

Statistics Commission



**ACCESS TO NATIONAL
STATISTICS ON TRANSPORT
VIA THE WEB**

**Statistics Commission Report No 6
January 2002**

Statistics Commission

**Statistics Commission
Report No. 6**

Access to National Statistics on Transport via the Web

Prepared for the Commission by

Steer Davies Gleave

**Access to National Statistics on
Transport via the Web**

Revised Report

January 2002

Prepared for:

Statistics Commission
10 Great George Street
London
SW1P 3AE

Prepared by:

Steer Davies Gleave
28-32 Upper Ground
London
SE1 9PD

- 1. INTRODUCTION..... 1**
 - Background and Aims 1
 - Methodology..... 1
- 2. RESULTS..... 5**
 - Introduction..... 5
 - Indicators..... 5
 - Summary by Grade of Query 5
 - Analysis by Experience Level of Data-shopper 7
 - Success Rate of Individual Queries 8
 - Sources 9
 - National Statistics Website 10
 - DLTR Website 10
 - Comments 11
- 3. RECOMENDATIONS 12**
 - Introduction..... 12
 - National Statistics Website (StatBase)..... 12
 - DLTR Website 13

File Ref: p:\1 - central admin\publications\commission documents\final documents for web\reports 1-9 & 11-\006 - transport stats via web.doc

1. INTRODUCTION

Background and Aims

- 1.1 The Statistics Commission wished to test the accessibility of National Statistics on transport via the Internet and consequently commissioned Steer Davies Gleave to conduct a small scale “mystery shopping” style study.
- 1.2 The study was designed not just to test how easy it is to obtain statistics via the Web, but also to obtain useful qualitative data which could be used to inform recommendations for improving the access to transport statistics via the National Statistics and other relevant government Websites.

Methodology

- 1.3 The overall idea of the study was to devise some realistic tests of the accessibility of transport related statistics. Thus, a mix of expert and novice “data shoppers” were employed to try and answer a pre-defined list of questions. The 33 questions used are shown in Table 1.1. It will be seen that there is room for ambiguity in some of these, a consequence of the desire to pose some questions which realistically might be made by members of the general public, as well as other questions of the type likely to be asked by transport professionals.
- 1.4 Note that the questions were not designed necessarily to be ones that could be answered using National Statistics but rather were designed to reflect the range of types of questions which different people might be expected to pose.
- 1.5 This table also shows that the queries have been “graded” into three levels of complexity:
- (1) simple – one dimensional questions
 - (2) intermediate – two dimensional questions incorporating some comparative data such as by year, country or demographic sub-group
 - (3) complex – more complicated queries which require a degree of analysis in the statistics sought.
- 1.6 A set time limit was set for trying to answer each query, with more time allotted to the more complex questions.
- 1.7 Seven different individuals were used to try and answer the queries, three of whom were classified as “novices” because they had only used the Internet for personal use and were not transport professionals, and four who are professional employees of Steer Davies Gleave (two who are research executives, one a senior transport consultant and the Research Manager).

