

trivia

The Julian and Gregorian calendars

When you work with very old dates you can run into a problem with the Julian and Gregorian calendars.

In 1582 it was noticed that the seasons had drifted by 10 days because the Julian system (named after Julius Caesar who adopted it in 45 BC) had incorrectly miscalculated a year as being $365 \frac{1}{4}$ days. A year is actually slightly shorter than this.

Pope Gregory XIII decreed that, in order to put things right, the world had to lose 10 days to make up for all of those extra leap years.

The reformed Gregorian calendar also adopted a new leap year rule to keep things on track in future years.

It took nearly 200 years for everybody to get on board with the Gregorian calendar. The Catholic countries of Spain, Italy and Portugal went for it at once but England and parts of America didn't convert until September 14th 1752.

This means that in Spain the dates October 5th 1582 to October 14th 1582 never actually existed. In England it was the dates Sept 3rd 1752 to September 14th 1752.

The strangest case of all was Sweden who decided to "phase it in gradually" between 1700 and 1740 meaning that their calendar was out of step with the rest of the world for 40 years.

If you work with historical data from this era you have to be very careful indeed.

Lesson 4-2: Understand date serial numbers

Excel stores dates in a very clever way. Understanding Excel's date storage system empowers you to use date arithmetic. You can use date arithmetic to compute the difference between two dates (in days) or to shift date ranges by a given time interval.

How Excel stores dates

Dates are stored as simple numbers called *date serial numbers*. The serial number contains the number of days that have elapsed since 1st January 1900 (where 1st January 1900 is 1).

The world began in 1900

An interesting shortcoming of Excel is its inability to easily work with dates before 1900. Excel simply doesn't acknowledge that there were any dates before this time. If you work with older dates you will have to work-around this limitation.

In Excel every time is a date, and every date is a time

This one is an eye opener! We've already realized that 5th January 1900 is stored as the number 5. What would the number 5.5 mean? It would mean midday on 5th January 1900.

It is possible to format a date to show only the date, only the time, or both a time and a date.

When you enter a time into a cell without a date, the time is stored as a number less than one. Excel regards this as having the non-existent date of 00 January 1900!

When you enter a date into a cell without a time, the time is stored as midnight at the beginning of that day.

- 1 Create a new blank workbook and put the numbers 1 to 5 in cells A1:A5.
- 2 Type the formula **=A1** into cell B1 followed by the **<Enter>** key, and then AutoFill the formula to the end of the list.

AutoFill was covered in: *Lesson 2-14: Use AutoFill for text and numeric series.*

	A	B
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5

trivia

The peculiar case of the Excel date bug and Lotus 1-2-3

Here's the Gregorian leap year rule © 1582 Pope Gregory XIII:

Every year that is exactly divisible by four is a leap year, except for years that are exactly divisible by 100; the centurial years that are exactly divisible by 400 are still leap years.

This means that the year 1900 wasn't a leap year but 2000 was (causing many millennium software bugs).

The designers of Lotus 1-2-3 weren't paying enough attention to Pope Gregory's rules. Their Date() function thought that 1900 was a leap year and thus recognised the mythical date February 29th 1900.

Because Excel needed to be compatible with Lotus 1-2-3, Microsoft had to replicate the Lotus bug when they designed Excel.

Try entering 29th Feb 1900 into a worksheet and Excel will gladly accept it.

This bug has the effect of introducing a one-day error into any date arithmetic you may do that spans 29th February 1900.

- 3 Apply a date format to column A that will show a four digit year.

This was covered in: *Lesson 4-1: Format dates.*

	A	B
1	01 January 1900	1
2	02 January 1900	2

This reveals that the numbers 1 to 5 represent the dates 1-Jan-1900 to 05-Jan-1900.

- 4 Change the date format in column A so that it shows both dates and times.

Setting a custom format was covered in: *Lesson 4-1: Format dates.*

Choose the custom format: **dd mmm yyyy hh:mm**. (If it isn't shown in the list you will need to type it into the *Type* box).

Notice that when you enter a date without a time, the time is set to midnight at the beginning of that day.

	A	B
1	01 Jan 1900 00:00	1
2	02 Jan 1900 00:00	2


- 5 Change the time in cell A2 to 12:00.

Notice that the number in cell B2 has changed to 2.5 showing that times are stored by Excel as the decimal part of a number.

	A	B
1	01 Jan 1900 00:00	1
2	02 Jan 1900 12:00	2.5

- 6 Compute the number of days that occurred between 01 Jan 1900 and 01 Jan 2000.

Now that you have a good grasp of Excel's serial numbers, this task is easy.

1. Enter the two dates in cells D1 and D2, one beneath the other.
2. Click in cell D3 to make it the active cell.
3. Click: Home→Number→Comma Style. 

You'll be learning more about the comma style in: *Lesson 4-3: Format numbers using built-in number formats.*

4. Subtract one date from the other by entering the formula: **=D2-D1** into cell D3.

	D
1	01/01/1900
2	01/01/2000
3	36,525.00

You now know that 36,525 days occurred during the twentieth century (actually 36,524 due to the Lotus 1-2-3 bug – see sidebar).

- 7 Close the workbook without saving.