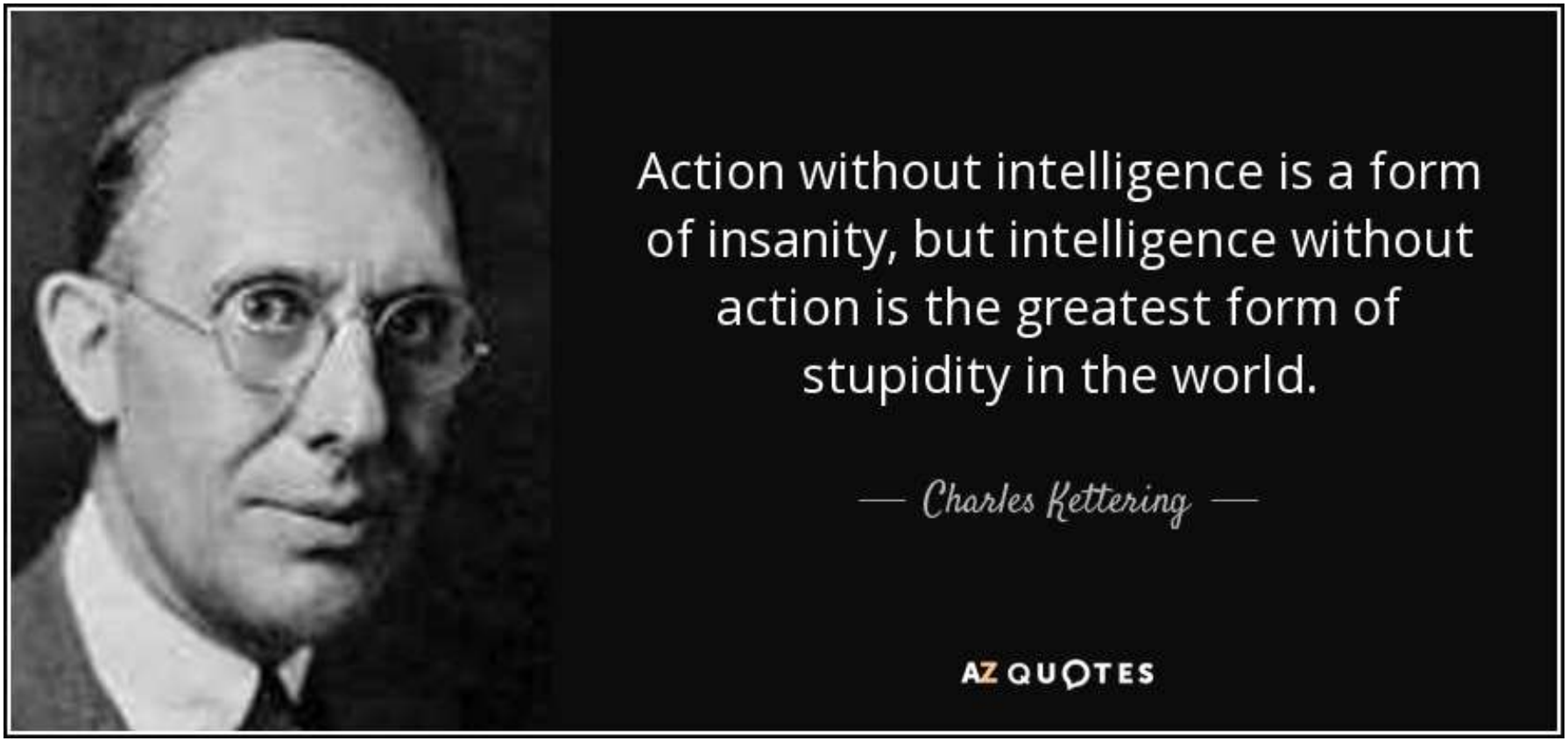
Public Health England

Protecting and improving the nation's health

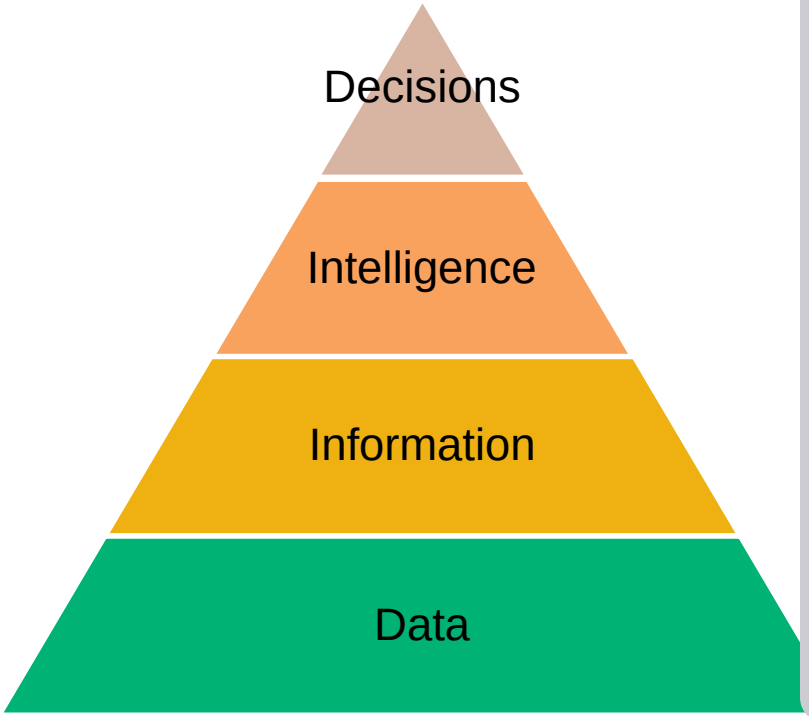
Intelligent design: creating a blueprint for future intelligence provision

Matt Hennessey, Associate Director



Our challenge is to build a unified intelligence function that is neither insane nor stupid! All roads lead to a decision or behaviour change

Data driven decisions: core components



Enablers	Skills	Knowledge	Delivery
Software	Negotiation	Wider system awareness	Media and board presentation
Hardware	Consultancy	Political acumen	Needs assessments
IG	Information and evidence synthesis	Epidemiology and statistics	Annual reports
	Interpretation	Subject specific knowledge e.g. health economics,	JSNA
	Advanced analysis	Effective report writing and data transformation	CBA
	Basic analysis		Tables and charts
	Information management		Cleaned datasets
	Database and IT maintenance		

Sweating the assets: Analytical resource

Are we at a crossroads?

