Single Queue Multi-Server Modeling and Simulation of Tim Hortons

Project Report

Course: ENGG 815

Modeling, Simulations, and Computer Aided Processes

Department of Industrial Systems Engineering

University of Regina

Submitted To

Dr. Wei Peng

Submitted By

Ashifa Najumudeen (200379781),

Pijoe Jack Peter (200391997),

Shubham Sharma (200398667)

April 22, 2018

1. SUMMARY:

Queuing theory is the study of waiting in a line and to predict how long the customer waits in the queue and making smart decisions to wait in the queue for the short time. Customers must wait in the queue and nobody wants to wait in the line, and so we must find the strategies to wait for a shorter time in the queue. Queuing theory is the mathematical analysis of queues and waiting times in a stochastic system. These stochastic systems have been used extensively to analyze the production and service processes demonstrating random variability in service times and arrival times.

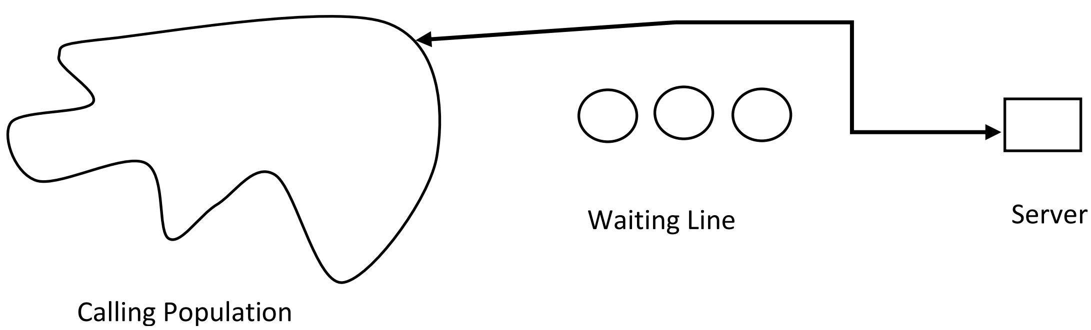
The need for queue arises when the short-term demand for service exceeds the capacity and most often it is caused by random variation in service times and inter-arrival time of the customer. If there is long term demand for service, then the capacity of the queue will overflow. In this project, we studied the queuing system of Tim Hortons and simulated the same using Pro-Model. The model is a single queue with two servers system G/G/2. We collected data during rush-hour, analyzed data, then simulated the behavior of the system using Pro-Model, validated the simulation model using T-test and implemented the simulation model, analyzed the outputs and conducted the cost analysis of the system.

TABLE OF CONTENTS

1. Description of study system:

Queuing system is used to control the queues. In this project, we have used the Tim Hortons queuing system to analyze the data and find a way to reduce the waiting time of the customers. Queue systems have come a long way from simple physical activities to some digital applications. The principle of queuing system depends on three criteria: fairness, engaging queuing and explained waiting. Capacity problems are very common in industry and one of the main drivers of process redesign, need to balance the cost of increased capacity against the gains of increased productivity and service. Queuing and waiting time analysis is particularly important in-service systems. The effect of large costs of waiting will lost sales due to waiting.

Queuing models provide the analyst with powerful tools for designing and evaluating performance of the queuing system. Typical measures of the system performances are: 1. Server utilization, length of waiting lines, and delays of customers, 2. For relatively simple system, compute mathematically and 3. For realistic models, simulation is usually required. Simulation is used in the analysis of queuing models. The following is the simple but typical queuing model:



Some examples of the real-world queuing systems are: commercial queuing system, transportation service systems, business-internal service systems, social service system, etc. The components of basic queuing process are:

- The calling population: The population from which the customers or job originate. The size of the calling population can be finite or infinite. It can either be homogenous (only one type of customers or jobs) or heterogenous (numerous different kinds of jobs or customers).
- The arrival processes: It determines how, when and where the customer/jobs arrive to the system. From which one important characteristic is the customers'/jobs' inter-arrival times. To specifically mention that the arrival process requires data collection of inter-arrival times and statistical analysis.
- The queue configuration: It mainly specifies the number of queues (single or multiple lines to several service stations), their location, their effect on customer behavior and the maximum size (number of jobs the queue can hold).
- The service mechanism: It can involve one or several services facilitates with one or several parallel service servers. The service provided by a server is characterized by its service time. Specification is required and typically involves gathering of data and statistical analysis. Most analytical queuing models assume exponentially distributed service times, with some generalizations.
- The queue discipline: It specifies the order by which jobs are being served in the queue. Most commonly used principle is the FIFO (First In First Out). It can entail prioritization based on customer data.

1.1. Extenuating Effects of Long Queues:

- Obscuring the queue from arriving customers: It is done to reduce the queue length by diverting the customers to some area, e.g. Amusement parks diverts people to buy tickets outside the park, banks broadcast news on TV at various stations along the queue.
- Use the customer as resource: It is using the customer to fill in their form by themselves, e.g. Patient filling out medical history from while waiting for the patients.

- Making the customers' wait comfortable and distracting their attraction: While the customers wait for service, they are distracted by another things, eg. Complementary drinks given by the restaurant, computer games, food courts, at airports.
- Explain reason for the wait: The customers are explained for the delay in service and can be convinced by the authority of administration, but in most cases the customers are not satisfied of the long queue.
- Provide negative estimates of the remaining wait time: The customers are informed about their waiting time in advance so that the wait time seems shorter if an estimate is given.

Be fair and open about the queuing disciplines used: Announce in advance the queuing principles followed to the customer to avoid the difficulties.

2. Description of the simulation model:

Simulation is the emulation of a dynamic system using a computer model to evaluate and improve the performance of a system. Simulation provides a practical method for doing system experimentation.

Queuing simulation is a method used to study how a system with a limited amount of resources distribute their resources to elements that are waiting to be served, while the waiting elements may exhibit distinct variability in demand, i.e. arrival times and require distinct processing time.

