The Thrills of Roller Coasters: Using Data to Inspire Investigations

Author(s)	Hollylynne Lee, Gemma Mojica, Heather Barker, Christina Azmy, and Taylor Harrison
Author Affiliation and Location	HI-RiSE: A Hub for Innovation and Research in Statistics Education Friday Institute for Educational Innovation North Carolina State University, Raleigh, NC hirise.fi.ncsu.edu This lesson plan and associated research was developed as part of the ESTEEM project, funded by the National Science Foundation (DUE-1625713).
Contact Information	Hollylynne Lee hollylynne@ncsu.edu Gemma Mojica gmmojica@ncsu.edu
Abstract to Lesson Plan	This lesson builds on research of how real-world data investigations can be used to engage students in mathematics classes to learn key concepts in data and statistics. The lesson uses a free online data visualization tool and a data set of 157 roller coasters from U.S. amusement parks from the past 100 years. Attributes like top speed, material used, seating arrangement, length of track, and height quickly engage students in using their real-world and scientific understandings to make sense of trends and patterns in the data.
Introduction	Students use real data to analyze the maximum height of 31 older roller coasters and then make predictions about reasonable height expectations for all older coasters. Students become familiar with the structure of a data set and the relationship between a case, a collection of cases, and graphical representations of certain attributes. Students understand the key measures used to describe a univariate distribution in the context of real data and distribution. Students then explore a larger data set with 157 cases and 15 attributes. Students pose their own questions after exploring the data set and use the investigative cycle to investigate. Students compare two distributions to make claims about whether wooden or steel coasters are different. While not all students have ridden roller coasters before, it is likely they have seen them in the media. The Internet is full of videos made by roller coaster enthusiasts. Recommended videos to help students understand the specific context are included. Many students love the idea of the legal thrill they can get from roller coasters and get excited seeing them in action. The data in this lesson was curated by the ESTEEM project team using the following websites: rcdb.com, wikipedia.com, and ultimaterollercoaster.com
Learning Objectives	Students will be able to: • Identify patterns and causal relationships they see in data • Pose questions about the relationships between different characteristics of roller coasters • Analyze the data to answer their questions using statistical software • Obtain, evaluate, and communicate information about their investigation to their classmates

Appropriate Grade Levels	Content: Engineering, Math, Statistics, and Physical Science Grades 6-8, 9-12
Group Size/# of Students	2-3 students working together on one laptop
Approximate Time of Lesson	Two 60-minute class periods
Resources Needed for Students	A laptop or computer for every 2-3 students to share Internet access (to use data and online software linked below) Handout for completion (Roller Coaster Investigation) Handout about definitions of variables (Roller Coaster Data: Attributes and Definitions)
Resources Needed for Educators	Internet Access (to play video, and use data and online software linked below) Computer and projector for display **Background Preparation for Teacher** 1. This lesson addresses the statistical investigation cycle, which includes posing questions, collecting data, analyzing data, and making inferences. To read more about the statistical investigation cycle and statistical habits of mind, see: https://s3.amazonaws.com/fi-courses/tsdi/unit_2/Essentials/Habitsofmind.pdf 2. Another brief reading describes statistical habits of mind and how they can be used in the four phases of a statistical investigation. Think about how you might help your students develop these habits when teaching with this lesson. See: https://s3.amazonaws.com/fi-courses/tsdi/unit_2/Essentials/Habitsofmind.pdf 3. This lesson also introduces students to the Common Online Data Analysis Platform (CODAP), a free online software tool for creating data visualizations. Learn how to use CODAP with this brief introductory video: https://youtu.be/aD5tLWld98w 4. Launching data-rich tasks involves hooking students into the context and ensuring they activate their real-world understandings. You can watch a video of a teacher introducing this lesson here: https://youtu.be/aXolxokHRxU 5. You can also watch a video of how a teacher chooses students to share their roller coaster investigation and orchestrates a discussion using student work: https://youtu.be/ETNF_542DvU
Apps/Websites Needed	CODAP, Common Online Data Analysis Platform, https://codap.concord.org https://tinyurl.com/31USCoasters https://tinyurl.com/157UScoasters

