**Medication Errors and Effects**

**Data Description:**

This dataset talks about the medication error reports, submitted to the FDA, to support the post-marketing safety surveillance program for drug and therapeutic biologic products. The structure of the AERS database adheres to the international safety reporting guidance issued by the International Conference on Harmonisation (ICH E2B). Adverse events and medication errors are coded to terms in the Medical Dictionary for Regulatory Activities (MedDRA) terminology.

**Data Dictionary:**

1. **Demographics\_data\_2015.csv**

| Column Name | Description |
| --- | --- |
| Primaryid | Unique Id in the table |
| Caseid | Unique Id for every case |
| Caseversion | Version of the Case |
| I F Code |  |
| Event Date | Date the adverse event occurred or began. (YYYYMMDD format). If a complete date is not available, a partial date is provided. |
| Event Date Unparsed | Date adverse event occurred or began. Dates are in raw form. |
| Manufacture Received Date | Date manufacturer first received initial information. In subsequent versions of a case, the latest manufacturer received date will be provided (YYYYMMDD format). If a complete date is not available, a partial date will be provided. See the NOTE on dates at the end of this section. |
| Manufacture Received Date Unparsed | Date manufacturer first received initial information. In subsequent versions of a case, the latest manufacturer received date will be provided (YYYYMMDD format). If a complete date is not available, a partial date will be provided. See the NOTE on dates at the end of this section. Dates are in raw form. |
| Init Fda Dt | Init Fda Dt |
| Init Fda Dt Unparsed | Init Fda Dt Unparsed |
| FDA Received Date | Date FDA received Case. In subsequent versions of a case, the latest manufacturer received date will be provided (YYYYMMDD format). |
| FDA Date Unparsed | Date FDA received Case. In subsequent versions of a case, the latest manufacturer received date will be provided. Dates are in raw form. |
| Report Type | Code for the type of report submitted. "EXP" means expedited (15 day), "PER" is periodic, and "DIR" is direct. |
| Auth Num | Auth Num |
| Manufacturer Number | Manufacturer's unique report identifier. |
| Manufacturer Filing Report | Coded name of manufacturer sending report; if not found, then verbatim name of organization |
| Lit Ref | Lit Ref |
| Patient's Age | Numeric value of patient's age at event |
| Patient's Age Unit | Unit abbreviation for patient's age. "DEC" is decade, "YR" is year, "MON" is month, "WK" is week, "DY" is day, and "HR" is hour. |
| Age Grp | Age Group |
| Sex | Gender |
| E-Submission | Indicates whether (Y/N) this report was submitted under the electronic submissions procedure for manufacturers. |
| Patient's Weight | Weight. Numeric value of patient's weight. |
| Patient's Weight Unit | Unit abbreviation for patient's weight. "KG" is kilograms, "LBS" is pounds, and "GMS" is grams. |
| Report Date | Date report was sent (YYYYMMDD format). If a complete date is not available, a partial date is provided. |
| Report Date Unparsed | Date report was sent. If a complete date is not available, a partial date is provided. Dates are in raw form. |
| Manufacturer Notified | Whether (Y/N) voluntary reporter also notified manufacturer (blank for manufacturer reports). |
| Occupation Code | Abbreviation for the reporter's type of occupation in the latest version of a case. "MD" is physician, "PH" is pharmacist, "OT" is other health professional, "LW" is lawyer, and "CN" is consumer. |
| Reporter Country | Country |
| Occurrence Country | Country where it occurred |
| Serialid |  |
| Enigma Serial ID |  |

