

What are aggregate data?

David Rawnsley
UKDS Senior Data Technical Officer


Online tools for accessing and analysing
census data from the UK Data Service: census
aggregate data

December 1st 2020



Quiz time

- A – 2,467
- B – 27,229
- C – 176,632
- D – 390,870
- E – 1,156,221

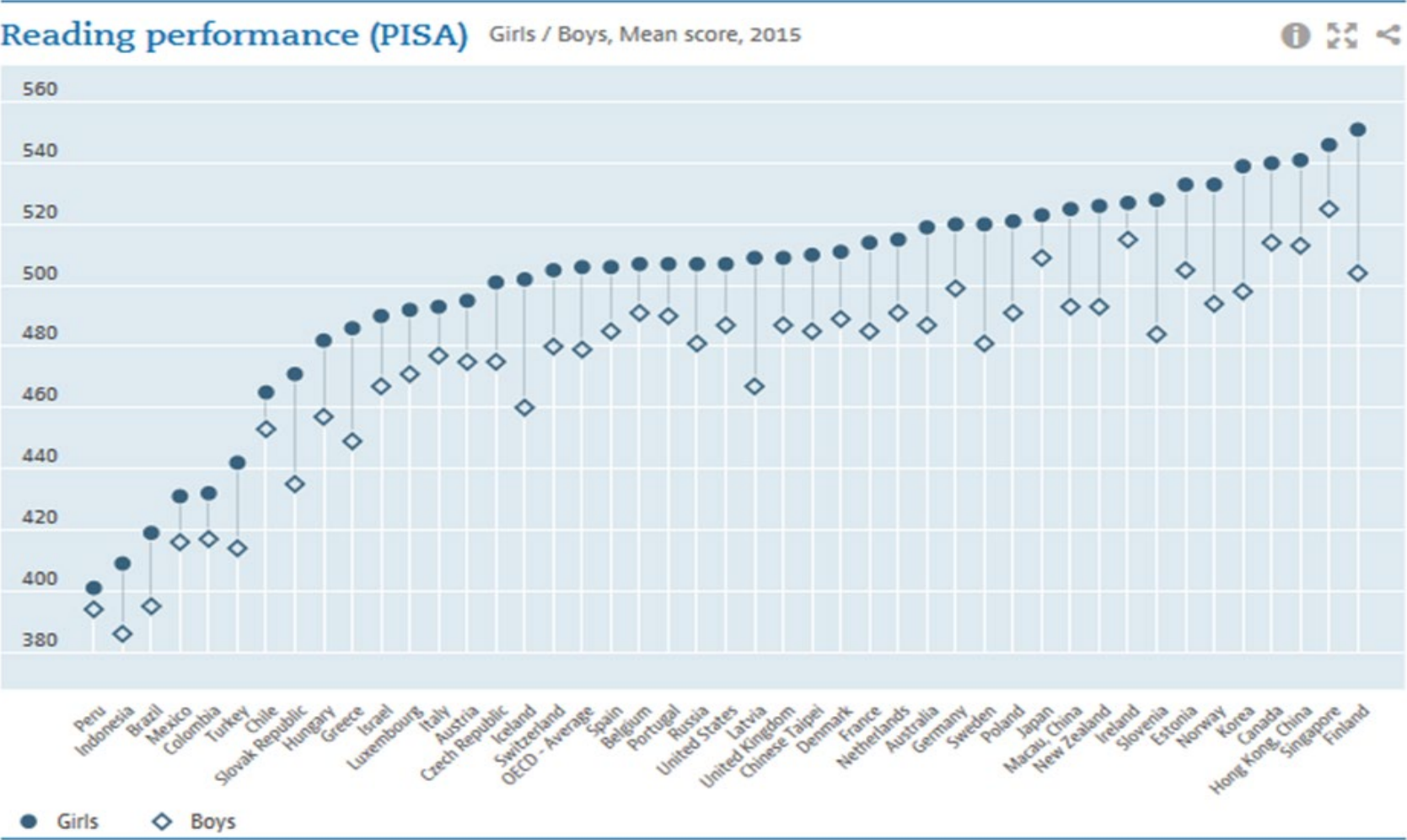


In 2011 how many people in the UK identified their religion as Jedi?

