# Reflections on a new data journalism paper at AUT University

Allan Lee, Jeanz conference 2015

Data journalism

- Full paper by March 2016

- Preparing the course/student entry

- Course content

- Text book and blog

- Assessment

- Student work

- Student feedback

- Changes for next year
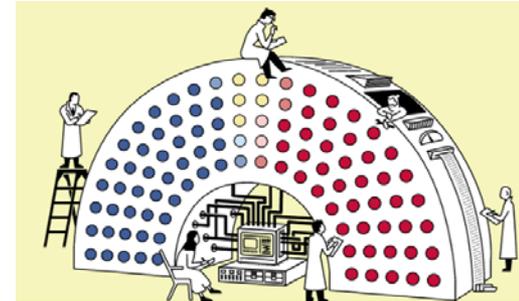
Data journalism

Preparing the course

Harkanwal Singh,
New Zealand Herald
data editor

Jeff Kelly Lowenstein,
a Fulbright Scholar
from Columbia College,
Chicago

The students

14 students – BCS journalism majors, plus two minors

Entry to the paper required sitting a maths test

Course content

## Weekly 3-hour class structure:

Part A (45 mins)
Discuss reading/Blog entries/Quiz
Student presentations – critiques of data journalism projects
(marked by their colleagues)

Part B (45 mins)
Lecture session/guest speaker

Part C (60 mins)
Workshop time – software training, working through textbook,
exercises
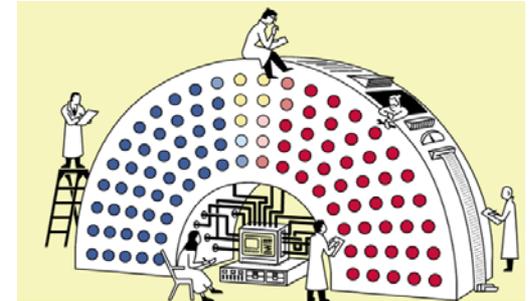
Course content

Lecture sessions:

Best of data journalism worldwide
Intro to data journalism tools and processes
Statistics essentials for the data journalist (Robin Hankin)
Exploring data.govt.nz/where to find data (Joanne Alexander)
Data visualisation/Statschat (Thomas Lumley)
Data journalist/intro to programming (Harkanwal Singh)
Sources of data
Data analysis/handling survey data
Social media as data source (Verica Rupar)
Open government and FOI (Greg Treadwell)
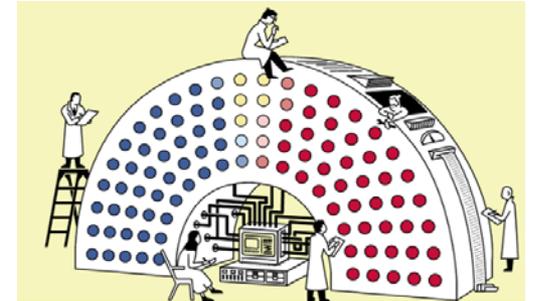
Course content

Workshop sessions:

Finding data – online databases
Excel spreadsheets + Google Sheets
Testing data integrity/identifying dirt
Cleaning data with Open Refine
Analysing data
Filtering and summarising data
Generating summary statistics (sums, averages, percentages, ratios, running rates)
Pivot tables in Excel
PDF converters – Cometdocs and Zamzar
Charting in Excel and with web tools (bar graphs, line charts, pie charts, data maps, scatterplots)
Google Fusion Tables/ .kml files + heat maps
Visualising with Datawrapper
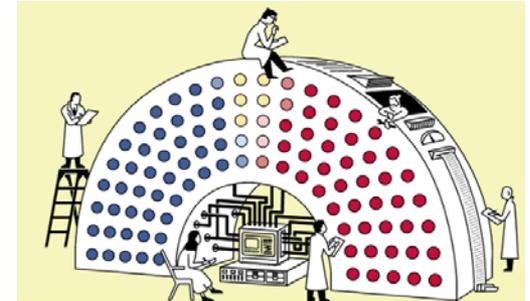
Course content