The Thrills of Roller Coasters: Using Data to Inspire Investigations

Detailed Lesson Activity

Lesson Activity	**Day 1: 60 minutes total** **5 minutes** Have laptops and handouts already on desks. Students sit in pairs with a laptop. They need to log on and access the Internet. As students are getting settled, they can complete questions 1-3 on the handout (experience with coasters and what makes them thrilling/scary).  Describe "Four Phases of a Statistical Investigation" and how this is very similar to the scientific method. **5 minutes** To engage students in the content, play the video of a first person view of a rider on the JackRabbit wooden roller coaster at KennyWood Park, built in 1921 and still in operation. Have a short classroom discussion of what students may have noticed about the experience of riding the coaster in the video. https://www.youtube.com/watch?v=tWh9ofIm-B8 **10 minutes** Have students open the following CODAP document: tinyurl.com/31uscoasters Explain the structure of the data table with each row constituting a "case." Each case contains information about one roller coaster, and there are 31 older coasters in this data set from the U.S. Take one minute to look at data in the table (show how to scroll to see variable names) and the handout with variable definitions. Connect these variables with what students may have listed as things that make coasters *thrilling or scary.* Which variables might relate to an exciting roller coaster ride? How about a scary ride? Pose the following question to focus the work: ***How tall do older roller coasters tend to be?*** The data has already been collected, but to answer our question about typical height of older coasters, what variable can we use? (Max Height) **15 minutes** Click on the Graph icon on the toolbar, and a window with 31 randomly placed dots should appear. There is no structure to this graph yet. We only see 31 cases. You can click on a dot and the associated case will highlight in the table. Drag Max Height to the x-axis. This creates a dotplot to examine how the max heights of these 31 coasters are distributed. Have students click on a dot of interest to them in the graph. What does each dot represent? Why is it in the position it is on the number line?

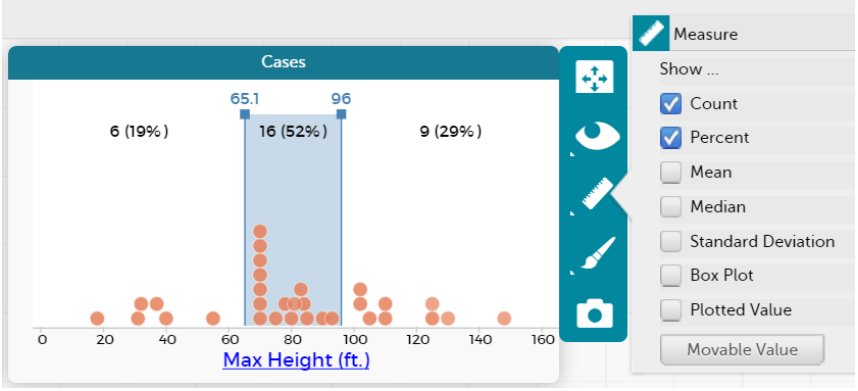
Do all the coasters have the same height or similar heights? (No) This is why we need statistics. When data varies, we need ways to describe the data to look for tendencies and patterns. There is not a formula or rule that says all coasters will be a certain height. Thus, in order to answer our question, we need to have ways of describing how the data is distributed.

Let's start by looking at some interesting cases. Minimums and maximums often draw our attention. Have students complete question 4 and 5 on the handout to find out more about the coasters with minimum and maximum height. Have groups share findings.

10 minutes

Let's return to our focus on typical roller coasters. Are the extreme coasters typical in height? No. So we need a way of describing "typical"!

Show students how to add two moveable lines (bottom of ruler menu) to view a shaded region. Have them drag the lines so that they have shaded a clump of data that they think represents typical height of coasters. Then demonstrate how to add count and percent and have students write down information on the handout for question 6.



This is an informal way to eyeball what *feels* typical when you see a distribution.

Show students how to add the mean (and median) under the ruler. These are two ways that we can measure a single value to describe typical height.

10 minutes

Give students an opportunity to create graphs to look at other attributes that may interest them (about 3-4 minutes). This should be playful as they examine any trends in variables that interest them. (Many will be attracted to top speed, max drop, and duration of ride.)

Ask students to report out what they explored and found out about these older roller coasters.

