Written by: Kimani Roach

Unbound Education

Lesson plan: Data Detectives, Credit Decisions Through Data Investigation

Inspired by: Angela Davis (Ms.)

Lesson Overview

This lesson engages students to analyze a secondary dataset using a combination of CODAP and Desmos to explore and visualize relationships between two quantitative variables. Students investigate the credit information of individuals applying for credit cards at a financial institution to determine whether meaningful relationships exist within the data. The students begin by examining a scatter plot displaying applicants' credit limit and credit card balance. Through a Notice and Wonder activity, they identify patterns, trends, clusters, possible outliers, and the overall direction of the association represented in the scatter plot. This discussion serves as the foundation for generating additional questions about the dataset and encourages students to distinguish between statistical and non-statistical questions suitable for investigation. Using both CODAP and Desmos, students will investigate their questions by creating and analyzing scatter plots of different variable pairs. As they explore the data, students will describe the strength and direction of associations, interpret the correlation coefficient (r), and use technology to construct and evaluate linear models. They will interpret the meaning of the slope and y-intercept in the context of the data and use the models to make predictions and answer real-world questions about consumer credit. Throughout the lesson, students engage in authentic data analysis while developing statistical reasoning, mathematical modeling skills, and data literacy. They learn how mathematics can be used as a powerful tool to investigate real-world situations, identify relationships between variables, and communicate evidence-based conclusions. The number of class periods required to complete this lesson will depend on the level of student agency provided during the question-generation process and the depth of students' data exploration and discussion. **This lesson is designed to be completed within one 90-minute session, however teachers are allowed to make different decisions based on their respective context.**

NC 8th grade Standard Alignment

NC.8.SP.1 Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Investigate and describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

NC.8.SP.2 Model the relationship between bivariate quantitative data to:

• Informally fit a straight line for a scatter plot that suggests a linear association.

• Informally assess the model fit by judging the closeness of the data points to the line.

Written by: Kimani Roach

NC.8.SP.3 Use the equation of a linear model to solve problems in the context of bivariate quantitative data, interpreting the slope and y-intercept.

Virginia 8th grade Standard Alignment

8.PS.3 The student will apply the data cycle (formulate questions; collect or acquire data; organize and represent data; and analyze data and communicate results) with a focus on scatterplots.

(a): Formulate questions that require the collection or acquisition of data with a focus on scatterplots.

(c): Organize and represent numeric bivariate data using scatterplots with and without the use of technology.

d): Make observations about a set of data points in a scatterplot as having a positive linear relationship, a negative linear relationship, or no relationship.

(e): Analyze and justify the relationship of the quantitative bivariate data represented in scatterplots.

(f): Sketch the line of best fit for data represented in a scatterplot.

Learning Objectives

By the end of this lesson, students will be able to:

1. Construct and interpret scatter plots to examine the relationship between two quantitative variables in a consumer credit Card dataset.
2. Identify and describe patterns in scatter plots, including positive or negative associations, clustering, outliers, and whether the relationship appears linear or nonlinear.
3. Draw and evaluate a line of best fit by informally assessing how closely the line represents the data.
4. Interpret the slope and y-intercept of a linear model in the context of consumer credit data.
5. Use a linear model to make predictions and answer questions about relationships between variables in a real-world financial context.
6. Communicate conclusions by using evidence from graphs and data to justify their interpretations and predictions.

Written by: Kimani Roach

Essential Question(s):

How can scatter plots and linear models help us make sense of real-world data?

How can data help financial institutions understand customer behavior?

The use of the Data cycle

This lesson engages students in the full data cycle (formulate questions; collect or acquire data; organize and represent data; analyze data and communicate results) The Data Cycle in a Nutshell as is shown in figure 1. While the figure shows an ordered cycle the process is not linear and involves moving back and forth between the phases. While the lesson plan is designed to showcase one possible flow of the data cycle/data investigation process (Data investigation Pathways NC Standards Share.pdf - Google Drive), teachers can choose to take their own paths within their respective classrooms. For help in thinking about how to engage in the various aspects of the data cycle, consider the following flowchart that shows possible connections between the elements of the data investigation process based on the type of analysis you want to design for your students.

