Brief Intro

WHAT IS YOUR NAME/WHERE FROM?

WHAT IS YOUR INTEREST IN DATA VISUALIZATION?

WITH WHAT INFORMATION DO YOU WANT/EXPECT FROM THIS WORKSHOP?
Definitions

WHAT DO WE MEAN WHEN WE TALK ABOUT DATA VISUALIZATIONS?
infographics
media
design
big
analytics
plots
information
interactive

science
cart
graphics
multi
data
visualization
Discussion Questions

• What is information science and how is it different from data analytics, data science, big data and data visualization?
• How are infographics, graphics, and data visualizations different?
• What do data science and data analytics have to do with data visualization?
• How are static and interactive data visualizations different?
Terminology

Structuring the parts

**Data**
- Information Science
- Data Science
- Big Data

**Design**
- Information visualization
- InfoVis
- Infographics
- Multi-Media
- Static Data visualization
- Interactive Data Visualization
Data Visualization in Theory and Practice

WHY USE DATA VISUALIZATIONS?
What do you think the advantages and disadvantages are for using interactive data visualizations?
Amplified Cognition
- Utilize working memory
- Card, Mackinlay & Schneierman (1999)
- Liu, Nersessian & Stasko, 2008

Distributed Cognition – Theory for information visualization
- “Push the boundary of the unit of cognitive analysis beyond a human individual”
- “Cognition is a property of interaction”
- Liu, Nersessian & Stasko, 2008
Practical Reasons for using Data Visualizations

- Explore
  - Explore Data
- Answer
  - Answer questions
- Support
  - Support an argument
- Explain
  - Explain – just the facts
Strengths of Interactive Data Visualizations

Data modeling

1. Linking multiple, disparate data sources. Ideas?
2. Performs well for very large datasets.
3. Allows multiple perspectives of data.
4. Perfect for online publishing or for a computer environment.
Planning

STARTING A DATA VISUALIZATION PROJECT
What we want to think about here is:

1. how does what we know about interactive data visualizations inform our decision making around audience/goals/questions and,

2. how does our audience/Goals and Questions impact what kind of visualization we choose to use?
Brainstorm ideas.

Engage your curiosity.

Have fun thinking about your topic!

Make notes.

Share your ideas!

Who is your intended audience?
- General population
- Students
- Academics and scholars
- Administrators
- Adults/college/K-12

What are your goals for the data visualization?
- Collaboration with peers
- New perspectives on a set of data
- Teaching/instructional design
- Generating more questions
Generate Questions

What question(s) do you want to answer with the data visualization? Be specific. Do you WANT to answer a questions?

What data do you need to have to answer the questions or what data will you need to explore your topic? What it you don’t know?

Where will you get the data (all sources)?
- Personal research
- Web scraping
- Online databases
- Non-profit or government organizations
- Others?
IDENTIFY SCOPE

- Audience
- Goals
- Questions
Sample Questions

01. How did Haydn’s music change over time?
   - Musical forms?
   - Genres?
   - Instrumentation?
   - Tonality

02. How productive was Haydn over his lifetime?

03. How does Haydn compare to other composers of his time?
   - Across Geography?

04. What was going on in the world during the periods of his life and composing?
Identify Data Points

Work from a spread sheet or draw a chart on paper and identify each data point.

Figure out where you will get the information.
- Think through questions and fill out worksheet
- Get together in small groups and share your ideas
- Discuss how an interactive data visualization will help meet your goals.
- Work through listing specific data sources on worksheet
Extract, Transform, Load
• Demo tools:
  • Excel
  • Power BI
• Where and how to pull data
  • File types
  • Databases
  • Web pages
  • API’s
• Other Tools
  • Tableau
  • R
  • Databases such as SqlServer, mysql or MS Access
  • Online Tools such as DataBricks, Data.World, etc.
TRANSFORM

Cleaning data
Modeling

- Column Names
- Table Names
- Data Types
- Splitting columns where needed
- Adding index field
- Verify consistency of data
- Add columns for tracking who and when data added, active records, soft deletes, etc.
- Will this be a one-time process or will it be repeated as more data becomes available?
• Go ahead. Click the Button.
• It’s that easy!
Online Data Sources

https://www.propublica.org/datastore/
https://data.world/
https://www.kaggle.com/datasets
https://www.data.gov/
https://github.com/awesomedata/awesome-public-datasets
https://data.fivethirtyeight.com
https://www.ncdc.noaa.gov/cdo-web/
https://ieeepart.org/
https://www.r-project.org/
https://www.gutenberg.org/wiki/Main_Page
Imagine the Visualization

TRY OUT DIFFERENT PRESENTATION IDEAS
Who Taught Whom?

Classical music influence from Salieri to Stockhausen

Drag to navigate. Two-finger pinch to zoom.

Click and hold a node to highlight ancestors (teachers, their teachers, etc) and descendants (students, their students, etc).

Get Inspired!
How do you Visualize Equality?

The Equal Measures 2020 project aims to bring together and disseminate data on violence and gender-based violence on an ongoing basis around the world. As a step towards this goal, the DIVA 2020 Global Gender Violence Survey asked 518 gender-based violence experts to rate 12 issues that should be prioritized in the quest for better data.

Get Inspired!

The Colors of The Starry Night
Click any color(s) below to highlight painting by color, or highlight the painting to see the color composition.

The most common color is **Silver** which comprises 21.8% of the painting.
Get Inspired.
Where to start?

COMPARISONS
CAUSALITY, MECHANISM, STRUCTURE, EXPLANATION
MULTIVARIATE ANALYSIS
INTEGRATION OF EVIDENCE
DOCUMENTATION
CONTENT

Standards

1) data ink and graphical redesign,
2) Chart junk
3) Data Ink
4) Multifunctioning graphical elements,
5) High resolution data graphics and

Interaction

1) aggregate
2) zoom
3) Specific data selection
4) Multiple comparisons
5) Comparisons
6) Changes over time

Select a Tool

Microsoft Excel
- Data tab add-on (for older versions)
- Power Query (add on extension)
  - https://www.excelcampus.com/install-power-query/

MS Power BI
- Desktop version is free with professional email (not gmail, yahoo etc.)
- Online Version & Community workspace
  - Community workspace is not free but licenses may be acquired through employer if it is set up
  - https://www.powerbi.com
Select a Tool

Tableau

- Free Version for students and academic pricing
  - https://www.tableau.com/academic
  - https://www.tableau.com

Other Tools

Data Bricks
Data.World
Kaggle.com
Create and Publish

From Power BI

1. Open desktop version of power bi and click on the “Publish” icon.
2. Select “my workspace”
3. Sign if asked. Use same credentials used to download power BI desktop
4. Go to “my Workspace” and select the “reports” tab
5. open the visualization you want to publish
6. Click on the “File” tab and select “Publish to Web”
7. Copy and paste the embedded code snippet into your blogpost or web page.
Now What?

Practice
Use the worksheet to plan different projects
Work with different tools to see what works best for you.
Most tools have an online community of users who help each other learn how to use it.

Design
Collaborate
Email me if I can help answer any questions.
Share your work back with me so I can learn from what you have learned.

cathygreen@okstate.edu