

## Introduction to Stata 2017-18

### 04. How to Work with Dates

1. Preliminary – Excel File of Sample Dates .....	2
2. From <i>Excel Dates</i> to <i>Stata Dates</i> .....	4
3. From <i>Excel Strings</i> to <i>Stata Dates</i> .....	5

**Before you begin, RECALL:** Stata stores dates as integer with the integer value = # elapsed days since January 1, 1960. To appreciate this, consider the following code fragment:

```
. * Create a new variable called day0 and set it equal to January 1, 1960
. * Create a new variable called day1 and set it equal to January 1, 1961 (yes, one year later)
. generate day0=mdy(1,1,1960)
. generate day1=mdy(1,1,1961)

. * Print out day0 and day1 with no formatting gives correct #'s of elapsed days since January 1, 1960
. list day0 day1
```

```
1. |-----+-----|
   | day0   day1 |
   |-----+-----|
   |      0    366 |
   |-----+-----|
```

```
. * In general however, we want formatted dates so that we understand them!
. * There are lots of choices for formatting dates. Check the help for others
.
. format %td day0
. format %td day1
. list day0 day1
```

```
1. |-----+-----+-----|
   |      day0      day1 |
   |-----+-----+-----|
   | 01jan1960  01jan1961 |
   |-----+-----+-----|
```

### 1. Preliminary – Excel File of Sample Dates

To begin: Create the following simple excel file with 4 variables: 2 formatted as date (dob\_d and dsurvey\_d) and 2 formatted as text/string (dob\_s and dsurvey\_s)

A	B	C	D
dob_d	dsurvey_d	dob_s	dsurvey_s
6/9/1956	12/15/1999	6/9/1956	12/15/1999
10/1/1953	9/15/2017	10/1/1953	9/15/2017
10/14/2002	10/4/2017	10/14/2002	10/4/2017

**Key:** To create this file, I did the following:

1. Launch excel
2. Enter variable names in row 1
3. Format initially empty columns “A” and “B” as **date**. To do this:
  - a) highlight entire column, and then
  - b) from the menu bar: **FORMAT > CELLS > DATE**
4. Format initially empty columns “C” and “D” as **text**. To do this:
  - a) highlight entire column, and then
  - b) from the menu bar: **FORMAT > CELLS > TEXT**
5. Enter data
6. Finally, save. Here, I saved my excel file to my desktop, naming it **dates\_practice.xlsx**

### Next, launch Stata and Import

- . \* --- Import Excel Data from menu bar: **FILE > IMPORT > EXCEL SPREADSHEET** ---\*
- . \* --- **DON'T FORGET!**: Check box for “Import first row as variable names”

The screenshot shows the Stata 'Import Excel' dialog box. The 'Excel file' field contains the path '/Users/cbigelow/Desktop/dates\_practice.xlsx'. The 'Worksheet' dropdown is set to 'Sheet1 A1:D4'. The 'Cell range' field is 'A1:D4'. The 'Import first row as variable names' checkbox is checked. The 'Variable case' dropdown is set to 'preserve'. Below the dialog box, a preview of the data is shown, displaying the first four rows of the Excel file, including the variable names in the first row.

Stata shows you the associated command in the command window (Yours will look slightly different, of course):

```
. import excel "/Users/cbigelow/Desktop/dates_practice.xlsx", sheet("Sheet1") firstrow
```

### Now, still in Stata - follow along ...

```
. set more off
. * First, let's save the stata data set as mydates.dta. To do this from menu bar: FILE > SAVE AS  --*
. save "/Users/cbigelow/Desktop/mydates.dta"
file /Users/cbigelow/Desktop/mydates.dta saved
```

```
. * --- command DESCRIBE to obtain information on storage type ---*
. describe dob_d dsurvey_d dob_s dsurvey_s
```

variable name	storage type	display format	value label	variable label	KEY:
dob_d	int	%td		dob_d	So far so good. dob_d is an integer variable
dsurvey_d	int	%td		dsurvey_d	Also good. dsurvey_d is an integer variable
dob_s	str10	%10s		dob_s	As expected. dob_s is a string variable
dsurvey_s	str10	%10s		dsurvey_s	Also expected. dsurvey_s is a string variable

