



National Audit Office

Applying data analytics to performance and financial audit: the NAO experience

Michael Kell
Chief Economist, NAO

Overview

- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion

The NAO

- Established in 1866
- Independent of executive, reports to Parliament, no judicial authority
- 800 staff
- Financial audit: 370 accounts certified, covering £1.6trn; set standards for local government audit
- Performance audit: 65 VfM reports, 40-50 PAC hearings
- Investigations: 40-50 reports

The case for data analytics

Productivity

Doing our work faster

Quality

Fewer mistakes, greater consistency

Insight

Creating new insights

Impact

Better communication

Overview

- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion

Webscrapping – OFSTED data on pupil premium

1. Programmatically navigate to school page

Ofsted raising standards improving lives

Register / Login

High contrast version

Find an inspection report About our inspection reports GOV.UK/Ofsted Parent View

Nelson Castercliff Community Primary School
Unique Reference Number (URN): 119281

Marsden Hall Road North
Nelson
Lancashire
BB9 8JJ

Telephone number: 01282617627
Fax number: 01282693526

Download the latest school inspection report

Overall effectiveness **Inadequate**
Inspection date 7 May 2014
First publication date 3 Jul 2014

Data Dashboard

Report	Inspection date	First publication date
Section 8 inspection report	9 Jun 2015	1 Jul 2015
Section 8 inspection report	20 Jan 2015	11 Feb 2015
Section 8 inspection report	11 Sep 2014	2 Oct 2014
School inspection report	7 May 2014	3 Jul 2014
Section 8 inspection report	15 Jan 2013	5 Feb 2013

2. Identify link to latest inspection report within HTML and download

Ofsted raising standards improving lives

Nelson Castercliff Community Primary School
Marsden Hall Road North, Nelson, Lancashire, BB9 8JJ

Inspection dates 7-8 May 2014

	Previous inspection:	This inspection:	Requires improvement
Overall effectiveness	Inadequate	Inadequate	3
Achievement of pupils	Inadequate	Inadequate	4
Quality of teaching	Inadequate	Inadequate	4
Behaviour and safety of pupils	Inadequate	Inadequate	4
Leadership and management	Inadequate	Inadequate	4

Summary of key findings for parents and pupils

This is a school that requires special measures.

- The school's work to safeguard pupils is inadequate. Leaders' and governors' checks on safeguarding fail to meet government requirements.
- Pupils in all year groups across the school make inadequate progress. Standards by Year 6 in English and mathematics have been very low for several years and show too little sign of improving.
- Pupils' key numeracy and literacy skills are poor. They do not use and apply them well enough in other subjects.
- Teaching is inadequate. Expectations are too low. The most able pupils are not challenged enough. Pupils' work and progress are not properly checked.
- Pupils' achievement is held back in a range of subjects because the tasks that teachers plan do not build successfully on pupils' previous knowledge.
- Pupils' behaviour requires improvement because in lessons, pupils' attitudes to learning are not always good. They do not always listen to their teachers and pupils' work is often poorly presented.
- A record of decline in the school's performance since the last inspection reflects the inadequate ability of the leadership to make the necessary improvements.
- Leaders have failed to ensure adequate procedures are in place for checking and improving the performance of teachers.
- Methods for tracking pupils' progress are weak, consequently, both teaching and achievement have fallen to inadequate levels.
- Governors do not hold leaders to account sufficiently, or challenge them robustly enough about the school's performance, including the impact of the pupil premium funding.

The school has the following strengths:

3. Search text for refs to pupil premium

ensure adequate for checking and nce of teachers.

Methods for tracking pupils' progress are weak, consequently, both teaching and achievement have fallen to inadequate levels.

■ **Governors do not hold leaders to account sufficiently, or challenge them robustly enough about the school's performance, including the impact of the pupil premium funding.**

■ The recently appointed headteacher has accurately evaluated the school's strengths and weaknesses and has started to make improvements.

4. Copy all paragraphs out for further analysis

urn	School name	refs to 'pupil premium'
137816	Eston Park Academy	<ul style="list-style-type: none"> The performance of students known to be eligible for free school meals varies too much between subjects. The academy does not use its pupil premium funding effectively to ensure these students achieve as well as they should, particularly in English. The proportion of students known to be eligible for the pupil premium (additional funding for those pupils known to be eligible for free school meals, children from service families and those children that are looked after) is above average. The pupil premium funding is used to provide additional support and resources. However, these are not targeted precisely enough on the students for whom the funding is intended. As a result, the achievement of these students is too variable. In English, for example, students known to be eligible for free school meals underachieve in comparison with other students and with similar students nationally. The governing body manages the academy's finances efficiently, although the impact of pupil premium funding and the link between pay progression and performance are not being checked closely enough. The governing body makes sure that safeguarding requirements are met.

...then simply repeat (7,000 times!)

Webscrapping – Readability of Tax Guidance

GOV.UK



Income Tax

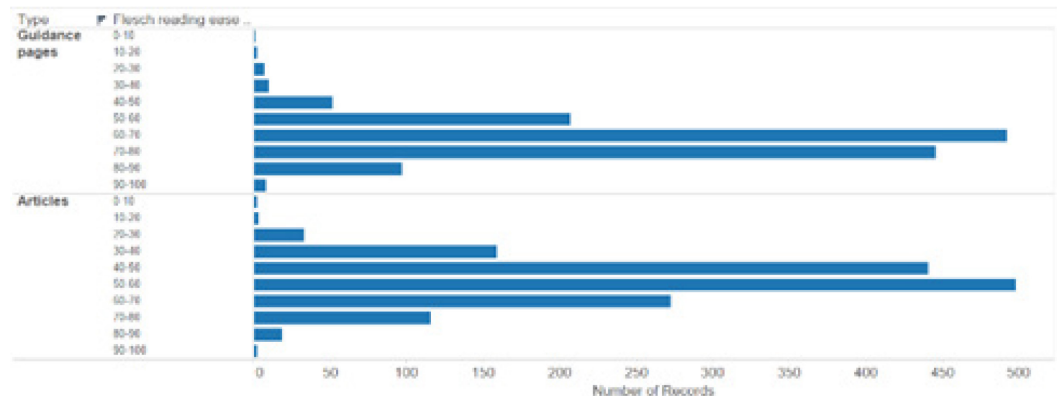
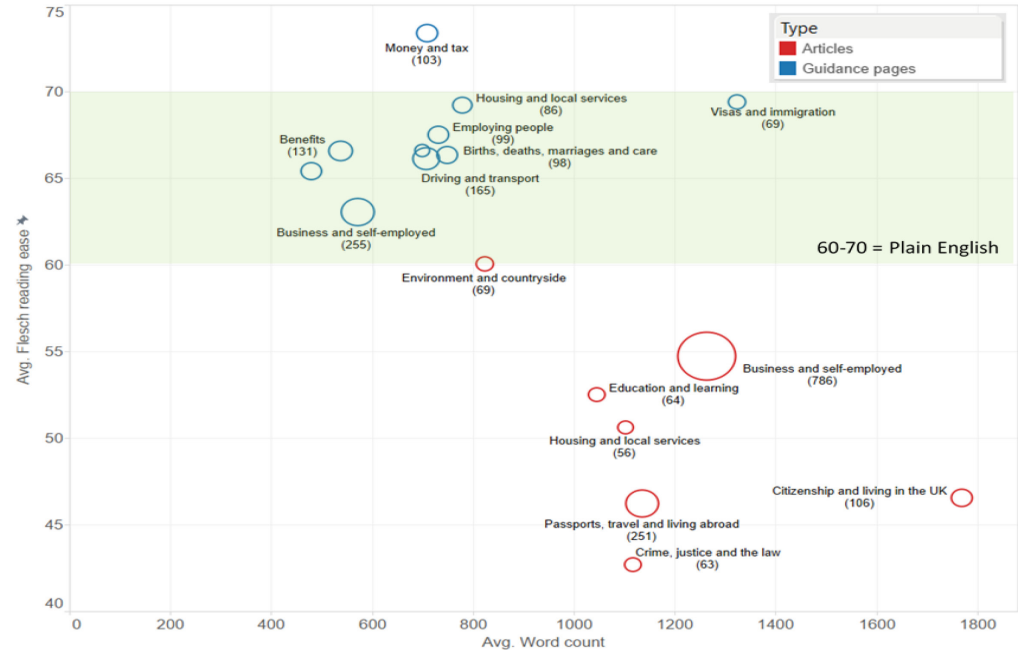
1. Overview
2. [How you pay Income Tax](#)
3. [Tax-free and taxable state benefits](#)
4. [Work out if you need to pay Income Tax](#)
5. [Check you're paying the right amount](#)

