

Contents

Background Information.....	2
What is the reference period for the estimates of station usage?	2
When will the next dataset be released?	2
Where can historic estimates of station usage data be found?.....	3
Why are there no estimates of station usage for 2003-04?	3
Where can data prior to 1997-98 be found?	3
Data Scope and Methodology.....	4
How is the data for the estimates of station usage collected?	4
The data for a particular station is not correct	4
Why does it take so long to produce the estimates of station usage when ticket sales data is available overnight?	5
What changes have been made to the methodology this year?	5
What change was made to the methodology for demand at London Stations in 2015-16?.....	6
Why did the London methodology change in 2015-16 have an effect on stations outside of the London Travelcard Area, such as Birmingham New Street?	7
How is usage of group stations estimated?	7
Why are the entries and exits recorded as identical in previous years?.....	7
Why are journeys for certain operators excluded from the data?.....	8
What ticket types are included in the data and which category do these tickets belong in?	8
How are sales of different ticket types converted into entries and exits?	8
Does the data include people who use the station but do not travel?	8
Does the data include those people who travel on the train without purchasing a ticket?	9
Why is an explanation of the change in station usage not been included for every station?	9
Can a further breakdown of the estimates of station usage be produced?	9
Which geographical boundary definitions have been used?.....	10
Is it possible to provide an alternative geographic breakdown?	10
Why are there some stations with no data?.....	10
Why are there some stations missing when the estimates of station usage has data for all stations?	10
Using the Estimates of Station Usage Dataset	11
Where are the columns in the dataset defined?	11
Is it possible to search for a specific station?	11
Is it possible to list the stations in alphabetical order?	11
Related Statistics.....	11
Where can information on station footfall be found?.....	11
Where can information about station car park usage be found?.....	11

Background Information

The latest in the series of **Estimates of Station Usage**, for all stations in Great Britain, was published by **Office of Rail and Road (ORR)** on 1 December 2017. This is available on the ORR website at www.orr.gov.uk/statistics/published-stats/station-usage-estimates

The **estimates of station usage** consist of estimates of the total numbers of people:

- Travelling from or to the station (**entries and exits**); and
- Interchanging at the station (**interchanges**)

Additional information includes geographical location, ticket type and changes from the previous year.

Alongside these frequently asked questions, a number of documents are available on the ORR website, including:

- Estimates of Station Usage 2016-17 - Key Facts (PDF)
- Estimates of Station Usage 2016-17 (Excel)
- Time Series of Estimates of Station Usage 1997-98 to 2016-17 (Excel)
- Methodology Report for the Estimates of Station Usage 2016-17 (PDF)
- Infographics (PDF)

If you have any further enquiries or feedback about the Estimates of Station Usage, the ORR Information and Analysis Team can be contacted at rail.stats@orr.qsi.gov.uk

What is the reference period for the estimates of station usage?

The estimates of station usage are based on the financial year. Each year runs from 1 April to 31 March.

When will the next dataset be released?

ORR aims to publish the estimates of station usage as soon as possible after the end of the data reference period. It is important to take time to implement the detailed methodology and carry out a series of quality assurance checks to ensure the final data are as accurate as possible.

Publication of estimates of station usage 2017-18 is likely to be in December 2018. For the most up to date information, please refer to the [ORR publication schedule](#).

Where can historic estimates of station usage data be found?

Estimates of station usage are available on an annual basis from 1997-98 onwards. These are published on the [ORR website](#).

From 2013-14, we have also published a collated dataset giving a time series of estimates of station usage from 1997-98 to the most recent data. This is also available on the [ORR website](#).

Why are there no estimates of station usage for 2003-04?

As a result of improvements made to the methodology, the estimates of station usage for 2003-04 were not produced. For further information about the changes made, please see the methodology report for the 2004-05 release.

Where can data prior to 1997-98 be found?

ORR does not hold any estimates of station usage data prior to 1997-98.

The [National Archives](#) or the [National Railway Museum](#), which holds railway documentation from pre-privatisation, may be able to provide further assistance.

Data Scope and Methodology

For more detailed information about the methodology used to produce the estimates of station usage, please refer to the methodology report, which can be found on the [ORR website](#).