TABLE 1.1: QUESTIONS POSED

Theme	Questions (complexity grading*)
Cars & car travel	<ol style="list-style-type: none"> 1. How many cars are there in the UK? (1) 2. How many people have regular access to a car? (2) 3. What is the average amount spent on buying a car? (2) 4. What is the area (County) with the highest / lowest car ownership? (2) 5. How many miles per year does the average driver drive? (2) 6. What volume of pollutants do cars in the UK generate now, and 10 years ago? (2)
Roads	<ol style="list-style-type: none"> 7. How much is spent on roads? (1) 8. How has spending on roads changed over the last 10 years? (2) 9. How have average road speeds changed over the last 10 years? (2) 10. Which region suffers from the most road congestion? (3) 11. How many people are killed on British local, single carriageway, dual carriageway and motorways each year? (2)
Walking & cycling	<ol style="list-style-type: none"> 12. How far does the average person walk per year? (1) 13. How does the amount people walk vary by age? (2) 14. How does the amount people cycle vary by age? (2) 15. What are the numbers of pedestrians killed and injured by where accident took place (pavement/street/minor road/main road) over the last 5 years (3)
Public transport	<ol style="list-style-type: none"> 16. How has government spending on public transport changed over the last 10 years? (2) 17. How have taxi, bus and rail fares changed over the last 10 years? (3) 18. How do public transport mode shares vary between rural, urban and metropolitan areas? (2) 19. What is the average age of buses in the UK? (1) 20. How does public transport use vary by income/social class? (2) 21. Which Train Operator are customers the most & least satisfied with? (1) 22. Which Train Operator has the most & least reliable services? (1) 23. What has been the effect of the Hatfield crash on numbers of rail passengers, road traffic and road deaths? (3)
Air travel	<ol style="list-style-type: none"> 24. What proportion of the population makes at least one flight a year? (2) 25. What is the air/rail /coach/car share of London-Scotland travel? (3) 26. How many people use Heathrow airport per year? (1)
International comparisons	<ol style="list-style-type: none"> 27. How does spending on roads compare across the EU? (2) 28. How does the annual cost of running a car vary across the EU? (2) 29. How does the cost of rail fares vary across the EU (in terms of p per kilometre)? (3)
Freight	<ol style="list-style-type: none"> 30. What is the busiest port (for freight) in the UK? (1) 31. What % of freight traffic between Britain and the rest of the EU goes via the Channel Tunnel? (2) 32. What proportion of domestic freight is carried by rail, now (1) and compared with 5 and 10 years ago? (2) 33. What volume of livestock is moved in Britain in a year? (2)

1.8 The time allotted for each query was also adjusted to take into account the level of experience of the data-shopper. This not only reflects the fact that an experienced researcher is likely to get to the answer more quickly, it also reflects the level of expectation – a novice or a member of the general public would perhaps not expect to find the statistic they are after quite as quickly as a professional. The time made available for each query is shown in Table 1.2.

TABLE 1.2: TIME AVAILABLE TO ANSWER QUESTIONS (MINUTES)

	Simple	medium	Complex
Project Manager	5	7	13
Experienced	7	11	20
Novice	11	17	31

1.9 The broad mix of individuals used ensured that although the results of the study cannot be said to be statistically significant or representative of the whole (target) population, they do represent a realistic range of consumers.

1.10 One important consequence of this is that each individual started off with their own strategy for tackling the queries. These strategies included:

- starting with a Web browser such as Yahoo, MSN or Google;
- starting by going directly to a likely web address;
- starting from the National Statistics Website;
- starting from the DTLR Website.

1.11 Each data-shopper kept a record of their actions, any data they obtained, and also provided qualitative feed-back of their experience. The forms used to record this information are shown in Appendix A.

1.12 Not all of the queries were attempted by all of the data-shoppers. The study was designed in this way partly to spread the number of people conducting the research wider (for greater diversity) and partly to reduce the impact of learning – particularly for the novices. However, the Project Manager did tackle all the questions providing a certain amount of directly comparable data across queries.

1.13 The order in which the queries was tackled was randomised. The way this was done was to stratify the sample by query category and to randomly allocate the appropriate number of queries within category to each data-shopper. Having created each data-shoppers' list of questions, the list was then sorted using an Excel generated random number.

1.14 In total, 105 “data-shops” or tests were undertaken. This break down by grade of query as follows:

TABLE 1.3: NUMBER OF TESTS

Grade	Tests	Queries	Tests per query
Simple	8	35	4.4
Intermediate	20	56	2.8
Complex	5	14	2.8

1.15 The simple queries were tested slightly more to maintain a reasonable balance across all tests.

2. RESULTS

Introduction

2.1 The results of the survey fall into two categories:

- *indicators* – simple numerical analyses of the results of the data-shopping, useful for obtaining an initial overview and for comparisons,
- *comments* – qualitative information which helps to illustrate the problems encountered and what makes the difference between a rewarding and a frustrating experience.

