STUDENT PROFILING
Lesson 2: Being Typical

Australian Curriculum: Mathematics (Year 10)

ACMSP248: Determine quartiles and interquartile range
ACMSP249: Construct and interpret box plots and use them to compare data sets
ACMSP250: Compare shapes of box plots to corresponding histograms and dot plots
ACMSP251: Use scatter plots to investigate and comment on relationships between two numerical variables

Lesson abstract

Students compare a single student's data to a sample of 550 students and use a range of statistical tools to explore whether their student is 'typical'.

Mathematical purpose (for students)

What new things can we learn about our subject by comparing their responses to a wider range of responses?

Mathematical purpose (for teachers)

To develop a 'statistical toolkit' for use in data comparison and analysis, and to explore what it means to be 'typical'.

Suggested presentation Two lessons of approximately one hour each.

Vocabulary encountered

- box and whisker plot
- histogram
- interquartile range
- outlier

Lesson materials

- Spreadsheet *2a Census@School 2014*
- PowerPoint *2b Census@School*
- TinkerPlots file *2c Census@School 2014* (optional)
- Students will need the *Data Table* worksheet and *Selected Responses* cards used in Lesson 1

We value your feedback after these lessons via http://tiny.cc/resource-feedback

Mathematics by Inquiry is an initiative of, and funded by, the Australian Government Department of Education and Training

Australian Academy of Science

aamt

Looking at the data

Give students the *2a Census@School 2014* spreadsheet and the same *Selected Responses* card they used in Lesson 1. The spreadsheet contains Census@School responses of 550 different Year 10-12 students. Allow the class a few minutes to explore and come to grips with the spreadsheet. If your classroom has access to TinkerPlots, students can also explore the *2c Census@School 2014* TinkerPlots file.

Ask students to find and highlight their subject's responses. If they find this difficult, prompt them to sort the responses by birthday or postcode.

Observe that the dataset may include the responses of students who were creative with their answers. Can the class identify any such responses? One way of finding these might be identifying participants who have repeatedly provided unusual answers, such as the female born on 20 July 1999 who reports speaking nine languages, sleeping 16 hours a night, and taking a two hour ferry trip to school every day. Discuss as a class whether to remove these responses or to take them into consideration in other ways.

If students have not completed the "Useful ways to represent data" and "Useful ways to analyse data" columns of the *Data Table* worksheet from Lesson 1, discuss this with them and complete it now.

Meaningful analysis

Teacher Notes

- In this section students look at how they might possibly analyse the data in the spreadsheet. Question 28b ("How much money did you earn/receive last week from your main source of income?") is used as an example because it contains numerical data, for which it is possible to calculate the mean, median and mode.

Ask students to look at the answers for Question 28b. Have them find the mean (=AVERAGE()), median (=MEDIAN()) and mode (=MODE()). The mean income is $74, the median is $35 and the mode is $0. What do these answers tell us about our dataset? Can students imagine different situations in which each of the values might be useful?

Discuss how you might analyse Questions 1 (sex) and 23 (favourite sport). These questions each contain categorical data. Would you expect there to be a relationship between students' answers to these questions? A two-way table with a selection of responses is included below and in Slide 2 of PowerPoint *2b Census@School*:

	Basketball	Cricket	Dancing	AFL	Soccer	Hockey	Martial arts	Netball	None	Skateboarding/ Rollerblading	Swimming	Tennis
Male	22	10	1	45	36	5	6	0	12	10	6	7
Female	19	0	41	6	18	13	4	62	30	4	21	12

Ask students: *what's interesting about this table?* For example, without this table we may not have noticed that only one male listed dancing as his favourite sport – so we have found something new by **contextualising** his response.

Discuss: What other calculations could you make with this dataset? You could take percentages horizontally (e.g. percentage of males who play basketball) or percentages vertically (e.g. percentage of basketball players who are male). What is the difference? Can students imagine different situations in which each value might be useful?

Ask students to select two other questions for which they think their subject's responses might be relevant. Ask them to conduct the analysis and discuss how their subject's responses compare to the mean/median/mode.

Graphing data

Teacher Notes

- In this section students look at ways of graphing the data in the spreadsheet. Questions 18, 26 and 30 have been used as examples because each set of responses can be best displayed as a different type of graph.

Show students the prepared graphs included in slides 4-8 of the PowerPoint, or have them create their own using the directions in the Teacher Notes below. Working through one question at a time, discuss how the graph

represents the data. Students can locate their subject in the graph. Does this give any new information about the subject?

