Mkt Analytics Final Project

Summer Lieberman
Jessica Maffe
Owen Fernandez
James Santagati

Intro To Data

Data: *Superstore Sales*

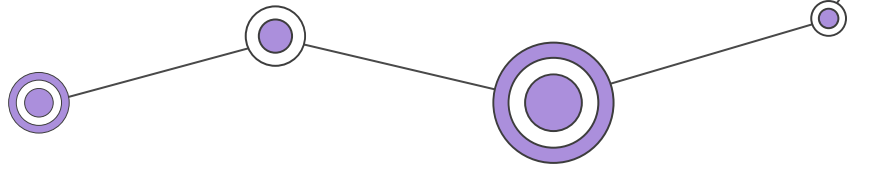**Machine Learning Models Used:**
- *A/B Testing*
- *Forecasting*

```
df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 9800 entries, 0 to 9799
Data columns (total 18 columns):
 #   Column         Non-Null Count   Dtype
---  ------         --------------   -----
 0   Row_ID         9800 non-null    int64
 1   Order_ID       9800 non-null    object
 2   Order_Date     9800 non-null    object
 3   Ship_Date      9800 non-null    object
 4   Ship_Mode      9800 non-null    object
 5   Customer_ID    9800 non-null    object
 6   Customer_Name  9800 non-null    object
 7   Segment        9800 non-null    object
 8   Country        9800 non-null    object
 9   City           9800 non-null    object
 10  State          9800 non-null    object
 11  Postal_Code    9789 non-null    float64
 12  Region         9800 non-null    object
 13  Product_ID     9800 non-null    object
 14  Category       9800 non-null    object
 15  Sub_Category   9800 non-null    object
 16  Product_Name   9800 non-null    object
 17  Sales          9800 non-null    float64
dtypes: float64(2), int64(1), object(15)
memory usage: 1.3+ MB
```

Python A/B Testing

- Segment V. Segment had no significance
- A/B Testing for segment: Consumer
  - Greatest significance state: New York
  - Least significant state: Illinois
- A/B Testing for segment: Corporate
  - Greatest significance state: Indiana
  - Least significance state: Texas
- A/B Testing for segment: Home office
  - Greatest significance state: Florida

```python
segments = df['Segment'].unique()

for segment in segments:
    print(f"\nA/B Testing for Segment: {segment}")
    segment_data = df[df['Segment'] == segment]
    states = segment_data['State'].unique()

    greatest_significance_state_greater = None
    greatest_significance_p_value_greater = 1
    least_significant_state_lesser = None
    least_significant_p_value_lesser = 1

    p_values_greater = []
    p_values_lesser = []
```

