

```

1 local b4={}; for k,_ in pairs(_ENV) do b4[k]=k end
2 local all,any,firsts,new,many,map,o,push
3 local rows,seconds,slots,sort,thing,things
4 local EGS, NUM, SYM = {}, {}, {}
5 -----
6 function NUM.new(i,at,s)
7     return new(i,{at=at,txt=s,w=s:find("-" and -1 or 1,all={},
8         ok=false, lo=math.huge, hi=-math.huge)}) end
9
10 function NUM.add(i,x)
11     if x ~= "?" then
12         i.ok = false
13         i.all[1 + #i.all] = x
14         if x < i.lo then i.lo = x end
15         if x > i.hi then i.hi = x end end
16     return x end
17
18 function NUM.dist(i,a,b)
19     if a=="?" and b=="?" then a,b=1,0
20     elseif a=="?" then b=norm(num,b); a=b>.5 and 0 or 1
21     elseif b=="?" then a=norm(num,a); b=a>.5 and 0 or 1
22     else a, b=norm(num,a), norm(num,b) end
23     return math.abs(a-b) end
24
25 function NUM.norm(i,x)
26     return i.hi - i.lo<1E-9 and 0 or (x - i.lo)/(i.hi - i.lo) end
27 -----
28 function SYM.new(i,at,s)
29     return new(i,{at=at,txt=s,_all={}}) end
30
31 function SYM.add(i,x)
32     if x ~= "?" then i._all[x] = 1+(i._all[x] or 0) end
33     return x end
34
35 function SYM.all(i)
36     if not i.ok then sort(i._all); i.ok=true end; return i._all end
37
38 function SYM.dist(i,a,b)
39     return a=="?" and b=="?" and 1 or a==b and 0 or 1 end
40 -----
41 function EGS.new(i)
42     return new(i,{rows={}, head=nil, all={}, x={}, y={}}) end
43
44 function EGS.add(i,t)
45     local add,now = function(col) return col:add(t[col.at]) end
46     if i.head
47     then i.rows[1+#i.rows] = map(i.all,add)
48     else i.head=t
49         for n,x in pairs(t) do
50             now = push(i.all, (x:find("[A-Z]" and NUM or SYM):new(n,x))
51             if not x:find"." then
52                 push((x:find"+" or x:find"-") and i.y or i.x,now) end end end end
53
54 function EGS.clone(i,init, j)
55     j = EGS:new()
56     j:add(i.head)
57     for _,row in pairs(init or {}) do j = egs1(j, row) end
58     return j end
59
60 function EGS.dist(i,r1,r2, d,n,norm)
61     d,n = 0, (#i.x)+1E-31
62     for _,col in pairs(i.x) do
63         inc = col:dist(r1[col.at], r2[col.at])
64         d = d + inc^2 end
65     return (d/n)^.5 end
66
67 function EGS.far(i,r1,rows, fun,tmp)
68     fun = function(r2) return {r2, i:dist(r1,r2)} end
69     tmp = sort(map(rows,fun), seconds)
70     print(1)
71     print(o(tmp))
72     return table.unpack(tmp[#tmp*.9//1] ) end
73
74 function EGS.half(i,rows)
75     local some,nth,sth,c,cosine,ls,rs
76     rows = rows or i.rows
77     some = #rows > 512 and many(rows,512) or rows
78     nth = i:far(any(rows), some)
79     sth,c = i:far(nth, some)
80     function cosine(r, a,b)
81         a,b = i:dist(r,nth),i:dist(r,sth);return {(a^2+c^2-b^2)/(2*c),r} end
82     ls,rs = i:clone(), i:clone()
83     for n,pair in pairs(sort(map(rows,cosine), firsts)) do
84         egs1(n <= #rows//2 and ls or rs, pair[2]) end
85     return ls,rs,l,r,c end
86 -----
87
88 function any(t) return t[math.random(#t)] end
89
90 function firsts(a,b) return a[1] < b[1] end
91
92 function many(t,n, u) u={};for j=1,n do t[1+#t]=any(t) end; return u end
93
94 function map(t,f,u) u={};for _,v in pairs(t) do u[1+#u]=f(v) end; return u end
95
96 function new(k,t) k._index=k; return setmetatable(t,k) end
97
98 function o(t, u)
99     if type(t)~="table" then return tostring(t) end
100     local key=function(k) return string.format("%.5s",k,t[k]) end
101     u = #t>0 and map(t,o) or map(sort(slots(t)),key)
102     return "{..table.concat(u,"")..}" end
103
104 function push(t,x) table.insert(t,x); return x end
105
106 function rows(file, x)
107     file = io.input(file)
108     return function()
109         x=io.read(); if x then return things(x) else io.close(file) end end end
110
111 function slots(t, u) u={};for k,_ in pairs(t) do u[1+#u]=k end; return u end
112
113 function sort(t,f) table.sort(t,f); return t end
114
115 function seconds(a,b) return a[2] < b[2] end
116
117 function thing(x)
118     x = x:match("%s*(-)%s*$")
119     if x=="true" then return true elseif x=="false" then return false end
120     return tonumber(x) or x end
121
122 function things(x,sep, t)
123     t={};for y in x:gmatch(sep or "[^,]+") do t[1+#t]=thing(y) end; return t end
124 -----
125 --for row in rows("../data/auto93.csv") do print(o(row)) end
126 local i=EGS:new()
127 i:half()
128 for row in rows("../data/auto93.csv") do i:add(row) end
129 for k,v in pairs(_ENV) do if not b4[k] then print("?",k,type(v)) end end

```