**E-commerce Sales Analysis**

* Reading the data and Cleaning
  + We need to install Pyarrow to read feather data
  + We drop rows from the data where any one of the value is null.
  + Then we drop duplicated rows where even Order ID (these are not the products in same order but duplicated rows)
* Finding month with maximum sales
  + First we split month from Order Date column and create a new column for ease
  + We observe that each column has object datatype which we need to change to int and float according to the need
  + Now we need total sales (money spent) for each product which we get by multiplying quantity ordered with price of each product
  + Now we plot a bar graph for visualization to see which month has highest sales
  + Conclusion **-> December has the highest sales**

**A graph of blue bars

Description automatically generated**

* Finding city with maximum sales
  + Similar to retrieving month from order date, we split the purchase address to find city and create a new column
  + We explore pie chart here to find ratio of sales each city accounts for.
  + Conclusion -> San Francisco, Los Angeles and New York have highest sales

A pie chart with different colored circles with Crust in the background

Description automatically generated

* Determining most sold products
  + We group the data by product and sum the total units ordered and mean of their selling price

A graph of blue bars

Description automatically generated with medium confidence

* + We see AAA batteries were most frequently order over 30000 times, whereas LG dryer and Washing machine are least bough items.
  + Let’s compare their respective selling price

A graph with numbers and lines

Description automatically generated

* + We see that most sold products are one of the cheapest products, whereas the inverse is not true where the most expensive product is not least sold one.
  + We now take the top 5 selling products to understand any trend

A graph of different colored lines

Description automatically generated**A graph of blue bars

Description automatically generated**

* + We see the top 5 products account for same ratio of total sales throughout the year.
  + Except for 9th month, where total sales decreased but sales of USB-C charging cable increased.
* Determine any trend of products being bought together
  + Firstly, we drop unique Order ID so we only have those orders who at least have 2 products.
  + We now join the products for each order ID to form a comma separated list of products like this

A screenshot of a computer

Description automatically generated

* + We now use the above data to plot a pie chart showing top 5 pair of products sold together

A pie chart with different colored circles

Description automatically generated