1. Overview

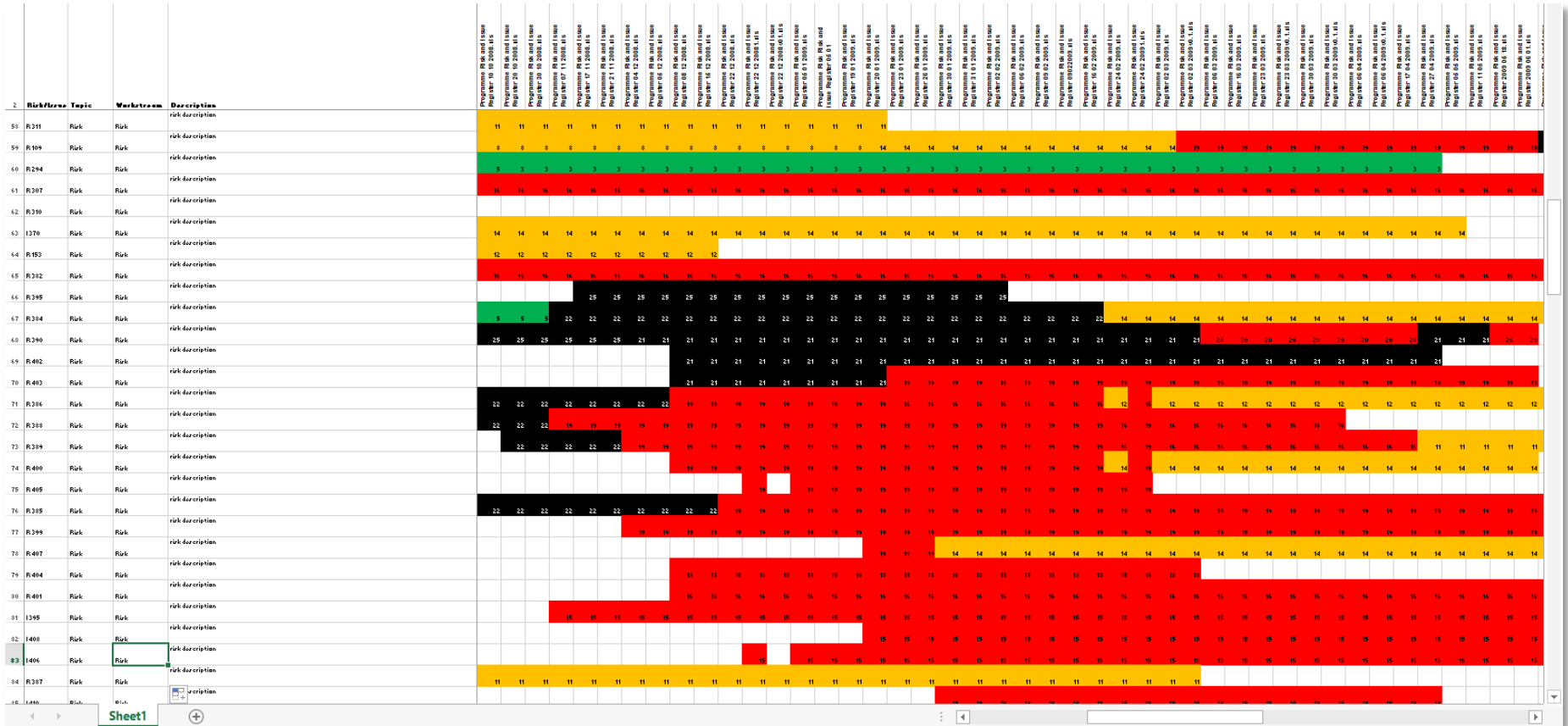
Income Tax is a tax you pay on your income. You don't have to pay tax on all types of income.

You pay tax on things like:

- money you earn from employment
- profits you make if you're self-employed - including from services you sell through [websites or apps](#)



Creating insight with risk logs



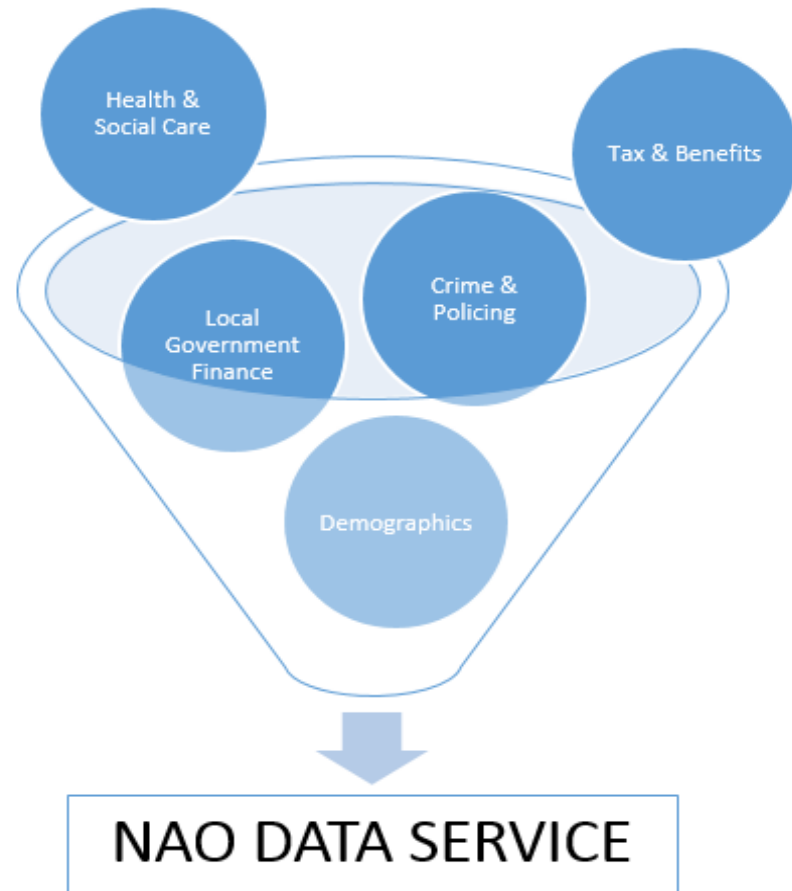
NAO Data Service



Client datasets as published -
insular and often hard to use



Client data within service -
linked and analytics ready



NAO Data Service

Entry point to service is through dedicated Share Point site...

The screenshot shows a web browser window displaying the NAO Data Service SharePoint site. The address bar shows the URL <http://naodataservice.nao.gsi.gov.uk/>. The page has a blue header with the SharePoint logo and navigation links. The main content area is titled 'Home' and includes a welcome message and a list of data themes on the left sidebar. The 'Featured Report' section displays a horizontal bar chart titled 'Local Gov Revenue: Spend per head in the Inner London Boroughs' for the year 2011. The 'Contact Details' section lists the project team and data leads. The 'Announcements' section shows recent updates.

Home

Welcome to the NAO Data Service. To start exploring the data, select the theme you are interested in from the menu on the left.

Alternatively, if you know which dataset you'd like to use, simply select a spreadsheet from the Report Templates folder.

*** The Data Service has now been expanded to include data about Crime, Policing and Justice. These datasets add to existing data within the service about Local Authority Finance, Health and Social Care and Demographics. ***

*** Datasets relating to Tax and Benefits will be added shortly. ***

Featured Report

Local Gov Revenue: Spend per head in the Inner London Boroughs

Borough	2011 Spend per head
Camden	~100
Greenwich	~80
Hackney	~100
Hammer-smith and Fulham	~150
Islington	~100

Year

2013
2012
2011
2010
2009
2008
2007

Level 02

EDUCATION SERVICES
HIGHWAYS AND TRANSPORT SERVICES

Contact Details

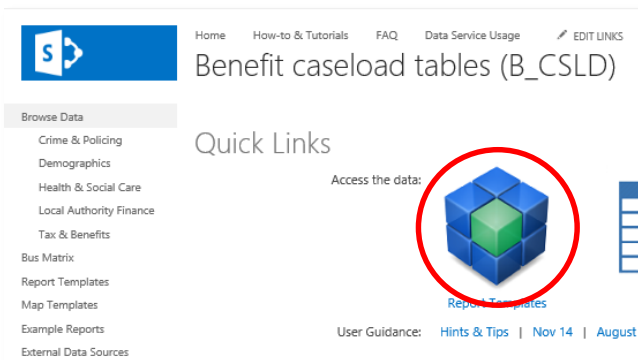
- Project team: Ben Coleman, Joanna Taylor, Heather Reeve-Black
- Data leads by theme:
 - Local Authority Finance: Mark Burkett
 - Health & Social Care: Will Palmer (Health), Mark Burkett (Social Care)
 - Crime & Policing: Rob Sabbarton
- Technical faults: Z3-Data Service Support

Announcements

Title	Modified
Minor revisions to historic recorded crime data by ONS - to be added to the data service shortly	15 Apr
Labour market stats revised by ONS - to be added to service shortly	26 Mar

NAO Data Service

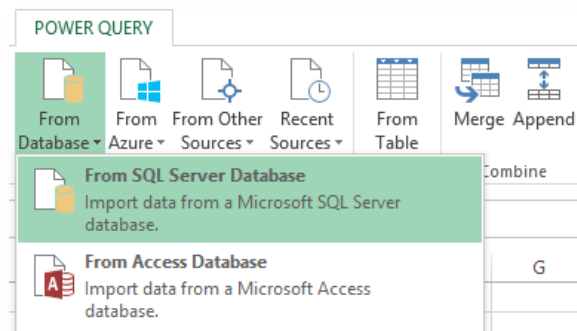
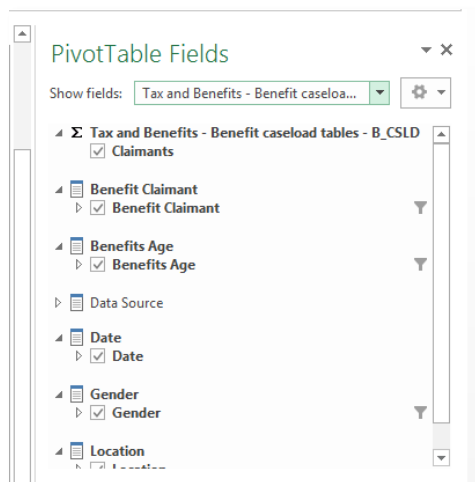
Access to data is via cube icon...



