[User Note1]

Yorkshire and Humber Regional Assembly

Analysis using the 2001 Census Journey to Work Data

Contents[User Note2]

Submission Reference	115487-00
Description	Final Issue

	Prepared by	**Reviewed by**	**Approved by**
Name			
Initials			
Date			

1 Introduction

Arup has used analysis of the 2001 Census Journey to Work data as part of our commission for the Yorkshire and Humber Regional Assembly (the Assembly) to prepare a Strategic Public Transport Framework for the Yorkshire and Humber Region. The use of the data allowed information on the origin and destination of journeys, made predominantly in the peak period when transport capacity is under most pressure, to reflect the main movements by mode within the region.

The Census data is a useful information source on the distribution of journey to work trips at any level from the very local to the regional, pan-regional (e.g. Northern Way) or national level. It is possible to present statistics and illustrated analysis for trips to or from a particular area, how those trips are made (mode) and what the catchment area is for journeys to a particular destination by different modes.

The use of the data has wider applications to support local and regional transport and land use planning, particularly in illustrating baseline information for specific locations and a more general representation of transport use in a defined area. Arup was asked by the Assembly to outline the potential for using this data to support further work relating to regional planning and transport and to provide further information on its content. This report sets out a brief summary of what the data contains (Section 2), the way it has been analysed thus far (Section 3) and finally offers some examples of further applications relevant to the work of the Assembly (Section 4).

2 About the Data

The 2001 Census Journey to Work data is supplied by the Office for National Statistics (ONS). It forms part of a wide dataset associated with the National Census undertaken every ten years. Data is available for trips to and from England, Scotland and Wales. The original data was made available by ONS in Spring 2004, at ward level then at output area level. Output areas are the smallest aggregation of socio-demographic data available from ONS and correspond to a small number of households within a single administrative area.

The Output Areas (OAs) for England and Wales were defined in January 2003. They are the small area building blocks for the release of 2001 Census results and are used more widely in the Neighbourhood Statistics produced by ONS. The Output Areas nest within ward and parish (community in Wales) boundaries legally in place at the end of 2002.

There are approximately 175,000 Output Areas in England and Wales, created around a target population of 125 resident households. Almost 80% of the resulting boundaries comprise between 110 and 139 households, with only 5% between 40 households (the confidentiality threshold) and 99 households, many of which are a single parish (or community in Wales).

As with most data available from the ONS relating to the Census, information is made available in comma separated value format on CD. The data exclude origin-destinations (ODs) for which there is no recorded movement and contain an adjustment for small counts in order that information about an identifiable individual is not disclosed.

The Journey to Work data provides origin and destination OA for all regular journeys to work recorded in the Census, split by mode, for all people aged between 16 and 74 in employment. Separate records are provided for full time students. The categories of mode specified in the data are as follows:

- Works mainly from home;
- Underground, tram or light rail;
- Train;
- Bus or coach;
- Taxi;
- Car driver;
- Car passenger;
- Motorcycle or moped;
- Bicycle;
- On foot;
- Other.

The data records include the relevant address at different geographic levels as follows:

- Output area;
- Ward;
- District;
- County;
- Region.

There are over six million records covering the UK making this a substantial database for processing and extracting data. However, this includes every incidence of recorded travel between any OA pair throughout the UK. In practice the majority of applications to which the data are relevant means that a regional subset is sufficient and this can be defined in Access or other database software in terms of all trips to or from a specific region. For more detailed analysis the same would be true for trips to or from a particular county, district or ward.

The supplied data present a separate file for Scotland (which was incorporated into the main dataset) and also for residents whose UK address was different from that one year before the census. This migration data only reflects total changes in population and doesn't include corresponding OD information.

Because the data are tied to administrative boundaries it is possible to display any quantity and area of data using geographical information systems (GIS). The output area alphanumeric codes are unique identifiers which can be used to represent individual records or aggregations.

3 Methods of Analysis

The work relating to the Strategic Public Transport Framework for the Yorkshire and Humber Region was our first application of the data and required a degree of learning in order to understand the limitations on using and presenting the data. The full dataset is difficult to manipulate and requires a reasonably high technical specification for the computer used for the processing. Furthermore, to present something reasonably meaningful from the national dataset means that anything below county, or perhaps district, level is too fine a level of detail. As with traditional strategic modelling it is necessary to define different levels of aggregation around an area of interest. For example, a study of Derby may contain data for each ward in the city, each district for the remainder of Derbyshire, at county level for the rest of the region and then aggregated by region beyond that.

