatter plots, bar graphs, and geographical maps, were created using data visualization tools to enhance interpretability.

**Hardware & Software Used**:

- **Hardware**: Standard computing equipment (PC/laptop)
- **Software**: Power BI for visualization, Python (pandas, scikit-learn, matplotlib, seaborn) for data analysis and modeling.

**Testing Technologies Used**: Model validation was conducted using training and testing data splits, with performance metrics such as $R^2$ and MSE providing insights into model accuracy.

**Contribution of the Project**: The project contributes to a deeper understanding of the economic impact of gender inequality. It underscores the need for targeted interventions to reduce gender disparities and provides actionable insights that can inform policy-making. By highlighting correlations and potential predictive outcomes, the project supports efforts to achieve SDG 5 and SDG 8, advocating for equitable and sustainable economic growth.

# 1. Objective and Scope of the Project

**1.1 Objective:** The primary objective of this project is to analyze socio-economic data to understand the impact of gender inequality on economic growth and to propose data-driven strategies to promote gender equality and enhance economic development. The specific objectives are:

- To collect and analyze socio-economic data from reliable sources related to gender inequality and economic growth.
- To identify key areas where gender disparities are most pronounced and how they affect economic performance.
- To understand the temporal and spatial trends of gender inequality in relation to economic indicators.
- To develop predictive models that illustrates the potential economic outcomes of reducing gender inequality.
- To propose actionable strategies and policy recommendations to address gender disparities and promote inclusive economic growth.
- To assess the potential impact of these strategies on achieving SDG 5, SDG 8 and fostering sustainable economic development.

**1.2 Scope**: The scope of the project includes:

- Collecting, preprocessing, and analyzing socio-economic data from trusted sources.
- Developing predictive models to forecast potential economic outcomes with improved gender equality.
- Creating comprehensive visualizations and dashboards using Power BI and Python to display key findings.
- Highlighting policy recommendations that align with sustainable development goals (SDG 5 and SDG 8).

# 2. Theoretical Background and Definition of Problem

## 2.1 Theoretical Background

The theoretical background for this project involves understanding the interplay between gender inequality and economic growth, grounded in socio-economic theories, empirical studies, and global frameworks like the **Sustainable Development Goals (SDGs)**.

1. **Gender Inequality and Economic Growth**:

   - Gender inequality refers to disparities between men and women in access to resources, opportunities, and rights. This inequality manifests in education, healthcare, labor force participation, political representation, and income.
   - Economic growth, often measured by GDP or GNI per capita, is directly influenced by human capital, productivity, and resource allocation. When gender disparities limit women's contributions, economies fail to realize their full potential.

2. **Relevant Frameworks**:

   - **Human Capital Theory**: Educated and healthy populations contribute more effectively to economic output. Gender gaps in education and health reduce the overall human capital available for growth.
   - **Gender and Development Approach**: Emphasizes that empowering women leads to more equitable and sustainable development outcomes. Gender inequality is not just a social issue but also an economic barrier.
   - **Intersectionality Theory**: Highlights how overlapping social categories (e.g., caste, ethnicity, and gender) exacerbate disparities, which is particularly relevant in diverse regions like India.

3. **Empirical Evidence**:

   - Studies show that reducing gender inequality can lead to higher productivity, increased labor force participation, and greater innovation.

- Countries with higher gender equality indices often exhibit higher GDP per capita and better social outcomes.

4. **Global Context and SDGs**:

   - **SDG 5** focuses on achieving gender equality and empowering all women and girls.
   - **SDG 8** emphasizes promoting inclusive and sustainable economic growth, productive employment, and decent work for all.
   - The interconnected nature of these goals highlights the importance of addressing gender inequality to drive economic progress.

This project incorporates several critical technologies and concepts that are essential for effective data analysis and decision-making. Below are the theoretical foundations of the key technologies used:

**Data Analysis**: Data analysis refers to the process of inspecting, cleansing, transforming, and modeling data with the goal of discovering useful information, informing conclusions, and supporting decision-making. For this project, descriptive and inferential statistical techniques were used to identify trends, correlations, and outliers. Tools like Python libraries (pandas, NumPy) facilitated data manipulation, while visualization libraries (matplotlib, seaborn) helped represent data in graphical formats to enhance understanding.

**Concepts of Data Preprocessing**: Preprocessing involves preparing raw data for analysis by handling missing values, standardizing formats, and removing outliers. Techniques such as normalization and data transformation ensure the data is in a suitable format for modeling and visualization.

**Modeling**: Predictive modeling is an essential aspect of data analysis that uses statistical algorithms to forecast outcomes based on historical data. In this project, Linear Regression and K-Nearest Neighbors (KNN) models were implemented using scikit-learn. Linear Regression was chosen for its simplicity in modeling relationships between dependent and independent variables, while KNN provided an alternative approach to regression by considering the k-nearest data points to predict outcomes.

**Dashboards**: Dashboards are interactive panels that display visual data representations to make analysis more accessible and actionable. Power BI was used to create interactive dashboards in this project, allowing users to explore trends, compare indicators, and view geographic distributions of gender inequality and economic growth. Dashboards support dynamic data exploration and provide stakeholders with an intuitive way to interpret complex data relationships.

**Evaluation Metrics**: To validate model performance, metrics such as $R^2$ scores and Mean Squared Error (MSE) were utilized. These metrics provided insights into the predictive power and accuracy of the implemented models.

Understanding these technologies and concepts was pivotal in achieving reliable analysis and presenting clear, actionable insights on the impact of gender inequality on economic growth.

## 2.2 Definition of the Problem

Gender inequality remains a persistent and systemic issue that significantly hinders economic development.

1. **Key Aspects of the Problem**:

    - **Education**: Unequal access to education for girls and women reduces their skill levels and earning potential.
    - **Labor Force Participation**: Women often face barriers to entering the workforce, unequal pay, and limited opportunities for career advancement.
    - **Health Disparities**: Maternal mortality rates, access to healthcare, and life expectancy discrepancies impact women's ability to contribute economically.
    - **Political Representation**: Limited involvement of women in decision-making processes results in policies that inadequately address their needs and contributions.
    - **Economic Indicators**: Countries with high levels of gender inequality often show lower GNI per capita, slower growth, and inefficient resource utilization.

2. **Specific Focus in India**:

   - In India, socio-cultural norms, discriminatory practices, and systemic barriers exacerbate gender disparities in key areas like education, employment, and healthcare.
   - Despite progress in some indicators, structural inequalities continue to limit women's economic participation and contribution to national growth.

3. **Research Gaps**:

   - While many studies highlight the existence of gender inequality, there is a need for data-driven insights into its quantifiable impact on economic growth, especially in diverse and populous countries like India.
   - Few studies provide actionable, predictive models that can demonstrate the economic benefits of reducing gender inequality.

# 3. System Analysis and Design vis-à-vis User Requirements

## 3.1 System Analysis

### 3.1.1 Understanding User Requirements

The project aims to address the needs of multiple stakeholders, including policymakers, researchers, and social activists. The key user requirements identified are:

- **Data-Driven Insights**:
  - Users need an analysis of how gender disparities affect economic growth.
  - Identify specific areas (e.g., education, health, workforce participation) where gender inequality is most pronounced.
- **Visual Interpretations**:
  - Users require clear and interactive visualizations of data trends and geographic disparities for better interpretability.
- **Predictive Insights**:
  - Policymakers and planners need models that predict economic benefits from reducing gender inequality.
- **Actionable Recommendations**:
  - Users expect actionable strategies and policy suggestions derived from the analysis.

### 3.1.2 Functional Requirements

The system should:

- Import and preprocess socio-economic data for gender-related and economic indicators.
- Perform statistical and correlation analysis to identify key trends and relationships.
- Develop predictive models to estimate the impact of improved gender equality on economic growth.
- Generate visualizations and dashboards to present findings in an accessible format.
- Support exporting reports and findings for external use.

### 3.1.3 Non-Functional Requirements

The system should:

- Be **user-friendly** and accessible to users with minimal technical expertise.
- Ensure **data integrity** and consistency during preprocessing and analysis.
- Provide **high performance** for processing large datasets efficiently.
- Be **scalable** for analyzing additional regions or broader datasets.
- Maintain **security** for sensitive data, especially when handling datasets with personal or sensitive information.

## 3.2 System Design

### 3.2.1 Architecture Overview

The system follows a modular architecture, consisting of the following components:

- **Data Input Module**:
  - Allows users to upload CSV files or connect to online datasets (e.g., Kaggle APIs).
  - Includes preprocessing functionalities to clean, filter, and structure data.
- **Analysis Module**:
  - Performs descriptive and correlation analysis to evaluate the relationship between gender inequality indices and economic indicators.
  - Implements statistical functions to calculate measures like mean, median, and correlation coefficients.
- **Prediction Module**:
  - Utilizes a Linear Regression model to predict economic outcomes (e.g., GNI per capita) based on gender inequality indices.
  - Provides $R^2$ and MSE metrics for model evaluation.
- **Visualization Module**:
  - Generates interactive visuals like scatter plots, bar charts, and geographical maps using Python libraries and Power BI.
  - Allows users to filter and explore data trends by regions or time periods.
- **Report Generation Module**:
  - Provides summarized findings in PDF or Excel format, including graphs, tables, and policy recommendations.

### 3.2.2 Technology Stack

- **Frontend**: Power BI (for visualizations)
- **Backend**: Python (pandas, scikit-learn, matplotlib, seaborn)
- **Data Storage**: Local storage for datasets and intermediate results
- **Development Tools**: Jupyter Notebook, Power BI Desktop

### 3.2.3 Data Flow

1. Users input data into the system.
2. The system cleans and preprocesses the data.
3. Analytical functions identify trends and correlations.
4. Predictive models estimate potential economic outcomes.
5. Visualizations and reports are generated for decision-making.

### 3.2.4 User Interface Design

- **Dashboard**: Displays interactive charts, heatmaps, and tables summarizing the analysis.
- **Data Upload Page**: Simple interface for uploading datasets with preprocessing options.
- **Prediction Tool**: Allows users to input hypothetical gender equality improvements to see potential economic gains.

# 4. System Planning (PERT Chart)

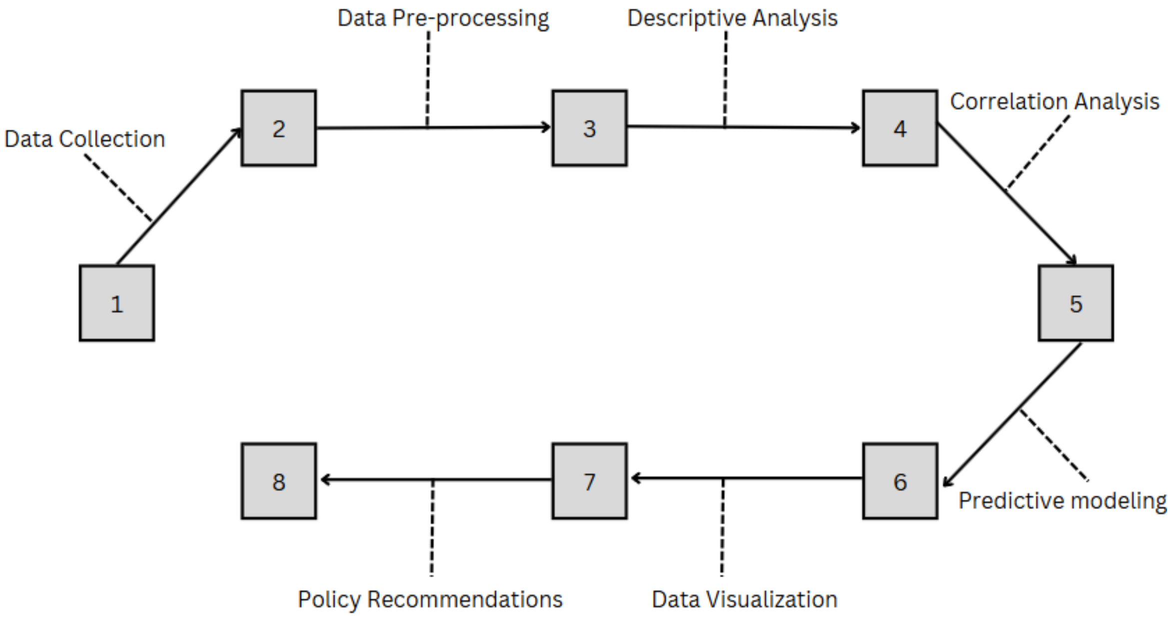



*Figure 1*

**Data Collection**:

- Gathered data relevant to the project..