How is the data for the estimates of station usage collected?

As Britain does not have a fully gated rail network, a complete recording of passenger flows through stations is not possible. As a result, use of stations has to be estimated and there are some limitations on the data.

The estimates of station usage dataset is derived from the Origin Destination Matrix (ODM), a comprehensive matrix of passenger flows between stations throughout Great Britain.

The ODM is largely based on data produced for the MOIRA2.2 rail planning tool, which is produced for the rail industry by Resonate. MOIRA2.2 is a matrix that provides an estimate of journeys in the rail network in Great Britain for the duration of a financial year (1 April to 31 March). It is predominantly based on information from LENNON, the rail industry's ticketing and revenue system. MOIRA2.2 includes all journeys associated with point to point flows and includes additional information to reflect travel on other products, such as London Travelcards and multi-zonal tickets in other major urban areas.

The production of the ODM requires some further adjustments to address known issues with the MOIRA2.2 data. These include an allocation of tickets sold to 'London Terminals,' allocation of demand between individual stations in group stations outside of London, such as Manchester group stations and a number of cases where adjustments are made to selected stations to account for specific known issues, for example Digby & Sowton.

A complete overview of the methodology can be found in the Executive Summary of the methodology report, with more detailed information in Chapter 2. This can be found on the [ORR website](#).

The data for a particular station is not correct

It is important to emphasise that these data are estimates of station usage.

As Britain does not have a fully gated railway system, a complete recording of passenger flows through stations is not possible. In the absence of comprehensive and robust count data, station usage is estimated using ticket sales.

More information on the limitations of the data and details of the methodology can be found in the methodology report on the [ORR website](#).

Why does it take so long to produce the estimates of station usage when ticket sales data is available overnight?

Although ticket sales data is available from LENNON overnight, the estimates of station usage data is predominantly based on the rail industry planning tool, MOIRA2.2. This data is not available until around three months after the end of the financial year. There are also a number of processes and additional datasets that contribute towards the final dataset. There is also an extensive quality assurance process.

More detailed information about the way the estimates of station usage are constructed can be found in the methodology report on the [ORR website](#).

What changes have been made to the methodology this year?

Improvements to the methodology are made in most years. This is in response to user feedback suggesting that users would prefer the data to be as accurate as possible, rather than maintain absolute consistency in the time series.

As a result, changes to the methodology mean that direct year-on-year comparisons for the stations affected are not valid.

The improvements to the methodology for the Estimates of Station Usage 2016-17, and a summary of the change, are as follows:

London Group Stations Allocation Update

An adjustment has been made to the methodology used to allocate journeys made with an origin and destination both recorded as London Group Stations. The previous methodology was overstating use at Kensington Olympia station.

To resolve this, journeys with an origin and destination stated as London Group Stations have been allocated to individual London Terminal stations using the underlying MOIRA2.2 matrix.

The only noticeable effect of this methodology change is at Kensington Olympia station, resulting in a decrease of usage. As overall usage at the other stations affected is high, the impact of the change of methodology is not noticeable.

Season Ticket Journey Allocation

The methodology used to allocation journeys made using season tickets from or to certain stations has been updated to reflect demand in 2016-17.

During the production of the Estimates of Station Usage 2014-15, an adjustment to the allocation of usage at stations around Southend was made. Analysis of LENNON data showed that season tickets issued for travel to or from Southend Victoria were actually being used to travel from alternative stations, as the price of a season ticket is the same.

Through consultation with train operating companies (TOCs) and analysis of LENNON sales data a number of other locations where it is thought this is occurring were identified and the same adjustment has been applied from 2015-16. The full list of stations this adjustment effects can be found in the methodology document.

For more information about these changes and the impact of improving the methodology, please refer to the methodology report on the [ORR website](#).

What change was made to the methodology for demand at London Stations in 2015-16?

