**Outpatient Data (Train and Test):**

**It consists of the claim details for the patients who were not admitted into the hospital, who only visited there. Important columns are explained below.**

**BeneID:** It contains the unique id of each beneficiary i.e patients.  
**ClaimID:** It contains the unique id of the claim submitted by the provider.  
**ClaimStartDt:** It contains the date when the claim started in yyyy-mm-dd format.  
**ClaimEndDt:** It contains the date when the claim ended in yyyy-mm-dd format.  
**Provider:** It contains the unique id of the provider.  
**InscClaimAmtReimbursed:** It contains the amount reimbursed for that particular claim.  
**AttendingPhysician:** It contains the id of the Physician who attended the patient.  
**OperatingPhysician:** It contains the id of the Physician who operated on the patient.  
**OtherPhysician:** It contains the id of the Physician other than AttendingPhysician and OperatingPhysician who treated the patient.  
**ClmDiagnosisCode:** It contains codes of the diagnosis performed by the provider on the patient for that claim.  
**ClmProcedureCode:** It contains the codes of the procedures of the patient for treatment for that particular claim.  
**DeductibleAmtPaid:** It consists of the amount paid by the patient. That is equal to Total\_claim\_amount — Reimbursed\_amount.

**Inpatient Data (Train and Test):**

**It consists of the claim details for the patients who were admitted into the hospital. So, it consists of 3 extra columns Admission date, Discharge date, and Diagnosis Group code.**

**AdmissionDt:** It contains the date on which the patient was admitted into the hospital in yyyy-mm-dd format.  
**DischargeDt:** It contains the date on which the patient was discharged from the hospital in yyyy-mm-dd format.  
**DiagnosisGroupCode:** It contains a group code for the diagnosis done on the patient.

**Beneficiary Data (Train and Test):**

**This data contains beneficiary KYC details like DOB, DOD, Gender, Race, health conditions (Chronic disease if any), State, Country they belong to, etc. Columns of this dataset are explained below.**

**BeneID:** It contains the unique id of the beneficiary.  
**DOB:** It contains the Date of Birth of the beneficiary.  
**DOD:** It contains the Date of Death of the beneficiary if the beneficiary id deal else null.  
**Gender, Race, State, Country:** It contains the Gender, Race, State, Country of the beneficiary.  
**RenalDiseaseIndicator:** It contains if the patient has existing kidney disease.  
**ChronicCond\_\*:** The columns started with “ChronicCond\_” indicates if the patient has existing that particular disease. Which also indicates the risk score of that patient.  
**IPAnnualReimbursementAmt:** It consists of the maximum reimbursement amount for hospitalization annually.  
**IPAnnualDeductibleAmt:** It consists of a premium paid by the patient for hospitalization annually.  
**OPAnnualReimbursementAmt:** It consists of the maximum reimbursement amount for outpatient visits annually.  
**OPAnnualDeductibleAmt:** It consists of a premium paid by the patient for outpatient visits annually.