Data Analysis Report for Assortment Office

DATE:	Apr 2, 2024
PRESENTED BY:	Jieda(Jay) Wang

Background

The Assortment Office faces a significant challenge in determining the optimal production/order quantities for each article (a specific product in a particular color) across different markets. Traditionally, this process involves manual estimations by buyers, which is time-consuming. To improve efficiency and accuracy, two machine learning models were developed to predict future sales in pieces per article, per market, on a weekly basis. Additionally, these models provide order suggestions to balance the risk between stockout and excess inventory at the season's end.

Objective

The primary objective of this analysis is twofold:

1. Analyzing Sales Predictions (Part 1): Evaluate the prediction accuracy of the two machine learning models to identify which model provides the best sales forecasts.

2. Analyzing Order Suggestions (Part 2): Analyze the order suggestions from both models to determine which one optimally suggests production/order quantities, minimizing the risk of stockout and overstock.

Methodology

Data Preparation

Data from five sources including actual sales, article hierarchy, and predictions from both models A and B were merged and enriched to facilitate comprehensive analysis.

Part 1: Analyzing Sales Predictions

RMSE and MAE Analysis:

Both Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were calculated to evaluate the models' prediction accuracy. Lower values indicate better performance.

Comparative Analysis of Model A and Model B Performance

Description	RMSE	MAE
Model A	174.626039	96.666689
Model B	196.249974	102.154063

Market-wise Performance Breakdown of Models A and B

Market	RMSE Model A	RMSE Model B	MAE Model A	MAE Model B
BE	279.054868	373.341860	191.117320	259.447588
CA	55.112532	67.047296	34.476325	41.857244
DE	242.270873	291.961097	148.736934	182.535889
GB	101.137653	83.010060	65.450142	53.485806
JP	84.792202	78.778071	44.198259	39.306093
SE	80.715632	71.937141	39.070782	33.899947
CN	145.779804	125.792453	88.595966	73.207357
PL	118.673830	96.352336	72.636408	59.867587
US	267.976236	220.341626	184.911500	143.356500

Monthly Analysis of Model A and Model B Predictive Accuracy

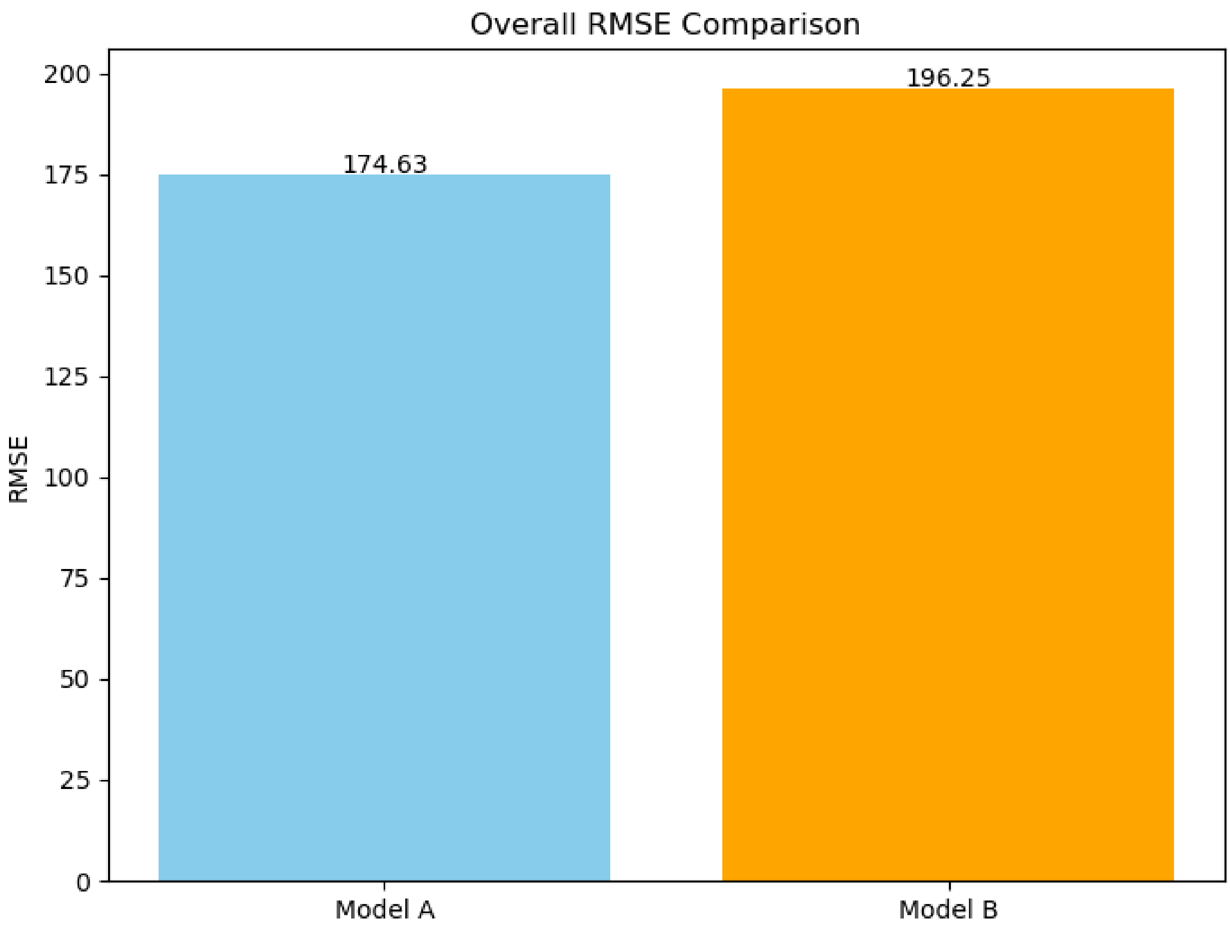
Year_Month	RMSE Model A	RMSE Model B	MAE Model A	MAE Model B
2018-03	32.460232	8.062258	31.666667	6.333333
2018-04	221.507366	422.193933	164.546667	259.070000
2018-05	286.655800	384.479973	190.527415	233.436554
2018-06	204.998534	228.347972	129.328735	137.103904
2018-07	232.860102	263.405037	133.399901	144.357910
2018-08	159.827763	170.335865	96.224310	97.892586
2018-09	124.196796	130.451439	75.945440	76.034039
2018-10	126.242023	133.258374	75.521791	76.116179
2018-11	146.025518	149.643340	75.021051	76.507915
2018-12	183.417230	184.785135	86.316150	87.412138
2019-01	191.358992	204.303978	80.192374	84.905835
2019-02	63.238843	71.300513	36.791667	41.600877
2019-03	54.687859	37.079708	30.190476	26.523810

