

```

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  --- strings
39  fmt = string.format
40
41  --- maths
42  big = math.huge
43  max = math.max
44  min = math.min
45  r   = math.random
46
47  --- column headers
48  function klassp(x) return x:find"!$" end
49  function lessp(x)  return x:find"$" end
50  function morep(x)  return x:find"+$" end
51  function nump(x)   return x:find"^[A-Z]" end
52  function ignorep(x) return x:find":$" end
53  function goalp(x)  return morep(x) or lessp(x) or klassp(x) end
54
55  --- tables
56  function push(t,x) table.insert(t,x); return x end
57  function sort(t,f) table.sort(t,f); return t end
58
59  --- meta
60  function new(k,t) k.__index=k; k.__tostring=o; return setmetatable(t,k) end
61  function map(t,f, u) u={};for k,v in pairs(t) do push(u,f(v)) end; return u end
62  function sum(t,f, n) n=0; for _,v in pairs(t) do n+=f(v) end; return n end
63  function slots(t, u)
64      u={}
65      for k,v in pairs(t) do k=tostring(k);if k:sub(1,1)~="_" then push(u,k) end end
66      return sort(u) end
67
68  --- print tables, recursively
69  function oo(t) print(o(t)) end
70  function o(t)
71      if type(t)~="table" then return tostring(t) end
72      local key=function(k) return fmt(":%s %s",k,o(t[k])) end
73      local u = #t>0 and map(t,o) or map(slots(t),key)
74      return '{ '..table.concat(u, " ")...' }' end
75
76  --- strings to things
77  function thing(x)
78      x = x:match"^(%s*)(.*)$"
79      if x=="true" then return true else if x=="false" then return false end
80      return tonumber(x) or x end
81
82  function things(x,sep, t)
83      t={}
84      for y in x:gmatch(sep or "[^,]+") do push(t,thing(y)) end
85      return t end
86
87  function rows(file, x)
88      file = io.input(file)
89      return function()
90          x=io.read(); if x then return things(x) else io.close(file) end end end
91
92  --- errors
93  fails=0
94  function azzert(test, msg)
95      print(test and "PASS: "or "FAIL: ",msg or "")
96      if not test then
97          fails=fails+1
98          if the.dump then assert(test,msg) end end end
99

```

```

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  --
132  function NUM.new(k,n,s)
133      return new(k,{n=0,at=n or 0,txt=s or "",has=SOME:new(),
134                  w=lessp(s or "") and -1 or 1, lo=big, hi=-big}) end
135  function NUM.add(i,x)
136      if x ~="?" then
137          i.n = i.n + 1
138          i.has:add(x); i.lo,i.hi = min(x,i.lo), max(x,i.hi); end end
139  function NUM.norm(i,x)
140      return math.abs(i.hi-i.lo)<1E-9 and 0 or (x-i.lo)/(i.hi - i.lo) end
141  function NUM.dist(i,x,y)
142      if x=="?" and y=="?" then return 1
143      elseif x=="?" then y=i.norm(y); x=y<0.5 and 1 or 0
144      elseif y=="?" then x=i.norm(x); y=x<0.5 and 1 or 0
145      else x,y = i.norm(x), i.norm(y) end
146      return math.abs(x-y) end
147  --
148  --
149  --
150  --
151  function COLS.new(k,row, i)
152      i = new(k,{all={},x={},y={}})
153      for at,txt in ipairs(row) do push(i.all, i:col(at,txt)) end
154      return i end
155  function COLS.add(i,t)
156      for _,col in pairs(i.all) do col:add( t[col.at] ) end
157      return t end
158  function COLS.col(i,at,txt, col)
159      if ignorep(txt) then return SKIP:new(at,txt) end
160      col = (nump(txt) and NUM or SYM):new(at,txt)
161      push(goalp(txt) and i.y or i.x, col)
162      if klassp(txt) then i.klass = col end
163      return col end
164  --
165  --
166  --
167  --
168  function ROWS.new(k,init, i)
169      i = new(k,{rows=SOME:new(), cols=nil})
170      if type(init)=="string" then for row in rows(init) do i:add(row) end end
171      if type(init)=="table" then for row in init do i:add(row) end end
172      return i end
173  function ROWS.add(i,row)
174      if i.cols then i.rows:add( i.cols:add(row) )
175      else i.cols = COLS:new(row) end end
176  function ROWS.dist(i,row1,row2, d)
177      function d(col) return col:dist(row1[col.at], row2[col.at])^the.p end
178      return (sum(i.cols.x, d)/ #i.cols.x)^(1/the.p) end
179  -----
180  --
181  --
182  --
183  --
184  local egs={}
185  function egs.nothing() return true end
186  function egs.the() oo(the) end
187  function egs.f1() print(1) end
188  function egs.f2() print(2) end
189
190  if the.help then print(help)
191  else local b4={}; for k,v in pairs(the) do b4[k]=v end
192      for _,todo in pairs(the.todo=="all" and slots(egs) or {the.todo}) do
193          for k,v in pairs(b4) do the[k]=v end
194          math.randomseed(the.seed)
195          if type(egs[todo])=="function" then egs[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)

```