

# Comoto Data Analyst Technical Prompt

## Section 1

### Data Cleaning

I de-duplicated the data and removed sales with a Quantity or Sale value of 0. I also confirmed the Sales values were equal to Quantity times Price. There wasn't much to be done beyond that, so far as I could tell.

## Section 2

### 90 Day Refund Window

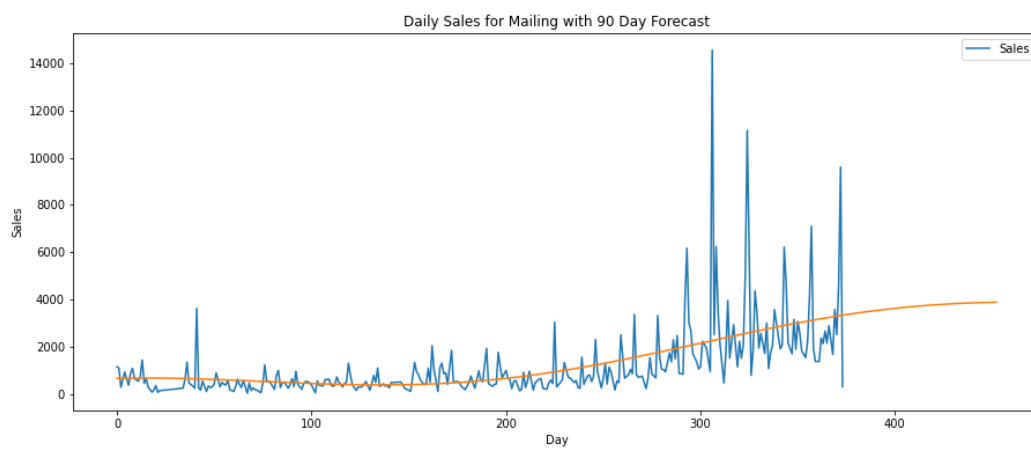
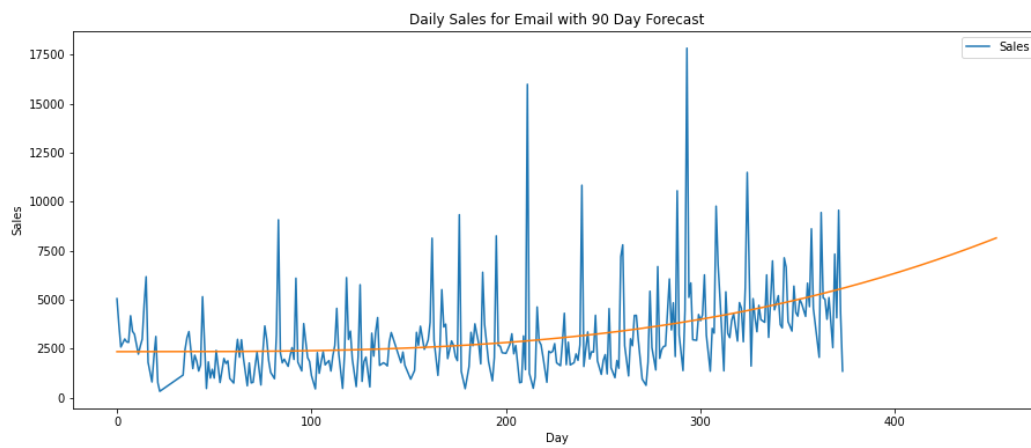
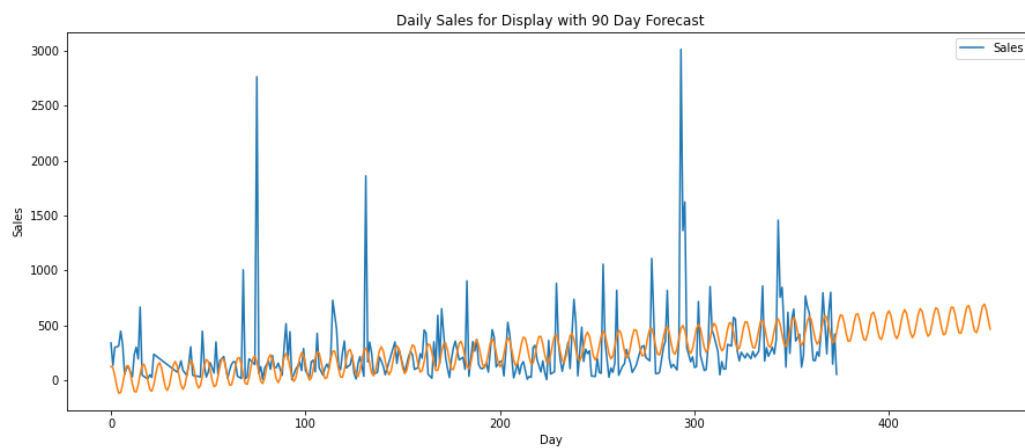
Refund SalesOrder values were not related to the SalesOrder value for the order they were issued for. Thus, it was impossible to link refunds to the exact order they were made on. Instead, I use a simple approximation: I subtracted the number of refunds a customer made during a 90 day period from the number of orders they made in the same window.