- Cannot afford to be all things to all people
- Maximising transferable **skills** in the system
- Maximising transferable **capacity** in the system
- Setting clear career paths; framework for development
- Understanding where and how to deploy specialist knowledge

Laying the foundations: data/information

- Big data; new sources, new formats
- Changes to Architecture and IG - cloud
- Opportunities of new technology for data transformation
- Embracing automation
- Managing the risks of automation – data integrity and quality

The new world of ‘Data Science’ – Managing your Munging

Say what?!

I need an oozie to link my Pig and my Hive

How big does my sandbox need to be?

Can you data wrangle with a Panda?

Fork me on GitHub

This slide is already out of date!

Deciphering the jargon!

Automating data analysis – Example HAY

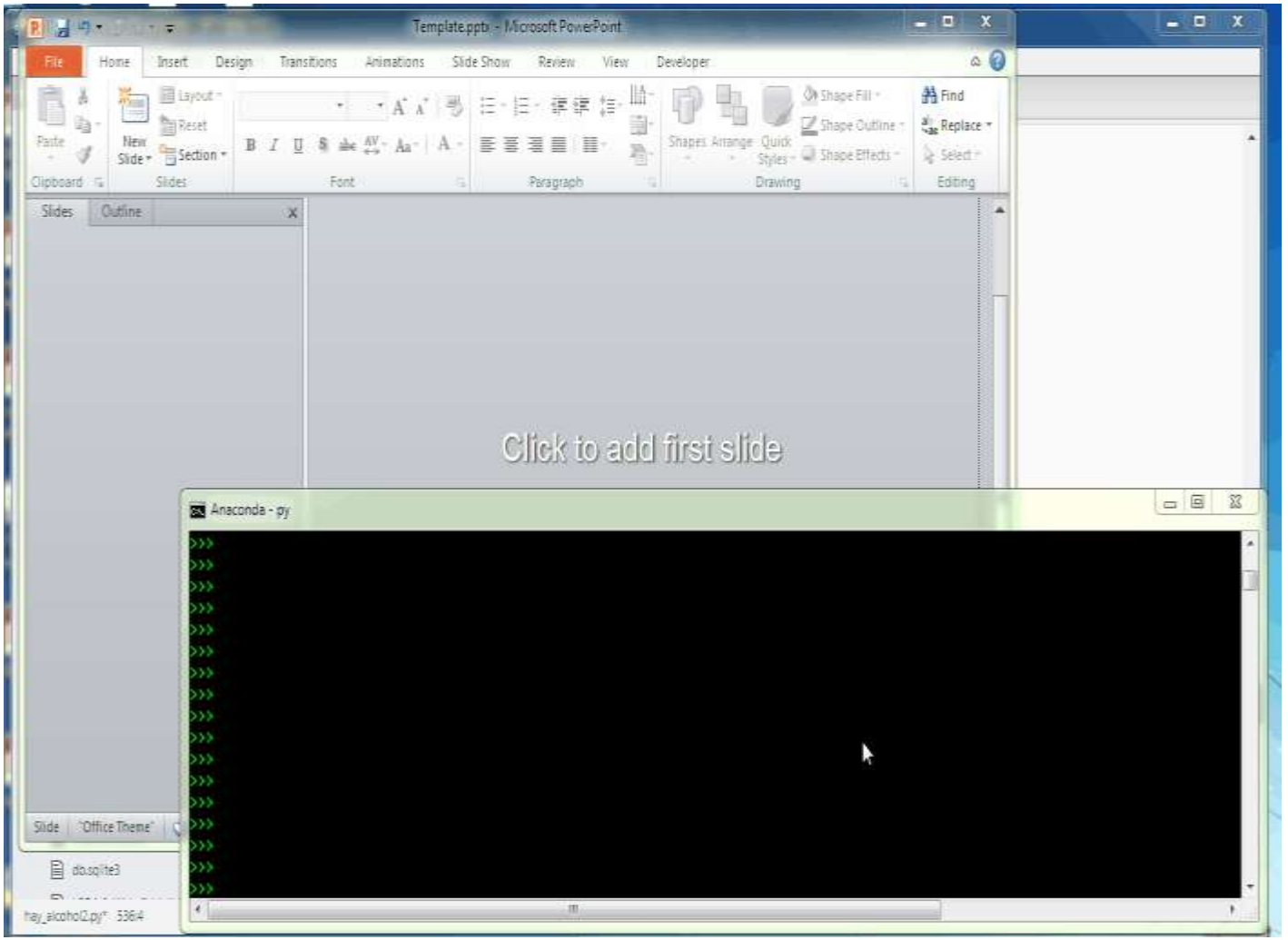