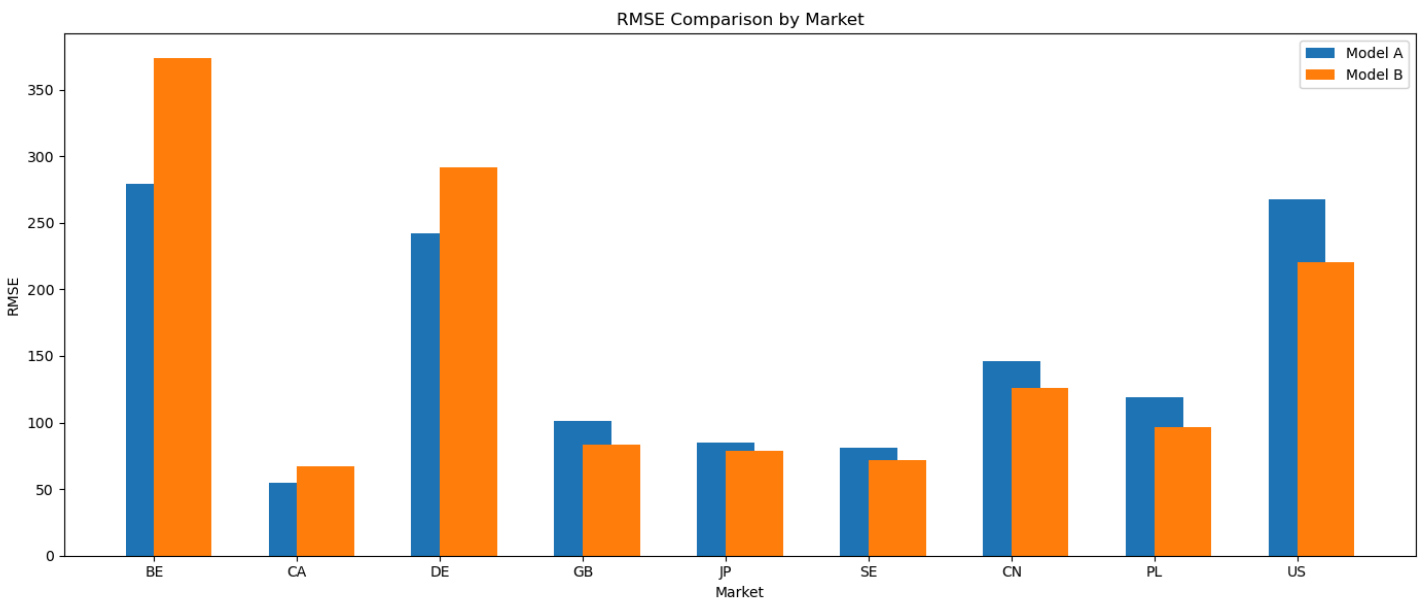
Departmental and Sectional Performance Evaluation of Models A and B

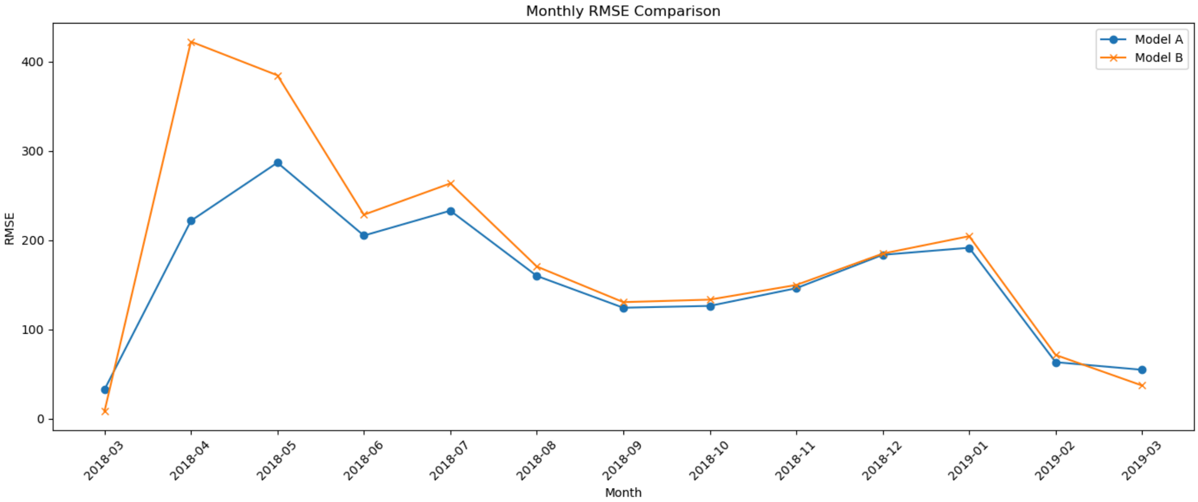
Department	Section	RMSE Model A	RMSE Model B	MAE Model A	MAE Model B
Department_B	Section_01	148.138697	148.628313	80.005353	78.271776
Department_C	Section_01	192.180225	217.859020	104.525802	112.056769
Department_D	Section_02	166.776208	179.731461	101.171794	103.967438
Department_E	Section_02	142.205458	161.149090	84.815238	89.116835
Department_F	Section_03	285.438353	329.859473	186.105363	198.726151
Department_G	Section_03	107.031939	123.172067	58.991628	63.910903

