Cloud Data Platform and Tools - Power BI Desktop Report

1. Introduction

Winter service provision is a core responsibility for local authorities in the UK, particularly in urban areas where snow and ice can significantly affect mobility, accessibility, and road safety. In a city such as Edinburgh, gritting routes and grit bin placement play a crucial role in maintaining safe movement by prioritising key roads for treatment and providing residents with access to de-icing materials. Ensuring that these static assets are effectively aligned with operational gritting routes remains a challenge due to the spatial complexity of urban infrastructure.

The increasing availability of open geospatial data, combined with cloud-oriented business intelligence tools, provides new opportunities to address these challenges more systematically. While Microsoft Power BI is not a dedicated geographic information system, it provides a flexible environment for integrating heterogeneous datasets and supporting exploratory spatial analysis (Microsoft, 2023).

This project develops an interactive Power BI dashboard to support winter service analysis for the City of Edinburgh. GeoJSON datasets describing grit bin locations and gritting routes are transformed and integrated to enable proximity-based assessment of winter service coverage. By examining the spatial relationship between pedestrian-focused assets and treated road networks, the dashboard highlights grit bins located beyond a defined distance threshold, indicating potential gaps in service provision.

An independently sourced road traffic collision dataset is included to provide additional contextual insight into road safety conditions across the city. Although not used to infer causality, this dataset supports broader situational awareness during adverse weather conditions.

2. Datasets

This section evaluates the datasets used in the dashboard, focusing on their data types, structure, and suitability for spatial analysis within a cloud-oriented analytical environment.

2.1 Data Types and Cloud Data Concepts

The grit bin locations and gritting routes datasets are supplied in GeoJSON format. GeoJSON is a semi-structured, object-based spatial standard designed primarily for data exchange and interoperability (IETF, 2016). Its nested structure aligns more closely with non-relational, BLOB style data than with traditional relational tables, making it unsuitable for direct analytical querying without transformation. Consequently, the GeoJSON datasets require flattening and normalisation before they can support filtering, aggregation, and distance-based analysis within Power BI. This reflects a common cloud analytics pattern in which raw object-based data is reshaped into a tabular semantic model for analytical consumption.

In contrast, the independently sourced road traffic collision dataset is provided as a CSV file, representing a structured relational format that aligns naturally with Power BI's internal tabular model and therefore requires comparatively limited preprocessing. This distinction highlights the differing analytical overheads associated with object-based versus relational data when consumed by business intelligence platforms.

