

Video: Basic statistics in MATLAB (3:30 min)

(00:00)

We've already set up the project. We've created a Lesson 5 folder and a Lesson 5 script to go with it. We've downloaded the data NYCDiseases.mat, and we've loaded it into the workspace. Let's start by calculating the average of the measles dataset. We'll create a new cell, and we'll use the mean function to calculate the average. We'll define a variable called measlesAver which is going to hold the overall average. The mean function calculates the average of an array. We apply it to measles, but we use the linear representation to put measles in a single column so we get a single value out, which is the overall average. We execute, and we see a new variable in the workspace.

(00:53)

Now let's output the average in the command window using the fprintf. We'll clean up the documentation a little bit first so we know what each cell is doing. To call the fprintf statement, I start with the function name fprintf and give the arguments in parenthesis. The first argument is the format specifier which says how to output. I begin with an informative message and follow it with the value of the average. In order to output a variable without actually typing the variable in, I use a %g. When MATLAB encounters a %g, it looks at the next variable in its argument list and outputs its value. Since I want to output the average of measles, I put the measlesAver next. When I save the cell and execute, I see the message in the command window.

(01:59)

The next example shows how to output values from a calculation without defining intermediate variables. In this example, I'm going to output the overall median, max, and min of measles cases. After typing the name of the function fprintf, I give a format specifier which contains a message very similar to what I want in the output. I begin by outputting a median (that's the first %g), and then I'll output the max and the min, which I'm enclosing in square brackets. I'll follow that by a \n so that the output will fall on a new line. After closing the quotes, we have to list the three values corresponding to the %g's. These are going to be the overall median, the overall max, and the overall min of the measles array. We execute the cell and we get an error message also indicated by the red bar. It looks like we're missing a closing brace. I'll correct the problem, save the cell, and re-execute. Now I see the message that I wanted.