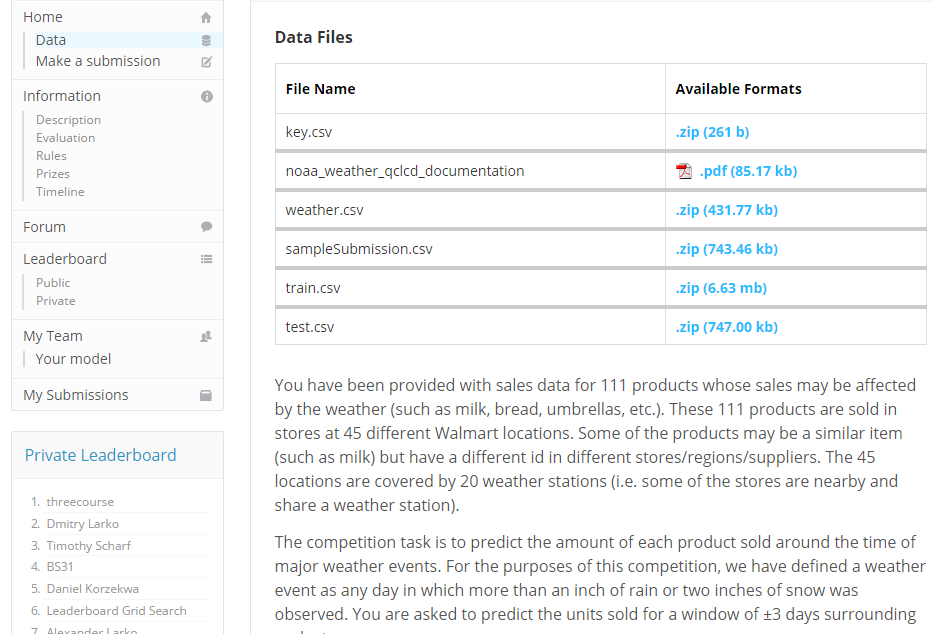
**Workflow Documentation**

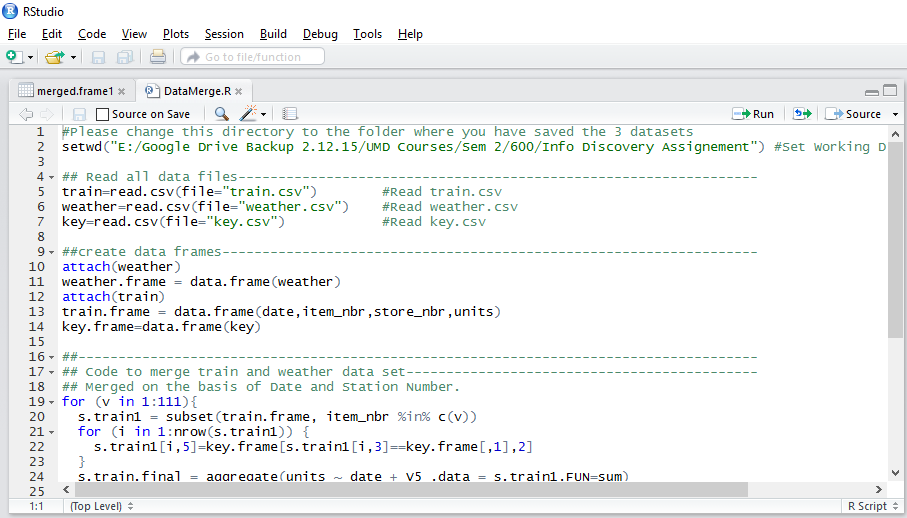
1. Go to the source of the focal dataset at Kaggle, link: <https://www.kaggle.com/c/walmart-recruiting-sales-in-stormy-weather>



1. Pick the below three mentioned files as our data sets. The weather.csv file is our focal data set, which we are going to merge with the train.csv and key.csv datasets.



1. Extract the three datasets: key.csv, weather.csv and merge.csv
2. The next step is merging the data –
3. We have used R programming to merge weather.csv, train.csv and key.csv.
4. The column “date” is common between weather and train datasets.
5. Also, we have used “station\_nbr” from the weather dataset, “store\_nbr” and “Item\_nbr” from the train dataset and, “store\_nbr” and “station\_nbr” from the key dataset to combine and create the merge.csv dataset which shows the number of items sold at a particular station on a given date and weather conditions.
6. The code with comments to merge the dataset is included the R Script – “DataMerge.R”



1. We need to open the above-mentioned file in R and Click on the “Run” Command to run each line of the R Script and generate our final “merge.csv” dataset.