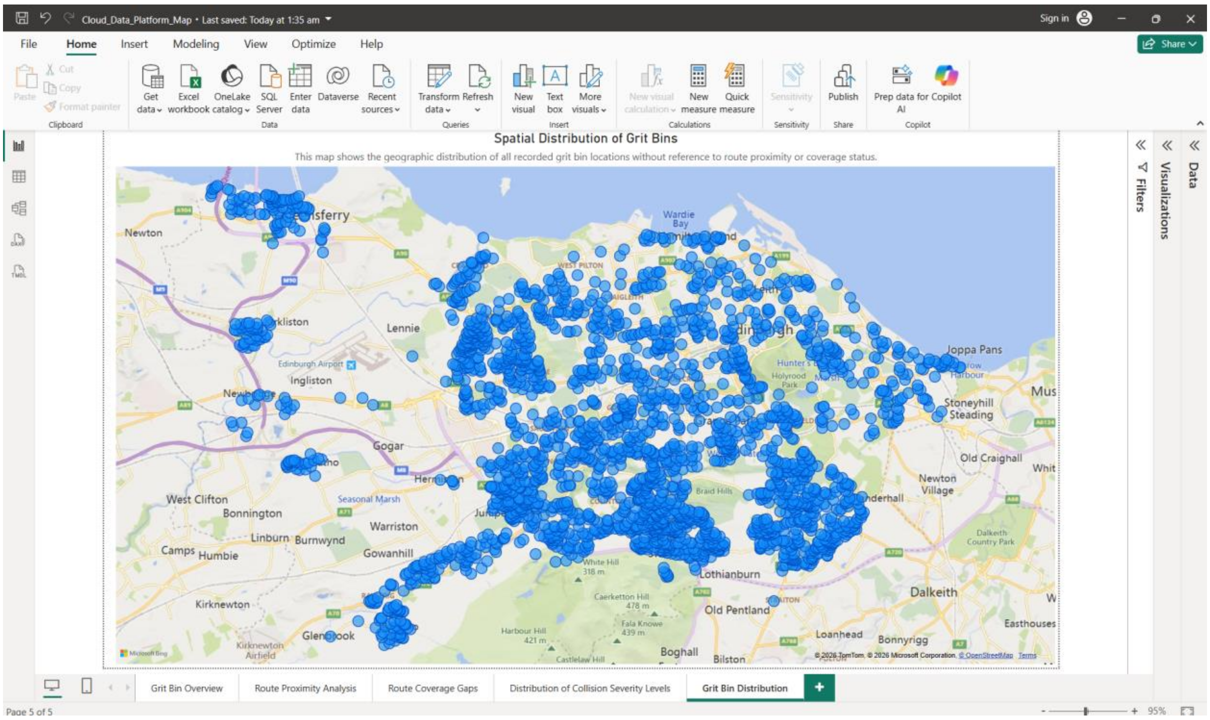
While Power BI is not a cloud storage platform, it functions as a cloud-oriented analytical consumer by operating on data already structurally optimised for analysis. Transforming object-based spatial data into a tabular model improves analytical accessibility and performance but introduces trade-offs, including the simplified geographic structure and loss of topology. Power BI lacks native topology, network awareness, and spatial indexing found in dedicated GIS or spatial databases such as PostGIS. While a spatial database such as PostGIS would preserve topology and enable network-constrained distance modelling, its use would introduce unnecessary architectural complexity for a citizen-facing exploratory dashboard. In this context, Power BI's simplified spatial model represents a proportionate analytical compromise, prioritising transparency and accessibility over exhaustive spatial precision.

2.2 Grit Bin Locations Dataset

The grit bin locations dataset provides point-based spatial data representing publicly accessible grit bins across the City of Edinburgh. Each record includes geographic coordinates alongside descriptive attributes identifying individual bin locations. As pedestrian-focused winter service assets, grit bins form the primary basis for assessing accessibility to treated road infrastructure.

An initial visualisation of the dataset, shown in Figure 1, provides an overview of city-wide coverage patterns before analytical modelling. This establishes a baseline for subsequent proximity analysis and highlights variations in asset density across different areas of the city.

Figure 1: *Spatial distribution of grit bin locations across the City of Edinburgh before analytical modelling*



2.3 Gritting Routes Dataset

The gritting routes dataset represents road segments prioritised for treatment during winter conditions and is modelled using line-based geometries. These routes reflect operational decisions made by the local authority to maintain mobility and safety on key transport corridors.

