Statistics: Math Academic 12

At-a-Glance

NOVA SCOTIA
NOUVELLE-ÉCOSSE

2026

Statistics: Math Academic 12 – At A Glance

Prepared by the Department of Education and Early Childhood Development

This is the most recent version of the current curriculum materials as used by teachers in Nova Scotia.

Learners will analyse how data is collected, interpreted and used to inform decision making in society.

Introduction to Statistics and Data Science

Rationale

Data can be used to analyze critical social issues and can influence decision making and policy in society. This course will prepare learners for a data-driven world and emphasize the importance of data literacy. Learners will study how data is collected (e.g. surveys, census, experiments), the difference between primary and secondary data, how to see bias in data as well as how data is interpreted to convey meaning. They will explore the difference between privacy settings such as cookies, location services, and targeted advertising. Learners will be provided with an opportunity to study the uses and misuses of data as well as comparing how different data sets are interpreted by various media sources. This outcome can be further developed across subsequent modules. Learners will also analyse how data is represented to interpret meaning and investigate the use of statistics in data science careers. An important exploration for learners is in coming to know more about the story that exists within data.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Personal and Career Development (PCD)

Indicators

- Identify how different forms of data are collected (COM, CT)
- Compare primary and secondary data sources used in observational studies and experiments. (COM, CT)
- Investigate how a variety of data sources and forms of data are used to convey meaning. (COM, CT)
- Evaluate the ethical considerations and bias in collecting and analysing data. (CZ, CT)
- Investigate how governments and organizations use data to inform decision making, policy, and/or legislation. (CZ, CT, PCD)

Concepts (and Guiding Questions)

Understanding Data Collection

- Why is data and statistical thinking important in our lives?
- How are various forms of data collected?
- Why does the method of data collection matter?
- How can I identify the information or data that I need to make informed decisions?

Primary versus Secondary Data

- How is data used for experiments and observational studies?
- What is the difference between primary and secondary data?
- How could primary and secondary data be collected on the same topic or for the same purpose?

Using Data to Convey Meaning

- What do different visual representations (e.g. graphs, charts, etc.) tell us about the data?
- How can the same data lead to different conclusions based on a given audience?
- How have different sources and forms of data been used to shape understanding?
- How can representations of data be used to inform or persuade an audience?

Ethical Considerations

- How can bias (e.g. selection, confirmation, measurement) affect data collection and interpretation?
- How can bias in data collection and analysis be identified?
- How can I gather data from various sources in ways that are authentic and accurate?
- What are ethical considerations surrounding the gathering and use of data about people?
- How can I identify uses and misuses of data?

Decision Making

- How do data and statistical sciences help us make decisions?
- How are various careers in data science part of data-based decision making?
- How have organizations justified decisions using data?
- How can data be used to advocate for change?
- What types of data do governments and organizations rely on to make decisions?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Learn about the data science process by considering a local question or issue.
  e.g. social media usage, sleep habits, sports performance data, local decision making.
- Explore how data is used to convey information.
  e.g. comparing visual representations of the same data set, the use of data across media articles.
- Determine how to identify and gather reliable data.
  e.g. recommending data needed to inform decision making, reviewing potential bias or gaps in data sets.
- Examine bias, privacy, and ethical considerations in data collection and use.
  e.g., explore how personal data is collected through cookies, location services, and apps; analyse how data is used in targeted advertising; review privacy settings and terms of service; question how consent is obtained and whether the data collected is necessary or used for other purposes.
- Review how data is used to inform decision making.
  e.g. selecting a site for a public building (e.g. hospital, school, library), making public health decisions (e.g. vaccination funding or clinics, accessing social services, etc.), school funding, transit route locations or proposed changes.

Learners will investigate ways that statistical analysis can be used to understand data

Introduction to Statistics and Data Science

Rationale

Statistics allows us to analyse data in a meaningful way. Learning about basic statistical analysis, descriptive statistics, and using tools and techniques to collect and analyse, learners will have the ability to interpret and present data. Algebraic reasoning plays a key role in statistical analysis, as learners use formulas to represent relationships within data. They interpret variables within statistical formulas and rearrange equations to solve for unknown values. Using measures of central tendency learners will explore relationships within data and use measures of variability to understand the spread of data. They will explore the difference between sample and population standard deviation and examine outliers and how these effect data spread.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Solve for unknown values of measures of central tendency and variability through algebraic reasoning (COM, CT)
- Investigate data using measures of central tendency (i.e. mean, median, mode) and variability (i.e. range, interquartile range, standard deviation) (CT, TF)
- Analyse how measures of central tendency and spread influence data interpretation. (CT, TF)
- Use measures of spread and univariate statistics to make informed decisions. (COM, CT, TF)

Concepts (and Guiding Questions)

Algebraic Reasoning

- How can algebraic formulas help me represent and analyse data?
- How can I rearrange formulas to solve for unknown values?
- How can solving for unknown values help me to better understand real-world data?

Measures of Central Tendency and Variability

- How can using measures of central tendency help me to understand a data set?
- What is the difference between measures of spread and measures of variability?
- How can I determine which measure of central tendency is most helpful in a given context?
- How can algebraic formulas be used to calculate measures of central tendency?

Influencing Data Interpretation

- What information from measures of spread and variability can help me better understand a data set?
- How can I use technology to explore measures of central tendency and variability?

