# Stata commands

## Basic command structure

▶ You can basically express every Stata command as a variant of the following generic structure:

[*prefix*:] **command** [*varlist*] [*if*] [*in*] [*weight*] [**using** "`filename`"][, *options*]

▶ The only mandatory part is the **command** itself.
▶ All terms in square brackets are optional. Don't type the brackets.
  - The *prefix* **by** can slice your data into different blocks and perform analysis within each block.
  - By specifying a *varlist* you apply to command to selected variables only.
  - Using *if*- and *in*-conditions allows you restrict the command to specific observations/rows of your data set.
  - **using** passes a `path` to your command. This is mainly used when using different data then currently stored.
  - *options* tailor your command to your specific needs.

# Example: `summarize`

- Having loaded the automobile data, we can pass the command `summarize` to the command bar.
- You will get summary statistics for all variables in the data set.
- You can restrict the command to specific variables by passing a list of variables, e.g., `mpg` and `price`:

  `summarize mpg price`
- To pass all variables to a command, you may use an **asterisk \*** as a wildcard or use _all:

  `summarize *`

  `summarize _all`

```
. summarize

    Variable         Obs        Mean    Std. dev.

        make           0
       price          74     6165,257    2949,496
         mpg          74     21,2973    5,785503
       rep78          69    3,405797    ,9899323
    headroom          74    2,993243    ,8459948

       trunk          74    13,75676    4,277404
      weight          74    3019,459    777,1936
      length          74    187,9324    22,26634
        turn          74    39,64865    4,399354
displacement         74     197,2973    91,83722

   gear_ratio         74    3,014865    ,4562871
      foreign         74    ,2972973    ,4601885
```
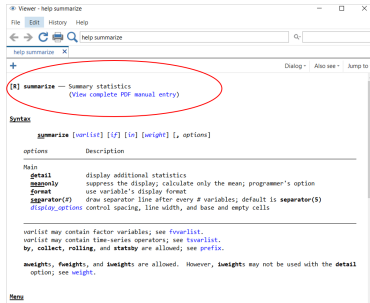
# Getting help

▶ Stata has so many commands that it is impossible to know them all.

▶ Also, even if you know a command you might want to get to know its options.

▶ Thus, Stata **Help Files** will sooner or later become your best friend.

▶ To open them type `help` (OR just `h` for short) in front of a command that you need help with.

▶ To make things easier, all help files look similar. Let's take a look at the help file for command `summarize`.
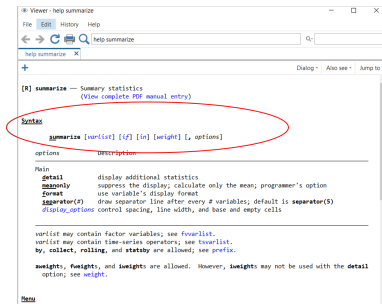
```
help summarize
```

# Help File: Title Section

- In the file, you will see the command name and a brief description at the very top.
- There is also a link to the **.pdf** of the Stata manual entry for the command in question - in this case `summarize`.
- The manual entries include details about methods and formulas used for estimation commands, and thoroughly explained examples.



Screenshots taken with permission of the publisher. *Stata Statistical Software: Release 19* (College Station, TX: Stat-aCorp, 2025).
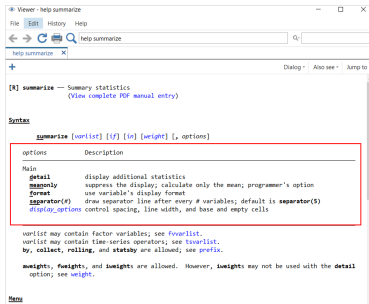
# Help File: Syntax Section

▶ The next section called **Syntax** provides various uses of the command and how to specify them.

▶ Words in **bold** font are required.

▶ The underlined part of the command is the minimal abbreviation required for Stata to understand it. For instance, we can use `su` instead of `summarize`.

▶ Words in *italic* font are to be substituted by the user. For example, *varlist* is a list of one or more variables that the user can specify.

▶ A comma `,` is almost always needed to initiate the list of `options`.



Screenshots taken with permission of the publisher. *Stata Statistical Software: Release 19* (College Station, TX: StataCorp, 2025).

# Help File: Options Section

- ▶ As the name indicates, **Options** are not required for Stata to process a command.
- ▶ Most Stata commands come with a variety of options that alter how the commands process or output the data.
- ▶ Options can also be abbreviated by typing the <u>underlined</u> portion.
- ! Always read this section carefully. Many options come with **defaults** that might influence your analysis. Only by understanding and correctly specifying the options, you will get Stata to do what you really want.



Screenshots taken with permission of the publisher. *Stata Statistical Software: Release 19* (College Station, TX: StataCorp, 2025).

## Help File: The Rest

- ▶ Below **Options** are further **Descriptions**, **Examples** of using the command, including video examples! (occasionally)
- ▶ Click on "Also see" to open help files of related commands.
- ▶ Some commands store results. `summarize` not only calculates means for you, but it also stores them for a while - until the next command uses the same storage device.
- ▶ A list of storage devices can be found in the **Help file** for usage.



Screenshots taken with permission of the publisher. *Stata Statistical Software: Release 19* (College Station, TX: StataCorp, 2025).
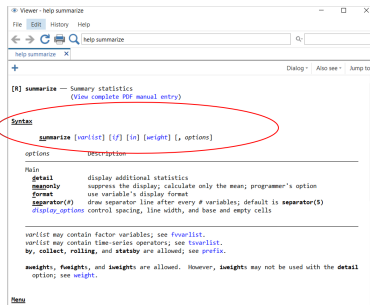
# What about `summarize`?

- ▶ Let's take a closer look at `summarize`. The **Help file** will give us all the information we need.

- ▶ The command will calculate and display summary statistics of variables for us.

- ▶ We can specify one or more variables for the command.

- ▶ By choosing the option `detail` we will get additional statistics. Which one? Scroll down in the help file to find out!

- ▶ Let's try it out: Let Stata calculate the mean of the variable *weight*. Also take a look at the median weight of the sample.

  ```
  summarize weight

  sum weight, d
  ```



Screenshots taken with permission of the publisher. Stata Statistical Software: Release 19 (College Station, TX: Stat-aCorp, 2025).

# Try it out

▶ **help** tabulate

▶ **help** generate

▶ **help** regress