Data Visualization

Making Charts & Graphs and Other Cool Stuff

CFDR Fall 2022 Workshop Series

BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Today's Presentation

1. Why charts?
2. What are the different types of charts (with dos and don'ts)?
3. What are the basic principles of chart design?
4. What are some charts to be cautious of?
5. What are some programs (other than Excel) I can use to create data visualizations?
6. What are some outlets for data visualization publication?
7. Do you always need a chart?

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Why – charts?

“

Designing good charts, however, presents more challenges than tabular display as it draws on the talents of both the scientist and the artist. You have to know and understand your data, but you also need a good sense of how the reader will visualize the chart’s graphical elements.”

~ Gary Klass

BGSU | Center for Family and Demographic Research

BOWLING GREEN STATE UNIVERSITY

Picture Superiority Effect

Information is better remembered in tests of recall and item recognition when presented as pictures rather than words

Fruit <

What are –

the different types of charts?

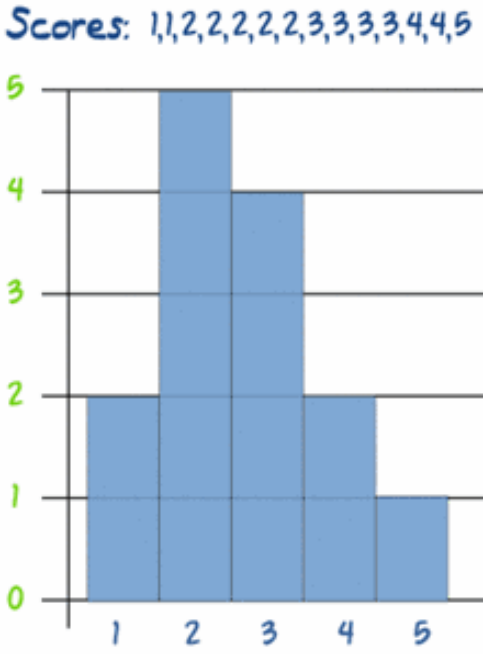
Histograms

A vertical bar chart that depicts the distribution of a set of data

Characteristics

- Bars represent the frequency of occurrence by classes of data.
- Enables you to see the shape of the data's distribution.
- Like a bar chart, but a histogram groups numbers into ranges.
- The horizontal axis is continuous like a number line (no gaps between columns).
- Great way to show results of continuous data:
  - Weight
  - Height
  - How much time
- A Frequency Histogram uses vertical columns to show how many times each score occurs.

Example

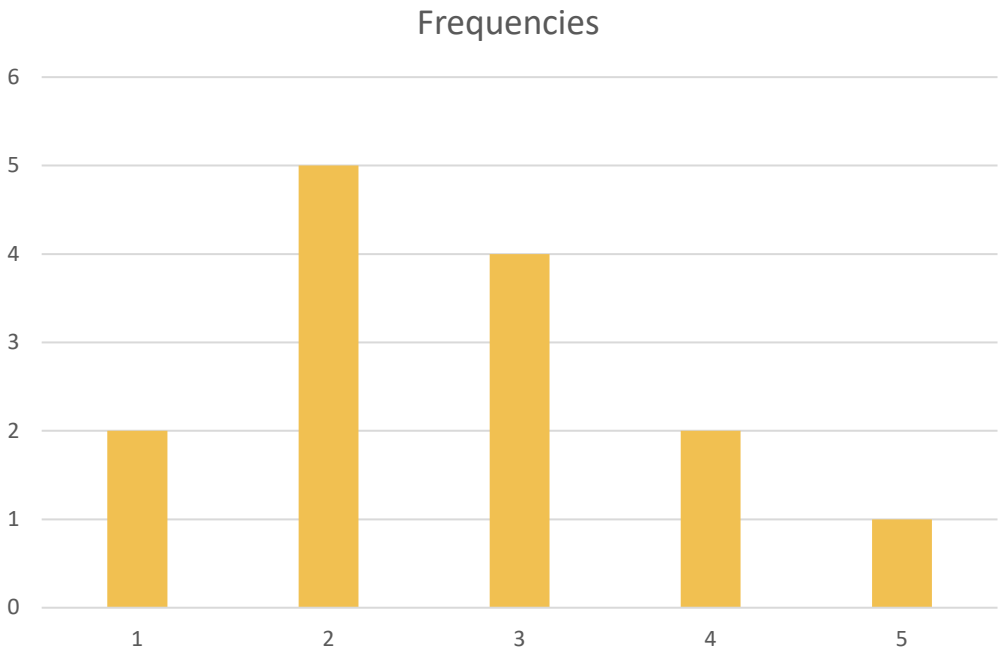


Source of chart: http://www.mathsisfun.com/data/histograms.html

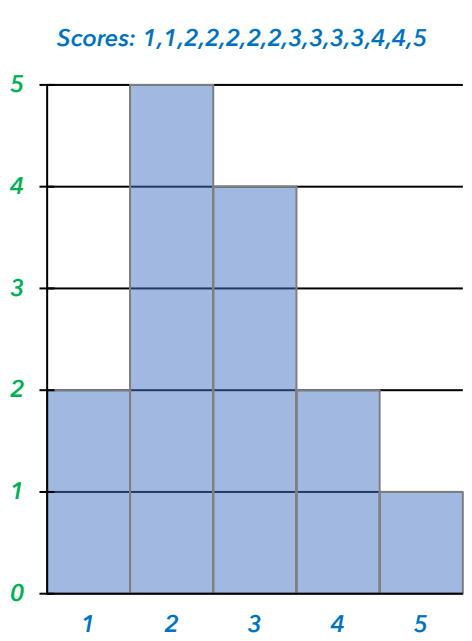
Histograms

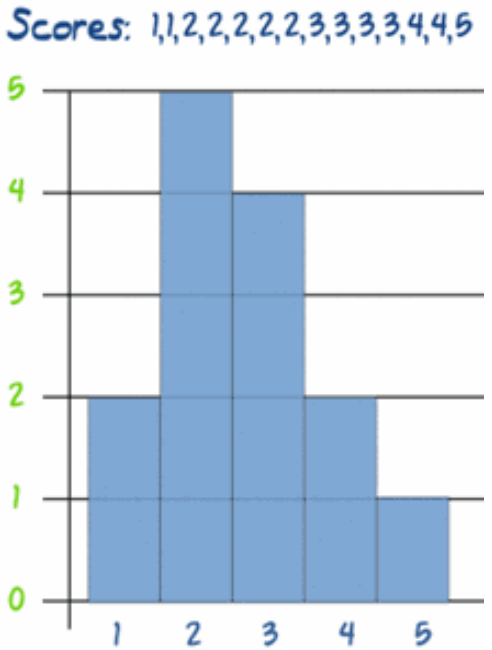
A vertical bar chart that depicts the distribution of a set of data

Clustered Column - Unformatted



Histogram - Formatted





Pie Charts

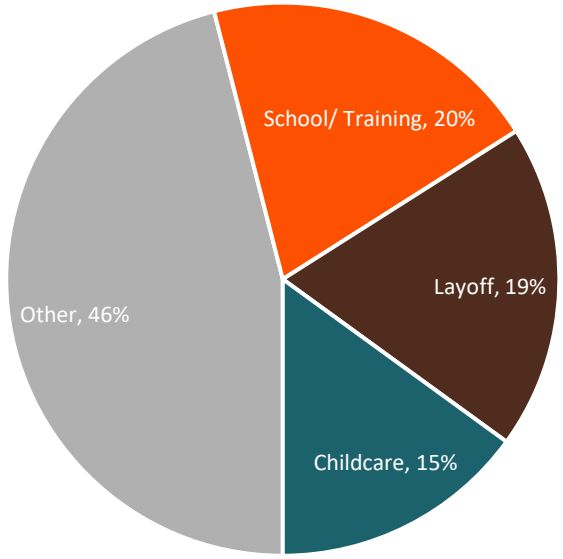
Generally used to show percentage or proportional data classified into nominal or ordinal categories

Characteristics

- Show the size of items in one data series, proportional to the sum of the items
- Useful for displaying data that are classified into **nominal** or **ordinal** categories
- Rules for pie charts:
  - **Avoid using pie charts**
  - Use only for data that add up to some meaningful total
  - Avoid comparisons across multiple pie charts
  - Five is the maximum number of slices, but two is better…

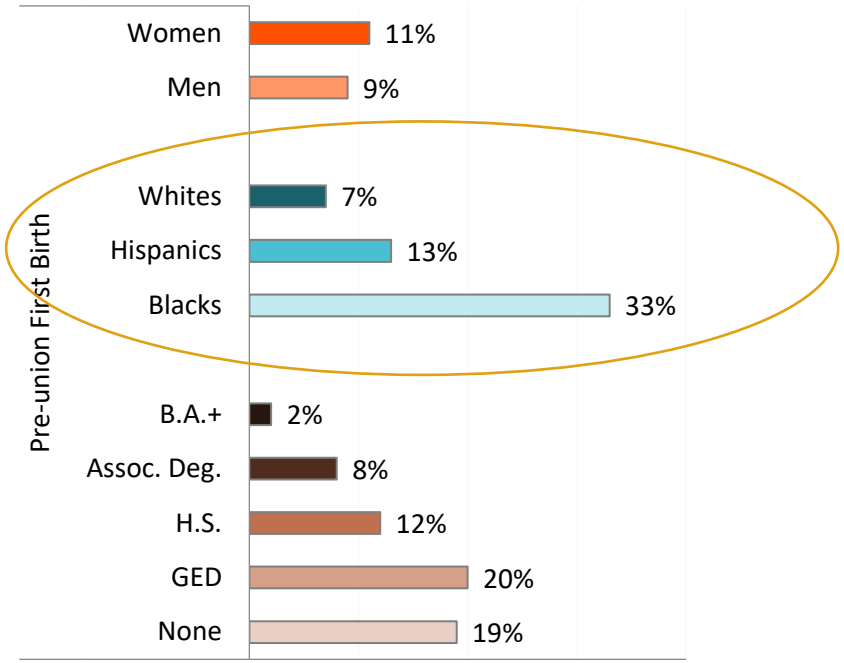
Example

Top Reasons for Fathers Leaving the Workforce in 2008



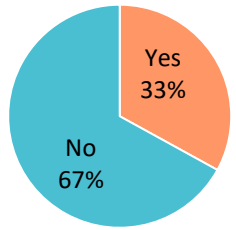
Bar Chart vs. Pie Charts

Prevalence of Pre-union First Birth across Demographic Characteristics

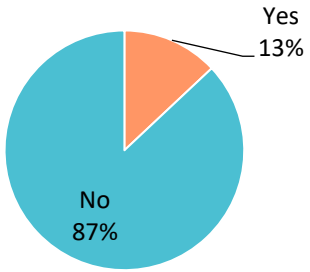


Prevalence of Pre-union First Birth by Race/Ethnicity:

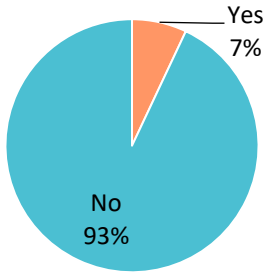
Blacks



Hispanics



Whites

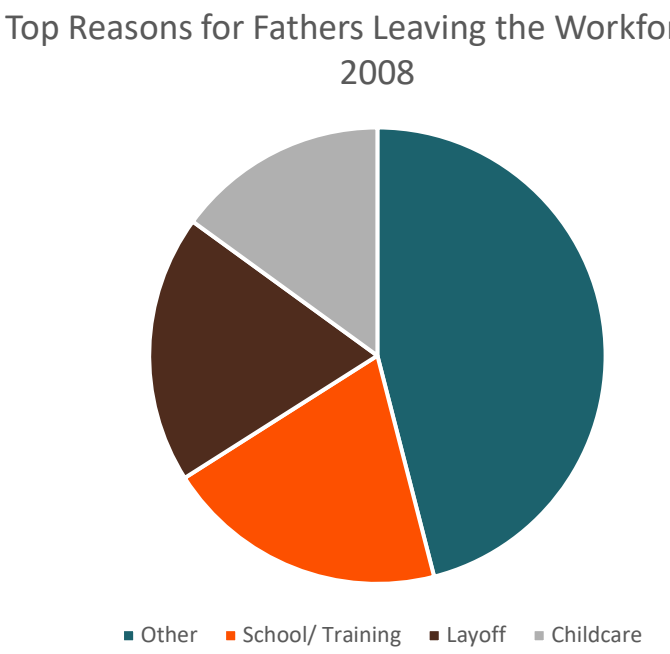


Source: National Longitudinal Survey of Youth 1997 (NLSY97), Rounds 1-13: 1997-2009 (weighted). U.S. Department of Labor, Bureau of Labor Statistics, NCFMR analyses of valid cases.