Tableau Initial State Visual By Segment

Compared Avg. Sales do not appear to match up with python

- Consumer: West Virginia
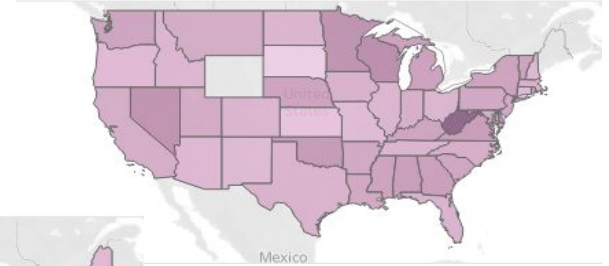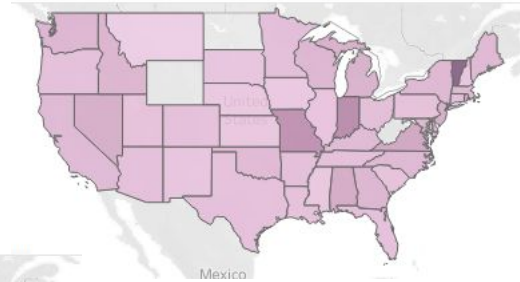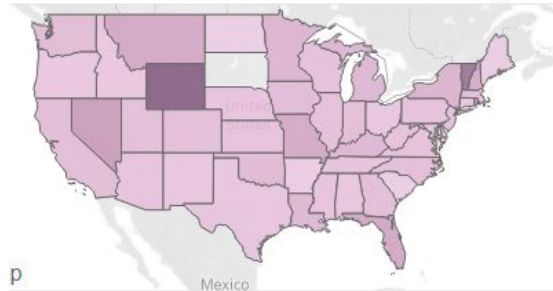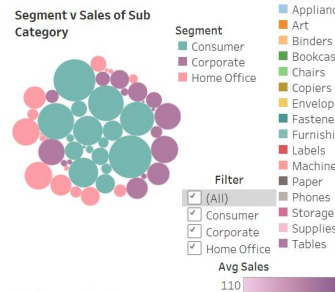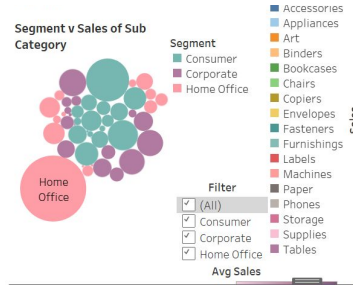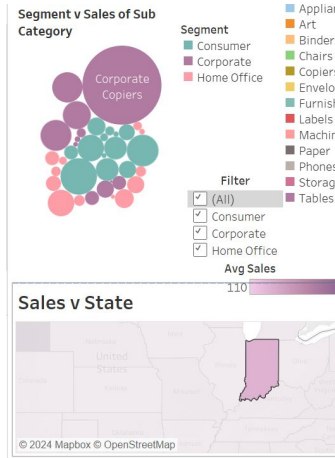- Corporate: Vermont
- Home Office: Wyoming

Tableau v Python State Comparison

Segment v Sales of Sub Category

Segment
- Consumer
- Corporate
- Home Office

Corporate Copiers

Filter
- ☑ (All)
- ☑ Consumer
- ☑ Corporate
- ☑ Home Office

Avg Sales
110

Sales v State

© 2024 Mapbox © OpenStreetMap

Segment v Sales of Sub Category

Segment
- Consumer
- Corporate
- Home Office

Home Office

Accessories, Appliances, Art, Binders, Bookcases, Chairs, Copiers, Envelopes, Fasteners, Furnishings, Labels, Machines, Paper, Phones, Storage, Supplies, Tables

Filter
- ☑ (All)
- ☑ Consumer
- ☑ Corporate
- ☑ Home Office

Avg Sales
110

Sales v State

© 2024 Mapbox © OpenStreetMap

Segment v Sales of Sub Category

Segment
- Consumer
- Corporate
- Home Office

Appliances, Art, Binders, Bookcases, Chairs, Copiers, Envelopes, Fasteners, Furnishings, Labels, Machines, Paper, Phones, Storage, Supplies, Tables

Filter
- ☑ (All)
- ☑ Consumer
- ☑ Corporate
- ☑ Home Office

Avg Sales
110

Sales v State

© 2024 Mapbox © OpenStreetMap

- Segment does dominate the significant state
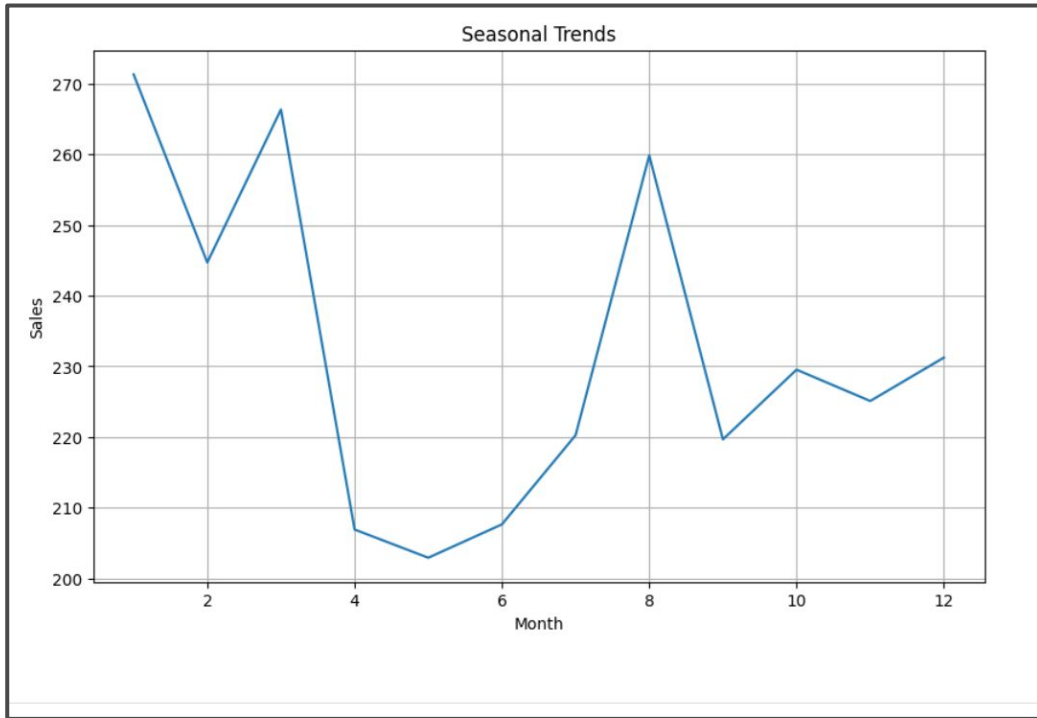- Potentially more significant sales

