



# Passenger Rail Usage: Quality and Methodology Report

Release Date: June 2018

# Contents



<b>Introduction</b>	<b>3</b>
<b>Methodology</b>	<b>4</b>
<b>Historical background</b>	<b>10</b>
<b>Relevance</b>	<b>12</b>
<b>Accuracy and reliability</b>	<b>13</b>
<b>Timeliness and punctuality</b>	<b>15</b>
<b>Accessibility and clarity</b>	<b>16</b>
<b>Coherence and comparability</b>	<b>17</b>

# Introduction

This is a report on the quality of the Passenger Rail Usage statistical release and data portal tables. It helps users to understand the quality of our statistics, and also ensures ORR is compliant with principle 4 of the Code of Practice for Official Statistics<sup>1</sup>.

The quality report covers the following areas:

- **Methodology** – detail on the various data sources and methodology used to compile the statistics;
- **Historic background** – a background to each statistic and details of changes throughout the time series;
- **Relevance of the data** – the users of the statistics, and our engagement;
- **Accuracy and reliability** – the accuracy of each statistic;
- **Timeliness and punctuality** – our timescales for the production, quality assurance and publication of each statistic;
- **Accessibility and clarity** – the format of our statistics and where they can be found;
- **Coherence and comparability** – comparisons to similar statistics published elsewhere.

Rail usage data on passenger services are supplied by a number of different sources and are measured using a range of metrics: passenger kilometres, journeys and revenue, and passenger train kilometres. These are covered in detail in the methodology section.

ORR publishes quality reports on other ORR statistics which can be found on the [statistical release](#) page.

---

<sup>1</sup> Principle 4: Sound methods and assured quality. Statistical methods should be consistent with scientific principles and internationally recognised best practices, and be fully documented. Quality should be monitored and assured taking account of internationally agreed practices. The Code of Practice can be accessed at: <http://www.statisticsauthority.gov.uk/assessment/code-of-practice/index.html>

# Methodology

## Passenger kilometres

- **Passenger kilometres** – the number of kilometres travelled by passengers on the network.

The data used to derive passenger kilometres are sourced from the LENNON (Latest Earnings Network Nationally Over Night) ticketing and revenue database and train operating companies (TOCs). Passenger kilometres are calculated by multiplying the number of passengers on a particular flow by the number of track kilometres between the two required stations. The track kilometres data are built into the LENNON system.

## Passenger journeys

- **Passenger journeys** – the number of passenger journeys made on the network.

The data used to derive passenger journeys are sourced from the LENNON database and TOCs. A passenger journey is defined as the number of legs required to get from origin to destination. For example, if your ticket is valid between London and Halifax and you change trains at Leeds, this would be classed as two journeys, one for the London to Leeds leg and one for the Leeds to Halifax leg.

## Passenger revenue

- **Passenger revenue** – the volume of passenger revenue includes all ticket revenue and miscellaneous charges associated with passenger travel on national railways, but not including government support or grants.

The data used to derive passenger revenue are sourced from the LENNON database. Passenger kilometres, journeys and revenue data are used to produce revenue per passenger kilometre and revenue per passenger journey data.

LENNON holds information on the vast majority of national rail tickets purchased in Great Britain and is used to allocate the revenue from ticket sales between TOCs. LENNON contains two datasets; pre-allocation (sales) and post-allocation (earnings). Passenger usage statistics in the statistical release are based on the post-allocation dataset so that kilometres, journeys and revenue data can be assigned to TOCs (for use in the NRT TOC key statistics). The pre-allocated dataset does not disaggregate data by TOCs. The pre-allocate dataset collects total kilometres, journeys and revenue by flow and then, based on pre-designated allocation factors, apportions the data to appropriate the TOCs. For example in the pre-allocated dataset, a passenger may

purchase a ticket for “ANY PERMITTED” route between London Terminals and Birmingham BR. LENNON then uses the allocation factors to split kilometres, journeys, and revenue between the relevant TOCs for which the journey could be made.

Allocations are created for each tickets group by ORCATS (Operational Research Computerised Allocation of Tickets to Services), dependent on sales levels. These allocations are principally used to apportion kilometres, journeys and revenue between TOCs.

ORCATS is a mathematical model which uses a similar logic to journey planning systems and identifies passenger ‘opportunities to travel’ from an origin station to a destination station using timetable information. It is used for real time reservation and revenue sharing on inter-available tickets between TOCs which divides ticket revenue in instances where a ticket or journey is on a flow that is operated by multiple TOCs. More information on the development of the LENNON database can be found in the [Origin – Destination Matrix 2016/17 Summary Report](#) which can be accessed via the [statistical release](#) page, under Regional Rail Usage.