A queueing model follows the structure of:

- Arrival process
- Service process
- Number of servers

2.1 Simulation system elements:

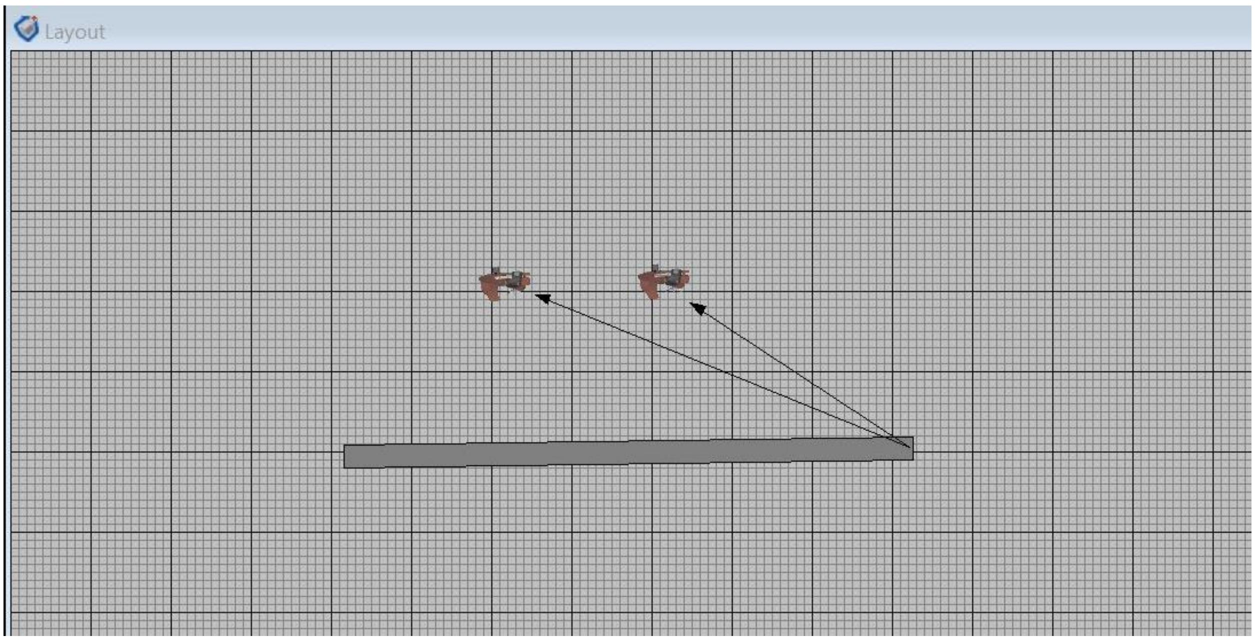
Entities, activities, resources, and controls.

Entities	Items processed through the system	• Human/animate (customers, patients, passenger, etc.) • Inanimate (Parts, documents, etc.) • Intangible (calls, e-mail, etc.)
Activities	Tasks performed in the system either directly or indirectly on the entities. Activities consume time and use resources	• Entity processing (check-in, treatment, inspection, fabrication, etc.) • Entity/resource movement (ride elevator, forklift travel, etc.) • Resource adjustment (Machine setups) • Maintenance & repair
Resources	May be required for performing activities. Have characteristics such as, capacity, speed, cycle time, reliability, etc.	• Human/animate (operator, doctors, etc.) • Inanimate (equipment, tooling, etc.) • Intangible (information, power, etc.)
Controls	Controls how, when, and where activities are performed	• Production routing sequence • Work schedules • Task prioritization

This project depicts the queuing behavior at Tim Hortons. It is G/G/2 system where first G represents the deterministic nature of interarrival time (Normal Distribution), Second G represents the service time which is having Uniform Distribution and 2 represents number of servers.

The system is simulated in Pro-model. The dataset consists of Arrival Time and Service Duration of 57 customers who visited Tim Hortons in Rush hour. And from this data inter-arrival time and service time is calculated.

Layout of the system:



Two server and single queue

Locations:

Locations

Icon	Name	Cap.	Units	DTs...	Stats	Rules...
	Desk1	1	1	None	Time Series	Oldest
	Desk2	1	1	None	Time Series	Oldest
	queue	20	1	None	Time Series	Oldest, FIFO

In this model, there are two desks for the customers and a queue from which they start moving to one of the desks. The customers move in the basis of first-in and first-out.

Entities:

The data processed by a model is known as an entity. The entities depict the worker and the speed at which they work.

Entities

Icon	Name	Speed (fpm)	Stats
	Worker	150	Time Series

Arrivals:

Arrivals

Entity...	Location...	Qty Each...	First Time...	Occurrences	Frequency
Worker	queue	1	0	Inf	N(54.9,9.5) sec

The arrival time indicates all the scheduled arrivals to the system. In this system, the frequency of a customer arrival is about 54.9 seconds. The location is where the entity arrives in the system. Here the entity arrives the queue.

Processing:

Process [3]

Entity...	Location...	Operation...
Worker	queue	
Worker	Desk1	Wait U(60,8.414) sec
Worker	Desk2	Wait U(60,8.414) sec

Routing for Worker @ Desk2

Blk	Output...	Destination...	Rule...
1	Worker	EXIT	FIRST 1

The main part in the system is the processing of this model. It defines the routing of entities through this system and the functions that take place in in each location. For this system, there are 3 entities, the location of the entity is 2 desks and the queue.

3. Description of the improved model:

In this system, we are improving the model by adding one more server to the system. So, this is a G/G/3 system where first G represents the deterministic nature of interarrival time (Normal Distribution), Second G represents the service time which is having Uniform Distribution and 3 represents the number of servers. The use of this improved model is to reduce the waiting time and service time. With the addition of a third server, the customers in the queue will not have to wait for a long time.

Locations:

Locations

Icon	Name	Cap.	Units	DTs...	Stats	Rules...
	Desk1	1	1	None	Time Series	Oldest
	Desk2	1	1	None	Time Series	Oldest
	Queue	10	1	None	Time Series	Oldest, FIFO
	Desk3	1	1	None	Time Series	Oldest

In this model, the only change done is to add one more server desk to the system and all the desks have the same capacity as before. The customers move in the basis of first-in and first-out.

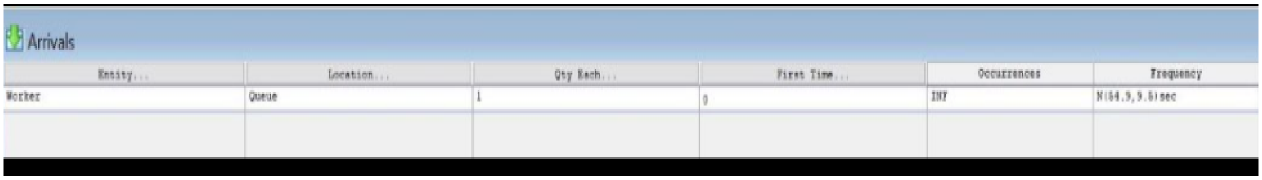
Entities:

Entities

Icon	Name	Speed (fpm)	Stats
	Worker	150	Time Series

The data processed by a model is known as an entity. The entities depict the worker and the speed at which they work.

