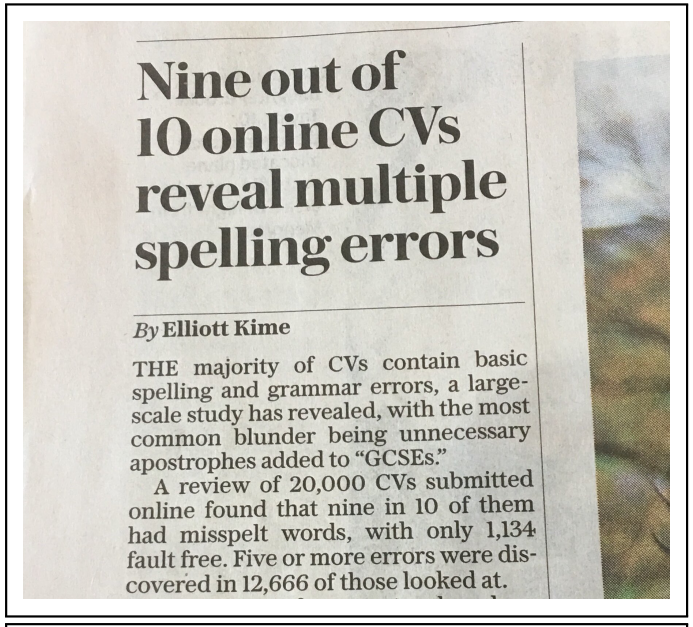
Assessing a Preliminary Year Data Literacy Module