Approx 1.8 million rows

Python used to merge daily .csv files

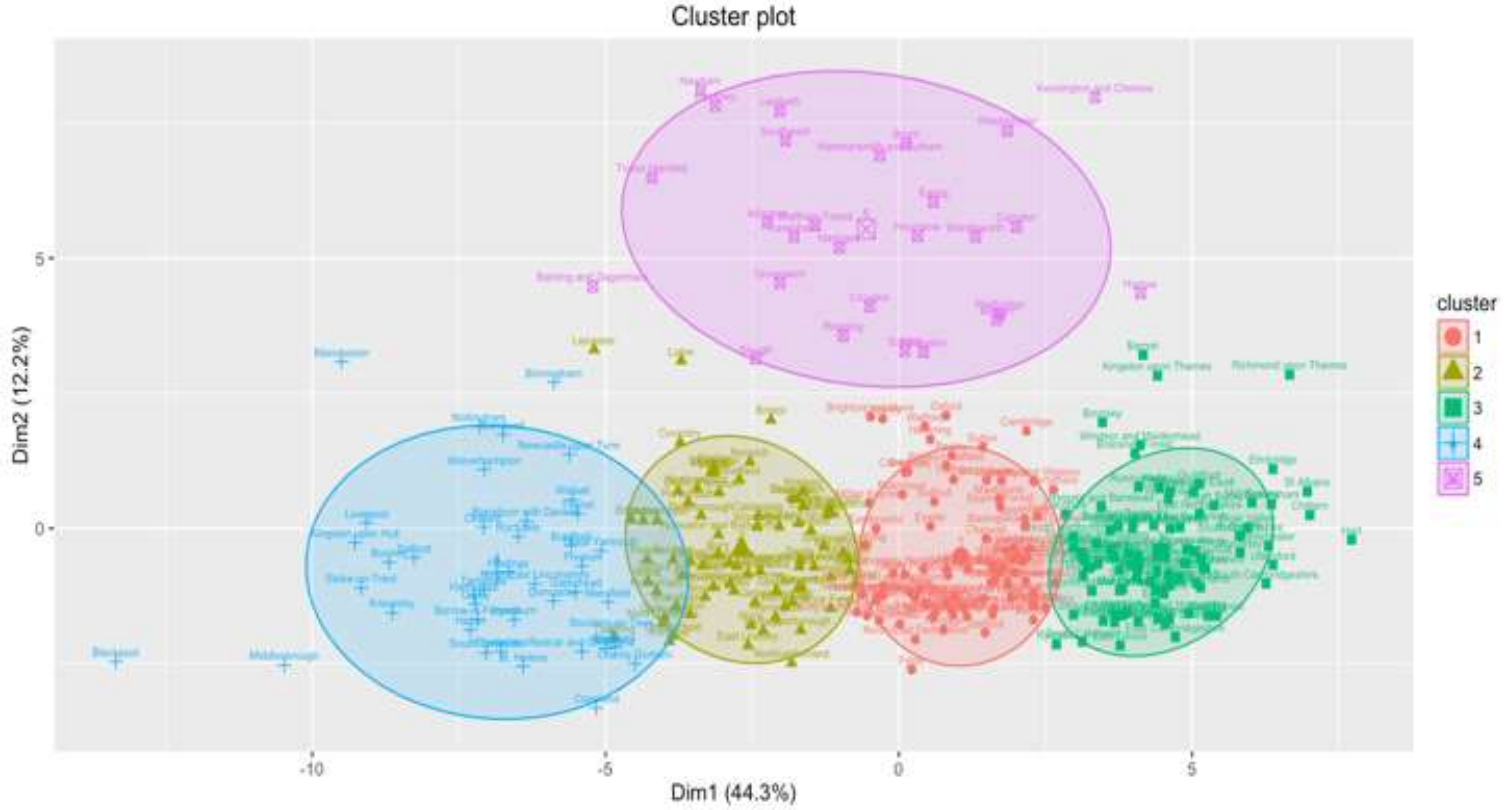
Pandas used to analyse dataframes – lightning quick to organise data in to groups, such as age, alcohol consumption, smoking status

Win32com acts as a bridge between Python and MS Office tools (Excel and Powerpoint)

Result – repeatable, high quality



New insight? Clustering of local authorities in health profiles – what do they tell us?



Department for Transport

Data Science
Public consultations

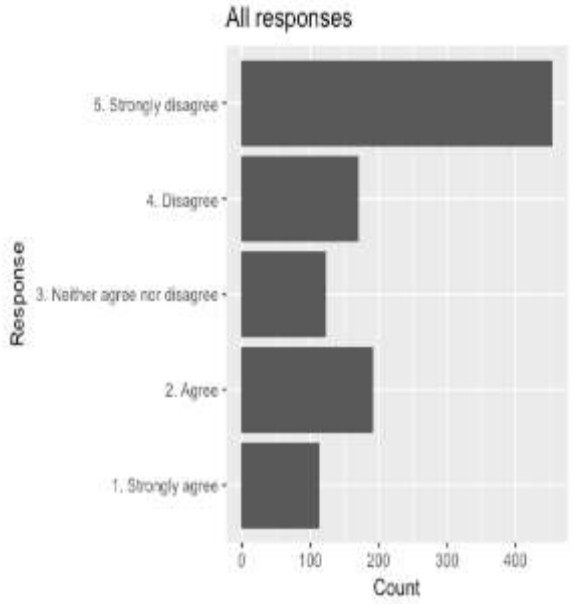
Tom Ewing David Pavitt

Basic stats and word counts

Q8. How strongly do you agree or disagree with our proposed environmental objective for the next regime?

Response strength

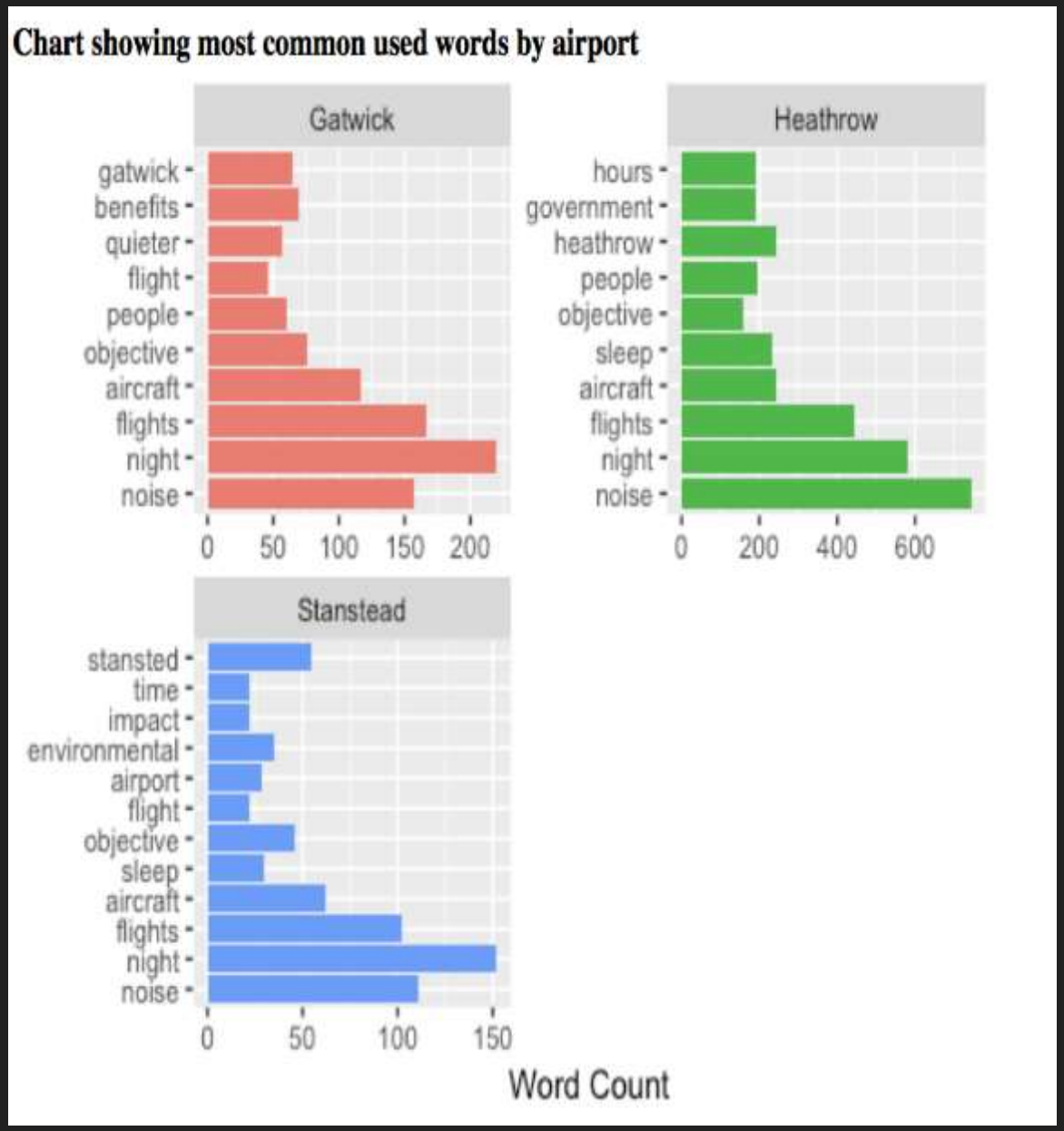
The average response to the question is 3.6 where 1 = strongly agree and 5 = strongly disagree.



