



Introduction to Stata



Structure

Split into sections

20-30 min per section

Time for questions between
each section

Hopefully, less boring

Purpose

- ◆ By the end of this webinar you should be able to:
 - ◆ Understand the basics of how Stata works and looks like
 - ◆ Do basic data management tasks (import, export, create variables)
 - ◆ Understand how to perform basic statistical tests (t-test, chi square, ANOVA)
 - ◆ Run different types of regression (linear, binary, interaction)

Technical Terms

- ◆ IV: independent variable (explanatory variable)
- ◆ DV: dependent variable (outcome variable)
- ◆ Syntax: the language Stata understands
- ◆ Commands: the syntax you type into Stata to tell the software to run some task



Module 1

Typing v. “point and click”

Updating

.Do files

Finding help

Doing “things” in Stata

- ◆ Two methods: typing commands or using menus (“point and click”)
- ◆ Typing commands is like programming, menus are like SPSS or Excel
- ◆ Benefits and drawbacks to both

.Do files

- ◆ The place where you type commands
- ◆ Can edit, run, and add comments
- ◆ Can also run version control
- ◆ The main place where you save commands

Menus

- ◆ Are in the main portion of Stata
- ◆ Relatively easy to use
- ◆ The menu tabs are very descriptive



Hist... ▾ 🔍 ✕



Filter commands



Comma...

There are no
items to
show.

Notes:

1. Unicode is supported; see [help unicode_advice](#).
2. New update available; type `-update all-`

Checking for updates...

(contacting <http://www.stata.com>)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 ([what's new](#))
Current update level: 21 Jan 2021 ([what's new](#))

Possible actions

[Install available updates](#) (or type `-update all-`)

Click to [edit automatic update checking preferences](#)

•
•
•
•
•
•
•
•
•
•
•

Command 🔍

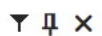


Different Windows

- ◆ Main Stata: Results Window



History



Filter commands here

Command _rc

1 swd C:\Users\faela\Docu... 199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

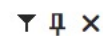
Install available updates (or type -update all-)

Click to [edit automatic update checking preferences](#)

.
. .
. .
. .
. .
. .
. .

Command

Variables



Filter variables here

Name Label

There are no items to show.

Properties



< >

Variables

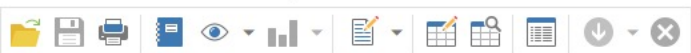
Name	
Label	
Type	
Format	
Value label	
Notes	

Data

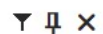
Frame	default
Filename	
Label	

Different Windows

- ◆ Main Stata: Results Window
- ◆ Command window (different than .Do files)



History



Filter commands here

#	Command	_rc
1	swd C:\Users\faela\Docu...	199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

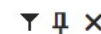
Install available updates (or type -update all-)

Click to [edit automatic update checking preferences](#)

.
. .
. .
. .
. .
. .
. .

Command

Variables



Filter variables here

Name	Label
------	-------

There are no items to show.

Properties



Lock < >

Variables

Name	
Label	
Type	
Format	
Value label	
Notes	

Data

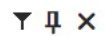
Frame	default
Filename	
Label	

Different Windows

- ◆ Main Stata: Results Window
- ◆ Command window (different than .Do files)
- ◆ Variables Window



History



Filter commands here

Command _rc

1 swd C:\Users\faela\Docu... 199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

Install available updates (or type -update all-)

Click to [edit automatic update checking preferences](#)

•
•
•
•
•
•
•
•

Command

Variables



Filter variables here

Name	Label
------	-------

There are no items to show.

Properties



Lock < >

Variables

Name	Label
Type	
Format	
Value label	
Notes	

Data

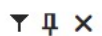
Frame	default
Filename	
Label	

Different Windows

- ◆ Main Stata: Results Window
- ◆ Command window (different than .Do files)
- ◆ Variables Window
- ◆ Review Window



History



Filter commands here

Command _rc

1 swd C:\Users\faela\Docu... 199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

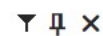
Install available updates (or type -update all-)

Click to edit automatic update checking preferences

•
•
•
•
•
•
•

Command

Variables



Filter variables here

Name Label

There are no items to show.

Properties



< >

Variables

Name	
Label	
Type	
Format	
Value label	
Notes	

Data

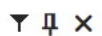
Frame	default
Filename	
Label	

Different Windows

- ◆ Main Stata: Results Window
- ◆ Command window (different than .Do files)
- ◆ Variables Window
- ◆ Review Window
- ◆ Properties Window



History



Filter commands here

Command _rc

1 swd C:\Users\faela\Docu... 199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

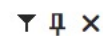
Install available updates (or type -update all-)

Click to edit automatic update checking preferences

•
•
•
•
•
•
•
•

Command

Variables



Filter variables here

Name Label

There are no items to show.

