Making a figure with Stata or Excel

Biostatistics 212

Lecture 6

Agenda

1. Figure basics
   - Why make a figure?
   - Types of figures
   - Elements of a figure
2. Steps in making a figure with Excel
3. Steps in making a figure with Stata

Why use figures?

- When a graphical display of information more effectively conveys the intended message than words.

- “A picture is worth a thousand words”

Types of Figures

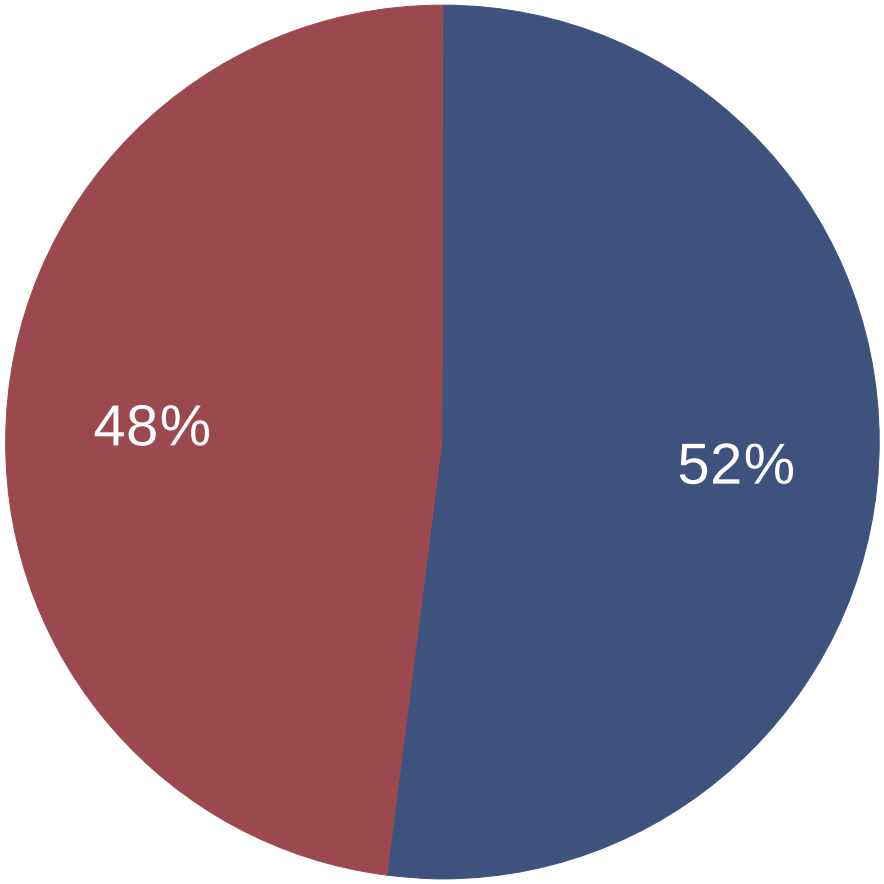
1. Photographs
2. Diagrams
3. Figures that Present Numerical Data

- Pie charts
- Scatter plots
- Bar graphs
- Line graphs

Figures that Present Numerical Data

- GOOD for presenting overall effects
- NOT GOOD for presenting specific measurements.

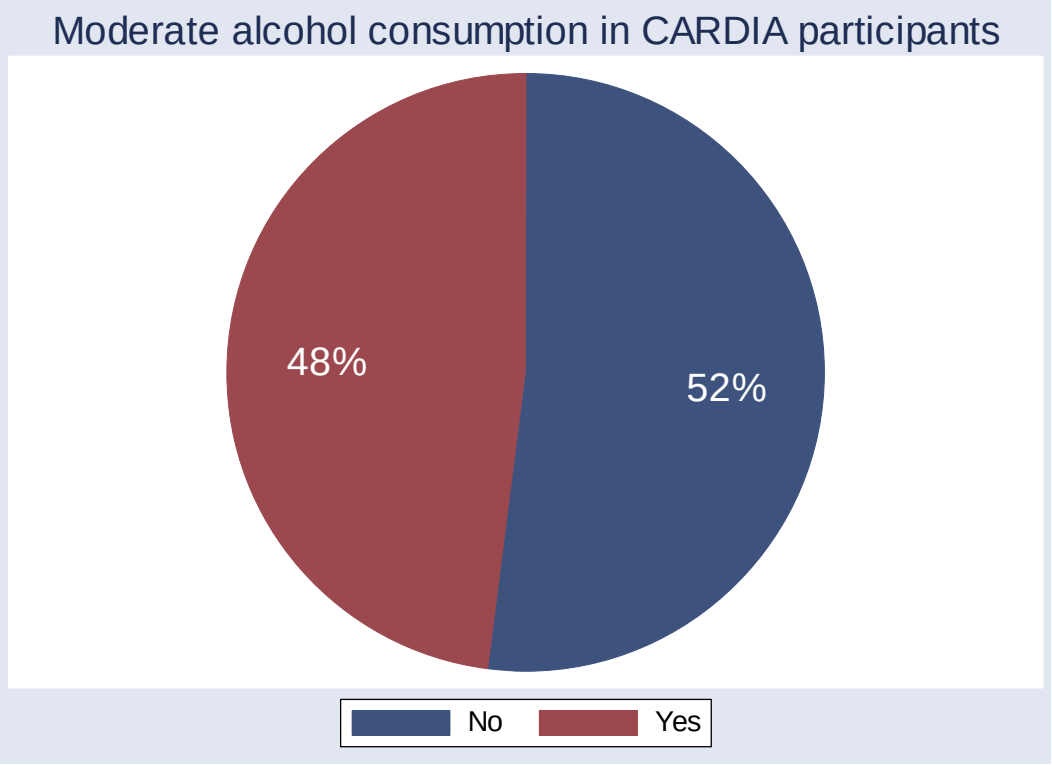
Moderate alcohol consumption in CARDIA participants



