

```

1  #!/usr/bin/env lua
2  -- vim : ft=lua et sts=2 sw=2 ts=2 :
3  -- Copyright (c) 2019,2020 Tim Menzies
4  -- All rights reserved.
5
6  -- Redistribution and use in source and binary forms, with or without
7  -- modification, are permitted provided that the following conditions are met:
8  -- 1. Redistributions of source code must retain the above copyright notice,
9  --    this list of conditions and the following disclaimer.
10 -- 2. Redistributions in binary form must reproduce the above copyright notice,
11 --    this list of conditions and the following disclaimer in the documentation
12 --    and/or other materials provided with the distribution.
13 --
14 -- THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS"
15 -- AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE
16 -- IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR PARTICULAR PURPOSE ARE
17 -- DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS BE LIABLE
18 -- FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL
19 -- DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR
20 -- SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER
21 -- CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY,
22 -- OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE
23 -- OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.
24
25 local b4={}; for k,_ in pairs(_ENV) do b4[k]=k end --used later (to find rogues)
26 local azzert,big,cli,fails,fmt,goalp,help,ignorep,klassp
27 local lessp,map,main,max,min,morep
28 local new,nump,o,oo,push,r,rows,slots,sort,sum,the,thing,things
29 local COLS, EGS, NUM, ROWS, SKIP, SOME, SYM = {}, {}, {}, {}, {}, {}, {}
30
31 function cli(want,x)
32   for n,got in ipairs(arg) do if got==want then
33     x = x==false and true or x==true and "false" or arg[n+1] end end
34   if x=="false" then return false else return tonumber(x) or x end end
35
36 help = [[
37 sl [OPTIONS]
38
39 OPTIONS:
40 -D      stack dump on assert fails      = false
41 -d F    data file                        = etc/data/auto93.csv
42 -h      show help                        = false
43 -k P    max kept items                   = 256
44 -S P    set seed                         = 10019
45 -t S    start up action (all= do all)    = nothing
46
47 KEY: F=filename P=posint S=string ]]
48
49 the = {dump = cli("-D", false),
50        data = cli("-d", ".etc/data/auto93.csv"),
51        help = cli("-h", false),
52        keep = cli("-k", 256 ),
53        seed = cli("-S", 10019),
54        todo = cli("-t", "nothing")}

```

```

55 --
56 -- FUNCTIONS
57 --
58 --
59 -- strings
60 fmt = string.format
61
62 -- maths
63 big = math.huge
64 max = math.max
65 min = math.min
66 r   = math.random
67
68 -- column headers
69 function klassp(x) return x:find("$" end
70 function lessp(x)  return x:find("-$" end
71 function morep(x)  return x:find"+"$" end
72 function nump(x)   return x:find"^[A-Z]" end
73 function ignorep(x) return x:find"$" end
74 function goalp(x)  return morep(x) or lessp(x) or klassp(x) end
75
76 -- tables
77 function push(t,x) table.insert(t,x); return x end
78 function sort(t,f) table.sort(t,f); return t end
79
80 -- meta
81 function new(k,t) k.__index=k; k.__tostring=o; return setmetatable(t,k) end
82 function map(t,f, u) u={};for k,v in pairs(t) do push(u,f(v)) end; return u end
83 function sum(t,f, n) n=0; for _,v in pairs(t) do n=n+f(v) end; return n end
84 function slots(t, u)
85   u={}
86   for k,v in pairs(t) do k=tostring(k);if k:sub(1,1)~="_" then push(u,k) end end
87   return sort(u) end
88
89 -- print tables, recursively
90 function oo(t) print(o(t)) end
91 function o(t)
92   if type(t)~="table" then return tostring(t) end
93   local key=function(k) return fmt("%s%s",k,o(t[k])) end
94   local u = #t>0 and map(t,o) or map(slots(t),key)
95   return '{'..table.concat(u, " " ).."}' end
96
97 -- strings to things
98 function thing(x)
99   x = x:match"^(%s*)(-)%s*$"
100   if x=="true" then return true elseif x=="false" then return false end
101   return tonumber(x) or x end
102
103 function things(x,sep, t)
104   t={}
105   for y in x:gmatch(sep or"([^\+])") do push(t,thing(y)) end
106   return t end
107
108 function rows(file, x)
109   file = io.input(file)
110   return function()
111     x=io.read(); if x then return things(x) else io.close(file) end end end
112
113 -- errors
114 fails=0
115 function azzert(test, msg)
116   print(test and "PASS: "or "FAIL: ",msg or "")
117   if not test then
118     fails=fails+1
119     if the.dump then assert(test,msg) end end end

```

```

120 -- CLASSES
121 --
122 --
123 --
124 -- SOME
125 function SOME.new(k,keep) return new(k,{n=0,_all={}, keep=keep or the.keep}) end
126 function SOME.add(i,x)
127   i.n = i.n+1
128   if #i._all < i.keep then push(i._all,x) ; return i._all
129   elseif #() < i.keep/i.n then i._all[r(#i._all)]=x; return i._all end end
130
131 -- SKIP
132 function SKIP.new(k,n,s) return new(k,{n=0,at=at or 0,txt=s or ""}) end
133 function SKIP.add(i,x) return x end
134
135 -- SYM
136 function SYM.new(k,n,s) return new(k,{n=0,at=n or 0,txt=s or "",has={}}) end
137 function SYM.add(i,x,inc)
138   if x ~= "?" then
139     inc = inc or 1
140     i.n = i.n + inc
141     i.has[x] = inc + (i.has[x] or 0) end end
142 function SYM.dist(i,x,y)
143   return (x=="?" and y=="?" and 1) or (x==y and 0 or 1) end
144
145 -- NUM
146 function NUM.new(k,n,s)
147   return new(k,{n=0,at=n or 0,txt=s or "",has=SOME:new(),
148     w=lessp(s or "") and -1 or 1, lo=big, hi=-big}) end
149 function NUM.add(i,x)
150   if x ~= "?" then
151     i.n = i.n + 1
152     i.has:add(x); i.lo,i.hi = min(x,i.lo), max(x,i.hi); end end
153 function NUM.norm(i,x)
154   return math.abs(i.hi-i.lo)<1E-9 and 0 or (x-i.lo)/(i.hi - i.lo) end
155 function NUM.dist(i,x,y)
156   if x=="?" and y=="?" then return 1
157   elseif x=="?" then y=i.norm(y); x=y<0.5 and 1 or 0
158   elseif y=="?" then x=i.norm(x); y=x<0.5 and 1 or 0
159   else x,y = i.norm(x), i.norm(y) end
160   return math.abs(x-y) end
161
162 -- COLS
163 function COLS.new(k,row, i)
164   i = new(k,{all={},x={},y={}})
165   for at,txt in ipairs(row) do push(i.all, i:col(at,txt)) end
166   return i end
167 function COLS.add(i,t)
168   for _,col in pairs(i.all) do col:add( t[col.at] ) end
169   return t end
170 function COLS.col(i,at,txt, col)
171   if ignorep(txt) then return SKIP:new(at,txt) end
172   col = (nump(txt) and NUM or SYM):new(at,txt)
173   push(goalp(txt) and i.y or i.x, col)
174   if klassp(txt) then i.klass = col end
175   return col end
176
177 -- ROWS
178 function ROWS.new(k,init, i)
179   i = new(k,{rows=SOME:new(), cols=nil})
180   if type(init)=="string" then for row in rows(init) do i:add(row) end end
181   if type(init)=="table" then for row in init do i:add(row) end end
182   return i end
183 function ROWS.add(i,row)
184   if i.cols then i.rows:add( i.cols:add(row) )
185   else i.cols = COLS:new(row) end end
186 function ROWS.dist(i,row1,row2, d)
187   function d(col) return col:dist(row1[col.at], row2[col.at])^the.p end
188   return (sum(i.cols.x, d)/ #i.cols.x)^(1/the.p) end

```

```

189 -- DEMOS
190 --
191 --
192 --
193 function EGS.nothing() return true end
194 function EGS.the() oo(the) end
195 function EGS.rand() print(r()) end
196 function EGS.f1() print(1) end
197 function EGS.f2() print(2) end
198
199 -- Start-up
200 if the.help then print(help) else
201   local b4={}; for k,v in pairs(the) do b4[k]=v end
202   for _,todo in pairs(the.todo=="all" and slots(EGS) or {the.todo}) do
203     for k,v in pairs(b4) do the[k]=v end
204     math.randomseed(the.seed)
205     if type(EGS[todo])=="function" then EGS[todo]() end end end
206
207 for k,v in pairs(_ENV) do if not b4[k] then print("?",k,type(v)) end end
208 os.exit(fails)

```