In addition to the LENNON data, TOCs also provide data to ORR detailing the number of passenger kilometres and journeys that are, recorded outside of the LENNON system. These include kilometres and journeys on tickets such as operator specific tickets and PTE multi-modal tickets. These are referred to as non-LENNON data.

## Passenger train kilometres

- **Passenger train kilometres** - the actual mileage in kilometres travelled by revenue earning passenger trains on the Network Rail infrastructure.

The passenger train kilometres are derived from Network Rail’s Track Access Billing System (TABS), which Network Rail use to bill train operators. TABS captures the actual train distance operated on Network Rail’s track.

Through its data portal ORR publishes quarterly and annual the passenger train kilometres data for both franchised and non franchised operators.

It is to be noted that Heathrow Express is not charged through TABS and hence the total figure for non franchised operators do not include Heathrow Express.

*Note: From 2015-16 Q3 we have replaced timetabled train kilometres (TTKM) with passenger train kilometres.*

## Sector definitions

The rail network is subdivided into three sectors – Long Distance, London and South East and Regional. At the time of privatization these three sectors were based on the franchise map and how the franchises fitted within the three British Rail sectors.

- **Long distance** – based on the British Rail InterCity services which were long-haul express services;
- **London and South East sector** – based on the British Rail Network South East services which principally operated commuter trains in the London area and the inter-urban services in South East England, although the network reached as far west as Exeter;
- **Regional sector** – based on the British Rail Regional services which were all other services not included in the other two sectors. In some cases, for example in the Network Rail monitor, services in Scotland (ScotRail only and not cross border services) are treated separately from regional services. For the statistics published in the themed statistical releases services in Scotland are defined as regional services.

Following the refranchising of some services in 2006 and 2007 a number of train operating companies operate services in more than one of the sectors. The measures of rail usage are assigned to sectors based on the service code of a train. For example all passenger journeys and kilometres on a long distance service from London to Scotland are assigned to the long distance sector, regardless of whether passengers embark or disembark from the train at an intermediate station. The current mapping of services to within each sector is listed below.

### *Long distance*

- CrossCountry;
- East Midlands Trains (services to and from London St Pancras);
- Great Western Railway (High speed services);
- Virgin Trains East Coast;
- Virgin Trains West Coast;

### *London and South East*

- Greater Anglia;
- c2c;
- Chiltern Railways;
- Great Western Railway (London and Thames Valley commuter services);
- Govia Thameslink Railway;
- London Overground;
- Southeastern;
- South Western Railway;
- TfL Rail;
- West Midlands Trains (Euston, St Albans and Bletchley services).

### *Regional*

- Arriva Train Wales;
- Caledonian Sleeper;
- East Midlands Trains (services in the East Midlands and in Central and Northern England);
- Great Western Railway (services in the west of England);
- Merseyrail;
- Northern;
- ScotRail;
- TransPennine Express (rail usage statistics only);
- West Midlands Trains (Other services).

## Franchised / Non-franchised operators

Franchised operators are those which operate under the terms of franchises let by Department for Transport. A franchised train operator is the successful winner of a competitive tender process that sets out the train services required to be operated (service level agreement). A franchised operator must operate within the parameters set out in their franchise and they have a fixed life (although this can be extended and a train operator can win successive terms).

- Greater Anglia
- Arriva Train Wales;
- Caledonian Sleeper;
- c2c;
- Chiltern Railways;
- CrossCountry;
- East Midlands Trains;
- Govia Thameslink Railway;
- Great Western Railway;
- London Overground;
- Merseyrail;
- Northern;
- ScotRail;
- Southeastern;
- South Western Railway;
- TfL Rail;
- TransPennine Express ;
- Virgin Trains East Coast;



- Virgin Trains West Coast;
- West Midlands Trains.

Non-franchised operators hold licenses to provide supplementary services on chosen routes. A non-franchised operator operates because they choose to and the law allows them to have open access to the railway network, subject to the approval of the ORR. These operators are usually given approval by the ORR to operate for a definite period, though it can be renewed. The current non-franchised operators are<sup>2</sup>:

- Hull Trains;
- Grand Central;
- Heathrow Express.

All freight operators are open access operators.