Properties



Lock < >

Variables

Name	
Label	
Type	
Format	
Value label	
Notes	

Data

Frame	default
Filename	
Label	

Let's take a look at Stata!

Module 2

Working with Data in Stata

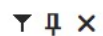
- ◊ Working Directory
- ◊ Opening Datasets
- ◊ Entering Data

Setting the Working Directory

- ◆ Where Stata looks for files
- ◆ Where Stata saves files
- ◆ `cd "E:\"`
- ◆ `pwd`



History



Filter commands here

#	Command	_rc
1	swd C:\Users\faela\Docu...	199

Checking for updates...
(contacting http://www.stata.com)

Update status

Last check for updates: 08 Oct 2021
New update available: 08 Jul 2021 (what's new)
Current update level: 21 Jan 2021 (what's new)

Possible actions

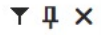
Install available updates (or type -update all-)

Click to [edit automatic update checking preferences](#)

•
•
•
•
•
•
•
•

Command

Variables



Filter variables here

Name	Label
------	-------

There are no items to show.

Properties



Lock < >

Variables

Name	
Label	
Type	
Format	
Value label	
Notes	

Data

Frame	default
Filename	
Label	

Calling in Data

- ◆ Stata looks in the working directory
- ◆ Can import data from other sources, you just have to specify the file path

```
use Module_2_Data.csv
```

```
import delimited "C:\Users\faela\Downloads\Module_2_Data.csv"
```


Entering Data Manually

◆ Rules of Thumb

- ◆ Web Survey tools provide you with survey responses in Excel or SPSS formats which you can then import into Stata
- ◆ You can also enter data directly
- ◆ Each row of data usually represents a survey respondent
- ◆ Each column of data usually represents a variable or survey question
- ◆ Variable names should only include letters or underscores _
 - ◆ Should start with a letter
 - ◆ No blank spaces
 - ◆ Each variable name should be unique
 - ◆ Keep variable names short
 - ◆ Stata is case sensitive

Entering Data cont.

- ◆ Be sure to give each respondent a unique ID
- ◆ Enter as many numeric responses as possible
 - ◆ For example, gender should be 1,0 instead of Male, Female
- ◆ Include raw data, un-summarized
- ◆ Only include one value per cell. Don't enter data like 120/80 for bp. Enter systolic bp as one variable and diastolic bp as another variable
- ◆ If you have missing values, you can leave them blank
- ◆ Date variables are easily computed if entered as 3 separate variables

Module 3

Data Management in Stata

Understanding and Creating Variables

- ◊ Variable types
- ◊ Creating categorical and dummy variables
- ◊ Dealing with missing values*

