Networks Assignment 3

Question 1:

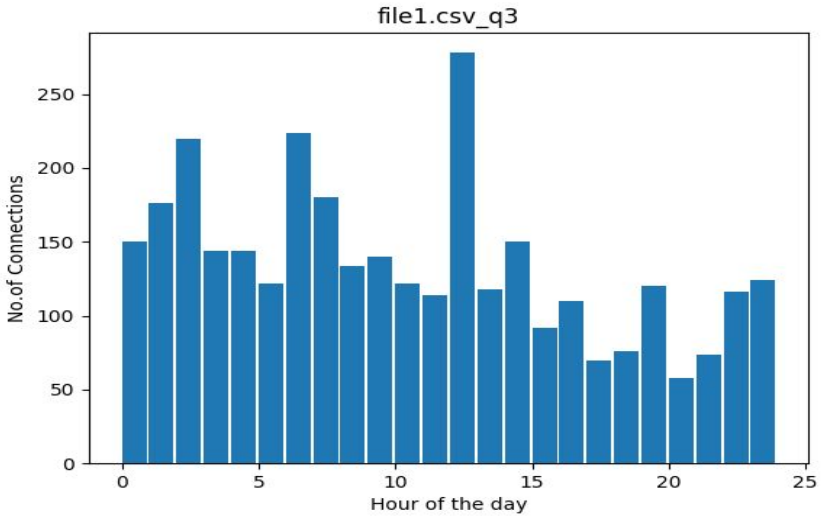
	No. of servers	No. of clients
File 1(03-01-11)	45	522
File 2(03-01-14)	50	939
File 3(03-01-18)	89	510

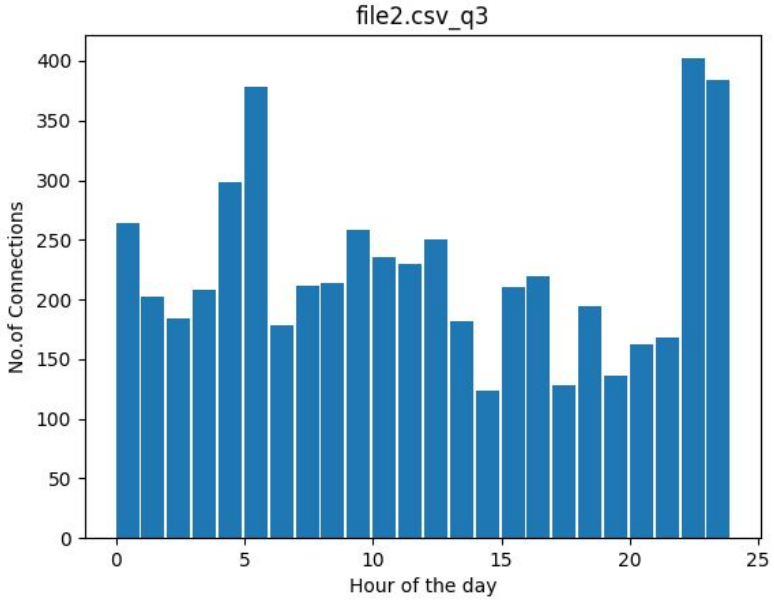
Question 2:

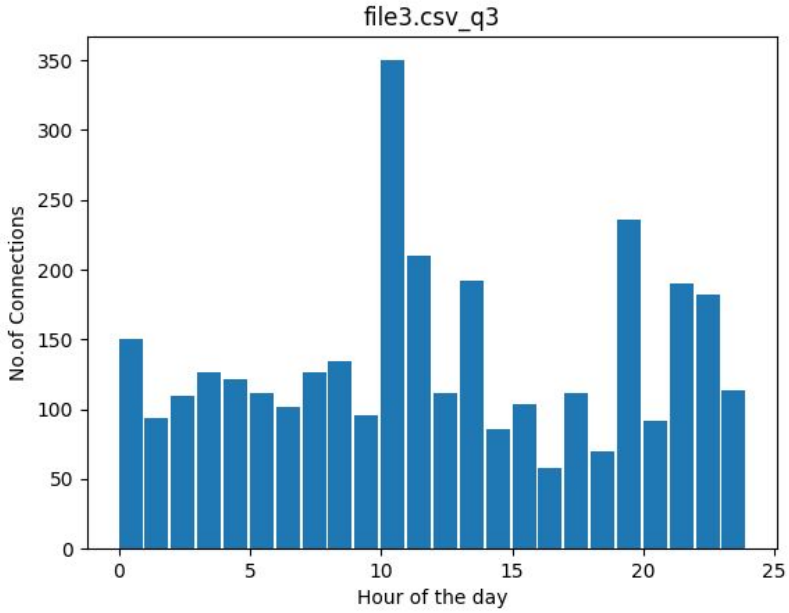
	No. of TCP flows
File 1	3256
File 2	5422
File 3	3280

Question 3:

The plots of No. of flows vs hour of the day for all 3 files are shown below.







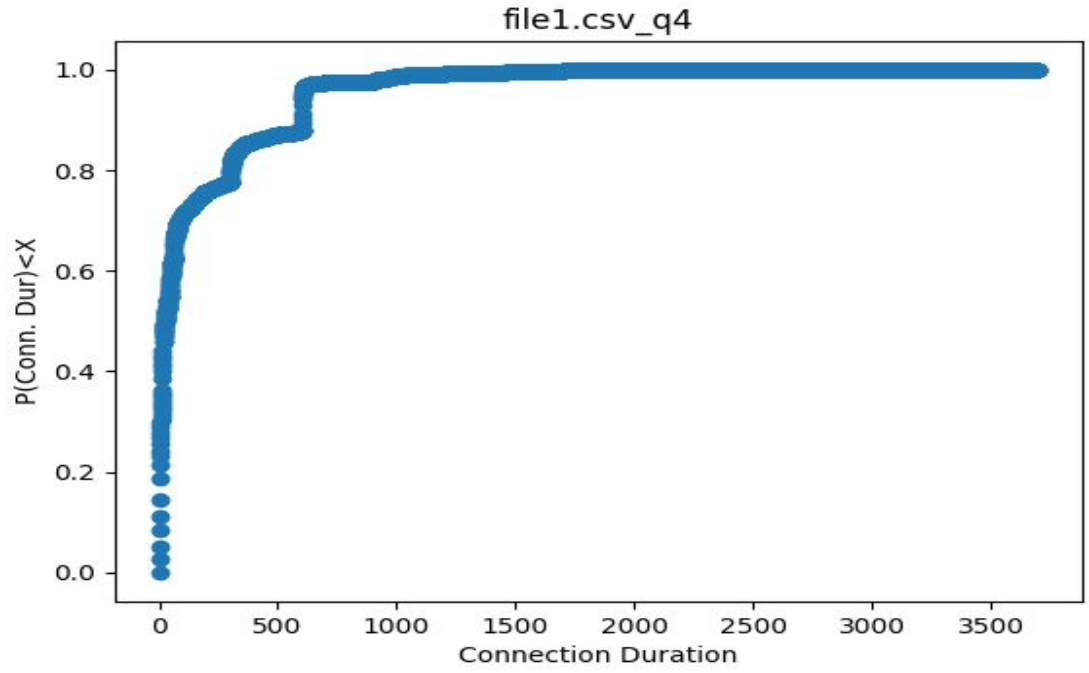
We can use these traffic profiles to detect if a system is under DoS attack.

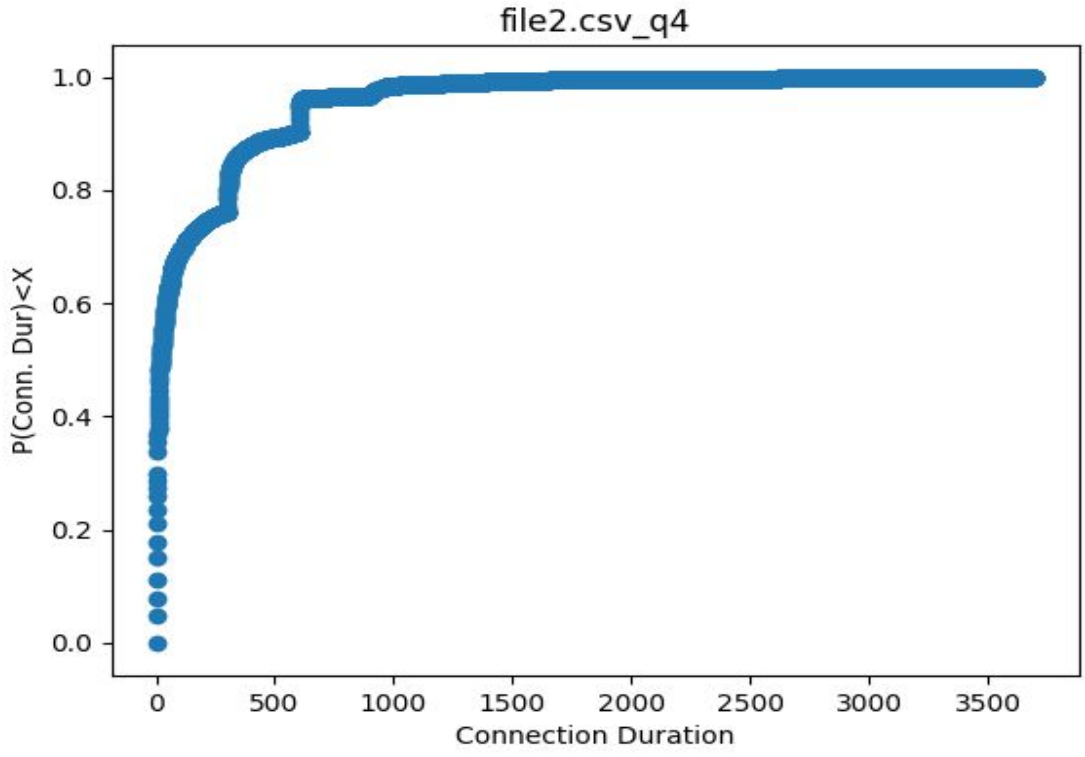
If a system is under DoS attack, The attacker establishes a large number of half-open or fully open TCP connections at the target host. The host can become so bogged down with these bogus connections that it stops accepting legitimate connections. So, if the system is under the attack, No. of TCP connections abruptly rises. As we can see in the plots of file 1(at hour 12), file 2(at hour 22) and file 3(at hour 10, 19) the no. of TCP connections suddenly increased(by 2-3 times).