Figures

"A picture is worth a thousand words"

How many words is this picture worth?



Figures

"A picture is worth a thousand words"

How many words is this picture worth?

48% of CARDIA participants consume alcohol moderately.

Worth = 7 words

Figures - Reconsider if…

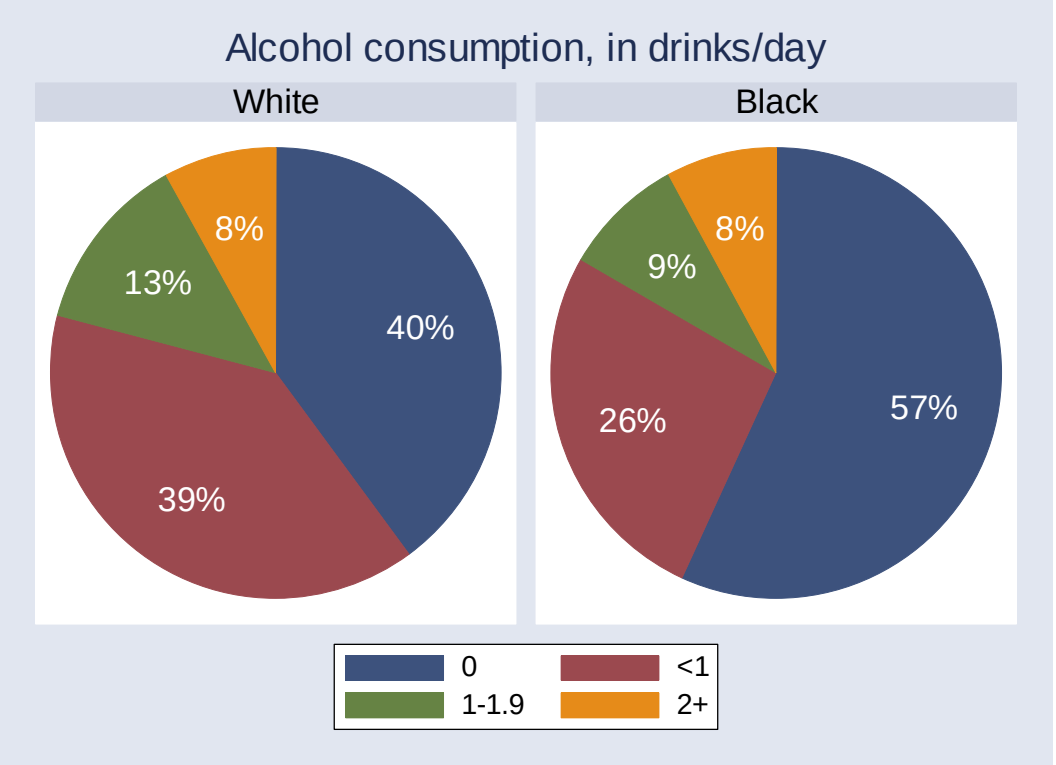
"Light on data"

- General guideline: Figures should have a minimum of 4 data points
- Writing a few sentences much easier then creating a figure.

Figures

"A picture is worth a thousand words"

How many words is this picture worth?



Figures

"A picture is worth a thousand words"

How many words is this picture worth?