Row Labels	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Greater Britain	6,391,800	5,175,400	5,804,800	6,188,900	5,186,200	4,202,800	4,202,800	5,101,500	5,881,000	5,881,200	5,762,500	5,423,500	5,423,500
England	4,808,700	4,236,400	4,479,500	4,195,700	4,138,600	4,804,900	4,384,000	4,754,800	4,506,000	4,404,400	4,754,000	4,444,500	4,729,400
London	712,100	702,500	734,300	734,600	738,500	741,200	738,000	705,800	721,600	721,600	721,600	721,600	721,600

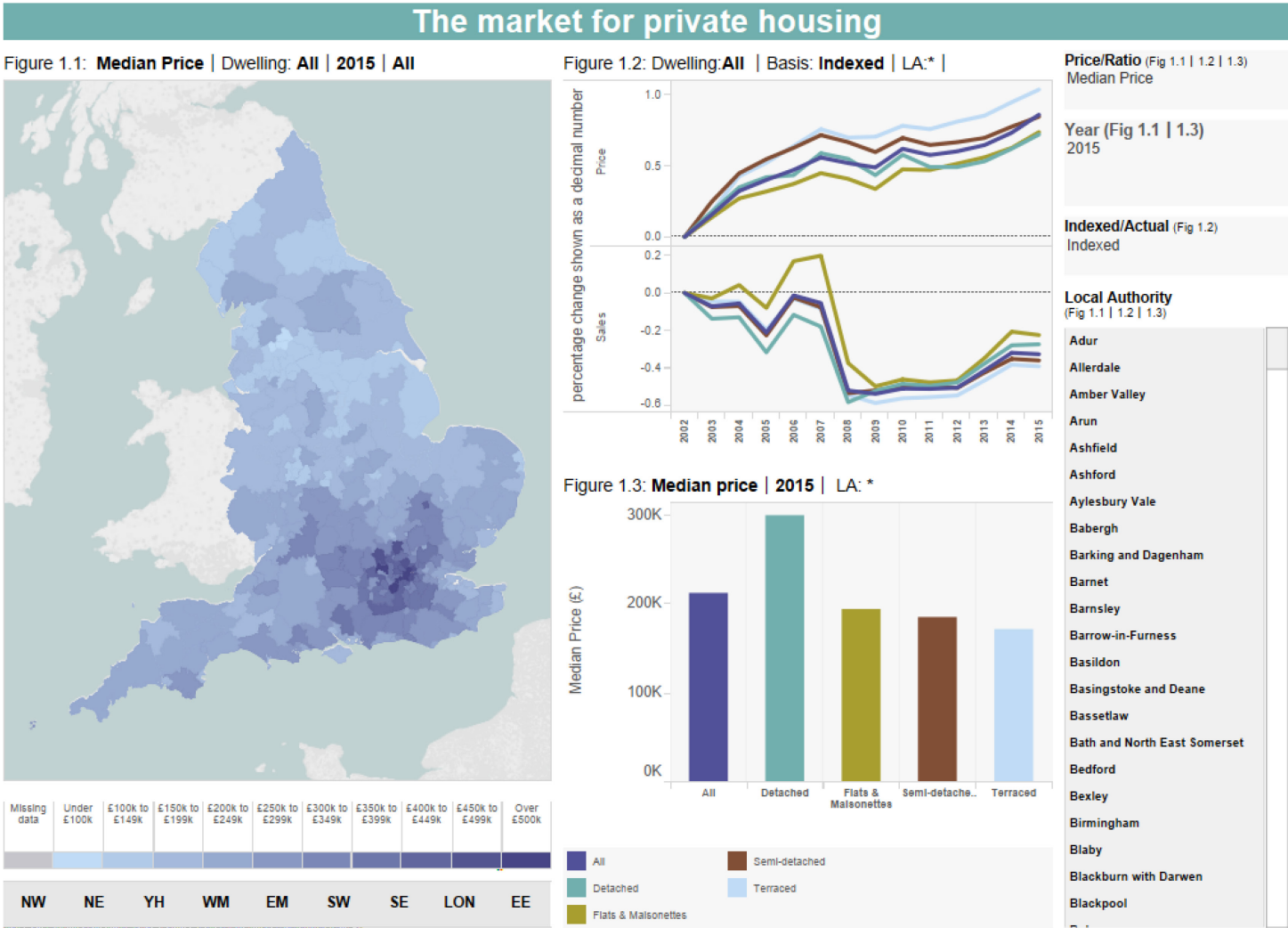
...providing access to data in pivot table form

...with the ability to easily customise & link to other data in service

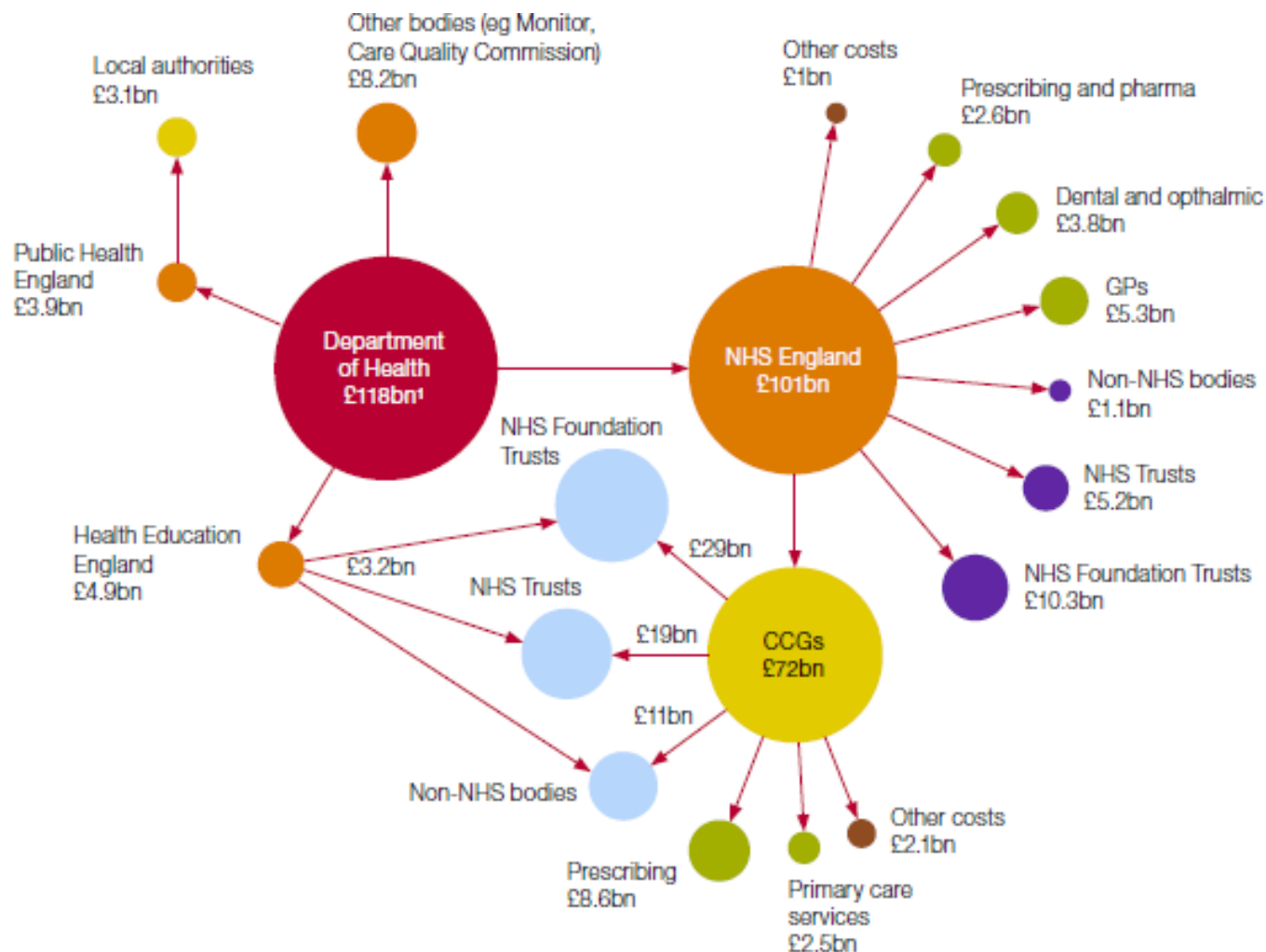


Alternatively, advanced users can connect to the SQL database directly

Interactive visualisations



Simple visualisations: bubble diagrams



Harvesting our past reports

Search Past NAO Reports

Search Criteria About

Enter search criteria below. Summary results appear below, detailed extracts and analytics to the right.