Arrivals:



Arrivals

Entity...	Location...	Qty Each...	First Time...	Occurrences	Frequency
Worker	Queue	1	0	INF	N(54.9,9.6)sec

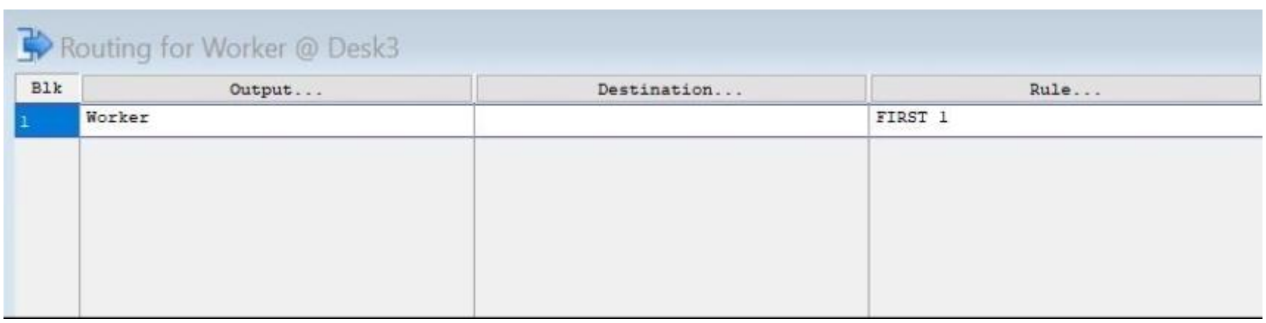
The arrival time indicates all the scheduled arrivals to the system. In this system, the frequency of a customer arrival is about 54.9 seconds. The location is where the entity arrives in the system. Here the entity arrives the queue.

Processing:

Process [4]

Entity...	Location...	Operation...
Worker	Queue	
Worker	Desk1	Wait U(60,8.414)sec
Worker	Desk2	Wait U(60,8.414)sec
Worker	Desk3	Wait U(60,8.414)sec

The main part in the system is the processing of this model. It defines the routing of entities through this system and the functions that take place in in each location.



Routing for Worker @ Desk3

Blk	Output...	Destination...	Rule...
1	Worker		FIRST 1

The queue is routed to a desk such that when one customer leaves the desk, the other person in the queue goes to the desk.

4. Comparison of the scenarios:

In this project we have discussed the problem of long length of queues resulting jockey and loss of brand name of a company because of the delay in services provided by that company. We took two situations for analyzing same calling population. In the first case 57 number of customers are served by 2 servers causing delay in the service and hence long queues and increased waiting time. This problem can be solved by increasing the number of servers and this is done in the second approach. In that part, three servers are used for the same data set and results are recorded and shown below. From the results we can clearly see that by increasing the number servers the delay time can be reduced.

5. Analysis of ProModel Results:

5.1. 2-Server System:

Scoreboard:

Scoreboard

Name	Total Exits	Average Time In System (Min)	Average Time In Operation (Min)	Average Cost
Worker	259.00	2.08	2.08	0.00

Once the processing is done, the results are produced. The scoreboard provides key statistics for the entities of our model such as the total number of exits, average time of a customer in the system, average time in operation and the average cost.

Location Summary

Name	Scheduled Time (Hr)	Capacity	Total Entries	Average Time Per Entry (Min)	Average Contents	Maximum Contents	Current Contents	% Utilization
Desk1	4.00	1.00	95.00	1.00	0.40	1.00	0.00	39.62
Desk2	4.00	1.00	165.00	0.99	0.68	1.00	1.00	68.14
queue	4.00	20.00	261.00	1.08	1.17	2.00	1.00	10.14

The location summary gives a total summary of all the entries, capacity of the system and the average time per entry which is the waiting time in the queue. The percentage of utilization of each queue is also found.

Entity Summary							
Name	Total Exits	Current Quantity In System	Average Time In System (Min)	Average Time In Move Logic (Min)	Average Time Waiting (Min)	Average Time In Operation (Min)	Average Time B
Worker	259.00	2.00	2.08	0.00	0.00	2.08	

The entity summary shows the total number of exits from the system and the average time of a customer in the system which is 2.08 minutes.

Location Costs:

Location Costs						
Name	Operation Cost	% Operation Cost	Resource Cost	% Resource Cost	Total Cost	% Total Cost
Desk1	16.31	39.56	0.00	0.00	16.31	39.56
Desk2	24.92	60.44	0.00	0.00	24.92	60.44
queue	0.00	0.00	0.00	0.00	0.00	0.00

The location costs indicate the operation cost in each desk, the operation cost % and the total cost of this system during a time interval.

5.2. 3-Server System:

Scoreboard:

Scoreboard				
Name	Total Exits	Average Time In System (Min)	Average Time In Operation (Min)	Average Cost
Customer	260.00	1.82	1.82	0.17

Once the processing is done, the results are produced. The scoreboard provides key statistics for the entities of our model such as the total number of exits, average time of a customer in the system, average time in operation and the average cost.

Location Summary								
Name	Scheduled Time (Hr)	Capacity	Total Entries	Average Time Per Entry (Min)	Average Contents	Maximum Contents	Current Contents	% Utilization
Cashier 1	4.00	1.00	87.00	1.01	0.37	1.00	0.00	36.56
Cashier 2	4.00	1.00	87.00	0.99	0.36	1.00	0.00	36.00
Queue	4.00	10.00	262.00	0.81	0.89	2.00	1.00	2.02
Cashier 3	4.00	1.00	87.00	1.00	0.36	1.00	1.00	36.23

The location summary gives a total summary of all the entries, capacity of the system and the average time per entry which is the waiting time in the queue. The percentage of utilization of each queue is also found.

Entity Summary							
Name	Total Exits	Current Quantity In System	Average Time In System (Min)	Average Time In Move Logic (Min)	Average Time Waiting (Min)	Average Time In Operation (Min)	Average Time
Customer	260.00	2.00	1.82	0.00	0.00	1.82	

The entity summary shows the average time of a customer in the system which is 1.82 minutes. This shows that there has been an improvement in the model because the average time of a customer in the system has decreased after the addition of a third server.

Location Costs:

Location Costs						
Name	Operation Cost	% Operation Cost	Resource Cost	% Resource Cost	Total Cost	% Total Cost
Cashier 1	14.62	33.64	0.00	0.00	14.62	33.64
Cashier 2	14.40	33.13	0.00	0.00	14.40	33.13
Queue	0.00	0.00	0.00	0.00	0.00	0.00
Cashier 3	14.45	33.23	0.00	0.00	14.45	33.23

The location costs indicate the operation cost in each desk, the operation cost % and the total cost of this system during a time interval.

5.3. Verification and Validation:

By no. of Exits:

Customer – 260

The total no. of exits is based on the mean of the inter-arrival rate:

= (60/54.9) * 60 * 4

= 262

Thus, the total no. of exits is almost equal to the above value.

Time in system:

Total no. of exits: 260

Average time in the system: 2.08

To find the average time in the real system: Mean waiting time – Mean service time

= 3:45 – 1:15

= 2:20 minutes which is close to the time in the system.