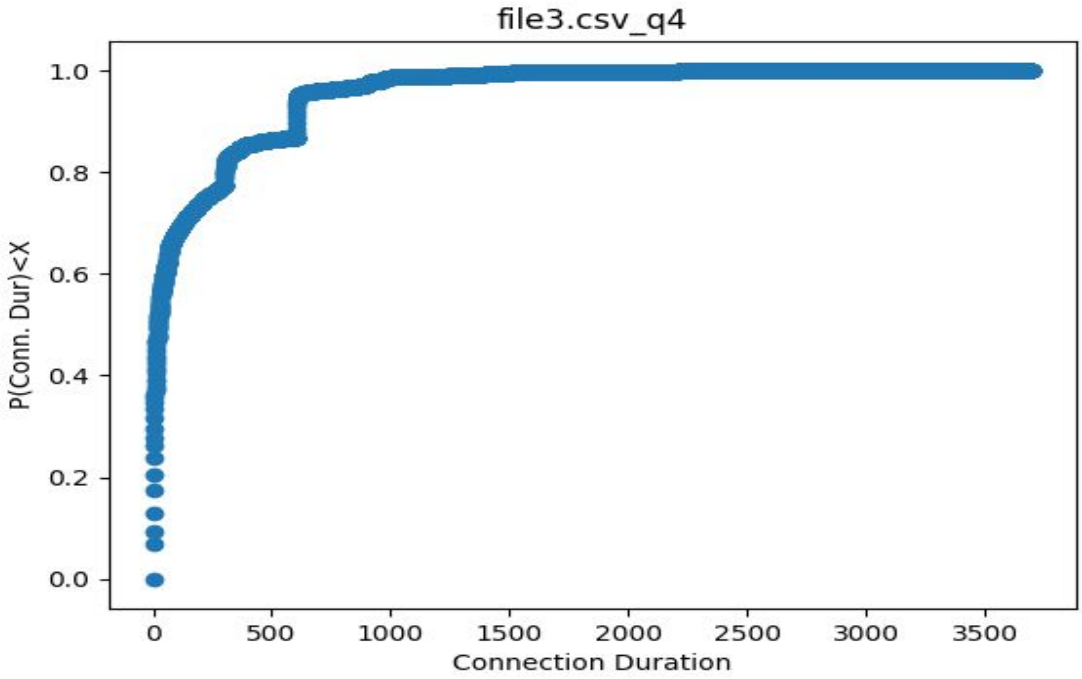
This may imply that the system is under DoS attack at that particular time.

Question 4:

The CDF plots of TCP connection durations for all 3 files are shown below.





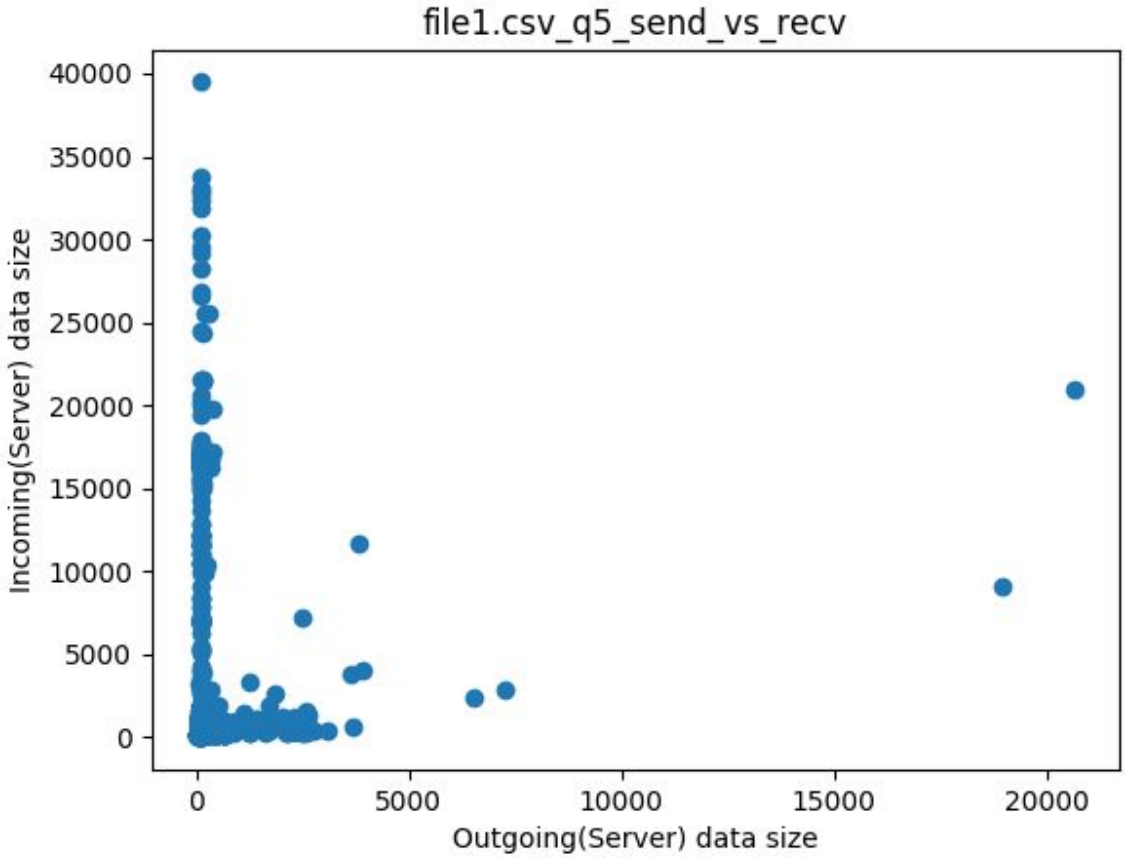


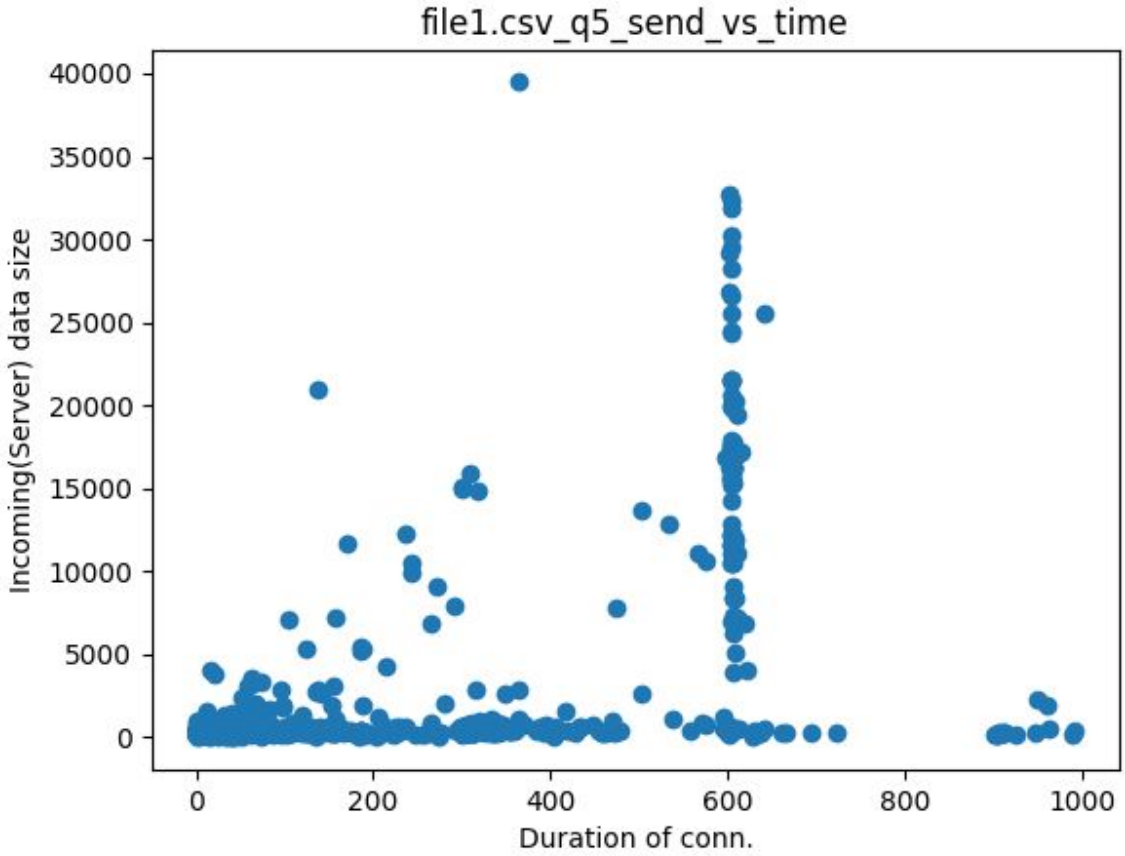
As we can see from the graphs above, most of the TCP connections are short in duration(more than 80% of the connections have duration < 500).

