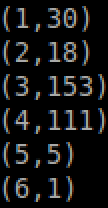
Task 1

1.the frequency of each degree value

D=group C by freq;

E=foreach D generate $0, COUNT($1);



2.the percentage of dead-end nodes

D=join A by nodeB left outer, B by $0;

E=filter D by $2 is null;

F=group E by $1;

G=group F all;

H=group B all;

I= foreach H generate (double)COUNT($1) as , (double)COUNT (G.$1) as b;

J=foreach I generate (float)((a/a+b));



3.he average degree of the graph

D=group C by freq;

E=foreach D generate $0 as deg, COUNT($1) as frequ, $0\*COUNT($1) as total;

F=group E all;

G= foreach E generate SUM(E.frequ) as a, SUM(E.total) as b;

avg=foreach G generate b/a;

dump avg;





Triangle Counting