

```

1  #!/bin/bash
2  files="/home/wayne/pub/cs146/awkcel"
3
4  # 1(a)
5  # command: use code from 1(b) to generate table, output to file named
6  # "orthologs_count.tsv"
7  # then use the following commands to find most/least ortholog pair
8  # least
9  tail -n +2 orthologs_count.tsv | sort -k 3 -n | head -n 1
10 # most
11 tail -n +2 orthologs_count.tsv | sort -k 3 -nr | head -n 1
12
13 # answer:
14 # most orthologs: rat and mouse
15 # least orthologs: yeast and alpaca
16
17 # 1(b)
18 # command that creates a table of the number of orthologs for every pair
19 species=(`echo $(head -n 1 "$files/orthologs.tsv") | tr "\t" " ")
20 num_species=${#species[@]}
21 printf "first\tsecond\torthologs\n"
22 for (( i=0; i<$num_species; i++));
23 do
24     for (( j=$((i+1)); j<$num_species; j++));
25     do
26         first=${species[$i]}
27         second=${species[$j]}
28         printf "%s\t%s\t" $first $second
29         ./awkcel "$first!=\"_\" && $second!=\"_\" {count++} END{print count}"
30         "$files/orthologs.tsv"
31     done
32 done
33
34 # output: see "orthologs_count.tsv"
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67

```

```

68 # 2(a)
69 # command:
70 ./awkcel 'date=="2011-07-01" {printf "Jul 1 2011:\nAMZN:%s\nMSFT:%s\n", AMZN, MSFT}
71 date=="2011-07-05" {printf "Jul 5 2011:\nAMZN:%s\nMSFT:%s\n", AMZN, MSFT}' \
72 "$files/historical.2011.tsv"
73
74 # closing prices:
75 # Jul 1 2011:
76 #   AMZN:209.49
77 #   MSFT:26.02
78 # Jul 5 2011:
79 #   AMZN:213.19
80 #   MSFT:26.03
81 # AMZN went up, MSFT also went up (barely)
82
83 # 2(b)
84 # command
85 ./awkcel '
86 BEGIN{prev=0; worst=0}
87 {change=AMZN-prev; prev=AMZN; if(change < worst){worst=change; day=date}}
88 END{printf "worst day: %s\nchange: %.2f\n", day, worst}
89 ' "$files/historical.2011.tsv"
90 # answer
91 # worst day: 2011-10-26
92 # change: -29.07
93
94 # 2(c)
95 # command
96 ./awkcel '
97 BEGIN{first=1;worst=0}
98 {
99     if (first==1) {
100         first=0
101         for (i=2; i<=NF; ++i) {
102             prev[i] = $i # first day, just log the prices
103         }
104     }
105     for (i=2; i<=NF; ++i) {
106         if ($i != "_" && prev[i] != "_") {
107             change = $i/prev[i] * 100 - 100 # percent change
108             if (change < worst) {
109                 worst = change
110                 day = date
111                 company = i # just storing the column number of that company
112             }
113         }
114         prev[i] = $i
115     }
116 }
117 END{printf "worst single day loss:\ncompany(column): %s\nday: %s\nchange:
%.2f%%\n", company, day, worst}
118 ' "$files/historical.2011.tsv"
119 # output:
120 # worst single day loss:
121 # company(column): 203
122 # day: 2011-06-16
123 # change: -66.70%
124 # finding out which company is at column 203
125 column=203
126 cat "$files/historical.2011.tsv" | gawk -F"\t" "{print $column; exit}"
127 # output: CSX
128
129 # final answer:
130 # worst single day loss:
131 # company: CSX
132 # day: 2011-06-16
133
134
135

```

```

136 # Spiral Galaxies
137 # command
138 ./awkcel '
139 BEGIN{one=0;zero=0;half=0}
140 {
141     P_SP = P_CW + P_ACW
142     if (P_SP == 1) {++one}
143     if (P_SP == 0) {++zero}
144     if (P_SP > 0.5) {++half}
145 }
146 END{ printf "number of galaxies that has P_SP:\nexactly 1: %d\nexactly 0: %d\n>0.5: %d\n", one, zero, half }
147 ' "$files/SDSS+GZ1+SpArcFiRe+SFR.tsv"
148 # answer
149 # number of galaxies that has P_SP:
150 # exactly 1: 3943
151 # exactly 0: 165832
152 # >0.5: 66906
153
154 # histogram
155 ./awkcel '{
156     P_SP=P_CW+P_ACW;
157     bin=int(10*P_SP);
158     n[bin]++;
159     numArcs[bin]+=totalArcLength;
160 }
161 END{
162     for(i=0;i<=10;i++)
163     print i, n[i], numArcs[i]/n[i];
164 }' "$files/SDSS+GZ1+SpArcFiRe+SFR.tsv"
165
166 # output:
167 # 0 525463 192.035
168 # 1 94299 329.484
169 # 2 39115 449.432
170 # 3 22636 510.712
171 # 4 15938 544.275
172 # 5 13520 565.314
173 # 6 11583 589.788
174 # 7 10382 614.252
175 # 8 12503 640.813
176 # 9 15490 691.263
177 # 10 4437 730.696
178

```