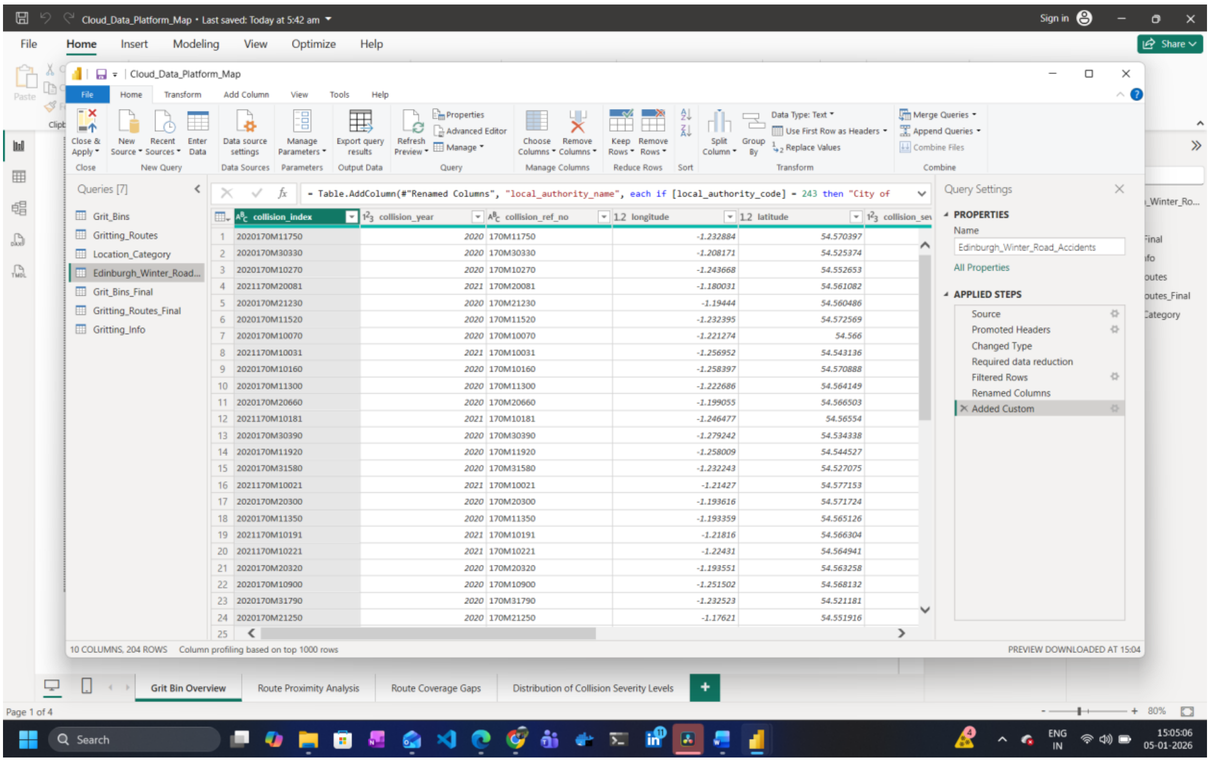
When integrated with the grit bin locations dataset, the gritting routes enable coverage assessment based on spatial proximity rather than visual inspection alone. This supports a more systematic evaluation of winter service provision and aligns with national guidance emphasising the prioritisation of key routes during adverse weather conditions (UK Roads Liaison Group, 2016).

2.4 Independent Dataset Integration

To meet the requirement for an independent dataset, a road traffic collision dataset for the City of Edinburgh was incorporated into the Power BI model. The dataset is provided in CSV format and contains geocoded collision records with associated attributes such as data and severity.

Before integration, the dataset was filtered to the City of Edinburgh and reduced to analytically relevant fields using Power Query, as illustrated in Figure 2. The dataset is not directly joined to the winter service data, avoiding unsupported causal interference while still providing useful contextual insight into road safety patterns.

Figure 2: *Power Query transformation of an independently sourced road traffic collision dataset filtered to the City of Edinburgh*



2.5 Dataset Summary

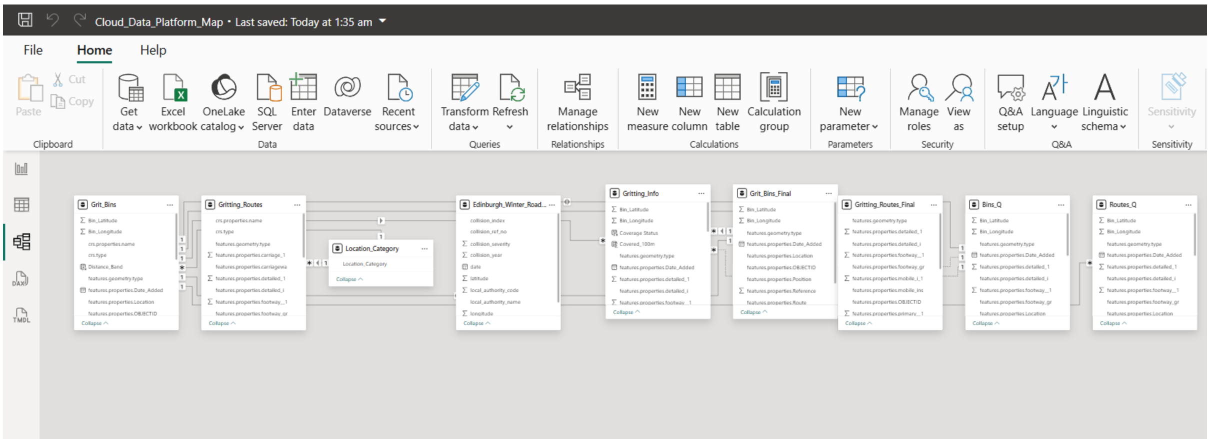
Table 1 summarises the datasets used in this project, including their geometry types, source formats, and analytical roles.

Table 1: *Summary of datasets used in the study*

Dataset	Geometry	Source Type	Analytical Purpose
Grit Bin Locations	Point	GeoJSON	Identify pedestrian winter service assets
Gritting Routes	Line	GeoJSON	Represent vehicular winter service coverage
Road Traffic Collisions	Point	CSV file	Provide independent road safety context

To support transparency and reproducibility, a Power BI data model diagram is included as Figure 3, illustrating the relationships between datasets and derived analytical measures in line with best practice for spatial analysis (Goodchild, 2010). Overall, the combination of object-based GeoJSON ingestion, relational restructuring, and semantic modelling demonstrates how heterogeneous cloud data types can be operationalised cohesively for decision support rather than exhaustive spatial precision.

Figure 3: *Power BI data model illustrating relationships between the unified Gritting_Info table, the Location_Category dimension, and the independent collision dataset*



3. Data transformation and Modelling

This section describes the transformation and modelling processes applied in Power BI to support spatial analysis of winter service provision.

3.1 GeoJSON Datasets transformation

Both GeoJSON datasets were imported using the JSON connector and flattened by expanding the underlying FeatureCollection structures. Geometry and attribute properties were separated and normalised, while redundant metadata was removed to improve clarity and performance. This transformation converts nested spatial objects into a relational structure suitable for Power BI's analytical engine, enabling filtering, aggregation, and distance-based calculations.

For grit bins, point geometries were retained and latitude and longitude values derived from coordinate arrays. For gritting routes, line geometries and coordinate sequences were preserved to support subsequent distance calculations. Data types were explicitly assigned, and columns were renamed using human-readable conventions.

3.2 Unified Spatial Fact Table

The transformed grit bin and gritting route datasets were appended into a single table named Gritting_Info, which serves as the central spatial fact table. A categorial attribute (Location_Type) distinguishes between bins and routes, enabling independent filtering within a unified structure and reducing structural complexity. This design aligns with dimensional modelling principles by consolidating heterogeneous spatial records into a single analytical fact table (Kimball and Ross, 2013).

To support clarity, a separate Location_Category dimension was created and linked to the fact table using a one-to-many relationship, with Location_Category acting as the primary key. This dimensional separation enables categorial indexing on location type, improving query efficiency, filter responsiveness, and overall performance within Power BI's VertiPaq storage engine. The resulting star style schema enhances aggregation efficiency while maintaining flexibility for interactive spatial analysis.

3.3 Proximity-Based Modelling

To assess winter service coverage, a proximity-based approach was implemented to calculate the distance between each grit bin and the nearest gritting route. Bins were classified as effectively covered if located within 50 meters of a gritting route, supporting clear spatial interpretation and interactive filtering within the Power BI dashboard. The 50 metre threshold was selected as a conservative proxy for short pedestrian access under winter conditions, balancing analytical clarity with the limitations of straight-line distance. Distances were derived using a minimum point-to-line method, measuring the shortest distance from each bin to the nearest route geometry, rather than relying on route vertices, which reduces spatial bias in proximity estimates (de Smith et al., 2009).