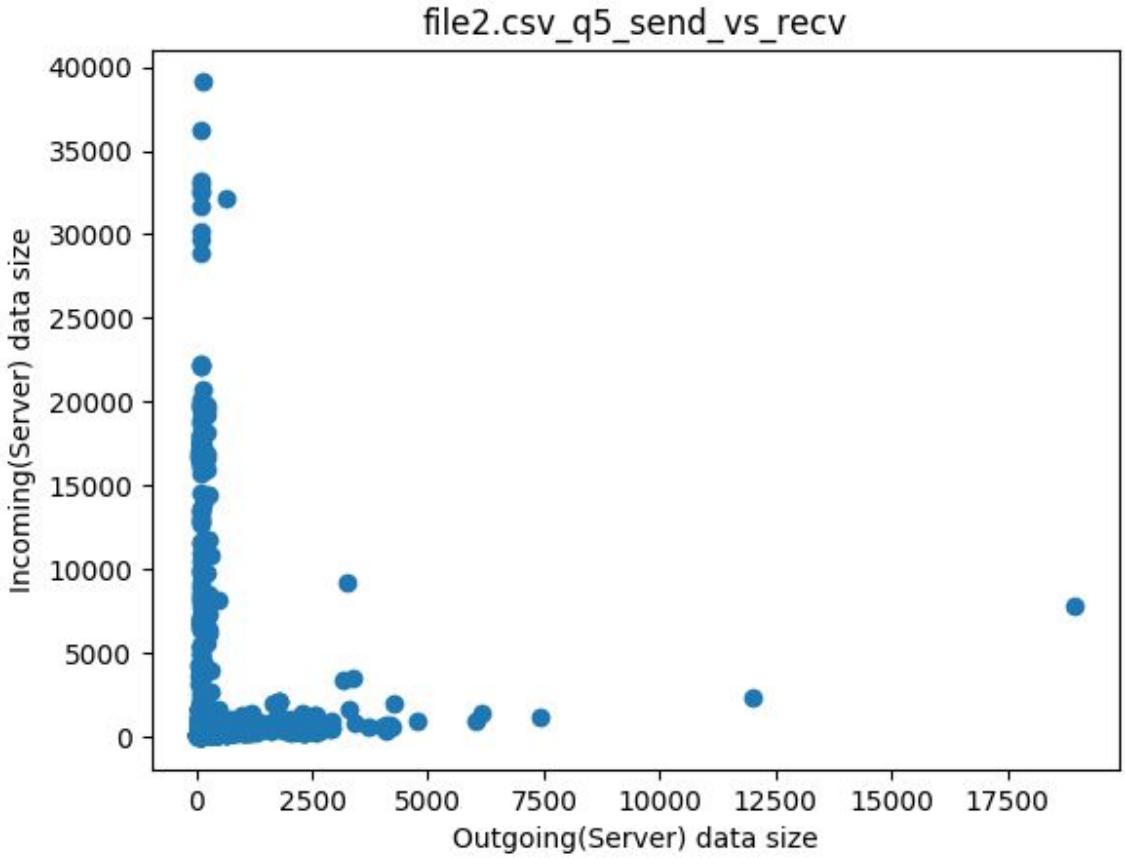
The reason behind this is the trace contains FTP connections on port 21 which are used to initiate the FTP connection, this contains only FTP commands, once the FTP connection is established there's no use of this TCP connection(on port 21) as the file transfer takes place on port 20. So, after the successful establishment of the FTP connection on port 20, the TCP connection(on port 21) is ended.

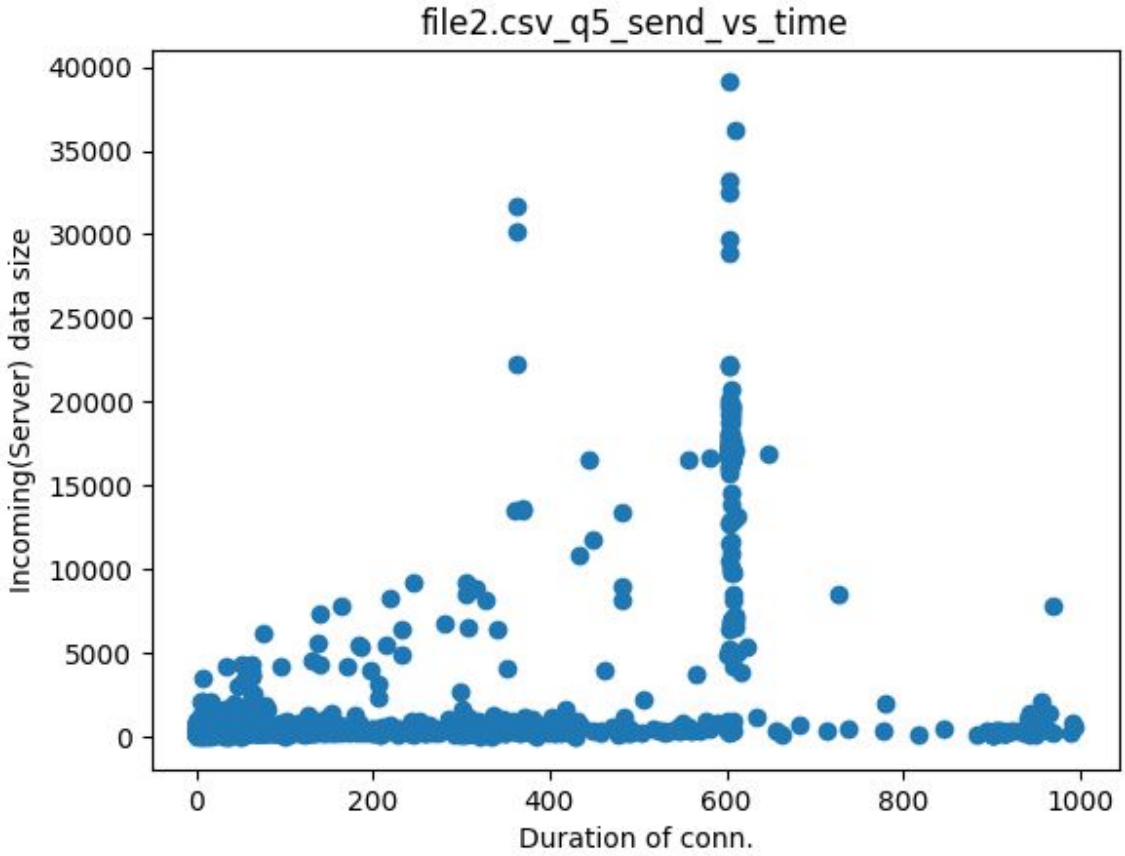
Question 5:

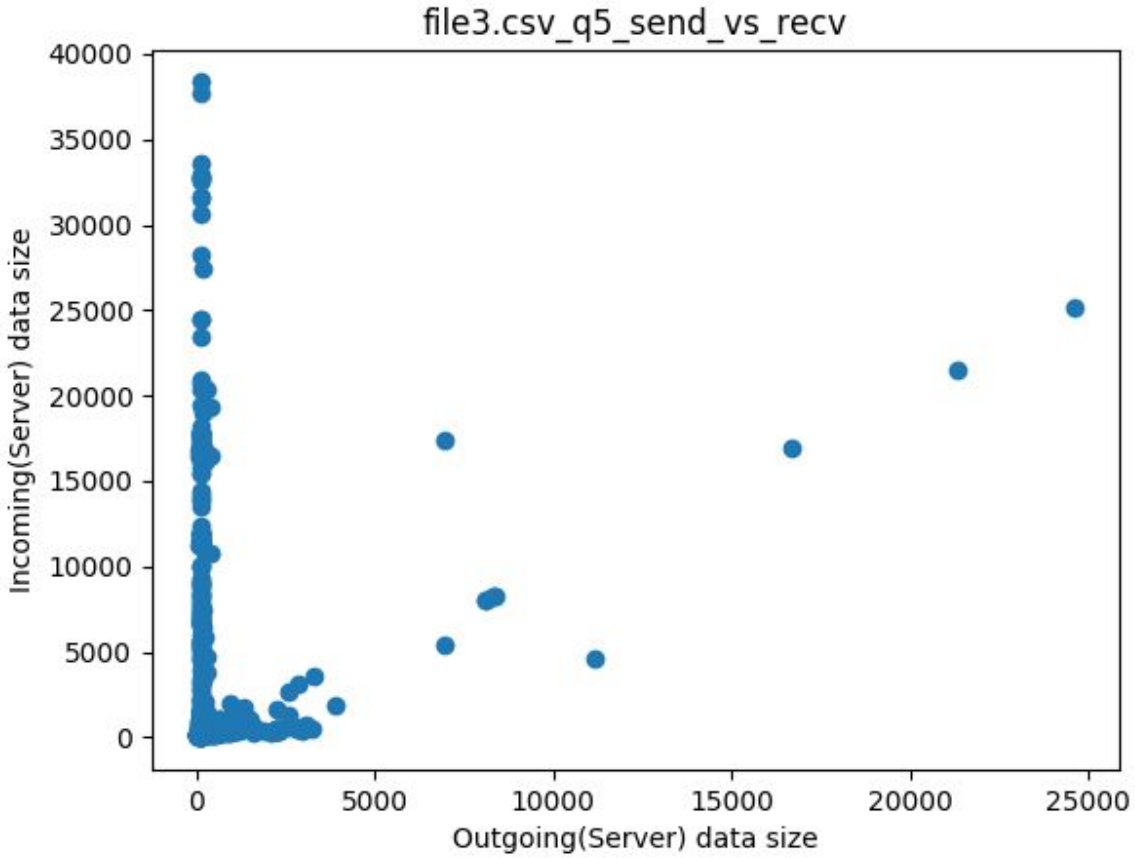
The plots given below are the send data vs connection duration and send data vs received data.

