

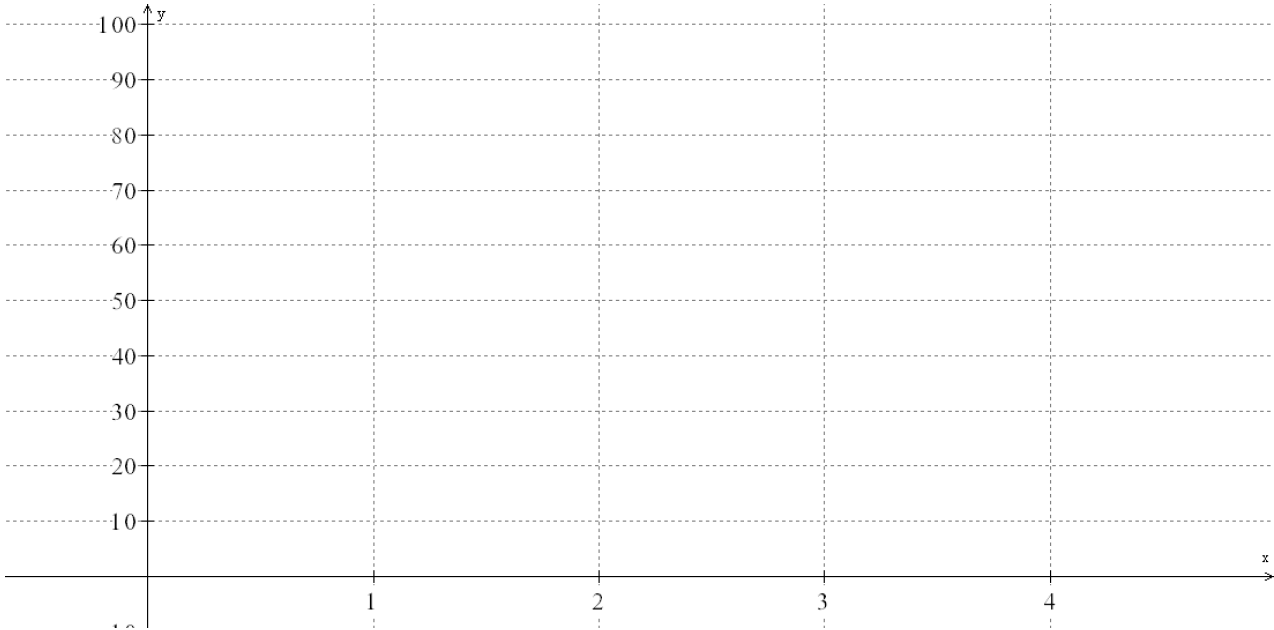
The Clash Royale Conundrum

Jonah is convinced that playing Clash Royale actually *helps* him focus better on schoolwork. Lucas, Norah, Tommy, and Sebastian are... skeptical. They decide to collect data over 12 weeks, recording the average hours per day Jonah spent playing Clash Royale and his test scores in AP Statistics.

The Data

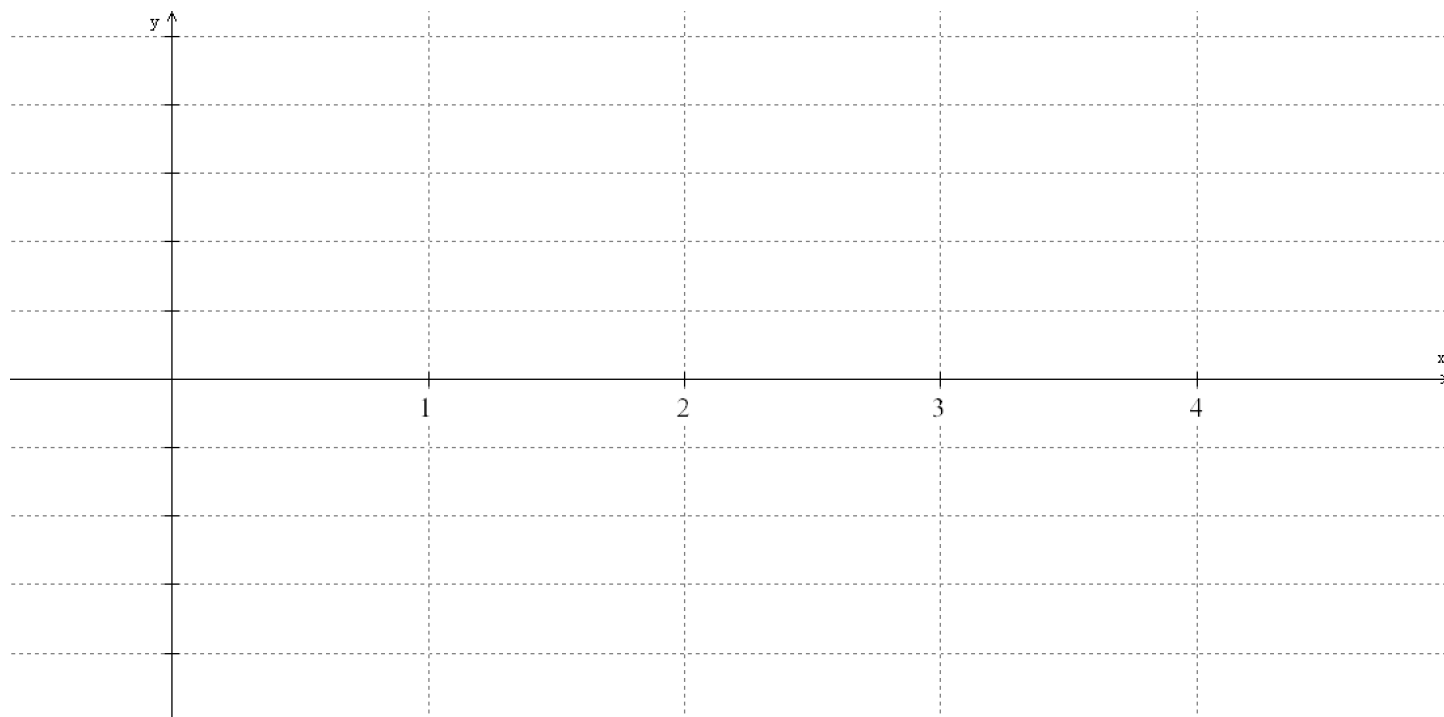
Week	Hours/Day Playing CR	Test Score (%)
1	0.5	94
2	1.0	91
3	1.5	89
4	0.8	93
5	2.5	82
6	3.0	78
7	1.2	88
8	2.0	85
9	3.5	74
10	2.8	80
11	1.8	86
12	4.0	68

- 1) Create a scatterplot with Hours/Day Playing CR on the x-axis and Test Score on the y-axis. Describe the relationship you observe (direction, form, strength).



- 2) Calculate the correlation coefficient (r). Show your work or describe the calculator steps you used.
- 3) Interpret the correlation coefficient in context. What does this value tell us about Jonah's "mental warm-up" theory?
- 4) Find the equation of the least squares regression line. Write it in the form: $\hat{y} = a + bx$ Use proper variable names in context (not just x and y).
- 5) Interpret the slope in context. Be sure to include units and explain what it means for Jonah's gaming habits.
- 6) Interpret the y-intercept in context. Does it make sense in this scenario? Why or why not?
- 7) Use your LSRL to predict Jonah's test score if he plays Clash Royale for 2.5 hours per day.
- 8) Calculate the predicted test score for Week 3 using your LSRL. Then calculate the residual for Week 3.
- 9) What does this residual tell us about Jonah's actual performance compared to what the model predicted in Week 3?
- 10) During Week 9, Jonah played 3.5 hours per day and scored 74%.
 - Calculate the residual for Week 9
 - Is this test score above or below what the model predicted?
 - Jonah claims his low score was because "the elixir golem nerf threw off my whole strategy." Based on the residual, was his performance better or worse than expected given his gaming time?

11) Create a residual plot (residuals vs. hours played). You may calculate all residuals first, or create the plot using your calculator's diagnostic features.



12) Analyze your residual plot:

- Does the plot show a random scatter around the horizontal axis?
- Are there any patterns or curves?
- What does this tell us about whether a linear model is appropriate?

13) Identify any potential outliers or influential points in your residual plot. What week(s) might these correspond to?

14) Calculate r^2 and interpret it in context. What percentage of the variation in Jonah's test scores can be explained by the hours he spends playing Clash Royale?

15) Jonah's table, concerned for his academic future, sees these results and immediately confiscate his iPad.
"Playing Clash Royale is CAUSING your grades to drop!" Norah exclaims.

Is this a valid inference based on this regression analysis? Explain your reasoning, addressing:

- Correlation vs. causation
- Possible lurking variables
- The limitations of this study

16) Suppose Jonah decides to quit Clash Royale entirely (0 hours/day) based on this analysis. Using your LSRL, predict his test score. Would you trust this prediction? Why or why not? (Hint: Think about extrapolation)

Jonah argues: "What if I'm only playing so much because I'm already stressed about my bad grades? Maybe bad grades cause MORE gaming, not the other way around!"

Question 17: This is an example of reverse causation. If Jonah's theory were true, what would we expect to see if we switched our explanatory and response variables? Would the relationship change? Discuss.