**Problem description:** In this recruiting competition, job-seekers are provided with historical invoice data for many outlets located in different regions. The candidates are required to forecast the sales at day level for a number of outlets. The expected outcome is: for each (date, outlet, item) triplet, forecast the sale. Candidates are also provided additional information on outlets and items in ‘*outlet\_master.csv*’ and ‘*item\_master.csv*’ respectively.

**Data:**

1. **sales\_train.csv.zip:**

Attributes:

1. ***outletCode:*** outlet identifier
2. ***invoice\_date:*** date of transaction
3. ***itemCode:*** item/sku identifier
4. ***sales\_in\_litres:*** item volume sold in litres. You arepredicting in litres.
5. **outlet\_master.csv**:

Attributes:

1. ***outletCode:*** outlet identifier
2. ***latitude:*** geographical latitude component of outlet
3. ***longitude:*** geographical longitude component of outlet
4. ***outletType:*** type of outlet
5. ***outletCategory:*** category of outlet
6. **item\_master.csv:**

Attributes:

1. ***itemCode:*** item/sku identifier
2. ***packet\_size:*** packet size of item/sku in liter
3. ***brand*** : brand the item is from
4. ***category:*** item category
5. ***class:*** item class
6. **test\_data.csv.zip:**

Attributes:

1. ***outletCode:*** mentioned above
2. ***invoice\_date:*** mentioned above
3. ***itemCode:*** mentioned above
4. ***actual\_sales\_in\_litres:*** You are given the actual volume sold in litres
5. ***predicted\_sales\_in\_litres:*** Predicted volume in litres

**Accuracy/Error metric:**

***Mean Absolute Percentage Error(MAPE):***

***Screenshot from 2020-11-23 00-59-32***

Where, M is MAPE, Ais ***actual\_sales\_in\_litres and*** F is ***predicted\_sales\_in\_litres***

**Instructions:**

1. You are required to put forecasted sales in the *predicted\_sales\_in\_litres column in test\_data.*
2. The actual sales information is already provided.
3. Based on the above error metric formula, calculate the error(%).
4. You can iterate 1-3 as many times you wish to reduce the error rate.
5. You need to submit the code with the least error rate achieved, the code must be able to run completely to produce sales forecast.
6. It is preferred to receive information on exploratory data analysis if you have done any. You can submit in EDA.doc/EDA\_docx/EDA.pdf in such case.