Data & Knowledge Management

Presented by

Clayton J. Hamilton, PharmD

Pharmacist Informatician – Regional Data Manager (VISN 20)

Clayton.Hamilton@va.gov

Objectives

Differentiate between structured versus unstructured EMR data

Summarize high level concepts for EMR data transactions and data warehousing

Describe key elements of a relational database.

Formulate simple SQL queries using a basic understanding of language syntax

*EMR = Electronic Medical Record

Why specialize in health informatics as a pharmacist?

Multiply your impact on a patient population

Improve health care quality and patient safety

Support clinical research projects

Lead the data revolution in health care

Databases & Healthcare – Case Study

Background

- VA CARES is an oncology pharmacist-led telehealth medication management program
- Primarily serves rural Veterans receiving oral antineoplastic therapies prescribed by non-VA providers

Problems

- Lots of pre-enrollment data needed along with status tracking
- The VA CARES team would like to track upcoming telehealth visits and other data across 3 medical centers
- Need to determine if there are cost-savings associated

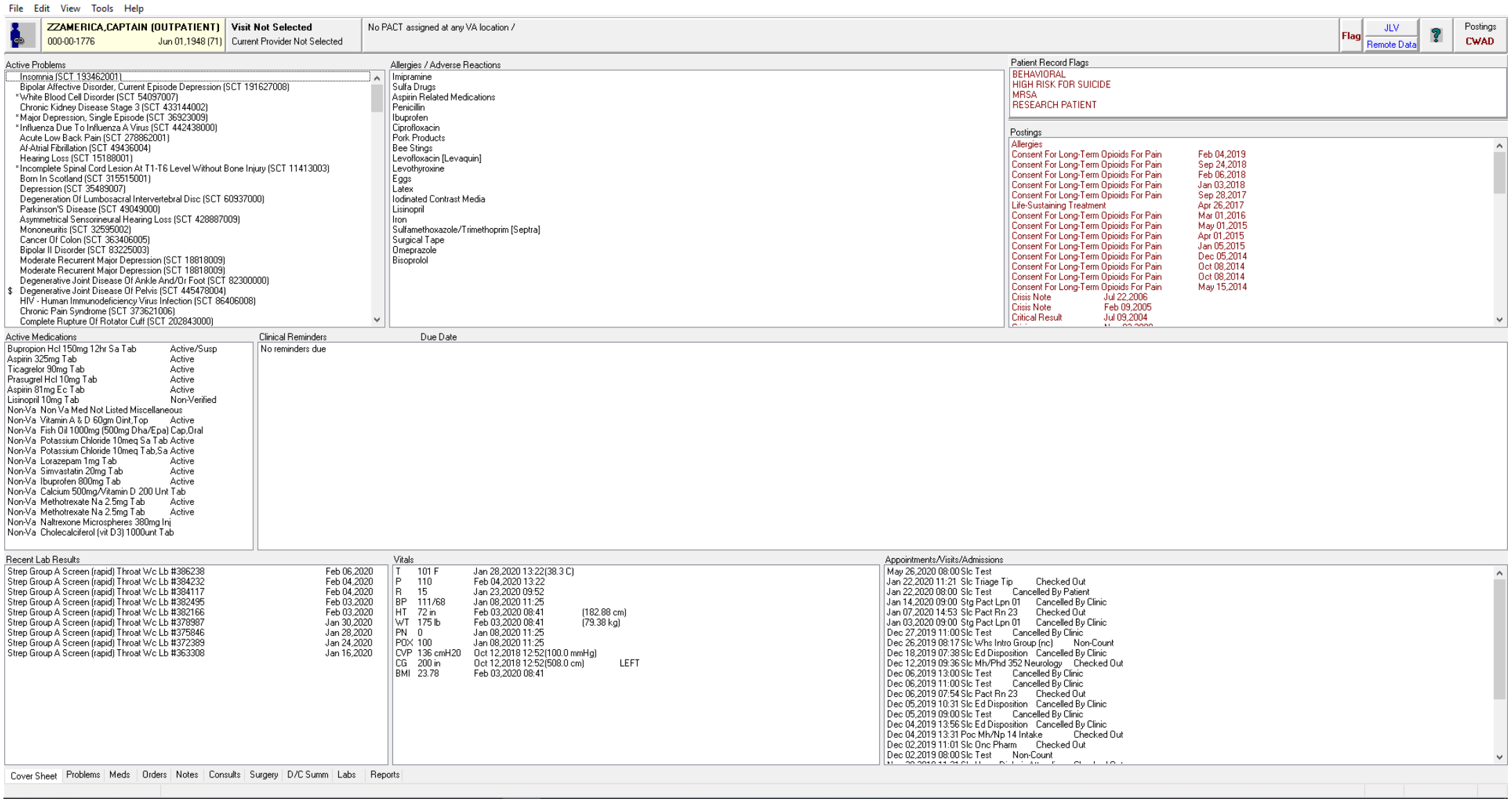
Solution

- Developed an internal patient enrollment app to track status and data needed from non-VA providers
- Developed a dashboard that tracks visit data across 3 medical centers and reminds clinicians to reach out
- Retrieved data captured in clinician notes to support implementation scientist's cost savings analyses

Passey D, Healy R, Qualls J, Hamilton C, Tilley E, Burningham Z, Sauer B, Halwani A. Development and implementation of a pharmacist-led telehealth medication management program for veterans receiving oral antineoplastic therapies through the MISSION Act, *American Journal of Health-System Pharmacy.* 2022 Jun;79(11):835-843.

Medical Data Pipeline - Overview

VA Computerized Patient Record System (CPRS)

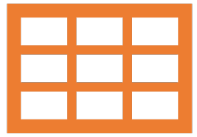


*GUI: Graphical User Interface

Transactional Data

- Any information entered into the electronic medical record (EMR)
- Single transactions are the base unit of the entire EMR
- **Records management**: The process of retaining transactions for future use
- Communicate completed work to other users

