Sales dataset

This dataset was extracted from the internet and it contain a table which gives the record of the sales of accessories in different countries with the order unit,profit and other vital information. The dataset contains date, day, month, year, customer age, age group, customer gender, country, state, product category, subcategory, product', order quantity,unit cost', unit price,profit, cost and revenue. The dataset contains 113036 rows and 18 columns. Feel free to explore the data in the csv file before analysing it. You can change the data type of the column as most times they are always object. The column name might not suit your work well,so you can change it to your desired name. The main purpose of this challenge is to improve our analysis skill and prepare us fully for more larger dataset out there.

So here are 10 fantastic questions you can try to answer using your convenient analytics tools such as python,R or SQL:

1. Check for any missing values and drop if any found.
2. Clean the dataset. Some of the entered values might be incorrect so we need to clean it also.
3. Check for duplicates and also remove any duplicate found.
4. Extract this columns to a new table as we won't be using the rest. The column needed includes month,year, customer age,gender,country,profit,order unit.
5. Which country have the highest profit in year 2013,2014,2015,2016.
6. Visualize using any convenient visualization tool to show the trend of the profit with different years.
7. Which month in the year 2015 have the highest order and which year have the highest order also. Use a bar chart to show the progression of the increase or decrease in profit from year 2013 upward.
8. Use a pie chart to show how the gender interact with the goods sold. The pie chart should show who orders the product between the male and the female.
9. What is the total profit in 2016 in Canada.
10. Give a short and concise conclusion from the analyses you made from the dataset.

© Joe-Flames 🔥🔥🔥