CENTER FOR FAMILY & DEMOGRAPHIC RESEARCH

Making Graphs with Excel

Summer 2014 Workshop Series

why-

CHARTS?

Picture Superiority Effect

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

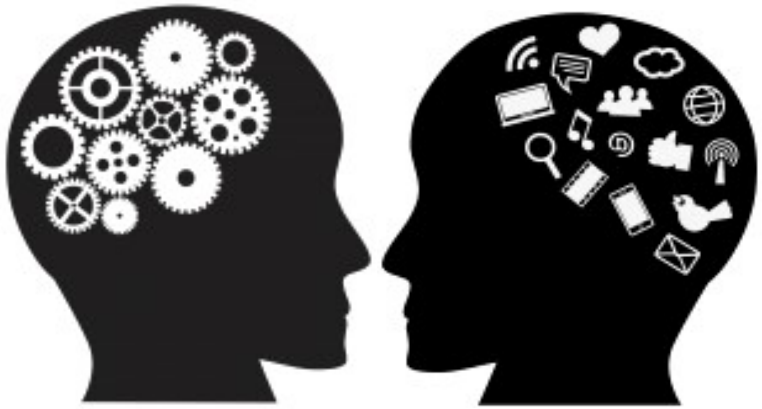
Fruit <

Why is it so difficult for–

SOCIOLOGISTS?

How do you process?

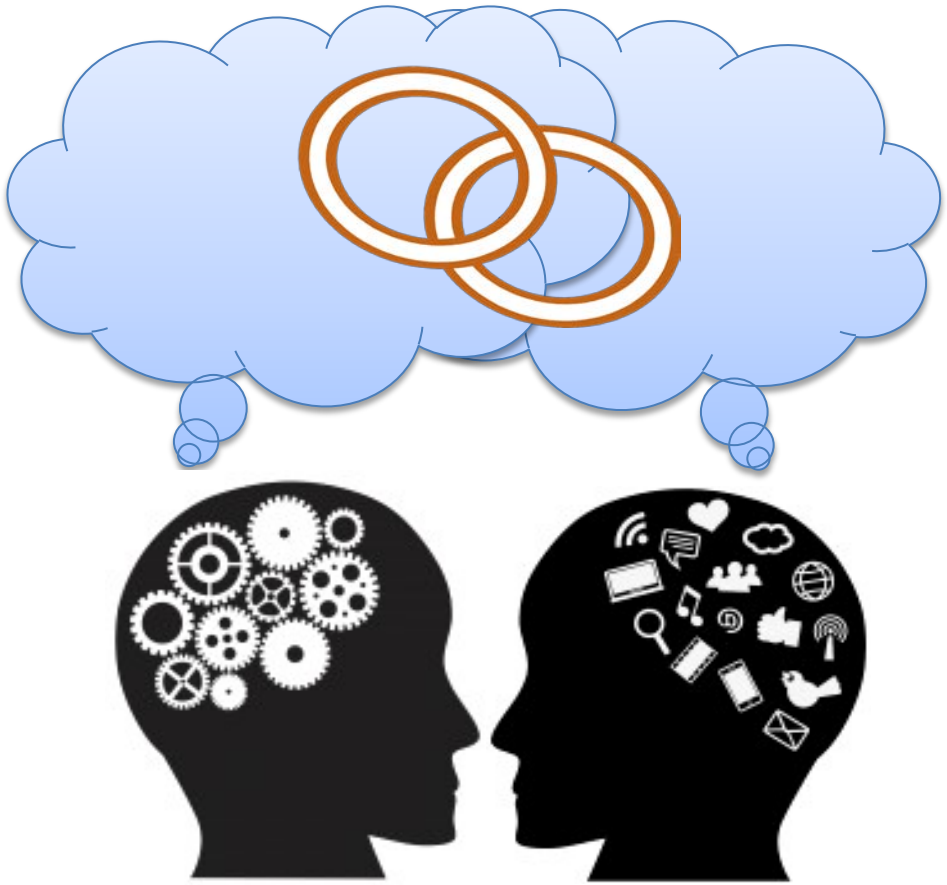
- **Analytical**
- **Logical**
- **Precise**
- **Repetitive**
- **Organized**
- **Details**
- **Scientific**
- **Detached**
- **Literal**
- **Sequential**



- **Creative**
- **Imaginative**
- **General**
- **Intuitive**
- **Conceptual**
- **Big picture**
- **Heuristic**
- **Empathetic**
- **Figurative**
- **Irregular**

I Propose we Marry the Two

The pun is intended!



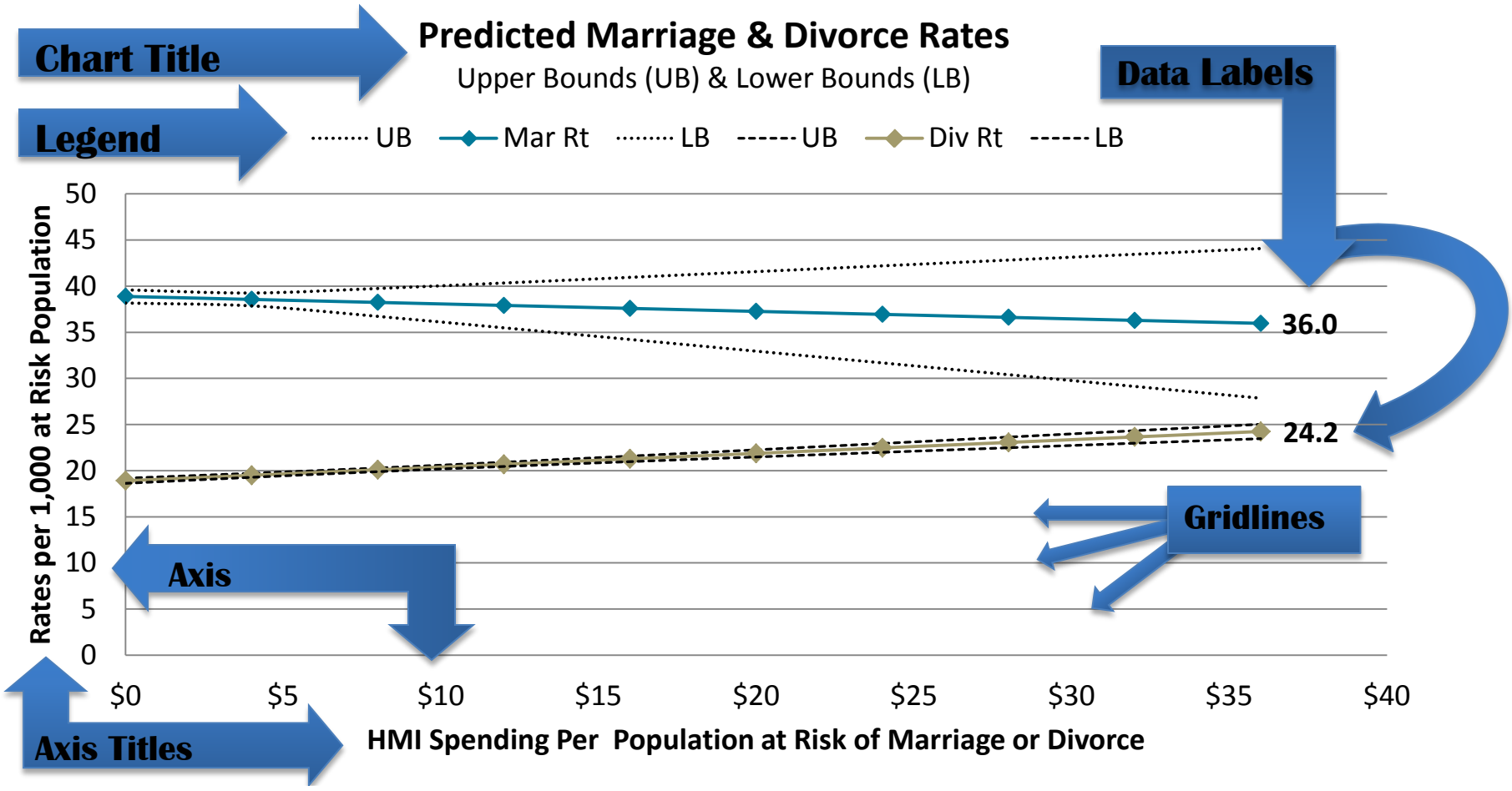
Organization of Presentation

- **Structure of an Excel Chart**
- **Different Types of Excel Charts**
- **Basic Principles of Chart Design**
- **Graphing Interaction Effects**
- **Creating a Chart with a Double Axis**

What makes up-

THE STRUCTURE OF AN EXCEL CHART?

Let's Dissect...