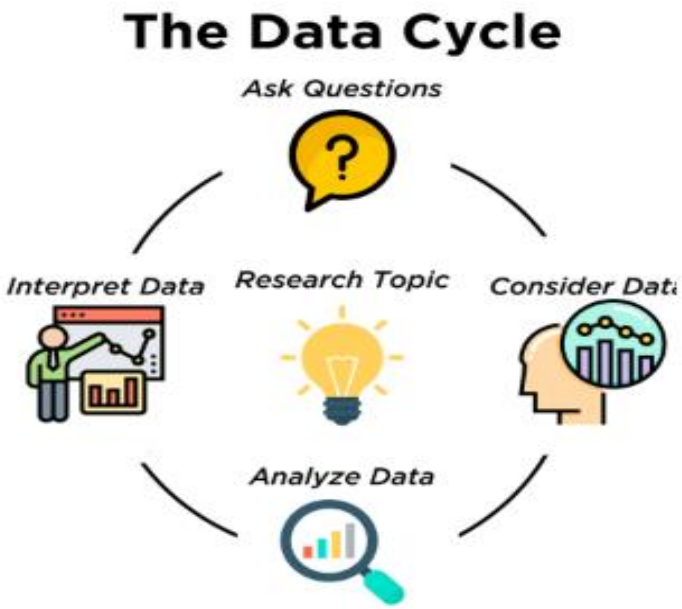


Figure 1. Bootstrap. (n.d.). *The data cycle*. Bootstrap: Data Science. https://www.bootstrapworld.org/materials/fall2025/en-us/lessons/data-cycle/pages/notes-data-cycle.html

Written by: Kimani Roach

Materials

CSV file with Credit Information

https://docs.google.com/spreadsheets/d/14hkFSr6OUD_LlenaFhVjJte8UycwSa08/edit?usp=sharing&ouid=109208274264053377689&rtpof=true&sd=true

CODAP of Credit Dataset

https://codap.concord.org/app/?v=3#shared=https%3A%2F%2Fcfm-shared.concord.org%2FEpWLq8C26ei01m94QzGg%2Ffile.json

Instructional Delivery Plan

Launch (15 - 20 minutes)

Figure 1: Scatterplot of credit Limit and Credit Balance of credit card applicants

The students will be given the following data visualization (scatterplot) of some applicants' credit balance and credit limit on their credit cards. The students will be asked to take some time to examine the data visualization and make a record of what they notice about the data visual and what do they wonder about it. They will be given an opportunity to turn to the student next to them and share one thing they notice and wonder after which the class will have a whole group discussion where the students will be given the opportunity to share their respective notice and wonder.

Written by: Kimani Roach

The goal of this is:

1. **For Students to observe and describe patterns in a scatter plot by identifying trends, clusters, outliers, and possible relationships between credit card balance and credit limit.**

After which the class will watch a video on credit and factors that are considered when financial institutions are considering credit (loans, credit cards, mortgage etc.) for applicants.

Video Suggestions

Video on no credit history, no credit score: Credit Scores and Reports 101 (Credit Card and Loan Basics 2/3)

Introduction to credit: Introduction to Credit - What You Need To Know

Consider the Data/ Creating a Statistical Question (20 – 30 minutes)

After students watch these videos, they will be given the opportunity to discuss credit and its role in us economy both by discussing its advantages and potential disadvantages. They will be given a dataset with the following scenario:

Scenario:

Carolina Community Bank has hired your team as junior data analysts. The bank wants to better understand the characteristics of people who apply for credit cards so it can make informed lending decisions and better serve its customers. Over the past several weeks, the bank has collected information from previous credit card applicants, including each applicant's **annual income, age, years of education, marital status, student status, gender, ethnicity, number of credit cards, credit rating, approved credit limit,** and **current credit card balance**.

Before making any changes to its lending policies, the bank has asked your team to analyze these data to identify patterns and relationships that may exist among the variables. Your job is **not** to decide who should receive a credit card. Instead, you will use statistical tools to investigate the data, answer meaningful questions, identify relationships between variables, develop linear models when appropriate, and communicate evidence-based conclusions that can help the bank better understand its customers.

As a junior data analyst, your first task is to explore the dataset and develop a statistical question that can be investigated using scatter plots and linear models.

Written by: Kimani Roach

https://codap.concord.org/app/?v=3#shared=https%3A%2F%2Fcfm-shared.concord.org%2FEpWLq8C26ei01m94QzGg%2Ffile.json

The students will be given the data and the opportunity to explore from the data students will come up with a question that they want to investigate after examining the Metadata to understand what each variable is representing. The question will then be fixed to make sure that it is a statistical question,

The goal of this is for:

- Students generate statistical questions from real-world data that can be investigated using bivariate data and scatter plots.