Not covered
Programming/writing code in
Python/Ruby/Javascript
More advanced statistical methods (eg correlation,
regression analysis, probability testing)
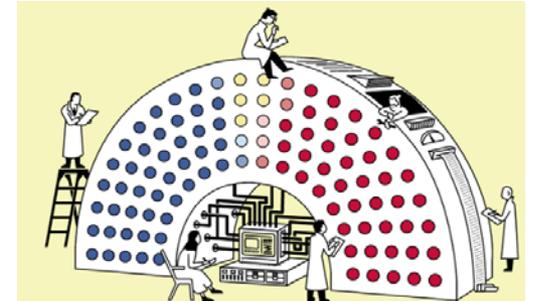
Data journalism

Course text

David Herzog,
Missouri School
of Journalism

Data journalism

Course blog

https://autdatajournalism.wordpress.com

Assessment

1. Group presentation (20%): a critique of a significant data journalism project/story (mostly in pairs)

2. Data story proposal (30%): story outline, identify databases, identify 4 sources including central characters, plan for data analysis and visualisation

3. Data story (50%): 1500 word story, minimum of 3 graphics/infographics/tables, 4 sources including central characters,
+ a 500 word reflection,
+ a 300 word data analysis report
+ spreadsheets and transcripts.
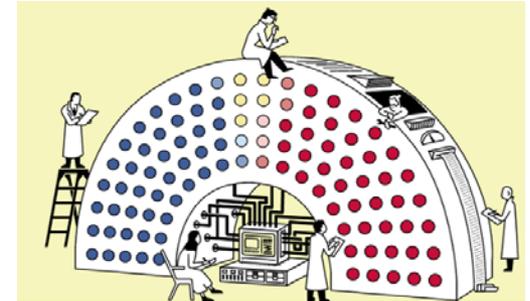
Examples of student work

Fewer young drivers on our roads
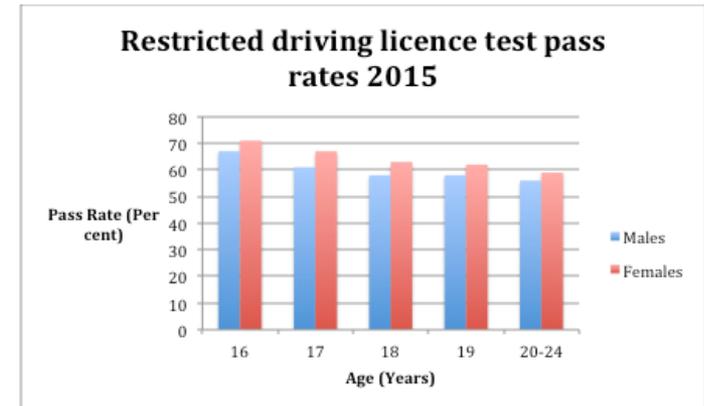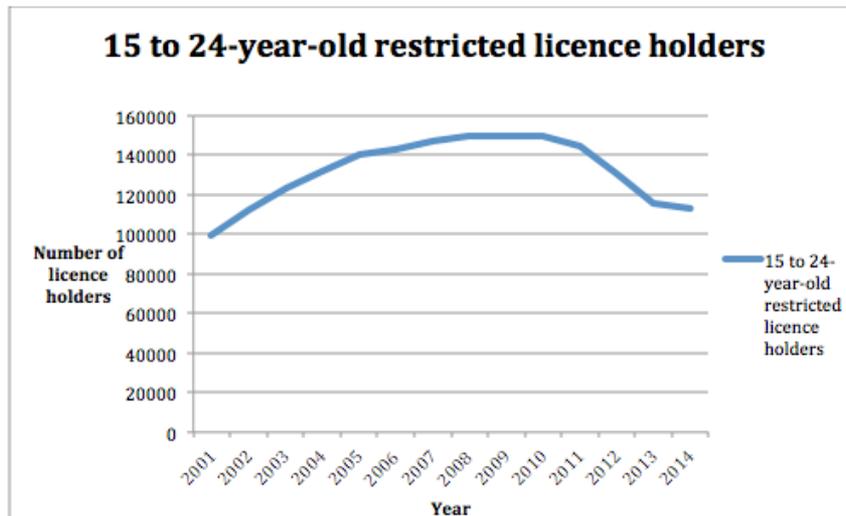
Examples of student work