Aggregate data are data about
populations, groups, regions or countries



Averaging individual level data



OECD Chart: Reading performance (PISA), Girls / Boys, Mean score, Annual, 2015

59



35



17



22



Totalling headcounts

Age: 16-24 years old – religion (England and Wales): Muslim – Unit: people

GEO_LABEL	GEO_TYPE	F103600	F103603	F103606	F103609	F103612	F103615	F103618	F103621	F103624	F103627	
		Age : Age 16 to 24 - Religion [E][S][W] : Sex : Total\ Sex - Unit : Persons										
Allerdale	Local Authorities	9230	6461	10	3	2	24	3	17	2143	567	
Barrow-in-Furness	Local Authorities	7570	4636	10	1	1	13	0	19	2419	471	
Carlisle	Local Authorities	11925	6903	36	37	2	45	8	40	4061	793	
Copeland	Local Authorities	7170	5307	16	5	3	22	1	6	1392	418	
Eden	Local Authorities	4758	2870	16	4	4	17	0	15	1486	346	
South Lakeland	Local Authorities	8822	5048	27	3	7	19	3	26	2986	703	

Office for National Statistics ; National Records of Scotland ; Northern Ireland Statistics and Research Agency (2017): 2011 Census aggregate data. UK Data Service (Edition: February 2017). DOI: <http://dx.doi.org/10.5257/census/aggregate-2011-2>

The Census of Population

- Every 10 years
- First modern census in 1841
- Data held secure for 100 years
- Digitised censuses from 1971 (1961 ongoing)
- Last ever census?

Census 2021

- England and Wales Census Day
 - Rehearsals have taken place already, good results
 - Address registers complete
 - Mainly online
- Northern Ireland's Census – – Sunday 21st March 2021
 - Mainly online supported by administrative data
- Scotland's Census – 2022
 - Date to be announced

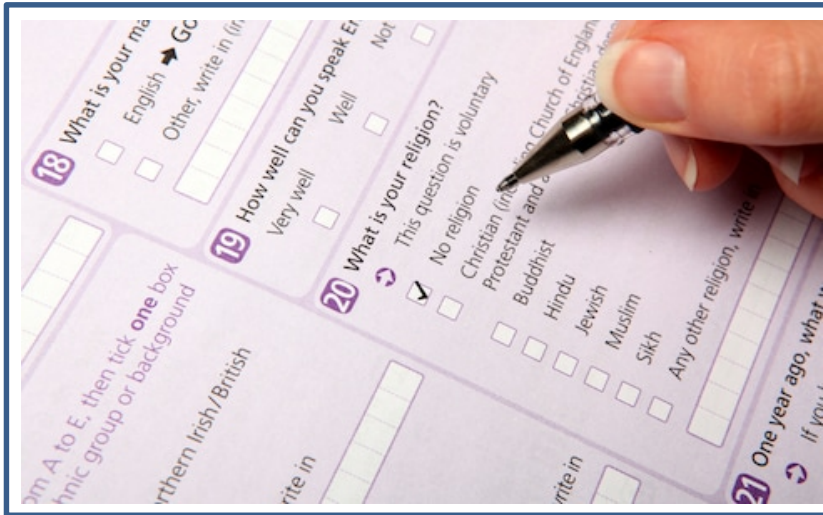
What can the census aggregate data tell us

- The most complete source of information about demographic and socio-economic characteristics of the UK population.
 - Population
 - Employment, occupation, socio-economic class, qualifications
 - Ethnicity, religion, national identity, language
 - Housing, tenure, residence type
 - Household and family composition, occupancy
 - Carers and caring
 - Long term health problems or disability
 - Travel to work, migration

What the census cannot tell us

- Wealth and income
 - Derived deprivation data
- Sexual orientation
- Personal identification
 - Data blurring and obfuscation