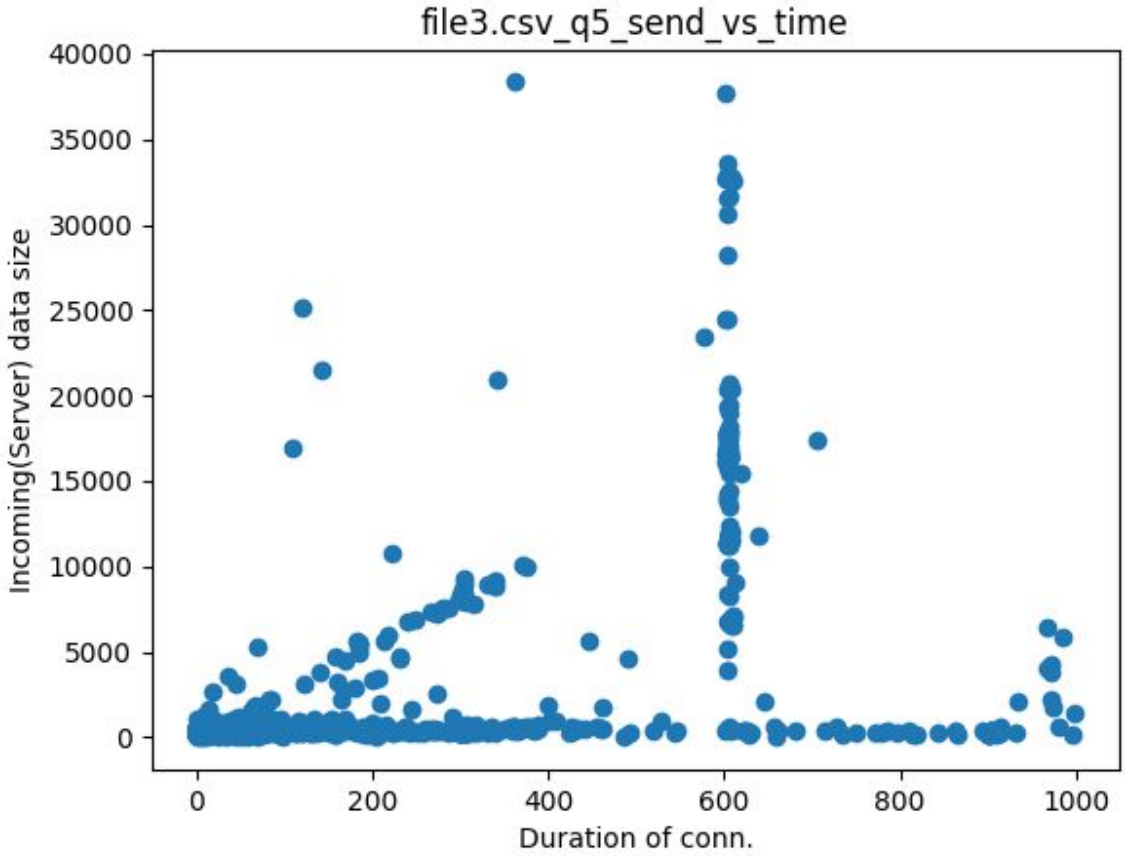

file1.csv_q5_send_vs_recv
Incoming(Server) data size
Outgoing(Server) data size
0
5000
10000
15000
20000
25000
30000
35000
40000
0
5000
10000
15000
20000


file1.csv_q5_send_vs_time
Incoming(Server) data size
Duration of conn.
0
5000
10000
15000
20000
25000
30000
35000
40000
0
200
400
600
800
1000


file2.csv_q5_send_vs_recv
Incoming(Server) data size
Outgoing(Server) data size
40000
35000
30000
25000
20000
15000
10000
5000
0
0
2500
5000
7500
10000
12500
15000
17500


file2.csv_q5_send_vs_time
Incoming(Server) data size
Duration of conn.
40000
35000
30000
25000
20000
15000
10000
5000
0
0
200
400
600
800
1000

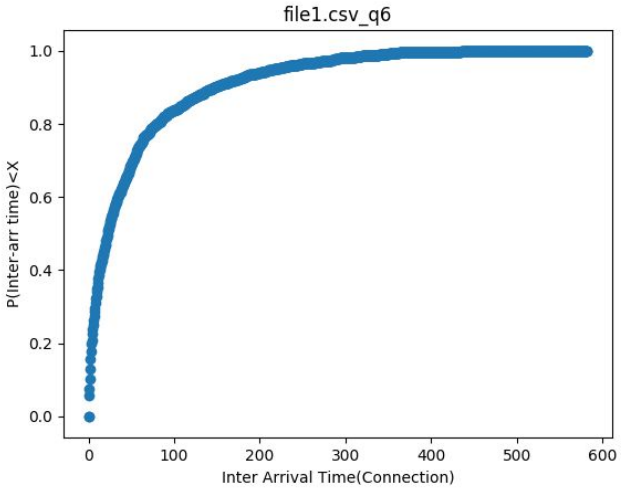

file3.csv_q5_send_vs_recv
Incoming(Server) data size
Outgoing(Server) data size
0
5000
10000
15000
20000
25000
30000
35000
40000

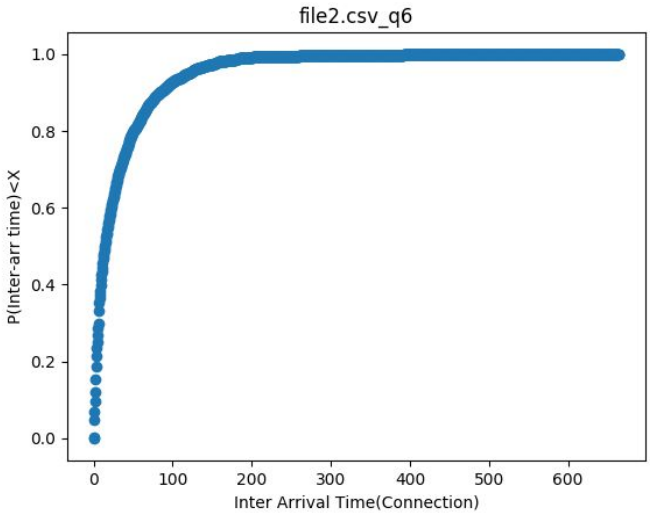


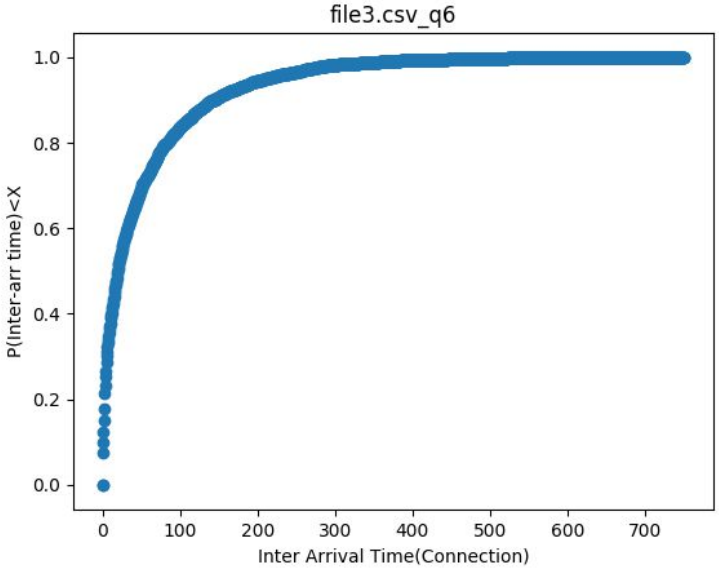
	Corr. coeff. of send vs time	Corr. coeff of send vs recv
File 1	0.7274	0.5421
File 2	0.1649	0.5671
File 3	0.4265	0.6267

Question 6:

The CDF plots of inter-arrival times of TCP connections are shown below.







