

Adam Abounnaaim

- An oddity I encountered in the data were the presence of negative values within some columns, such as for Potassium or Vitamins. I assume that all the values in the data set are in terms of grams or other units of measurement, as is usual for nutritional information, so a value of “-1” would mean “-1 grams of potassium” or “-1 cup of cereal,” which are illogical. I went through the data under this reasoning and performed a small cleanup, changing the -1s to 1s and presuming that the insertion of a negative value was just an error.
- G and K consume a massive amount of the total cereal count, with other manufacturers only occupying much smaller positions and contributing smaller numbers in comparison to them. Depending on what context or issue the data is being analyzed for, this observation could either be good news or bad news, such as an over-reliance on only two manufacturing centers and the need to increase production at others to prevent system failures should G and K experience major problems.
- The sodium content of all the cereals appears incredibly high in comparison to the content of other nutrients such as fiber or protein. This observation might prompt looking into low-sodium ideas or options, which could potentially be a whole new product of its own if demand for lower-sodium options is evident.
- Is manufacturing center A new? With how only one cereal is produced from there, if it is not new, then it is being seriously underutilized. This assertion depends on context, but this ties in with the observation that G and K are dominating the production of cereals, which is logistically unwise in my opinion.
- The range of the amount of nutrients present in the circles is incredibly wide, with some maximum values being in the double-digits while minimum values sit at 1 or even 0. This might be alright, but it makes me think that a standard could be set to have a certain amount of a nutrient present across all cereals possibly, though I’m uncertain what sort of purpose a decision like that would serve.

I loaded the data into Excel—since that is the application I am most comfortable analyzing spreadsheets within—and performed the small cleanup previously mentioned before. I also finished analyzing maximum-minimum values for all the columns and was able to make a pie chart of the percentage of total cereal count by manufacturer, from which I noticed that G and K occupy an excessively large share of the pie.

The only challenge and frustration I encountered was in trying to create a clustered-bar graph of the maximum-minimum values of the nutritional information, but every time I did so, the formatting would be thrown off somehow and it wouldn’t look the way I intended it to. Google searches were of no help as well unfortunately.