Visualization of interactions during the course of an epidemic

When examining the dynamics of an outbreak or event in public health, one should not be limited to constructing an epidemic curve to, since the curve does not fully inform, especially if there is an overlap of situations or an interaction between the people studied.

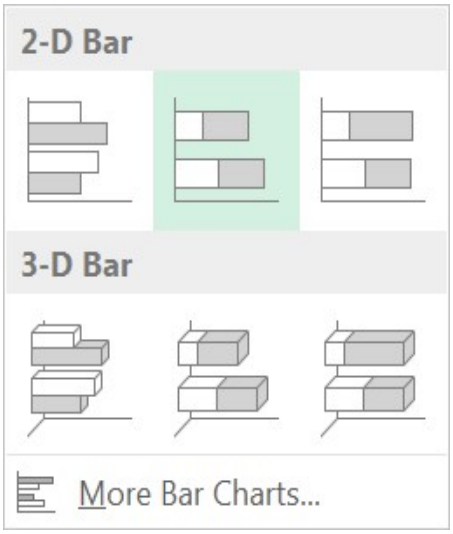
It is important to be able to display graphically when there is a common source of contagion or a critical period to identify, such as the hours or days elapsed since an exposure and its resulting effect. We are going to use as an example the information on exposure time in a dialysis unit of a hospital in which there was an HIV epidemic in 1992 in Colombia. This work was published in the Lancet simultaneously with an abridged version in the MMWR. Both documents are available on the internet.

We will get a visualization using stacked bars using Excel. First, we need to open a database called HURGV.xls (among the other attachments to the article, the labels are in Spanish- translations below).

1. We visualize the database and identify the following variables: the (ID) variable, which is the number of the patient and the dates of the start of attendance at the dialysis unit (Inicio -START) and end of attendance (FIN: END), the status of infection and the date of first diagnosis of HIV infection.

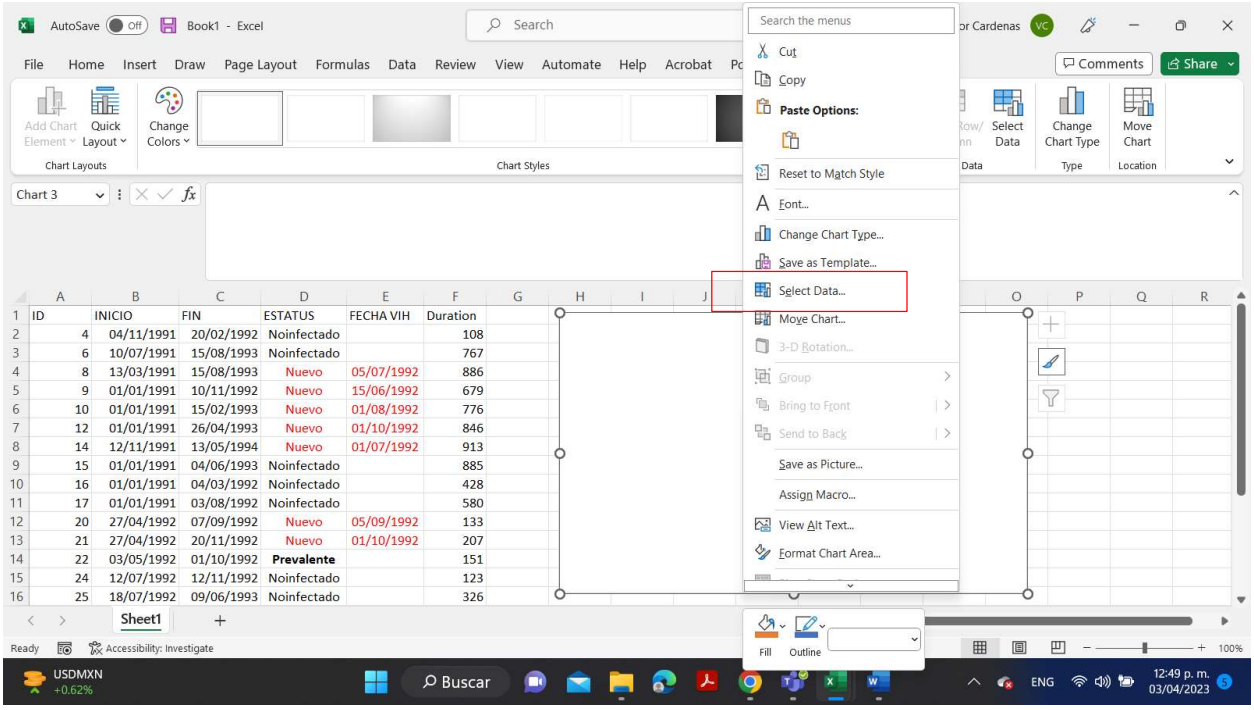
2. Let's compute a column called [Duracion: DURATION] [to calculate the period of attendance in days at the dialysis unit] using the function (*fx*) days in parentheses, which returns the number of days between two dates; It will look like this in the first row with data =DAYS(end_date; start_date) and then fill using autofill paste.

3. We insert the chart to select in anywhere on the sheet, and in this case we are choosing an stacked bar chart, which looks as follows:



