Census At School

Census At School

Task Description
Students will be introduced to one variable and two variable data analysis tools including quartiles, box plots, and correlation using scatter plots.
The class will collect data using the Census at School survey, and then compare their data to a larger sample of Canadian data.
A lesson is also provided to teach the prior learning needed to support this task.

Big Idea
- One and two variable data analysis can be used to answer questions and help draw conclusions
- Different data representations can be used to show a variety of viewpoints
- Data can be used to show connections between two variables

Duration : 2-3 periods (does not need to be done all at once)

Recommended Materials
Chromebooks for each group
Desmos activity
Census at school survey and printed census at school data collection sheet
Measuring tapes and rulers
Applications for Data analysis:
- Excel (online or chrome app): does everything needed. You can log in using your board login details.
- Geogebra Spreadsheet: does everything needed in a basic way
- Google Sheets: does most things but not boxplots
- Desmos Boxplot Generator: can be used to fill in the gap in Sheets. A little clunky to use but shows the code.
- Desmos: can be used for scatter plots and some statistical analysis
- Graphing calculators (or emulator)

Learning Goal(s)
- Students will identify a question of interest based on a selection of variables in the census at school survey
- Students will collect information
- Students will represent and analyse single-variable data in various ways including using box plots
- Students will represent and analyse two-variable data in various ways including using scatter plots, correlation and different regression models
- Students will draw conclusions about the data they have collected and compare to data sets from across Canada

Sample Success Criteria
- I can collect data about myself and others and organise it for analysis
- I can represent data using box plots, histograms and scatter plots
- I can analyse central tendency and spread of one-variable data using mean median, mode, range and quartiles
- I can analyse two-variable data using correlation to show the relationship between two variables
- I can compare my data to class data and Canadian data
- I can draw conclusions from the collected and analysed data

Overall Expectations
AA1. Social-Emotional Learning Skills
develop and explore a variety of social-emotional learning skills in a context that supports and reflects this learning in connection with the expectations across all other strands

Census At School

A1. Mathematical Processes
apply the mathematical processes to develop a conceptual understanding of, and procedural fluency with, the mathematics they are learning

D1. Collection, Representation, and Analysis of Data
describe the collection and use of data, and represent and analyse data involving one and two variables

D2. Mathematical Modelling
apply the process of mathematical modelling, using data and mathematical concepts from other strands, to represent, analyse, make predictions, and provide insight into real-life situations

Social-Emotional Learning (SEL) Skills

- Recognizing and Identifying Emotions That Support Mathematical Learning
- Recognizing Sources of Stress That Present Challenges to Mathematical Learning
- Identifying Resources and Supports That Aid Perseverance in Mathematical Learning
- Building Healthy Relationships and Communicating Effectively in Mathematics
- Developing a Healthy Mathematical Identity Through Building Self-Awareness
- Developing Critical and Creative Mathematical Thinking

** This overall expectation is to be included in classroom instruction, but not in assessment, evaluation, or reporting. See **further information** about approaches to instruction that support all students as they work to apply mathematical thinking, make connections, and develop a healthy identity as mathematics learners to foster well-being and the ability to learn mathematics.*

Mathematical Processes
There are opportunities for students to engage in the **mathematical processes** throughout this task. For example:

- **Problem Solving:** critical thinking, math identity, knowledge, collaboration, lived reality, creative thinking, confidence
- **Reasoning and Proving:** justification, proportional reasoning, algebraic reasoning, spatial reasoning, numbers, operations, geometric properties, measurement
- **Reflecting:** identify what is working, what isn't working, appropriate strategy, reasonableness of their answer
- **Connecting:** connect different mathematical concepts
- **Communicating:** share ideas, understandings and solutions, provide feedback, pose questions
- **Representing:** represent math relationships using pictures, diagrams, numbers and symbols
- **Selecting Tools and Strategies:** test, revise, confirm reasoning, remembering how they solved a problem

Prior Learning and Placement of Task during Semester

Prior Learning
Grade 8 Data Expectations

1. Before tackling the main task, a review of mean, median, mode, range, and an introduction to quartiles and box plots should take place. Here is a one-variable lesson plan to accomplish that.
2. Desmos Activity - Interpreting Boxplots (Modified)
3. Knowledge of creating scatter plots and analysing correlation may be reviewed/pre-taught, or this task could be used to cover this.

