MATH5007 Tri2a 2023: Assignment 3: Simulation

Name: Syed Muhammad Ahmed Zaidi

Student ID: 20972008

**Question 1: General (60 Marks)**

**Question 1 part 1 (10M)**

**Ans)** For Multi-Objective optimization problem, the first step is to find the best solution considering one objective function at a time. This gives us target values for the objectives that would be ideal to achieve. Considering the mining problem, the goals were to minimize the cost ,waste and accidents. However, since all goals cannot be achieved in an ideal way when considering all of them at the same time, we move to the second step which is to create a new variable (Q) which captures the deviation when we are willing to achieve a number that would satisfy each goal to the maximum. Then we create a new objective function (z) that minimizes this deviation. And finally we create constraints for each of the objective. Since all of them are minimization goals, we need to subtract the target values captured in the first step from the current values and divide them with the target to get a percentage value. This constraint should be < = to the value of Q.

**Question 1 Part 2 (10M)**

**Ans)** Simulation is a kind of a model that replicates the reality to find further insights into how things work. Discrete event oriented simulations is a type which focuses on one event a time. The simulation model is in a constant state until an ‘’event” occurs that changes the scenario allowing it to move from one state to another. This progression in state due to occurance of events is what is recorded by discrete event oriented simulation that allows a detailed analysis of how entities within the model are interacting with each other. An example of such could be a coffee shop in which the entities are the customers themselves which arrive into the shop based on a distribution. This arrival is called the source. When customers come in they wait for their turn to be served by the barista which is the queue. When the barista asks them about their order and takes time to serve them, this is referred to as the delay as it is the time taken to complete a specific task. The simulation model records all these events to evaluate metrics like wait time, queue size, and service efficiency so that they can be improved in reality.

**Question 1 Part 3 (15M)**

**Ans)** The advice to run a simulation before upgrading the production line in reality will have numerous benefits. As a successful business, it is always important to keep the costs low to be able to invest in areas to prosper further. By running a simulation model, we can reduce cost of wastage in reality by experimenting with the size and layout within a computer model. This will also save up significant amount of resources including time. A simulation model will further help mitigate risks and majority of potential issues with increased size can be solved within a computer model before implementing it in reality. It will also allow us to be flexible with our decisions. Since the stakes will be low, we can experiment with various configurations and production line setups, exploring in detail to find the best possible results.

The second part of the advice will have its own benefits and risk. Hiring an external company to create a model does allow third party professionalism to be included. Their experience and knowledge may surpass what an employee of the company might have. They may know some advanced tools and softwares that they can use to make the model much more robust to give accurate results. Asking an employee to do the task may mean that he/she has to be taken away from their own tasks. This along with them taking more time to build the model compared to a professional team, may mean an inefficient use of time. The involvement of an employee in this model may also lead to a biased approach as they may choose to overlook certain details that an external company may consider important for precised accuracy.

On the other hand, if the company is looking at saving costs, hiring an external company may not be the best choice since they are professionals and may charge huge sum of money. Employing someone internally to do the same task may mean that they being a part of the company may have detailed knowledge about the processes which external companies may not. This deep insights can only be known by someone who works within and hence may be overlooked if it is outsourced. Sharing the data to companies outside may also have risk attached to it since data security and confidentiality can always be compromised. Having the task completed by an internal employee may reducen this risk significantly.

Based on the above discussion, I believe the advice to run a simulation model before upgrading should definitely be followed. However, whether to go with an external company or an internal employee highly depends on the size of the business. If the business is new and growing, choosing a cost effective way would be a better option. However, if the business is already well established with a huge structure, employing an external team would allow more precision in results.

**Question 1 Part 4 (5M)**

**Ans)** Linear optimization is fairly easy compared to Non-linear Optimization. Even though the implementation is same, mathematically they are entirely different as Linear Optimization is like finding a cow on a flat land whereas Non-Linear is to find a cow on a mountainous terrain; not knowing whether its on the peek or down in the valleys. Thus, linear optimization is to solve a problem by finding the best possible solution when the relationship between the variables and proportional to each other. In other words, each variable will have a certain known impact on the outcome, for example the input of a certain resource having a known impact on the production depending on what the objective and constraints within the question is. One the other hand, Non-Linear Optimization are problems in which the relationships between variables are not straight but complex. For example a certain input of a resource men not necessarily double the output. It may even mean that adding to much of an input may cause the output to start reducing. Hence, finding the best outcome becomes harder as there are multiple valleys and peaks on the graph which takes advanced methods to understand whether it’s a local or a global best solution.

