**Dummy Data Creation Guide for Health Condition Records in Multiple Formats**

**Objective:**  
For this task, we will generate dummy data records for various datasets simulating patient information across different medical facilities.

**Data Formats:**  
The datasets will be stored in a variety of formats, such as CSV, Excel, and JSON.

**Fields Included:**  
These records will contain fields like email, full name, diagnosis, treatments, and relevant dates (e.g., admission and discharge).

**Purpose:**  
The purpose of this task is to create a consistent, representative sample of health condition records that can be used for testing and validation purposes in a project.

**Record Generation:**  
Initially, 600,000 records will be generated for developing and testing the ETL Pipeline, followed by 200,000 subsequent records for refining the ETL Pipeline and conducting thorough testing. Finally, 50,000 records will be utilized in the live environment, where the ETL pipeline will support actual business needs. The ETL job will be scheduled to run weekly on Windows.

**Example Records for Each Dataset:**

1. **Hospital\_A\_Patients.csv**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Email** | **Full\_Name** | **Gender** | **Phone** | **Treatment** | **Diagnosis** | **Address** | **Discharge\_Date** | **DOB** | **Admission\_Date** |
| john.doe@example.com | John Doe | M | 555-1234 | Medication A | Hypertension | 123 Elm St, City, State | 2024-01-10 | 1980-01-15 | 2024-01-01 |
| jane.smith@example.com | Jane Smith | F | 555-5678 | Insulin | Diabetes | 456 Oak St, City, State | 2024-02-05 | 1975-05-23 | 2024-02-01 |

1. **Clinic\_B\_Records.xlsx**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Date\_Out** | **ContactEmail** | **Condition** | **Therapy** | **DateOfBirth** | **ContactPhone** | **Name** | **Date\_In** | **Sex** | **ResidenceAddress** |
| 2024-02-05 | jane.smith@clinicb.com | Diabetes | Insulin | 1975-05-23 | 555-8765 | Jane Smith | 2024-02-01 | F | 789 Pine St, City, State |
| 2024-01-10 | john.doe@clinicb.com | Hypertension | Medication A | 1980-01-15 | 555-4321 | John Doe | 2024-01-01 | M | 321 Birch St, City, State |

1. **Medical\_Clinic\_C.csv**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **MailingAddress** | **TreatmentPlan** | **Name\_Last** | **DischargeDate** | **AdmitDate** | **EmailAddress** | **Gender** | **BirthDate** | **Name\_First** | **PhoneNumber** |
| 654 Maple St, City, State | Medication A | Doe | 2024-01-10 | 2024-01-01 | john.doe@medicalclinic.com | M | 1980-01-15 | John | 555-3456 |
| 987 Cedar St, City, State | Insulin | Smith | 2024-02-05 | 2024-02-01 | jane.smith@medicalclinic.com | F | 1975-05-23 | Jane | 555-7890 |

1. **Health\_Records\_D.xlsx**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Address** | **Diagnosis** | **DischargeDate** | **EmailID** | **AdmissionDate** | **Phone\_No** | **PatientName** | **DOB** | **Treatment** |
| 321 Oak St, City, State | Hypertension | 2024-01-10 | john.doe@healthd.com | 2024-01-01 | 555-6543 | John Doe | 1980-01-15 | Medication A |
| 654 Pine St, City, State | Diabetes | 2024-02-05 | jane.smith@healthd.com | 2024-02-01 | 555-9876 | Jane Smith | 1975-05-23 | Insulin |

1. **Patient\_Records\_E.csv**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **PhoneNum** | **DischargeDate** | **Diagnosis** | **FirstName** | **AdmitDate** | **LastName** | **MailingAddr** | **Sex** | **EmailAddress** | **BirthDate** |
| 555-1122 | 2024-01-10 | Hypertension | John | 2024-01-01 | Doe | 987 Elm St, City, State | M | john.doe@patientrecords.com | 1980-01-15 |
| 555-3344 | 2024-02-05 | Diabetes | Jane | 2024-02-01 | Smith | 654 Birch St, City, State | F | jane.smith@patientrecords.com | 1975-05-23 |

1. **Hospital\_F\_PatientData.xlsx**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Address** | **Date\_Out** | **Diagnosis** | **PhoneNumber** | **Name** | **Treatment\_Type** | **Date\_In** | **EmailAddress** | **DOB** | **Sex** |
| 123 Spruce St, City, State | 2024-01-10 | Hypertension | 555-2233 | John Doe | Medication A | 2024-01-01 | john.doe@hospitalf.com | 1980-01-15 | M |
| 456 Fir St, City, State | 2024-02-05 | Diabetes | 555-5566 | Jane Smith | Insulin | 2024-02-01 | jane.smith@hospitalf.com | 1975-05-23 | F |

1. **Clinic\_G\_PatientList.csv**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Email** | **Medication** | **AdmissionDate** | **Phone** | **Condition** | **Full\_Name** | **DischargeDate** | **Gender** | **BirthDate** | **Address** |
| john.doe@clinicg.com | Medication A | 2024-01-01 | 555-7788 | Hypertension | John Doe | 2024-01-10 | M | 1980-01-15 | 789 Elm St, City, State |
| jane.smith@clinicg.com | Insulin | 2024-02-01 | 555-9900 | Diabetes | Jane Smith | 2024-02-05 | F | 1975-05-23 | 123 Maple St, City, State |

1. **Patient\_Data\_H.xlsx**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Address** | **DateOfBirth** | **AdmissionDate** | **TreatmentPlan** | **Phone\_No** | **Name** | **DischargeDate** | **MainDiagnosis** | **Email** | **Sex** |
| 321 Fir St, City, State | 1980-01-15 | 2024-01-01 | Medication A | 555-3344 | John Doe | 2024-01-10 | Hypertension | john.doe@patientdatah.com | M |
| 654 Oak St, City, State | 1975-05-23 | 2024-02-01 | Insulin | 555-5566 | Jane Smith | 2024-02-05 | Diabetes | jane.smith@patientdatah.com | F |

1. **Medical\_Database\_I.json**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Name** | **Phone** | **DOB** | **DateAdmitted** | **Condition** | **Gender** | **Email** | **Drugs** | **MailingAddress** | **DateDischarged** |
| John Doe | 555-2233 | 1980-01-15 | 2024-01-01 | Hypertension | M | john.doe@medicaldb.com | Medication A | 789 Cedar St, City, State | 2024-01-10 |
| Jane Smith | 555-4455 | 1975-05-23 | 2024-02-01 | Diabetes | F | jane.smith@medicaldb.com | Insulin | 123 Pine St, City, State | 2024-02-05 |

1. **Patient\_Records\_J. json**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Phone** | **Email** | **MailingAddress** | **BirthDate** | **AdmissionDate** | **Disease** | **DischargeDate** | **Therapy** | **Name** | **Gender** |
| 555-6677 | john.doe@patientrecordsj.com | 456 Cedar St, City, State | 1980-01-15 | 2024-01-01 | Hypertension | 2024-01-10 | Medication A | John Doe | M |
| 555-8899 | jane.smith@patientrecordsj.com | 789 Spruce St, City, State | 1975-05-23 | 2024-02-01 | Diabetes | 2024-02-05 | Insulin | Jane Smith | F |