Transactional Data Structure



Structured data

Fits a pre-defined model

Much easier to process/query

Examples: dates, lab data, SSNs, phone numbers, NDCs

Semistructured data

Contains both structured and unstructured data

Example: cellphone photos

Unstructured data

No predefined model

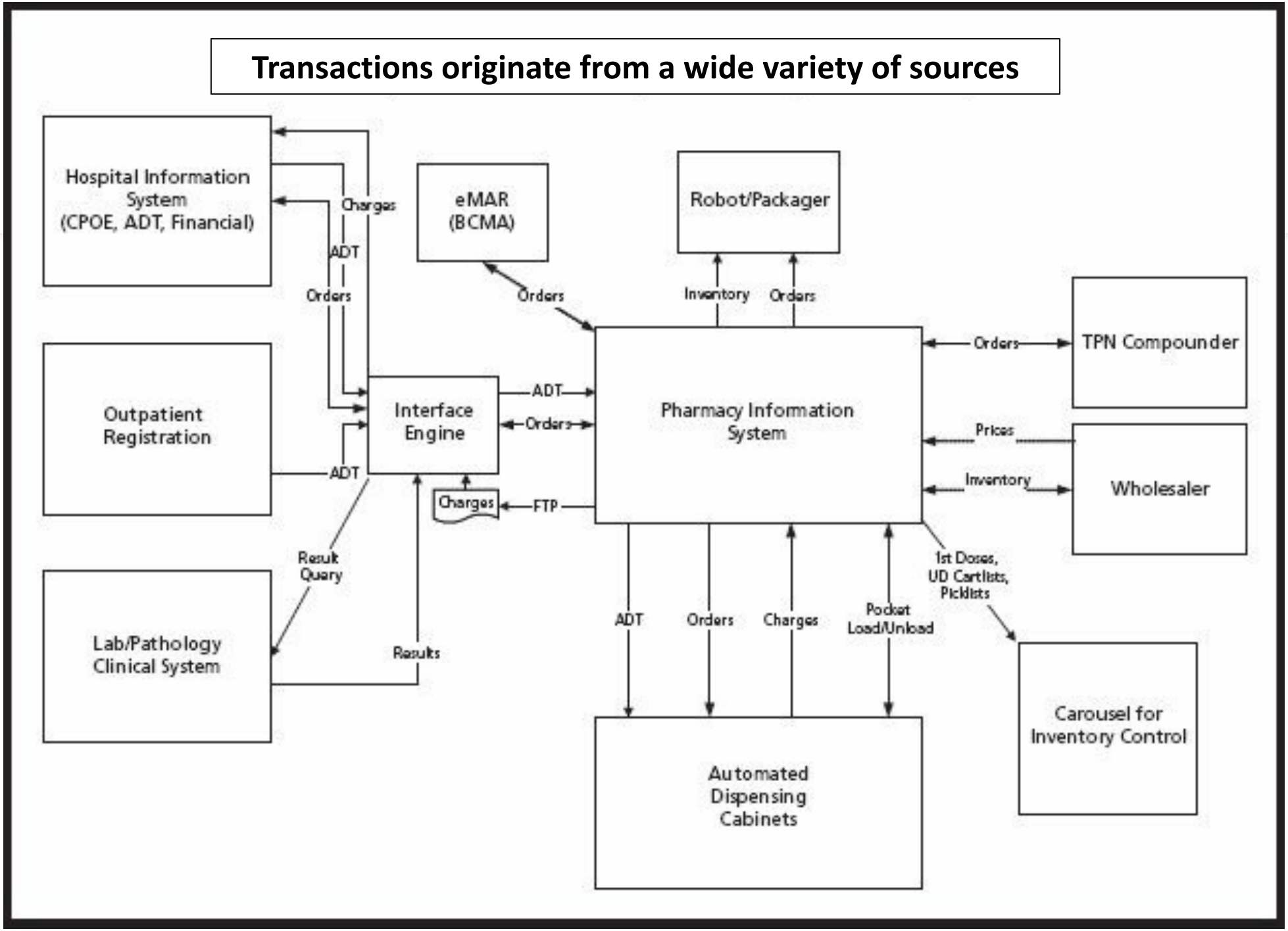
Much more process intensive to query

Example: raw text, video, audio recordings

Transactional Data: Structured Lab data entry

Collection Date/Time	Test	Result / Status	Flag	Units	Ref Range
Dec 09, 2019	25 OH VITAMIN D	63		ng/mL	30 - 100

