**Checklist for Health Condition Registry ETL**

**Objective**

Provide a comprehensive framework to integrate health condition registry data, ensuring consistency and accuracy through a well-defined ETL process.

**Checklist**

**☑ Understand the project objective and scope.  
☑ Design the database schema with tables: Patient, Condition, Treatment, Gender.  
☑ Set up SQL Server database.  
☑ Develop the Extract phase:  
 ☑** Merge datasets from various sources. **☑** Map fields to the master schema. **☑ Develop the Transform phase:  
 ☑** Verify data integrity. **☑** Handle missing data (imputation/placeholders). **☑** Standardize field formats (dates, phone numbers). **☑** Clean addresses and validate data. **☑** Check for duplicates and consistency. **☑ Develop the Load phase:  
 ☑** Create foreign keys for normalization. **☑** Prevent duplicates by checking unique identifiers (phone/email). **☑** Insert only unique records into the database. **☑ Schedule the ETL pipeline to run periodically.  
☑ Validate the ETL process by querying the database and checking data accuracy.  
☑ Document the ETL process, schema design, and usage instructions.**