Indicators

2.2 The indicators are:

- success rate* – the proportion of tests which were completely or largely successful;
- achievement rate* – the proportion of tests which were answered in full;
- failure rate* – the proportion of tests which elicited no relevant information at all;
- actions required – the average number of actions or steps required to achieve an answer;
- satisfaction rating – qualitative rating of the ease of tackling each question (on a scale out of 10);

2.3 The indicators based on proportions* are reported at an individual query level using the following scale:

0%	-
1-39%	*
40-59%	**
60-99%	***
100%	****

Summary by Grade of Query

2.4 In this section we summarise the survey results with breakdowns by grade/complexity of query, starting with the level of success or failure for the enquiries.

2.5 Overall, just over half the queries were "successful", though a little less than a third were answered in full from data available on the Web.

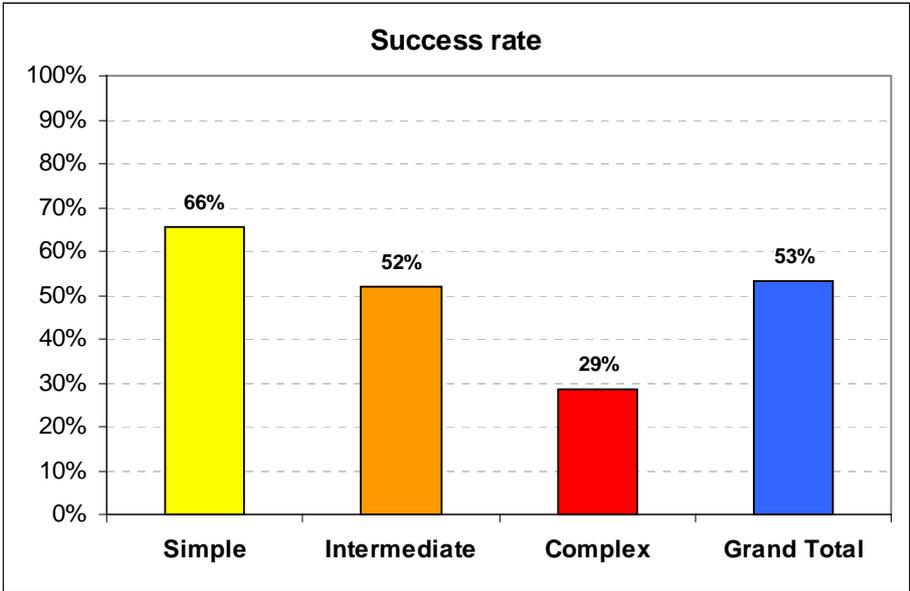
2.6 As would be expected, the success rate declined as the questions become more complex and indeed the expectation was that the complex queries would generally prove too much of a challenge (see Table 2.1 and Figure 2.1). On the other hand, a third of simple queries could not be answered and for a quarter of them no relevant information was found at all (this is denoted by the "failure rate" shown in table 2.1). Later on we examine in more detail which particular questions proved most difficult to answer.

TABLE 2.1: SUCCESS & FAILURE RATES

	Success rate	Achievement Rate	Failure rate
Simple	66%	49%	26%
Intermediate	52%	21%	27%
Complex	29%	7%	50%
Grand Total	53%	29%	30%

Note : Success =largely answered, Achieved = fully answered, failed = no relevant statistics

FIGURE 2.1: SUCCESS RATE BY GRADE OF QUERY



Note : % of queries largely answered

2.7 On average, successful queries were answered in 5 minutes (note that this does not include any time taken to log onto the internet), and 5.7 actions or steps were required to get to the answer (see Table 2.2). The minimum number of actions taken was 2, while the most was 17 (going up to 32 if unsuccessful queries are included).

TABLE 2.2: AVERAGE NUMBER OF ACTIONS AND TIME TAKEN FOR SUCCESSFUL QUERIES

	Average Actions	Average Time
Simple	5.4	3.8
Intermediate	5.2	5.8
Complex	10.0	6.3
Grand Total	5.7	5.0

2.8 Data-shoppers were asked to give a satisfaction rating designed to give an indication of not just the successfulness of each test but also the extent to which trying to answer the query was an enjoyable or frustrating experience. The overall average rating was 5.5 out of 10, declining with the complexity of the question broadly in line with the success rate, although the rating for complex queries may have been raised a little by lower levels of expectation.

TABLE 2.3: AVERAGE SATISFACTION RATING

	Average
Simple	6.2
Intermediate	5.6
Complex	3.9
Grand Total	5.6

Analysis by Experience Level of Data-shopper

2.9 Experience clearly makes a big difference to how easy it is to use the Web, as shown in Table 2.4 where it can be seen that the success rate for experts¹ was much higher than that for novices (also bear in mind that novice data-shoppers were given longer to complete their tasks).

TABLE 2.4: SUCCESS RATE AND SATISFACTION BY TYPE OF DATA-SHOPPER

	Success rate	Satisfaction rating
Expert	70%	6.9
Novice	44%	4.8

¹ Experts includes experienced researchers but not the Research Manager who deliberately used StatBase as the initial source in all cases, even though this was not always the best option.

Success Rate of Individual Queries

2.10 The level of success for each query is shown below. 4 stars means 100% success and no stars means 0% success.

TABLE: 2.5: SUCCESS RATES BY QUERY

	Success rating
1. How many cars are there in the UK?	***
2. How many people have regular access to a car?	****
3. What is the average amount spent on buying a car?	*
4. What is the area (County) with the highest / lowest car ownership?	***
5. How many miles per year does the average driver drive?	***
6. What volume of pollutants do cars in the UK generate now, and 10 years ago?	***
7. How much is spent on roads?	****
8. How has spending on roads changed over the last 10 years?	*
9. How have average road speeds changed over the last 10 years?	*
10. Which region suffers from the most road congestion?	-
11. How many people are killed on British local, single carriageway, dual carriageway and motorways each year?	****
12. How far does the average person walk per year?	**
13. How does the amount people walk vary by age?	*
14. How does the amount people cycle vary by age?	****
15. What are the numbers of pedestrians killed and injured by where accident took place (pavement/street/minor road/main road) over the last 5 years	*
16. How has government spending on public transport changed over the last 10 years?	***
17. How have taxi, bus and rail fares changed over the last 10 years?	****
18. How do public transport mode shares vary between rural, urban and metropolitan areas?	-
19. What is the average age of buses in the UK?	*
20. How does public transport use vary by income/social class?	*
21. Which Train Operator are customers the most & least satisfied with?	**
22. Which Train Operator has the most & least reliable services?	*
23. What has been the effect of the Hatfield crash on numbers of rail passengers, road traffic and road deaths?	**
24. What proportion of the population makes at least one flight a year?	***
25. What is the air/rail /coach/car share of London-Scotland travel?	-

26. How many people use Heathrow airport per year?	****
27. How does spending on roads compare across the EU?	*
28. How does the annual cost of running a car vary across the EU?	-
29. How does the cost of rail fares vary across the EU (in terms of p per kilometre)?	-
30. What is the busiest port (for freight) in the UK?	****
31. What % of freight traffic between Britain and the rest of the EU goes via the Channel Tunnel?	*
32. What proportion of domestic freight is carried by rail, now and compared with 5 and 10 years ago?)	****
33. What volume of livestock is moved in Britain in a year?	-

