Our Agenda:

- Visualization process
- Tools of the trade
The Visualization Process
1. Start with your goal
2. Choose the right chart
3. Perfect your design
1. Start with your goal

What is the purpose of your viz?

What are you trying to tell people?

What questions are you trying to answer?
1. **Start with your goal**

What’s the point?

![Bar chart showing awards vs. school]

- UCSF: 10
- UCLA: 4
- UC Davis: 7
- UCSD: 2
- Berkeley: 8
1. Start with your goal

What’s the point?

UCSF has more awards than any other UC
1. **Start with your goal**

Who is your audience?

- Your colleagues?
- Other researchers?
- The public?
1. Start with your goal

How are people consuming your message?

Mediated?

Unmediated?
Activity: identify the message and audience

The Pumpkin Spice Latte debuted at Starbucks in 2003 using a recipe with no actual pumpkin in it. The popular PSL flavor is actually a blend of traditional fall spices, typically found in Pumpkin Pie recipes. Starbucks altered the recipe in 2015 to include pumpkin, but even then the change was imperceptible and served only to “appease those who wanted to see real pumpkin on the list of ingredients.” - Wikipedia, IF.org

**Activity:** identify the message and audience

**Fig 2.** Temporal distribution of insect biomass.
Activity: identify the message and audience

2. Choose the Right Chart

Let your task be your guide. Ex: Sales per unit

Huh?

Got it!
<table>
<thead>
<tr>
<th>Task</th>
<th>Chart</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compare data across categories</td>
<td>Bar chart</td>
</tr>
<tr>
<td>Visualize trends in data over time</td>
<td>Line chart</td>
</tr>
<tr>
<td>Show proportions (not too many!)</td>
<td>Pie chart</td>
</tr>
<tr>
<td>Show geocoded data</td>
<td>Map</td>
</tr>
<tr>
<td>See relationship between different variables</td>
<td>Scatterplot, Heatmap</td>
</tr>
<tr>
<td>Display a project schedule</td>
<td>Gantt chart</td>
</tr>
<tr>
<td>Understand distribution of data</td>
<td>Histogram, Box-and-whisker plot</td>
</tr>
<tr>
<td>Show hierarchical data as a proportion of a whole</td>
<td>Tree map</td>
</tr>
<tr>
<td>Show survey responses (Likert scale)</td>
<td>Stacked bar</td>
</tr>
<tr>
<td>Qualitative Data</td>
<td>Various</td>
</tr>
</tbody>
</table>
Compare data across categories - bar chart

Sales per Unit

<table>
<thead>
<tr>
<th>Unit</th>
<th>Sales</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1000</td>
</tr>
<tr>
<td>B</td>
<td>2000</td>
</tr>
<tr>
<td>c</td>
<td>3000</td>
</tr>
<tr>
<td>D</td>
<td>4000</td>
</tr>
<tr>
<td>E</td>
<td>5000</td>
</tr>
</tbody>
</table>
Track changes over time - line chart

Company Sales, 2000-2004

Sales

Year
Show proportion/percentages - pie chart

Sales

A
25.0%

B
75.0%
Show geocoded data - map
See relationships between two variables - scatterplot
See relationships between multiple variables - heatmap - numerical data

http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0075348
See relationships between multiple variables - heatmap - categorical data

<table>
<thead>
<tr>
<th>Affect of Service</th>
<th>Student</th>
<th>Staff</th>
<th>Faculty</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependability in handling users' service problems</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees who are consistently courteous</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees who deal with users in a caring fashion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees who have the knowledge to answer users questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees who instill confidence in users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employees who understand the needs of their users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Giving users individual attention</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Readiness to respond to users questions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Willingness to help users</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Information Control</th>
<th>Student</th>
<th>Staff</th>
<th>Faculty</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>A library website enabling me to locate information on my own</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Easy-to-use access tools that allow me to find things on my own</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making electronic resources accessible from my home or office</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making information easily accessible for independent use</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modern equipment that lets me easily access needed information</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print and/or electronic journal collections I require for my work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The electronic information resources I need</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The printed library materials I need for my work</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Library as Place</th>
<th>Student</th>
<th>Staff</th>
<th>Faculty</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>A comfortable and inviting location</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A gateway for study, learning, or research</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community space for group learning and group study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Library space that inspires study and learning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quiet space for individual activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Extra Questions</th>
<th>Student</th>
<th>Staff</th>
<th>Faculty</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>A secure and safe place</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ability to navigate library web pages easily</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making me aware of library resources and services</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Services that help me manage and share my research data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The library provides access to archival materials</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2015 LibQUAL Results at a Glance

For each question the LibQUAL survey asks respondents:
- The minimum level of service they would accept
- Their desired level of service
- Their perceived level of service at the library

How we rate
- Weak: Library performance below minimum
- Doing Well: Library performance between minimum and desired
- Very Strong: Library performance above desired

Library
- All

School
- All

Compiled by Ariel Deardorff, UCSF Assessment and Data Management Librarian
January 25, 2016

Understand distribution of data - histogram, box-and-whisker
Hierarchical as a proportion of a whole - treemap

Where did people go within the Library?