Transactions originate from a wide variety of sources



Planning Transactional Data Extraction

1. Why do you need this information?
2. What resources do you have available?
3. When is your deadline and what is the project timeframe?
4. Who is your base cohort?
5. Where is your target data currently located and which data points do you need?

Methods to extract insights from transactional data

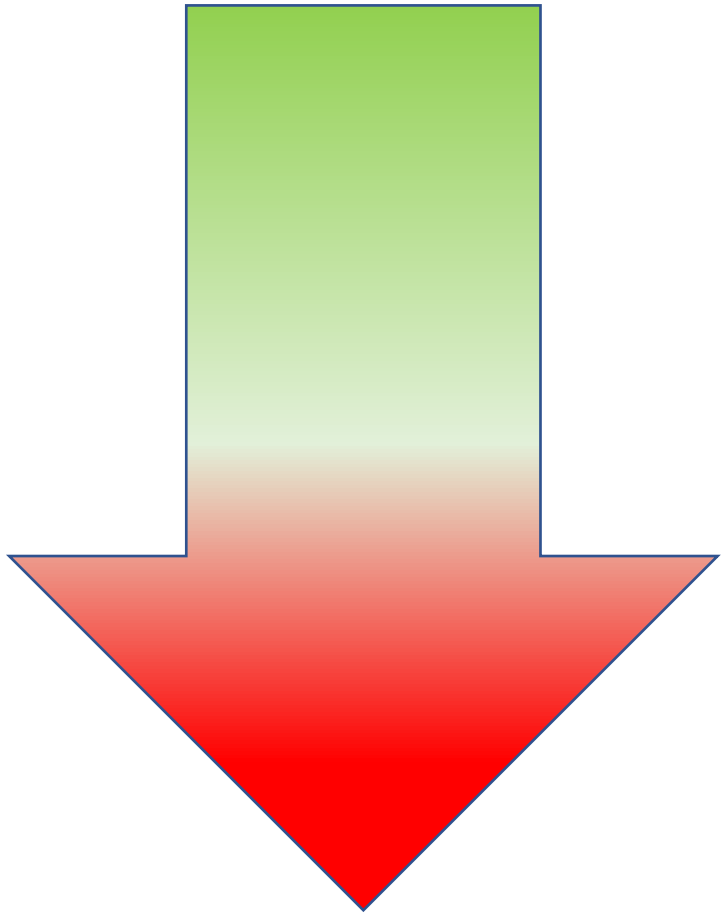
From least to most difficult:

Structured data:

1. Descriptive statistics
2. Statistical testing
3. Machine learning

Unstructured data:

