

## Abstract:

The point of this document is to highlight the changes made to the original data and the differences between the two spreadsheets.

## Changes Made to Original Data:

Firstly, some data rows from the original sheet had to be removed - lines 75<sup>1</sup>, 76<sup>2</sup> & 81<sup>3</sup> refer to footnotes for an explanation - said data was removed due to its irrelevance with the questions asked. Secondly, degrees were cleaned up for consistency - answers such as “comp sci” were submitted and were changed to “Computer Science.” Lastly, two unique sheets were created to store the cleaned data. Each sheet specializes in a different presentation of the “**What program are you currently studying?**” column. One sheet, “**Cleaned Up ChatGPT Responses Non-Generalized,**” keeps the specific program - i.e., “Computer Science” is used as the answer within the cell. However, the other sheet, “**Cleaned Up ChatGPT Responses Generalized,**” changes the specific degree and replaces it with the general department which it falls under.<sup>4</sup>

## Argument for Generalized Data:

The decision to generalize the data was made because many specific degrees may lead to convoluted and hard-to-read data. If specific data is used, over forty points must be graphed for degrees alone. However, if generalized data is utilized, only eleven data points would need to be graphed for degrees - roughly seventy-three percent less than specific data. We must decide, as a group, what spreadsheet to use: either generalized or non-generalized.

---

<sup>1</sup> Row 75 has gender as “Chinese.”

<sup>2</sup> Row 76 has gender as “Arab.”

<sup>3</sup> Row 81 has gender as “Tiny wang gang” and degree as “Rizzology.”

<sup>4</sup> Exceptions include “Occupational Health and Safety.”