

# Basic Commands for Data Cleaning

## Data Management

<b>use, sysuse, import excel, insheet using, infix using</b>	Load/Import Data
<b>save, export excel, outsheet</b>	Save/Export Data
<b>[, clear], [, replace]</b>	Most relevant options
<b>mkdir, [shell] rmdir [/s], rm</b>	Create/Remove folder & files
<b>order, move</b>	Organize vars.
<b>merge n:n keyvars using</b>	Merge datasets
<b>append using [, force]</b>	Append datasets

## Loops and Conditional Execution

<b>foreach x in/of list..{} , forv x=..{} , while x..{} if{} , else if{} , else {}</b>	Loops based on local macros
<b>command... if x==...</b>	General condition
<b>command... if x==...</b>	Row condition

## Other

<b>label var/data name "#"</b>	Descriptive labels
<b>label define.., label val..</b>	Value labels
<b>di operation/value</b>	Display on screen

## Data Manipulation

<b>gen, egen, replace, drop, keep, ren</b>	Generate, replace, drop, rename variable(s)
<b>sort</b>	Sort observations
<b>sample</b>	Keep random subset
<b>encode/decode</b>	String to/from encoded
<b>dstring, tostring</b>	String to/from numeric
<b>bysort groupvar:</b>	Operate within groups
<b>ustrpos(), substr()</b>	Parts of strings
<b>date(), mdy(), month()</b>	Time-date formatting
<b>format tvar %td</b>	Date display format
<b>runiform(), runiformint()</b>	Random numeric variables (set seed!)
<b>compress</b>	Memory optimization

## Basic Data Inspection

<b>browse, list</b>	See/Display data table
<b>describe</b>	Dataset summary
<b>summarize [varlist]</b>	Variable(s) sumstats
<b>table, tabulate</b>	Summary stats tables
<b>count if...</b>	Count complying obs.