Calculations were implemented using DAX to ensure proximity measures respond dynamically to user interaction. Additional measures summarised total bins, bins beyond the threshold, and the percentage within effective coverage, supporting headline indicators and coverage gap mapping (Russo and Ferrari, 2019). A random sample of bins was manually validated using direct coordinate comparison, increasing confidence in the analytical results. Manual verification step strengthens confidence in the analytical measures and mitigates risks associated simplified distance modelling.

3.4 Independent Dataset Transformation

The road traffic collision dataset was transformed using Power Query by filtering recording to the City of Edinburgh, removing non-essential attributes, and standardising data types. The data was retained as a standalone table to preserve analytical transparency.

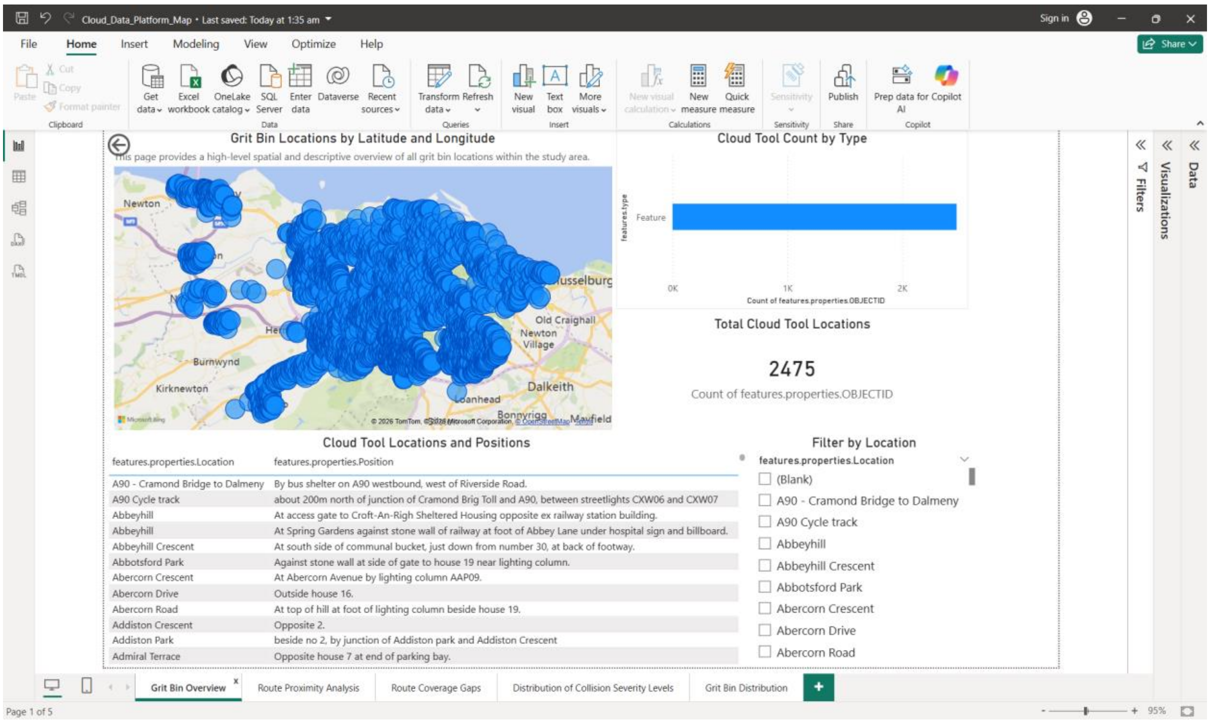
3.5 Data Model Structure

The model adopts a star style structure centred on the Gritting_Info fact table, with a Location_Category dimension linked via a one-to-many relationship, while the collision dataset remains disconnected to preserve analytical transparency. Figure 3 illustrates how these relationships support proximity-based modelling and analytical measures. Performance optimisation is achieved through dimensional separation and dictionary encoding within Power BI's VertiPaq engine, reducing cardinality and limiting filter propagation. Implementing proximity logic in DAX rather than Power Query ensures spatial metrics remain responsive to slicers and user interaction, supporting efficient aggregation and advanced interactive geospatial computation.

