Train data set :

Using cleaning.py script,we can understand that Outlet\_Id : out027,out019 is not providing Item weight for every item….

Using relation.py script ,we can understand that primary key is (Item\_Id, Outlet\_Id)

Using check\_toclean.py script ,we can understand that Item\_weight,Outlet\_size need to clean

It is clearly shows that Outlet\_Sales is highest for Out027 using shop\_sales.py

There are 16 different types of items and 2 types of fat contents are present (item\_fat\_count.py)

Item\_Id is repeated(item\_copy.py). Using Item\_Id we try to clean the data.