There are TWENTY questions, which tend to get harder as you go along.

A researcher believes cupcakes make people fall asleep. At noon, he feeds cupcakes to a random sample of cupcake-eaters until they are stuffed, then times how long it takes for the volunteers to fall asleep. He finds that 80% of them fall asleep within two hours.

**Question 1:** From this study alone, can you conclude that cupcakes cause obesity?

**Question 2:** How could the researcher improve his experimental design?

**Question 3:** Here is a frequency histogram of cupcake weights. What are the lower and upper quartiles of cupcake weights?

The scale the researcher uses to weigh the cupcakes is not completely accurate: it gives a measurement error of mean 0 and SD 1 gram.
Question 4: What is the probability the weight given by the scale for a certain cupcake differs from its true weight by more than 1 gram (in either direction)?

Question 5: Here is a boxplot of the number of sprinkles on each cupcake. Explain why the SD might be misleading as a measure of the spread of the number of sprinkles on each cupcake, and suggest a better measure of spread.

The following table gives the age, maximum number of cupcakes that person can eat, and age times maximum cupcakes for each person in the sample:

<table>
<thead>
<tr>
<th>Age</th>
<th>Max. cupcakes</th>
<th>Age times max. cupcakes</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>18</td>
<td>450</td>
</tr>
<tr>
<td>30</td>
<td>20</td>
<td>600</td>
</tr>
<tr>
<td>40</td>
<td>15</td>
<td>600</td>
</tr>
<tr>
<td>60</td>
<td>10</td>
<td>600</td>
</tr>
<tr>
<td>65</td>
<td>12</td>
<td>780</td>
</tr>
<tr>
<td>Mean</td>
<td>44</td>
<td>15</td>
</tr>
<tr>
<td>SD</td>
<td>15.937</td>
<td>?</td>
</tr>
</tbody>
</table>

Question 6: What is the SD of the number of cupcakes eaten?
Question 7: What is the correlation between age and number of cupcakes eaten?
Question 8: What is the regression line for predicting number of cupcakes eaten from age?

Question 9: How would you check that the Gaussian linear model holds for this data?

Question 10: Assuming the Gaussian linear model holds, what percentage of 50-year olds will eat more than 15.5 cupcakes?

Question 11: Based on this and larger samples, the researcher concludes that as you get older, your maximum number of cupcakes eaten goes down. Explain why this might not be the case.

Question 12: I roll a die twice. What is the probability that the first roll is odd or the second roll is even?

Question 13: I roll a die twice. What is the probability that the first roll is odd and the sum of the two rolls is odd?

Question 14: What are the advantages and disadvantages of a cross-sectional study?

Question 15: A certain population has mean weight 150 pounds with SD 15 pounds. The weights are not normally distributed. What is the smallest possible value of the lower quartile of weight for this population?

Question 16: A certain population has normally distributed heights with mean 66 inches. 90% of this population have heights below 69.9 inches. What is the interquartile range of heights for this population?

Question 17: Give numerical values for a data set consisting of variables A, B and C, such that A and B are correlated; A and C are correlated; but B and C have correlation 0.

Question 18: In a certain population, the log of weight has a normal distribution. Will the median weight be greater than, less than, or equal to the mean weight for this population?

Question 19: I roll four dice. What's the probability their sum is 14?

Question 20: I deal five cards from a shuffled deck. What's the probability of getting four aces?