5.4. T-Test:

Taking 20 samples from the data set to perform the T-test and use level of significance as 0.01.

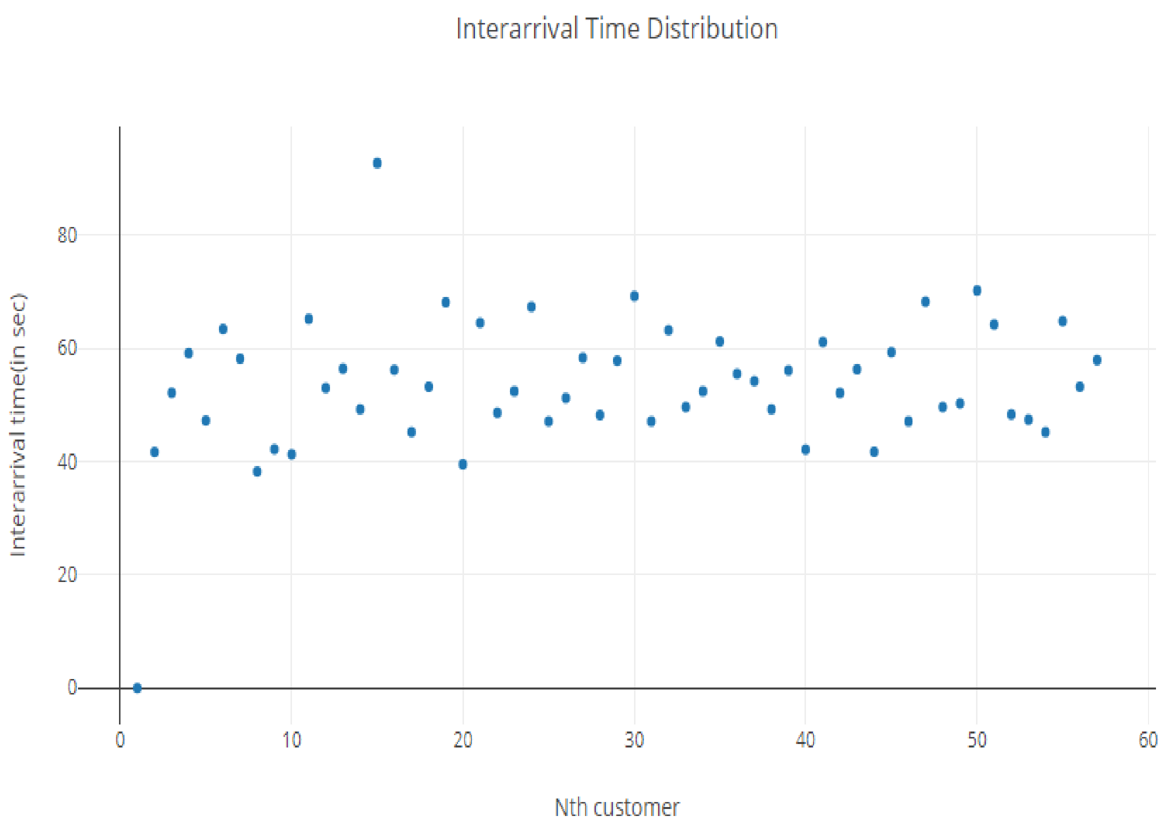
Mean of data:

$\bar{x} = (\Sigma\ x_i) / n$

Standard Deviation of data:

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}$$

For Interarrival time:



Calculating the mean (for population) = 51.06 sec

μ0 = 53.6 sec

Standard deviation = 17.437254514401

$t_0 = (\bar{x} - \mu 0)/ s /\sqrt{n}$

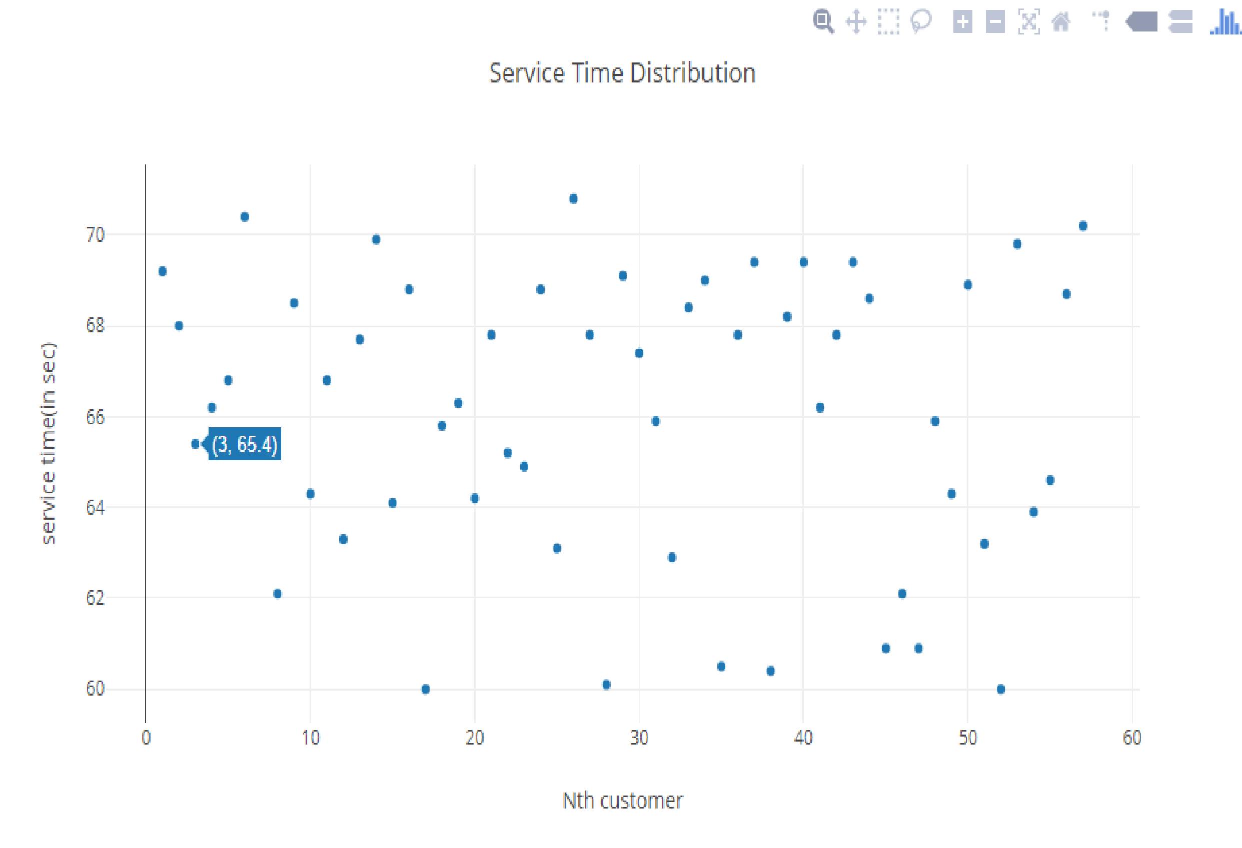
$= (51.06 - 53.6) \sqrt{20}/17.437254514401 = -0.65$

for $\alpha = 0.01$, $t_{19,0.005} = 2.84$

$t_0 < t_{19,0.005}$

Hence failed to reject the hypothesis.