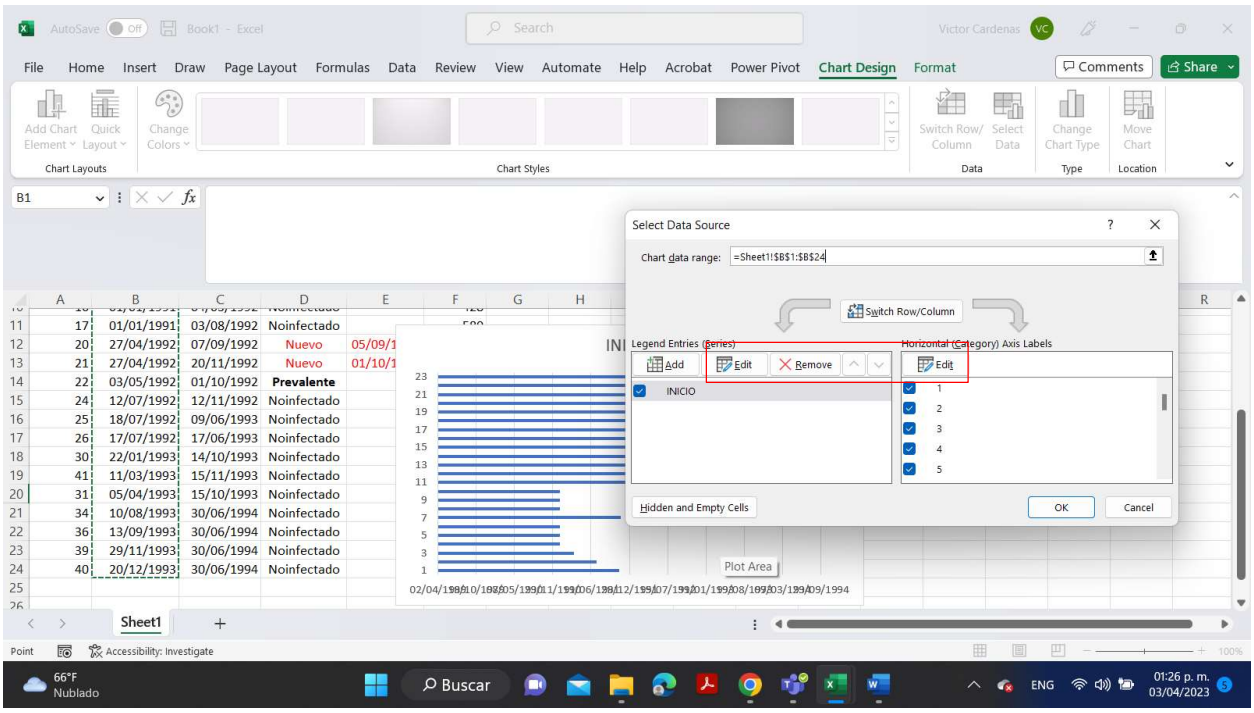
Depending on the version of Excel (2016 or earlier) it will look a little different.

4. Within the graph canvas, press the right mouse button to select the data.

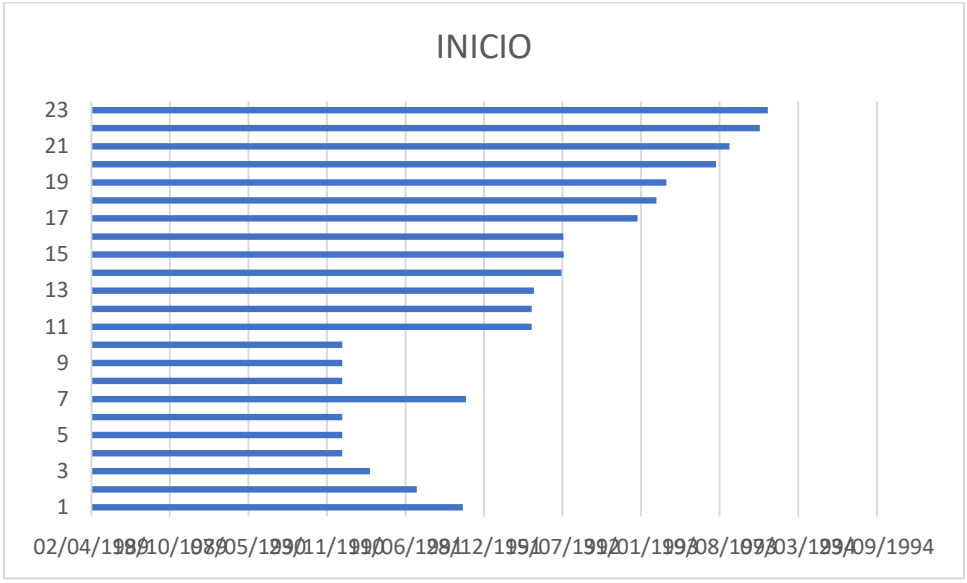


	A	B	C	D	E	F
1	ID	INICIO	FIN	ESTATUS	FECHA VIH	Duration
2	4	04/11/1991	20/02/1992	Noinfectado		108
3	6	10/07/1991	15/08/1993	Noinfectado		767
4	8	13/03/1991	15/08/1993	Nuevo	05/07/1992	886
5	9	01/01/1991	10/11/1992	Nuevo	15/06/1992	679
6	10	01/01/1991	15/02/1993	Nuevo	01/08/1992	776
7	12	01/01/1991	26/04/1993	Nuevo	01/10/1992	846
8	14	12/11/1991	13/05/1994	Nuevo	01/07/1992	913
9	15	01/01/1991	04/06/1993	Noinfectado		885
10	16	01/01/1991	04/03/1992	Noinfectado		428
11	17	01/01/1991	03/08/1992	Noinfectado		580
12	20	27/04/1992	07/09/1992	Nuevo	05/09/1992	133
13	21	27/04/1992	20/11/1992	Nuevo	01/10/1992	207
14	22	03/05/1992	01/10/1992	**Prevalente**		151
15	24	12/07/1992	12/11/1992	Noinfectado		123
16	25	18/07/1992	09/06/1993	Noinfectado		326

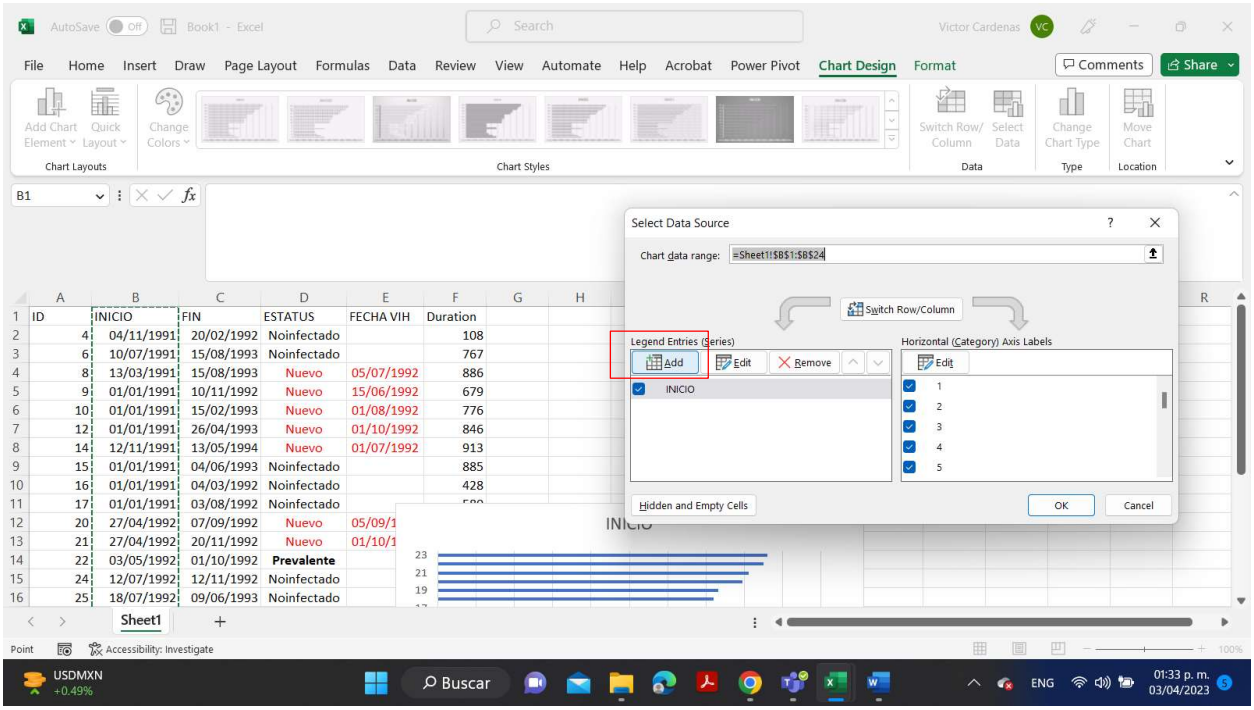
5. Select as the name of the series, the cell that has it, in this case the series is in B1 (inicio: START) of our Excel sheet called HURGV. Values range from B2 to B24.