3rd Floor - Hearst Room
3rd Floor - Living Room
3rd Floor - General

2nd Floor - Classrooms
2nd Floor - Computing Labs
2nd Floor - Tech Commons

4th Floor
4th Floor - General

4th

Study Rooms

5th Floor - East Asian Room
1st Floor

5th Floor

Survey responses (likert scale) - stacked bar chart
Qualitative Data - some examples

- Highlight quotes
- Icons
- for
- themes
- Diagram

Word cloud
2. Choose the Right Chart

Resources:

Data Visualization Catalogue

Visualizing Health

Makeover Monday

Qualitative Data Resources:

Qualitative Chart Chooser

Displaying Qualitative Data
Activity: Choose the Right Chart

Prompt: You have a dataset about the rates of X disease in San Francisco and want to show a general audience that the disease is more prevalent in poorer neighborhoods. What is your message? What kind of chart would you use?

Take 5 minutes to work on this on your own. Use a paper and pens to sketch your ideas.

We will discuss as a group.
Activity: Choose the Right Chart

Potential Answers:

- **Scatter plot**: Disease rate versus neighborhood income level
- **Bar chart**: Disease rate in each neighborhood, sort and highlight the poorest ones
- **Heatmap (map)**: The darker the color of the neighborhood the higher the rate of disease (highlight/annotate poorer neighborhoods)
3. Perfect the Design

Use title to tell a story

Emphasize the most important data

Use color, size, position or annotations to highlight and encode meaning

Midwest Sales have grown steadily since 2000

- NW
- SW
- Mid

Year:
- 2000
- 2001
- 2002
- 2003
- 2004
3. Perfect the Design

Make color work for you:

- Is color being used for sequential, diverging, or categorical data?
- Colorblind safe?
- Black and white printer safe?

Resource: ColorBrewer
3. Perfect the Design

Reduce “chart junk”
Activity: Perfect the Design

Awards vs. School

<table>
<thead>
<tr>
<th>School</th>
<th>Awards</th>
</tr>
</thead>
<tbody>
<tr>
<td>UCSF</td>
<td>10</td>
</tr>
<tr>
<td>UCLA</td>
<td>4</td>
</tr>
<tr>
<td>UC Davis</td>
<td>7</td>
</tr>
<tr>
<td>UCSD</td>
<td>3</td>
</tr>
<tr>
<td>Berkeley</td>
<td>8</td>
</tr>
</tbody>
</table>
Spreadsheet Tools

Pros

● Familiar
● Easy to get started
● Large user community

Cons

● Time consuming
● Static

Examples:

● Excel
● Google Sheets

Resources

● Stephanie Evergreen's Blog - Evergreen Data
● Juice Analytics Chart Chooser
Focus on: Excel

Library vs. UCSF Campus

- Learn & grow
- Progress
- Best friend
- Employees committed to quality
- Mission/Purpose
- Opinions count
- Development
- Cares about me
- Recognition
- Opportunity to do best
- Have materials & equipment
- Know what’s expected
Programming Languages

Pros

● Very customizable
● Unlimited designs
● Reproducible

Cons

● Steep learning curve

Examples:

● R - ggplot2
● Python - matplotlib, seaborn

Resources

● R Graphics Cookbook (ebook)
● Data Visualization with Python and Javascript (ebook)
● Ggplot2 cheat sheet
library(datasets)
data("iris")

install.packages("ggplot2")
library(ggplot2)

ggplot(iris, aes(Sepal.Length, Sepal.Width, color=Species)) + geom_point()
Interactive

Pros

- Interactive
- Unlimited designs
- Large user community

Cons

- Steep Learning Curve
- Sometimes $$$

Examples:

- Tableau (Public, Desktop)
- Qlik (Sense, View)
- Google data studio
- D3 - javascript library
- Cytoscape

Resources

- Tableau at UCSF
- Tableau Training
- Communicating Data with Tableau (ebook)
Focus on: Tableau

Live Interactive Dashboard

Building a Tableau dashboard
Activity: What tools do you love?
Feedback: What tools would you like to learn?
Questions?
References


Which chart or graph is right for you? Tableau. https://www.tableau.com/learn/whitepapers/which-chart-or-graph-is-right-for-you?signin=03caf125ca386fbc650118b0fce4b73f