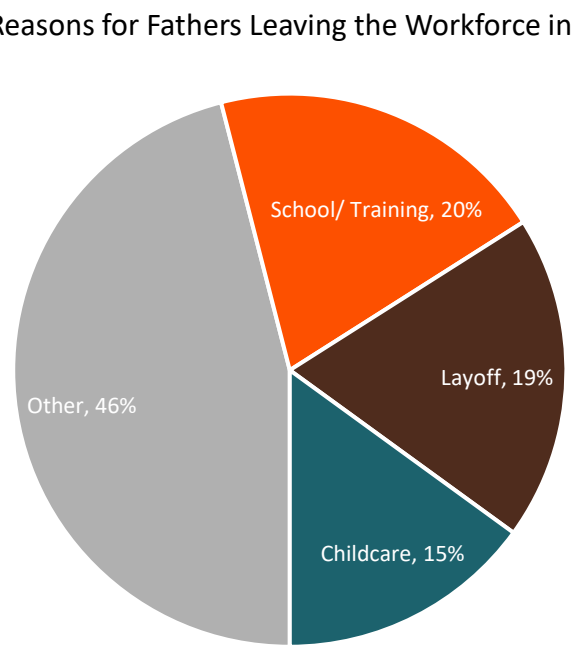
Pie Charts

Generally used to show percentage or proportional data classified into nominal or ordinal categories

Pie Chart - Unformatted



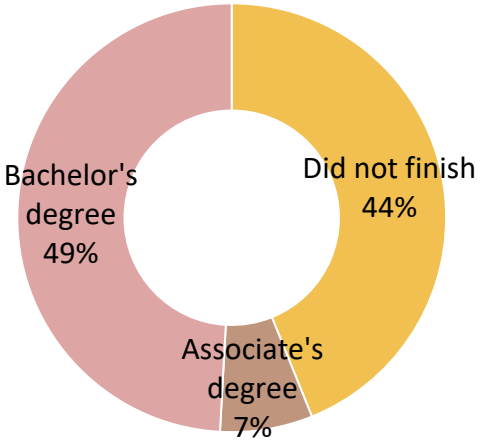
Pie Chart - Formatted



Other Types of Pie Charts

Doughnut

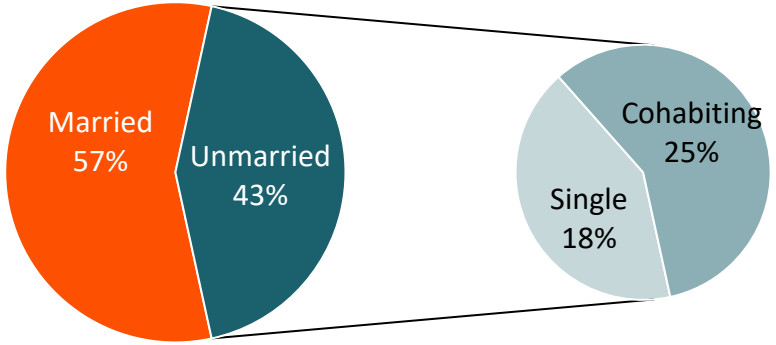
Percent of young adults who enroll in a 4-year program by degree earned by age 25



Source: National Longitudinal Survey of Youth 1997, Rounds 1-13: 1997-2009 weighted. U.S. Department of Labor, Bureau of Labor Statistics, NCFMR analyses of valid cases.

Pie-of-Pie

Percent of births by informal marital status of mother, 2005-2010

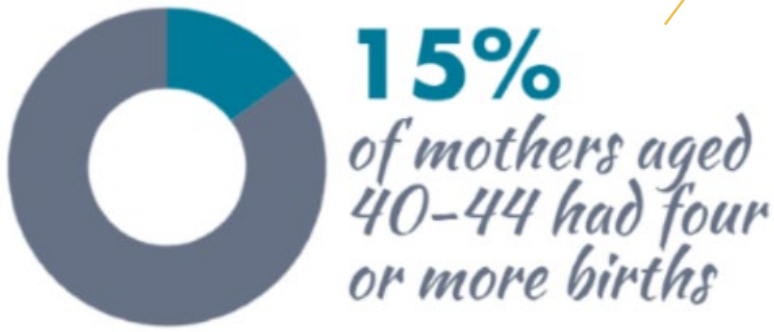


Source: NSFG 2006-2010

BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Another Way to Use Pie Charts

who had only one waited the longest, with a median age at first birth of 28. Conversely, mothers with four or more births began childbearing earlier, with a median age of 19.



- **Not** the **primary** form of distilling the information.
- Used as a **graphic element** to **re-emphasize** the focus of the Profile.

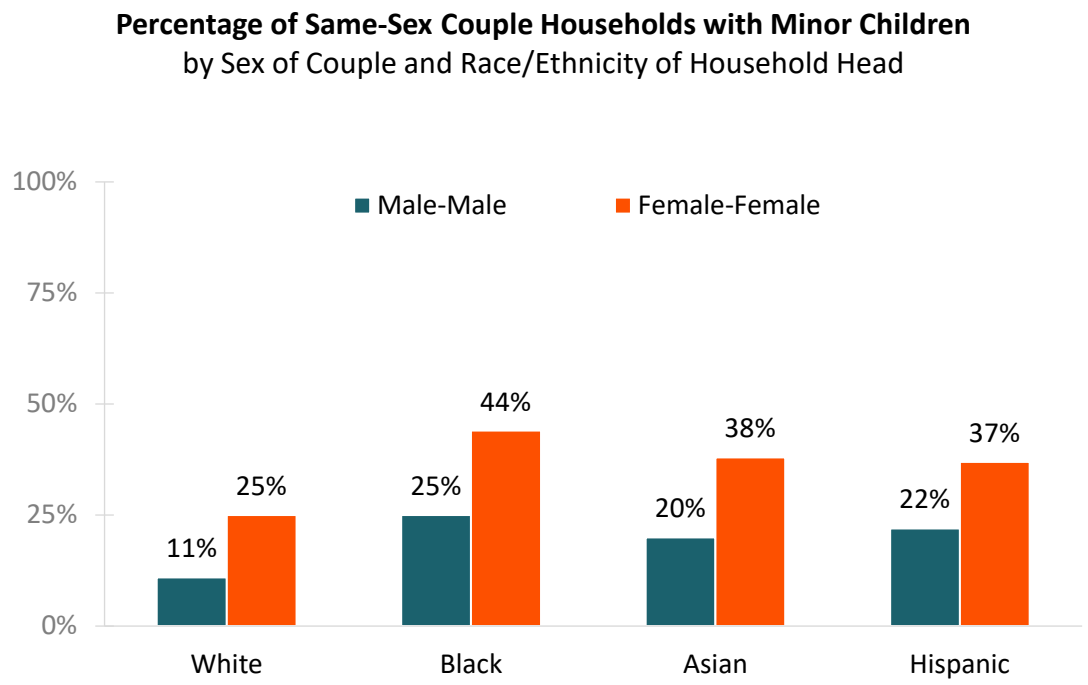
Column & Bar Charts

Useful for showing data changes over a period of time or for illustrating comparisons among items

Characteristics

- In column charts, categories are typically organized along the horizontal axis and values along the vertical axis.
- Allows user to compare values across categories. You can use a clustered column chart type when you have categories that represent:
  - Ranges of values (e.g., item counts)
  - Specific scale arrangements (e.g., a Likert scale)
  - Names that are not in any specific order (e.g., item names, geographic names, names of people)

Example – Side-by-Side Column Chart



Source: U.S. Census Bureau, American Community Survey, 1-Year Estimates, 2012

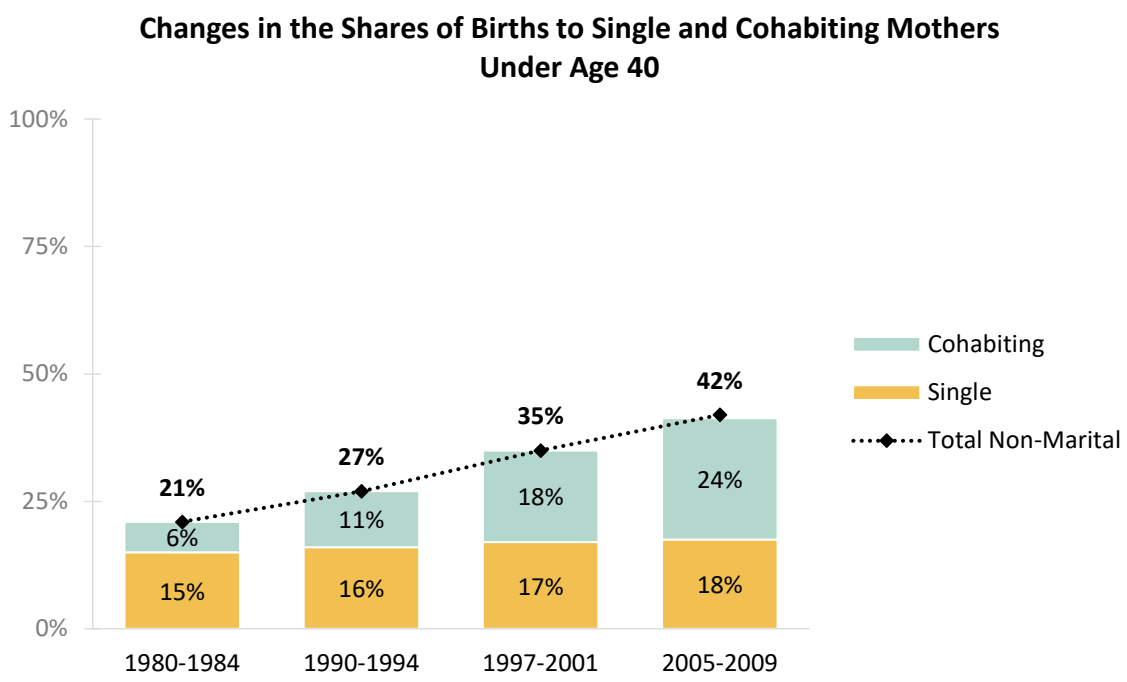
Column & Bar Charts

Useful for showing data changes over a period of time or for illustrating comparisons among items

Characteristics

- In column charts, categories are typically organized along the horizontal axis and values along the vertical axis.
- Allows user to compare values across categories. You can use a clustered column chart type when you have categories that represent:
  - Ranges of values (e.g., item counts)
  - Specific scale arrangements (e.g., a Likert scale)
  - Names that are not in any specific order (e.g., item names, geographic names, names of people)
- Stacked column: show the relationship of individual items to the whole

Example – Stacked Bars & Trend Lines



Sources: 1980-1984 data, Bumpass & Lu (2000) using NSFH, 1987/1988; 1990-1994 & 1997-2001 data, Kennedy & Bumpass (2008) using NSFG 1995 & NSFG 2002; 2005-2009, NCFMR analyses using NSFG 2006-2010.

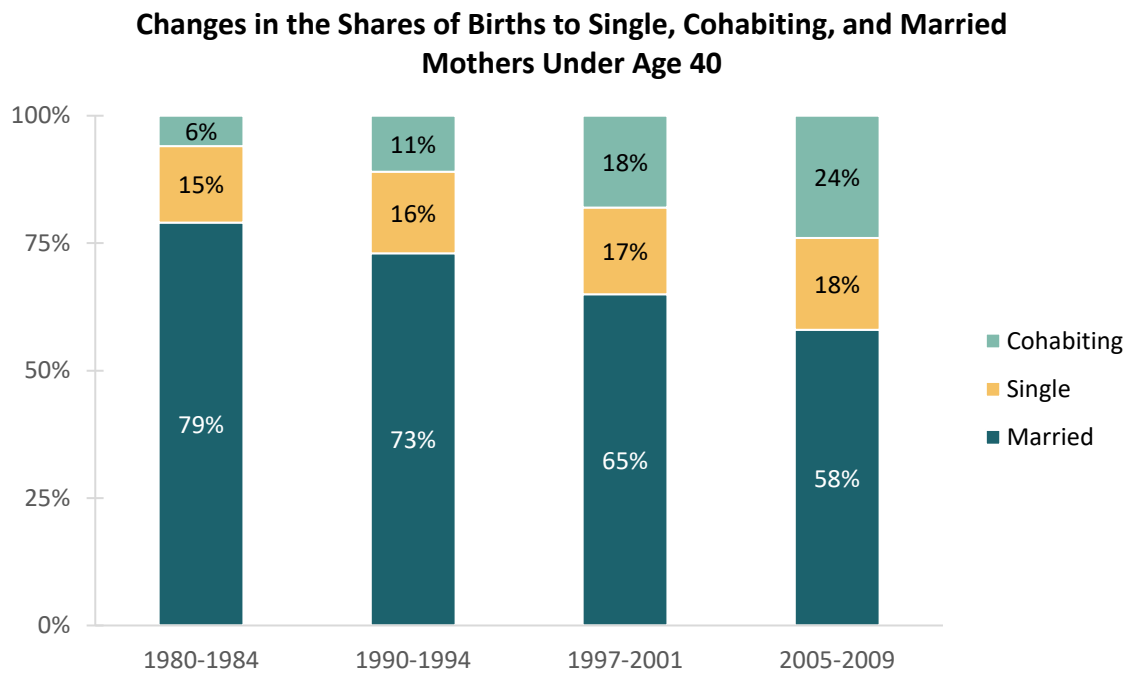
Column & Bar Charts

Useful for showing data changes over a period of time or for illustrating comparisons among items

Characteristics

- In column charts, categories are typically organized along the horizontal axis and values along the vertical axis.
- Allows user to compare values across categories. You can use a clustered column chart type when you have categories that represent:
  - Ranges of values (e.g., item counts)
  - Specific scale arrangements (e.g., a Likert scale)
  - Names that are not in any specific order (e.g., item names, geographic names, names of people)
- Stacked column: show the relationship of individual items to the whole
- 100% stacked column: compare the percentage that each value contributes to a total across categories.