Women having kids later in life
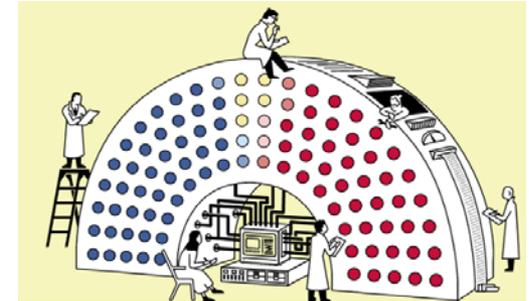
Total number of births per age group 1991 and 2014

- 1991
- 2014

Number of births 2014
- 0 to 5,000
- 5,000 to 10,000
- 10,000 to 15,000
- 15,000 to 20,000
- 20,000 to 25,000
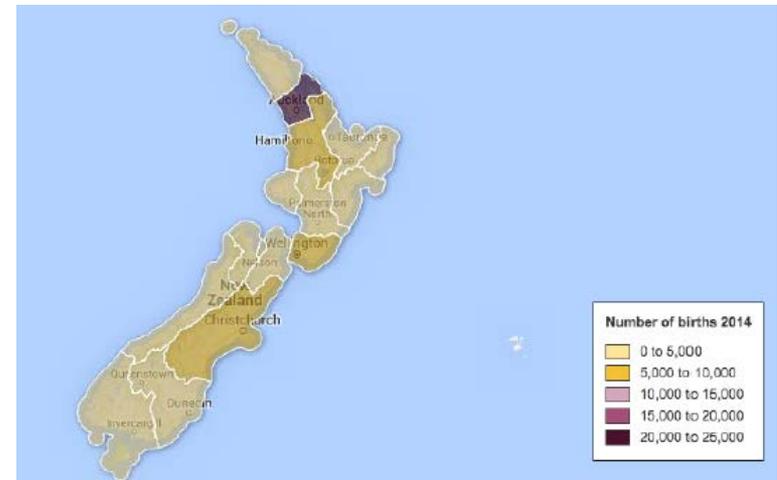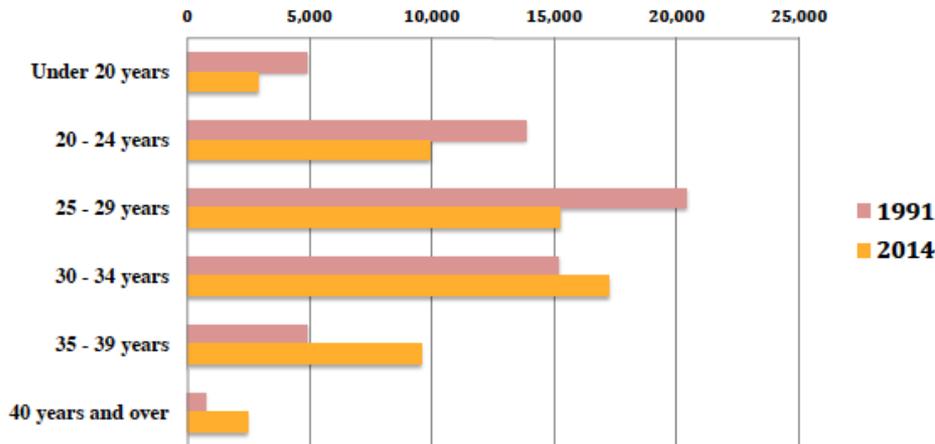
Examples of student work