Breakdown

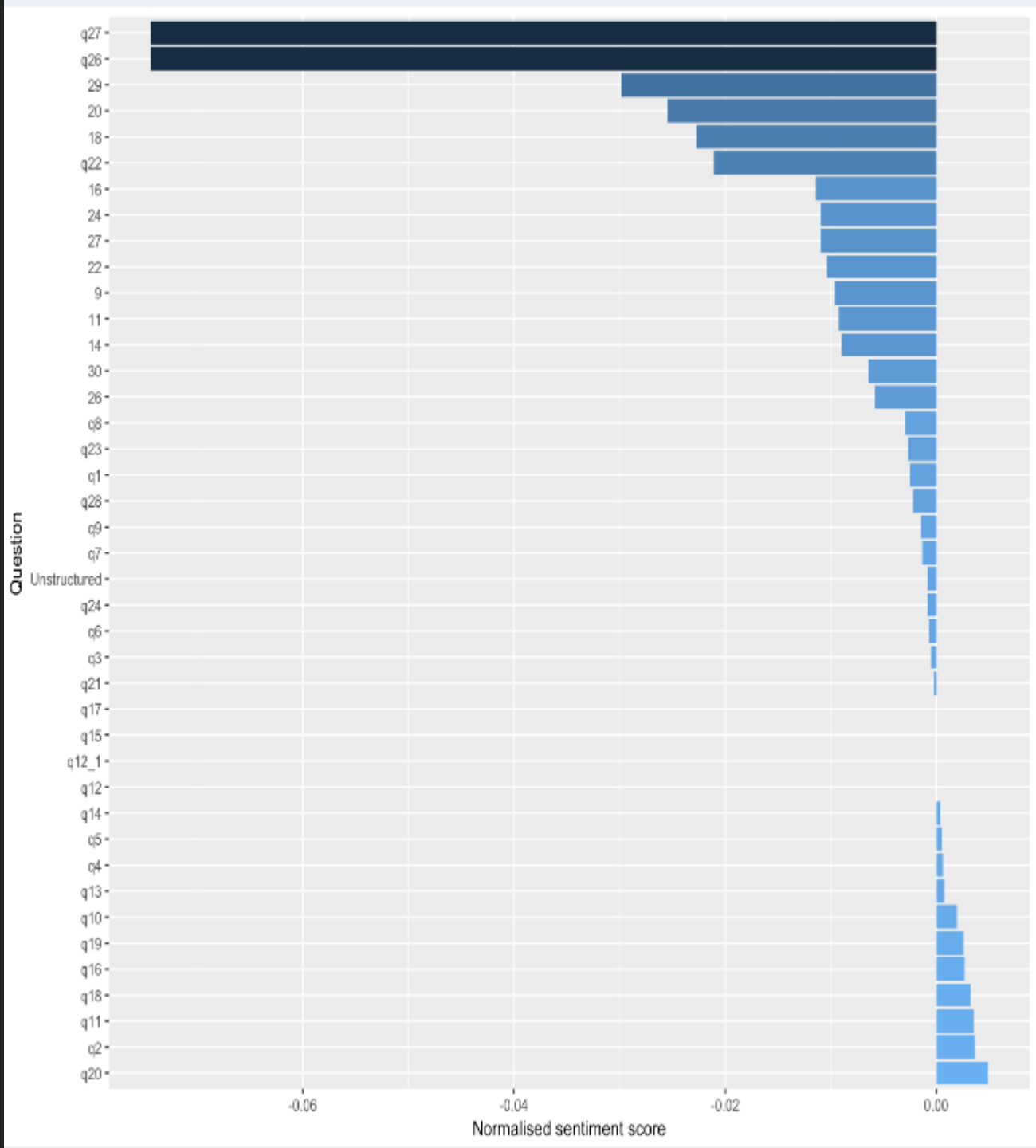
Agreement score by organisation type

Organisation Type	Average Score
None	3.7
Public Body	3.4
Noise or environmental group	4.1
Airport or Airline	1.9
Business or Trade Union	2.1
Other community group or charity	2.8
Local Authority	3.5



