Midterm 2 Exam

INFO 4602 Information Visualization
Evan Rice and Chloe Wolz

Overview

What data does the user see?

- In our original sketches from HW3 the user sees a heat map which conveys the risk of covid-19 in areas around the U.S. In our visual from HW4, the user sees an interactive prototype where they are able to customize their results based on age and pre-existing health conditions, in order to find a state and country that feels the most safe to them.
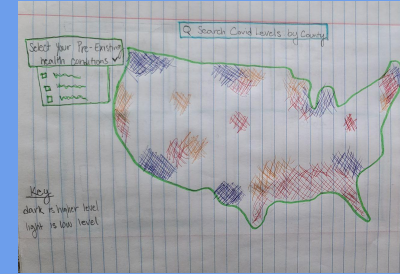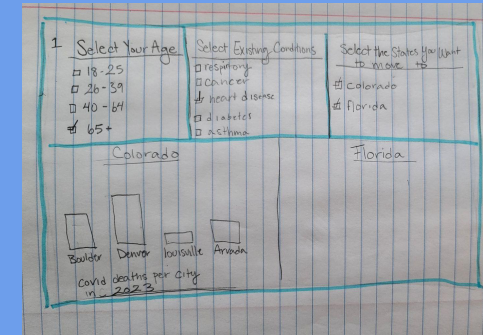
Why would someone use the visualization?

- Our visualizations would be used for people who are looking to move to a different state in the U.S, and have pre existing health conditions and is concerned about covid. The user would be able to compare states covid-19 statistics based on county, which they could use to make an informed decision about where they should moved based on their age, pre existing condition, states they are comparing, and the covid data based on country.

Sketch from HW4:

How has the visualization been designed?

- This visualization has been designed with the 65+ , at risk community in mind. We want to focus on creating a interface where 65+ adults can compare two states, in our example, Colorado and Florida, and see covid-19 statistics in counties in both states. This creates an easy and effective way for older adults with pre-existing conditions, to be able to feel like they can confidently move states with safety being a main priority.

What?

County	Residents with at least one dose	Residents with a completed primary series	Percent of total pop with at least one dose	Percent of total pop with a completed primary series	Residents 65+ with at least one dose	Residents 65+ with a completed primary series	Percent of 65+ pop with at least one dose	Percent of 65+ pop with a completed primary series	People 65+ with a bivalent booster dose	People 65+ with a bivalent booster dose	Percent of total pop with a bivalent booster dose	Percent of 65+ pop with a bivalent booster dose
Adams County	403191	370106	77.9	71.5	56854	53711	95	95	86897	28170	16.8	50.7
Alamosa County	10646	9487	65.6	58.4	2391	2240	95	95	2187	1001	13.5	44.4
Arapahoe County	514088	469111	78.3	71.4	88373	82621	95	93.3	131689	46543	20.1	52.6
Archuleta County	9367	8378	66.8	59.7	3713	3384	95	88.3	2634	1594	18.8	41.6

County	Percent of total pop with at least one dose
Adams County	77.9
Alamosa County	65.6
Arapahoe County	78.3
Archuleta County	66.8
Baca County	42.5
Bent County	28.7

What are your Data and Dataset Types?
- Our data and dataset types consist of two heatmaps, one bar graph and one data table. Our visualizations will incorporate raw data which can be accessed as data tables

What is the Dataset Availability?
- We are using data from the CDC website and all of our datasets are available for download using this link:
- Main link: https://stacks.cdc.gov
  - https://www.cdc.gov/surveillance/resp-net/dashboard.html
  - https://covid.cdc.gov/covid-data-tracker/#county-view?list_select_state=Colorado&data-type=CommunityLevels
  - https://data.cdc.gov/Public-Health-Surveillance/United-States-COVID-19-County-Level-Data-Sources/7pvw-pdbr
  - https://data.cdc.gov/Case-Surveillance/COVID-19-Case-Surveillance-Public-Use-Data-with-Ge/n8mc-b4w4

