



The Office of the National Coordinator for Health Information Technology



# Introduction to Health Care Data Analytics

## Module 4: Data Analysis Tools and Techniques

### Lecture b, Databases Part I

This material was developed through a collaboration between Bellevue College and the Veterans Health Administration, U.S. Department of Veterans Affairs, funded in part by the Department of Health and Human Services, Office of the National Coordinator for Health Information Technology award number 90WT0002. Except where otherwise noted, this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

---

---

---

---

---

---

---

---

---

---

## Data Analysis Tools and Techniques Learning Objectives

- Define data analytics terms
- Describe the process steps of data analytics and the tools used in each step
- Describe the role of the data analyst
- Identify tools and techniques used to analyze and interpret health care data effectively
- Describe key database concepts.
- Describe the various types of databases and how they are structured
- Describe key data warehouse concepts
- Describe enterprise data architecture as seen in health care organizations

---

---

---

---

---

---

---

---

---

---

## Overview



This lecture focuses on the specific tools and techniques used in data analysis as related to databases.

---

---

---

---

---

---

---

---

---

---

Image by Stuart Miles, 2014

3

### Why Is This Important?

- Ever need something more than a simple spreadsheet to organize your data?
- Ever have difficulty trying to find data?
- Have a need to search for answers to relevant clinical or business questions?

---

---

---

---

---

---

---

---

### Prevalence of Databases



Images by Sira Anamwong, 2016; Shutterstock: ©Shamleean, ©Visual3Dfocus, ©Chalkom; photostock, 2011

---

---

---

---

---

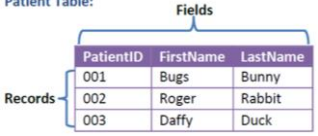
---

---

---

### What Is a Database?

- A database is a collection of data organized for a specific purpose and has the following properties:
- Consists of one or more tables
- A **row** in a table is called a **record**
- A **column** in a table is called a **field**




---

---

---

---

---

---

---

---

# What Is a Database Management System (DBMS)?

**Database Management System (DBMS)** – software package used to create, enter, and modify data and retrieve information from database



Capron & Johnston, 2004

7

---

---

---

---

---

---

---

---

# Types of Data Stored in a Table

Each field in a table has a data type.

Data Type	What is it?	Examples
Numeric	Integers Numbers with decimals Money (Currency)	458 2.567 \$45.00
Date and Time	Dates only Times only Dates and times	2016-03-24 09:51 2016-03-24 09:51
Character Strings	Different kinds of text strings, differentiated by how long they are	John Doe Congestive Heart Failure
Binary Strings	Different kind of binary data, including "true or false" types of data	Yes No True False

8

---

---

---

---

---

---

---

---

# Data Integrity

**Data integrity** is the degree in which data is accurate and reliable

- Integrity constraints are rules that data must follow.
- Example:
  - Month field (numeric) numbers 1 to 12 are acceptable values.
  - John enters a 13 into the field and receives an error message.



Capron & Johnston, 2004

Image by ©Stuart Miles, Shutterstock

9

---

---

---

---

---

---

---

---

## Metadata

- Data that describes the “properties or characteristics of data and the context of that data”
- Used to create programs, procedures, controls and queries to manipulate and manage the data in a database

Name	Type	Length	Description	Source
ID	Integer	10	Social Security Number	Registration
Name	Alphanumeric	30	Patient name	Patients
Appointment	Date and time	20	Appointment date and time	Appointments



Hoffer, Prescott & McFadden, 2007

Image by ©PlusONE, Shutterstock

10

---

---

---

---

---

---

---

---

## Data Dictionary

- A structure that stores metadata
- Used to control database operations, integrity and accuracy
- Contains the field name, data type, field size as well as validation rules which allows the DBMS to enforce integrity constraints



Image by ©Jatsomwang, Shutterstock

11

---

---

---

---

---

---

---

---

## Data Dictionary and Validation Rule Violation Example

Data Dictionary

Error Message or Alert

Validation rule (type of integrity constraint)

