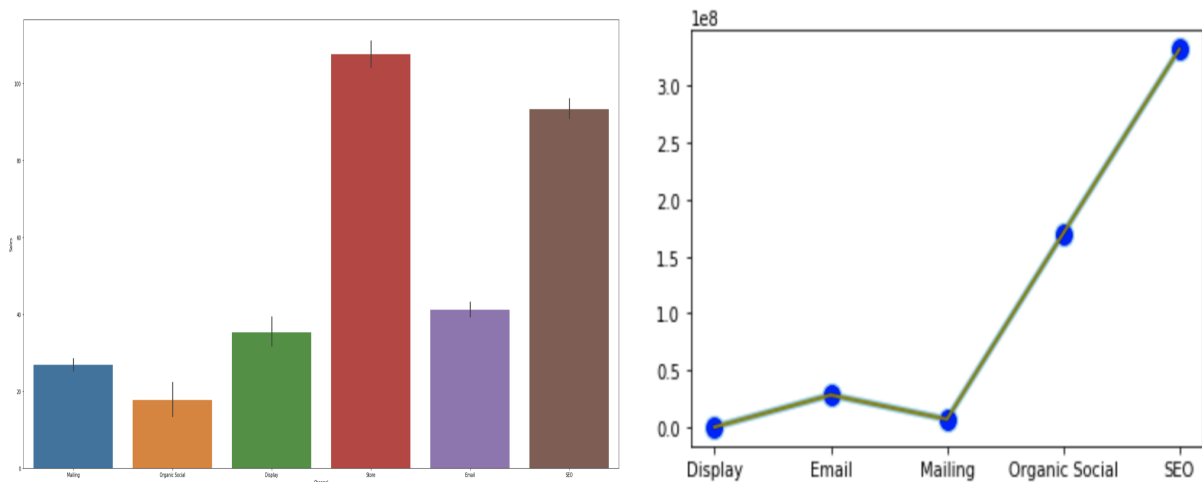


# Write Up

To calculate the non-refunded orders I have used the sales that are less than zero grouped by CustomerID and SalesOrder to depict the orders that are non-refunded.

I have applied both Linear Regression and LSTM to create the sales forecast and to check which models perform better for our dataset. For the Linear Regression model, we calculated Adjusted R-squared to check our model accuracy or how our model is performing better. We get the Adjusted R-squared as 62% for our model. Another model that I used is LSTM. Initially, I calculated the predictions on the  $y_{test}$  set which was approximately at 91%. Based on my evaluation I am choosing LSTM for my dataset as the prediction for this model is 91% which is higher than linear regression. LSTM is capable of remembering the past information and while predicting the future values, it takes this past information into account. This is important in our case because the previous sales data is crucial in predicting its future sales.



The first graph shows us the trends between the Channel and the Sales. We can observe from the graph that the Store channel has the highest number of sales followed by SEO which is the second largest generator for the sales. Whereas, Organic Social has the lowest number of sales which is little surprising. The second graph shows the graph between the actual sales and the predicted sales with respect to the channel. Using LSTM our model predicted very good predictions as in the graph we can clearly see the actual sales(blue line) and the predicted sales(green) are overlapping and sales are potentially increasing with time.

Sales forecasting is critical, because it helps us to find out how much income we can expect to generate in a certain period of time. The information in sales forecasts can be used to make intelligent decisions related to your sales process.

## Recommendations to potentially increase sales during your forecasted period:

- Keep track of the External factors like the Market and Competitor changes, Seasonality etc.
- We should not only focus on our highest sales generated channels but should also focus on the other channels for example: Organic Social which has a very high potential of increasing sales which is the lowest sales generated channel.
- Consistently track and clean the data and use some forecasting methods like multivariable analysis forecasting.