Reading the File:

1. Go to your terminal
2. Run this command: python3 triads.py --filename epinions69.csv
3. Make sure epinions69.csv will be in same folder as triads.py

Library Used:

* Tabulate
* Combinations
* Networkx

Problem: Calculate the following:

1.Number of edges in the network

2.Number of self-loops

3.Number of edges used to identify triads(referred to as TotEdges) [ this should be a. –b. ]

4.Number of positive (trust) edges (ignore self-loops)

5.Number of negative (distrust) edges(ignore self-loops)

6.Probability p that an edge will be positive: (number of positive edges) / TotEdges

7.Probability that an edge will be negative: 1 –p

**Approach Used**: To calculate all these problem from **1 to 5**, I am creating a function called graph\_analysis which is reading and iterating through the input file and splitting the reviewer and reviewee and trust and distrust values.

**6-7**: To calculate the probability for trust and distrust edges, just used the given formula.

8.Expected distribution of triad types (based on p and 1 –p applied to the number of triangles in the graph). Show number and percent.

a. Trust-Trust-Trust

b. Trust-Trust-Distrust

c. Trust-Distrust -Distrust

d. Distrust-Distrust-Distrust

e. Total

**Approach Used:**

**For calculating percentage:** Using probabilities that we calculated in 6-7, based on this we also calculated probabilities of TTT, TTD, TDD, DDD

**For calculating numbers:** For numbers I multiplied the above calculating percentages with the number of triangles present in the epinions file.

To show in the tabular form I used **tabulate** library, which can be installed using command: pip install tabulate.

9. Actual distribution of triad types. Show number and percent.

a. Trust-Trust-Trust

b. Trust-Trust-Distrust

c. Trust-Distrust -Distrust

d. Distrust-Distrust-Distrust

e. Total

**Approach Used:**

**For calculating percentage:** Apply this formula total count of the triangles in triad type by total number of triangles

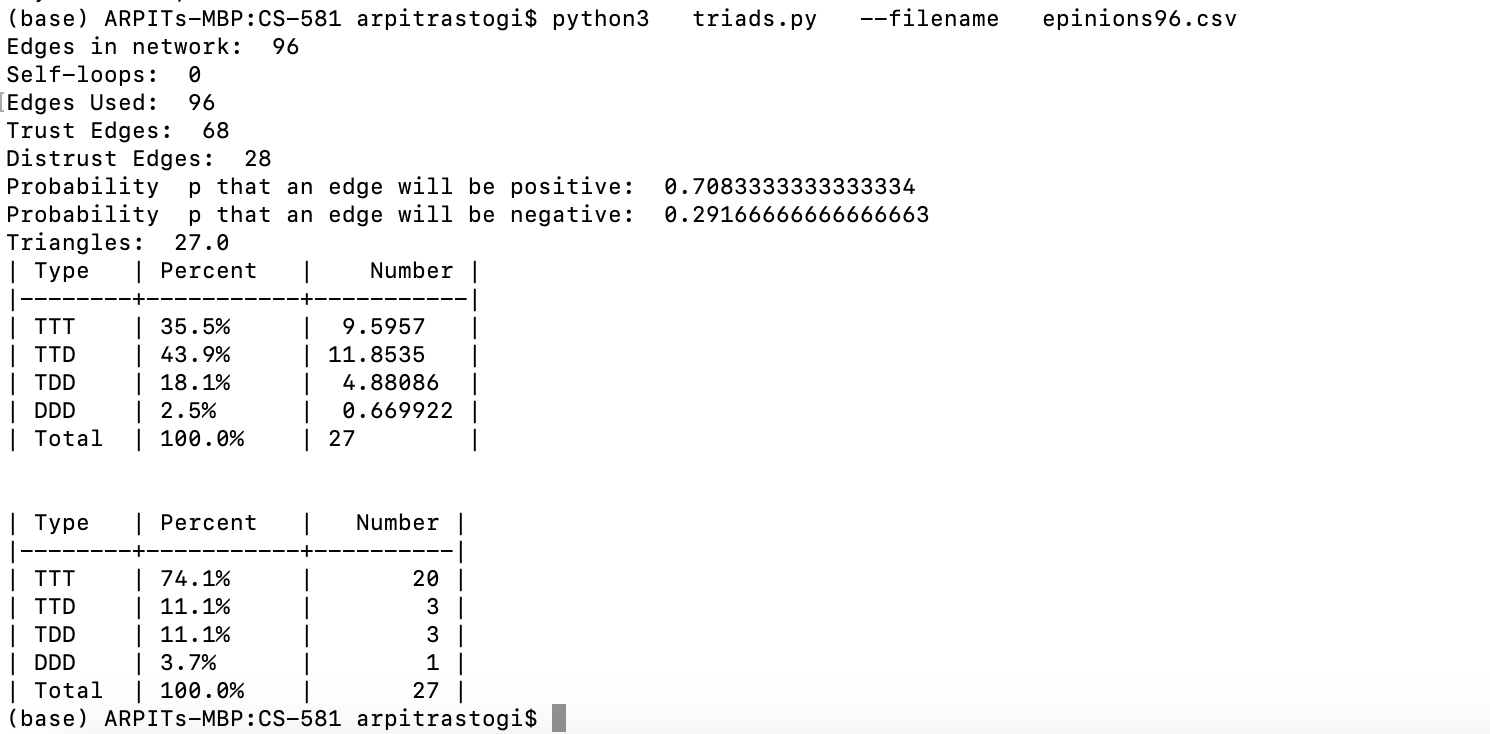
**For calculating numbers:** First calculated the total triangles using networkx.enumerate\_all\_cliques() and from that created a list of triangles (cliques of length 3), which I then processed one triangle at a time to find the type using the 3 edges and trust values.

**Challenges:** All of the questions from 1 to 7 are straight forward. So, didn’t take time to implement it.

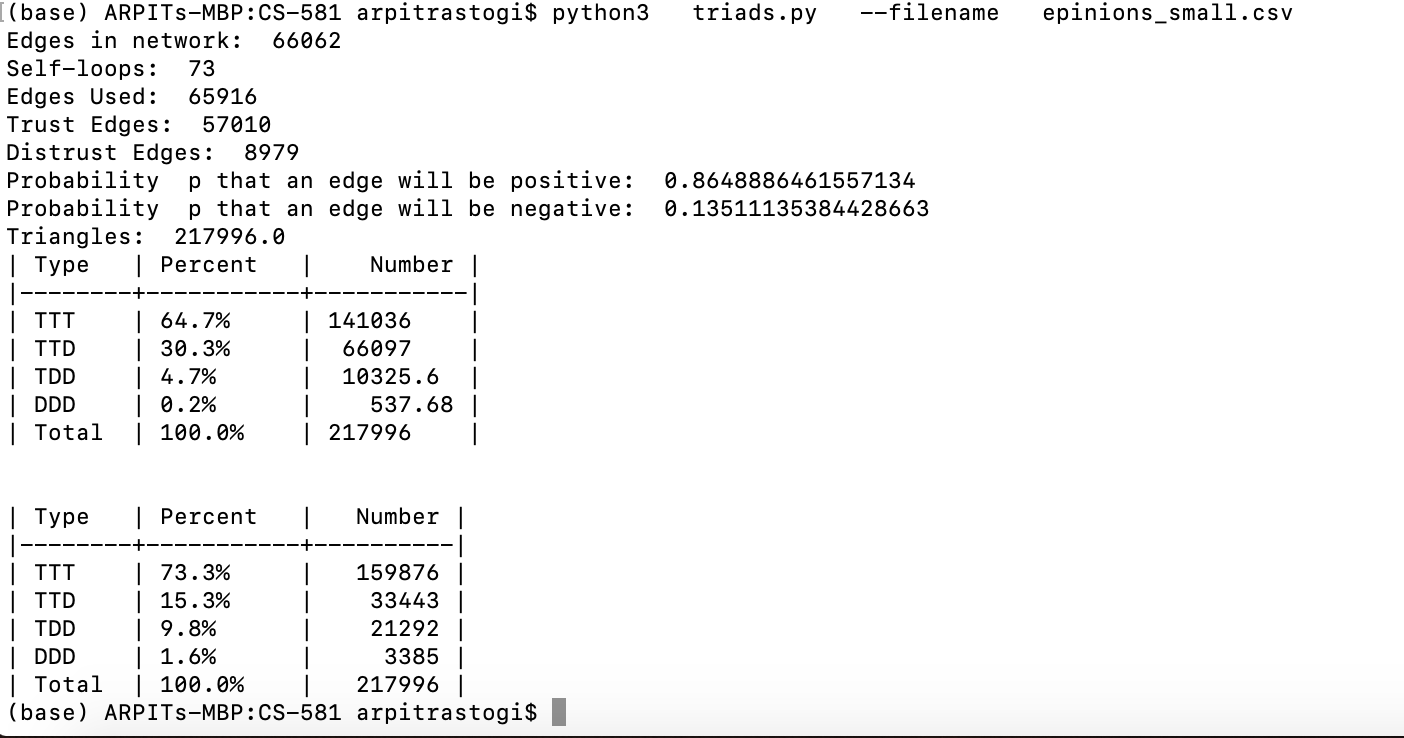
But the 8 question about finding Actual Distribution was little tricky for which I needed professor help.

**Difference between Actual and expected Value** is that expected values are calculated based on the probability that’s why there is a difference.

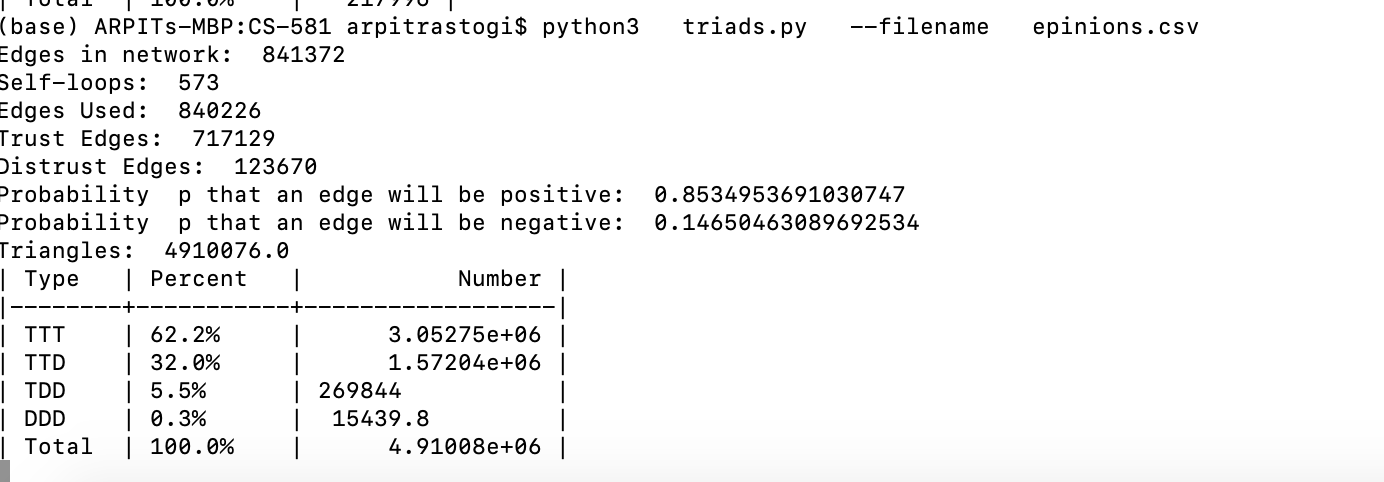
**Screenshot for epinions69.csv**

****

**Screenshot for epinions\_small.csv**

****

**Screenshot for epinions.csv**



This file tooks longer than 10 mins so I aborted the process.