MA0004 - Preliminary Mathematics II

Dr. Geraint Palmer

November 14, 2019

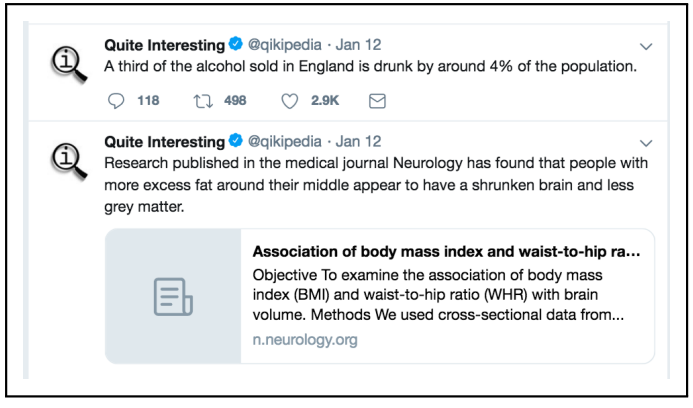
Module Aims


Nine out of 10 online CVs reveal multiple spelling errors

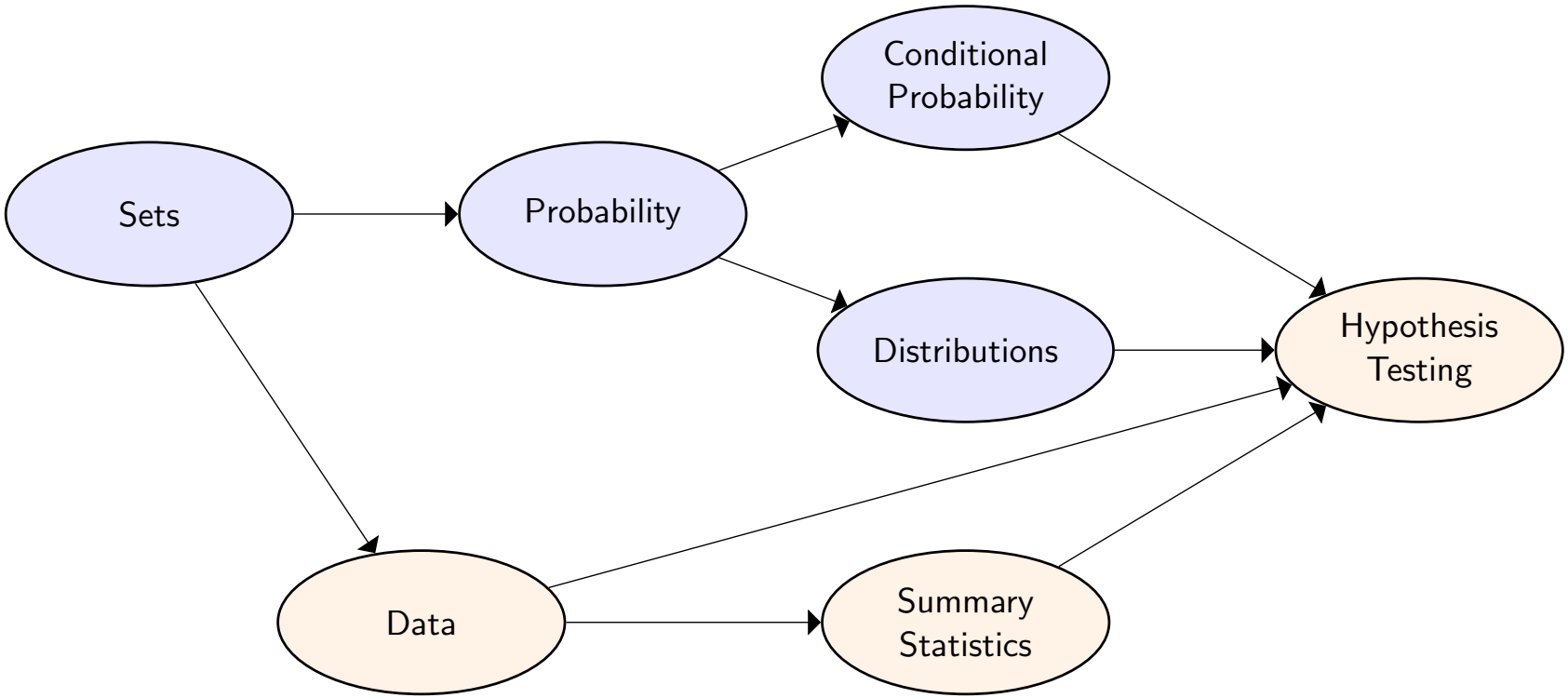
By Elliott Kime

THE majority of CVs contain basic spelling and grammar errors, a large-scale study has revealed, with the most common blunder being unnecessary apostrophes added to "GCSEs."

A review of 20,000 CVs submitted online found that nine in 10 of them had misspelt words, with only 1,134 fault free. Five or more errors were discovered in 12,666 of those looked at.

