**Data Dictionary**

**PscInspectionId:** ID associated with PSC inspection

**deficiency\_code:** Defect code associated with a deficiency. It is a 5-digit code.

Note: In csv files, a deficiency code starting with 0 might have only 4 digits. In this case, you will have to prepend a 0. For example, a deficiency code 01104 might appear as 1104 in csv, so you need to make it 01104 after loading.

**def\_text:** This is the information associated with a deficiency. It usually contains the following information:

* + Deficiency
  + Description of deficiency
  + Root cause analysis
  + Corrective action
  + Preventive action

**annotation\_id:** Unique ID associated with an annotator

**username:** Username of the annotator

**annotation\_severity:** Severity of the deficiency assigned by the annotator. The possible values are High, Low or Medium.

**InspectionDate**: Date on which the PSC inspection took place.

**VesselId:** Unique ID associated with the vessel which got inspected.

**PscAuthorityId:** Id associated with a regional Memorandum of Understanding (MoU). These are regional agreements among maritime authorities of different countries to enforce consistent safety, security, and environmental standards for ships entering their ports. For e.g. id 7 may be associated with Paris MoU (covers Europe) which means vessel was inspected at a port covered under Paris MoU.

**PortId:** Id associated with the port visited by the vessel.

**Age:** Age of vessel at the time of inspection.

**VesselGroup:** Type of vessel

**Note:**

1. In the train set, each set of annotations generally contains 3 records. It may contain more than 3 records if the same deficiency was shared with more than one team of annotators. In the scenario if it contains less than 3 records then it could be because an annotator skipped the annotation.
2. The test data has each row which is a unique record (based on combination of PscInspectionId and deficiency\_code).
3. You need to provide the predicted severity in a new column ‘predicted\_severity’. You have been provided with a sample submission file. Make sure that the column names remain the same as in the provided submission file.