The Christmas season has just come to a close and it is time to analyze the past month's sales. Upon viewing the inventory it is apparent record keeping is not what it should be.

First redundancies: I removed 3 out of 4 date columns. Also, I know this is a data set about Mountain Bikes, so I can remove the columns that label every row a mountain bike.

Then I fixed the sales orders. One was misnumbered, but context clues made the solution obvious. Then there was a duplicate entry that needed to be removed.

Null values are a little more complicated. Luckily I could cross reference more complete orders to fill out missing data points.

- Row 6 was missing unit cost, so I cross referenced it with Row 8, the sale of the same item.
- Row 9 was missing unit price so I cross referenced it with Row 2, the sale of the same item.
- Row 23 was missing quantity but the total cost is equal to the unit cost, meaning there is a quantity of 1.
- Row 22 is missing the product purchased, but it appears to be a Mountain-500 Black. What I can not decipher is the frame size. While this is not detrimental to sales records and their purpose, I would want to cross reference the inventory records just to be sure the information exists somewhere.
- While row 17 was missing an age group, an exact age was still provided. This was simply a matter of falling in line with the pre-existing age ranges provided. (This is not a redundant column as the age groups can be used for data visualization)

Lastly, because all of these mistakes are naturally human, I wanted to double check the math. Luckily total cost/revenue are both calculated using a function. The profit was manually entered which leaves room for typos. I will replace these entries with a function (Revenue-Cost). The profit Sum does not change after applying the function, pointing towards consistency.