Placement of Task during Semester
This activity can take place at any point of the semester and can be broken up to cover one-variable and two-variable data at separate times. Extension activities are provided to continue learning using these resources.

Census At School

Diagnostic Task & Look fors

The data analysis can be done as one bigger project or the one-variable and two-variable parts can be done separately, depending upon course organisation.

1. One-variable analysis
This diagnostic would work well at the end of the one-variable data lesson, provided above.

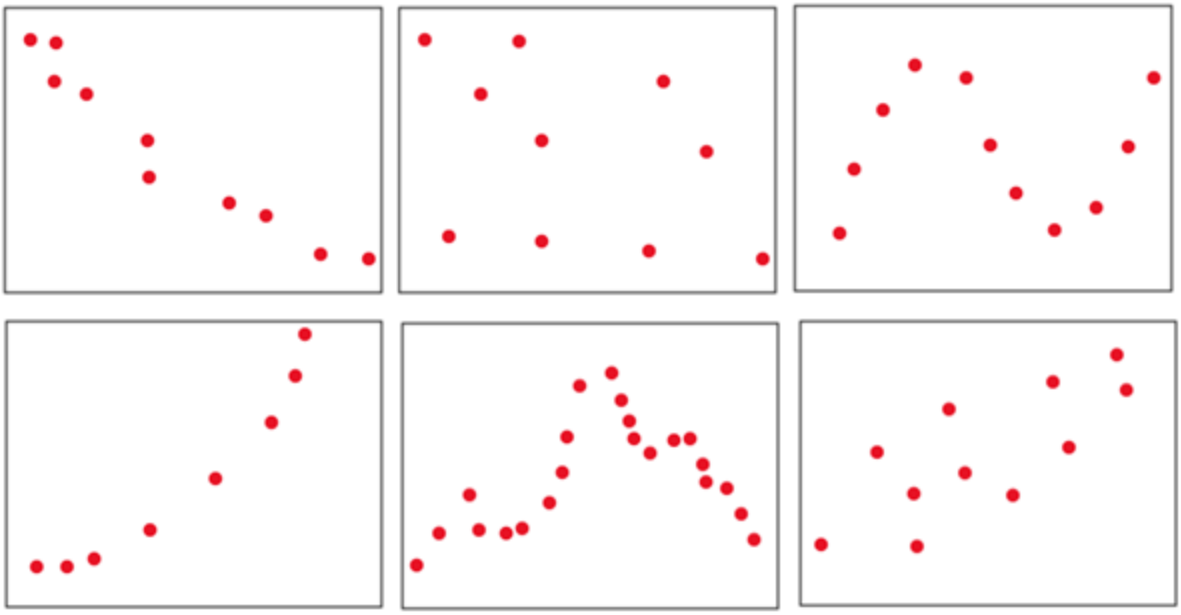
Task Slow Reveal Boxplots

- Present each slide one at a time giving students prompts provided in the teacher notes on the slides (these notes do not show while presenting). Give time for thinking and sharing before advancing the slides.
- Routine for slow reveal graphs - As more and more of the graph (plot) is revealed, students refine their interpretation and construct meaning. This routine increases access for students without sacrificing rigor or engagement.