Learners will investigate ways that statistical analysis can be used to understand data (cont.)

Using Univariate Statistics to make Decisions

- What is the role of univariate statistics in making decisions?
- How do quartiles and percentiles help me to understand data?
- What types of decisions can be made about various data sets?
- How can different measures of spread inform decision making?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Calculate measures of spread and variability to help interpret data.
  e.g. compare how different measures influence the interpretation of data, determine the impact of outliers on calculated data.
- Use graphical representations of univariate statistics to convey meaning.
  e.g. identify how a histogram vs a box plot shows spread and outliers differently.
- Explore different forms of graphical representations that help to express and interpret data.
  e.g. compare how two news sources present the same data using different graphs and statistics.
- Use data in relation to a local community issue to propose solutions.
  e.g. analyse statistics related to public health, education, sustainability, or a local community concern to propose possible solutions.

Learners will evaluate categorical and quantitative (discrete and continuous) data.

Introduction to Coding

Rationale

Understanding the distinctions between different types of data allows learners to make more informed decisions when interpreting and analysing data in everyday life. Learners will explore categorical and quantitative (discrete and continuous) data. Data may be presented visually in different ways which best match the type of data and the intended audience. Within this exploration, learners will use and compare multiple graphical representations of data such as bar graphs, histograms, box plots, dot plots (line plots), scatterplots, and frequency polygons to visualize, represent, and interpret data. Learners will also use discrete data to create box plots identifying the minimum, maximum, median, lower quartile, Q1, upper quartile, Q2, and determine the interquartile range and if outliers exist. An important aspect of this outcome is the comparison of different types of visualizations, to better make decisions about sets of data. Teachers are encouraged to explore how technology can be used to generate and compare multiple representations of data for analysis.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)

Indicators

- Identify examples of categorical and quantitative data. (COM, CT)
- Classify data sets based on their characteristics. (COM, CT)
- Use graphs and diagrams to represent data. (COM, CT)
- Analyse categorical and quantitative data (COM, CT)

Concepts (and Guiding Questions)

Types of Data

- Where are different types of data used?
- How is meaning conveyed in the data that I find?

Characteristics of Data

- What are the characteristics of discrete and continuous data?
- What is the difference between categorical and nominal data?
- How could I determine how to classify data?

Graphing Data

- Based on the data, how could I determine what type of graph to use?
- How do I determine which type of visual is used to interpret the data?
- How are outliers represented in various graphs?

Learners will evaluate categorical and quantitative (discrete and continuous) data. (cont.)

Visualization of Data

- What comparisons can I make between graphs generated and their intended use?
- How do dot plots and frequency polygons assist in the analysis of data?
- What summary statistical measures would best allow comparison of one data set to another?
- How do outliers skew data?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Explore the use of data across media.
  e.g. examine a data set that includes both categorical (e.g., product type) and quantitative (e.g., cost or rating) data.
- Choose specific types of visuals based on data type.
  e.g. organize data sets into categorical or quantitative (discrete or continuous) types; compare how different types of data are collected and used; determine what type of data is needed to answer specific questions.
- Use graphs and diagrams to compare categorical and quantitative data.
  e.g., explore data sets related to labour market information or data relating to health and wellness; create visual representations from the data.
- Use data to identify patterns, trends, distributions.
  (e.g. spread or concentrations) and draw conclusion: e.g. data related to the environment, social media use, or financial literacy.

Learners will construct a coding sequence for the analysis of measures of spread and univariate statistics.

Introduction to Coding

Rationale

Coding is an important skill to develop and use in data collection and analysis. Coding empowers learners to manipulate data efficiently, uncover patterns, and improve accuracy through cleaning techniques. By implementing summarization, visualization, and randomness control with seed commands, they can develop a solid foundation for analysis. Through these techniques learners will gain a better understanding of how to perform simple data cleaning to remove and fix errors, sort and rank data, and use loops.

Competencies

- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Develop basic coding sequences to collect data. (CT, TF)
- Apply coding to clean, smooth, and manipulate data sets. (CT, TF)
- Solve for measures of central tendency and variability. (COM, CT, TF)
- Justify conclusions using results of analysed data. (COM, CT)

Concepts (and Guiding Questions)

Developing Code

- How does coding help in gathering and exploring data sets?
- How do loops help in the gathering of data?
- How can I problem solve a coding error?
- How can I improve the overall performance and accuracy of my code?

Cleaning and Smoothing the Data

- What are key characteristics of effective coding sequences?
- What data moves are necessary when organizing and managing data?
- How do I resolve issues that might exist in raw data (e.g., missing values, errors, outliers)?
- How do decisions about cleaning data affect the results?

Solving for Univariate Statistics

- What sequences of code are needed to solve for measures of central tendency and variability?
- How can I identify errors in coding by reviewing results?
- How can I determine the accuracy and reliability of the results of my coding?

Justify Conclusions

- What inference can I make from the data set used?
- How does sample size impact the types of conclusions that can be made?
- How can I determine trends in data sets?
- How can I determine the reliability of my results?