1. Drug Information.csv

| Column Name | Description |
| --- | --- |
| Isr | Number that uniquely identifies an AERS report. Primary link field between data files. |
| Drug Sequence No. | Unique number for identifying a drug for a case. This field can be used as a primary link between different tables within this dataset. |
| Drug Role | Code for drug's reported role in event. "PS" is primary suspect drug, "SS" is secondary suspect drug, "C" is concomitant, and "I" is interacting. |
| Drug Name | Name of medicinal product. If a "Valid Trade Name" is populated for this case, then Drug Name = Valid Trade Name; if not, then Drug Name = "Verbatim" name, exactly as entered on the report. For the great majority of reports, there is a "Valid Trade Name." |
| Validated/Verbatim Name | "1" indicates that a validated trade name is used for the drug name and "2" indicates that a verbatim name is used. |
| Route | The route of drug administration. |
| Dose | Verbatim text for dose, frequency, and route, exactly as entered on report. |
| Dechallenge | Dechallenge code, indicating if reaction abated when drug therapy was stopped. "Y" is positive dechallenge, "N" is negative dechallenge, "U" is unknown, and "D" does not apply. |
| Rechallenge | Rechallenge code, indicating if reaction recurred when drug therapy was restarted. "Y" is positive rechallenge, "N" is negative rechallenge, "U" is unknown, and "D" does not apply. |
| Lot No. | Lot number of the drug. |
| Expiration Date | Expiration date of the drug. (YYYYMMDD format). If a complete date is not available, a partial date is provided. |
| Exp Dt Unparsed | Exp Dt Unparsed |
| NDA No | National Drug Administration number of the drug. |
| Quarter | Quarter |
| Serialid | Serialid |

1. **Drug Therapy Duration.csv**

| **Column Name** | **Description** |
| --- | --- |
|  |  |
| ISR | Individual Safety Report (ISR) number. |
| Drug Sequence | Drug Sequence Number |
| Start Date | Date therapy was started (or re-started) for this drug (YYYYMMDD) If a complete date not available, a partial date is provided. |
| Start Date Unparsed | Date therapy was started (or re-started) for this drug (YYYYMMDD) If a complete date not available, a partial date is provided, in raw data format. |
| End Date | Date therapy was stopped for this drug. (YYYYMMDD) If a complete date not available, a partial date will be provided. |
| End Date Unparsed | Date therapy was stopped for this drug. (YYYYMMDD) If a complete date not available, a partial date will be provided, in raw data format. |
| Duration | Duration of drug therapy |
| Duration Unit | Duration unit for drug therapy |
| Quarter | Quarter |
| Serialid | Serialid |

1. **Event Terms.csv**

| Column Name | Description |
| --- | --- |
| Isr | Number that uniquely identifies an AERS report. Primary link field between data files. |
| Preferred Term | Preferred Term" level medical terminology describing the event, using the Medical Dictionary for Regulatory Activities (MedDRA). The order of the terms for a given event does not imply priority. In other words, the first term listed is not necessarily considered more significant than the last one listed. |
| Quarter | Quarter |
| Serialid | Serialid |

1. **Patient Outcomes.csv**

| **Column Name** | **Description** |
| --- | --- |
| ISR | The number that uniquely identifies an AERS report. Primary link field between data files. |
| Outcome | The patient outcome |
| Outc Cod Definition | Outcome Code Definition. DE, Death | LT, Life-Threatening | HO, Hospitalization - Initial or Prolonged | DS, Disability | CA, Congenital Anomaly | RI, Required Intervention to Prevent Permanent Impairment/Damage | OT, Other |
| Quarter | Quarter |
| Serialid | Serialid |

1. **Preferred Term Indicators.csv**

| **Column Name** | **Description** |
| --- | --- |
| ISR No. | The number that uniquely identifies an AERS report. Primary link field between data files |
| Drug Seq | The drug sequence number for identifying a drug for an ISR |
| Preferred Term | "Preferred Term" level medical terminology describing the indication for use, using the Medical Dictionary for Regulatory Activities (MedDRA). |
| Quarter | Quarter |
| Serialid | Serialid |

1. **Report Sources.csv**

| Column Name | Description |
| --- | --- |
| Isr | Number that uniquely identifies an AERS report. Primary link field between data files. |
| Report Source | Code for an initial source of the report. |
| Rpsr Cod Definition | Report Source Code Definition. FGN, Foreign | SDY, Study | LIT, Literature | CSM, Consumer | HP, Health Professional | UF, User Facility | CR, Company Representative | DT, Distributor | OTH, Other |
| Quarter | Quarter |
| Serialid | Serialid |

**Guidelines for Trainees:**

Below are the expectations/guidelines from/for the Mock Project. You can use the below points to structure your thought process and plan it accordingly. Feel free to explore all the technologies learned during training whether on-premise or in the cloud.

Steps to follow:

* Gather data understanding via the data dictionary and domain knowledge.
* Perform ETL using a suitable tool.
* Identify dimensions and facts and perform data modeling.
* Dashboarding and story building
  + Identify suitable metrics and generate insights
  + Identify the story points in the data and come up with a compelling story
* Presentation `
  + Make a PPT to showcase all of the above in 15 minutes.