1. Chart Review
2. Rule-based text mining
3. Advanced Natural Language Processing techniques

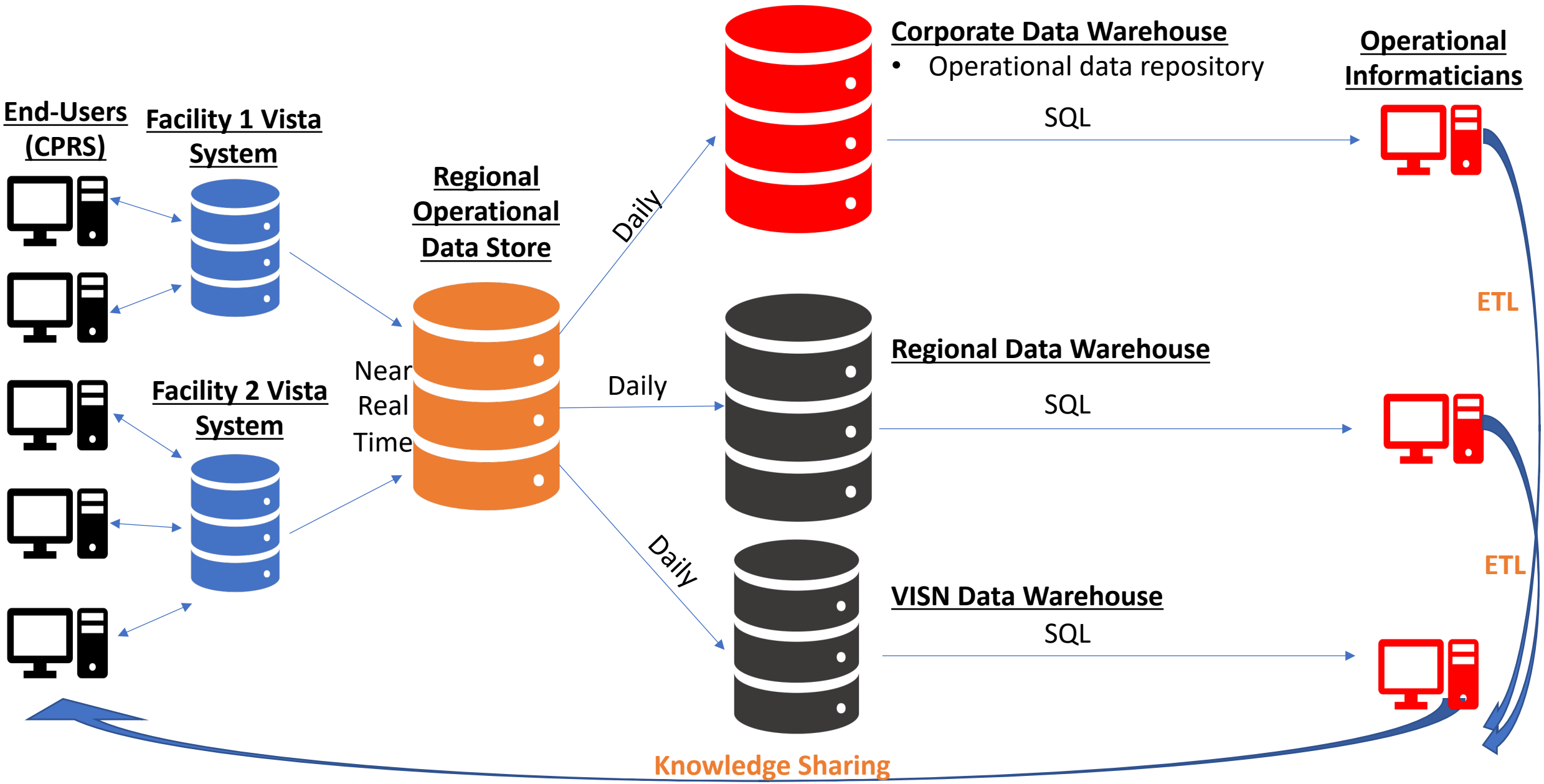


Where do all these transactions go?

