

# Report Template

## Individual Coursework 5DATA004C Data Science Project Lifecycle

Name: Ruchintha Dias

Student Id: 20232724 / 20844558

Dataset: [National Transport Report Tables — 2020\\_2021 - Inter Provincial Bus Routes – Normal Buses \(As at May 2017\)](#)

Link to Video presentation: [https://drive.google.com/file/d/1mMPykZ-Ro6dWbDe7EhzmE4Lq8wiJu4gn/view?usp=drive\\_link](https://drive.google.com/file/d/1mMPykZ-Ro6dWbDe7EhzmE4Lq8wiJu4gn/view?usp=drive_link)

Link to Streamlit app: <https://dspl-icw-nydnctkxdemoqffgb5xyzq.streamlit.app/>

Link to GitHub repository: <https://github.com/Ruchintha-Dias/DSPL-ICW/tree/main>

## Aims and Objectives

### Dataset Overview:

A National Transport Commission publication contains the complete data on Sri Lankan inter-provincial bus routes through its "National Transport Report Tables — 2020\_2021 - Inter Provincial Bus Routes – Normal Buses (As at May 2017)" dataset. The information system executes six fundamental functions to provide route numbers as well as origin point, and destination point data alongside numbers of buses and operating characteristics and travel distance details. The database originates from a National Transport Commission publication with Team Watchdog making it accessible on their GitHub repository.

### Project Aims:

1. The focus of interactive dashboard development targets essential inter-provincial bus transportation statistics and trends in Sri Lanka.
2. The system needs an easy-to-use interface that lets users control the dataset through adjustable parameters across connected visualization components.
3. The system needs to extract important findings through its ability to identify which routes transport the most buses while also determining their greatest distances and usual travel durations.

### Objectives:

1. A dataset appears on the display through Streamlit framework where users find multiple interactive visualization elements.
2. Methods of data visualization should be implemented to show operational metrics.
3. Users can assess data using selection options for routes and origin-destination-distance criteria together with travel time parameter adjustments.
4. Customers obtain details concerning operational efficiency in combination with route characteristics.

## Development Methodology

Agile Development Approach: A project structure following Agile development methodology enabled the creation of work through iterative development and constant feedback integration and adaptation processes. A project segmentation included the following stages:

1. The project team acquired data while understanding its structures as well as its content.
2. The team processed the data by resolving emptiness in values while standardizing field titles and maintaining data uniformity.
3. Dashboard Design: Planned the layout and user interface of the Streamlit application.

4. Implementation: Developed the Streamlit app with interactive elements and visualizations.
5. The team performed a full examination to verify application functionality together with user-friendly features.
6. The delivery of the application involved public link availability through deployment.
7. Flexibility in development occurred through this methodology while immediate feedback integration became possible.

## Requirements

### Functional Requirements

#	Title	Description
FR-01	Data Loading	The system will retrieve data from a remote CSV file located on GitHub.
FR-02	Interactive Filters	The system must allow users to interact with filters through multiple selection options that control route numbers and origins and destinations.
FR-03	Range Sliders	The interface includes range sliders that enable users to apply distance along with travel time restrictions for their dataset searches.
FR-04	Visualizations	The application shows different charts such as bar graphs combined with scatter plot elements and heatmap illustrations to display multiple data points.
FR-05	Summary Metrics	Present key metrics such as total number of routes, average distance, and average travel time.

### Non Functional Requirements

#	Title	Description
NFR-01	Performance	Make sure the dashboard permits users to open and interact with it at a fast rate.
NFR-02	Usability	A user-friendly interface needs to be created which provides intuitive functionality.
NFR-03	Accessibility	The application enables users to access its functions through a public URL which does not need authentication.
NFR-04	Maintainability	Structure the code for easy maintenance and future improvements.
NFR-05	Scalability	Future-proof the application design by building its ability to process bigger datasets.

## Test cases

#	TC1	Title:	<i>Data Loading</i>
Description	Verify that the dataset is loading correctly from the GitHub repository.		
Steps and input data	<ol style="list-style-type: none"> <li>1. Launch the Streamlit app.</li> <li>2. Observe the initial loading process.</li> </ol>		
Dependencies	GitHub		
Expected result	The dataset should load without errors, and the dashboard should display the data accordingly.		
#	TC2	Title:	<i>Filter Functionality</i>
Description	Ensure that multiple choice filters for route numbers, origins, and destinations work as expected.		
Steps and input data	<ol style="list-style-type: none"> <li>1. Select specific route numbers from the multiple choice filter.</li> <li>2. Observe the changes in the visualizations.</li> <li>3. Repeat for origins and destinations.</li> </ol>		
Dependencies	Streamlit		
Expected result	Visualizations should update to reflect the selected filters, and only show relevant data.		
#	TC3	Title:	<i>Range Slider Functionality</i>
Description	Check the distance and travel time range sliders to filter the data.		
Steps and input data	<ol style="list-style-type: none"> <li>1. Adjust the distance slider to a specific range.</li> <li>2. Observe the changes in the visualizations.</li> <li>3. Repeat for the travel time slider.</li> </ol>		
Dependencies	Streamlit		
Expected result	Visualizations should update to reflect the selected ranges, and only show data within the specified distance and travel time.		
#	TC4	Title:	<i>Visualization Accuracy</i>
Description	Verify that the graphs accurately represent the filtered data.		
Steps and input data	<ol style="list-style-type: none"> <li>1. Apply specific filters and range selections.</li> <li>2. Examine bar charts, scatter plots, and heat maps.</li> <li>3. Cross-reference with raw data to ensure accuracy.</li> </ol>		
Dependencies	Streamlit		
Expected result	Visualizations should accurately represent filtered data, with accurate values and labels.		
#	TC5	Title:	<i>Summary Metrics Calculation</i>
Description	Ensure that the summary metrics display correct calculations based on the filtered data.		
Steps and input data	<ol style="list-style-type: none"> <li>1. Apply specific filters and range selections.</li> <li>2. Observe the summary metrics that appear at the top of the dashboard.</li> <li>3. Manually calculate the desired metrics based on the filtered data.</li> </ol>		
Dependencies	Streamlit		

Expected result	<i>The summary metrics should match the manually calculated values, reflecting accurate computations based on the filtered dataset.</i>
-----------------	---

## Test log

TC	Date	Executed by	Actual result	Pass/Fail	Notes
TC1	2025/05/06	Ruchintha Dias	The dataset was loaded successfully without errors.	Pass	None
TC2	2025/05/06	Ruchintha Dias	Filters were applied correctly; visualizations were updated accordingly.	Pass	None
TC3	2025/05/06	Ruchintha Dias	The range slider filtered the data as expected; the visualizations reflected the changes.	Pass	None
TC4	2025/05/06	Ruchintha Dias	The visualizations accurately represented the filtered data; the values matched the raw data.	Pass	None
TC5	2025/05/06	Ruchintha Dias	Summary metrics show accurate calculations based on filtered data.	Pass	None