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

Source

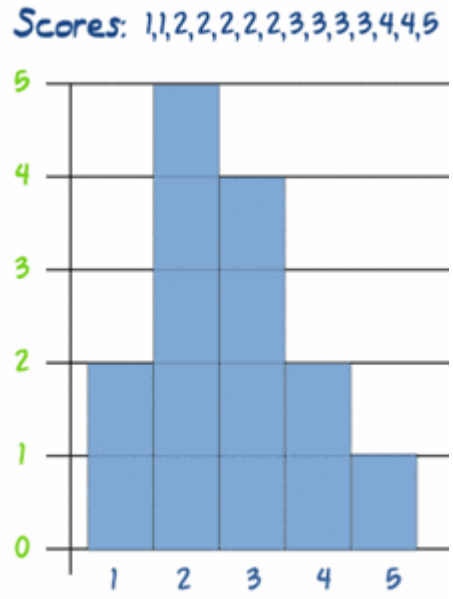
What are-

THE DIFFERENT TYPES OF CHARTS?

Histograms

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

Histograms, example



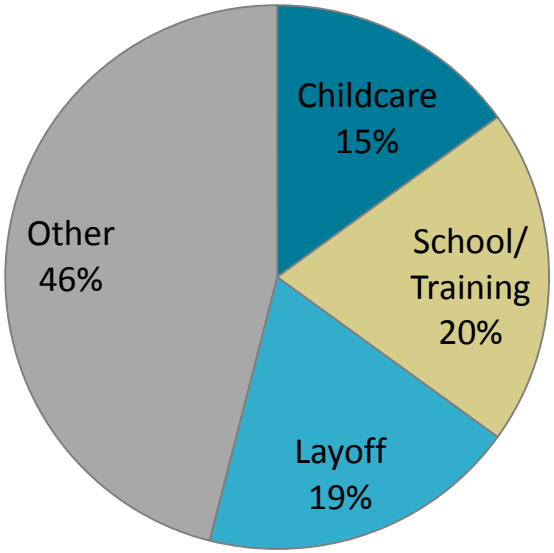
Pie Charts

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

Pie Charts, examples

Simple Pie

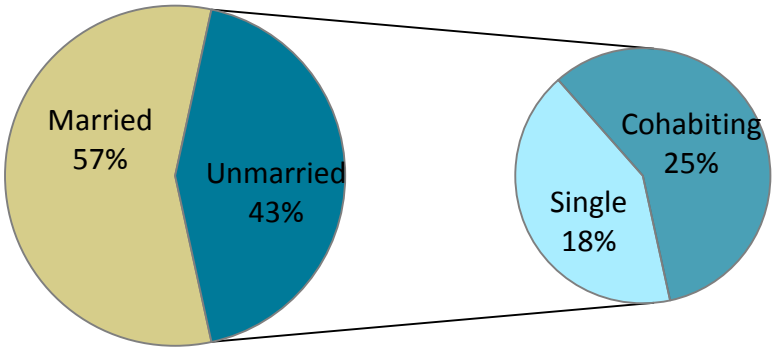
Top Reasons for Fathers Leaving the Workforce in 2008



Source: Survey of Income and Program Participation, 2008 March Supplement

Pie-of-Pie

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

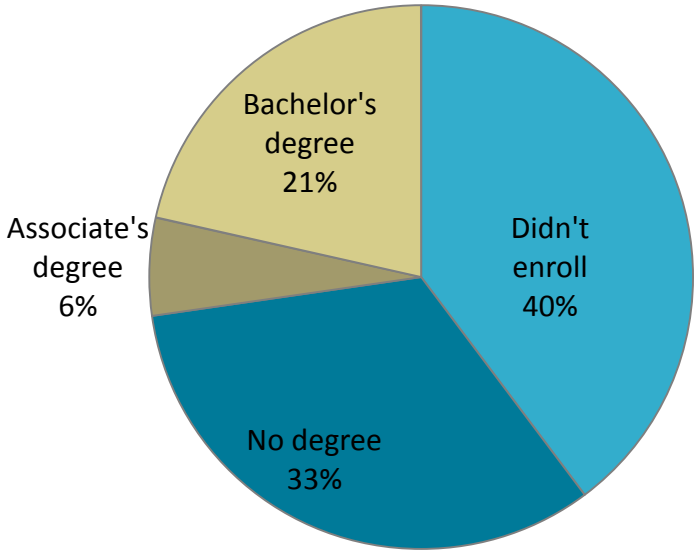


Source: NSFG 2006-2010

Pie Charts, examples

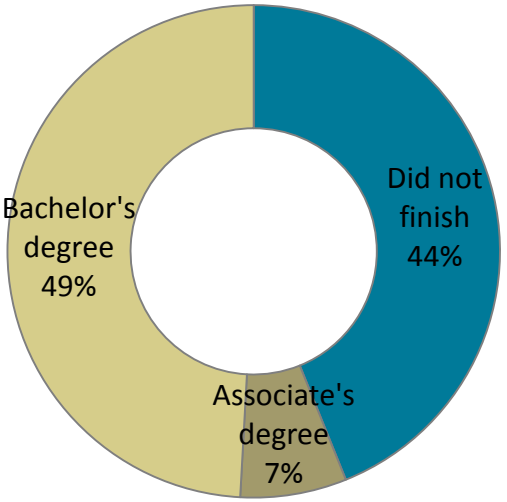
Simple Pie

College experiences of young adults (by age 25)



Doughnut

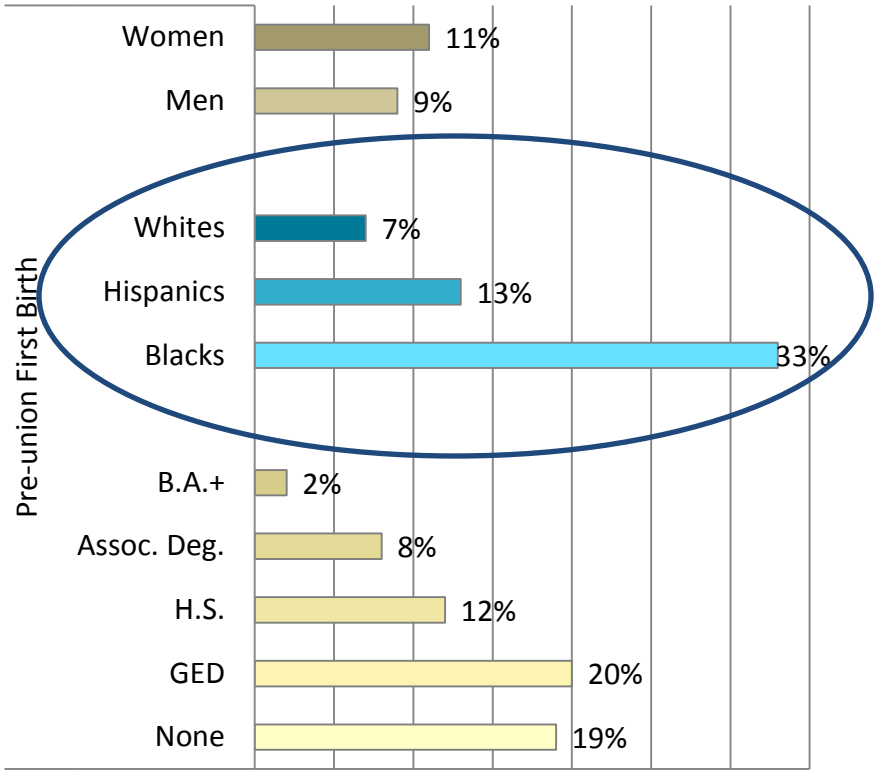
Percent of young adults who enroll in a 4-year program by degree earned by age 25



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.

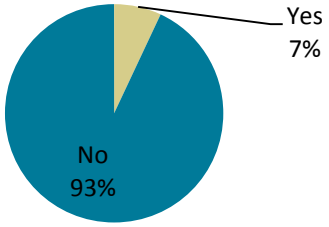
Bar Chart, example

Prevalence of Pre-union First Birth across Demographic Characteristics

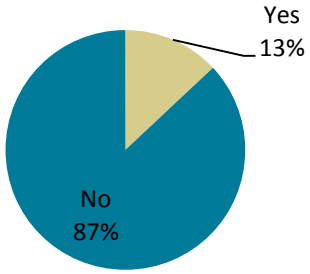


Prevalence of Pre-union First Birth by Race/Ethnicity:

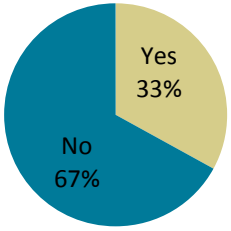
Whites



Hispanics



Blacks



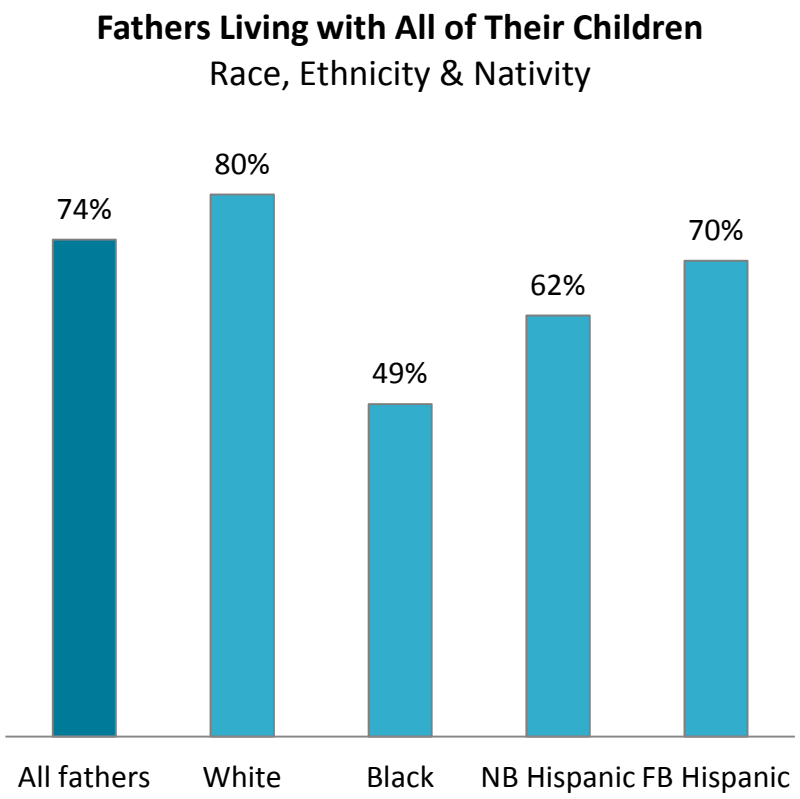
Source: National Longitudinal Survey of Youth 1997 (NLSY97), Rounds 1-13: 1997-2009 (weighted). U.S. Department of Labor, Bureau of Labor Statistics, NCFMR analyses of valid cases.

Column & Bar Charts

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

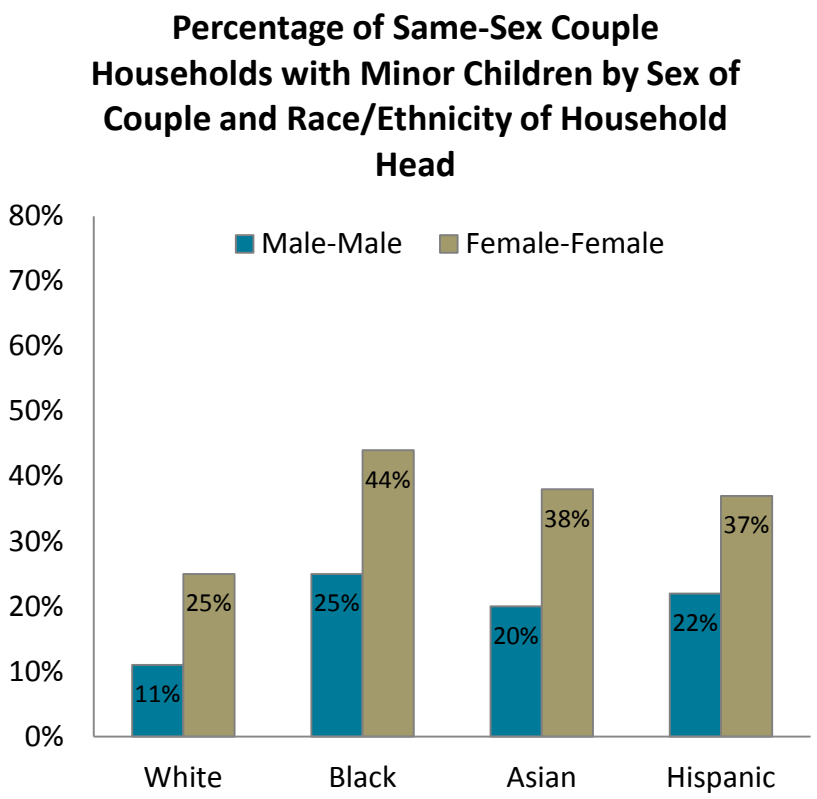
Column Charts, examples

Simple



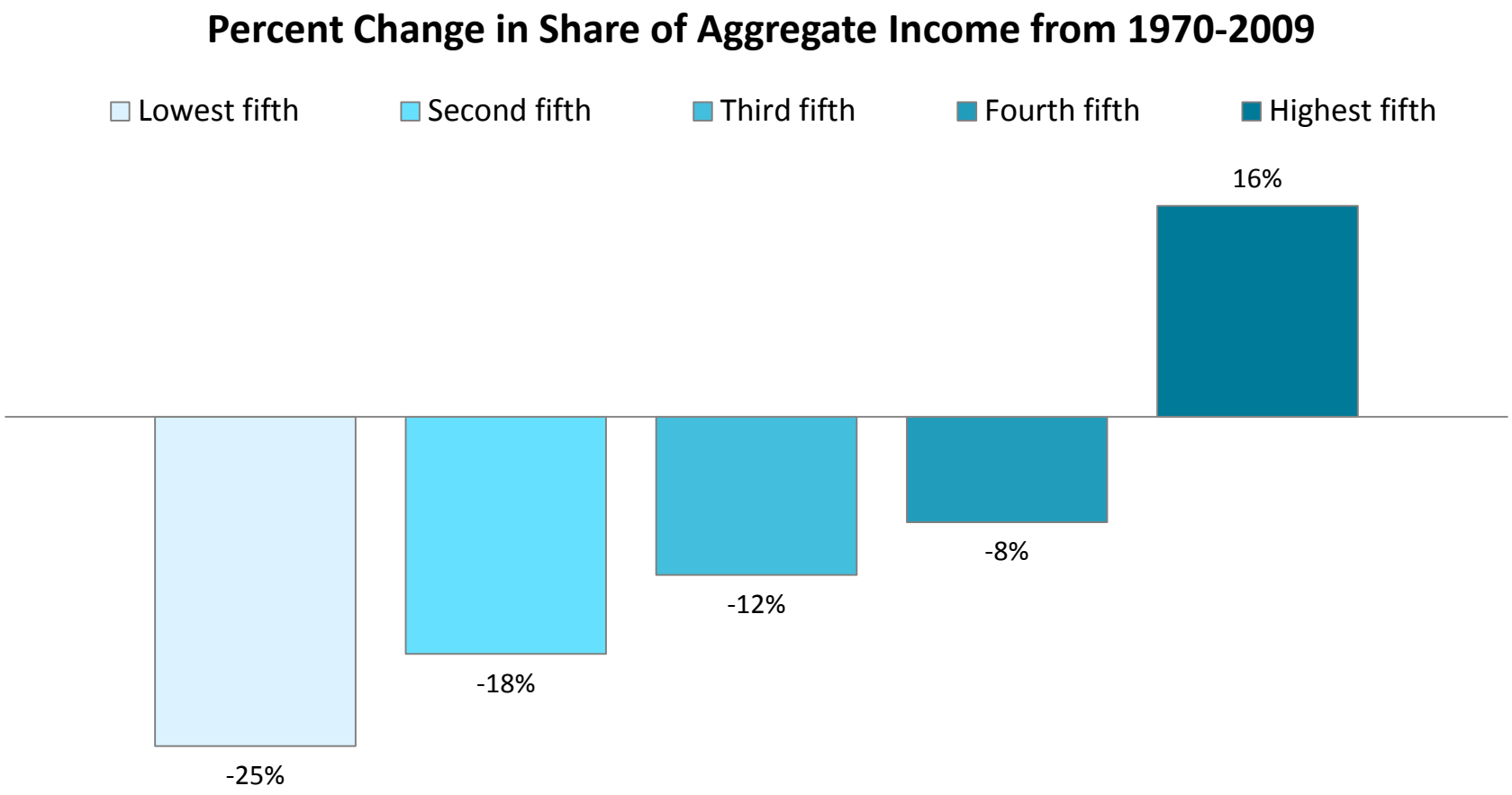
Source: NSFG 2006-2010

Side-by-Side



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

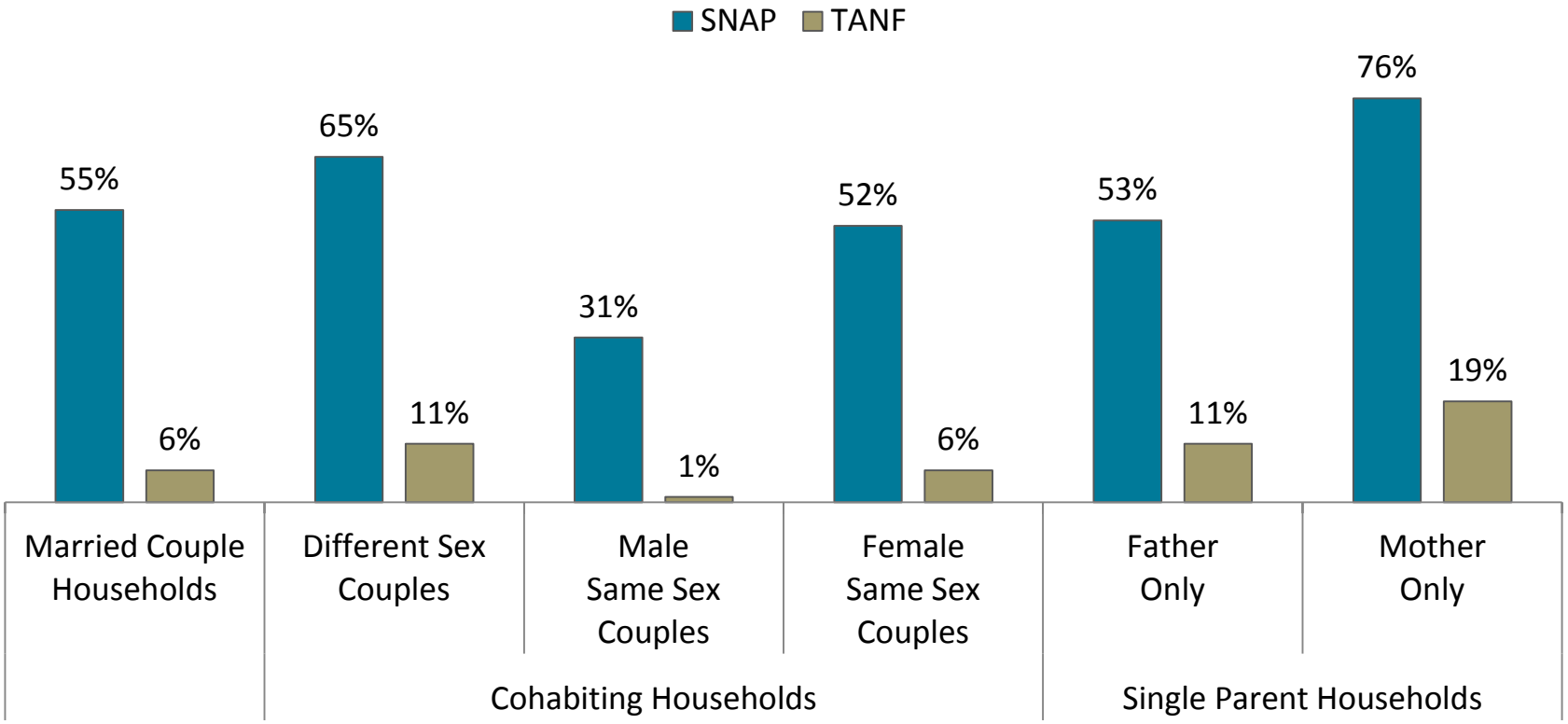
Column Charts, examples



Source: U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplements

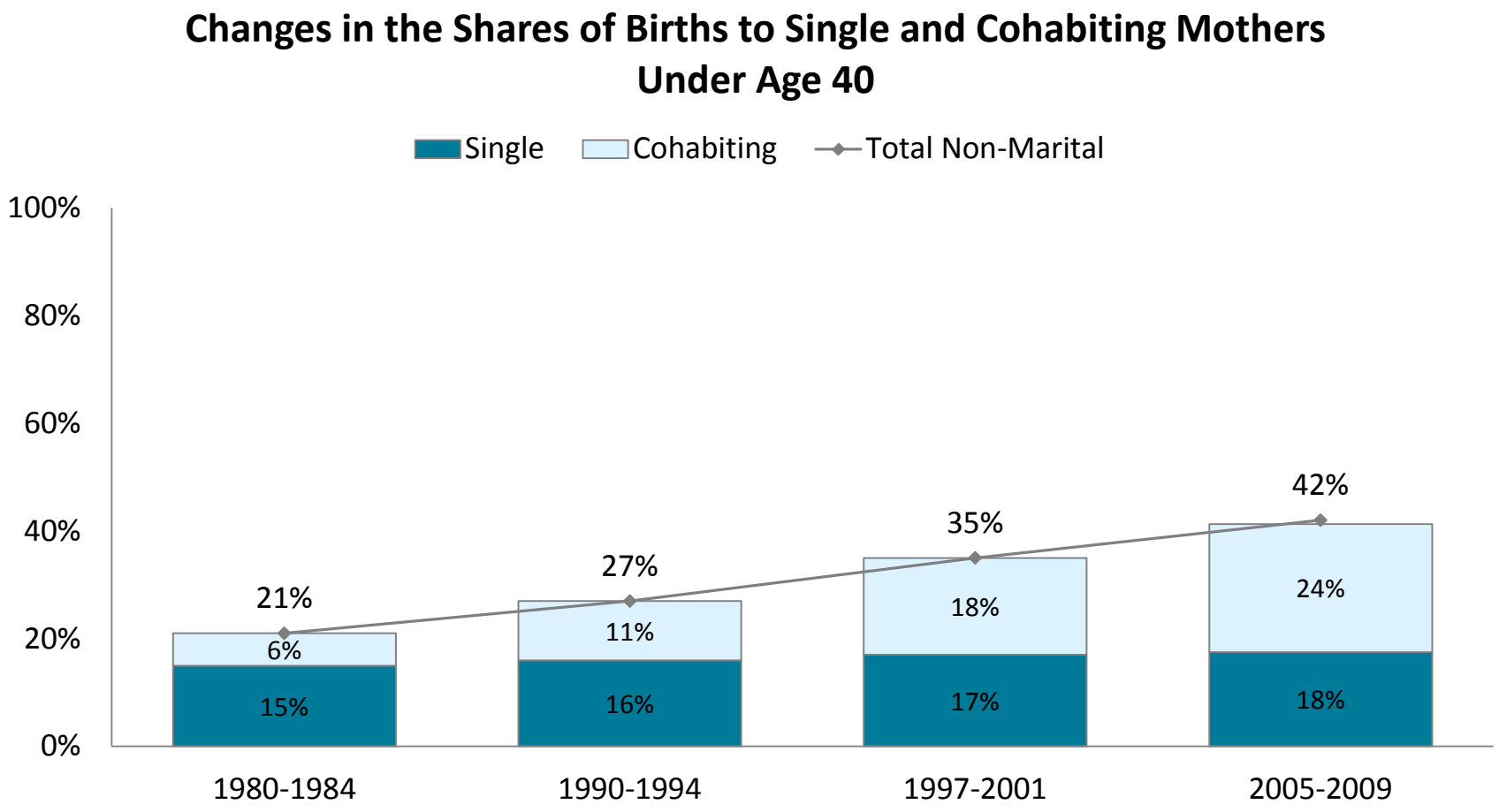
Column Charts, examples

Public Assistance Participation among U.S. Children in Poverty
by Family Structure, 2010



Source: U.S. Census Bureau, American Community Survey, 1-Year Estimates, 2010

Column Charts, examples



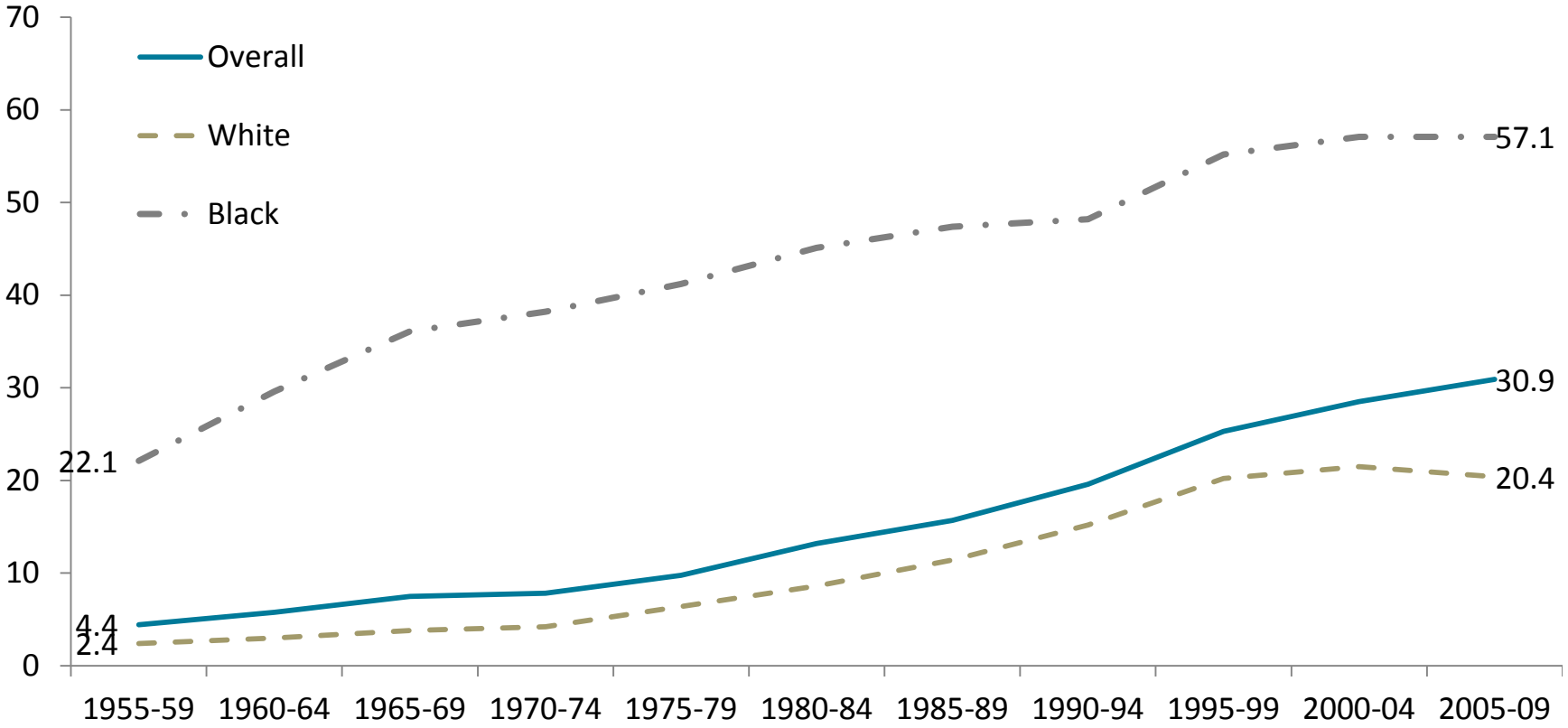
Sources: 1980-1984 data, Bumpass & Lu (2000) using NSFH, 1987/1988; 1990-1994 & 1997-2001 data, Kennedy & Bumpass (2008) using NSFG 1995 & NSFG 2002; 2005-2009, NCFMR analyses using NSFG 2006-2010.

Line Charts

Ideal for showing trends over time

Line Charts, examples

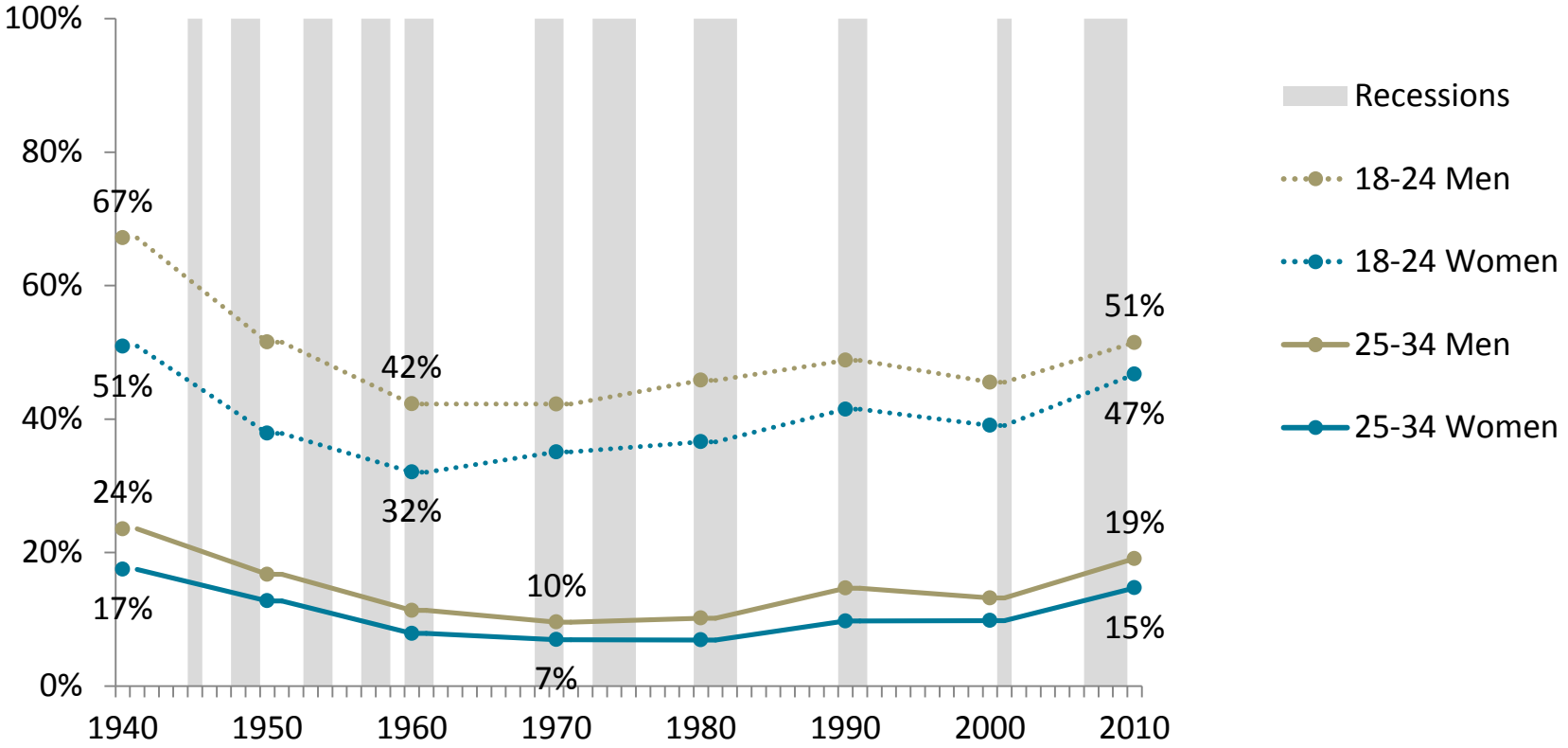
Share of Married Mothers Experiencing a Premarital Birth, by Race and Marriage Cohort



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).

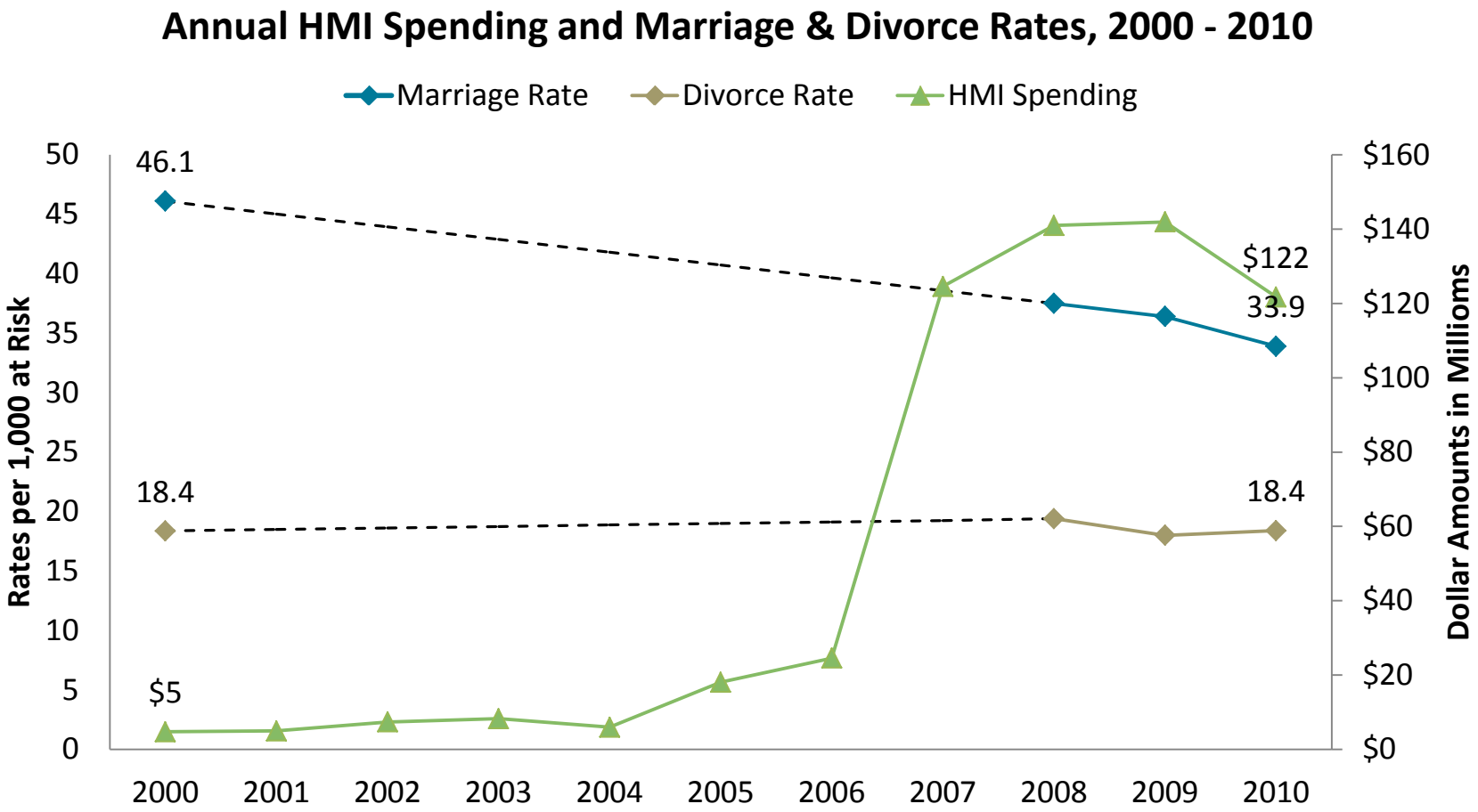
Line Charts, examples

Young Adults Living in a Parent's Household and Economic Recession Years by Sex and Ages, 1940-2010



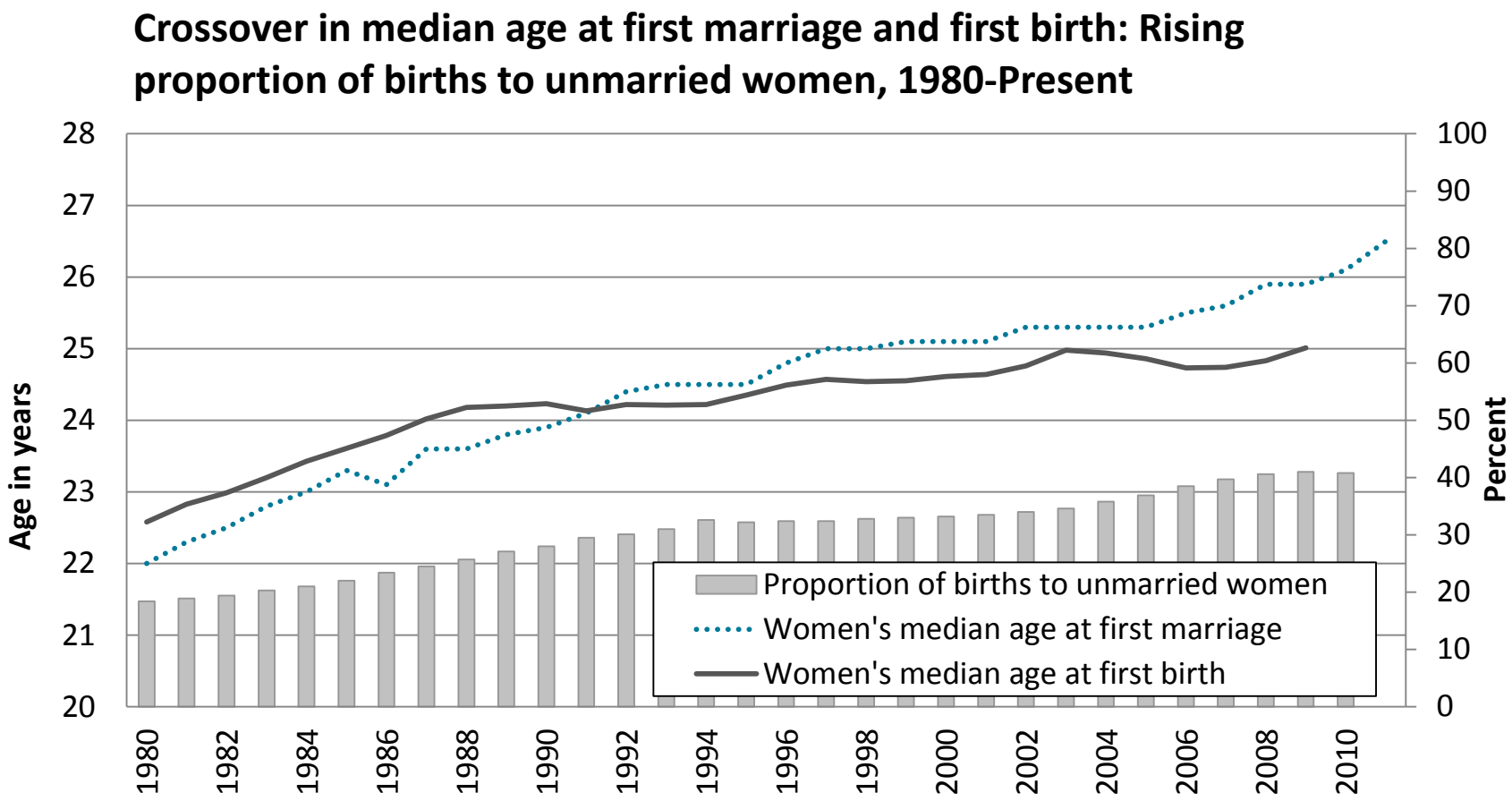
Source: U.S. Census Bureau, Decennial Census, 1940-2000 (IPUMS); U.S. Census Bureau, American Community Survey, 1-year estimates 2010 (IPUMS)

Line Charts, examples



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.

Line Charts, examples



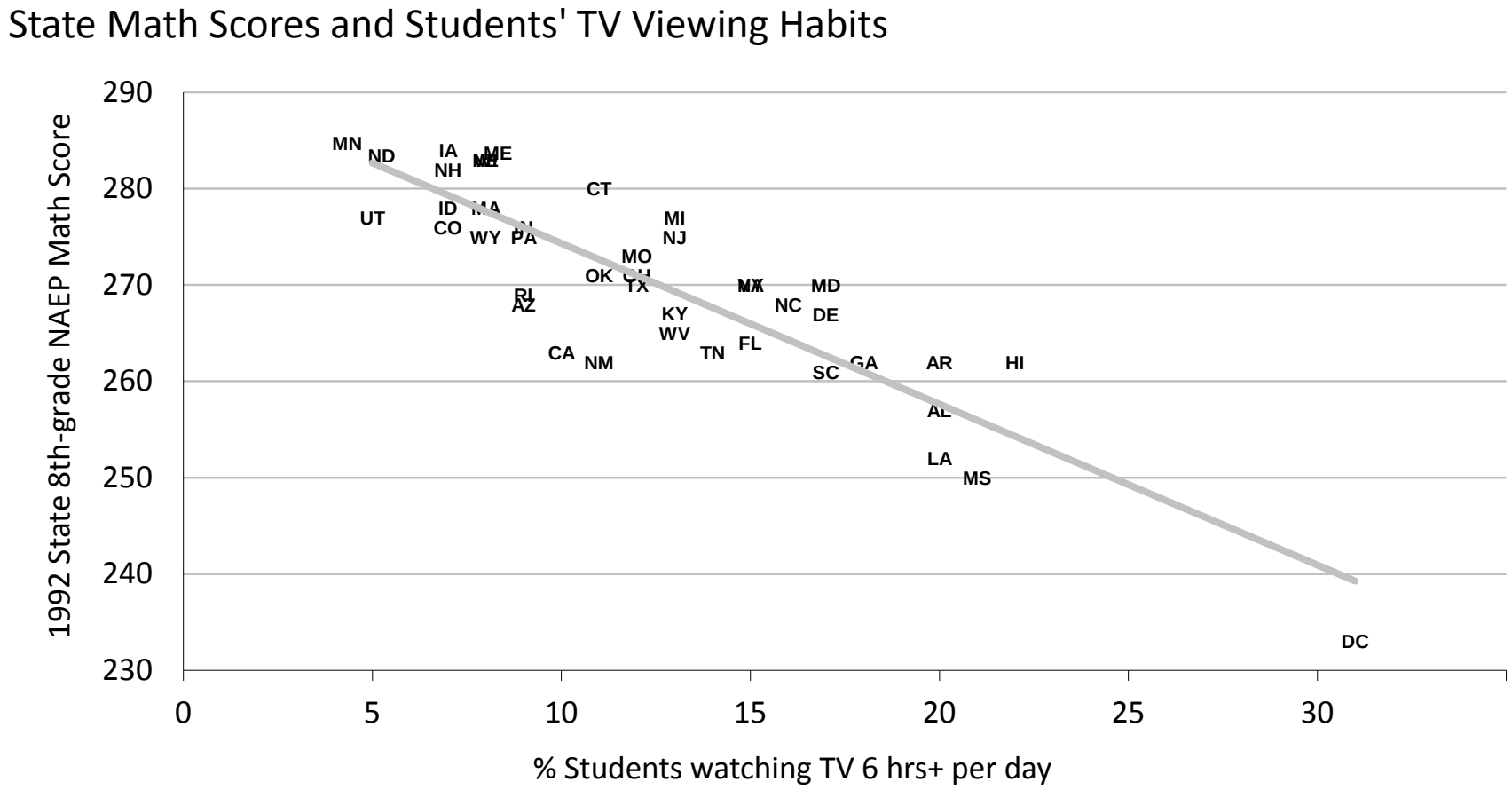
Sources:

1. U.S. Census Bureau, Current Population Survey, March and Annual Social and Economic Supplements, 2012 and earlier.
2. Centers for Disease Control and Prevention. National Center for Health Statistics. Vital Stats. http://www.cdc.gov/nchs/vitalstats.htm. [March 2013].
3. Martin JA, Hamilton BE, Ventura SJ, et al. Births: Final data for 2009. National vital statistics reports; vol 60 no 1. Hyattsville, MD: National Center for Health Statistics. 2011.
4. Hamilton BE, Martin JA, Ventura SJ. Births: Preliminary data for 2010. National vital statistics reports web release; vol 60 no 2. Hyattsville, MD: National Center for Health Statistics. 2011.

Scatter Plots

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

Scatter Plots, example

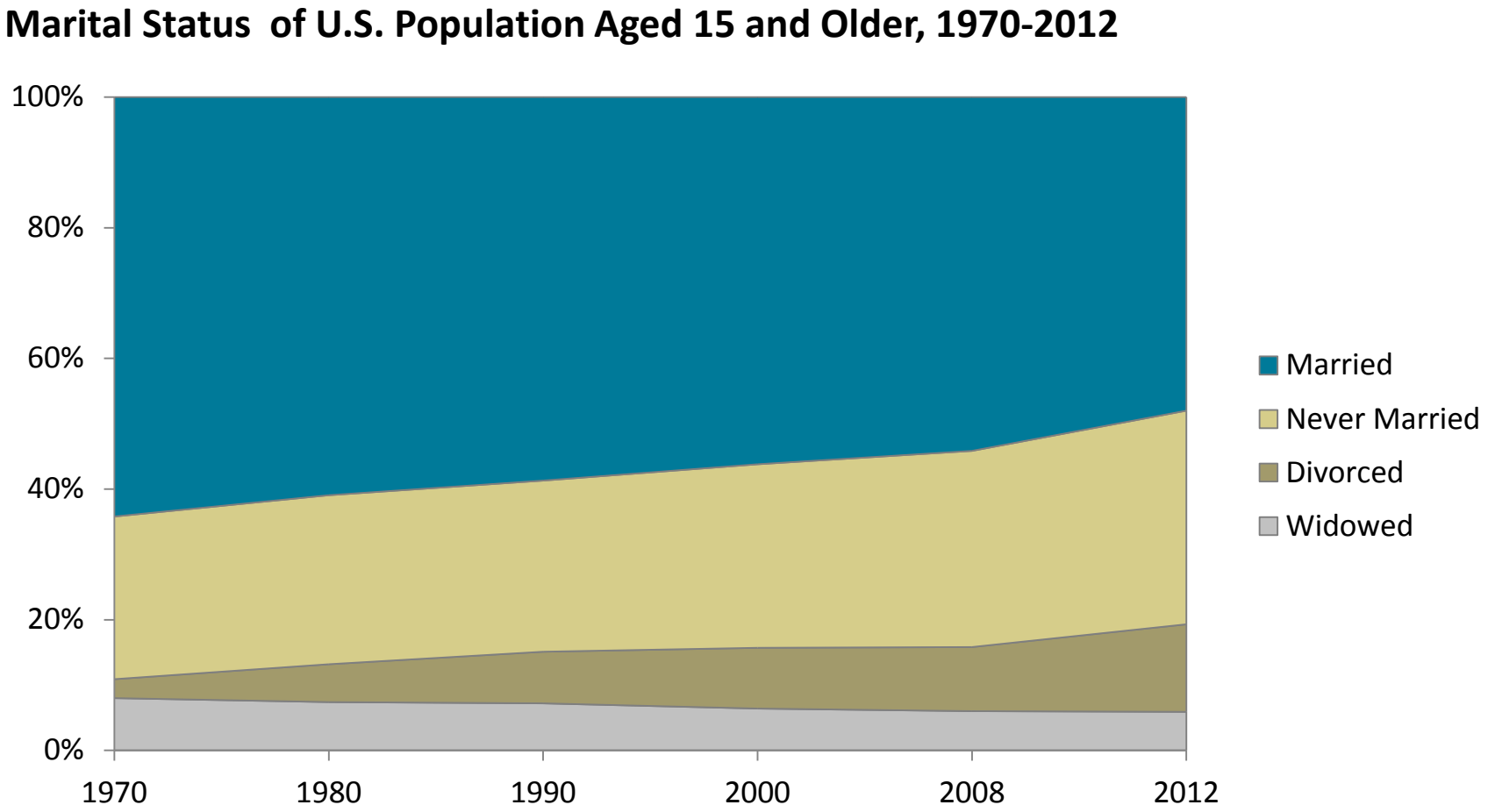


Source: National Center for Educational Statistics, 1994

Area Charts

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

Area Charts, example



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)

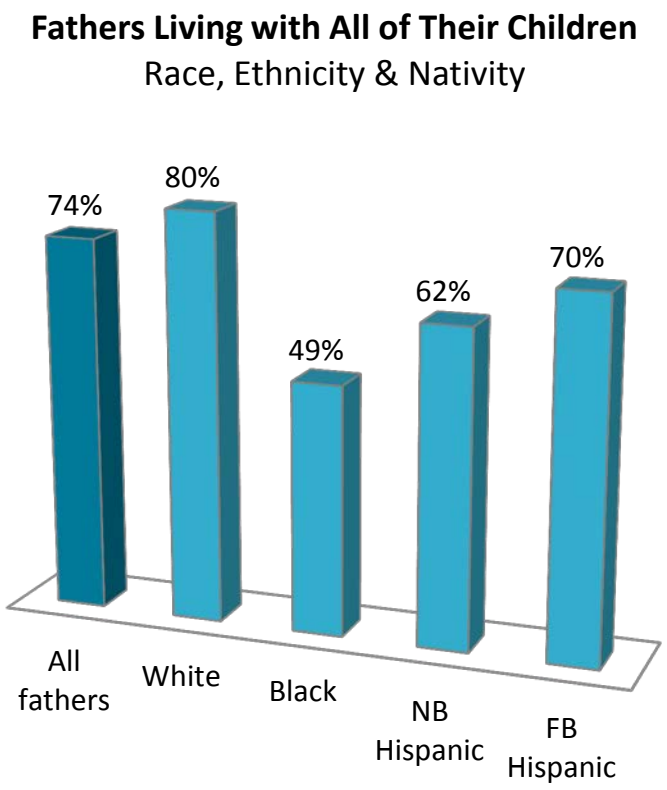
What are some-

BASIC PRINCIPLES OF CHART DESIGN?

1. Simplify

- **Minimize ink-to-data ratio**
- **Remove unneeded chart elements**
  - **Gridlines**
  - **Chart borders**
  - **Axis titles**
  - **Legends**
  - **Markers & data labels**
  - **Decimal points (in axis & data labels)**
  - **Trend lines**
  - **NO 3D charts!!!!!!!!!!!!!!!!!!**
- **Sort data in a meaningful way**

Example of a 3D Chart:



2. Color vs. Black & White

- **When in doubt → black & white**
- **Color *can* help tell a story**
- **Color = branding (e.g. CFDR, NCFMR, BGSU)**
  - **Use a cohesive and consistent color palette**
  - **Be mindful of how audience will view**
    - **Excel vs. Word vs. PDF**
    - **Color vs. B&W print copy**

3. Do NOT Use Distorted Charts

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

4. Present Related Charts Simultaneously

- **One-after-another or side-by-side if possible**
- **Emphasizes importance of appropriate axis and scales**

5. Know Your Audience

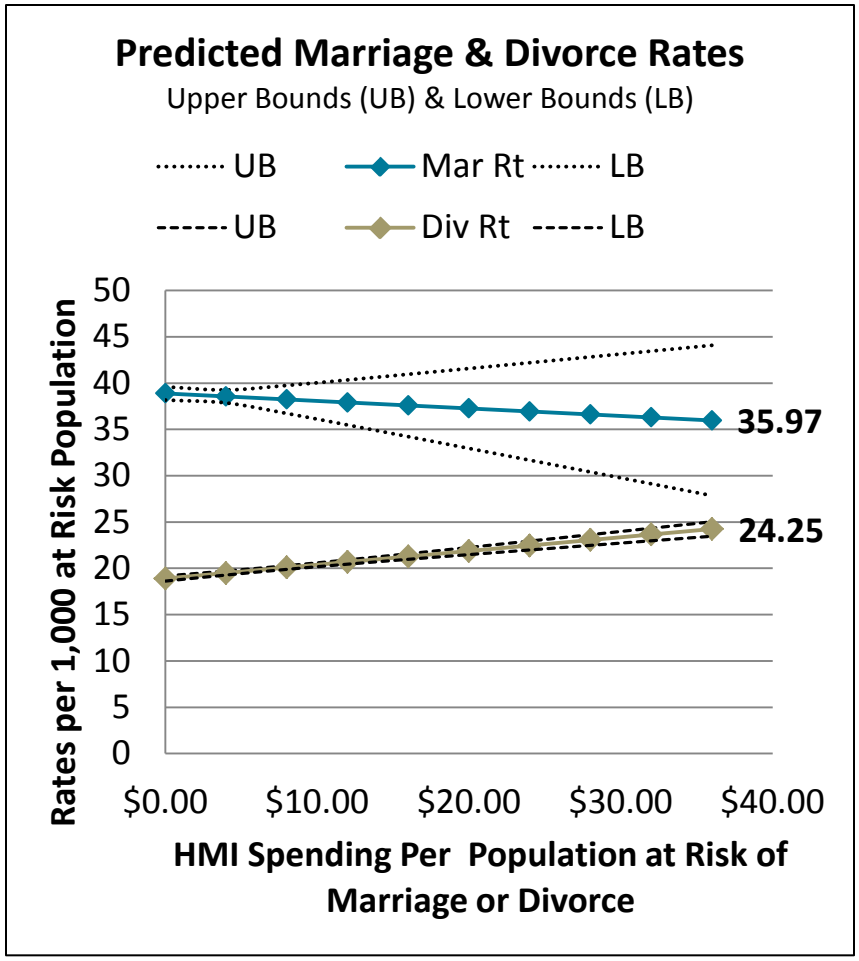
- **Academics vs. lay folks**
- **Undergraduate students vs. graduate students**
- **Graduate students vs. professors**
- **PAA presentation vs. job talk**

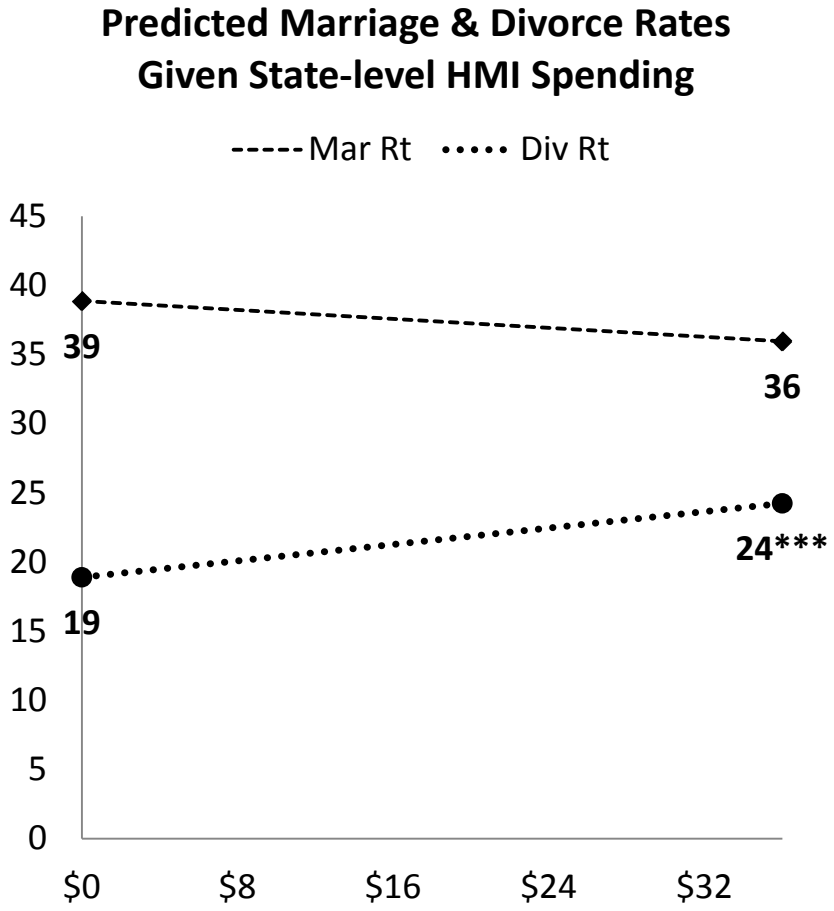
6. TMC = TMI

- **Too many charts (TMC) is as bad as too much information (TMI) → Do NOT overload your audience!**

Let's apply some principles:

Which is easier to understand?





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

7. Do you need a chart?

$117 million

The amount annual HMI spending in the U.S. increased from 2000 – 2010

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

How do I-

CHART INTERACTION EFFECTS?

Logistic Regression Predicting Ever Marrying

- **An interaction between a categorical and continuous predictor** (DeMaris 2004, p 143)**:**

$$E(Y) = \beta_0 + \delta_1 Black + \beta_1 Parity + \Upsilon_1 Black*Parity$$

  - **The subpop consists of only White and Black women**
  - **Black is a dummy variable**
  - **Parity indicates number of live births, range 0-15**
  - **Analyses is weighted**

Logistic Regression Predicting Ever Marrying, cont.

- **Stata Output for Full Model:**

```
. svy, subpop(blkwht): logistic evermar black PARITY PARITYblk, coef
(running logistic on estimation sample)

Survey: Logistic regression

Number of strata   =        56                  Number of obs      =     12279
Number of PSUs     =       152                  Population size    =  61754741
                                                Subpop. no. of obs =      8568
                                                Subpop. size       =  45835139
                                                Design df          =        96
                                                F(   3,     94)    =    186.25
                                                Prob > F           =    0.0000
------------------------------------------------------------------------------
             |             Linearized
     evermar |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       black |  -.4698438   .1172022    -4.01   0.000    -.7024885   -.2371992
      PARITY |   1.458909   .0707637    20.62   0.000     1.318444    1.599374
   PARITYblk |  -.9253343   .0978554    -9.46   0.000    -1.119576   -.7310928
       _cons |  -.8652098   .0616793   -14.03   0.000    -.9876423   -.7427772
------------------------------------------------------------------------------
```

Logistic Regression Predicting Ever Marrying, cont.

- **Table of Results**

Logistic Regression Predicting Ever Marrying

	Model 1 (Zero-Order)		Model 2		Model 3 (Full)	
	Coef.	SE	Coef.	SE	Coef.	SE
Black	-0.854	0.325***	-1.589	0.113***	-0.470	0.117***
Parity	1.040	0.054***	1.150	0.053***	1.459	0.071***
Black X Parity					-0.925	0.098***
Constant			-0.679	0.06 ***	-0.865	0.062***

Logistic Regression Predicting Ever Marrying, cont.

- **Equation for Full Model**

  $E(Y) = \beta_0 + \delta_1 Black + \beta_1 Parity + \Upsilon_1 Black*Parity$

- **Equation for Black Women**

  $E(Y) = \beta_0 + \delta_1 + \beta_1 Parity + \Upsilon_1 *Parity$

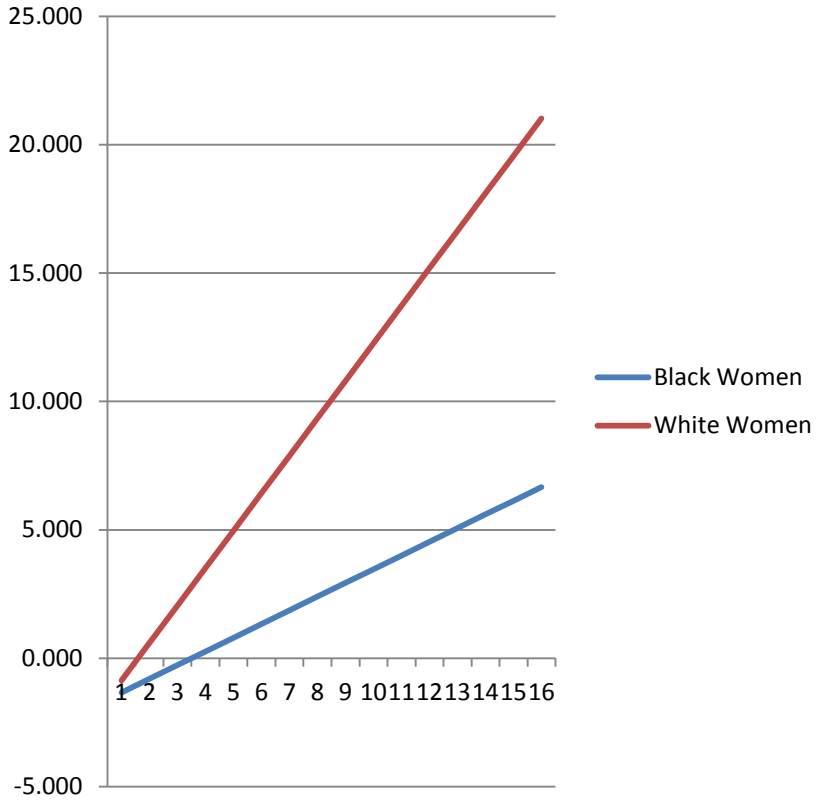
- **Equation for White Women**

  $E(Y) = \beta_0 + \beta_1 Parity$

- **Now, Plug and Play in Excel!**

Logistic Regression Predicting Ever Marrying, cont.

Unformatted



Formatted