**Question 1 Part 5 (a) (5M)**

**Ans)** As per my understanding of the unit and what I’ve learnt from it, I believe that understanding the context of the question is the most important part when carrying out optimization. In the early stages of doing AMPL questions, I used to dive in quickly to find the required data and start creating the .dat file. However, right in the middle when I got to get stuck it was hard to look back as I did not understand what the context of the question is. An error in the model becomes way more daunting when you don’t have a clue of what each SET or PARAMETER is actually representing. Hence, reading the question carefully to understand what is happening and why we are carrying out the optimization in the first place can allow much more efficiency further into the model creation. Also having a better understanding of the question details also gives a sense of accomplishment once the results show something relatable. Having the right answer showing some random figures is great but having the understanding of what that actually means to that question helps to feel achieved as the question is ‘solved.’

**Question 1 Part 5 (b) (5M)**

**Ans)** Definitely proud of making the progress of being able to write models. Its been great three months of learning specially about a very unique field of study. Ive always been keen to find solutions to compelx problems and having to know such softwares like AMPL and ANYLOGIC to solve real world issues by creating mathematical models, is a great skill to have. I would love to put in more time and practice to make these skills and everyday thing rather than just for educational use. I believe the unit Goal to learn these have been achieved but on a personal level my goal is to be so fluent with these skills that I can divide everyday problems into objective functions, variables and contraints and then solve them mathematically in the mind. If required for something more extensive, may even use these softwares to get the optimal solutions.

**Question 1 Part 5 (c) (5M)**

**Ans)** MORE PRACTICE! This is undoubtedly something I wanted to do but as an international student doing a Post Graduate degree, working to earn a living, paying the fees and also taking care of your family back home is way too hard. All comes back to Time Management. Hence, if I were to do something differently would be to allocate more time to learn with more practice. Experiment with the software, syntax, inputs, just to see what other differences can be explored. I would try not to take the easiest path to reach solutions but to create more challenges, maybe add more constraints to a given question, just as a way to get a better grip on the skillset and also be fluent enough to be able to use these on daily life challenges.

**Question 1 Part 5 (d) (5M)**

**Ans)** For AMPL, majority of it was straight forward to understand, however it used to get very challenging when implementing it youself. For example, when Sir Torsten Reiners used to walkthrough the question in the lecture, It used to feel like it’s the easiest thing to do. However, implementing it yourself without any assistance used to get a bit difficult. Especially when questions used to have further parts. Having to get done with a part and getting that sense of achievement, but the second part (like an additional constraint) requiring to make changes in the model (which used to cause errors) were hard to accept. However, as per the course overview, I thinking the most important and relevant part to the unit would be to break every problem down into further categories like objective function, variables and constraints. This is something implementable in the simplest problems of real life which actually helps in finding solutions much more efficiently.

**Question 2: Larger Simulation Model (40M)**

**Question 2 Part 1 (10M)**

**Ans)** The simulation is about a support centre which can be contacted through 3 ways including Calls, Emails and Faxes which are ranked respectively according to their priority. These requests can be ranked according to their complexity. There are a total of three different levels, with each level transferring the requests to the next one if its beyond a certain level of complexity. Level 1 requests can also be transferred directly to Level 3 considering it being harder to solve for Level 2. These requests can then further be divided into 2 more categories which can be either software or hardware for which the support centre has specialized staff. There are in total 4 categories of employees, 1 being those who are training to become Level 1 staff. Then are those that are working in the above levels of support. There are difference in the number of requests received and also the availability of workers depending on the time as night shifts will have fewer shifts as well as the number of working employees.

From the simulation there are a number of metrics that are being reported back. Firstly the source shows the number of requests being created. The Level 1 box shows the quantity of workers currently working and also a bar that shows their current utilization. The Lost Calls, icons show how many calls were not attended due to exhaustive utilization. Moving forward the requests are represented with the use of different shapes and colors and the directions show whether they are moving forward to Level 2 or directly to Level 3. These boxes now also show the distribution of employees with their specialization in software or hardware. Other things being the same regarding the utilization bar and animated staff working. The direction then finally move towards a icon showing total number of resolved requests. The user-friendly interface allows handles to be able to change arrival rate for requests, night arrival multiplier and new employees being hired by the company. The simulation page also reports back the current time and the type of shift that is in progress. This is also along with an animation of workers currently being trained along with their number.

Going further, the statistics page shows numerous visualizations to find further insights about the mode. First being a pie chart reporting the percentages for the status of requests including, finished, lost, in process and waiting. Further a time series plot showing the mean waiting time for the complete day for different types of requests. The Resources panel reports information about the current status of all staff members. The 2nd pie chart reports the percentages of requests that were resolved at different levels of the support centre. Lost Call ratio and Total utilization of resources are also evident with different visualizations. Lastly, multiple histograms show the processing time of requests in total and then breaking them down into different levels.

**Question 2 Part 2 (15M)**  
**Ans)** When running the simulation on the default arrival rate of 20 Calls, 15 Emails and 7 Faxes, we can see that the utilization of resources by level 1 and level 2 has been under 50%. Which is a great sign showing that the resources are more than the demand which is the reason why there has never been a lost call. However, the trend does show that the resources stay low in the level 3 showing an almost 100% utilization for both software and hardware. However, since there is a promotion system with people getting trained and moving to the next level, on 18th of Jan the resources suddenly double up which allows the utilization of level 3 to drop down to a very impressive 50-60%.

If I were to increase the calls by just 5, emails by 25 and and faxes by 18. All reaching round 25, it shows that the support centre starts to feel a bit pressurized. This can be from the fact that there were no lost calls in the last scenario however just with a little increase there has been a total of 16 lost calls. The utilization of workers staff in Level 1 and Level 2 stay low with around half however, this time even with the increased staff members the utilization of Level 3 staff members was near full exhaustion. This being said, I did observe that due to the increase the speed of the overall number of requests satisfied was much more quicker. In the first scenario, the 49,268 requests were entertained until 1st of Feb but with a faster arrival rate the requests were finished by 22nd Jan however this does take into account that 11% are still in waiting while 16 of them were lost as well.

To test the limit I wildly increased the arrival rates of all three to 100. This showed that there is indeed a point in which the call centre would break down to performing really bad. The results showed that around 31% of the requests went to lost cause with no none to attend them due to the high utilization of resources at that point. Out of the total 49400 requests, 29240 are still waiting which comes up 59% of the total requests. The utilization of all three levels are way to high touching around the 100% mark. The wait time is around 144.68 minutes on average which is something no customers want to do.

From the given bars, the maximum increase for arrival rates is 100. Considering that option we were able to observe that the support centre starts performing horrendously. However, with these configurations the support centre is still performing. There were still a 4971 requests that were solves. Hence, I was not able to find the point where they were not able to perform at all but yes with highly increased values of arrival rates, the support centre did collapse to a terrible performance.

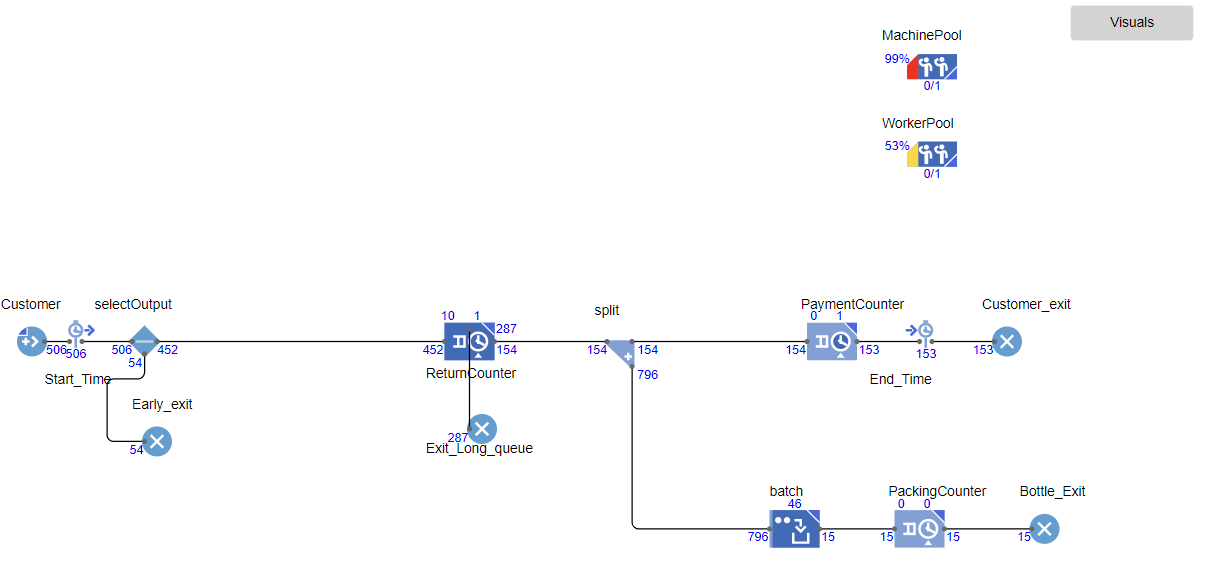
**Question 2 Part 2 (15M)**

**Ans)** From my understanding about the model, this is a zoomed version of what is exactly happening in the Level 1 of the support centre. From the above, we can understand the structure of the box which shows how the number of requests are being shown and also how the workers are categorized while they are responding to a request and while they are not.

Coming to the model logic, it shows that whenever a request is received, it goes through a timer timemesurestart the starts recording the time from this position till the others. It then comes across a select output that considers whether the request is a call or email/faxes and through this it is sent to different directions. Both are sent to a queue, however the calls have an instant delay time of a maximum of 2 hours. If the calls go past a certain delay time, the customers will hang up and the request would go into a sink called expired. If not, they are moved forward. Here we see the first TimeMeasureEnd that records the time from when the request for call was started until the support responses. There are two ‘’holds” attached here which I believe reflects the waiting for support to pickup. The requests then further flow down towards the TimeMeasureStart1, which starts recording another time interval between the call or email/faxes being responded till a further point. The requests are then forwarded to a sieze where they are held while a resource (support staff) is linked to that request. This is then further transferred to the delay called ‘processing’. Now moving forward, if the request is of the complexity that Level1 can solve, it uses the select output that sends it to delay three where the resource is released. Here we find another timeMeasureEnd1, that captures the finishing time of the request from its start just before the sieze. Finally, this request is then sent to sink with the remark of done. However, the select output does also pave way towards another way where the complexity is higher than what Level1 can solve. This is why there is another selectoutput that sends the request further to either Level2 or Level3 based on how the support operator ranked it as. They can be seen with a route delay 1 and 2 further down the model flow.

**Question 3: Simulation Model (140 Marks)**

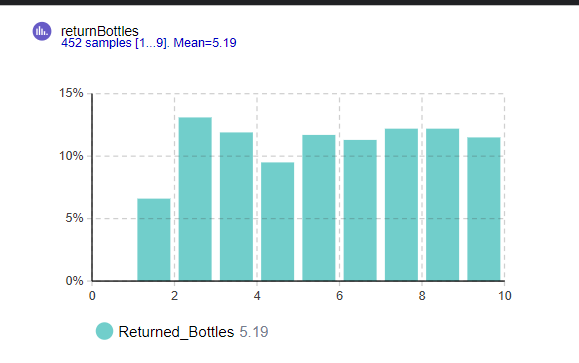
**Question 3 Part (a) (40M)**

**Ans)** 

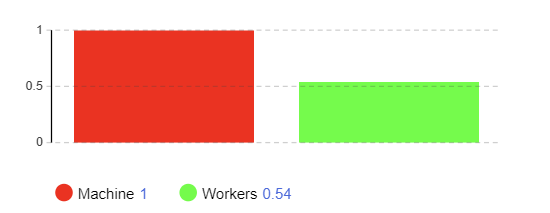
Based on the model shown above. The utilization of resources including MachinePool and WorkerPool is coming out to be 99% and 53% respectively. The number of workers who left the store due to not having bottles amount to 54 whereas customers who left due to a long queue is amounting to 287. Lastly the total number of packed batches are 15.

**Question 3 Part (b) (20M)**

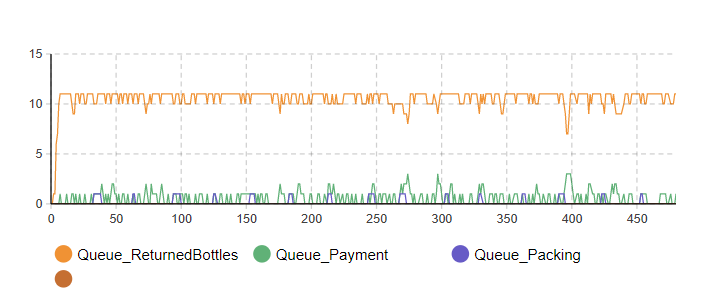
**Ans)** Find screenshots of visualizations below :



*Figure 1 : Histogram for Returned Bottles*



*Figure 2 : Bar Chart for Resource Utilization*



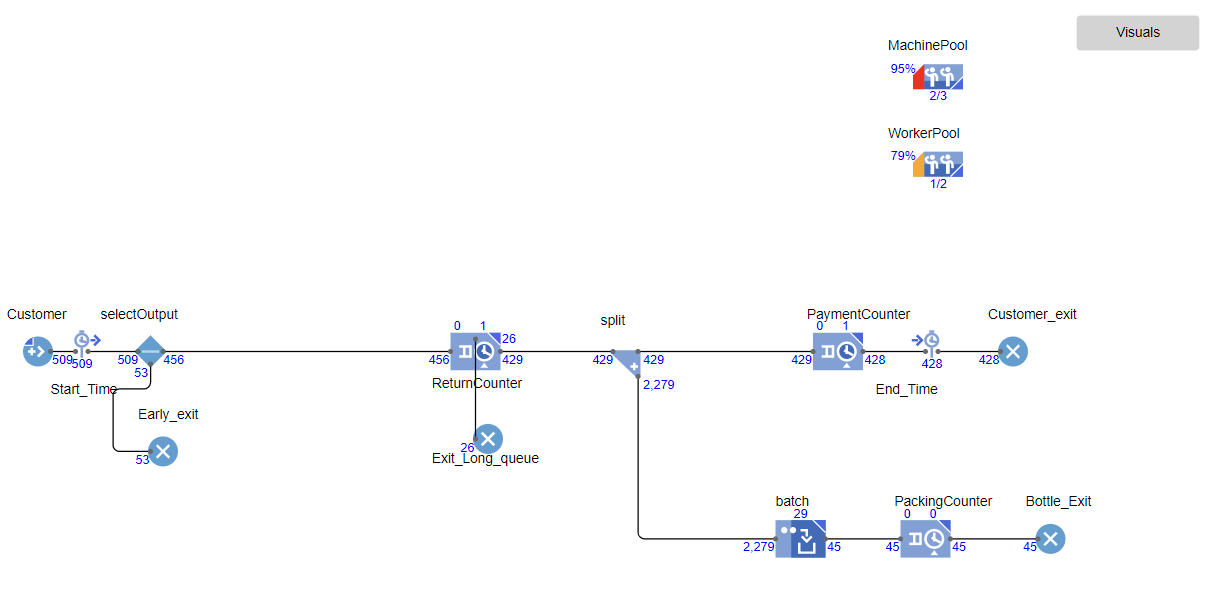
*Figure 3 : TimePlot for Size of the Queues*

**Question 3 Part (c) (5M)**

**Ans)** The results show a highly exhaustive utilization of both resources including the bottle machine as well as the worker. This does not show a good sign since the work being done by both resources is almost 100% however, there is still long queues around each counter. At the same time, due to a very long wait, the business is losing a lot of customers who left the store without returning the bottles. The ratio for customers leaving the store due to the wait time is more than how many customers are actually being paid which is a very negative sign therefore I do not think it’s a possible scenario for the shop.

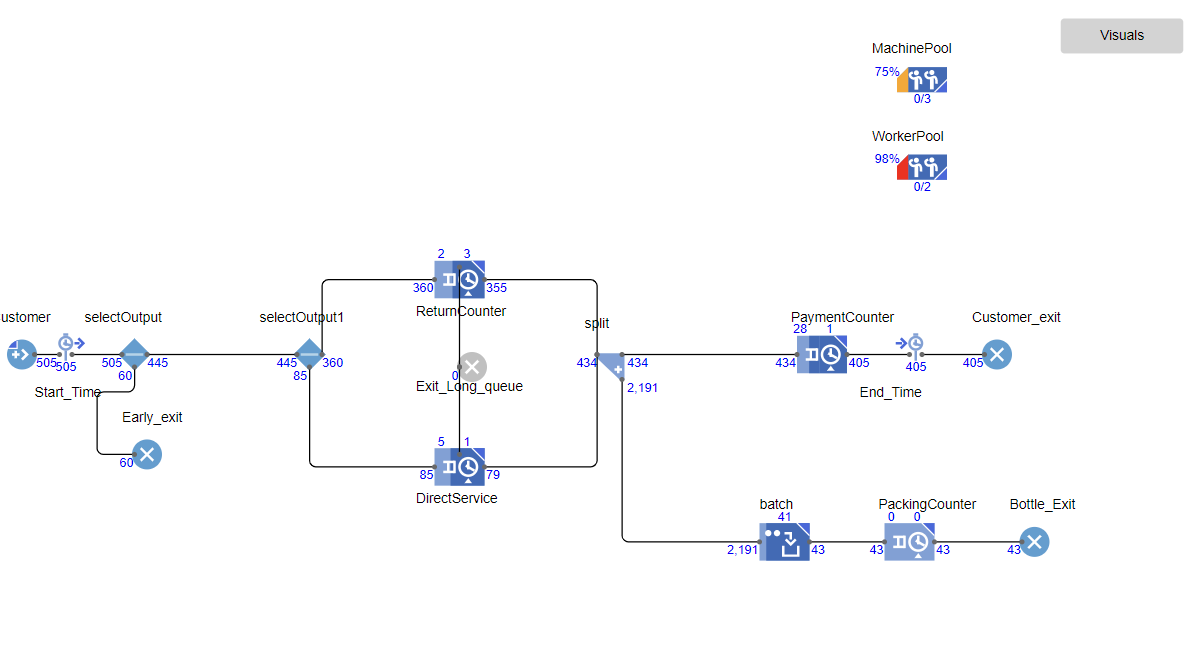
**Question 3 Part (d) (5M)**

Ans) By running multiple experiments in which I increased the number of machines and workers one by one, I can suggest that based on the current model setup, the ideal number of machines and workers would be MachinePool = 3 and WorkerPool = 2. This is because in this scenario, I tried to balance good results with cost effectiveness. I increased the resources enough that they can satisfy a great deal of demand however do not impact too much on the cost. Based on this, the machines are being exhaustively utilized but not compared to having only one. The end results show that they are being utilized 95% whereas the workerpool at 79%. The amount of customers leaving the store due to a long queue has substantially decreased to only 26 which is also a very low number in proportion to the customers that got paid which amounts to about 428. I can potentially increase the machines to 4 to eliminate the issue of people leaving the store due to a long queue. However, this may increase the cost as bottle machines wouldn’t be cheap. Considering this, letting 26 customers go seems like a better option. Below is a screenshot of the results.



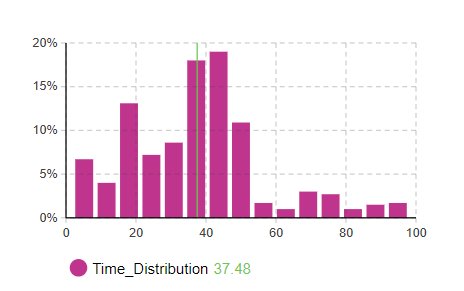
**Question 3 Part (e) (20M)**

**Ans)** Based on the results after making the above given changes, we can see that the MachinePool is being utilized at 75% whereas the WorkerPool is still exhaustively being consumed at 98%. This does not show a good sign as this may mean we require atleast another worker to be able to reducen this down to complete all tasks. From the Other results we do see that there are no exits due to long queues, however the queues still show a long wait time particularly at the payment counter. There are still 28 people in the queue to get their money back. This shows that the machines are performing well (since no long queues at return counter) and so are workers in the initial phase (no long queues at direct service) but it just gets a bit too crowded when the move to the payment section. Therefore I believe increasing one more worker may solve the queues at payment as well.



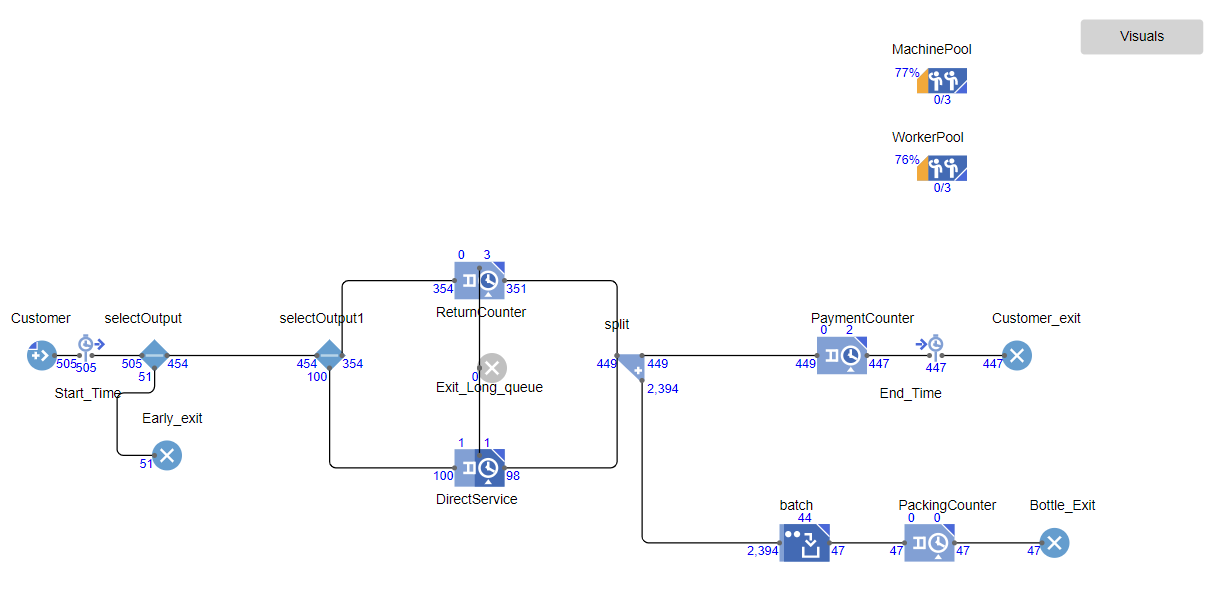
**Question 3 Part (f) (20M)**

**Ans)** After implementing the time measurement tools, we get to understand that there is a wide spread of customers taking different time from starting their process by entering the store till making the payment and making an exit. On the y axis we see the percentage of customers whereas on the x axis we can see the amount of time they took from starting to ending point. From this we can understand that a big chunk of customers fall between the range of 20 mins to 50 Mins. This is the reason why the mean is coming out to be 37.48 Mins. The visualization also shows some percentage that left the store very quickly while some even stayed near 90 mins mark. The obersvation taken from this is that the store is working quiet efficiently with majority taking a time interval of 37 mins however, there are some instances when the store was not able to perform well which kept some customers waiting for more than 90 mins. This could be the peak time when there were huge queues even though the workers were working at full utilization, they were not able to get them out that quickly.



**Question 3 Part (g) (5M)**

**Ans)** As the model was already performing well with the last experiment of 3 Machines and 2 workers, we will not need to make to many changes to reduce wait time. The only long queue being formed is at the payment counter. As we were able to find that there were 28 customers in queue of the payment counter, therefore to reduce this we do not need to increase the machines as firstly, they are not working at exhaustive utilization and secondly, they are not required at this counter. The resource to increase is workers but by only 1. This will allow to wait time to reduce as there will no longer be queues being formed at payment counter as well. A screenshot of the model results are shown with 3 MachinePool and also 3 WorkerPool.



**Question 3 Part (h) (15M)**

**Ans)** Answered on Model

**Question 3 Part (i) (10M)**

**Ans) Change #1 :**

We can possible add time measurement tools at different stages of the model. For example, we can add one from start till giving the bottles, or from giving bottles till the exit. This will allow to break the model down into phases and analyze them individually. This way improvement can be found only for areas that are not performing well while performing no changes to areas already working at good efficiency.

**Change #2 :**

We can potentially set an automation at preference of whome to submit the bottles to. At the moment the model implements a system that customers prefer 20% direct service and 80% machine service. However, this may mean that there will be times when machine customers have a long queue due to its preference but the direct service customers have no one at the counter. Therefore, implementing an automation that if the queue reaches a certain number at machine, they should automatically be sent to direct service irrespective of their preference. This will reduce the overall queue sizes and also improve efficiency.

**Change #3:**

Just to make the model more detailed reflecting real life even more, we can potentially create two services at the payment counter. This would mean one is a machine that gives customers, while the other is a worker. Just the same way there would be a select output that considers the preference of people who wants to get paid through a machine or by a direct worker. This would reduce pressure on the worker resource and may even mean that we no longer need another extra worker. However another resource (maybe called an ATM) can be created which would handle paying the customers autonomously.