How do you Create a Population Pyramid in Excel?

Authored by

stats writer

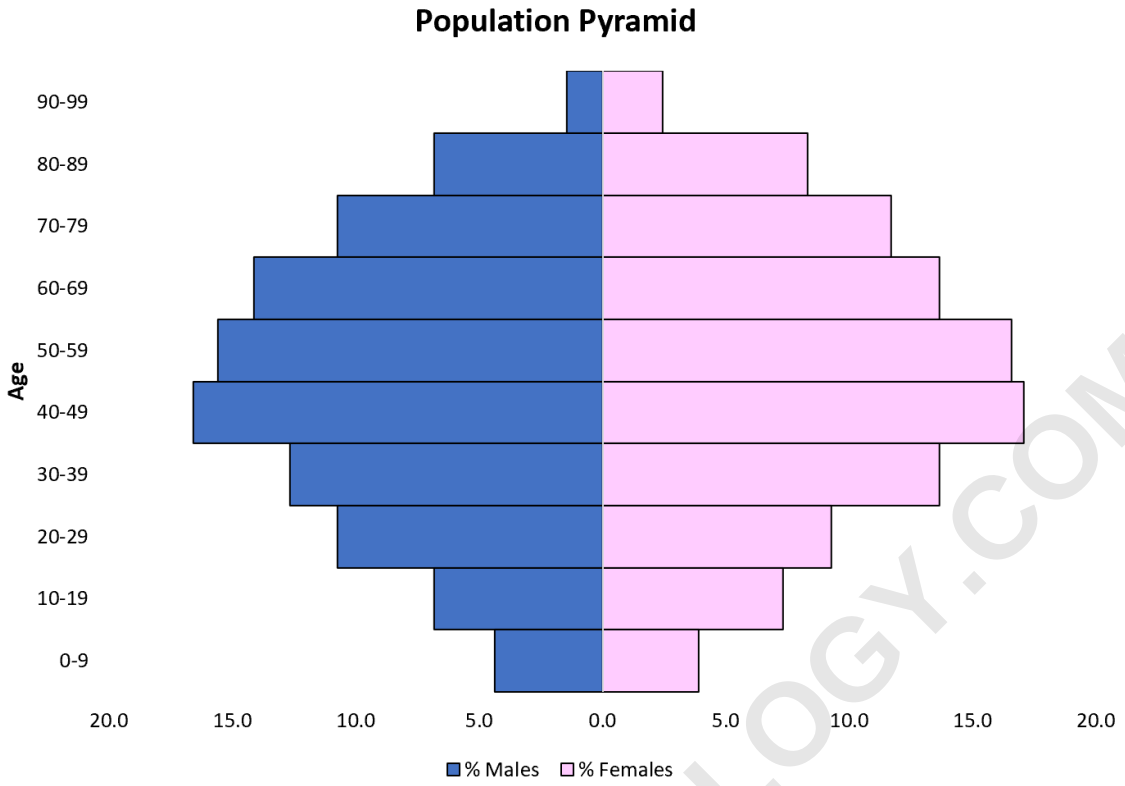
December 26, 2025

Generating a population pyramid in Excel is a powerful way to visualize demographic data. This specialized chart requires careful data preparation and specific formatting adjustments to achieve the characteristic visual structure. Unlike standard bar charts, the pyramid necessitates using two separate data series, representing male and female populations, where one series must be displayed as negative values to facilitate the back-to-back bar presentation typical of this structure.

The process begins by organizing raw data into distinct columns for age groups, male counts, and female counts. Subsequent steps involve calculating the percentage distribution for each age cohort, which standardizes the visualization across different total population sizes. Once the data is properly formatted, a **Stacked Bar Chart** is introduced, followed by crucial customization steps. These customizations include eliminating the spacing between bars and manipulating the primary horizontal axis (the x-axis) to display negative percentage values as positive numbers, ensuring clarity and accuracy in the final demographic representation.

A **population pyramid** is fundamentally a graphical illustration that details the distribution of a population by both age cohorts and gender. It serves as an invaluable tool for analysts and demographers seeking to understand the underlying composition of a population, predict future trends in population growth, and assess dependency ratios. By following this comprehensive tutorial, you will master the techniques required to construct a professional and informative population pyramid using Microsoft **Excel**.