Learners will construct a coding sequence for the analysis of measures of spread and univariate statistics. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Use code to collect and organize data.
  e.g., collect information (e.g., screen time, spending habits, music preferences) and use code (e.g., spreadsheets or basic scripts) to organize and sort responses.
- Clean and preparing data sets.
  e.g., work with a messy data set (e.g., budget tracking, daily expenses, or fitness data); use code to clean, organize, and prepare the data for analysis.
- Explore how changes in data affect results.
  e.g., use code to modify a data set (e.g., add an unusually high expense or score) and observe how measures of central tendency and spread change.

Learners will evaluate how simulations are used to support data analysis in coding.

Introduction to Coding

Rationale

Simulations play a key role in data analysis by allowing learners to model real-world scenarios, test hypotheses, and observe patterns in a controlled environment. Learners will use simulations to model a contextual situation mathematically and use the models for analysis. For coding-based simulations, learners could generate random data sets using seed commands to model statistical distributions. Learners will understand how different factors influence outcomes and simulate real-world scenarios like customer behavior or scientific experiments to draw insights from data. In these types of simulations learners can create visuals to describe how a quantitative variable change over time and use a simulation to summarize numerical data in relation to the context of the problem.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Interpret pre-existing code that models univariate data. (COM, CT, TF)
- Select simulation parameters to test different conditions. (CT, TF)
- Investigate the role of loops in a coding program. (CT, TF)
- Analyse simulations that model non-fictitious data. (COM, CT, TF)

Concepts (and Guiding Questions)

Using Pre-existing Code

- How was coding used to model univariate statistics?
- How can coding assist in modeling data sets?
- What coding was done to store lists of data for modeling?

Simulation Parameters

- What is the role of simulation in exploring the sampling distribution of a statistic?
- How can I identify a parameter to test which will help me get more information from a data set?
- What other information might make interpretations of my data sets more informative?

Using Loops in Coding

- What is the purpose of a loop in a program?
- How does increasing the number of trials affect the results?
- How does looping improve efficiency and reliability in simulations?

Learners will evaluate how simulations are used to support data analysis in coding. (cont.)

Data Simulations

- What type of graph can I choose to examine different data sets?
- How do visuals help me to interpret the results?
- How can coding help in visualizing and describing variation and statistical parameters?
- How can I use simulation results to inform predictions or decisions?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Use simulations to explore univariate data.
  e.g., simulate repeated trials that use probability to determine if outcomes are fair over time; examine assumptions in a model, identify missing variables, and consider how simulations may oversimplify complex real-world situations.
- Examine how coding structures (loops, variables, randomness) impact the results of simulations.
  e.g. use loops to run repeated trials, generate random values to simulate uncertainty, and decide how to store results for analysis and visualization.
- Compare simulated data to real-world data.
  e.g. use simulations to model real data sets (e.g. sports statistics, environmental data, or school-related data) and evaluate the effectiveness of the model.

Learners will investigate probabilistic modeling through problem-solving.

Using Data to Explore Variation

Rationale

Probabilistic modeling helps learners understand uncertainty, make predictions, and solve real-world problems using data-driven approaches. Probabilistic modeling is widely used across various industries to make predictions, optimize decisions, and assess uncertainty. Some applications include finance and risk analysis, within healthcare, and in weather forecasting. Learners will explore theoretical and experimental probability calculations. They will learn how to solve problems involving independent and dependent events, mutually exclusive and non-mutually exclusive events, sampling with and without replacement, and conditional probability. They will also explore the relationships among the fundamental counting principle, the Addition Law of Probability, and binomial probability distributions in calculating probability. Learners will use visualizations to model and make sense of probability situations.

Competencies

- Communication (COM)
- Critical Thinking (CT)

Indicators

- Compare outcomes of theoretical and experimental probability. (COM, CT)
- Evaluate sources of variation between theoretical and experimental probability, including sample size, randomness, and external factors. (COM, CT)
- Solve probability problems involving (COM, CT)
    - independent and dependent events
    - sampling with and without replacement
    - mutually exclusive and not mutually exclusive events
    - conditional probability
    - binomial probability distribution
- Use tree diagrams, lists, and tables to solve problems involving (COM, CT)
    - relative frequency
    - combined events
    - fundamental counting principle

Concepts (and Guiding Questions)

Experimental and Theoretical Probability

- Why might the outcomes of theoretical probability and experimental probability be similar or different?
- What is the influence of sample size on theoretical probability?
- What can the probability of an event tell us about how likely an event will happen?
- How can I use probability to decide if something will always, sometimes, or never happen?

Variation in Probability

- What is the role of probabilistic models for describing variation?
- What strategies can I use to describe the impact of the sampling variation of a statistic (e.g., the sample mean, external factors, etc.)?
- What effect does the sample size have on variation?

Learners will investigate probabilistic modeling through problem-solving. (cont.)

Solving Probability Problems

- How can I determine how to solve various probability questions with the data and context I have?
- How can I tell if events are dependent or independent and how does this impact the way I calculate probability?
- How can I tell if events using sampling with and without replacement and how does this impact the way I calculate probability?
- How can I tell if events are mutually exclusive and not mutually exclusive and how does this impact the way I calculate probability?
- How does the binomial probability distribution predict the likelihood of an event?

