# 単語数や行数を数えるコマンド

**■文字数を数えるwcコマンド**

wcコマンドはファイルの行数、単語数、文字数を数えるフィルタコマンドです。

単語数は、空白または改行で区切られている前提でカウントされます。

●wcコマンドの書式

|  |
| --- |
| wc 【ファイルのパス】 |

次にwcコマンドをカレントディレクトリの「mail.csv」に対して実行した結果を示します。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ wc mail.csv  8 16 321 mail.csv |

結果は行数、単語数、文字数(バイト数)の順に表示されます。引数に複数のファイルを指定した場合は、個別の結果の後に合計が表示されます。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ wc mail.csv maillist.org.txt  8 16 321 mail.csv  12 12 302 maillist.org.txt  20 28 623 total |

行数、単語数、文字数(バイト数)のいずれかだけを表示したい場合は下記のオプションを使います。

|  |  |
| --- | --- |
| オプション | 説明 |
| -c | 文字数(バイト数)を表示 |
| -l | 行数を表示 |
| -w | 単語数を表示 |

例えば、行数だけを表示するには下記のように指定します。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ wc -l mail.csv  8 mail.csv |

**■標準入力からデータを読み込む**

wcコマンドはフィルタコマンドなので引数でファイルのパスを指定しなければ標準入力からデータが読み込まれます。他のコマンドと組み合わせると、色々と便利に使えます。

例えば、指定したディレクトリのファイルの数を表示するには、ファイルの一覧を表示するlsコマンドの出力をパイプ「|」で渡します。ファイルの数を調べるには、行数と単語数のどちらを使えばよいでしょうか？試しにホームページの一覧を表示する「ls ~」コマンドの出力を、パイプ「|」を介してwcコマンドに繋げてみます。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ ls ~ | wc  16 17 155　　　　　　←左から順に行数、単語数、文字数 |

この例では、行数と単語数が1だけ違います。通常、lsコマンドは画面に結果を表示するとき、画面のサイズに応じて横にいくつか並べて表示します。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ ls ~　　↓1行に表示するファイル数は画面の幅で決まる  archivefile.tar.bz2 bin document mail.csv maillist.org.txt My New test |

ところが、結果をパイプ「|」を介して別のコマンドに渡した場合は、1行に一つずつファイルが並べられます。結果をcatコマンド(もしくは「cat –n」)に渡すと確認できます。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ ls ~ | cat  archivefile.tar.bz2  bin  cal.txt  document  mail.csv  maillist.org.txt  My New  test  test1.txt  test2.txt  test3.txt  test4.txt  test5.txt  test.csv  test.txt  yahoo.html |

それでは、なぜ行数と単語数が異なっていたのか。本来なら行数が16行、単語数も16個のはずです。それは「My New」というディレクトリ名は間にスペースがあるため、2つの単語としてカウントされたわけです。スペースがある名前はこれだけですから、結果として単語数が行数より1多くなっていたわけです。したがって、ディレクトリ内の、ファイルの数だけを求めるには「-l」オプションを使用します。

|  |
| --- |
| [ec2-user@ip-172-31-8-35 ~]$ ls ~ | wc -l  16　←ホームディレクトリ内のファイルの数 |