Data Familiarization

1. allergies.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| START | Date (DD-MM-YYYY) | Date when the allergy was first recorded |
| STOP | Date (DD-MM-YYYY) | Date when the allergy was resolved (blank if ongoing) |
| PATIENT | String (Text) | Unique identifier for the patient |
| ENCOUNTER | String | Unique identifier for the encounter when allergy was noted |
| CODE | String | Standard medical code for the allergy – SNOMED-CT |
| DESCRIPTION | String | Human readable description of the allergy |

1. careplans.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Id | String | Unique identifier for the care plan |
| START | Date (DD-MM-YYYY) | Date when the care plan started |
| STOP | Date (DD-MM-YYYY) | Date when the care plan ended (blank if ongoing) |
| PATIENT | String | Unique identifier for the patient |
| ENCOUNTER | String | Unique identifier for the encounter |
| CODE | String | Standardized code for the type of care plan - SNOMED |
| DESCRIPTION | String | Human-readable description of the care plan |
| REASONCODE | String | SNOMED-CT code for the reason/condition needing care plan |
| REASONDESCRIPTION | String | Human readable reason for the care plan |

1. conditions.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| START | Date (DD-MM-YYYY) | Date when the condition began |
| STOP | Date (DD-MM-YYYY) | Date when the condition ended (blank if ongoing) |
| PATIENT | String | Unique identifier for the patient |
| ENCOUNTER | String | Unique identifier for the encounter |
| CODE | String | Standard medical code (SNOMED) for the condition |
| DESCRIPTION | String | Human-readable condition name |

1. devices.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| START | DateTime | Date when the device record began |
| STOP | DateTime | Date when the device record ended (blank if ongoing) |
| PATIENT | String | Unique identifier for the patient |
| ENCOUNTER | String | Unique identifier for the encounter |
| CODE | String | Standard medical code ()SNOMED for the device |
| DESCRIPTION | String | Human-readable description of the device |
| UDI | String | Unique Device Identifier (UDI) for the device |

1. encounters.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Id | String | Encounter ID |
| START | DateTime | Date and time when the encounter began |
| STOP | DateTime | Date and time when the encounter ended |
| PATIENT | String | Unique Identifier for Patient |
| ORGANIZATION | String | Unique Organization ID |
| PROVIDER | String | Unique Provider ID |
| PAYER | String | Unique Payer ID |
| ENCOUNTERCLASS | String | Classification of the encounter |
| CODE | String | Standard medical code ()SNOMED for the encounter |
| DESCRIPTION | String | Human-readable encounter description |
| BASE\_ENCOUNTER\_COST | Number (Decimal) | Base cost of the encounter |
| TOTAL\_CLAIM\_COST | Number (Decimal) | Total claim cost for the encounter |
| PAYER\_COVERAGE | Number (Decimal) | Amount covered by payer |
| REASONCODE | String | Standard medical code for the reason of encounter |
| REASONDESCRIPTION | String | Reason description |

1. imaging\_studies.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Id | String | Imaging study ID |
| DATE | DateTime | Date and time when the imaging study was performed |
| PATIENT | String | Unique Identifier for Patient |
| ENCOUNTER | String | Unique Identifier for Encounter |
| BODYSITE\_CODE | String | Standard medical code (SNOMED) for the body site imaged |
| BODYSITE\_DESCRIPTION | String | Body site description |
| MODALITY\_CODE | String | Code for the imaging modality |
| MODALITY\_DESCRIPTION | String | Imaging modality description |
| SOP\_CODE | String | Code for the SOP (Service-Object Pair) class |
| SOP\_DESCRIPTION | String | Human readable SOP class description |

1. immunizations.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| DATE | DateTime | Date and time of administration |
| PATIENT | String | Unique Identifier for Patient |
| ENCOUNTER | String | Unique Identifier for Encounter |
| CODE | String | Standard code (SNOMED) for the immunization/vaccine |
| DESCRIPTION | String | Human-readable description of the immunization/vaccine |
| BASE\_COST | Decimal | Cost of the immunization |

1. medications.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| START | DateTime | Start date and time of the medication coverage or dispensing period |
| STOP | DateTime | End date and time of the medication coverage or dispensing period |
| PATIENT | String | Unique Identifier for Patient |
| PAYER | String | Unique Identifier for Payer\ |
| ENCOUNTER | String | Unique Identifier for Encounter |
| CODE | String | Standard code (RxNorm) for the medication |
| DESCRIPTION | String | Human-readable description of the medication |
| BASE\_COST | Decimal | Cost of the medication per unit |
| PAYER\_COVERAGE | Decimal | Amount covered by the payer |
| DISPENSES | Integer | Number of times the medication was dispensed |
| TOTALCOST | Decimal | Total cost of the medication |
| REASONCODE | String | Standard code (SNOMED) for the reason for prescription |
| REASONDESCRIPTION | String | Human-readable description of the reason for prescription |

1. observations.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Column Name | Data Type | Description | Coding System |
| DATE | DATETIME | Date and time the observation was recorded | – |
| PATIENT | STRING | Unique identifier for the patient | UUID |
| ENCOUNTER | STRING | Unique identifier for the encounter during which the observation was made | UUID |
| CODE | STRING | Clinical code for the observation (eg, height, weight, BMI, pain score) | LOINC |
| DESCRIPTION | STRING | Human-readable description of the observation | LOINC Description |
| VALUE | STRING | Result value of the observation (numeric or text depending on TYPE) | – |
| UNITS | STRING | Units of the measurement (eg, cm, kg, kg/m2, {score}) | UCUM (Unified Code for Units of Measure) |
| TYPE | STRING | Type of value (numeric, text) | – |