**Data Pre-processing**:

- Cleaning the collected data by removing missing or erroneous values.
- Transform or standardize data to ensure compatibility between datasets.
- Establish relationships between datasets.

**Descriptive Analysis**:

- Summarized the data using measures such as mean, median, and standard deviation.
- Provide a high-level understanding of the data's key features.

**Correlation Analysis**:

- Identify relationships between variables .
- Use statistical measures like Pearson or Spearman correlation coefficients.

**Predictive Modeling**:

- Develop models to predict outcomes.
- Apply machine learning algorithms or statistical techniques.

**Data Visualization**:

- Create charts and graphs to present insights (e.g., maps, scatter plots, bar charts).
- Make results easier to interpret for stakeholders.

**Policy Recommendations**:

- Based on insights, suggest actionable strategies to reduce inequality and enhance economic growth.
- Address key findings such as improving education or workforce participation.

**System Outcome**:

- Deliver final recommendations and visualization tools (e.g., interactive dashboards).
- Present conclusions to stakeholders.

# 5. Methodology adopted, System Implementation & Details of Hardware & Software used System Maintenance &Evaluation

## 5.1 Methodology Adopted

### 1. Data Collection

- **Sources**: Datasets sourced from Kaggle, including Gender Inequality Index (GII), Gross National Income (GNI), and other socio-economic indicators.
- **Criteria**: Focused on datasets relevant to gender disparities (education, health, labor force participation) and economic growth.

### 2. Data Preprocessing

- **Inspection**: Checked for null values and inconsistencies.
- **Cleaning**: Removed incomplete records, normalized data ranges, and ensured uniform data formats.
- **Feature Selection**: Retained relevant features such as GII, maternal mortality, and GNI.

### 3. Descriptive Analysis

- Used statistical measures (mean, median, variance) to summarize key indicators.
- Identified regional trends and variations over time.

### 4. Correlation Analysis

- Employed heatmaps to evaluate relationships between GII and economic indicators like GNI and maternal mortality.
- Highlighted strong correlations to guide predictive modeling.

### 5. Predictive Modeling

- **Model Used**: Linear Regression.

- ❖ **Steps**:
  - ➢ Split the data into training (80%) and testing (20%) sets.
  - ➢ Trained the model to predict economic outcomes (e.g., GNI per capita) using GII as the primary feature.
- ❖ **Evaluation Metrics**:
  - ➢ $R^2$ Score: Assessed model fit.
  - ➢ Mean Squared Error (MSE): Measured prediction accuracy.

**6. Data Visualization**

- ❖ **Tools**: Power BI, Matplotlib, Seaborn.
- ❖ **Outputs**:
  - ➢ Scatter plots for GII vs. GNI.
  - ➢ Heatmaps for correlation analysis.
  - ➢ Geographical maps showing regional disparities.

**7. Policy Recommendations**

- ❖ Formulated actionable strategies based on findings to reduce gender disparities and enhance economic growth.

## 5.2 System Implementation

The project was implemented in modular phases:

1. **Data Ingestion**:

   - ➢ CSV files were read into Python using Pandas.
   - ➢ Automated scripts were created for data cleaning.

2. **Analysis and Modeling**:

   - ➢ Python libraries such as Pandas and Scikit-learn were used for analysis and modeling.

- Regression results were validated using test data.

3. **Visualization**:

   - Power BI was employed to create dashboards, while Matplotlib and Seaborn were used for static visualizations in Python.

4. **Output Delivery**:

   - Findings and recommendations were compiled into reports and presentations.

## 5.3 Details of Hardware & Software Used

### 5.3.1 Hardware Used

- **Processor**: Intel Core i5 or higher.
- **RAM**: 8 GB minimum (16 GB recommended for faster processing).
- **Storage**: At least 256 GB SSD for storing datasets and generated outputs.

### 5.3.2 Software Used

- **Operating System**: Windows 10/11.
- **Development Environment**: Jupyter Notebook for Python scripting.
- **Libraries**:
  - Pandas, Numpy (Data manipulation).
  - Scikit-learn (Modeling).
  - Matplotlib, Seaborn (Visualization).
- **Visualization Tool**: Power BI Desktop.

## 5.4 System Maintenance

1. **Regular Updates**:

   - Periodically refresh datasets to include the latest data for more accurate analysis.
   - Update libraries and tools to the latest versions for compatibility.

2. **Error Handling**:

   - Incorporate error-handling scripts to manage missing or corrupt data entries.

3. **Documentation**:
   - Maintain detailed documentation of data sources, preprocessing steps, and analysis workflows.
4. **Scalability**:
   - Ensure that the system can handle additional features or larger datasets without performance degradation.

## 5.5 System Evaluation

1. **Performance Metrics**:
   - Evaluate the model's accuracy using updated data periodically (e.g., quarterly).
   - Monitor $R^2$ scores and MSE to ensure the model remains reliable.
2. **User Feedback**:
   - Collect feedback from stakeholders to improve usability and refine policy recommendations.
3. **Impact Assessment**:
   - Assess the real-world application of recommendations on reducing gender disparities.
   - Re-evaluate strategies based on measurable socio-economic improvements.
4. **Visualization Usability**:
   - Ensure dashboards remain intuitive, with clear and actionable insights.
   - Update visualization templates based on user feedback.

# 6. Detailed Life Cycle of the Project

## 6.1 DFD Level-0

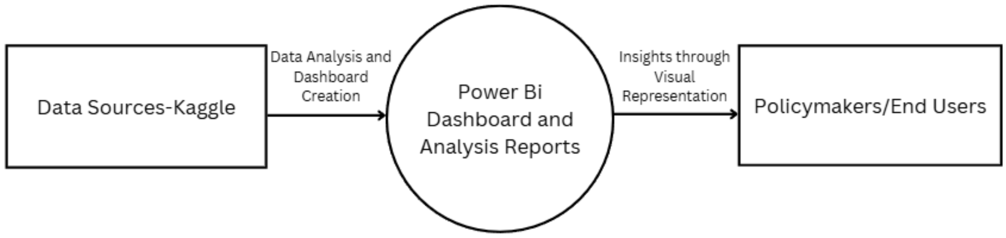



*Figure 2*

- Input: Data is sourced from platforms like Kaggle.
- Process: Data is analyzed, and a Power BI dashboard is created to generate analytical reports.
- Output: Insights are visually represented and shared with policymakers or end users for decision-making.

## 6.2 DFD Level-1

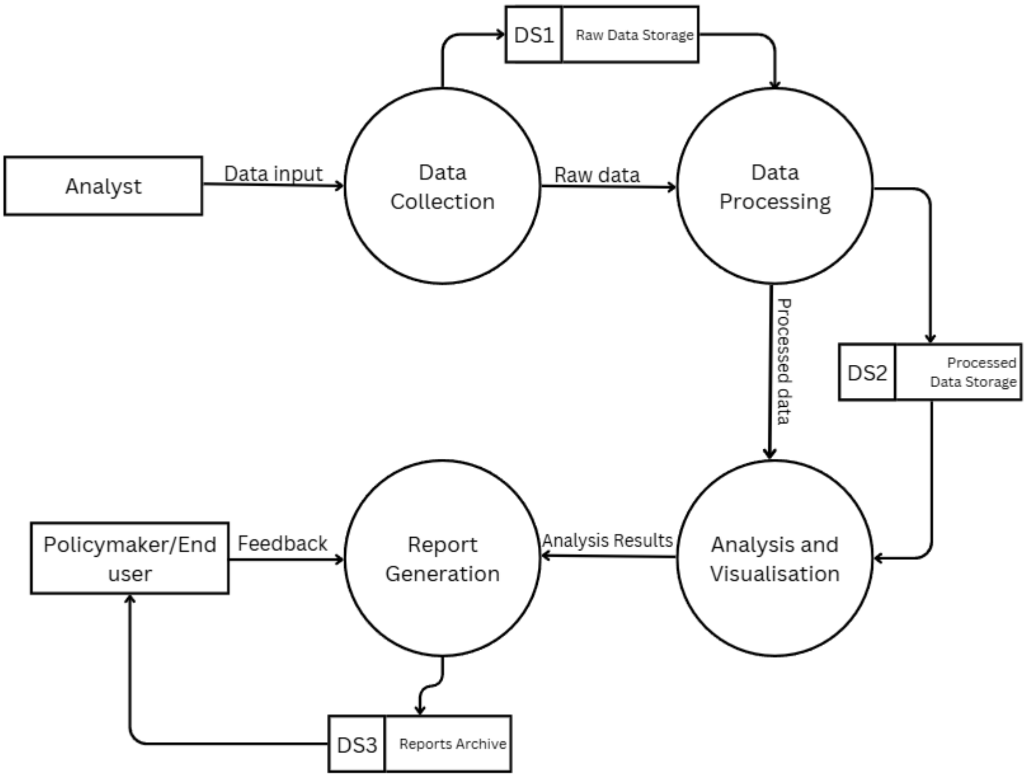



*Figure 3*

- Data Collection: Analysts provide data input, stored as raw data (DS1).
- Data Processing: Raw data is cleaned and transformed into processed data (DS2).
- Analysis and Visualization: Processed data is analyzed and visualized in dashboards/reports.
- Report Generation: Reports are archived (DS3) and shared with policymakers or end users, who provide feedback for refinement.

# Coding and Screenshots of Project

## Code:

## Working with data

```
In [ ]: import pandas as pd
        import numpy as np

        import matplotlib.pyplot as plt
        import seaborn as sns

        from sklearn.model_selection import train_test_split
        from sklearn.preprocessing import StandardScaler

        from sklearn.linear_model import LinearRegression
        from sklearn.neighbors import KNeighborsRegressor

        from sklearn.metrics import mean_squared_error, r2_score, mean_absolute_error
```

```
In [ ]: # Loading the dataset
        df = pd.read_csv(r'/kaggle/input/gender-inequality-index/Gender_Inequality_Index.cs
```

```
In [ ]: # Inspecting the dataset
        print(df.info())
```

```
In [ ]: print(df.describe())
```

```
In [ ]: # Handling missing values
        df = df.dropna()
```

```
In [ ]: print(df.info())
```

```
In [ ]: df.head()
```