Types of Variables

- ◇ Categorical
- ◇ Ordinal
- ◇ Continuous/Ratio

Variable Types and Storage

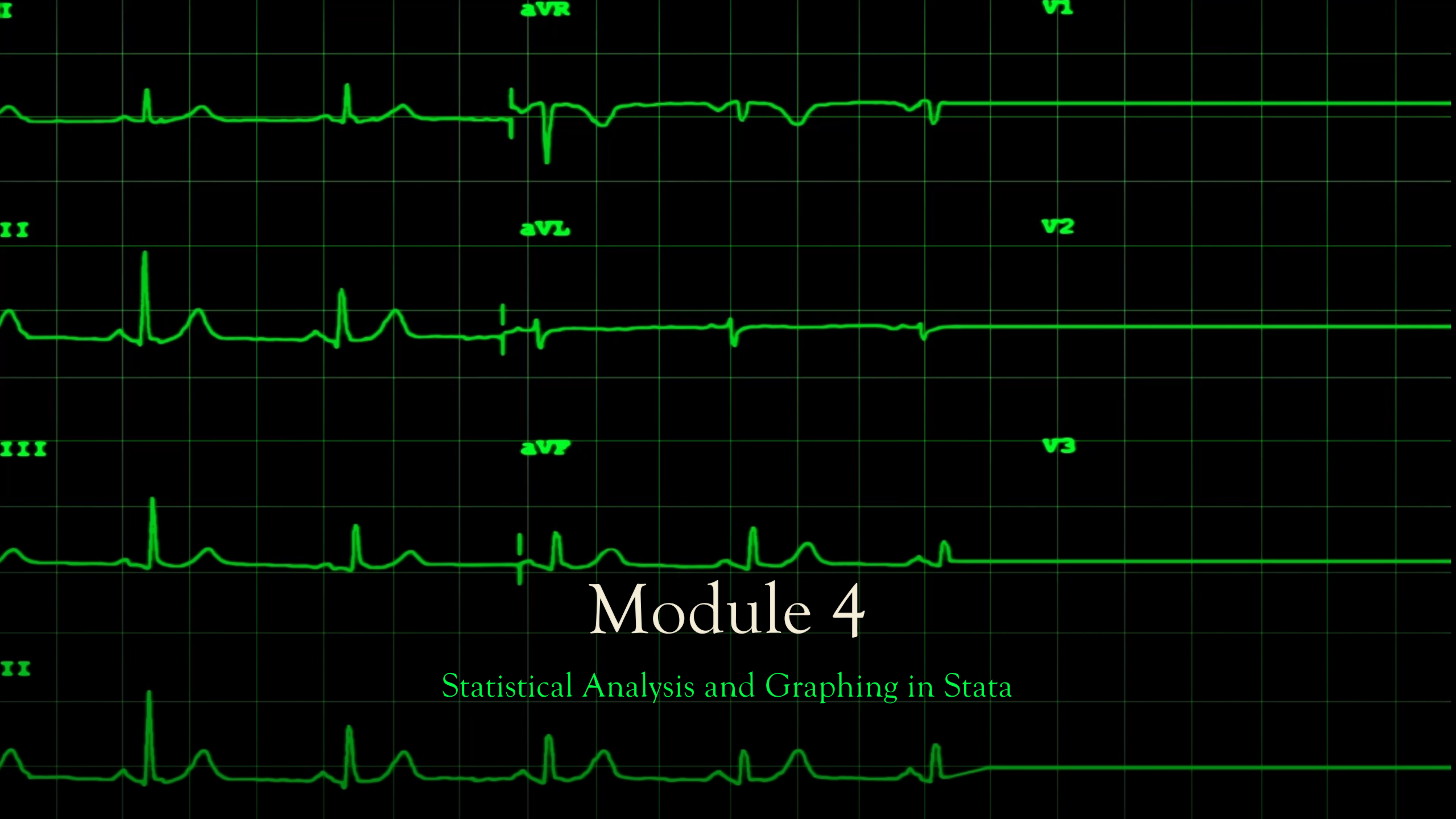
- ◇ String, Numeric
- ◇ String stored as str# (str4)
- ◇ Numeric stored as byte, int, long, float, double

Label Values of Categorical Variables

- ◆ `label define label_name value1 "label1" value2 "label2" ...`
- ◆ `label values varname label_name`

Create New Variables

- ◆ Data – Create or change variables – create new variable
- ◆ Enter the name of the variable and enter the values in the contents window
- ◆ Usually easier to code manually



Module 4

Statistical Analysis and Graphing in Stata

Tests and Regression

- ◆ Tests to run/how to run the tests
- ◆ Graphs to run/how to create the graphs
- ◆ Linear regression

Correlation v. Causation

- ◆ We can find relationships, difficult to say whether the relationship is causal

Parametric v. Non-Parametric Tests

01

Parametric assumes the sample being analyzed is from an underlying normal distribution

02

Non-Parametric tests are used when this assumption is violated

03

For the purpose of this course we will assume data are normal

Parametric v. Non-Parametric

Parametric Test	Non-Parametric Test
Pearson Correlation Coefficient	Spearman Correlation Coefficient
Independent Samples t-test	Mann-Whitney/Wilcoxon Rank Sum
Paired Samples t-test	Wilcoxon Signed Rank Test
One Way ANOVA	Kruskal Wallis Test
Repeated Measures ANOVA	Friedman's Test
Chi-square Test	Fisher's Exact Test

Parametric Tests

Test	Type of Data
Pearson Correlation Coefficient	DV= Continuous, IV=Continuous
Independent Samples t-test	DV= Continuous, IV= Categorical (2 groups)
Paired Samples t-test	DV= Continuous, repeated, IV= Categorical (2 groups)
One-way ANOVA	DV= Continuous, IV= Categorical (2 or more groups)
Repeated Measures ANOVA	DV= Continuous, repeated, IV= Categorical (2 or more groups)
Chi-square Test	DV= Categorical, IV= Categorical

Graphing

Number of Variables	Variable Type(s)	Graph(s)
1	Categorical	Bar chart, pie chart
1	Continuous	Histogram, box plot, qq plot
2	Both categorical	Clustered or paneled bar chart
2	One categorical, one continuous	Side by side box plots
2	Both continuous	Scatterplot
3	Two continuous, one categorical	Scatterplot with markers