Using Tree Diagrams, Lists, and Tables

- How is the fundamental counting principle used to find probabilities?
- How can I use tree diagrams or relative frequency tables to help in finding the probabilities of specific events?
- How does a tree diagram help determine the probability of combined events?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Explore the use of probabilistic modeling in various scenarios.
  e.g. compare small vs. large trials to see how variability decreases; analyse how real-world factors (e.g. bias, environmental conditions, human error) impact results.
- Solve different probability problems.
  e.g. independent and dependent events, mutually exclusive and non-mutually exclusive events, sampling with and without replacement, and conditional probability.
- Compare when to use a tree diagram, a frequency table or a list to display data.
  e.g. illustrating the number of possible outcomes, analysing the effect the sample size has on any conclusions drawn.
- Using tree diagrams to model multi-step probability scenarios
  e.g. Illustrating number of possible outcomes and analysing the effect of sample size on results.

Learners will evaluate the probability of data relationships among various sets of data.

Using Data to Explore Variation

Rationale

Evaluating the probability of relationships between data sets allows learners to uncover connections, trends, and patterns within their analyses. Venn diagrams are useful in probability calculations because they provide a visual understanding of relationships among data sets. Learners will explore the notation used when creating Venn diagrams including universal set, subsets, complements, union, intersection and disjoint. They will use set notation to identify and describe the shaded region in Venn diagrams. By using Venn diagrams involving probability, learners will be able to draw and support conclusions.

Competencies

- Communication (COM)
- Critical Thinking (CT)

Indicators

- Use Venn diagrams to visually represent and summarize data, including intersections, unions, and complements. (COM, CT)
- Express data using Venn diagrams. (COM, CT)
- Apply set notation (i.e., $\cup$, $\cap$, $\subseteq$, $\subset$, and ') to describe relationships between data sets. (COM, CT)
- Justify conclusions based on probability and statistical reasoning using a Venn diagram. (COM, CT)

Concepts (and Guiding Questions)

Introduction to Venn Diagrams

- How is data that intersect, overlap, or remain distinct represented through a Venn diagram?
- What is the difference in the notation of union and intersection?

Representing Data in Venn Diagrams

- What strategies can be used to translate set notation into a visual representation?
- How can I decide where to place elements in a Venn diagram?
- How can I determine an unknown value in a diagram by using available information?

Set Notation for Venn Diagrams

- What strategies would I use to translate a visual representation into set notation?
- How can I use set notation to describe the different regions in a Venn diagram?

Making Conclusions

- What conclusions can be made using a Venn diagram?
- How does organizing data in a Venn diagram help us explain our reasoning and support conclusions in contextual probability problems?

Learners will evaluate the probability of data relationships among various sets of data. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Use Venn diagrams to help solve probability problems.
  e.g. draw and label Venn diagrams using notation and/or shading, solve a variety of contextual problems using Venn diagrams.
- Represent relationships between data sets using Venn Diagram.
  e.g. collect data to represent in Venn diagrams to show overlaps, unions, and complements.
- Explore set relationships using notation and visual models.
  e.g. using symbols ($\cup$, $\cap$, $\subseteq$, $\subset$, ') alongside Venn diagrams to describe how data sets relate; translate between diagrams, words, and symbolic notation.
- Support conclusions using Venn diagrams and probability reasoning.
  e.g. making and defending claims where two events are related, independent, or mutually exclusive based on diagram structure and calculated probabilities.

Learners will solve a variety of problems involving permutations and combinations

Using Data to Explore Variation

Rationale

Permutations and combinations are two distinct concepts in probability theory. Studying permutations and combinations is essential in mathematics because it helps develop problem-solving skills and a deeper understanding of counting principles. These concepts allow us to calculate possible arrangements and selections in various scenarios, making them useful in areas like probability, statistics, and real-world decision-making. Permutations and combinations have applications that span across various careers and activities including in cybersecurity and machine learning, experimentation, risk modeling, and workflow automation.

Competencies

- Communication (COM)
- Critical Thinking (CT)

Indicators

- Solve counting and probability problems using factorial notation. (COM, CT)
- Identify when to use permutations and combinations. (COM, CT)
- Use permutations and combinations in problem solving. (COM, CT)

Concepts (and Guiding Questions)

Factorial Notation

- How does factorial notation help us represent and calculate repeated multiplication in counting problems?
- What real-life or mathematical contexts (e.g., arranging items, scheduling tasks) can be used to demonstrate the relevance of factorial notation?
- How does understanding the pattern of factorials help explain why $4!=4(3!)$?

Permutations Versus Combinations

- How can I tell if the order of items matters in a counting problem?
- How can I count the number of ways to arrange a set of objects, and does it change when I only use a part of the set?
- How can I find the number of ways to choose items when the order doesn't matter?
- How do I decide whether to use a permutation or a combination when solving problems?
- How does the importance of order influence how different outcomes are counted?
- How can I choose the right strategy—permutation or combination—to solve a variety of counting problems?

Problem-Solving

- How can I determine whether a problem involves arrangements (i.e. permutations) or selections (i.e. combinations)?
- What strategies can help organize my thinking when solving complex counting problems (e.g., tree diagrams, lists, formulas)?
- How do restrictions (like two variables must be together) change the total number of arrangements?
- How can I check whether my answer makes sense in the context of the problem?