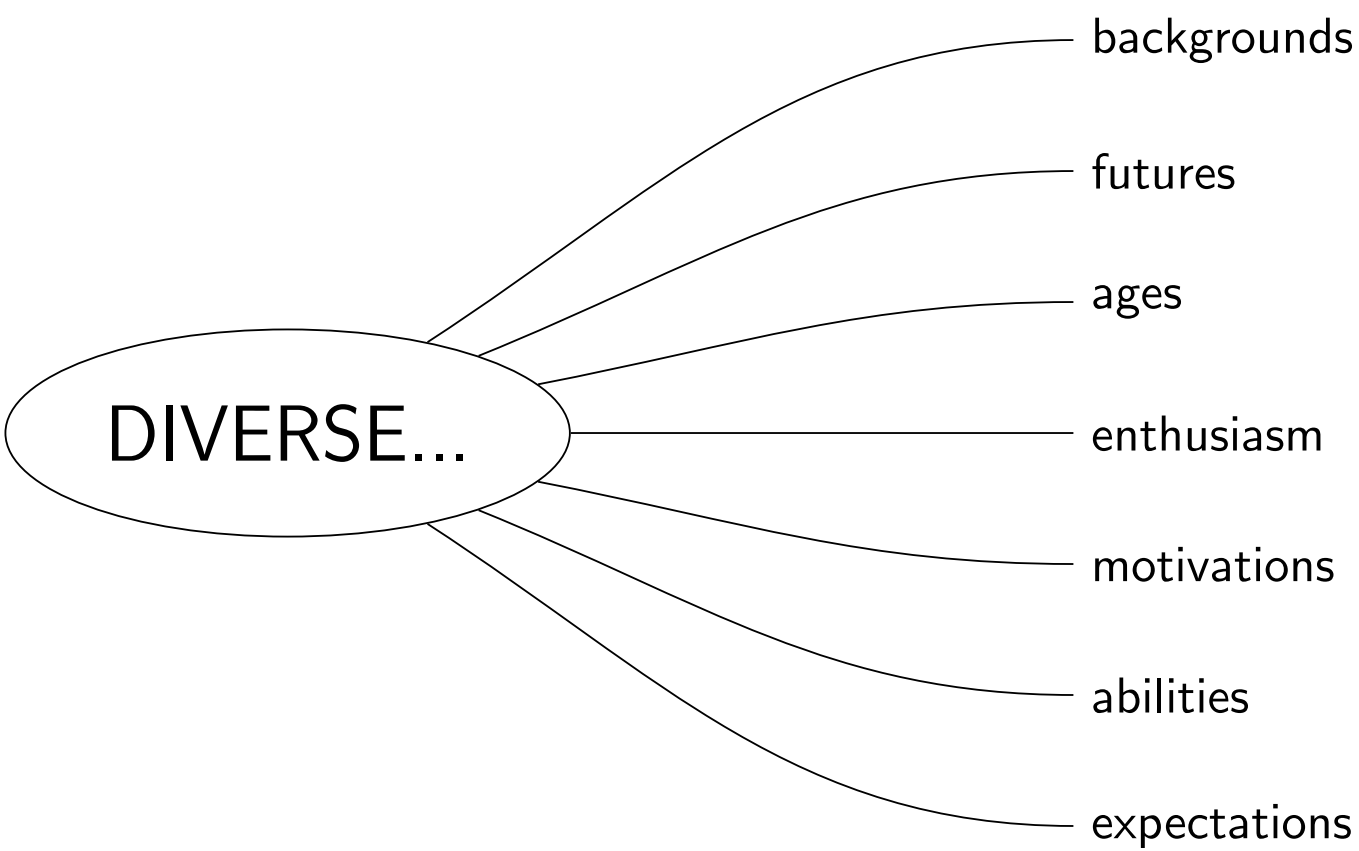


Course Contents



Class Demographics

Foundation year biomed, biochem, medicine, chemistry, and dentistry students



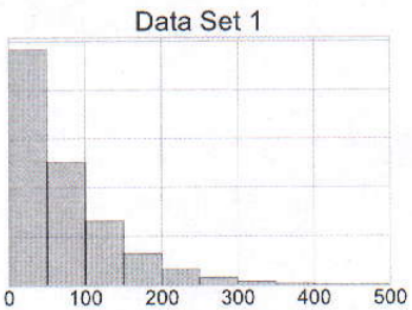
Group Coursework (15%)

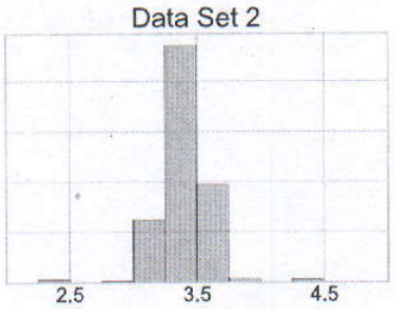
- Explore data
- Create of a piece of data literature
- Begin to develop "data wisdom"
- Group work & organisation
- Freedom to investigate topics outside of taught content

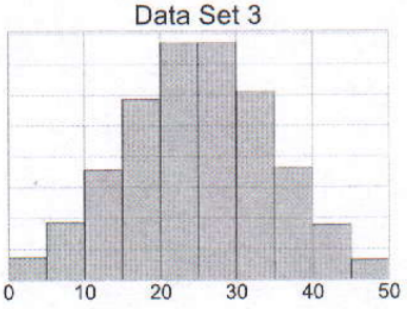
Exam (85%)

- Read data literature
- Some conceptual questions
- Rote probability questions
- Hypothesis tests by hand

3. Consider the three data sets, 1, 2 and 3, shown in the histograms below:







(i) Match the following standard deviations to the data sets: $\sigma_A = 10$, $\sigma_B = 0.2$, $\sigma_C = 75$.

(ii) Which data set has the greatest range?

(iii) For each data set, give the most appropriate measure of centrality, and explain why.

[5]

8. Six people were chosen at random to trial a new performance enhancing drug. They were asked to run 200m, once before taking the drug, and then again after taking the drug. The table below gives the times of the six runners:

Person	Speed Before	Speed After
1	43s	40s
2	48s	38s
3	45s	46s
4	43s	37s
5	44s	38s
6	40s	33s

(a) Draw a box and whisker plot for each of the speeds before and after taking the drug, clearly labelling the first quartile, median, and third quartile.

