REPORT

# Food Demand Forecasting

(Genpact)

NAME- MANAV MISHRA

UNIVERSITY ROLL NO- 2013654

**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.

**Introduction :**

### 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.

**Motivation :**

1.The core motive for making this project was primarily an extreme interest towards an interesting area of research. Also Food Demand Forecasting is a hot topic for research.

2.Use of electronic media is increasing day by day.

1. Time is money or even more valuable than money therefore instead of

spending time in reading

Example — Analyzing a product based on it’s reviews and comments.

**TOOLS AND LIBRARY:**

1. **Sklearn:**Scikit-learnis a free machine learning library for Python. It features various algorithms like support vector machine, random forests, and k-neighbours, and it also supports Pythonnumerical and scientific libraries like NumPy and SciPy
2. **Seaborn:** Seabornis a library that uses Matplotlib underneath to plot graphs. It will be used to visualize random distributions.
3. **Pandas:** It is a library available for python programming generally used for data analysis or while dealing with data.
4. **Matplotlib:** It is a library available for python programming used to visualize the data such as to plot the graphs, bar charts, etc.
5. **NumPy:** It is a library available for python programming used for dealing with matrices and arrays and to apply mathematical operations and functions on our dataset in the present in the form of an array.
6. **Math:** The math module is used to access mathematical functions in the Python. All methods of this functions are used for integer or real type objects, not for complex numbers. To use this module, we should import that module into our code.

METHODOLOGY:

Now I imported the libraries discussed above > Then I read all the files using pandas library and stored them into dataframe of the respective name>Then I merged meal\_info and fulfilment\_center\_info with test and train file>Then I checked null values of my train file, there were no null values so I moved on>Then by using info() method I checked all the object values present in my train file>Then by using dummy matrices I converted all the object type values to int or float type and merged them into my train file and dropped the original columns >since id of each row is unique so it doesn’t affects the prediction so I removed it>Then I saw the relation between weeks and num order by using pyplot which gave me the graph of no. of orders per week>Then to condense the data of weeks I converted into month, year and quarter>Then I dropped the week column from my train file>Then we tried to find the outliers in various column by using boxplot>Then by analyzing pearson correlation values I dropped number region code column>then i stored dependent variable in y and independent in x>then I split my train dataset into two parts one for training purpose and other for testing to model in the ratio of 8:2>then I used randomforest regressor to train my data model and used r2 score to calculate the approx. value of the data model>Then I applied this data model on test file and recorded all the values of num order into a csv file.

Data Dictionary:

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

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)

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/…)

**Leaning Outcome :**

Demand forecasting helps businesses make informed decisions that affect everything from inventory planning to supply chain optimization. With customer expectations changing faster than ever, businesses need a method to accurately forecast demand.  
If you’re looking for an e-commerce fulfillment solution to help you improve demand forecasting, learn more about how Ship Bob helps you replenish stock and deliver the experience that customers want. Request a pricing quote of our fulfillment services below