A total of ten supplementary files are provided. These materials contain sufficient detail to enable full replication of the procedures used in the ERP report and to ensure transparency in the analysis.

**1. output.csv**

This file contains the complete dataset required for the project. The data were extracted from the original records of the New York City (NYC) Taxi & Limousine Commission (TLC).

**2. encoded\_output.csv**

This file includes data for all locations and all 29 days of February 2024. It reports the number of pick-ups for each time slot together with the corresponding demand levels.

**3. assignment\_matrix.csv**

This file is used to generate the alluvial plot. It contains the class identifiers assigned to each location.

**4. Attributes.docx**

This document serves as metadata, providing detailed descriptions of all variables employed in the project as well as in the original dataset.

**5. Analysis.R**

This script contains the procedures for exploratory data analysis (EDA) and the plotting of metrics used in model selection.

**6. MarkovChainMixtureEM.R**

This script provides the structure for fitting a mixture of Markov Chain models via the Expectation–Maximization (EM) algorithm. The code was originally developed by my supervisor, Dr. Mark Muldoon, and is gratefully acknowledged here.

**7. testing.R**

This script applies the dataset to the model structure and conducts five independent runs in order to identify the most appropriate model.

**8. TransitionMatrixHeatmap.R**

This script generates heat maps of the gamma transition matrices corresponding to the four classes reported in Section 5.

**9. AlluvialCategories.R**

This script produces the alluvial plot, which provides valuable support for model selection.

**10. DiscussionCode.R**

This script reports the proportions of each class, supporting the interpretation and discussion of results.