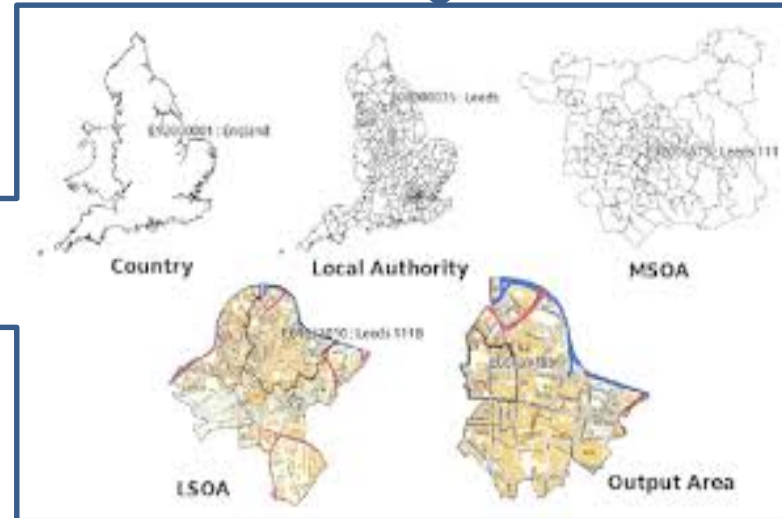
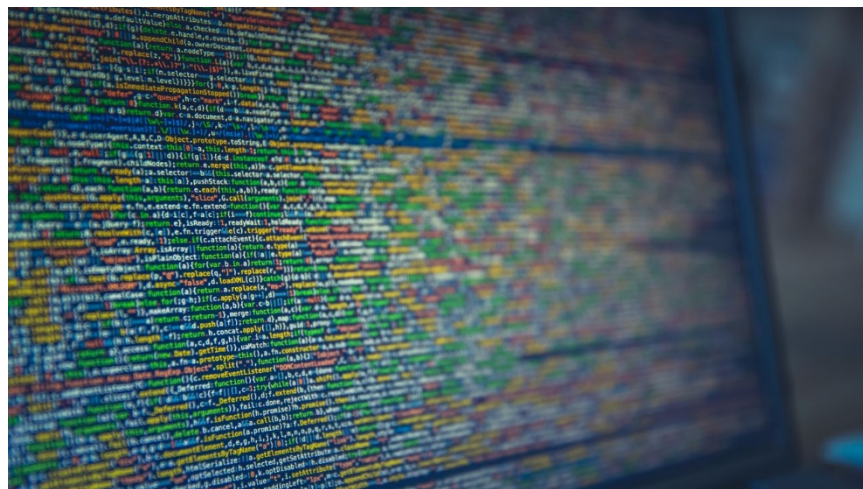
How are census aggregate data produced



Wiltshire Census 2011 Selected Statistics Profile Tool
 Produced by Wiltshire Council, May 2013 www.wiltshirecouncil.gov.uk/opa/opa-census

Wiltshire Council
 where everyone matters

Population	Profile A		Profile B		Profile C	
	Area	Trowbridge	Wiltshire	South West	Area	South West
Code	Community Area	Local Authority	Area	Local Authority	Area	Local Authority
	EW0600054	E06000054	E12000009	E12000009	E12000009	E12000009
	Number	%	Number	%	Number	%
Total usually resident population (1)	41,649	100.0	470,881	100.0	2,288,835	100.0
Total males	20,387	48.8	231,704	49.2	2,590,608	49.0
Total females	21,262	51.1	239,177	50.8	2,698,227	51.0
Age 0-4	2,793	6.7	28,373	6.0	298,094	5.8
Age 5-9	2,477	5.9	27,199	5.8	272,869	5.2
Age 10-14	2,856	6.8	29,276	6.2	296,892	5.8
Age 15-19	2,585	6.2	29,167	6.2	328,077	6.2
Age 20-24	2,366	5.7	24,720	5.2	333,166	6.3
Age 25-29	2,359	6.1	24,471	5.4	307,159	6.1
Age 30-34	2,078	6.2	20,078	5.3	296,791	5.8
Age 35-39	2,909	7.0	30,389	6.5	322,384	6.1
Age 40-44	3,286	7.8	36,130	7.7	376,838	7.1
Age 45-49	3,037	7.3	38,817	7.8	389,583	7.4
Age 50-54	2,873	6.9	32,704	7.0	360,433	6.8
Age 55-59	2,344	5.6	28,863	6.1	323,198	6.1
Age 60-64	2,096	6.2	31,248	6.6	361,507	6.8
Age 65-69	2,016	4.8	25,135	5.3	297,440	5.6
Age 70-74	1,810	3.9	18,935	4.2	236,462	4.5
Age 75-79	1,325	3.2	18,189	3.4	195,668	3.7
Age 80-84	947	2.3	12,887	2.8	150,083	2.9
Age 85-89	585	1.4	7,725	1.6	99,900	1.9
Age 90-94	295	0.8	3,343	0.7	41,548	0.8
Age 95-99	67	0.1	794	0.2	10,843	0.2
Age 100+	8	0.0	100	0.0	1,492	0.0
Lives in a household	41,286	99.1	470,881	97.2	2,179,084	97.8
Lives in a communal establishment (2)	363	0.9	13,027	2.8	113,851	2.2
Schoolchild or full-time student aged 4 and over at their non-term-time address	476	1.1	9,360	2.0	71,157	1.3
Total Children (Ages 0-17)	9,566	23.0	103,400	22.0	1,058,613	20.0
Ages 0-2	1,722	4.1	16,688	3.6	178,165	3.4
Pre-school (ages 3-4)	1,071	2.6	13,375	2.4	117,909	2.2
Primary school (ages 5-10)	2,979	7.2	32,700	6.9	328,054	6.2
Secondary school (ages 11-15)	2,682	6.4	29,895	6.3	306,530	5.8
College/6th form (ages 16-17)	1,112	2.7	12,432	2.6	129,939	2.4
Working age population (WAP Ages 16-64)	26,499	63.4	294,523	62.5	3,323,813	62.8
Retirement age population (RAP Ages 65+)	6,794	16.3	83,488	18.2	1,023,444	19.8