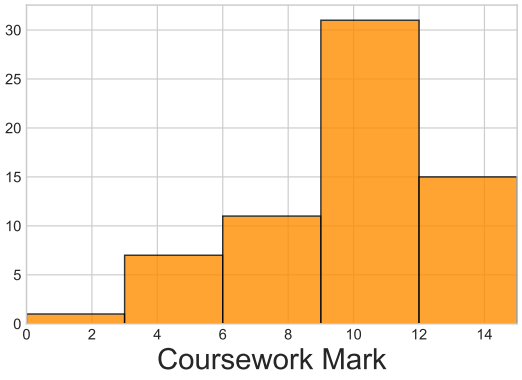
[7]

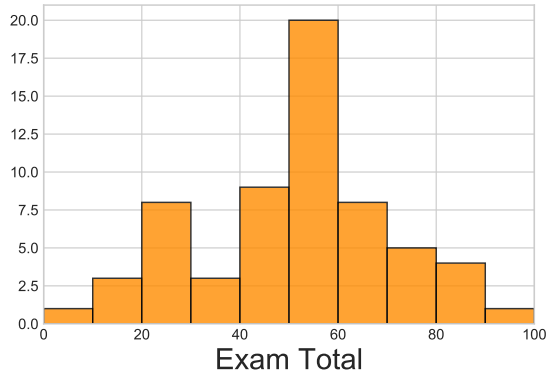
(b) Use an appropriate data visualisation to show the relationship between the speeds before and after taking the drug. Is there a positive, negative, or no relationship?

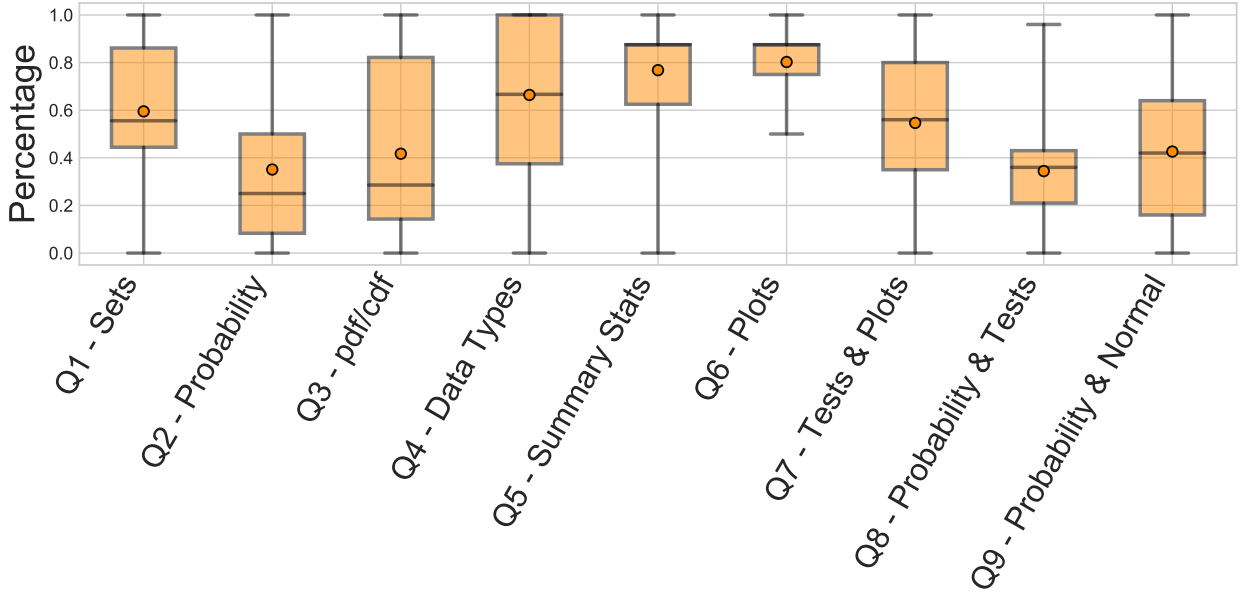
[4]