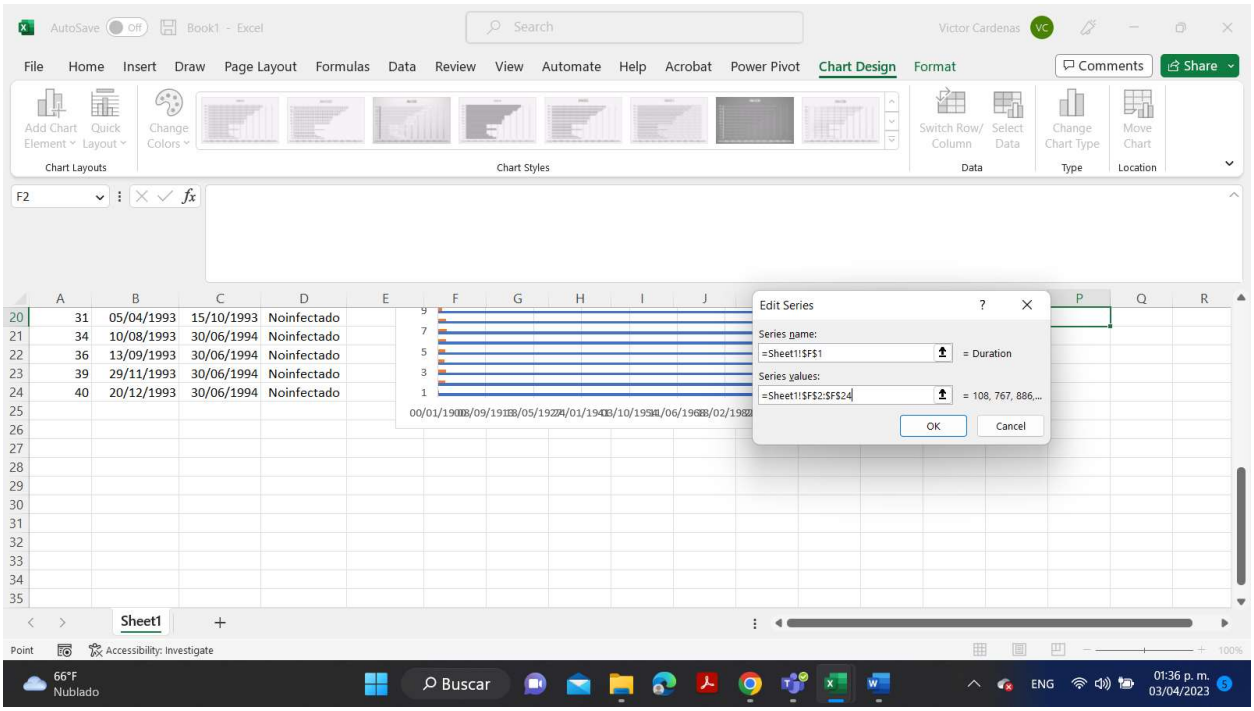


6. We see that we have filled in the start date at the end of the bars, but still, it is not what we want to do, since it is intended to display is the duration -that is the time at risk- on the time scale.

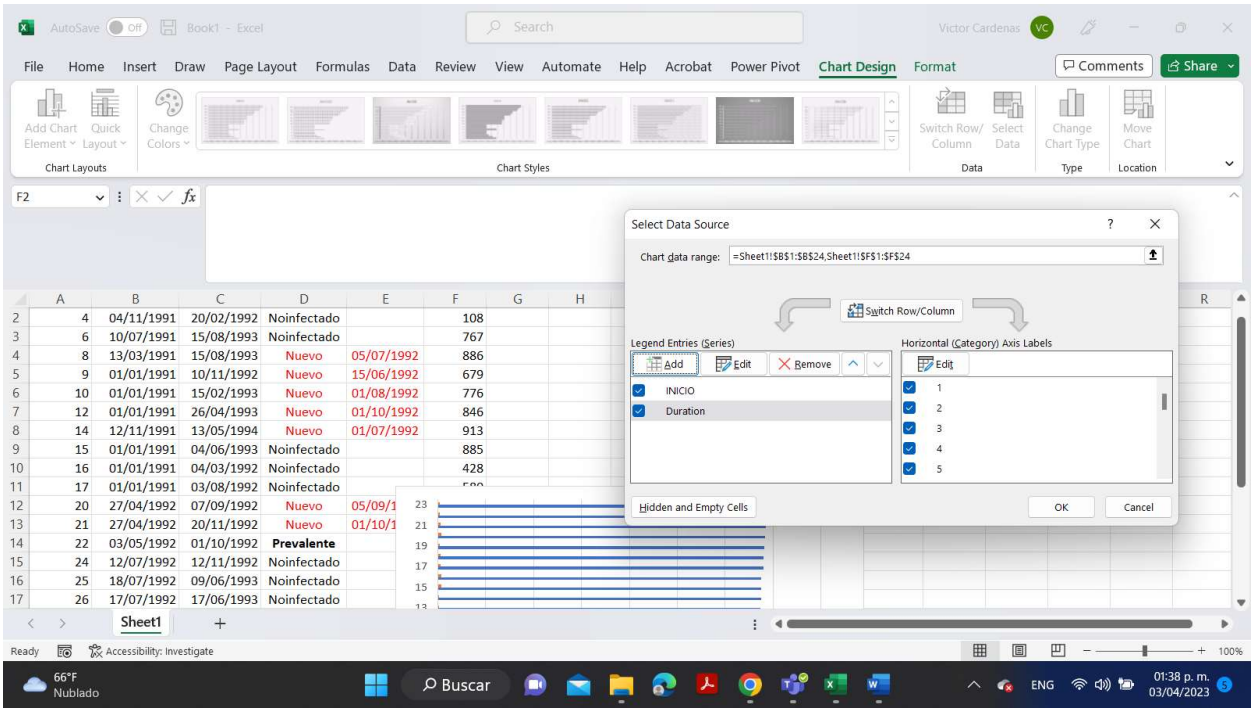


7. Therefore, we are going to add the data of the DURATION variable, for which, within the graph canvas, press the right mouse button to select the data and add the variable