Sample Questions: How well does a person's credit limit predict their credit card balance?

Note: the teacher can give the students some guidelines to help them come up with questions that are both statistical and meet the standards of exploring associations between variables. Some of these guidelines could include:

Guidelines for creating statistical questions

1. **Your question should compare or connect two variables.**

- Ask about how one piece of information relates to another.

2. **Your question should be answered using the data in the dataset.**

- Make sure the information needed to answer your question is included in the table or graph.

Your question should expect different answers, not just yes or no.

- Good statistical questions look for patterns, trends, or differences because the data vary.

<u>Explore, Visualize and do Statistic (10-15 minutes)</u>

Once the students come up with their different statistical questions, the next step is to have them explore the dataset, for this section they will be allowed to utilize two additional software that will be introduced to them, Desmos and CODAP. This plan will outline how you could explore both options and integrate the two of them however, teacher can decide to only focus on one of them if that is what they see fit.

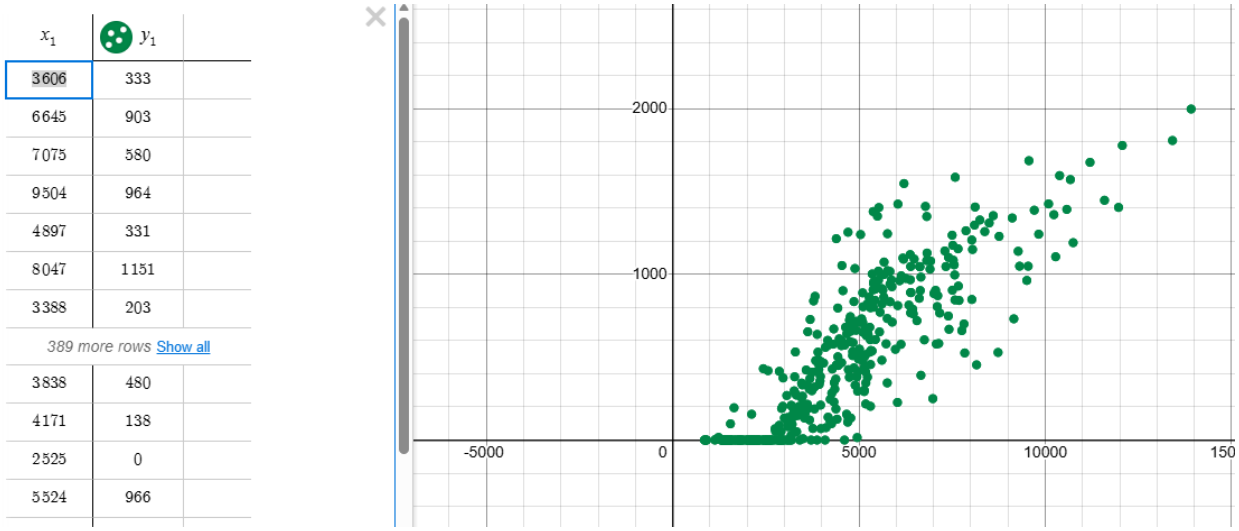
Sample Questions: How well does a person's credit limit predict their credit card balance?

Written by: Kimani Roach

DESMOS

Using access to the CSV link, the students will merge the two variables that they are focusing on, in the case Credit Limit and Credit Balance, students will copy and paste the data for both variables into DESMOS using **Credit limit as the (independent/Predictor/ x – variable)** and **Credit Balance as the (dependent/ response/ y – variable).** As shown in figure two, after which the students will be allowed to explore the data visual (Scatter) plot and the pattern/ trends that exist within the data.

