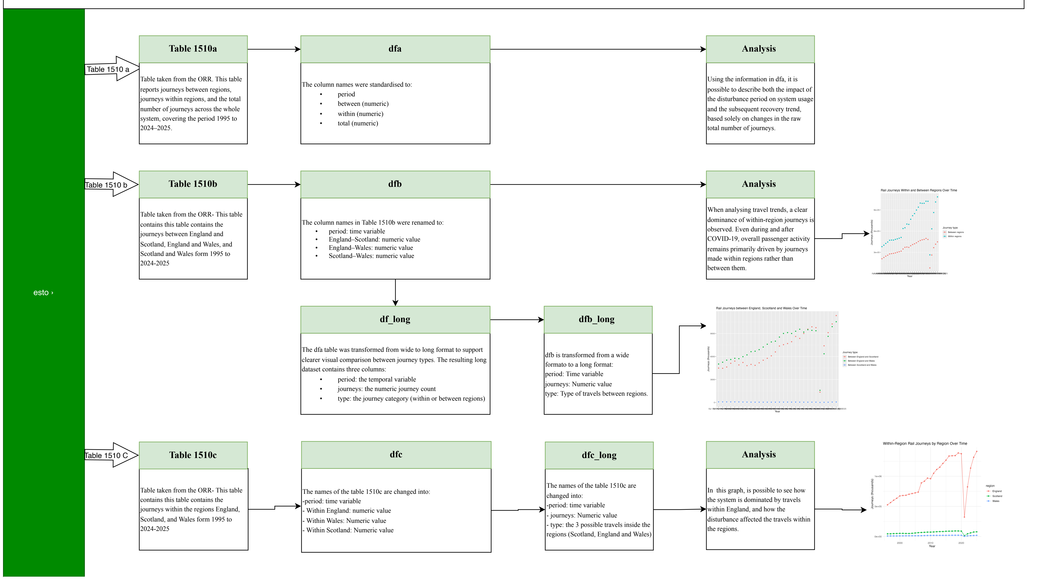
Methodological Note on Exploratory Analysis and Regional Scope

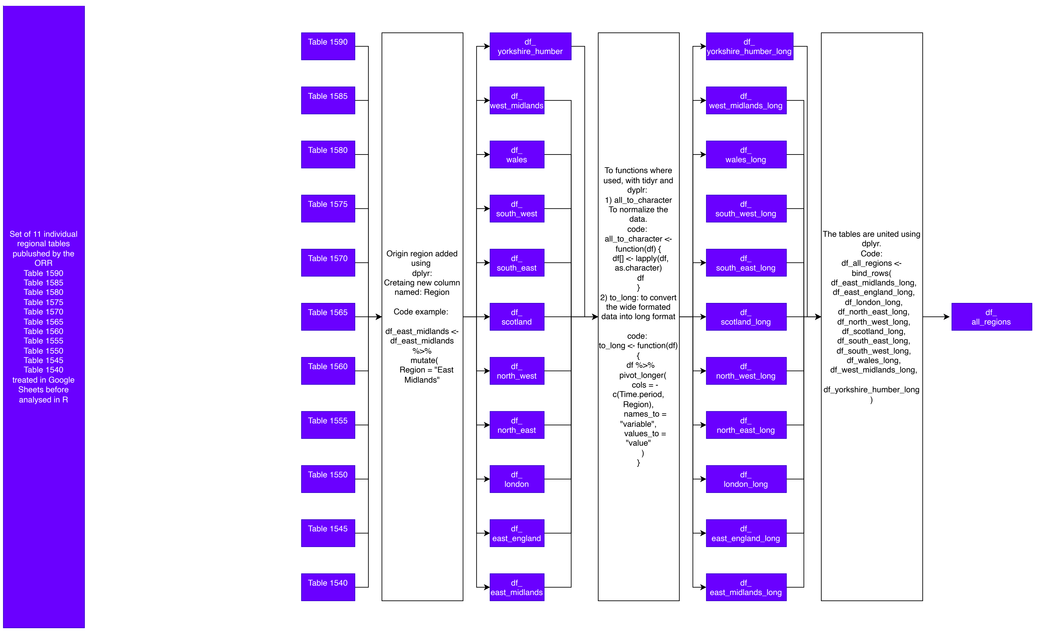
Although datasets from the ORR 1510 family were not explicitly used in the final analytical models, they played an important role during the early stages of the research by supporting contextual understanding of how the UK railway system has evolved over time. These tables functioned as an initial exploratory reference that helped assess how passenger journey data could be structured and interpreted, and they contributed to identifying a potential methodological limitation: regional aggregation can obscure important internal dynamics of the system.

Early exploratory analysis based on a limited subset of regions revealed that not all regions behaved uniformly over time. Differences in scale, temporal evolution, and responses to external shocks—most notably the COVID-19 pandemic—suggested that regions exhibited distinct structural trajectories rather than a single homogeneous pattern. This insight motivated the hypothesis that regional behaviour was heterogeneous and that different regions experienced and responded to disruption in systematically different ways. As a result, the analysis transitioned from exploratory, partial views toward a comprehensive regional approach. All eleven operational regions were subsequently incorporated into the final analysis and form the core empirical basis of the study. For this reason, selected intermediate analyses are included in an illustrative manner, not as final results but as part of the analytical reasoning process that shaped the research design.

In particular, the complete data-cleaning and preparation pipeline for the regional dataset is presented in full, as regional aggregation proved to be the most complex and methodologically critical stage of the workflow. Showing this process helps clarify the analytical choices made, the assumptions required when working with aggregated data, and the limitations inherent in regional-level representations of a highly interconnected transport system.



First insight: Regions within the United Kingdom Railway system behave differently, but the system usage appears to be concetraded within England.
Second insight: The predominat type of travel appears to be within regions, a trend mantiend hisotircyly, despite this, is poosyble to see different sgrent of inter-regionals connections and hisotricall shits over time.



First insight: Although regions behave differently, railway usage is heavily concentrated in England and dominated by a small number of hubs, particularly London, making region-level analysis too coarse on its own.

Second insight: Even when connections are represented symmetrically (A - B = B - A), their importance is asymmetric: the same link can represent a minor share of journeys for one region and a dominant share for another.