Vaccination rates on the increase

Percentage of infants vaccinated by given ethnic group

Percentage of infants vaccinated by district health board from 2009 - 2015
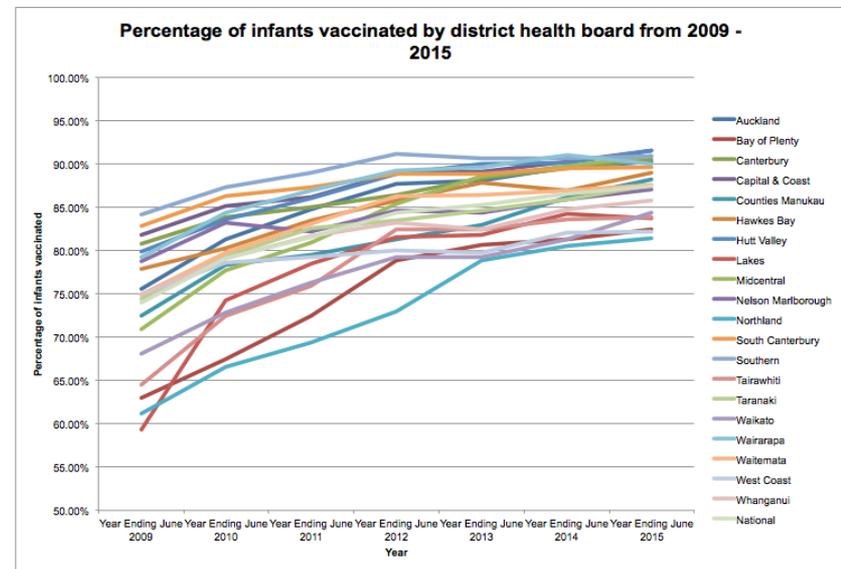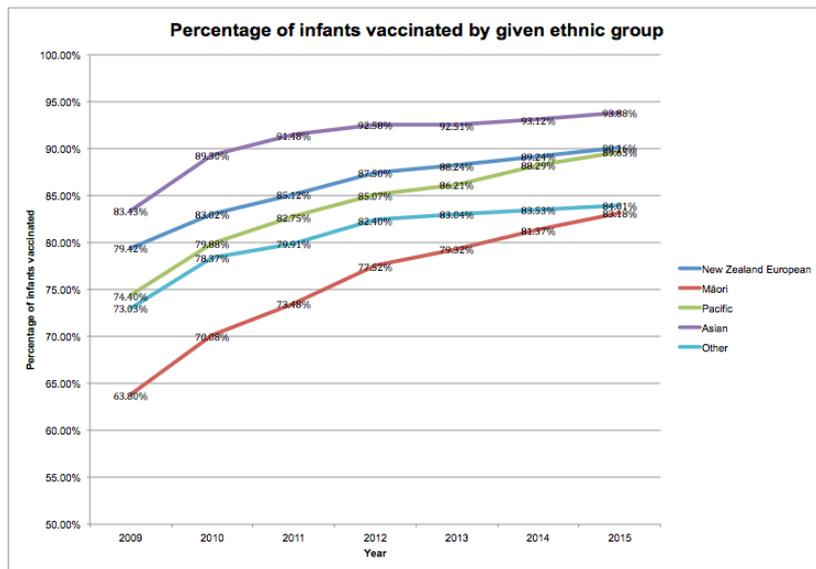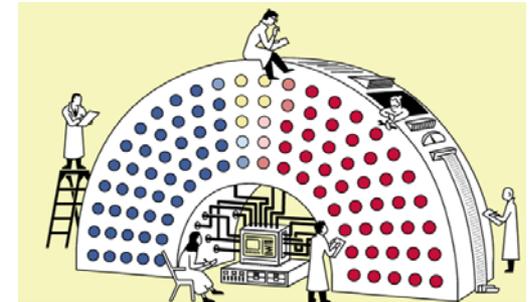
Examples of student work

Baby names reflect greater diversity
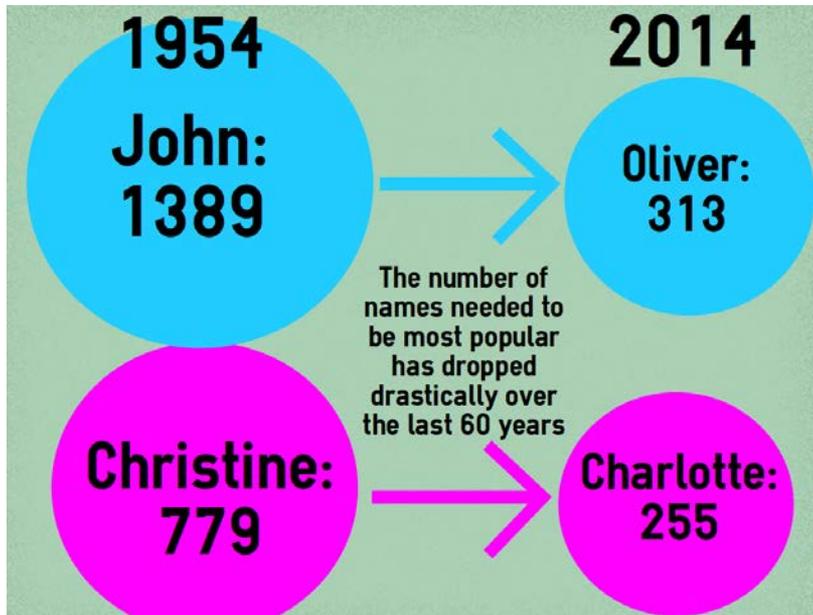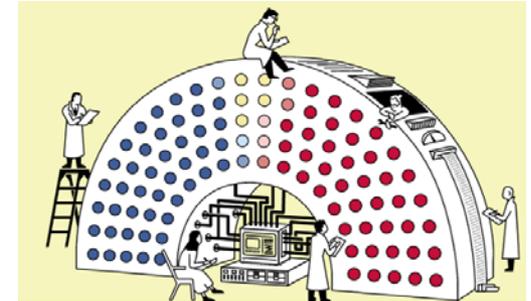
Comparing Top 100 Most Popular Baby Names in New Zealand between 1954 and 2014

This graph maps the popularity of the four most popular boys and girls names in New Zealand in 1954 and 2014 between those 60 years. The graph only highlights the time the names spent in the Top 100, and as such does not map their popularity once they have left.

Oliver   Jack   Charlotte   Olivia   David   John   Christine   Susan
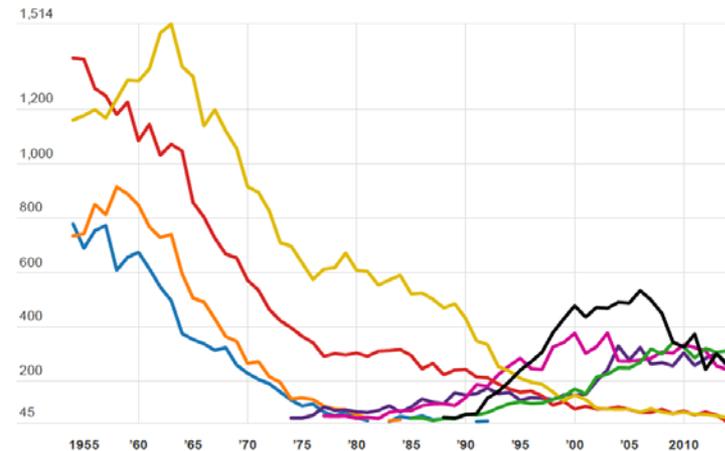
ource: New Zealand Department of Internal Affairs Get the data   Created with Datawra