	White	Black
Drinks/day	n=1935	n=1727
0	40%	57%
0.1-0.9	39%	26%
1-1.9	13%	9%
2+	8%	8%

Worth = 1 small table

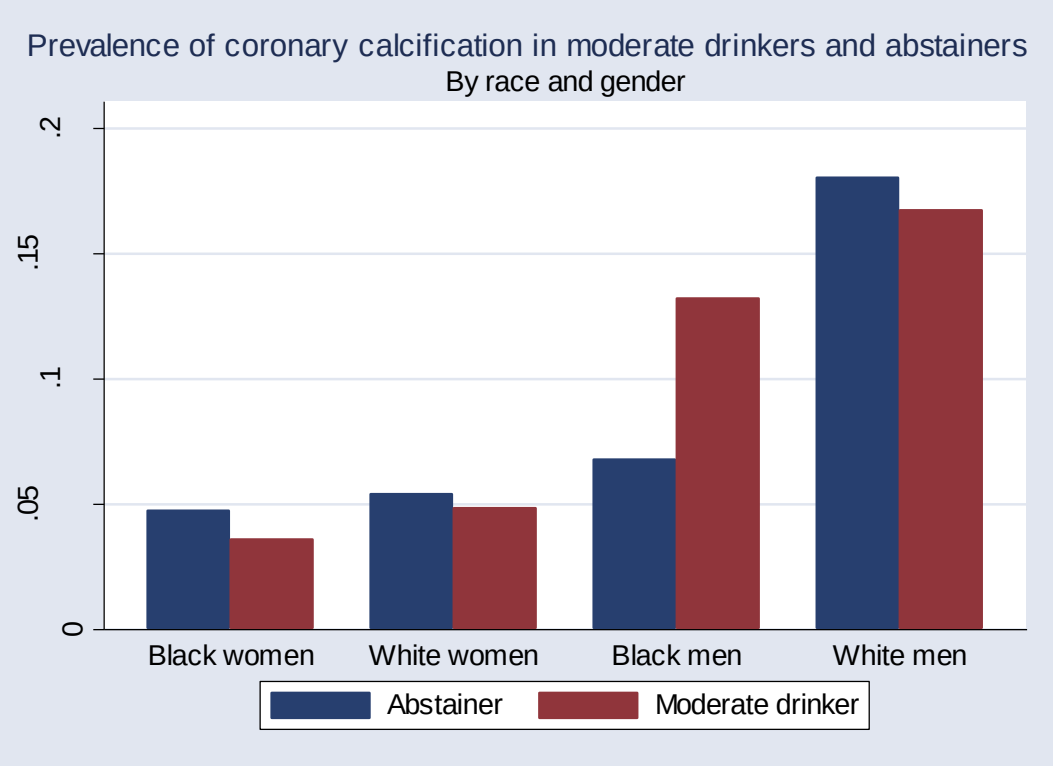
Figures versus Tables

- Numerical figures should be used when the overall pattern is more important than the actual values.
- Tables are preferred when the exact values are needed.

Figures

- “A picture is worth a thousand words”

How many words is this picture worth?



Figures

- “A picture is worth a thousand words”

How many words is this picture worth?

	% with CAC	
	Abstainer	Mod drinker
Black women	.047	.036
White women	.054	.049
Black men	.068	.132
White men	.180	.167

“Do I really need this figure?

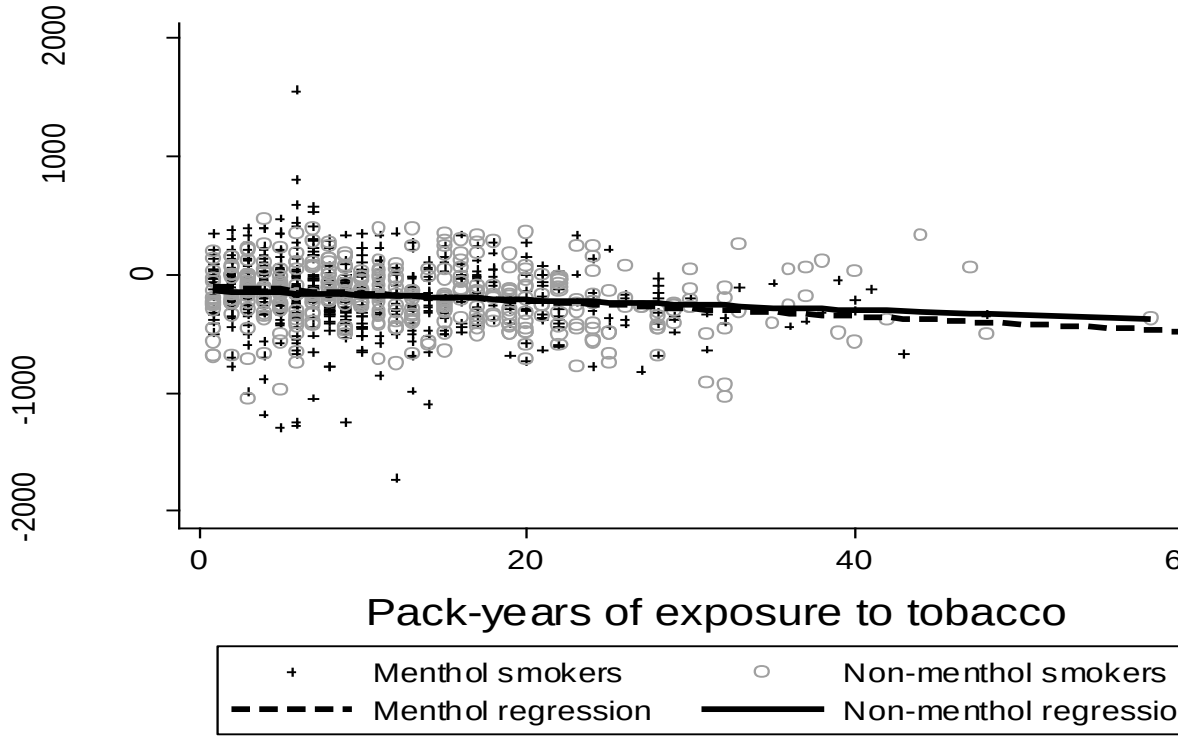
Always ask yourself:

- “Do I really need this figure?“
- “Does it make my point?”

Figures

- "A picture is worth a thousand words"

How many words is this picture worth?



Figures

- “A picture is worth a thousand words”

How many words is this picture worth?

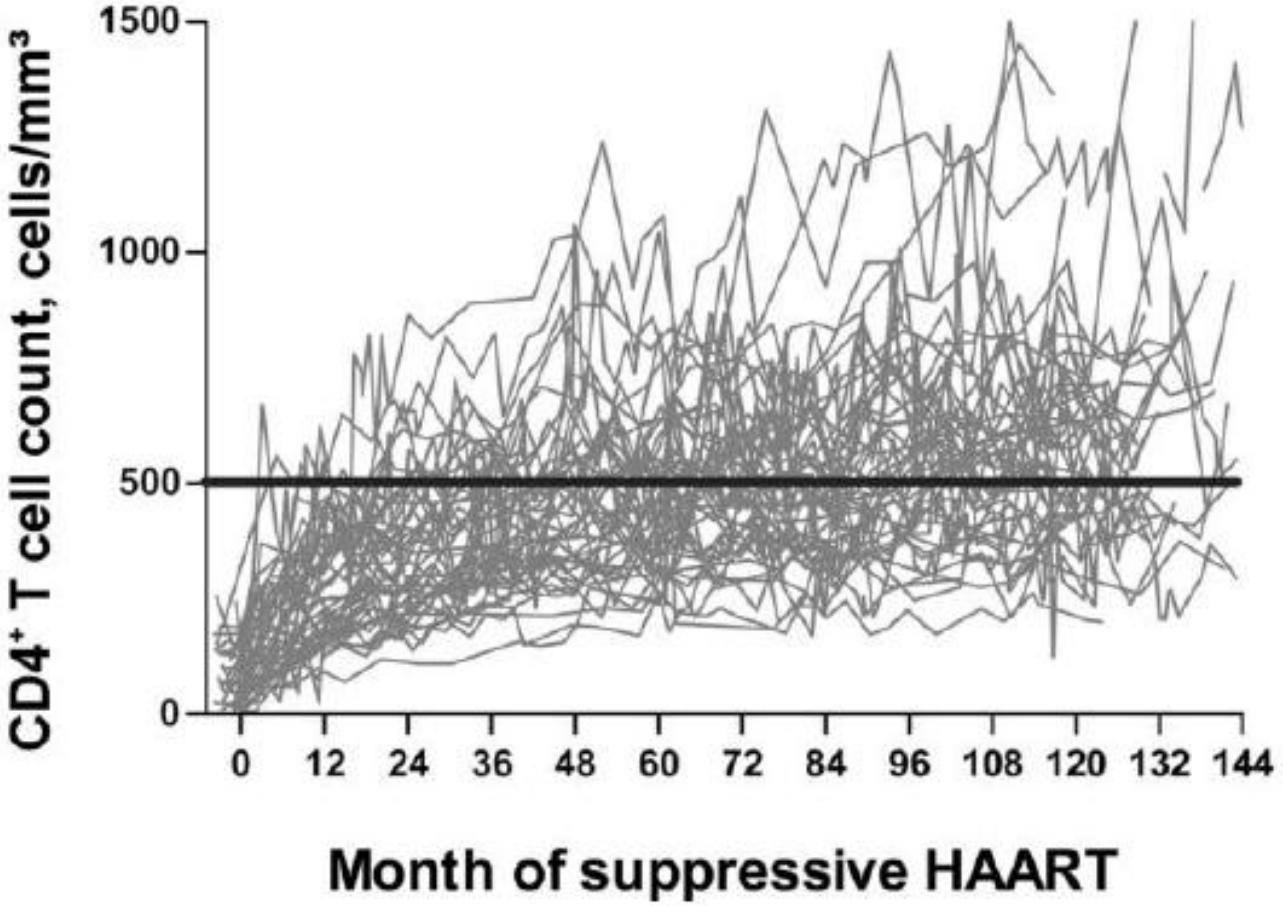
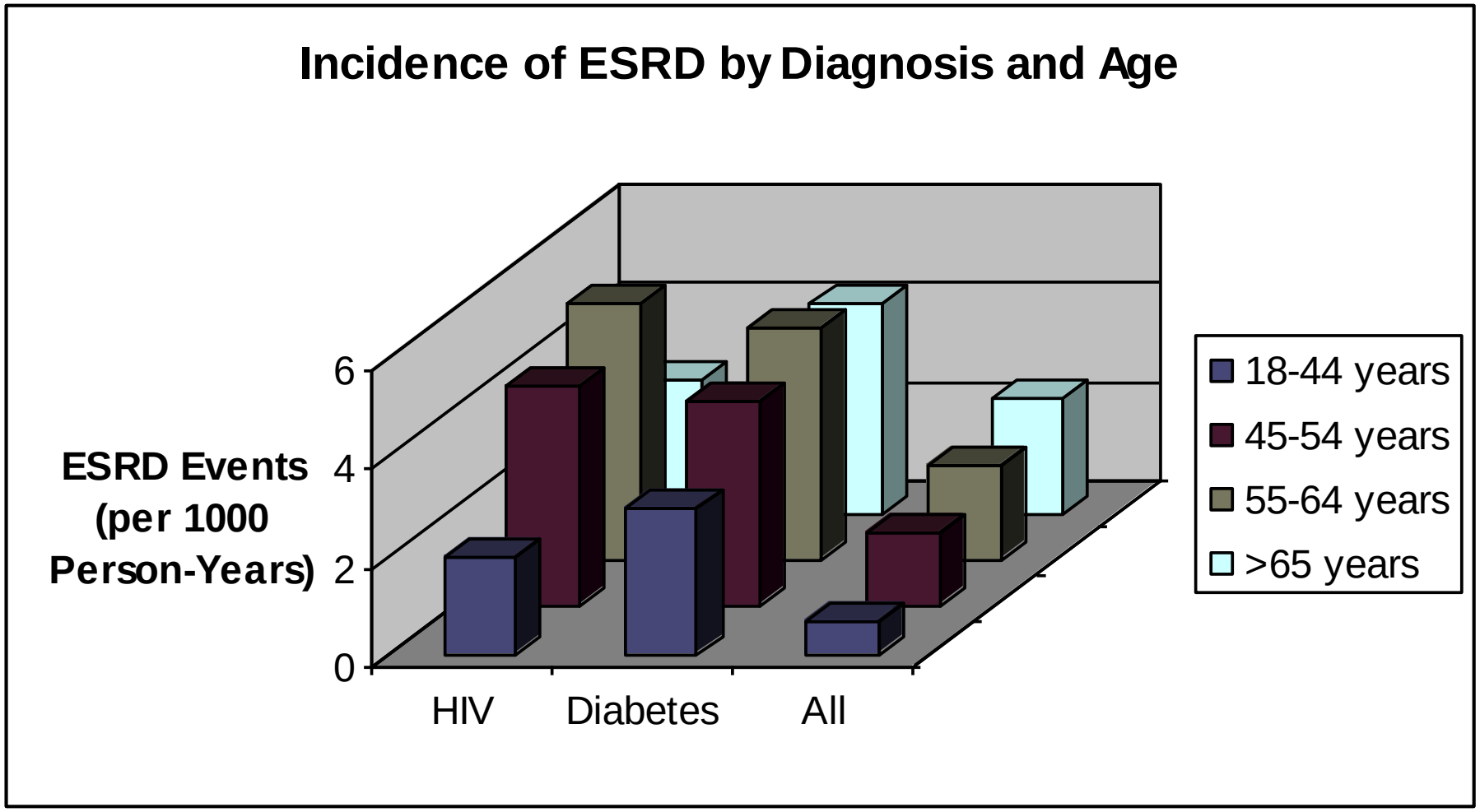


Figure 1. Peripheral CD4⁺ cell counts in patients who had maintained viral suppression for 10 continuous years. Only those who had a pretherapy CD4⁺ cell count <200 cells/mm³ are shown ($n = 48$). A significant subset of individuals appeared to have their CD4⁺ cell counts plateau below normal levels (defined here as 500 cells/mm³).

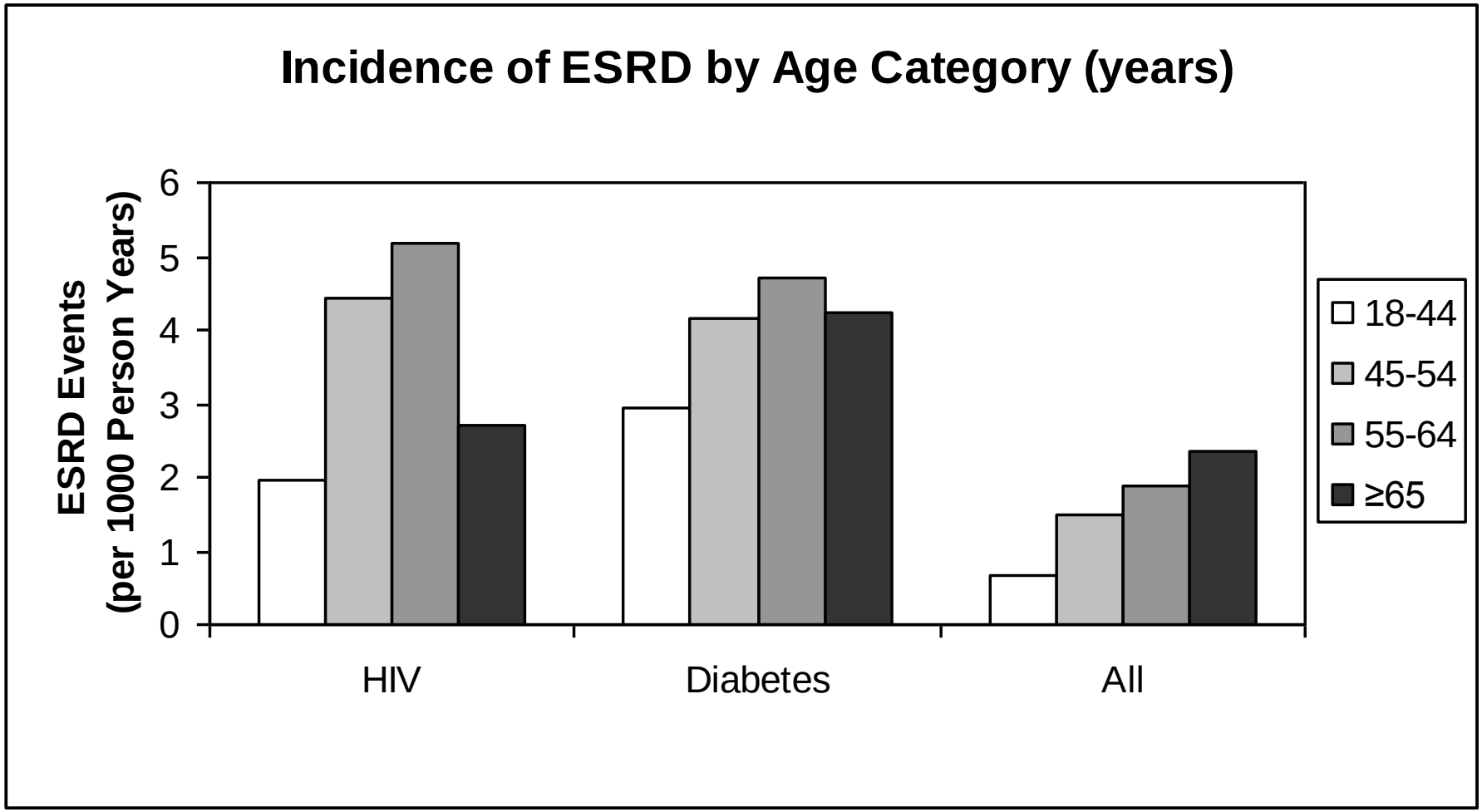
Figures: Reconsider if...

1. Light on data
2. Pie chart
3. 3-D graph

3-D Graphs


Incidence of ESRD by Diagnosis and Age
ESRD Events (per 1000 Person-Years)
0
2
4
6
HIV
Diabetes
All
18-44 years
45-54 years
55-64 years
>65 years

Without 3-D


Incidence of ESRD by Age Category (years)
ESRD Events
(per 1000 Person Years)
0
1
2
3
4
5
6
HIV
Diabetes
All
18-44
45-54
55-64
≥65

Figures - Reconsider if…

1. Light on data
2. Pie charts
3. 3-D graphs
4. Lines that cross
5. Too complex (takes 5 minutes to orient yourself to the figure)

Agenda

1. Figure basics
   - Types of figures
   - Why make a figure?
   - Elements of a figure
2. **Steps in making a figure with Excel**
3. Steps in making a figure with Stata

Steps in making an Excel figure

- First, ask yourself:

  "What is the purpose of the figure?"
- Sketch the Figure, with title
  - Try several versions
  - Point should be clear at a glance
  - Requires some artistic vision…

Steps in making an Excel figure

- Make a dummy TABLE
  - Contains the data for the figure

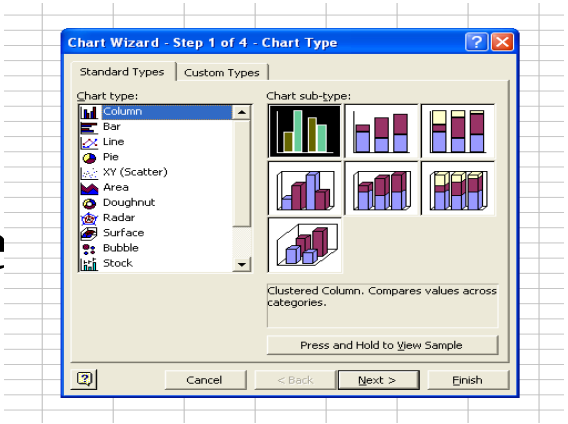
Steps in making an Excel figure

- Make a dummy TABLE
  - Contains the data for the figure
- Write a .do file to fill in the table

Steps in making an Excel figure

- Make a dummy TABLE
  - Contains the data for the figure
- Write a .do file to fill in the table
- Copy and paste from log file into the Table

Steps in making an Excel figure



- Make a dummy TABLE
  - Contains the data for the figure
  - Doesn't have to look nice
- Write a .do file to fill in the table
- Copy and paste from log file into the Table
- Use the Chart Wizard to create the Figure

Steps in making an Excel figure

- Make a dummy TABLE
  - Contains the data for the figure
  - Doesn't have to look nice
- Write a .do file to fill in the table
- Copy and paste from the log file or the results window into the Table
- Use the Chart Wizard to create the Figure
- *Format*, *format*, *format* until it looks nice

Excel Demonstration…

Pay attention to…

- Formatting
  - Keep it simple.
  - The time-consuming part of making a figure is usually related to formatting.

Pay attention to…

- Labeling
  - Your figure should be understandable by itself (without reading the manuscript).
  - All axes should be labeled.
  - Include important p-values

Pay attention to…

- The Figure Legend
  - Title, explanations, p-values, etc
  - Separate section in manuscript or at bottom of page – depends on journal

Stata commands intro

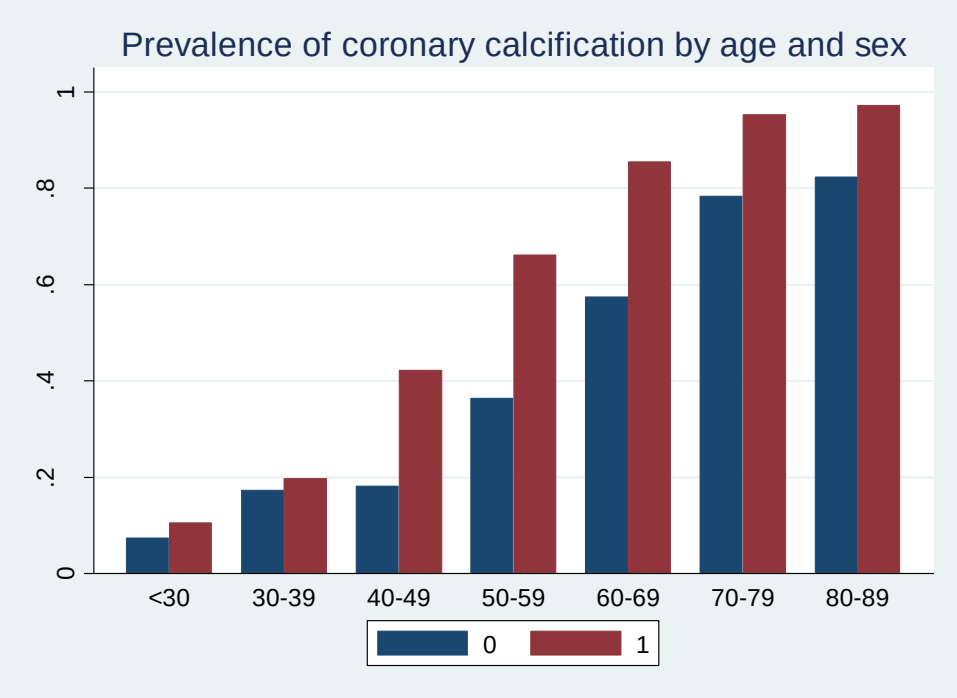
- Pie charts
  - `graph pie, over(`*`catvar`*`)`
- Bar graphs
  - `graph bar (mean)` *`yvar1 yvar2`*`, over(`*`catvar1`*`) over(`*`catvar2)`* `asyvars`
- Box plots
  - `graph box` *`contvar1 contvar2`*`, over(`*`catvar1`*`) over(`*`catvar2`*`)`
- Scatter plots
  - `twoway (`*`graphtype yvar xvar`*`) (`*`graphtype yvar xvar`*`)`
  - `scatter, line, connect, lowess, lfit, qfit,` etc