Report Section:
All

Year:
All

Search term:
Procurement

Number of matches found:
2452

Top reports matched:

Report	# matches
Improving government procurement	162
Police procurement	124
The Government Procurement Card	100
Procuring new trains	91
The procurement of consumables by NHS acute and Foundation trusts	87
Major Projects Report 2014 and the Equipment Plan 2014 to 2024	75
Major Projects Report 2015 and the Equipment Plan 2015 to 2025	69
Reducing the cost of procuring Fire and Rescue Service vehicles and	69

Matches Analytics

Matching report extracts shown below. Use controls to further filter, sort or download your results.

Show 5 rows Copy CSV PDF

Showing 1 to 5 of 2,452 entries

Previous 1 2 3 4 5 ... 491 Next

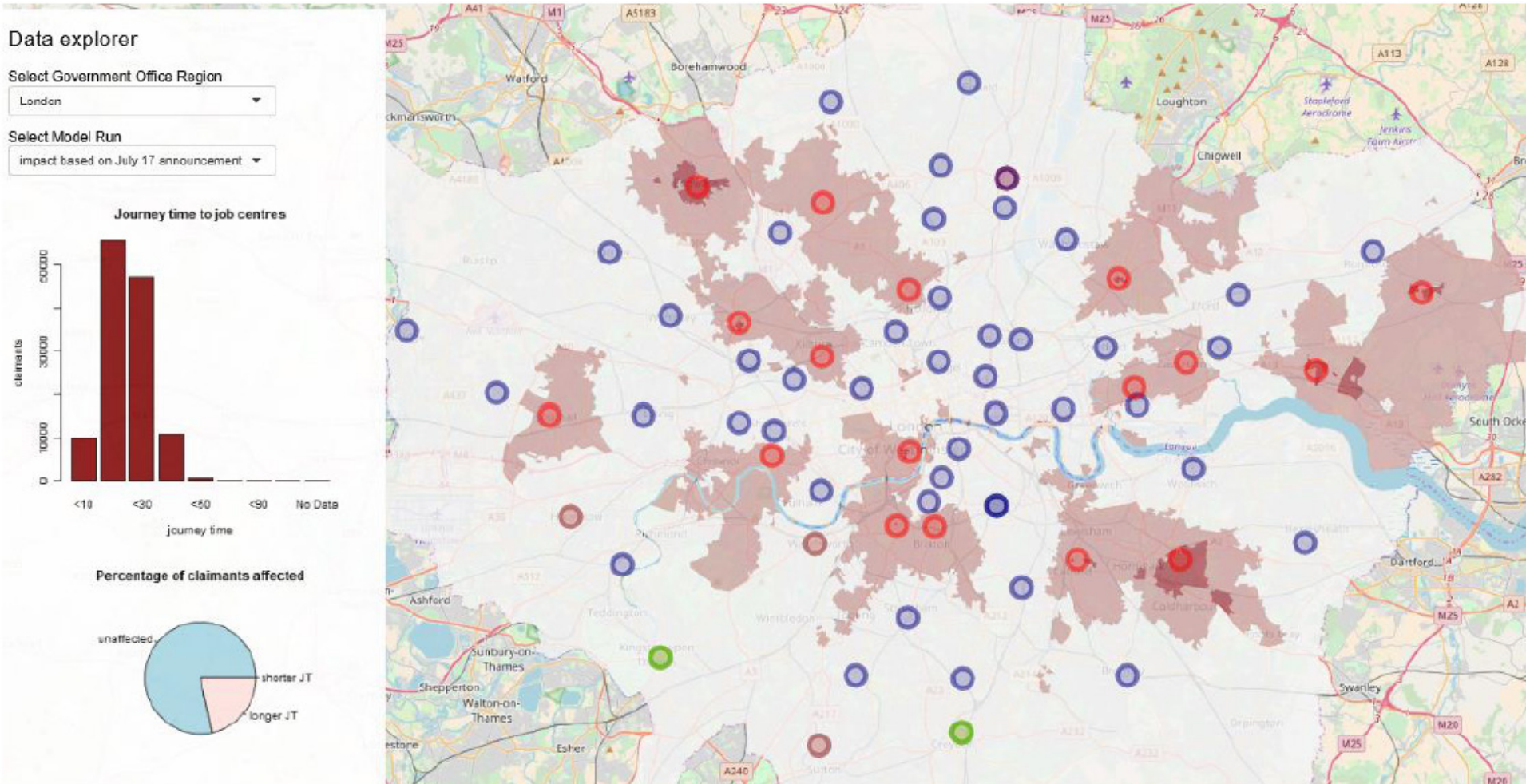
Year	Report	Section	Subsection	Paragraph Text
All	All	All	All	All
2015-16	Disposal of public land for new homes	summary	Land ownership	In the case of the site at Blackwall Reach, Swan Housing Association noted that although GLA and Tower Hamlets have taken steps to acquire the relevant titles prior to tendering the development, there is a title issue yet to be resolved at the time of our interview over land that will be needed in the final phase Case study 1. Similarly, the site at Manor Kingsway is owned by several governmental bodies. During the procurement process, HCA has indicated that there was an agreement between all the landowners committing them to support the development. However, the developer later discovered that as this was only a memorandum of understanding. It has since encountered difficulties in negotiating with landowners Case study 2.
2015-16	Disposal of public land for new homes	appendices	Observations	The land is owned by several governmental bodies. During the procurement process, HCA indicated that there was an agreement between all the landowners committing them to support the development. However, Kier Partnership Homes later discovered that this is no more than a memorandum of understanding and noted that some of the landowners seemed to be reluctant participants, leading to some difficulties in the negotiation process, which significantly delayed development. Kier Partnership Homes recommended that the agreement could be better drawn up to align the interest of all the third party landowners. Additionally, there should be a single public body empowered to represent all the public sector landowners.
2015-16	Early review of the Common Agricultural Policy Delivery Programme	main content	The CAP Delivery Programme	1.11 Procurement and development of the system took place during 2013 and 2014. The Programme was working towards immovable deadlines the service had to be ready to accept BPS claims by 15 May 2015 and to make BPS payments to customers between 1 December 2015 and 30 June 2016. Failure to meet these deadlines would incur disallowance penalties. Figure 3 sets out the key events. 3.24 GDS committed to reducing overall costs and improving delivery confidence. However, no formal estimates of cost reductions were provided and the Major Projects Authority's delivery confidence assessments did not improve. The Memorandum of Understanding agreed between

100%

Topic Modelling (Machine learning)