Images by K Brandt, 2016

12

---

---

---

---

---

---

---

---

# Data Models

- Used to capture the relationship among data within a database and are used in the conceptualization and design of the database
- Basic building blocks of all data models are **entities, attributes** and **relationships**



Hoffer, Prescott & McFadden, 2007; Rob & Coronel, 2004 Image by ©Singham, Shutterstock 13

---

---

---

---

---

---

---

---

---

---

# Entity, Attributes and Instances

- Entity:** person, place, object, event or concept, or thing (table)
- Attribute:** a characteristic of an entity (column or field in a table)
- Instance:** each row or record in a table

Logical/Relational	Physical Implementation
Entity	Table
Attribute (Field)	Column
Instance (Record)	Row

Hoffer, Prescott & McFadden, 2007 14

---

---

---

---

---

---

---

---

---

---

# Relationships

- Describes how two or more entities or tables are related within the database (Cardinality)
- Depicted in data model
- Several methods to notate relationships, i.e. Crow's Foot Model
- Relationships can be one to one (1:1), one to many (1:M) and many to many (M:M)

Common Relationships using Crow's Foot Notation

Relationship	Symbol
One to One	↔
One and Only One	⌈—⌋
One to Many	⌈—⌋
One to One or Many	⌈—⌋
Many to Many	⌈—⌋

Image by K. Brandt, 2016 15

---

---

---

---

---

---

---

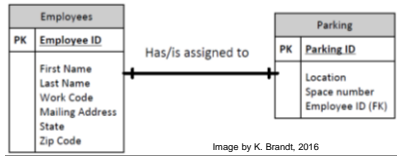
---

---

---

### Relationships – One to One (1:1)

- One to One: A single entity instance is related to a single instance in another entity
- Example: An employee is assigned one parking place. A parking place is assigned to one employee.




---

---

---

---

---

---

---

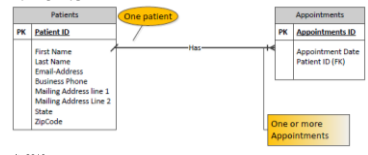
---

---

---

### Relationships – One to Many (1:M)

- A single entity instance in one entity type is related to many entity instances in another entity type.
- Most common
- Example: One patient has one or more appointments




---

---

---

---

---

---

---

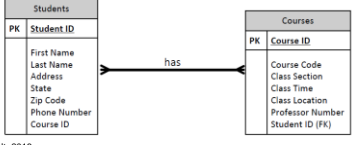
---

---

---

### Relationships – Many to Many (M:M)

- Many entity instances in one entity type are related to many instances in another entity type
- For more efficient design, uses a lookup table to link the many to many tables
- Example: Many students are registered for many classes or many classes contain many students.




---

---

---

---

---

---

---

---

---

---

# Building a Data Model

- Need type of data used, how they are used and in what time frames as well as the business rules of the organization
- Business rules:
  - Precise, non-ambiguous descriptions of policies, procedures or principles within a specific organization's environment
- Examples of business rules:
  - A patient can make many appointments to one or more clinics.
  - One customer can place many orders.



Rob & Coronel, 2004

Image by @docstockmedia, Shutterstock

19

---

---

---

---

---

---

---

---

---

---

# Summary

- A database :
  - Is a collection of data organized for a specific purpose
  - Is designed for use by individuals to whole organizations
  - Uses specific types of data with data integrity rules enforcement
- Data within the database is defined through metadata and data dictionary
- Conceptual database model:
  - Is a database structure using entities, attributes and instances
  - Depicts the relationships between entities
  - Follows business rules

20

---

---

---

---

---

---

---

---

---

---

# Data Analytics Tools and Techniques References – Lecture b

**References**  
 Capron, H. L., & Johnson, J. A. (2004). *Computers: Tools for an information age* (8th ed.). Upper Saddle River, NJ: Prentice Hall.  
 Hoffer, J. A., Prescott, M. B., & McFadden, F. R. (2007). *Modern database management* (8th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.  
 Joos, J., Nelson, R., & Smith, M. J. (2010). *Introduction to computers for healthcare professionals* (5th ed.). Boston, MA: Jones & Bartlett.  
 Rob, P., & Coronel, P. (2004). *Database systems: design, implementation & management* (6th ed.). Boston: Course Technology.  
 Shelly, G. B., Cashman, T. J. & Rosenblatt, H. J. (2003). *Systems analysis and design* (5th ed.). Boston: Course Technology.