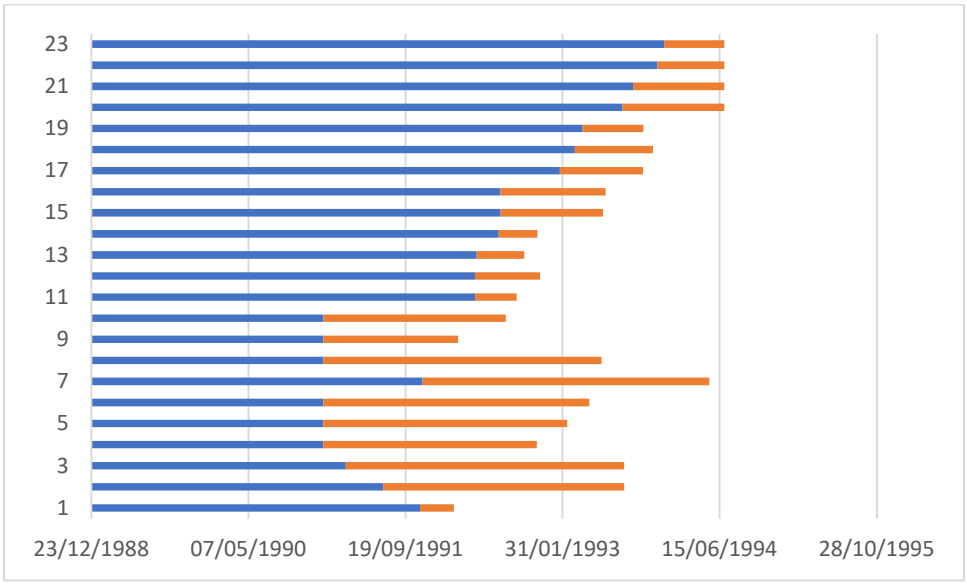




We have added as title the first cell of the F column, and the array of cells F2:F24 as data values.