Spatial Analysis

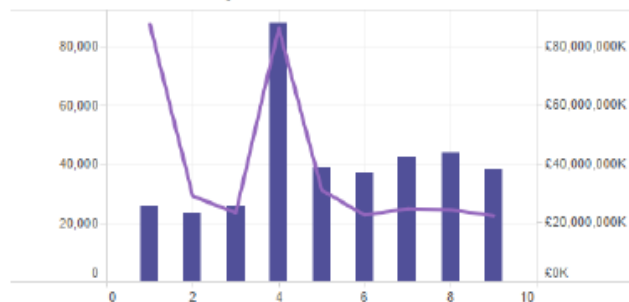


Overview

- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion

Journal Analytics Dashboards

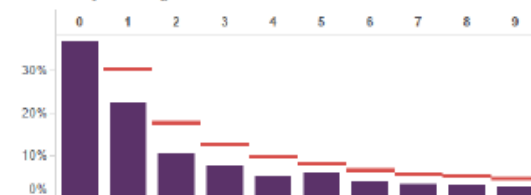
Number Lines & Total Value by Period



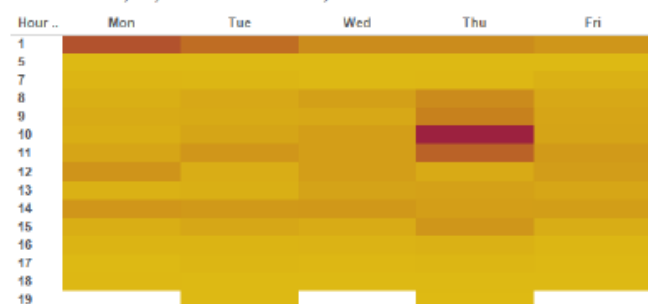
Number Lines & Total Value by Preparer



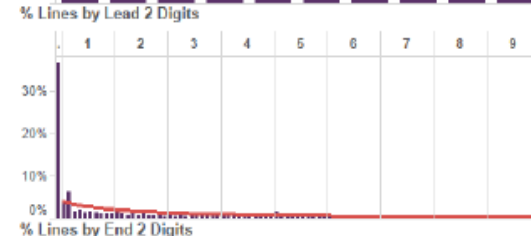
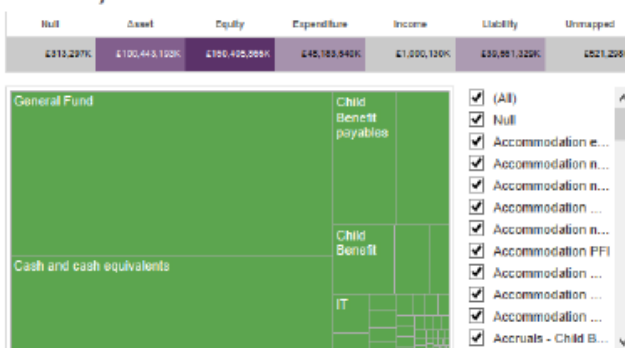
% Lines by Lead Digit



Number Lines by Day of Week & Time of Day



Total Value by FS Element & Line



Round Number Values

Multiple	Value
0.01	282,356
1	48,897
10	19,458
100	8,726
1000	3,323
10000	1,631
100000	478
1000000	70

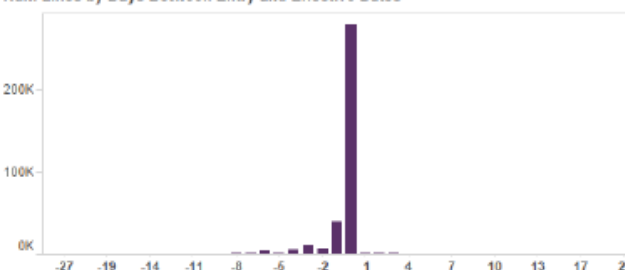
Most Frequent Values

Amount (bin)	Value
£0.00	132,990
-£115.00	14,871
-£100.00	3,947
£50.00	1,945
£30.00	1,293
£25.00	1,279
-£10.00	1,175
£100.00	948
-£25.00	940
-£20.00	927

Highest Value Journals

AmtL_For_View (bin)	Value
£31,208,187,257	2
£28,542,772,040	2
£2,333,003,101	2
£1,954,213,760	1
£1,894,582,001	1
£1,887,804,500	1
£1,865,843,028	1
£1,826,252,436	1
£1,787,030,773	1
£1,778,725,074	1

Num Lines by Days Between Entry and Effective Dates



Type of journals to view: Credits & Debits (ABS)

Specify value range: 0 to 31208187257

Search line description: [Input field]

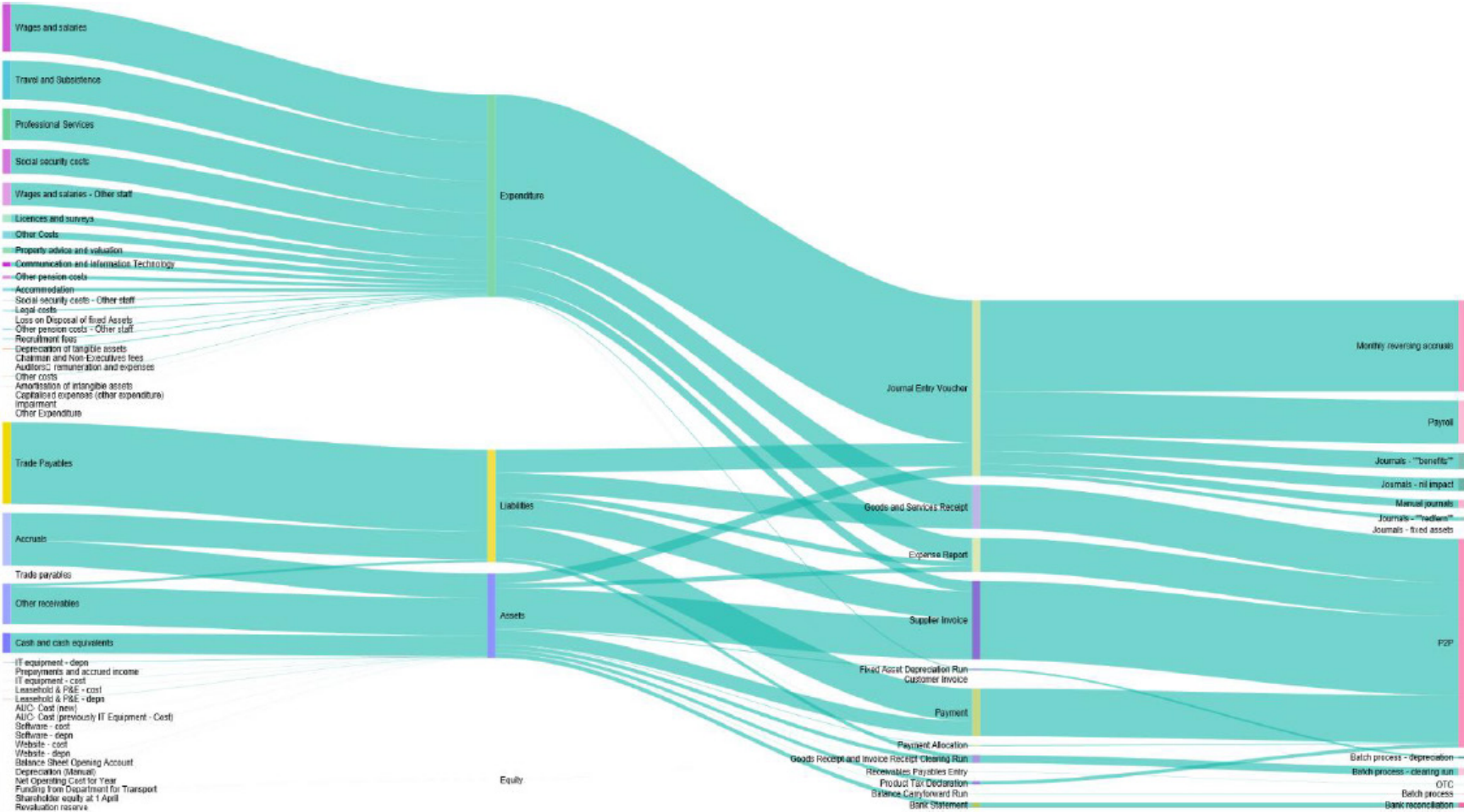
Description length: 0 to 50

Checksum value: 350,528,653,703

Number of lines selected (click button to view): 363,039

[Green Arrow Button]

Sankey diagrams



Improved testing of payroll

Payroll AnalyticsInstructionsData EntryAnalyticsPayroll TestsException ReportSummary Report Creation

Search Income tax lines in payroll

income

Materiality Base

1

Materiality

1

Performance Materiality

1

Employer Pension Contributions as a percentage (e.g. 15% would be 0.15)

0.12

GL Basic Pay Figure

1

GL Income Tax

1

GL National Insurance Employer

1

GL National Insurance Employee

1

Staff Costs Prediction

Generated on 2017-02-02

Instructions

This is a staff costs prediction based on the pay details provided on the data entry page, month period, and allow you to compare the ledger values to the prediction, calculating methodology.

Calculations are done at the employee level, and comparisons are made on an employee.

Based on the payroll data provided, the total payments made in year will be compared to expectation.

Please note that as this prediction is based on your full payroll listing, there should not be a ledger. Where this page produces large variances, you should consider whether there are pay rates that may account for this change. Tax rates are set as per HMRC guidance for [Income Tax rates 2016-17](https://www.gov.uk/income-tax-rates) [National Insurance rates 2016-17](https://www.gov.uk/national-insurance-rates-letters)

Basic pay prediction based on

[1] "56,190,236"

Basic pay from Payroll Data - this is

[1] "4,519,500"

Tax Calculations based on

Income Tax

[1] "11,080,302"

Payroll AnalyticsInstructionsData EntryAnalyticsPayroll TestsException ReportSummary

Ratio 1 to be applied

1.251.52

Ratio 2 to be applied

1.251.52

Download Exception Report

Payroll Exception Testing

NAO Payroll Data Analytics Application

Generated on 2017-02-02

Instructions

A series of tests have been run between the data input on the data entry page and the payroll data. The results of the tests are shown on this page, sample from the exception lists for each test, due to size)

Ratio Tests

Test whether fails payment ratio 1

[1] "All passed Ratio Test"

SmartExporter

SmartExporter Desktop Edition

/Home/Data Requests/Create Data Request/5. Fields and Filters

Home

- 1. Introduction
- 2. Connection
- 3. Select Tables
- 4. Quick-Filters
- 5. Fields and Filters
- 6. Joins
- 7. Options
- 8. Finish

Specify which data to include in the table

Verify Check size Preview

	Name	Row Count	Size (MB)	Description
<input checked="" type="checkbox"/>	BKPF	0	0	Accounting Document Header
<input checked="" type="checkbox"/>	BSEG	0	0	Accounting Document Segment
<input checked="" type="checkbox"/>	LGIR DT	0	0	Document type descriptions in Journal


Fields Filters

Fields - BSEG

These are the fields that will be retrieved from the SAP® system while running the Data Request.

Search: ☒ Show selected fields

	Name	Description
<input checked="" type="checkbox"/>	BUKRS	Company Code
<input checked="" type="checkbox"/>	BELNR	Accounting Document Number
<input checked="" type="checkbox"/>	BUZEI	Number of Line Item Within Accounting Document
<input checked="" type="checkbox"/>	BUZID	Identification of the Line Item
<input checked="" type="checkbox"/>	AUGBL	Document Number of the Clearing Document
<input checked="" type="checkbox"/>	SHKZG	Debit/Credit Indicator
<input checked="" type="checkbox"/>	GSBER	Business Area
<input checked="" type="checkbox"/>	DMBTR	Amount in Local Currency
<input checked="" type="checkbox"/>	WRBTR	Amount in document currency
<input checked="" type="checkbox"/>	SGTXT	Item Text
<input checked="" type="checkbox"/>	BEWAR	Transaction Type
<input checked="" type="checkbox"/>	KOSTL	Cost Center
<input checked="" type="checkbox"/>	AUFNR	Order Number
<input checked="" type="checkbox"/>	VBELN	Billing Document
<input checked="" type="checkbox"/>	VBEL2	Sales Document
<input checked="" type="checkbox"/>	DOCEN	Sales Document Item

 **Certificate**
SAP INTEGRATION CERTIFICATION

SAP SE hereby confirms that the ABAP Interface software for the product:
SmartExporter 4.0

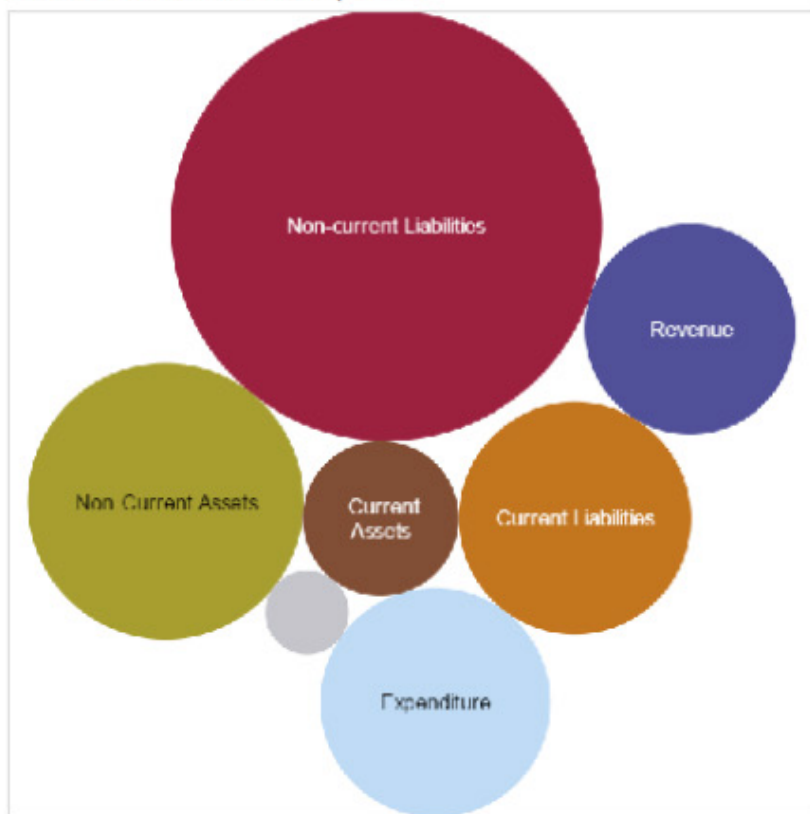
Efficiency in eliminations

The image displays a complex financial dataset in Excel, with multiple overlapping spreadsheets. The top spreadsheet, titled 'BalanceSheetDR - Excel', shows a summary of values with a yellow highlight on a cell containing '129,735'. Below it, another spreadsheet shows a 'Total No Match' cell. The background shows a large, detailed financial statement with many rows and columns of data, including various financial metrics and codes.

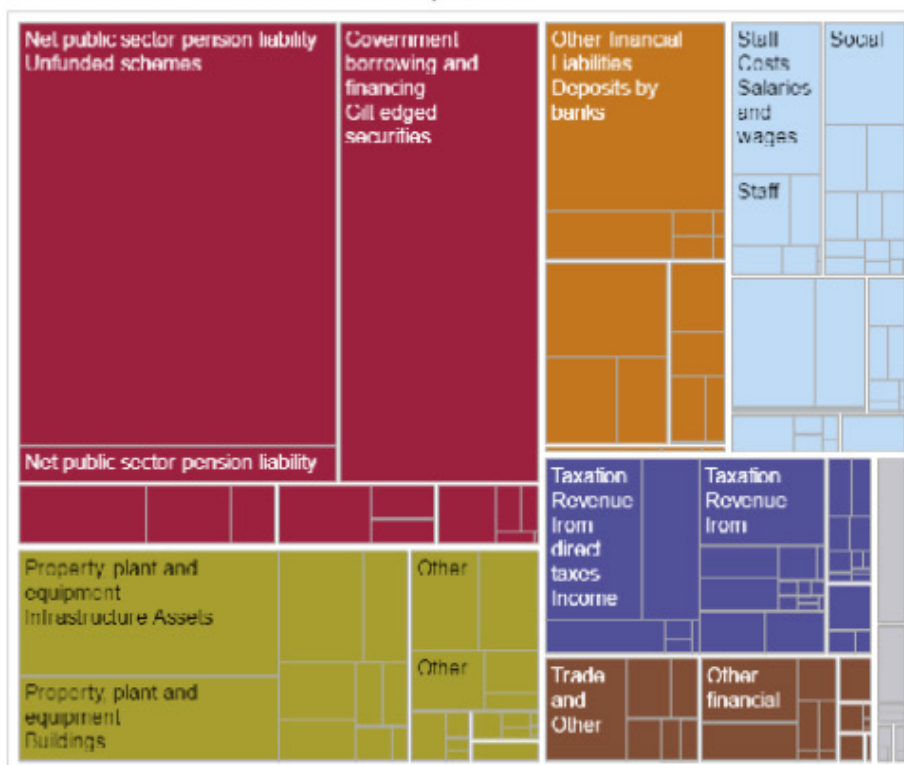
```
1 library(data.table)
2 match <- read.csv("H:/WGA/MR/Match14.csv", header=TRUE, colClasses=c("factor", "factor", "factor", "factor"))
3 match["id1"] <- paste0(match$org, match$CPID, match$matrel, collapse= NULL)
4 match["id2"] <- paste0(match$CPID, match$org, match$matrel, collapse= NULL)
5 suspense <- read.csv("H:/WGA/MR/Suspensemap.csv", header=TRUE)
6 mapped <- merge(match, suspense, by.x = "matrel", by.y = "MR.no", all.x = TRUE)
7 mapped <- data.table(mapped)
8 mapped[, "transid" := paste(matrel, mval, sep=":a")]
9 setkey(mapped, transid, org, CPID)
10 mapped[, transmatch := chmatch(id1, id2, nomatch= 7777), by= transid]
11 mapped[, "transid2" := (1:N), by= transid]
12 nomatch <- mapped[transmatch == 7777]
13 setkey(nomatch, id1)
14 nomatchunique <- unique(nomatch)
15 mismatches <- mapped[transmatch != 7777]
16 write.csv(mismatches, "H:/WGA/mismatches14.csv")
17 nomatchsum <- nomatchunique[, sum(mval), by=SusACC]
18 write.csv(nomatchsum, "H:/WGA/nomatchsum14.csv")
19 multiple <- read.csv("H:/Janetst1.csv", header=TRUE)
20 multiple <- data.table(multiple)
21 setkey(multiple, UniqueID)
22 multi <- unique(multiple)
23 test2 <- multi[, sum(mval), by=SusACC]
24 write.csv(test2, "H:/jan0test1.csv")
```