## Relation betwen various dataset features

```
In [ ]: numerical_columns = ['GII', 'Maternal_mortality', 'Adolescent_birth_rate',
                             'Seats_parliament', 'F_secondary_educ',
                             'M_secondary_educ', 'F_Labour_force','M_Labour_force']

        # Filtering the DataFrame to include only numerical columns
        df_numerical = df[numerical_columns]

        # Calculating the correlation matrix
        correlation_matrix = df_numerical.corr()

        # correlation matrix
        plt.figure(figsize=(12, 8))
        sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
        plt.title('Correlation Matrix')
        plt.show()
```

```
In [ ]: #Pairplot
        sns.pairplot(df[['GII', 'Maternal_mortality', 'Adolescent_birth_rate', 'Seats_parli
                         'F_secondary_educ', 'M_secondary_educ', 'F_Labour_force','M_Labour

        plt.show()
```

```
In [ ]: #Box plot
        plt.figure(figsize=(12, 8))
        sns.boxplot(x='Human_development', y='GII', data=df)
        plt.title('GII by Human Development Category')
        plt.show()
```

```
In [ ]: #Scatter plot
        plt.figure(figsize=(12, 8))
        sns.scatterplot(x='Maternal_mortality', y='GII', data=df)
        plt.title('GII vs Maternal Mortality')

        plt.figure(figsize=(12, 8))
        sns.scatterplot(x='Adolescent_birth_rate', y='GII', data=df)
        plt.title('GII vs Adolescent Birth Rate')
        plt.show()
```

```
In [ ]: plt.figure(figsize=(12, 8))
        sns.scatterplot(x= 'F_secondary_educ', y='GII', data=df)
        plt.title('GII vs F_secondary_educ')
        plt.show()

        plt.figure(figsize=(12, 8))
        sns.scatterplot(x='F_Labour_force', y='GII', data=df)
        plt.title('GII vs F_Labour_force')
        plt.show()

        plt.figure(figsize=(12, 8))
        sns.scatterplot(x='Seats_parliament', y='GII', data=df)
        plt.title('GII vs Seats_parliament')
        plt.show()
```

```
In [ ]: # GII vs F_secondary_educ
        plt.figure(figsize=(12, 8))
        sns.regplot(x='F_secondary_educ', y='GII', data=df)
        plt.title('GII vs F_secondary_educ')
        plt.show()

        # GII vs F_Labour_force
        plt.figure(figsize=(12, 8))
        sns.regplot(x='F_Labour_force', y='GII', data=df)
        plt.title('GII vs F_Labour_force')
        plt.show()

        # GII vs Seats_parliament
        plt.figure(figsize=(12, 8))
        sns.regplot(x='Seats_parliament', y='GII', data=df)
        plt.title('GII vs Seats_parliament')
        plt.show()
```

```
In [ ]: #Bar plot
        plt.figure(figsize=(12, 8))
        sns.barplot(x='Human_development', y='GII', data=df)
        plt.title('Average GII by Human Development Category')
        plt.show()
```

```
In [ ]: df[['GII', 'Maternal_mortality', 'Adolescent_birth_rate', 'Seats_parliament',
            'F_secondary_educ', 'M_secondary_educ', 'F_Labour_force']].hist(figsize=(12, 10
        plt.show()
```

```
In [ ]: df.sort_values(by=['GII'],ascending=True,inplace=True)
        #plt.plot(x,y)
        plt.figure(figsize=(40,27))
        sns.barplot(x='Country', y='GII', data=df, palette='viridis',width=0.5)
        plt.title('GII Country-wise')
        plt.xlabel('Country')
        plt.ylabel('Gender Inequality Index (GII)')
        plt.xticks(rotation=90)
        plt.show()
```

# Adding GNI and HDI to the Dataset

```
In [ ]: df_gni = pd.read_csv('/kaggle/input/gross-national-income-per-capita/Gross National
```

```
In [ ]: df_gni.head()
```

```
In [ ]: df_gni.info()
```

```
In [ ]: df_gni  =df_gni [['Country','HDI Rank (2021)', 'Gross National Income Per Capita (2
```

```
In [ ]: df_gni.head()
```

```
In [ ]: # Merge the datasets on the 'Country' column
        df_combined = pd.merge(df, df_gni, on='Country', how='inner')
```

```
In [ ]: df_combined.head()
```

```
In [ ]: df_combined.info()
```

```
In [ ]: df_combined.tail()
```

```
In [ ]: df_combined.rename(columns={'Gross National Income Per Capita (2021)': 'GNI'}, inpl
        df_combined.rename(columns={'HDI Rank (2021)': 'HDI'}, inplace=True)
```

```
In [ ]: numerical_columns = ['GII', 'Maternal_mortality', 'Adolescent_birth_rate',
                             'Seats_parliament', 'F_secondary_educ',
                             'M_secondary_educ', 'F_Labour_force','M_Labour_force','GNI','H
```

```
df_numerical = df_combined[numerical_columns]
correlation_matrix = df_numerical.corr()

plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

```
In [ ]: numerical_columns = ['GII', 'Maternal_mortality', 'Adolescent_birth_rate',
                             'Seats_parliament', 'F_secondary_educ', 'F_Labour_force']

        df_numerical = df_combined[numerical_columns]

        correlation_matrix = df_numerical.corr()

        plt.figure(figsize=(12, 8))
        sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
        plt.title('Correlation Matrix')
        plt.show()
```

```
In [ ]: numerical_columns = ['GII','GNI','HDI', 'Maternal_mortality', 'Adolescent_birth_rat
                             'Seats_parliament', 'F_secondary_educ', 'F_Labour_force']

        df_numerical = df_combined[numerical_columns]

        correlation_matrix = df_numerical.corr()

        plt.figure(figsize=(12, 8))
        sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
        plt.title('Correlation Matrix')
        plt.show()
```

```
In [ ]: # Scatter plot for GII vs GNI
        plt.figure(figsize=(10, 6))
        sns.scatterplot(data=df_combined, x='GII', y='GNI', hue='Country',style='Country',
        plt.title('Scatter Plot of GII vs GNI')
        plt.xlabel('Gender Inequality Index (GII)')
        plt.ylabel('Gross National Income (GNI)')
        plt.legend(title='Country',bbox_to_anchor=(1.05, 1), loc='upper left')
        plt.show()
```

```
In [ ]: # Pair plot for multiple variables
        sns.pairplot(df_combined, vars=['GII', 'GNI', 'HDI'], hue='Country')
        plt.suptitle('Pair Plot of GII, GNI, and HDI', y=1.02)
        plt.show()
```

```
In [ ]: def remove_outliers(df, column):
            Q1 = df[column].quantile(0.25)
            Q3 = df[column].quantile(0.75)
            IQR = Q3 - Q1
            lower_bound = Q1 - 1.5 * IQR
            upper_bound = Q3 + 1.5 * IQR
            df_out = df[(df[column] >= lower_bound) & (df[column] <= upper_bound)]
            return df_out
```

```
df_no_outliers = df_combined.copy()
for column in ['GII', 'Maternal_mortality', 'Adolescent_birth_rate',
                     'Seats_parliament', 'F_secondary_educ',
                     'M_secondary_educ', 'F_Labour_force','M_Labour_force','GNI','H
    df_no_outliers = remove_outliers(df_no_outliers, column)
```

In [ ]:
```
df_numerical = df_no_outliers[numerical_columns]

correlation_matrix = df_numerical.corr()

plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

In [ ]:
```
numerical_columns = ['GII', 'Maternal_mortality', 'Adolescent_birth_rate',
                     'Seats_parliament', 'F_secondary_educ', 'F_Labour_force','GNI'
df_numerical = df_no_outliers[numerical_columns]


correlation_matrix = df_numerical.corr()

plt.figure(figsize=(12, 8))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm')
plt.title('Correlation Matrix')
plt.show()
```

# Linear Regression Model

In [ ]:
```
df_scenario1 = df_combined[['GNI', 'GII']]
X1 = df_scenario1[['GII']]
y1 = df_scenario1['GNI']
X1_train, X1_test, y1_train, y1_test = train_test_split(X1, y1, test_size=0.2, rand


model1 = LinearRegression()
model1.fit(X1_train, y1_train)


y1_pred = model1.predict(X1_test)


mse1 = mean_squared_error(y1_test, y1_pred)
r2_1 = r2_score(y1_test, y1_pred)

print(f"Scenario 1 - Mean Squared Error: {mse1}")
print(f"Scenario 1 - R^2 Score: {r2_1}")


plt.scatter(y1_test, y1_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
```

```
plt.title('Scenario 1 - Actual vs Predicted GNI')
plt.show()
```

In [ ]:
```
df_scenario2 = df_combined[['GNI', 'GII', 'HDI', 'Maternal_mortality', 'Adolescent_

X2 = df_scenario2.drop('GNI', axis=1)
y2 = df_scenario2['GNI']

X2_train, X2_test, y2_train, y2_test = train_test_split(X2, y2, test_size=0.2, rand

model2 = LinearRegression()
model2.fit(X2_train, y2_train)

y2_pred = model2.predict(X2_test)

mse2 = mean_squared_error(y2_test, y2_pred)
r2_2 = r2_score(y2_test, y2_pred)

print(f"Scenario 2 - Mean Squared Error: {mse2}")
print(f"Scenario 2 - R^2 Score: {r2_2}")

plt.scatter(y2_test, y2_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
plt.title('Scenario 2 - Actual vs Predicted GNI')
plt.show()
```

In [ ]:
```
df_model = df_combined[['GII', 'GNI', 'HDI']]
X = df_model[['GII', 'HDI']]
y = df_model['GNI']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_sta

model = LinearRegression()
model.fit(X_train, y_train)


y_pred = model.predict(X_test)
```

# Linear Regression Model Evaluation

In [ ]:
```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
rmse = np.sqrt(mse)
print(f"R^2 Score: {r2}")
print(f"Mean Squared Error: {mse}")
print(f"Mean Absolute Error: {mae}")
print(f"Root Mean Squared Error: {rmse}")
```

In [ ]:
```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
```

```
print(f"Mean Squared Error: {mse}")
print(f"R^2 Score: {r2}")
```

In [ ]:
```
# Plotting the results
plt.scatter(y_test, y_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
plt.title('Actual vs Predicted GNI with Linear regression model')
plt.show()
```

In [ ]:
```
residuals = y_test - y_pred
plt.figure(figsize=(10, 6))
plt.scatter(y_pred, residuals, edgecolors=(0, 0, 0))
plt.hlines(y=0, xmin=y_pred.min(), xmax=y_pred.max(), colors='r', linestyles='--')
plt.xlabel('Predicted GNI')
plt.ylabel('Residuals')
plt.title('Residual Plot( Linear regression model)')
plt.show()
```

# KNN Models

In [ ]:
```
df_scenario1 = df_combined[['GNI', 'GII']]

df_scenario1.dropna(inplace=True)
X1 = df_scenario1[['GII']]
y1 = df_scenario1['GNI']

X1_train, X1_test, y1_train, y1_test = train_test_split(X1, y1, test_size=0.2, rand

scaler1 = StandardScaler()
X1_train_scaled = scaler1.fit_transform(X1_train)
X1_test_scaled = scaler1.transform(X1_test)

knn1 = KNeighborsRegressor(n_neighbors=5)
knn1.fit(X1_train_scaled, y1_train)

y1_pred = knn1.predict(X1_test_scaled)

mse1 = mean_squared_error(y1_test, y1_pred)
r2_1 = r2_score(y1_test, y1_pred)

print(f"Scenario 1 - Mean Squared Error: {mse1}")
print(f"Scenario 1 - R^2 Score: {r2_1}")

plt.scatter(y1_test, y1_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
plt.title(' Actual vs Predicted GNI')
plt.show()
```

In [ ]:
```
df_scenario2 = df_combined[['GNI', 'GII', 'HDI', 'Maternal_mortality', 'Adolescent_

df_scenario2.dropna(inplace=True)
```

```
X2 = df_scenario2.drop('GNI', axis=1)
y2 = df_scenario2['GNI']

X2_train, X2_test, y2_train, y2_test = train_test_split(X2, y2, test_size=0.2, rand

scaler2 = StandardScaler()
X2_train_scaled = scaler2.fit_transform(X2_train)
X2_test_scaled = scaler2.transform(X2_test)

knn2 = KNeighborsRegressor(n_neighbors=5)
knn2.fit(X2_train_scaled, y2_train)

y2_pred = knn2.predict(X2_test_scaled)

mse2 = mean_squared_error(y2_test, y2_pred)
r2_2 = r2_score(y2_test, y2_pred)

print(f"Scenario 2 - Mean Squared Error: {mse2}")
print(f"Scenario 2 - R^2 Score: {r2_2}")

plt.scatter(y2_test, y2_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
plt.title('Scenario 2 - Actual vs Predicted GNI')
plt.show()
```

In [ ]:

```
df_knn = df_combined[['GNI', 'GII', 'HDI']]

df_knn.dropna(inplace=True)

X = df_knn[['GII', 'HDI']]
y = df_knn['GNI']
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_sta

scaler = StandardScaler()
X_train_scaled = scaler.fit_transform(X_train)
X_test_scaled = scaler.transform(X_test)

knn = KNeighborsRegressor(n_neighbors=5)
knn.fit(X_train_scaled, y_train)

y_pred = knn.predict(X_test_scaled)

mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)

print(f"Mean Squared Error: {mse}")
print(f"R^2 Score: {r2}")

# Plotting the results
plt.scatter(y_test, y_pred)
plt.xlabel('Actual GNI')
plt.ylabel('Predicted GNI')
```

```
plt.title('Actual vs Predicted GNI (KNN model)')
plt.show()
```

# KNN Model Evaluation

```
In [ ]: mse = mean_squared_error(y_test, y_pred)
        r2 = r2_score(y_test, y_pred)
        mae = mean_absolute_error(y_test, y_pred)
        rmse = np.sqrt(mse)
        mape = np.mean(np.abs((y_test - y_pred) / y_test)) * 100

        print(f"Mean Squared Error: {mse}")
        print(f"R^2 Score: {r2}")
        print(f"Mean Absolute Error: {mae}")
        print(f"Root Mean Squared Error: {rmse}")
        print(f"Mean Absolute Percentage Error: {mape}%")
```

```
In [ ]: mse = mean_squared_error(y_test, y_pred)
        r2 = r2_score(y_test, y_pred)
        mae = mean_absolute_error(y_test, y_pred)
        rmse = np.sqrt(mse)

        print(f"R^2 Score: {r2}")
        print(f"Mean Squared Error: {mse}")
        print(f"Mean Absolute Error: {mae}")
        print(f"Root Mean Squared Error: {rmse}")
```

```
In [ ]: residuals = y_test - y_pred
        plt.figure(figsize=(10, 6))
        plt.scatter(y_pred, residuals, edgecolors=(0, 0, 0))
        plt.hlines(y=0, xmin=y_pred.min(), xmax=y_pred.max(), colors='r', linestyles='--')
        plt.xlabel('Predicted GNI')
        plt.ylabel('Residuals')
        plt.title('KNN Residual Plot')
        plt.show()
```

## Output:

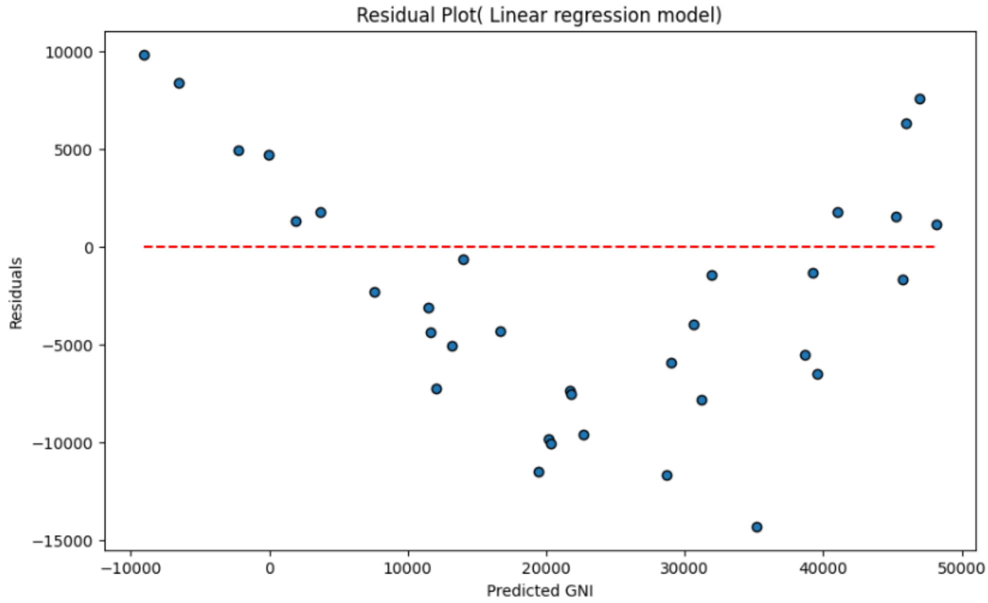



*Fig. 4*

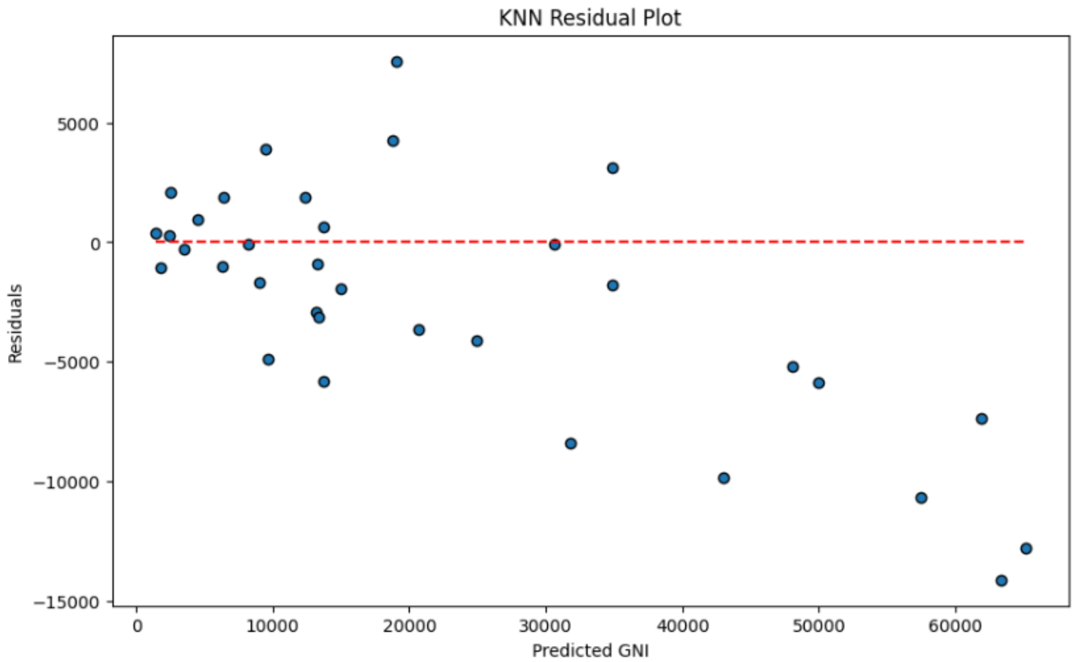



*Fig. 5*

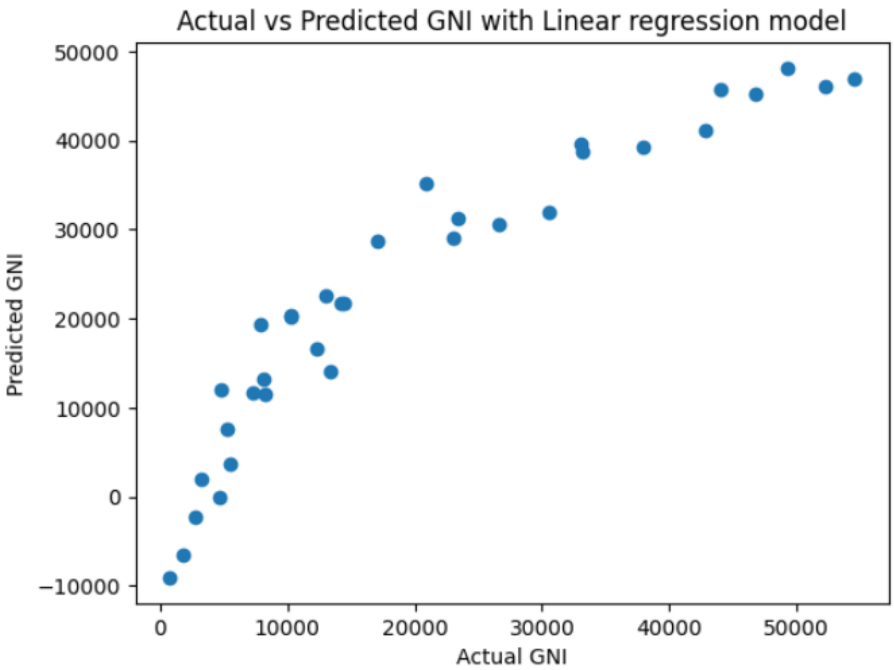



*Fig. 6*

```
Mean Squared Error: 29240778.60148078
R^2 Score: 0.8885951161213896
```

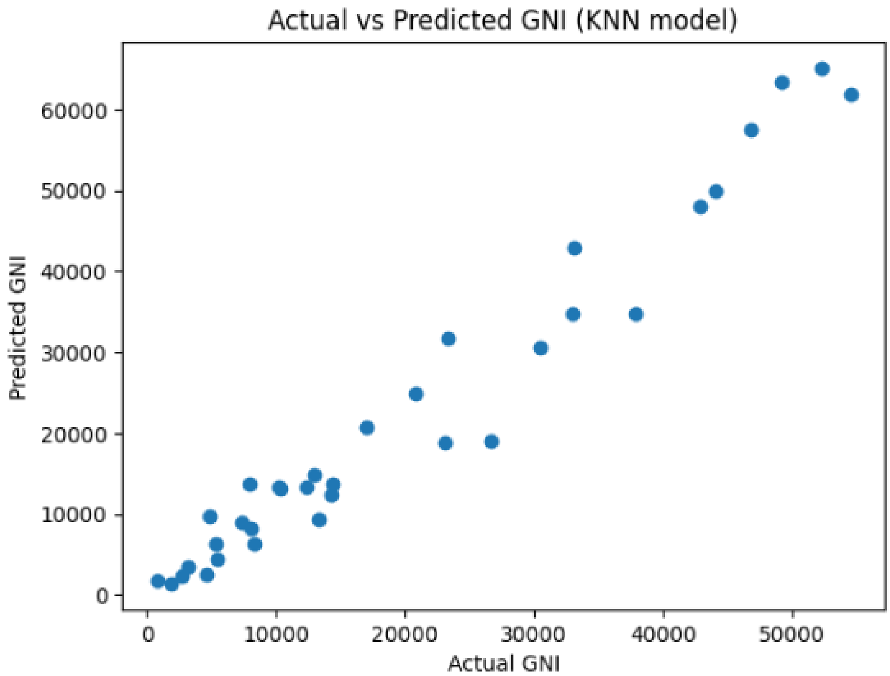



*Fig. 7*

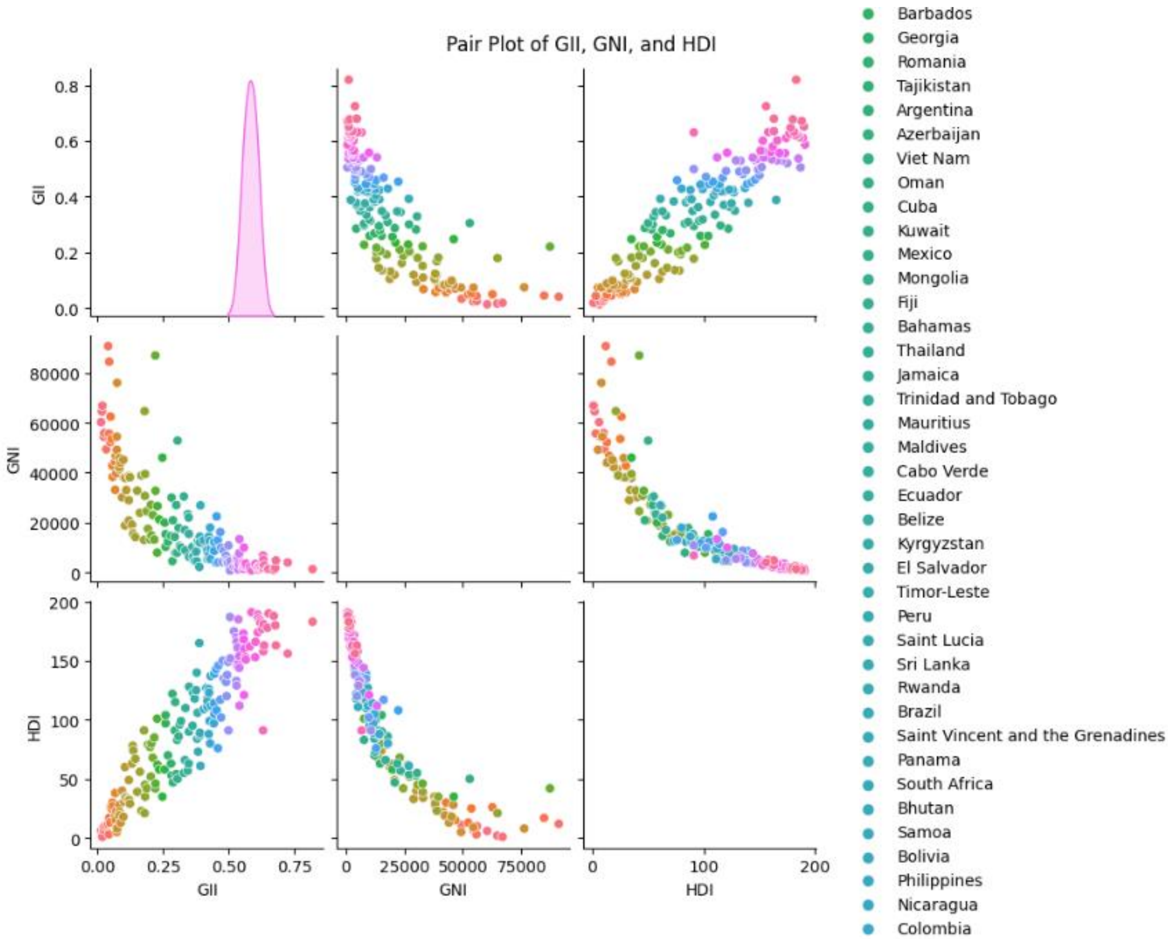

Pair Plot of GII, GNI, and HDI
GII
GNI
HDI
0.0
0.2
0.4
0.6
0.8
0
20000
40000
60000
80000
0
50
100
150
200
0.00
0.25
0.50
0.75
0
25000
50000
75000
100
Barbados
Georgia
Romania
Tajikistan
Argentina
Azerbaijan
Viet Nam
Oman
Cuba
Kuwait
Mexico
Mongolia
Fiji
Bahamas
Thailand
Jamaica
Trinidad and Tobago
Mauritius
Maldives
Cabo Verde
Ecuador
Belize
Kyrgyzstan
El Salvador
Timor-Leste
Peru
Saint Lucia
Sri Lanka
Rwanda
Brazil
Saint Vincent and the Grenadines
Panama
South Africa
Bhutan
Samoa
Bolivia
Philippines
Nicaragua
Colombia