4. Data Visualisation and Dashboard Design

This dashboard was designed as an interactive interface to support both city-wide overview and localised investigation of winter service coverage.

Figure 4: *Overview of the Power BI dashboard for winter service analysis in the City of Edinburgh*



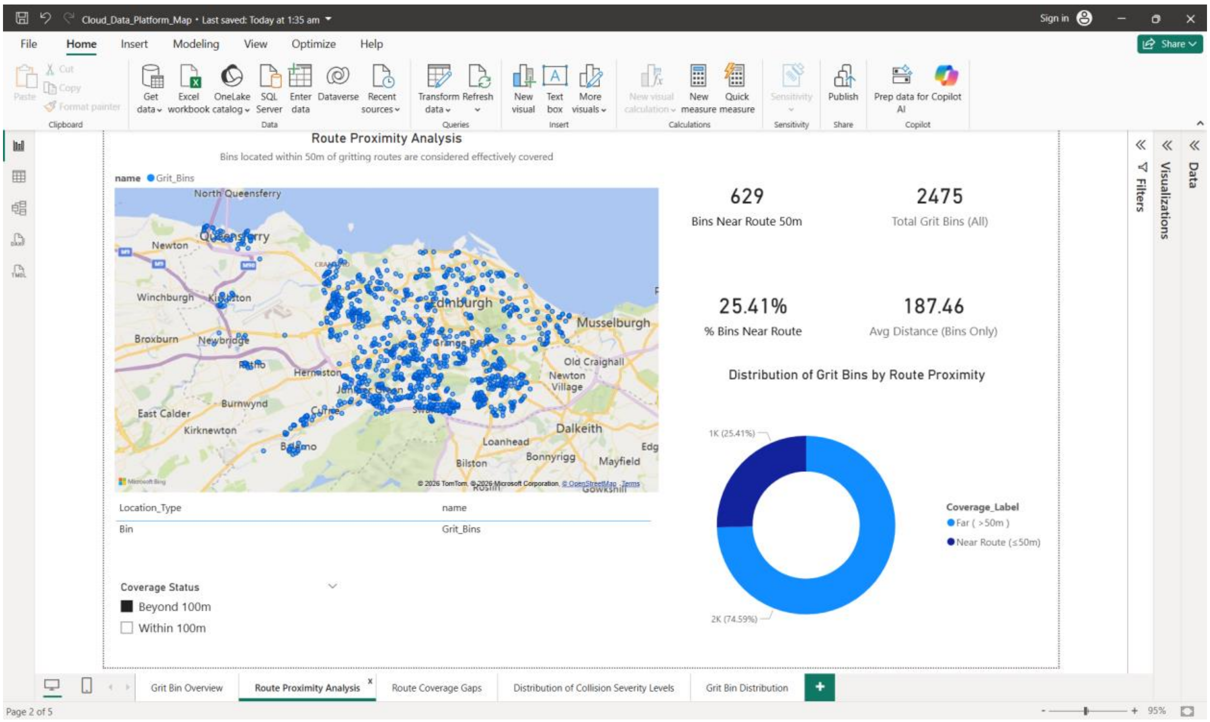
4.1 Dashboard Overview

A central map visualisation displayed gritting routes and grit bin locations, supported by summary indicators that provide immediate quantitative context. This layout enables users to move between high-level assessment and neighbourhood-level analysis.

4.2 Spatial Coverage Assessment

Grit bins are colour-coded based on proximity classification, allowing uncovered locations to be identified quickly using pre-attentive visual cues. Interactive filtering, zooming, and selection allow users to isolate specific areas or categories of interest (Few, 2012; Keim et al., 2008).

Figure 5: *Grit bins classified by proximity to the nearest gritting route*



4.3 Summary Indicators and Quantitative Context

Indicator cards display the total number of grit bins, the number beyond the distance threshold, and the percentage within coverage. These visuals provide low cognitive load summaries and respond dynamically to filters.

