

## DATA MANAGEMENT

### SHOWING OUTCOMES WITH DATA

Danny Limoges, PharmD  
Vaccine & 340B Coordinator  
St. Bernard Hospital and Health Care Center

*The presenter has no conflicts of interest to declare.*



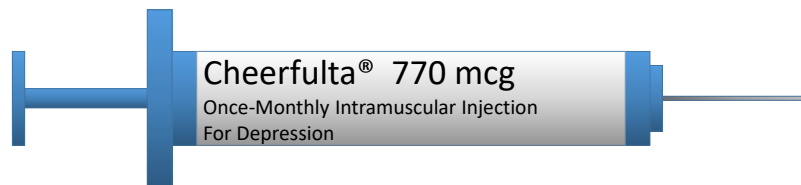
## Objectives

- Explain the importance of standardized data collection
- Identify available data sources and reporting resources
- Describe examples of using charts and graphs to show outcomes with data



## Standardized Data Collection

Consider a scenario:

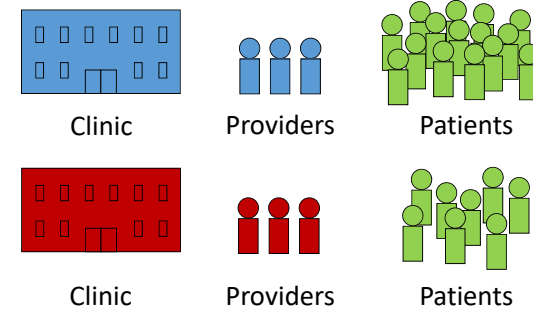


A new weight-based dosing antidepressant,  
approved for Major Depressive Disorder.



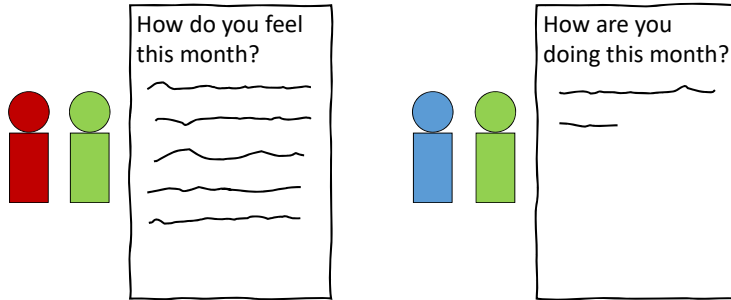
## Standardized Data Collection

Consider a scenario:



## Standardized Data Collection

Consider a scenario:



How do you feel this month?

How are you doing this month?

## Standardized Data Collection

In what ways did the **form** impact how standard the data collection was?

## Standardized Data Collection

In what ways did the **environment** impact how standard the data collection was?

## Standardized Data Collection

How might the results of this study be skewed based on these factors?

## Standardized Data Collection

### Definitions

#### Primary Data

Information that is commissioned and gathered for the purpose of the study.

#### Secondary Data

Information that was not gathered expressly for the study.



## Standardized Data Collection

### Four Characteristics of standardized data:

- Complete
- Generalizable
- Reliable
- Valid



## Standardized Data Collection

### Complete

How does incomplete data collection alter the results of a study?

What pieces of data are required to accurately study the question proposed?

Is the data recorded frequently enough to eliminate bias?

Examine a sample of the data set before attempting the study if possible.



## Standardized Data Collection

### Generalizable

In what context is the data recorded?

Location?

Demographics?

Time Period?

Methods of measurement?

Does the context support the degree of generalization intended?



## Standardized Data Collection

### Reliable

Are the results repeatable?

What factors may have influenced the repeatability of the data?

Do the methods of recording data encourage consistency?



## Standardized Data Collection

### Valid

What is the accuracy of the method of measure?

Is there researcher bias?

Does the data mean what you think it means?



## Data Sources and Reporting Resources

Standard Data Sources:

- Medical Records
- Internal Electronic Databases
- External/commercial billing databases
- Patient registries
- Government databases

Kauffman YS, Witt DM. The Essential Guide to Pharmacy Residency Research. American Society of Health-System Pharmacists; 2020. p. 62



## Data Sources and Reporting Resources

Public Databases:

- <https://healthdata.gov/>
  - State Drug Utilization Data
  - COVID-19 Impact and Community Vulnerability
  - De-identified Inpatient Discharges
- World Health Organization
- openFDA

10 Best Healthcare Data Sets (Examples). Archer-soft.com. Published February 19, 2020. Accessed August 9, 2021. <https://archer-soft.com/blog/10-best-healthcare-data-sets-examples>



## Data Sources and Reporting Resources

### Public Databases:

- Data and Tools of the National Center for Health Statistics
- data.gov
- Medicare.gov
- HCUP – inpatient stays and ambulatory surgery in US Hospitals <https://hcup-us.ahrq.gov/databases.jsp>

10 Best Healthcare Data Sets (Examples). Archer-soft.com. Published February 19, 2020. Accessed August 9, 2021. <https://archer-soft.com/blog/10-best-healthcare-data-sets-examples>



## Data Sources and Reporting Resources

### Accessibility of Data Sources

#### Factors indicating good availability of data:

- Existing report
- Existing data fields / measures
- Already-standardized and defined measures
- Open platform (Excel/Access)



## Data Sources and Reporting Resources

### Accessibility of Data Sources

#### Factors indicating more time needed to gather data:

- Custom report to gather data
- Custom data fields / measures
- Measures not previously standardized / well-defined
- Closed / limited platform (EMR / Administrative Data)



## Data Sources and Reporting Resources

What programs do you use for data or creating reports?



## Data Sources and Reporting Resources

- Most accessible reporting resources
  - Microsoft Excel
  - Microsoft Access
  - Power BI



## Data Sources and Reporting Resources

- Most accessible reporting resources
  - Microsoft Excel
  - Microsoft Access
  - Power BI



## Data Sources and Reporting Resources

- Larger Scale
  - SPSS / SAS / Tableau
- Programming-focused (high control, automation)
  - SQL query
  - Python in a Jupyter Notebook  
(with Pandas / Plotly libraries)



Which chart/graph type do you wish you knew how to use better?

- A. Bar graph
- B. Pie Chart
- C. Line Graph
- D. Box and Whisker plot
- E. Scatter Chart
- F. Heat Map



## Chart Vs. Graph

- Use a graph if...
  - Data shows pronounced trends
  - Paints an interesting picture
  - Reveals relations between variables

Ng KH, Peh WC. Preparing effective illustrations. Part 1: graphs. Singapore Med J. 2009;50(3):245-249.

## Chart Vs. Graph

- Use a graph if...
  - Data shows pronounced trends
  - Paints an interesting picture
  - Reveals relations between variables

Ng KH, Peh WC. Preparing effective illustrations. Part 1: graphs. Singapore Med J. 2009;50(3):245-249.

## The Thesis of a Graph

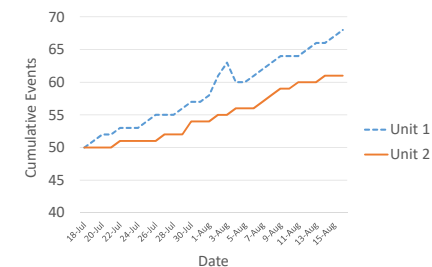
A graph must communicate a clear thesis within several seconds.

Examples:

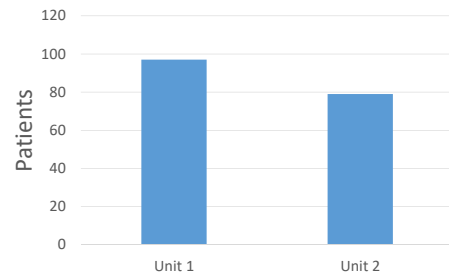
- **A** rises faster over time than **B**.
- Most patients are in groups **A** and **B**.
- Treatment **A** performs significantly better than **B**.

## Line Graph

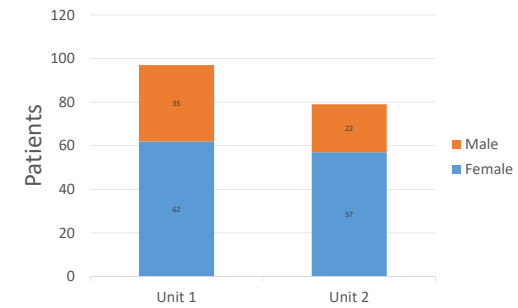
Related groups changing over time, compared to one another.



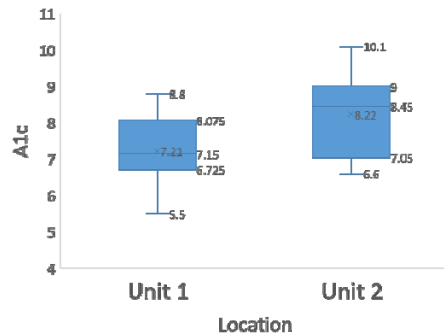
## Bar Graph



## Stacked Bar Graph

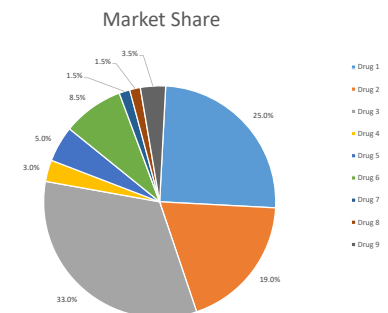


## Box and Whisker Plot



## Pie Chart

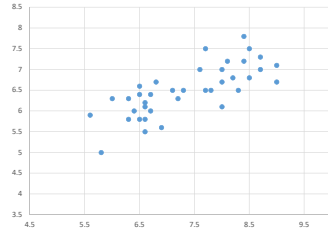
- Consider whether this form is best
- Avoid 3D Effects
- Group large segments together
- Ensure labels are clear





