About Food Demand Forecasting Challenge

Demand forecasting is a key component to every growing online business. Without proper demand forecasting processes in place, it can be nearly impossible to have the right amount of stock on hand at any given time. A food delivery service has to deal with a lot of perishable raw materials which makes it all the more important for such a company to accurately forecast daily and weekly demand.

Too much inventory in the warehouse means more risk of wastage, and not enough could lead to out-of-stocks — and push customers to seek solutions from your competitors. In this challenge, get a taste of demand forecasting challenge using a real dataset.

Data Science Resources

- Are you a complete beginner? If yes, you can check out our latest 'Intro to Data Science' course to kickstart your journey in data science.
- Refer this comprehensive guide on Time Series Classification with its implementation in python.
- You can also refer this guide on Forecasting and Modeling that will be helpful in solving this problem.

Rules

- One person cannot participate with more than one user accounts.
- This is proprietary dataset, you can only use for this hackathon (Analytics Vidhya Datahack Platform) not for any other reuse
- You are free to use any tool and machine you have rightful access to.
- You can use any programming language or statistical software.
- You are free to use solution checker as many times as you want.

FAQs

1. Are there any prizes/AV Points for this contest?

This contest is purely for learning and practicing purpose and hence no participant is eligible for prize or AV points.

2. Can I share my approach/code?

Absolutely. You are encouraged to share your approach and code file with the community. There is even a facility at the leaderboard to share the link to your code/solution description.

3. I am facing a technical issue with the platform/have a doubt regarding the problem statement. Where can I get support?

Post your query on discussion forum at the thread for this problem, discussion threads are given at the bottom of this page. You could also join the AV slack channel by clicking on 'Join Slack Live Chat' button and ask your query at channel: practice_problems.

Registration Fee

Free

Problem Statement

Your client is a meal delivery company which operates in multiple cities. They have various fulfillment centers in these cities for dispatching meal orders to their customers. The client wants you to help these centers with demand forecasting for upcoming weeks so that these centers will plan the stock of raw materials accordingly.

The replenishment of majority of raw materials is done on weekly basis and since the raw material is perishable, the procurement planning is of utmost importance. Secondly, staffing of the centers is also one area wherein accurate demand forecasts are really helpful. Given the following information, the task is to predict the demand for the next 10 weeks (Weeks: 146-155) for the center-meal combinations in the test set:

- Historical data of demand for a product-center combination (Weeks: 1 to 145)
- Product(Meal) features such as category, sub-category, current price and discount
- Information for fulfillment center like center area, city information etc.

Data Dictionary

1. **Weekly Demand data (train.csv):** Contains the historical demand data for all centers, test.csv contains all the following features except the target variable

Variable	Definition
id	Unique ID
week	Week No
center_id	Unique ID for fulfillment center
meal_id	Unique ID for Meal
checkout_price	Final price including discount, taxes & delivery charges
base_price	Base price of the meal
emailer_for_promotion	Emailer sent for promotion of meal
homepage_featured	Meal featured at homepage
num_orders	(Target) Orders Count

2. **fulfilment_center_info.csv:** Contains information for each fulfilment center

Variable	Definition
center_id	Unique ID for fulfillment center
city_code	Unique code for city
region_code	Unique code for region
center_type	Anonymized center type
op_area	Area of operation (in km^2)

3. **meal_info.csv:** Contains information for each meal being served

Variable	Definition
meal_id	Unique ID for the meal
category	Type of meal (beverages/snacks/soups….)
cuisine	Meal cuisine (Indian/Italian/…)

Evaluation Metric

The evaluation metric for this competition is *100*RMSLE* where RMSLE is Root of Mean Squared Logarithmic Error across all entries in the test set.

Public and Private Split

Test data is further randomly divided into Public (30%) and Private (70%) data.
- Your initial responses will be checked and scored on the Public data.
- The final rankings would be based on your private score which will be published once the competition is over.