

Data Engineer

Reports to: Head of Data Science and External Data

Contract: Full time, Permanent

Location: London or Manchester.

Base salary: £50 – 55K TBC

Benefits

- 30 days annual leave plus bank holidays
- Employee assistance programme
- Hybrid working (minimum two days per week in the office)
- Access to the Civil Service Pension Scheme, including Alpha (currently 28.9% employer contribution)

About Transport Focus

Transport Focus is the voice of Britain's transport users. We use powerful insight and evidence to challenge, influence, and improve the user experience across the transport sector. Key to our success is our strong stakeholder network and our ability to turn passenger and road user feedback into meaningful change.

We are an executive non-departmental public body sponsored by the Department for Transport (DfT). We are led by a board of non-executive directors appointed by DfT, the Scottish and Welsh Governments, and the London Assembly, and run by a senior leadership team based in Manchester and London.

This is a pivotal period for the organisation as we prepare to take on new responsibilities linked to the Government's railway reforms.

Role outline

Our Data Engineer will be key to the operation and ongoing development of Transport Focus's new Data Service. Initially, you will work closely with contract staff who have implemented the infrastructure and established the initial data feeds and pipelines. You will then take ownership of those feeds and pipelines, operate them reliably going forward and lead development and testing of robust, consistent, and reproducible new data models as the Data Service expands. The role will be central to ensuring that data from multiple transport modes is processed in a consistent way, enabling reliable reporting, meaningful insight, and future analytical capability, including AI-driven analysis.

What you will do

- Design and implement data models and transformation logic aligned to agreed metrics definitions and governance standards.
- Take ownership of existing data feeds and pipelines once they have been developed, ensuring they continue to operate reliably.
- Add, configure, and evaluate new data feeds, APIs and source datasets as the service expands across bus, rail, road, and potentially, other transport modes.
- Work with the Head of Data Science & External Data to prioritise new data feeds, changes to existing feeds and improvements to curated datasets.
- Develop and maintain data pipelines to ingest, transform, and prepare multi-modal transport datasets and survey data.

Data Engineer

- Ensure consistent data structures, naming conventions and transformation approaches across rail, bus, and road datasets.
- Implement validation and data quality rules to ensure accuracy, completeness, and reliability of outputs.
- Maintain clear source-to-metric traceability, ensuring that outputs can be linked back to source datasets.
- Manage routine enhancements to the data service, including new reporting requirements and improvements to existing data structures.
- Translate business-defined metrics into implementable data logic, ensuring alignment between definition and output.
- Support the development of curated and semantic data layers to enable reporting, dashboards, and insight generation.
- Ensure data processing is reproducible, traceable, and aligned to governance and audit requirements.
- Identify gaps between metrics requirements and available data, and support pragmatic, phased implementation approaches.
- Manage variability and inconsistency in source data, including differing formats, update frequencies, and levels of completeness across transport datasets.
- Apply version control and change management to data models, pipeline logic, and transformation rules so that changes to metric definitions or source data structures are controlled and documented.

What you will deliver

- Business as usual operation of existing pipelines, including monitoring, issue-resolution and change management.
- New data pipelines supporting ingestion, transformation and preparation of transport and survey datasets.
- Consistent, documented data models aligned to agreed metrics definitions and governance frameworks.
- Validated and quality-assured data outputs supporting reporting and performance management.
- Curated data layers structured to support reporting tools, dashboards, and insight products.
- Data processing logic that is reproducible, traceable, and aligned to governance and audit requirements
- Documented transformation logic, data models, and processing rules to ensure resilience and ongoing maintenance.
- A consistent engineering approach applied across all datasets, enabling scalable and maintainable delivery.

Who you will work with

- **Data Programme Lead** (a temporary role): alignment to metrics definitions, governance frameworks, and business requirements
- **DevOps Platform Engineer** (a temporary role): alignment to platform architecture, pipelines, and deployment processes

Data Engineer

- **Transport Data Engineer** (a temporary role): responsible for initial implementation of the data feeds and pipelines you will be responsible for
- **Head of Data Science and External Data + Data Analyst (permanent roles):** together, responsible for provision of curated, dependable, and well-documented datasets to support reporting, analysis, modelling, and future AI-enabled insight.
- **Other Transport Focus stakeholders:** alignment to reporting requirements, usability, and interpretation of outputs.

Tech stack

- Cloud Platform: Microsoft Azure
- Logging platform: Graphana
- Primary Database: PostgreSQL
- Data Processing / ETL: Go (Golang)
- Version Control: GitHub/Azure git
- Future Web/API Development: .NET (ASP.NET Core)

As the platform evolves, we expect additional technologies to be introduced where appropriate, potentially including infrastructure-as-code, orchestration, and analytics tooling, together with logging tools (Graphana and Loki).

Skills and Experience

Essential:

- The ability to take over and maintain data pipelines developed by others.
- Designing and implementing automated data transformation pipelines and data model.
- Be comfortable designing solutions that operate effectively at scale and can accommodate long-term data growth.
- Processing large, complex datasets in a scalable and efficient manner
- Defining schemas, data structures, and transformation logic
- Implementing data validation, quality rules, and error handling
- Working with structured and semi-structured data from multiple sources
- Debugging and optimising complex data processing workflows
- A strong understanding of how data modelling supports metrics definition, reporting and insight generation.
- The ability to translate business requirements into robust, implementable data logic.
- Experience designing reproducible and traceable data processing approaches.
- Awareness of how data structures support downstream analytics and AI-driven use cases.
- The ability to document data models, transformation logic, and processing rules clearly.
- The ability to apply pragmatic judgement where data is incomplete or inconsistent, ensuring outputs remain usable and clearly understood.
- SQL, Python and Cloud-based data tooling proficiency
- Experience with modern cloud-native data engineering tooling, including Infrastructure as Code (Terraform), data transformation frameworks (e.g. dbt), (CI/CD pipelines would be advantageous.)

Desirable:

Data Engineer

- Experience working with transport data (e.g. outputs from the Department for Transport's Bus Open Data Service, the Rail Delivery Group's Darwin, Network Rail's TRUST, and National Highways' Datex II), or comparable transport datasets, is desirable.

Our commitment to equality, diversity, and inclusion

Transport Focus is committed to representing all transport users. Our team members will play a large part in inspiring the organisation to think differently in terms of approach to this commitment. We must show that our work is representative, relevant, and useful. Diversity and inclusion is key to being able to do this effectively. We believe that everyone in society deserves the same chances in life. Having a diverse and inclusive organisation means we can tap into different perspectives and experiences to generate new ideas and initiatives, which will make our work more relevant and useful. We want Transport Focus to be a place for everybody and anybody to work. We aim to build teams that draw upon the widest range of experiences, talent, and ideas, and for that reason welcome applications from the broadest range of backgrounds and communities.