Interactive data visualisations: WGA

Overall WGA Balances | 2015



Breakdown of WGA Balances | 2015



Overview

- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion

Leverage senior sponsorship



Bring in specialists and protect their time



Reach out to junior staff



Focus on what helps

- Data gathering and presentation, rather than analysis
- Identify tools/techniques that meet specific needs
 - Back catalogue analyser
 - Webscraping
 - Bubble charts
 - Interactive visualisations
- Spread the word

Stay joined up with “traditional” analysis



Keep track of skills

Screenshot of the Skills Passport for data analytics

2. Using the following criteria, please complete the table based on your experience with each of these software packages:

0 = never heard of it

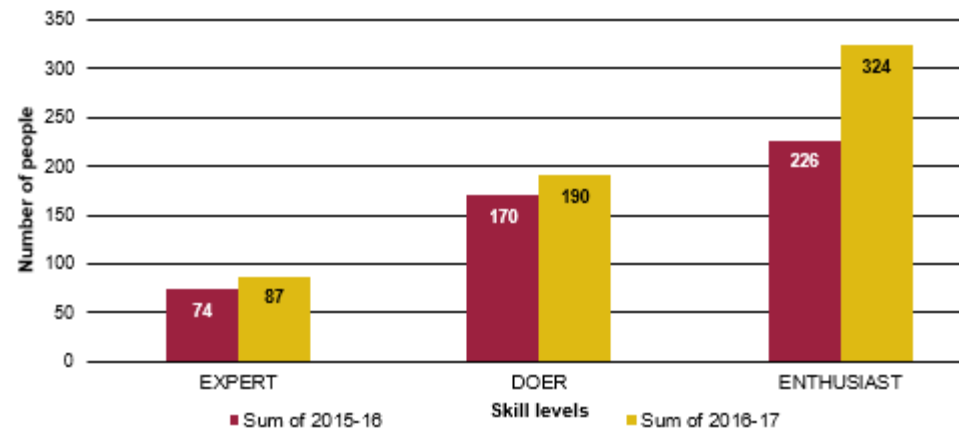
1 = heard of it, but never used it

2 = have used it for simple applications

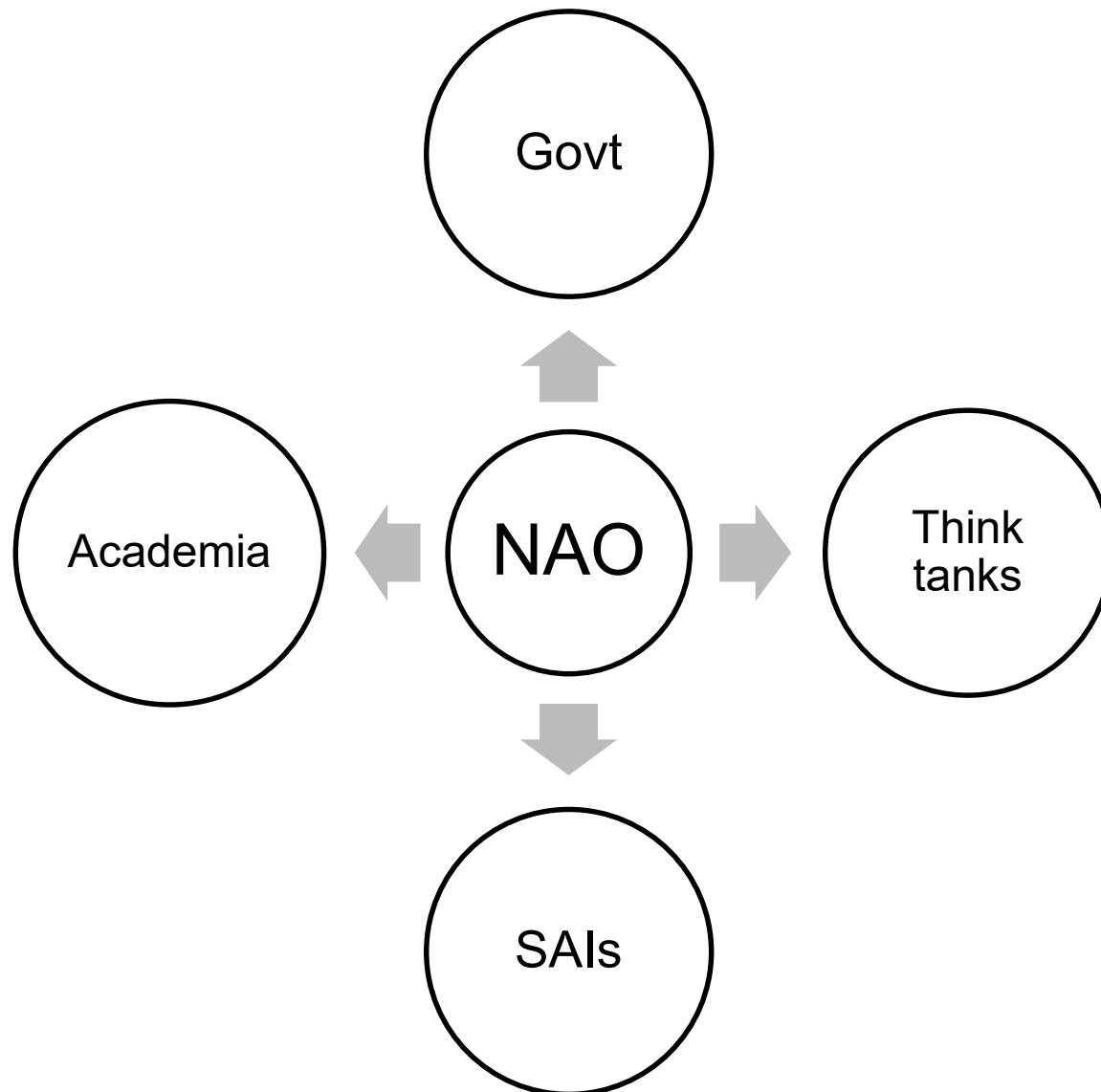
3 = have used extensively and can provide advice to others

	R	Python	SQL (Not relevant to Data visualisation)			
Data from Web (eg. webscraping, website automation, connecting to APIs, connecting to RSS feeds)	<input type="text"/>	<input type="text"/>	<input type="text"/>			
Text analysis (eg. text search, word clouds, n-grams, sentiment analysis, topic modelling)	<input type="text"/>	<input type="text"/>	<input type="text"/>			
Data visualisation (eg. making charts/graphics, interactive dashboards, developing web applications)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Data wrangling (eg. un/pivoting data, filtering, linking, appending, calculating new vars, grouping)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>
Machine learning (eg. predictive modelling, automatic classification algorithms, pattern recognition)	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>	<input type="text"/>

Other (please briefly specify software/language and how you used it)



Look outwards



Overview

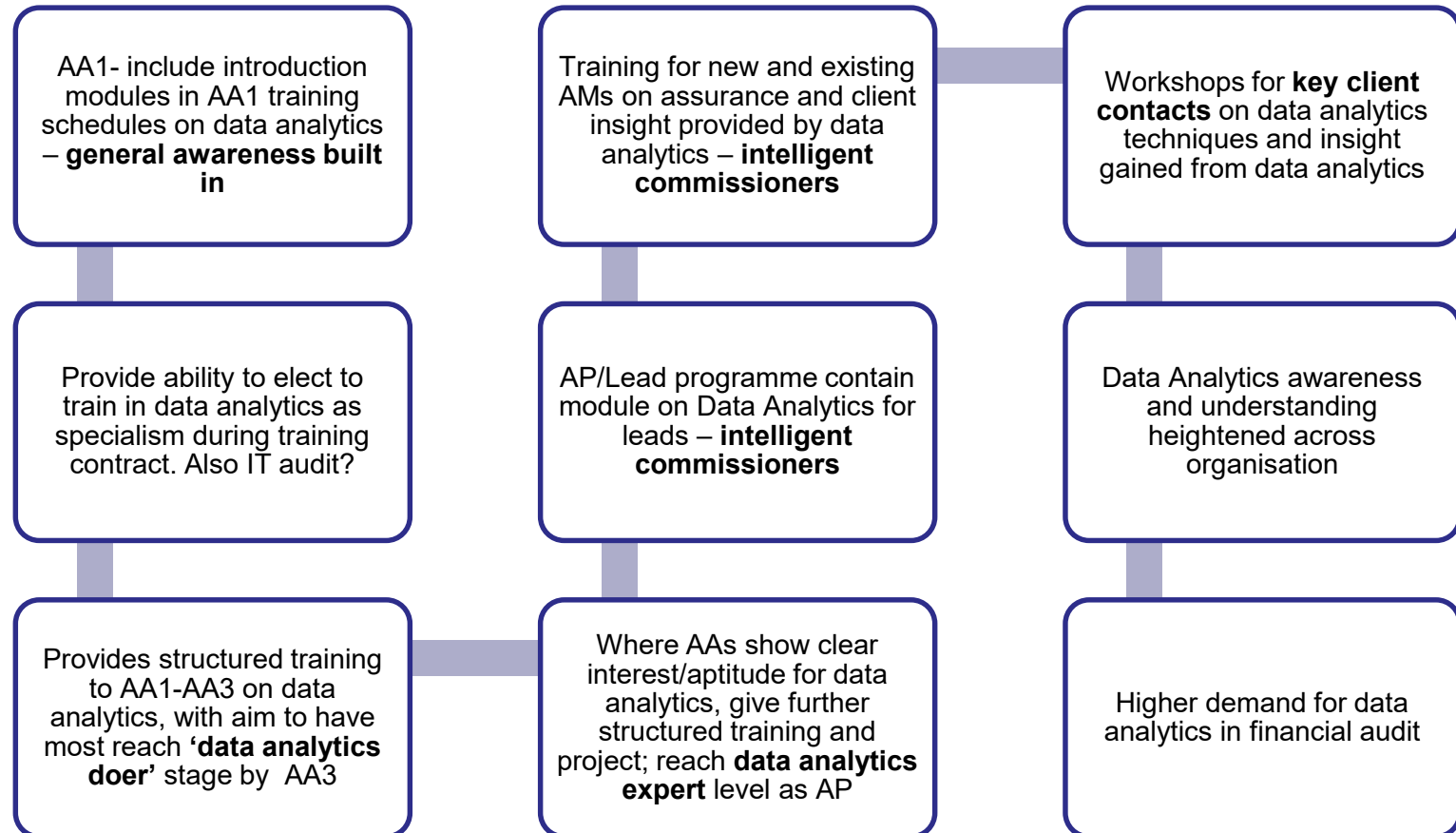
- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion



What's worked

- Find senior people who get it
- Keep on keeping on
- Make sure FA and PA data analytics stay integrated
- Think long term, think pipeline

Financial audit training



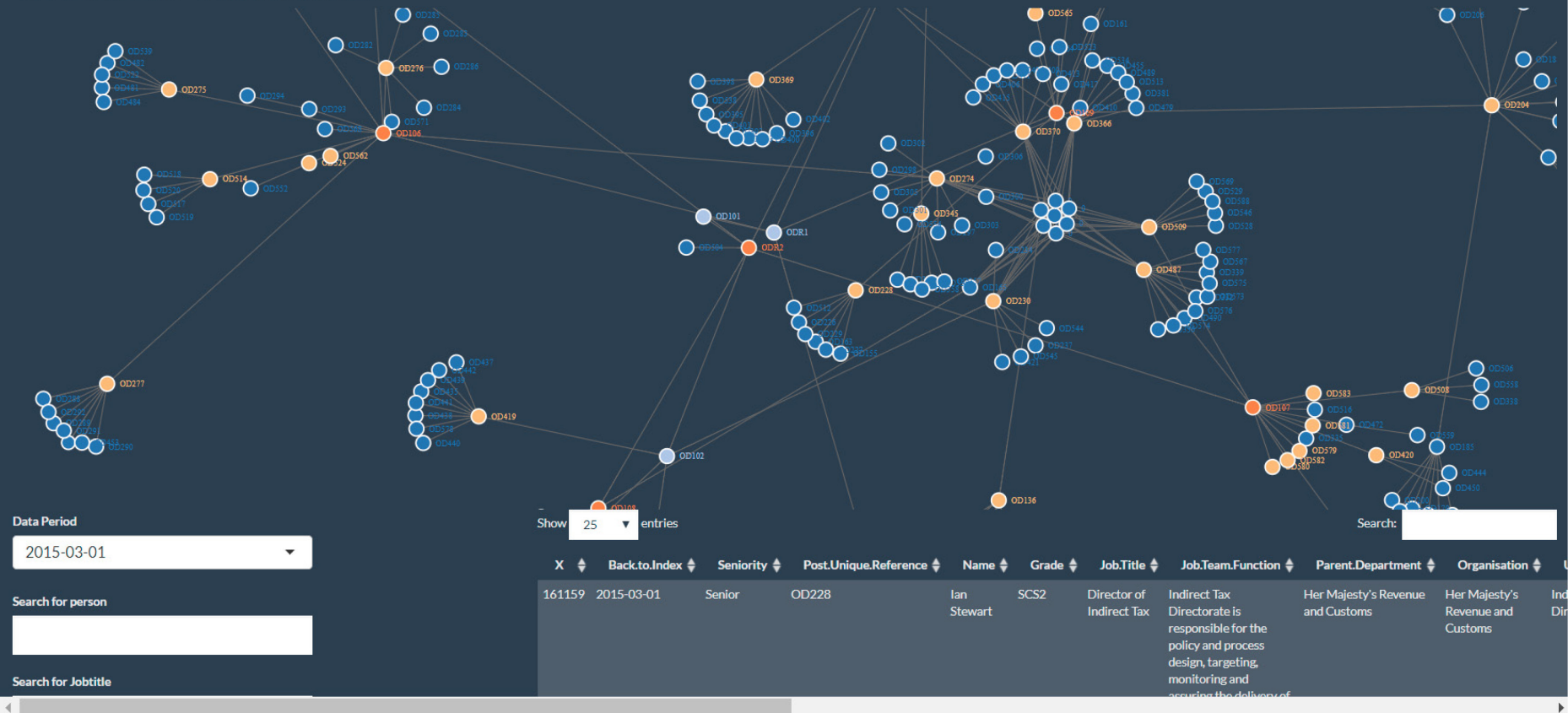
Overview

- Background on the NAO
- Examples of Data analytics
 - Performance audit
 - Financial Audit
- How we've done it
 - Performance audit
 - Financial Audit
- Discussion

Spare slides

Department Organograms

Civil Service Organograms



Parliament Monitor

NAO Parliament Monitor



Select function:

Home

Search questions

PAC

View members

Constituency map

Questions from PAC members

House of Commons MPs

Select question types

Questions from PAC members



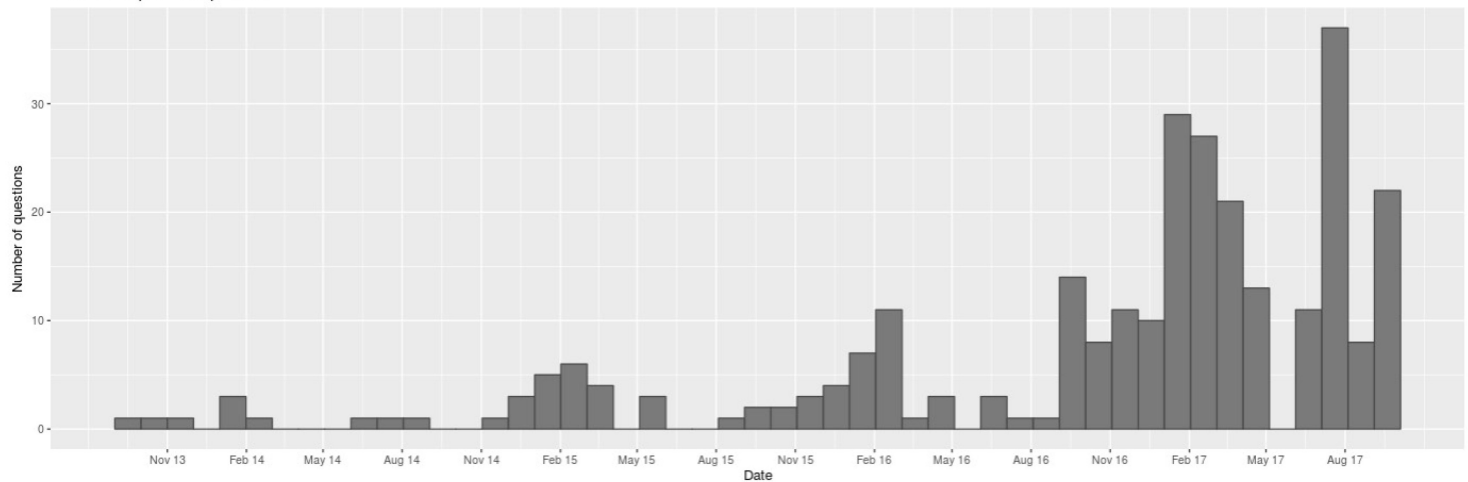
National Audit Office

Number of results per PAC member:

10

This page may take a few seconds to load - please be patient!

Number of questions per month



The table below lists the most recent results for PAC members, as listed on data.parliament.uk. Refine your results using the filters in the column headers.

5. Comprehensive L&D offer

- **External training providers** – to provide R training for initial core of users, plus later on package writing in R etc
- **Training courses provided by NAO staff** - data service training, Powerpivot, webscraping, APIs, data visualisation, coding in R.
- **E-training** - initially using free web provided R / data analytics resources moved onto more structured DataCamp courses
- **Portfolio of skills** - Encouraging staff to think about analytical skills as a “portfolio” – and to complete a “skills passport” – (next slide for more detail)
- **Sharing knowledge** - network meet “show and tell” sessions to share what is possible and blogs / updates to spread new uses / methods / tools Background on the NAO