Note: success = largely successful

2.11 Those queries which caused the least and most problems are highlighted below:

Queries with 100% success	Queries with 0% success
2. How many people have regular access to a car?	10. Which region suffers from the most road congestion?
7. How much is spent on roads?	18. How do public transport mode shares vary between rural, urban and metropolitan areas?
11. How many people are killed on British local, single carriageway, dual carriageway and motorways each year?	25. What is the air/rail /coach/car share of London-Scotland travel?
14. How does the amount people cycle vary by age?	28. How does the annual cost of running a car vary across the EU?
17. How have taxi, bus and rail fares changed over the last 10 years?	29. How does the cost of rail fares vary across the EU (in terms of p per kilometre)?
26. How many people use Heathrow airport per year?	33. What volume of livestock is moved in Britain in a year?
30. What is the busiest port (for freight) in the UK?	
32. What proportion of domestic freight is carried by rail, now and compared with 5 and 10 years ago?)	

2.12 A particular factors which appeared to create difficulties was the inclusion of a spatial element, for example, comparisons between different parts of the country, or between the UK and other EU countries.

Sources

2.13 The two main sources which provided information were the National Statistics / StatBase Website and the DTLR Website. These two accounted for nearly 80% of relevant statistics obtained, with the DTLR site being the top source (see Table 2.6).

2.14 The other Websites which provided useful data were:

- [www.greengas.u-net.com/transport info-html](http://www.greengas.u-net.com/transport%20info-html)
- www.rail-reg.gov.uk/complain/bullet3/annexb.htm
- www.seanet.co.uk
- www.sra.gov.uk

TABLE 2.6: SOURCES USED FOR SUCCESSFUL ENQUIRIES

	% of successful enquiries
NS	24%
DTLR	54%
NS & DTLR	6%
Other	17%

National Statistics Website

2.15 The National Statistics site was accessed altogether in 47 out of the 105 queries (45%). It was used slightly more in unsuccessful queries (49%) than successful ones (41%). The NS logo was clearly evident when using this site.

2.16 In virtually all cases it was the StatBase sub-site which was used, either because that was where the user was directed to, or in the case of experienced users, that was their preferred starting point.

DTLR Website

2.17 The DTLR Website was used in 51 of the 105 queries, and of these 61% were successful.

2.18 The NS logo was not evident on this site.

Comments

- 2.19 Data-shoppers commented on the factors which they thought contributed to making their search easier or more difficult. Some of these were specific to a particular enquiry while others are more general in nature. A summary of the positive and negative open ended comments is provided below.

Positive comments	Negative comments
<ul style="list-style-type: none"> ➤ Easy once found DTLR site ➤ Great site! (DTLR) ➤ Hot link from StatBase to DTLR site ➤ Good once I'd found the pdf version of Social Trends ➤ StatBase very user friendly 	<ul style="list-style-type: none"> ➤ Not available in StatBase had to go to DTLR site ➤ Search facility in StatBase not very helpful ➤ Couldn't get access to historic issues of Social trends ➤ Didn't get directed to NTS even though it would have had the answer ➤ Didn't get directed to DTLR Traffic Speed reports even though it would have had the answer ➤ Some DTLR publications not accessible from StatBase ➤ Focus on Roads looked useful but couldn't access content ➤ Having selected a link to one table in Social Trends it was not possible to go straight to another table or section ➤ found no sources for European comparisons ➤ found no references to pollution ➤ found lots of articles on congestion but no statistics

- 2.20 Some of the greatest causes of frustration which are worth highlighting were:

- knowing that a report or statistic was available but not being directed towards it, or it not being listed in searches;
- finding a statistic which was close to what was needed but not being able to re-direct or narrow down the search without starting from the beginning again;
- this was especially frustrating when the statistic being searched for was actually in the same document but it was not possible to quickly move from one statistic to another in the same topic area, or from one topic area to another;
- finding data on the topic of interest but in the wrong format (for example, car ownership per household rather than the number of cars owned).

