**I. Understanding Data**

Complete the following table.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable Name** | **Question** | **Categorical or Continuous?** | **Scale** |
|  | Eye color | Cat. Quant. | N O I R |
|  | Number of siblings | Cat. Quant. | N O I R |
|  | State where you were born | Cat. Quant. | N O I R |
|  | Did you eat breakfast this morning? | Cat. Quant. | N O I R |
|  | Did you vote in the last election? | Cat. Quant. | N O I R |
|  | Shoe size | Cat. Quant. | N O I R |
|  | Seconds you held your breath | Cat. Quant. | N O I R |
|  | Approximate number of miles your parents/family live from Purchase | Cat. Quant. | N O I R |
|  | Total $ in cents in your purse, wallet, bag, etc. | Cat. Quant. | N O I R |
|  | College major | Cat. Quant. | N O I R |
|  | Year in school (freshman, soph…) | Cat. Quant. | N O I R |
|  | Care about learning from this class | Cat. Quant. | N O I R |
|  | Worried about doing well in this class | Cat. Quant. | N O I R |

**II. Understanding and graphing frequencies**

Have you ever played Crocodile Dentist, pictured here? You take turns pushing down a tooth until the crocodile chomps down on someone and that person is then out of the game. Ever wonder if the chomps are randomly distributed across the teeth? In other words, might you wonder whether some teeth are more “dangerous” than others? How would we figure this out?

Did you say we need to count the frequencies of chomps for each tooth? If so, you’re in luck because I spent quite a while collecting this data for you! So let me tell you what I did. First, to make things easier, I gave each tooth a name, starting from “A” all the way to “J” because there were ten teeth. Then, I conducted 35 trials pushing down one tooth at a time, starting from “A” and going in order. When it chomped me, I tallied which tooth it chomped down on. After 35 trials, I had the following data:

|  |  |
| --- | --- |
| Tooth “Name” | # of Chomps |
| A | 5 |
| B | 5 |
| C | 6 |
| D | 6 |
| E | 7 |
| F | 2 |
| G | 1 |
| H | 1 |
| I | 2 |
| J | 0 |

1. We have one variable, “tooth name” which is on what kind of measurement scale? \_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
2. Then, we have the frequency for each tooth. Frequency is always what kind of measurement scale? \_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
3. **Now, re-create this frequency table in Excel either by copy/pasting it or typing it in by hand.**
   1. A frequency table will list the variable in the first column. In this case, the “tooth” varies. So we title the column based on the variable name and then below the title we list the different “values” the variable can have (in this case the letters A-J). Then, in our other columns, we include information about the frequency that we see each value of the variable.
   2. This frequency table is showing what we call “raw frequency” which represents the literal number of times that value was observed.
   3. We also sometimes report “relative frequency” also known as the percentage. In other words, if we had watched the Crocodile Dentist 20 times and it had chomped on Tooth A 2 times, then it would have chomped on Tooth A 2 out of 20 times (in math terms: 2/20 🡪 .1 🡪 10% of the time)
4. Can you easily tell which teeth are more dangerous from the Frequency Table? We can tell a bit, but a visual chart would make it even easier. To create a frequency chart, highlight your Frequency Table in Excel and then click “Insert” and then “Recommended Charts.” Scroll through the options to see how they each would visualize and represent the data. For frequencies, bar charts are often the most useful, so that’s what we’ll choose:
5. Here, we can quickly see which teeth are the most likely to bite us. Which teeth are the most likely to bite you? \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Which teeth seem somewhat “safer?” \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_
6. The default graph is nice, but it is not formatted in APA style, so we need to make some adjustments. You can click on “Add Chart Element” to then add a label for your x-axis (Tooth Name), and your y-axis (# of Chomps).
7. I’m also going to edit my title by clicking on it to make it more informative (# of Chomps Per Tooth in Crocodile Dentist going from A to J).
8. Finally, APA style does not include those horizontal lines. So right-click directly on one of those lines (or double-click) until each horizontal line is showing you a circle at each end. Then a side-bar should open up that will let you change the line to “No line.”

Screenshot showing the "format major gridlines" toolbar on the right-hand-side of an Excel spreadsheet. An arrow shows you to click where it says "No line." 
The limited view of the Excel worksheet to the left of the toolbar shows how the graph lines are selected, which is indicated by them looking blue and having little circles on the ends.