Sentiment

Across the different consultation questions



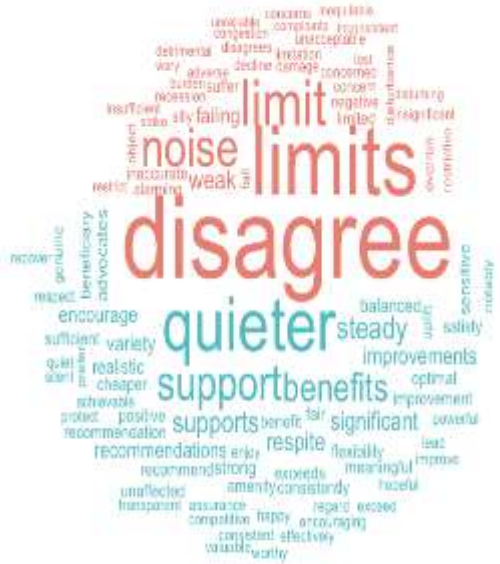
Sentiment

Word cloud v2.0

Select Question

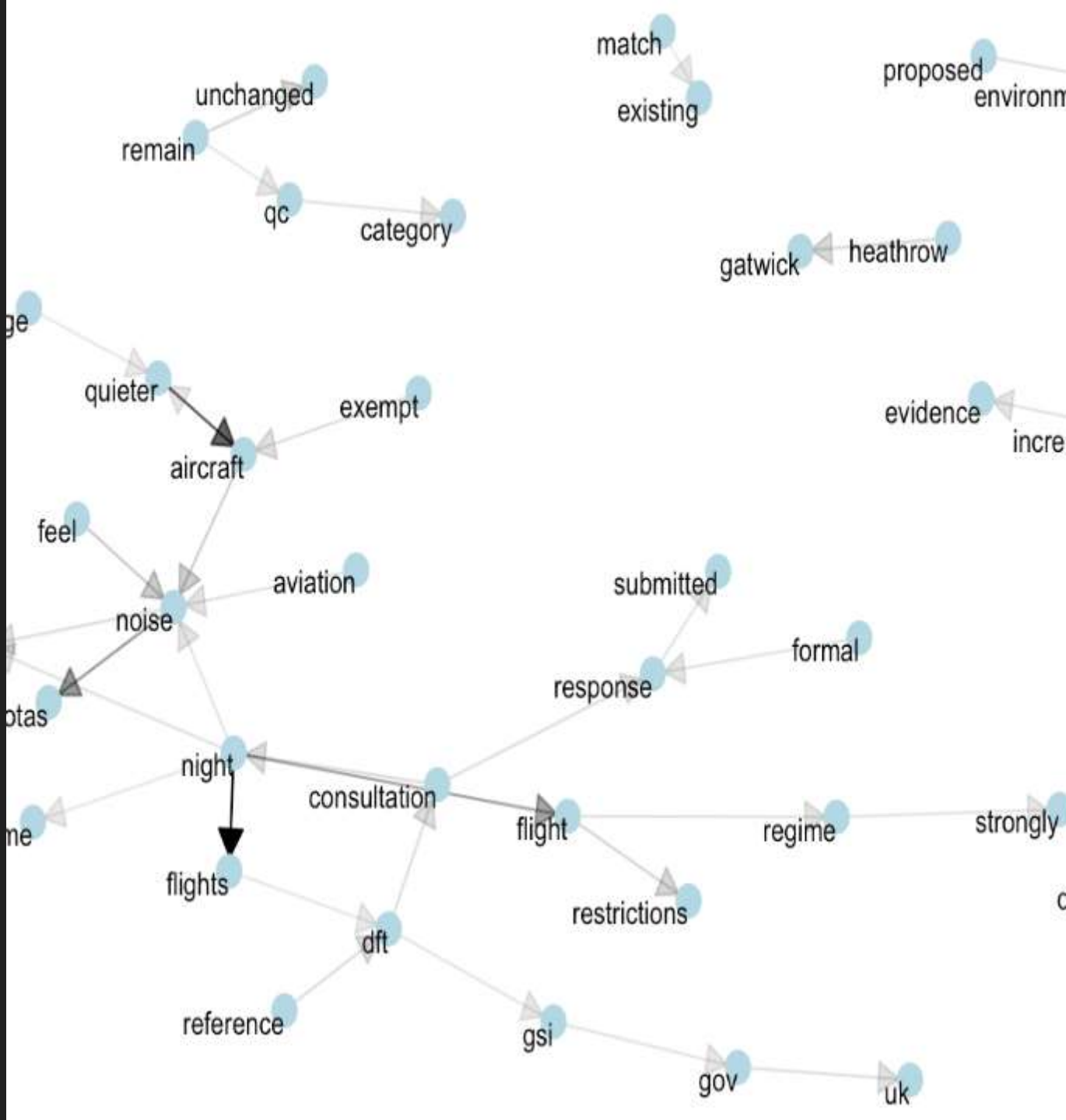
q9

negative



positive

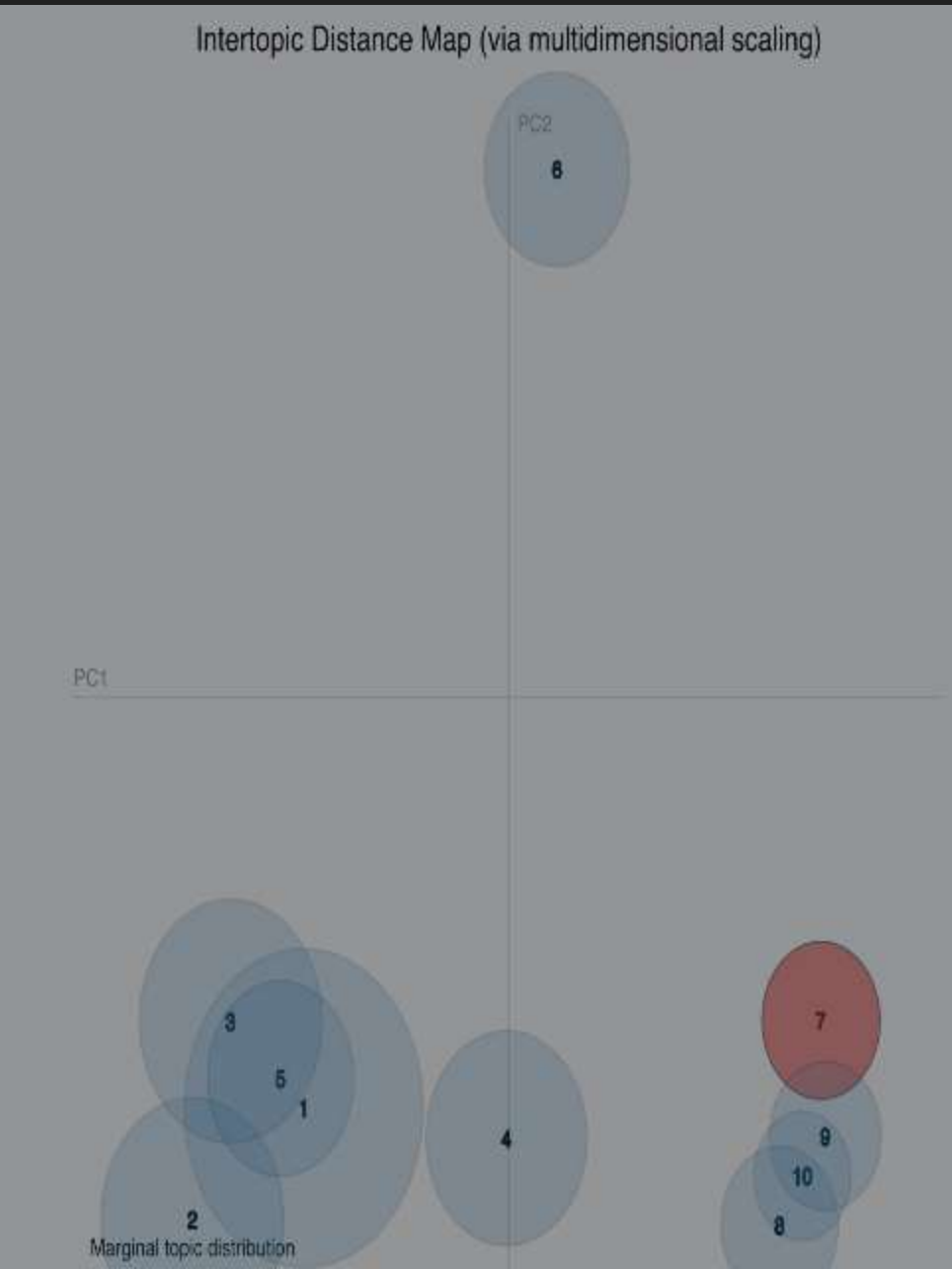
Relationships between words

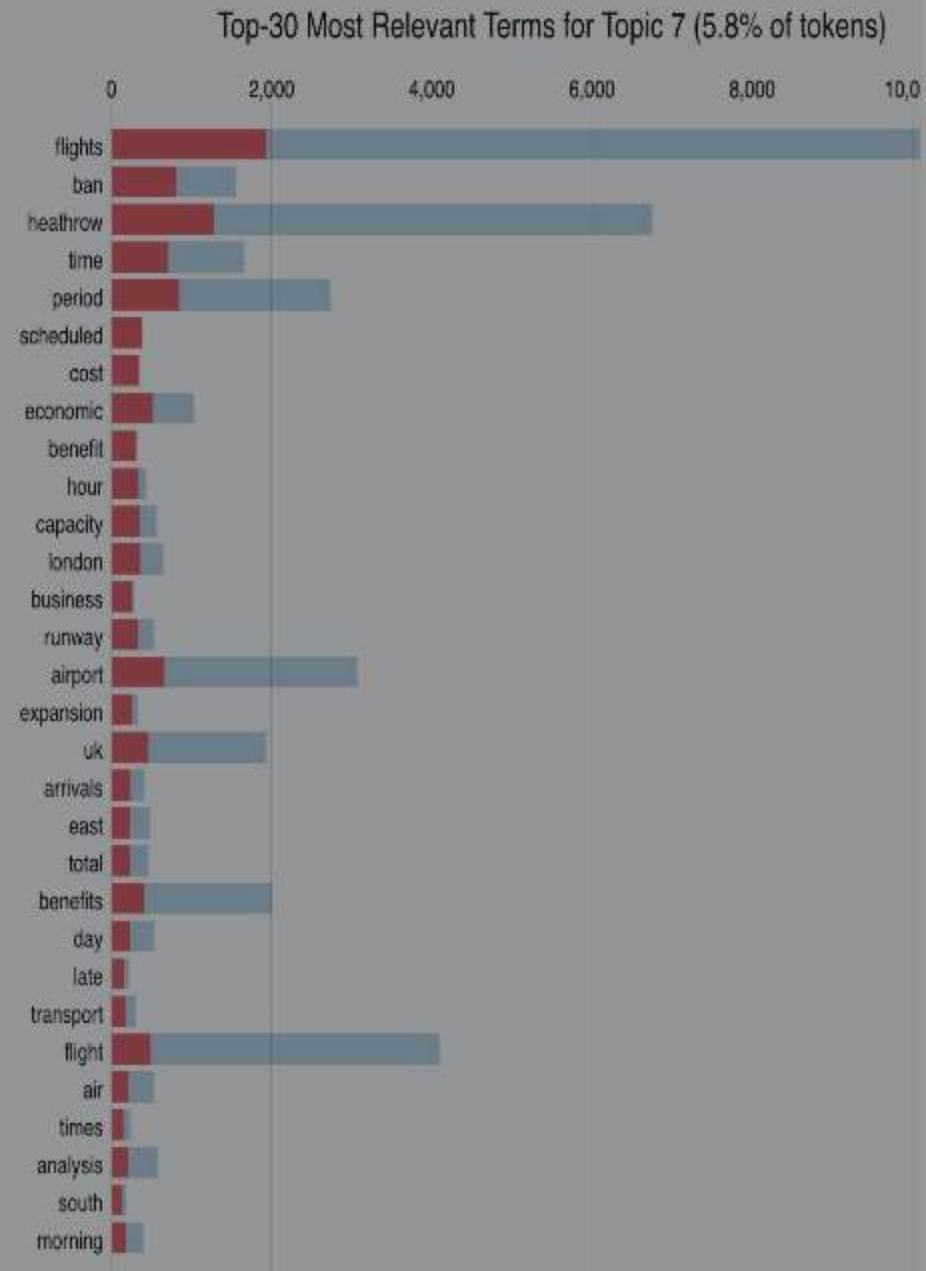


Topic modelling

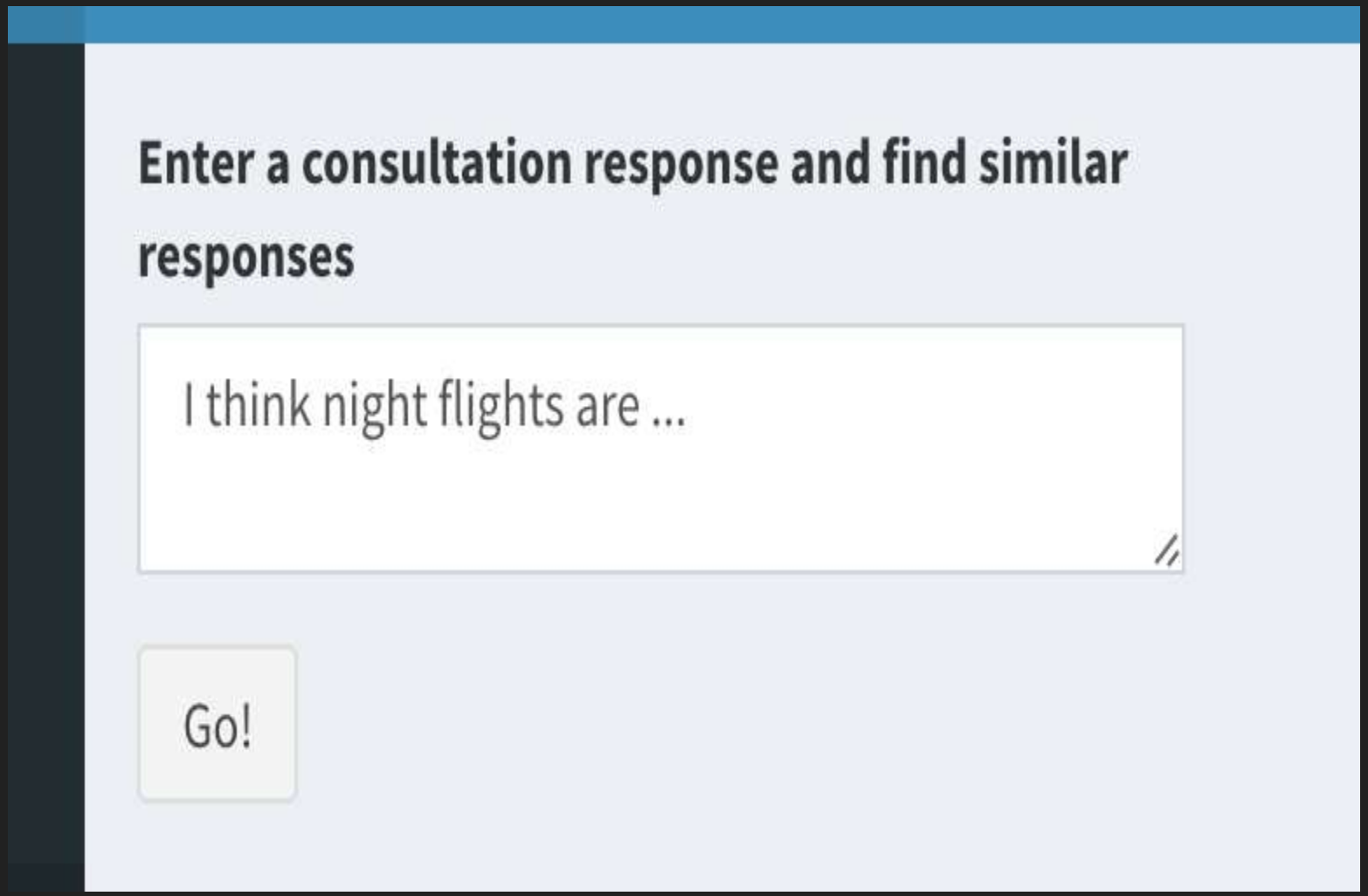