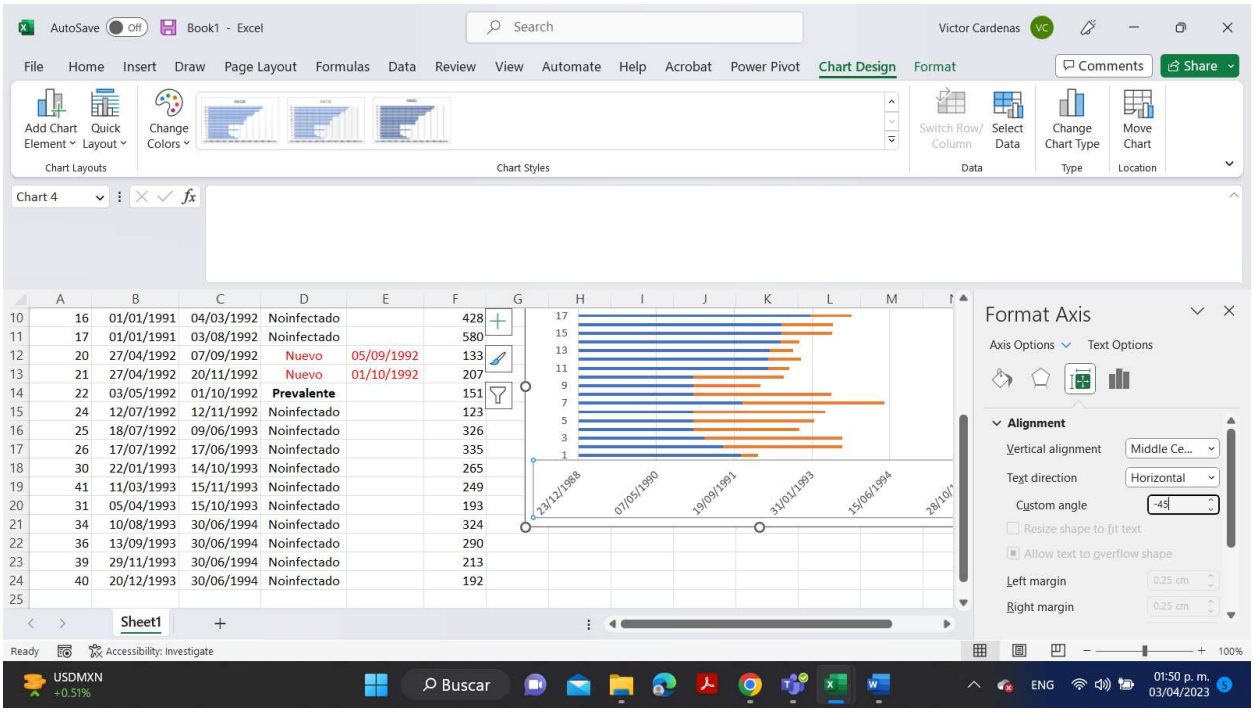


8. If the stacked bar chart is selected, the graph will look like this:

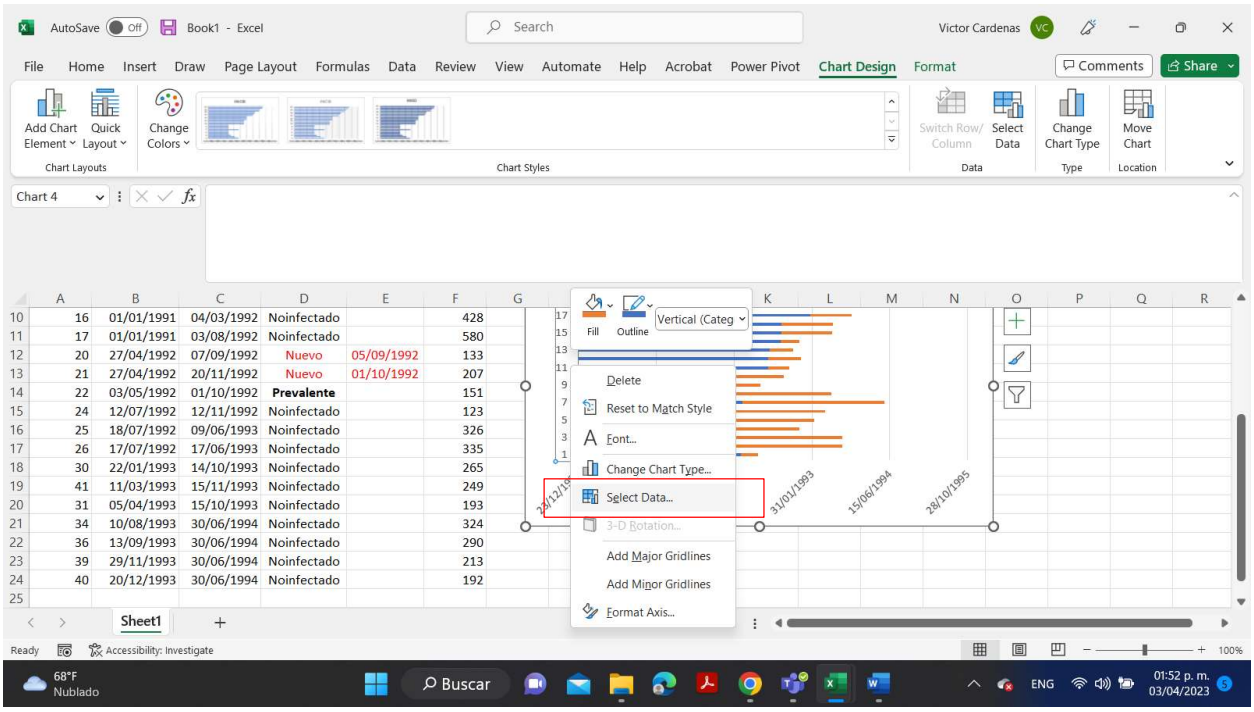


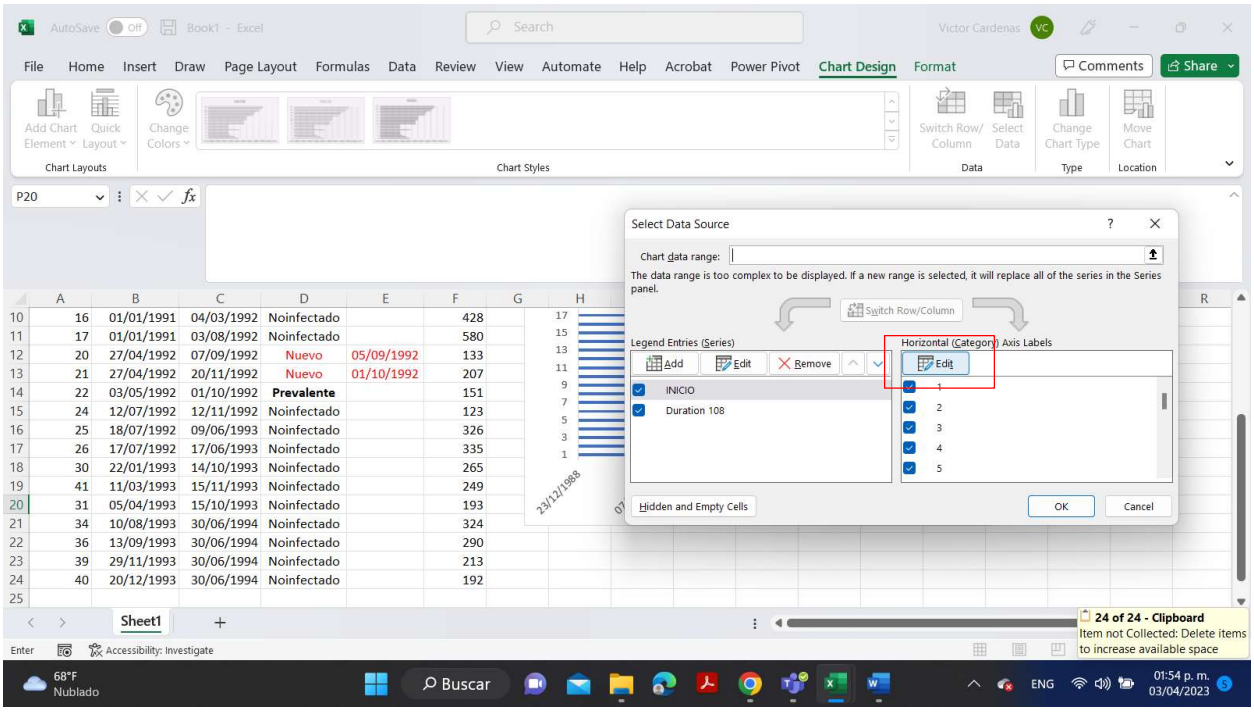
We have 23 plotted subjects, one for each stacked bar. We see that bar 1 corresponds to the subject with ID 4 who started on November 4 and ended on February 20, 1992, with a duration of 108 days, which is the true focus of our observation when putting together this graph, so that the orange bars are the ones we are interested in.

