HealthPulse Project Architecture

Internal and External Sources

Data processing

Store, Retrieve and Manage Data

Data Integration Services

Data Processing and Analytics Services

Consumer Services

CSV Data

- Data Cleaning
- Data Integration
- Data Transformation
- Data Reduction
- Data Discrealization

Database

Apache Kafka

PySpark ⇔ Python

Tableau

Customized Dashboard and apps

Data Quality Check

Data Validation/Migration Check

Data Processing

User Case Testing