**Images**  
 Slide 3: Miles, Stuart. (2014). Database Magnifier Shows Bytes Magnification And Computing. Retrieved from <http://www.freepressphotos.net/images/database-magnifier-shows-bytes-magnification-and-computing-photo-p294453>  
 Slide 5: Anamwong, Sira (2016). Shopping online stock photo. Retrieved from <http://www.freepressphotos.net/images/shopping-online-photo-p388857>

21

---

---

---

---

---

---

---

---

---

---

## Data Analytics Tools and Techniques References – Lecture b – Continued

Slide 5:

Shamleen (n.d.) Written word online college application on blue keyboard button. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic.mhtml?id=383615734&src=lb-42520639>

Visual3Dfocus (n.d.) Digital library - Colored books inside computer in the design of the information related to online education and training. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic.mhtml?id=3847316598&src=lb-42520639>

Chaikom (n.d.) Illustration of buildings in the city. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic-436509937/stock-vector-illustration-of-buildings-in-the-city.html?src=WUlvKIVedbogToAMml7zTw-1-35>

Photostock (2011). Smiling young doctor working on his desk stock photo. Retrieved from [http://www.freedigitalphotos.net/images/Healthcare\\_g355-Smiling\\_Young\\_Doctor\\_Working\\_On\\_His\\_Desk\\_p34030.html](http://www.freedigitalphotos.net/images/Healthcare_g355-Smiling_Young_Doctor_Working_On_His_Desk_p34030.html)

Slide 6, 12, 14-18: Brandt, K. (2016).

Slide 9: Miles, Stuart. (n.d.) Data integrity meaning honourable knowledge and facts. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic-222100363/stock-photo-data-integrity-meaning-honourable-knowledge-and-facts.html?src=YA3Rl6UE7Iq-danzQYng-1-d>

Slide 10: PlusONE (n.d.) Blue tag metadata texts with other related keywords in word tag cloud design for web concepts. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic-233145829/stock-photo-blue-tag-metadata-texts-with-other-related-keywords-in-word-tag-cloud-design-for-web-concepts.html?src=7nYjkQ1ZzGHjdtXcQPP6A-1-3>

22

## Data Analytics Tools and Techniques References – Lecture b – Continued 2

Slide 11: jehsomwang (n.d.) Infographic set vector. Purchased from Shutterstock. Retrieved from [http://www.shutterstock.com/pic-153454277/stock-vector-infographic-set-vector.html?src=pp-same\\_artist-173366498-Yl6kBzWxts-R\\_DNCxtAV\\_w-2](http://www.shutterstock.com/pic-153454277/stock-vector-infographic-set-vector.html?src=pp-same_artist-173366498-Yl6kBzWxts-R_DNCxtAV_w-2)

Slide 13: Singkham (n.d.) Business woman drawing entity relation diagram (ERD) and database design. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic.mhtml?id=178462571&src=lb-42520639>

Slide 19: Docstockmedia (n.d.) Rules business concept puzzle with female hand and text. Purchased from Shutterstock. Retrieved from <http://www.shutterstock.com/pic.mhtml?id=350938112&src=lb-42520639>

23

## Introduction to Health Care Data Analytics: Data Analytics Tools and Techniques Lecture b

*This material was developed through a collaboration between Bellevue College and the Veterans Health Administration, U.S. Department of Veterans Affairs, funded in part by the Department of Health and Human Services, Office of the National Coordinator for Health Information Technology award number 90WT0002.*

Version 1.0/Fall 2016

24