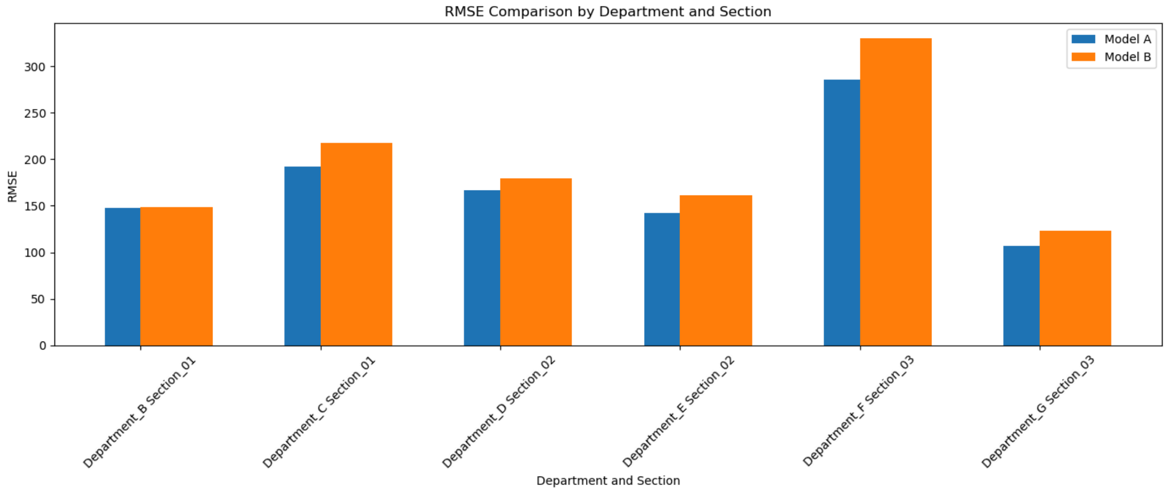
Visualizations:

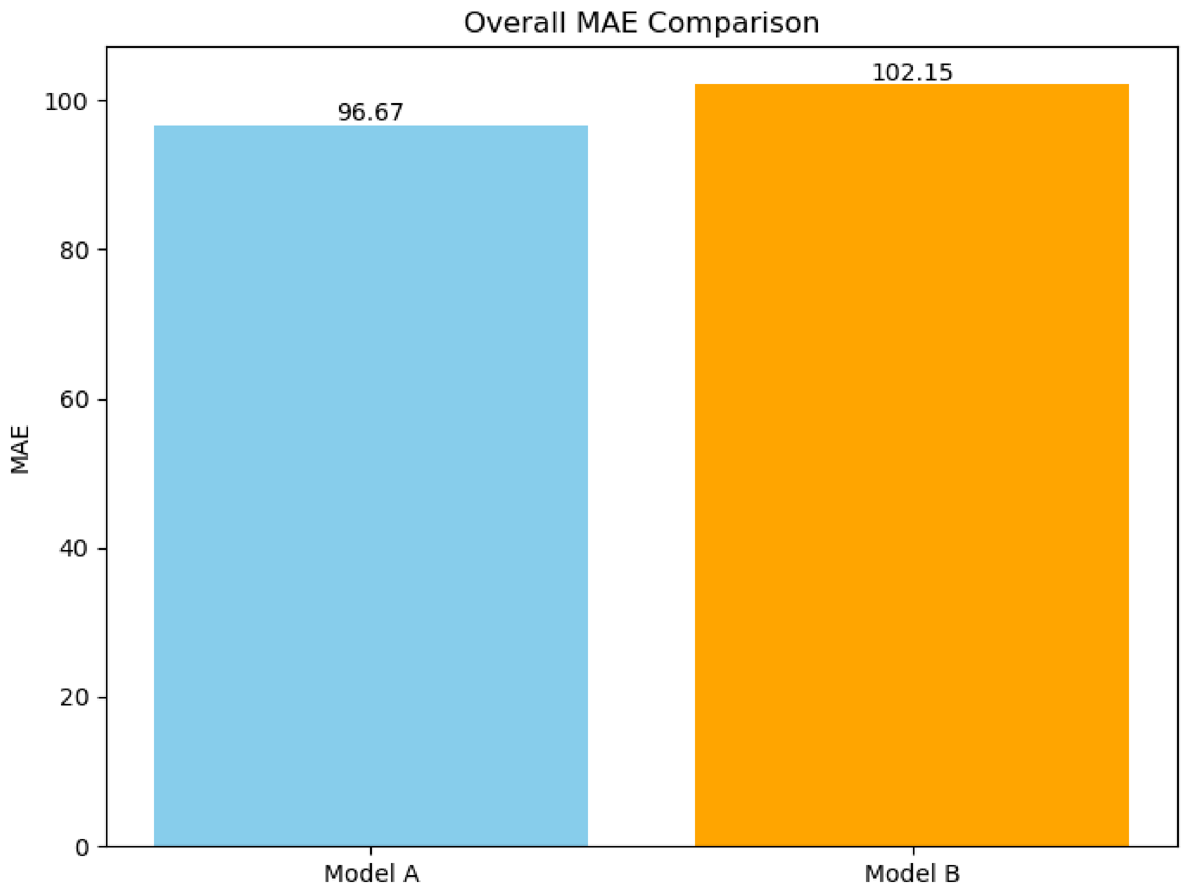
Histograms and line plots were used to compare RMSE and MAE values across markets, months, and department/sections for both models.