The short answer: Data Warehouse(s)

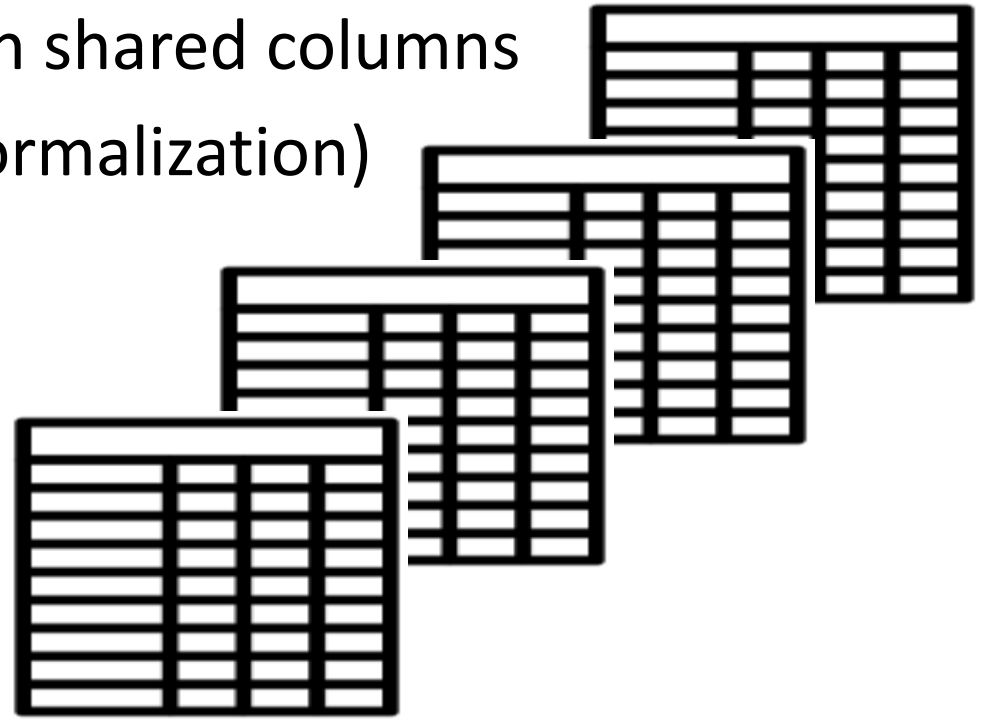
- Consists of many servers and databases
- Efficiently gathers, transforms, and stores health data
- Access to the data warehouse grants users the ability to provide accurate management information and supporting data analysis

VA Data Warehousing Architecture



The Relational Database

- A database is a repository of data
- Allows data to be added, modified, and queried
- Structured data is typically stored in tables similar to spreadsheets
- Tables can be joined/linked together through shared columns
- Goal is to store a piece of data only once (normalization)