This tutorial provides detailed instructions on the data manipulation and charting features necessary to replicate the advanced population pyramid shown below:



Data Preparation for Demographic Visualization

Before initiating the charting process, it is essential to structure your raw demographic data appropriately within the Excel worksheet. A population pyramid requires two primary components: categorical variables defining the age cohorts and quantitative variables representing the population count for each gender within those cohorts. Proper organization ensures that subsequent calculations and chart generation steps proceed smoothly and accurately.

You must establish three initial columns: one for the **Age Group** (the labels for the vertical axis), one for the raw population count of **Males**, and one for the raw population count of **Females**. Ensure that the age groups are arranged in ascending order, typically starting with the youngest cohort (e.g., 0-4 years) at the top of your dataset. This foundational structure is critical because Excel will plot these categories sequentially, corresponding to their order in the table.

While the initial data setup involves absolute population counts, visualizing these figures as percentages is highly recommended. Using percentages allows for easy comparison across different populations or time periods, as the data is normalized relative to the total population size. The following steps will demonstrate how to transform these raw counts into the necessary percentage format required for the final chart.

	A	B	C	D	E	F
1	Age	Males	Females			
2	0-9	9000	8000			
3	10-19	14000	15000			
4	20-29	22000	19000			
5	30-39	26000	28000			
6	40-49	34000	35000			
7	50-59	32000	34000			
8	60-69	29000	28000			
9	70-79	22000	24000			
10	80-89	14000	17000			
11	90-99	3000	5000			
12						
13						
14						
15						
16						
17						
18						
19						

Step 1: Inputting Raw Demographic Data

The first concrete action involves inputting the specific population figures into your spreadsheet. Begin by creating dedicated columns for the age brackets and the corresponding counts for males and females. Accuracy here is paramount, as all subsequent steps rely entirely on the validity of these initial figures. If your data source provides aggregated population figures rather than counts by gender, you will need to perform an initial calculation to segregate these totals.

For optimal visualization, the age groups should be consistently defined (e.g., 5-year intervals) and clearly labeled. These labels will form the core categories along the central vertical axis of the completed population pyramid. It is standard practice to organize the data such that the 'Male' column appears immediately before the 'Female' column, although the primary calculation step (Step 2) will transform this arrangement significantly.

Note the structure in the provided image. We have ten distinct age categories, each with a recorded count for both male and female populations. This data arrangement is the starting point for calculating relative proportions, which is the key to creating the symmetrical visual effect of the pyramid.

Step 2: Calculating Percentage Distribution for Visualization

To successfully construct a population pyramid, you cannot use the raw counts directly. Instead, you must calculate the percentage that each gender/age cohort represents relative to the total

population. Furthermore, to force the male bars to display on the left side of the vertical axis, you must convert the calculated male percentages into **negative values**. This mathematical manipulation tricks the charting mechanism in Excel into displaying the bars in opposite directions, creating the hallmark pyramid shape.

First, calculate the total population count by summing all the male and female counts across all age groups. Then, calculate the percentage for each cell. For example, the percentage for a specific male age group is calculated as: (Male Count / Total Population) * 100%. Ensure that when calculating the percentages, you use an **absolute reference** for the total population cell (e.g., $G$12) so that this denominator remains fixed when copying the formula down the column.

Crucially, once the male percentages are calculated, you must multiply this result by -1. For instance, if the formula for calculating the male percentage in cell D2 is `=(B2/$G$12)`, the modified formula used for the chart data set should be `=-(B2/$G$12)`. The female percentages remain positive. This step results in two new columns--one with negative male percentages and one with positive female percentages--which will serve as the actual source data for the stacked bar chart.