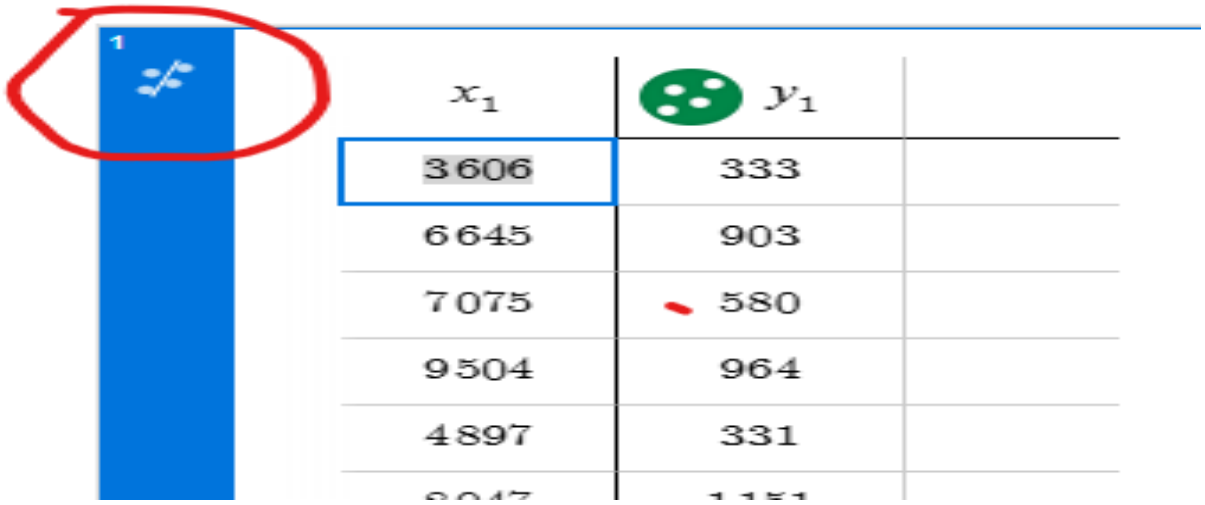
Figure Two: Scatter plot showing pattern between Credit Limit and Credit Balance



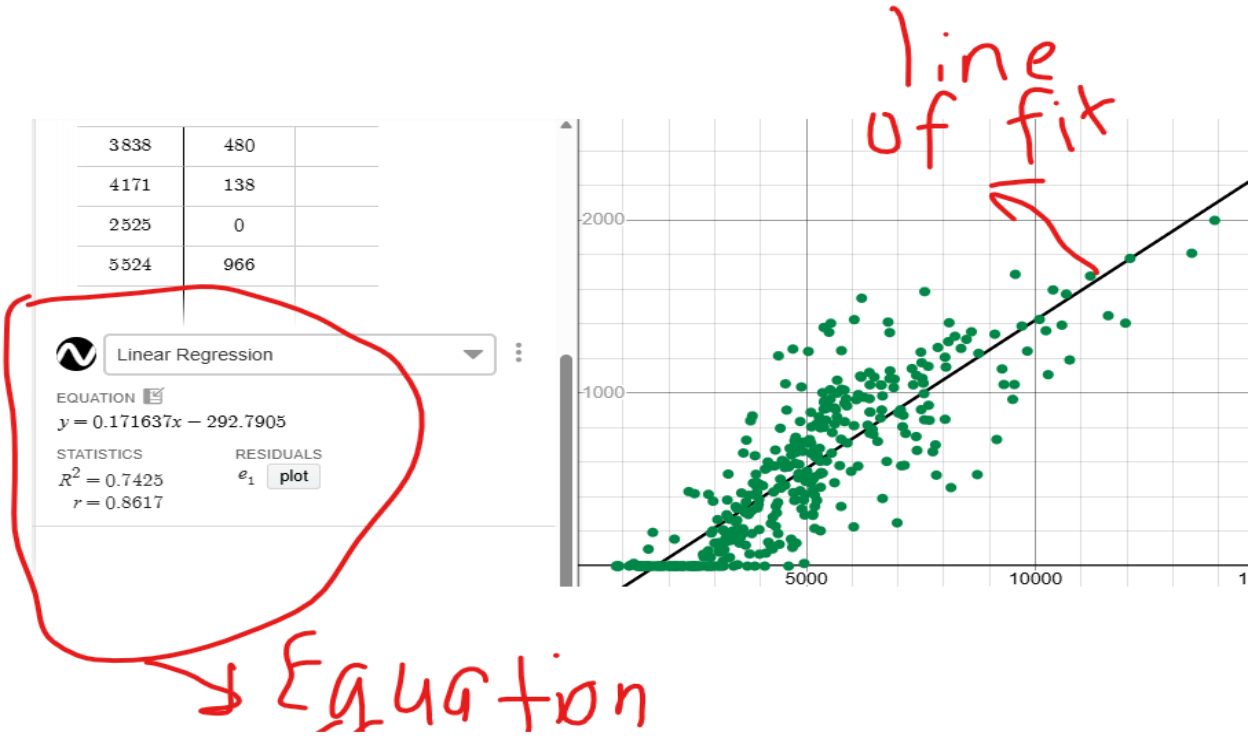
Once students have examined the scatter plot of the variables they are investigating, they can go ahead and run their regression model to fit a linear Regression model to the dataset and perform a model assessment using the regression feature as is shown below:

Step 1: to run the linear model to the three dots on the top left-hand corner of the software once the table is added, once you click those three dots the straight line is automatically fitted to the data and a formal equation that can be used to model the line is given

Written by: Kimani Roach

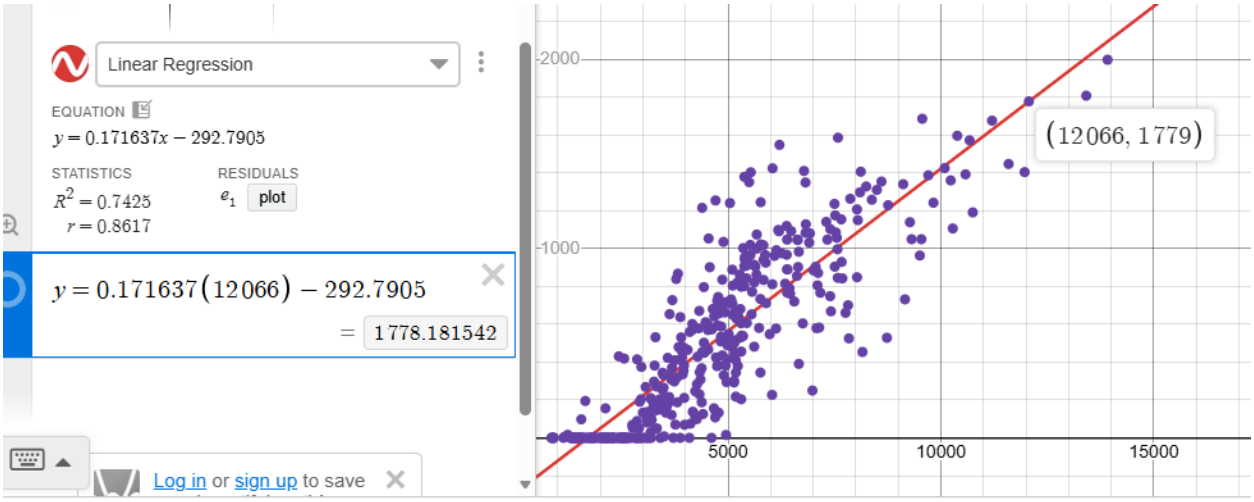


x_1	y_1
3606	333
6645	903
7075	580
9504	964
4897	331



Ther students could then start using the model to make some predictions and see how far the model was or how well it predicted what would have happened, for example, if an applicant has a credit limit of $12,066 what will their predicted credit balance be and how well does it compares to the actual balance?

Written by: Kimani Roach



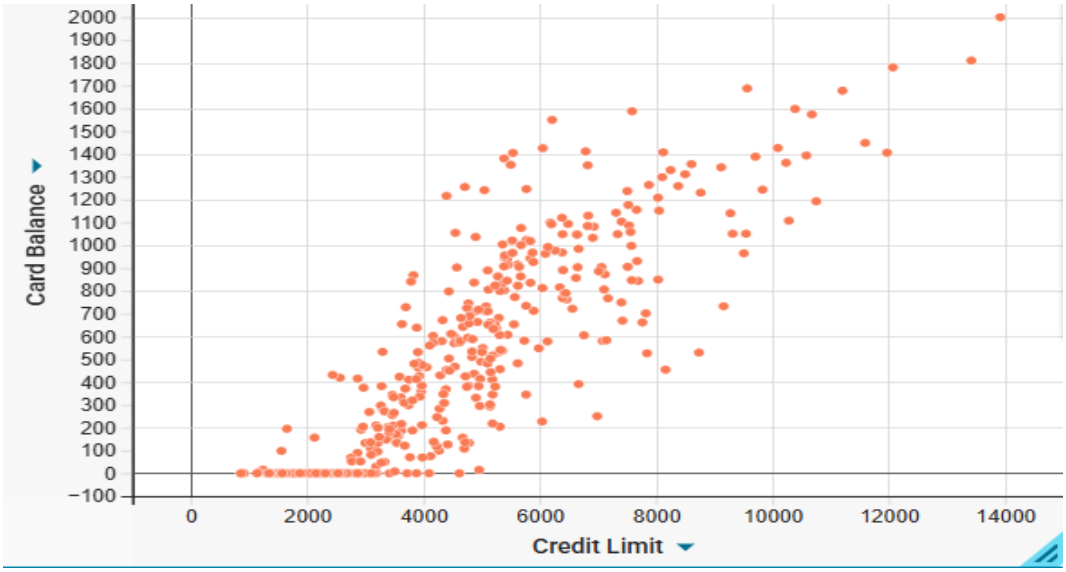
Note: the predicted balance would be $1778,18, however the actual balance would be $1779 which means the models almost perfectly predicted what would happen in this instance and was only off by roughly 82 cents.

