Data Challenge: Demand Forecasting

Background

The goPuff Operations team needs to staff locations appropriately to efficiently fulfill orders and keep our customers happy.

You are provided a subset of order data for some initial model building aimed at **forecasting orders for each of the next 14 days**. The plan is to present this model to another data scientist on your team and your manager. The main goal is to present your basic model and make recommendations for future improvements. Your recommendations for future work could consider both increasing model complexity and using additional data.

Instructions

Perform the analysis in a jupyter notebook (or similar) and add commentary that allows a reviewer to understand your approach, analytical observations, key decisions, and code. If you prefer to write-up your observations in a separate document (with embedded figures), that's OK as long as you provide code files as well.

Evaluation

Everyone

Your submission will be evaluated based on the following criteria:

1. Do the insights and recommendations make sense? Are they well supported by the analysis?
2. Does the candidate demonstrate knowledge of the business problem?
3. Did the candidate use appropriate techniques? Does the candidate have a strong understanding of the techniques employed?
4. Is the code organized, understandable, and correct?

The submission *will not* be evaluated based on the performance of the model, the aesthetic quality of visualizations, or the number of techniques employed.

Expect to spend ~3 hours on the challenge. We don't expect the analysis to be perfect.

Data

You are provided with one data file with order counts by location for a few obscured locations:

orders_by_location.csv

Data dictionary

Feature Name	***Explanation***
location_id	Unique identifier for each warehouse
location_name	Name of warehouse
date	date
order_count	Number of orders

Please note: location names have been changed, so no need to introduce location-specific features to the model