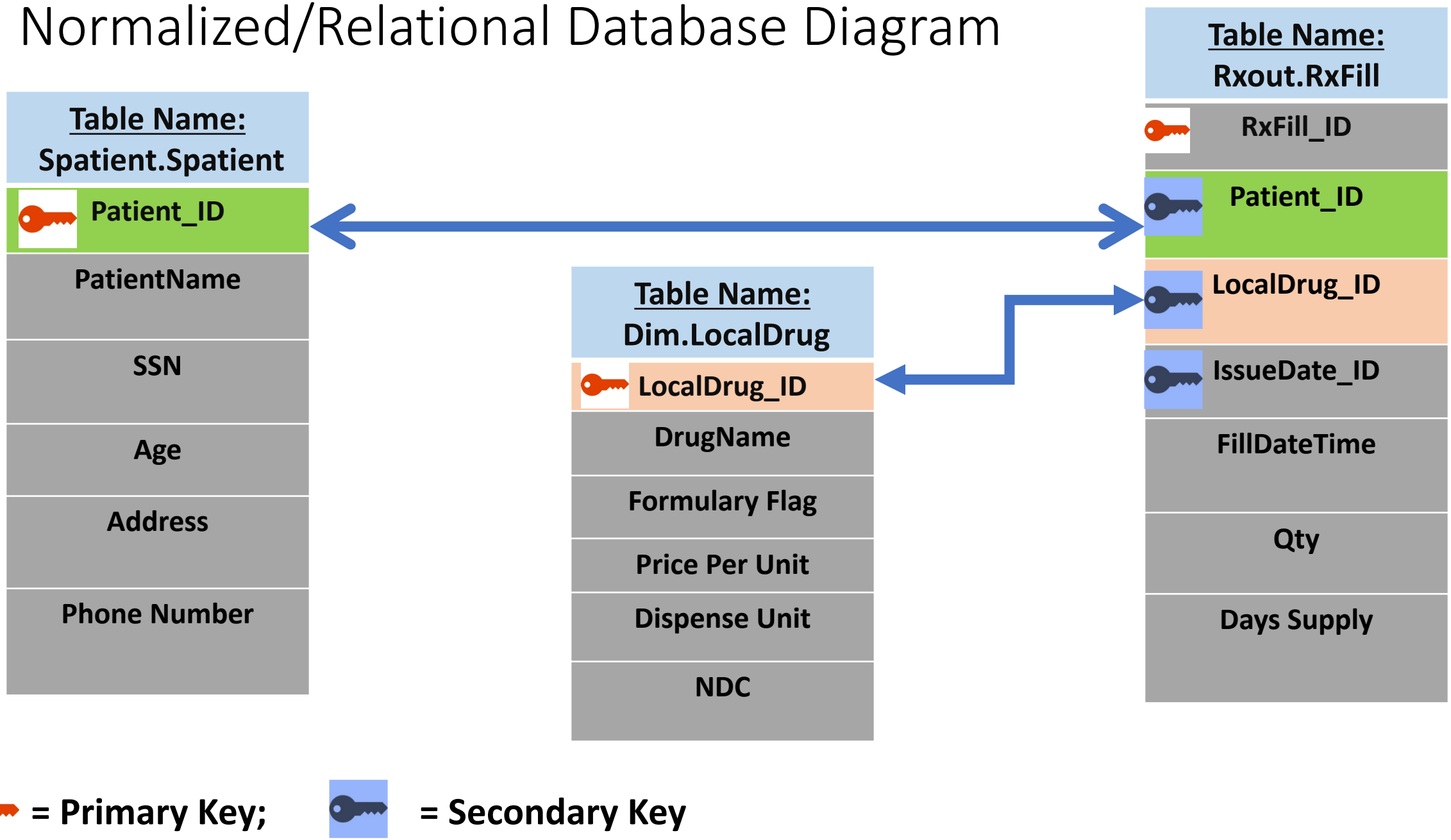


Denormalized Data

Table Name: Rxout.RxFill									
RxFill_ID	PatientName	Age	DrugName	IssueDate	FillDate	Qty	DaysSupply	FormularyFlag	PricePerUnit
1	Smith,John	55	Atorvastatin	2021-01-01	2021-01-05	30	30	Y	0.05
2	Smith,John	55	Atorvastatin	2021-01-01	2021-02-10	30	30	Y	0.05
3	Smith,John	55	Metformin	2021-01-01	2021-01-05	60	30	Y	0.01
4	Smith,John	55	Metformin	2021-01-01	2021-02-11	60	30	Y	0.01

- Note how the gold-colored columns contain duplicate data
- This duplication becomes non-trivial when this data model is applied to millions of patient records
- Increases required storage space and data processing time