3. RECOMENDATIONS

Introduction

- 3.1 Overall it is clear that there is a vast amount of information on transport available via the Web and that very considerable effort has been put into making it accessible to the general public and professionals alike.
- 3.2 Nevertheless, this research paints a rather mixed picture of how easy it is to access statistics and in this section the aim is to identify positive suggestions for how further improvements might be made to the accessibility of transport statistics.
- 3.3 The two main sources for transport information are the National Statistics Website (and specifically StatBase) and the DTLR Website and since these are very different in their content and structure they are dealt with separately. A general point though is that the connectivity between StatBase and the DTLR site could be improved. Although there are some hot-links in StatBase these could be extended, while the DTLR site makes little or no reference to National Statistics or StatBase once you've moved beyond the introductory pages.

StatBase

- Given that the DTLR site is the number one source for transport statistics one suggestion is that whenever a key word such as “transport”, “car”, “road”, “rail” etc. is used in a search the DTLR site automatically results as a “hit”. This might be extended so that for transport related sections/tables within Regional and Social trends the reader is also pointed to the DTLR site as a useful source of further information.
- A more specific recommendation along the same lines is the addition of links to Transport Statistics GB, a DTLR publication which often has more detailed information than Social Trends or Regional Trends, but which is rarely mentioned in StatBase (it is mentioned as a “flagship product” in the introductory pages in the NS site, though without a hot-link, but then does not appear much, if at all, in StatBase – it is, for example, missing from the “publication datasets” list, as is Transport Trends).
- For rail related queries links to the SRA (Strategic Rail Authority) and Railtrack sites might also usefully be added, with something along the same lines potentially also set up for other modes.
- The search facility in StatBase appeared to be quite limited in scope. This might be addressed by ensuring it comes up with the DTLR site for transport queries, though alternatively the search facility might be extended so it searches the DTLR site itself.
- The most frustrating thing of all when searching the Web is coming to a complete dead end. This happened quite a few times during this exercise even though the data being sought was available. This might be addressed by referring to or adding links to related sources. These might include Transport Statistics GB mentioned above, also Transport Trends, Regional Trends and Social Trends.

For example, when searching for car ownership data in StatBase one data-shopper was directed to a table in Social Trends. What would be helpful in an instance like this is to have additional links not just to the one data table but to other car ownership related data which is available in Regional Trends and Transport Statistics GB (a simpler alternative would be simply to add a comment such as "additional information on this topic can be found in ...").

- The Product Information pages (for example for Regional Trends) were useful and included hotlinks to pdf reports where available. One improvement though would be to have the Chapter Headings converted to live links to the relevant parts of the electronic version of the report, or to a list of the electronic data tables available from that section.
- The colour scheme used in the NS site does not highlight clearly enough the live links (they are coloured light blue with no underline) making it quite easy to miss them, particularly with some monitors. We would recommend at least underlining all hot-links if not choosing a more distinctive colour.
- Certain types of query proved more challenging. For example, while time-series data was readily available, geographical/spatial comparisons were usually more difficult. One specific area where we believe it would be worth providing more links and sources is international and specifically European comparisons.

DTLR Website

- Our main recommendation for the DTLR site concerns adding reference to National Statistics since at the moment it is not at all evident whether a publication on the site is a National Statistic or not. Thus, while in the introduction to the DTLR site there is a useful introductory paragraph to National Statistics mentioning the National Statistics logo as a mark of assurance, once you move onto the Transport Statistics section there appears to little or no use of the NS logo. This is in spite of the fact that, for example, Transport Statistics GB is published under joint DTLR and National Statistics branding.
- A minor point is that while there are a number of useful links in the Transport Statistics Main Index Page, they are on a page which is slightly confusingly referred to as "Links to other foreign transport statistics web sites" (linked sites includes the SRA, Railtrack, Highways Agency and StatBase, as well as international transport statistics sites).
- In addition, it would also be useful to have a page of links to major private sector organisations such as the AA and RAC, Train Operating Companies and bus companies. This would enable the site to become the standard first port of call for transport related information.