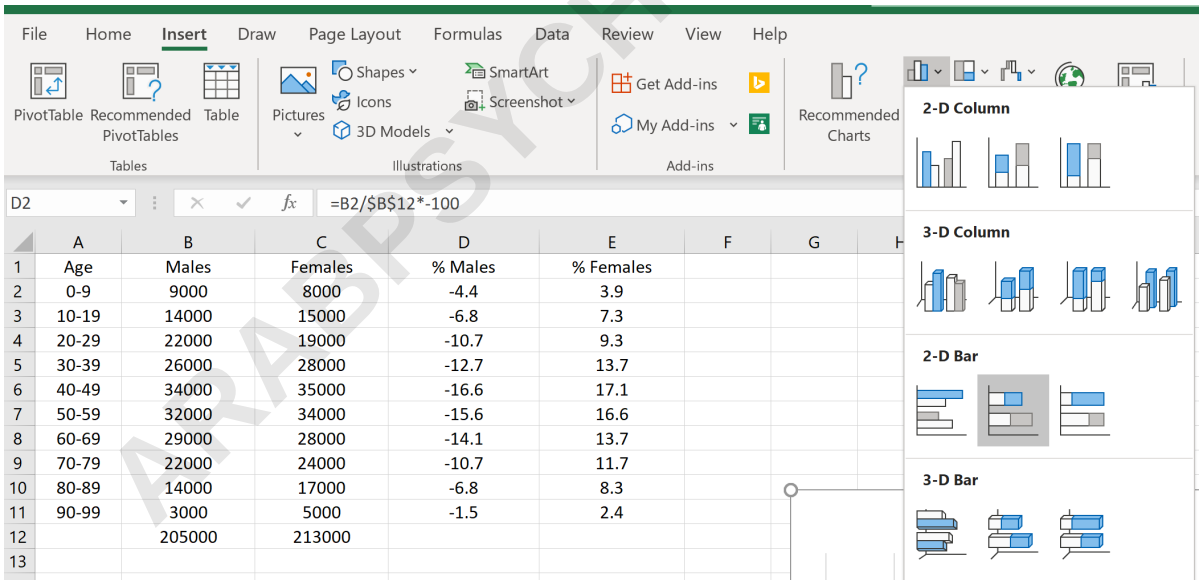
	A	B	C	D	E	F
1	Age	Males	Females	% Males	% Females	
2	0-9	9000	8000	-4.4	3.9	
3	10-19	14000	15000	-6.8	7.3	
4	20-29	22000	19000	-10.7	9.3	
5	30-39	26000	28000	-12.7	13.7	
6	40-49	34000	35000	-16.6	17.1	
7	50-59	32000	34000	-15.6	16.6	
8	60-69	29000	28000	-14.1	13.7	
9	70-79	22000	24000	-10.7	11.7	
10	80-89	14000	17000	-6.8	8.3	
11	90-99	3000	5000	-1.5	2.4	
12		205000	213000			
13						
14						
15	**Formulas Used:**					
16	Age	Males	Females	% Males	% Females	
17	0-9	9000	8000	=B17/$B$12*-100	=C17/$B$12*100	
18	10-19	14000	15000	=B18/$B$12*-100	=C18/$B$12*100	
19	20-29	22000	19000	=B19/$B$12*-100	=C19/$B$12*100	
20	30-39	26000	28000	=B20/$B$12*-100	=C20/$B$12*100	
21	40-49	34000	35000	=B21/$B$12*-100	=C21/$B$12*100	
22	50-59	32000	34000	=B22/$B$12*-100	=C22/$B$12*100	
23	60-69	29000	28000	=B23/$B$12*-100	=C23/$B$12*100	
24	70-79	22000	24000	=B24/$B$12*-100	=C24/$B$12*100	
25	80-89	14000	17000	=B25/$B$12*-100	=C25/$B$12*100	
26	90-99	3000	5000	=B26/$B$12*-100	=C26/$B$12*100	
27		=SUM(B17:B26)	=SUM(C17:C26)			
28						
29						
30						

Step 3: Creating the Initial 2-D Stacked Bar Chart

With the data correctly transformed into negative and positive percentages, you are ready to insert the initial chart. The key chart type for a population pyramid is the **Stacked Bar Chart**, specifically the 2-D version, as this configuration allows the display of the male and female data series side-by-side along a shared vertical axis of age categories.