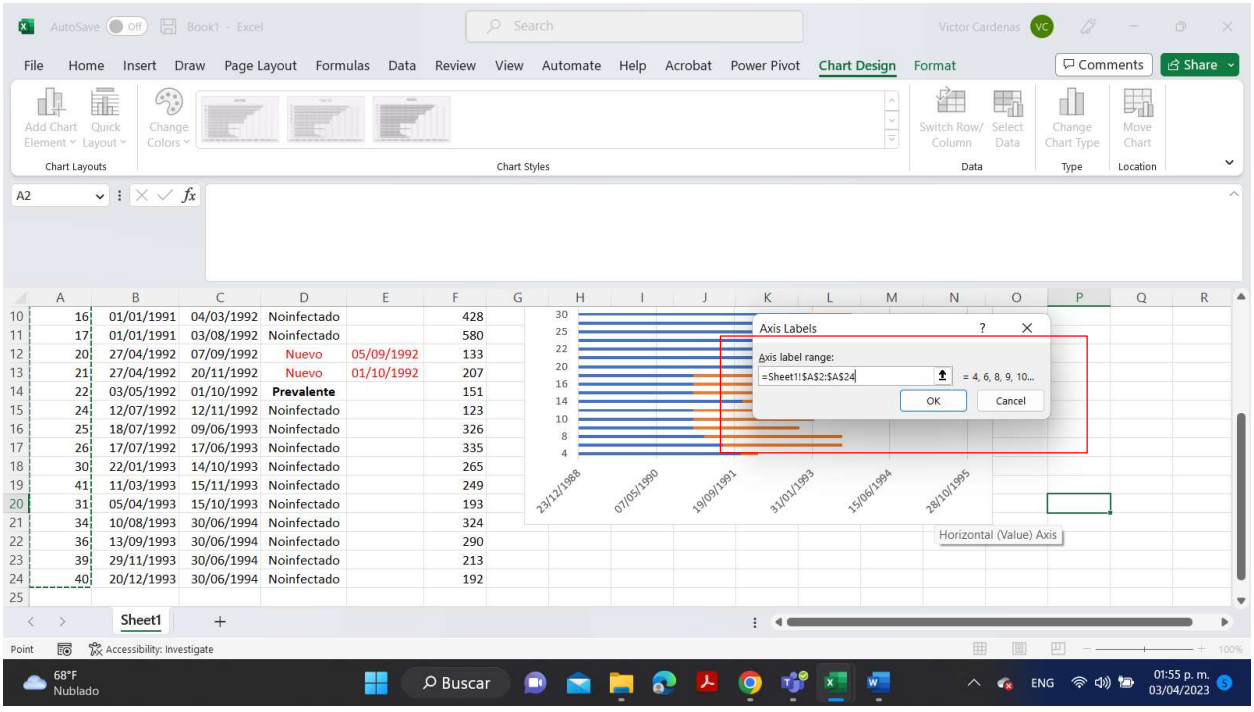
You can shorten the time periods on the X axis, in order to clarify the intent of the graph. To do this, with the right mouse button choose FORMAT OF THE GRAPH AREA, a work area will appear where it can be modified (this is optional).

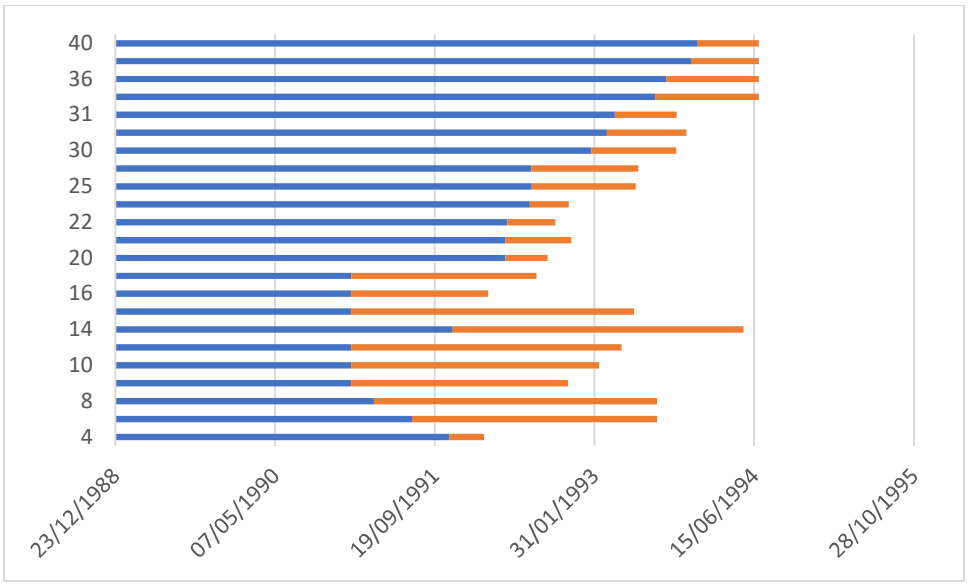


9. On the Y axis, we can place the ID of each subject, right clicking on the Y axis, choosing “Select data…”, then when the submenu appears, choose Series1 and Edit the axis labels; which will allow in a following submenu to select the labels of the axis placing those of the ID variable.