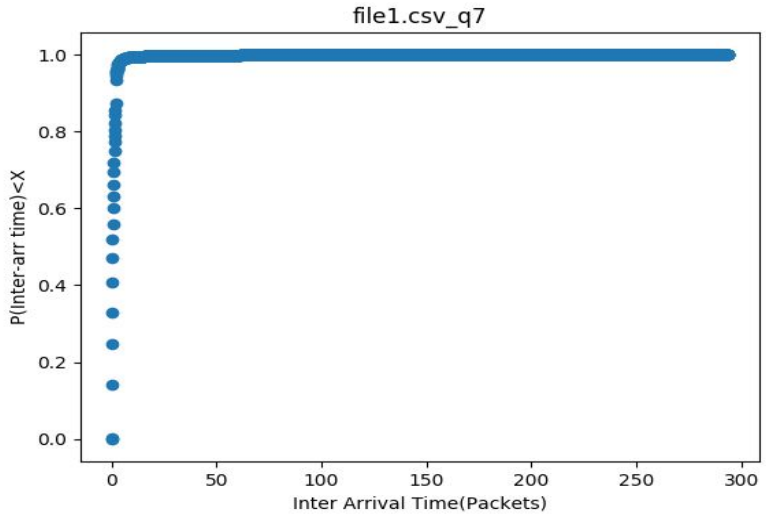
The inter-arrival time varies from 0 - 800 secs.

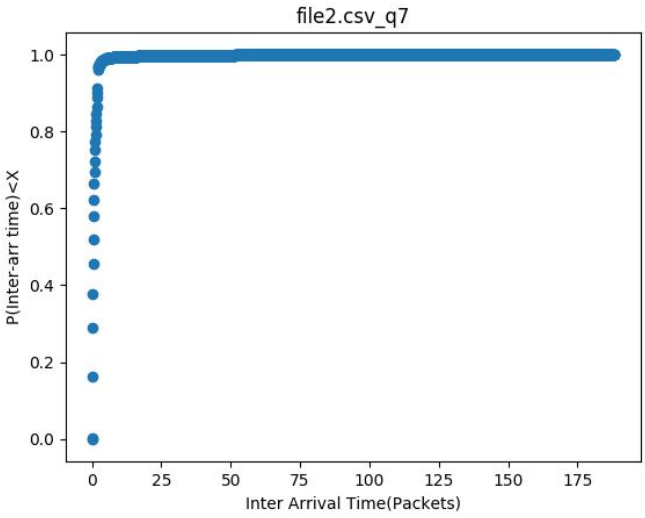
	Mean	Median
File 1	51.84	22.31

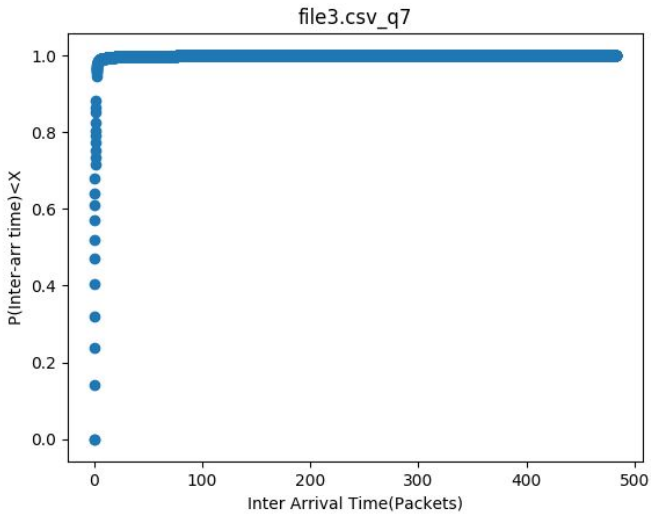
File 2	31.51	14.39
File 3	51.38	19.66

Question 7:

The CDF plots of inter-arrival times of packets sent to servers in the network are shown below.







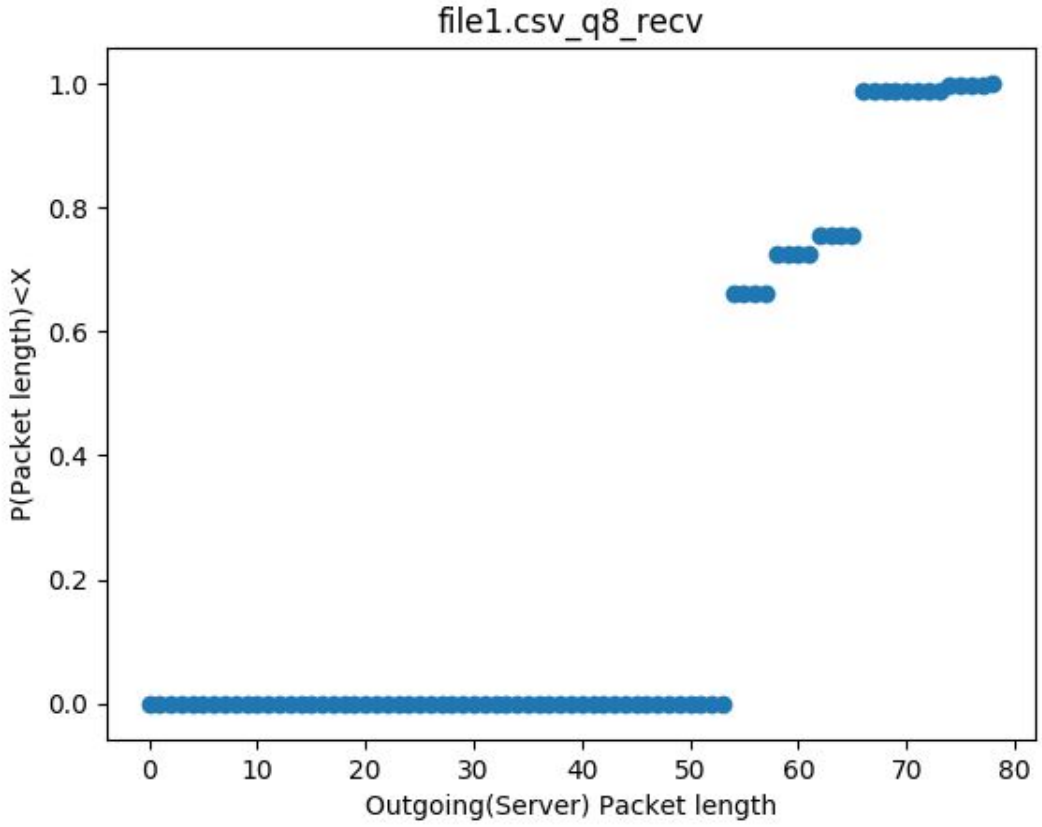
	Mean	Median
File 1	1.133	0.550
File 2	0.957	0.457
File 3	1.175	0.554

As you can see, most of the points(almost 95-99%) have inter-arrival time < 10-25 secs. The reason behind this is because the no. of clients are almost 10 times the no. of servers, all of these clients send packets to different servers at a time making inter-arrival time very less.

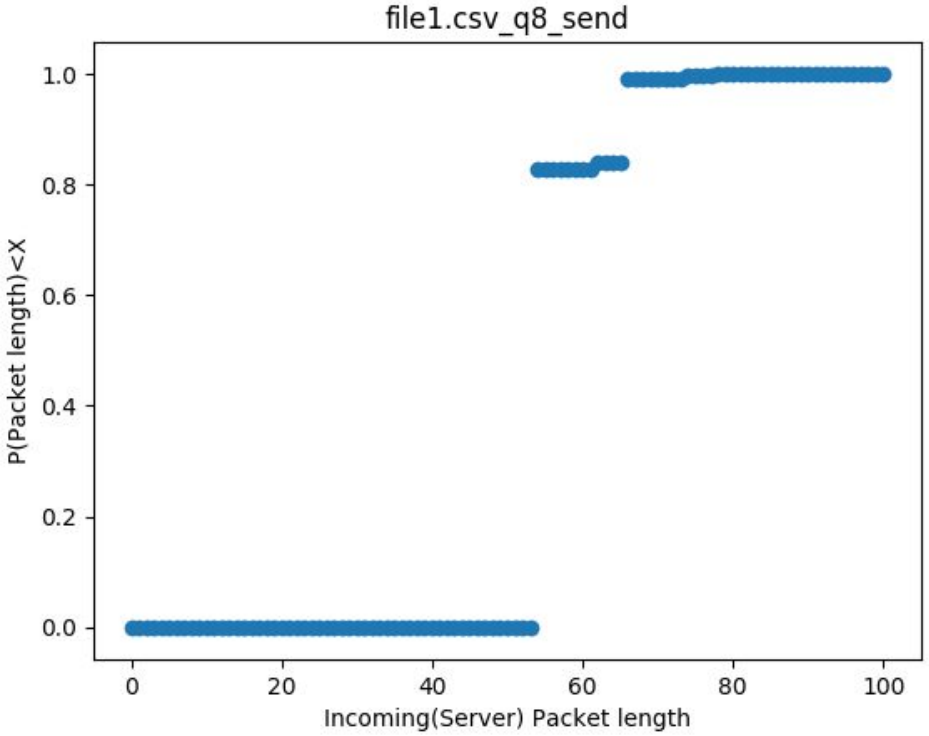
Even though most of the packets are sent at a time, the inter-arrival time has a wide range(300-500 secs) in all 3 scenarios because…

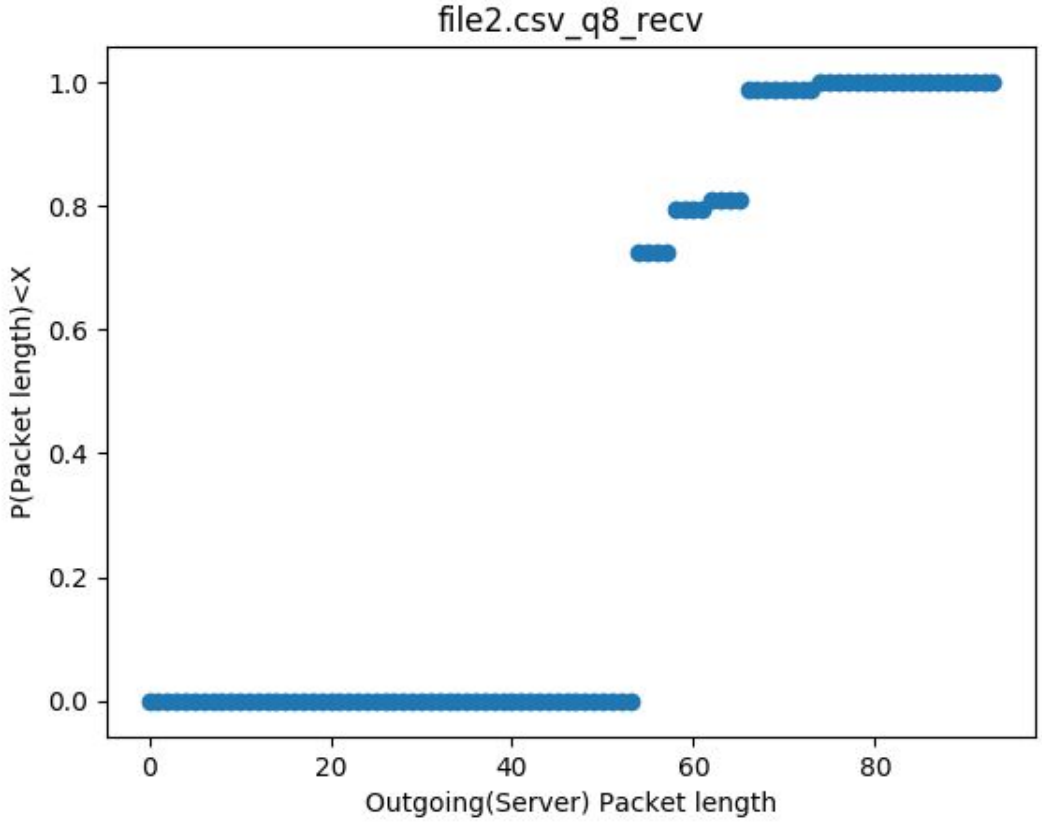
Question 8:

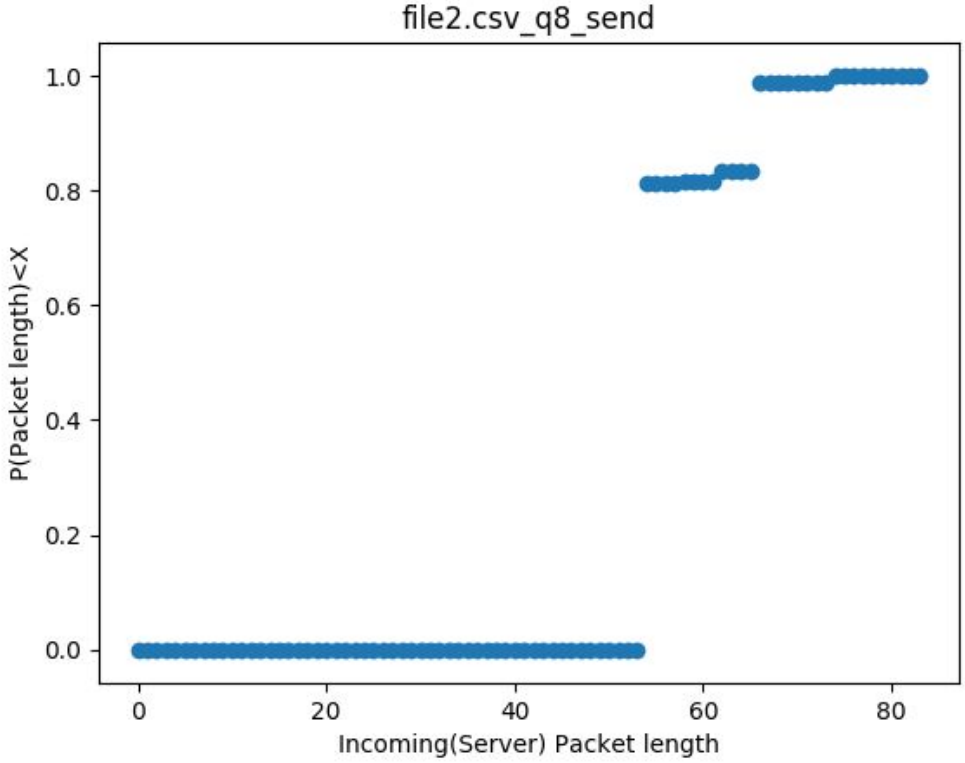
The CDF plots of packet lengths send/received are shown below.

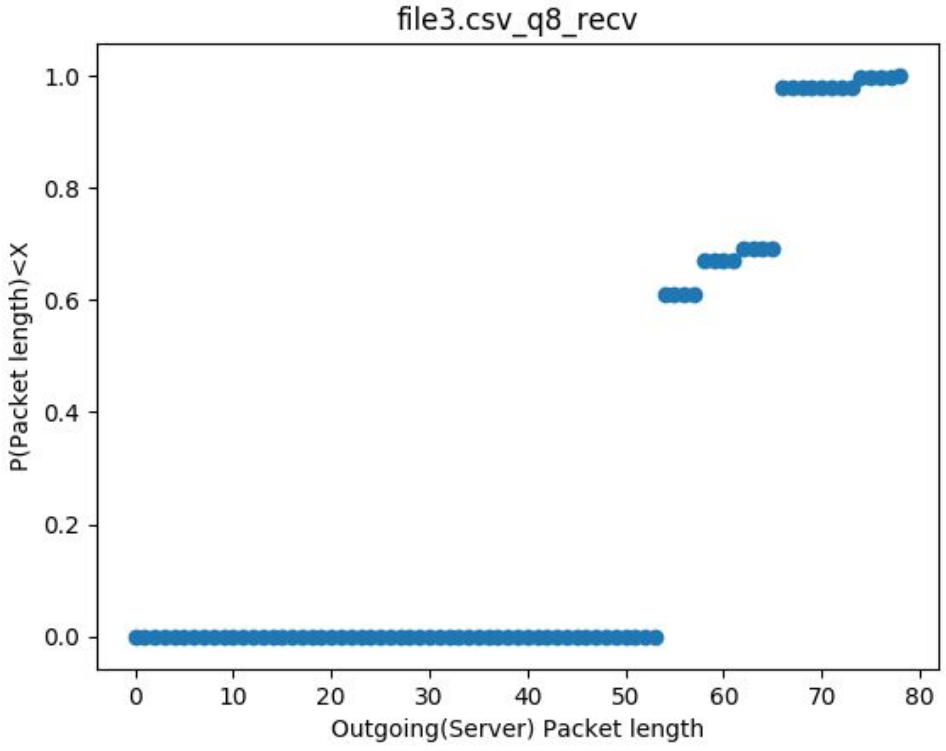


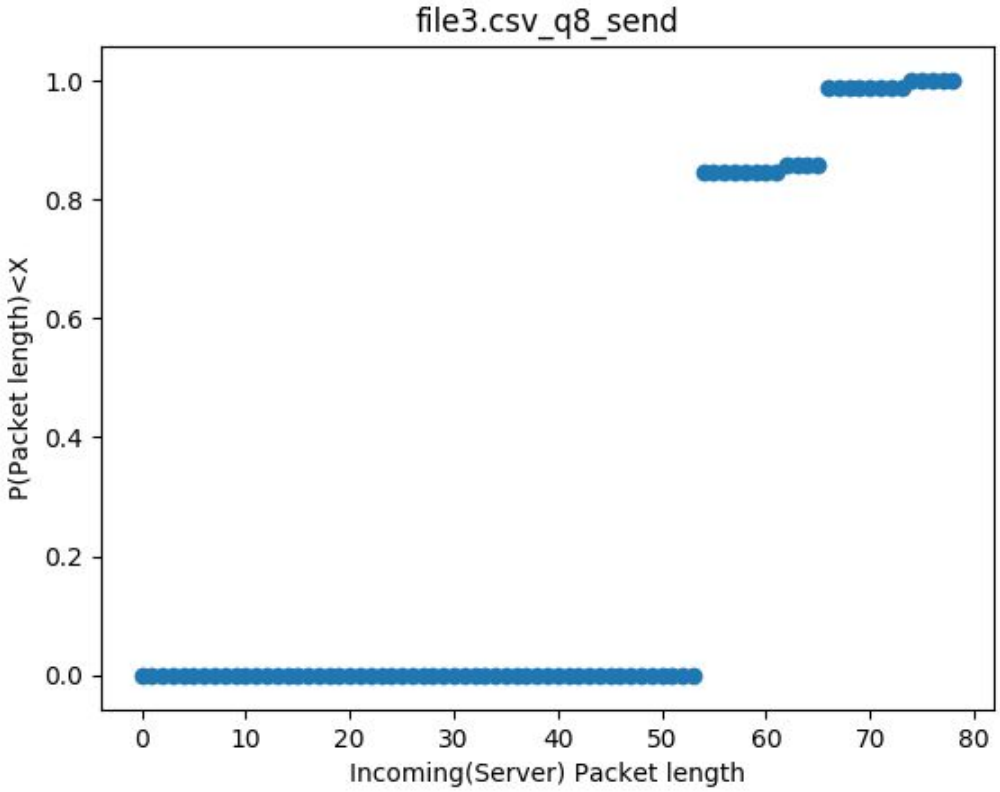
b


file1.csv_q8_send
P(Packet length)<X
Incoming(Server) Packet length
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80
100


