Miles A. Peña

DSC 500 Introduction to Data Science

Matthew Metzger

October 21, 2023

Concession Data

I created a pivot table with the data as it was provided and noticed that there were a few categories that were duplicates due to the fact that there were some spelling errors. I first used spell checker on Excel in order to ensure that there were no misspellings and re-did the pivot table. I then corrected the dates to be in the same format and corrected the dates like 01/41/22 as well as the dates that were reflecting January 2023. After the dates were corrected, I verified the min and max of sale price for each item to ensure that every item was charged correctly and corrected those that weren’t. I used the find and select tool in the sales spreadsheet to double check that all the prices were correct for each item based on the min and max from the pivot table. I then clicked on the category column in order to see what categories were present to ensure that there were no duplicates due to misspellings. There were two misspellings, corrected these and re-did the pivot table with all the clean data.

I used =VLOOKUP() in order to look up data from ‘Item Look Up Table’ column 2 and bring over the cost of goods to the sales transactions worksheet. I used absolute cell referencing when referring to the item look up table in order to be able to copy the =VLOOKUP formula down the column. Once I had all the costs in the sales transaction worksheet, on the next column I used a formula that subtracted the cost of goods from the sales price in order to find the profit of the item. I copied this all the way down the column. At the bottom of this column I took the =SUM() of all rows in this column to find the total profit for all sales. I then sorted the cells by date to make it easier to find the total profit per day. To do this, I used =SUM() and took all the rows that pertained to a specific day and this this for each date displayed. I used =MAX() in order to find the most profitable item and =MIN() to find the least (which ended up breaking even so then I went through and looked through the list for the one that created the lowest amount of profit without being zero).

To find the total calories I once again used =VLOOKUP() but this time column 3 in order to bring the calorie count from ‘Item Look Up Table’. I again used absolute cell referencing to be able to copy the formula all the way down without altering the range. I created a new column in the sales transactions worksheet titled Total Calories to place this information into. At the bottom of this column I took the =SUM() of all rows in this column to find the total calories for all sales. I then took the =SUM() of each date to find the total amount of calories sold per day. I did two separate =MAX() and =MIN() formulas. On one I did just including all values to find the item with most caloric content which was nachos and the one with the least caloric content which was bottled water. I also found the =MAX() and =MIN() of the dates in order to find which day had the most caloric content sold and which had the least.