5 minutes

End the class period by telling students that tomorrow they will examine a larger set of data that includes more recently built roller coasters. Ask students what they think they will discover about roller coasters if they include more coasters, and coasters that have been built more recently. Do they anticipate any trends or patterns? Why?

Day 2: 60 minutes total

20 minutes
Ask students to remind the class what they observed about max height with older coasters and any trends they anticipated finding with a larger data set.

Let's look at more recent roller coasters and a data set with many more cases and variables!
Open http://tinyurl.com/157UScoasters

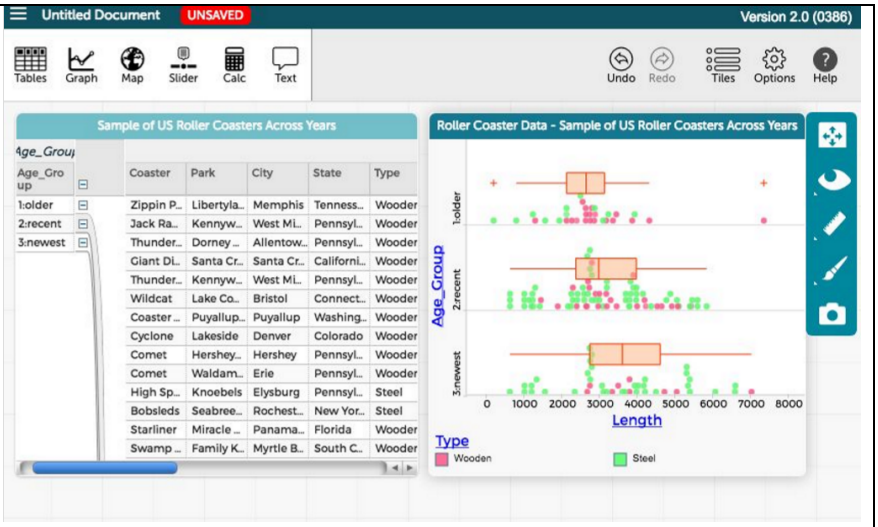
Look at the handout of all the attributes. Be sure students understand the different attributes.

Students are typically curious about roller coasters they may have heard of or ridden. At the end of the data table, the latitude and longitude values are included for each roller coaster location. Opening the map (globe icon on toolbar) will show the location of each of the roller coasters. Some locations (e.g., an amusement park) have more than one coaster in the data set. Use the marquee tool (dashed square) to select a location and see which coasters in the data table are at that location.

Give students time to use the tools in CODAP to examine attributes they are interested in. Be sure they learn how to place variable names on the y-axis and to also change the color of the cases in a graph by dragging a variable name to the center of a graph. These techniques help students' visualize relationships between variables.

Encourage them to pose a question about roller coasters that has to do with what they're interested in. The questions should focus on the relationship between at least two variables. For example:

- Does the height of the coaster affect the top speed?
- Does the design type determine how many loops are in a coaster?
- Is there a difference in heights between coasters built with wood or steel?
- Is there a difference in maximum heights between older and more recent coasters?
- Has track length changed much over the years? (See sample graph below used to explore this question. Though the median length of tracks has increased, the newest coasters also have much more variability in length than older ones.)

	 Be sure students record their question on a display or board. They need to make sure the questions can be answered using the variables in the data set. Each group should share their question and which variables they will need to use to answer their question. **15 minutes** Instruct students that they will take time to continue to explore the data set. They can try to answer their question. Tell them that they will have to use the data set to answer the question. They will have to use a graph and other statistical measures (e.g., mean, median) to answer their question. They need to create something they can share in front of the class about their investigation. (Note: To make this a more involved project, students could be asked to write a report or create a slideshow explaining their findings.) **20 minutes** Have groups share what they investigated. If possible let students project the graphs they make in front of the class. **5 minutes** Reflection/Assessment ● What did you learn about how data and graphs can be used to answer questions? ● What did you learn about roller coasters?
Final Product/Assessment	The final product is an informal student presentation. You may also want to collect the handouts to see what students recorded. Since this lesson is meant to introduce students to the statistical investigation process the assessment should be low stakes.

Appendices: The handouts to use during class are included on the next pages.