Normalized/Relational Database Diagram



= Primary Key; = Secondary Key

Pharmacy Relational Database Diagram Example

Retrieving Structured Medical Data

Examples with practice database

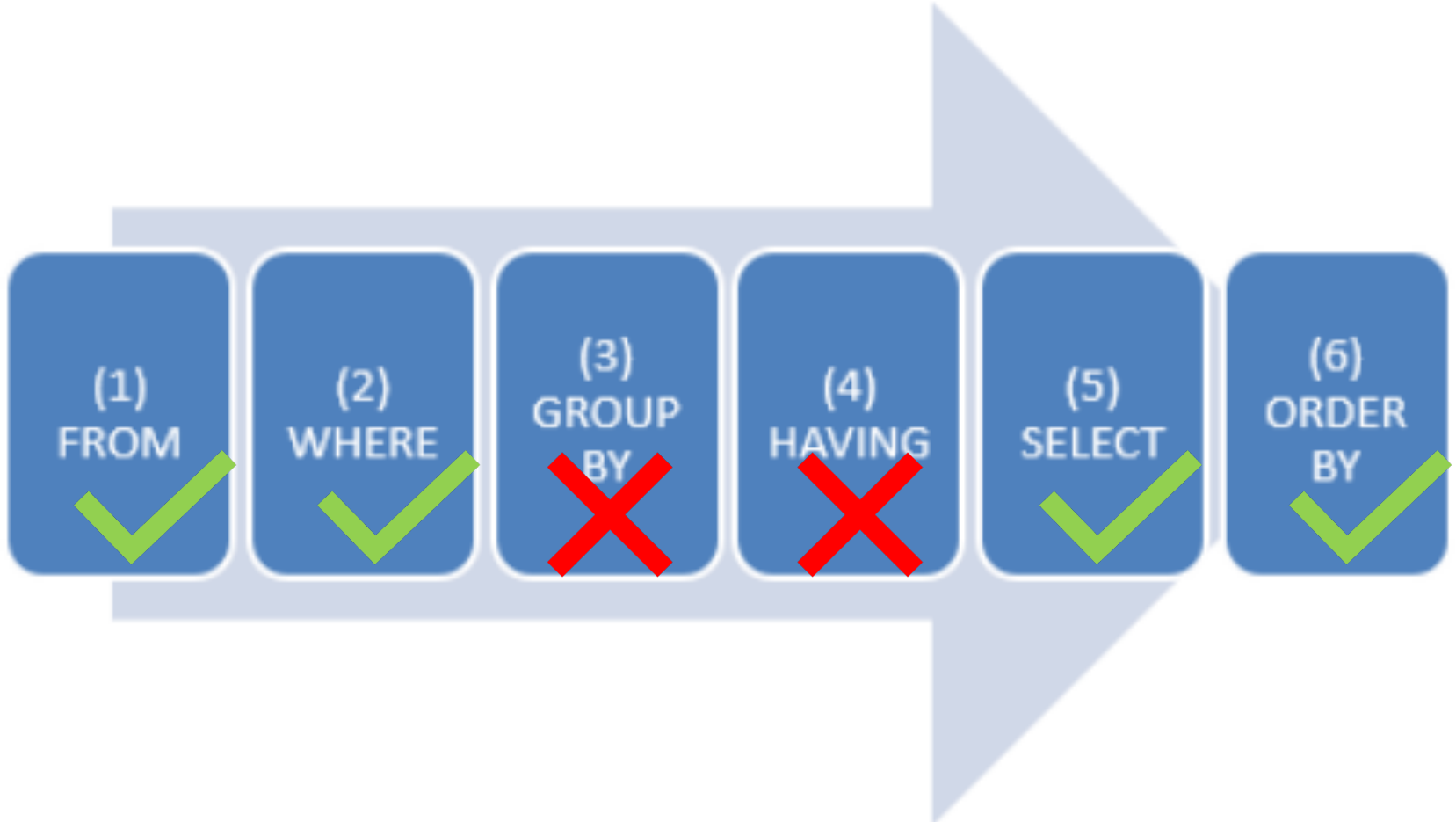
What is SQL?

- Structured Query Language
- Used to communicate with data within a database
- Code is written in an user-interface that is connected to the database
- Critical data operations:
  - Accessing
  - Updating
  - Inserting
  - Manipulating
  - Modifying

Basic structure of a SQL query

```
Select [Column1Name]
        , [Column2Name]
From [DatabaseName].[Schema].[Table]
Where ….[insert expressions for column filtering here]
Order by [ColumnName];
```

Logical Processing of SELECT



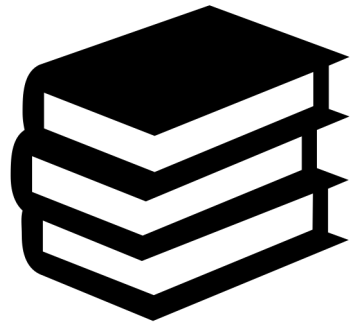
FROM Clause

Indicates which database table(s) the query will retrieve data from

Format:

FROM Database.Schema.TableName

```
Select *
From Pharm.mockpharmacydata
Where DispensedDate >= '2019-10-01’
Order by DispensedDateTime;
```



WHERE Clause Comparison Operators

Comparison	Symbol
Equal to	=
Greater than	>
Lesser than	<
Greater than or equal to	>=
Less than or equal to	<=
Not equal to	<>

```
Select *
From Pharm.mockpharmacydata
Where DispensedDate >= '2019-10-01’
Order by DispensedDateTime;
```

SELECT Clause

Indicates which database column(s) will appear in the results set

Tip: **Select *** means “select all columns”

```
Select [Column1]
      ,[Column2]
      ,……
From Pharm.mockpharmacydata
Where DispensedDate >= '2019-10-01’
Order by DispensedDateTime;
```

