Deleted all 1/3/2023 since all transactions took place in January 2022. The 2023 data is irrelevant and will generate incorrect result. Formatted all the dates to the following format Month/Date/Year to make all transactions’ date format the same. Also, for the transaction that took place on 1/41/2022 I changed the date to 1/4/2022. We could make a guess and assume the correct date should be 4/1/2022 by removing 1 in front. But I feel more comfortable to make it 1/4/2022. Since all the transactions of 1/41/2022 happened between 1/2/2022 and 1/6/2022. It is my guess that 1/41/2022 belongs to 1/4/2022. Highly likely when the data was transferred from one format to another this error appeared. On the other hand, we all know there is only maximum 31 days in a month and there are only 12 months a year. Thus, if we move 4 before the slash then it becomes 14/1/2022. That is not possible. If we move 1 behind slash then the date becomes 1/4/12022. That is not possible either. If removing 2 at the end of the year then the date becomes 1/4/1202. The year then falls into 13th century.

Found incorrectly spelled the following Items as Licorce Rope, Gumy Bears, Popcrn, Pizzaa, Hmburger, and Popsile. Corrected misspelling to reflect correctly items. Also, in Category found and corrected misspelled the following Hot Fod, Beerages, and Beverags. In addition, corrected misspelled word Cand to Candy and Frozen Treat to Frozen Treats in category which removed incorrect categories and added the sales numbers to correct categories. Formatted all raw numbers as currency to reflect dollar denominated prices.

Noticed on 1/8/2022, for item Popsicle, Chocolate Bar and Bottled Water the prices were expressed in percentage term as 300%, 200% and 300% respectively. Corrected all the percentage numbers to prices as follows; $3, $2, and $3 respectively. It was not difficult to figure out the correct price because all other transactions for the same items had the same price either on the same or other dates. These three huge numbers would create positive skewness to the entire data set sample.

Found 2 out of 19 beer items sold for $2 and $3. Corrected both to have price of $4 for each because all other remaining 17 beer items sold for $4 on the same or different date. Seems like the lower prices were assigned mistakenly. After correcting all the misspelling, I was able to sort each item and category which helped me further to find anomalies and correct them. For example, item Soda and Licorice Rope had reported as $25 and $20 respectively. All the same items on the same day had been sold for $2.50 and $2.00 respectively. So, corrected to the same prices.

In dollar term Popcorn, Beer and Bottled Water were the most profitable items with total profit sales of $72, $38, and $32.50 for all days. However, in percentage term the most profitable Items slightly different. Popcorn, Bottled Water and Popsicle having profit margins of 90%, 83% and 83% respectively. Beer got kicked out by Popsicle. The total profit margin was calculated as sum of total profit of each item for all the days divide by the sum of total sale of each item. Gummy Bears is at the break-even level not earning any profit and not losing any money. Nachos is the least earning profitable item both in dollar and percentage term, $6 and 17% respectively. The remaining items are still earning good return ranging from $12-$27 and 50%-67% in profit margin. On January 6th and 7th, the highest caloric items were sold totaling 12,565 and 10,910 calories respectively. January 4th had the lowest caloric items sold totaling 4,585 calories in total per day.