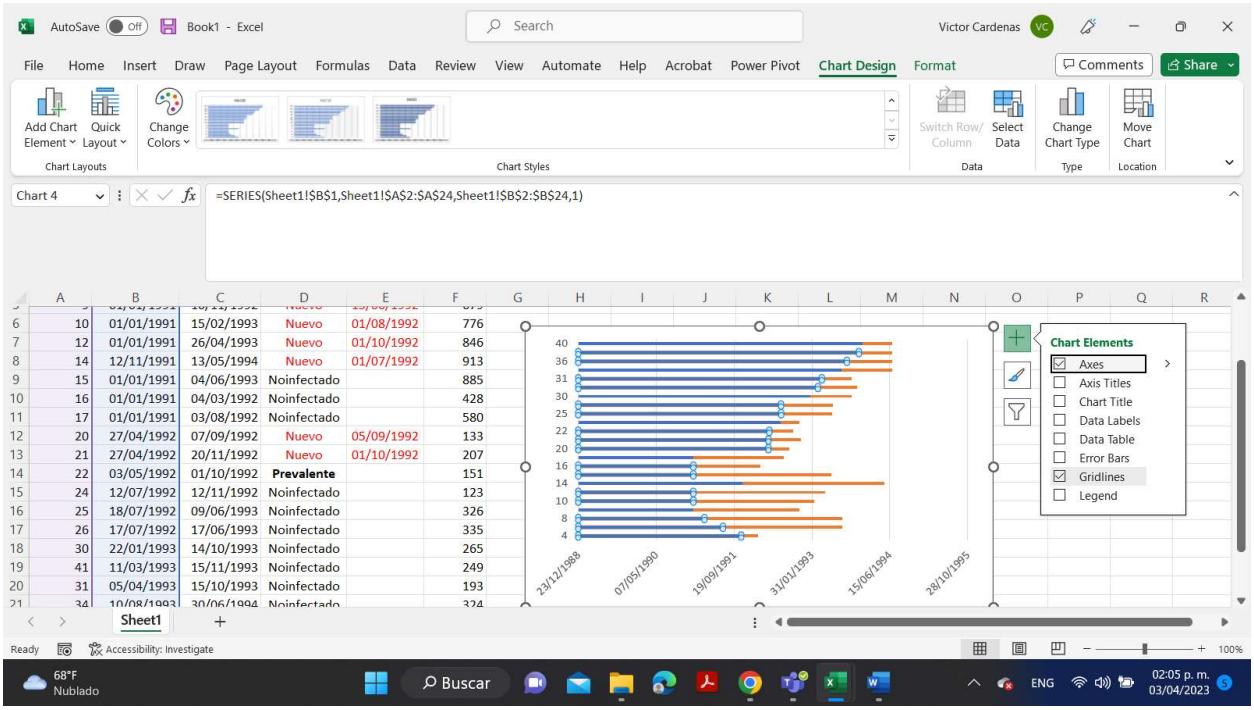


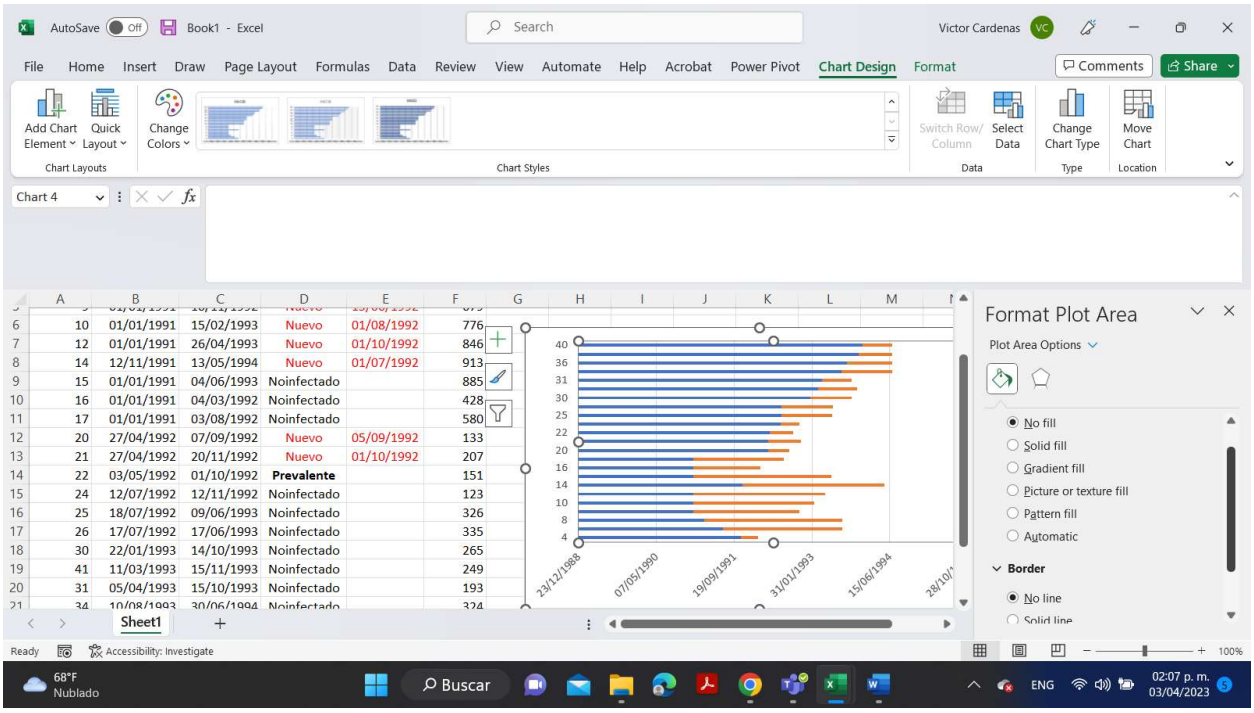

Select Data Source
Chart data range:
The data range is too complex to be displayed. If a new range is selected, it will replace all of the series in the Series panel.
Switch Row/Column
Legend Entries (Series)
Add
Edit
Remove
INICIO
Duration 108
Horizontal (Category) Axis Labels
Edit
1
2
3
4
5
Hidden and Empty Cells
OK
Cancel
Sheet1
24 of 24 - Clipboard
Item not Collected: Delete items to increase available space
68°F
Nublado
Buscar
ENG
01:54 p. m.
03/04/2023


Axis Labels
Axis label range:
=Sheet1!$A$2:$A$24
= 4, 6, 8, 9, 10...
OK
Cancel
23/12/1988
07/05/1990
19/09/1991
31/01/1993
15/06/1994
28/10/1995
Horizontal (Value) Axis
Sheet1
68°F
Nublado
Buscar
ENG
01:55 p. m.
03/04/2023

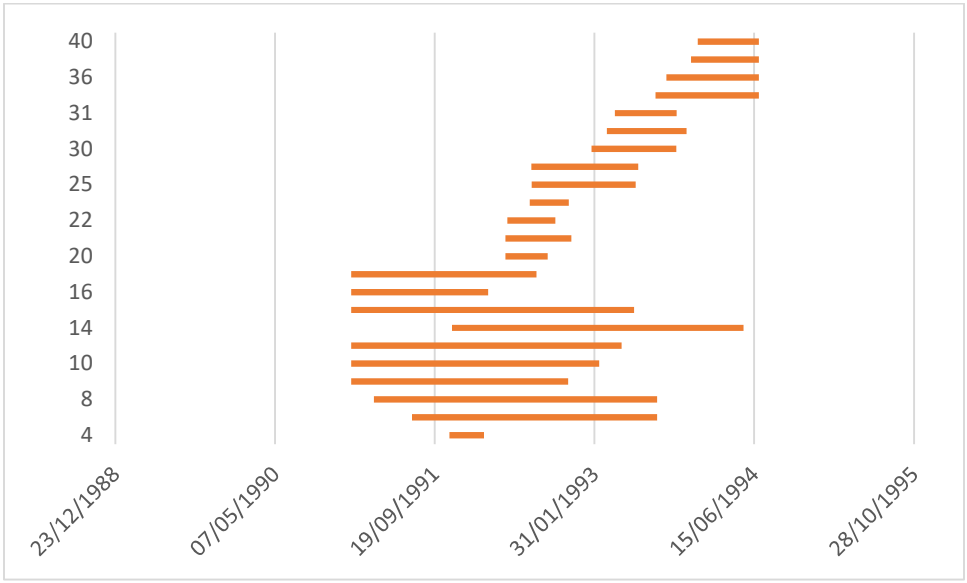


10. Now we are going to make the blue bars invisible, for which we are going to select the series displayed on bars and clicking the left button will choose "Format data series...", the menu will appear displayed, where in the series option, choosing the paint bucket, we will select the “No fill” option.