*Fig. 8*

## KNN Model Evaluation

```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
rmse = np.sqrt(mse)
mape = np.mean(np.abs((y_test - y_pred) / y_test)) * 100

print(f"Mean Squared Error: {mse}")
print(f"R^2 Score: {r2}")
print(f"Mean Absolute Error: {mae}")
print(f"Root Mean Squared Error: {rmse}")
print(f"Mean Absolute Percentage Error: {mape}%")
```

```
Mean Squared Error: 29240778.60148078
R^2 Score: 0.8885951161213896
Mean Absolute Error: 3962.3409343311755
Root Mean Squared Error: 5407.47432739914
Mean Absolute Percentage Error: 26.312548221206722%
```

*Fig. 9*

## Linear Regression Model Evaluation

\+ Code \+ Markdown

```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
rmse = np.sqrt(mse)
print(f"R^2 Score: {r2}")
print(f"Mean Squared Error: {mse}")
print(f"Mean Absolute Error: {mae}")
print(f"Root Mean Squared Error: {rmse}")
```

```
R^2 Score: 0.8297065498094043
Mean Squared Error: 44697439.653824426
Mean Absolute Error: 5657.288166900759
Root Mean Squared Error: 6685.614381178773
```

*Fig. 10*

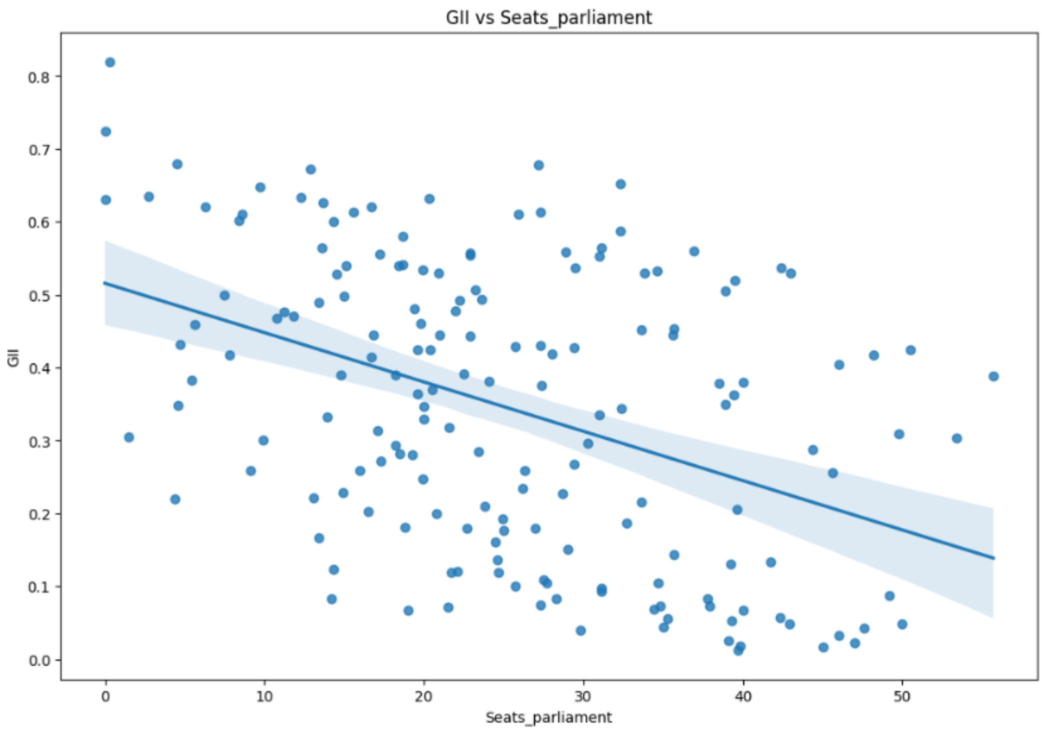

GII vs Seats_parliament
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
GII
0
10
20
30
40
50
Seats_parliament

*Fig. 11*

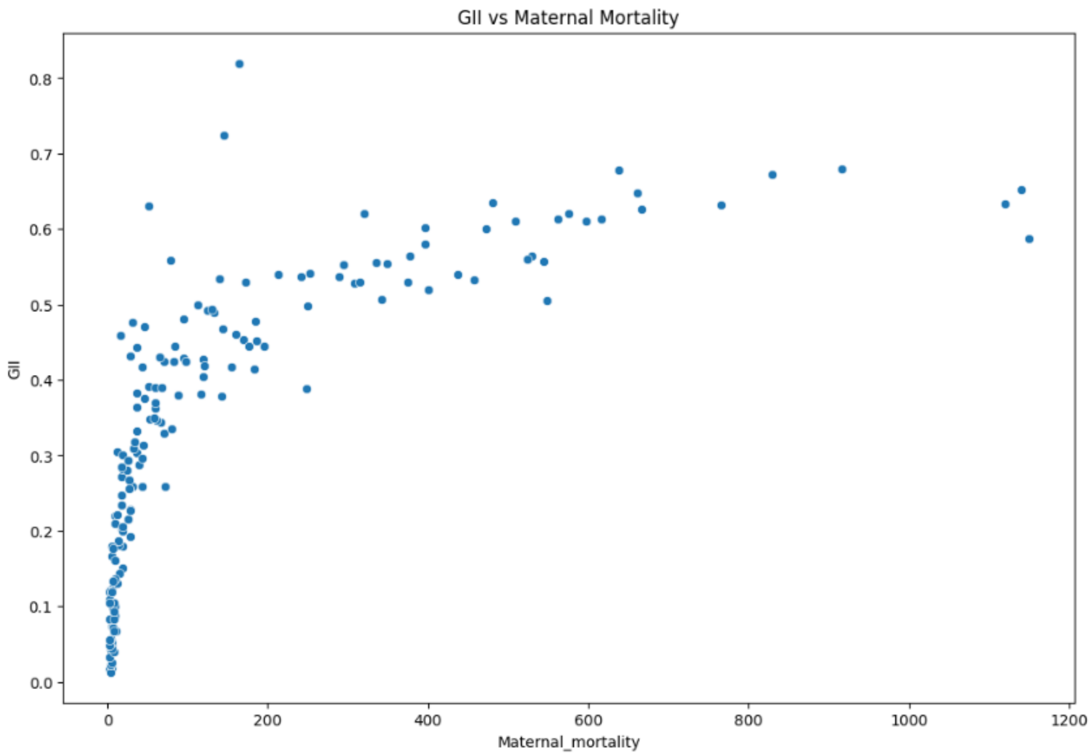
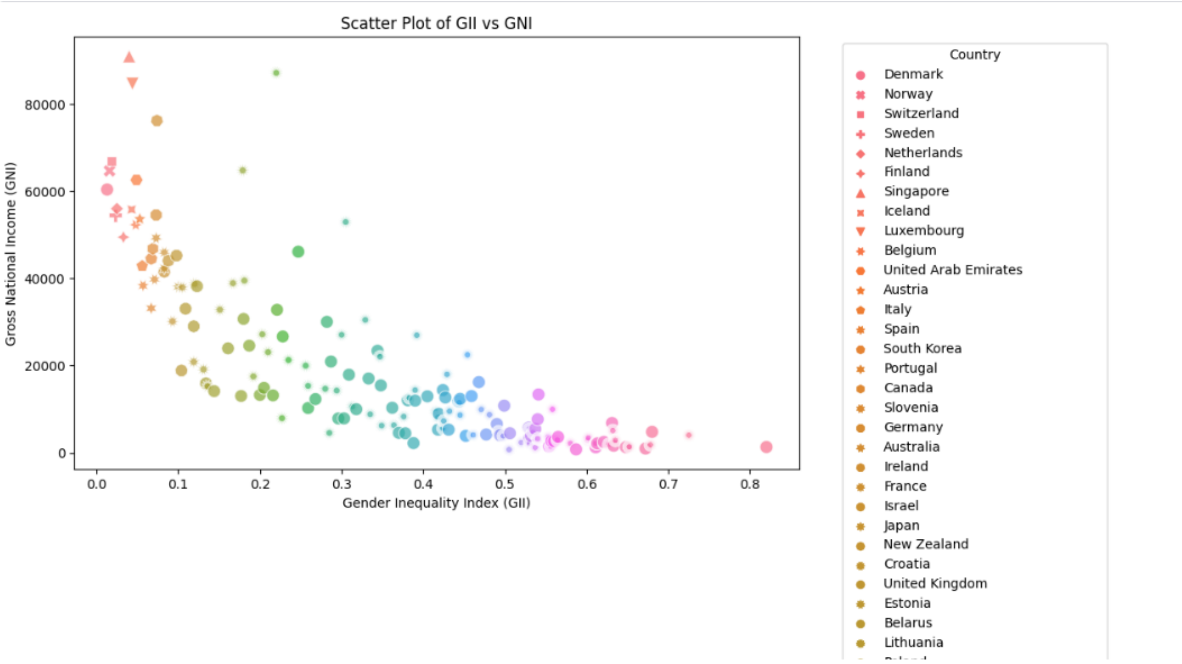

GII vs Maternal Mortality
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
GII
0
200
400
600
800
1000
1200
Maternal_mortality

*Fig. 12*



Scatter Plot of GII vs GNI
Gross National Income (GNI)
80000
60000
40000
20000
0
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Gender Inequality Index (GII)
Country
Denmark
Norway
Switzerland
Sweden
Netherlands
Finland
Singapore
Iceland
Luxembourg
Belgium
United Arab Emirates
Austria
Italy
Spain
South Korea
Portugal
Canada
Slovenia
Germany
Australia
Ireland
France
Israel
Japan
New Zealand
Croatia
United Kingdom
Estonia
Belarus
Lithuania