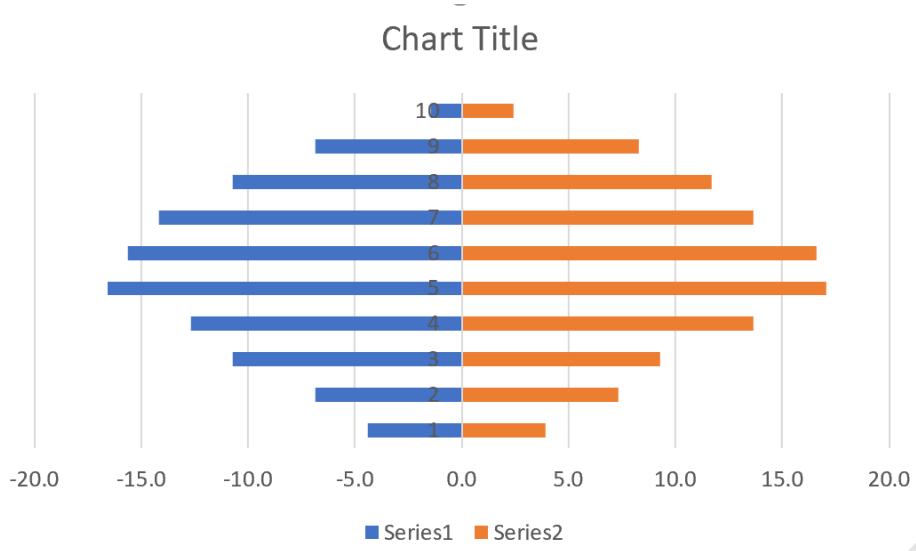
The process for insertion is straightforward. First, **highlight the data range** containing only the Age Group labels and the two percentage columns (e.g., highlight cells C2:E11, assuming C contains the labels and D/E contain the negative male and positive female percentages, respectively). Next, navigate to the **Insert** tab on the Excel ribbon. Within the **Charts** group, select the option for the **2-D Stacked Bar Chart**. Ensure you select the horizontal bar chart option, not the vertical column chart.

Upon selection, Excel will automatically generate a chart. At this stage, the chart will appear incomplete. The bars will have gaps between the age categories, and the horizontal axis (x-axis) will show negative numbers, which is technically correct based on the input data but visually confusing for a demographic visualization. These issues will be resolved in the subsequent formatting steps.



	A	B	C	D	E
1	Age	Males	Females	% Males	% Females
2	0-9	9000	8000	-4.4	3.9
3	10-19	14000	15000	-6.8	7.3
4	20-29	22000	19000	-10.7	9.3
5	30-39	26000	28000	-12.7	13.7
6	40-49	34000	35000	-16.6	17.1
7	50-59	32000	34000	-15.6	16.6
8	60-69	29000	28000	-14.1	13.7
9	70-79	22000	24000	-10.7	11.7
10	80-89	14000	17000	-6.8	8.3
11	90-99	3000	5000	-1.5	2.4
12		205000	213000		
13					

The following chart will automatically appear, requiring significant modification:



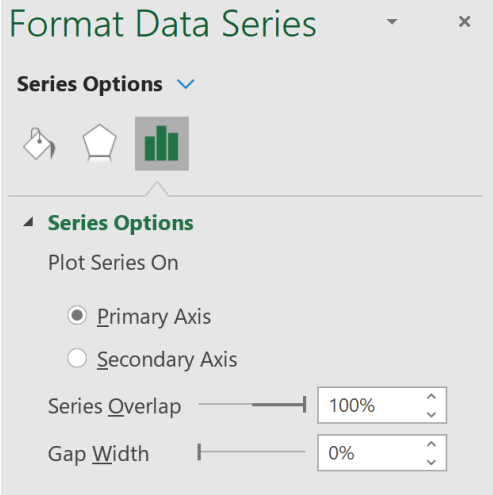
Step 4: Essential Modifications to Achieve the Pyramid Look

The initial chart lacks the continuous, cohesive shape of a true population pyramid. Several critical formatting adjustments are necessary to transform the standard stacked bar chart into the desired demographic visual. These modifications focus on adjusting the bar spacing and enhancing the visual aesthetics of the chart elements.

