

```

1  #!/usr/bin/env lua
2  local b4={}; for k,_ in pairs(_ENV) do b4[k]=k end --used later (to find rogues)
3  local azzert,big,cli,fails,fmt,goalp,help,ignorep,klassp
4  local lessp,map,main,max,min,morep
5  local new,nump,o,oo,push,r,rows,slots,sort,sum,the,thing,things
6  local COLS, NUM, ROWS, SKIP, SOME, SYM = {}, {}, {}, {}, {}
7  -----
8  function cli(want,x)
9      for n,got in ipairs(arg) do if got==want then
10         x = x==false and true or x==true and "false" or arg[n+1] end end
11         if x=="false" then return false else return tonumber(x) or x end end
12
13  help = [[
14  ./sl.lua [OPTIONS]
15
16  OPTIONS:
17      -D      stack dump on assert fails      = false
18      -d F    data file                       = etc/data/auto93.csv
19      -h      show help                       = false
20      -k P    max kept items                  = 256
21      -S P    set seed                        = 10019
22      -t S    start up action (all= do all)    = nothing
23
24  KEY: F=filename P=posint S=string
25  ]]
26
27  the = {dump = cli("-D", false),
28         data = cli("-d",  "etc/data/auto93.csv"),
29         help = cli("-h",  false),
30         keep = cli("-k",  256 ),
31         seed = cli("-S",  10019),
32         todo = cli("-t",  "nothing")}
33  -----
34  --
35  --
36  --
37  --
38  --
39  --- strings
40  fmt = string.format
41
42  --- maths
43  big = math.huge
44  max = math.max
45  min = math.min
46  r = math.random
47
48  --- column headers
49  function klassp(x) return x:find("$" end
50  function lessp(x) return x:find"$" end
51  function morep(x) return x:find"+$" end
52  function nump(x) return x:find"[A-Z]" end
53  function ignorep(x) return x:find"$" end
54  function goalp(x) return morep(x) or lessp(x) or klassp(x) end
55
56  --- tables
57  function push(t,x) table.insert(t,x); return x end
58  function sort(t,f) table.sort(t,f); return t end
59
60  --- meta
61  function new(k,t) k.__index=k; k.__tostring=o; return setmetatable(t,k) end
62  function map(t,f, u) u={};for k,v in pairs(t) do push(u,f(v)) end; return u end
63  function sum(t,f, n) n=0; for _,v in pairs(t) do n=n+f(v) end; return n end
64  function slots(t, u)
65      u={}
66      for k,v in pairs(t) do k=tostring(k);if k:sub(1,1)~="_" then push(u,k) end end
67      return sort(u) end
68
69  --- print tables, recursively
70  function oo(t) print(o(t)) end
71  function o(t)
72      if type(t)~="table" then return tostring(t) end
73      local key=function(k) return fmt("%.8s %s",k,o(t[k])) end
74      local u = #t>0 and map(t,o) or map(slots(t),key)
75      return ('%.table.concat(u," ")..")" end
76
77  --- strings to things
78  function thing(x)
79      x = x:match"%s*(-)%s*$"
80      if x=="true" then return true elseif x=="false" then return false end
81      return tonumber(x) or x end
82
83  function things(x,sep, t)
84      t={}
85      for y in x:gmatch(sep or"([^\,]+)") do push(t,thing(y)) end
86      return t end
87
88  function rows(file, x)
89      file = io.input(file)
90      return function()
91          x=io.read(); if x then return things(x) else io.close(file) end end end
92
93  --- errors
94  fails=0
95  function azzert(test, msg)
96      print(test and "PASS: "or "FAIL: ",msg or "")
97      if not test then
98          fails=fails+1
99      if the.dump then assert(test,msg) end end end

```

```

100 --
101 --
102 --
103 --
104 function SOME.new(k,keep) return new(k,{n=0,_all={}, keep=keep or the.keep}) end
105 function SOME.add(i,x)
106     i.n = i.n+1
107     if #i._all < i.keep then push(i._all,x) ; return i._all
108     elseif r() < i.keep/i.n then i._all[r(#i._all)]=x; return i._all end end
109
110 --
111 --
112 --
113 function SKIP.new(k,n,s) return new(k,{n=0,at=at or 0,txt=s or""}) end
114 function SKIP.add(i,x) return x end
115
116 --
117 --
118 --
119 function SYM.new(k,n,s) return new(k,{n=0,at=n or 0,txt=s or"",has={}}) end
120 function SYM.add(i,x,inc)
121     if x == "?" then
122         inc = inc or 1
123         i.n = i.n + inc
124         i.has[x] = inc + (i.has[x] or 0) end end
125 function SYM.dist(i,x,y)
126     return (x=="?" and y=="?" and 1) or (x==y and 0 or 1) end
127
128 --
129 --
130 --
131 function NUM.new(k,n,s)
132     return new(k,{n=0,at=n or 0,txt=s or"",has=SOME.new(),
133                  w=lessp(s or "") and -1 or 1, lo=big, hi=big}) end
134 function NUM.add(i,x)
135     if x == "?" then
136         i.n = i.n + 1
137         i.has.add(x); i.lo,i.hi = min(x,i.lo), max(x,i.hi); end end
138 function NUM.norm(i,x)
139     return math.abs(i.hi-i.lo)<1E-9 and 0 or (x-i.lo)/(i.hi - i.lo) end
140 function NUM.dist(i,x,y)
141     if x=="?" and y=="?" then return 1
142     elseif x=="?" then y=i.norm(y); x=y<0.5 and 1 or 0
143     elseif y=="?" then x=i.norm(x); y=x<0.5 and 1 or 0
144     else x,y = i.norm(x), i.norm(y) end
145     return math.abs(x-y) end
146
147 --
148 --
149 --
150 function COLS.new(k,row, i)
151     i = new(k,{all={},x={},y={}})
152     for at,txt in ipairs(row) do push(i.all, i:col(at,txt)) end
153     return i end
154 function COLS.add(i,t)
155     for _,col in pairs(i.all) do col:add( t[col.at] ) end
156     return t end
157 function COLS.col(i,at,txt, col)
158     if ignorep(txt) then return SKIP:new(at,txt) end
159     col = (nump(txt) and NUM or SYM):new(at,txt)
160     push(goalp(txt) and i.y or i.x, col)
161     if klassp(txt) then i.klass = col end
162     return col end
163
164 --
165 --
166 --
167 function ROWS.new(k,init, i)
168     i = new(k,{rows=SOME:new(), cols=nil})
169     if type(init)=="string" then for row in rows(init) do i:add(row) end end
170     if type(init)=="table" then for row in init do i:add(row) end end
171     return i end
172 function ROWS.add(i,row)
173     if i.cols then i.rows:add( i.cols:add(row) )
174     else i.cols = COLS:new(row) end end
175 function ROWS.dist(i,row1,row2, d)
176     function d(col) return col:dist(row1[col.at], row2[col.at])^the.p end
177     return (sum(i.cols.x, d)/ #i.cols.x)^(1/the.p) end
178
179 -----
180 --
181 --
182 --
183 local eggs={}
184 function eggs.nothing() return true end
185 function eggs.the() oo(the) end
186 function eggs.rand() print(r()) end
187 function eggs.f1() print(1) end
188 function eggs.f2() print(2) end
189
190 if the.help then print(help) else
191     local b4={}; for k,v in pairs(the) do b4[k]=v end
192     for _,todo in pairs(the.todo=="all" and slots(eggs) or (the.todo)) do
193         for k,v in pairs(b4) do the[k]=v end
194         math.randomseed(the.seed)
195         if type(eggs[todo])=="function" then eggs[todo]() end end end
196
197 for k,v in pairs(_ENV) do if not b4[k] then print("?",k,type(v)) end end
198 os.exit(fails)

```