Contents of Appendices:

APPENDIX A: FORMS

APPENDIX B: OURCOMES TABLE

APPENDIX C: RESULTS (too bulky to include here but available in hard copy from the Statistics Commission)

APPENDIX A

Forms

Access to National Statistics: Record Sheet

Query reference

Your initials

No.	Action	Outcome	N S	Time	Attachments
				
				
				

Access to National Statistics: Summary Sheet

Query reference

Your initials

Did to what extent do you feel you managed to answer the question posed ?

1. Completely
2. Mostly
3. Only partly
4. Not at all

How long did you spend on (trying to) answer the query?

Do you think this is a reasonable amount of time bearing in mind the nature of the question?

1. Yes
2. Perhaps a little longer than it should be
3. A lot longer than it should be

Overall, on a scale of 1 to 10 (1=extremely difficult to 10=extremely easy), how easy did you find it to answer the question posed ?

What in particular made it easy or difficult?

Were there any Web-sites which you found particularly useful? (if so, which)

Were there any Web-sites which provided useful information but which were more difficult to use (in this instance) than they needed to be?

APPENDIX B

Outcomes Table

Query	Completed	Largely completed	Partially completed	Failed	Grand Total	Successful
1. How many cars are there in the UK?	4		1	1	6	4
2. How many people have regular access to a car?	0	3			3	3
3. What is the average amount spent on buying a car?	0	1		2	3	1
4. What is the area (County) with the highest / lowest car ownership?	1	1		1	3	2
5. How many miles per year does the average driver drive?	1	1		1	3	2
6. What volume of pollutants do cars in the UK generate now, and 10 years ago?	0	2		1	3	2
7. How much is spent on roads?	4	1			5	5
8. How has spending on roads changed over the last 10 years?	0	1	2		3	1
9. How have average road speeds changed over the last 10 years?	1		1	1	3	1
10. Which region suffers from the most road congestion?	0		2	1	3	0
11. How many people are killed on British local, single carriageway, dual carriageway and motorways each year?	0	3			3	3
12. How far does the average person walk per year?	2			2	4	2
13. How does the amount people walk vary by age?	1		1	1	3	1
14. How does the amount people cycle vary by age?	2	1			3	3
15. What are the numbers of pedestrians killed and injured by where accident took place (pavement/street/minor road/main road) over the last 5 years	0	1	1	1	3	1
16. How has government spending on public transport changed over the last 10 years?	1	1		1	3	2
17. How have taxi, bus and rail fares changed over the last 10 years?	1	1			2	2
18. How do public transport mode shares vary between rural, urban and metropolitan areas?	0		1	1	2	0
19. What is the average age of buses in the UK?	1		2	1	4	1
20. How does public transport use vary by income/social class?	1		1	1	3	1

File Ref: p:\1 - central admin\publications\commission documents\final documents for web\reports 1-9 & 11-\006 - transport stats via web.doc

21. Which Train Operator are customers the most & least satisfied with?	0	2		2	4	2
22. Which Train Operator has the most & least reliable services?	1			3	4	1
23. What has been the effect of the Hatfield crash on numbers of rail passengers, road traffic and road deaths?	0	1		1	2	1
24. What proportion of the population makes at least one flight a year?	0	2	1		3	2
25. What is the air/rail /coach/car share of London-Scotland travel?	0			2	2	0
26. How many people use Heathrow airport per year?	2	2			4	4
27. How does spending on roads compare across the EU?	0	1	1	1	3	1
28. How does the annual cost of running a car vary across the EU?	0		1	2	3	0
29. How does the cost of rail fares vary across the EU (in terms of p per kilometre)?	0			2	2	0
30. What is the busiest port (for freight) in the UK?	3	1			4	4
31. What % of freight traffic between Britain and the rest of the EU goes via the Channel Tunnel?	1		2		3	1
32. What proportion of domestic freight is carried by rail, now and compared with 5 and 10 years ago?)	3				3	3
33. What volume of livestock is moved in Britain in a year?	0		1	2	3	0
Grand Total	30	26	18	31	105	56