Topic 1	Topic 10	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9
planes	agree	flights	flights	regime	email	noise	aircraft	agree	government
west	strongly	sleep	ban	environmental	www	pollution	quota	section	noise
gatwick	proposal	pm	heathrow	objective	visit	db	noise	aircraft	disagree
noise	disagree	hours	time	length	security	government	airlines	strongly	quotas
flying	movement	ban	period	proposed	scanned	sleep	summer	proposals	suggestion
fly	comments	stansted	scheduled	benefits	response	people	reduction	disagree	proposal
communities	aircraft	residents	cost	period	symanteccloud	residents	winter	limits	gatwick
affected	additional	disturbance	economic	opportunity	cloud	levels	movements	additional	people

Topic modelling - clustering and visualisation





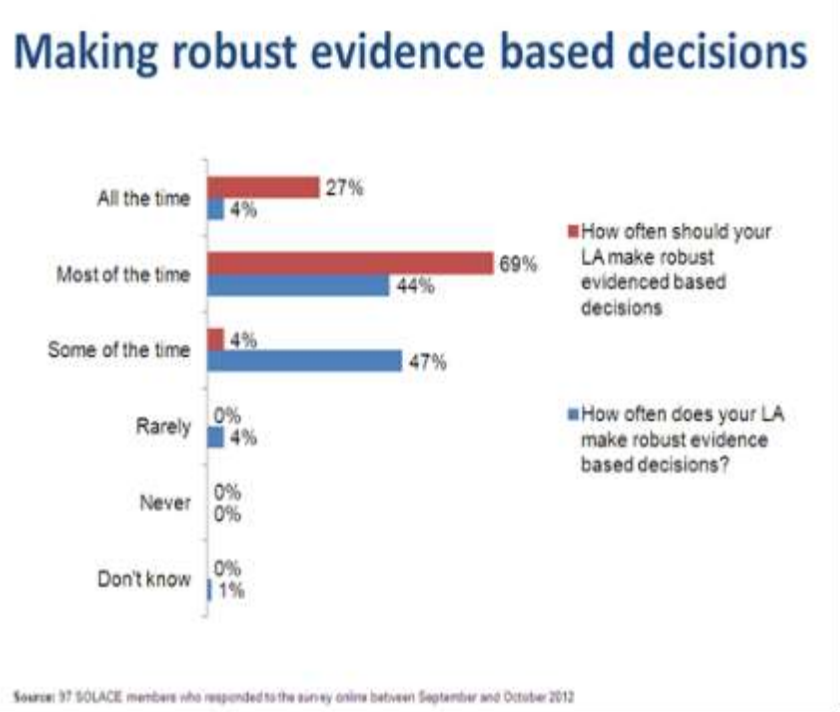
Topic modelling - similarity, machine learning?