file2.csv_q8_recv
P(Packet length)<X
Outgoing(Server) Packet length
0.0
0.2
0.4
0.6
0.8
1.0
0
20
40
60
80


file2.csv_q8_send
P(Packet length)<X
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
Incoming(Server) Packet length


file3.csv_q8_recv
P(Packet length)<X
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50
60
70
80
Outgoing(Server) Packet length



As we can see from the above plots, there's no packet with length < 50, then the CDF suddenly rises to a value of around 0.6-0.8 and remains constant for a significant range. This shows the clustering of packet sizes at 50(actually at 54). This can also be observed in the packet trace files, as most of the packets(TCP layer) have length 54. So, there's a sudden rise in CDF at packet length 54 in all these graphs.

Question 11:

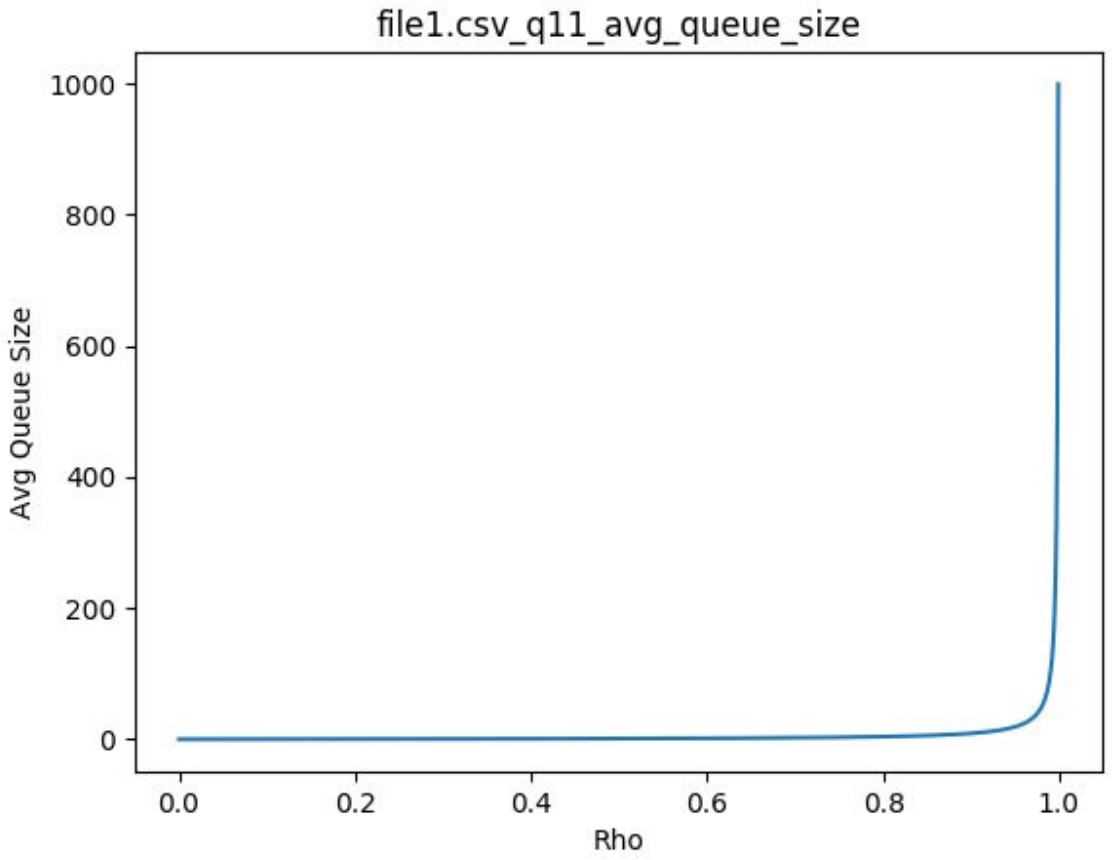
	Lambda
File 1	0.8839387
File 2	1.057798
File 3	0.8739079

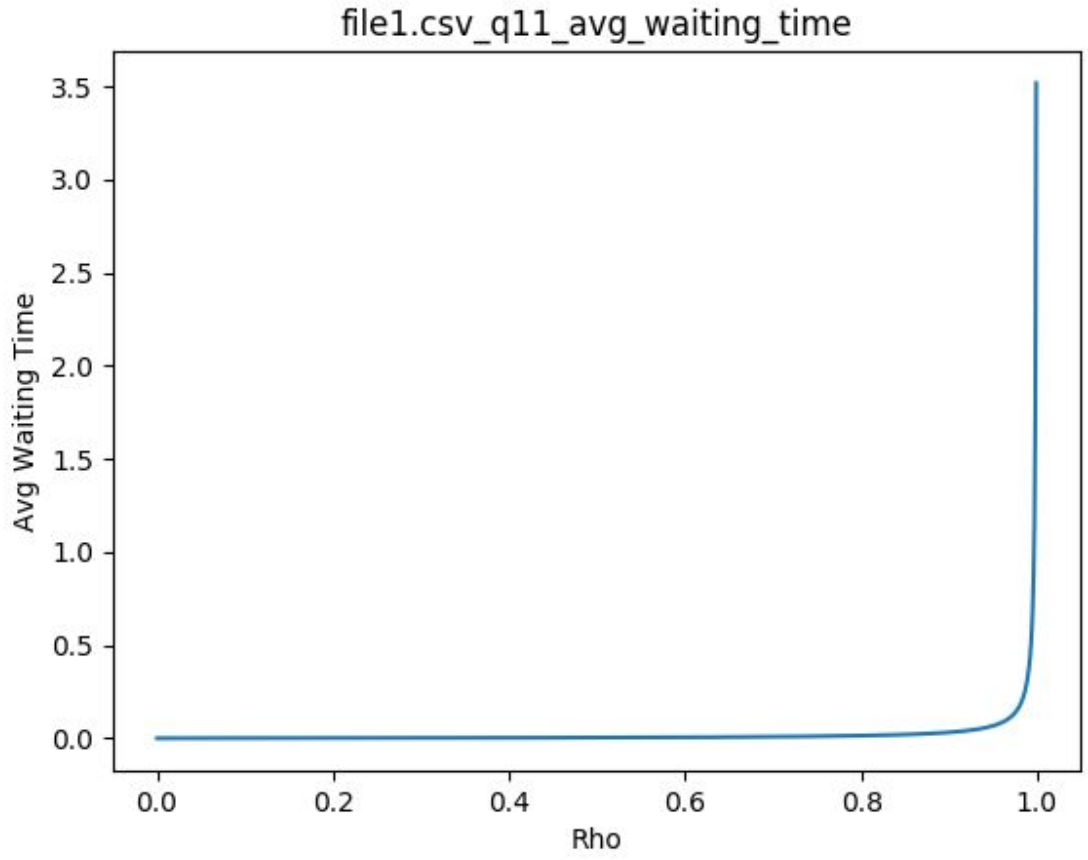
The values in the table above are the values of lambda, the parameter of the exponential fit we got from the R script.

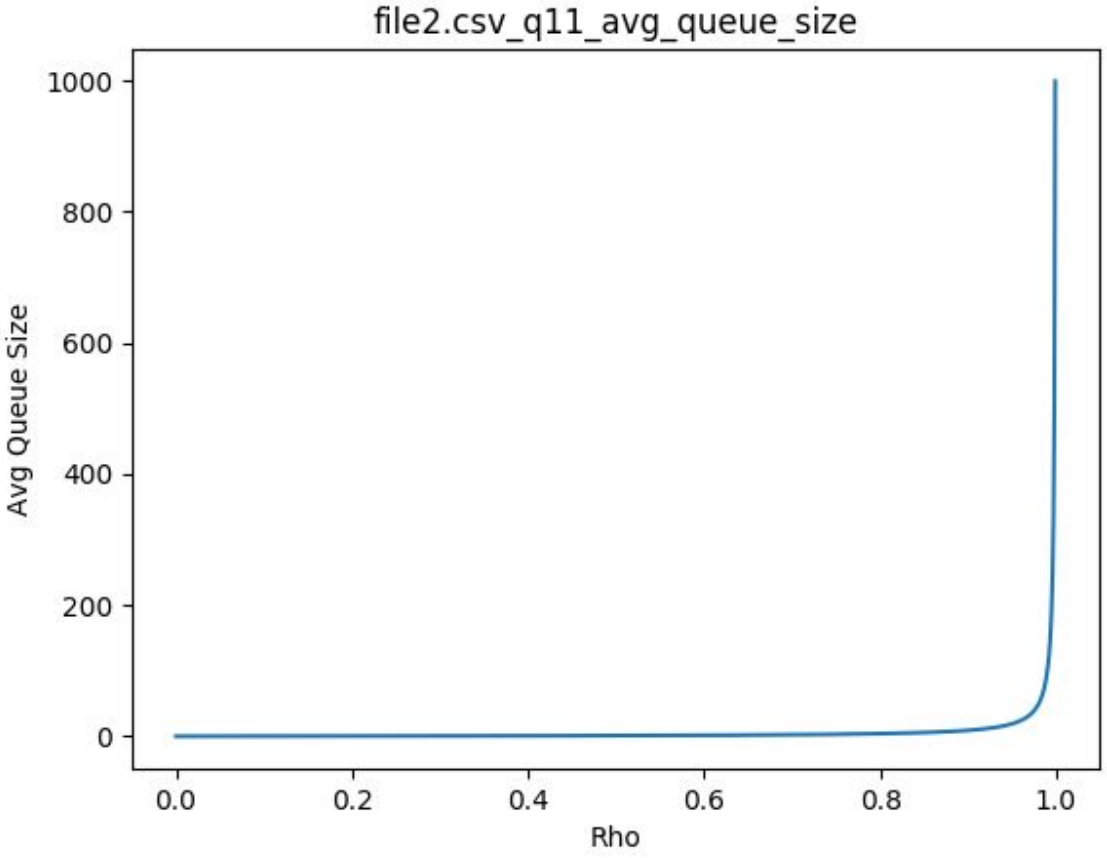
The average packet size(from question 8) is around 56.345 bytes.
Given, the link can transmit data of 128 kbps.
So, the value of mu = (128x1000)/(56.345x8) = 283.964 packets/sec.

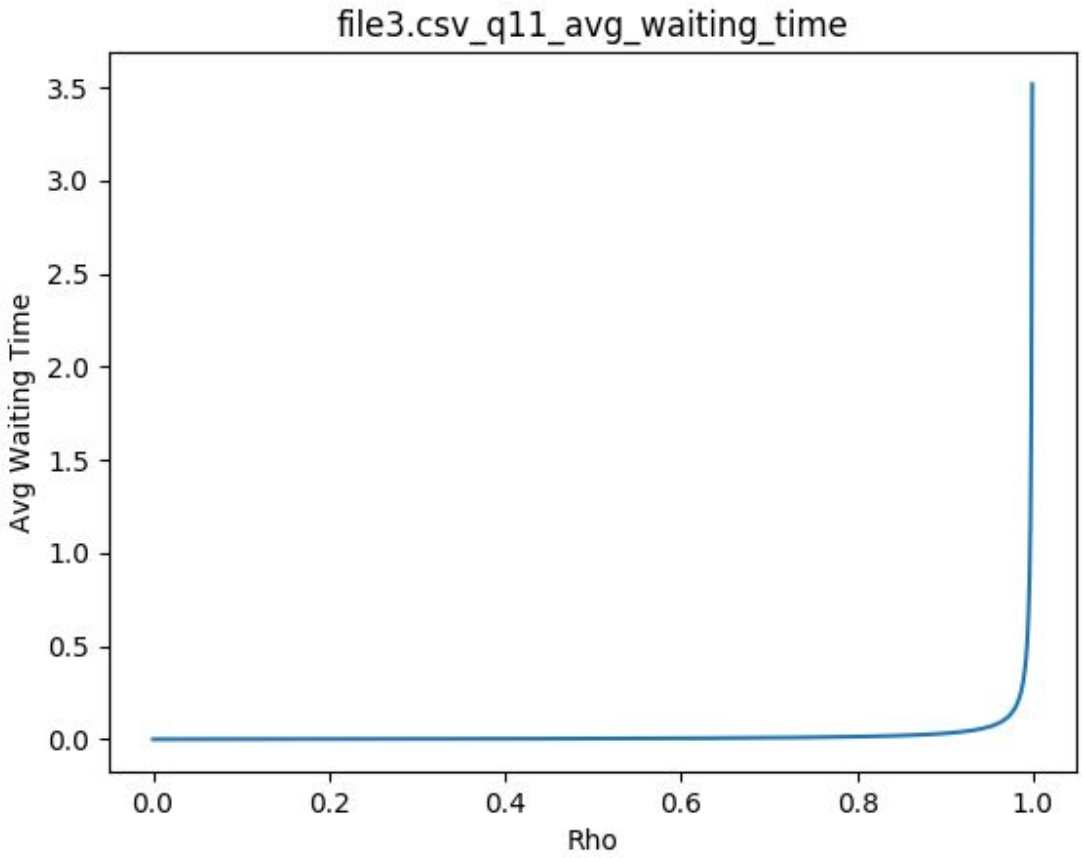
Rho = lambda/mu.

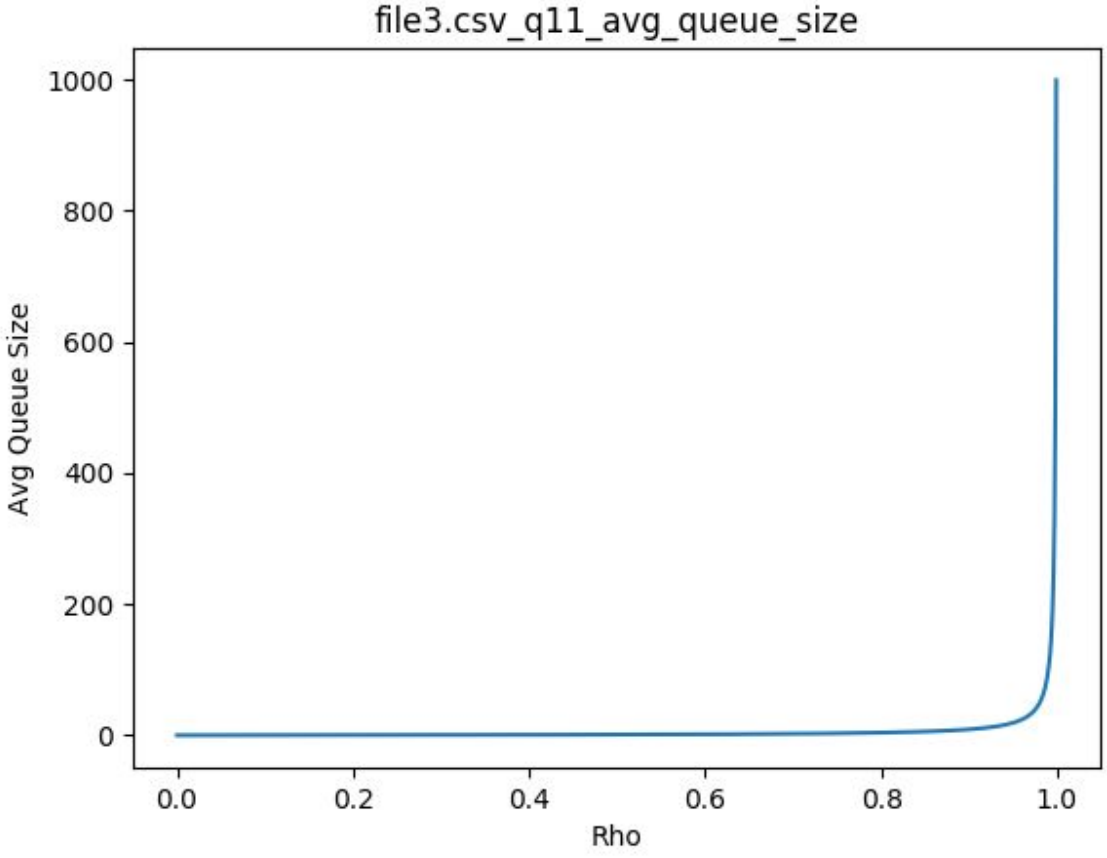
	Utilization factor(rho)
File 1	0.00311
File 2	0.00372
File 3	0.00307




file1.csv_q11_avg_waiting_time
Avg Waiting Time
Rho
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0


file2.csv_q11_avg_queue_size
Avg Queue Size
Rho
0
200
400
600
800
1000
0.0
0.2
0.4
0.6
0.8
1.0


file3.csv_q11_avg_waiting_time
Avg Waiting Time
Rho
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0


file3.csv_q11_avg_queue_size
Avg Queue Size
Rho
0
200
400
600
800
1000
0.0
0.2
0.4
0.6
0.8
1.0

As given in the assignment statement, the utilization factor is very less.

	Average queue size	Average waiting time
File 1	0.003119702	1.098×10^{-5}
File 2	0.00373389	1.315×10^{-5}
File 3	0.003079454	1.084×10^{-5}