

PROJECT PLAN

ANALYSIS OF GLOBAL SOCIO-ECONOMIC INDICATORS USING WORLD BANK DATA (2015)

1. Project Overview

Objective:

To explore socio-economic trends and correlations using 2015 world statistics to identify key factors influencing GDP and development metrics globally. The study aims to provide insights for students, researchers, and socio-economic development initiatives.

Key Research Questions:

- How does GDP correlate with literacy rates, internet penetration, and life expectancy?
- Is there a correlation between population growth and literacy/internet penetration?
- Which regions have the highest and lowest GDP per capita?
- Which regions have the highest literacy and internet penetration rates?
- Is there any correlation between exportation and GDP?
- What is the relationship between women's literacy rates and GDP?

Dataset: World Bank Data (2015) curated by GaryHoov, summarizing countries' GDP, population, literacy rates, life expectancy, and more.

Audience: Students, development professionals, educational institutions, and socio-economic researchers.

PHASE	TASKS	DEADLINE	DELIVERABLE
Planning	Define scope, objectives, and research questions.	2024-12-04	Project charter
Data Preparation	Data cleaning, preprocessing, and validation. Identify missing values and inconsistencies.	2024-12-07	Cleaned dataset
Exploratory Analysis	Conduct descriptive statistics, data visualization, and initial observations.	2024-12-08	Insights and trends
Correlation Analysis	Investigate correlations between GDP and literacy, internet penetration, life expectancy, etc.	2024-12-08	Correlation matrix, charts
Regional Analysis	Compare GDP per capita, literacy rates, and internet penetration by region.	2024-12-08	Regional comparison report
Exportation Analysis	Explore relationship between exportation and GDP.	2024-12-08	Insights on export-GDP correlation
Women's Literacy Analysis	Examine impact of women's literacy on GDP.	2024-12-08	Analysis report
Reporting	Compile results into a report and prepare a presentation for stakeholders.	2024-12-09	Final report and presentation slides

2. Methodology

Tools:

- Python: Data processing, visualization, and statistical analysis (pandas, matplotlib, seaborn).
- Pearson
- Excel: Data validation and manual adjustments.

Steps:

1. Data Cleaning:

- Address missing values, duplicates, and outliers.
- Convert categorical variables and ensure consistent units.

2. Exploratory Data Analysis (EDA):

- Visualize distributions, trends, and outliers.
- Generate summary statistics.

3. Correlation Analysis:

- Compute linear correlation.
- Visualize relationships using scatterplots and heatmaps.

4. Regional Analysis:

- Aggregate data by regions (e.g., continents).
- Analyze and visualize GDP per capita, literacy, and internet rates.

5. Export-GDP Analysis:

- Explore regression between export values and GDP.
- Include trade-to-GDP ratio analysis.

6. Women's Literacy & GDP:

- Compare women's literacy rates with GDP across countries.
- Conduct regression analysis for deeper insights.

7. Synthesis:

- Summarize key trends and insights.
- Draft implications for socio-economic development strategies.

3. Risk Assessment

Risk	Mitigation
Missing/incomplete data	Use imputation techniques or supplement with external sources if possible.
Misinterpretation of correlations	Clearly distinguish correlation vs causation in reporting.
Regional disparities	Use normalized metrics (e.g., per capita values) for fair comparisons.
Time constraints	Prioritize high-impact analyses; use iterative improvement for details.

4. Deliverables

- Dataset Preparation Report: Document on data cleaning and transformations.
- Analysis Outputs: Charts, correlation matrices, and regional comparisons.
- Final Report: Preparation of report and presentation of analysis.
- Presentation Deck: Visual summary for stakeholders.

5. Timeline

Duration: 10 days

Start Date: 02.12.2024

End Date: 12.12.2024

6. Stakeholder Communication Plan

Audience	Meeting Time	Communication	Content
Project Team	3 classes	Slack/Meetings/WApp	Progress updates, issues, next steps
Students/Research	2 classes + 2 off-class meetings	Report/Presentation/WApp	Key findings, analysis and insights.
Development Initiatives	3 off-class meetings	Report/Presentation	Preparation of report and presentation of analysis.