CODAP

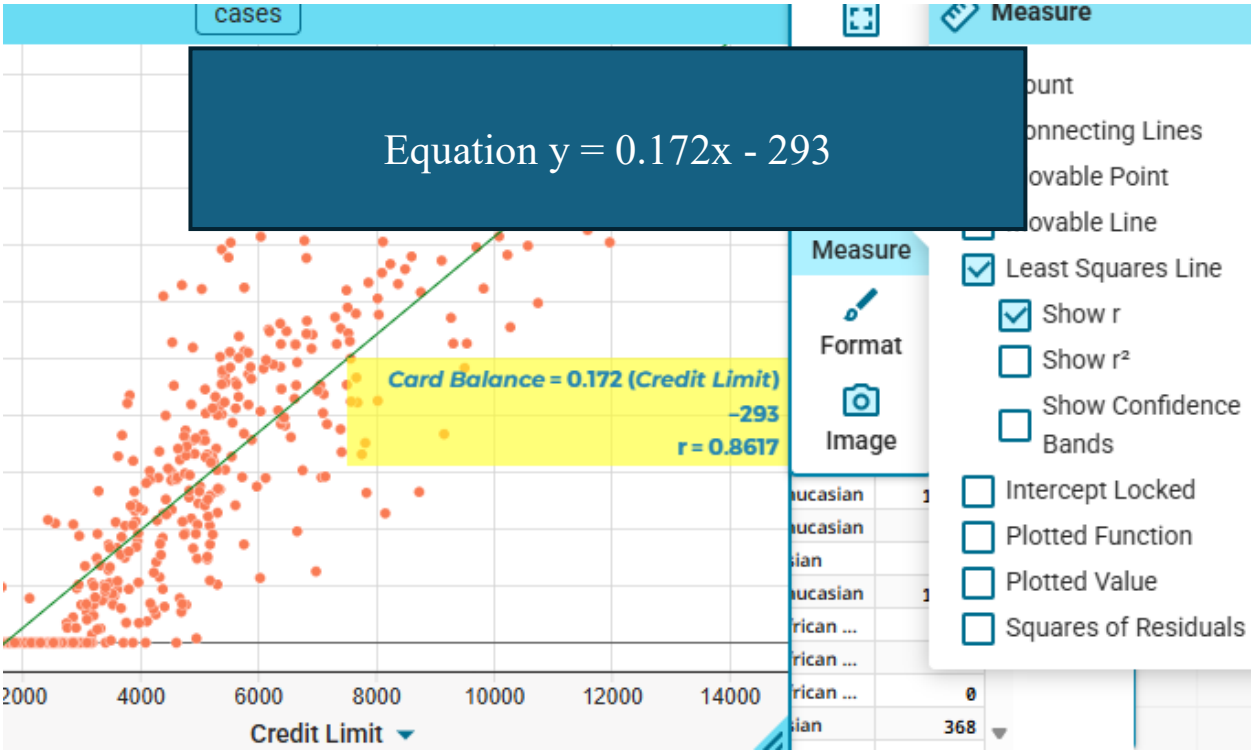
The Approach for CODAP id similar except, the students do not need to use a CSV as the dataset will be loaded into the software for immediate use of the variables the students are focusing on, in the case Credit Limit and Credit Balance, students will copy and paste the data for both variables into DESMOS using **Credit limit as the (independent/Predictor/ x – variable)** and **Credit Balance as the (dependent/ response/ y – variable),** this mean that credit limit will be dragged to the x – axis of the graph while credit Balance to the y – axis as is shown in figure three below, after which the students will be allowed to explore the data visual (Scatter) plot and the pattern/ trends that exist within the data.