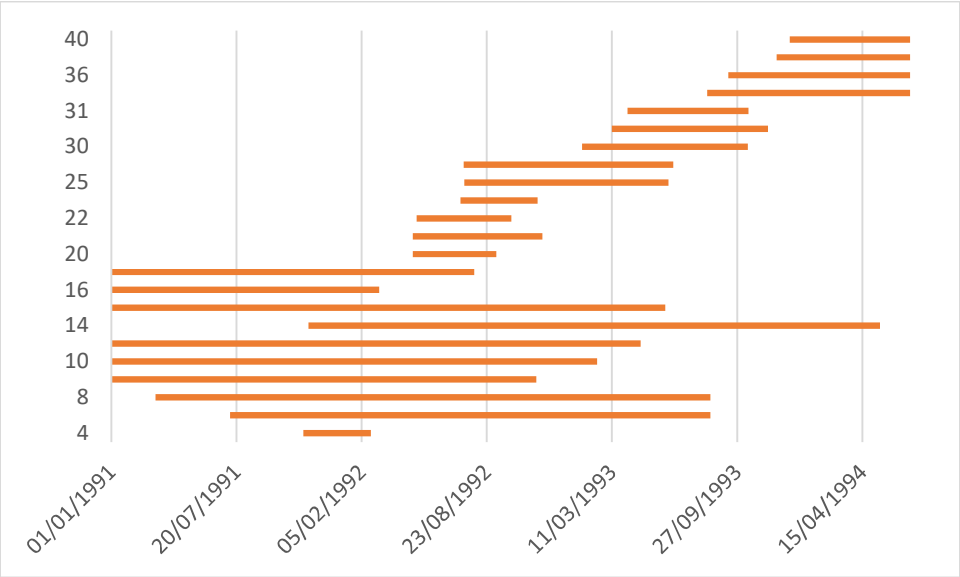


11. The graph should look like this:



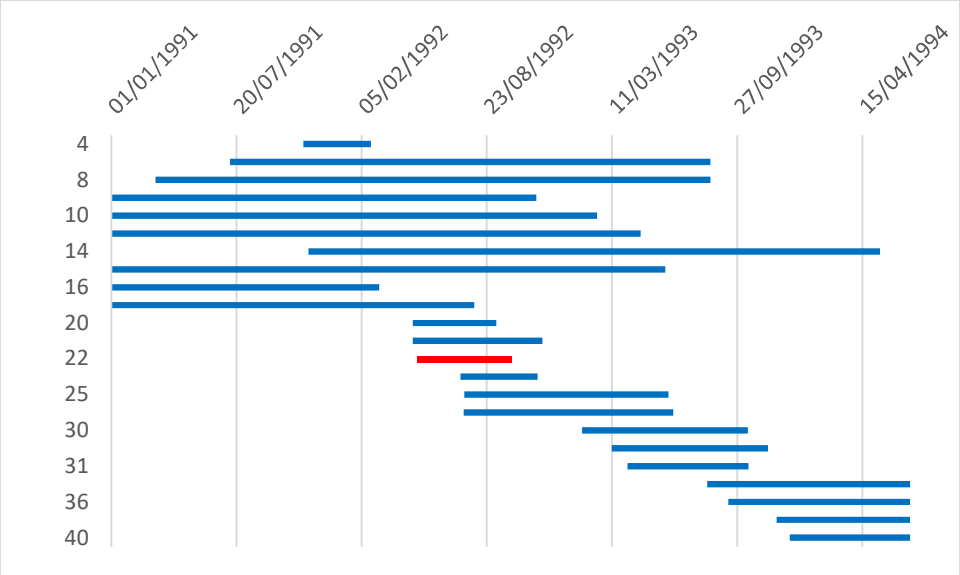
12. Let's place some limit the dates, to make sure they fall between January 1, 1991 and June 30, 1994, when the cases occurred. For this, the dates are selected, after the right click, choose "format the axis..." and in the axis options we place the maximum and minimum limits.

Observation: to obtain the numbers that go in the minimum and maximum limits you must convert the dates to the number format (33,239 and 34,515). You can figure out the range of number values by selecting the cells with those values, minimum and maximum in the desired range, and change the format to number.



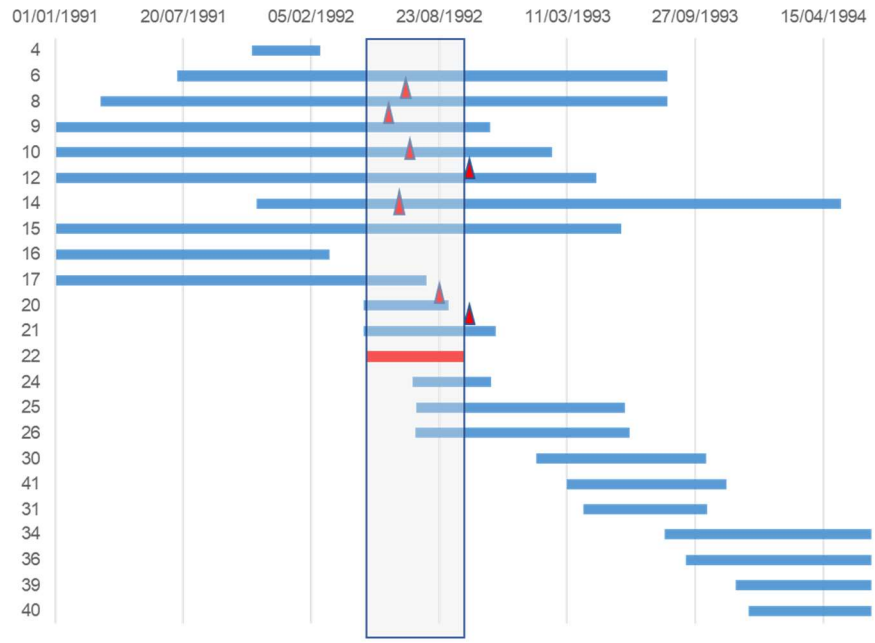
13. To achieve the proposed graph, we are going to change the color of case 22 to red and the others to blue, we will also place the first case on top, so we must go back to the Y axis, select it and enable the submenu with a right click to give axis format, there in the bar labels option select “Categories in reverse order”

14. So that the dates can also be displayed above, we will go to the submenu with the right click, format the axis, and in axis options, label position choose low. The graph will look like this:



15. Selecting the graph, we can copy it and paste it as a picture in Power Point so we can add symbols, text boxes or others that allow these to be part of the figure, this would be the final result.

Figure 6. Chronic dyalisis patients in a hemodialysis unit between January 1991, and June 1994, in whom serum for HIV antibody was available for testing, and HIV serostatus by date of serum collection



Date of serum collection with new infection –"Cases" ▲ Attending the dialysis unit (person-time) ▬

Source: Reference 3.