**Table names:**Deployment\_data.csv : the Data we got after deploying the models on the 14 days in the future:

hybrid\_data\_split.csv: The data and predictions of the predicting the total workorders and the workorders for each facility.

lr\_data\_columns.csv: the columns\features left for the Linear Regression model after performing Feature Selection for it.

lr\_split\_hy\_fi.csv: Linear Regression Feature importances.  
  
merged\_big\_df\_split\_rd.csv: the final preprocessed data where we have the WO\_vol, Duration and the risks (the target columns) for each facility of each workorder in any date.  
  
split\_out\_data.csv: Data after the first preprocessing that was for predicting the workorders for each facility for 14 days (14 steps/days in the future for each facility was the target columns) in the future.

xgb\_data\_columns.csv: the columns\features left for the XGB model after performing Feature Selection for it.

xgb\_split\_hy\_fi.csv: XGB Feature importances.  
  
**train & test percentage:** 80% train, 20% test, without shuffling  
  
**model names:**for predicting the total workorders and for each facility workorders we’ve used a Hybrid XGB + LR model  
for predicting the WO\_vol, we’ve used an XGB regressor  
for predicting the two risks and the duration bin, we’ve used 3 XGB classifiers  
  
**methods used in the deployment:**

for predicting the total and for-each-facility workorders, I moved all the steps from the last row of the data to be lags and then used them as X and predicted the 14 steps for each facility and that was the y

and for the WO\_vol, risks and Duration predictions, I used the number of work orders for each facility to generate the facility and workorider Id, and then added them to the X data, then I inputted it to the models and got the ys that was WO\_vol, risks and Duration  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
  