*Fig. 13*

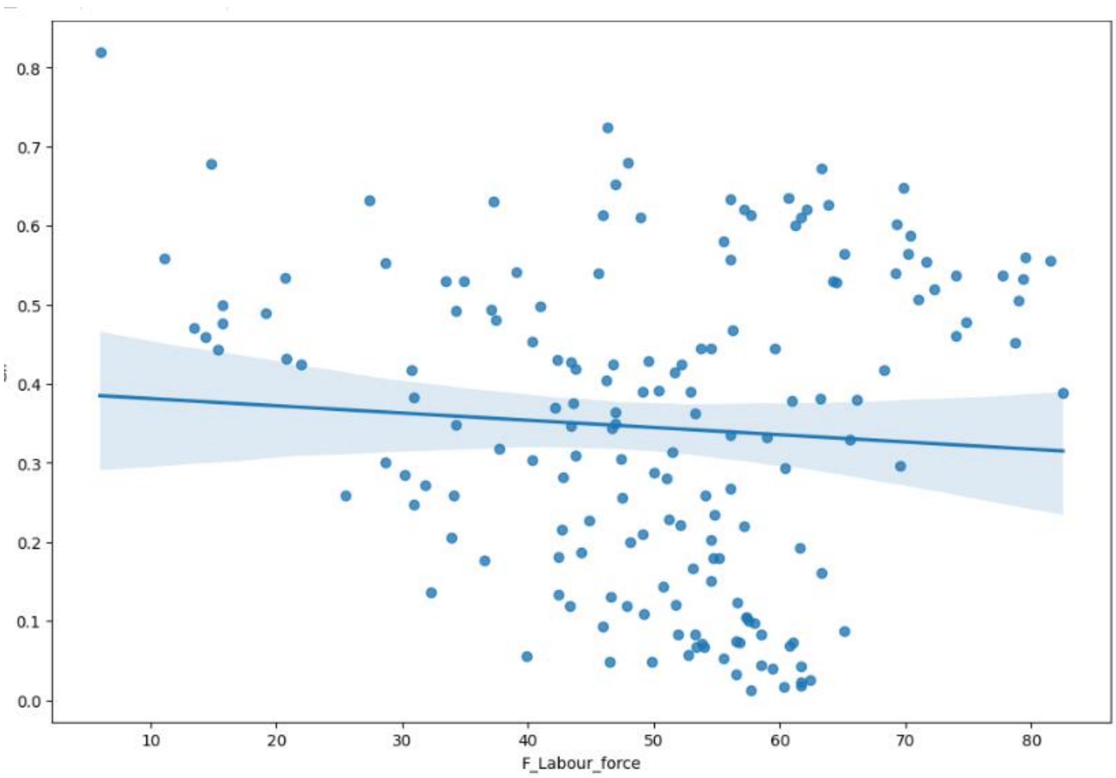

0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
10
20
30
40
50
60
70
80
F_Labour_force

*Fig. 14*

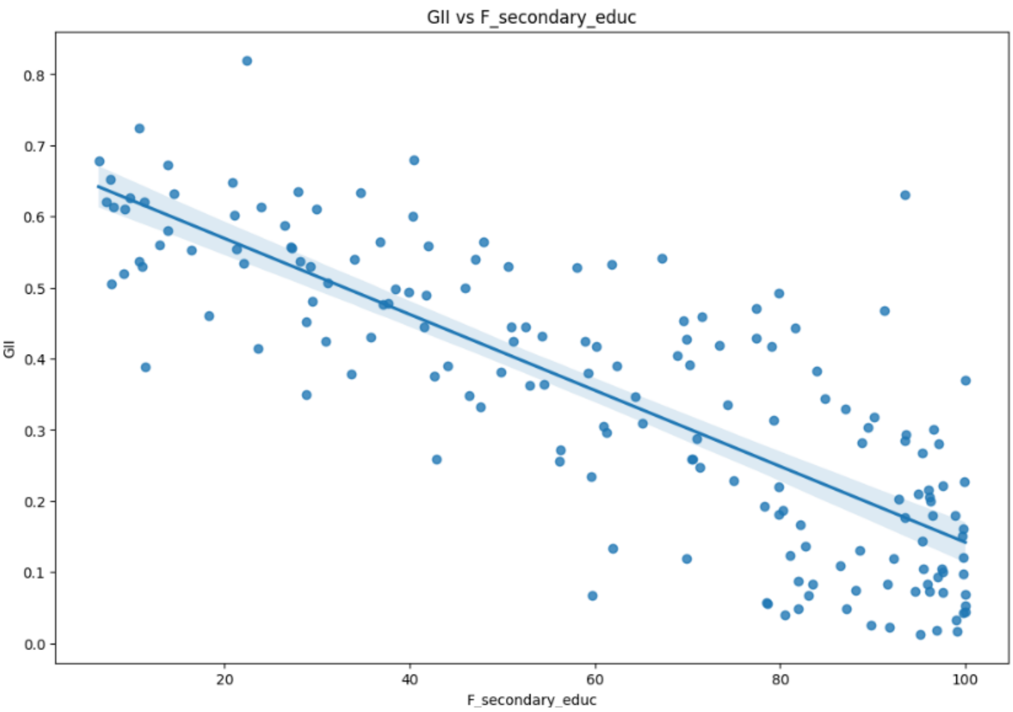

GII vs F_secondary_educ
GII
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
20
40
60
80
100
F_secondary_educ

*Fig. 15*

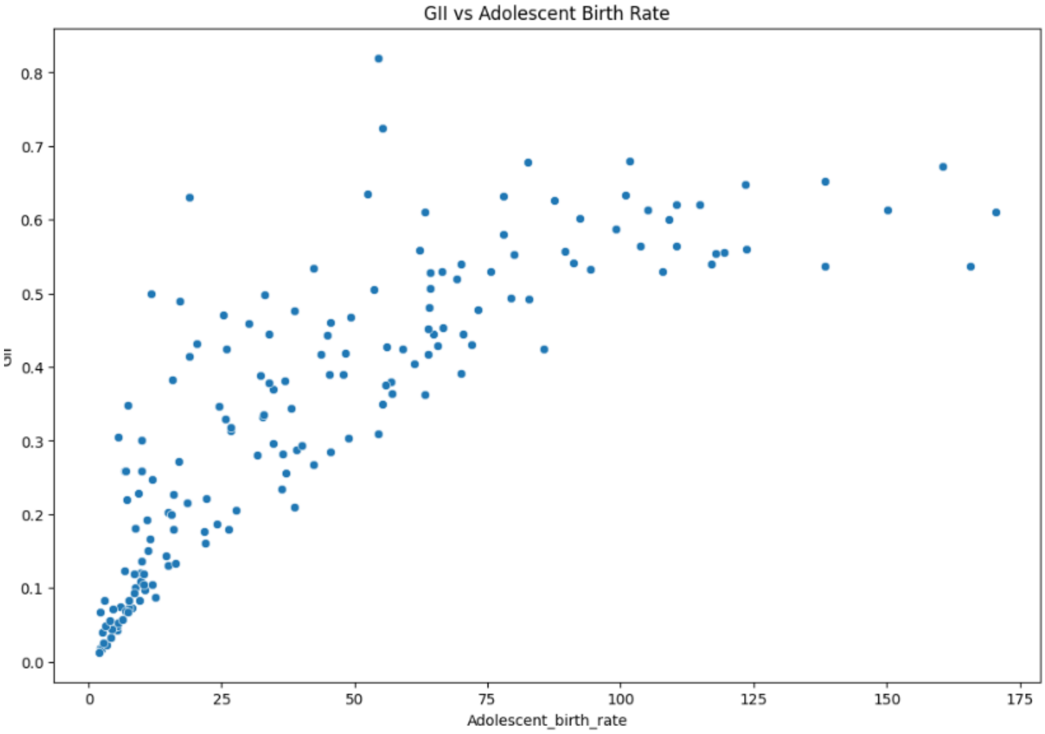

GII vs Adolescent Birth Rate
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
GII
0
25
50
75
100
125
150
175
Adolescent_birth_rate

*Fig. 16*

```python
df_combined.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 166 entries, 0 to 165
Data columns (total 13 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   Country                166 non-null    object
 1   Human_development      166 non-null    object
 2   GII                    166 non-null    float64
 3   Rank                   166 non-null    float64
 4   Maternal_mortality     166 non-null    float64
 5   Adolescent_birth_rate  166 non-null    float64
 6   Seats_parliament       166 non-null    float64
 7   F_secondary_educ       166 non-null    float64
 8   M_secondary_educ       166 non-null    float64
 9   F_Labour_force         166 non-null    float64
 10  M_Labour_force         166 non-null    float64
 11  HDI                    166 non-null    float64
 12  GNI                    166 non-null    float64
dtypes: float64(11), object(2)
memory usage: 17.0+ KB
```

*Fig. 17*

```python
df_combined.head()
```

| | Country | Human_development | GII | Rank | Maternal_mortality | Adolescent_birth_rate | Seats_parliament | F_secondary_educ | M_secondary_educ | F_Labour_force | M_Labour_force | HDI Rank (2021) | Gross National Income Per Capita (2021) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **0** | Denmark | Very high | 0.013 | 1.0 | 4.0 | 1.9 | 39.7 | 95.1 | 95.2 | 57.7 | 66.7 | 6.0 | 60364.78595 |
| **1** | Norway | Very high | 0.016 | 2.0 | 2.0 | 2.3 | 45.0 | 99.1 | 99.3 | 60.3 | 72.0 | 2.0 | 64660.10622 |
| **2** | Switzerland | Very high | 0.018 | 3.0 | 5.0 | 2.2 | 39.8 | 96.9 | 97.5 | 61.7 | 72.7 | 1.0 | 66933.00454 |
| **3** | Sweden | Very high | 0.023 | 4.0 | 4.0 | 3.3 | 47.0 | 91.8 | 92.2 | 61.7 | 68.0 | 7.0 | 54489.37401 |
| **4** | Netherlands | Very high | 0.025 | 5.0 | 5.0 | 2.8 | 39.1 | 89.8 | 92.7 | 62.4 | 71.3 | 10.0 | 55979.41100 |

*Fig. 18*

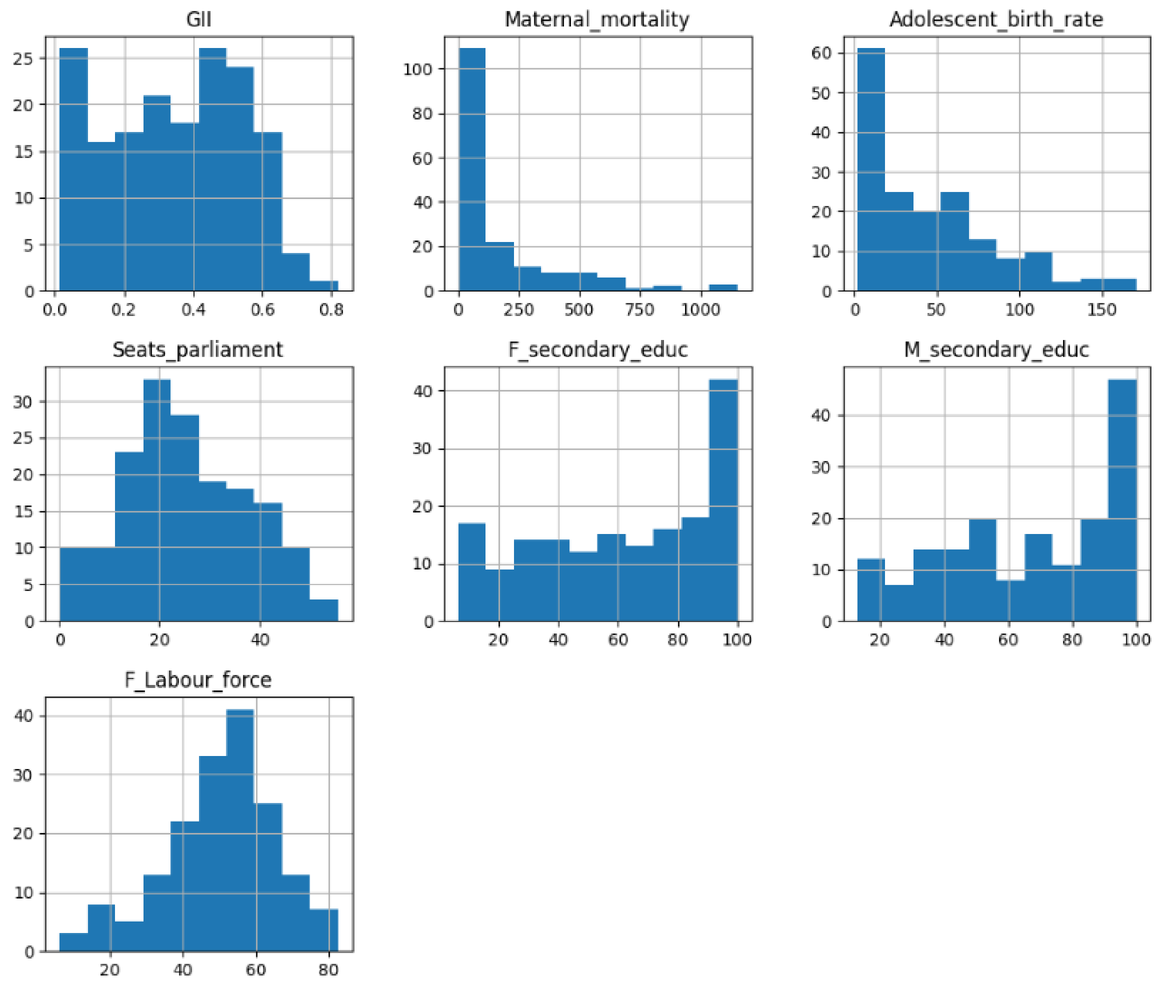

GII
Maternal_mortality
Adolescent_birth_rate
Seats_parliament
F_secondary_educ
M_secondary_educ
F_Labour_force