The Tricky Issue of Census Geography

- Building block is the Output Area
- These are used to create Super Output Areas (SOA)
 - Two types – Lower layer SOA (LSOA) and Middle layer SOA (MSOA)
 - In Scotland – Data Zones (DZ) and Intermediate Geographies (IG)
 - In Northern Ireland – just LSOA
- None of the above relate to anything ‘real’.
- Regions, Counties, Local Authorities, Wards and Electoral Divisions
- <http://geoconvert.digitalresources.jisc.ac.uk/>

A UK Census?

- Different nations, different governments, different ideas.
 - Office for National Statistics
 - National Records of Scotland (formerly GROS)
 - Northern Ireland Statistics and Research Agency
- Different Geographies
- Different Questions
- Different Outputs

Accessing census aggregate data

- Infuse - <http://infuse.ukdataservice.ac.uk/>
 - Data from 2001 and 2011 censuses
- Casweb - <http://casweb.ukdataservice.ac.uk/>
 - Data from 1971, 1981, 1991 and 2001 censuses
 - Boundary data with 1991 & 2001 censuses
- DKAN - <https://www.statistics.digitalresources.jisc.ac.uk/>
 - Currently only 2011 but currently loading 2001 and plans for censuses back to 1971

Activity 1

InFuse - <http://infuse.ukdataservice.ac.uk/>

Worksheet -

https://stats2.digitalresources.jisc.ac.uk/training/handout_1.pdf

10 minutes.

Deprivation data

- No questions on income or wealth.
- Derived from room occupancy, house ownership/tenancy, car availability, employment status etc...
- A number of recipes
 - Carstairs
 - Townsend
 - Index of Multiple Deprivation/Indices of Deprivation
- <https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019>

Matching Geographies

- GeoConvert - <http://geoconvert.digitalresources.jisc.ac.uk/>
 - Uses address points to calculate an area population
 - Clever proportioning
 - Caveats
 - Think about what you are attempting to do.
 - Postcodes change
- Supporting documentation - <http://geoconvert.digitalresources.jisc.ac.uk/help/documentation.html>

Activity 2

InFuse - <http://geoconvert.digitalresources.jisc.ac.uk/>

Postcode file -

<http://stats2.digitalresources.jisc.ac.uk/training/postcodes.csv>

Data file - <http://stats2.digitalresources.jisc.ac.uk/training/data.csv>

Census bulk data

- DKAN - <https://www.statistics.digitalresources.jisc.ac.uk/>
- Whole tables for geographical areas
- Easy to search
- Metadata
- Expanding all the time

Activity 3

DKAN - <https://www.statistics.digitalresources.jisc.ac.uk/>

Find the “Age by Ethnic group by Long-term health problem or disability by Sex (England and Wales) 2011” table.

