Unit 1 Investigative Report: Solving Your Own Case

Mission Objective

To demonstrate mastery of Unit 1 concepts by conducting a complete data investigation using the Data Cycle, analyzing data with CODAP, and presenting clear, evidence-based conclusions.

Mission Background

Congratulations, Data Detectives!

You have successfully completed your initial training! You have mastered the art of identifying data types, asking statistical questions, and analyzing numerical data's shape, center, and spread using both manual techniques and our powerful digital toolkit, CODAP.

Now, it's time for your Final Mission as a lead Data Detective! You will choose a "case" (a dataset) to investigate and formulate a relevant statistical question. You will then analyze the data using CODAP and present your findings in a comprehensive "Case Report."

This is your chance to show off all your detective skills!

The Case Files (Data Options)

You will choose ONE of the following datasets for your investigation. All of these are pre-loaded in CODAP.

1. Roller Coasters: Contains data about roller coasters (ex. Height, Speed, Type, Material, etc.).
2. Mammals: Contains data about various animals (ex. Sleep, Life Span, Diet, Type, etc.).
3. Granola Bars: Contains nutritional data about different brands of granola bars (ex. Calories, Carbs, Fiber, Sugar, etc.).

This work © 2026 by
ThinkData Ed and Data Science Education Center

Mission Steps (The Data Cycle in Action)

Your Case Report will follow the phases of the Data Cycle. You will need to provide evidence for each step.

Phase I: Pose a Statistical Question	☐ Choose one statistical question that can be answered using your chosen dataset. ☐ Your question must involve at least one numerical variable. ☐ Your question can involve a categorical variable if you want to explore relationships between groups.

Phase II: Consider the Data	☐ Identify the dataset you are using.
	☐ Briefly describe the source of the data. Is it primary data you collected, or secondary data collected by someone else?
	☐ Explain what information is provided in the dataset.
	  ☐ How many observations are there?
	  ☐ How many variables were reported?

Phase III: Analyze the Data (Using CODAP!)	☐ Use CODAP to create at least TWO different graphs that help you answer your statistical question. These could be:
	  ☐ Dot plots
	  ☐ Histograms
	  ☐ Boxplots
	☐ For each numerical variable relevant to your question, you must:
	  ☐ Describe its shape (symmetric, skewed right, skewed left, bimodal).
	  ☐ Determine the most appropriate measure of center (mean or median) and justify your choice based on the shape.
	  ☐ Determine the most appropriate measure of spread (IQR or MAD) and justify your choice based on the shape.
	  ☐ Identify any unusual or extreme observations.
	☐ Take clear screenshots of your CODAP graphs to include in your report.

Phase IV: Interpret the Results	☐ Write a clear conclusion that directly answers your statistical question. ☐ Use evidence from your analysis (your graphs, measures of center, measures of spread, and observations about shape/relationships) to support your answer. ☐ Explain what story the data tells you about your chosen "case."

This work © 2026 by
ThinkData Ed and Data Science Education Center

Mission Case Report (Submission Requirements)

You will submit a digital Case Report (ex. a Google Doc, Google Slides, Powerpoint, etc.)

Your report must include the following sections:

1. Detective Name: (Your Name)
2. Case File Title: (A creative and/or relevant title for your investigation)
3. My Statistical Question: (from Phase I)
4. My Data Source: (from Phase II)
5. My Analysis & Evidence: (from Phase III)
    a. Graph 1 & Observations:
        i. Screenshot of your first CODAP graph.
        ii. Description of the distribution's shape, best center (with justification), best spread (with justification), and any unusual observations.
    b. Graph 2 & Observations:
        i. Screenshot of your second CODAP graph.
        ii. Description of the distribution's shape, best center (with justification), best spread (with justification), and any unusual observations.
    c. ***OPTIONAL***: Additional graphs/observations if needed to answer your statistical question.
6. My Interpretation & Conclusion: (from Phase IV)
    a. Your answer to the statistical question.
    b. How your analysis (graphs, measures of center/spread, shape) supports your answer.

This work © 2026 by
ThinkData Ed and Data Science Education Center

Mission Rubric: The Data Detective's Performance Review

Criteria	Exceeds Expectations (4 points)	Meets Expectations (3 points)	Approaching Expectations (2 points)	Needs Improvement (1 point)
Data Cycle: Pose Questions (Statistical Question)	The question is clear, concise, and requires analysis of a numerical variable. It is insightful and perfectly suited to the chosen dataset.	The question is clear and requires analysis of a numerical variable. It is a good fit for the chosen dataset.	The question is somewhat unclear or does not clearly require a statistical investigation of a numerical variable.	A statistical question is not provided, or it is not appropriate for the dataset.
Data Cycle: Consider Data	The data source is clearly identified with detailed description of observations, variables, and whether data is primary or secondary.	The data source is identified with adequate description of observations and variables.	The data source is identified but the description is incomplete or missing key details.	The data source is not identified, or no description is provided.
Data Cycle: Analyze Data	Creates 2+ highly appropriate graphs in CODAP that effectively visualize the data and provide compelling evidence to answer the question. Describing Shape, Center, & Spread: Accurately describes the shape of the distribution and correctly selects the most appropriate measure of center (mean/median), and selects the most appropriate measure of spread (IQR/MAD). Provides a clear, correct justification for these values based on the distribution's shape.	Creates 2 graphs in CODAP that are appropriate for the data and provide evidence to answer the question. Describing Shape, Center, & Spread: Accurately describes the shape of the distribution and correctly selects the appropriate measure of center, with a clear justification. Selects the appropriate measure of spread (IQR/MAD) and provides a clear justification.	Creates 1-2 graphs in CODAP that are somewhat appropriate, but do not fully support the investigation or are incorrectly labeled. Describing Shape, Center, & Spread: Describes the shape of the distribution with some errors, or selects the correct measure of center but provides a weak justification. Selects a measure of spread but the justification is weak or incorrect.	Graphs are not created, or they are not relevant to the question. Describing Shape, Center, & Spread: The shape of the distribution is not described, or the measure of center is incorrect. The measure of spread is not selected, or it is incorrect.
Data Cycle: Interpret Data	The conclusion directly and thoroughly answers the statistical question. It is supported with specific, compelling evidence from the analysis.	The conclusion directly answers the statistical question and is supported with evidence from the analysis.	The conclusion is somewhat vague or does not fully answer the question, and the supporting evidence is weak.	A conclusion is not provided, or it does not relate to the original question or the analysis.
Report Formatting & Clarity	The Case Report is exceptionally well-organized, easy to read, and includes all required sections and screenshots. Language is precise and professional.	The Case Report is well-organized and includes all required sections and screenshots. The writing is clear.	The Case Report is somewhat disorganized or is missing a key section. The writing is occasionally unclear.	The Case Report is incomplete, very disorganized, or difficult to read.
			TOTAL SCORE	_____ / 20

This work © 2026 by
ThinkData Ed and Data Science Education Center