PUBLIC TRANSPORTATION ANALYSIS EFFICIENCY

Phase 4: Development part 2

Submitted by:

Kurukunda Srinivas

Au723921243028

Srinusb000@gmail.com

Using visualization libraries (e.g., Matplotlib, Seaborn) to create histograms, scatter plots, and correlation matrices of water quality analysis.

Content:

In this project, we delve into the analysis of a public transportation system's efficiency, a vital component of urban mobility. By gathering data from various sources, including schedules, GPS tracking, and passenger feedback, we have meticulously prepared and integrated the information to ensure data quality. Utilizing IBM Cognos, we've designed interactive dashboards and reports that visually represent on-time performance, passenger feedback, and service efficiency metrics. These user-friendly tools empower stakeholders to explore data, identify trends, and make informed decisions. Complementing these visuals, Python plays a crucial role in advanced data analysis, calculating service punctuality rates and conducting sentiment analysis. Furthermore, custom Python-generated visualizations enrich the insights presented in our reports. Automation, integration, and version control ensure that our analysis remains up-to-date and accessible. By documenting our methods and deploying the project's outcomes, we aim to contribute to the continuous improvement of the transportation system's efficiency, ultimately benefiting the community it serves.

## Stage 1: Project Planning and Data Collection

1. Define Project Scope: Start by clearly defining the scope of your project. What specific aspects of the public transportation system do you want to analyze? Is it on-time performance, passenger feedback, route efficiency, or a combination of these factors?

2. Identify Data Sources: Determine where you'll collect data. Possible sources include government databases, transportation companies, public surveys, and GPS tracking systems on vehicles.

3. Data Gathering: Collect the necessary data, which could include: On-time performance data: Arrival and departure times, delays. Passenger feedback: Surveys, comments, ratings. Service efficiency metrics: Vehicle maintenance records, fuel consumption, route details.

4. Data Quality: Ensure that the collected data is clean, complete, and reliable. Address issues like missing values, duplicates, and outliers.

## Stage 2: Data Integration and Cleaning

5. Data Integration: Combine data from various sources into a unified dataset. Ensure that data from different sources is mapped to a common set of variables.

6. Data Cleaning: Clean the dataset by addressing missing values, outliers, and inconsistencies. Use data preprocessing techniques such as imputation, filtering, and data normalization.

## Stage 3: IBM Cognos Dashboard and Reports

7. IBM Cognos Setup: Install and configure IBM Cognos Analytics. Create data connections to your integrated dataset.

8. Dashboard Design: Create interactive dashboards to visualize key performance indicators (KPIs) and metrics. Examples of widgets you might include:

- Cards displaying on-time performance rates.

- Line charts showing trends in on-time performance.

- Heatmaps showing passenger feedback sentiment by route.

9. Report Creation: Generate detailed reports that provide insights into different aspects of the transportation system, such as route efficiency, driver performance, and maintenance schedules.

10. Interactive Features: Incorporate interactive features like filters and drill-down capabilities to allow users to explore data in-depth.

11. Scheduling and Sharing: Set up automated report scheduling to ensure that stakeholders receive regular updates. Share dashboards and reports with relevant decision-makers.

## Stage 4: Advanced Data Analysis with Code

12. Python Integration: Utilize Python for more advanced data analysis and insights:

13. Service Punctuality Rate: Calculate service punctuality rates by analyzing the on-time performance data. This might involve calculating the percentage of on-time arrivals and departures.

14. Sentiment Analysis: Conduct sentiment analysis on passenger feedback using natural language processing (NLP) libraries like NLTK, spacy, or textblob. Determine passenger satisfaction levels and common complaints.

15. Data Visualization with Python: Create additional visualizations using Python to complement your Cognos reports. These could include custom charts, word clouds, and geographical maps to visualize route efficiency.

16. Data Exploration: Use Python for exploratory data analysis (EDA) to uncover patterns, correlations, and anomalies within your dataset. This can help you identify root causes of inefficiencies.

## Stage 5: Integration and Automation

17. Automate Analysis: Develop Python scripts or workflows to automate data extraction, cleaning, and analysis tasks. Schedule these tasks to run at regular intervals for ongoing monitoring.

18. Dashboard Integration: Embed Python-generated visualizations in your IBM Cognos dashboards if needed. This can provide a unified interface for stakeholders.

19. Version Control: Use version control systems like Git to manage code changes, ensuring you can track and revert modifications.

## Stage 6: Documentation and Deployment

20. Documentation: Create comprehensive documentation that includes:

- Explanations of your analysis methods.

- Code documentation.

- Definitions of key metrics.

- User guides for accessing and using the dashboards and reports.

21. Deployment: Deploy your dashboards, reports, and Python code in a production environment accessible to relevant stakeholders. Ensure that the data is updated as frequently as needed.

22. User Training: Provide training to end-users and stakeholders on how to effectively use the dashboards, reports, and analysis tools. Ensure they understand the insights and how to make data-driven decisions.

23. Continuous Improvement: Regularly update and refine your dashboards, reports, and analysis as new data becomes available, and as transportation policies and needs evolve. Gather feedback from users to make improvements.

This detailed project plan outlines the steps from project planning to deployment for analyzing the efficiency of a public transportation system. By following these steps, you can effectively utilize IBM Cognos and Python for a comprehensive analysis that benefits transportation stakeholders and improves service quality.