Examples of student work

Kiwis are having smaller families
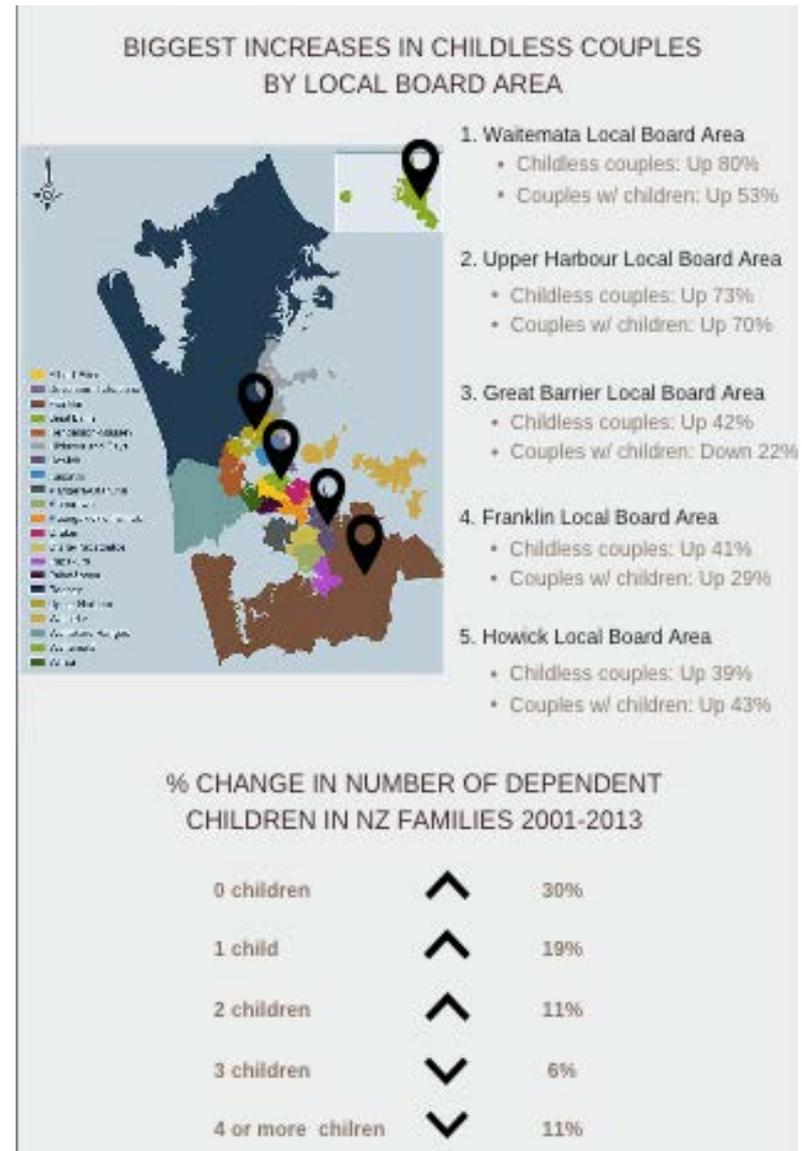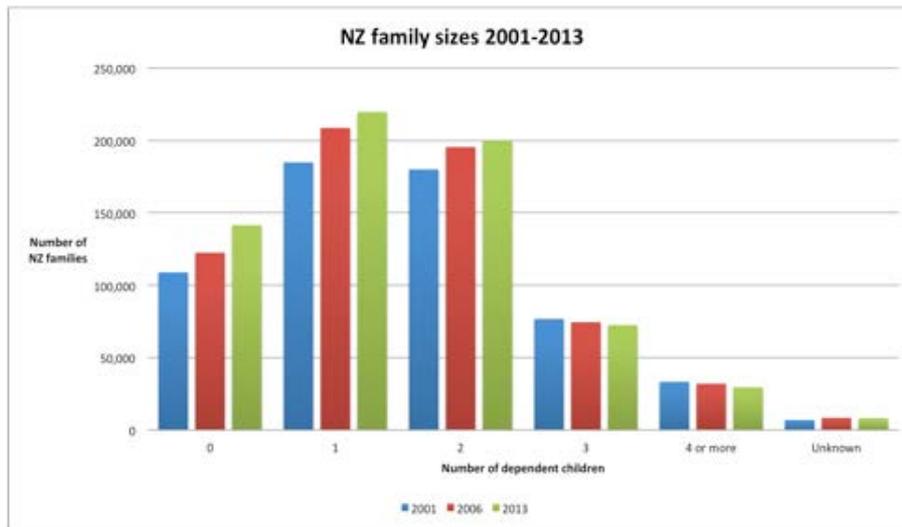
Student feedback

Paper ratings for JOUR703 Data Journalism - Overview

Legend: Strongly Agree + Agree | Neutral | Disagree + Strongly Disagree

Categories (top to bottom):
- Appropriate Workload
- Good Teaching
- Learning Resources
- Appropriate Assessment
- Clear Goals and Standards
- Generic Skills
- Overall Experience

