PHASE 3 - DEVELOPMENT

1. Define Analysis Objectives:

- Clearly outline the specific objectives of your public transportation efficiency analysis. For example, you might aim to assess punctuality, ridership trends, operational costs, or customer satisfaction. Be as specific as possible in defining your goals.

2. Data Source Collection:

- Identify the source of your transportation data. It could be public transit authorities, government agencies, or a combination of sources.

- Acquire the necessary datasets that align with your analysis objectives. You may require data on routes, schedules, ridership, operational costs, maintenance, and more.

3. Data Quality Assessment:

- Before diving into data processing, perform an initial assessment of the collected data to understand its quality and structure.

- Check for data completeness, accuracy, consistency, and any potential issues or anomalies. Ensure you have the metadata or data dictionary to understand what each variable represents.

4. Data Preprocessing:

- Clean the collected data to ensure its quality and accuracy. This involves tasks like:

  - Handling missing data: Decide whether to impute missing values or remove incomplete records.

  - Standardizing data formats: Ensure data is in a consistent format (e.g., date formats, units of measurement).

  - Removing duplicates: Eliminate duplicate records that may skew analysis results.

  - Handling outliers: Determine how to deal with extreme values that may impact the analysis.

5. Data Integration:

- If your data comes from multiple sources, integrate it into a cohesive dataset. Ensure that the data is in a compatible format for analysis.

- Perform data transformation, such as joining tables, merging datasets, or aggregating data if necessary.

6. IBM Cognos Setup:

- Install and configure IBM Cognos if it's not already set up within your organization.

7. Data Loading into IBM Cognos:

- Load the preprocessed dataset into IBM Cognos for analysis. Ensure that the data is structured appropriately, and the data types are correctly defined.

8. Data Modeling:

- Define the structure of your data model within IBM Cognos. This may include creating dimensions, measures, and relationships.

9. Data Exploration:

- Explore the dataset within IBM Cognos to better understand its content and relationships.

- Identify relevant variables and patterns that are essential for your analysis objectives.

10. Data Visualization in IBM Cognos:

- Use IBM Cognos to create visualizations based on your analysis objectives. Common visualization types include charts, graphs, maps, and dashboards.

- Create interactive visualizations that allow you to explore data insights more effectively.

11. Analysis and Interpretation:

- Conduct the actual analysis using the visualizations and data models you've created in IBM Cognos.

- Extract meaningful insights, trends, and patterns from the data.

- Formulate conclusions and make data-driven recommendations based on your analysis.

12. Documentation and Reporting:

- Document your analysis process, including the steps taken, data sources used, and any assumptions made.

- Create comprehensive reports or presentations to communicate your findings to stakeholders, making it clear how the analysis aligns with the defined objectives.

13. Data Privacy and Compliance:

- Ensure compliance with data privacy regulations and ethical considerations throughout the analysis process, especially when handling sensitive transportation data.