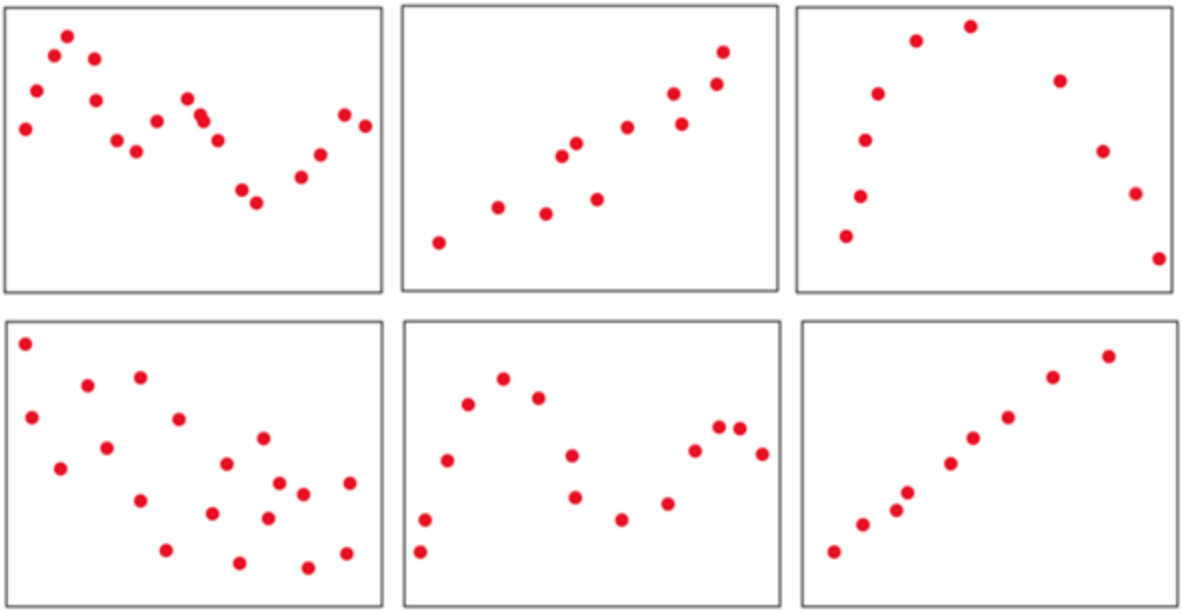
Look Fors

- Do students know the key features of and how to interpret a boxplot?
- Do students make connections between boxplots and measures of central tendency and dispersion?
- Do students use vocabulary such as quartile, median, minimum, maximum, range, interquartile range, box plot (or box and whisker)?
- Do students make accurate comparisons between similar data sets represented in boxplots?
- Do students draw reasonable conclusions about familiar data represented in boxplots?

2. Two-variable analysis (initially developed for task Lumber Pricing Task)
Take a look at the following scatterplots.
Which ones would be appropriate to use a line of best fit, a curve of best fit and which would not?
For those that show linear correlation, rank the scatter plots from weakest to strongest correlation.



Census At School



Discuss and compare your choices.

Look fors

- Understands that a often a 'line' of best fit may not be appropriate
- Use of vocabulary: correlation, strong/weak, linear/non-linear
- Can select appropriate regression type
- Can evaluate strength of regression
- Communication of idea

Educator Notes:

- Teachers can decide whether to let students have individual think time first or jump straight into group discussion
- In class student copy - option to cut and physically manipulate
- Jamboard option for online learning - Please have students drag and drop the frames into the order they decide to rank the images. Alternative option: Educator may choose to put into a cardsort in Desmos

Task

Part 1: Data Collection

The data analysis can be done as one bigger project or the one-variable and two-variable parts can be done separately, depending upon course organisation.

It is suggested that you email parents the information at the following link to inform them that their child's data is being submitted anonymously to the census at school project and they should email you with questions:
https://censusatschool.ca/parents/

Teachers will need to create an account and register each class (this takes 2 min). This will generate a class code and password for the students to use when taking the survey.
https://censusatschool.ca/

Census At School

Students will fill in the census survey online using the code and passwords created. The survey covers a variety of different topics, including questions that involve measuring various body dimensions. These measurements could be measured ahead of time using the measurement recording slips provided. (Alternatively, the survey questions can be found here.) The census at school site also includes a measurement guide if you wish to use it.

After the students have completed and submitted their surveys you can access your submitted class data (when logged in), Canadian summary data and some international data here. This information can be viewed online, or downloaded as a Microsoft Excel file or CSV file for importing into other spreadsheet applications. You may need to establish a method for distributing this information to the students, such as downloading the information and transferring it to a shared Google sheet.

Suggestion: teachers may want to look over the class data to verify for possible errors (for example measuring in inches instead of cm). This can also be a point of discussion with the class.
Possible discussion points:

- What errors are possible when we complete the survey questions? How could this influence the results?
- What can be done to reduce these errors?
- What are some possible sources of bias in this survey?How could this influence the results?
- What can be done to reduce these sources of bias?
- Can you find good examples in the survey where sources of errors or bias were minimized?