*Fig. 19*

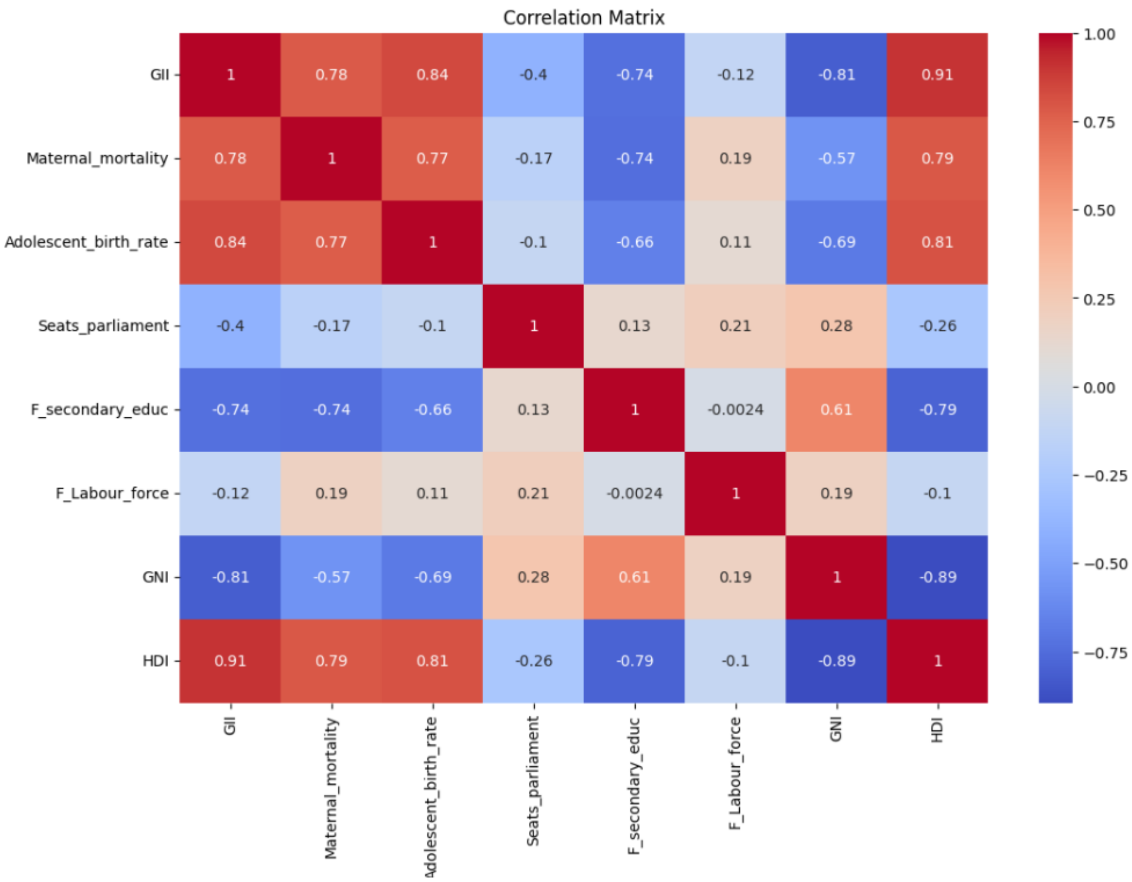

Correlation Matrix
GII
Maternal_mortality
Adolescent_birth_rate
Seats_parliament
F_secondary_educ
F_Labour_force
GNI
HDI
1 0.78 0.84 -0.4 -0.74 -0.12 -0.81 0.91
0.78 1 0.77 -0.17 -0.74 0.19 -0.57 0.79
0.84 0.77 1 -0.1 -0.66 0.11 -0.69 0.81
-0.4 -0.17 -0.1 1 0.13 0.21 0.28 -0.26
-0.74 -0.74 -0.66 0.13 1 -0.0024 0.61 -0.79
-0.12 0.19 0.11 0.21 -0.0024 1 0.19 -0.1
-0.81 -0.57 -0.69 0.28 0.61 0.19 1 -0.89
0.91 0.79 0.81 -0.26 -0.79 -0.1 -0.89 1
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75

*Fig. 20*

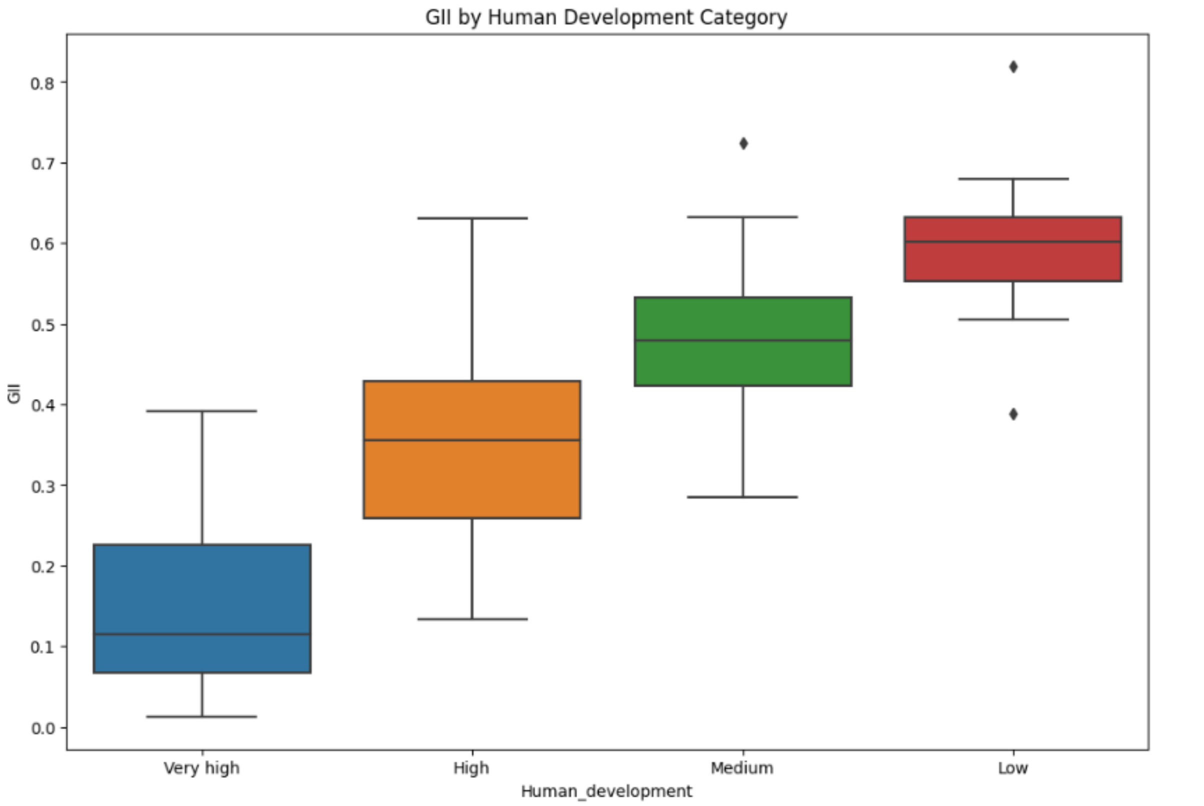

GII by Human Development Category
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
GII
Very high
High
Medium
Low
Human_development