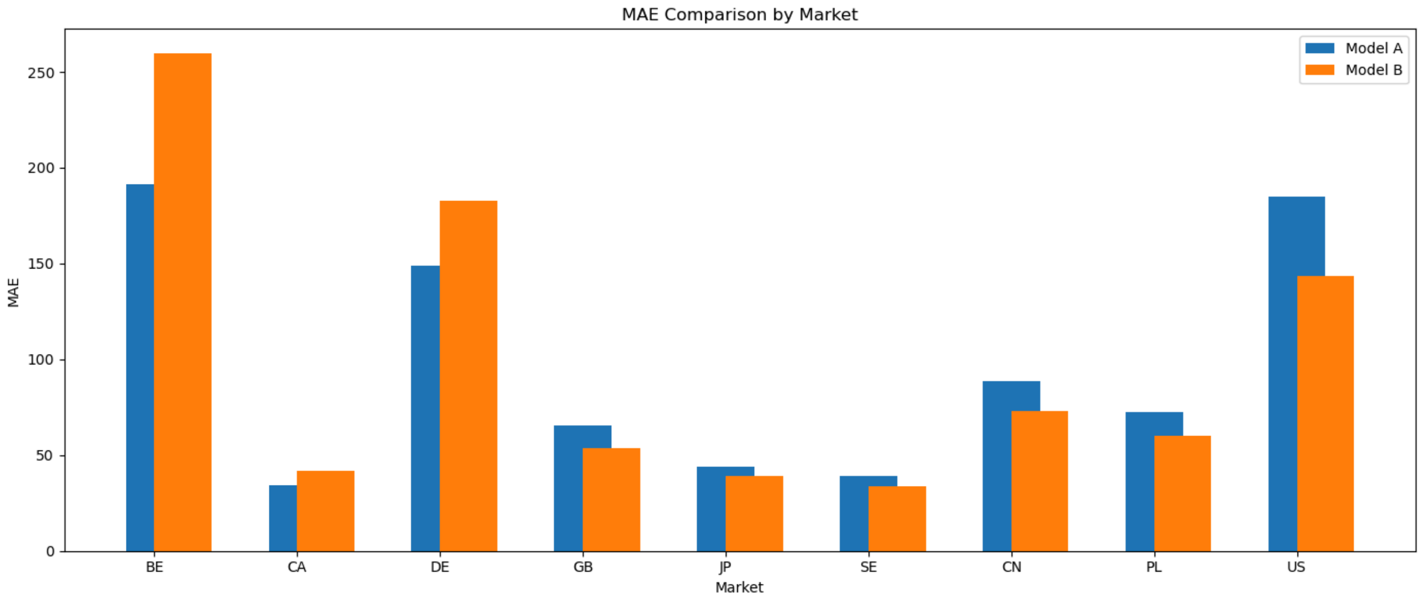


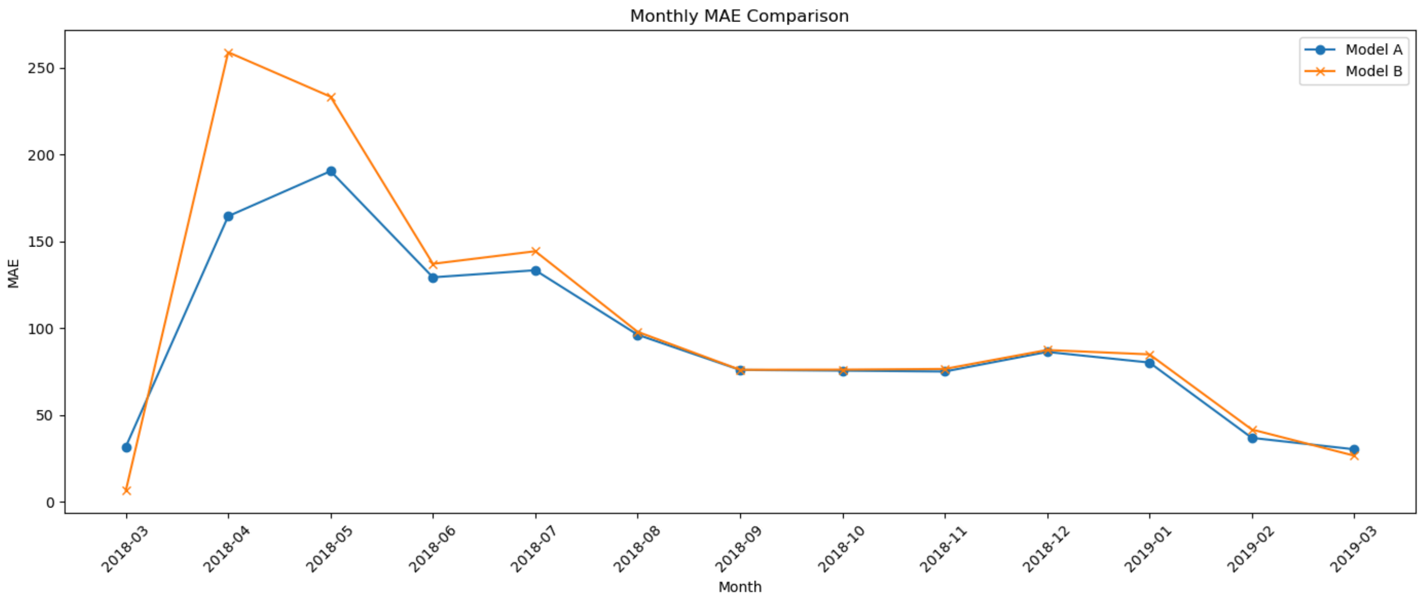


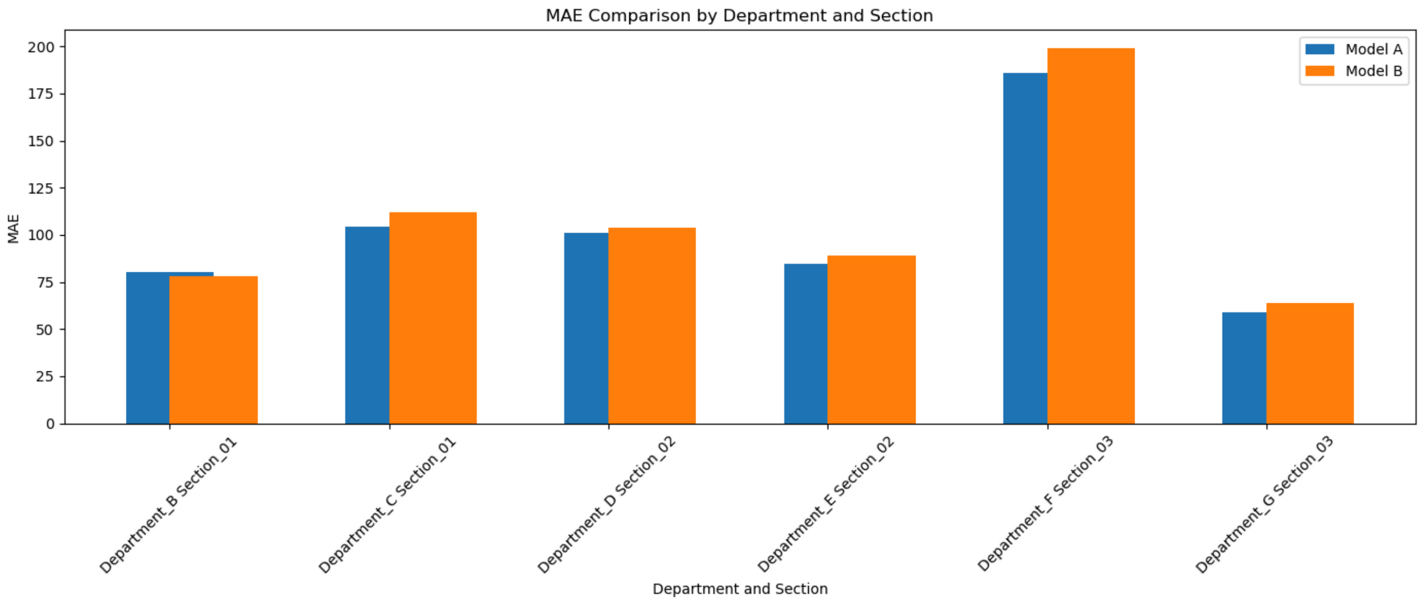

Monthly RMSE Comparison
Model A
Model B
400
300
200
100
0
RMSE
2018-03
2018-04
2018-05
2018-06
2018-07
2018-08
2018-09
2018-10
2018-11
2018-12
2019-01
2019-02
2019-03
Month


RMSE Comparison by Department and Section
Model A
Model B
300
250
200
150
100
50
0
RMSE
Department_B Section_01
Department_C Section_01
Department_D Section_02
Department_E Section_02
Department_F Section_03
Department_G Section_03
Department and Section


Overall MAE Comparison
96.67
102.15
100
80
60
40
20
0
MAE
Model A
Model B


MAE Comparison by Market
Model A
Model B
MAE
Market
BE
CA
DE
GB
JP
SE
CN
PL
US