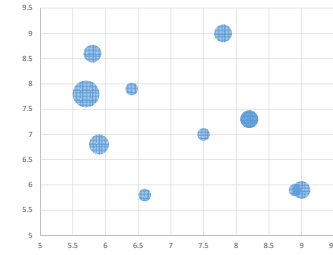
## Scatter Chart

When many observations create a shape and it may be important to depict clustering patterns.



## Bubble Chart

Depicting three dimensions of data.



## Heat Map

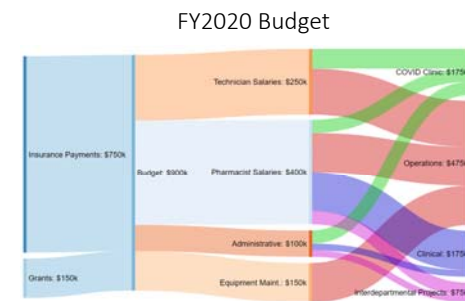
Depicting patterns in large amounts of data.

Clinical Interventions	Hospital	2020 Q1	2020 Q2	2020 Q3	2020 Q4	2021 Q1	2021 Q2
PICU	1	150	165	170	162	168	175
MICU	1	208	150	180	212	220	225
Floor 2	1	130	132	117	145	180	141
Floor 3	1	104	105	121	145	140	128
Floor 4	1	129	109	150	115	113	115
Floor 5	1	101	123	148	120	119	120
Floor 1	2	91	87	118	112	97	96
Floor 2	2	116	70	86	89	110	116
Floor 3	2	82	92	96	135	120	80
Floor 4	2	86	114	100	85	90	106
Floor 5	2	112	94	104	116	105	84
Floor 1	3	55	66	62	64	59	70
Floor 2	3	66	50	60	75	71	51
Floor 3	3	40	45	44	40	54	50
Floor 4	3	63	50	75	75	77	56

## Sankey Diagram

Complex progressions:

- Characteristics of patient groups
- Clinical cohorts
- Budgets



Generated with fictional data by Danny Limoges, PharmD using [www.sankeymatic.com](http://www.sankeymatic.com)

## References

- Kauffman YS, Witt DM. The Essential Guide to Pharmacy Residency Research. American Society of Health-System Pharmacists; 2020.
- 10 Best Healthcare Data Sets (Examples). Archer-soft.com. Published February 19, 2020. Accessed August 9, 2021. <https://archer-soft.com/blog/10-best-healthcare-data-sets-examples>
- Ng KH, Peh WC. Preparing effective illustrations. Part 1: graphs. Singapore Med J. 2009;50(3):245-249.
- Randle T. What do you mean I'm not supposed to use Pie Charts?! Geckoboard.com. Published August 17, 2015. Accessed August 2, 2021. <https://www.geckoboard.com/blog/pie-charts/>
- Bogart S. Sankey Diagram Generator. Accessed August 2, 2021. <https://sankeymatic.com/>



## Questions?



## Learning Assessment Questions

### Data Management – Showing Outcomes Using Data

1. Which of the following represents the most standardized method of data collection?
  - a. Providers write a free text description about the side effects of a medication every time patients voice a complaint about medication.
  - b. Blood pressures are taken by nurses, using different brands of blood pressure cuff, and are recorded in the EHR.
  - c. Unvalidated claims data from hospitals in two different countries are examined to understand which medications physicians prefer.
  - d. Patients are asked to state two adjectives that describe their pain level, which are entered in a free-text field in a form.
2. Which of the following best describes a well-made graph?
  - a. Surprising and persuasive
  - b. Large and detailed
  - c. Clear and accurate
  - d. Elegant and colorful
3. Which of the following would be best depicted as a graph rather than a data table?
  - a. Yearly data on the budget allocated to a certain department, which has been the same for 5 years.
  - b. A series of hundreds of data points, scattered in a seemingly random pattern.
  - c. Several dozen demographics about a patient population.
  - d. Hundreds of data points of individual patients, describing patient age and a1c that form a pattern.
4. Which of the following programs would work best to summarize a large, unfiltered file of lab results (1 gigabyte).
  - a. Excel
  - b. Word
  - c. SPSS
  - d. Power BI

Answer key:  
1. B 2. C 3. D 4. C