What are the Attribute Types?
- The attributes types consist of categorical and numerical data for the majority of our results. We will also be using some ordinal and continuous data to help tell our story effectively.
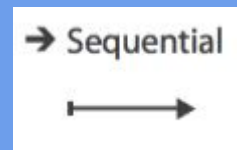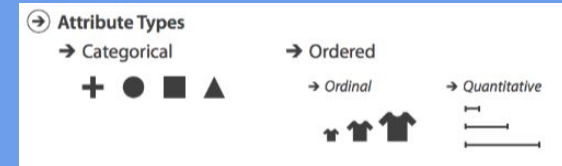
What is the Ordering Direction?
- Our ordering direction will be based on time or from least to greatest numerical values.
- Right to left
- Color intensity from light to dark

Attribute Types
→ Categorical
→ Ordered
→ Ordinal
→ Quantitative

→ Sequential

Why?

- **Why is the visualization being used by the user?**
  - The visualization is being used by the user in order to help find make an informed decision about what state they should move to as older adults who have existing health conditions.
- **What is the purpose?**
  - To help the stakeholder, in our case someone who is over the age of 65, choose the safest place to live between Colorado and Florida based on the current community Covid-19 levels.
- **Why did we create this visualization?**
  - To present and identify the filtered data in a comprehensible way and format easy for the direct stakeholder to understand and interpret.
- **How is the user navigating?**
  - They would be using their mouse if we make it an interactive visualization with mouse over. Most likely there will be a dashboard which they will be able to view the key covid-19 highlights between each county from each state.
- **What is the user looking for?**
  - Concrete answers to the research question: Where would be the better place to live if you are aged 65 or older between those two states, based on covid-19.
- **What does a user seek to learn about the data?**
  - The user seeks to learn more about the places they might live or currently live in. It gives an abstract of the larger picture of covid-19 within the area and can help users understand more about the community they are a part of. They are seeking to see where is the safest place to live comparing Colorado and Florida without covid-19 impacting their life.
- **What are the attribute types and ordering direction?**
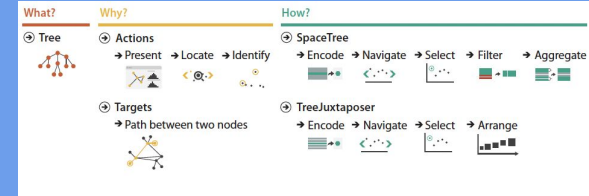 - Attribute types and ordering direction can be found on the previous slide.

How?

How can a vis idiom be constructed out of a set of design choices?

- **How to encode data?**
  - In this project we will encode data to show Covid rates for people ages 65+ for Colorado and Florida. This will be done by changing our data into a new format using a scheme. For example, we have data from multiple sources for two different states. We can concatenate the data sources and encode their counties to represent their state numerically.
  - Lots of visual encoding. We will use position, shape, color, size, and value to encode the data to more effectively present the data to the stakeholder.
- **How to manipulate Data?**
  - Manipulate the data to clean it up, simplify it, make it easier to use to answer questions related to our project. We will remove null values and values that don't make sense. We will concatenate the data from our different sources to be able to answer the research question.
- **How to facet data?**
  - We can take any of the datasets and facet by row or column to see the distribution of the data in its different facets at a glance. This might be helpful in visualizing covid age related information or vaccine immunization information across counties.
- **How to reduce data?**
  - To reduce our data to contain only what we need to present an answer to our stakeholder, we will need to remove useless information that doesn't relate to answering the question. This could be columns for smoking habits, drinking habits, blood type, etc.. which are found in our datasets. We also can remove age related data for ages which do not relate to the research question.
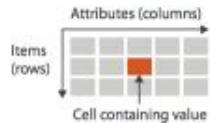
Analyzing What-Why-How