0
50
100
150
200
250


Monthly MAE Comparison
Model A
Model B
MAE
Month
2018-03
2018-04
2018-05
2018-06
2018-07
2018-08
2018-09
2018-10
2018-11
2018-12
2019-01
2019-02
2019-03
0
50
100
150
200
250


MAE Comparison by Department and Section
Model A
Model B
MAE
Department and Section
Department_B Section_01
Department_C Section_01
Department_D Section_02
Department_E Section_02
Department_F Section_03
Department_G Section_03
0
25
50
75
100
125
150
175
200

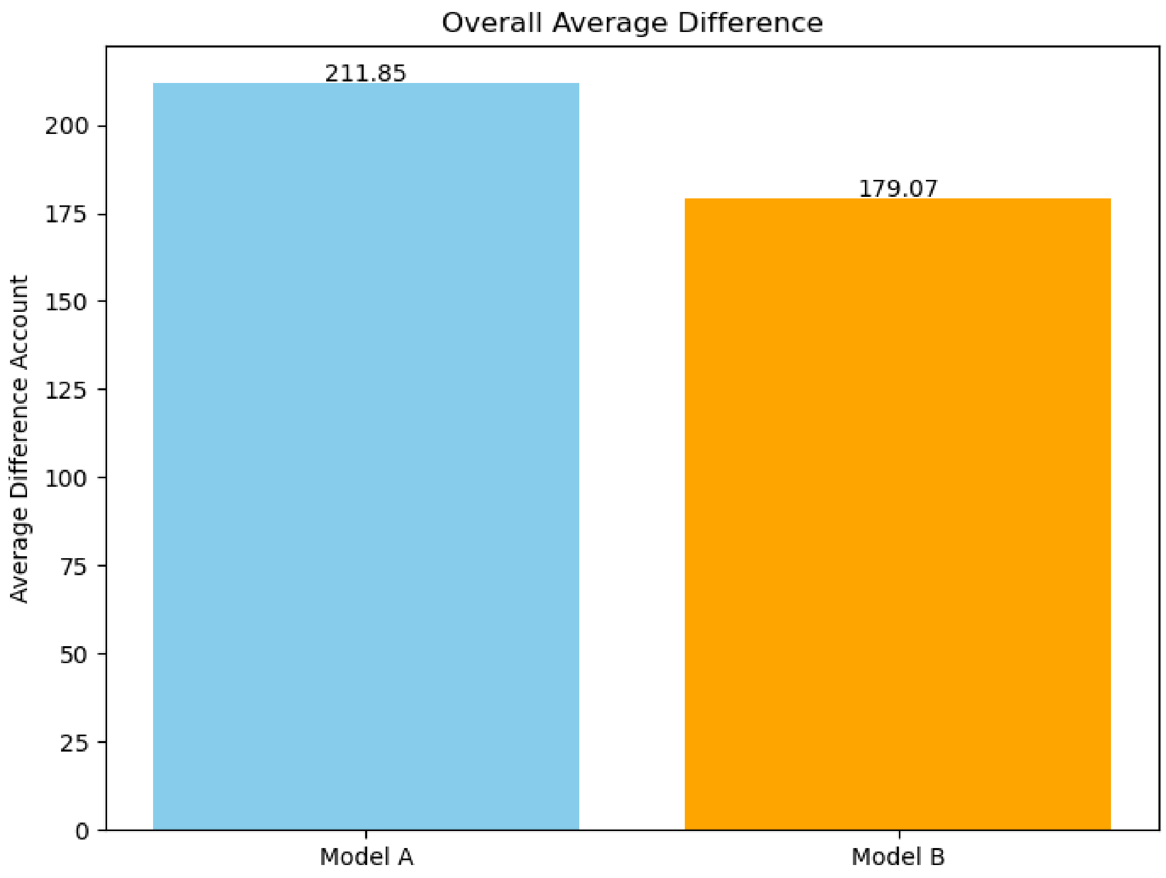
Part 2: Analyzing Order Suggestions

Order Suggestions vs. Actual Sales:

The absolute difference between suggested orders and actual sales was analyzed to assess each model's ability to match supply with demand.

Model A and Model B Average Difference

Description	Value
Model A average difference	211.84600549089012
Model B average difference	179.07303791436934

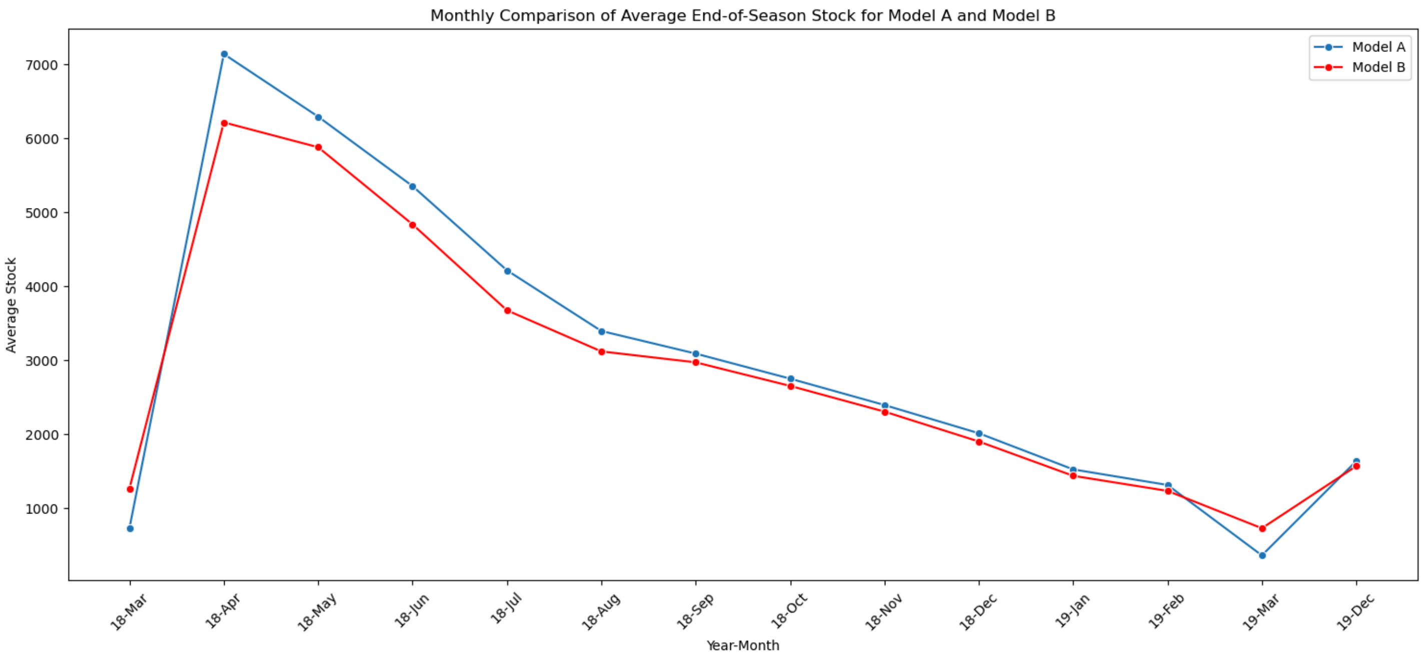


Stock and Profit Analysis:

The average end-of-season stock levels and cumulative profit for both models were calculated and compared monthly and quarterly. This analysis aimed at evaluating the models' effectiveness in balancing stock levels to avoid stockout and overstock.

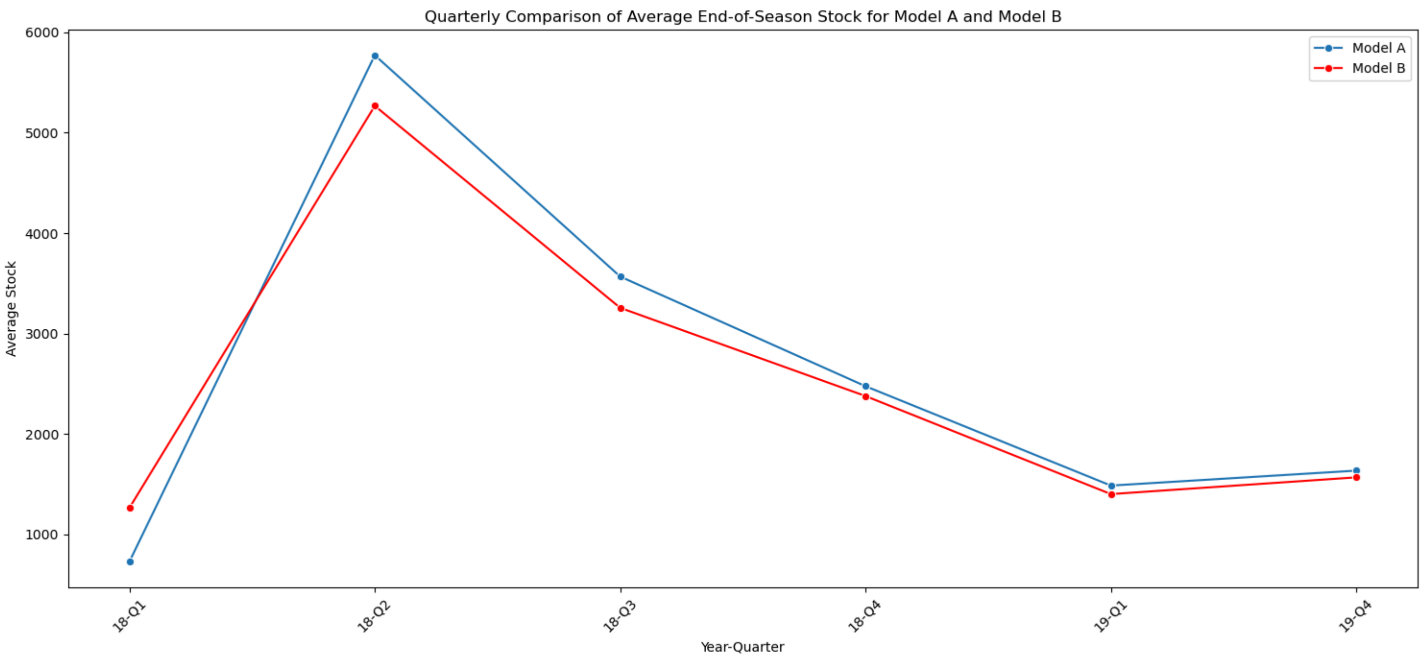
Monthly Comparison of Average End-of-Season Stock for Model A and Model B

