C LIBRARY - <FLOAT.H>

The **float.h** header file of the C Standard Library contains a set of various platform-dependent constants related to floating point values. These constants are proposed by ANSI C. They allow making more portable programs. Before checking all the constants, it is good to understand that floating-point number is composed of following four elements −

Component	Component Description
S	sign +/ −
b	base or radix of the exponent representation, 2 for binary, 10 for decimal, 16 for hexadecimal, and so on...
e	exponent, an integer between a minimum $\mathbf{e_{min}}$ and a maximum $\mathbf{e_{max}}$.
p	precision, the number of base-b digits in the significand.

Based on the above 4 components, a floating point will have its value as follows −

```
floating-point = ( S ) p x b^e

or

floating-point = (+/-) precision x base^exponent
```

Library Macros

The following values are implementation-specific and defined with the #define directive, but these values may not be any lower than what is given here. Note that in all instances FLT refers to type **float**, DBL refers to **double**, and LDBL refers to **long double**.

Macro	Description
FLT_ROUNDS	Defines the rounding mode for floating point addition and it can have any of the following values − • -1 − indeterminable • 0 − towards zero • 1 − to nearest • 2 − towards positive infinity • 3 − towards negative infinity
FLT_RADIX 2	This defines the base radix representation of the exponent. A base-2 is binary, base-10 is the normal decimal representation, base-16 is Hex.
FLT_MANT_DIG DBL_MANT_DIG LDBL_MANT_DIG	These macros define the number of digits in the number $intheFLT_RADIXbase$.

FLT_DIG 6 DBL_DIG 10 LDBL_DIG 10	These macros define the maximum number decimal digits $base - 10$ that can be represented without change after rounding.
FLT_MIN_EXP DBL_MIN_EXP LDBL_MIN_EXP	These macros define the minimum negative integer value for an exponent in base FLT_RADIX.
FLT_MIN_10_EXP -37 DBL_MIN_10_EXP -37 LDBL_MIN_10_EXP -37	These macros define the minimum negative integer value for an exponent in base 10.
FLT_MAX_EXP DBL_MAX_EXP LDBL_MAX_EXP	These macros define the maximum integer value for an exponent in base FLT_RADIX.
FLT_MAX_10_EXP +37 DBL_MAX_10_EXP +37 LDBL_MAX_10_EXP +37	These macros define the maximum integer value for an exponent in base 10.
FLT_MAX 1E+37 DBL_MAX 1E+37 LDBL_MAX 1E+37	These macros define the maximum finite floating-point value.
FLT_EPSILON 1E-5 DBL_EPSILON 1E-9 LDBL_EPSILON 1E-9	These macros define the least significant digit representable.
FLT_MIN 1E-37 DBL_MIN 1E-37 LDBL_MIN 1E-37	These macros define the minimum floating-point values.

Example

The following example shows the usage of few of the constants defined in float.h file.

```c
#include <stdio.h>
#include <float.h>

int main()
{
   printf("The maximum value of float = %.10e\n", FLT_MAX);
   printf("The minimum value of float = %.10e\n", FLT_MIN);

   printf("The number of digits in the number = %.10e\n", FLT_MANT_DIG);
}
```

Let us compile and run the above program that will produce the following result –

```
The maximum value of float = 3.4028234664e+38
The minimum value of float = 1.1754943508e-38
The number of digits in the number = 7.2996655210e-312
```