Learners will solve a variety of problems involving permutations and combinations (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Explore factorial notation in counting scenarios.
  e.g., investigate how factorials represent total arrangements (e.g. seating plans, rankings, scheduling); compare growth of possibilities as the number of items increases.
- Compare situations where order does vs does not matter.
  e.g. investigate situations where order is important (e.g. lineups, codes, tournament placements); compare total possible outcomes and consider how restrictions impact arrangements. Explore scenarios where order does not matter (e.g. forming groups, choosing committees, selecting items); identify how the number of choices affects possible combinations.

Learners will apply coding that simulates and analyses probability and variation

Using Data to Explore Variation

Rationale

Coding bridges the gap between theory and application, allowing learners to interact with probability in a dynamic, hands-on way. By creating simulations, learners can explore randomness, observe trends, and analyse statistical patterns in context. Understanding probability can be enhanced through coding and the use of simulations. Learners can further understand these concepts by being involved in a variety of simulations and coding experiences where they explore probability, permutations, and combinations. Coding allows learners to simulate different scenarios, run experiments, and analyse trends, making abstract probability concepts more tangible.

Competencies

- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Select coding sequences to find probabilities. (CT, TF)
- Use coding to explore variability in probability experiments. (CT, TF)
- Interpret results of probability experiments. (COM, CT, TF)

Concepts (and Guiding Questions)

Using Coding for Probabilities

- How can I use coding to describe how sampling a population with N objects works?
- How can the number of trials change the probabilities?

Using Simulations to explore Permutations and Combinations

- How can I determine the likelihood of various events?
- How can I use code in solving permutations and combination problems?
- What are the impacts of the changes I make in probability experiments?

Interpreting Results

- How can simulations help me in the analysis of results?
- What did I learn from using spreadsheets to find theoretical and experimental probabilities?

Learners will apply coding that simulates and analyses probability and variation (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Use spreadsheets to find experimental and theoretical probability of a data set.
  e.g. the probability of an event occurring.
- Explore variability in probability experiments.
  e.g., run repeated trials of simulated events to compare results across different sample sizes; modify section of code to explore variability, examine how outcomes vary between small and large numbers of trials; investigate how results stabilize over time.
- Explore simulation data using tables or graphs.
  e.g. identify patterns, trends, and outliers; explain reasoning using appropriate mathematical language; justify conclusions based on evidence.

Learners will evaluate relationships between data variables using statistics in context.

Modeling Data

Rationale

Evaluating relationships between data variables helps learners uncover patterns, make predictions, and draw meaningful conclusions. Being able to distinguish the difference between correlation and causation is often critical when making decisions about data. By understanding this difference, a critical evaluation of claims about the data can be made. Learners will explore bivariate statistics as they model data collected and analyse the results. Learners will use coding to explore lines of best fit, residuals, and correlation coefficients to determine how well a mathematical model can fit the data. They will collect, clean, and smooth data and then glean insights from their findings. Where possible, the use of non-fictitious data is encouraged as it provides additional context for learners while exploring data sets.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Identify correlation and causation in data sets. (CZ, COM, CT)
- Investigate bivariate data using scatterplots, correlation coefficients, and regression. (CZ, COM, CT, TF)
- Prove if correlations exist within given sets of data. (CZ, COM, CT, TF)
- Evaluate the authenticity of media claims that indicate correlation in bivariate data sets. (CZ, COM, CT)

Concepts (and Guiding Questions)

Correlation and Causation

- What story can be told about the data in the model?
- How could I determine if correlation and causation exist in a data set?
- How can residual analysis help determine the strength of a correlation?

Bivariate Data

- What is bivariate data?
- How can scatterplots be used to represent bivariate data?
- What is the purpose of the line of best fit?

Proving Correlations

- What is the difference between a positive and negative correlation?
- What would it mean if a scatterplot had a moderately positive correlation?

Evaluating Claims of Correlations

- Why does correlation not imply causation using contextual examples?
- How can I use a data set to find data that can be correlated?

Learners will evaluate relationships between data variables using statistics in context. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Distinguish between correlation and causation.
  e.g., analyse scenarios where correlated variables do not imply one causes the other; identify how variables may influence the data.
- Test whether correlations exist in data sets.
  e.g., collect or analyse paired data to determine if a relationship exists; compare multiple data sets to evaluate which show stronger or weaker correlations; consider the impact of sample size and variability.
- Explore media claims involving data.
  e.g., review graphs and headlines in media articles; question whether claims about relationships are supported by the data; identify misleading representations, omitted variables, or exaggerated conclusions.

Learners will evaluate the validity of inferences made from data sets.

Modeling Data

Rationale

Evaluating the validity of inferences from data sets is a critical skill in data analysis, statistics, and decision-making. Information involving correlations may be presented as causal in order to influence and lead people towards a specific conclusion. This can happen through advertisements, articles, and various forms of media and debates. Thinking critically about the validity of statements and claims is a skill that will help learners as they navigate an interconnected world. To further develop the ability to distinguish between correlation and causation, it is important to understand how different types of variables and contexts can affect interpretation of data. Learners will explore meaning when variables are causal, related, or have a connection. They will also investigate confounding, mediating, and lurking variables, direct vs. random correlations, and spurious correlations which are crucial for analysing relationships in data and avoiding misleading conclusions.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)

Indicators

- Question various inferences made from data (CZ, COM, CT)
- Analyse how evidence supports or weakens an inference. (CZ, COM, CT)
- Compare alternative explanations for relationships in data. (CZ, COM, CT)

Concepts (and Guiding Questions)

Inferences Made from Data

- How can I identify types of variables in data? (i.e. confounding, mediating, lurking, spurious)
- How can I identify what type of claim is being made from data sets?
- How could each type of variable influence any inferences that are made?
- How can I determine if a relationship appears to be present among variables?