Teacher Notes

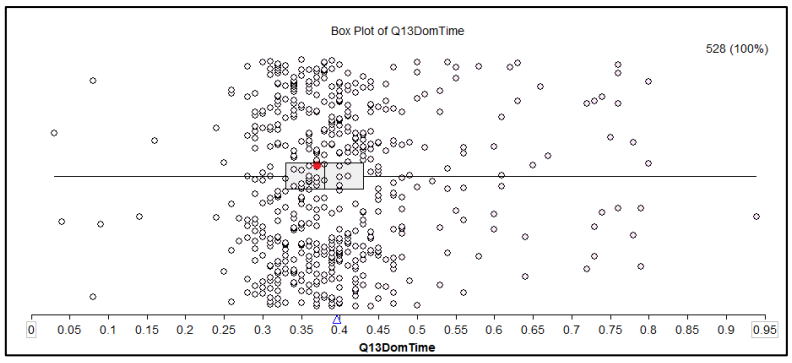
- Question 18 (slide 4) – a column graph
  - To create a graph for Question 18, the function "=COUNTIF(U1:U551, "Yes")" was used in Column U and then applied to Columns U-AH (all breakfast responses). This shows how many people answered "Yes" to each category of food. These results were then graphed as a column chart.
- Question 26 (slide 5) – box and whisker plots
  - Columns AP-AY were selected concurrently and graphed as a box and whisker plot. Excel plots the outliers. This may be a useful class discussion.
- Question 30 (slides 6-8) – histograms
  - Columns BQ-BY were each independently selected and graphed as a histogram. The axis bin width was set to 100 for each histogram.

Ask students to choose two other questions which they think might be relevant for their subject and consider the best way to represent the data for the question in order to contextualise their subject. For example, if a subject claims an unusual number of hours of sleep on a school night (Question 19), the student may choose to generate a box and whisker plot to see how extreme the response is.

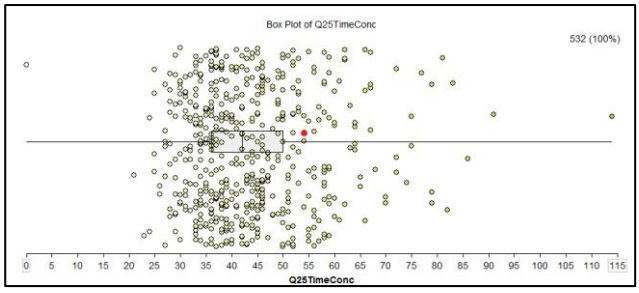
How typical is my subject?

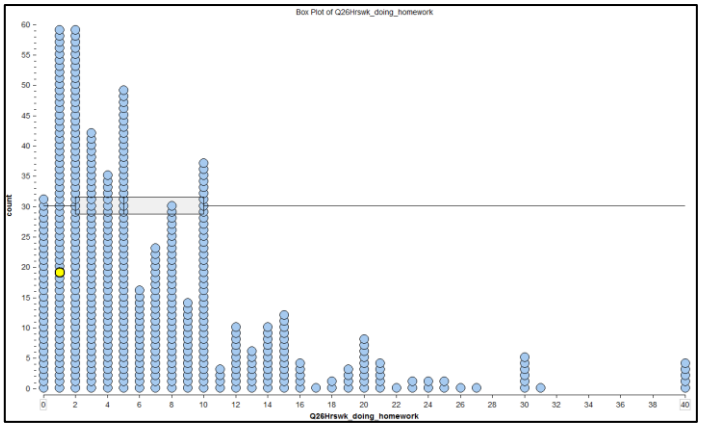
Students revisit their description of their subject's life from Lesson 1 and add or amend details based on their new findings. The following example (continuing the story of Subject #28 from Lesson 1) is given on Slides 9-12 if you need to illustrate the process for students:

- *As a tennis player, Hanif needs a good reaction time. The box and whisker plot shows that his reaction time is slightly better than average – his dominant hand reaction time was 0.37 seconds and the mean was 0.46 seconds.*

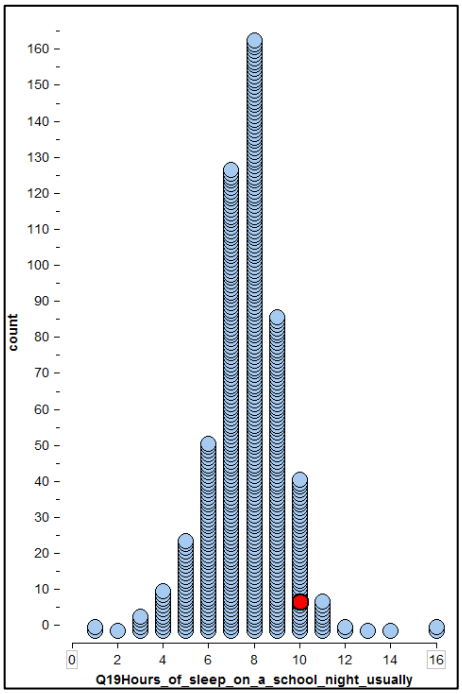


- *Hanif took a lot longer than other students to do the concentration game – his time was in the top quartile. He does one hour of homework a week, which is less than 75% of students. Maybe he has a hard time focusing on homework.*





All that tennis wears him out so he sleeps a lot – he spends more time sleeping than 87% of students do.



Reflection

Students take turns to present their new findings to the class. Would they describe their subject as **typical**? Why/why not?