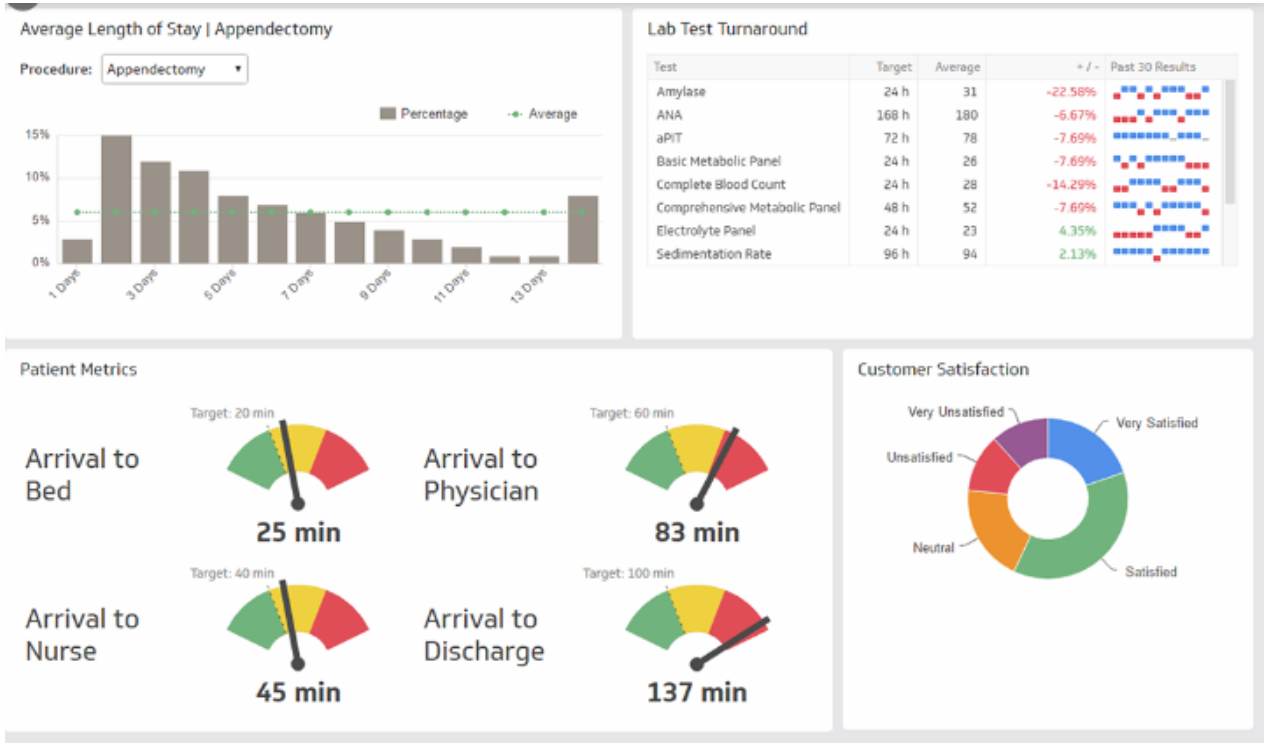
ORDER BY Clause

Indicates which column(s) the results will be sorted by

```
Select [Column1]
        ,[Column2]
        ,……
From Pharm.mockpharmacydata
Where DispensedDate >= '2019-10-01’
Order by DispensedDateTime;
```

Clinical Dashboards and Performance Reports

- Electronic reports can be built to share data pulled using SQL
- Many software applications exist:
  - SQL Server Reporting Services (SSRS)
  - Power BI (Business Intelligence)
  - Pyramid Analytics
  - Tableau



Key Points

- Pharmacists are well positioned to merge clinical and data knowledge
- Understanding the fundamentals of how data is captured and retrieved allows for rapid analysis of your patient population
- Structured Query Language (SQL) allows you to interact with a database
- Data can be presented to leadership/clinicians to support decision-making

Next step (optional)

- Follow the steps outlined in my GitHub repository to create your first database
- Follow the DataManipulationScript.sql to practice executing and writing queries
- Find open-source data online and import into your database.
  - Humana Insurance Cost Transparency for example
- Review the data, ask questions about the data, attempt to answer those questions using SQL