1. Week 11 - Correlation - Answers

Answers

When you have completed all of the lab content, you may want to check your answers with our completed version of the script for this week. **Remember**, looking at this script (studying/revising it) does not replace the process of working through the lab activities, trying them out for yourself, getting stuck, asking questions, finding solutions, adding your own comments, etc. **Actively engaging** with the material is the way to learn these analysis skills, not by looking at code written by someone else…

Lab activity 1: Interpreting correlation

1. Below are scatterplots that show the relationship between 'how much you know about correlation and how attractive you appear to members of the opposite (&/or same) sex'. Choose the type of correlation (strength and direction) displayed in each graph.

Answer

a. strong positive correlation
b. null correlation
c. moderate positive correlation
d. perfect negative correlation

2. Suppose it was observed that there is a correlation of r = -.81 between a driver's age and the cost of car insurance. This correlation would mean that, in general, older people pay more for car insurance.

Answer

FALSE Explanation: The correlation coefficient is negative and therefore infers a negative correlation. As such, older people pay less for car insurance: as age increases, car insurance costs decrease.

3. Suppose that there is a correlation of $r = .87$ between the length of time a person is in prison and the amount of aggression the person displays on a psychological inventory administered at release. This means that spending a longer amount of time in prison *causes* people to become more aggressive.

Answer

FALSE Explanation: This is a bit of trick question as it has the sneaky 'cause' word in. The correlation coefficient is a positive number, suggesting a positive relationship between length of time in prison and aggression. However, causation cannot be inferred from correlation and therefore we cannot know whether time spent in prison CAUSES aggression, and rather we suggest a relationship between the two: as length of time in prison increases, aggression increases.

4. A significant correlation was found between having great hair and performance in correlation labs. The correlation coefficient was .7. How much variance in correlation lab performance can the 'greatness' of your hair explain?

Answer

c 49% The 'coefficient of determination' or 'R-squared' tells us the proportion or variance in one variable that can be predicted if we know the other variable. We can determine this by squaring the r. Therefore, $.7^2 = .49$, $R^2 = .49$.

Lab activity 2: Constructing scatterplots and calculating correlations

You can download the RMd-script that contains the code to complete lab activities 2 and 3 here: 122_wk11_labActivities2_3.Rmd.

Question 3a: What can you tell from the scatterplot about the direction of the relationship?

Answer

There is a negative association between 'Home' and 'TV'. This means that the longer a child spends watching TV, the shorter they will read at home.

Question 4a: What is the correlation coefficient (Pearson's r)?

Answer

$\boldsymbol{r = -.65}$

Question 4b: What is the p value?

Answer

$p < .001$

Question 4c: Is the correlation significant at the p < .05 level?

Answer

Yes, because the p-value is smaller than .005

Question 4d: What are the degrees of freedom you need to report?

Answer

23

Step 5. How much variance in 'time spent reading' can be accounted for by 'time spent watching TV'?

Answer

42%

Step 6. Write a few sentences in which you report this result, following APA guidelines.

Answer

Something along the lines of: A Pearson's correlation coefficient was used to assess the relationship between time spent reading at home and time spent watching TV at home. There was a significant negative correlation, $r(23) =$ -.65, $p < .001$. As time spent watching TV increased, time spent reading at home decreased.

Lab activity 3: Hazardous alcohol use and impulsivity

Question 1a: How many variables does it have?

Answer

3

Question 2a: What can you tell from the scatterplot about the direction of the relationship?

Answer

There is a positive association between 'hazardous alcohol use' and 'impulsivity'. This means that as a participant's score on 'hazardous alcohol use' goes up, their score on 'impulsivity' also goes up.

Question 3a: What is the correlation coefficient (Pearson's r)?

Answer

$r = .54$

Question 3b: What is the p value?

Answer

$p = .014$

Question 3c: Is the correlation significant at the $p < .05$ level?

Answer

Yes

Question 3d: What are the degrees of freedom you need to report?

Answer

18

Question 3e: How much variance in 'impulsivity' can be accounted for by 'hazardous alcohol use'?

Answer

29%

Question 3f:. Give three logically possible directions of causality, indicating for each direction whether it is a plausible explanation in light of the variables involved (and why). No, this is not a trick question —-I know that correlation does not infer causation, but think critically! New studies/ideas are constructed by thinking what the previous study doesn't tell us about what could be happening with the variables of interest.

Answer

Just really looking for reasoning here.
Examples:

- **Being more impulsive may make people consume more alcohol.**
- **Consuming more alcohol may make people more impulsive.**
- **An outgoing personality might influence both your level of impulsivity and you are more likely to be socialising in the pub and consuming alcohol. So the same 'third factor' may influence both our variables of interest.**

Step 4. Write a few sentences in which you report this result, following APA guidelines.

Answer

Something along the lines of: A Pearson's correlation coefficient was used to assess the relationship between alcohol use and impulsivity. There was a significant positive correlation, $r(18) = .54$, $p < .014$. People who reported to consume more alcohol, scored higher on the impulsivity scale.