Part 2: One-Variable Data Analysis

Students will work through the mathematical modelling process outlined in Curriculum Expectation D2 to identify a question of interest, analyse data and draw conclusions to answer their question.

Using the census at school questions, students should select one variable they wish to investigate further and pose a question of interest based on that census topic. For example, *"Do students in my class exercise more than the Canadian average?"*

Possible discussion notes

- What makes a good question?
- What questions could we answer using this data?
- Which questions are interesting to you?
- Pick a question and create a hypothesis about that question.
- Students may rely on stereotypes when making hypotheses. This would be a good opportunity, with caution, to challenge and discuss. The safety of all students in the class would have to be considered.
- Make a prediction as to what you think the results may show.
- What tools could be used to analyse and model your data?
- At this stage, it may be useful to discuss quantitative vs qualitative data and categorical vs discrete vs continuous data, so that students choose an appropriate variable type for the desired analysis. Refer to the Glossary.

Students will use technology (but it can be done by hand) to organize the results of their census question of focus, calculate the one-variable data statistics (mean, median, mode, range, quartiles and interquartile range), and display the information graphically (the main focus should be the new box plot, but a histogram would probably be the students first choice and the automatic graph generated by the spreadsheet).

Students will use their analysis to draw conclusions, make predictions and answer their question of interest.

Possible Discussion Notes

- What does the data tell you? Was your hypothesis correct? How has your thinking changed?
- Students will use their analysis to draw conclusions and answer their question of interest.
- How can the model be used to answer the question of interest?
- How well does the model fit the context?
- What are the potential limitations of the model?
- What predictions can be made based on the model?

Students can share their findings in groups or as a class.

Part 3: Two-Variable Data Analysis

Students will work through the mathematical modelling process outlined in Curriculum Expectation D2 to identify a question of interest, analyse data and draw conclusions to answer their question.

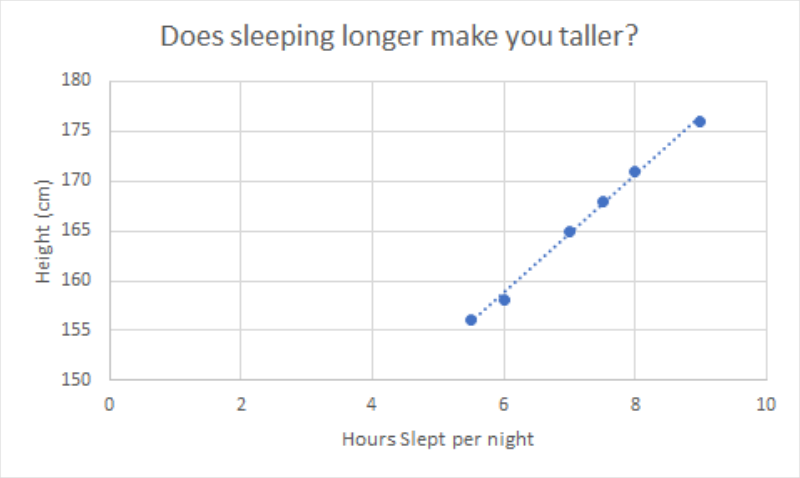
Using the census at school questions, students should select two variables they wish to investigate further and compare. They will pose a question of interest based on that census topic. For example, *"Do taller students exercise more?"*

- What makes a good question?
- What questions could we answer using this data?
- Which questions are interesting to you?
- Pick a question and create a hypothesis about that question.
- Make a prediction as to what you think the results may show.
- What tools could be used to analyse and model your data?
- At this stage, it may be useful to discuss quantitative vs qualitative data and categorical vs discrete vs continuous data, so that students choose an appropriate variable type for the desired analysis. Refer to the Glossary.

Students will use technology (but it can be done by hand) to organise the results of their census question of focus, display the two-variable data as a scatterplot, investigate the most appropriate regression line, and the strength of correlation (for linear models).

- What does the data tell you? Was your hypothesis correct? How has your thinking changed?
- Students will use their analysis to draw conclusions and answer their question of interest.
- How can the model be used to answer the question of interest?
- How well does the model fit the context?
- What are the potential limitations of the model?
- What predictions can be made based on the model?
- At this stage, it may be useful to discuss correlation vs causation. Refer to Glossary. Correlation does not imply causation, especially with small sample sizes. Just because the scatter plot may show a strong correlation (pattern), does not mean that there is causation (link).