ROLLER COASTER INVESTIGATION Name _____

1. My experience with roller coasters is…..

2. What physical aspects of roller coasters make them thrilling or scary?

3. What are some physical aspects that engineers who build coasters may have to take into consideration when designing coasters?

****

We have data about 31 older roller coasters from amusement parks in the U.S. Let's use this data to investigate some properties of coasters.
http://tinyurl.com/31UScoasters

Posing an Investigative Question: How tall do older roller coasters tend to be?

Let's graph the Max Height (drag label from table to x-axis in graph) and find out how we can describe the height of roller coasters. Each dot in the graph represents a case of a roller coaster and corresponds to one row in the table. Click on a dot and notice the table. Explore the coasters by clicking on dots or selecting rows in the table.

Let's explore the data to find out about some extreme coasters.

4. Name of coaster with the smallest maximum height: _____

Smallest maximum height: _____

What else can you find in the data table about the coaster with the smallest height?

5. Name of coaster with the largest maximum height: _____

Largest maximum height: _____

What else can you find in the data table about the coaster with the largest maximum height?

6. Add two moveable lines to form a shaded region and move it so it covers a clump of the coasters you consider to be typical in height. (Use the ruler menu beside the graph window.)

What range of data does your shaded area cover? _____

How many coasters are in the shaded region? _____

What percent of the data are in your shaded region? _____

7. Find the following measures about the height of these older coasters:

Range: _____

Median: _____ Mean: _____

8. Based on this sample of data, what do you think is the typical height of *all* older roller coasters in the U.S.?

********Let's Look at More Coasters!*********

Open a new data file with a larger sample of U.S. roller coasters, both older and newer, at http://tinyurl.com/157UScoasters

9. Explore the data with variables that interest you. Use different graphs to explore any patterns and trends for these coasters. Write down some of the things you noticed. Are any of the things you noticed something that you think engineers have to think about when designing coasters?

10. Pose a question about roller coasters that has to do with the variables you are interested in. The question should focus on the relationship between at least two variables. Record your question here:

11. Continue to explore the data set and try to answer the question you posed above, using the data. Use at least one graph and other statistical measures (e.g., mean, median) to answer your question. Be ready to share what you did with the class. You may have to recreate your graph for the entire class.

Extension:

12. Are there any differences between wooden and steel roller coasters? What does the data tell us? Write down what you explored and what you noticed.

Roller Coaster Data: Attributes and Definitions

#	Attribute Name	Description	Units
1	Coaster	Name of the roller coaster	
2	Park	Name of the park where the roller coaster is located	
3	City	City where the roller coaster is located	
4	State	State where the roller coaster is located	
5	Type	Material of track (steel or wooden)	
6	Design	How a passenger is positioned in the roller coaster	
	Design Types:	Bobsled – designed like a bobsled run – without a fixed track. The train travels freely through a trough	
		Flying – a roller coaster ridden while parallel with the track	
		Inverted – a roller coaster that uses trains traveling beneath, rather than on top of, the track. Unlike a suspended roller coaster, an inverted roller coaster's trains are rigidly attached to the track.	
		Pipeline – a coaster where riders are positioned between the rails instead of above or below	
		Sit Down – a traditional roller coaster ridden while sitting down	
		Stand Up – a coaster ridden while standing up instead of sitting down	
		Suspended – a roller coaster using trains that travel beneath the track and pivot on a swinging arm from side to side, exaggerating the track's banks and turns	
		Wing – a coaster where pairs of riders sit on either side of a roller coaster track in which nothing is above or below the riders	
7	Opened	Year when roller coaster opened	
8	Top Speed	Maximum speed of roller coaster	mph
9	Max Height	Highest point of roller coaster	ft
10	Drop	Length of largest gap between high and low points of roller coaster	ft
11	Length	Length of roller coaster track	ft
12	Duration	Time length of roller coaster ride	seconds
13	Inversions?	Whether or not roller coaster flips passengers at any point (Yes or No)	
14	# of Inversions	Number of times roller coaster flips passengers	
15	Age Group:	1: Older (Built between 1900-1979)	
		2: Recent (1980-1999)	
		3: Newest (2000-current)	
Data Sources:		RCBD.com Wikipedia.com UltimateRollerCoaster.com	