1. organisations.csv

|  |  |  |
| --- | --- | --- |
| NAME | TYPE | DESCRIPTION |
| Id | String | Unique identifier for the hospital record |
| NAME | String | Name of the hospital or healthcare organization |
| ADDRESS | String | Street address of the hospital |
| CITY | String | City where the hospital is located |
| STATE | String | State abbreviation where the hospital is located |
| ZIP | String | ZIP code of the hospital location |
| LAT | Float | Latitude coordinate of the hospital location |
| LON | Float | Longitude coordinate of the hospital location |
| PHONE | String | Contact phone number of the hospital |
| REVENUE | String | Reported revenue of the hospital |
| UTILIZATION | Integer | Measure of hospital utilization (eg, number of patients served, services) |

1. patients.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Type | Description |
| Id | String (UUID) | Unique identifier for each patient record |
| BIRTHDATE | Date | Patient’s date of birth |
| DEATHDATE | Date (nullable) | Patient’s date of death (if applicable) |
| SSN | String | Social Security Number |
| DRIVERS | String | Driver’s license number |
| PASSPORT | String | Passport number |
| PREFIX | String | Name prefix (eg, Mr, Mrs, Ms) |
| FIRST | String | First name of patient |
| LAST | String | Last name of patient |
| SUFFIX | String | Name suffix (eg, Jr, Sr, PhD) |
| MAIDEN | String | Maiden name (if applicable) |
| MARITAL | String | Marital status (eg, M = Married, S = Single) |
| RACE | String | Patient’s race classification |
| ETHNICITY | String | Ethnicity classification (eg, hispanic, nonhispanic) |
| GENDER | String | Patient gender (M/F) |
| BIRTHPLACE | String | Region of patient’s birth |
| ADDRESS | String | Street address |
| CITY | String | Current city of residence |
| STATE | String | State of residence |
| COUNTY | String | County of residence |
| ZIP | String | Zip code of residence |
| LAT | Float | Latitude coordinate of residence |
| LON | Float | Longitude coordinate of residence |
| HEALTHCARE\_EXPENSES | Currency | Total healthcare expenses incurred by patient |
| HEALTHCARE\_COVERAGE | Currency | Healthcare coverage/insurance amount available |

1. payers.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Id | String | Unique identifier for the payer/insurance provider |
| NAME | String | Name of the payer |
| ADDRESS | String | Address of the payer’s headquarters |
| CITY | String | City of the payer’s headquarters |
| STATE\_HEADQUARTERED | String | State of the payer’s headquarters |
| ZIP | String | ZIP code of the headquarters |
| PHONE | String | Phone number of the payer |
| AMOUNT\_COVERED | Decimal | Total amount covered by the payer |
| AMOUNT\_UNCOVERED | Decimal | Amount not covered by the payer |
| REVENUE | Decimal | Total revenue generated by the payer |
| COVERED\_ENCOUNTERS | Integer | Number of encounters covered |
| UNCOVERED\_ENCOUNTERS | Integer | Number of encounters not covered |
| COVERED\_MEDICATIONS | Integer | Number of medications covered |
| UNCOVERED\_MEDICATIONS | Integer | Number of medications not covered |
| COVERED\_PROCEDURES | Integer | Number of procedures covered |
| UNCOVERED\_PROCEDURES | Integer | Number of procedures not covered |
| COVERED\_IMMUNIZATIONS | Integer | Number of immunizations covered |
| UNCOVERED\_IMMUNIZATIONS | Integer | Number of immunizations not covered |
| UNIQUE\_CUSTOMERS | Integer | Count of unique customers served by the payer |
| QOLS\_AVG | Decimal | Average Quality of Life Score (QOLS) for customers |
| MEMBER\_MONTHS | Integer | Number of member-months |

1. payer\_transitions.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| PATIENT | String | Unique Identifier for Patient |
| START\_YEAR | Integer | Year coverage started |
| END\_YEAR | Integer | Year coverage ended |
| PAYER | String | Unique Identifier for the Payer |
| OWNERSHIP | String | Name of the payer organization |

1. procedures.csv

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| DATE | DateTime | Date and time of the procedure or service |
| PATIENT | String | Unique Identifier for Patient |
| ENCOUNTER | String | Unique Identifier for Encounter |
| CODE | String | Standard code (SNOMED CT) representing the procedure or service |
| DESCRIPTION | String | Human-readable description of the procedure or service |
| BASE\_COST | Decimal | Cost associated with the procedure or service |
| REASONCODE | String | Standard code (SNOMED CT) representing the reason for the procedure/service |
| REASONDESCRIPTION | String | Human-readable description of the reason for the procedure/service |

1. providers.csv

|  |  |  |
| --- | --- | --- |
| Name | Type | Description |
| Id | String | Unique identifier for the provider record |
| ORGANIZATION | String | Identifier of the organization |
| NAME | String | Full name of the healthcare provider |
| GENDER | String | Gender of the provider (eg, M = Male, F = Female) |
| SPECIALITY | String | Medical specialty or area of practice of the provider |
| ADDRESS | String | Street address of the provider’s practice location |
| STATE | String | State abbreviation of the practice location |
| ZIP | String | ZIP code of the provider’s practice location |
| LAT | Float | Latitude coordinate of the provider’s practice location |
| LON | Float | Longitude coordinate of the provider’s practice location |
| UTILIZATION | Integer | Number of healthcare encounters, visits, or services attributed to provider |