4.4 Independent Dataset Visualisation

The road traffic collision dataset is presented in a table visualisation, allowing sorting and filtering by attributes such as severity and location. This supports situational awareness without implying causal relationship. While no causal inference is implied, the collision dataset enables users to visually inspect whether coverage gaps coincide with historically higher incident densities, supporting situational awareness. In an operational context, this enables stakeholders to cross-reference identified service gaps against historically hazardous locations without conflating correlation with causation.

4.5 Critical Reflection on Visual Design

The dashboard enables effective exploratory spatial analysis but has limitations. Straight line distance simplifies accessibility, and the absence of temporal filtering limits seasonal analysis. Network-based or road-constrained alternatives were considered but rejected due to lost topology after GeoJSON flattening and reduced interactive performance within Power BI's in-memory engine (Van Oort, 2006).

4.6 Dashboard Structure and Page Design Rationale

Although the brief references a single-page dashboard, multiple pages were used to reduce visual overcrowding and improve interpretability. Each page addresses a distinct analytical objective while maintaining a consistent data model, interaction design, and filtering logic across the report.

5. Conclusion

This project demonstrates how semi-structured geospatial data can be transformed into an interactive analytical model using cloud-oriented business intelligence tools. By integrating GeoJSON-based gritting routes and grit bin locations within Power BI, the dashboard supports proximity-based assessment of winter service coverage and highlights potential gaps in pedestrian access to treated roads.

The inclusion of an independent road traffic collision dataset provides valuable contextual insight while maintaining analytical integrity. Within a modern cloud data platform architecture, Power BI operates as a semantic and visualization layer consuming pre-ingested object and relational data rather than acting as a primary storage service. This project reflects such an architecture by transforming exchange-optimised GeoJSON objects into an analytical star schema optimised for interactive querying. Overall, the dashboard functions as a decision support artefact rather than a descriptive map, enabling stakeholders to identify spatial service gaps and prioritise intervention under constrained operational resources.

References

de Smith, M.J., Goodchild, M.F. and Longley, P.A. (2009) *Geospatial analysis: a comprehensive guide to principles, techniques and software tools*. 3rd edn. Leicester: Matador.

Few, S. (2012) *Show me the numbers: designing tables and graphs to enlighten*. 2nd edn. Burlingame, CA: Analytics Press.

Goodchild, M.F. (2010) ‘Twenty years of progress: GIScience in 2010’, *Journal of Spatial Information Science*, 1, pp. 3–20.

Internet Engineering Task Force (IETF) (2016) *RFC 7946: The GeoJSON format*. Available at: https://www.rfc-editor.org/rfc/rfc7946 (Accessed: 08 January 2026).

Keim, D.A., Mansmann, F., Schneidewind, J., Thomas, J. and Ziegler, H. (2008) ‘Visual analytics: scope and challenges’, *Lecture Notes in Computer Science*, 4950, pp. 76–90.

Kimball, R. and Ross, M. (2013) *The data warehouse toolkit: the definitive guide to dimensional modelling*. 3rd edn. Hoboken, NJ: Wiley.

Longley, P.A., Goodchild, M.F., Maguire, D.J. and Rhind, D.W. (2015) *Geographic information science and systems*. 4th edn. Hoboken, NJ: Wiley.

Microsoft (2023) *Power BI documentation*. Available at: https://learn.microsoft.com/power-bi (Accessed: 08 January 2026).

Russo, M. and Ferrari, A. (2019) *The definitive guide to DAX: business intelligence with Microsoft Power BI, SQL Server Analysis Services, and Excel*. 2nd edn. Redmond, WA: Microsoft Press.

UK Roads Liaison Group (2016) *Well-managed highway infrastructure: a code of practice*. London: UK Roads Liaison Group.

Van Oort, P. (2006) ‘Spatial data quality: from description to application’, *International Journal of Geographical Information Science*, 20(2), pp. 133–152.

Word count (Main body text + Section headings + In-text citations):1995