1. **What issues do you see with my research methods in collecting that Crocodile Dentist data?**
2. **There are a few issues you may have come up with, but maybe you wondered if the data only came out that way because we always started from left to right (what I called teeth A through J). Maybe you wondered what our data would look like if we instead went from right to left (J through A). You’re in luck because I did that, too, for another 35 trials! Here’s that data added to our previous frequency table.**

|  |  |
| --- | --- |
| Tooth “Name” | J to A # of Chomps |
| A | 0 |
| B | 1 |
| C | 1 |
| D | 4 |
| E | 8 |
| F | 7 |
| G | 5 |
| H | 5 |
| I | 2 |
| J | 2 |

* You’ll notice that we have updated the title for our frequency column. Now that we have two kinds of raw frequencies we’re examining, it is better to use a clear, precise title so we do not get confused about which frequencies we’re looking at.

1. Try making a chart again on your own following the same steps as before. Edit your chart like we did before to make it APA style. Then paste it here.
2. [Pretend you had not seen the first chart.] Interpret this graph. What conclusions are you drawing about which teeth are more or less dangerous and why?
3. Now, let’s put our information together across the two ways we tested the Crocodile Dentist (left-to-right and right-to-left). We are going to add the frequencies from both ways to get the total frequencies for each tooth. If you place both frequency columns side-by-side, you can use a formula to have Excel calculate their sum. In the next column to their right, you would type: =SUM(1stcell:lastcell) where you either enter in the first and last cell values of the data you want to add OR drag and highlight the cells. You can also do this below to create a Total Trials row.

|  |  |  |  |
| --- | --- | --- | --- |
| Tooth “Name” | A to J # of chomps | J to A # of Chomps | Total Chomps |
| A | 5 | 0 | 5 |
| B | 5 | 1 | 6 |
| C | 6 | 1 | 7 |
| D | 6 | 4 | 10 |
| E | 7 | 8 | 15 |
| F | 2 | 7 | 9 |
| G | 1 | 5 | 6 |
| H | 1 | 5 | 6 |
| I | 2 | 2 | 4 |
| J | 0 | 2 | 2 |
| TotalTrials | 35 | 35 | 70 |

1. Now that we have those numbers, we can calculate the Total Relative Frequencies. You can similarly have Excel do this for you. It’s the Total Raw Frequency DIVIDED BY the total number of trials (so in this case, 70). You can use a formula to get this value for each tooth =cell/70
2. Now make a new frequency table and paste it below that shows just the Total Relative Frequencies of each tooth.
3. Now paste a new graph showing the total relative frequencies of each tooth chomping.
4. What would you conclude now that we have graphed it?

1. If you would want more information to draw stronger conclusions, tell me what data we would need to collect.

**More on Graphs**

For the following assignment, you will need to make charts on the computer using Microsoft Excel. You can use color, but make sure the graph is still comprehensible when printed in black and white. Make sure to read the instructions for each chart—some questions specify the data you should be depicting.

How do you decide which kind of chart to use? Here are some guidelines to use in addition to what we have discussed in class:

1. If a chart has a categorical variable and a continuous variable, a bar chart may be the most appropriate. Vertical bars are preferred over horizontal bars the majority of the time.
2. If a chart has two continuous variables (it may also have a third categorical variable in some cases), then a line graph or histogram may be the most appropriate.
   1. If the purpose of the graph is to show how one variable changes as a function of the other, then a line graph may be the most appropriate.
   2. If the purpose of the graph is to show the distribution of scores and you want to represent the frequency or percent of times each variable occurs, then a histogram may be most appropriate.
3. If a chart has a categorical variable and what percentages of each add up to make 100% of another variable, then a pie chart may be most appropriate. The purpose of a pie chart is to illustrate the proportions making up something.

**Additional tips**:

1. The dependent variable (outcome variable) is plotted on the Y-axis (this variable will always represent a quantity of something), while the predictor variable is plotted on the X-axis.
2. If you have three or more variables, you will need a legend and additional bars or lines

1. YouTube has tons and tons of tutorials for how to make graphs in Excel, so make sure you check there (or elsewhere on the internet) first if you need some help! Part of this assignment involves working independently and being resourceful in figuring out how to create an appropriate graph

1. Aesthetics matter! The charts you create should be easy to look at and understand (clear labeling, large and clear fonts, etc.)

PRO-TIP: You can try to copy/paste these tables straight into Excel rather than typing them out.

PRO-TIP: Use Google to find tutorials or help on any chart features you cannot figure out on your own. The truth is out there