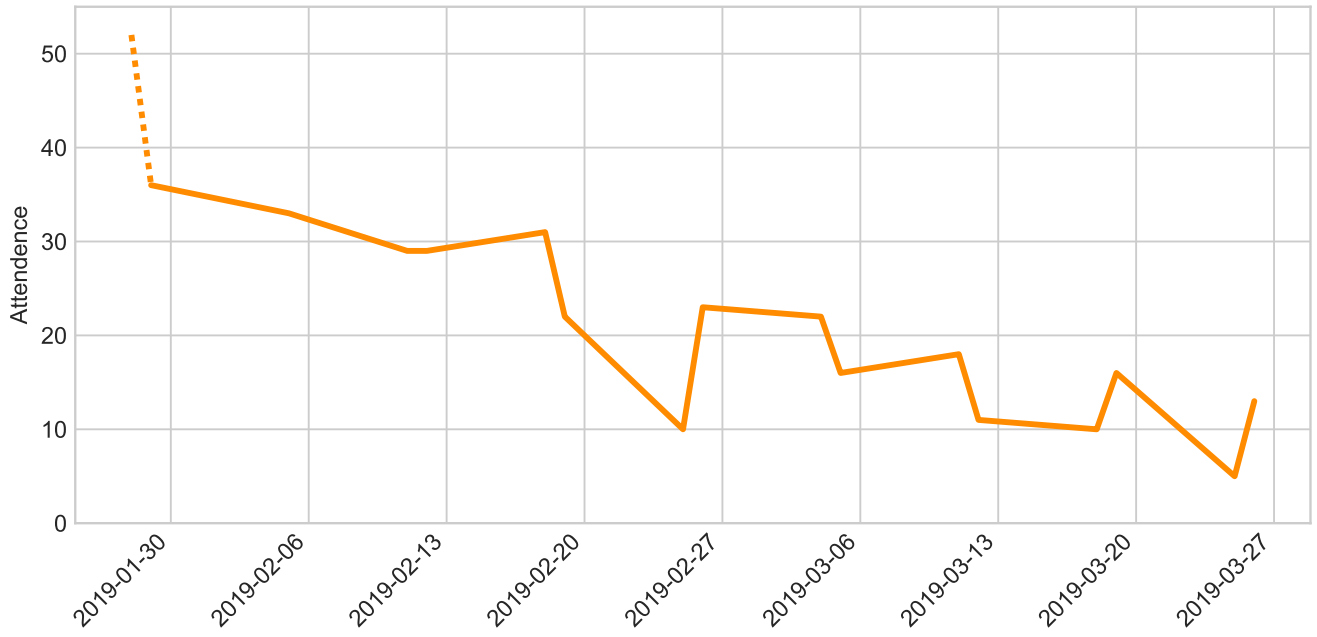
(c) Use an appropriate hypothesis test to investigate whether the drug has lowered the speed at which the runners can run 200m.

- State the null and alternative hypothesis. [2]
- Choose a confidence level and carry out the test. [9]
- State all conclusions. [3]


30
25
20
15
10
5
0
0
2
4
6
8
10
12
14
Coursework Mark


20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0
20
40
60
80
100
Exam Total


Percentage
1.0
0.8
0.6
0.4
0.2
0.0
Q1 - Sets
Q2 - Probability
Q3 - pdf/cdf
Q4 - Data Types
Q5 - Summary Stats
Q6 - Plots
Q7 - Tests & Plots
Q8 - Probability & Tests
Q9 - Probability & Normal



Question	% N/A (raw)	% Agree (valid NSS)	Avg Score(Stats)
15. Overall, I am satisfied with the quality of the module.	0%	50%	3.22

GAISE Report

1. Teach statistical thinking.
   - Teach statistics as an investigative process of problem-solving and decision-making.
   - Give students experience with multivariable thinking.
2. Focus on conceptual understanding.
3. Integrate real data with a context and purpose.
4. Foster active learning.
5. Use technology to explore concepts and analyze data.
6. Use assessments to improve and evaluate student learning.

https://www.amstat.org/asa/files/pdfs/GAISE/GaiseCollege_Full.pdf

Individual CW (50%)

- Read data literature
- Create a piece of data literature
- Apply probablistic ideas
- Connect theory and practise

Group CW (50%)

- Explor data
- Create of a piece of data literature
- Develop “data wisdom”
- Group work & organisation
- Encouraged to investigate topics outside of taught content

Option A

You have the following items: a pack of cards, coins, 6, 8, 10 and 12 sided die, and a roulette wheel. Invent a casino game and write a report selling this game to a casino.
You should include:

- a detailed description of the game
- all probabilities involved
- PDFs and CDFs of game outcomes
- expected winnings and reccommendations for prizes and charges

The casino is more likely to buy games that will make them a profit, whilst still being attractive to customers.