What ‘Units’ are used?

What is the definition of a “Long-term health problem or disability”?

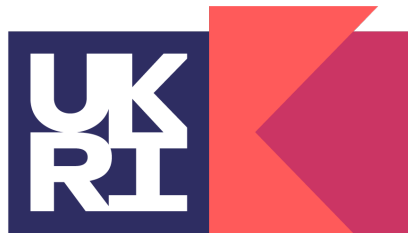
For England and Wales what is the total number of men in the Asian /Asian British category that considered themselves to have a Long-term health problem or disability?

Resources



What is the UK Data Service?

- A comprehensive resource funded by the Economic and Social Research Council (ESRC)
- A single point of access to a wide range of secondary social science data
- Support, training and guidance




**Economic
and Social
Research Council**



Online resources

- Guides
- Videos
- Webinars
- Data Skills Modules

About us	Get data	<u>Use data</u>	Manage data	Deposit
Advice and training			Advice and training	
Guides			Locate resources and get the help you need	
Video tutorials				
Data in use				
Citing data				
Secure Lab				
Teaching with data				
Student resources				
Data Skills Modules				
Secondary analysis				

Methods and Software Guides

For example:

- **Software guides**
 - What is Stata?
 - What is SPSS 20 for Windows?
 - Using R to analyse key UK surveys
- **Survey data**
 - What is weighting?
 - What is complex sample design?
 - What are hierarchical files?



Contents

1. Introduction.....	4
1.1. What is R?.....	4
1.2. The pros and the cons of R.....	5
2. Using R: essential information.....	7
2.1. Installing and loading user-written packages.....	9
2.2. Getting help.....	10
2.3. Interacting with R: command line vs graphical interface.....	12
2.4. Objects in R.....	13
2.5. Deleting object from the R environment using the rm() function.....	14
2.6. Saving in R.....	14
3. R Studio.....	17
4. Opening UK Data Service datasets in R.....	19
4.1. Which version of UK Data Service datasets can I use in R?.....	19
4.2. Variables types.....	21
5. Essentials of data manipulation.....	23
5.1. Identifying and selecting variables and observations.....	23
5.2. Creating and recoding variables.....	24
5.3. Renaming variables and categories.....	25
5.4. Missing values.....	26
6. Descriptive statistics using R - continuous variables.....	29
6.1. Distribution graphs.....	30
6.2. Significance testing of correlation.....	33
6.3. Tables of summary statistics.....	33
6.4. Bar charts and plots of summary statistics.....	35
6.5. Saving and importing a graph in a Word document.....	38
6.6. Weighted descriptive statistics.....	39
7. Categorical variables: contingency tables.....	41
7.1. One way frequency tables.....	41
7.2. Creating bar plots of one-way frequency tables.....	41
7.3. Two way or more contingency tables.....	42
7.4. Test of association between categorical variables.....	45
7.5. Univariate and bivariate graphs for categorical variables.....	46
8. Plotting simple maps in R.....	48
8.1. Acquiring and downloading the data.....	48

Videos

"Can you show me
how to do that?"



SHARE 

YouTube: <https://www.youtube.com/user/UKDATASERVICE>

Data Skills Modules

- Survey data
- Aggregate data
- Longitudinal data
- Glossary

<https://www.ukdataservice.ac.uk/use-data/data-skills-modules.aspx>

Surveys unit 1: What are survey data?

[View all modules](#)

[Menu](#) | [Glossary](#)

Surveys

Unit 1: What are survey data?

This unit will explore survey data.