There were two changes to the methodology that have had a significant effect on the usage at stations in London in 2015-16:

- *In-boundary London Travelcard Methodology*

In previous years, London Travelcard journeys had been allocation to stations using London Area Travelcard Survey (LATS) data from 2001. For the production of the MOIRA2.2 dataset, Oyster Clicks Model (OCM) data from Transport for London (TfL) has been used to allocate journeys made wholly within the London Travelcard Area to individual London stations.

This has in general re-allocated some journeys that would have been to London Terminal stations to stations outside of Zone 1.

- *London Terminals Demand Allocation*

For the 2015-16 data onwards, the MOIRA2.2 input data has been disaggregated by individual London Terminal where possible, such as where a ticket is bought to a specific terminal rather than to the generic 'London Terminals.' This provides an improved reflection of journey origins and destinations.

As a result of the changes made to the estimates of London station usage, direct comparisons between entries, exits and interchanges for 2015-16 and previous years are not valid. An estimate of the effect of the methodology change has been included in the 2015-16 dataset to help users identify where the methodology change is affecting results.

Why did the London methodology change in 2015-16 have an effect on stations outside of the London Travelcard Area, such as Birmingham New Street?

Some small changes to the estimates of station usage have occurred to stations outside of the London Travelcard area because of the complex series of interlinked methods used to compile these statistics.

These changes are small and have no material effect on the estimates for the stations where this has occurred.

How is usage of group stations estimated?

Group stations are a small number of stations that are treated as a group for an origin or destination, rather than individual stations. For example, where a ticket identifies the origin or destination as Manchester BR, passengers could use Manchester Piccadilly, Manchester Victoria, Manchester Oxford Road, Salford Central or Deansgate.

Current industry data does not distinguish between the component stations and therefore a split of entries and exits between these stations has been estimated during the production of the origin destination matrix.

More information on the methodology used to estimate usage at group stations can be found in the methodology report on the [ORR website](#).

Why are the entries and exits recorded as identical in previous years?

Entries and exits have been recorded as identical since a change to the methodology in 2008-09.

This is a direct result of integrating the rail industry's principle planning tool, MOIRA, into the production process for the estimates of station usage. This brought substantial benefits as MOIRA includes an estimate of revenue and journeys made using zonal products sold by PTEs and provides a more complete representation of travel on the national rail network. The previous exclusion of these journeys was a substantial deficiency of the estimates of station usage.

However, MOIRA does not disaggregate single journeys. Therefore, in estimating passenger journeys, all ticket sales have been split equally between the two directions of travel.

The split between entries and exits is not shown the 2016-17 dataset.

Why are journeys for certain operators excluded from the data?

Sales for Heathrow Express and Eurostar services are not included in the rail industry systems. Therefore travel using tickets from these operators will not be included in the estimates of station usage.

Consequently, the usage at stations as St. Pancras, Ashford International and Ebbsfleet will not be a true reflection of the total usage at these stations.

What ticket types are included in the data and which category do these tickets belong in?

Journeys and revenue are sub-divided into the following four ticket types in the base demand matrices:

- Full – all walk-up undiscounted single or return tickets, whether or not issued with a status discount, such as a child fair or with a railcard.
- Reduced – all walk-up discounted single or return tickets, whether or not issued with a status discount
- Advance – all advance-purchase tickets
- Seasons – all multi-use tickets

These four ticket types are further divided by First and Standard Class.

How are sales of different ticket types converted into entries and exits?

Ticket transactions are converted into an estimate of the number of journeys made by applying a series of ticket type journey factors.

Single and return tickets unambiguously translate into one and two journeys respectively. The number of journeys made using season tickets is estimated using a historic estimate of the monthly use of a season ticket. Ticket periods of other lengths are converted into a number of journeys using a proportion of the monthly journey factor.

More detailed information can be found in the methodology report on the [ORR website](#).

Does the data include people who use the station but do not travel?

This is defined as 'station footfall' and includes those individuals entering the station without the intention of travelling on the rail network. This includes use of shopping or restaurant facilities.

As the estimates of station usage are based on tickets sales, the data is representative of those individuals that are intending to travel on the rail network. Therefore, station footfall is not included in the data.

Network Rail publishes information on [station footfall at Network Rail managed stations](#) only.

Does the data include those people who travel on the train without purchasing a ticket?