Supporting an Inference

- How do I know if there is enough evidence to justify an inference?
- How reliable is the data?
- How does sample size, variability, and bias influence any inferences made?

Alternative Explanations

- What relationships appear within the data?
- How could relationships be explained, when considering data that may be confounding, mediating, lurking, or spurious?
- How could the data support more than one inference?
- How strongly does the data support each possible inference?

Learners will evaluate the validity of inferences made from data sets. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Explore media sources used to support claims.
  e.g., examine news articles, social media posts, or reports that use data to make claims; analyse how data is presented to shape audience perspectives or conclusions
- Explore whether data provided is sufficient and relevant to support a claim.
  e.g. identify contradictory evidence or gaps in the data; assess how sample size, variability, or bias strengthens or weakens the reliability of any inferences made.
- Explore explanations for relationships in data.
  e.g. generate and evaluate multiple explanations for an observed relationship; consider how confounding, mediating, or lurking variables could be used to make a claim; determine whether a relationship may be causal or spurious by comparing competing explanations; determine which explanations are most accurately supported by the data.

Learners will implement a process for accurate and reliable data collection and analysis.

Modeling Data

Rationale

Accurate and reliable data collection is the foundation of meaningful analysis. Data can be collected using a variety of methods and sources. Defining clear objectives and determining what data is needed is the first step in this process. Learners will consider how to avoid collecting irrelevant information as they work through their process. They will learn how to choose appropriate collection methods such as surveys, experiments, or observations, and ensure data is collected consistently, avoiding biased sampling, and checking for errors. It is important to use reliable tools and technology to gather accurate data. In this outcome, learners can collect their own data for analysis using a variety of methods including sensors and other technology.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Determine a process for collecting data through a variety of methods. (CZ, COM, CT, TF)
- Use coding to clean and smooth data. (CZ, COM, CT, TF)
- Evaluate the quality and integrity of cleaned data to ensure accuracy and reliability. (CZ, COM, CT, TF)
- Justify decisions made from data sets. (CZ, COM, CT)

Concepts (and Guiding Questions)

Process to Collect Data

- How does the data I need influence the way(s) I choose to collect it?
- What decisions can I make to avoid introducing potential bias into my results?

Using Coding to Clean Data

- What are the differences between cleaning and smoothing data?
- What process can I use to clean and smooth the data that I have?

Evaluating Data

- How does cleaned data compare to raw data?
- How can I determine the integrity of cleaned data sets?

Justifying Decisions

- How does the story of the data along with my own analyses lead me to make an informed conclusion?
- What conclusions can I make from my data set and how can I support this with evidence?

Learners will implement a process for accurate and reliable data collection and analysis. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Evaluate the quality and integrity of data.
  e.g., compare data sets before and after cleaning to assess impact; identify sources of bias or error; evaluate whether the data set is complete and representative; determine how data limitations affect reliability

Learners will determine ways to make decisions using data distributions.

Hypothesis Testing and Distributions

Rationale

Data distributions provide critical insights that help make informed decisions across various fields. Within measures of central tendency, metrics like mean, median, and mode help summarize the general behavior of a data set, guiding strategic decisions. Analysing variability like standard deviation and variance indicate data spread, helping assess risks and uncertainties. Using distributions, learners can estimate the likelihood of outcomes in different scenarios. In quality control or anomaly detection, understanding distribution tails helps define acceptable performance levels. Regression and statistical forecasting rely on data distributions to make accurate projections in business, healthcare, and science. Learners will expand their knowledge by using mean and standard deviation to explore normal distribution and the empirical rule. The range rule will show learners how to estimate the standard deviation. As well, a comparison of normal distribution curves to analyse the spread of data will be explored. Finally, learners will calculate the z-score, confidence interval, and margin of error and determine the confidence level for a sample.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)

Indicators

- Investigate the use of normal distribution, z-scores, and confidence intervals. (CZ, COM, CT)
- Compare normal distribution curves of various sets of data. (CZ, COM, CT)
- Solve problems involving z-score, confidence intervals, and margin of error. (CZ, COM, CT)
- Evaluate the validity of data claims using normal distribution, z-scores, and confidence intervals. (CZ, COM, CT)

Concepts (and Guiding Questions)

Normal Distribution, Z-scores, Confidence Intervals

- How does mean and standard deviation help construct a normal distribution curve?
- How do I recognize and assess patterns in data?
- When is it appropriate to use the z-score formula?
- What can I learn from the confidence level and margins of error?

Comparing Normal Distribution

- How does the spread of the data affect the shape of a normal distribution curve?
- What effect does changing the mean have on the normal distribution curve?
- What effect does changing the standard deviation have on the normal distribution curve?

Problem-Solving

- When is it necessary to use the z-score formula?
- How does sample size affect the formula used to calculate a confidence interval?
- How can I determine the margin of error?

Learners will determine ways to make decisions using data distributions. (cont.)

