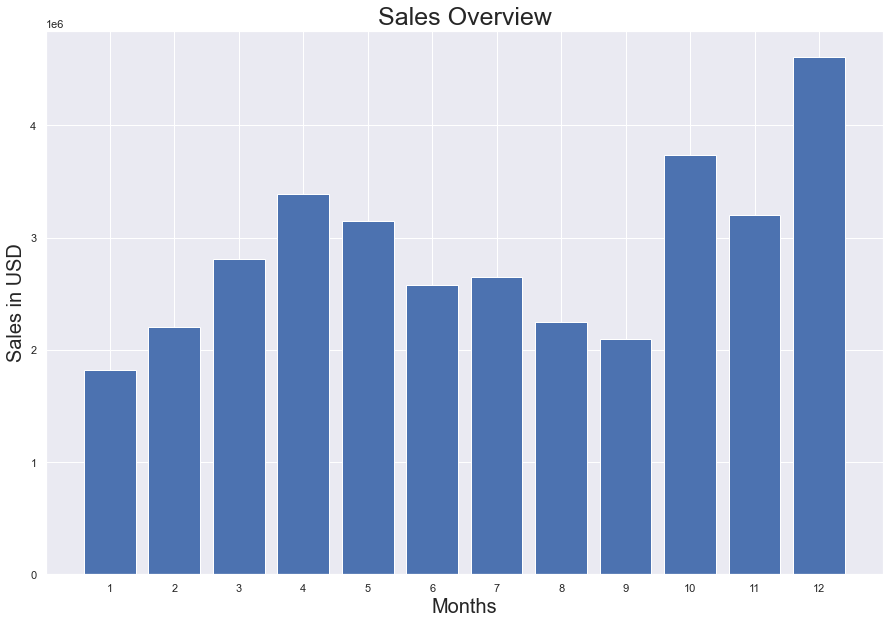
**Sales Analysis Report**

* Questions tackled:

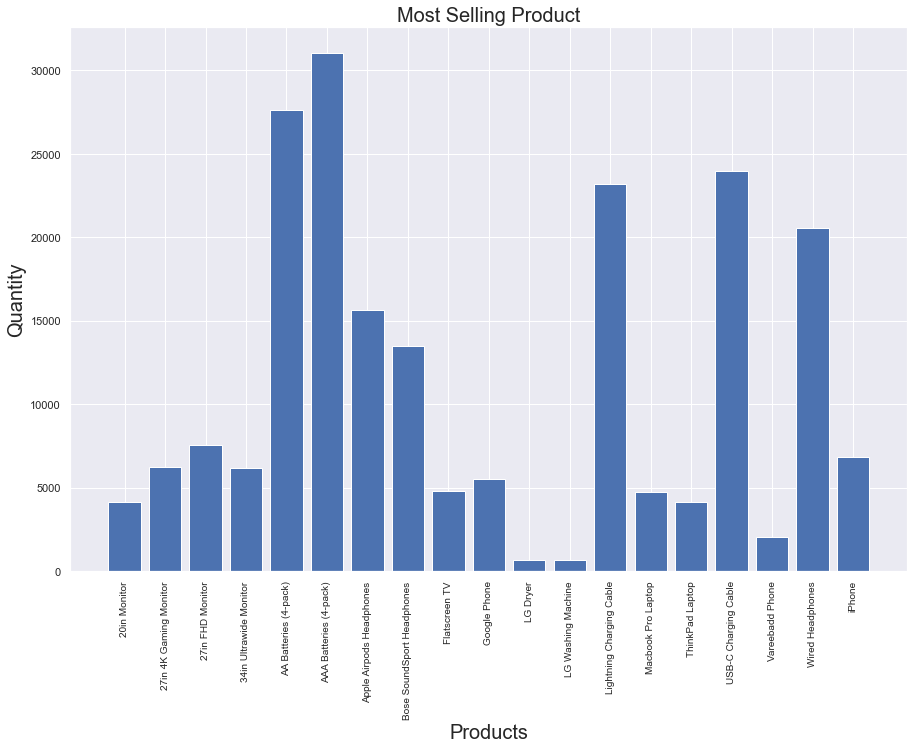
1. What was the best month for sales? How much was earned in that month?

**Ans: December was the best month for sales. Earning is 4613443.34 USD.**

****

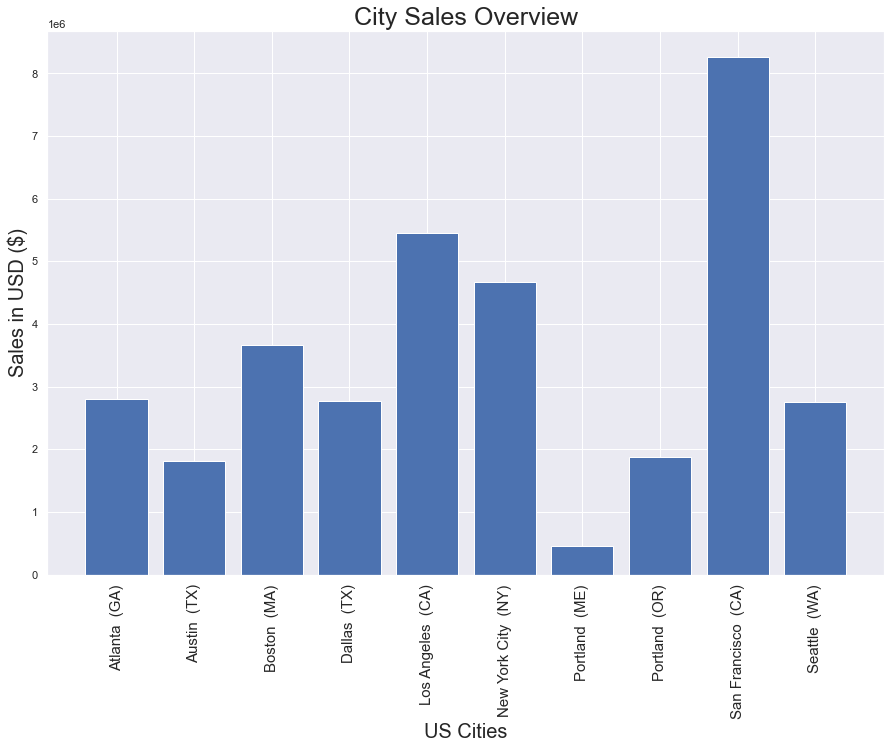
1. What product sold the most?

**ANS: AAA and AA batteries (4-Pack) are the most sold products because their price is comparatively very less than others.**

****

1. What city sold the most product?

**ANS: San Francisco has sold most products with sales of 8262203.91 USD.**

****

* Problems faced:

1. **There were chunks of data like, different months data in different csv files.**

**Sol:** Solved by adding each csv in different variable and made list of the variable and concinnated them using concat () function of pandas.

1. **There was no individual column for month.**

**Sol:** As the format of the Order Date was MM/DD/YYYY I first converted the Order Date column into string so that I can use string slicing and through that I sliced only MM by

data['Month'] = data ['Order Date'].str [0:2] and made a column with months.

1. **Month’s column was in string format.**

**Sol:** In order to calculate the months, we need to convert it into an integer, but as it was in string format, I used type casting from str to int by pandas’ function called astype ()

1. **Found that there were NANs in the data.**

**Sol:** Used the function isna () and any () to find out where exactly the NANs are and dropped all the data which contains NANs using dropna () function.

1. **Found that there were ‘Or’ in the Months Column.**

**Sol:** As now, the format of the Order Date was ‘Or’ I used string slicing and through that I updated the columns which were not ‘Or’.

data = data [data['Order Date'].str[0:2] != ‘Or’].

1. **Found that the Quantity Ordered and Price Each Columns are of string type**

**Sol:** Converted them into integer using to\_numeric ()

1. **There was no Sales Column, adding one.**

**Sol:** Sales is product of Quantity Ordered and Price Each.

data['Sales'] = data['Quantity Ordered'] \* data['Price Each']

1. **Need every data calculated for creation of Sales column.**

**Sol:** Used groupby () to solve this problem.

1. **According to the question we need to create City column.**

**Sol:** Made City column using apply ()