

# Setting Parameters/Variables in Pig

In Pig, parameters can be set using the `%declare` command and `%default` command:

## **%DECLARE:**

Scripts can store variables by declaring them using the `declare` command as shown in the example below. The variable name should be preceded by '\$' when being used after declaration.

### **Example:**

```
%declare MyVariable 'useful_data';  
A = LOAD '/data/mydata/$MyVariable';  
B = FILTER A BY $0>'5'
```

## **%DEFAULT:**

Default command can be used to give default values to variables in case a particular value is not specified.

### **Example:**

```
%default DATE '20090101';  
A1 = load '/data/mydata/$DATE';
```

# Specifying parameters in the command line:

Command line arguments can be specified using the `--param` option as shown below:

### **Example:**

Assuming we have a script 'myscript.pig' with the following contents:

```
A= LOAD '$data' USING PigStorage() AS (f1:int, f2:int, f3:int);
```

This script can be executed by indicating the 'data' parameter in the following way:

```
>pig -f myscript.pig --param data=Somedata.txt
```

## Specifying parameters using a parameter file:

Parameters can also be specified using a file. This is useful when we have many parameters to be passed to a script.

### Example:

```
>pig -param_file myparams.txt myscript.pig
```

Contents of 'myparams.txt':

```
data=mydata
```

## Order of precedence for parameters:

1. Highest precedence is given to parameters defined using a declare preprocessor.
2. Second highest precedence is for the parameters defined as command line arguments (Using -param option).
3. Third highest precedence is for parameters that are defined in a parameter file.
4. Lowest precedence is given to parameters defined using the default command.

For more details, refer:

[https://pig.apache.org/docs/r0.7.0/piglatin\\_ref2.html](https://pig.apache.org/docs/r0.7.0/piglatin_ref2.html)