Eliminating the Gap Width between Bars

To create the characteristic solid pyramid shape, the spacing between the bars, known as the Gap Width, must be removed. This ensures that the bars for contiguous age groups touch each other seamlessly, emphasizing the distribution across the entire lifespan.

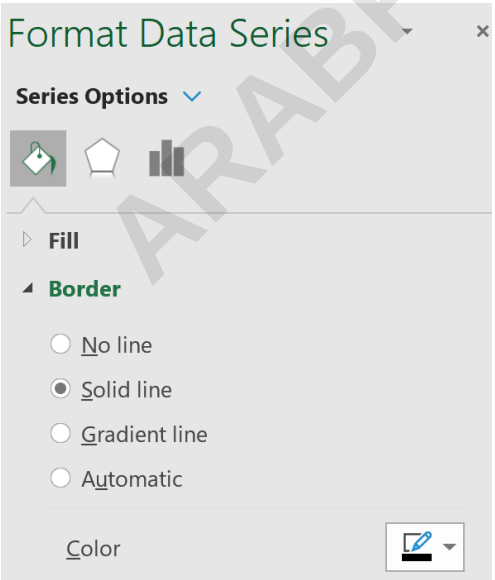
Accessing Format Data Series: Right-click on any bar in the chart (either the male or female series). From the context menu, select the option labeled **Format Data Series...**
Adjusting Gap Width: In the task pane that appears, locate the Series Options settings. Change the value for **Gap Width** to **0%**. This action causes all bars representing different age groups to align directly against one another.



Applying Borders for Definition

Since the bars are now touching, adding a thin border line helps delineate the boundaries between the adjacent age groups, significantly enhancing readability and professional appearance. This border should ideally be a contrasting color, such as black.

Selecting Border Options: Ensure the data series is still selected. In the **Format Data Series** pane, click on the paint bucket icon (Fill & Line).
Defining the Border: Expand the **Border** section. Select **Solid line**.
Setting Color: Change the **Color** option to black. Repeat this process for the other data series (male or female) to ensure both sides of the pyramid have defined boundaries.



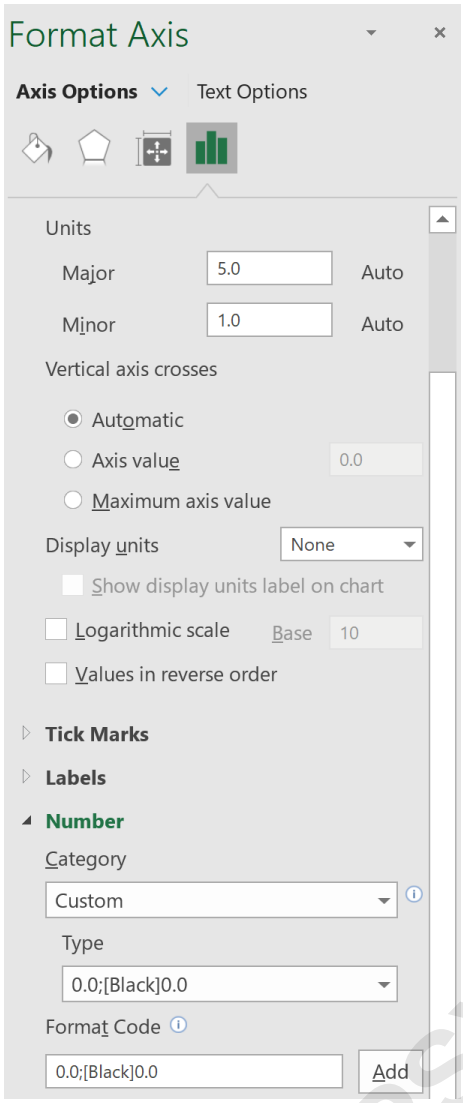
Step 5: Formatting the Axes for Clarity

The axis formatting is perhaps the most crucial step for presenting the population pyramid correctly, as it addresses the visual confusion caused by the necessary negative input data and the positioning of the age labels.

Displaying the Horizontal Axis Labels as Positive Numbers

Even though the male population is represented by negative values in the data source, the chart must visually display these values as positive percentages (e.g., 5.0% instead of -5.0%). This transformation is achieved using a custom number format code within Excel.