Written by: Kimani Roach

Figure 3: Scatter plot showing pattern between Credit Limit and Credit Balance



Once the students have examined the pattern within the scatter lot, they will attached the best fit line by going to the right-hand side of the graph and select the following**: measure** followed by **Least Square Line** as is shown below. When you examine the visual you will notice that unlike DESMOS, CODAP does not directly give you the equation for that line of best fit, however you can write it with what is given to you example:



Written by: Kimani Roach

Note: DESMOS and CODAP will give different values as CODAP rounds the values to different places.

Communication Findings

Once students have explored the data and engage with it to aid their various investigations they can start to communicate their findings. A good start would be to make note of the things they notice as they dive into the data as this will help them to communicate the story that the data is telling:

The pattern within the scatter plot shows that there is a strong positive relationship/ association between applicants' credit limits and credit balance.

The equation for the line of best fit is $y = 0.172x - 293$. To help students make sense of this equation, they can draw on their prior knowledge of linear equations in the form $y = mx + b$.

- **m(the slope) tells us how much the response variable (y) is predicted to change for every 1-unit increase in the explanatory variable (x).**
- **b(the y-intercept or constant term) tells us the predicted value of the response variable (y) when the explanatory variable (x) is 0.**

Interpret the equation within context

Slope = 0.172: For every increase of $1 in the credit limit of the applicants, their predicted credit balance increases by $0.172.

Y-intercept = -293: When the credit limit $= 0$, the model predicts that the credit balance $= -\$293$. Students should also discuss whether this value is meaningful in the context of the data, since the intercept is not always realistic if $x = 0$ is outside the range of the data. A good question to ask to help them with this is**: Does a predicted credit limit of -$293 make sense? Why or why not?**

Note: While the slope provides a meaningful description of the relationship between the two variables, the y-intercept is not always meaningful, especially when $x = 0$ is outside the range of the observed data or when a negative prediction is not realistic in context.

Written by: Kimani Roach

This is actually a great opportunity to create a teaching moment for the student to understand that **mathematical models are approximations** and that every part of a model should be interpreted within the context of the data, rather than accepted at face value.

Conclusion/ answer Statistical Question (2 – 5 minutes)

Once the students have communicated their findings they will proceed to concluding their investigations and answering their statistical question within a very simple communicable format. The example is shown below:

Sample Statistical Question: How well does a person's credit limit predict their credit card balance?

Conclusion: The data show a **strong positive relationship** between a person's credit limit and credit card balance on their credit card history. This means that people with higher credit limits tend to have higher credit balances. The line of best fit models this relationship well.

References

Bootstrap. (n.d.). The data cycle. Bootstrap: Data Science. https://www.bootstrapworld.org/materials/fall2025/en-us/lessons/data-cycle/pages/notes-data-cycle.html

Common Online Data Analysis Platform (CODAP). (n.d.). Shared credit dataset. https://codap.concord.org/

Frank, E., Harrell, F. E., & others. (1993). Credit approval dataset. UCI Machine Learning Repository. https://archive.ics.uci.edu

North Carolina Department of Public Instruction. (2023). North Carolina Standard Course of Study: Mathematics Grade 8. https://www.dpi.nc.gov

Written by: Kimani Roach

North Carolina Department of Public Instruction. (2023). K–12 Mathematics Unpacking Documents. https://www.dpi.nc.gov

Acknowledgements

Acknowledgement

The author of this lesson would like to acknowledge **Miss Davis, teacher of Economics and Personal Finance**, whose dedication as a Finance and Economics educator inspired the development of this lesson. Her work in helping students understand consumer credit, credit scores, and the factors that influence creditworthiness demonstrate how real-world financial contexts can make mathematics meaningful and relevant.

The author also gratefully acknowledges **Dr. Travis Weiland** for his leadership and contributions to data science education. His work promoting statistical inquiry, data investigations, and authentic data experiences has greatly influenced the ideas and instructional approach incorporated throughout this lesson.

To Cite This Document

Roach, K. (2026). Investigating Consumer Credit Using Scatter Plots and Linear Models. Unbound Education. https://www.unboundeducation.org