## Rail usage data

Data within the LENNON ticketing and revenue database was changed following refranchising in 2006 and 2007 as the Rail Delivery Group (RDG) wished to keep a consistent time series. However, because ORR is able to extract data at a route level the data for multi-sector TOC's can be assigned to the relevant sectors. There is a difference between the rail usage and performance data as TransPennine Express services are classed as regional services with rail usage data.

---

<sup>2</sup> Wrexham and Shropshire ceased operation during January 2011

# Historical background

## Passenger train kilometres

Network Rail bills train operators (both passenger and freight) for all train movements on their infrastructure. Billing was previously done by separate systems for passenger and freight operators (PABS & BIFS respectively). From 1 April 2009, billing has been done by a single system, known as TABS (Track Access Billing System).

## Passenger Kilometres, Passenger Journeys and Passenger Revenue

Historically (1986/87 to 2002/03) passenger usage data was calculated from CAPRI (Computer Analysis of Passenger Revenue Information) which was the rail industry's former central ticketing system. Data for this period can now be found in the ORR [data portal](#) as well as in the archived Strategic Rail Authority (SRA) National Rail Trends publications available from [Archived National Rail Trends](#). CAPRI was, however, unable to correctly record sales of certain products, including some operator-specific tickets and Passenger Transport Executive (PTE) multi-modal tickets. Following a review by SRA and with significant assistance of TOCs robust estimates for these products were included in the passenger kilometres and journey data from the beginning of 2001-02, although they were backdated to the beginning of 1999-00. Passenger revenue data were not included in these estimates and as such the data remained unaffected.

Since 2003/04 the rail industry's central ticketing and revenue system, LENNON, is the basis for passenger kilometres, journeys and revenue data.

Since 2010-11 Q1, non-Lennon data have been provided to ORR each quarter by each of the TOCs. Merseyrail provide their non-Lennon passenger kilometres and journeys data every 6 months during quarter 2 and quarter 4. Hence their passenger kilometres and journeys data for quarters 1 and 3 are estimated by applying the percentage change on the Regional kilometres and journeys data captured by LENNON between the current quarter and the same quarter in the previous year. For example, if the Regional journeys increased by 2.85% in 2015-16Q1 compared to 2014-15 Q1, non-Lennon journeys for Merseyrail in 2015-16 Q1 would be an additional 2.85% of non-Lennon journeys recorded for Merseyrail in 2014-15Q1. The same method is applied to estimate the passenger kilometres for Merseyrail in quarters 1 and 3. The data are hence provisional for quarters 1 and 3 and are updated in the following quarters when Merseyrail provide their data for the 6 months.

London Overground provide their quarters' passenger kilometres and journeys data captured from the Train Load weight system, which is their total passenger kilometres and journeys data for the quarter. Hence their non-Lennon usage data is derived by deducting their data captured by LENNON from the total passenger kilometres and journeys figures provided for the quarter.

Prior to 2010-11, non-Lennon data were provided annually to ORR at the end of the financial year and were apportioned out to each financial quarter. The apportionment of the additional kilometres and journeys data to each quarter was based on the split of LENNON data. For example, if 26% of LENNON journeys were in Q1, 26% of the additional non-Lennon journeys would be assigned to Q1.

Data for non-franchised operators were not collected prior to 2008-09 Q3. Since this time LENNON data for non-franchised operators includes Hull Trains and Grand Central but not Heathrow Express who are not included within the LENNON database. Similarly Heathrow Express do not supply ORR with non-Lennon data. Between 2008-09 Q4 and 2010-11 Q4, the data also included Wrexham and Shropshire but they ceased trading on 28 January 2011.

Since 2012-13 Q2 passenger kilometres, journeys and revenue data by ticket type for ordinary fares have been disaggregated further based on Advance, Anytime (Peak), Off Peak (including Super Off Peak) and Other (non-Lennon tickets and promotional tickets). It is not possible to provide a time series prior to 2011-12 for this level of disaggregation because LENNON only maps the data back to 2011-12 Q1.

# Relevance

*The degree to which the statistical product meets the user in both coverage and content*

Measures of rail usage are key indicators of the levels of rail use in Great Britain. They provide a clear indication of the number of passengers using and journeys made on the network, providing an indication of the levels of demand for rail travel. This can help in both short-term and long-term planning for the industry and wider stakeholders, both at a National level and within the rail sectors.

Passenger revenue data provides data an insight into revenue levels within the industry as well as the levels of revenue generated through each ticket type, which can provide an indication of changes in ticket purchasing trends.