Accessing Format Axis: Right-click directly on the horizontal axis (x-axis) labels (the percentage numbers). Select **Format Axis...**
Entering Custom Format: In the task pane, click on the **Number** category options.
Under the **Format Code** input box, type the specific custom code: `0.0;0.0` (or `0%;0%` if displaying as percentages). This code dictates that positive numbers (0.0) and negative numbers (0.0, styled in black) should both be displayed identically, without the minus sign.
Click **Add** to apply this custom format to the axis. The negative signs will immediately disappear, providing a clean percentage display.

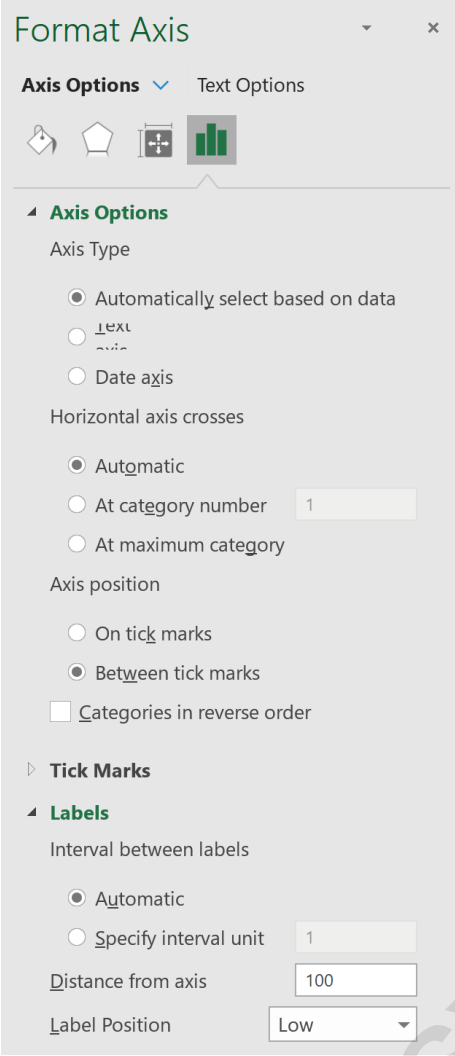


Relocating the Vertical Age Axis Labels

By default, the vertical axis (y-axis) labels, which represent the age cohorts, are usually positioned in the center of the chart where the horizontal axis intersects. For a population pyramid, these labels should ideally be moved to the left edge of the chart area to maximize space and cleanly separate the labels from the data bars.

Accessing Format Axis (Vertical): Right-click on the vertical axis (the age group labels). Select **Format Axis...**
Adjusting Label Position: In the task pane, navigate to the **Labels** section. Under **Label Position**, change the setting from "Next to Axis" or "Automatic" to **Low**.

Setting the label position to **Low** instructs Excel to place the age labels at the furthest left boundary of the chart, thereby positioning them effectively outside the plotted area of the bars.



Step 6: Finalizing Aesthetics and Presentation

The final stage involves cleaning up the visual elements to ensure the chart is professional, informative, and ready for presentation. Attention to detail in the title, color scheme, and removal of superfluous elements significantly impacts the overall readability of the demographic visualization.

Start by providing a clear, descriptive title for the graph (e.g., "Population Structure by Age and Gender, 2024"). Adjust the colors of the two data series to clearly distinguish between males and females, following conventional standards if possible (e.g., blue for males, red/pink for females). These colors should be selected from the **Fill & Line** options in the Format Data Series pane.

A crucial element to remove are the primary vertical grid lines. These lines, which extend from the central vertical axis across the graph, often clutter the visualization and detract from the pyramid shape. Simply click on any of the vertical grid lines and press the **Delete** key to remove them. You may also consider adding data labels to the end of each bar if precise percentage values are

required, although this can sometimes make the chart overly dense.

Upon completing these detailed steps--from data transformation to axis manipulation and aesthetic refinements--the resulting visualization is a clean, accurate, and standardized population pyramid.

The final, polished result should accurately reflect the demographic structure: