WEEK 6 REPORT – VISUALIZATION

Group Breakout Room Number: Jeffery 156  
Date and Time of Tutorial Session: Wednesday October 9th, 4:00-5:30pm  
Group Members: Katarzyna Fraser, Hanna Asin, Ella Pustil, Annika Tran,   
Graded out of 10 marks.

This assignment is based on cyberbullying data from a national survey carried out in 2020 asking 1000 randomly selected young people between the ages of 12 and 18 years old about both their offline and online behaviour.

**Q1** For boys, consider how responses to the intervention question change with age group.

* 1. Create a contingency table in RStudio and fill in the table below to show the number of *boys* who selected each response to the intervention question for each age group (2 marks)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | Agree | Somewhat agree | Somewhat disagree | Disagree | Pref not to say |
| 12 | 27 | 15 | 4 | 1 | 4 |
| 13 | 32 | 20 | 3 | 1 | 5 |
| 14 | 44 | 29 | 4 | 3 | 4 |
| 15 | 41 | 22 | 11 | 3 | 4 |
| 16 | 22 | 31 | 11 | 3 | 7 |
| 17 | 24 | 20 | 2 | 6 | 3 |
| 18 | 29 | 25 | 12 | 5 | 6 |

* 1. Which age group of boys had the largest *frequency* of "Agree" responses to the intervention question? (1 mark)  
       
     The age group of 14 had the largest frequency of agree responses

See code:

largest\_frequency\_age = which.max(boys\_table[, "Agree"])

* 1. Which age group of boys had the largest *proportion* of "Agree" responses relative to all responses? (1 mark)

Group 12 had the largest proportion of agree responses

See code:

boys\_proportions = prop.table(boys\_table, 1)

**Q2** For 18-year-olds, how does gender identity relate to willingness to intervene?

1. For 18-year-olds, create a bar plot that compares intervention responses (i.e., "Agree", “Somewhat agree”, etc.) for all levels (i.e., categories) of gender identity. Gender identity should be on the x-axis. Note that you can use the table command in R to create a table that can then be used with the barplot command. In the plot, include a legend and a descriptive figure caption under the plot for full marks. You can choose whether to make a stacked barplot or a grouped (i.e., non-stacked) barplot (3 marks).

A graph of different colored bars

Description automatically generated

1. For 18-year-olds, how does the willingness to intervene to mean comments, posts, or pictures that participants judge to be unfair or morally wrong differ between boys and girls? (1 mark)

The willingness to intervene to mean comments, posts, or pictures that they judge as unfair or morally wrong differs significantly. The girls were the most likely to intervene in comparison to the boys, who were less likely to intervene

1. For 18-year-olds, what challenge, from a statistical point of view, do you see in interpreting the responses for the gender diverse group? If you were to redo the survey, what changes might you make to address this challenge? (2 marks)

The main challenge in interpreting responses from the gender diverse group is likely due to the small sample size. When a group (gender diverse in this case) has significantly fewer respondents, it can lead to statistical bias and a lack of good representation. This may cause it to be difficult to draw meaningful conclusions or generalize the findings to a broader population. With fewer data points, any patterns or trends observed may be due to chance rather than any actual relationship, which overall reduces the reliability (how trustworthy/accurate it is) and validity (measures what it aims to measure) of the results.

**Notes:**

* Only one group member submits the report
* The report must be a Word .DOC, .DOCX or .PDF file
* Make sure everyone in the group has a copy of the report
* Double check what you have submitted‼
  + view it on OnQ to make sure everything is there and visible
* Lastly, everyone in the group needs to submit their own version of the R script.