1. United Nations life expectancy data

Life expectancy at birth is a measure of the average a living being is expected to live. It takes into account several demographic factors like gender, country, or year of birth.

Life expectancy at birth can vary along time or between countries because of many causes: the evolution of medicine, the degree of development of countries, or the effect of armed conflicts. Life expectancy varies between gender, as well. The data shows that women live longer that men. Why? Several potential factors, including biological reasons and the theory that women tend to be more health conscious.

Let's create some plots to explore the inequalities about life expectancy at birth around the world. We will use a dataset from the United Nations Statistics Division, which is available here.

A data.frame: 6 × 7

	Country.or.Area	Subgroup	Year	Source	Unit
	<chr>	<chr>	<chr>	<chr>	<chr>
1	Afghanistan	Female	2000-2005	UNPD_World Population Prospects_2006 (International estimate)	Years
2	Afghanistan	Female	1995-2000	UNPD_World Population Prospects_2006 (International estimate)	Years
3	Afghanistan	Female	1990-1995	UNPD_World Population Prospects_2006 (International estimate)	Years
4	Afghanistan	Female	1985-1990	UNPD_World Population Prospects_2006 (International estimate)	Years
5	Afghanistan	Male	2000-2005	UNPD_World Population Prospects_2006 (International estimate)	Years
6	Afghanistan	Male	1995-2000	UNPD_World Population Prospects_2006 (International estimate)	Years

2. Life expectancy of men vs. women by country

Let's manipulate the data to make our exploration easier. We will build the dataset for our first plot in which we will represent the average life expectancy of men and women across countries for the last period recorded in our data (2000-2005).

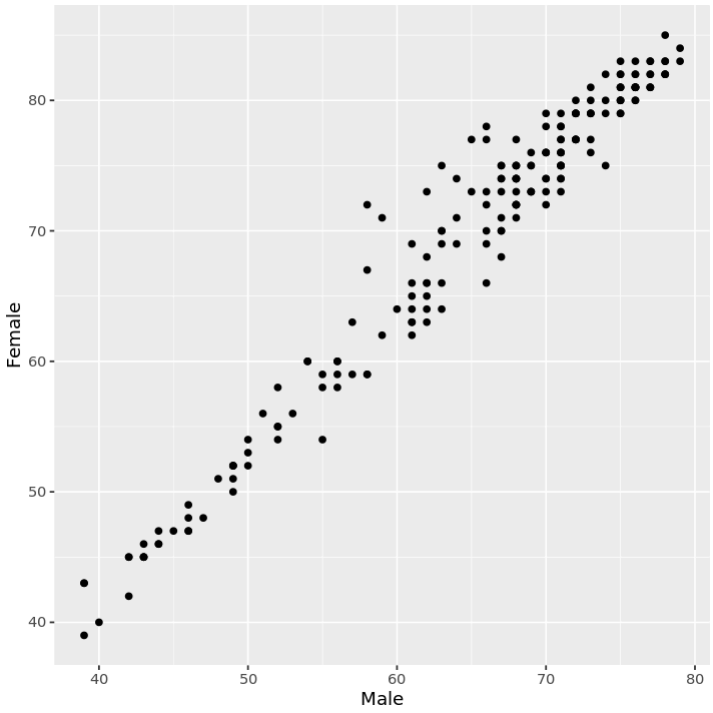
A data.frame: 6 × 3

	Country.or.Area	Female	Male
	<chr>	<int>	<int>
1	Afghanistan	42	42
2	Albania	79	73
3	Algeria	72	70
4	Angola	43	39
5	Argentina	78	71
6	Armenia	75	68

3. Visualize I

A scatter plot is a useful way to visualize the relationship between two variables. It is a simple plot in which points are arranged on two axes, each of which represents one of those variables.

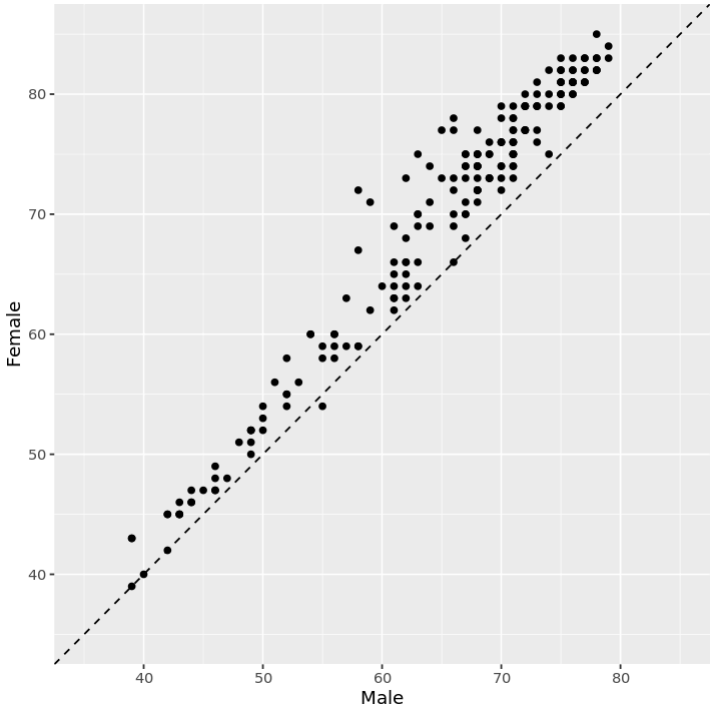
Let's create a scatter plot using `ggplot2` to represent life expectancy of males (on the x-axis) against females (on the y-axis). We will create a straightforward plot in this task, without many details. We will take care of these kinds of things shortly.



4. Reference lines I

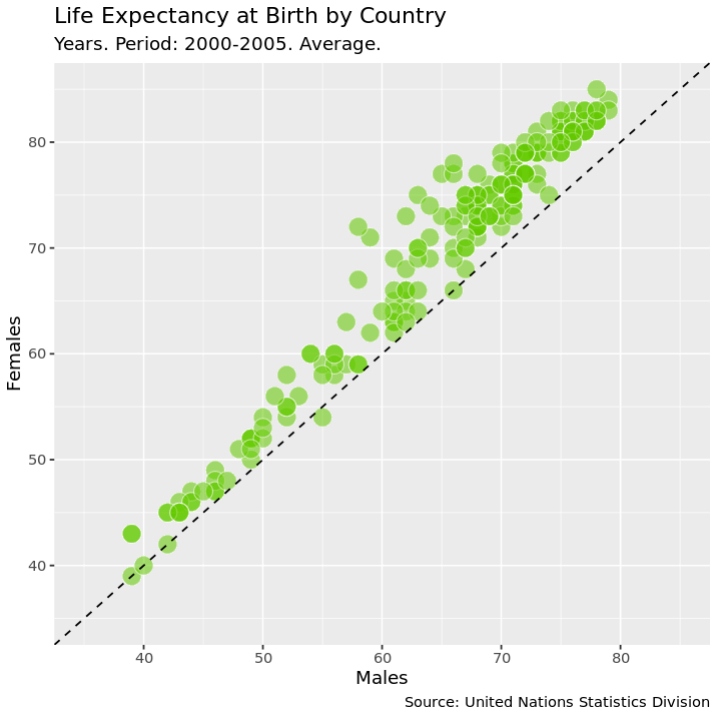
A good plot must be easy to understand. There are many tools in `ggplot2` to achieve this goal and we will explore some of them now. Starting from the previous plot, let's set the same limits for both axes as well as place a diagonal line for reference. After doing this, the difference between men and women across countries will be easier to interpret.

After completing this task, we will see how most of the points are arranged above the diagonal and how there is a significant dispersion among them. What does this all mean?



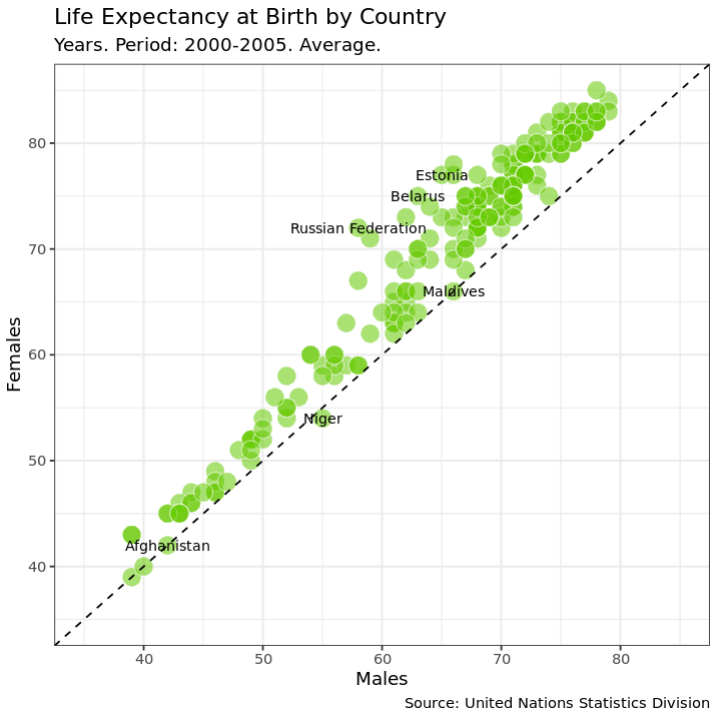
5. Plot titles and axis labels

A key point to make a plot understandable is placing clear labels on it. Let's add titles, axis labels, and a caption to refer to the source of data. Let's also change the appearance to make it clearer.



6. Highlighting remarkable countries I

Now, we will label some points of our plot with the name of its corresponding country. We want to draw attention to some special countries where the gap in life expectancy between men and women is significantly high. These will be the final touches on this first plot.



7. How has life expectancy by gender evolved?

Since our data contains historical information, let's see now how life expectancy has evolved in recent years. Our second plot will represent the difference between men and women across countries between two periods: 2000-2005 and 1985-1990.

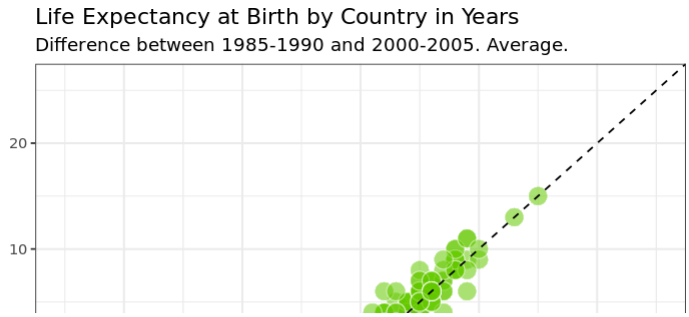
Let's start building a dataset called `subdata2` for our second plot.

A data.frame: 6 × 10

	Country.or.Area	Source	Unit	Value.Footnotes	Fe
	<chr>	<chr>	<chr>	<int>	<i
1	Afghanistan	UNPD_World Population Prospects_2006 (International estimate)	Years	NA	41
2	Albania	UNPD_World Population Prospects_2006 (International estimate)	Years	NA	75
3	Algeria	UNPD_World Population Prospects_2006 (International estimate)	Years	NA	67

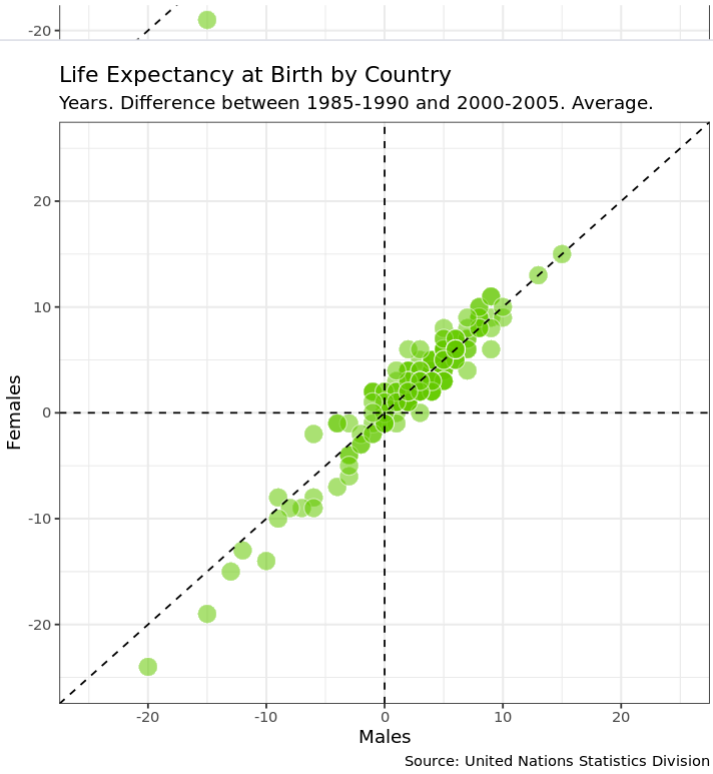
8. Visualize II

Now let's create our second plot in which we will represent average life expectancy differences between "1985-1990" and "2000-2005" for men and women.



9. Reference lines II

Adding reference lines can make plots easier to understand. We already added a diagonal line to visualize differences between men and women more clearly. Now we will add two more lines to help to identify in which countries people increased or decreased their life expectancy in the period analyzed.



10. Highlighting remarkable countries II

As we did in the first plot, let's label some points. Concretely, we will point those three where the aggregated average life expectancy for men and women increased most and those three where decreased most in the period