**Python Function to Read a Mixed Numeric/String Table from a Text File**

By Bill Menke, June 26, 2025

A commonly encountered scientific data format is a text file containing a table of data, in which each data entry is delimited by a tab, comma or other known character, and in which each column is either numerical or textual (i.e., strings). Sometimes, the first row of the table contains textual column names.

Several well-known Python packages (e.g. pandas) include methods for reading such a table. However, some of these methods produce output that can be difficult to use. I present here is a stand-alone function as an alternative. It is more tuned to the tables commonly encountered in environmental science.

```
import numpy as np

# readtable:  reads a table of mixed numerical and string columns from a text file
#             (a column contains strings if at least one of its rows is non-numeric)
#             (a table entry of no length counts as a string)
# by Bill Menke, June 2025
# Input:
#        pathmane: pathname of text file
#        delimiter: character between columns, e.g. "\t" or ","
#        hascolnames: 1 = first row of table is column names, 0 = no column names
present
#        (colnames are string entries in the first row of the table)
# Output:
#        T the table as a list of M columns, where each column is either
#             a Nx1 array of floats or a len-N list of character strings
#        N: number of rows
#        M: number of columns
#        isnum: Mx0 array of integers, 1: col is numerical, 0: col is strings
#        colname: list of colnames
#        status: 1 on success, 0 on fail
# How to access row n of column m of table T:
#        col = T[m];
#        if( isnum[m] == 1 ):
#            value = col[n,0];
#        else:
#            string = col[n];
def readtable(pathname,delimiter,hascolnames):
    T = [];
    isnum=np.zeros((0,));
    colname = [];
    status=0;
    N=0;
    M=0;
    # open file
    try:
        fdr = open(pathname);
    except:
        print("error: reaadtable: can open %s" % (pathname,) );
        return(T,N,M,isnum,colname,status);

    # part 1: determine size
    N=0;
    while True:
        line = fdr.readline();
        if not line:
            fdr.close();
            break;
        Nline = len(line);
        j = [];
```

```
    for i in range(Nline):
        if( line[i] == delimiter ):
            j.append(i);
    Mj = len(j)+1;
    if( N==0 ):
        M=Mj;
    if( Mj != M ):
        print("error: reaadtable: wrong number of cols, line %d" % (Mj,) );
        return(T,N,M,isnum,colname,status);
    N=N+1;

# part 2: read into list of lists
fdr = open(pathname);
L = [];
NL = 0;
while True:
    line = fdr.readline();
    Nline = len(line);
    if not line:
        fdr.close();
        break;
    L.append([]);
    Nline = len(line);
    j = [];
    for i in range(Nline):
        if( line[i] == delimiter ):
            j.append(i);
    j.append(Nline-1);
    Nj = len(j);
    l=0;
    for i in range(Nj):
        r=j[i];
        L[NL].append(line[l:r]);
        l=j[i]+1;
    NL=NL+1;

# part 3: deal with coilnames
if( hascolnames == 1 ):
    if( N==0 ):
        print("error: reaadtable: no colnames present" );
        return(T,N,M,isnum,colname,status);
    for i in range(M):
        colname.append((L[0])[i]);
    L = L[1:N];
    N=N-1;
else:
    for i in range(M):
        colname.append("");
if( N==0 ):
    print("error: reaadtable: no data present" );
    return(T,N,M,isnum,colname,status);

# part 4 determine numerical columns
isnum = np.ones((M,),dtype=int);
for m in range(M):
    for n in range(N):
        try:
            v = float((L[n])[m]);
        except:
            isnum[m]=0;
            break;

# part 5 build table
```

```
    for m in range(M):
        if( isnum[m] == 1 ):
            c = np.zeros((N,1));
            for n in range(N):
                c[n] = float((L[n])[m]);
        else:
            c = [];
            for n in range(N):
                c.append((L[n])[m]);
        T.append(c);

    status=1;
    return(T,N,M,isnum,colname,status);
```

---- The data file mydata.txt -----

```
C0    C1    C2
1     2     4
2     5     4
3     9     4
4           5
5     A     4
6     3     4
7     B     4
8     5     4
9     1     4
10    0     6
```

--- end of the data file ----

--- Example ----

```
pathname = "mydata.txt";
delimiter = "\t";
hascolnames = 1;
T,N,M,isnum,colname,status = readtable(pathname,delimiter,hascolnames);
print("Table is %d by %d" % (N,M) );
for m in range(M):
    if( isnum[m] ):
        print("column %d (numeric) with colname <%s>" % (m,colname[m]) );
        c = T[m];
        for n in range(N):
            print( "%d %.2f" % (n, c[n,0]) );
    else:
        print("column %d (string) with colname <%s>" % (m,colname[m]) );
        c = T[m];
        for n in range(N):
            print( "%d <%s>" % (n, c[n]) );
```

--- Output of Example ----

```
Table is 10 by 3
column 0 (numeric) with colname <C0>
0 1.00
1 2.00
2 3.00
3 4.00
4 5.00
```

5 6.00
6 7.00
7 8.00
8 9.00
9 10.00
column 1 (string) with colname <C1>
0 <2>
1 <5>
2 <9>
3 <>
4 <A>
5 <3>
6 <B>
7 <5>
8 <1>
9 <0>
column 2 (numeric) with colname <C2>
0 4.00
1 4.00
2 4.00
3 5.00
4 4.00
5 4.00
6 4.00
7 4.00
8 4.00
9 6.00