Validity of Claims

- How does increasing the confidence level affect the margin of error?
- How does increasing the sample size affect the margin of error?
- Where can I find an example of a claim that accurately uses margin of error?
- How can I use understanding of normal distribution, z-scores, and confidence intervals to determine if a claim is valid?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Explore the use of probability distributions.
  e.g., determine if data follows a normal distribution; use z-scores to interpret how far values fall from the mean; examine how confidence intervals provide a range of likely values beyond a sample.
- Solve problems within distributions.
  e.g. normal distribution, z-score calculations, confidence interval, and margin of error.
- Explore data and claims from media sources.
  e.g. assess whether a claim falls within a reasonable range based on a confidence interval; use z-scores to determine if values are typical or unusual; determine if conclusions are reasonable based on the variability of data and the margin of error.

Learners will investigate how hypothesis testing is used to make informed statistical decisions

Hypothesis Testing and Distributions

Rationale

Hypothesis testing is a fundamental statistical method used to evaluate claims and make informed decisions based on data. It helps determine whether results are statistically significant or likely due to random chance. Students will explore the four key steps of hypothesis testing: Stating the null and alternative hypotheses, selecting the appropriate parameter test, calculating the test statistic, and making a decision based on a comparison with a critical value or p-value. They will learn to distinguish between one-tailed and two-tailed tests and understand when each is appropriate based on the nature of the hypothesis. Learners will also apply z-tests and t-tests, using the Central Limit Theorem and characteristics of the sample (e.g., known vs. unknown standard deviation, sample size) to determine which test is appropriate. In addition, learners will investigate Chi-squared tests, including the Goodness of Fit and the Test for Independence, to analyze categorical data and examine relationships between variables. Learners will explore Type I and Type II errors, understanding the implications of incorrect conclusions in hypothesis testing.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Investigate the four steps of hypothesis testing. (CZ, COM, CT)
- Compare the use of one-tailed and two-tailed hypothesis tests. (CZ, COM, CT)
- Apply z-tests and t-tests based on the Central Limit Theorem and sample characteristics. (CZ, COM, CT, TF)
- Investigate Chi-squared as a statistical hypothesis test for categorical data analysis. (CZ, COM, CT, TF)

Concepts (and Guiding Questions)

Steps in Hypothesis Testing

- Why are the 4 steps of hypothesis testing important in making statistical decisions?
- How do I write a null and alternative hypothesis for a given situation?
- What does the test statistic tell us about our data and the original claim?
- How do I decide whether to reject or fail to reject the null hypothesis?
- What is meant by a Type I error and a Type II error in hypothesis testing? Can I provide a theoretical example?

One-Tailed and Two-Tailed Tests

- What information do I need to decide between a one-tailed and a two-tailed test?
- How do I know if the null hypothesis should be rejected in a one-tailed test? in a two-tailed test?
- How does the level of significance affect whether the null hypothesis should be rejected?
- How do rejection regions differ between one-tailed and two-tailed tests?
- In which contextual situations would each type of test be appropriate?

z-tests and t-tests

- How does the Central Limit Theorem assist in choosing a z-test or a t-test?
- How does sample size and knowing the population standard deviation influence our choice of test?
- How do I interpret the results of a z-test or t-test in the context of a real-world question?
- How can I use the confidence interval formula to check decisions made in a hypothesis test?

Chi-Squared

- What types of data and questions are best suited for Chi-squared tests?
- How do the Chi-squared Goodness of Fit and Test for Independence differ?
- What are degrees of freedom, and why do they matter in a Chi-squared test?
- How can I tell if the difference between observed and expected data is statistically significant?

Learners will analyse the relationship between p-values and significance levels (α), in hypothesis testing.

Hypothesis Testing and Distributions

Rationale

Understanding the relationship between the p-value and the significance level (α) is essential in hypothesis testing. The p-value represents the strength of evidence against the null hypothesis based on sample data, while the significance level is a predetermined threshold used to decide whether to reject the null hypothesis.

Although related, these two concepts serve different roles: the p-value is calculated from the data, and α is set by the researcher in advance. Decisions are made by comparing the two—if the p-value is less than α, the null hypothesis is rejected; if it is greater than or equal to α, it is not rejected. Learners will engage in problem-solving and simulations to explore this relationship and understand how it guides statistical decision-making. This conceptual understanding will allow them to interpret results more confidently and apply hypothesis testing in meaningful, real-world contexts.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Technological Fluency (TF)

Indicators

- Investigate how coding is used to determine significance levels. (CZ, COM, CT, TF)
- Determine the significance levels necessary to accept or reject a null hypothesis. (CZ, COM, CT)
- Justify conclusions within hypothesis testing. (CZ, COM, CT, TF)

Concepts (and Guiding Questions)

Coding and Significance Levels

- How can I use coding or technology (e.g., Python, R, calculators) to calculate p-values and compare them to significance levels?
- What are the advantages of using code or simulations to test hypotheses?
- How can I model different significance levels in a simulation to see how they affect decisions?

Accepting or Rejecting the Null Hypothesis

- How does changing the significance level affect whether I reject or fail to reject the null hypothesis?
- What is a common threshold for α, and why might I choose different levels in different situations (e.g., 0.01 vs. 0.05)?
- How do I use the comparison between p-value and α to decide what actions to take?

Justifying Conclusions

- How can I make decisions to accept or reject the null hypothesis using the p-value or significance level?
- How can I explain our decision to reject or not reject the null hypothesis using statistical evidence?
- How can I communicate our findings clearly to someone who doesn't have a background in statistics?
- What are the risks of making a wrong decision in hypothesis testing (e.g., Type I or Type II errors), and how do they affect our conclusions?

Learners will analyse the relationship between p-values and significance levels (α), in hypothesis testing. (cont.)

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Use spreadsheets or statistical software to calculate p-values from sample data.
  e.g. simulate sampling distributions to observe how p-values change; compare outputs from code to manual calculations.
- Use data and simulations to explore p-values and significance level.
  e.g. compare p-values to different significance levels (e.g., 0.05, 0.01) to make decisions; analyse how changing the significance level affects conclusions; explore the relationship between p-values and decision thresholds.

Learners will create a communication form to express statistical data and relating findings.

Hypothesis Testing and Distributions

Rationale

This outcome focuses on the application, analysis, and communication of statistical data through the creation of a purposeful communication form. Throughout the course, learners have engaged with a variety of statistical concepts, coding activities, data collection techniques, and simulations. Here, they will design and carry out an investigation that involves analysing a data set which is either collected or sourced. Learners will determine how to communicate their findings in a format appropriate to a chosen audience, which could include peers, school community, or broader communities. Emphasis is placed on clarity, accessibility, and audience awareness in presenting statistical insights. This project provides students with an opportunity to apply their learning in an authentic and meaningful way, demonstrating both analytical and communication skills.

Competencies

- Citizenship (CZ)
- Communication (COM)
- Critical Thinking (CT)
- Personal and Career Development (PCD)
- Technological Fluency (TF)

Indicators

- Use new or collected data for analysis and communication. (CZ, COM, CT, TF)
- Investigate the effectiveness of different methods for communicating statistical information. (CZ, COM, CT)
- Formulate options to share information to specific audiences and purposes. (CZ, COM, CT)
- Develop a communication form for a specific audience and purpose. (CZ, COM, CT, PCD)

Concepts (and Guiding Questions)

Using Data

- What types of data are available or relevant to the topic I want to explore?
- How can I collect or find reliable data that helps answer a meaningful question?
- What steps do I need to follow to organize and analyse data effectively?
- What patterns, trends, or relationships can I identify in my data?

Effective Communication

- What are some common ways to communicate data (e.g., graphs, reports, infographics, presentations)?
- What types of communication forms make sense to use for my data? (i.e. video, graph, social media campaign, infographic, etc.)
- Which formats or tools help make data easier to understand for different audiences?
- How do visuals (e.g., charts, tables, maps) support the story I want to tell with my data?
- How can I evaluate whether a method of communication clearly and accurately conveys statistical information?

Learners will create a communication form to express statistical data and relating findings. (cont.)

Sharing to an Audience

- Who is my intended audience, and what do they need or want to know?
- What is the purpose of my communication—informing, persuading, raising awareness?
- How can technology assist us in the communication of statistical ideas?
- How can I tailor my language, visuals, and format to suit my audience?
- What are some possible ways I could present my findings, and what are the pros and cons of each?

Communicating Results

- How can I organize my findings in a way that makes sense to my audience?
- What communication tools or platforms will best support my message?
- How can I make sure my communication is clear, accurate, and engaging?
- How will I check that my final product effectively meets the needs of my audience and reflects my analysis?

Sample Topics for Investigation

As part of exploring this outcome, learners can:

- Outline a meaningful statistical question and data investigation.
  e.g., develop a clear, measurable question; determine what data is needed to answer the question; select appropriate methods for data collection; use coding and analysis tools to organize, explore, and interpret the data

Skills

Apply: Carry out, use, or complete a procedure/technique.

Analyse: Consider the nature or structure of something by deconstructing it into its component parts to understand or explain it.

Classify: Identify attributes and select criteria for groupings and subgroupings; sort based on selected criteria and consider the grouping(s); incorporate a new item in a group; offering a rationale for the choice based on relationships.

Compare: Make observations; identify similarities and differences; identify relationships and offer an interpretation; communicate the findings.

Determine: Decide or conclude through reasoning or observation.

Develop: Strengthen skills, knowledge, or understanding through building on acquired or existing abilities

Evaluate: Offer a critical review of something by considering the strengths and limitations.

Express: Convey thought, ideas, or feelings through a selected communication form or medium.

Formulate: Identify a topic of interest; brainstorm ideas; choose, prioritize, and refine ideas; evaluate choices.

Implement: Carry out a process, strategy, or plan.

Identify: Distinguish specific elements based on their unique features.

Interpret: Use knowledge and understanding to draw conclusions from given information Investigate: carry out an inquiry to examine and better understand a concept.

Investigate: Carry out an inquiry to examine and better understand a concept.

Justify: Support a conclusion based on valid evidence. Question: Consider an argument, perspective or issue in a way that uncovers assumptions, biases, or underlying relationships.

Prove: Use a sequence of logical steps to establish existence, truth, or validity of something by evidence or argument.

Select: Locate several relevant and dependable details to support an answer.

Solve: Find a solution, explanation, or answer.

Use: Put knowledge or skills into action or service.