Python Forecast
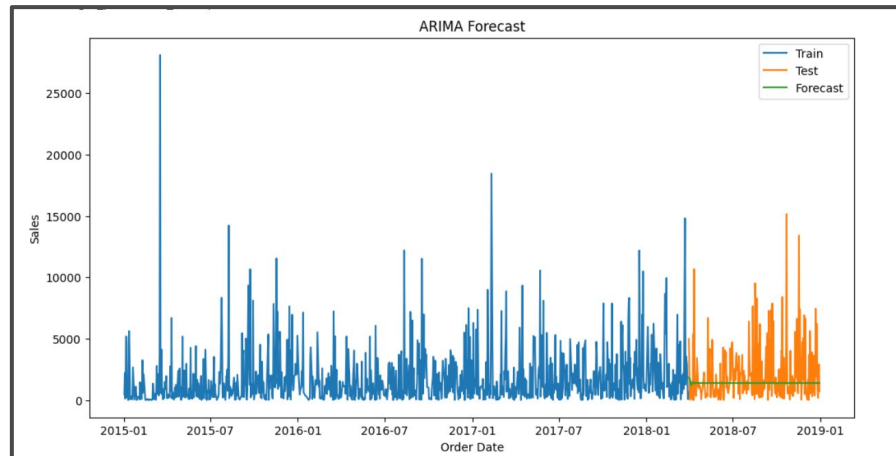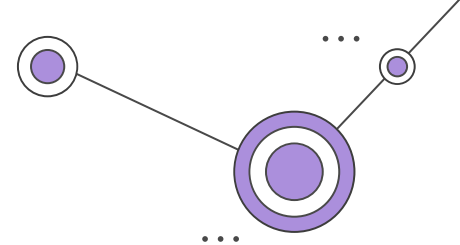
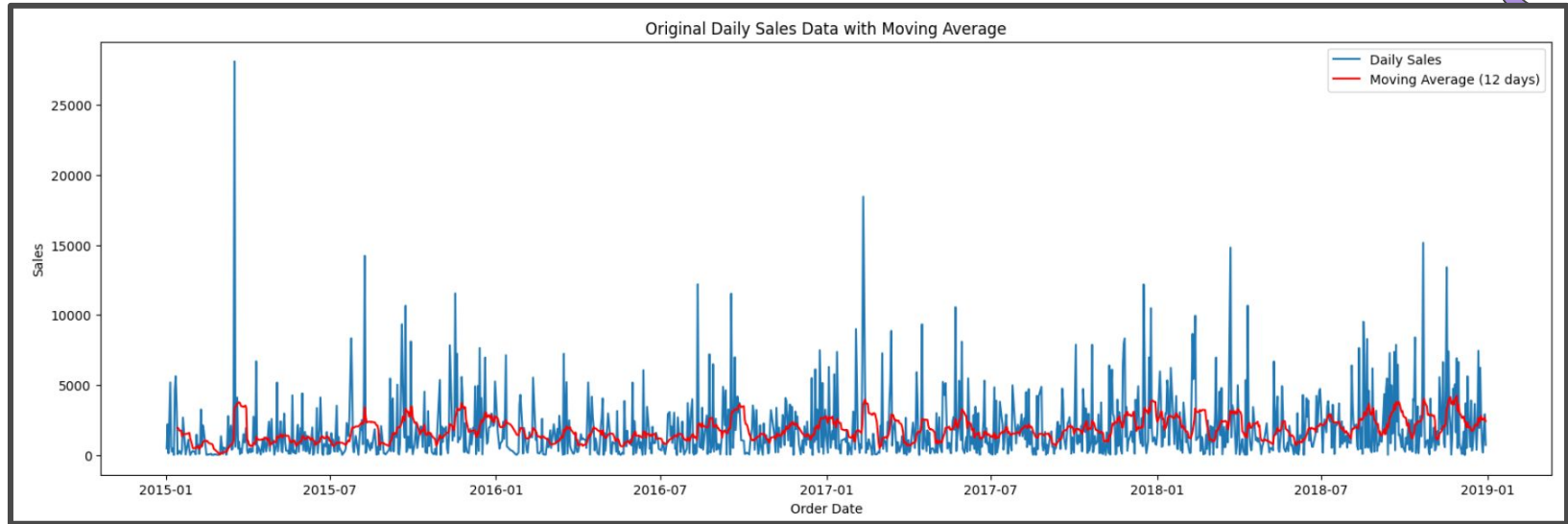Original code we ran: could not interpret much
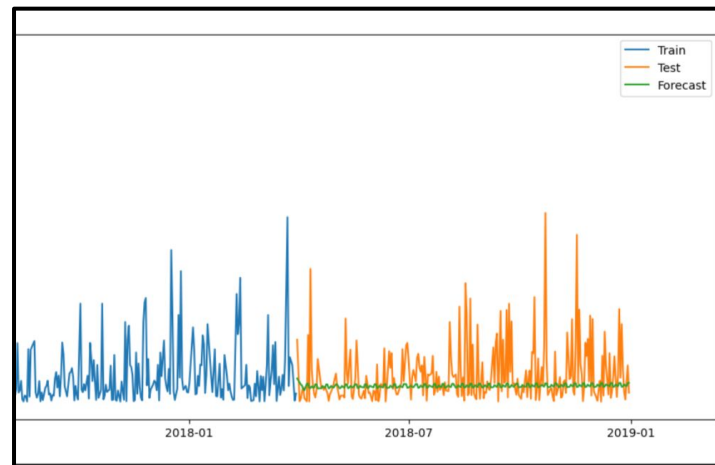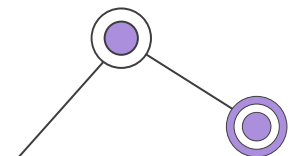
Refined code: tiny spikes in the green line show an increase in sales

Python Seasonality

Ran a code to determine any seasonality
Results: Sales were highest in January, March, and August
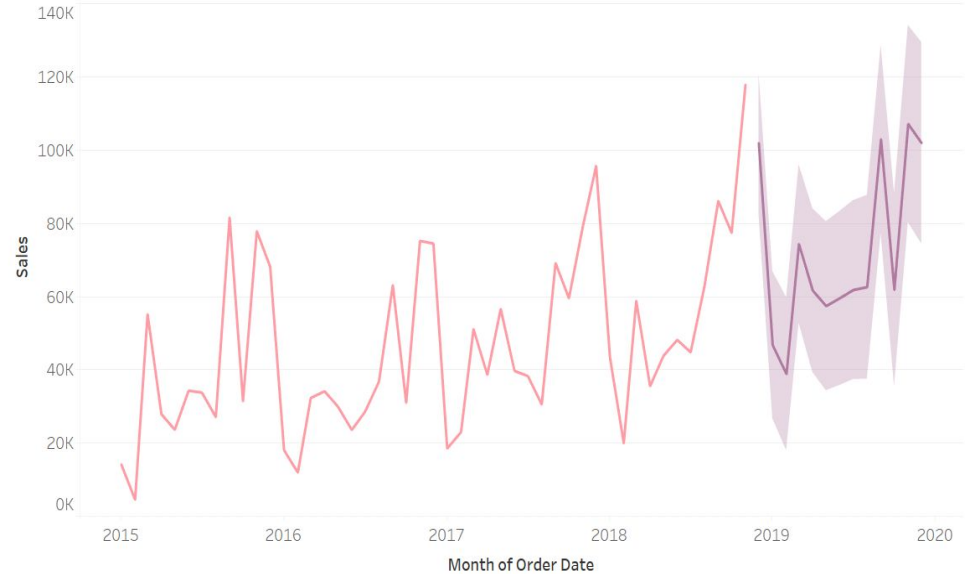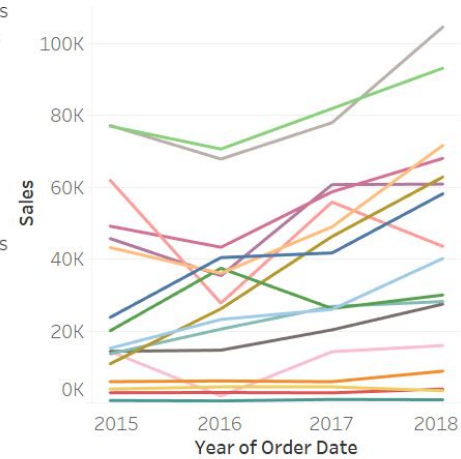Shows a similar pattern of seasonality

Python Moving Average

Original Daily Sales Data with Moving Average

Ran a code to determine moving average, which shows a similar pattern of seasonality

Tableau Forecast

- The forecasting results of our Tableau matched our Python findings
- In both you can find spikes in sales during similar seasons/months