Performance data published on our data portal is used by a range of individuals for planning, analysis, decision making and data validation.

More detailed information on users of ORR statistics and meeting the needs of users is available on our [user engagement webpage](#).

# Accuracy and reliability

*The proximity between an estimate and the unknown true value.*

## Passenger kilometres, journeys and revenue

Passenger kilometre, journeys and revenue data are sourced directly from LENNON (LENNON data) and train operating companies (non-LENNON data). The LENNON system automatically records the majority of ticket sales with non-LENNON data capturing the tickets sold through outlets not linked to the LENNON database. Combining the two sets of data provides a best estimate of passenger ticket sales from which passenger kilometres, journeys and revenue can be derived. The LENNON database is primarily an accounting tool and therefore faces limitations when being used for statistical reporting. With all large data sources there may be input errors which are more likely to occur in the journeys, rather than revenue data. Due to the size and complexity of the dataset we are unable to validate each and every entry.

Known problems in the data capture include those relating to travelcards, return and single journey tickets, multiple tickets, rail staff passes, ticketless travel and other rail systems. Full details of these limitations can be found in Annex B of the [Origin Destination Matrix 2016/17 Summary Report](#).

It must also be noted that revenue, in addition to ticket revenue, includes other miscellaneous charges associated with passenger travel on national railways, such as car parking charges. There may be differences between the actual values and published statistics resulting from tickets involving travel on London transport, the receipts for which are apportioned.

The passenger kilometres and journeys statistics, which are based on the LENNON data and non-LENNON data received from TOCs, may also differ from actual passenger kilometres and passenger journeys because the data does not make adjustment for unused tickets and passengers cutting journeys short (i.e. alighting a train before the destination station stipulated on their ticket).

Whilst the LENNON database is updated overnight there may be circumstances in which passenger revenue figures would be revised and therefore change the revenue figures in LENNON. These are likely to occur within the first two weeks after the end of each period. Therefore to minimise the risk using incorrect data, ORR waits at least one month before extracting the data from LENNON.

It should also be noted that rail usage sector data do not align with rail performance sector data in some cases. This is because for rail usage, each TOC within LENNON is assigned to one

particular sector whilst in rail performance, sectors are assigned based on the route code and not TOC. This ensures a consistent time series with data prior to re-franchising.

The passenger journey totals should not be compared with those published as part of ORR's Regional Rail Usage release as they are calculated on a different basis. For example, a journey from London to Halifax, which may involve two trains (one from London to Leeds and another from Leeds to Halifax), would be treated as two journeys in the rail usage passenger journeys total as it would involve two different services. However, in Regional Rail Usage, it is treated as one journey reflecting the origin and destination shown on the ticket. Please see [Passenger journeys in Great Britain](#) which explains the differences in more detail.

Data extracted from the LENNON database is converted into a standard format and subject to a series of quality assurance checks before publication. Similarly all non-LENNON data that are received from TOCs are subject to a series of quality assurance checks before publication. We check the data is provided in the correct format, there are no inconsistencies in the data and trends over time are similar, to ensure accurate data is published.

## Passenger train kilometres

Actual train distance operated is captured on Network Rail's Track Access Billing System (TABS). Actual train distances are also published by Network rail as passenger mileage and freight in the Network Rail Annual Return.

Any arising issues are flagged with NR (for Passenger train kilometres) or the LENNON support helpdesk (for passenger kilometres, journeys and revenue) who must confirm the anomalies or correct the data and re-submit. Explanations from the data suppliers regarding data anomalies are included within our commentary to explain the data and trends.

These data are then prepared for publication. The process includes quality assuring the tables and charts produced and providing supporting commentary regarding the key trends, methodology and quality measures. These reports are subject to peer review.

The final stage of the quality assurance process is a sign off by the statistics Head of Profession confirming the data meets the quality standards and are fit for publication.

## Timeliness and punctuality

*Timeliness refers to the time gap between publication and the reference period. Punctuality refers to the gap between planned and actual publication dates.*

The data contained within the passenger rail usage statistical release are published quarterly on the ORR data portal approximately 60 days after the end of the quarter.

ORR has memorandums of understanding (MoUs) with data suppliers detailing the scope and timeliness of each dataset supplied. This ensures consistent and timely data is received each period. The MoUs are reviewed on an annual basis.

More detailed information on timeliness and effectiveness of the statistical output is available on our [user engagement webpage](#).

## Accessibility and clarity

***Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information. Clarity refers to the quality and sufficiency of the metadata, illustrations and accompanying advice.***

All passenger rail usage data can be accessed on the [Data Portal](#) free of charge.

Department for Transport (DfT) also publish rail usage data, based on data published by ORR, on their website and can be accessed at [DfT Rail usage, infrastructure and performance](#).

Passenger rail usage data, particularly passenger revenue data, is viewed as commercially sensitive. Due to this commercial sensitivity, lower levels of disaggregation are not possible; however train operating companies and RDG do occasionally publish other comparable data on a one-off basis.

***The procedures and policy used to ensure sound confidentiality, security and transparent practices.***

ORR is fully compliant with the Statistics and Registration Service Act 2008 and principle 4 of the Code of Practice for Official Statistics.

More information is available on our [user engagement webpage](#).



# Coherence and comparability

*Coherence is the degree to which data that are derived from different sources or methods, but refer to the same topic, are similar. Comparability is the degree to which data can be compared over time and domain.*

Passenger kilometres, journeys and revenue data are based on a combination of data sources. The data used to derive passenger kilometres, journeys and revenue data are derived from a combination of LENNON data and data provided directly from train operating companies (non-LENNON data). The two sources of passenger kilometres and journey data are required to capture the full volume of passenger kilometres and journeys on the rail network as the LENNON database does not include all methods of and types of products sold.

Non-LENNON data are supplied directly by train operating companies who are coherent in use of the LENNON database and the products which it is able to capture. As a result, TOCs are fully aware of the types of product which are not captured and the resultant passenger kilometres and journeys which should be reported through the non-LENNON data submission to ORR. This product knowledge minimises the chances of double reporting occurring.

There are a number of known issues in the comparability of passenger journeys data. These are covered in the 'Accuracy and reliability' section of this report.

## Comparability to European data

The objective nature of passenger usage data means that comparable data can be obtained across the majority of European countries. Eurostat are the statistical office of the European Union and comparable data on the number of passenger journeys and passenger kilometres are available from the [Eurostat database](#).

The data for the UK on the Eurostat database differs from the figures published as part of the passenger rail usage statistical release as Eurostat data includes data from Northern Ireland Railways, Eurostar and estimates of usage on Le Shuttle and Heathrow Express services.

## Length of Comparable Time Series

Measure	Time Series	Data Portal Table
Passenger train kilometres Quarterly Annual	2010-11 Q1 2010	<a href="#">Table 12.13</a>
Passenger kilometres Annual	1947	<a href="#">Table 12.2</a>
Passenger journeys by sector Quarterly Annual	1994-95 Q1 1994-95	<a href="#">Table 12.3</a>
Passenger kilometres by ticket type Quarterly Annual	1994-95 Q1 1986-87	<a href="#">Table 12.4</a>
Passenger Journeys Annual	1950	<a href="#">Table 12.5</a>
Passenger journeys by sector Quarterly Annual	1994-95 Q1 1994-95	<a href="#">Table 12.6</a>
Passenger journeys by ticket type Quarterly Annual	1994-95 Q1 1986-87	<a href="#">Table 12.7</a>

<p>Passenger revenue by sector</p> <p>Quarterly Annual</p>	<p>1995-96 Q1 1994-95</p>	<p><a href="#">Table 12.8</a></p>
<p>Passenger revenue by ticket type</p> <p>Quarterly Annual</p>	<p>1996-97 Q1 1986-87</p>	<p><a href="#">Table 12.9</a></p>
<p>Revenue per passenger kilometre and revenue per passenger journey</p> <p>Quarterly Annual</p>	<p>1995-96 Q1 1986-87</p>	<p><a href="#">Table 12.10</a></p>
<p>Passenger kilometres by train operating company</p> <p>Quarterly Annual</p>	<p>2011-12Q1 2011-12</p>	<p><a href="#">Table 12.11</a></p>
<p>Passenger journeys by train operating company</p> <p>Quarterly Annual</p>	<p>2011-12Q1 2011-12</p>	<p><a href="#">Table 12.12</a></p>



© Crown copyright 2018

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit [nationalarchives.gov.uk/doc/open-government-licence/version/3](http://nationalarchives.gov.uk/doc/open-government-licence/version/3) or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information you will need to obtain permission from the copyright holders concerned.

This publication is available at [orr.gov.uk](http://orr.gov.uk)

Any enquiries regarding this publication should be sent to us at [orr.gov.uk](http://orr.gov.uk)