X-axis: 0% 20% 40% 60% 80% 100%
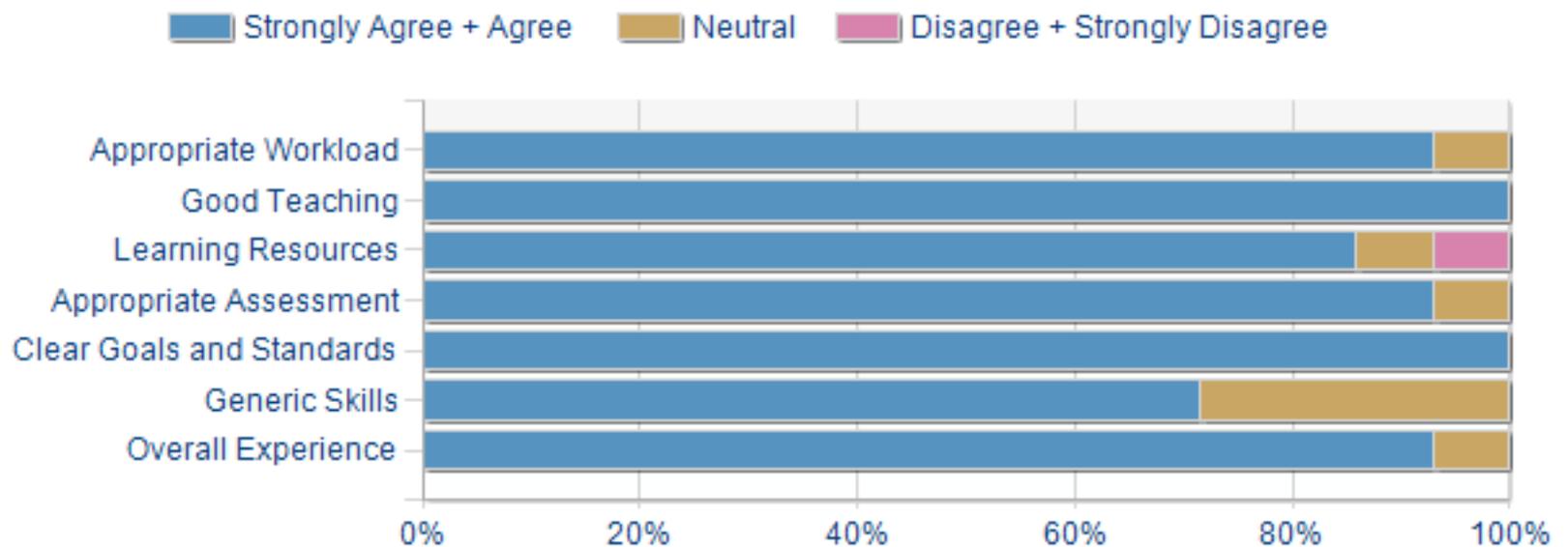
Student feedback

Best aspects of paper: learning how to find stories in data, learning to use Excel, how to formulate graphs, guest speakers, whole new skill set, class discussions, sparked my interest to pursue data journalism further

Things to improve: class could be shorter, more time on Datawrapper, more time on Google Fusion tables, more revision at the end, more guidance on the presentation assignment, too much required in the final assignment

3-hour class: Class structured well/good mix of activities (11)

Coding: most would have liked a practical session

Assessment: All said positive things. Proposal helped with final project. One said proposal could be due earlier, and one didn't like Week 13 deadline

Student presentations: all enjoyed these, learned a lot and found discussions interesting. But they could have wrapped up earlier

Changes for next year

Cover extra apps for infographics (eg Piktocharts) and mapping (CartoDB)

Spend more time on Google Sheets

Add in some practical sessions on coding

Review the statistics content with Robin Hankin

Provide extra readings for the non-journalism students

Get through student presentations by mid-semester break

Reduce final story word count (1500-1200), and require 3 not four sources

Bring in Excel revision exercises earlier