Sales Over Time W/ Forecast
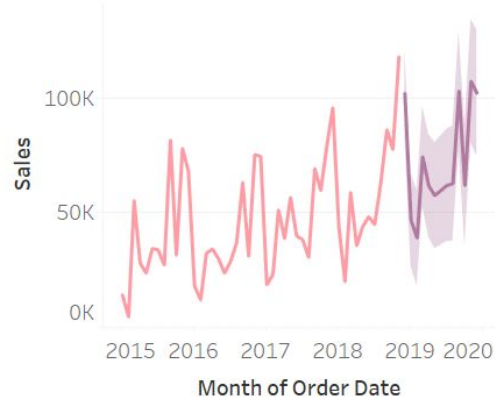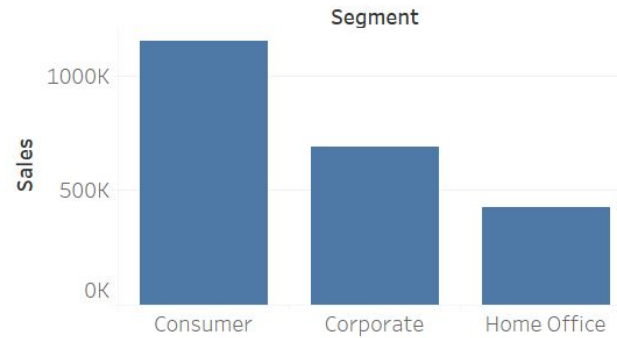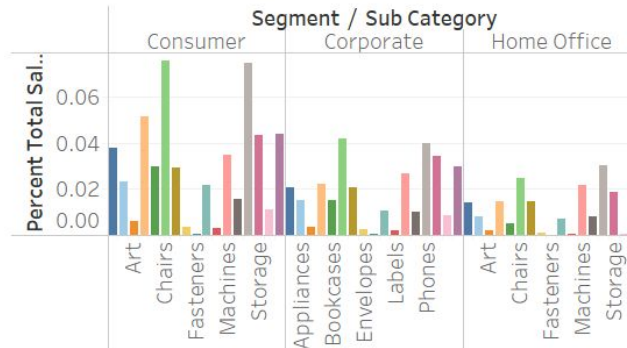
Dashboard 1

Dashboard 2

Dashboard 3
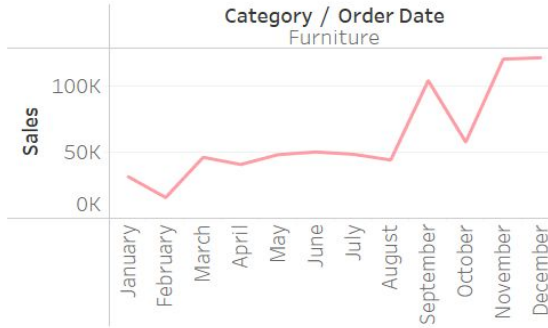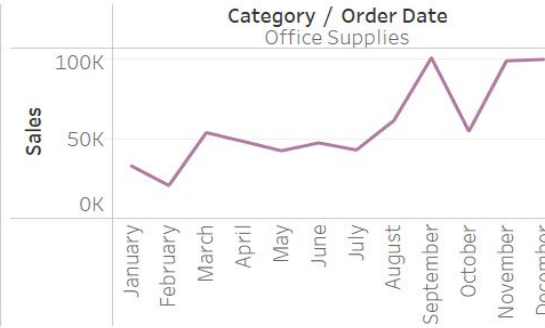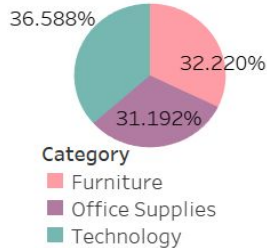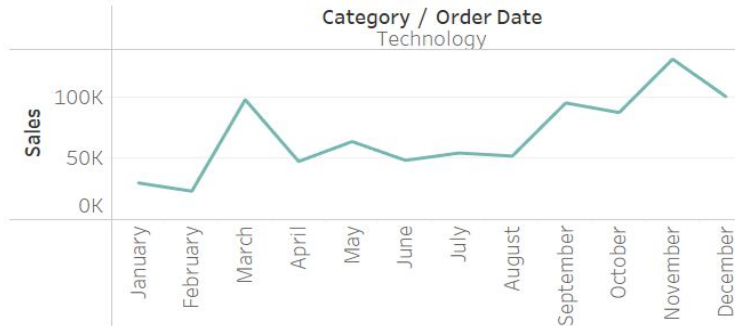
Recommendations

- Allocate resources and budget to both results of A/B testing and states with more average sales
  - Develop individual strategy based on best segment and product
- Seasonal trends big in Fall and Winter requires more production and resources
  - Back to school messaging for furniture & office supplies
  - Research why March is such a big month
- Adjust strategies to focus more on phones and chairs based on product