Census At School



Students can share their findings in groups or as a class.

Look Fors:

- Can students pose a question of interest that relates to the Survey?
- Do students create graphical displays (histogram, box plot, scatter plot)?
- Do students calculate one and two variable statistics?
- Do students draw appropriate conclusions that connect their work to their questions of interest?

Assessment

Observations during the lesson focusing on the process expectations and success criteria

- I can connect the mathematical process skills to the math I am doing
- I can explain how using the mathematics process skills can help me solve a problem

A1. Mathematical Processes

apply the mathematical processes to develop a conceptual understanding of, and procedural fluency with, the mathematics they are learning

Ideas:

- Reflect on the success criteria
- Have students create a one page summary of their group's work
- Assign individual students to extract a different class sample from the data and repeat some or all of the one-variable and two-variable tasks
- Collect evidence for your portfolio

This task can be included in the **Student Portfolio of Process Expectations** (Google Slides)

Additional Resources

Shortcut to data results page for Census at School which includes canadian and international data, teacher and student instructions, privacy information, parent information, and other lesson ideas.
https://censusatschool.ca/data-results/

Census At School

Additional learning resources and lesson ideas using Census at School data:
https://censusatschool.ca/learning-activities/grade-9-to-12/

Measures of Spread (terms and definitions) from Statistics Canada
https://www150.statcan.gc.ca/n1/edu/power-pouvoir/ch12/5214876-eng.htm

Geogebra Spreadsheet Tutorials. Shows students how to use the one variable and two variable statistics functions in geogebra.
https://www.geogebra.org/m/XUv5mXTm#chapter/400292

Slow Reveal Graphs: An Instructional Routine to Promote Sensemaking about Data

Desmos Graphing Calculator: Linear Regressions
Google Sheets: How to Find Slope in Google Sheets
How to create a scatterplot in Google Sheets

Glossary:

minimum: the lowest value in a set of data

maximum: the highest value in a set of data

range: the difference between the highest and lowest values in a set of data. The range is a measure of the dispersion of the data. For example, in the data set 8, 32, 15, 10, the range is 24, that is, 32–8.

box plot: a graphic representation of the spread of a data set.
A rectangle (box) shows the spread of the central half of the distribution, with the first quartile on the left edge, the third quartile on the right edge, and the median as a line within the box. Lines (whiskers) extend from the sides of the box to the lowest and highest values that are not outliers. Potential outliers are marked with a symbol beyond the whiskers. Also known as box-and-whisker plot.

quartile values: Values that divide a sequenced data set into four parts, each of which represents 25% of the data falling within that range. For example, 25% of the data is below the point that defines the first quartile, which is the middle number between the smallest number in the data set and the median.

median: one of the measures of central tendency. The median is the middle value of an ordered list. For example, 14 is the median for the set of numbers 7, 9, 14, 21, 39. If there is an even number of data values, then the median is the average of the two middle values.

lower quartile: the number dividing the first and second quartile

upper quartile: the number dividing the third and fourth quartile

interquartile range (IQR): the difference between the lower quartile and the upper quartile

Census At School

qualitative (categorical) data: Non-numerical data that can be organized by categories, such as type of pet, or eye colour.

quantitative data: data that is numerical and acquired through counting or measuring; for example, number of sides of a three-dimensional object or amount of rainfall in a season.

discrete data: a form of quantitative data. Discrete data is data that can be counted, such as the number of pets, the number of siblings, or the number of buttons.

continuous data: a form of quantitative data. Continuous data is data that can be measured but not counted, such as time, height, and mass.

scatter plot: a graph designed to show a relationship between corresponding numbers from two sets of data measurements associated with a single object or event; for example, a graph of data about students' marks and the corresponding amounts of study time. Drawing a scatter plot involves plotting ordered pairs on a coordinate grid.

correlation: Correlation is a statistical measure (expressed as a number) that describes the size and direction of a relationship between two or more variables. A correlation between variables, however, does not automatically mean that the change in one variable is the cause of the change in the values of the other variable.

causation: Causation indicates that one event is the result of the occurrence of the other event; i.e. there is a causal relationship between the two events. This is also referred to as cause and effect.