Example ➡

Idiom What-Why-How

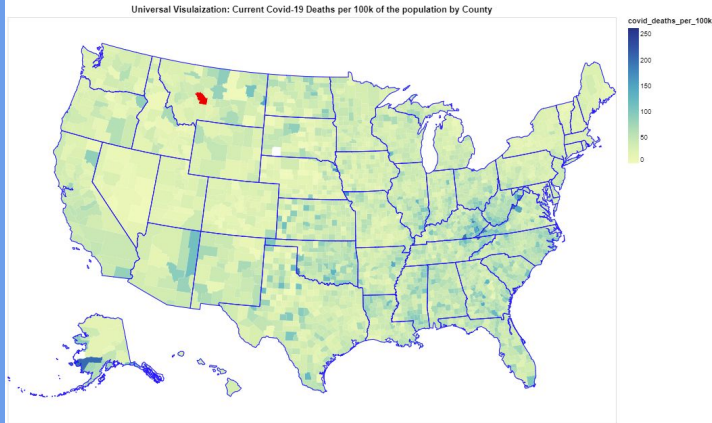Total COVID-19 Vaccine vs 65+ COVID-19 Vaccine
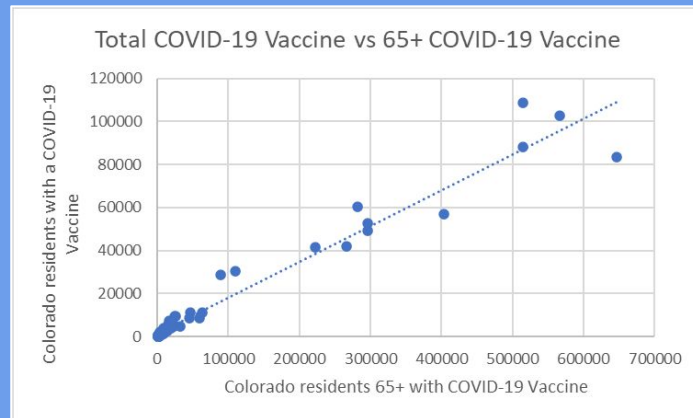
Idiom	Scatterplots
What: Data	Table: two quantitative value attributes.
How: Encode	Use dots to represent values using vertical and horizontal spatial positioning for two different numeric variables.
Why: Task	Visualize the correlation of two or more measures at the same time, including trends, outliers, correlation, and distribution.
Scale	People: Hundreds of thousands

Stakeholder: Retired Couple, Ages 65+, Who Have Existing Health Conditions and are Looking for a State and County to Reside in With Low Covid-19 Death Rates
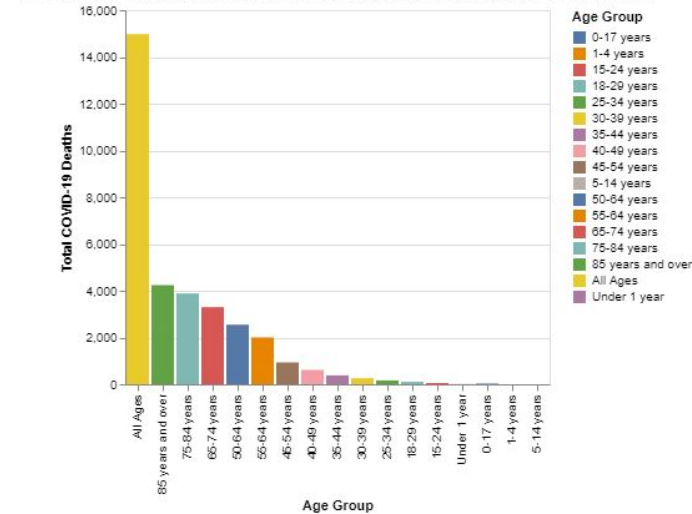
Universal Visualization:

Personal Visualization: