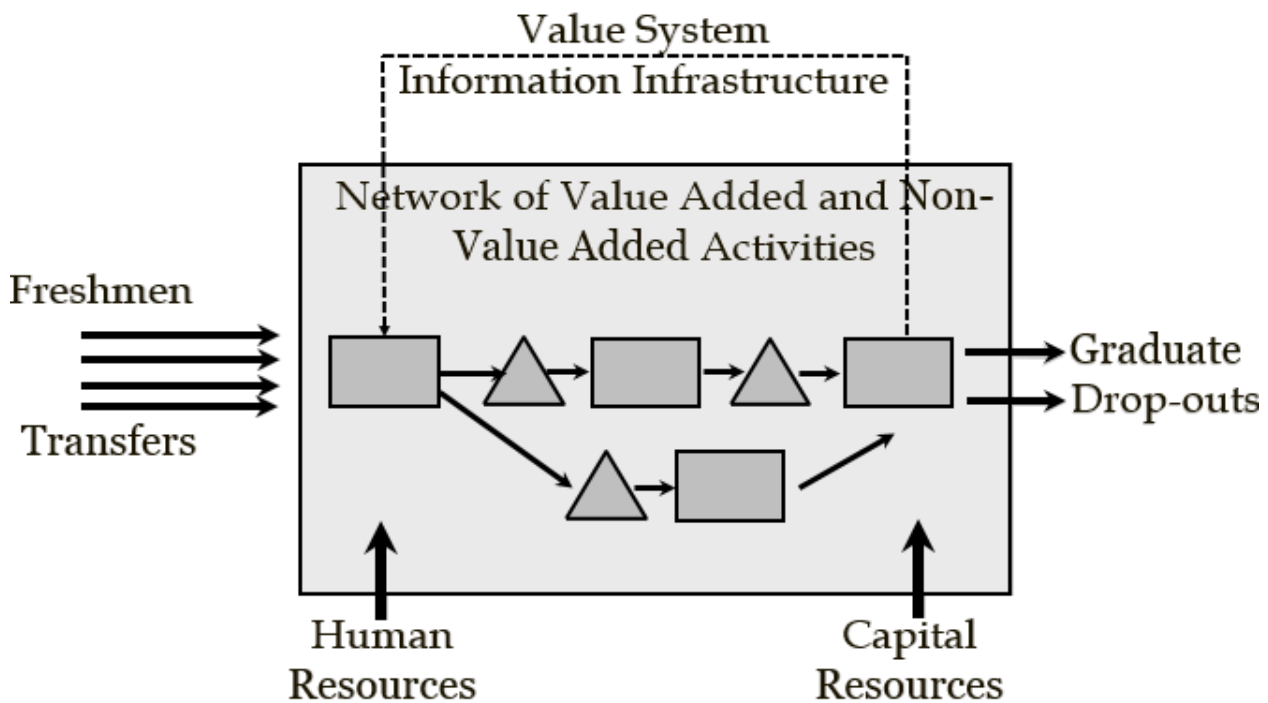


Business Process Flow Engineering



Problem-Based Learning Study Guide

Ardavan Asef-Vaziri

Systems and Operations Management

David Nazarian college of Business and Economics

Chapter 1

Process Flow Analysis

The Little's Law: The Core Concept in Business Processes Engineering

Eyes must be washed; to see things differently. Sohrab Sepehri, Persian Poet, 1928 – 1980.

The key problems of this chapter can be accessed at

<http://www.csun.edu/~aa2035/CourseBase/Process/Assignments/AssignmentAndPractice/ProcessFlowKeyProblems.pptx>

The recorded lectures – if any- and the excel worksheets- if any- are embedded inside the PowerPoint slides. By clicking on them the mp3/mp4 files and excel pages will open

It has been reported that U.S. residents spend close than 40 billion hours per year waiting in lines. Even at a cost of \$15 per hour, this adds up to \$600 billion per year. Add that to the cost, the frustration, and the irritation people feel when waiting in lines. In this chapter, we study the relationship between Throughput (the rate at which people or product enter and leave a system), Flow Time (the amount of time they spend in the system), and Inventory [the number of flow units (people or products) in the system].

Problem1. The Coffee Shop. A manager of a local coffee shop close to the Reseda and Plummer intersection has realized that during busy hours, the entrance door on average opens 2 times per minutes to let in 2 customers per minute. In a stable system, if 2 customers enter per minute, two customers should leave per minute. In the short run, we may have variations, but over a long period, input should be equal to output. It is possible that in 5 minutes, 10 customers come in and 12 customers go out. However, over a long period (even a day), the output cannot exceed input, because how can the difference be generated? Similarly, the output cannot be less than input, because, over a long period, there will be no room in the coffee shop (or even in a large stadium). It is possible that in 5 minutes, 10 customers come in and 8 customers go out. However, over a full day, input cannot exceed output. In stable systems, flow units come in as input, leave the system as output, and input per unit of time is equal to output per unit of time. In our example, if, on

average, two customers come in per minute, then two customers, on average, should leave the system.

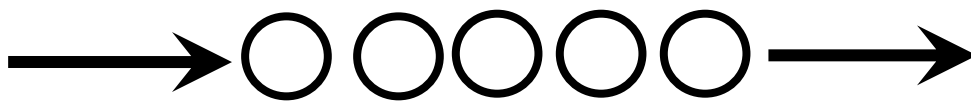
a) What is the throughput of the coffee shop?

Every minute, 2 customers enter and two customers leave. $R = 2$ per minute.

Throughput (R) is the average flow rate in a stable system where the average input is equal to average output over an extended period. Throughput is expressed as a number with a time unit attached to it (e.g. per minute, per hour, per day, per month, etc.).

Inventory (I) is the number of flow units in the system (e.g. customers in a coffee shop, students at CSUN, cars for sale on a lot, etc.).

During the corresponding hours in the coffee shop, on average, there are 5 customers in the store (system), 4 are waiting in line (buffer) to order, and one is with the server.



b) What is the inventory (I) in the coffee shop?

On average, there are 5 customers in the coffee shop. Inventory is 5. $I = 5$.

c) How long on average is a customer in your coffee shop (inventory expressed in the unit of time)?

Flow Time (T) is the time it takes an input to become an output. It is the time a flow unit spends within a system.

Before a customer steps in the system, there are 5 other customers in the system. At the instant she steps into the line for service, one fully served customer leaves the system. That is, there are always 5 people in the system. Our incoming customer at the beginning has 4 people in front of them, then 3 in front and 1 behind, then 2 in front and 2 behind, then 1 in front and 3 behind, then no one in front and 4 behind her. At the instance when she steps out of the system, just in the fraction of second stepping out, she looks over her shoulder.

How many people are behind her? 5.

At what rate did they come in? 2 per minute.

How long does it take 5 people to come in if they arrive at the rate of 2 per minute?

1 minute

2 customers

How many minutes (T) 5 customers

$$T = (1 \times 5) / 2 \rightarrow T = 2.5 \text{ minutes.}$$

On average, a customer spends 2.5 minutes in the coffee shop; flow time (T) is 2.5 minutes. Each customer enters the coffee shop, spends 2.5 minutes on average and then leaves. In the above computation, the flow time (T) is defined in minutes because R was in minutes. R carries a time unit with it, i.e., 2/minute., $2(60) = 120/\text{hour.}$, and, if a day is 8 hours, it can also be expressed as $2(60)(8) = 960/\text{day}$. **However, remember: inventory (I), does not carry a time unit, it is always a number.**

Now, let us generalize

1 time unit R flow units

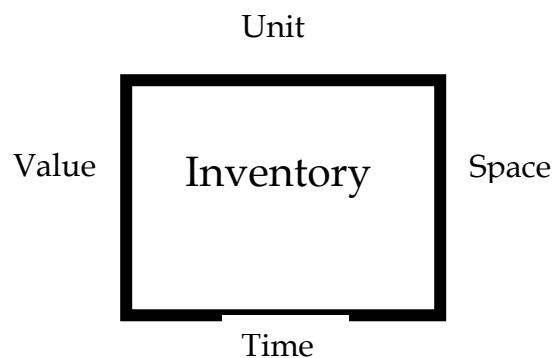
How many time units (T) I flow units

$$T = I/R \rightarrow RT = I.$$

The Little's Law is expressed as Throughput x Flow Time = Inventory.

$$R \times T = I \text{ or } T = I/R \text{ or } R = I/T$$

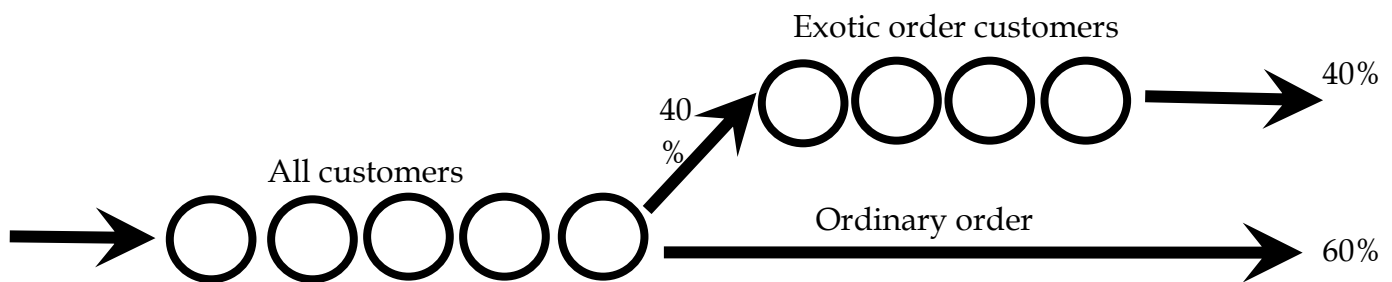
A Fundamental Insight. Note, that the Little's Law, $T = I/R$, is nothing more than a unit conversion, converting numbers into time. It turned **5 units** of inventory into **2.5 minutes** of inventory. Suppose we have **100** units of item A, and **1000** units of item B. What item do we have more of? In the count dimension, item B has a higher inventory. Suppose we use 4 units of item A per day ($R_A = 4/\text{day}$), and 200 units of item B per day ($R_B = 200/\text{day}$). In the time dimension, we have ($T = 100/4$), **25** days inventory of item A, and ($T = 1000/200$) **5** days inventory of item B. On the time dimension, the inventory of item A is larger than the inventory of item B. It takes more time to consume the inventory of item A, compared to that of consuming item B. In addition to measuring inventory in count and time dimensions, we can measure it on the third dimension of value. How much money is invested in the inventory of each item? There is even a fourth dimension. How much space (in a warehouse)?





If you have had difficulty understanding this problem, we encourage you to watch the recorded lecture at <https://youtu.be/bEM6z27bdWA> The PowerPoint slides are at [Process Flow Measures Basics](#).

Problem 2. The Coffee Shop – Extended. Let us go back to our prototype example. Suppose after the initial waiting line in the coffee shop, to pay or get standard items such as a black coffee, there is a second waiting line for exotic hot and cold beverages such as lattes, cappuccinos, etc. Suppose R is still 2 customers per minute, and still, on average, there are 5 customers in the first line, waiting to pay for their order and get their non-exotic coffee or other items already available on the shelves (made-to-stock, or MTS items). In addition, suppose 2 out of the 5 customers place their order and then go on to wait in the exotic order line which needs preparation (make-to-order or MTO items). Therefore, 40% of the customers place exotic orders. A pictorial representation of the process is illustrated below.



a) What is the flow time of the first line?

Each customer spends 2.5 minutes in the first line (as determined in the first problem).

b) What is the throughput of the customers in the exotic order line?

The throughput of the second line is 40% of 2 customers per minute. That is $0.4(2) = 0.8$ customers per minute, or $0.8(60) = 48$ customers per hour, or $0.8/60 = 0.13333$ customers per second. **Note that it is not 40% of 5, but 40% of 2.** Indeed, 40% of those 5 people will go to the second line, but there are two points to mention. First, 40% of the inventory will go to the second line, but not suddenly. They will go at the rate of $0.4(2) = 0.8$. Second, there are always 5 people in the first line, no matter the rate at which they go into the second line.

c) What is the inventory (I) of the exotic order line?

Inventory of the second line, as given in the graph, is 4. It is given by the problem. It is always 4, no matter whether we state R in minutes, seconds, or hours.

d) What is the flow time of a person who orders regular coffee?

We have already done this computation. It is $T = I/R = 5/2 = 2.5$.

e) What is the flow time of a person who orders exotic coffee?

The person who orders exotic coffee has already spent 2.5 minutes in the first line. For the time in the second line, we again apply Little's Law where $R=0.8$ and $I=4$.

$R \times T = I \Rightarrow 0.8 \times T = 4 \Rightarrow T=5$ minutes (since T is 0.8 per minute.).

T (first line) = 2.5, T (second line) = 5 minutes.

T (exotic order) = 2.5 + 5 = 7.5 minutes.

f) What is the flow time of a prototype customer (flow unit)?

In general, a prototype flow unit does not exist in reality. It is a melted version, a weighted average, of all of the flow units. In this example, it is a customer who is 60% a person who gets an MTS item and 40% a customer who gets an MTO item.

There are several ways to answer this question. We show three of them:

Procedure 1 (Micro) -

60% simple order: $T = 2.5$,

40% exotic order: $T=2.5+5= 7.5$.

An average customer is 60% a simple order person (2.5 minutes.), and 40% an exotic order person (7.5 minutes).

$T = 0.6(2.5) + 0.4(7.5) = 4.5$ minutes. We can use the SUMPRODUCT function in excel if the summation of weights is equal to 1. If it is not, then we will use SUMPRODUCT and then divide it by the SUM of the weights.

Procedure 2 (Micro/Macro) -

Everyone goes through the first process and spends 2.5 minutes.

60% spend no additional time and leave. 40% spend 5 additional minutes.

$0.6(0)$ ordinary order customers + $0.4(5)$ exotic order customers = 2 minutes

2.5 (every customer) + 2 (exotic orders) = 4.5 minutes.

Procedure 3 – (Macro- the most logical)

The throughput of the system is 2 per minute. There are 9 flow units in the system (5 in the first and 4 in the second line).

$$R \times T = I \rightarrow 2 \times T = 9 \rightarrow T = 4.5 \text{ minutes}$$

The throughput of the coffee shop (system) is 2 per minute, or 120 per hour, or 720 per day (assuming 6 busy hours per day), or 1/30 per second. However, inventory in the system is **always 9**.



If you have had difficulties, understanding this problem we encourage you to go through problems E1 and E3 on the following slides. [Assignment Process Basics Set1 Problems](#)

Problem 3. David Nazarian College of Business and Economics

(DNCBE). Over the past 10 years, on average, there were 1600 incoming students per year at DNCBE. On average, the headcount of students enrolled over the same period was 7200.

a) On average, how long does a student spend in DNCBE (average time to graduation)?

$$RT=I \rightarrow 1600T = 7200 \rightarrow T = 4.5 \text{ years}$$

While at first glance, this number appears to be better than the surrounding programs, we should note that it is for transfer and freshmen students combined. In fact, 40% of the students are freshmen, and the rest are transfer students. In addition, according to the data from CSUN *by*, the time to graduation (the duration of time from the first day of starting the college education to the day of graduation) for freshmen students on average, is 2.25 times of the transfer students.

b) How long does it take freshman students to graduate? How long does it take transfer students to graduate?

T = Time to graduation for a transfer student (60% of all students).

2.25T = Time to graduation for a freshman (40% of all students).

Average time to graduation for all students = 4.5 years.

All students comprise 40% freshman and 60% transfer, and this can be represented by the following equation:

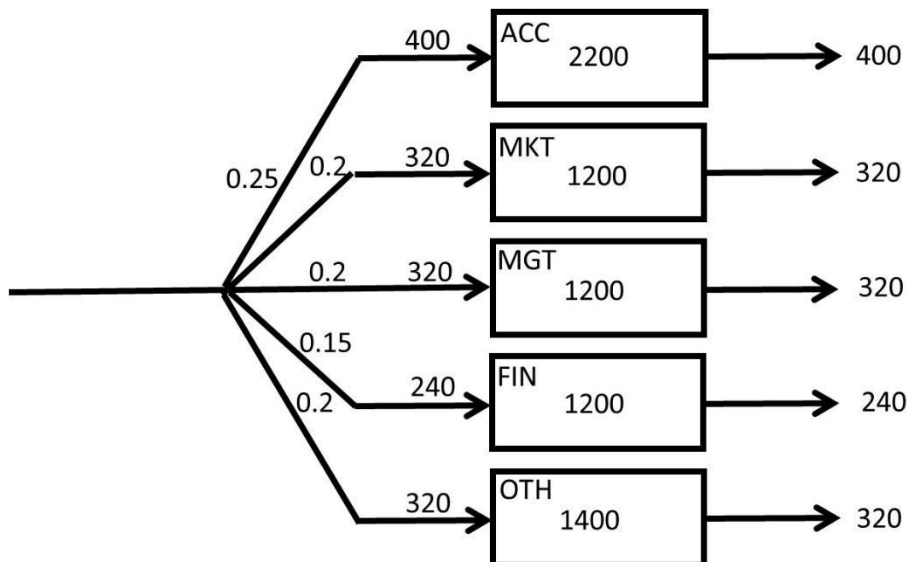
$$0.6T + 0.4(2.25T) = 4.5 \text{ years} \Rightarrow 0.6T + 0.9T = 4.5 \text{ years} \Rightarrow 1.5T = 4.5 \text{ years} \Rightarrow T = 3 \text{ years}$$

Time to graduation for a transfer student = 3 years.

Time to graduation of for a freshman student = $3(2.25) = 6.75$ years.

Out of the 1600 incoming students per year, on average, 25% are accounting majors, 20% marketing, 20% management, 15% finance, and the rest are other majors. The business school has 2200 accounting students, 1200 marketing students, 1200 management students, and 1400 students in all other majors except finance. The rest of the 7200 headcounts of the college students are in the finance department.

c) On average, how many accounting students graduate each year?



If 25% of the 1600 students enter and exit per year, then the throughput for accounting department is: $R_{ACC} = 0.25(1600) = 400$ per year

d) How long, on average, does it take a finance student to graduate (what is the flow time of a finance student)?

$$R_{FIN} = 0.15(1600) = 240 \text{ /year}$$

$$I_{FIN} = 7200 - (2200 + 1200 + 1200 + 1400) = 1200 \text{ students.}$$

$$T_{FIN} = 1200 / 240 = 5 \text{ years}$$

e) On average, which major takes the longest time to graduate and which major takes

the shortest time to graduate?

Major	%	R	I	T
ACC	0.25	400	2200	5.5
MKT	0.2	320	1200	3.75
MGT	0.2	320	1200	3.75
FIN	0.15	240	1200	5
OTH	0.2	320	1400	4.375
	1	1600	7200	

Longest = Accounting (5.5 years)

Shortest = Marketing and Management (3.75 years).

f) Using part (a), prove that your computations in part (g) are correct.

We have time to graduation for all the five groups and their relative weights.

$\text{SUMPRODUCT } 0.25(5.5) + 0.2(3.75), 0.2(3.75), 0.15(5), 0.2(4.375) = 4.5$.

We had the information that 40% of the incoming students are freshmen, and the rest are transfer students. Sadly, we have the additional information that there are 12.5% dropouts, where 10% are freshmen and the rest transfers. The time to graduate for freshmen students, on average, is 2.25 times that of the transfer students. The average time that the dropouts spend in DNCBE is 1.5 years. The average, time for transfer students that dropout is 0.75 years.

How long does it take freshman students to graduate?

T: time to graduation for a transfer student @60%-2.5% = 57.5%.

2.25T: time to graduation for a freshman student @40%-10% = 30%.

1.5: Time to drop-out for a freshman@10%

0.75: Time to drop-out for a transfer@2.5%

$0.025(0.75) + 0.1(1.5) + 0.575T + 0.3(2.25T) = 4.5$

$0.16875 + 0.575T + 0.675T = 4.5$

$1.25T = 4.5 - 0.16875 = 4.33125$

$T = 3.465$

Time to graduation for a transfer student = 3.465

Time to graduation for freshman student = 2.25T

Time to graduation for freshman student = $2.25(3.465) = 7.8$



If had difficulty understanding this problem, we encourage you to watch the recorded lecture of a very similar problem posted on <https://youtu.be/gFNYXGye4Jo> The PowerPoint slides are on the Process Key Problems posted at the beginning of this chapter.

Problem 4. CSUN&UCLA Fresh Juice Stands. A recent CSUN graduate has opened up a cold beverage stand, CSUN-STAND, in Venice Beach. She works 8 hours a day and observes that, on average, there are 320 customers visiting the CSUN-STAND every day. She also notes that, on average, a customer stays in line at the stand for 6 minutes.

a) How many customers, on average, are waiting at CSUN-STAND?

- A) 4
- B) 2.75
- C) 3.75
- D) 3.25
- E) 4.25

She is thinking about running a marketing campaign to boost the number of its customers. She expects that the number of customers will increase to 480 per day after the campaign. She wants to keep the line short at the stand and hopes to have only 2 people waiting, on average. Thus, she decides to hire an assistant.

b) What is the average time a customer will wait in the system after all these changes?

- A) 4 min.
- B) 3 min.
- C) **2 min.**
- D) 1 min.
- E) none of the above

c) After the marketing campaign, a recent UCLA graduate opens up a competing cold beverage stand. The UCLA graduate is not as efficient as the CSUN graduate is, so customers must stay an average of 10 minutes at “UCLA-STAND.” Suppose there is an

average of 3 customers at UCLA-STAND. The total number of customers for both CSUN- and UCLA- STANDs remains at 480 per day, as it was after the marketing campaign. Now it is divided between the CSUN- STAND and UCLA- STAND. How many are fewer customers visiting CSUN-STAND per day?

- A) 177 customers per day
- B) 144 customers per day**
- C) 166 customers per day
- D) 155 customers per day
- E) 133 customers per day



If you have had difficulty understanding the problem, we encourage you to watch the recorded lecture of a problem very similar to this problem, recorded at

https://youtu.be/QjS_K1zcmw0 The PowerPoint slides are on

the Process Key Problems posted at the beginning of this chapter.

Problem 5. Academic Technology Help Desk.

Original Process: A help-desk administrator at CSUN receives 2000 emails per month requesting assistance. Assume there are 20 working days per month. On average, 100 unanswered emails are waiting in the mailbox of Administrator A (Admin-A).

What is the average flow time for handling a request?

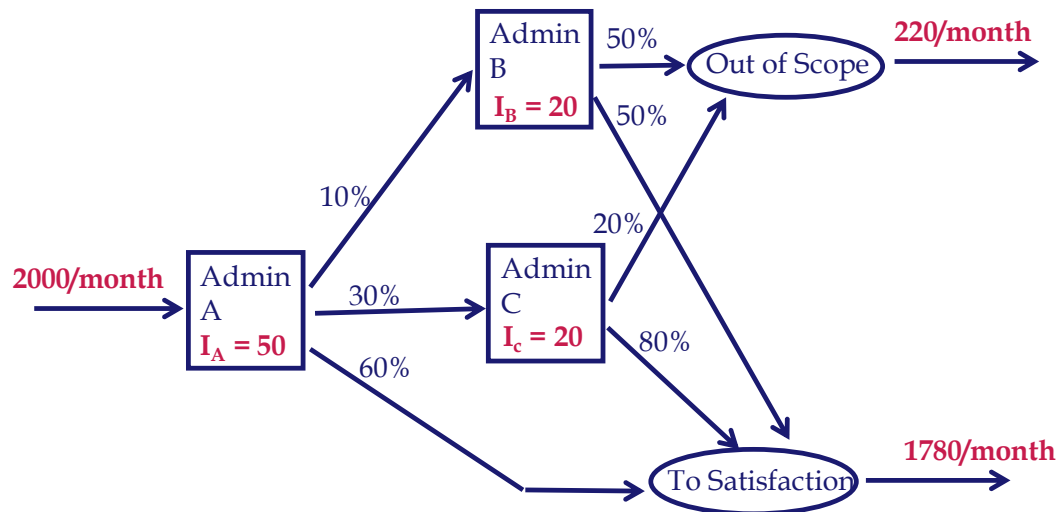
$$R = 2000 / 20 = 100 \text{ per day.}$$

$$RT = I \rightarrow 100T = 100 \rightarrow T = 1 \text{ day}$$

New Process: To help reduce the number of unanswered emails, CSUN hired two more administrators to the help-desk department. In the new process, Administrator A (Admin-A) responds to 60% of the emails **to-satisfaction**, without further investigation. 10% of the remaining emails are forwarded to Administrator B (Admin-B), and the rest to Administrator C (Admin-C). On average, 50 emails in the email box of Admin-A, 20 in the email box of Admin-B, and 20 are waiting in the email box of Admin-C. Additionally, Admin-C responds to 80% of the emails sent to him to-satisfaction, and he considered the rest of the emails as **out-of-scope** of the help-desk responsibilities. Admin-B responded to half of the emails sent to him to-satisfaction, and the rest of the emails out-of-scope.

- a) Compute average flow time
- b) Compute average flow time at Admin-A
- c) Compute average flow time at Admin-B
- d) Compute average flow time at Admin-C

- e) Compute average flow time of an out-of-scope response
- f) Compute average flow time of a response to-satisfaction



a) Compute Average Flow Time

$$I = I_a + I_b + I_c = 50 + 20 + 20 = 90$$

$$R = 2000 \text{ per month or } 2000/20 = 100 \text{ per day or } 100/8 \text{ (working hours)} = 12.5 \text{ per hour}$$

I is always 90. It does not carry a time unit.

$$T = I/R = 90/12.5 = 7.2 \text{ hours. It is in hours because } R \text{ was calculated per hour.}$$

b) Compute Average Flow Time for Admin-A:

$$\text{Throughput } R_a = 12.5 \text{ emails/per hour}$$

$$\text{Average Inventory } I_a = 50 \text{ emails}$$

$$T_a = 50/12.5 = 4 \text{ hours with Admin-A}$$

c) Compute Average Flow Time for Admin-B:

$$\text{Throughput } R_b = 0.1 (12.5) = 1.25 \text{ emails/hour}$$

$$\text{Average Inventory } I_b = 20 \text{ emails}$$

$$T_b = 20/1.25 = 16 \text{ hours with Admin-B}$$

d) Compute Average Flow Time for Admin-C:

Throughput $R_C = 0.3 (12.5) = 3.75$ emails/hour

Average Inventory $I_C = 20$ emails

$T_C = 20/3.75 = 5.33$ hours with Admin-C

Additional Information Derived from the Problem:

One flow unit at a very macro level = Email

2000 flow units/per month at very micro level = each specific email

In the Original Process, there are two flow units: To-satisfaction and Out-of-scope

In the New Process, there are five flow units: To-satisfaction (Admin-A), To-satisfaction (Admin-B), To-satisfaction (Admin-C), Out-of-scope (Admin-B), and Out-of-scope (Admin-C)

To-satisfaction (Admin-A): A

To-satisfaction (Admin-B): A, B

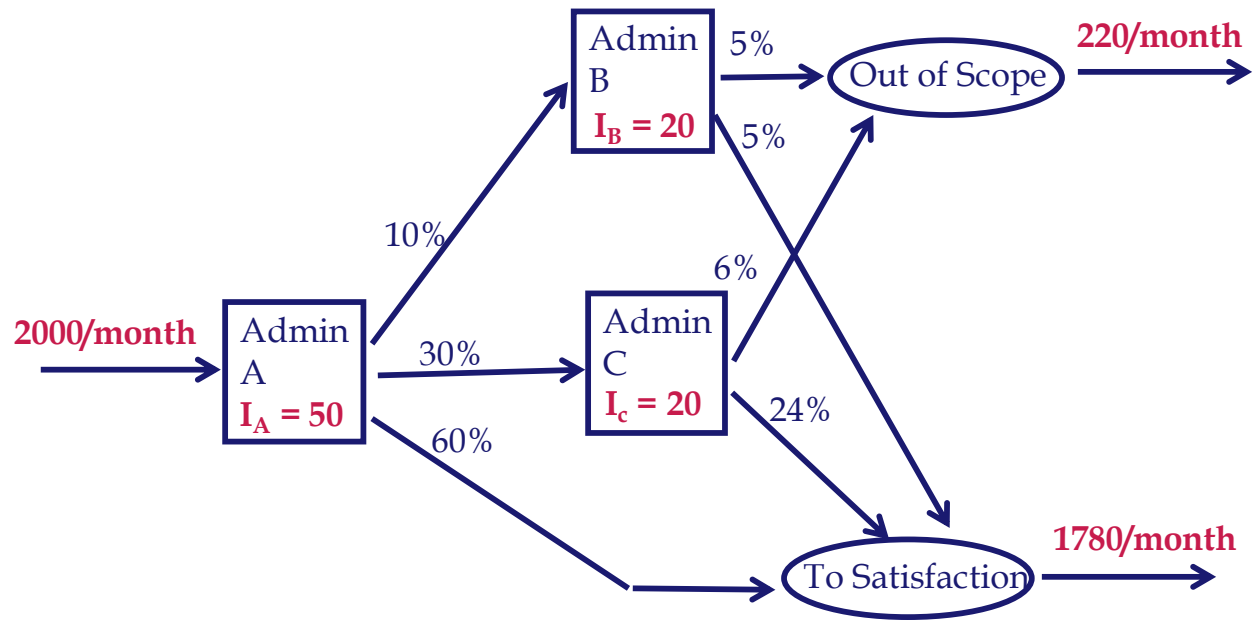
To-satisfaction (Admin-C): A, C

Out-of-scope (Admin-B): A, B

Out-of-scope (Admin-C): A, C

$T_A = 4$ hours, $T_B = 16$ hours, $T_C = 5.33$ hours.

We also need percentages of each of the five flow units



e) Compute Average Flow Time of an Out-of-scope Email:

Out-of-scope (Admin-B): A, B $\rightarrow 4+16 = 20$ Out-of-scope -A: 5%

Out-of-scope (Admin-C): A, C $\rightarrow 4+5.33 = 9.33$ Out-of-scope -B: 6%

$$[0.05(20) + 0.06(9.33)] / (0.05+0.06) = 14.18$$

f) Compute Average Flow Time of a Response to-satisfaction:

To-satisfaction -A: A $\rightarrow 4$ @ 60%

To-satisfaction -B: A, B $\rightarrow 4+16=20$ @ 5%

To-satisfaction -C: A, C $\rightarrow 4+5.333=9.333$ @ 24%

$$\frac{0.6}{0.89}(4) + \frac{0.05}{0.89}(20) + \frac{0.24}{0.89}(9.33) =$$

$$= 6.34$$

Check our computations:

Average flow time of an application

$$0.89(6.34) + 0.11(14.18) = 7.2$$



If you had difficulties understanding this problem, we encourage you to go through a very similar problem. The PowerPoint slides are in the Process Key Problems posted at the beginning of this chapter. Part1 of the recorded solution is at <https://www.youtube.com/watch?v=TauGBb5xbVs> The rest of the recorded solution is at <https://youtu.be/pyFq8JDHljM>

Problem 6. Northridge Hospital. On Average 80 patients per hour arrive at a hospital emergency room (ER).

All patients first register through an initial registration process. On average, there are 9 patients in the registration waiting line (RgBuff). The registration process takes 6 minutes. A triage nurse practitioner then examines the patients.

On average, there are 2 patients waiting in the triage waiting line (TrBuff). The triage classification process takes 5 minutes.

On average, 91% of the patients are sent to the Simple-prescription process and the remainder is sent to Hospital-admission.

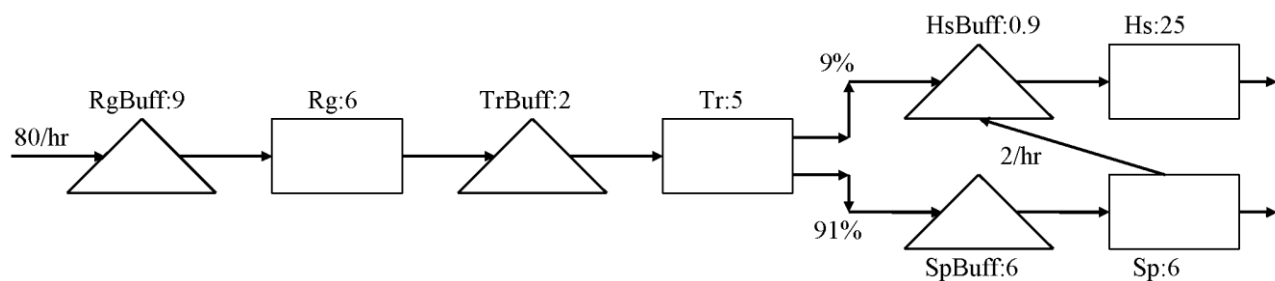
On average, 6 patients are waiting in the simple prescription waiting line (SpBuff) in front of this process.

A physician spends 6 minutes on each patient in the Simple-prescription process.

Unfortunately, on average, 2 patients per hour are sent to the Hospital-admission buffer (HsBuff) after being examined for 6 minutes in the Simple prescription process.

On average, 0.9 patients are waiting in the HsBuff.

A physician spends 25 minutes on each patient in the Hospital-admits process.



	R/hr	R/min	I	T
RgBuff	80	1.33333	9	
RgProc	80	1.33333		6
TrBuff	80	1.33333	2	
TrProc	80	1.33333		5
SpBuff	72.8	1.21333	6	
SpProc	72.8	1.21333		6
HsBuff	7.2	0.12	0.9	
HsProc	7.2	0.12		25

	R/hr	R/min	I	T
RgBuff	80	1.33333	9	
RgProc	80	1.33333		6
TrBuff	80	1.33333	2	
TrProc	80	1.33333		5
SpBuff	72.8	1.21333	6	
SpProc	72.8	1.21333		6
HsBuff	9.2	0.15333	0.9	
HsProc	9.2	0.15333		25

7.2/hour, and 0.12/minute on the left table are incorrect. Why?

Because 7.2 is 9% of 80. However, the input in HsBuffer and HsProc is not 7.2.

Is it more or less?

It is more because 2 patients per hour are directed from SpProc (simple prescription process) to HsBuff (hospital admission waiting line).

Therefore, the numbers adjusted on the right table are $7.2+2=9.2$ per hour or 0.15333 per minute.

Now if we look at the rows of the table, each contains two elements of the Little's law, and finding the third one is trivial. However, we need to note that since flow time (T) is stated in minutes, we cannot use R per hour, but R per minute. The rest is just $T=I/R$ and $I=RT$, as summarized in the following table:

	A	B	C	D	E	F
1		R/hr	R/min	I	T	
2	RgBuff	80	1.33333	9	6.75	=D2/C2
3	RgProc	80	1.33333	8	6	=C3*E3
4	TrBuff	80	1.33333	2	1.5	=D4/C4
5	TrProc	80	1.33333	6.66667	5	=C5*E5
6	SpBuff	72.8	1.21333	6	4.94505	=D6/C6
7	SpProc	72.8	1.21333	7.28	6	=C7*E7
8	HsBuff	9.2	0.15333	0.9	5.86957	=D8/C8
9	HsProc	9.2	0.15333	3.83333	25	=C9*E9

Now we have everything needed to compute the average flow time.

Now, that is not to add the times, because 100% of flow units do not go through each and

every process.

It is instead required to add the inventories, which is equal to 43.68.

Now we have a black box, and for the time being, we do not care what is happening inside it. However, we know that 80 flow units per hour enter and exit this black box, and on average, there are 43.68 flow units inside the box. Then by virtue of the Little's law,

$$RT=I \rightarrow 80T = 43.68 \rightarrow T=0.546.$$

0.546 What?

Hour, because R is stated in terms of hours. If you want it in minutes, just multiply it by 60.

Alternatively, you may have R= 1.33333 per minute, instead of 80 per hour. Therefore,

$$1.33333T=43.68 \rightarrow 32.76 \text{ minutes.}$$

What is the flow time of the patients with a simple prescription? In addition, what is the flow time of the patients who are hospital admitted? There is only one type of simple prescription patient. Nevertheless, the hospital admitted patients are of two types; those admitted directly, and those who wrongly go through the simple prescription process and then are redirected to hospital admission.

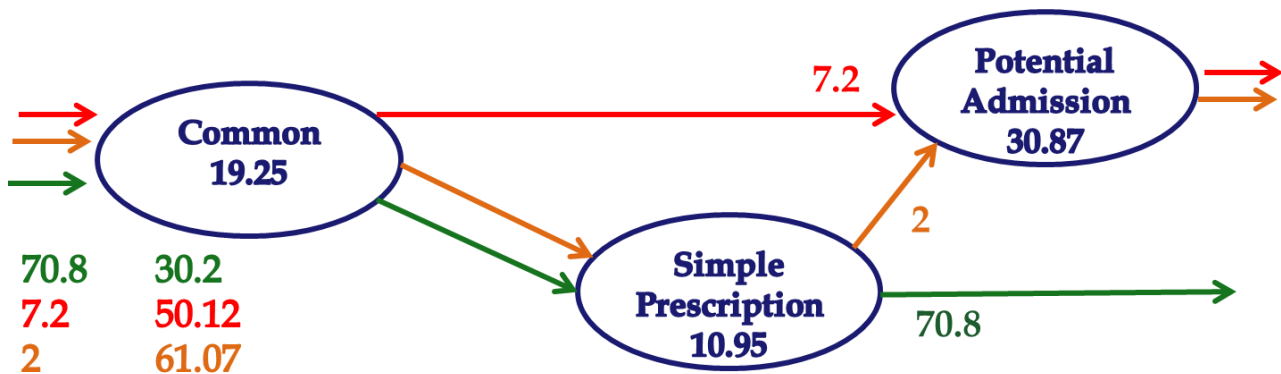
	R/hr	R/min	I	T	
RgBuff	80	1.33333	9	6.75	
RgProc	80	1.33333	8	6	
TrBuff	80	1.33333	2	1.5	
TrProc	80	1.33333	6.66667	5	19.25
SpBuff	72.8	1.21333	6	4.94505	
SpProc	72.8	1.21333	7.28	6	10.9451
HsBuff	9.2	0.15333	0.9	5.86957	
HsProc	9.2	0.15333	3.83333	25	30.8696

$$\text{Common: } 6.75+6+1.5+5 = 19.25$$

$$T_{SP} = 19.25+10.95 = 30.2$$

$$T_{PA1} = 19.25 + 30.87 = 50.12 \dots\dots (7.2 \text{ PA patients out of } 9.2 \text{ PA patients})$$

$$T_{PA2} = 19.25 + 10.95 + 30.87 = 61.07 \text{ (2 PA patients out of } 9.2 \text{ PA patients)}$$



$$T_{PA} = 50.12(7.2/9.2) + 61.07(2/9.2) =$$

$$T_{PA} = 50.12(0.782609) + 61.07(0.217391) = 52.5$$

We already have the average flow time

$$T = I/R = 43.68 / (80/60) = 32.76$$

$$T = 30.20(70.8/80) + 50.12 (7.2/80) + 61.07 (2/80)$$

$$T = 32.76$$



If you are looking for more challenging problems, we encourage you to go through the problems posted at [Assignment Process Basics Set2 Problems](#)

Problem 7. SAMOAK Industries. SAMOAK family has been in industrial developments for close to a century. They got the idea of smooth flow from Henry Ford and took it to a new dimension of time-based competition. The following data represents the inputs and outputs at one of their plants over a period of 6 weeks. Each column of the following tables represents 7 days of a week, a total of 42 days. Assume that the plant is working 24 hours 7 days a week. Each column of these matrices is corresponding to a week, where the first row is the first day of the week. Analyze these data, and estimate the average inventory, average flow time, and capacity of this process.

Input					
2	0	2	3	1	5
2	3	1	3	2	2
1	4	6	3	2	0
0	0	1	3	7	5
2	1	2	5	0	5
1	3	1	4	3	0
0	3	2	2	1	6

Output					
1	0	3	2	2	2
2	3	2	2	1	4
1	2	3	2	4	1
1	2	4	4	2	1
1	1	1	4	4	4
2	3	1	4	4	4
0	1	3	3	1	5



Note. Sometimes have data in a long column in an excel sheet and want to transform it into a matrix (table) to put into a report. Sometimes we have a table in a report and like to transform it into a column of data in an excel sheet. To learn how to do these using functions such as ROWS, COLUMNS, INT, MOD, INDEX, and MATCH, you may watch the lecture at https://www.youtube.com/watch?v=G_M3i2XVKmo&t=27s The excel file is at [Descriptive Statistics-exl](#) on tab 6. TurnArrayToMatrix. The excel file of this problem is at [Prepare for the Game](#) Tab 2. 2. AveInvFlow. On this tab, I have also shown how to transform the long column of data into a table.

The input and output data are in columns B and C of the Excel file, respectively. Since we need to compute inventory, the inventory of the first day is the difference between the input and output on the first day, i.e., $2-1=1$. The inventory of day 42 is the summation of the input in the first 42 days minus the summation of the output in the first 42 days, i.e., $99-97=2$. We compute the summation of input in each day, in column D using the function $=\text{SUM}(\$B\$2:B2)$ on the first day. If we copy this formula down to day 42, it will appear as $=\text{SUM}(\$B\$2:B43)$. However, for output data, we cannot copy this formula to the next column, because it will appear as $\text{SUM}(\$B\$2:C2)$ in the first day, and $\text{SUM}(\$B\$2:C43)$ in the last day. We can fix this problem by replacing $=\text{SUM}(\$B\$2:B2)$ by $=\text{SUM}(B\$2:B2)$. That is to replace $\$B\2 with $B\$2$. Now if we copy $=\text{SUM}(B\$2:B2)$ from column D to column E, we will have $=\text{SUM}(C\$2:C2)$, and in the last row of column D we will have $=\text{SUM}(C\$2:C43)$. Now we can compute inventory in each day, in column F, by subtracting the Total output from total input in each day. The formulas are shown below

	A	B	C	D	E	F
1	Day	Input	Output	SUM-In	SUM-Out	I
2	1	2	1	2	1	1
3	2	2	2	4	3	1
4	3	1	1	5	4	1
41	40	5	4	93	88	5
42	41	0	4	93	92	1
43	42	6	5	99	97	2
44				=SUM(B\$2:B43)	=SUM(C\$2:C43)	=D43-E43

Descriptive Statistics for output and inventory are shown below

Output				Inventory			
Mean	2.31	=AVERAGE(C2:C43)		1.31	=AVERAGE(F2:F43)		
Min	0	=MIN(C2:C43)		0	=MIN(F2:F43)		
Max	5	=MAX(C2:C43)		5	=MAX(F2:F43)		
StdDev	1.32	=STDEV.S(C2:C43)		1.39	=STDEV.S(F2:F43)		
CV	0.57	=B49/B46		1.06	=F49/F46		
Count	42	=COUNT(C2:C43)		42	=COUNT(F2:F43)		
95% CM	0.40	=CONFIDENCE.NORM(0.05,B49,B51)		0.42	=CONFIDENCE.NORM(0.05,F49,F51)		

Because the average output is $R=2.31$, and average inventory is $I=1.31$, the estimated flow time is $2.31T=1.31 \rightarrow 0.57$ day or 13.6 hours. If we exclude the days when output was 0, $SUM(RANGE)/COUNTIF(RANGE, ">0")$, we will get another approximation of throughput, which is 2.43. In that case, the average flow time reduces to 0.54, or 13 hours. The actual flow time measured directly was 0.52 days or 12.53 hours.

However, over the 6 weeks, they are almost equal.

LFT-Game1 Questions.

Here is an example of what I expect in your Game 1 first 50 days report.

1. **Demand.** You can provide two or perhaps four numbers regarding the demand in the past 50 days and your expectations for near future.
2. **Capacity.** You can provide 3 numbers representing the capacity of the machines.
3. **Inventory.** You can provide one number as the average inventory in the first 50 days.
4. **Flow time.** You can provide one or perhaps two numbers as your estimate of the average flow time in the first 50 days.

Chapter 2

Capacity Planning

The First Encounter with the Bottleneck

The Achilles' Heel in the Business Processes

*Everything comes to us that belongs to us if we create the capacity to receive it.
Rabindranath Tagore, 1862-1941.*

The key problems of this chapter can be accessed at

<http://www.csun.edu/~aa2035/CourseBase/Capacity/CapacityKeyProblems.pptx>

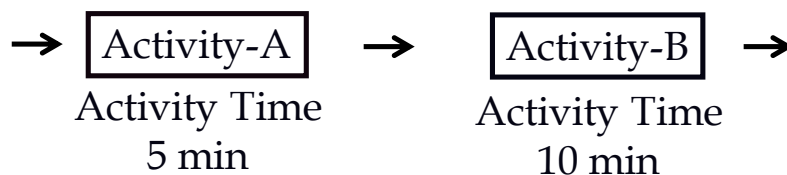
The recorded lectures – if any- and the excel worksheets- if any- are embedded inside the PowerPoint slides. By clicking on them the mp3/mp4 files and excel pages will open



Whenever you feel that you do not follow the text, you may watch my lecture on capacity at [Capacity Lecture Recorded](#). The PowerPoint slides can be accessed at [Capacity PowerPoints](#).

In this chapter, we will discuss both the capacity of operation and the capacity of a process. We also define terms such as Bottleneck, Utilization, and Throughput.

Problem 1. A process is composed of two sequential activities A and B.



a) Compute the Theoretical Flow Time.

The **theoretical flow time** of this process is $= 5+10= 15$. Theoretically, it takes a flow unit (a patient, a product, etc.) 15 minutes to flow from input form into output form.

b) How do we compute the Flow Time of the process?

We do not know the **Flow Time** of this process. The flow time is the **activity** times plus the buffer times which is represented as **Rectangle** (representing value-added activities) times and **triangle** times (representing non-value-added buffers or waiting lines). **Buffer or Waiting** times are usually 10, 20, and 100 times of the activity times. We will discuss flow time in future chapters.

Activity Time (or **Work Content**) of Activity-A = 5 min. We also refer to it as **T_p of Activity-A**.

Activity Time (or **Work Content**) of Activity-B = 10 min. We also refer to it as **T_p of Activity-B**.

If each activity is completed using a single resource, for example, if Resource-1 is assigned to Activity-1, and Resource-2 is assigned to Activity-2, then the **Unit Load of Resource-1 is 5 minutes**, and the Unit Load of Resource-2 is 10 minutes. We also refer to 5 minutes as **T_p or Unit Load of Resource-1** and 10 minutes as **T_p or Unit Load of Resource-2**.

T_p: When applied to activity, is called Work Content, Activity Time, or Activity Load.

T_p: When applied to a resource, is called Unit Load or Resource Load.

c) Compute Effective Capacity of each resource.

The Effective Capacity (or simply Capacity) of a Single Resource-Unit is equal to 1 divided by its unit load, or $1/T_p$. The Unit Load of Resource-1 is 5 minutes. If Resource-1 can complete 1 Activity-A in 5 minutes, how many units it can complete in 1 minute?

Unit	Time
------	------

1	5
---	---

X	1
---	---

$$X = (1 \times 1)/5 = 1/5$$

In 1 minute, it can complete $1/5$ Activity-A. That is $1/T_p$.

The capacity of the single resource-unit of **Resource-1** $= 1/T_p = 1/5$ per minute or $60(1/5) = 12$ per hour. Capacity of the single resource-unit of Resource-2 $= 1/T_p = 1/10$ per minute or $60(1/10) = 6$ per hour. If we had two units of Resource-1, we would have referred to them as the **Resource-Pool** of Resource-1.

d) What is the Capacity of the process?

The Capacity of the process = $R_p = \min \{1/5, 1/10\} = 1/10$ per minute or $R_p = \min \{12, 6\} = 6$ per hour.

A chain is as strong as its weakest link. The bottleneck is the resource-unit (resource-pool) with the minimum effective capacity.

The capacity of a process is equal to the capacity of its bottleneck.

Capacity is the maximum stable flow rate. During periods of heavy congestion, throughput reaches the capacity, but in most of the time $\text{Throughput} < \text{Capacity}$.

e) What is the Cycle time of this process?

Cycle time (CT) is the duration of time that the process needs between completing two consecutive flow units. It is not the flow time. It depends on the number of sequential operations and, indeed, on the capacity of the process (bottleneck). The higher the capacity, the shorter the cycle time. If the capacity is 6 per hour, then the process has the capability to send a product out (or complete a service) every $60/6 = 10$ minutes; cycle time is 10 minutes.

We can achieve the same outcome using 10-minute cycles by using $\text{Cycle time} = 1/\text{Capacity}$. Capacity was 6 per hour or $1/10$ per minute. CT (Cycle time) is then equal to $1/6$ hour or $1/(1/10) = 10$ minutes.

In a synchronized operation we will have $\text{Throughput} = \text{Demand}$. The Demand (that our process gets), which is equal to our Throughput, is always less than Capacity.

Suppose that the throughput in our problem is $R = 5$ per hour (or $1/12$ per minute).

f) Compute Capacity Utilization.

Capacity Utilization (or simply Utilization) of a Resource-Pool = $\text{Throughput} / \text{Capacity of a resource-pool} \rightarrow U = R/R_p$

Process Utilization (or simply Utilization) of the process

$U = \text{Throughput} / \text{Capacity of the bottleneck resource-pool} \rightarrow U = 5/6$

g) Compute Takt Time.

Cycle time is the amount of time that the process **needs** between completing two consecutive flow units. It depends on capacity; $CT = 1/R_p$.

Takt time (TT) is the amount of time that the process **has** between completing two consecutive flow units. Takt Time is related to Throughput (Demand). If the throughput is 5 per hour, then the process has a $60/5 = 12$ minutes. Takt time is 12 minutes; this means the system has 12 minutes to send the next flow unit out. In general, Takt time = $1/\text{Throughput}$. Throughput is 5 per hour or $1/12$ per minute.

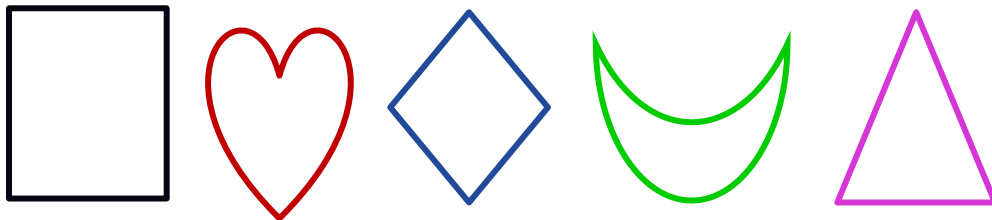
$TT = 1/(\text{Throughput}) = 1/5$ hour or $60(1/5) = 12$ minutes, or $1/(1/12) = 12$ minutes.

Demand = Throughput \leq Capacity

Takt Time \geq Cycle Time

$R \leq R_p$, where $TT \geq CT$

Cross-Training (Pooling). If we cross-train Resource-1 and Resource-2 so that each resource can perform both Activities A and B, then things may change. Usually, if we assign only one task or a limited number of tasks to a resource, the resource becomes specialized and efficient. But if several tasks are assigned to a resource, efficiency goes down. To illustrate this reality, consider drawing 20 of each of the following shapes. First, try drawing 20 of the first shape under each other (in the same color), then 20 of the second shape in the next column adjacent to your already drawn shapes, and then continue for the third, fourth, and fifth shapes. Now you have 20 rows and 5 columns of drawings of 5 shapes each in its assigned color. Now let's prepare the same thing differently.



This time, first draw the 5 shapes (each in its corresponding color), then go to the next row and repeat what you did in the first row, and then repeat for the next 18 rows.

Which procedure took more time? Drawing all 20 instances of each shape together or switching between the shapes? That is the difference between specialty and flexibility. Cross-training and pooling lead to flexibility. Flexibility means longer processing times (larger T_p). However, it does not necessarily mean lower capacity (smaller R_p).

Returning to our process with two activities, suppose that when both activities are assigned to both resources, the efficiency of both workers goes down. Perhaps that is due to setup for switching from one activity to the other. Suppose that when task 1 is assigned to each of the two resources, in the new situation, it takes 6 minutes instead of 5 minutes. Also, suppose that when task 2 is assigned to each of the two resources, it takes 12 minutes instead of 10 minutes.

h) Compute capacity under cross-training.

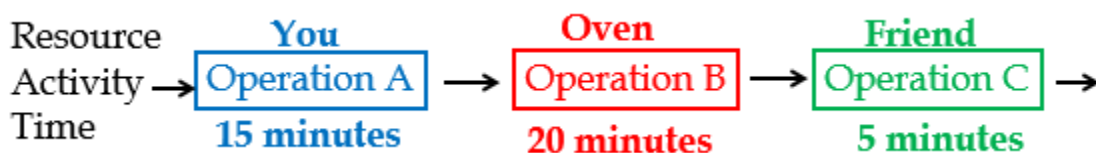
Now we have a Resource-Pool that contains two Resource-Units. The Unit Load of the Resource-Pool = $6+12 = 18$ minutes. We also refer to it as **Tp of the Resource-Pool**. Again, note that there is a different naming for Tp. When Tp is referring to an activity, we call it the **activity time** (or work content). When Tp is referring to a resource, we call it the **Unit Load of the Resource** (or load of a flow unit on a resource).

Given the unit load of 18 minutes, the capacity of a single Resource-Unit = $1/18$ flow units. But we also need a unit of time. Since Tp is 18 minutes, the capacity of a single resource-unit is $1/Tp = 1/18$ per minute.

The capacity of a single Resource-Unit = $60(1/18) = 3.333$ flow units per hour.

However, there are 2 resource-units in the resource-pool. Therefore, the capacity of the resource-pool is $2(1/18) = 1/9$ per minute, or $2(3.333) = 6.666$ per hour. While the total unit load increased from $5+10$ to 18, the capacity did not decrease, but instead, it increased by about 11%. That is the power of **cross-training, pooling, and centralization**.

Problem 2. We are making batches of muffins. There are three sequential activities: A (preparation), B (baking), and C (packaging and labeling). There are three resources: you and your friend (human resources), and the oven (capital resource).



To produce each batch of muffins, you prepare the material and put the batch in the oven (there is only a single oven, and it can just bake one batch at a time). Then your friend takes the batch out and completes packaging and labeling. The processing time for each operation is given above. Estimating the processing times is the subject of motion and time studies. This system works for 4 hours, that is, $4 \times 60 = 240$ minutes.

Compute the capacity of each resource, the capacity of the process, and the utilization of each resource if throughput is at 100% capacity of the process, the cycle time, and the takt time.

Capital Resources — Fixed Assets such as land, buildings, facilities, machinery, oven, etc.

Human Resources — People such as operators, assemblers, engineers, waiters, chefs, customer-service representatives, you, your friend, etc. Each activity may require one or more resources, and each resource may be allocated to one or more activities. A resource, a baker, may be used by several activities such as mixing, kneading and forming a dough. An activity like loading an oven may require multiple resources such as a baker and an oven.

Resource-Unit — An individual resource (chef, mixer, oven), or a combination of different individual resources, for example, an operating room.

Capacity (per hour): You: $60/15 = 4$, Oven: $60/20 = 3$, Friend: $60/5 = 12$

Process Capacity = Min {4, 3, 12} = Capacity of the bottleneck = 3 per hour.

Each hour we produce 3 units. Starting from the second unit, every 60 minutes, 3 units may enter, pass, and leave the process. We refer to this $60/3 = 20$ as the average inter-arrival time, average inter-departure time, or cycle time.

We computed the capacity/hour. We could also have computed the capacity per minute:

Capacity/minute: You: $1/15$ per minute, Oven: $1/20$ per minute, Friend: $1/5$ per minute.

Process Capacity = Min $\{1/15, 1/20, 1/5\} = 1/20$.

The capacity of the bottleneck = $1/20$ per minute. Each minute we produce $1/20$ units. In 20 minutes, we can send out, or take in, one product. Twenty minutes is the inter-arrival time, inter-departure time, and cycle time.

a) How long does it take to produce a batch of muffins?

In formal terms, the question is asking what the flow time in this process is. We cannot yet determine this because we also need the waiting times.

b) How long does it theoretically take to produce a batch of muffins?

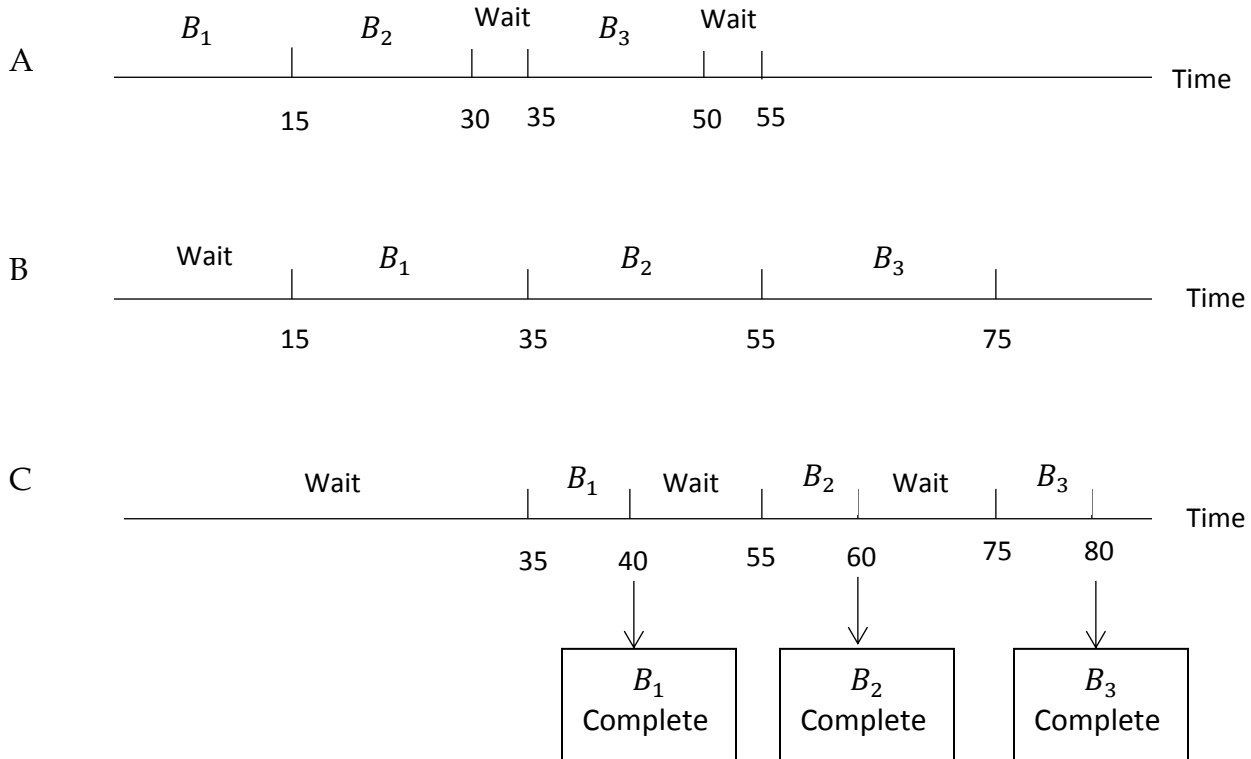
In formal terms, the question is asking what the theoretical flow time in this process is.

Theoretical Flow Time = $15 + 20 + 5 = 40$

c) How often does a batch of muffin could enter (exit) this process?

In formal terms, what is the **cycle time** of this system?

The following shows the timeline for each activity – A, B, and C – in a three-batch baking process, where B_1 is batch 1, B_2 is batch 2, and B_3 is batch 3. Please note that we do not allow inventory in the process, and therefore Activity A may need to wait for the availability of the oven (Activity B) before preparing for the next batch.



You prepare a batch and place it in the oven at minute 15. You then start the next batch and complete it at minute 30. The oven is still baking the first batch. It will be done at minute $15+20 = 35$.

At minute 35, your friend can take the first batch out of the oven, and after 5 more minutes, at minute 40, he is done. The **first batch exits at minute 40**. The oven is the **bottleneck** since you and your friend need 15 and 5 minutes, respectively, but the Oven needs 20 minutes. At minute 35 you can put the second batch in the oven. Your friend takes it out of the oven at minute $35+20 = 55$ and then sends it out of the process at minute 60. That is $60 - 40 = 20$ minutes after the first batch. Similarly, the third batch exits at minute 80.

Therefore, the capability of the process regarding the time between exits of two consecutive batches, the cycle time, is?

Cycle time = Max {15, 20, 5} = 20. The oven is the bottleneck.

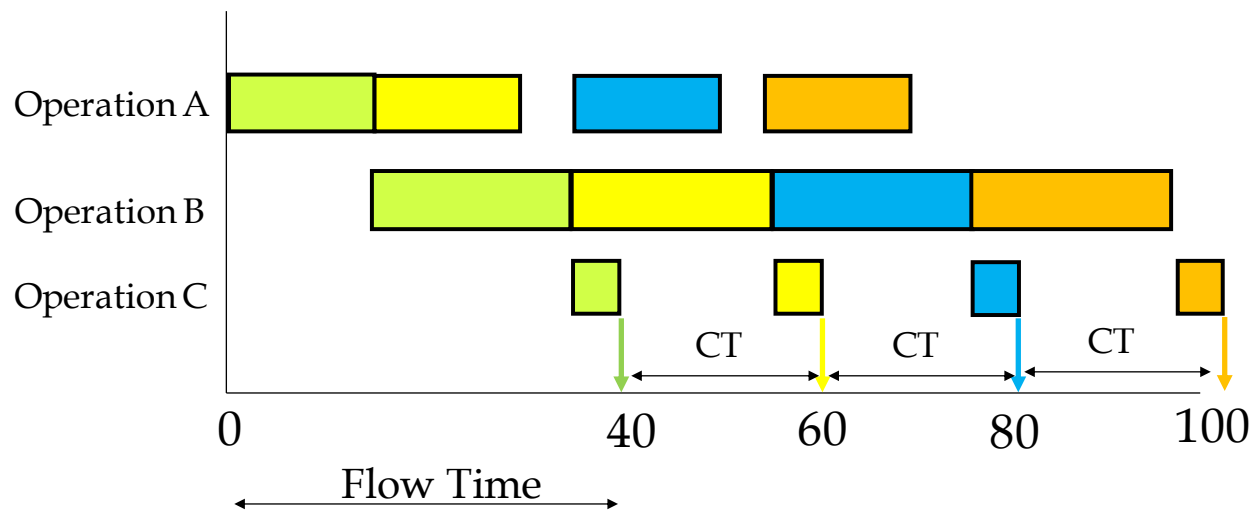
Or we can do it in a long way:

- Capacity = $\text{Min} \{1/15, 1/20, 1/5\} = 1/20$
- Cycle time = $1/R_p = 1/(1/20) = 20 \text{ min.}$

Cycle Time: Starting from 0 vs. Continual

d) How many batches can you produce per day?

Case 1, Starting at 0. We have 4×60 minutes. It takes you 40 minutes to produce the first batch. In the remaining $240 - 40 = 200$ minutes, given the cycle time is 20 minutes, we produce 1 batch per 20 minutes, and that gives us 10 batches in 200 minutes. We produce $1 + 10 = 11$ per 4 hours. We could also have said that in the first 40 minutes, we produced 1 batch and in the next 200 minutes we produced $1/20$ batch per min. That is $1 + 200(1/20) = 11$ per 4 hours.



Case 2, Continual. Suppose we are not producing muffins, but something else, such that at the start of each day there is work-in-process (WIP) from the previous day in the system. For example, suppose it is a small portion of a painting and you can make that part ready today, and put it into the oven at the start of the next day. What is the capacity (or maximum Throughput)?

The flow time is 40 minutes. The cycle time is 20 minutes. Therefore, capacity is $1/20$ per minute. In 4 hours, it is $240(1/20) = 12$.

By now we should know the terms Flow Time, Cycle Time, and Capacity.

e) What is the Utilization of the oven if the process works at full capacity?

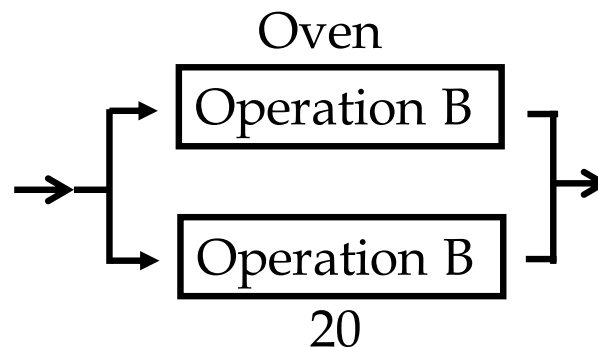
The oven is always working. Every 20 minutes, 1 batch comes, and 1 batch leaves. Utilization of the oven is 1 or 100%.

In every 20 minutes, you only work 15 minutes. Your utilization is $15/20 = 0.75$ or 75%. In every 20 minutes, your friend only works 5 minutes. Your friend's utilization is $5/20 = 0.25$ or 25%.

Utilization: You: $3/4 = 0.75$, Oven: $3/3 = 1$, Friend: $3/12 = 0.25$.

Increasing the capacity? Relaxing the bottleneck by increasing the level of resources.

Increasing the level of resources? Adding a resource-unit to a resource-pool. Our oven resource-pool contains only one single resource-unit. By increasing its level, we mean adding a second oven.



Resource-Pool – A collection of **interchangeable** resource-units that can perform an identical set of activities. Processing time = $T_p = 20$ minutes.

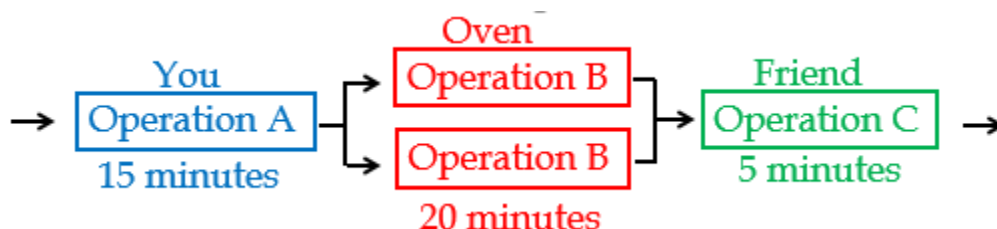
Resource-Pool contains 2 Resource-Units; $c = 2$, and unit load is 20 mins, $T_p = 20$ mins.

The capacity of a resource-unit = $1/20$ per minute or $60(1/20) = 3$ per hour.

Capacity of a resource-pool = $2(1/20) = 1/10$ per minute, or $60(1/10) = 6$ per hour.

f) What is the cycle time (after how many minutes a product can exit this system)?

If each minute produces 0.1 unit, then how many minutes are required to produce one unit? $1/0.1 = 10$. Cycle time is 10 minutes.



g) Compute capacity of each of the three resource-pools.

Capacity of Resource-Unit (batch/min) Y: 1/15, O:1/20, F: 1/5

Capacity of Resource-Pool (batch/min) Y: 1/15, O:2/20, F: 1/5

Process Capacity = Capacity of the bottleneck = 1/15 per minute. Now, you are the bottleneck.

Each minute the system produces 1/15 units. In 15 minutes, we can send out or take in one product.

Inter-arrival time = inter-departure time = cycle time = 15 min.

h) Compute the utilizations, if the process can work at 100% capacity.

Utilizations are:

$$U-Y = (1/15)/(1/15) = 1,$$

$$U-O = (1/15)/(1/10) = 0.67,$$

$$U-F = (1/15)/(1/5) = 0.33.$$

Capacity of each resource-pool in one hour is **Y: $60(1/15) = 4$, O: $60(2/20) = 6$, F: $60(1/5) = 12$.**

Process Capacity = 4 per hour. Every hour the system produces 4 units. In 15 minutes, we can send out or take in one product. Interarrival time, inter-departure time, cycle time = 15 minute.

Utilizations are: $U-Y = 4/4 = 1$, $U-O = 4/6 = 0.67$, $U-F = 4/12 = 0.33$.

Two Ovens plus Cross-Functional Workers. Suppose you and your friend are flexible resources who can do both activities 1 and 3. Suppose the unit loads (T_p) remain the same.

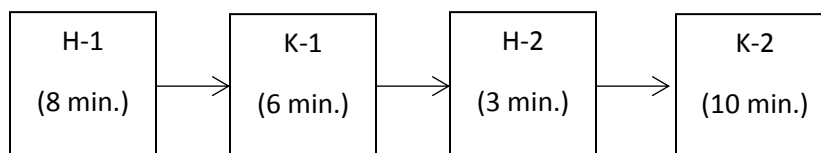
Resource	Human Resources	Capital Resources
Time	5+15	20
Capacity	$60(2/20) = 6$	$60(2/20) = 6$

Cross-functional workers and resource-pooling are great operational strategies. However, in this specific example, we need to be careful. If the process works at 100% capacity, the utilization of all resources is 100%. This strategy is very risky as a small variation can reduce the capacity significantly. We will later show that no system can produce at 100% capacity and the more bottlenecks, the lower the actual throughput.

Resource Pooling—Combining separate resource-pools into a single more flexible pool that is able to perform several activities. To transform specialized resources into general-purpose resources, we can cross-train workers and use general-purpose machines.

Resource Pooling is a powerful operational tool that can significantly affect **not only capacity and throughput** but also **flow time**.

Problem 3. Four consecutive activities: Two human resources (H-1 and H-2) and two capital resources (K-1 and K-2). Flow units pass the four resources in the following sequence: H-1, K-1, H-2, and K-2. Unit loads $Tp_1=8$, $Tp_2=3$, $Tp_3=6$, $Tp_4=10$.



Compute the capacity of each resource per minute, the capacity of the process, and utilization at 100% capacity. Then increase the capacity by increasing the level of the bottleneck and by cross-training. In cross-training, assume that Tp remains the same for all resources.

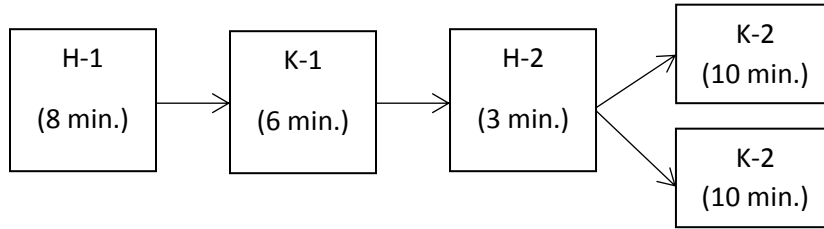
Specialization and one resource-unit at each activity:

	A	B	C	D	E	F	G
1	Cap. Per Minute						
2	Resource	H-1	K-1	K-2	H-2		
3	Tp	8	6	10	3		
4	Rp/min	0.125	0.16667	0.1	0.33333333	=1/B3	
5	Proc. Rp	0.1	0.1	0.1	0.1	=MIN(\$B4:\$E4)	
6	U@100%Cap	0.8	0.6	1	0.3	=\$B5/B4	

The same computations in a different way:

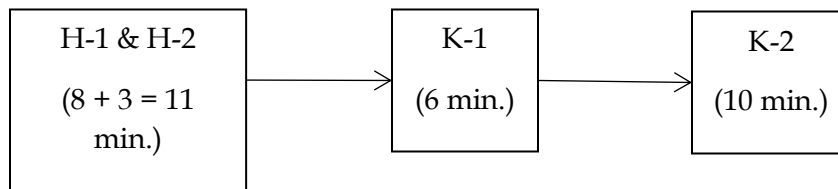
Cap. Per hour				
Resource	H-1	K-1	K-2	H-2
Tp	8	6	10	3
Rp/hr	7.5	10	6	20
Proc. Rp	6	6	6	6
U@100%Cap	0.8	0.6	1	0.3

Specialization and two resource-units in K-2 single pool:



Cap. Per hour. Increase the level of the bottleneck. Two K-2.				
Resource	H-1	K-1	2(K-2)	H-2
Tp	8	6	10	3
Rp/hr.	7.5	10	12	20
Proc. Rp	7.5	7.5	7.5	7.5
U@100%Cap	1	0.75	0.625	0.375

Cross-Training at H (resources H1 and H2 are now in resource-pool H), and a single resource-unit at each K:



Cap. Per hour. One K-2, but Cross Functions H-1&H-2=> H				
Resource	H	K-1	(K-2)	
Tp	11	6	10	
Rp/hr	10.9091	10	6	
Proc. Rp	6	6	6	
U@100%Cap	0.55	0.6	1	

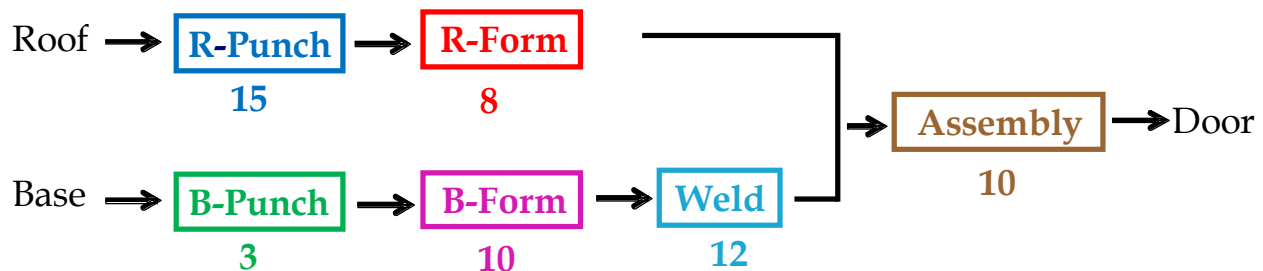
Do you know the general relationship between theoretical flow time and cycle time?

The number of sequential operations is $\geq (\text{Flow Time})/(\text{Cycle Time})$

The minimal number of sequential operations is $= (\text{Flow Time})/(\text{Cycle Time})$ rounded up.

Problem 4. This problem is based on the prototype example in our reference book MBPF. MamossaAssaf Inc. fabricates garage doors. Roofs are punched in a roof punching press (15 minutes per roof) and then formed in a roof forming press (8 minutes

per roof). Bases are punched in a base punching press (3 minutes per base) and then formed in a base forming press (10 minutes per base), and the final base is welded in a base-welding machine (12 minutes per base). The base subassembly and the roof then go to the final assembly, where they are welded together (10 minutes per garage) on an assembly welding machine to complete the garage. Assume there is one operator at each station.



This problem is among our key problems. It can be accessed at

<http://www.csun.edu/~aa2035/CourseBase/Capacity/CapacityKeyProblems.pptx>



The recorded lectures – if any- and the excel worksheets- if any- are embedded inside the PowerPoint slides. By clicking on them the mp3/mp4 files and excel pages will open

a) What is the Theoretical Flow Time? (The minimum time required to produce a garage from start to finish.)

Flow Time → Theoretical Flow Time →

Roof Path: $15 + 8 = 23$

Max = $25 + 10 = 35$

Base Path: $3 + 10 + 12 = 25$

Critical Path = $\text{Max}(23, 25) = 25$

Theoretical Flow Time = 35

b) What is the capacity of the system in terms of garages per hour?

R-Punch: $1/15$ per minute or 4 per hour

R-Form: $1/8$ per minute or 7.5 per hour

B-Punch: 1/3 per minute or 20 per hour

B-Form: 1/10 per minute or 6 per hour

Welding: 1/12 per minute or 5 per hour

Assembly: 1/10 per minute or 6 per hour

Therefore, R-Punch is the bottleneck, and the Process Capacity is 4 flow units per hour.

c) If you want to increase the process capacity, which is the resource-pool that you would put additional resource-units toward?

Obviously, the R-Punch is the best solution, because increasing capacity of any other resource is a mirage.

d) Compute utilization of all the resources at the full process capacity. Although in the real life, it is impossible for a process to perform at its 100% capacity, assume that the throughput is equal to the process capacity. Throughput = 4. Compute utilization of all resources.

R-Punch Utilization = $4/4 = 100\%$.

R-Form Utilization = $4/7.5 = 53.33\%$.

B-Punch Utilization = $4/20 = 20\%$.

B-Form Utilization = $4/6 = 66.67\%$.

Welding Utilization = $4/5 = 80\%$.

Assembly Utilization = $4/6 = 66.67\%$

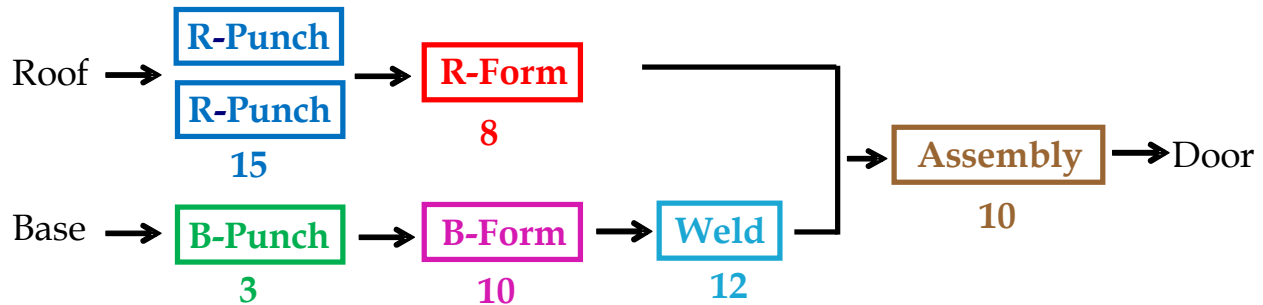
e) What is the utilization of the most utilized and least utilized resources?

The most utilized resource is obviously R-Punch. Its utilization is 100%.

However, utilization of the bottleneck in the real world is never 100% because no process can have a throughput equal to its capacity. In reality, utilization of all the resources will be less than what we have computed for this problem. **This process can never produce 4 flow units per hour continually.**

What is the utilization of the least utilized resource? Interestingly, our other Punch machine, B-Punch is the least utilized resource. $U=20\%$.

f) Relaxing the bottleneck by lifting its level (i.e., adding a new R-Punch resource). Suppose we double the capacity of the bottleneck by adding the same capital and human resources. What is the new capacity of the system?



R-Punch: A single resource-unit, $1/15$. Two resource-units, $1/15 + 1/15 = 2/15$ per minute or 8 units per hour. In general, $R_p = c/T_p$.

R-Form: $1/8$ per minute or 7.5 per hour

B-Punch: $1/3$ per minute or 20 per hour

B-Form: $1/10$ per minute or 6 per hour

Welding: $1/12$ per minute or 5 per hour

Assembly: $1/10$ per minute or 6 per hour

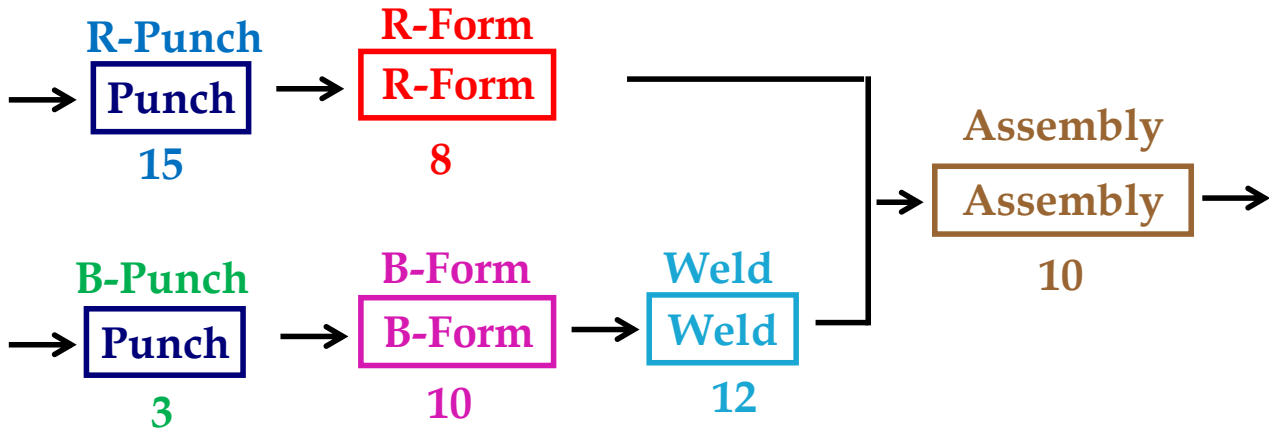
Process Capacity is 5 per hour.

g) We doubled the capacity of the bottleneck, but the capacity of the system increased by only 25%. This situation is an example of what managerial experiment?

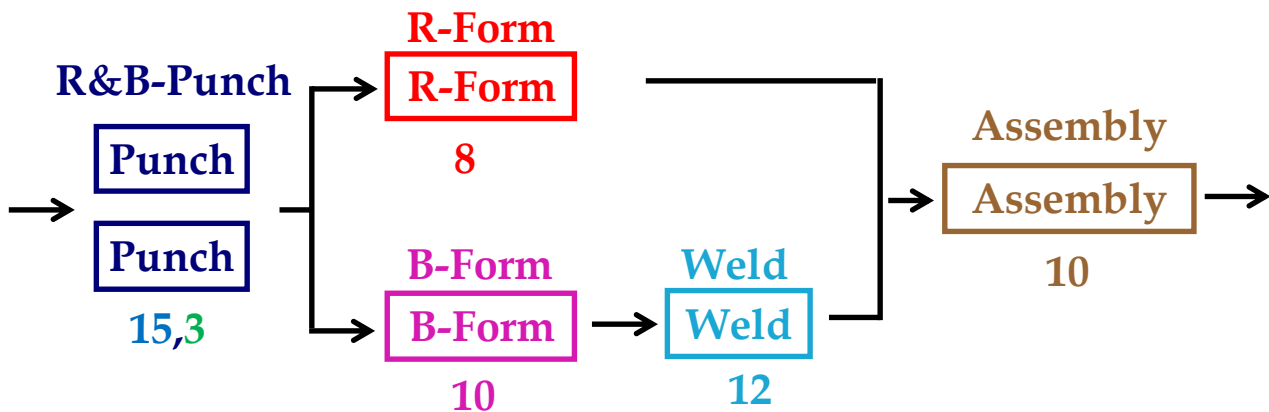
- 1) When we relax a bottleneck, the bottleneck shifts to another operation. The bottleneck shifts from R-Punch to Welding.
- 2) Diminishing Marginal Return. By increasing input, the output will not increase linearly.

h) Pooling and Cross-Training. Now, suppose we return to the original situation in which we have a single machine and a single operator at each operation. Suppose we pool R-Punch and B-Punch machines and we cross-train their operations. We form a new resource-pool named Punch in which both R-Punch and B-Punch operations are done in this resource-pool. What is the new capacity of the system?

In the new process, similar resources in the Punch resource-pool have been assigned to R-Punch and B-Punch activities. The process can be represented as follows:



We have pooled two special-purpose resources of R-Punch and B-Punch into a resource-pool of Punch, which can perform both R-Punch and B-Punch activities. In practice, when two special-purpose resources are pooled into a more general-purpose resource-pool, we observe an increase in the unit loads. However, for the purpose of simplicity, we assume that 15 minutes and 3 minutes remain the same and the total unit load of Punch resource-pool stays at $15 + 3 = 18$. We also use the following representation.



The capacity of the process is computed as follows:

Punch: $2/18$ per minute or 6.67 per hour

R-Form: $1/8$ per minute or 7.5 per hour

B-Form: $1/10$ per minute or 6 per hour

Welding: 1/12 per minute or 5 per hour

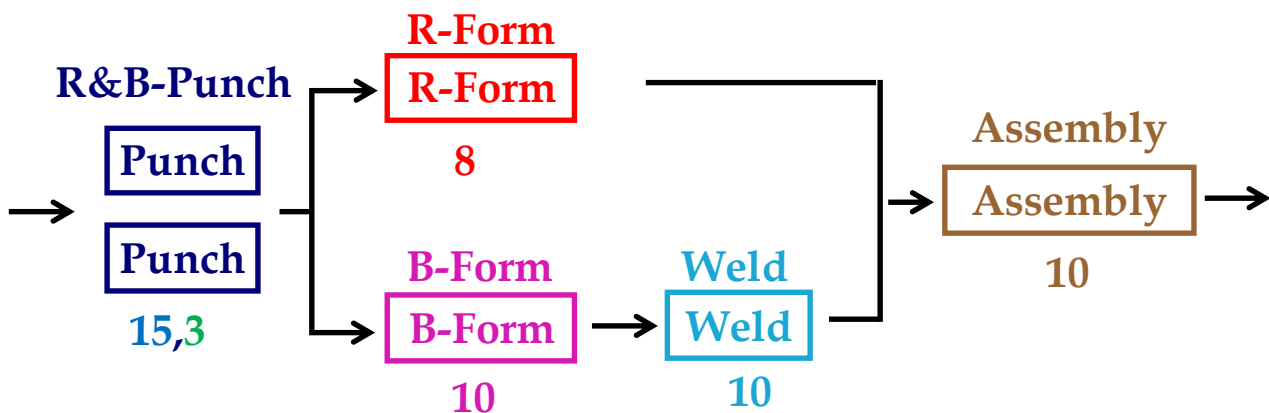
Assembly: 1/10 per minute or 6 per hour

Process Capacity is increased to 5 per hour, without adding any new resource.

i) This situation is an example of what managerial experiment?

- 1) Cross-training and pooling can increase the capacity.
- 2) Usually, the cost of cross-training and pooling is lower than the cost of adding the second resource-unit.

j) Productivity Improvement – Method, Training, Technology, and Management. Now suppose by (i) investing in improved jigs and fixtures (**technology**), (ii) implementing a better **method** of doing the job, (iii) **training** our human resources, and (iv) better managing human, capital, and information resources, we can reduce the welding time from 12 minutes to 10 minutes. **What is the new capacity of the system?**

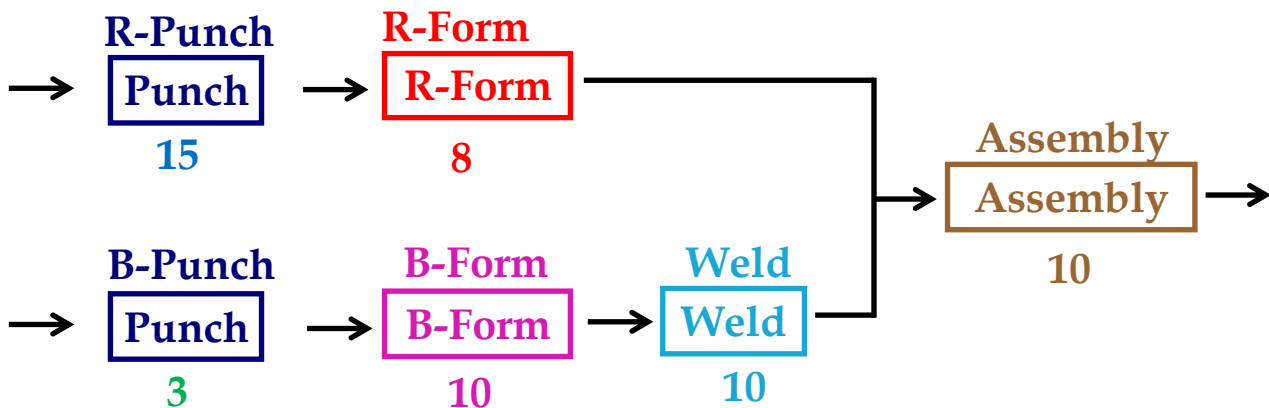


Process Capacity is now increased to 6 per hour.

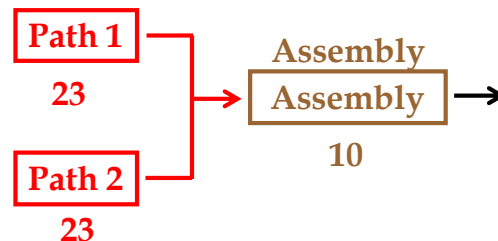
k) Why is it impossible to work at 100% of capacity?

There are 3 bottlenecks. This is a risky situation. Any of the bottlenecks could cause the throughput of the system to fall below 6 per hour. Suppose the input to a bottleneck resource is not ready and the resource stays idle waiting for input for 5 minutes. Assume the resource is a physician, and the patient has not shown up yet. Alternatively, suppose Punch fails to provide input to B-Form for 1 hour, or B-Form fails to feed Weld, or Weld fails to feed Assembly – that hour of capacity perishes. The more bottlenecks exist in the system, the higher the probability of not meeting the capacity. In addition, the system has three bottlenecks, but they are back-to-back, that is, one bottleneck feeds the next one.

1) Is flow time at risk?



Both paths to the last bottlenecks are critical. They can both increase the flow time.



Lessons learned throughout this problem:

1. When a bottleneck resource is relaxed, the bottleneck shifts to another resource.
2. By doubling the bottleneck resource, the capacity usually does not double. This could be interpreted as diminishing marginal return situation.
3. One other way to increase capacity is cross-training (for human resources) and pooling (for capital resources).
4. Usually, the cost of cross-training and pooling is lower than the cost of adding the second resource-unit.
5. One other way to increase capacity is to reduce the unit load. (i) Investing in improved jigs and fixtures (technology), (ii) implementing a better method of doing the job, (iii) training our human resources, and (iv) better managing human, capital, and information resources.
6. Processes cannot work at 100% capacity. Capacity is perishable – It is lost if the input is not ready. The more the bottleneck resources, the lower the utilization.

7. Convergence points are important in managing the flow time. The more convergence points, the higher the probability of the flow time exceeding the average flow time.

Problem 5. SAMOAK Industries. SAMOAK family has been in industrial developments for almost a century. They got the idea of smooth flow from Henry Ford and took it to a new dimension of time-based competition. The following data represents the inputs and outputs at one of their plants over a period of 6 weeks. Each column of the following tables represents 7 days of a week, a total of 42 days. Assume that the plant is working for 24 hours during 7 days a week. Each column of these matrices is corresponding to a week, where the first row is the first day of the week. Analyze these data and estimate the average inventory, average flow time, and capacity of this process.

Input						Output					
2	0	2	3	1	5	1	0	3	2	2	2
2	3	1	3	2	2	2	3	2	2	1	4
1	4	6	3	2	0	1	2	3	2	4	1
0	0	1	3	7	5	1	2	4	4	2	1
2	1	2	5	0	5	1	1	1	4	4	4
1	3	1	4	3	0	2	3	1	4	4	4
0	3	2	2	1	6	0	1	3	3	1	5

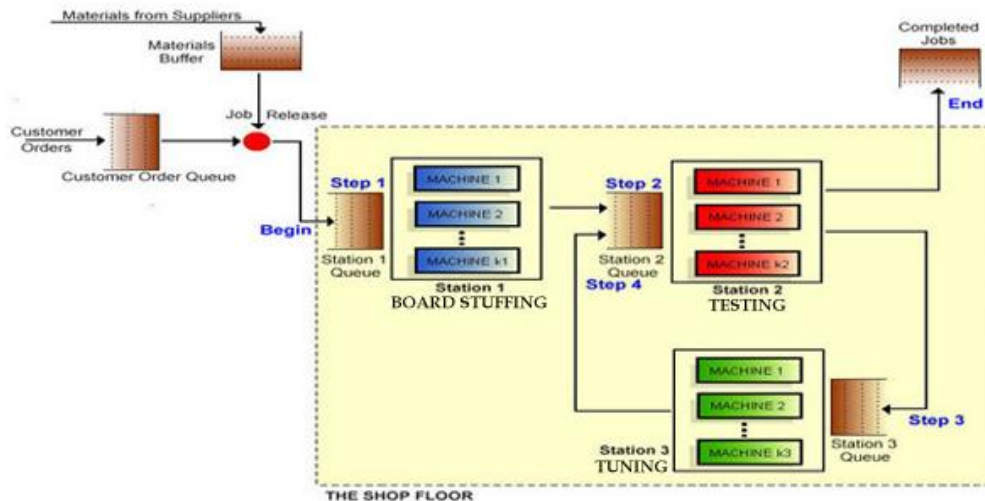
a) What is the capacity of this process?



The Excel file of this problem is at [Prepare For The Game](#).

The input and output data are in columns A and B of the Excel file, respectively.

If we look at the daily output, we see 5 units of outputs on day 42. We may conclude that the capacity is at least 5. However, in several other days, such as days 25–27, although the beginning inventory is greater than 0 and the ending inventory is also greater than 0, we have not produced 5 units. Therefore, we can just conclude that the capacity of the process is about 5 units. The process blueprint is shown below. The incoming flow units go to the first station, next to the second, and last to the third. From the third station, they go back to the second station, and after processing in this station, they will leave the system in the form of output.



The utilization data for the three stations are shown in [Prepare for the Game](#). You may prepare a descriptive statistic table similar to what we provided below.

Station 1		Station 2		Station 3		Output	
Mean	0.51269	Mean	0.18279	Mean	0.20619	Mean	2.31
Standard Error	0.0443	Standard Error	0.01592	Standard Error	0.02473	Min	0
Median	0.4525	Median	0.158	Median	0.152	Max	5
Standard Deviation	0.28713	Standard Deviation	0.10317	Standard Deviation	0.16028	StdDev	1.32
Range	1	Range	0.385	Range	0.699	CV	0.57
Minimum	0	Minimum	0	Minimum	0	Count	42
Maximum	1	Maximum	0.385	Maximum	0.699	90%CM	0.14
Count	42	Count	42	Count	42		

Compute the capacity of each station.

$$U = R/R_p$$

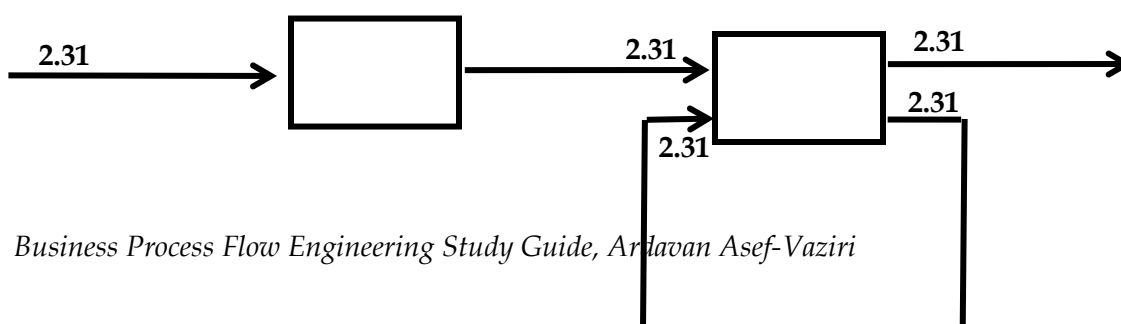
Average U values at the three stations are 0.513, 0.813, and 0.206. Average R is 2.31.

Capacity of station 1 is $0.513 = 2.31/R_{p1} \rightarrow R_{p1} \approx 4.5$.

Capacity of station 3 is $0.206 = 2.31/R_{p2} \rightarrow R_{p3} \approx 11.2$.

What about station 2?

Capacity of station 3 is NOT $0.183 = 2.31/R_{p23} \rightarrow R_{p3} \neq 11.2$. Why?





Capacity of station 3 is $0.183 = 2 \times 2.31 / R_{p23} \rightarrow R_{p3} \neq 22.4$.

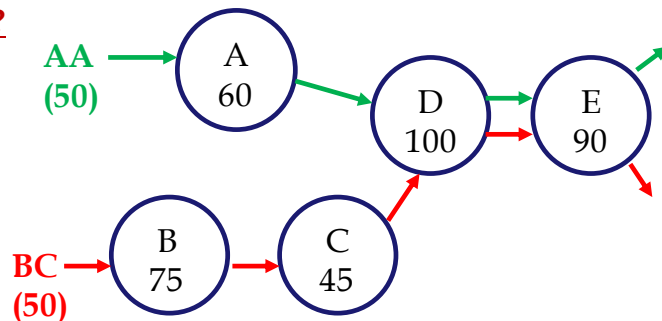
Station 1 is the bottleneck. You may compute the T_p of each station by using $T_p = 1/R_p$. You need to be careful about two things.

1. If there is more than one machine in a station, say if there are c machines in a station, flow time is c/R_p .
2. The flow time is in days because R_p is per day. You need to either divide R_p by 24 or multiply T_p by 24.

Problem 6. The following graph shows a production process for two products: AA and BC. Station D and E are flexible and can handle either product. Regardless of the type of the product, Station D can finish 100 units per day. Station E can finish 90 units per day. Station A works only for Product A and has a capacity of 60 units per day. Station B and C are only for Product BC and have a capacity of 75 and 45 units per day, respectively. The demand for each product is 50 units per day.

a) Which stations are the bottlenecks?

- A) Stations A and C
- B) Stations B and C
- C) Stations C and D
- D) Stations D and E
- E) Stations C and E



b) If the system can work at the processing capacity, which of the following is NOT true?

- A) The utilization of machine A is at least 75%
- B) The utilization of machine B is at least about 53%
- C) The utilization of machine B is at most 60%
- D) The utilization of machine D is 90%
- E) All of the above

E \rightarrow We can produce at most 90 AA and BC.

C → We can produce at most 45 BC.

We may produce all combinations from 50AA and 40 BC to 45AA and 45 BC.

A) We produce at least 45 AA: $45/60 = 75\%$

B) We produce at least 40 BC: $40/75 = 53.33\%$

C) $45/75 = 60\%$

D) $90/100 = 90\%$

For more problems, the reader may look at

[Assignment Capacity Problems](#)

[Assignment Capacity Problems – More](#)



The reader is referred to the following link for better understanding of flow time in a nondeterministic world:

<https://www.youtube.com/watch?v=wqjGsLsadOo&t=870s>

In this video, we will discuss the impact of interrelated activities, convergence points, and resource dependability on flow time. Flow time analysis also has extensive applications in project management. The reader is invited to watch the following video for better understanding of nondeterministic project scheduling:

<https://www.youtube.com/watch?v=7IEfN5OqtQ0&t=24s>

Chapter 4

Predictive Analytics I

Time Series Analysis and Regression

If you can look into the seeds of time, and say which grain will grow and which will not speak then unto me. William Shakespeare, 1564-1616.

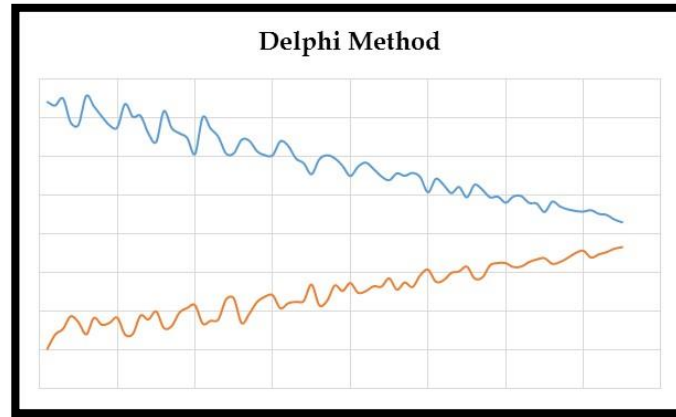
Introduction. Marketing, finance, and operations are the three key building blocks of manufacturing and service systems. Planning, organizing, budgeting, executing, and controlling are the main responsibilities of the three key managers. Operations Managers need forecasting for capacity planning, inventory management, scheduling, etc. Financial Managers need forecasting for investment analysis, revenue and cost analysis, cash flow planning, etc. Marketing Managers need forecasting for pricing, sales force planning, promotions, etc.

All of the marketing, finance, and operation functions need forecasting to develop integrated resources and processes in order to nourish multi-dimensional and structurally integrated capabilities to understand the revolving business eco-system, in order to create value and re-shape the business organization towards achieving the goals. While the three key managers may be interested in forecasting different variables, they have a common interest in sales forecast and investment plans. They are all interested in long-term and short-term forecasts for strategic, tactical, and operational decisions.

In general, there are two forecasting methods, **quantitative** and **qualitative**.

Quantitative forecast uses the historical data such as previous sales, revenues, production mix, production volumes, and financial reports. The two main types of quantitative forecasting are time series analysis and regression analysis.

Qualitative forecasting. Qualitative forecast relies on the subjective opinion and intuition of experts in the field. An example of the qualitative forecasting technique is the **Delphi method** which relies on a panel of experts who answer questionnaires in several rounds. After each round, a facilitator panel provides an anonymous summary of the experts' forecasts in the most recent round, as well as the reasons provided for each class of judgments. The experts are encouraged to revise their earlier responses in the light of the categorized responses provided by the others in the opposite opinion sides. It is believed that during this process the gaps between different opinions will decrease, and the whole group converges towards the most logical conclusions. The Delphi method is pictorially represented in the following figure.



We devote the rest of this chapter to quantitative forecasting. While our variable of interest throughout the example is the volume of sales, the ideas, concepts, and methods can be applied to any other variable.

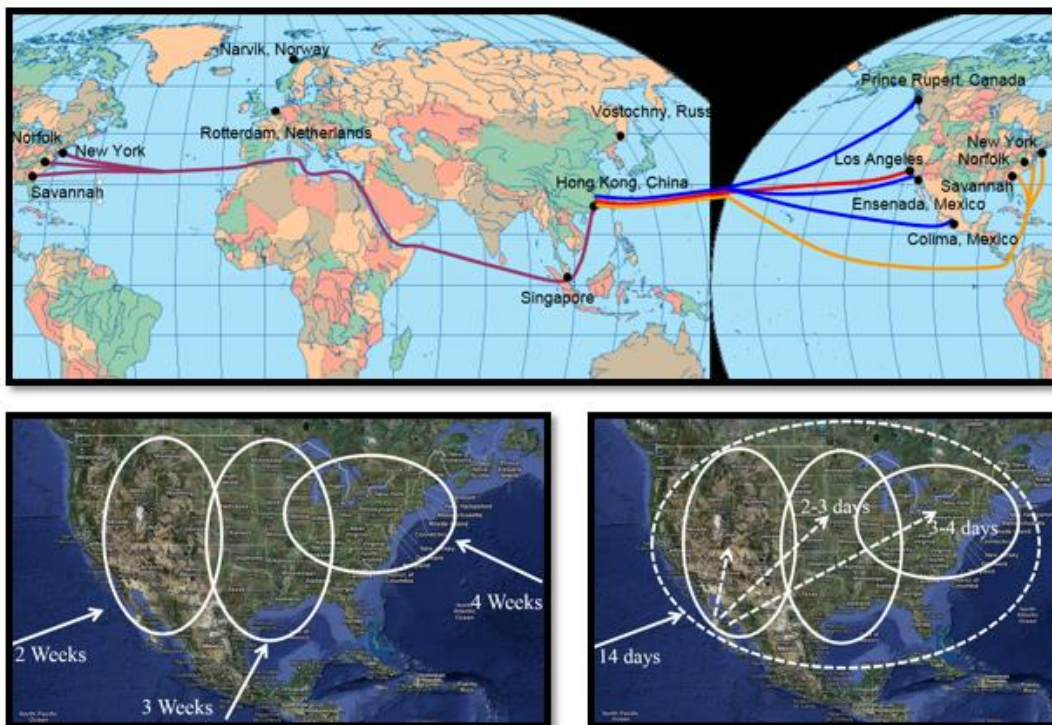
Characteristics of Forecasting Techniques. All forecasting techniques have three main characteristics in common.

1. **Forecasts are usually wrong.** Since the world is not deterministic – at least to us- all forecasts are almost always incorrect. Forecasts provide the average value for the variable of interest -sales or demand in our case. Demand is a random variable usually following Poisson or Normal distribution. Thus, besides the average demand, we should always accompany our forecast with a measure of variability- standard deviation, variance, or coefficient of variations. If our average forecast for next year is Y , and the standard deviation of Y is X , the coefficient of variations of our forecast is then defined X/Y . The lower the coefficient of variations, the better the forecast.
2. **Forecasts for aggregate items are more accurate than individual item forecasts.** Aggregate forecasts reduce the amount of variability. Forecasts for the total US economy is more accurate than the forecasts for manufacturing, agriculture, and service sectors. A forecast for manufacturing is more accurate than a forecast for the car industry, a forecast for the car industry is more accurate than a forecast for Ford Motors, a forecast for Ford Motors is more accurate than a forecast for Mustangs, and a forecast for a Mustang is more accurate than the forecast for a Mustang convertible. Aggregate forecasts reduce the amount of variability with respect to the average forecast. We can intuitively understand that the forecast for the summation of two products is more accurate than the forecast for each individual product, because the high demand for one product can be compensated by the low demand for the other, and the summation of the demand for both products will show less variability compared to the average forecast. From a mathematical point of view, the variance of the sum of two variables is equal to the sum of the variances of the two variables. Therefore, the standard deviation of the sum of two variables is less than the sum of the standard deviation of each of the two variables. If the standard deviations of two

models of Ford Motors cars are equal, and are equal to σ for each, then the variance for the summation of these two variables is $= \sigma^2 + \sigma^2 = 2\sigma^2 \rightarrow \text{SQRT}(2) \sigma$. That is, standard deviation of a variable defined as the summation of two other variables, is less than twice of the standard deviation of each variable.

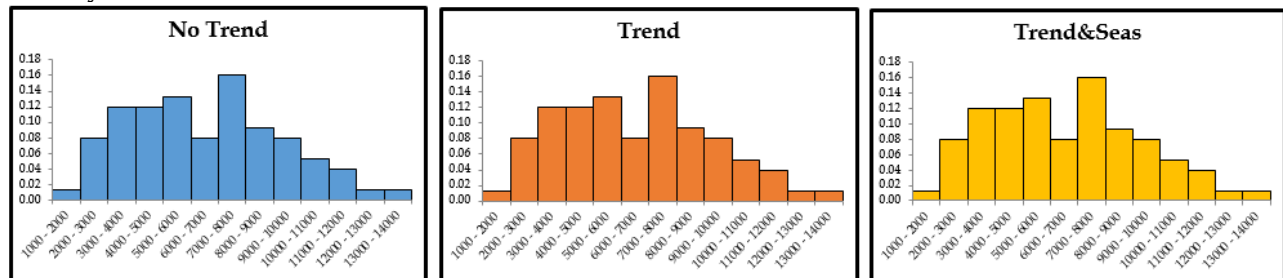
- Long-term forecasts are less accurate than short-term forecasts. Forecast accuracy diminishes as we look further into the future. As time passes, we get better information and make better predictions.

How enterprises may benefit from forecasting techniques characteristics. The combined ports of Los Angeles and Long Beach, also known as San Pedro Bay Ports, are ranked 5th in the world in respect to container handling, after the Port of Singapore and three ports in China. More than 50 percent of containers coming to the U.S. pass through San Pedro Bay ports and more than 1/3 of the containerized products consumed in all other states pass through these ports. The total value of the trade is about \$300 billion, creating around \$30 billion in state and local taxes, and 3 million full-time equivalent jobs. San Pedro Bay ports and Southern California need to retain their competitive edge, otherwise, other potential routes will take business from Southern California and its ports. What are the competing edges of Southern California ports? Deepwater facilities for post Panama ships, which may contain more than 8,000 containers; state of the art on-dock facilities to transfer containers between ship and train; intermodal transfer between ship, truck, and train; consolidation and distribution facilities for trans-loading from 20-foot containers and 40-foot containers to 56-foot containers, which are allowed to move on California roads, but as important as this capability is and maybe more important than this capability, are the common characteristics of all forecasting techniques.



If we want to transfer the load from the Far East to the East Coast, it will take 4 weeks. From the Far East to West Coast takes 2 weeks, and from the Far East to the mid-U.S. takes something between 2 to 4 weeks. Now, if I am going to ship loads from the Far East to East Coast, I should forecast the demand of the East Coast 4 weeks in advance. If I am going to ship from the Far East to West Coast, I should estimate the demand for the West Coast 2 weeks in advance. Estimates of the West Coast, which requires a forecast of 2 weeks, is more accurate than for the East Coast, which requires a 4-week forecast. Shorter time provides more accuracy. Look at the other property. All of the forecasts for the East Coast, West Coast, and mid-U.S. are less accurate than the forecast for the total demand in the U.S. So instead of forecasting for the East Coast alone for 4 weeks, the West Coast alone for 2 weeks, and the mid-U.S. for 3 weeks, I forecast the total demand for the U.S. for 14 days, 2 weeks in advance. Then when I send the container here, in one day I may transfer it to anywhere in California, in 2-3 days I may transfer it to somewhere in the mid-U.S., and in 3 to 4 days I may transfer it to somewhere in the East Coast. Now instead of estimating the demand of the East Coast alone, which is less accurate than the demand for the whole U.S., and instead of forecasting it for 4 weeks from now, I can forecast it for 14 days plus 3 days, which is 17 days from now. The forecast for the whole U.S. between 14 days and 17 days in advance is much more accurate than the forecast for the East Coast of 4 weeks in advance and the forecast for the mid-U.S. of 3 weeks in advance.

Time Series Analysis. The Histograms for three sets of data are shown below. They look exactly the same.

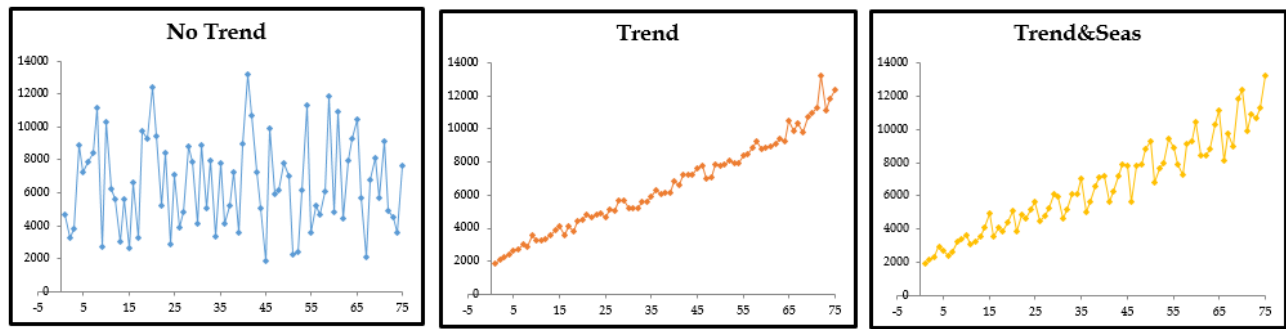


The descriptive statistics of these 3 sets are exactly the same with all decimal points.

Statistics	Set 1	Set 2	Set 3
Mean	6561.5467	6561.5467	6561.5467
Max	13229.0000	13229.0000	13229.0000
Min	1888.0000	1888.0000	1888.0000
Range	11341.0000	11341.0000	11341.0000
StdDev	2769.2898	2769.2898	2769.2898
CV	0.4220	0.4220	0.4220
Range/Mean	1.7284	1.7284	1.7284
Skew	0.3132	0.3132	0.3132

But, they are times series. They have happened over a set of equal time intervals. While the descriptive statistics as a snapshot of each set are exactly the same, the sets are very

different if we draw them as they have been observed along a time-related axis. Time series analysis is a quantitative forecasting technique applying a series of historical data collected in a regular interval (e.g., daily, weekly, monthly, and hourly, yearly).



Time series is based on analysis of the past data to identify some systematic components and extend them into the future for better forecasting. In order to gain the capability to extend them into the future, and has the ability to show some random components. This is done with respect to any variable of interest, but because we are talking about forecasting for demand, a value of interest is always demand, unless stated otherwise. Time series may contain three systematic components.

- **Level.** Where the demand is.
- **Trend.** To what direction the demand will move. Upward or downward?
- **Seasonality.** Predictable fluctuations. Does the demand peak (or valley) around specific times?

We can identify and quantify systematic components, but there is no way to predict or control random components. This is why forecasts are never precise.

There are two well-known time series analysis methods, (i) Moving Average, and (ii) Exponential Smoothing. In order to set the stage for moving average, let's compare two extreme points. On extreme left, we may decide to have our forecast for the next period to be equal to the actual for this period. On the extreme right, we may set our forecast for the next period equal to the average of all the data for all earlier periods. Two extreme points, one relies only on one piece of data, and the other relies on all pieces of data.

Problem 1- All Periods vs. Last Period. In some of the examples that I use, the input data are random. That means the input data changes each time we hit enter on the keyboard. Therefore, one can extract an almost infinite number of problems out of a single problem. However, in order not to allow the data to continually change when we are solving a problem, we usually copy the random data and paste it by-value into a different part of the Excel sheet. We then work on the paste-by value-data, which are now fixed at specific values. In the current example, random data are in column A, and we have pasted them by value in column E. In this example, the random data is in column A, where we have pasted it by value in column E. You may click on the figure below to access the Excel sheet.

	A	B	C	D	E	F	G	H	I
1									
2	Changing	Fixed	t	t	At	Ave	1-P	2-P	3-P
3	10	10	1	Jan	10	29.17			
4	10	10	2	Feb	10	29.17	10		
5	13	13	3	Mar	13	29.17	10	10.0	
6	10	10	4	Apr	10	29.17	13	11.5	11.00
7	10	10	5	May	10	29.17	10	11.5	11.00
8	18	18	6	Jun	18	29.17	10	10.0	11.00
9	19	19	7	Jul	19	29.17	18	14.0	12.67
10	44	44	8	Aug	44	29.17	19	18.5	15.67
11	36	36	9	Sep	36	29.17	44	31.5	27.00
12	56	56	10	Oct	56	29.17	36	40.0	33.00
13	49	49	11	Nov	49	29.17	56	46.0	45.33
14	75	75	12	Dec	75	29.17	49	52.5	47.00
15			13			29.17	75	62.0	60.00

Suppose we have the data for the past 12 periods (days, weeks, months, etc.) and we want to forecast for the next period. In this chapter we may refer to actual values in period t interchangeably as A_t (where t is a subscript) or simply as A_t . Whenever possible, we avoid subscript in our notations. Similarly, we refer to the forecast values in period t interchangeably as F_t and F_t .

Given the past data, now we want to forecast for the next period. On the extreme right, we may stay very conservative, and assume that the forecast for the next period is equal to the average of the past 12 periods. Therefore,

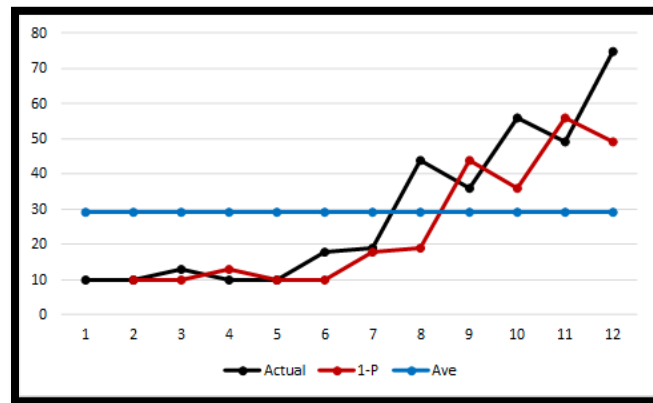
$$F_{13} = (A_{12} + A_{11} + A_{10} + \dots + A_3 + A_2 + A_1) / 12.$$

In that case, in cell F3 we may type = **AVERAGE (E3:E14)**. However, since we need to copy it down to all cells, we may push F4, and the formula will change to = **AVERAGE (\$E\$3:\$E\$14)**. Now we can copy it down to all periods. Following this conservative procedure, our forecast for period 13 is 29.17.

Last Period Forecast. From the extreme left, we may assume that the forecast for the next period is equal to the actual for this period.

$$F_{t+1} = A_t$$

Under the last period forecast, we can forecast starting from period 2. The forecast for period 13 is equal to the actual in period 12, and that is $F_{13} = A_{12} = 75$. The two extreme approaches are pictorially compared in the following figure. As we can see, one forecast is absolutely smooth with no response to the current changes, and the other is purely reflecting the most recent changes.



The last period forecast and the average of all periods forecast are simple, straightforward, and inexpensive. We can always use them as bases to compare the quality of other forecasting techniques. If the quality of other forecasting techniques are lower, then we can always use one of the last period forecasts or the average of all periods forecast- the better of the two.

However, as in many other situations, perhaps neither extreme right nor extreme left provide the best solution. Staying somewhere on the continuum between the two extremes may provide a better solution. That is the core idea behind the moving average approach.

One step to the right from the last period forecast is to compute the average of the last two periods. We may average every two consecutive periods and assume it as the forecast for the next period. In that case, our first forecast is for period 3 where $F_3 = MA_2^2 = (A_2 + A_1)/2$. Then we follow the same logic for period 4, and so on. We may also move one more step towards a smoother and less aggressive forecast, and compute 3-period-MA. In that case, we cannot forecast for periods 1-3, and our first forecast is for period 4, where $F_4 = MA_3^3 = (A_3 + A_2 + A_1)/3$. For example, forecasts for period 8, using 2-period and 3-period-MAs are $F_8 = MA_2^7 = (A_7 + A_6)/2$, and $F_8 = MA_3^7 = (A_7 + A_6 + A_5)/3$, respectively. Using 2- and 3-period-MAs, our forecasts for period 13 are $F_{13} = (A_{12} + A_{11})/2 = (49 + 75)/2 = 62$, and $F_{13} = (A_{12} + A_{11} + A_{10})/3 = (49 + 75 + 56)/3 = 60$, respectively.

Moving Average. This forecasting technique uses the average of the most recent actual data in a certain period to forecast the next period. The moving average calculated from the current period becomes the next period's forecast, a revolving average covering n periods -but not all the periods. As time passes, the forecast is renovated based on the most recent set of data. Moving Average (MA) calculates the forecast using the newest pieces of data and removes the oldest data.

Before writing the general moving average formula, let's compute a 4-period-MA for period 10. We need to start from A_{10} and include 4 pieces of the actual data. What is the index of the oldest piece of data? Is it $10 - 4 = 6$? No. Let's write the equation, $MA_{10}^4 = (A_{10} + A_9 + A_8 + A_7)/4$. Therefore, the index of the last piece of data is $10 - 4 + 1$. In a 25-periods-MA, in period 100, if the youngest piece of data belongs to period 100, the oldest piece of data belongs to what period? $100 - 25 + 1 = 76$

This knowledge can lead us to the general moving average forecasting formula as

$$F_{t+1} = MA_n = (A_t + A_{t-1} + \dots + A_{t-n+1}) / n$$

The Excel formulas in row 6 of the table for columns F to I are as follows:

F6	=AVERAGE(\$E\$3:\$E\$14)
G6	=E5
H6	=AVERAGE(E4:E5)
I6	=AVERAGE(E3:E5)

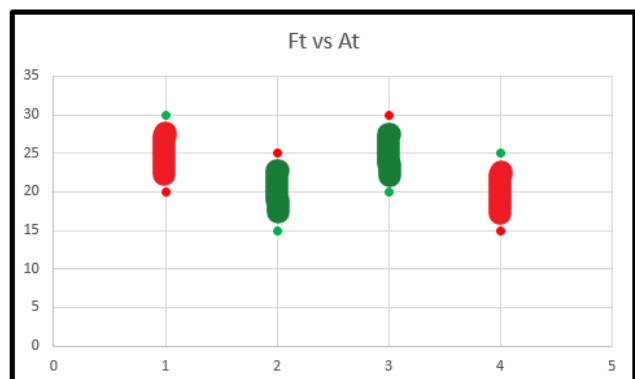
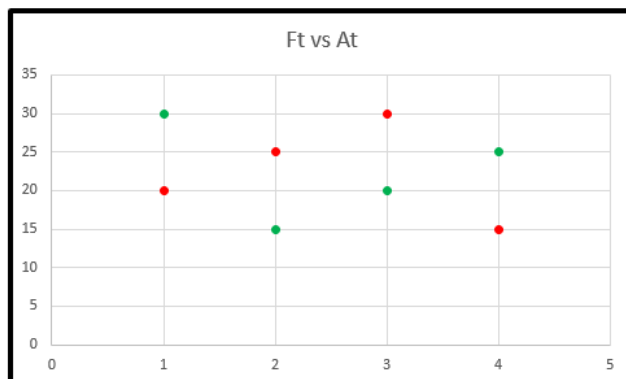
Using all, 1-, 2-, 3- period-MAs, our forecast for period 13 came out 29.17, 75, 62, and 64, respectively. Which one should we follow?



If you found the previous material easy, you may quickly go through the power point file at [Forecasting \(A\) Moving and Forecasting Measures](#). If you need help, the lecture on these slides is recorded and its available on the third slide. You may watch it, fast forward it, or rewind it if needed.

Measuring Forecast Accuracy and Variability. It's now the time to talk about defining the measure of effectiveness to find the suitability of the forecasting technique to our specific set of data. How could I measure the suitability of a forecasting technique for my specific set of data? How can I compare the quality of several forecasting techniques for my specific set of data?

Suppose the green dots in the following figures are our actual data, and the red dots are our forecasts based on a given method. As we can see, two of the forecasts under-estimate the demand and the other two over-estimate the demand. For simplicity, suppose the difference between actual and forecast in all periods have the same value with the exception that two of the values are positive and the other two are negative.



One measure of effectiveness to evaluate if a forecasting technique fits our data is to measure the gaps between actual and forecast of $E_t = A_t - F_t$. If we add the errors shown in the figure, the two negative values cross out the two positive values, and we get an

average of zero. If the average gap is 0, one may conclude that the forecasting method is perfect.

Going back to our 12-period data set, in the last period forecast, we need to compute $E_t = A_t - F_t$, for $t = 2$ to 12 (we do not have forecast for period 1). For 2-period-MAs we need to compute errors for all periods $t=3$ to 12, and for 3-period-MAs we need to compute $E_t = A_t - F_t$, for periods 4-12. Then we need to compute the summation or averages of the deviations for each method which could be misleadingly close to 0. Negative and positive components may cross each other out, while there may be significant gaps in both positive and negative sides. Since the variations on the negative side (forecasts being greater than actual) are undesirable too, the first task is to remove the negative signs. There are two well-known procedures to remove negative signs; absolute value, and square. If I have a value of -2, its absolute value is +2, and its square is +4. We can compute the absolute value of columns F to I, or square them, and then we can average absolute deviations or squared deviations for each of the procedures over the number of periods where E_t exists for the corresponding procedure. That would be the averages of 12, 11, 10, and 9 absolute, or squared errors for average, last period, 2-period and 3-period-MAs, respectively. We formally define Mean Absolute Deviation (MAD) and Mean Squared Error (MSE) as follows.

$MAD = AVERAGE(E)$ over the periods for which both A_t and F_t exist.

$MSE = AVERAGE(E^2)$ over the periods for which both A_t and F_t exist.

MAD for 1-period-MA is computed over 11 periods, while 3-period-MA is computed over 9 periods. If we compute MAD or MSE to compare two or more forecasting procedures, then they must be computed only over the periods for which actual data and forecasts are available for all procedures. Therefore if we are comparing 1- and 3-period-MA, we need to compare them over periods 4 to 12. MAD and MSE are computed for all the methods over periods 4-12, in the following Excel table. **By clicking on the table, it will take you to the Excel sheet where you can check the different formula in the corresponding cells.**

	A	C	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1								E				E ²				E /A					
2	Changing	t	Actual	Ave	1-P	2-P	3-P	E (Ave)	E (1P)	E (2P)	E (3P)	E ² (Ave)	E ² (1P)	E ² (2P)	E ² (3P)	E /A(Ave)	E /A(1P)	E /A(2P)	E /A(3P)		
3	24	1	10	29.17				19.17													
4	10	2	10	29.17	10			19.17	0												
5	10	3	13	29.17	10	10.0		16.17	3	3.00											
6	38	4	10	29.17	13	11.5	11.00	19.17	3	1.50	1.00	367.36	9.00	2.25	1.00	1.92	0.30	0.15	-0.10		
7	44	5	10	29.17	10	11.5	11.00	19.17	0	1.50	1.00	367.36	0.00	2.25	1.00	1.92	0.00	0.15	-0.10		
8	13	6	18	29.17	10	10.0	11.00	11.17	8	8.00	7.00	124.69	64.00	64.00	49.00	0.62	0.44	0.44	0.39		
9	20	7	19	29.17	18	14.0	12.67	10.17	1	5.00	6.33	103.36	1.00	25.00	40.11	0.54	0.05	0.26	0.33		
10	20	8	44	29.17	19	18.5	15.67	14.83	25	25.50	28.33	220.03	625.00	650.25	802.78	0.34	0.57	0.58	0.64		
11	26	9	36	29.17	44	31.5	27.00	6.83	8	4.50	9.00	46.69	64.00	20.25	81.00	0.19	0.22	0.13	0.25		
12	57	10	56	29.17	36	40.0	33.00	26.83	20	16.00	23.00	720.03	400.00	256.00	529.00	0.48	0.36	0.29	0.41		
13	72	11	49	29.17	56	46.0	43.33	19.83	7	3.00	3.67	393.36	49.00	9.00	13.44	0.40	0.14	0.06	0.07		
14	54	12	75	29.17	49	52.5	47.00	45.83	26	22.50	28.00	2100.69	676.00	506.25	784.00	0.61	0.35	0.30	0.37		
15		13		29.17	75	62.0	60.00	MAD (Mean Absolute Deviation)				MSE (Mean Square Error)				MAPE (Mean Absolute Percentage Error)					
16								19.31	10.89	9.72	11.93	493.73	209.78	170.58	255.70	0.78	0.27	0.26	0.25		
17										Best				Best					Best		

There are several points to mention. First, The MAD computation was easier when it was first implemented about a century ago. However, it is much more difficult to work with an

absolute component in mathematical expressions. It is not difficult to deal with squatted functions in mathematical expressions. Second, 1.25MAD is an estimate for the standard deviation of the forecast, and MSE is also an estimate for the standard deviation of the forecast. Third, sometimes we may assign a higher weight to a positive gap (where the forecast is less than the actual demand) than a negative gap (where the forecast is estimated at more than what the actual demand is). In the latter case, we are over stock, while in the first case we have lost sales. Usually, the cost of overstock is less than the cost of lost sales. In these cases, we may assign a coefficient of greater than 1 to negative values of $E_t = A_t - F_t$. Alternatively, we may use the newsvendor problem insight – discussed in ROP discussions- to find a good greater than 1 coefficient. Fourth, Besides MAD and MSE, there is a third method where we refer to it as a relative mean absolute deviation or R-MAD. In this third method, instead of averaging $|E_t|$ values, we average $|E_t| / A_t$ values. For example, a $|E_t|$ of 10 states there were 10 units of deviations between A_t and F_t . If A_t is 200, then 10 relative to 200 is a .05 (or 5% gap). Therefore, instead of the absolute gaps, the relative absolute gaps (relative to the demand) are taken into account. Fifth, there is still a fourth measure which is the maximum absolute deviation- where the preferred method is the method with a minimal maximum absolute gap between actual and forecast.

Mean absolute deviation serves two important purposes. I can use it to evaluate a single forecasting method or to compare two or more forecasting methods to identify the best one. MAD also provides an estimate of the standard deviation of demand forecast. The standard deviation of the forecast is equal to 1.25 MAD . Therefore, we can have F_{t+1} as the average demand for the next period, and 1.25 of the most recent MAD as the standard deviation of the forecast.



Now we are prepared to watch the lecture on excel based moving average embedded in the excel file at [1.MovingAverage2018.xlsx](#)

After understanding all the functions in the Excel sheet, please set column E equal to column A, and then go to an empty cell and click delete bottom several times. Each time you hit delete, your data changes and there are two interesting observations. First, the three methods do not necessarily get the same results, as one may prefer a 2-period-MA and the other a 3-period-MA. Second, if we only consider one of the three measures, it will not always point to a specific method as the best method. For example, as we change data, for one set of data, in the framework of MAD measure of effectiveness, each time we may find a different method (average, 1-, 2, or 3-period-MA) preferable. That means the choice of the method depends on the nature of the data.

The tracking signal is defined as the summation of the difference between actual and forecast, so we keep the positive and negative signs.

If actual is greater than forecast, the sign is positive, otherwise, it is negative divided by MAD. And because MAD by itself is a summation of the absolute value of the difference between actual and forecast divided by the number of periods, then if I replace MAD with this equation, I will get this formula. Note that MAD is always positive because it is a

summation of the actual forecast with absolute value signs divided by the number of periods. Therefore, MAD is always positive, but the numerator of tracking signal can be positive or negative. So tracking signal by itself could be positive or negative.

$$TS = \text{SUM}(Et) / \text{MAD}$$

We also define an LCL and a UCL for TS. In some textbooks, it is assumed that tracking signal should change between -4 and +4. But one can mathematically prove that these numbers are not correct.



Assignment Moving Average and Forecasting Measures Problems

Problems 1-6 are very easy problems. You only need to go through them if you still have difficulties understanding the moving average concepts.

Problems 7-13 are interesting problems, you definitely need to go through them.

Age of Data. What is the age of data in a 5 period moving average? The most recent piece of data is 1 period old, the second piece of data is 2 periods old, the third piece of data is 3 periods old, the fourth piece of data is 4 periods old, and the fifth piece of data is 5 periods old. On average, data is $(1+2+3+4+5)/5 = 15/5 = 3$ periods old. We could have reached the same number by stating that the newest piece of data is 1 period old, the oldest piece of data is 5 periods old, therefore, on average, the data is $(1+5)/2 = 3$ periods old. In general, in an n-period moving average, the data is $(1+n)/2$ periods old.

Weighted Moving Average. A weighted moving average allows to assign different weights to the data used in a moving average. Usually the weights go down as the age of the data in the moving average is increased. For example, usually the weight of the most recent piece of data, i.e., this period's data is higher than the weight of all other data contributing in the moving average. The sum of the weights must be equal to 1 (100%). If it is not, we need to transform it to 1 by defining each new weight to be equal to the old weight divided by sum of the weights. Forecasting based on an n period weighted moving average is stated as follows

$$F_t = w_1A_{t-1} + w_2A_{t-2} + w_3A_{t-3} + \dots + w_nA_{t-n}$$

Example: Determine the forecast for period 13 in the following data using four period moving average where the weights are $W_1=0.4$, $W_2=0.3$, $W_3=0.2$, and $w_4=0.1$ the 9th period, using the weights of 0.5 (the most recent period), 0.3 (the second most recent period), and 0.2 (the third most recent period).

w	t	At	Ft					
0.1	1	10						
0.2	2	10						
0.3	3	13						
0.4	4	10						
	5	10	10.9	=SUMPRODUCT(\$A\$3:\$A\$6,C3:C6)				
	6	18	10.6	=SUMPRODUCT(\$A\$3:\$A\$6,C4:C7)				
	7	19	13.5	=SUMPRODUCT(\$A\$3:\$A\$6,C5:C8)				
	8	44	16	=SUMPRODUCT(\$A\$3:\$A\$6,C6:C9)				
	9	36	27.9	=SUMPRODUCT(\$A\$3:\$A\$6,C7:C10)				
	10	56	33.2	=SUMPRODUCT(\$A\$3:\$A\$6,C8:C11)				
	11	49	43.9	=SUMPRODUCT(\$A\$3:\$A\$6,C9:C12)				
	12	75	48	=SUMPRODUCT(\$A\$3:\$A\$6,C10:C13)				
			59.5	=SUMPRODUCT(\$A\$3:\$A\$6,C11:C14)				

Use solver to find the best weights were we should satisfy the $w_4 \geq w_3 \geq w_2 \geq w_1$ constraint.

Exponential Smoothing. Exponential Smoothing (ES): is another forecasting technique where forecast for the next period is the same as forecast for this period plus a fraction of the gap between the actual and the forecast of this period. The fraction is defined by coefficient α which is between 0 and 1. Suppose we are in period 3, suppose $A_3=10$, $F_3=12$, and $\alpha=0.5$. Then our forecast for period 4 is $F_4 = F_3 + \alpha(A_3 - F_3) \rightarrow F_4 = 12 + 0.5(10 - 12) \rightarrow F_4 = 11$.

In general the exponential smoothing formula is stated as

$$F_{t+1} = F_t + \alpha (A_t - F_t)$$

The forecast for the first period is usually set to the actual for the first period.

We can manipulate the formula and write it in a different form

$$F_{t+1} = F_t + \alpha (A_t - F_t)$$

$$F_{t+1} = F_t + \alpha A_t - \alpha F_t$$

$$F_{t+1} = (1 - \alpha)F_t + \alpha A_t$$

We usually use the latter.

Exponential Smoothing is a weighted moving average were all the past data are present. The weight of data decreases as their age increases.

When comparing moving average with weighted moving average, there is a distinct difference between the two methods.

- In a moving average, the periods are all weighted equally and therefore given equal importance.
- In a weighted moving average, there is more “weight” given to the most recent data and less “weight” given to older data. The equations below clearly show these differences.

4-Period-Moving-Average (4-Period)

$$F_{t+1} = (A_t + A_{t-1} + A_{t-2} + A_{t-3})/4$$

We can write it as

$$F_{t+1} = A_t/4 + A_{t-1}/4 + A_{t-2}/4 + A_{t-3}/4$$

In other words

$$F_{t+1} = 0.25A_t + 0.25A_{t-1} + 0.25A_{t-2} + 0.25A_{t-3}$$

Therefore, a moving average is also a weighted moving average where the weights are equal as opposed to an actual weighted moving average which may look like

$$F_{t+1} = 0.35A_t + 0.30A_{t-1} + 0.25A_{t-2} + 0.1A_{t-3}$$

Note that the sum of all coefficients for each technique is equal to 1.

Example: Raw Materials for Production. In order to meet market demand for a rare disease medicine, the company Drugs-R-Us needs to purchase enough raw materials to support production requirements. However, the Warehouse Manager wants to keep inventory as low as possible while still supporting production. This will minimize tax expenses from storing surplus inventory at the end of each month. The table below shows the forecast for material purchases using the exponential smoothing method with an alpha of 0.4. It is noteworthy to mention that the forecast in period 2 was obtained using the Naive Method.

		$\alpha=0.4$				
t	At	Ft		E = At-Ft		MAD
1	130	130.0	=B3	0.00	=ABS(B3-C3)	0.00 =AVERAGE(\$E\$3:E3)
2	150	130.0	=(1-\$C\$1)*C3+\$C\$1*B3	20.00	=ABS(B4-C4)	10.00 =AVERAGE(\$E\$3:E4)
3	80	138.0	=(1-\$C\$1)*C4+\$C\$1*B4	58.00	=ABS(B5-C5)	26.00 =AVERAGE(\$E\$3:E5)
4	190	114.8	=(1-\$C\$1)*C5+\$C\$1*B5	75.20	=ABS(B6-C6)	38.30 =AVERAGE(\$E\$3:E6)
5	140	144.9	=(1-\$C\$1)*C6+\$C\$1*B6	4.88	=ABS(B7-C7)	31.62 =AVERAGE(\$E\$3:E7)
6	150	142.9	=(1-\$C\$1)*C7+\$C\$1*B7	7.07	=ABS(B8-C8)	27.53 =AVERAGE(\$E\$3:E8)
7	120	145.8	=(1-\$C\$1)*C8+\$C\$1*B8	25.76	=ABS(B9-C9)	27.27 =AVERAGE(\$E\$3:E9)
8	140	135.5	=(1-\$C\$1)*C9+\$C\$1*B9	4.55	=ABS(B10-C10)	24.43 =AVERAGE(\$E\$3:E10)
9	100	137.3	=(1-\$C\$1)*C10+\$C\$1*B10	37.27	=ABS(B11-C11)	25.86 =AVERAGE(\$E\$3:E11)
10	60	122.4	=(1-\$C\$1)*C11+\$C\$1*B11	62.36	=ABS(B12-C12)	29.51 =AVERAGE(\$E\$3:E12)
11	11	97.4	=(1-\$C\$1)*C12+\$C\$1*B12	86.42	=ABS(B13-C13)	34.68 =AVERAGE(\$E\$3:E13)
12	150	62.9	=(1-\$C\$1)*C13+\$C\$1*B13	87.15	=ABS(B14-C14)	39.05 =AVERAGE(\$E\$3:E14)
13		97.7	=(1-\$C\$1)*C14+\$C\$1*B14			
Forecast for next period:				$\mu=97.7$		$\sigma=48.8$
				=C15		=1.25*G14

Another forecast was generated for comparison. The exponential smoothing results below were generated using a α -value of 0.8.

		$\alpha=0.8$				
t	At	Ft		E = At-Ft		MAD
1	130	130.0	=B3	0.00	=ABS(B3-C3)	0.00 =AVERAGE(\$E\$3:E3)
2	150	130.0	=(1-\$C\$1)*C3+\$C\$1*B3	20.00	=ABS(B4-C4)	10.00 =AVERAGE(\$E\$3:E4)
3	80	146.0	=(1-\$C\$1)*C4+\$C\$1*B4	66.00	=ABS(B5-C5)	28.67 =AVERAGE(\$E\$3:E5)
4	190	93.2	=(1-\$C\$1)*C5+\$C\$1*B5	96.80	=ABS(B6-C6)	45.70 =AVERAGE(\$E\$3:E6)
5	140	170.6	=(1-\$C\$1)*C6+\$C\$1*B6	30.64	=ABS(B7-C7)	42.69 =AVERAGE(\$E\$3:E7)
6	150	146.1	=(1-\$C\$1)*C7+\$C\$1*B7	3.87	=ABS(B8-C8)	36.22 =AVERAGE(\$E\$3:E8)
7	120	149.2	=(1-\$C\$1)*C8+\$C\$1*B8	29.23	=ABS(B9-C9)	35.22 =AVERAGE(\$E\$3:E9)
8	140	125.8	=(1-\$C\$1)*C9+\$C\$1*B9	14.15	=ABS(B10-C10)	32.59 =AVERAGE(\$E\$3:E10)
9	100	137.2	=(1-\$C\$1)*C10+\$C\$1*B10	37.17	=ABS(B11-C11)	33.10 =AVERAGE(\$E\$3:E11)
10	60	107.4	=(1-\$C\$1)*C11+\$C\$1*B11	47.43	=ABS(B12-C12)	34.53 =AVERAGE(\$E\$3:E12)
11	11	69.5	=(1-\$C\$1)*C12+\$C\$1*B12	58.49	=ABS(B13-C13)	36.71 =AVERAGE(\$E\$3:E13)
12	150	22.7	=(1-\$C\$1)*C13+\$C\$1*B13	127.30	=ABS(B14-C14)	44.26 =AVERAGE(\$E\$3:E14)
13		124.5	=(1-\$C\$1)*C14+\$C\$1*B14			
Forecast for next period:				$\mu=124.5$		$\sigma=55.3$
				=C15		=1.25*G14

Since the MAD value for $\alpha = 0.4$ is smaller than the $\alpha = 0.8$, the better forecast is obtained with the smaller alpha. This is not a general rule, that means we do not know whether smaller or a larger α is the best fit to our data. However, it is a general rule that the smaller α -value is less reactive and is better for long range planning purposes. The larger α -value is more reactive and best utilized for short range forecasts. Indeed, if we set $\alpha = 0$, that means we pay no attention to reality, and stay without original forecast. In contrast, $\alpha = 1$ means that we put our full trust on the most recent piece of data.

Question $\alpha = 1$ is the same as what moving average? One-period moving average.

These data sets were generated using 2 α -values and it was determined $\alpha = 0.4$ is better than $\alpha = 0.8$. To determine the best α -value for the best exponential smoothing forecast,

the following table was generated using the “What-If-Analysis” option in Excel and the starting with $\alpha = 0.4$, and $MAD = 39$.

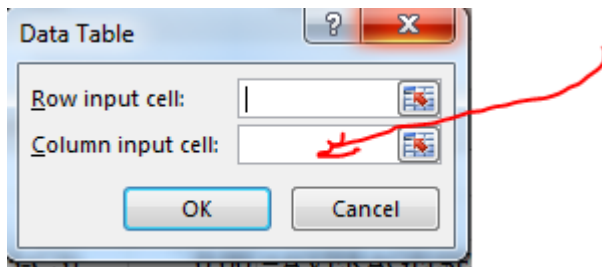
Type 0 in one cell and 0.1 in the cell below it.

0			
0.1			
	44.26		
0		=J1	
0.1		=J5+\$J\$2	
0.2		=J6+\$J\$2	
0.3		=J7+\$J\$2	
0.4		=J8+\$J\$2	
0.5		=J9+\$J\$2	
0.6		=J10+\$J\$2	
0.7		=J11+\$J\$2	
0.8		=J12+\$J\$2	
0.9		=J13+\$J\$2	
1		=J14+\$J\$2	

J	K	L	M
0			
0.1			
	44.26		
0		=J1	
0.1		=J5+\$J\$2	
0.2		=J6+\$J\$2	
0.3		=J7+\$J\$2	
0.4		=J8+\$J\$2	
0.5		=J9+\$J\$2	
0.6		=J10+\$J\$2	
0.7		=J11+\$J\$2	
0.8		=J12+\$J\$2	
0.9		=J13+\$J\$2	
1		=J14+\$J\$2	

Mark from J3 to K14. That is a table with two column and 12 rows where current MAD is in K4, and potential α values are typed in column J from J4 to J14.

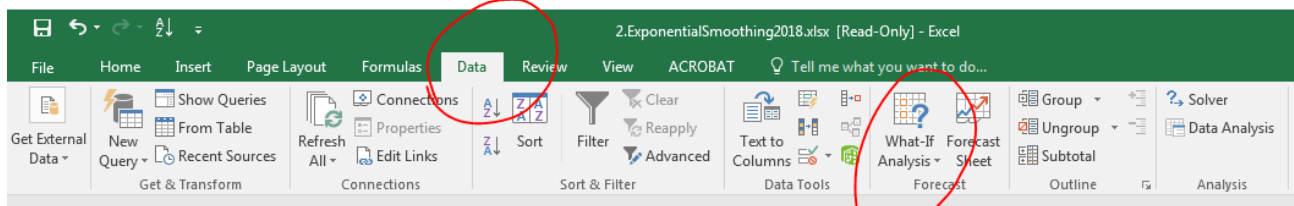
Go to what if analysis, and click on data table.



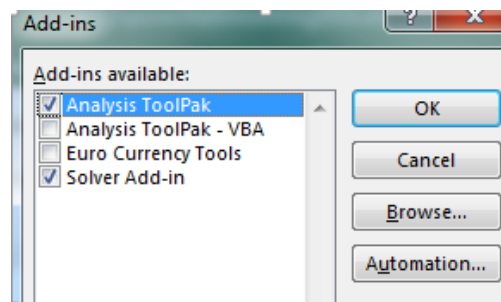
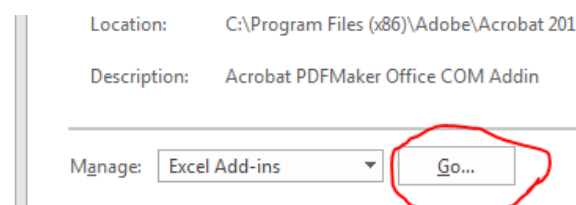
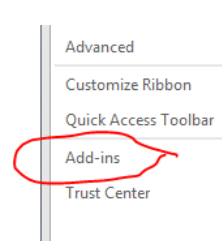
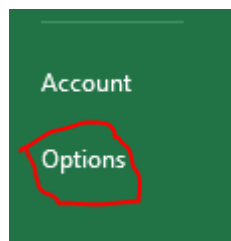
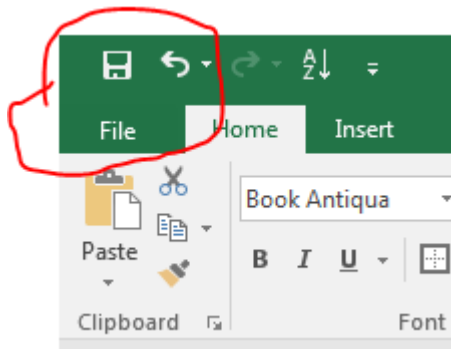
Your data is not typed in a row, they are typed in a column. Put your mouse in “Column input cell”. After having the mouse in “Column input cell”, click on the cell containing the α value. The table is automatically filled.

Chapter 4. Predictive Analytics I

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		$\alpha=$	0.4							0				
2	t	At	Ft		$ E = At - Ft $		MAD			0.1				
3	1	130	130.0	=B3	0.00	=ABS(B3-C3)	0.00	=AVERAGE(SES3:E3)						
4	2	150	130.0	=(1-\$C\$1)*C3+\$C\$1*B3	20.00	=ABS(B4-C4)	10.00	=AVERAGE(SES3:E4)			39.05			
5	3	80	138.0	=(1-\$C\$1)*C4+\$C\$1*B4	58.00	=ABS(B5-C5)	26.00	=AVERAGE(SES3:E5)		0	34.9167		=J1	
6	4	190	114.8	=(1-\$C\$1)*C5+\$C\$1*B5	75.20	=ABS(B6-C6)	38.30	=AVERAGE(SES3:E6)		0.1	36.3043		=J5+\$J\$2	
7	5	140	144.9	=(1-\$C\$1)*C6+\$C\$1*B6	4.88	=ABS(B7-C7)	31.62	=AVERAGE(SES3:E7)		0.2	37.2378		=J6+\$J\$2	
8	6	150	142.9	=(1-\$C\$1)*C7+\$C\$1*B7	7.07	=ABS(B8-C8)	27.53	=AVERAGE(SES3:E8)		0.3	37.9191		=J7+\$J\$2	
9	7	120	145.8	=(1-\$C\$1)*C8+\$C\$1*B8	25.76	=ABS(B9-C9)	27.27	=AVERAGE(SES3:E9)		0.4	39.0548		=J8+\$J\$2	
10	8	140	135.5	=(1-\$C\$1)*C9+\$C\$1*B9	4.55	=ABS(B10-C10)	24.43	=AVERAGE(SES3:E10)		0.5	40.181		=J9+\$J\$2	
11	9	100	137.3	=(1-\$C\$1)*C10+\$C\$1*B10	37.27	=ABS(B11-C11)	25.86	=AVERAGE(SES3:E11)		0.6	41.3821		=J10+\$J\$2	
12	10	60	122.4	=(1-\$C\$1)*C11+\$C\$1*B11	62.36	=ABS(B12-C12)	29.51	=AVERAGE(SES3:E12)		0.7	42.7226		=J11+\$J\$2	
13	11	11	97.4	=(1-\$C\$1)*C12+\$C\$1*B12	86.42	=ABS(B13-C13)	34.68	=AVERAGE(SES3:E13)		0.8	44.2571		=J12+\$J\$2	
14	12	150	62.9	=(1-\$C\$1)*C13+\$C\$1*B13	87.15	=ABS(B14-C14)	39.05	=AVERAGE(SES3:E14)		0.9	46.0448		=J13+\$J\$2	
15	13		97.7	=(1-\$C\$1)*C14+\$C\$1*B14						1	48.1667		=J14+\$J\$2	
16	Forecast for next period:				$\mu=$	97.7		$\sigma=$	48.8					
17					=C15			=1.25*G14						



If you do not see Analysis tool pack or solver in your data tab, you may follow the steps below to access them.





It is now the time to go through the slides at [Forecasting \(B\) Exponential Smoothing](#)
The lecture on these slides is on the second slide.

[Assignment Exponential Smoothing Problems](#)

[Moving Average.excel](#)

[Exponential Smoothing.excel](#)

[Simple Linear Regression.repository.youtube](#)

[Simple Linear Regression](#)

[Multiple Regression.repository.youtube](#)

[Forecasting \(B2\) Regression](#)

[Regression Problems](#)

[Regression.excel](#)

[ECON-FIN-STAT Integration.youtube](#)

[ECON-FIN-STAT Integration.youtube.exl](#)



Please also watch the video inside the following excel file.

[Visualization-exl](#) Start from the first page, built a similar page, to make sure that you have learned some of the basic excel functions.

Build your own worksheets similar to each of the worksheets in the excel file. There will be a quiz on this subject.

[Line & Scatter Chart-exl](#)

[RowToMatrToRow-exl](#)

[DotPlotne & Cathegorical-Histogram-exl](#)

Chapter 3

Throughput Analysis

Throughput first, Inventory Second, Living in Cost World Last – First Leave the Cost World

I do not know what I may appear to the world; but to myself, I seem to have been only like a boy playing on the seashore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, while the great ocean of truth lay all undiscovered before me. Isaac Newton, 1643-1727

The key problems of this chapter can be accessed at

<http://www.csun.edu/~aa2035/CourseBase/Throughput/ThroughpuKeyProblemst.pptx>

The recorded lectures – if any- and the excel worksheets- if any- are embedded inside the PowerPoint slides. By clicking on them the mp3/mp4 files and excel pages will open

In this chapter, we will discuss Throughput and review Capacity. We will also review concepts such as Bottleneck and Utilization, and introduce Capacity Waste Factor (CWF).

Throughput is the number of flow units that pass through the process per unit of time. We assume that we produce on demand and Throughput is equal to Demand. In a synchronized system, Throughput = Demand, that is $R=D$.

Steps in computing Throughput.

1. Observe the process for a number of periods.
2. Measure the number of flow units that are processed per unit of time.
3. Compute the average number of flow units per unit of time.

An indirect way of computing Throughput: $R = I/T$.

Throughput is the average flow rate. Capacity is the maximum sustainable flow rate. In periods of heavy congestion, Throughput is equal to Capacity for a short period, in all other times, Throughput < Capacity. Throughput cannot be equal to Capacity for a long period, i.e. Utilization cannot be equal to 1 except for short periods.

Imagine a freeway where all the cars are driving exactly 65 mph, and the distance between each car is only 1 inch. As long as everyone maintains a speed of exactly 65 mph, that is fine. However, can they do that? What happens if just one car hits the breaks? How long does it take to clean the freeway? Do cars move on the freeway easier when Utilization is 1 and they are moving bumper to bumper, or when 50% of the freeway is empty $U = 0.5$, or when $U = 0.25$? How much may traffic jam accidents create in each of these situations?

Cycle Time is defined in relation to Capacity and is a measure of internal capability.
 $CT = 1/R_p$.

Takt Time is defined in relation to Demand or Throughput and is a measure of external demand. Takt means pace in German. If demand = 12/hour, then we assume the Throughput = 12/hour.

Takt Time = $1/\text{Demand} = 1/\text{Throughput} = 1/R$.

Takt Time = $1/12$ hour or $60(1/12) = 5$ minutes.

$R = \min \{R_a, R_p\}$, since $R_a < R_p$, \rightarrow Takt time = $1/R_a > \text{Cycle Time} = 1/R_p$

Average inter-arrival time = $T_a = 1/R_a = 1/R = \text{Average inter-departure time}$

Sometimes, a manager may state that his machines have takt time of X minutes. This statement is incorrect. Machines have cycle time. Capacity is related to cycle time. Takt time is a measure of external demand; it has nothing to do with the internal capacity. In a synchronized system, Takt time is the time each station has to send one flow unit to the next station. Cycle time is the time the bottleneck (s) need to send the next product out. TT is always greater than CT since R_p is always greater than R .

Chapter 5 is on **flow time minimization**. Chapter 4 is on **throughput maximization**.

Chapter 4 and 5 are both **"time" minimizations**. Why?

By trying to increase throughput, we will minimize Takt time. Therefore, Chapter 5 is on **flow time minimization**. Chapter 4 is on **takt time minimization**.

Problem 1. On the last column of the left table below, we have work content (activity time) of five sequential activities (T_p of each activity).

Activity Time or work content of an activity (T_p). The amount of time the activity needs to be completed once on a specific resource. This includes a share of all distractions such as maintenance, repair, setup, etc. They are all inside T_p ; it is the effective unit load, not the theoretical unit load.

Activity	Resource	Work Content (mins)
Mailroom	Mailroom Clerk	1
Data Entry	Data-entry Clerk	5
Initial Processing	Claims processor	6
Inspection	Claims Supervisor	2.5
Final Processing	Claims processor	2

Resource	Unit Load(min)
Mailroom Clerk	1
Data-entry Clerk	5
Claims processor	8
Claims Supervisor	2.5

The third and fifth activities are done on a single resource, therefore regarding the resources, we have four Tp values.



Unit Load of a Resource Unit (Tp) – The amount of time the resource works to process each flow unit. That is a summation of the work content (activity time) of all activities assigned to the resource. This includes a share of all distractions such as maintenance, repair, setup, etc. It is an effective unit load, not the theoretical unit load.

Unit Load of Resource-3 = Work Content of Activity-3 + Work Content of Activity-5

Tp of Resource-3 = Tp Activity-3 + Tp Activity-5

Tp of Resource-3 = Unit Load of Resource-3 = 6+2 = 8

Effective Capacity of a Resource Unit is computed as $R_p = 1/\text{unit load} = 1/T_p$ in the same unit of time as Tp is defined.

If in a **Resource Pool** we have two resource units, then the Effective Capacity of a Resource Pool is $1/T_p + 1/T_p = 2/T_p$. If in a resource pool we have c resource units, then the Effective Capacity of a Resource Pool is c/T_p . If there are c Resource Units in the Resource Pool then the capacity of the resource pool is $R_p = c/T_p$, per the time unit in which Tp is defined.

Resource	Unit Load: T_p min/claim	Effective capacity of Resource Unit: $R_p = 1/T_p$ claims/min	Number of the Resource Units in the Resource Pool: c	Effective capacity of Resource Pool: $R_p = c/T_p$ claims/min
Mailroom clerk	1	1	1	1
Data-entry clerk	5	0.2	8	1.6
Claims processor	8	0.125	12	1.5
Claims supervisor	2.5	0.4	5	2

Unit load of claim processor is 8 min. Capacity is $1/8 = 0.125$ units per minute. If we multiply 0.125 by 60 (minutes per hour), the capacity is then 7.5 per hour. The capacity of a single resource unit is $1/T_p$ per the time unit as T_p stated. If each resource unit has a capacity of $1/8$ and if we have 12 resources units ($c=12$) in our resource pool of claim processor, then the capacity of the resource pool is $c/T_p = 12/8 = 1.5$ per minute or 90 per hour. Following the same computations, capacity per minute of all four resource pools is shown in the fourth column.

Bottleneck – The resource pool with the minimum effective capacity.

The effective capacity of a process: The effective capacity of the bottleneck; i.e. 1 flow unit per minute or 60 per hour.

The bottleneck is the mailroom clerk. One way of increasing capacity is to cross-train Claims Supervisor to help Mailroom clerk! An excellent example of how companies can increase capacity, especially in peak periods, by cross-training, is UPS.

This is done at UPS headquarters whereas all managers have been trained to work at the line in peak times during high seasons.

100% Utilization is a High Risk. Suppose the throughput of the system is 400 per day, (capacity is 480 per day). All the components of the previous table remain the same in the following table. Utilizations are computed in the last column.

Resource Pool	Unit Load: T_p min/claim	Effective capacity of Resource Unit: $1/T_p$ claims/min	# of Resource Units in the Resource Pool: c_p	Effective capacity of the Resource Pool: $R_p = c/T_p$ claims/min	Scheduled availability min/day	Effective capacity of Resource Pool: claims/day	Utilization
Mailroom clerk	1	1	1	1	480	480	83%
Data-entry clerk	5	0.2	8	1.6	480	768	52%
Claims processor	8	0.125	12	1.5	480	720	56%
Claims supervisor	2.5	0.4	5	2	480	960	42%

Theoretical Flow Time = $1+5+8+2.5 \rightarrow FT = 16.5$

Capacity = 1/min. \rightarrow Cycle Time = 1/1 minute $\rightarrow CT = 1$.

Demand = 5/6 per minute \rightarrow Throughput = 5/6 per minute.

Takt Time = 1/Demand = 1/Throughput = 5/6 per minute

Takt Time = $1/(5/6) = \rightarrow TT = 1.2$ minutes



If you have difficulty in understanding this problem, you may first try to see if you can solve these problems [Assignment Waiting Line Problems Set 0](#). You may then access the PowerPoints slides of the above pages at [Throughput](#) and its recorded lecture at [Throughput.Recorded](#). The first 25 minutes of the lecture and its corresponding slides cover Problem 1. The last 25 minutes of the lecture and its corresponding slides cover Problem 2.

Problem 2. Suppose all the assumptions remain the same. However, there are two types of billing: Physician claims 60%, Hospital claims 40%. Unit loads of Physician claims are shown in column 2 and Hospital claims are shown in column 3. Unit load of a prototype flow unit is 0.6 (Physician claim unit load) + 0.4 (Hospital claim unit load). Therefore, the unit load of the data entry clerk is $0.6(5) + 0.4(6) = 5.4$ minutes. Unit loads of the remaining resources are computed following the same way. Now the effective capacity of the process is reduced to 400 per day. 400 units of a product which of 60% Physician claim and 40% Hospital claim; that is 240 Physician claims and 160 Hospital claims per day.

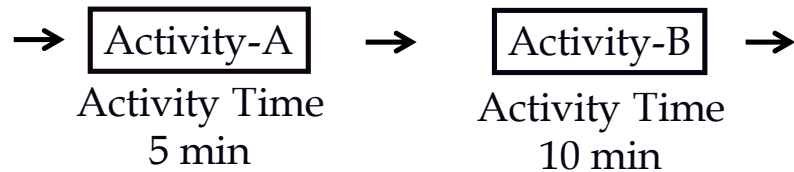
Resource Pool	UL- Physician min/claim	UL- Hospital min/claim	UL (60%- 40%) mix min/claim	Effective capacity of Resource Unit: $1/TP$ claims/min	# of Resouce Units in the Resource Pool: cp	Effective capacity of the Resource Pool: $R_p = cp/TP$ claims/min	Effective capacity of the Resource Pool: $R_p = cp/TP$ claims/day
Mailroom clerk	1	1.5	$0.6(1)+0.4(1.5) = 1.2$	0.83	1	0.83	400.00
Data-entry clerk	5	6	5.4	0.19	8	1.48	711.11
Claims processor	8	8	8	0.13	12	1.50	720.00
Claims supervisor	2.5	4	3.1	0.32	5	1.61	774.19

$R = R_a \rightarrow$ Takt time = $1/R_a$

Average inter-arrival time = $T_a = 1/R_a = 1/R$ = Average inter- exit time

In a synchronized system, Takt time is the time each station has to send one flow unit out to the next station.

Problem 3. Capacity Waste Factor (CWF). Reconsider the following example where Activities A and B are handled by resources 1 and 2 respectively.



T_p of Activity-A referred to as work content of activity-A, is 5 min., and T_p of Activity-B referred to as work content of Activity-B, is 10 min. T_p of Resource-1 referred to as unit load of resource-1, is 5 min., and T_p of Resource-2 referred to as unit load of Resource-2, is 10 min.

Now suppose Resource-A has a Capacity Waste Factor (CWF) of 40%, and Resource-B has a Capacity Waste Factor (CWF) of 25%. That means out of what we have as T_p of Resource-1, 25% of it is waste, and the rest is the Theoretical Unit Load.

Unit Load $(1-CWF)$ = Theoretical Unit Load.

Theoretical unit load of Resource-1 = $ThUL = T_p(1-CWF) = 5(1-0.4) = 3$ minutes.

Theoretical unit load of Resource-2 = $ThUL = T_p(1-CWF) = 10(1-0.25) = 7.5$ minutes.

Alternatively, if we had 3 and 7.5 as theoretical unit loads of Resource-1 and Resource-2, their unit load, T_p , could have been computed as

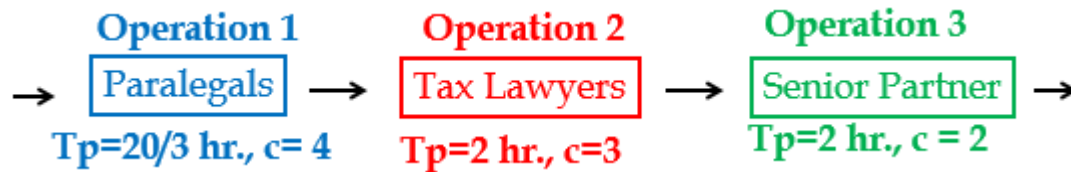
$T_p = ThUL / (1-CWF)$

Resource-1 $T_p = 3 / (1-0.4) = 5$ mins.

Resourc-2 $T_p = 7.5 / (1-0.25) = 10$ mins.

Note that if we do not consider the CWF, then we are talking about the theoretical capacity. If we take CWF into account, then we are talking effective capacity or simply capacity.

Problem 4. A law firm processes shopping centers and medical complexes contracts. There are four Paralegals, three Tax lawyers, and two Senior Partners. The unit loads of the resources to handle one standard contract is given below. Assume 8 hours per day and 20 days per month.



It takes a Paralegal 20 hours to complete 3 contracts. That is $20/3 = 6.667$ hours to complete a contract. It takes a Tax lawyer 2 hours to complete a contract. It takes Senior Partner 2 hours to complete a contract.

a) Compute the Flow Time of a contract.

$6.667 + 2 + 2 = 10.667$. This is not the flow time. It is the Theoretical Flow Time.

Flow Time = Theoretical Flow Time + Waiting times

Flow Time = 10.67 + Waiting times. We do not know the waiting times yet.

Compute the Capacity of each of the three Resource **Pools**.

A Paralegal can complete 1 contract in $20/3 = 6.667$ hours.

How many contracts in one hour? $1/6.667 = 0.15$

How many contracts can all the Paralegals complete in one hour?

There are 4 Paralegals: $c = 4$

Four Paralegals $4(0.15) = 0.6$ contracts per hour.

We could have also said that $T_p = 6.6667$ hours.

The capacity of one resource unit is $1/T_p$.

The capacity of one resource unit is $1/6.667 = 0.15$.

Capacity of all resource units: $R_p = c/T_p$ where $c = 4$ and $T_p = 6.667$

$R_p = 4/6.667 = 0.6$ per hour

Capacity of the resource pool is **0.6 contracts per hour**.

It is $8(0.6) = 4.8$ contracts **per day**

A Senior Partner can complete 1 contract in 2 hours.

How many contracts in one hour? $1/2 = 0.5$

How many contracts can all the Senior Partners complete in one hour?

There are two Senior Partners: $c = 2$

Therefore, Senior Partners $2(0.5) = 1$ contract per hour

We could have also said that $T_p = 2$.

The capacity of one resource unit is $1/T_p$.

The capacity of one resource unit is $1/2 = 0.5$.

Capacity of all resource units: $R_p = c/T_p$ where $c=2$ and $T_p = 2$

$R_p = 2/2 = 1$ per hour

The capacity of the resource pool is **one contract per hour**.

It is $8(1) = 8$ contracts **per day**

c) Compute the capacity of the process.

	Unit Load 50%SH 50%MD	Capacity of a Resource Unit /hr	# Of Resource Units	Capacity of the Resource Pool/hr	Cap of the R-Pool / day
Paralegal	6.667	0.15	4	0.6	4.8
Tax lawyer	2	0.5	3	1.5	12
Senior partner	2	0.5	2	1	8

d) Compute the cycle time?

4.8 units in 8 hours.

Cycle time = $8/4.8 = 1.67$ hours

Alternatively, $1/0.6 = 1.67$ hours

d) Compute the average inventory.

Let us look at the utilization of the three stations

Station	Capacity	Throughput	Utilization
Station 1	4.8	4.8	$4.8/4.8 = 1$
Station 2	12	4.8	$4.8/12 = 0.4$
Station 3	8	4.8	$4.8/8 = 0.6$

Chapter 4. Throughput

On average one person with a resource in Station 1, 0.4 people with a resource in Station 2, and 0.6 people with a resource in Station 3

Inventory with the processors is $1 + 0.4 + 0.6 = 2$

On average there are two flow units with the processors; Inventory in the processors (I)

Now let us look from another angle; from the Little's Law point of view

$RT = I \rightarrow R = 4.8$ per 8 hours or 0.6 per hour

$T = 10.67$ hours $\rightarrow I = 0.6(10.67) = 6.4$

6.4 vs 2? Where is my mistake?

$$1(4) + 0.4(3) + 0.6(2) = 6.4$$

e) There are 150 cases in November. Can the company process all 150 cases?

$150/20 = 7.5$ per day $\rightarrow 4.8$ (Capacity) < 7.5 (Demand).

f) If the firm wishes to process all the 150 cases available in November, how many professionals of each type are needed?

The capacity of a paralegal is 0.15 units per hour, therefore the daily capacity of this resource unit is $0.15(8) = 1.2$ units. Similarly, the daily capacity of a tax lawyer and a senior partner are $0.5(8) = 4$ units.

of paralegals required = $7.5/1.2 = 6.25$

of tax lawyers required = $7.5/4 = 1.875$

of senior partners required = $7.5/4 = 1.875$

These could be rounded up to 7, 2 and 2

We need 7, 2, and 2. We have 4, 3, and 2. We may hire three additional paralegals.

Alternatively, we may hire just two and for a total of six paralegals.

How much overtime for 0.25 paralegal who works eight hrs. /day?

$0.25(8) = 2$ hours total overtime.

There will be 6 paralegals; overtime pf each = $2/6 = 1/3$ hour

Alternatively, 20 minutes per paralegal. PLUS some safety Capacity.

Now suppose Throughput is 3.6 contracts per day.

Compute the average inventory.

Station	Capacity	Throughput	Utilization
---------	----------	------------	-------------

Chapter 4. Throughput

Station 1	4.8	3.6	$3.6/4.8 = 0.75$
Station 2	12	3.6	$3.6/12 = 0.3$
Station 3	8	3.6	$3.6/8 = 0.45$

Managerial Observation: Note that the utilization of bottleneck resource is not necessarily 100%.

On average

0.75 person with a resource in Station 1

0.3 person with a resource in Station 2

0.45 person with a resource in Station 3

Station	Rp	R	U	c	Ip
Station1	4.8	3.6	0.75	4	$4 * 0.75 = 3$
Station2	12	3.6	0.3	3	$3 * 0.3 = 0.9$
Station3	8	3.6	.45	2	$2 * 0.45 = 0.9$

$$I = 3 + 0.9 + 0.9 = 4.8$$

Let us check it through the Little's Law

$$\text{Theoretical Flow Time} = 6.66667 + 2 + 2 = 10.66667$$

$$RT = I \rightarrow 3.6(10.66667/8) = I \rightarrow I = 4.8$$

$$(3.6/8)(10.66667) = 4.8$$

Now suppose there are 16.8 contracts waiting in different waiting lines? What is the Flow Time?

$$RT = I \rightarrow 3.6 * T = 16.8 + 4.8 = 21.6 \rightarrow 3.6T = 21.6 \rightarrow T = 6 \text{ days.}$$



If you have difficulty in understanding this problem, you may access the PowerPoints slides of the problem at [Throughput Part 2a-Problem](#) and the recorded lecture at [Throughput Part2a-Problem.Recorded](#). The first 25 minutes of the lecture, and the last slides, covers Problem 1, and the last 5 minutes covers an extension of the problem. We encourage you to watch, at least, the last 5 minutes.

Capacity Waste Factor (CWF). The effective capacity of a resource unit is $1/T_p$. Unit load T_p is an aggregation of the productive as well as the wasted time.

T_p includes share of each flow unit of capacity waste and detractors such as

Resource breakdown

Maintenance

Quality rejects

Rework and repetitions

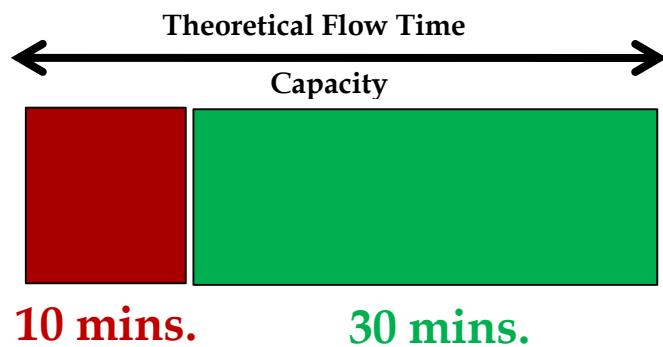
Setups between different products or batches

We may want to turn our attention to waste elimination and segregate the wasted capacity. Theoretical capacity is the effective capacity net of all capacity distracts.

Activity time

Capacity is computed based on the Unit Load

Theoretical Flow Time is computed based on Activity Time



Then what is the Flow Time?

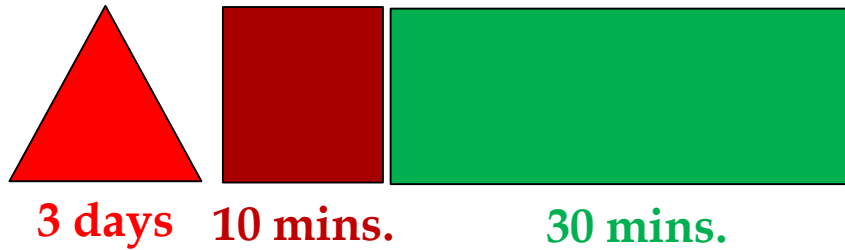
Unit Load, Activity time, Capacity, Theoretical Flow Time

Flow Time $T_i + T_p$

Flow time includes time in buffers

Capacity does not care about buffer times





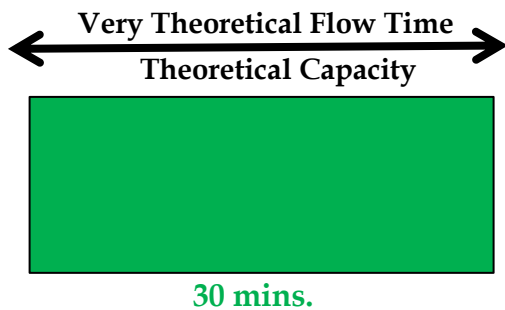
Flow Time Capacity does not care about buffer times

Theoretical Unit Load, Theoretical Activity Time

Theoretical Capacity is computed based on the Theoretical Unit Load (ThUL)

Theoretical Flow Time is NOT computed based on Theoretical Activity Time

Very Theoretical Flow Time is computed based on theoretical Activity Time ThUL
(1+CWF) = Unit Load (Tp),



Theoretical Unit Load, Theoretical Activity Time, Theoretical Capacity,

Very Theoretical Flow Time

An operating room (a resource unit) performs surgery every 30 min, $T_p = 30$ min. T_p includes all the distracts. We also refer to it as the Unit Load.

Effective capacity is $1/30$ per min or $60/30 = 2$ per hour.

On average, $1/3$ of the time is wasted (cleaning, restocking, changeover of nursing staff and fixing of malfunctioning equipment).

Capacity Waste Factor (CWF) = $1/3$.

Theoretical Unit load = $T_p \cdot (1 - CWF) = 30(1 - 1/3) = 20$ min.

$T_p = \text{Unit Load} = \text{ThUnit Load} / (1 - CWF) = 20 / (1 - 1/3) = 30$

Theoretical Capacity = $c / \text{ThUnit Load}$

Effective Capacity = Capacity = $c/\text{Unit Load}$.

Theoretical Capacity = 1/20 per minute or 3 per hour.

Effective Capacity = Theoretical Capacity (1-CWF)

Problem 5. A law firm processes (i) shopping centers and (ii) medical complexes contracts. The time requirements (unit loads) for preparing a standard contract of each type along with some other information is given below. In November 2012, the firm had 150 orders, 75 of each type. Assume 20 days per month and 8 hours per day. CWF at the three resources is 25%, 0%, and 50%, respectively.

	Unit Load Shopping (hrs /contract)	Unit Load Medical (hrs /contract)	No. Of Professionals
Paralegal	4	6	4
Tax lawyer	1	3	3
Senior partner	1	1	2

a) What is the effective capacity of the process (contracts / day)?

Paralegal: Theoretical Unit Load (50%Sh 50% Med): $0.5(4) + 0.5(6) = 5$ hrs.

Theoretical Capacity = 1/5 per hr.

Capacity Waste Factor (CWF) = 0.25

Unit Load = $T_p = 5/(1-0.25) = 20/3$ hrs.

Effective Capacity = Capacity = $1/(20/3) = 3/20$ per hr.

Tax Lawyer: Theoretical Unit Load $0.5(1) + 0.5(3) = \text{two}$ hrs.

CWF = 0

Theoretical Unit Load = $T_p = 2$ hrs.

Theoretical Capacity = 1/2 per hr.

Effective Capacity = Capacity = 1/2 per hr.

Senior Partner: Theoretical Unit Load $0.5(1) + 0.5(1) = \text{one}$ hrs.

Theoretical Capacity = $1/1 = 1$ per hr.

CWF = 0.5

Chapter 4. Throughput

Unit Load = $T_p = 1/(1-0.5) = 2$ hrs.

Effective Capacity = Capacity = $1/2$ per hr.

	ThUnit Load SH (hrs)	ThUnit Load MD (hrs)	Unit Load 50%SH 50%MD	Theoretical Capacity of a Resource Unit /hr	CWF	Unit Load 50%SH 50%MD	Capacity of a Resource Unit /hr	# Of Resource Units	Th Capacity of the Resource Pool/hr	Capacity of the Resource Pool/hr	Th Cap of R- Pool / day	Cap of the R- Pool / day
Paralegal	4	6	5	0.2	0.25	6.667	0.15	4	0.8	0.6	6.4	4.8
Tax lawyer	1	3	2	0.5	0	2	0.5	3	1.5	1.5	12	12
Senior partner	1	1	1	1	0.5	2	0.5	2	2	1	16	8



The PowerPoints slides of this problem are on the second part of at [Throughput Part 2a-Problem](#). We encourage you also to open the following problem, put it in slideshow mode, and solve it [Throughput Part 2b](#).

Chapter 6.

Deterministic Inventory Management Models

How can it be that mathematics, being after all a product of human thought which is independent of experience, is so admirably appropriate to the objects of reality? Albert Einstein, 1879-1955.

Importance of Inventory. Poor inventory management *hampers operations, diminishes customer satisfaction, and increases operating costs*. A typical firm probably has about **25% of its current assets** in inventories or about **90% of its working capital** (the difference between current asset and current liabilities). For example, 20% of the budgets of hospitals are spent on medical, surgical, and pharmaceutical supplies. For all hospitals in the U.S., it adds up to \$150 billion annually. The average inventory in the U.S. economy is about \$1.13 trillion, and that is for \$9.66 trillion of sales per year. In the virtue of the Little's Law, $9.66T=1.13$; each dollar spend in U.S. economy spends at least $1.13/9.66 = 0.115$ year or about 1.38 months in inventory. We used the term "at least" because cost of goods sold (CGS) is less than sales revenue. If we assume that, the CGS is $2/3$ of the sales revenue, or 6.44 trillion. Then each dollar spend in U.S. economy spends about $1.13/6.44 = 0.172$ year or more than two months in different forms inventory (raw material, work in progress, finished goods, goods in transport, etc.)

There are *two types* of inventory counting systems; **Periodical** and **Perpetual**.

In **periodical inventory system**, the available inventory is counted at the beginning of each period (end of the previous period). The required amount for the current period is computed, and the difference is ordered to satisfy the demand during the current period. You may imagine it as a one-bin system: there is one bin in which a specific raw material, part, component, or products is stored. We can look and see how full the bin is, and how much is empty. Each time, we only order enough to refill the single bin. The quantity that is ordered each time is variable, it depends on how much is needed to fill the bin, but the timing of order is fixed. The Re-order point (ROP) – when we reorder, is defined in terms of time. It is the beginning of the period. The *advantage* is that the timing is fixed. In additions, we can order for many items at the same time. Our ordering costs may go down because of ordering for several items at the same time. The *disadvantage* of

this system is that during the whole period, we have no information about inventory, because we only check it in the end of the current period, which is the beginning of the next period.

Perpetual Inventory Systems. In perpetual inventory system, when inventory reaches reorder point, we order a specific quantity. As opposed to the periodic inventory system, the quantity of order is fixed, where the timing of the order is variable. We usually order an economic order quantity, which we will discuss later, when inventory on hand reaches ROP. The ROP is defined in terms of quantity, or inventory on hand (or inventory position). You may imagine it as a two-bin system. Whenever the first bin gets empty, we order enough raw material, parts, components, or products to fill it. While waiting to get what we have ordered, we start using the inventory of the second bin. The benefit of this system is that it keeps track of inventory continuously.

Economic Order Quantity. Inventory models are perfect examples of applying mathematical models to real world problems. In this section, we discuss how to compute economic order quantity (EOQ). The EOQ computation is an example of trade-off in operations management. Trade-off between ordering cost and carrying cost.

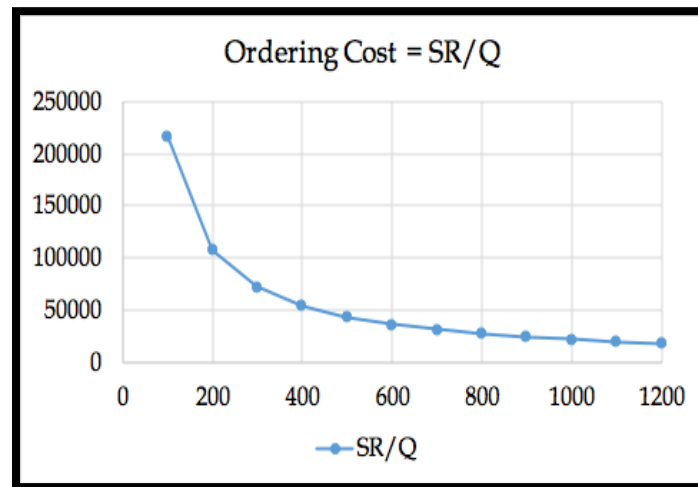
Problem 1: Q and EOQ. Consider a computer distribution firm with four retail stores in Northridge, Topanga, Sherman Oaks Galleria, and Glendale Americana. Each store at each mall sells an average of 40 laptops per day. Assume 30 working days per month. The cost of each laptop computer is \$800. Each time a store places an order to get a set of products; the ordering cost (cost of placing and order plus transportation cost, which is independent of the volume of order) is \$1500 *per order*. The carrying cost (including financial, physical, and obsolescence costs) of storing *one unit of product for one year* is 15% of the cost. That is $0.15(800) = \$120$ *per unit per year*. Assume a year is 360 working days, and a month is 30 working days.

The manager of Northridge-Store orders every 5 days, and manager of Topanga-store orders once a month. Which one do you follow?

Since the manager of Northridge-store orders every 5 days, she needs to place $360/5 = 72$ orders per year. Each time $40(5) = 200$ units.

The ordering cost is independent of the volume ordered, and it is \$1500 *per order*. That is $72(1500) = \$108000$. If the number of orders was not an integer, for example if we had ordered every 7 days and each time $7(40) = 280$ units, the number of orders would have come out to 51.43. In that case, the manager still places 52 orders. The cost of 51.43 orders, about \$77145, will count towards this year's costs, and the cost 0.57 order, about \$855, accounts for the next year's ordering costs. In our basic inventory model, one basic assumption is that everything remain the same from year to year.

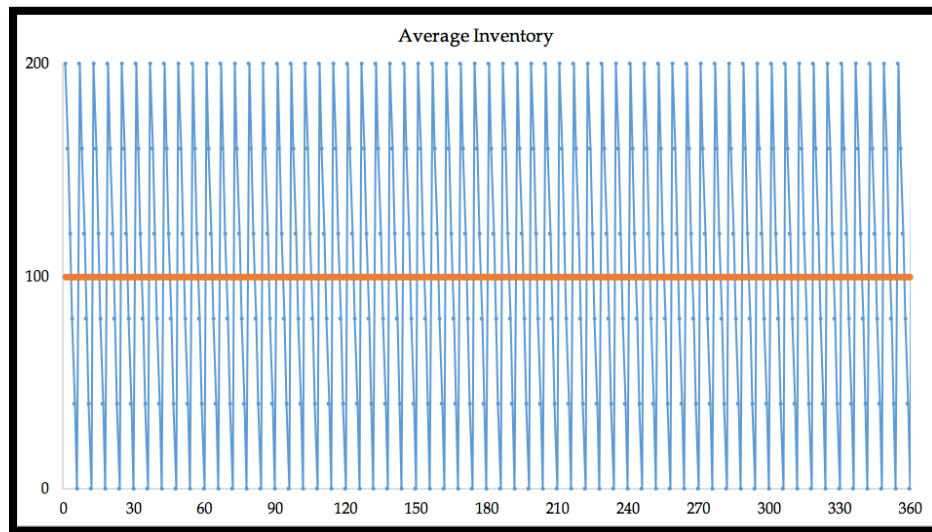
Manager of Topanga-store orders every month. She needs to place 12 orders per year. Each times $40(30) = 1200$; total of $12(1500) = \$18000$ ordering cost. As order size, Q , goes up, number of orders, R/Q goes down. The cost per order, S , is constant, and does not depend on the order quantity. The following curve shows the relationship between ordering quantity and ordering costs. As order size, Q goes up, the number of orders, and therefore the ordering cost, SR/Q , comes down.



You are the manager of the Glendale-store; do you follow Topanga-store or Northridge-store policy?

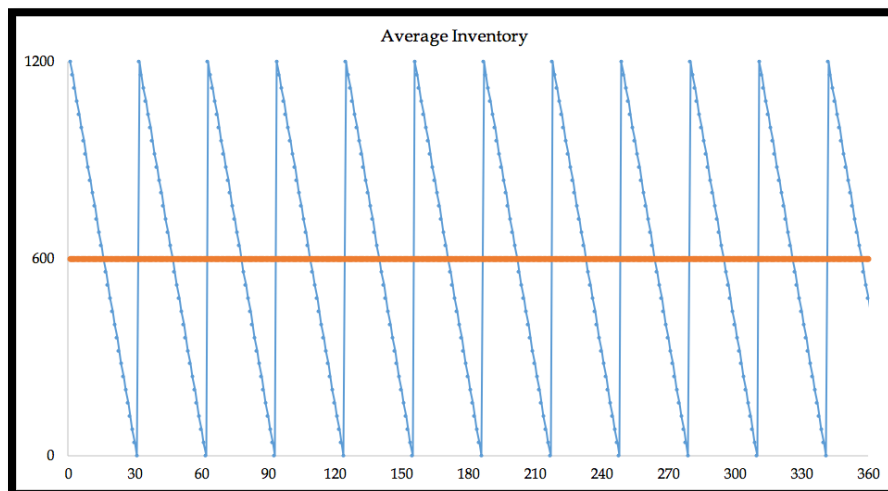
We do not know; we need carrying cost.

The manager of Northridge-Store orders 200 laptops per order. Therefore, there will be a maximum of 200 units, which gradually goes down at the rate of 40 units per day and reaches zero at the end of day 5 (start of day 6). Exactly at the same time, a new order of 200 units will arrive. The average inventory is $(200+160+120+80+40+0)/6 = 100$. Since the pattern is linear (decreases at a constant rate), we can just get the first and last number and average them $(200+0)/2 = 100$. The same pattern of changes in inventory level, as shown below, is repeated every month.

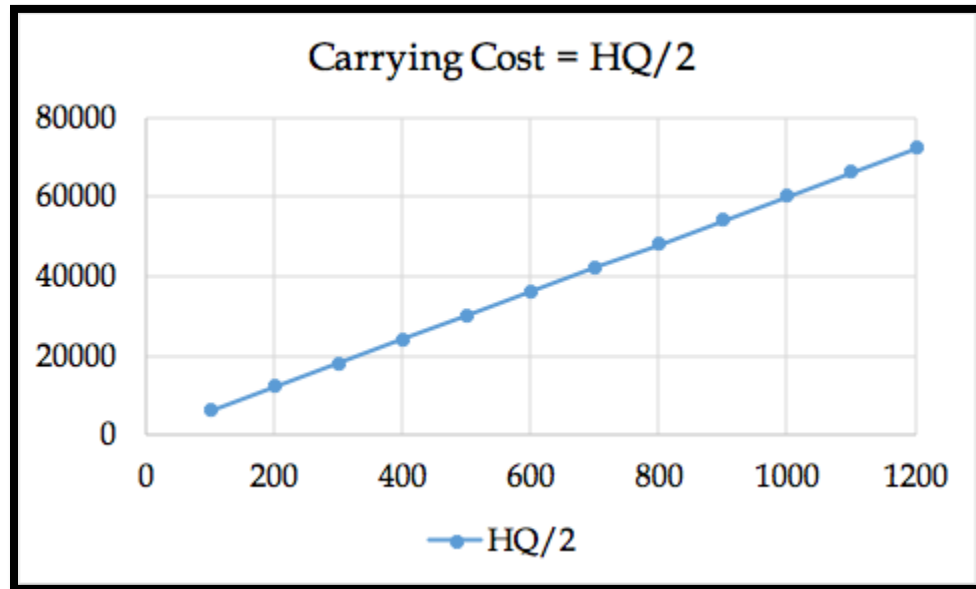


The average inventory in the first cycle (5 days) is $200/2 = 100$. Since the pattern is repeated, the average inventory in each of the following cycles is also 100. In general if each time we order Q units, and if the inventory at the end of the period is zero, then the average inventory is $(Q+0)/2 = Q/2$. We refer to $Q/2$ (half of the order size) as cycle inventory. If we keep a safety stock of I_s , the cycle inventory is still $Q/2$, while the average inventory is $Q/2 + I_s$. Why? We will show it later.) Cost of carrying one unit of inventory for one year is \$120. Since the average inventory 100 units, thus inventory carrying cost is $120(100) = \$12000$.

The manager of Topanga-store orders $40(30) = 1200$ units once a month. Therefore, there will be a maximum of 1200 units which gradually goes down at the rate of 40 units per day and finally at the end of day 30 (start of day 31) it reaches 0. The average inventory is then $1200/2 = 600$. The same pattern as shown below is repeated every month.



The average inventory during a single month is $1200/2 = 600$ units, the same pattern is repeated every month, thus, the average inventory per year is 600. Cost of carrying one unit of inventory for one year is \$120, which means the inventory carrying cost is $120(600) = 72000$. As order size goes up, maximum inventory and average inventory go up. Since carrying cost per unit per year is constant, as order quantity goes up, inventory carrying (or holding) costs go up.



Order Quantity	# of Orders	Ordering Cost	Average Inventory	Carrying Cost	Total Cost
Q	R/Q	SR/Q	I=Q/2	HQ/2	SR/Q+HQ/2
200	72.0	108000	100	12000	120000
1200	12.0	18000	600	72000	90000

Out of these two practices, we follow the Topange-store's policy because its total cost is smaller.

We can compute ordering costs (OC), carrying costs (CC) and total costs (TC) for alternative Q values.

Review of the parameters.

Demand per year = $D = R = 360(40) = 14400$ units.

Ordering cost per order = $S = \$1500$ per order.

Carrying cost (holding cost) = $\$120$ per unit per year

Order quantity = Q

of orders = R/Q

Ordering cost = SR/Q

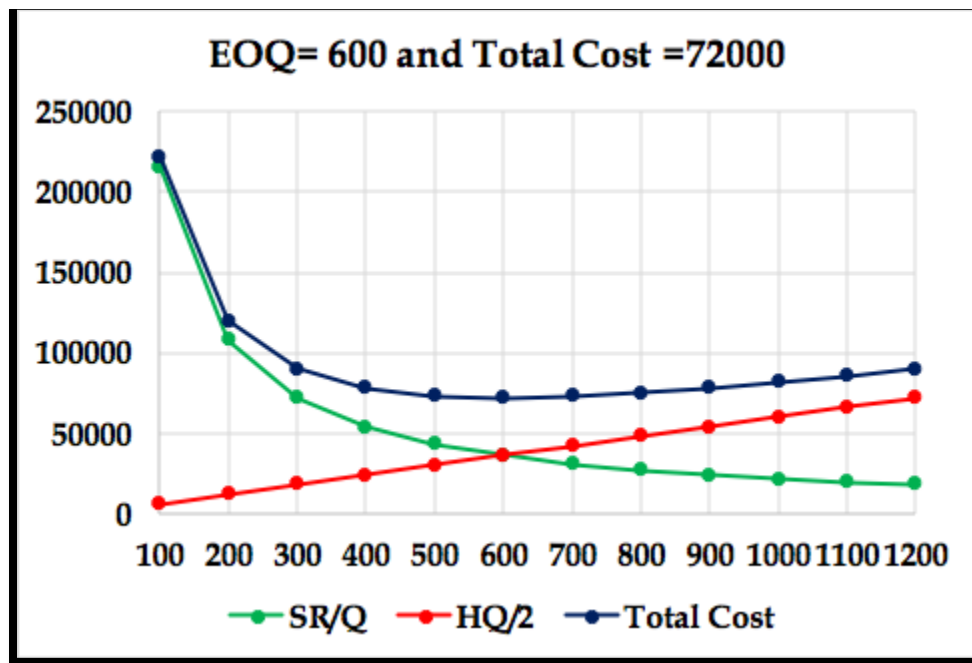
Average Inventory = $Q/2$

Carrying cost = $HQ/2$

Order Quantity	# of Orders	Ordering Cost	Average Inventory	Carrying Cost	Total Cost	Flow Time
Q	R/Q	SR/Q	$I=Q/2$	$HQ/2$	$SR/Q+HQ/2$	I/R
100	144.0	216000	50	6000	222000	1.25
200	72.0	108000	100	12000	120000	2.5
300	48.0	72000	150	18000	90000	3.75
400	36.0	54000	200	24000	78000	5
500	28.8	43200	250	30000	73200	6.25
600	24.0	36000	300	36000	72000	7.5
700	20.6	30857.1	350	42000	72857.14286	8.75
800	18.0	27000	400	48000	75000	10
900	16.0	24000	450	54000	78000	11.25
1000	14.4	21600	500	60000	81600	12.5
1100	13.1	19636.4	550	66000	85636.36364	13.75
1200	12.0	18000	600	72000	90000	15

It can be shown that Economic Order Quantity (EOQ) is at the point where ordering cost and carrying cost equate. That is

$$1500(14400/Q) = 120Q/2 \rightarrow Q^2 = 2*1500*14400/120 \rightarrow EOQ = 600.$$



EOQ can be computed independently. We chose to remember it through equity of costs, since it is easier and makes us independent of memorizing the EOQ formula. While memorizing things or just having the formula and plugging in the numbers may look easier, understanding the logic behind formulas, even just a small piece of it, adds more value to our knowledge. For us, deriving EOQ using equality of the two costs is enough. One may derivate is independently, by derivation the total cost term of $SR/Q + HQ/2$ with respect to Q and set the derivative equal to zero.

What we learned. Here we have some points to mention.

As Q goes up, SR/Q goes down.

As Q goes up, $HQ/2$ goes up.

In the above model, we considered inventory holding costs and ordering costs. There are two other inventory related costs.

Purchasing Costs are defined as costs for purchased items, and pure variable costs (materials and supplies, direct energy, but not human resource costs) for a production item. In our model, we assumed that purchasing costs in independent of order quantity. Over a period we need to order D (or R) units, if purchasing price is P (our pure variable cost is V), out total purchasing (production) cost is PR (or VR). It does not depend on Q (the quantity we order each time). We will later present inventory discount models where the purchase price (or variable costs) depends on the quantity ordered (or produced).

We assumed that the purchase price of the product is independent of the volume ordered. Note that no matter how many units we order each time, over one year we need to order $D=R=14400$ units.

In the formula, instead of demand (or throughput) per year, we can have demand (or throughput) per month or per day. In all cases, S remains, as it is, \$1200 in this example. However, we need to change H to $120/12 = \$10$ per month, or (if a year is 360 days) to $120/360 = 0.333$ per day. In the Littlefield game, it may provide more insight if instead of demand per year and carrying cost per year, we implement demand per month and carrying cost per month.

We do not teach OM as isolated islands. Recall the Little's Law. Throughput is equal to demand and we can show it by D or R . Average inventory is $Q/2$. In the virtue of the Little's law, $RT = Q/2$, therefore, the flow time is $Q/2R$ year. Since the optimal order quantity is 600 units, and demand (or throughput) is 40 per day, therefore, $40T = 600/2$, that is flow time $T = 7.5$ days.

EOQ is a mathematical formula for a portion of real world. In this model, we assume that we only have one single product. Demand is known, and is constant throughout the year. For example, we know that we need 5,000 units of product per year, and if a year is 50 weeks, then $1/50$ of this number is needed every week, that is 100 per week. If a week is 5 days, $1/5$ of whatever we need per week we need per day. Demand is known and it is constant. Every day, every minute, every hour, we have the same demand as any other minute, hour, or day.

Each order is received in a single delivery. When we order, we have a waiting period or a lead time. This could be one day, 2 days, and 3 days. It is known, and it is fixed. After lead-time, we receive the inventory that we have ordered. If lead-time is three days, as soon as our inventory reaches a level that we need for 3 days, then we order. Because demand is fixed and constant, at the second our inventory reaches 0, we get the product, and we replenish.

There are only two costs involved in this model: **ordering cost**, cost of ordering and receiving the order; and **holding** or **carrying costs**, costs to carry an item in inventory for one year. Unit cost of product does not play any role in this model because we do not get a quantity discount. It does not matter if we order one unit or one million units, the price is the same.

Shortage costs. These costs include lost profit and loss of goodwill. If we have stockouts, we may lose potential profits as well as customer loyalty. Models have been developed to include these costs in the EOQ models. We do not discuss these models. However, we will discuss shortage costs (underage costs) in our re-order point model.



The excel file of this problem can be accessed at [Order-Quantity-Props The name of the worksheet is 1.EOQ-Q.](#)



If you have the slightest difficulty on this subject, I encourage you to watch my recorded lecture at [Inventory Model: Basic Recorded](#) This lecture includes more examples



The PowerPoint slides of the lecture can be accessed at [Inventory Model: Basic](#). More assignment problems can also be accessed at [Assignment Inventory Basics Problems](#) For this part you may only solve Problems 1-2.

Problem 2: Centralization and Pooling. In this example, if all the parameters are the same for all four retail stores, $S = \$1200$, $R = 14400$ per year (1200 per month or 40 per day), $H = \$120$ per unit per year (\$10 per unit per month, 33.333 cents per day). As we computed above, under these parameters, it is at the benefit of each store to order 600 at a time.

Under optimal policy of $Q = 600$, how many orders a store places per year?

$R = 14400$, $Q = 600 \rightarrow$ # of Orders per year $= 14400/600 = 24$ times.

What is the length of each cycle? Cycle is the time interval between two orders.

24 orders per year, how often do we order? $1/24$ year.

$1/24$ year is $360(1/24) = 15$ days.

Alternatively, we could have said, demand per day is 40, we order 600, therefore, $600/40 = 15$ days.

What is the flow time?

$R = 14400$ per year, or 1200 per month, or 40 per day.

$I =$ Max Inventory divided by 2. $I = Q/2 = 600/2 = 300$. We also refer to $Q/2$ as cycle inventory. When we do not carry safety stock, cycle inventory is equal to average inventory. They differ in presence of safety stock.

Flow time

$T = 300/14400 = 3/144$ year.

Alternatively, $T = 300/1200 = 1/4$ month.

Alternatively, $T = 300/40 = 7.5$ days.

Alternatively, we could have said the time interval between two orders is 15 days. The first unit ordered will leave almost immediately; therefore, it will spend 0 days in the system. The last unit of an order will leave after 15 days (at the end of the cycle when the next order is about to arrive). Therefore, a unit, on average, spends $(0+15)/2 = 7.5$ days in the system.

Let us summarize. Each store orders 600 units at each order. They all together order $4(600) = 2400$ per order. Average inventory among all four stores is therefore, $2400/2 = 1200$. Throughput is 4×14400 per year, or 4×1200 per month, or 160 per day. Flow time, as computed earlier, can be recomputed as $1200/160 = 7.5$ days. Each store orders in 15 days intervals.

Centralization and Pooling. We refer to the preceding ordering system of the four retail stores as Decentralized system. Now let us consider a centralized system. In a centralized system, the demand or throughput is $4(14400) = 57600$ laptops per year. Inventory carrying costs remains the same. Ordering costs usually increases but not in the direct proportion of the demand, i.e., it does not become four times. For example, a truck, instead of having a round trip to a single store, needs to pass through multiple stations of all four stores. In this example for the purpose of simplicity, we assume that the ordering cost remains the same at \$1500 per order. Now let us compute EOQ for all warehouses.

Therefore, all stores order 1200 per order that is 300 per store.

Average inventory in all stores is 600 units compared to 1200 in the decentralized case. That is 50% reduction in inventory.

Average flow time is $600/(4 \times 40) = 3.75$ compared to 7.5 days in the decentralized case. That is 50% reduction in flow time.

Total cost of decentralized systems for a single warehouse is

$$1500(14400/600) + 120(600/2) = 3600 + 3600 = 7200$$

Total cost of the decentralized systems for all four warehouses is $4(7200) = 28800$ per year.

Total cost of the centralized systems for all four warehouses per year is

$$1500[(4 \times 14400)/1200] + 120(1200/2) = 7200 + 7200 = 14400. \text{ The total costs has decreased by } 50\%$$

	Decentralized Each Store	Decentralized All Stores	Centralized All Stores	Centralized Each Store	% Improvement Decent./Cent.
R/Day	40	160	160	40	
R/month	1200	4800	4800	1200	
R/Year	14400	57600	57600	14400	
S/Order	1500	1500	1500	1500	
H/Unit/Year	120	120	120	120	
EOQ	600	2400	1200	300	
Cycle Length	15	15	7.5	7.5	
# of Orders	24	24	48	48	
TC	72000	288000	144000	36000	50%
I	300	1200	600	150	50%
T days	7.5	7.5	3.75	3.75	50%

	A	B	C	D	E	F	G	H	I
1									
2		Decentralized Each Store	Decentralized All Stores	Centralized All Stores	Centralized Each Store	% Improvement Decent./Cent.			
3	R/Day	40	160	160	40				
4	R/month	1200	4800	4800	1200				
5	R/Year	14400	57600	57600	14400				
6	S/Order	1500	1500	1500	1500				
7	H/Unit/Year	120	120	120	120				
8	EOQ	600	2400	1200	300				
9	Cycle Length	15	15	7.5	7.5				
10	# of Orders	24	24	48	48				
11	TC	72000	288000	144000	36000	50%			
12	I	300	1200	600	150	50%			
13	T days	7.5	7.5	3.75	3.75	50%			
14									
15		=S\$17	=S\$17*S\$16	=S\$17*S\$16	=D3/S\$16				
16		=B3*S\$18	=C3*S\$18	=D3*S\$18	=E3*S\$18			Stores	4
17		=B4*S\$19	=C4*S\$19	=D4*S\$19	=E4*S\$19			R/Store/Day	40
18		=S\$20	=S\$20	=S\$21	=S\$21			Days/Month	30
19		=S\$22	=S\$22	=S\$23	=S\$23			Months/Year	12
20		=SQRT((2*B5*B6)/	=S\$16*B8	=SQRT((2*D5*D6)/	=D8/S\$16			S-Decent.	1500
21		=B8/B3	=C8/C3	=D8/D3	=E8/E3			S-Cent.	1500
22		=B5/B8	=C5/C8	=D5/D8	=E5/E8			H/Unit/Year-Decent.	120
23		=S\$20*(B5/B8)+S\$	=I16*B11	=S\$20*(D5/D8)+S	=D11/S\$16	=D11/C11		H/Unit/Year-Cent.	120
24		=B8/2	=C8/2	=D8/2	=E8/2	=D12/C12			
25		=(B12/B5)*360	=(C12/C5)*360	=(D12/D5)*360	=(E12/E5)*360	=D13/C13			
26									

Why we need to reduce inventory? There are times where inventory may be at the detriment of a company. For instance, a company with a large work in process and finished goods inventory may discover that the market is shifting from one product to another product. In this case, the company will have a large amount of work in process and finished goods inventory of a product when customers have already shifted to another product. Thus, the company will have *two choices*.

One choice is to **fire-sell** all inventories and finished goods what they have, which involves the selling of goods at extremely discounted prices. A drawback to fire selling is that it may turn into a significant loss because the inventory is potentially sold at 30 percent, 20 percent, or even 10 percent of their actual value. The second way is that they could sell their finished goods and at the same time, turn their work in process into sellable finished goods. Unfortunately, this means that there would be a lot of delay in entering the product into the market, and the company could lose a substantial portion of market share. Therefore, in both of these alternatives, they lead to loss. Thus, what is the message? We need to reduce our inventory as much as possible; we need to have minimal inventory.

Inventory adversely affects all competitive edges (Price/Quality/Variety/Time).

Inventory has cost (physical carrying costs, financial costs).

Inventory has risk of obsolescence (due to market changes, due to technology changes).

Inventory leads to poor quality (feedback loop is long).

Inventory hides problems (unreliable suppliers, machine breakdowns, long changeover times, too much scrap).

Inventory causes long flow time, not-uniform operations

We try to reduce inventory.

(a) By reducing EOQ

$$EOQ = \sqrt{\frac{2RS}{H}}$$

To reduce EOQ we may **↓R, ↓S, ↑H**

Two ways to reduce average inventory

- Reduce S

- Postponement, Delayed Differentiation

- **Centralize**

S does not increase in proportion of Q

EOQ increases as the square route of demand.

- **Commonality**, modularization and standardization is another type of Centralization

If centralization reduces inventory, why not everybody does it?

- Higher shipping cost
- Longer response time
- Less understanding of customer needs
- Less understanding of cultural, linguistics, and regulatory barriers

These disadvantages may reduce the demand

Please solve the rest of the problems at [Assignment Inventory Basics Problems](#)

Problem 3: Inventory Model with Price discount.

[Inventory Discount Model](#)

[Inventory Discount Model Recorded](#)

[Assignment Inventory Discount Model](#)

Two Mathematical Proofs- Total Cost and Flow Time of Q and EOQ.

Inventory Classification. There are three types of inventory. Input inventory is composed of raw materials, parts, components, and sub-assemblies that we buy from outside. In-process inventory are parts and products, sub-assemblies and components that are being processed; part, products, sub-assemblies, and components that are there to decouple operations. For example, assume that operation B follows operation A. In order to not completely have operation B dependent on operation a, we may put a little bit of inventory between those two. In addition, the third type of in-process inventory is when we realize that if we buy at large volumes, we get lower expense due to economies of scale. Then we have output inventory. We need to have some inventory, because when customers come, we cannot ask them to wait (at least most of the time). The best strategy is to have a low flow time in which I can deliver manufactured product and give it to the

customer. However, we are not there yet. Therefore, I should have inventory on shelf when customers come so that I can satisfy demand. Sometimes, the demand in one season is high and in another season is low. In this case, I should produce in low season and put it in inventory to satisfy demand in high season. Another type of output inventory is pipeline or transit inventories. This refers to products that are in a pipeline from manufacturing plants to warehouses, distribution centers, or retailers. The huge volumes of inventory on our highways are pipeline or in transit inventories.

Understocking is not good because we will not have enough products to satisfy demand. Customers will be dissatisfied, which will lead to a loss of sales. The customer may go to another vendor forever.

Overstocking is not good either because it has three types of costs: financial costs, physical costs, and obsolescence costs. Financial costs: Instead of having our money in a city or in a profitable business, we put it in inventory. Physical cost: Our inventory should be put in safe keeping somewhere. Thus, we either lease a warehouse or allocate a portion of our shop to a physical location of these products. Finally, we may have obsolescence cost: If we purchase for a large amount of inventory for a product that eventually gets low consumer demand, we may never be able to sell them. This renders the product obsolete.

A Classification Approach: ABC Analysis. In ABC analysis, the question is which type of inventory counting system is preferred? Is it periodical or perpetual? Perpetual is always better but more expensive because we need an automated system to continuously count our inventory. Therefore, we may conduct an ABC analysis.

Example: Here are our 12 parts (see Figure 3). Here is the annual demand of each part. Here is the unit cost of each part. If we multiply them, we will get the annual value of all items in our inventory system in our warehouse. If we sort them in non-increasing order, we will see that two items, which is 2 divided by 12 (15 to 20 percent of items), form 67 percent of the annual value. In addition, here, 7 items divided by 12 (about 55 percent of items), form 6 percent of the value. These are group C. These are group A and obviously, these are group B. For group A, we use perpetual. For group C, we may use periodical. For group B, we can use one of the two options.

Item Number	Annual Demand	Unit Cost	Annual \$ Value
1	2500	330	825000
2	1000	70	70000
3	1900	500	950000
4	1500	100	150000
5	3900	700	2730000
6	1000	915	915000
7	200	210	42000
8	1000	4000	4000000
9	8000	10	80000
10	9000	2	18000
11	500	200	100000
12	400	300	120000

Item Number	Annual Demand	Unit Cost	Annual \$ Value	% of Total	Classification
8	1000	4000	4000000	67%	A
5	3900	700	2730000		A
3	1900	500	950000	27%	B
6	1000	915	915000		B
1	2500	330	825000		B
4	1500	100	150000	6%	C
12	400	300	120000		C
11	500	200	100000		C
9	8000	10	80000		C
2	1000	70	70000		C
7	200	210	42000		C
10	9000	2	18000		C

Example 1a. Here we have 12 different video games along with their list of annual demand and unit cost. Using these values, we can find the annual dollar value, the percentage of their value, and then classify them as A, B, or C to see which type of inventory counting system is preferred.

Item Number	Annual Demand	Unit Cost	Annual \$ Value
1	200	7000	1400000
2	600	2000	1200000
3	1200	500	600000
4	300	6000	1800000
5	2000	700	1400000
6	4000	915	3660000
7	1000	1000	1000000
8	1500	4000	6000000
9	6000	10	60000
10	400	5000	2000000
11	3000	200	600000
12	8000	300	2400000

Total cost of any Q.

$$TC = \frac{SR}{Q} + \frac{HQ}{2}$$

Total Cost of EOQ? The same as above, but can also be simplified

+=

===

=

Flow time when we order of any Q?

Throughput = R, average inventory $I = Q/2$

$$RT = Q/2 \rightarrow T = Q/2R$$

Flow time when we order of EOQ?

Total Cost of EOQ? The same as above, but can also be simplified

$$I = EOQ/2$$

===

$$T = I/R$$

$$T = = =$$