The approach used in our work for the Assembly was to cordon the data to only those trips with an origin or a destination within the region. Even this database contains over 560,000 different records, each representing one or more journey between a separately identified OD pair of output areas.

Other than this cordoning of the data the remaining functionality and detail (for example by mode) was retained. This made it possible to look at movements in total, by individual mode and by car or public transport to or from each county, district, ward or output area in the region and for any user-specified aggregation of these subsets. In relation to the purpose of the study this meant that any geographical extent/boundary including city centres, wider city areas, city regions or other areas could be defined and travel to work statistics presented. Presentation of the data in tabular form was possible for each or all modes corresponding to district or sub-district level analysis.

Example 1: Extract from table showing the matrix of all Journeys to Work to and from districts in the Yorkshire and Humber Region

ORIGIN (Home) \ Destination (Place of Work/Study)		Bradford	Calderdale	Kirklees	Leeds	Wakefield	Craven	Hambleton	Harrogate	Richmondshire	Ryedale	Scarborough	Selby	Barnsley	Doncaster	Rotherham	Sheffield	York	East Riding of Yorkshire	Kingston upon Hull City of	North East Lincolnshire	North Lincolnshire
		Yorkshire and The Humber																				
Bradford	Yorkshire and The Humber	149960	4676	4235	21578	1514	3322	81	933	33	24	18	114	168	114	96	261	321	60	39	3	57
Calderdale		7068	64062	6582	4521	834	99	15	75	9	12	0	54	102	60	42	129	69	24	30	9	12
Kirklees		8906	8361	123936	17632	6209	45	45	234	9	21	12	126	1658	246	294	916	231	132	84	18	27
Leeds		15763	2044	6086	270754	9155	303	270	3788	69	108	48	1717	618	519	348	924	1893	327	246	60	150
Wakefield		1924	585	5674	21193	97019	30	54	306	33	27	15	2632	2580	1207	480	895	438	255	93	18	90
Craven		3984	129	102	941	63	16897	99	390	36	15	27	39	21	3	6	3	42	6	3	3	3
Hambleton		93	18	15	710	57	45	28532	2016	776	473	63	149	6	27	9	15	2492	96	21	0	15
Harrogate		1374	120	264	9794	474	321	1768	57593	234	132	90	365	51	48	63	123	1635	69	39	3	21
Richmondshire		21	15	9	147	36	21	1765	381	17470	24	18	6	3	0	3	12	117	9	3	3	27
Ryedale		60	9	9	264	39	9	481	180	15	18716	1136	111	12	21	3	12	2311	435	54	0	9
Scarborough		45	3	9	204	42	57	336	111	36	2061	37894	51	12	15	6	24	432	702	129	6	24
Selby		312	78	249	6350	2951	45	282	528	21	204	66	19868	115	462	57	120	4547	1200	204	18	84
Barnsley		453	168	1918	3092	6289	3	21	93	24	12	15	342	59632	2293	5025	7135	81	57	39	18	102
Doncaster		225	117	264	2070	2704	9	87	99	21	21	27	1344	1837	90217	6406	3724	429	997	213	144	2284
Rotherham		207	78	387	1158	876	6	18	66	6	3	9	129	3024	4881	65512	23141	93	111	51	45	225
Sheffield		312	99	654	2366	849	6	27	81	3	9	21	87	3245	1956	9974	183071	177	78	75	36	186
York		360	105	156	4515	522	21	1661	1592	35	1703	285	1682	60	252	39	126	70205	1762	375	12	84
East Riding of Yorkshire		210	87	162	1726	771	12	216	243	12	971	1237	1733	54	761	117	180	5307	91712	35158	528	1087
Kingston upon Hull City of		72	24	51	348	96	3	39	75	9	60	48	63	12	114	24	51	237	13856	77794	456	720
North East Lincolnshire		6	9	36	93	33	0	3	6	15	0	6	30	12	207	39	75	33	222	729	55624	4385
North Lincolnshire		54	21	69	333	216	9	24	15	6	6	12	117	30	1957	207	333	81	770	1302	2955	56237
Total Yorkshire and The Humber		**191409**	**80808**	**150867**	**369789**	**130749**	**21263**	**35824**	**68805**	**18872**	**24602**	**41047**	**30759**	**73252**	**105360**	**88750**	**221270**	**91171**	**112880**	**116681**	**59959**	**65829**
Total East Midlands		462	231	522	1815	750	39	174	231	69	48	75	171	765	3688	4690	20257	339	402	462	5910	4237
Total North East		294	150	261	1411	339	36	5125	646	1704	443	555	156	153	234	105	273	940	237	150	42	192
Total North West		2452	2245	2034	3979	672	2444	240	420	493	57	87	144	294	264	243	1029	448	258	237	63	174
Total West Midlands		294	96	288	600	180	15	51	108	81	3	33	15	285	141	174	594	126	114	120	51	69
Total London		219	117	222	363	75	27	33	45	45	12	36	33	63	60	78	330	69	63	72	24	66
Total East of England		183	93	159	417	72	21	36	45	45	39	18	24	54	87	289	450	153	72	123	48	63
Total South East		210	81	168	447	99	27	30	93	90	12	39	21	69	123	81	315	102	84	120	27	60
Total South West		183	54	123	261	48	45	48	90	39	3	9	12	18	51	54	174	72	45	105	21	24
Total Other (Wales, Scotland, NI)		126	61	117	291	66	12	45	42	117	15	18	30	33	48	63	183	54	63	45	15	54
Total		**195832**	**83936**	**154761**	**379373**	**133050**	**23929**	**41606**	**70525**	**21555**	**25234**	**41917**	**31365**	**74986**	**110056**	**94527**	**244875**	**93474**	**114218**	**118115**	**66160**	**70768**