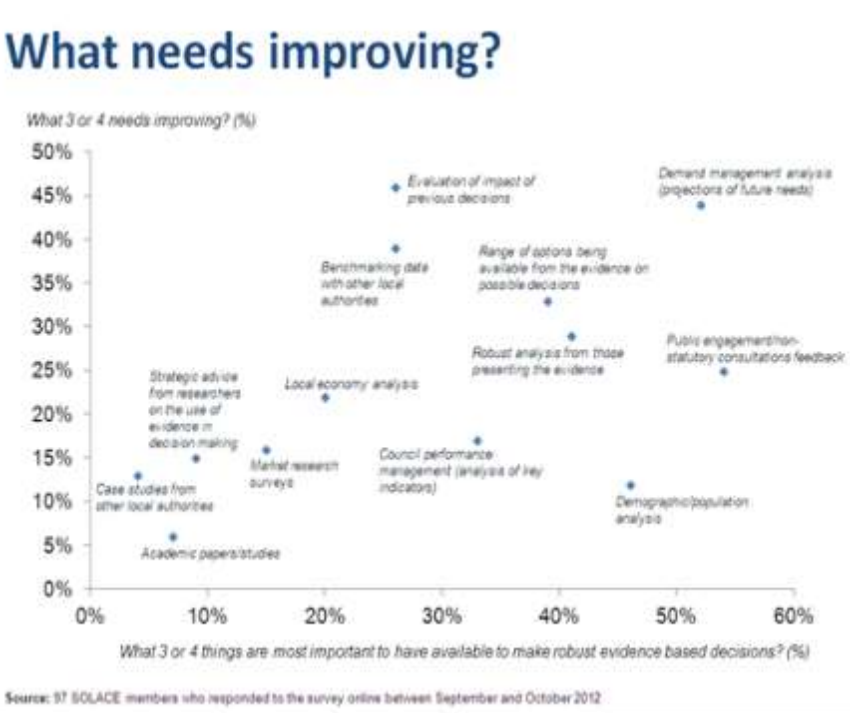


Mobilising our knowledge: intelligence to decision

SOLACE REPORT: The Role of Evidence in Local Government Decision Making

* Evidence is “data in context” – context being established by a hypothesis or belief

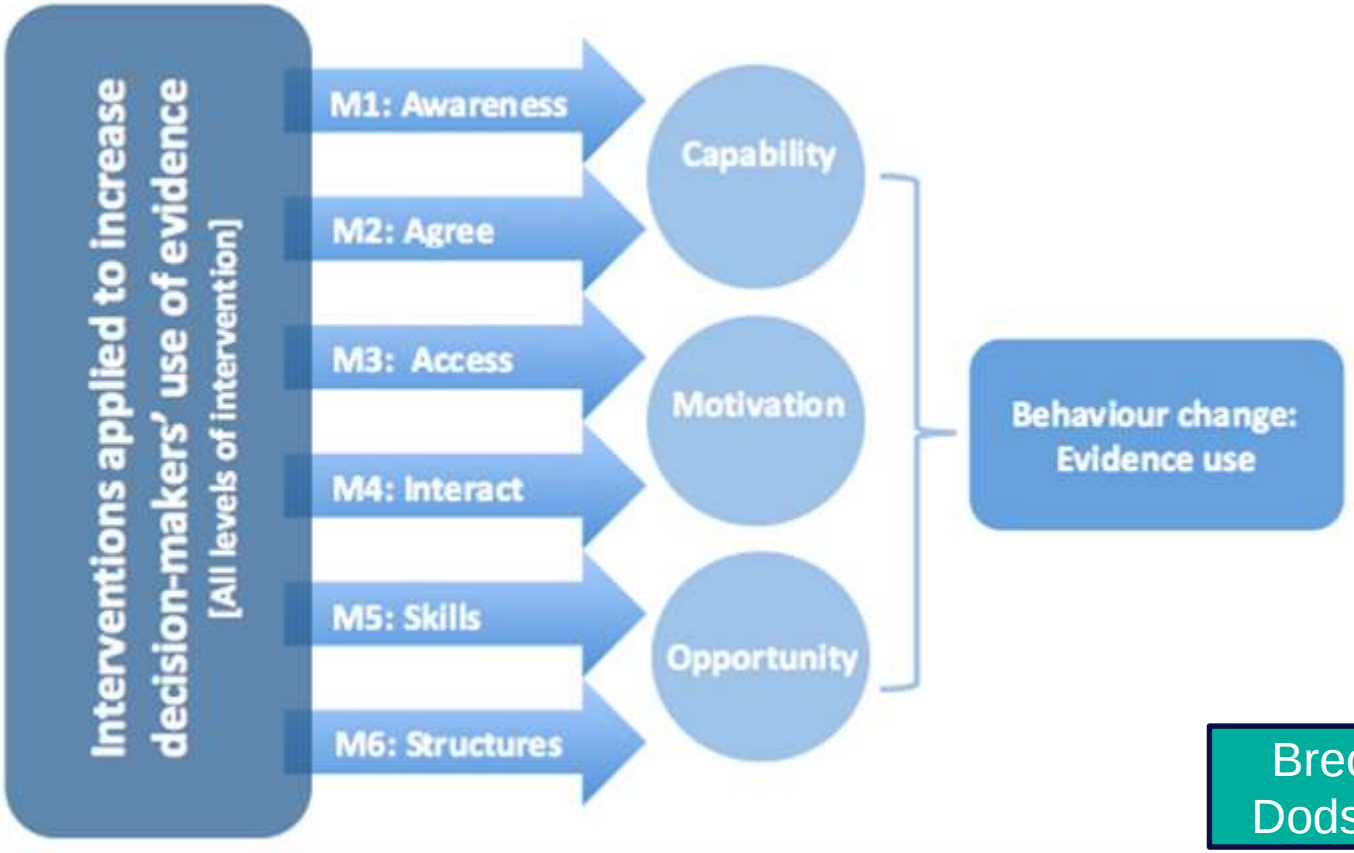




Models for a system or place-based approach?

Population hub:	
Shared investment	
Consultancy	
Digital platform	

How do we get people to act on our intel?



Breckon and Dodson model

Spend time with your 'audience'