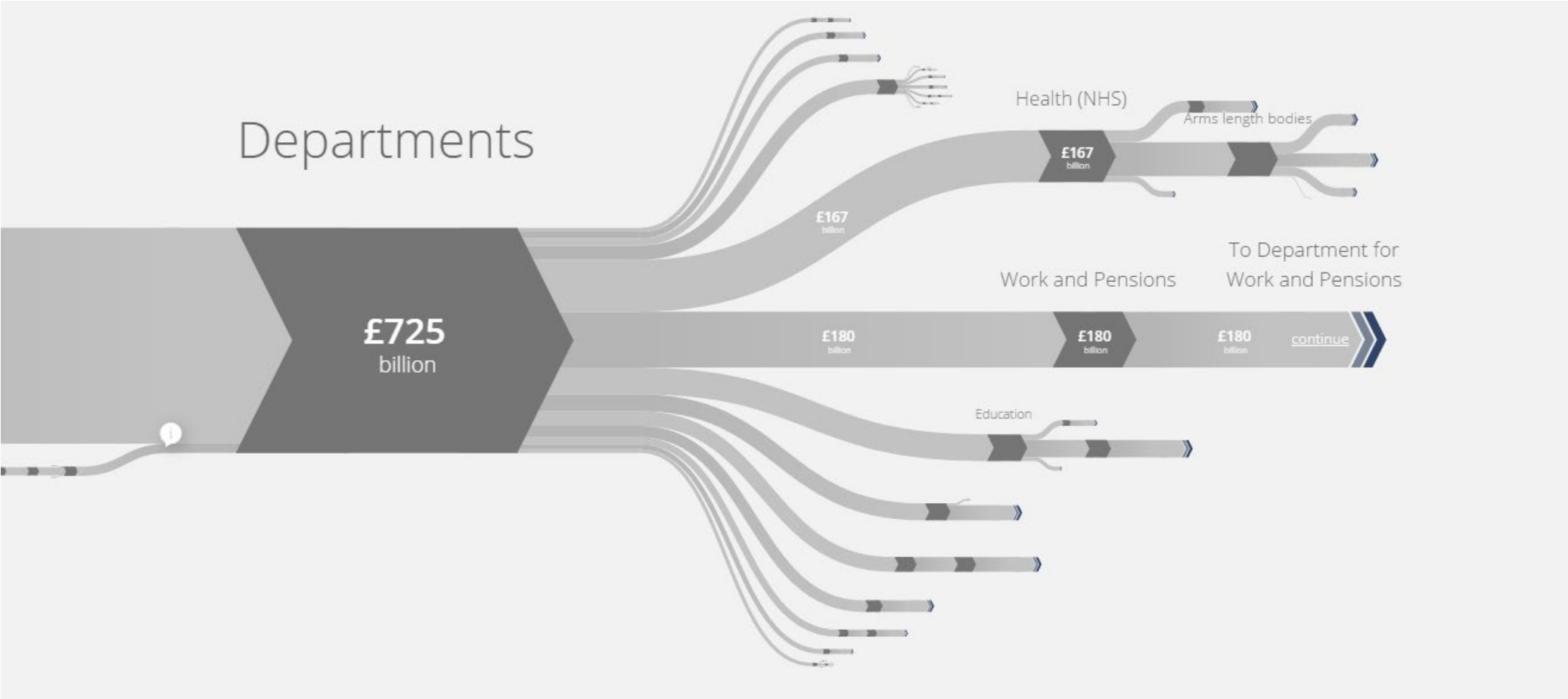
You will look at what the data are, some key concepts relating to survey data and how to find datasets that you can use.

You can work through this unit in order using the next button, or use the menu to view the list of sections and jump to the part you are most interested in.



NEXT >

Aggregate data: shaping policy



<https://uk.wikibudgets.org/w/united-kingdom-budget-2015>

Thank you

web: <https://ukdataservice.ac.uk/> YouTube: <https://www.youtube.com/user/UKDATASERVICE>

email: help@ukdataservice.ac.uk

twitter: @UKDataService

Visit the
UK Data Service



The screenshot shows the UK Data Service website homepage. The browser address bar displays <https://ukdataservice.ac.uk>. The navigation menu includes links for 'About us', 'Get data', 'Use data', 'Manage data', 'Deposit data', and 'News and events'. The main header features the UK Data Service logo, social media icons for LinkedIn, Facebook, Twitter, YouTube, and Instagram, and a 'Register / Logi' button. The main content area has a dark blue background with the text 'Explore the UK's largest collection of social, economic and population data resources' and a search bar with a 'Search data' dropdown and a magnifying glass icon. Below this, there are two columns: 'How to cite data correctly' with a video thumbnail titled 'Why cite data? #CiteTheData' and 'Guides and resources' with a list of links: 'Dataset guides', 'Topic guides', 'Methods and software guides', and 'Guides to exploring online'. A 'See more' link is at the bottom of the guides list. A 'Video tutorials' section is visible on the right side of the page.