For Service time:



Calculating the mean (for population) = 66.152 sec

μ0 = 66.059

Standard deviation = 2.673943470646

$t_0 = (\bar{x} - \mu 0) / s / \sqrt{n}$

$= (66.152-66.059) \sqrt{20} / 2.673943470646 = 0.1555$

for $\alpha = 0.01$, $t_{19,0.005} = 2.84$

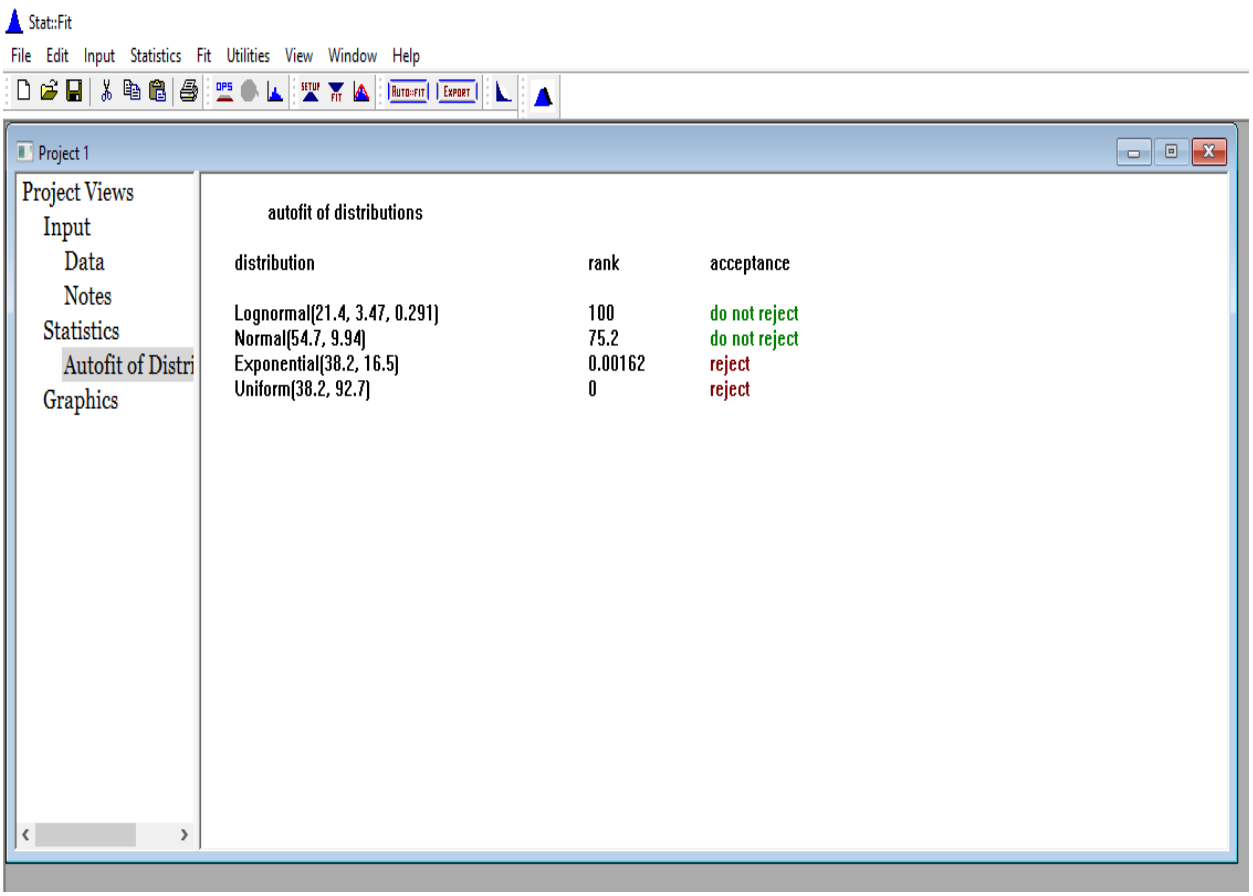
$t_0 < t_{19,0.005}$

Hence failed to reject the hypothesis.

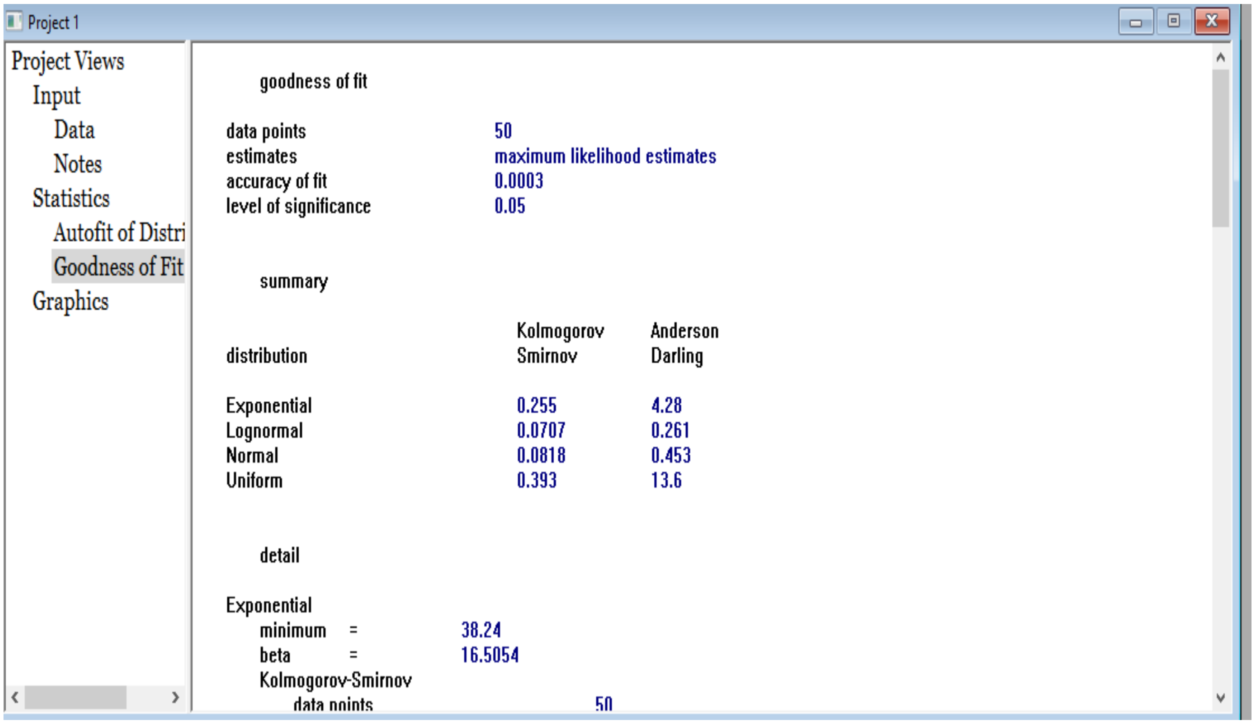
5.5. Goodness of Fit Test and Chi-square Test:

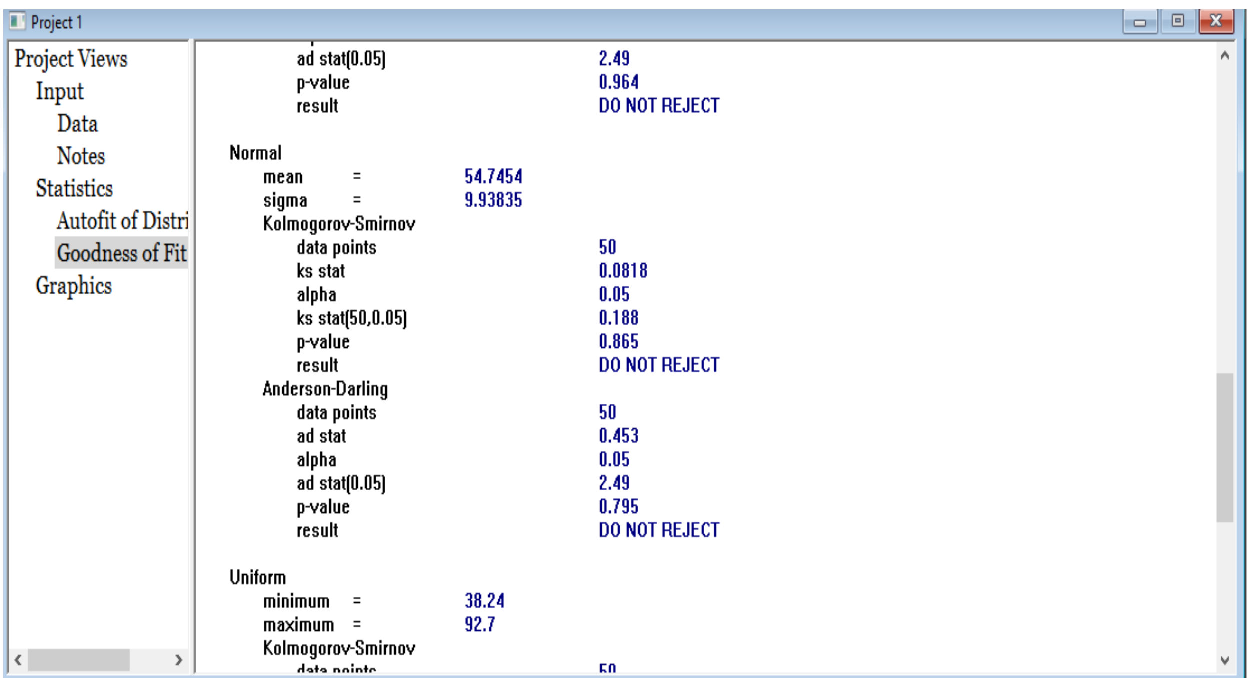
The autofit and goodness of fit test for inter-arrival time and service time is done to get the probability distribution of both.

Autofit test of Inter-arrival time is as shown:



Hence Normal Distribution is the result of Autofit test. Now for confirmation checking Chi-Square test results as shown:

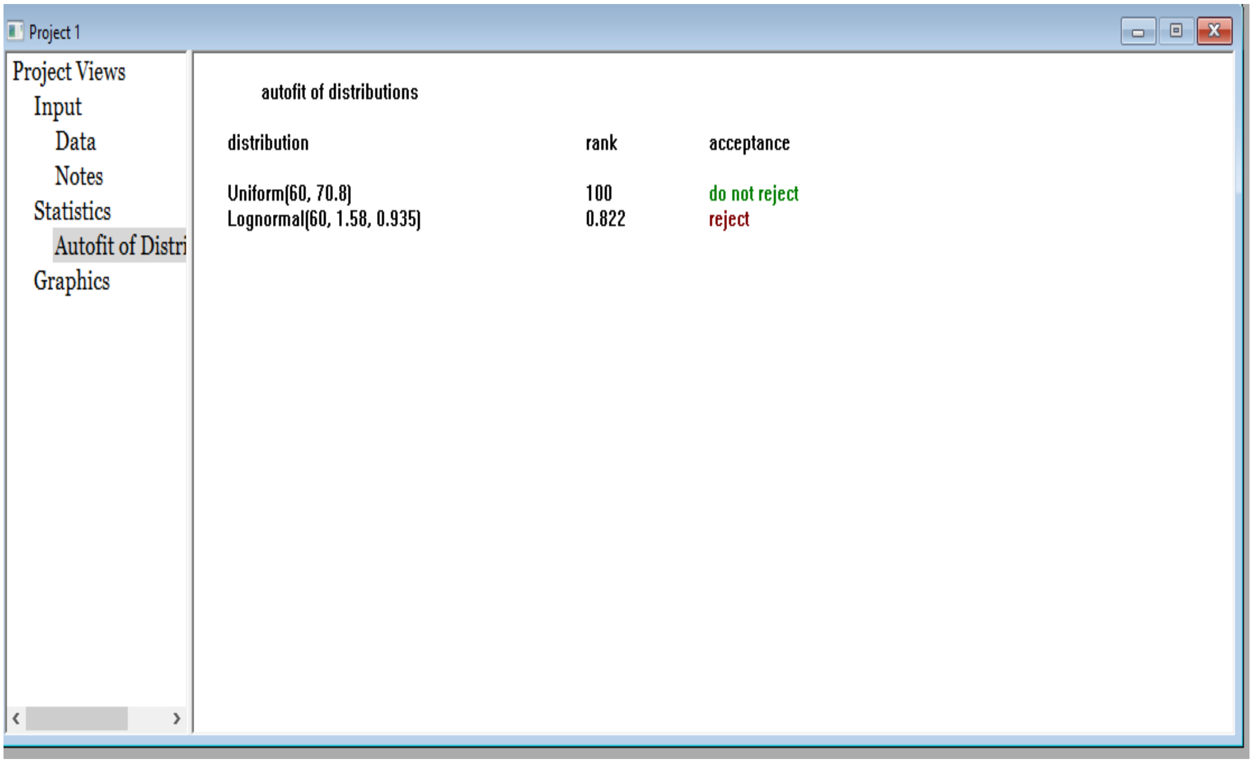




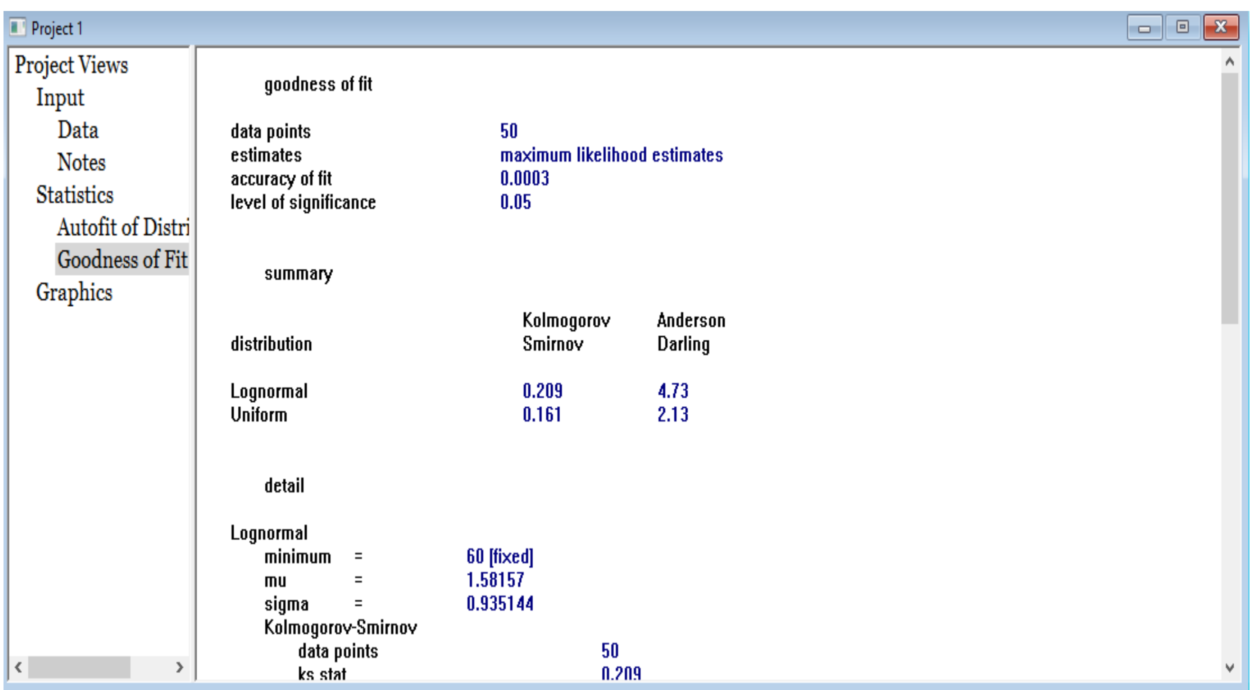
Hence Normal Distribution for inter-arrival time is Justified.

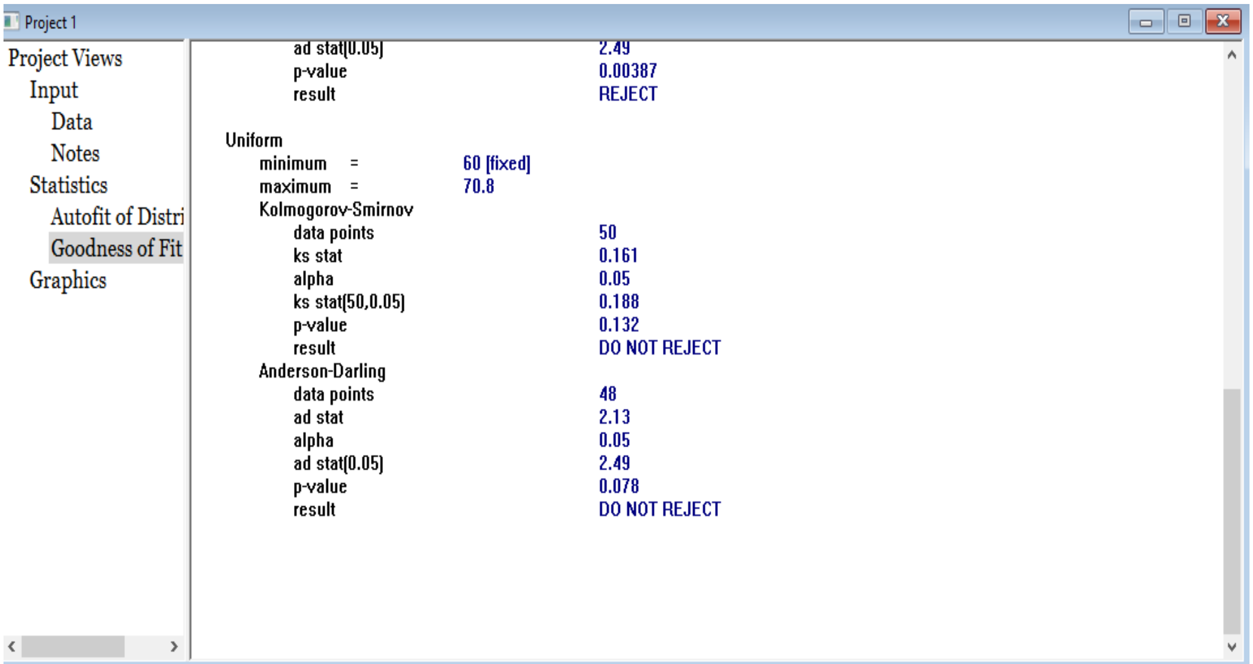
Calculating the Autofit and Goodness of fit test of Service Time:

Results of autofit test is shown below:



Hence autofit test shows the distribution of service time as Uniform distribution. Using Chi square test for confirmation. The results are shown below:





Hence the Uniform distribution for the service time is justified.

6. Recommendation:

Most of the Tim Hortons fast food chains have only a 2-server system which increases the waiting time of customers during rush hours. Sometimes customers wait in the line for a very long time and they decide to leave. To reduce the waiting time in the queue, I would recommend that they add one more server to the system which decreases the waiting time and the average time of a customer in the system is decreased. In the above system, customers in a 2-server system had to wait for more than 2 minutes. When an extra server was added to this system, the average time reduced to less than minutes. Hence, I would recommend the addition of a server to the current 2-server system.

7. Conclusion:

The above results state that a 2-server system has a longer waiting time and a longer average time in the system for a customer before they exit the queue. The average waiting time of a customer in the system was 2.08 minutes. With the addition of a 3rd server, the customers didn’t have to wait for a longer time and the average time in the system reduced to less than 2 minutes. Hence, I conclude that a 3-server system is highly efficient and reduces the average waiting time of the overall system.

8. Appendix:

The data set for this whole system are as follows:

Customers	Inter-Arrival Time	Service Time	Arrival Time	Service		Waiting Time
				Start	End	
1	00:00.0	01:09.2	12:00:00 PM	12:00:00 PM	12:01:09 PM	1:09
2	00:41.7	01:08.0	12:00:42 PM	12:01:09 PM	12:02:17 PM	1:35
3	00:52.1	01:05.4	12:01:34 PM	12:02:17 PM	12:03:22 PM	1:48
4	00:59.1	01:06.2	12:02:34 PM	12:03:22 PM	12:04:28 PM	1:54
5	00:47.3	01:06.8	12:03:21 PM	12:04:28 PM	12:05:34 PM	2:13
6	01:03.4	01:10.4	12:04:24 PM	12:05:34 PM	12:06:44 PM	2:20
7	00:58.1	01:05.3	12:05:22 PM	12:06:44 PM	12:07:49 PM	2:27
8	00:38.2	01:02.1	12:06:00 PM	12:07:49 PM	12:07:51 PM	1:51
9	00:42.2	01:08.5	12:06:42 PM	12:07:51 PM	12:08:59 PM	2:17
10	00:41.3	01:04.3	12:07:23 PM	12:08:59 PM	12:10:03 PM	2:40
11	01:05.2	01:06.8	12:08:28 PM	12:10:03 PM	12:11:09 PM	2:41
12	00:53.0	01:03.3	12:09:21 PM	12:11:09 PM	12:12:12 PM	2:51
13	00:56.4	01:07.7	12:10:17 PM	12:12:12 PM	12:13:19 PM	3:02
14	00:49.2	01:09.9	12:11:06 PM	12:13:19 PM	12:13:28 PM	2:22
15	01:32.7	01:04.1	12:12:38 PM	12:13:28 PM	12:14:32 PM	1:54
16	00:56.2	01:08.8	12:13:34 PM	12:14:32 PM	12:15:22 PM	1:48
17	00:45.2	01:00.0	12:14:19 PM	12:15:22 PM	12:16:22 PM	2:03
18	00:53.2	01:05.8	12:15:12 PM	12:16:22 PM	12:17:27 PM	2:15

19	01:08.1	01:06.3	12:16:20 PM	12:17:27 PM	12:18:33 PM	2:13
20	00:39.5	01:04.2	12:17:00 PM	12:18:33 PM	12:19:37PM	2:37
21	01:04.5	01:07.8	12:18:04 PM	12:19:37PM	12:20:44 PM	2:40
22	00:48.6	01:05.2	12:18:52 PM	12:20:44 PM	12:21:49 PM	2:57
23	00:52.4	01:04.9	12:19:42 PM	12:21:49 PM	12:22:53 PM	3:11
24	01:07.3	01:08.8	12:20:49 PM	12:22:53 PM	12:24:01 PM	3:12
25	00:47.1	01:03.1	12:21:36 PM	12:24:01 PM	12:25:04 PM	3:28
26	00:51.2	01:10.8	12:22:27 PM	12:25:04 PM	12:26:14 PM	3:47
27	00:58.3	01:07.8	12:23:25 PM	12:26:14 PM	12:27:21 PM	3:56
28	00:48.2	01:00.1	12:24:13 PM	12:27:21 PM	12:28:21 PM	4:08
29	00:57.8	01:09.1	12:25:10 PM	12:28:21 PM	12:29:30 PM	4:20
30	01:09.2	01:07.4	12:26:19 PM	12:29:30 PM	12:30:37 PM	4:18
31	00:47.1	01:05.9	12:27:06 PM	12:30:37 PM	12:31:42 PM	4:36
32	01:03.2	01:02.9	12:28:06 PM	12:31:42 PM	12:32:44PM	4:38
33	00:49.6	01:08.4	12:28:55 PM	12:32:44PM	12:33:52 PM	4:57
34	00:52.4	01:09.0	12:29:47 PM	12:33:52 PM	12:34:01 PM	4:14
35	01:01.2	01:00.5	12:30:48 PM	12:34:01 PM	12:35:01 PM	4:13
36	00:55.5	01:07.8	12:31:43 PM	12:35:01 PM	12:36:08 PM	4:25
37	00:54.2	01:09.4	12:32:37 PM	12:36:08 PM	12:37:17 PM	4:40
38	00:49.2	01:00.4	12:33:26 PM	12:37:17 PM	12:38:17 PM	4:51
39	00:56.1	01:08.2	12:34:22 PM	12:38:17 PM	12:39:25 PM	5:03
40	00:42.1	01:09.4	12:35:04 PM	12:39:25 PM	12:40:34 PM	5:30
41	01:01.1	01:06.2	12:36:05 PM	12:40:34 PM	12:41:40 PM	5:35
42	00:52.1	01:07.8	12:37:52 PM	12:41:40 PM	12:42:47 PM	4:45
43	00:56.3	01:09.4	12:38:48 PM	12:42:47 PM	12:43:56 PM	5:08

44	00:41.7	01:08.6	12:39:29 PM	12:43:56 PM	12:45:04 PM	5:35
45	00:59.3	01:00.9	12:40:28 PM	12:45:04 PM	12:46:04 PM	5:36
46	00:47.1	01:02.1	12:41:27 PM	12:46:04 PM	12:47:02PM	5:29
47	01:08.2	01:00.9	12:42:35 PM	12:47:02PM	12:48:02PM	5:26
48	00:49.6	01:05.9	12:43:24 PM	12:48:02PM	12:49:07PM	5:11
49	00:50.2	01:04.3	12:44:41 PM	12:49:07PM	12:50:11 PM	5:30
50	01:10.2	01:08.9	12:45:51 PM	12:50:11 PM	12:51:19 PM	5:28
51	01:04.2	01:03.2	12:46:55 PM	12:51:19 PM	12:52:22 PM	5:27
52	00:48.3	01:00.0	12:47:43 PM	12:52:22 PM	12:53:22 PM	5:39
53	00:47.4	01:09.8	12:48:30 PM	12:53:22 PM	12:54:31 PM	6:01
54	00:45.2	01:03.9	12:49:15 PM	12:54:31 PM	12:55:34 PM	6:19
55	01:04.8	01:04.6	12:50:19 PM	12:55:34 PM	12:56:38 PM	6:19
56	00:53.2	01:08.7	12:51:15 PM	12:56:38 PM	12:57:46 PM	6:31
57	00:57.9	01:10.2	12:52:12 PM	12:57:46 PM	12:58:57 PM	6:45