## 2. From Excel Dates to Stata Dates

This is easy. Stata can import excel dates and then store them as Stata dates.

```
. * --- List the data using default display format (%td) ---*
. list dob_excel dsurvey_excel
```

```
+-----+
| dob_excel   dsurvey~1 |
+-----+
1. | 09jun1956   15dec1999 |
2. | 01oct1953   15sep2017 |
3. | 14oct2002   04oct2017 |
+-----+
```

```
. * ---- Want a different display format? ---*
```

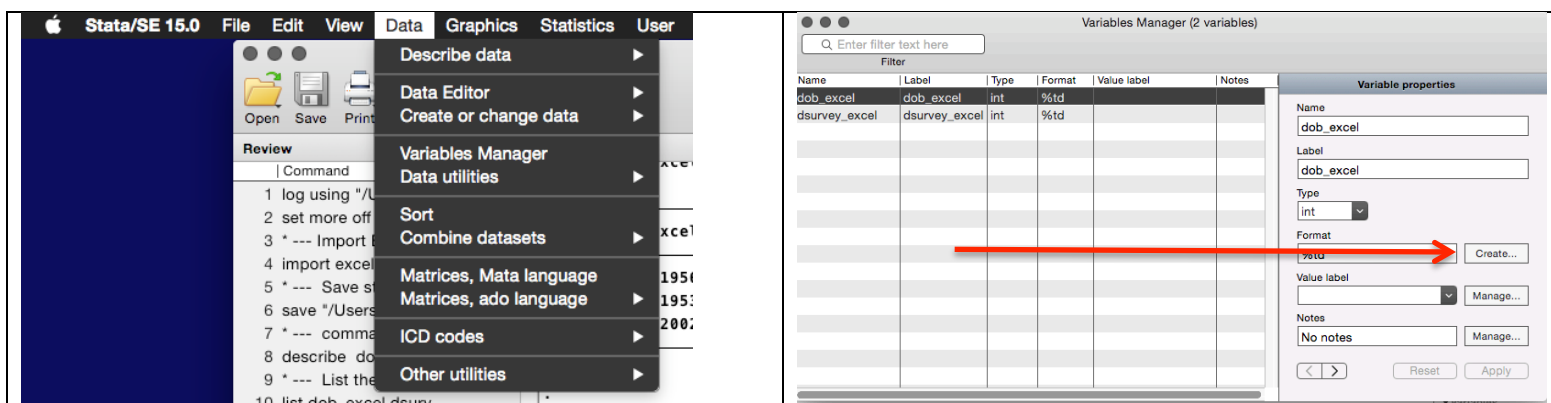
```
. * HOW TO:
```

```
. * To change the display format to something else, from the menu bar: DATA > VARIABLES MANAGER
```

```
(1) Highlight the first variable (dob_excel). Then, at right, under Format, click on CREATE
```

```
(2) From the drop down menu, be sure to select TYPE OF DATA: daily. Then click on your format selection.
```

```
REPEAT for dsurvey_excel
```



Stata shows you the associated commands (note that I chose different display formats for each variable)

```
. format %tdMonth_DD,_CCYY dob_excel
. format %tdNN/DD/CCYY dsurvey_excel
```

```
. * ---- Let's do one more check that all is well by calculating age at date of survey ---*
```

```
. generate age = (dsurvey_d - dob_d)/365.25
```

```
. list dob_d dsurvey_d age
```

```
+-----+
| dob_d       dsurvey_d   age |
+-----+
1. | 09jun1956   15dec1999   43.5154 |
2. | 01oct1953   15sep2017   63.9562 |
3. | 14oct2002   04oct2017   14.97331 |
+-----+
```

Hooray!

### 3. From *Excel Strings* to *Stata Dates*

Thank goodness. Stata has a nifty tool (“MDY”) for converting strings (strings that are dates, that is) to dates. Happily, it works with all kinds of strings that are dates ...

```
. * KEY
. * generate NEWDATEVAR=date(SOURCESTRINGVAR, "MDY")
. generate dobnew=date(dob_s, "MDY")
. generate dsurveynew=date(dsurvey_s, "MDY")

. describe dobnew dsurveynew
```

variable name	storage type	display format	value label	variable label
dobnew	float	%9.0g		
dsurveynew	float	%9.0g		

```
. format %td dobnew
. format %td dsurveynew
. generate agenew=(dsurveynew-dobnew)/365.25
. list dobnew dsurveynew agenew
```

	dobnew	dsurvey~w	agenew
1.	09jun1956	15dec1999	43.5154
2.	01oct1953	15sep2017	63.9562
3.	14oct2002	04oct2017	14.97331

Huzzah.