**DATA CLEANING, ANALYSIS AND VISUALIZATION IN EXCEL**

**DATA CLEANING PHASE**

|  |
| --- |
| 1. Copy Data into a new sheet for cleaning. |
| 2. Put filter on all headers to make it easy for data viewing. |
| 3. In the Gender Column we can see that there are four unique variables (M, W, Men and Women). Use find and replace function to replace M with Men and W with Women. After this we will have only two unique variables (Men, Women) under the Gender column. |
| 4. Under Channel we saw a channel named Amenazon. After a bit of research and after talking to the Data Owners we realized that it was Amazon that has been wrongly spelt as Amenazon. So, we filtered for Amenazon and used the Find and replaced function to change it to Amazon. |
| 5. Under the column size we saw Men as a size. After Engaging the Data owner, we were made aware that there was no size as Men but that it should have read Medium. We used the find and replace function to change Men to M which stands for Medium. |
| 6. Under the column Quantity I realised that we had the words One and Two as part of the variables. This will prevent us from aggregating that column, so I used the find and replace function to change it to numbers instead of the words that we have. |
| 7. The column Amount was wrongly spelt as Amenount so I changed it to Amount. |
| 8. I want to categorize the ages to make it easy for aggregation. I will create 3 categories namely Young below age 30, Adult between ages 30-50 years and Old, ages above 50. I would create an IF statement based on data under the Ages column. |
| 9. I will need to derive the month name from the date column to aid monthly analysis. I will use the text function to derive this. Text(G2,"mmmm"). |
| 10. I checked if there are any null values or blank cells in our dataset. I used both the count and filter functions to check this. There were no blank or null values.  **DATA ANALYSIS PHASE**   |  | | --- | | The following are questions that I wish to answer using the cleaned data: | | 1. I want to compare the relationship between orders and sales monthly. | | 2. I want to know which gender purchased more. | | 3. I want to see the total returns and how they compare with deliveries. | | 4. I also want to know which states has the highest sales. | | 5. I also want to know sales by Gender and Age group. | | 6. I also want to know sales by channel   |  | | --- | | **Visuals Creating Phase** | | 1. I created pivot tables for all the information I needed to answer my questions stated above. | | 2. I created visuals that most appropriately tells the story I want to tell. | | 3. After creating the visuals, I created a slicer to be able to filter through all the various visuals created to help users navigate to what ever details they want for any visual. | | 4. I pinned all these different visuals to a dashboard to tell a story | | |