Year	Month	Avg Stock A	Avg Stock B
18	3	729.333333	1264.333333
18	4	7143.226667	6218.460000
18	5	6295.532115	5882.026632
18	6	5355.949809	4837.972703
18	7	4216.612167	3675.809390
18	8	3395.899138	3119.462414
18	9	3089.795603	2971.510586
18	10	2750.977313	2653.004060
18	11	2393.949983	2305.252947
18	12	2012.172930	1900.753428
19	1	1522.946936	1436.304933
19	2	1312.260377	1232.271698
19	3	359.333333	726.666667
19	12	1634.978405	1567.945183



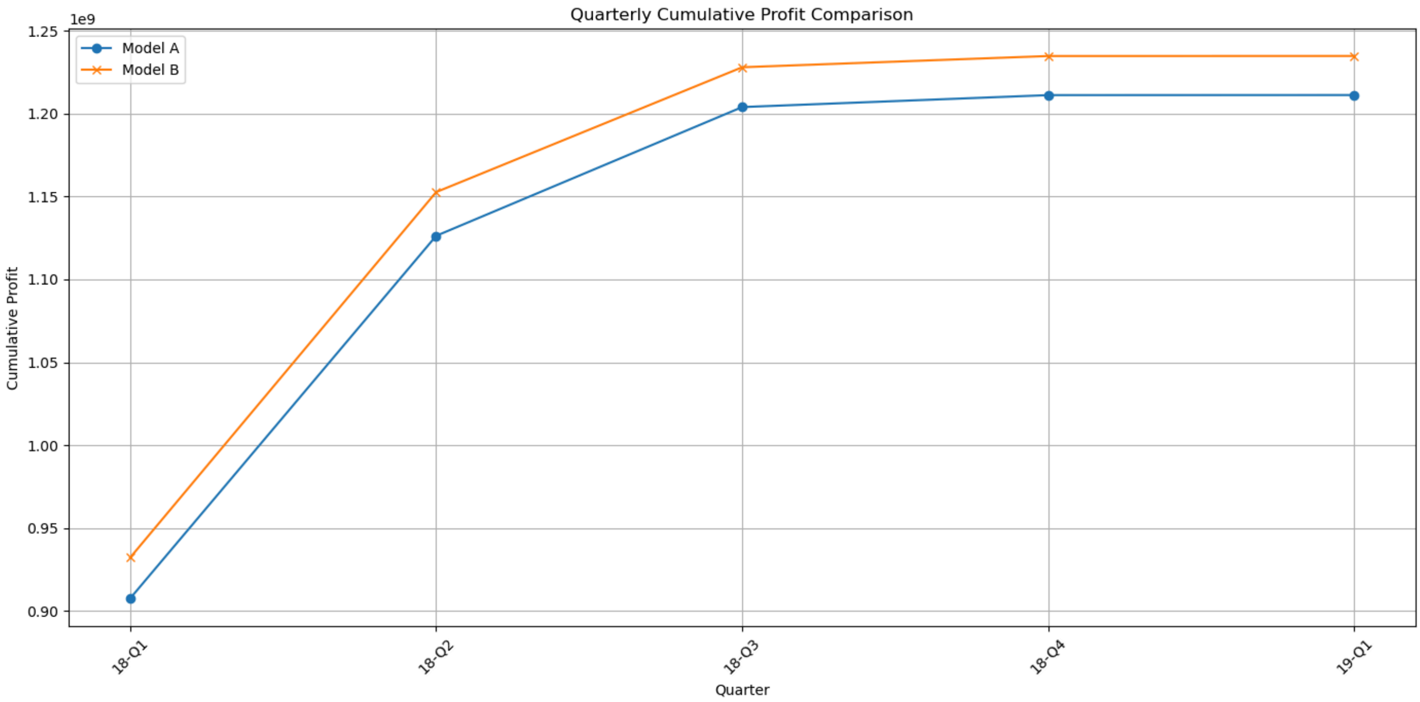
Quarterly Comparison of Average End-of-Season Stock for Model A and Model B

Year	Quarter	Avg Stock A	Avg Stock B
18	1	729.333333	1264.333333
18	2	5771.368374	5267.270011
18	3	3567.393796	3256.041359
18	4	2475.406663	2377.550162
19	1	1486.008717	1401.312578
19	4	1634.978405	1567.945183



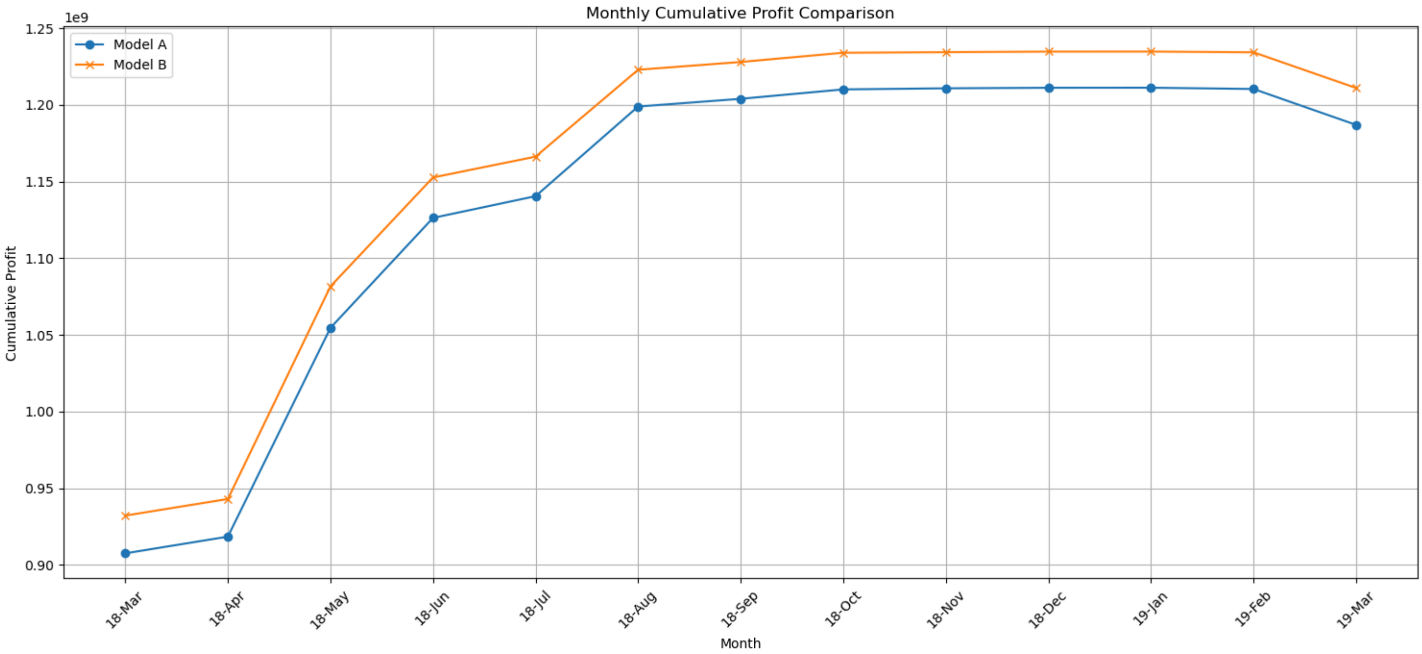
Quarterly Cumulative Profit Comparison

Quarter	Cumulative Profit A	Cumulative Profit B
18-Q1	9.076189e+08	9.322179e+08
18-Q2	1.126321e+09	1.152696e+09
18-Q3	1.203976e+09	1.227991e+09
18-Q4	1.211181e+09	1.234754e+09
19-Q1	1.211220e+09	1.234791e+09



Monthly Cumulative Profit Comparison

Month	Cumulative Profit A	Cumulative Profit B
18-Mar	9.076189e+08	9.322179e+08
18-Apr	9.184951e+08	9.430376e+08
18-May	1.054620e+09	1.081638e+09
18-Jun	1.126321e+09	1.152696e+09
18-Jul	1.140495e+09	1.166266e+09
18-Aug	1.198958e+09	1.222921e+09
18-Sep	1.203976e+09	1.227991e+09
18-Oct	1.210118e+09	1.233995e+09
18-Nov	1.210825e+09	1.234399e+09
18-Dec	1.211181e+09	1.234754e+09
19-Jan	1.211220e+09	1.234791e+09
19-Feb	1.210393e+09	1.234303e+09
19-Mar	1.186999e+09	1.211046e+09



Findings

Sales Predictions

Model A appears to have a competitive edge in prediction accuracy, which is critical for forecasting sales. Specific strengths of Model A include:

- Lower Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), which are critical metrics for forecasting accuracy. This suggests Model A has a better grasp of the data trends and less deviation from actual sales figures.
- In the monthly MAE comparison, Model A consistently shows closer predictions to actual sales, indicating a more stable prediction performance over time.
- Model A also excels in certain markets (for instance, Model A might perform particularly well in European markets such as Germany and Sweden, as indicated by lower MAE and RMSE in those regions).

Order Suggestions

On the other hand, Model B stands out when it comes to order suggestions with several strong points:

- Model B aligns more closely with actual sales, suggesting that its order suggestions may lead to a better match with customer demand.
- It maintains more optimal stock levels, as indicated by the lower average end-of-season stock. This can reduce costs associated with excess inventory and storage.
- Model B achieves higher cumulative profits, an indication that it balances stock availability against cost-effectiveness efficiently.

Recommendations

In light of the detailed analysis:

- **For sales predictions**: Firmly recommend Model A for its robust prediction accuracy, particularly in markets where it has demonstrated strength. This model will be beneficial for regions or products with a high cost of overestimating or underestimating demand.
- **For order suggestions**: Advise the adoption of Model B, which aligns closely with actual sales and maximizes profits by reducing stock-related issues. This model is especially valuable in markets where the cost of stockouts is high, or the product lifecycle is short, and overstock can lead to significant markdowns.

Integrated Approach

If possible, an integrated approach that leverages the strengths of both models would be ideal:

- Utilize Model A for its precise sales predictions to set targets and prepare the supply chain for upcoming demand.
- Employ Model B for its superior order suggestions to fine-tune the inventory levels, thus ensuring availability while maximizing profitability.

In scenarios where differentiated strategies are possible, tailoring the use of models by market specifics will optimize outcomes—applying Model A in markets where prediction accuracy has shown to drive better results and Model B in markets where inventory optimization impacts profits significantly.

Conclusion

The combination of Model A's accuracy in sales predictions and Model B's effectiveness in order suggestions can create a powerful toolset. This dual-model strategy aligns with business objectives by enhancing the precision of sales forecasts and optimizing order quantities, which in turn supports sustainability by reducing waste through better inventory management.