Option B

Your boss was supposed to attend a presentation on homelessness, but was unable to attend. They managed to obtain the presentation slides, and have asked you to write a one page report summarising the key points of the presentation. You should comment on:

- intetesting summary statistics
- significant relationships between variables
- extreme or unusual observations
- general trends

Ensure your writing is clear and accurate, referencing specific slide numbers and plots when discussing each finding.

Option C

Write a report on the "birthday paradox". It must include:

- a description of the problem and any assumptions
- a mathematical explanation of its solution

Then extend the problem in some way, for example:

- find some data and perform analyses on it - does this support the statement of the paradox?
- how does it apply to real life, e.g. sports teams or film casts - find data to support this
- change the problem, e.g. born in the same month / day of week

Group Coursework

Write a data analysis report on the following data, taken as a sample.

	A	B	C	D	E	F	G
1	***Triage Category***	***Gender***	***Age***	***Length of Stay***	***Weight on Admission***	***Weight on Discharge***	***Ambulance***
2	2	M	72	2.5	82.25	79.21	FALSE
3	2	M	72	2	81.09	79.06	FALSE
4	1	M	87	3	81.78	78.91	FALSE
5	3	M	89	6.5	81.29	80.12	TRUE
6	4	M	96	15.5	75.17	72.81	FALSE
7	2	M	58	8	78.92	76.81	FALSE
8	3	F	97	1.5	74.26	74.43	TRUE
9	5	F	71	6.5	68.67	67.67	TRUE
10	4	F	91	2.5	72.12	72.56	FALSE
11	4	F	65	6.5	67.53	68.01	FALSE

The report should highlight, with evidence, any interesting and useful insights the data set could offer. Evidence includes appropriate comments or measures of centrality, spread, and shape, appropriate data visualisations, and hypothesis tests.

Week	Session	Activity			
1	Monday (1 hour)	Introduction			
	Tuesday (2 hours)	Epidemiology activity			
2	Monday (1 hour)	Sets lesson			
	Tuesday (2 hours)	T/F	Sets activity	Sets tutorial	
3	Monday (1 hour)	Probability lesson			
	Tuesday (2 hours)	T/F	Probability activity	Probability tutorial	
4	Monday (1 hour)	Conditional probability lesson			
	Tuesday (2 hours)	T/F	Assessment workshop (assessment 1)		
5	Monday (1 hour)	Data lesson			
	Tuesday (2 hours)	T/F	Module feedback	Data tutorial	
6	Monday (1 hour)	Data visualisation lesson			
	Tuesday (2 hours)	T/F	Data viz activity	Data viz tutorial	
7	Monday (1 hour)	Distributions lesson			
	Tuesday (2 hours)	T/F	Critique data analysis reports		Assessment 1 hand in date
8	Monday (1 hour)	Stats tests lesson (χ^2-test)			
	Tuesday (2 hours)	T/F	Assessment workshop (assessment 2)		
9	Monday (1 hour)	Stats tests lesson (t-tests etc.)			
	Tuesday (2 hours)	T/F	Stats tests activity	Stats tests tutorial	
10	Monday (1 hour)	Data pitfalls video discussion			
	Tuesday (2 hours)	T/F	Assessment workshop (assessment 2)		
11	Monday (1 hour)	*Slack*			
	Tuesday (2 hours)	Assessment workshop (assessment 2)			
GSW	Monday (1 hour)	*Slack*			
	Tuesday (2 hours)	*Slack*			Assessment 2 hand in date

Worries / Places to Improve

Question	% N/A (raw)	% Agree (valid NSS)	Avg Score(Stats)
12. I have felt part of a community of staff and students whilst undertaking this module.	0%	39%	2.94

- self-selecting / assigned groups?
- regular reporting on projects?
- fair division of marks?
- communication under low attendance?
- use of electronic quiz software?
- class size / split?