This is defined as 'ticketless travel'

As the estimates of station usage are based on ticket sales, journeys associated with ticketless travel are not included in the data. This is more likely to be an issue on some flows and where ticketless travel is significant. As more stations have been gated over time and train operating companies (TOCs) focus on revenue protection activities, this is likely to be less of an issue than in the past.

There is a strong argument that it is not appropriate to include ticketless travel in the dataset as its purpose is to record genuine journeys on the rail network. The inclusion of ticketless travel could distort business cases for new investment where these are reliant on the estimates of station usage data.

It is worth noting that ticketless travel also includes an element of individuals who are legitimately travelling for free, such as the British Transport Police or rail industry employees.

Why is an explanation of the change in station usage not been included for every station?

Where there have been large changes in usage (increase or decrease) at particular stations ORR has tried to add contextual information to provide explanation for these changes, including where any methodological enhancements have had an effect. However, as there are a number of factors that will affect usage/demand at each station, it is not always possible to identify clear reasons for changes. Therefore the explanations provided for changes should not be seen as comprehensive, and may not include other information which some might consider to have driven changes in station usage at particular stations.

Can a further breakdown of the estimates of station usage be produced?

These estimates are annual figures. It is not possible to produce a further temporal breakdown, such as by week or time of day. The database only provides details on when a ticket was purchased rather than when the passenger actually travelled. Therefore, journeys cannot be assigned to specific days or times.

It is not possible to provide a breakdown of the data by traveller type, such as business, leisure or commuter, as there is no way of identifying this information within the data. Some information on the reasons for travel by rail can be found in the [National Travel Survey](#), which is published by the Department for Transport.

Which geographical boundary definitions have been used?

In 2016-17, three alternative geographical definitions have been included in for each station: Region, Local Authority District and Westminster Parliamentary Constituency.

The boundary definitions are as at December 2015.

More information on these boundary definitions, as well as the boundary files, can be found on the [Office for National Statistics Open Geography Portal](#).

Is it possible to provide an alternative geographic breakdown?

ORR is not able to provide bespoke geographical breakdowns of the estimates of station usage data.

However, as both the local authority and the coordinates (Ordnance Survey Northings and Eastings) of each station have been included within the dataset, it is possible to create an alternative geographical definition, using either a lookup file or a Geographical Information System (GIS), such as [QGIS](#) or [ESRI ArcGIS](#).

More information on geographical definitions in the United Kingdom, including look up files and boundary files to download, is available on the [Office for National Statistics Open Geography Portal](#).

Why are there some stations with no data?

In general, a station is recorded in the estimates of station usage if there is a record of ticket sales with the station as an origin or destination. However, it is known that there are some national rail stations for which sales are not recorded in the sales database, such as Heathrow stations, or the station is only used on a very limited number of days, such as Manchester United Football Ground.

To be consistent with the station count in the 2016-17 publication [Rail infrastructure, assets and environmental statistical release](#), these stations have been included in the station list but have not been populated with any data.

Why are there some stations missing when the estimates of station usage has data for all stations?

The estimates of station usage dataset includes all stations served by national rail services in 2016-17. Data for any stations opened after 31 March 2017 will be included in the next publication.

If you require information for a non-national rail station, please contact the relevant station operator or owner, who may be able to provide further assistance.

Using the Estimates of Station Usage Dataset

Where are the columns in the dataset defined?

A full description of the column headings can be found in the 'Column Descriptions' tab of the estimates of station usage spreadsheet.

For more information about these fields, please refer to the methodology report that can be found on the [ORR website](#).

Is it possible to search for a specific station?

For information on how to search for a specific station, or how to use filters, please refer to the 'How to filter' tab of the estimates of station usage spreadsheet.

Is it possible to list the stations in alphabetical order?

For information on how to sort the data by alphabetical order, or by any other column, please refer to the 'How to filter' tab of the estimates of station usage spreadsheet.

Related Statistics

Where can information on station footfall be found?

Network Rail publishes information on [station footfall at Network Rail managed stations](#) only.

Where can information about station car park usage be found?

For information about station car park usage, please contact the relevant train operating company.