*Fig. 21*

# Dashboard Screenshots:

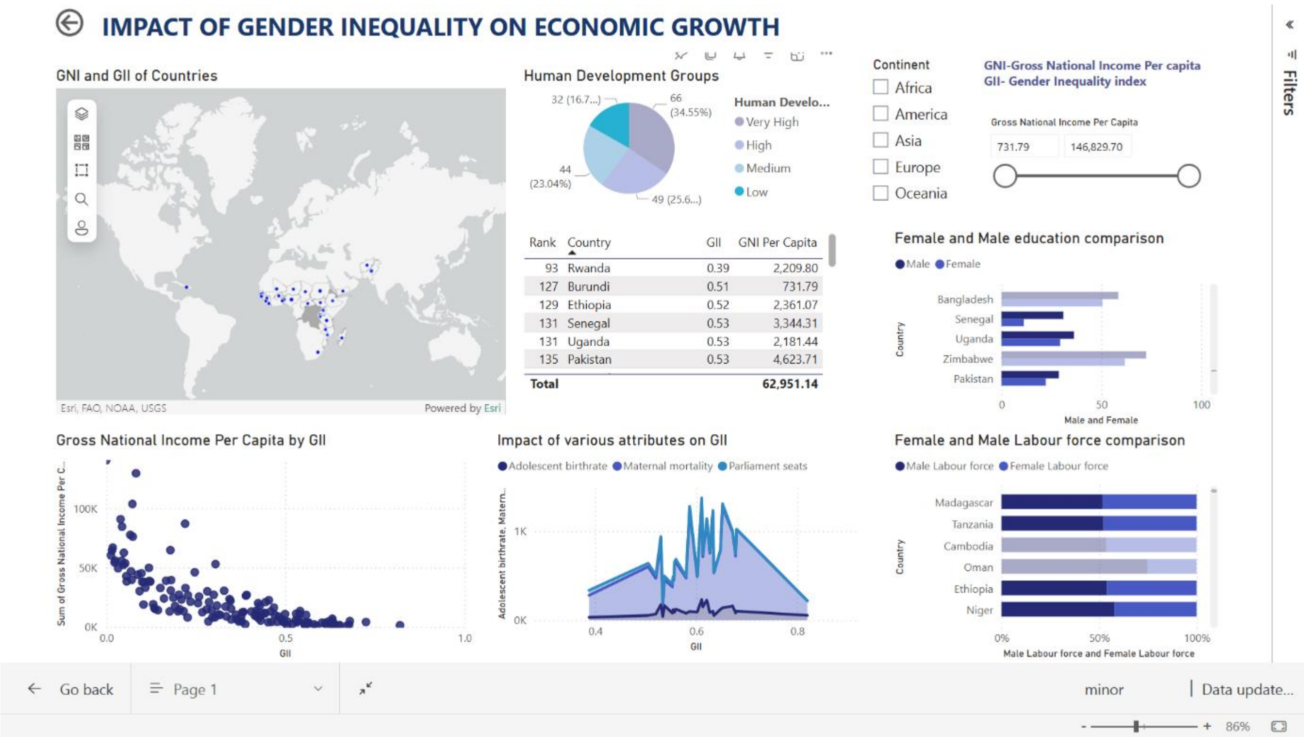



*Fig. 22*

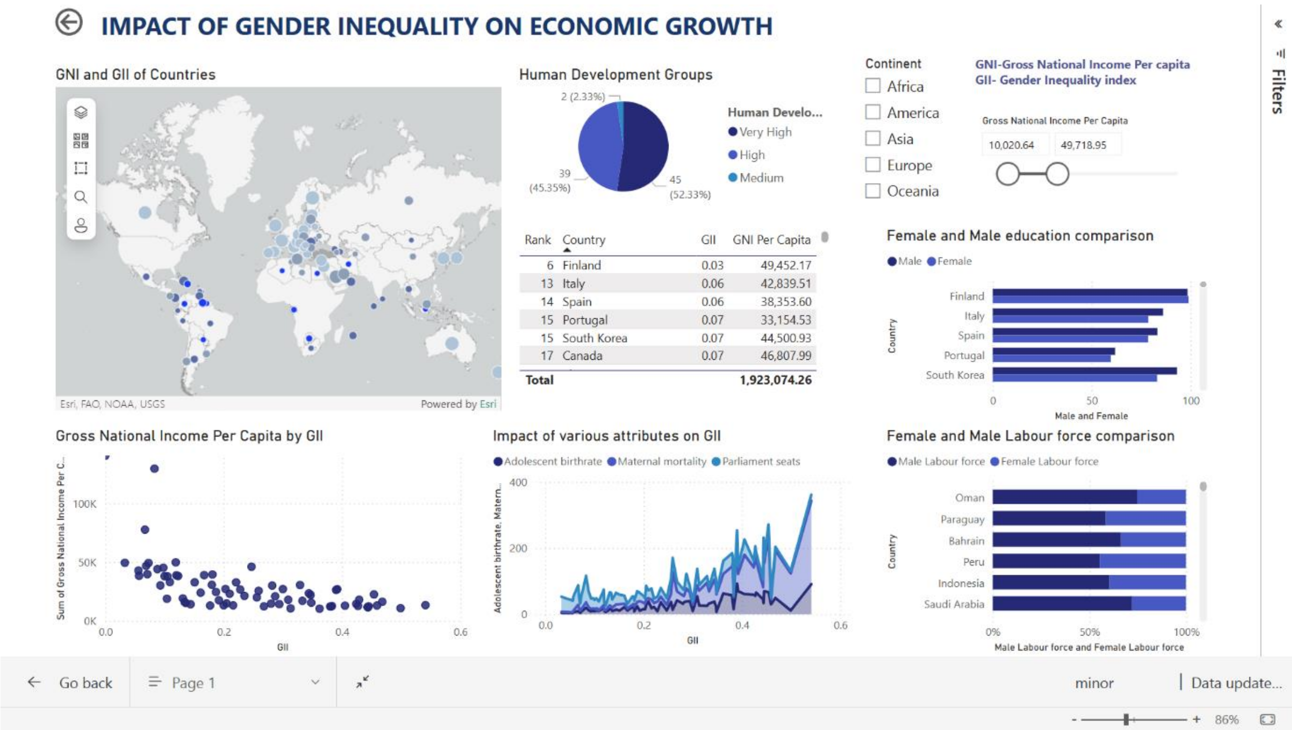



*Fig. 23*


IMPACT OF GENDER INEQUALITY ON ECONOMIC GROWTH
GNI and GII of Countries
Human Development Groups
Human Develo...
Very High
42 (100%)
Continent
America
Asia
Europe
Oceania
GNI-Gross National Income Per capita
GII- Gender Inequality index
Gross National Income Per Capita
32,820.56
146,829.70
Rank Country GII GNI Per Capita
1 Denmark 0.01 60,364.79
2 Norway 0.02 64,660.11
3 Switzerland 0.02 66,933.00
4 Sweden 0.02 54,489.37
5 Netherlands 0.03 55,979.41
6 Finland 0.03 49,452.17
Total 1,911,335.60
Female and Male education comparison
Male Female
Denmark
Norway
Switzerland
Sweden
Netherlands
Country
Male and Female
Gross National Income Per Capita by GII
Sum of Gross National Income Per C...
GII
Impact of various attributes on GII
Adolescent birthrate Maternal mortality Parliament seats
Female and Male Labour force comparison
Male Labour force Female Labour force
Qatar
United Arab ...
Kuwait
Bahrain
Saudi Arabia
Singapore
Male Labour force and Female Labour force
Go back
Page 1
minor
Data update...
86%
Filters

*Fig. 24*


IMPACT OF GENDER INEQUALITY ON ECONOMIC GROWTH
GNI and GII of Countries
Human Development Groups
Human Develo...
Very High
High
Medium
Low
Continent
Africa
America
Asia
Europe
Oceania
GNI-Gross National Income Per capita
GII- Gender Inequality index
Gross National Income Per Capita
731.79
146,829.70
Rank Country GII GNI Per Capita
7 Singapore 0.04 90,918.64
Total 90,918.64
Female and Male education comparison
Male Female
Netherlands
Finland
Singapore
Iceland
Luxembourg
Country
Male and Female
Gross National Income Per Capita by GII
Sum of Gross National Income Per C...
GII
Impact of various attributes on GII
Adolescent birthrate Maternal mortality Parliament seats
Female and Male Labour force comparison
Male Labour force Female Labour force
Singapore
Male Labour force and Female Labour force
Go back
Page 1
minor
Data update...
86%
Filters

*Fig. 25*


IMPACT OF GENDER INEQUALITY ON ECONOMIC GROWTH
GNI and GII of Countries
Human Development Groups
Continent
Africa
America
Asia
Europe
Oceania
GNI-Gross National Income Per capita
GII- Gender Inequality index
Gross National Income Per Capita
731.79
146,829.70
Rank Country GII GNI Per Capita
7 Singapore 0.04 90,918.64
11 United Arab Emirates 0.05 62,573.59
15 South Korea 0.07 44,500.93
22 Israel 0.08 41,523.74
22 Japan 0.08 42,274.29
35 Cyprus 0.12 38,188.42
Total 888,979.50
Female and Male education comparison
Gross National Income Per Capita by GII
Impact of various attributes on GII
Female and Male Labour force comparison

*Fig. 26*

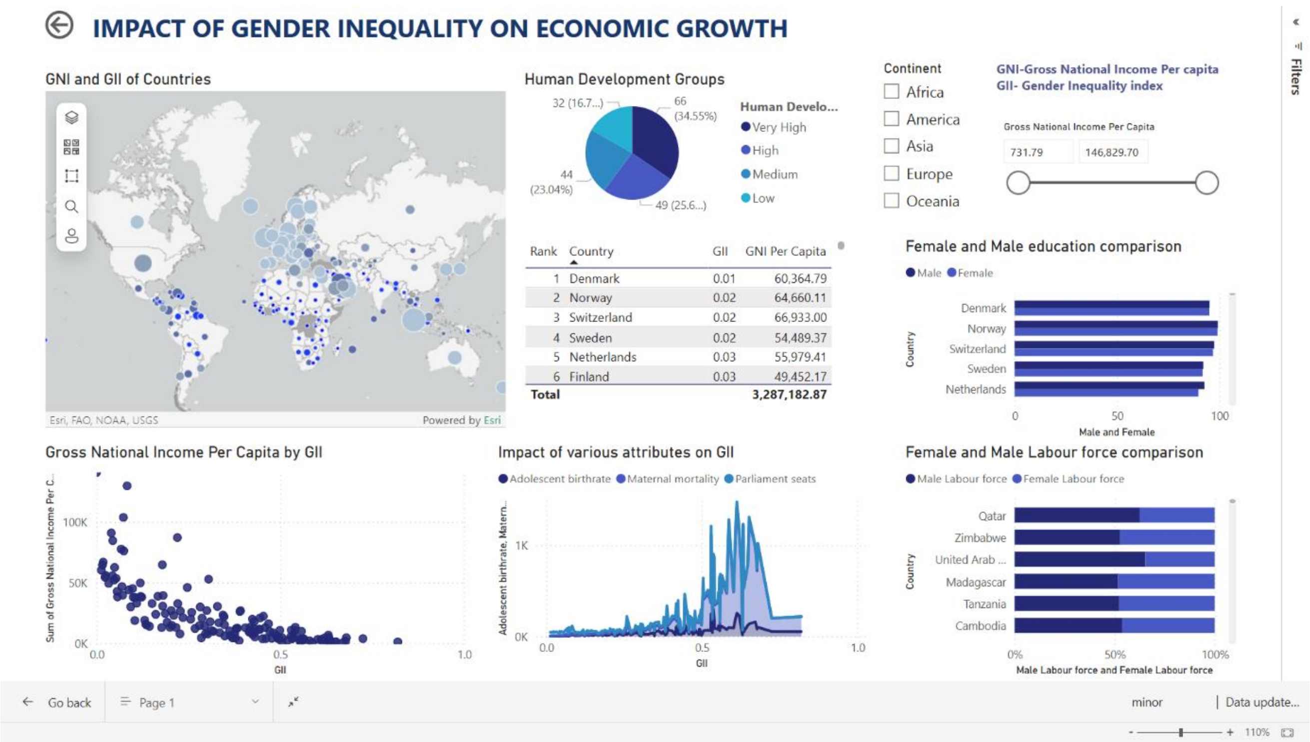

IMPACT OF GENDER INEQUALITY ON ECONOMIC GROWTH
GNI and GII of Countries
Human Development Groups
Continent
Africa
America
Asia
Europe
Oceania
GNI-Gross National Income Per capita
GII- Gender Inequality index
Gross National Income Per Capita
731.79
146,829.70
Rank Country GII GNI Per Capita
1 Denmark 0.01 60,364.79
2 Norway 0.02 64,660.11
3 Switzerland 0.02 66,933.00
4 Sweden 0.02 54,489.37
5 Netherlands 0.03 55,979.41
6 Finland 0.03 49,452.17
Total 3,287,182.87
Female and Male education comparison
Gross National Income Per Capita by GII
Impact of various attributes on GII
Female and Male Labour force comparison

*Fig. 27*

# Conclusion and Future Scope

## Summary of Findings:

- Gender Inequality Index (GII) and Economic Growth:
    - Negative Correlation: There is a strong negative correlation between the Gender Inequality Index (GII) and Gross National Income (GNI) per capita. Countries with higher gender inequality tend to have lower economic growth.
- Maternal Mortality and Economic Growth:
    - Inverse Relationship: Countries with high maternal mortality rates generally show lower GNI per capita and HDI. Reducing maternal mortality can lead to significant improvements in economic growth and development.
- Adolescent Birth Rate:
    - Economic Burden: Higher adolescent birth rates are linked to lower economic growth. Early pregnancies often disrupt female education and labor force participation, limiting economic potential.
- Female Education:
    - Increased female education levels correlate positively with higher GNI per capita and HDI. Educated women are more likely to participate in the labor force and contribute to economic growth. Investing in female education is a crucial driver for economic development.
- Female Labor Force Participation:
    - Economic Boost: Higher female labor force participation is associated with higher GNI per capita and HDI. Gender equality in the workforce enhances economic productivity and growth.
    - Policy Implications: Policies promoting gender equality in employment can significantly boost economic outcomes.

## Future Work:

- ❖ Policy Impact Analysis
  - ➢ Case Studies: Conduct case studies of specific countries or regions that have implemented successful gender equality policies to identify best practices and lessons learned.
  - ➢ Policy Simulation: Use economic modeling to simulate the potential impacts of proposed gender equality policies on economic growth and development.
- ❖ Technological Integration
  - ➢ Real-time Data Collection: Implement real-time data collection methods using mobile technology to gather up-to-date information on gender inequality indicators.

## Conclusion:

The findings highlight the critical role of gender equality in driving economic growth and development. Reducing gender inequality, improving female education, and increasing female labor force participation can lead to substantial economic benefits for countries. Addressing maternal mortality and adolescent birth rates is also essential for achieving sustainable economic growth.

# References


1. Data Sources- Data is taken from Kaggle

   https://www.kaggle.com/datasets/gianinamariapetrascu/gender-inequalityindex/data

   https://www.kaggle.com/datasets/iamsouravbanerjee/gross-national-income-percapita

2. To study about gender inquality index and other issues relalted to gender gap https://www.drishtiias.com/daily-updates/daily-news-analysis/india-s-progress-ingender-equality

3. Pandas Documentation. (n.d.). Pandas: Python Data Analysis Library. Retrieved from https://pandas.pydata.org/

4. NumPy Documentation. (n.d.). NumPy: Scientific Computing with Python. Retrieved from https://numpy.org/

5. Scikit-Learn Documentation. (n.d.). Scikit-Learn: Machine Learning in Python. Retrieved from https://scikit-learn.org/

6. Matplotlib and Seaborn Documentation. (n.d.). Visualization Tools in Python. Retrieved from https://matplotlib.org/ and https://seaborn.pydata.org/