For our analysis modal split information was particularly interesting, and this included aggregating all public transport modes for comparison against the corresponding number of car journeys (as driver or passenger). This sort of analysis, and the corresponding information for individual modes, was presented for different areas in each district as appropriate, for example the city centre, remaining suburban areas and outlying rural areas. These spatial definitions were used to distinguish between different movement corridors and to reflect the relative mode share of public transport within or outside of the central areas. This included an understanding of how public transport contributes to the geographic spread of journeys to the main centres and how well bus and rail cater for trips with more dispersed destinations, for example out of town employment sites and business parks traditionally associated with trips by car.

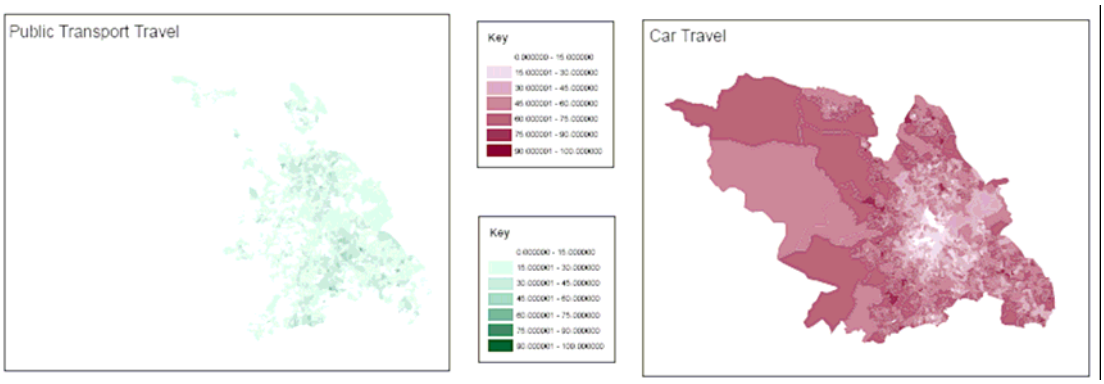
From the information provided by ONS it is not possible to define a route for each journey but by comparison with strategic and local highway networks, and public transport services and key interchange points, it is possible to gain an understanding of the use being made of each transport network. In particular it is possible to reflect the balance of short local journeys and longer distance journeys in particular corridors. For our strategic study it was not possible to look into these statistics at a local level.

Of further relevance to public transport is the understanding of corridors that exhibit significant bi-directional flow, particularly in the same assumed time period, to identify where bus and rail service may be most efficient. A relatively even bi-directional movement is preferred in order to match the service frequency and capacity to the demand and to maintain a reasonably high occupancy in both directions. An understanding of the current service specification for strategic movements then helps to illustrate those sections of the network that appear to performing well or poorly in relation to mode share.

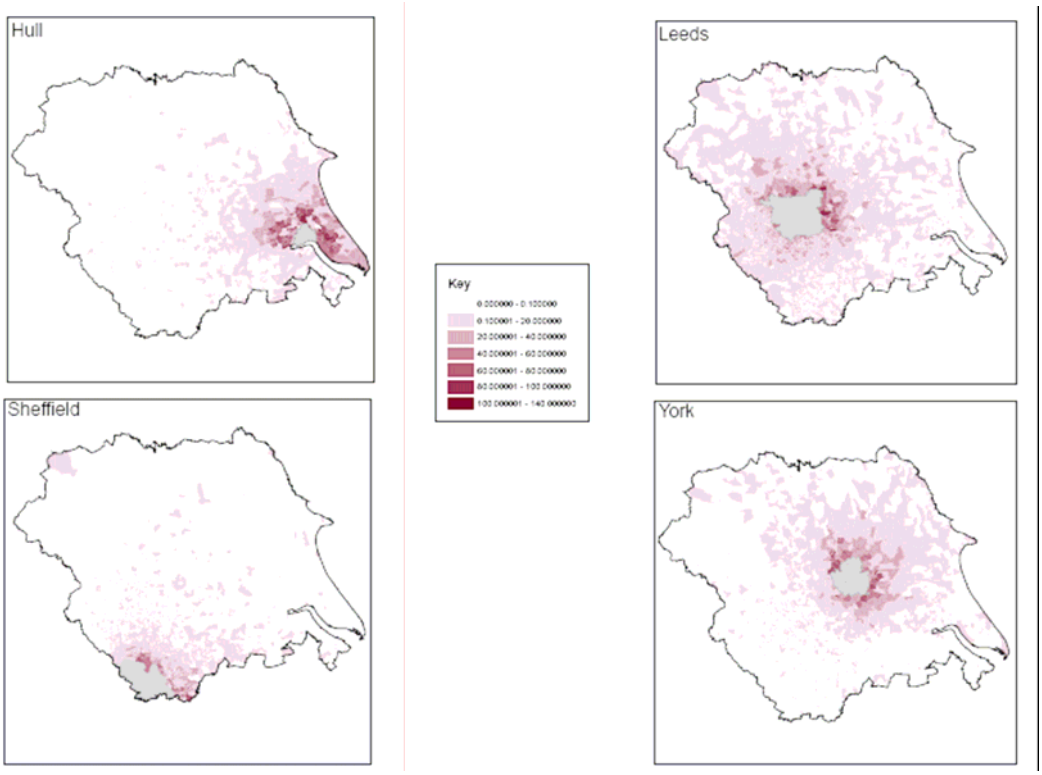
The data also lends itself to different types of aggregation and analysis in order to represent different aspects of the journey to work. The focus of the analysis was defined both for the origin and destination end of the trip to illustrate the distribution of trips to or from particular areas. These describe the relative competitiveness of employment areas within the region and identify key cross boundary issues that may or may not be important for current or future public transport accessibility.

This type of analysis is illustrated in the following examples, showing mode share for local (district level) journeys, geographic distribution of car trips to the main centres in the region and both the mode share and percentage of journeys to work in a specific destination district from all other districts.

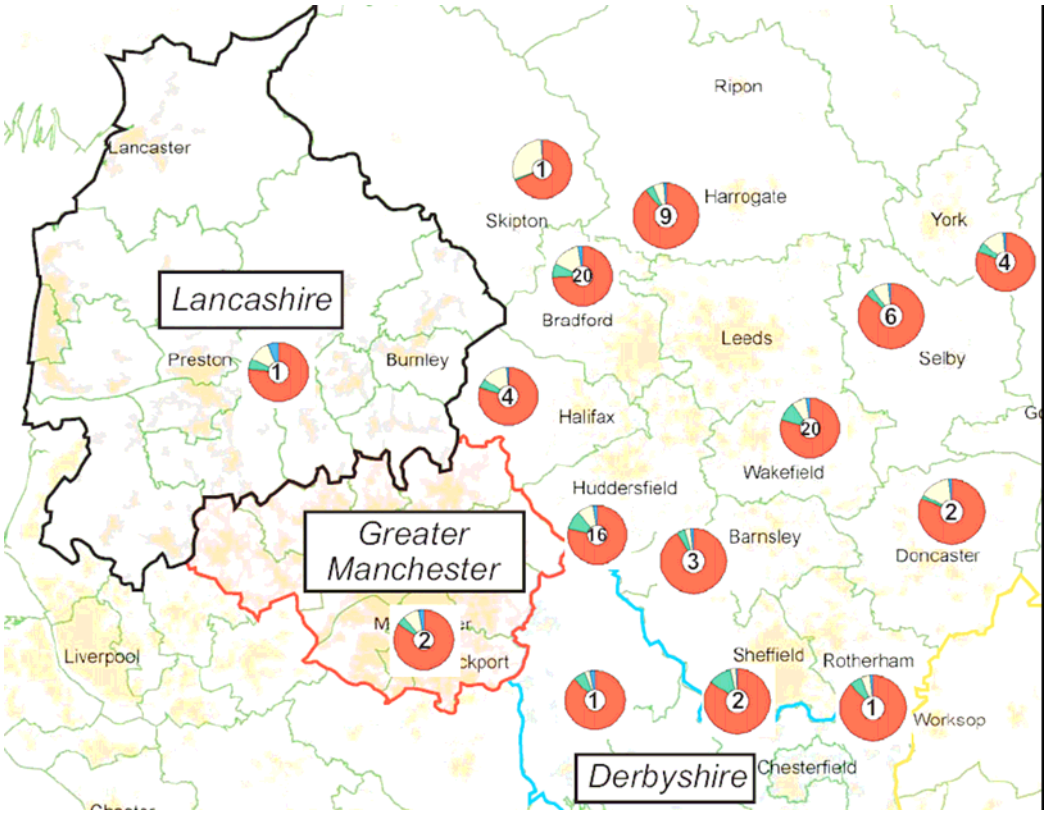
Example 2: Percentage mode share for car and public transport to Sheffield city centre from the rest of Sheffield district



Example 3: Distribution of Journey to Work trips by car to each of the four main regional centres.



Example 4: Distribution and mode share for Journey to Work trips with a destination in Leeds (extract).



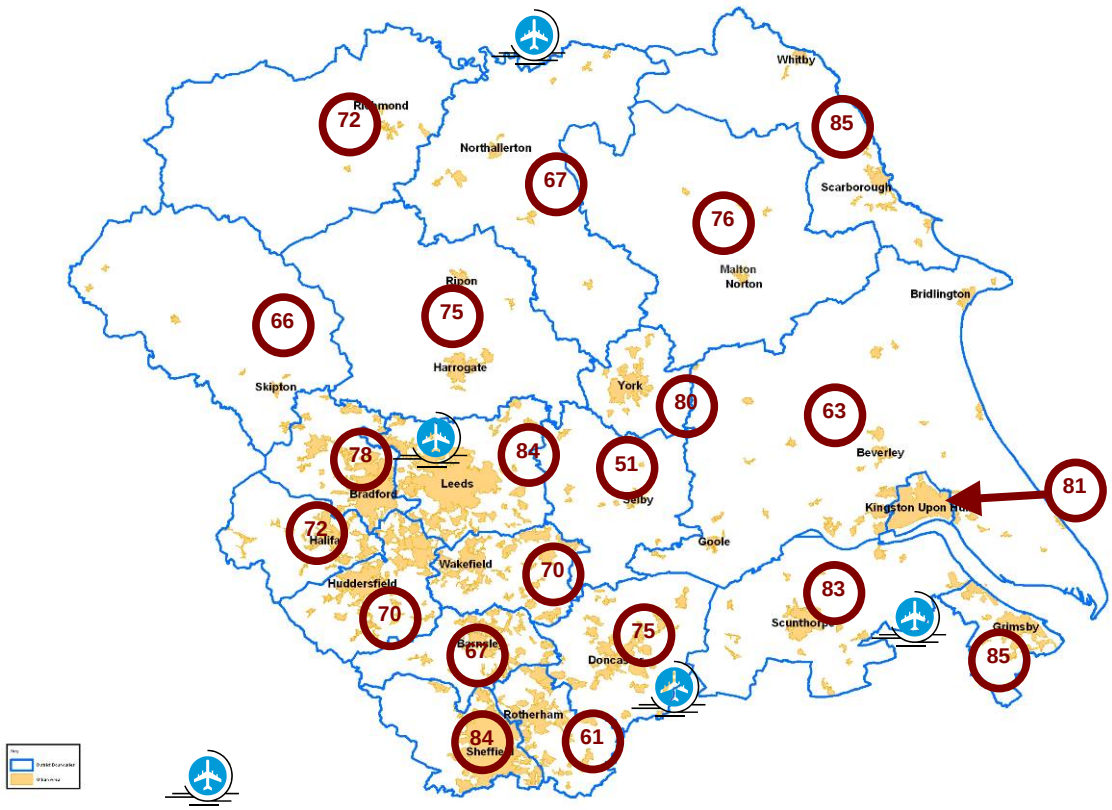
It is also possible to link this information to journey distances, based on typical travel distances between district centres or other specific locations. This illustrates the relevant distance thresholds for a certain percentage of journeys by specific modes. When linked to average speed of travel by each mode this can also be used to roughly represent journey time catchments and approximate isochrones.

All of the data is held within a database, in this case using Microsoft Access. This allows a variety of tables to be constructed based on queries that select certain data subsets according to user defined criteria. This preserves the overall dataset but allows relatively easy extraction and display of information relating to a specific mode, area or movement.

In general an attempt has been made to display the information spatially as far as possible so that it is easy to understand the relevant regional and sub-regional interactions. As illustrated in Example 4 because the data extends beyond the region (either a trip origin or a trip destination may be outside of the Yorkshire and Humber Region) this can also represent the relative importance of cross boundary links and movements with high volumes.

Comparison between districts is also easy to represent as illustrated in Example 5. Although the strategic study only considered this at the district level it would be equally possible to analyse the data at the ward or output area level.

Example 5: Percentage of Journeys to Work with an origin and destination in the same district (extract).



4 Application

The 2001 Census Journey to Work data is an important source for spatial transport and land use planning, at a strategic or local level. It links easily to GIS allowing useful presentation of the analysis and with appropriate manipulation of the data can be used to present aggregate or location-specific information.

The analysis performed to date has concentrated on the distribution and share for public transport and private modes at a district level or associated with key regional centres. However, this can readily be repeated for more specific local areas with corresponding reference to highway or public transport networks. This can be spatially very specific and would be easily represented against an annotated Ordnance Survey base or other scalable mapping.

At the local level this information can be used as the basis of distribution and mode of access for trips to individual development sites or employment areas and give indications of relative accessibility by mode. For example, this data could be used and shown graphically with employment land review data. With more detailed specification of the movements being analysed this can be extended to consider particular movement corridors and individual routes. This is particularly true for strategic networks such as motorways, trunk roads and railways. Further analysis can distinguish between local and regional journeys anticipated for particular corridors.

As well as providing data for employment locations the distribution of trips from a particular residential origin area can also be analysed. This can be used as a base distribution for new or increased residential development or to indicate the relative sustainability (in terms of potential use of public transport) between alternative sites. Analysis of individual modes can further support the presentation of local or regional policies, for example on cycling and walking, by relating plans to observed data.

This links to a potential application in respect of Local Plan proposals, Regional Spatial and Economic Strategies and for issues associated with physical geography such as landscape, topography, urbanisation, etc.

The ability to link the data to other GIS based information, related to the same administrative boundary data allows reference to other planning and socio-demographic information. Whilst the individual journey OD information cannot necessarily be linked to individual population the output area level allows for a considerable level of detail, for example in respect of social inclusion, accessibility and employment related analysis.

The relevance of the Journey to Work data also extends to supplementing accessibility planning and modelling, analysing access to health facilities or in relation to indices of deprivation for example. Combining both Census data analysis and other accessibility tools (for example Accession) could be used to indicate how theoretical journey time relates to mode share and catchment distribution of journeys to work.

For new transport infrastructure proposals it would be possible to use the Census data to analyse the catchment areas of any new schemes and how they might impact on existing mode shares, particularly on established public transport services.