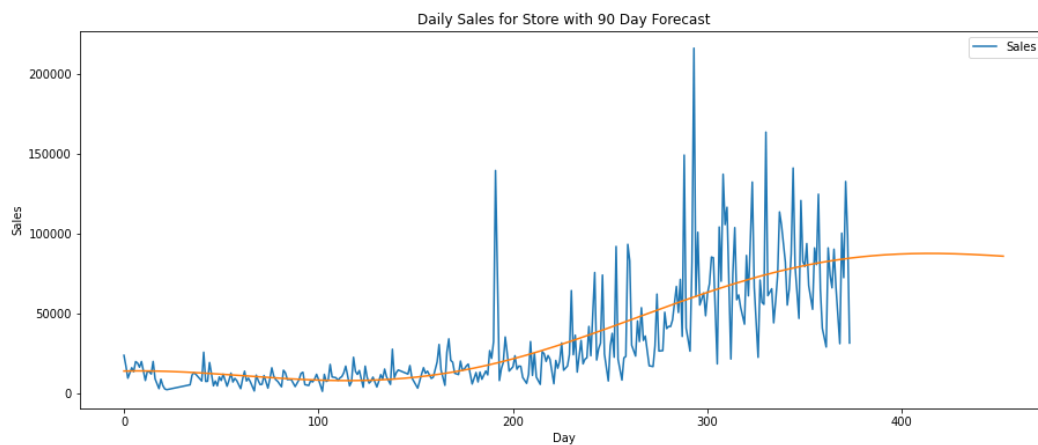
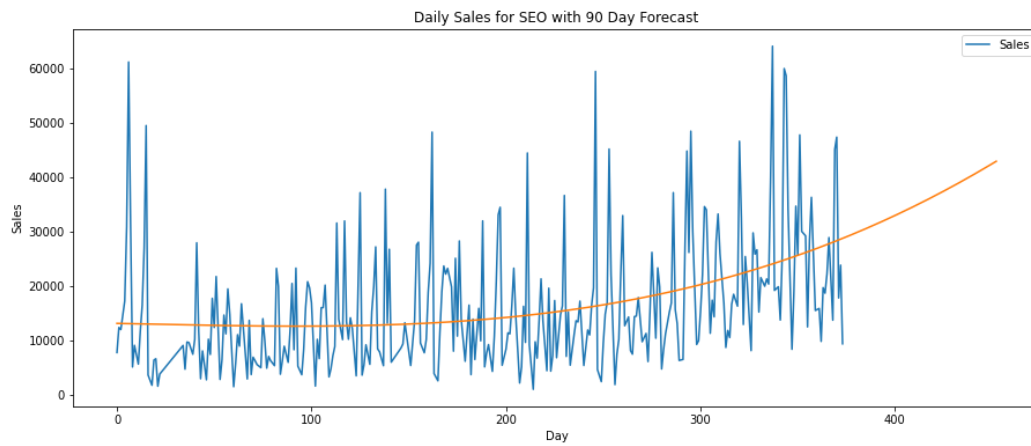
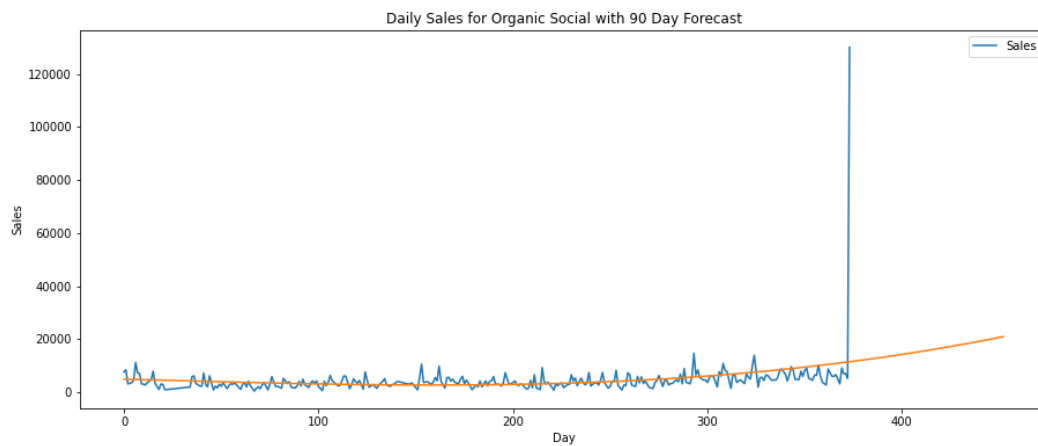
## Section 3

### Forecasting

Given the limited nature of the dataset and its relatively small time window (less than 1 years' worth of data isn't enough to find an accurate estimate for the seasonal effect), I took a simple approach and fit a sinusoidal curve with a linear trend that reasonably projected sales for certain channels over the 90 day forecast window. I could have used a more complex model like SARIMAX for the forecast, but I felt there wasn't enough data to justify it.