Example – Stacked Bars & Trend Lines

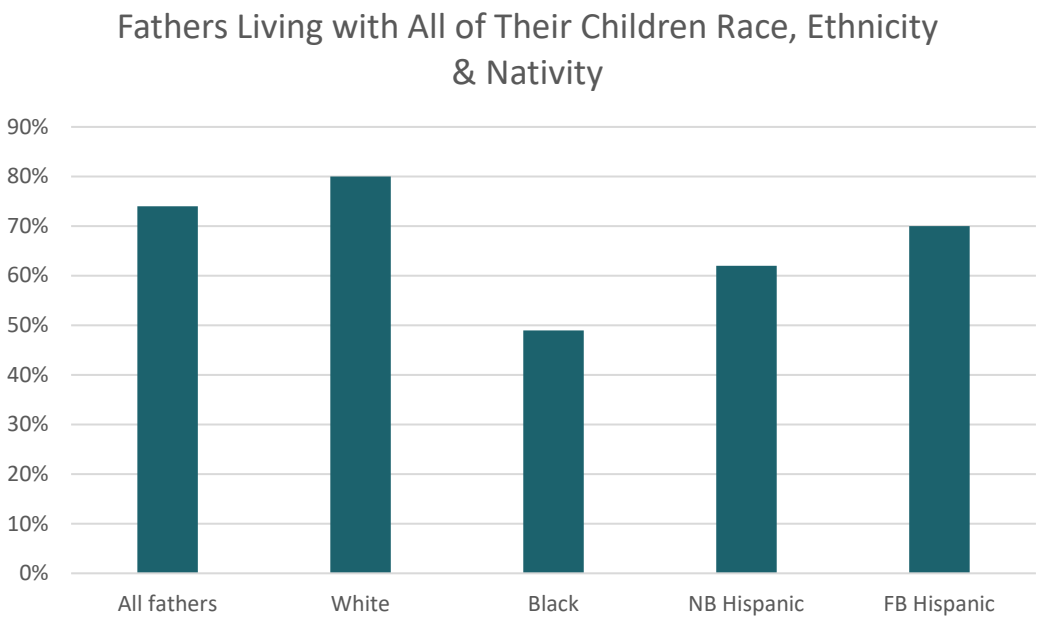


Sources: 1980-1984 data, Bumpass & Lu (2000) using NSFH, 1987/1988; 1990-1994 & 1997-2001 data, Kennedy & Bumpass (2008) using NSFG 1995 & NSFG 2002; 2005-2009, NCFMR analyses using NSFG 2006-2010.

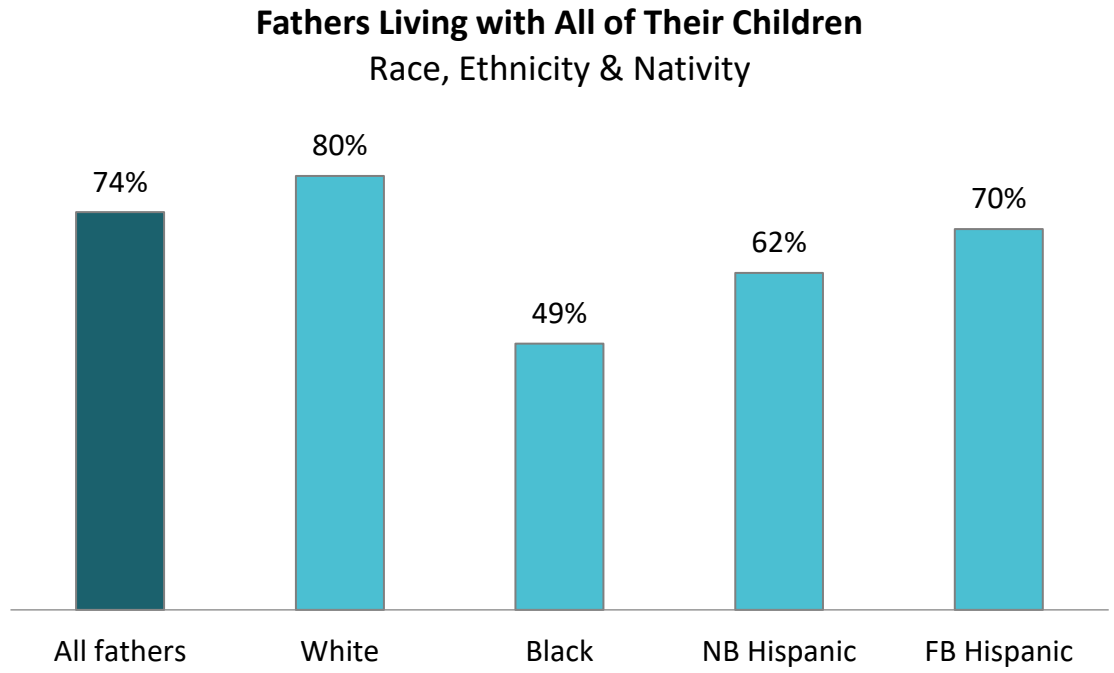
Column & Bar Charts

Useful for showing data changes over a period of time or for illustrating comparisons among items

Simple Column Chart - Unformatted



Example – Simple Column Chart



Source: NSFG 2006-2010

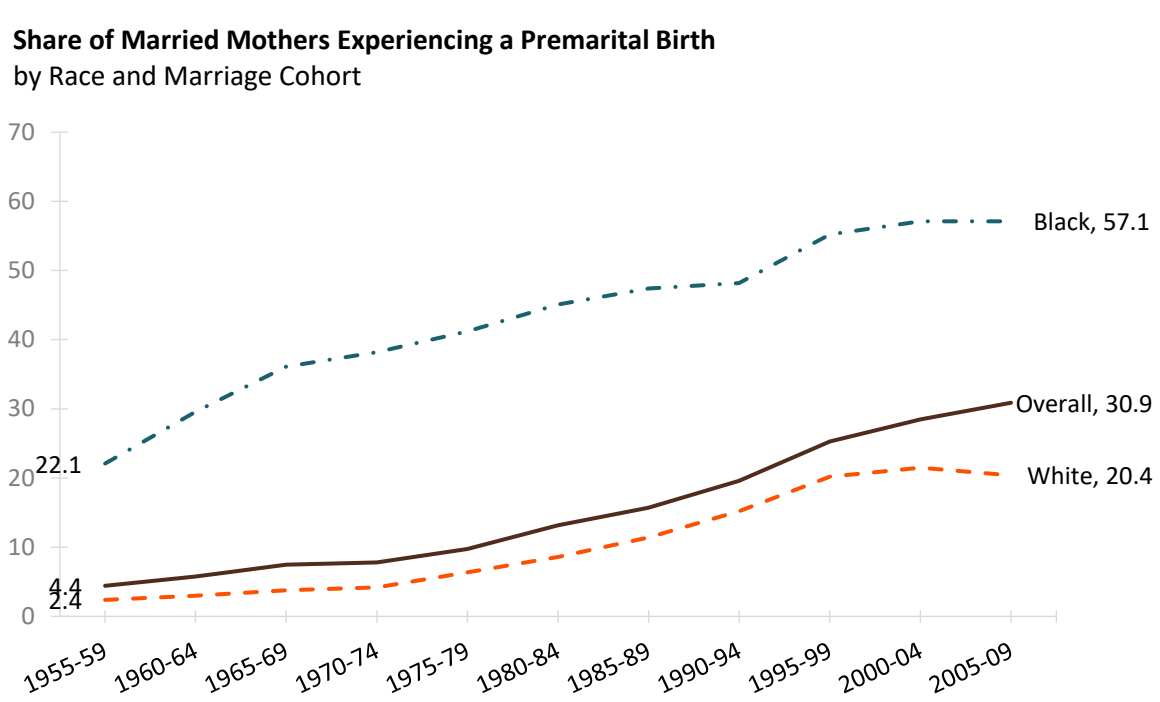
Line Charts

Ideal for showing trends over time at **equal intervals**

Characteristics

- Time is almost always displayed on the X-axis from left to right.
- Spacing between markers in the x-axis should be proportional
- If you can't easily see the pattern of each series, you may have too many
- Don't use a legend – directly label the series, instead

Example – Simple Line Chart



Source: The Integrated Fertility Survey Series (IFSS) is a project of the Population Studies Center and the Inter-university Consortium for Political and Social Research at the University of Michigan, with funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD, grant 5R01 HD053533; Pamela J. Smock, PI).

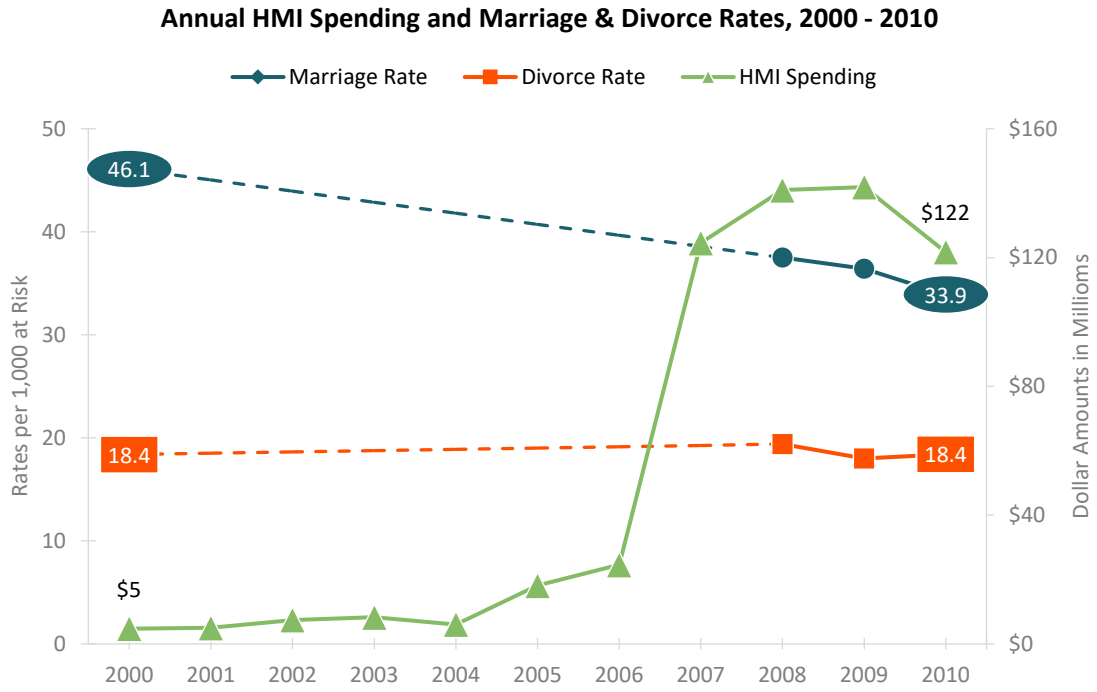
Line Charts

Ideal for showing trends over time at **equal intervals**

Characteristics

- Time is almost always displayed on the X-axis from left to right.
- Don't use a legend – directly label the series, instead
- If you can't easily see the pattern of each series, you may have too many
- Spacing between markers in the x-axis should be proportional
- Beware of scaling effects
- When displaying fiscal or monetary data over-time, it is often best to use deflated data (e.g., inflation-adjusted or % of GDP)

Example –Line Chart with a Double Axis



Sources: CDC/NCHS, National Vital Statistics System, 2000; Glass & Levchak, 2010, NCFMR County-Level Marriage & Divorce Data, 2000; U.S. Census Bureau, Decennial Census, 2000; U.S. Census Bureau, American Community Survey, 1-Year Estimates, 2008 – 2010; HMI Spending data – Hawkins et al., 2013.

Scatter Plots

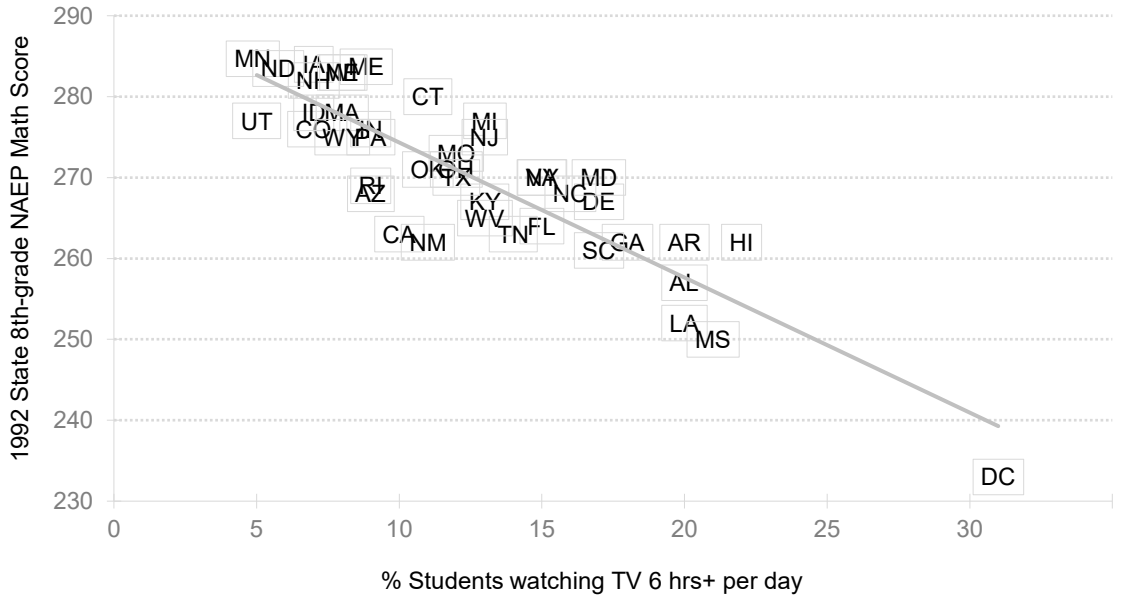
Commonly used to show the relationship between two variables e.g., correlation

Characteristics

- Use two interval-level variables
- Fully define the variables with the axis titles
- If there is an implied causal relationship between the variables, place the independent variable on the X-axis and the dependent variable on the Y-axis
- Scale the axes to maximize the use of the plot area for displaying the data points
- It's a good idea to add data labels to identify the cases
- In scatter plots, use empty circles as markers to let the reader see the overlapping points

Example – Scatter Plot

State Math Scores and Students' TV Viewing Habits



Source: National Center for Educational Statistics, 1994

Line Graph vs. Scatter Plots

Can be used for trend data at UNEQUAL intervals

Women's Marriage to Divorce Ratio, 1970-2015

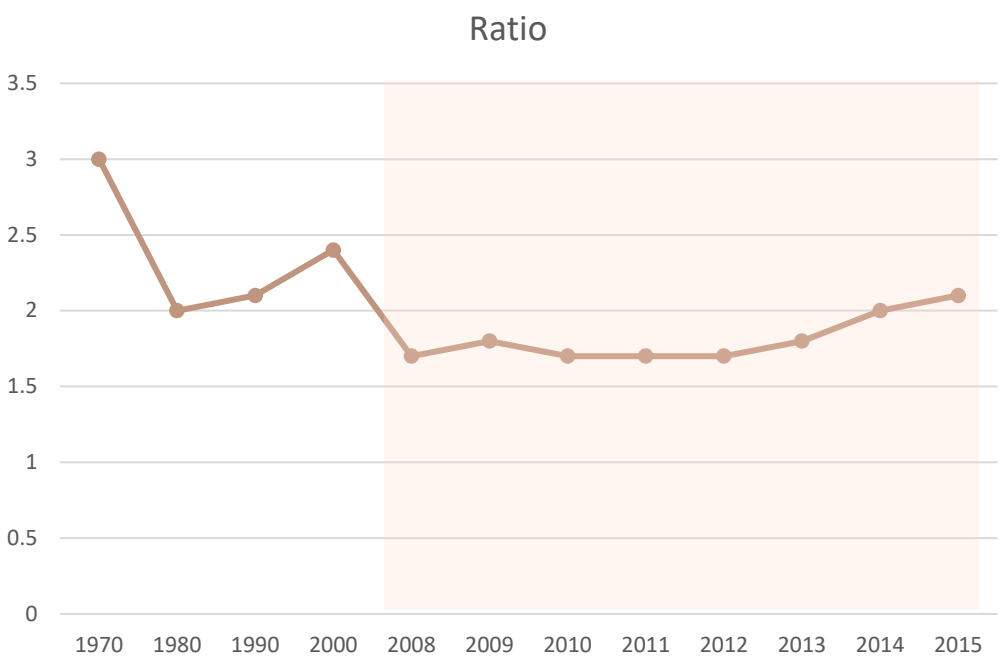
Year	Ratio
1970	3.0
1980	2.0
1990	2.1
2000	2.4
2008	1.7
2009	1.8
2010	1.7
2011	1.7
2012	1.7
2013	1.8
2014	2.0
2015	2.1

Sources: 1970-2000, National Center for Health Statistics; 2008-2015, U.S. Census Bureau, American Community Survey 1-yr est.

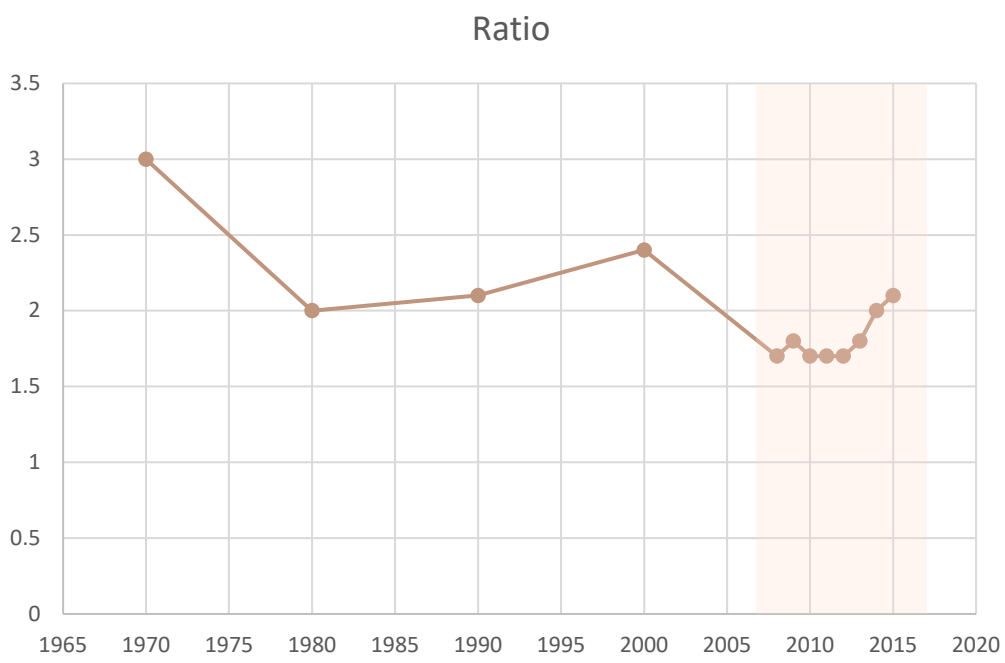
Line Graph vs. Scatter Plots

Can be used for trend data at UNEQUAL intervals

Line with Markers



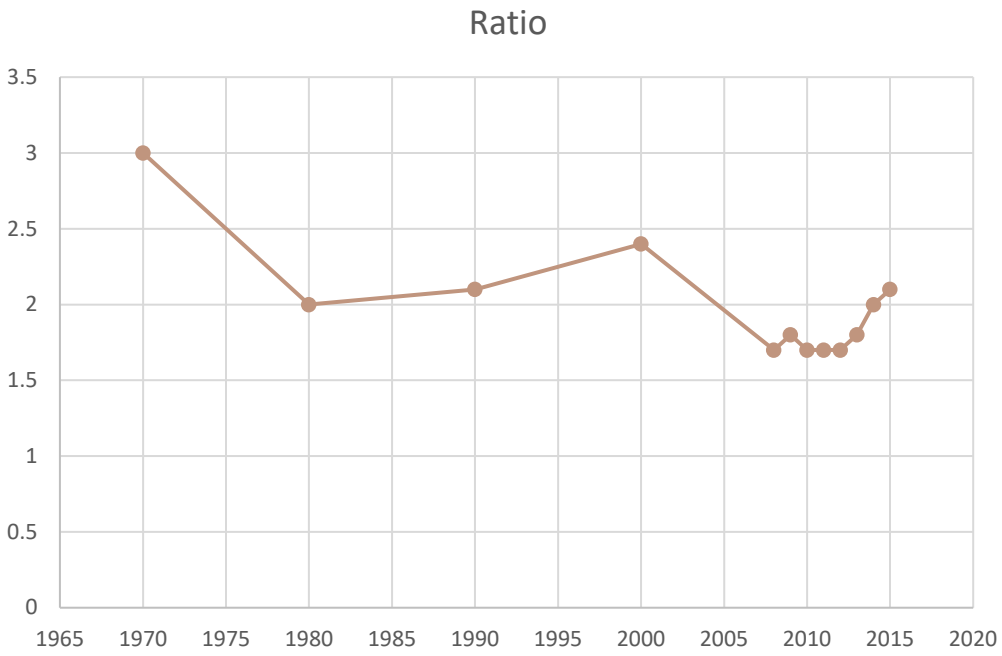
Scatter with Lines and Markers



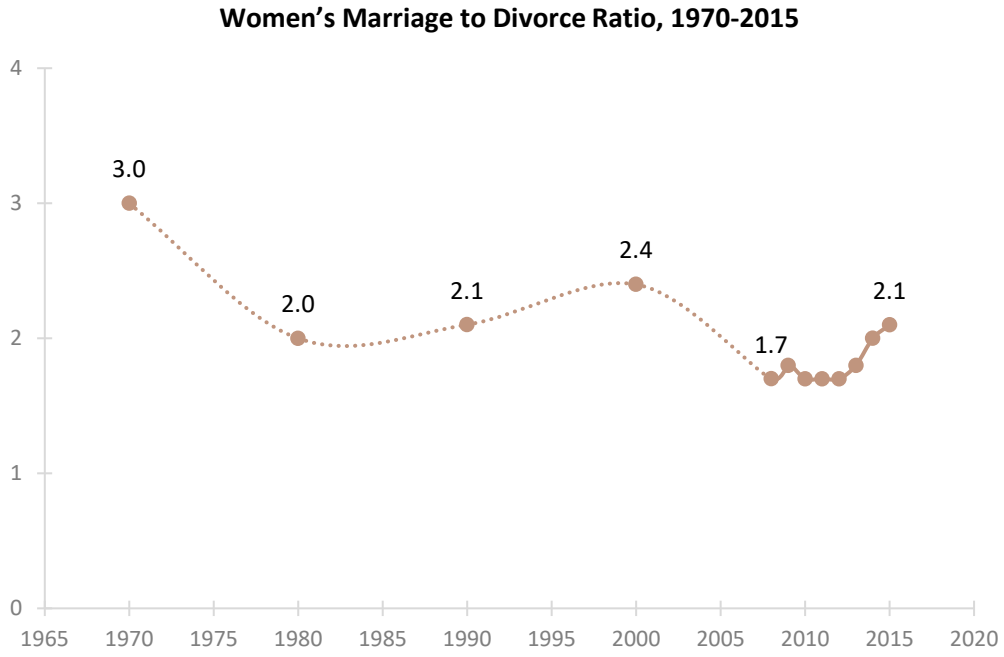
Line Graph vs. Scatter Plots

Can be used for trend data at UNEQUAL intervals

Scatter with Lines and Markers - Unformatted



Scatter with Lines and Markers - Formatted



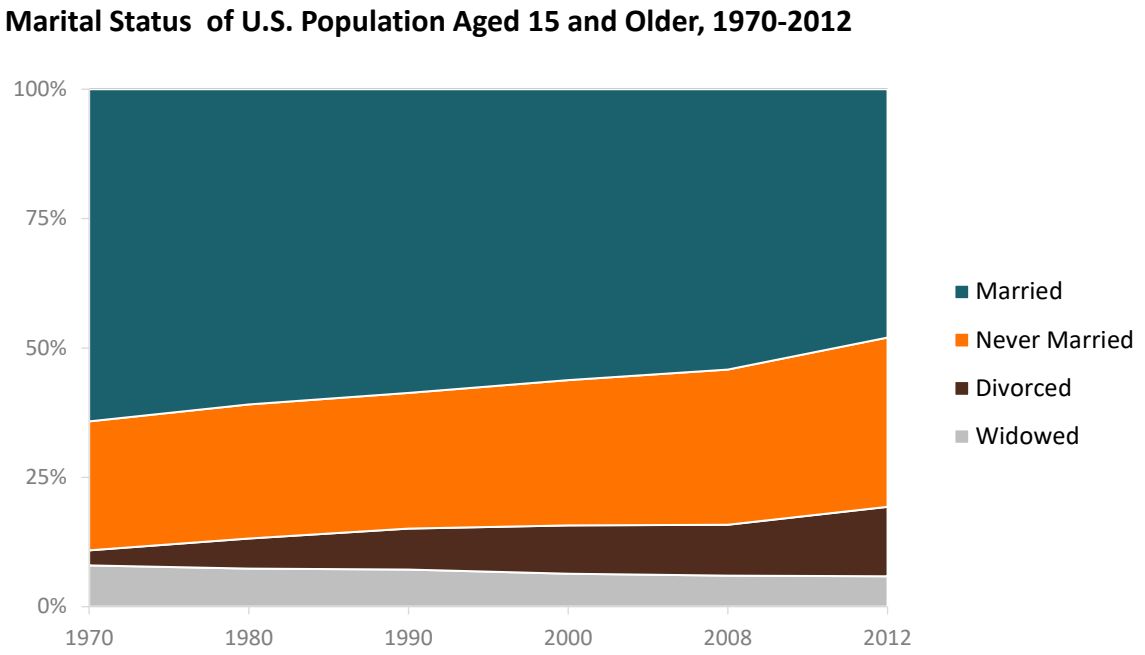
Area Charts

Show percentage or proportional data classified into nominal or ordinal categories over time

Characteristics

- Use these in moderation. Fall victim to the same visualization problems as pie- and doughnut-charts…pies and doughnuts are also area charts!
  - We avoid them "whenever possible because visual perception in humans can only compare areas as rough estimates" (Few, 2012; p275).
- The primary issue with area charts is our brains tend to read them as line charts—disregarding the "area" aspect of the chart.
  - The example here isn't as problematic as a regular ole area chart because it is a 100% stacked chart.

Example – Stacked Area Chart



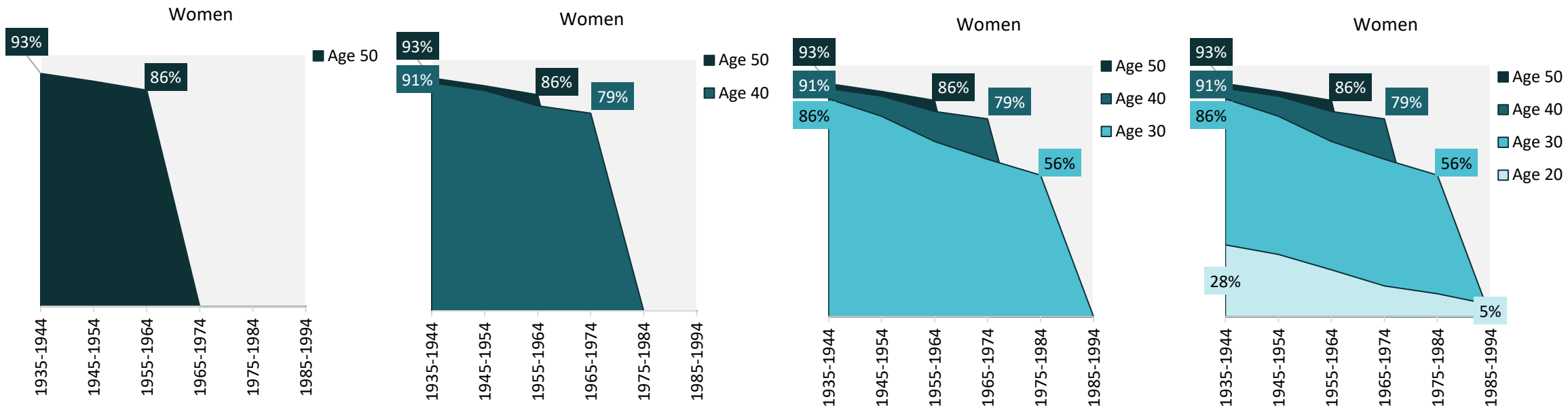
Source: 1970-2000 data, U.S. Census Bureau, Current Population Survey, March and Annual Social and Economic Supplements. 2008 and 2012 data, U.S. Census Bureau, American Community Survey, (IPUMS)

Area Charts

Show percentage or proportional data classified into nominal or ordinal categories over time

- **Not** stacked.
- Our brains will disregard the area underneath and behind.

Example – Area Chart



Source: NCFMR analysis of Survey of Income and Program Participation, 2021

What are the –

basic principles of chart design?

1. Simplify

- Sort data in a meaningful way
- Minimize ink-to-data ratio → remove unneeded chart elements
  - Gridlines
  - Chart borders
  - Axis titles
  - Legends
  - Markers and data labels
  - Decimal points (in axis and data labels)
  - Trend lines
  - NO 3D CHARTS

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

1. Simplify

NO 3D CHARTS!!!

BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

2. Color vs. Black & White

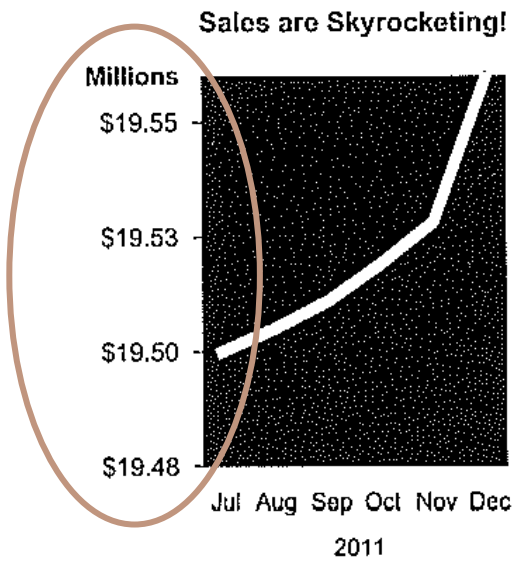
ACCESSIBILTY is your #1 Concern!

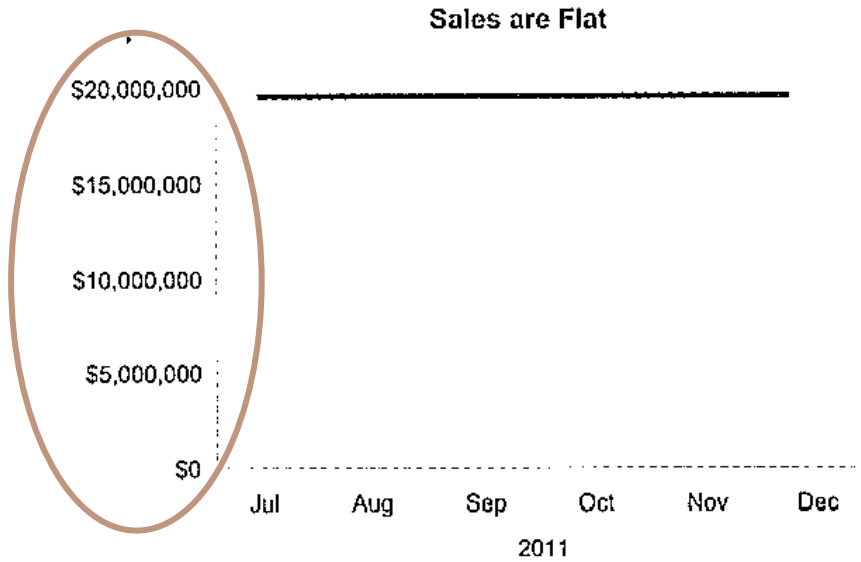
- When in doubt → Black & white
- BUT color can help tell a story
  - Color = branding
    - Use a cohesive and consistent color palette
    - Be mindful of how your audience will view your chart(s)
      - Excel vs. Word vs. PDF
      - Color vs. B&W print copy
      - Colorblind audience

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

3. Do NOT Use Distorted Charts

- Do NOT misrepresent your data!
  - Use appropriate and consistent axis and scales





BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

4. Present Related Charts Simultaneously

- One-after-another or side-by-side if possible
  - Use appropriate axis and scales

Figure 3. Remarriage Rates by Age Groups and Gender, 1990

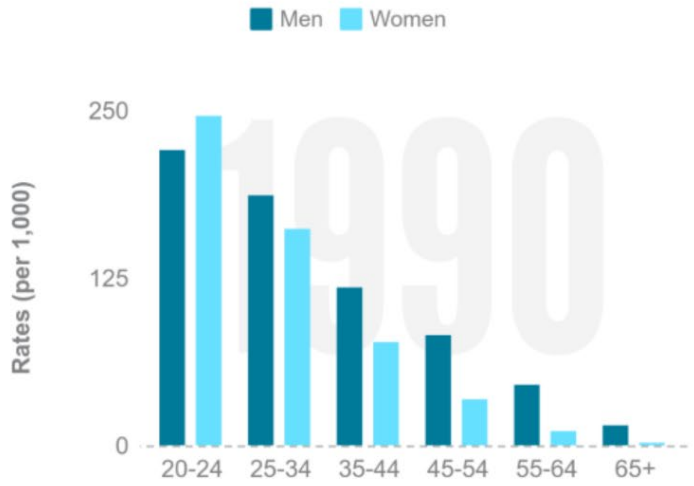
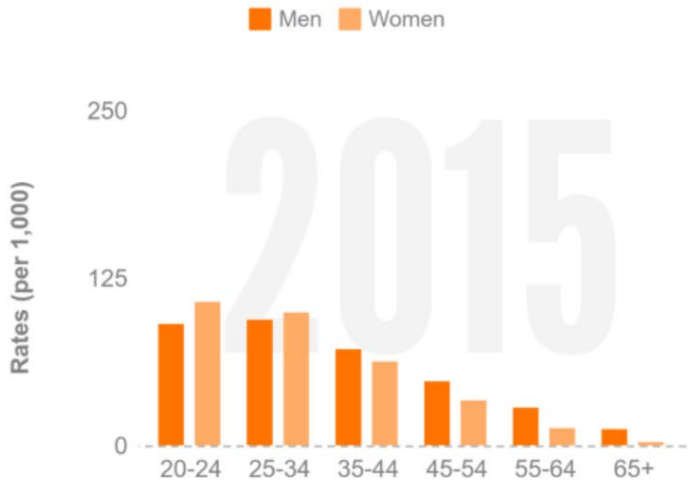


Figure 4. Remarriage Rates by Age Groups and Gender, 2015



Source: Centers for Disease Control and Prevention, National Center for Health Statistics, Vital Stats and U.S. Census Bureau, American Community Survey, 2015

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

5. Know Your Audience

- Academics vs. lay people
- Undergraduate students vs. graduate students
- Graduate students vs. professors
- PAA presentation vs. job talk

BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

6. TMC = TMI

- Too many charts (TMC) is as bad as too much information (TMI)…

→ Do **NOT** overload your audience!

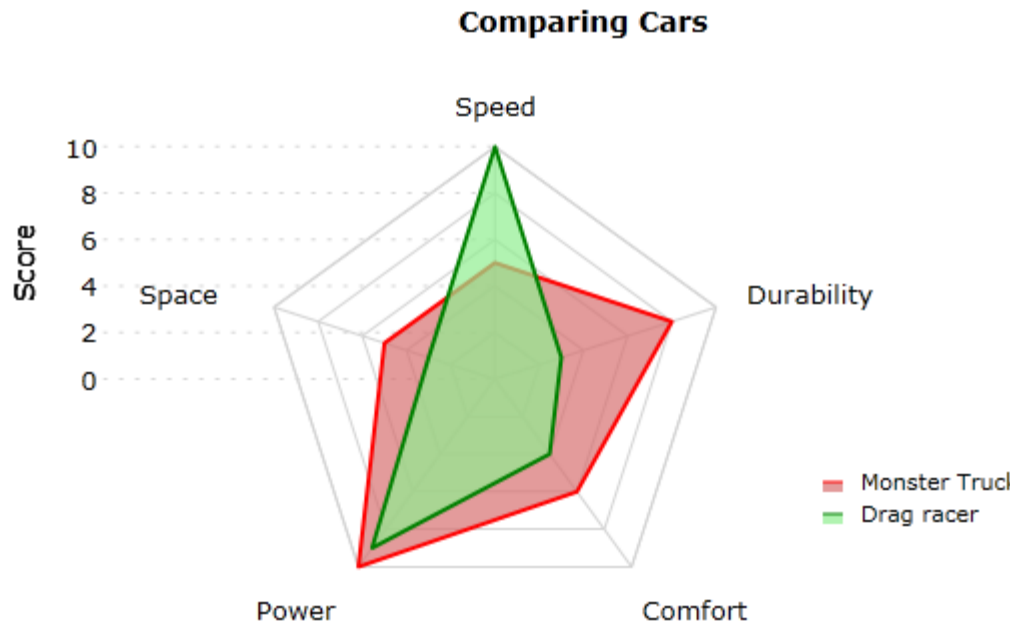
BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

What are some –

charts to be cautious of ?

Charts to be Cautious of…

- ✓ Pie/ Doughnut Charts
- ✓ Area Charts
- Radar Charts

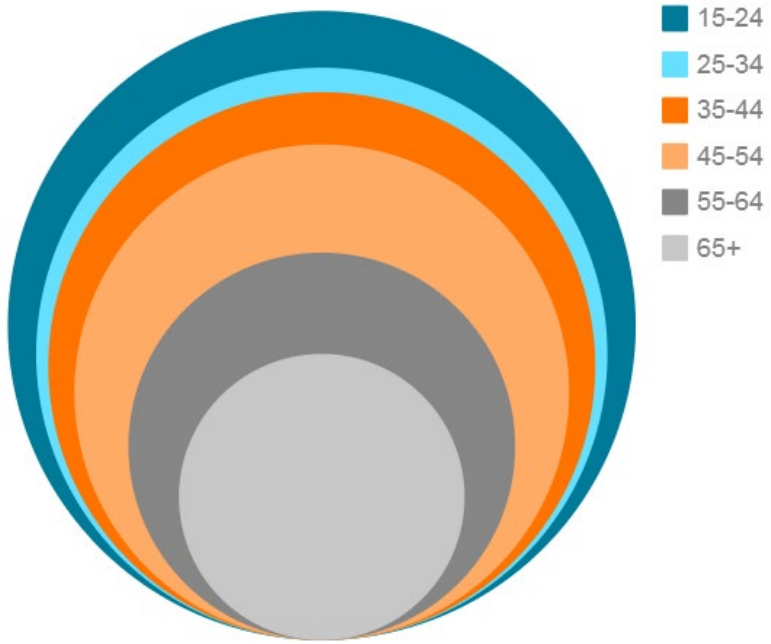


Example *Radar chart by* [Visiblox]

Charts to be Cautious of…

- ✓ Pie/ Doughnut Charts
- ✓ Area Charts
- ✓ Radar Charts
- Circle Charts

Divorce Rates per 1,000 by Age Groups, 2015

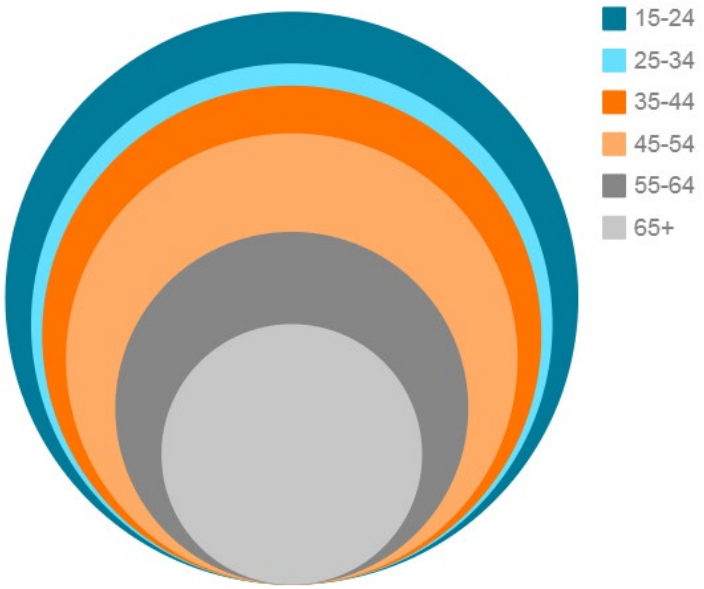


Source: U.S. Census Bureau, American Community Survey, 2015 1-year est.

Charts to be Cautious of…

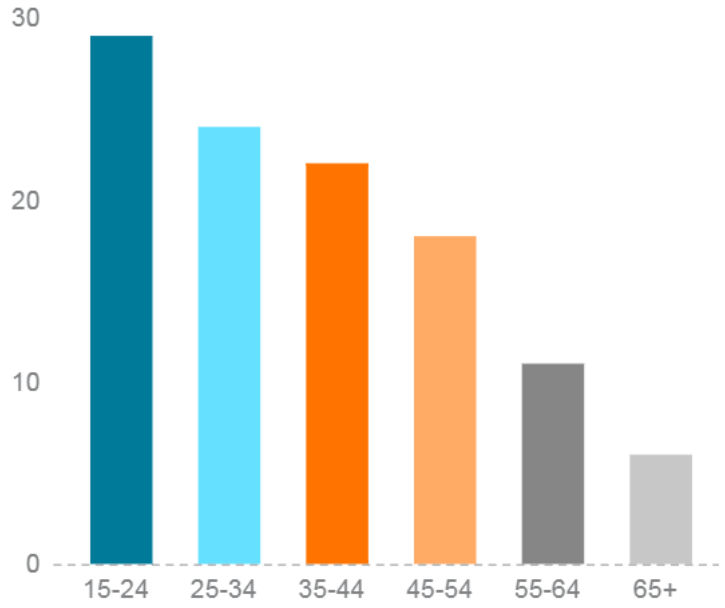
Divorce Rates per 1,000 by Age Groups, 2015

Circle Chart



Vs.

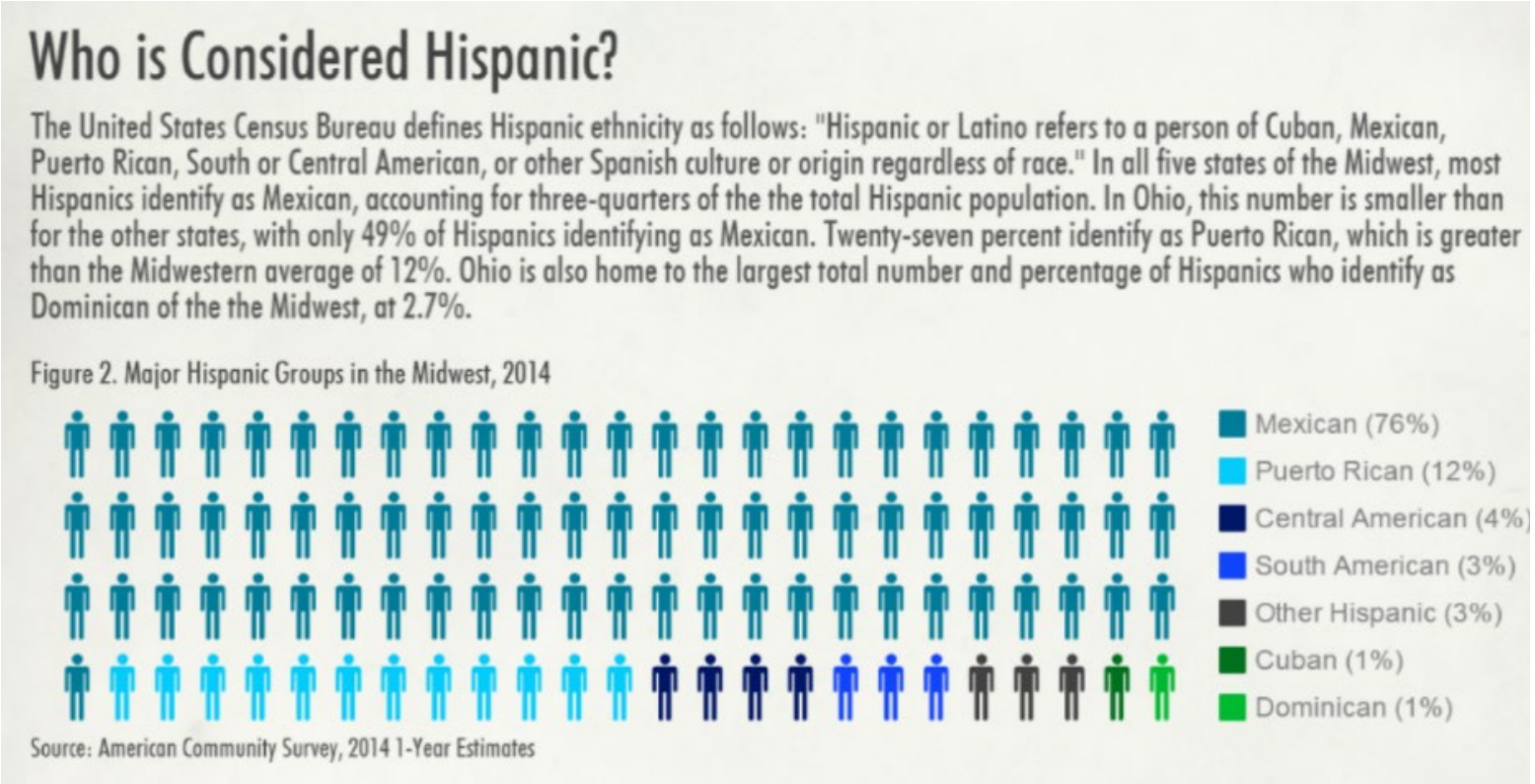
Column/ Bar Chart



Source: U.S. Census Bureau, American Community Survey, 2015 1-year est.

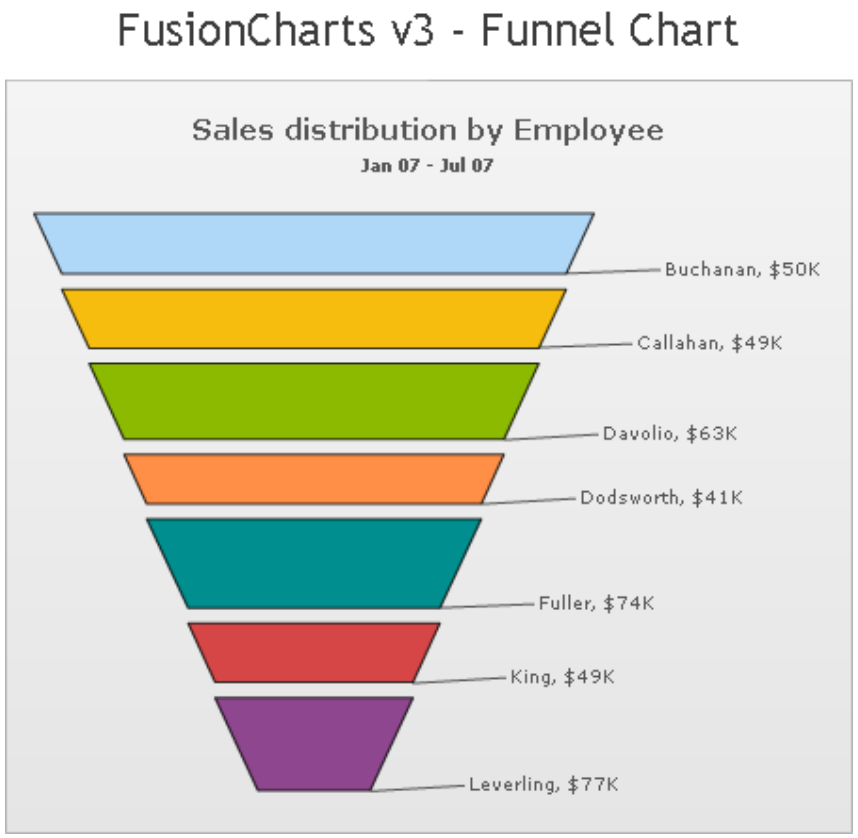
Charts to be Cautious of…

- ✓ Pie/ Doughnut Charts
- ✓ Area Charts
- ✓ Radar Charts
- ✓ Circle Charts
- Unit Charts



Charts to be Cautious of…

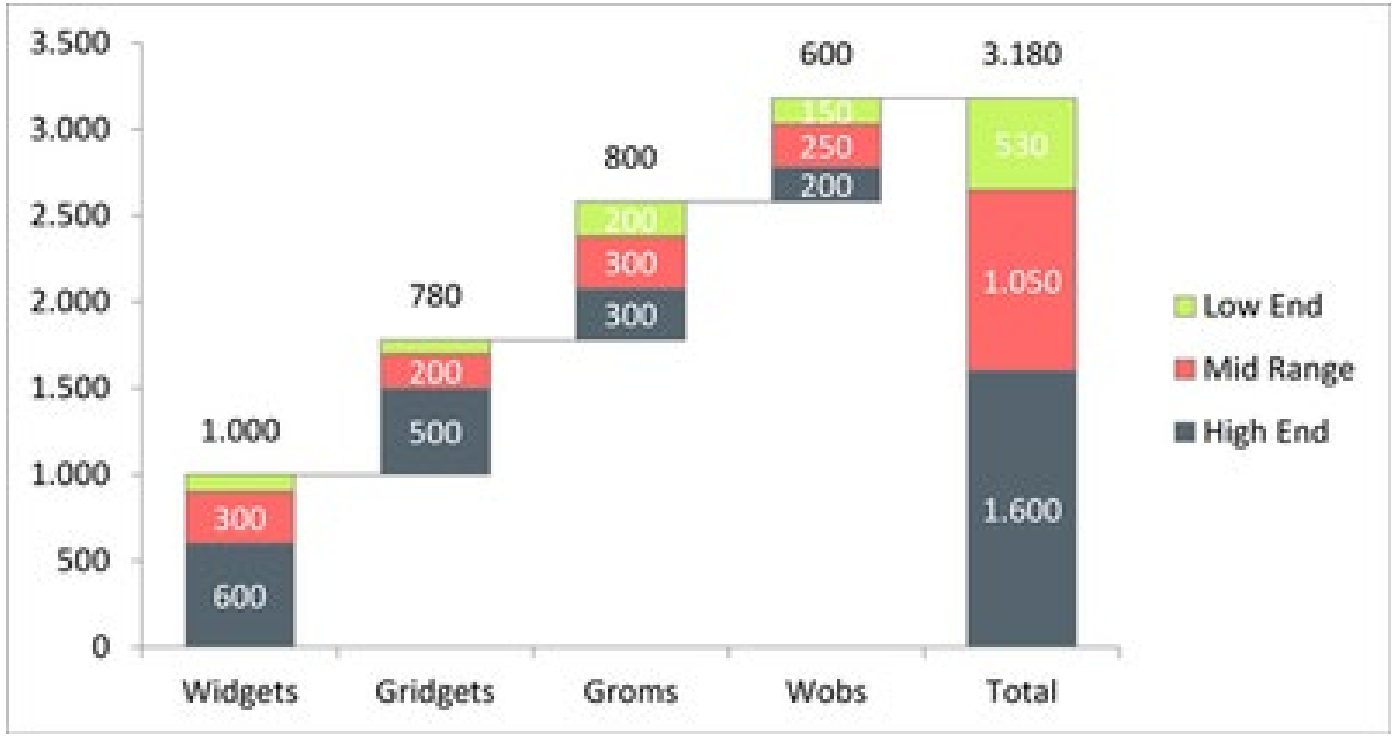
- ✓ Pie/ Doughnut Charts
- ✓ Area Charts
- ✓ Radar Charts
- ✓ Circle Charts
- ✓ Unit Charts
- Funnel Charts



Source: https://peltiertech.com/bad-graphics-funnel-chart/

Charts to be Cautious of…

- ✓ Pie/ Doughnut Charts
- ✓ Area Charts
- ✓ Radar Charts
- ✓ Circle Charts
- ✓ Unit Charts
- ✓ Funnel Charts
- Waterfall Charts



Source: http://community.powerbi.com/t5/Desktop/Stacked-waterfall-chart/td-p/188067

What are some programs –

(other than Excel)

I can use to create data visualizations?

Other Programs for Creating Data Visualizations

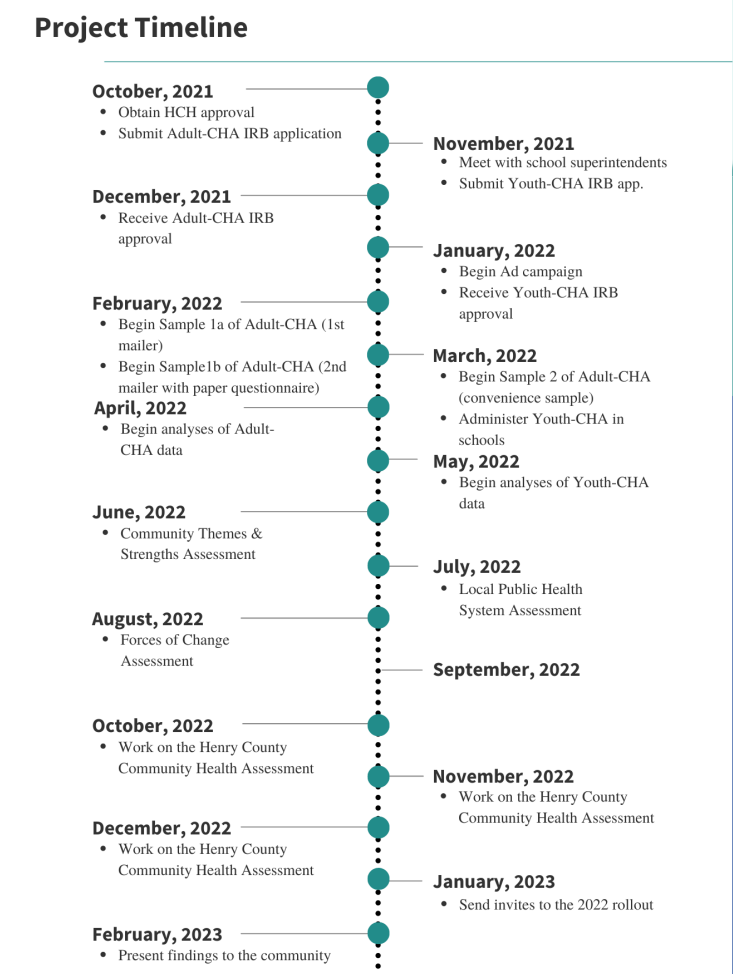
- Piktochart - https://piktochart.com/

https://create.piktochart.com/output/a9c32242cba8-falcon-flames-copy-2

Other Programs for Creating Data Visualizations

- Piktochart - https://piktochart.com/
- Canva - https://www.canva.com/

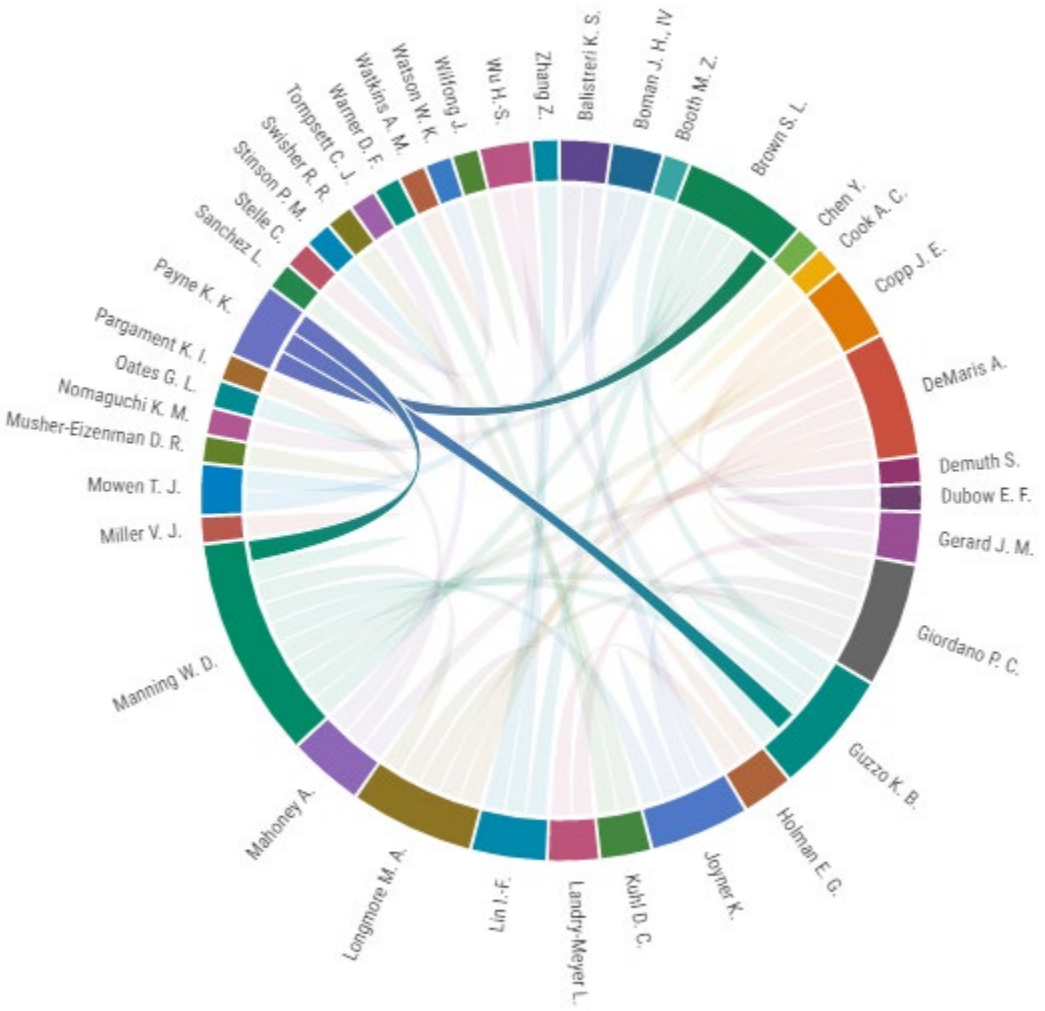
https://www.canva.com/design/DAEufkYzvg8/4T8JqpdoMkIRIbb5m2ETrA/edit?utm_content=DAEufkYzvg8&utm_campaign=designshare&utm_medium=link2&utm_source=sharebutton



Other Programs for Cr Visualizations

- Piktochart - https://piktochart.com/
- Canva - https://www.canva.com/
- Flourish - https://flourish.studio/

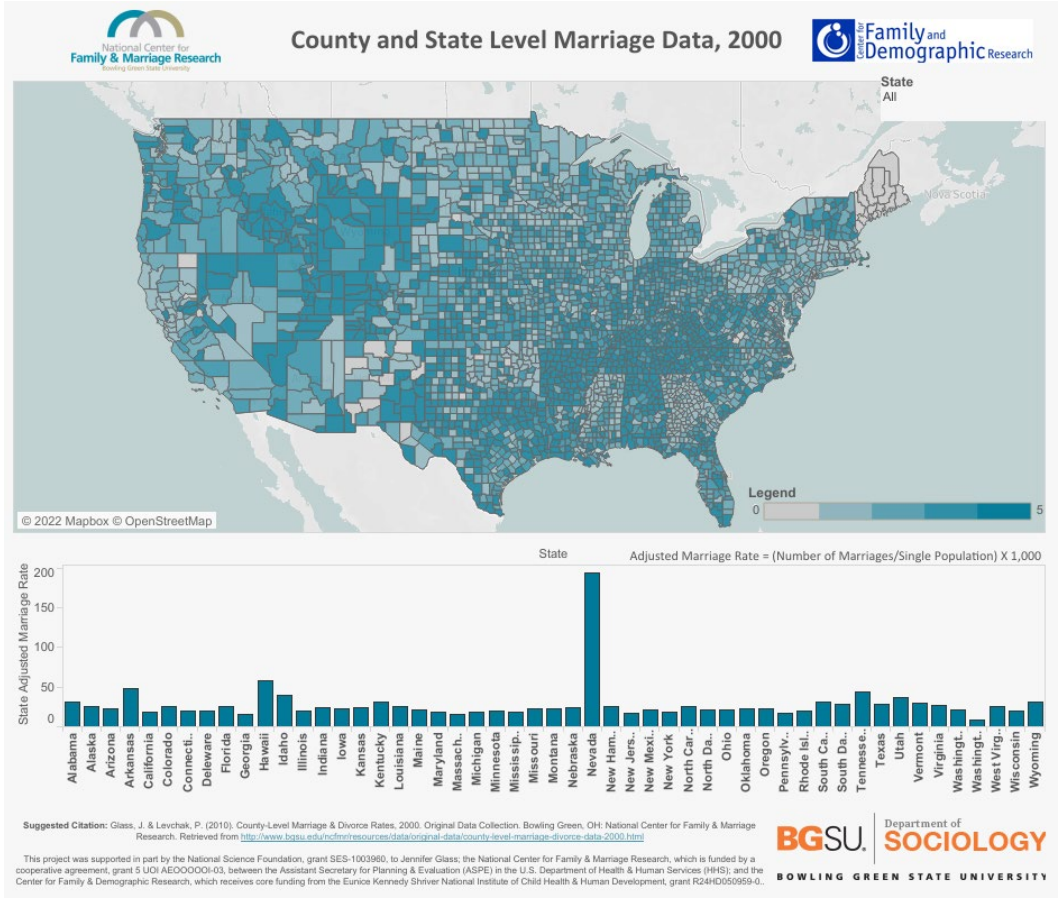
https://public.flourish.studio/visualisation/3166201/



Other Programs for Creating Data Visualizations

- Piktochart - https://piktochart.com/
- Canva - https://www.canva.com/
- Flourish - https://flourish.studio/
- Tableau - https://public.tableau.com/app/discover

https://public.tableau.com/views/CSLevelMarriageRates2000_1/MapBarChart?:language=en-US&:display_count=n&:origin=viz_share_link



Other Programs for Creating Data Visualizations

Piktochart - https://piktochart.com/

Canva - https://www.canva.com/

Flourish - https://flourish.studio/

Tableau - https://public.tableau.com/app/discover

ArcGIS

Stata

R

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

What are some –

outlets for data visualization publication?

Outlets for “Publication”

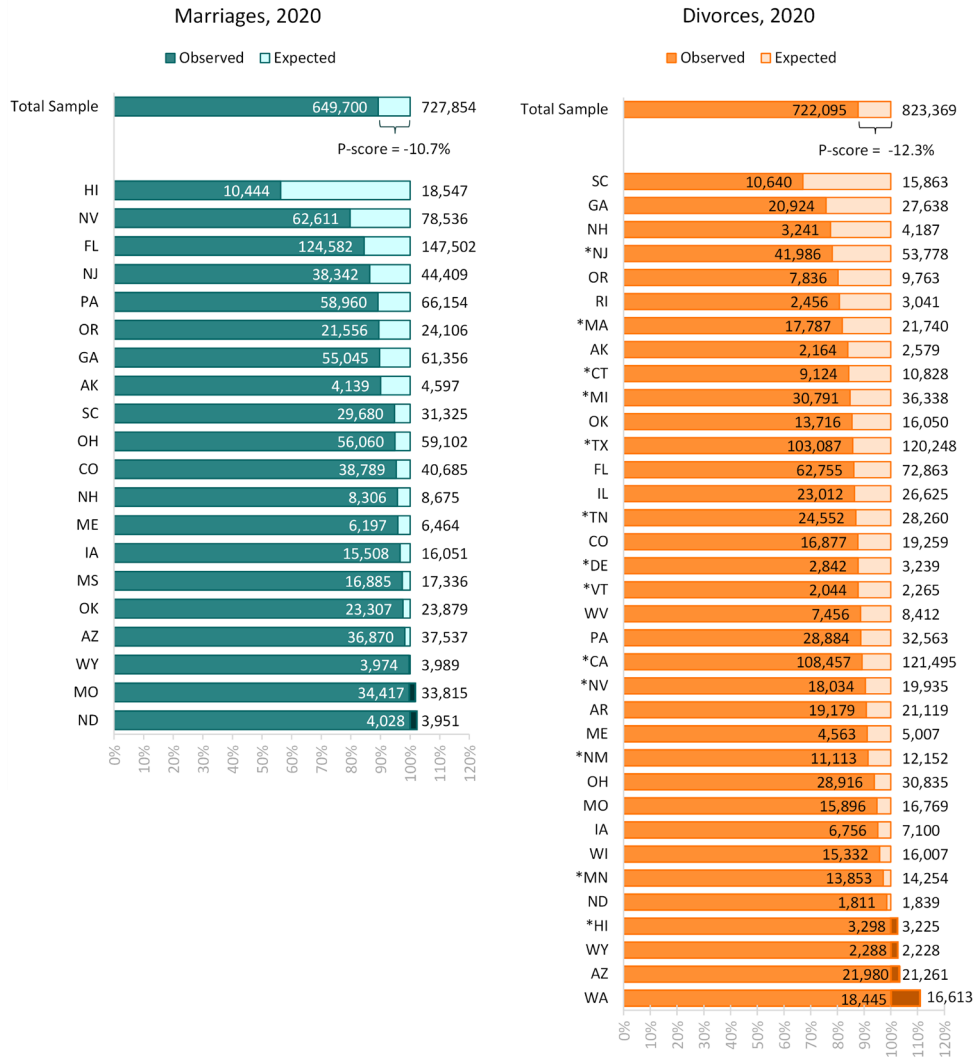
- Public sites in which others can view your work
  - Flourish - https://flourish.studio/
  - Tableau - https://public.tableau.com/app/discover

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Outlets for “Publication”

- Public sites in which others can view your work
  - Flourish - https://flourish.studio/
  - Tableau - https://public.tableau.com/app/discover
- Socius – Peer reviewed

https://journals.sagepub.com/doi/full/10.1177/23780231221090192



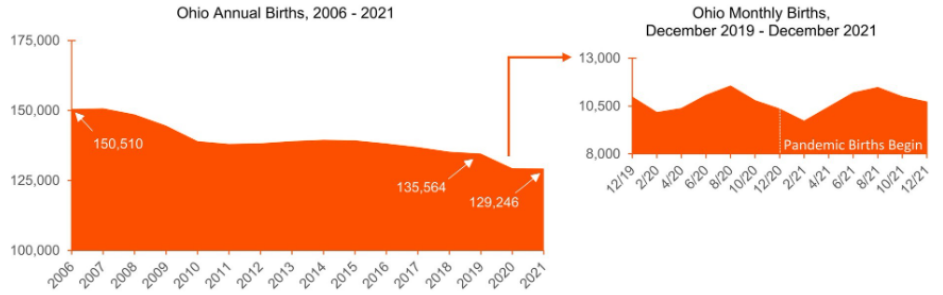
Outlets for "Publication"

- Public sites in which others can view your work
  - Flourish - https://flourish.studio/
  - Tableau - https://public.tableau.com/app/discover
- Socius – Peer reviewed
- NCFMR

https://www.bgsu.edu/ncfmr/resources/data/family-profiles/julian-manning-ohio-florida-births-prior-during-covid-pandemic-fp-22-24.html

Ohio

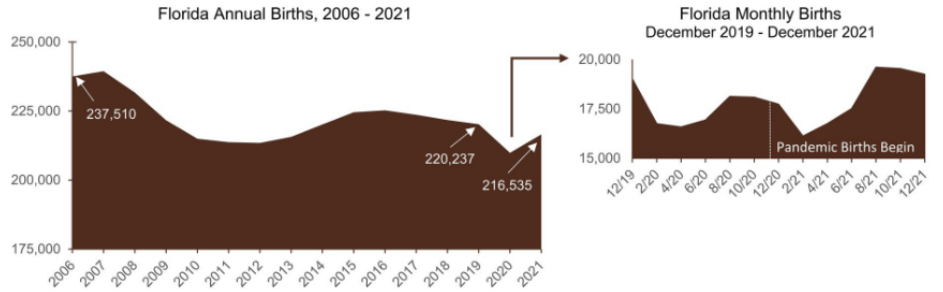
Figure 1. Births in Ohio Prior to and During the COVID-19 Pandemic



Source: NCFMR analyses of Ohio Department of Health (https://odh.ohio.gov/). These data were provided by the Ohio Department of Health. The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions.

Florida

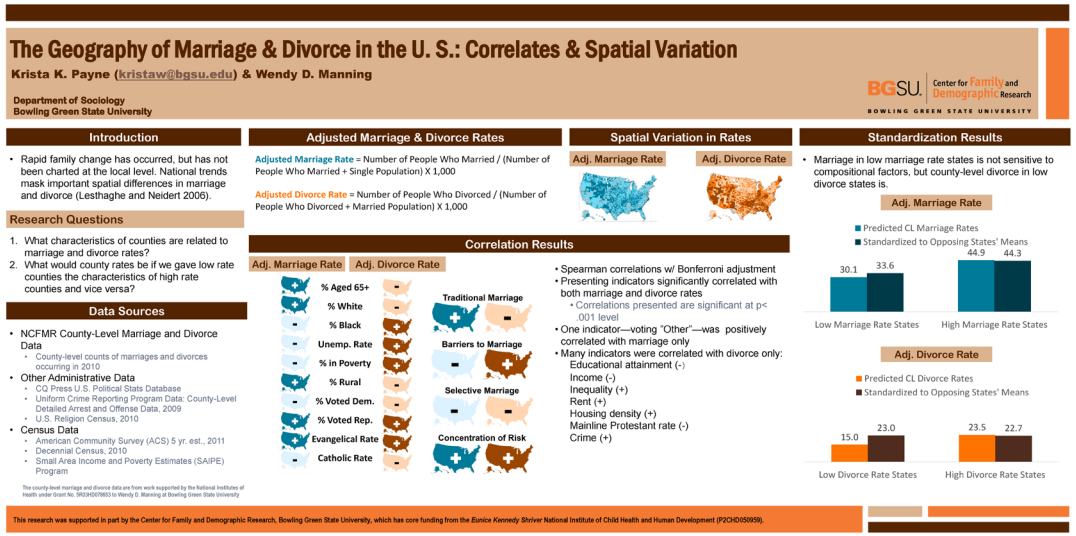
Figure 2. Births in Florida Prior to and During the COVID-19 Pandemic



Source: NCFMR analyses of Florida Department of Health (https://www.floridahealth.gov/)

Outlets for “Publication”

- Public sites in which others can view your work
  - Flourish - https://flourish.studio/
  - Tableau - https://public.tableau.com/app/disco
- Socius – Peer reviewed
- NCFMR
- Conference poster presentations



Do you –

always need a chart?

7. Do you need a chart?

$117 million

Increase in annual HMI spending in the U.S. from 2000 – 2010

Sources: U.S. Census Bureau, American Community Survey, 2008-2011; HMI spending data– Hawkins et al., 2013.

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

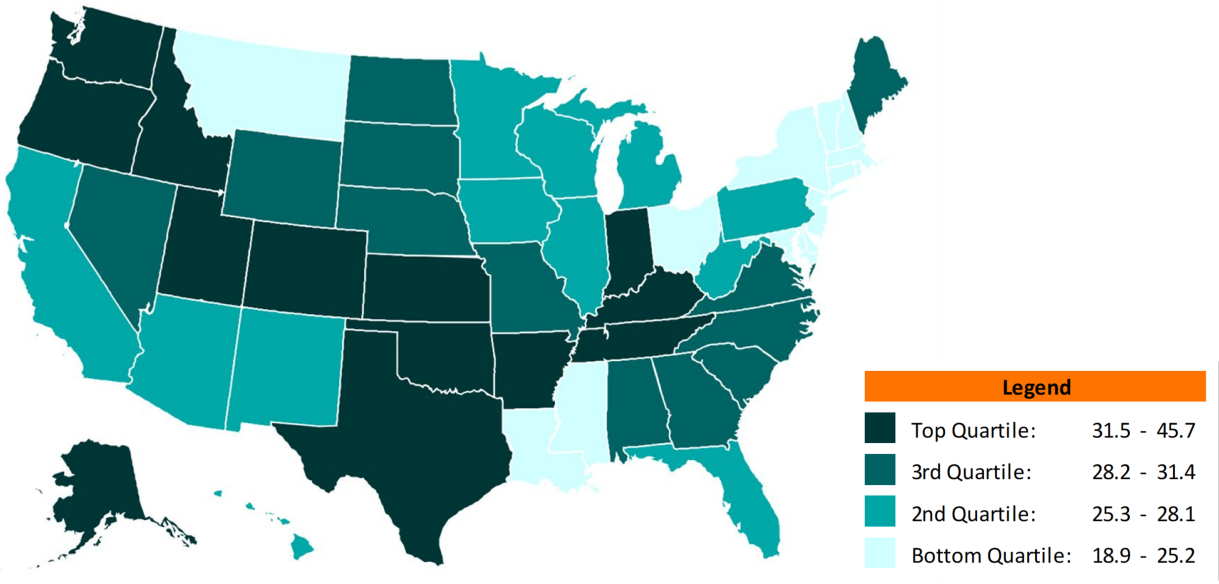
7. Do you need a chart?

Source: U.S. Census Bureau, American Community Survey, 2015 1-yr est.

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

7. Do you need a chart?

Geographic Variation of Women's Adjusted Marriage Rate Among States, 2020



Source: NCFMR analyses of U.S. Census Bureau, American Community Survey, 1-year Experimental PUMS, 2020

BGSU | Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Other Helpful Sites

https://datavizproject.com/

http://circos.ca/

https://www.rapidtables.com/convert/color/hex-to-rgb.html

https://www.rawgraphs.io/

Stata

- https://www.trentonmize.com/teaching/dmv

Colors

- https://www.colorhexa.com/00338d-to-ffffff
- https://mycolor.space/
- https://designsystem.digital.gov/design-tokens/color/overview/

Accessibility

- http://colorsafe.co/
- https://contrastchecker.com/
- https://webaim.org/resources/contrastchecker/

BGSU Center for Family and Demographic Research
BOWLING GREEN STATE UNIVERSITY

Krista Westrick-Payne, PhD.

A PUBLIC UNIVERSITY FOR THE PUBLIC GOOD ®