Steps in making a Stata figure

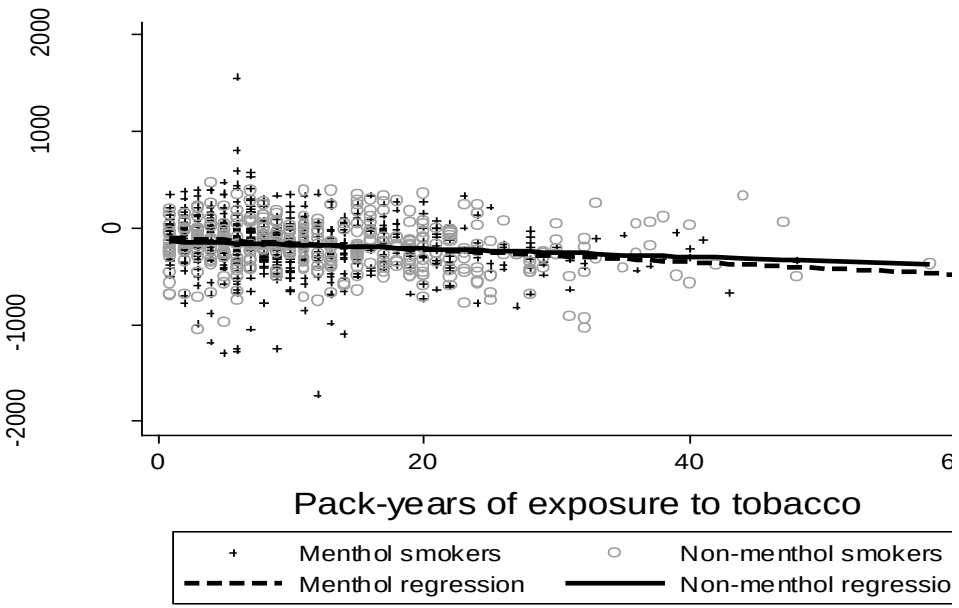
1. Sketch the figure
2. Write a do file
   - Compose the command using the dialog box
     - Get the syntax down, multiple iterations
   - Transfer to your do file and edit
3. Produce graph and edit more?
   - Graph editor function
   - May need additional Stata commands for p-values, figure legend, etc

Stata Demo

- Pie charts
- Bar charts
- Twoway plots
- Formatting options




```
graph bar (mean) anycac ///
        , over(male) ///
          over(agecat) ///
          asyvars ///
          ytitle(Prevalence of coronary calcification) ///
          title("Prevalence of coronary calcification by age and sex")
```




```
twoway (scatter dfev1 cumpy10 if menthol1==1, msymbol(plus) msize(small) mcolor(black)) ///
       (scatter dfev1 cumpy10 if menthol1==0, msymbol(circle_hollow)) ///
       (line m cumpy10 if menthol1==1, sort clcolor(black) clpat(dash) clwidth(thick)) ///
       (line nm cumpy10 if menthol1==0, sort clcolor(black) clpat(solid) clwidth(thick)) ///
       , ytitle(Change in FEV1 (milliliters), size(large)) yscale(titlegap(5)) ///
         xtitle(Pack-years of exposure to tobacco, size(large)) ///
         xscale(titlegap(3)) ///
         legend(order(1 "Menthol smokers" 2 "Non-menthol smokers" 3 "Menthol regression" ///
                      4 "Non-menthol regression")) ///
         scheme(s1mono) ///
         graphregion(fcolor(none) lcolor(none) ifcolor(none) ilcolor(none)) ///
         plotregion(fcolor(none) lcolor(none) ifcolor(none) ilcolor(none))
```

Stata vs. Excel for Figures

- Excel
  - Flexible and intuitive point-and-click figures
    - Easy to create and modify
    - Flexible, more options, error bars, adjusted estimates, good for bar graphs, etc
  - But…
    - Requires an extra step – copy/pasting to Excel
    - Harder to reproduce
    - Much harder to do scatter plots

Stata vs. Excel for Figures

- Stata
  - Can create very customizable figures using 1 complex Stata command
    - Easy to recreate – simple do file
    - No error
    - Scatter plots are MUCH easier with Stata
  - But…
    - Harder to create the first time? - no point and click
    - A little less flexible?
    - Difficult to format: Graphic Editor helps address this

Summary tips for figures

- Only use a Figure if:
  - There is an important message to convey
  - The visual will be more compelling and clear
- Try using both Stata and Excel
  - Stata will be harder at first, but often worth it
  - Browner book, Stata book both helpful
- Document, label, and be creative

Today's Lab

- You will create a pie chart, a box plot, and a scatter plot using stata.
- The focus will be on bringing the figure to publication grade.