I used non-linear least squares to fit the curves for each channel, and I used RMSE in conjunction with a simple grid search to evaluate and select the best fit for each curve. The resulting forecasts:



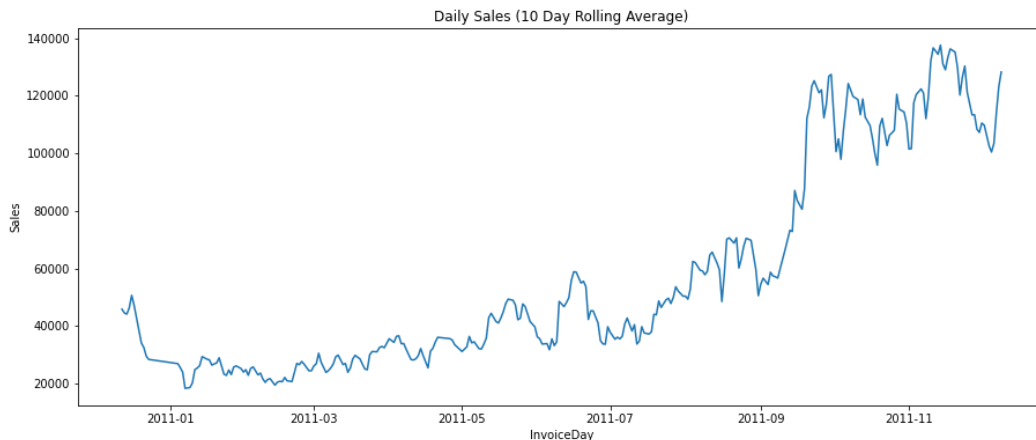


Though the final models don't all reflect the seasonal trend as accurately as I'd hoped, the better fitting ones show that the approach has promise and that the models could improve given more tuning.

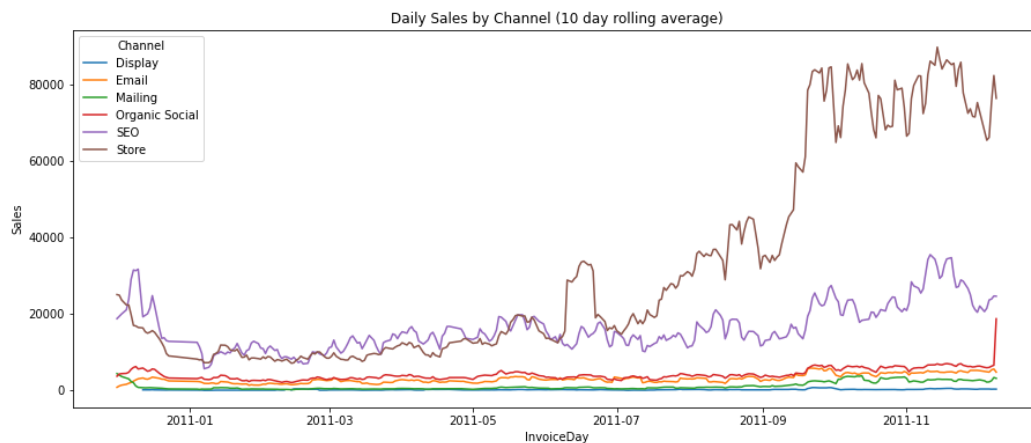
Aggregating the forecast results shows a 650% increase in total year-over-year sales for the 90 day window. However, I think that's optimistic, and I would have expected to see something in the 300-500% range.

## Section 4

Sales for the year have grown considerably, and sales continue to trend upwards. Year over year growth for the first week of December is up 294%, and the 10 day rolling average for total daily sales shows a clear upwards trajectory.



Breaking sales down by channel, we can see that this growth has largely been driven by growth in sales from stores.



Stores accounted for 66.15% of total sales in Q4, up from 37% in Q1. While growth in other channels remained modest, it's notable that mailing sales also saw a large increase over the year, increasing from 1.96% of total sales in Q1 to 2.63% of total sales in Q4.

While sales for Q1 of 2012 are forecasted to be up 779% over sales in Q1, this forecast is rosy and we should expect to see actual growth somewhere in the range of 300-500%.

It would be well worth examining what drove the explosion in store sales, as would be investigating the recent success of mail sales.

## Section 5

### Recommendations to Business Leaders

We should not expect sales to continue to increase during the forecasted period for Q1 of 2012, as we should see a seasonal drop off throughout the winter and early spring. However, to mitigate that drop off, I recommend business leaders do the following:

1. Focus on increasing the number of big spenders. The top 5% of customers account for 50% of sales, and attracting more customers like them will help boost overall sales numbers.
2. Focus on growing SEO. It's the second biggest sales channel, accounting for 28.4% of total sales, but it saw one of the lowest rates of growth throughout the year.