Descriptive Statistics II
Graphs
Seven (Old) Quality Tools

Analyse Overview

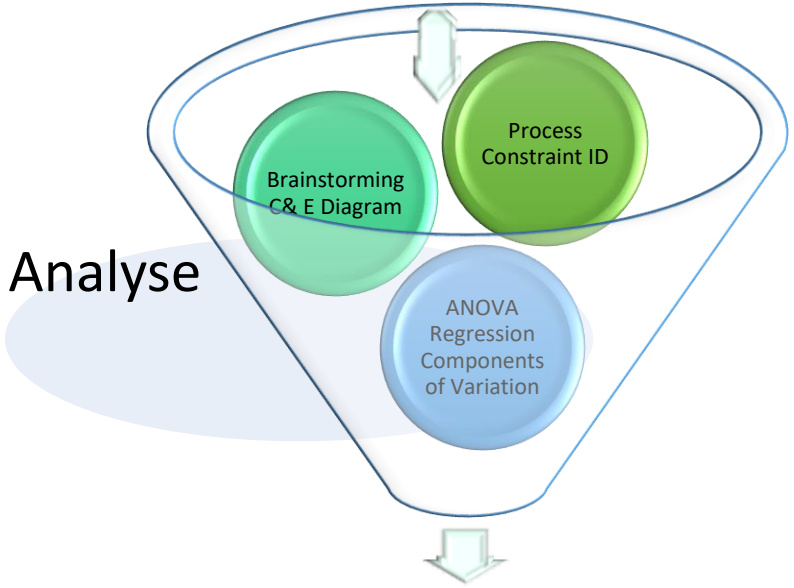
- In the “Analyze” phase, the goal is to identify some critical cause and effect relationships that explain most of the relationships between the key process input variables and the key process output variables.

- This starts by analyzing the process map to locate the constraint of bottleneck in the process.

- Efforts should be concentrated on identifying many cause and effect relationships and narrowing them to find a few key ones.

Analyze Overview

Define, Measure



Analyse

Improve, Control

- Use Brainstorming to identify potential root causes to gain insight into the key cause and effect relationships.
- Once they are prioritized use quantitative analysis to verify the relationships before proceeding to improve
- The DMAIC funnel summarizes this search during the Analyze phase

Analyze Roadmap

- **Conduct Value Analysis-**Identify value add, non value add and business non value added steps

- **Calculate Process Cycle Efficiency (PCE)-** Compare to world class benchmarks to help determine how much improvement is needed.

- **Analyze the Process Flow-** Identify bottleneck points and constraints in a process and assess their impact on process throughput and the ability to meet customer demands.

- **Analyze the data collected in measure** and generate theories to explain potential causes- Use brainstorming, FMEA, C&E diagrams etc,

Analyze Roadmap

- **Narrow the search-** Use ABC part stratification and various brainstorming, selection, and prioritization techniques to narrow the search for root causes and significant cause and effect relationships.

- **Use data to verify root causes-** Data is analyse using various statistical tools such as ANOVA and regression to verify significant relationships

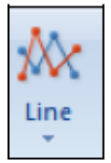
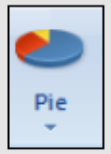
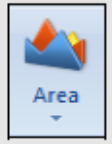
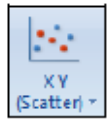
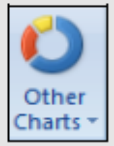
- **Gaining additional Insight-** Gain awareness of the impact of complexity and queuing on process performance

- **Analyze Gate Review-** A meeting is held with the Project Sponsor, Business Leadership and other interested parties to ensure key cause and effect relationships and root causes have been identified and continues alignment of the project with management expectations.

Analyze Tools

Excel Charts Types

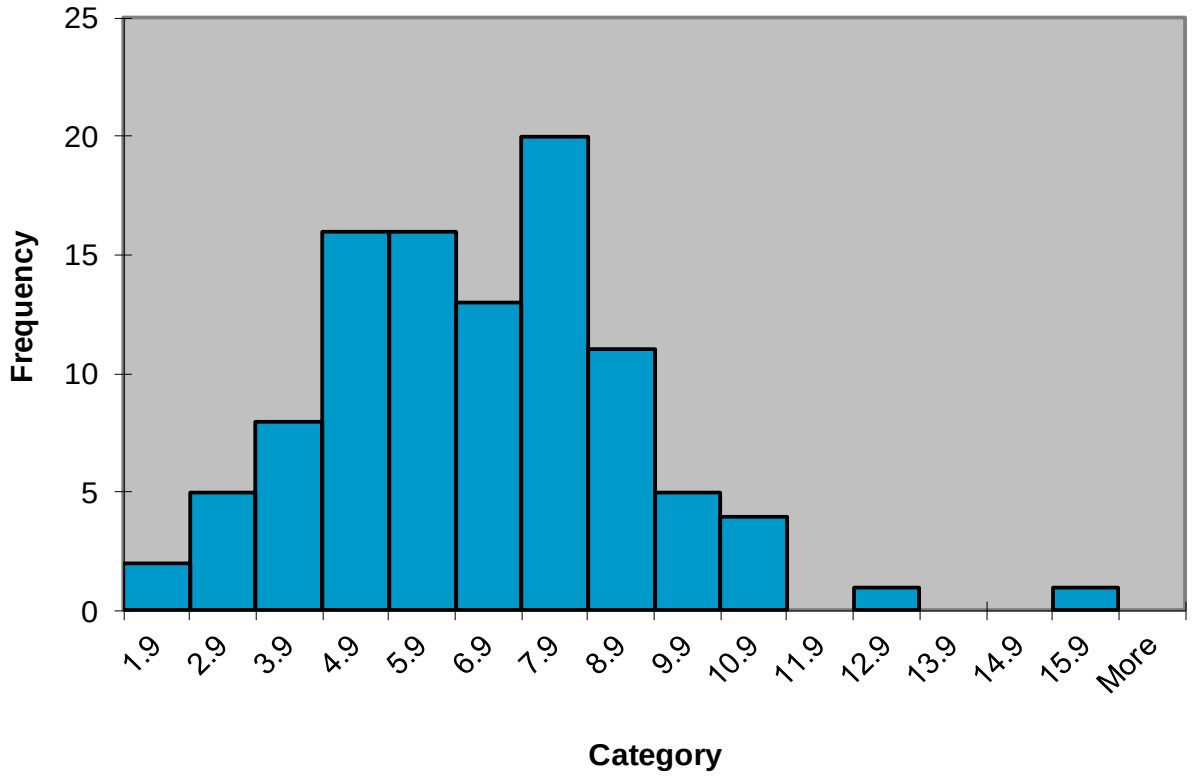
Table 5-1: Chart Types

	**Column**	Column charts are used when you want to compare different values vertically side-by-side. Each value is represented in the chart by a vertical bar. If there are several series, each series is represented by a different color.
	**Line**	Line charts are used to illustrate trends over time. Each value is plotted as a point on the chart and is connected to other values by a line. Multiple items are plotted using different lines.
	**Pie**	Pie charts are useful for showing values as a percentage of a whole. The values for each item are represented by different colors.
	**Bar**	Bar charts are just like column charts, except they display information in horizontal bars rather than in vertical columns.
	**Area**	Area charts are the same as line charts, except the area beneath the lines is filled with color.
	**XY (Scatter)**	Scatter charts are used to plot clusters of values using single points. Multiple items can be plotted by using different colored points or different point symbols.
	**Other Charts**	Select from Stock, Surface, Doughnut, Bubble, or Radar-type charts. You can also make a combination chart by selecting a different type of chart for only one of the data series.

Seven (Old) Quality Control Tools

1. **Histogram:** A graphic display of the number of times (frequency) a value occurs.
2. **Pareto diagram:** A bar chart that organizes the data from largest to smallest to direct attention on the important items (usually the biggest contributors).
3. **Cause and effect diagram:** A schematic tool that resembles a fishbone that lists causes and sub-causes as they relate to a concern, also known as Fishbone diagram or Ishikawa diagram.
4. **Process flow diagram:** A graphical illustration of the actual process.
5. **Check sheet:** A form used to collect, organise, and categorise data so it can be easily used for further analysis.
6. **Scatter diagram:** A graphical tool that plots one characteristic against another to understand the relationship between the two.
7. **SPC control chart:** A graph of time-ordered data that predicts how a process should behave.

Histogram



Histogram

- A **histogram** is used to display continuous data.
- Divide the range of the observations into non-overlapping intervals, usually of equal length.
- A rule of thumb for judging the number of intervals to use is by taking the square root of the number of observations.
- For example with 24 observations, 4.8990 or approximately 5 intervals (called classes) could be used. This is small so we will use 8.

Histogram Example I
On-Time Delivery Analysis

Variable **x** = Number of days Delivery is Late

(Each data point represents one shipment.)

Raw Data:

0	2	3	4	1	0	0	1
3	0	3	1	1	0	0	0
2	2	0	0	0	1	2	0
4	1	0	1	0	0	0	1
1	0	0	0	0	1	3	1

N = 40 shipments

Organising the Data

Step 1

Form a Data Array: ***Sort the data in numerical order***

Raw Data:

02	**3**	**4**	**1**	**0**	**0**	**1**
30	**3**	**1**	**1**	**0**	**0**	**0**
22	**0**	**0**	**0**	**1**	**2**	**0**
41	**0**	**1**	**0**	**0**	**0**	**1**
10	**0**	**0**	**0**	**1**	**3**	**1**

Data Array

Low	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1
	1	1	1	1	1	1	2	2
	2	2	3	3	3	3	4	4

High

Organising the Data
Step 2
Construct a Frequency Distribution

- **Frequency Distribution**
  - A mathematical function showing the number of instances in which a variable takes each of its possible values.

- **Ungrouped Frequency Distribution**
  - When the variable has only a *few* different values
  - Number of data values may be high or low

- **Grouped Data Frequency Distribution**
  - When the variable has *more than a few* different values
  - Number of data values is *high*

Grouped Frequency Distribution

- These are the numbers of newspapers sold at a local shop over the last 10 days:
  - 22, 20, 18, 23, 20, 25, 22, 20, 18, 20

- Let us count how many of each number there is:

Papers Sold	Frequency
18	2
19	0
20	4
21	0
22	2
23	1
24	0
25	1

- It is also possible to **group** the values. Here they are grouped in 5s:

Papers Sold	Frequency
15-19	2
20-24	7
25-29	1

Grouped Data Frequency Distribution

- Given these guidelines, the number of categories, k, for a histogram are as follows:

n	**k**	
1-7	**don't bother**	
8-15	**3 +/- 1**	These are "rule-of-
16-31	**4 +/- 1**	thumb". We should
32-63	**5 +/- 1**	not under- or over-
64-127	**6 +/- 2**	resolve the data.
128-255	**7 +/- 2**	
255+	**8 +/- 2**	

Histogram Example I
On-Time Delivery Example

Use ungrouped Frequency Distribution since the variable takes on only a few different values.

Data Array

Low	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
	0	0	0	1	1	1	1	1
	1	1	1	1	1	1	2	2
	2	2	3	3	3	3	4	4

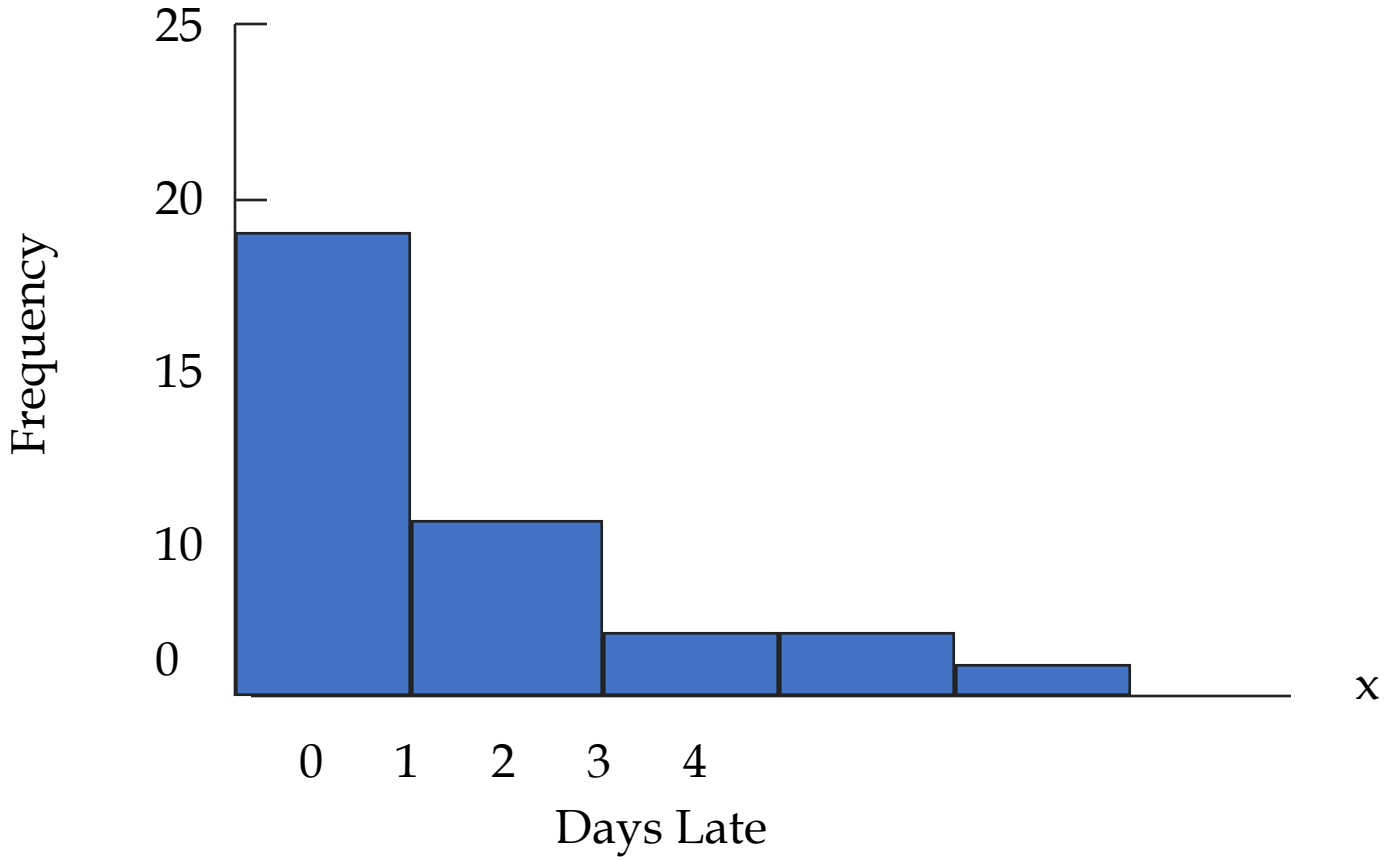
High

x	Frequency
0	19
1	11
2	4
3	4
4	2

Frequency Distribution

n= 40 values

Histogram Example I
On-Time Delivery Chart

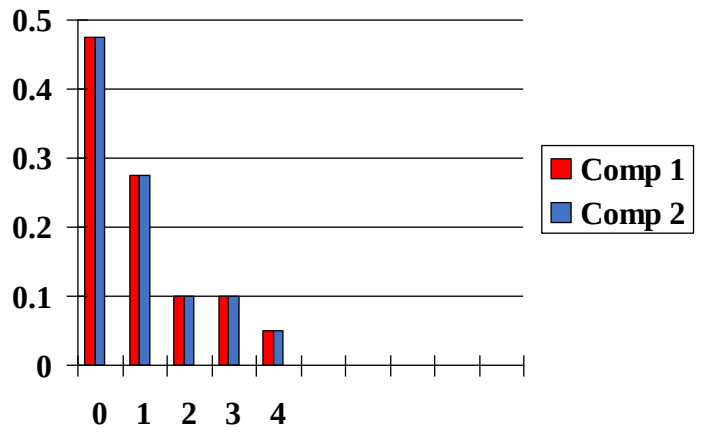


Comparing Two Suppliers: On-Time Delivery Distributions

	Supplier 1	Supplier 2
x	Frequency	Frequency
0	19	190
1	11	110
2	4	40
3	4	40
4	2	20
	40	400

Comparing Two Companies: On-Time Delivery Distributions

		1		2
X	F	RF	F	RF
0	19	.475	190	.475
1	11	.275	110	.275
2	4	.100	40	.275
3	4	.100	40	.275
4	2	.050	20	.050
	40		400	



Relative Frequency Distributions are useful for comparing two or more data sets which have different volumes of data.

Histogram Example II

The following data represents the number of poor quality apples that were rejected each day as part of a fruit and vegetable growers quality control program.

32	23	26	26	41	36	26	38
45	36	27	39	42	39	25	33
14	39	24	34	35	23	29	31

Histogram Example II

- To find the interval width, we take the range of the data and divide this by the number of intervals (8).

- In the example the minimum is 14 and the maximum is 45, so the interval size is:

$$\frac{45-14}{8} = 3.875 \approx 4$$

- So the first interval would be from 14 to less than 18 with midpoint 16, the second from 18 to less than 22 with midpoint 20, etc

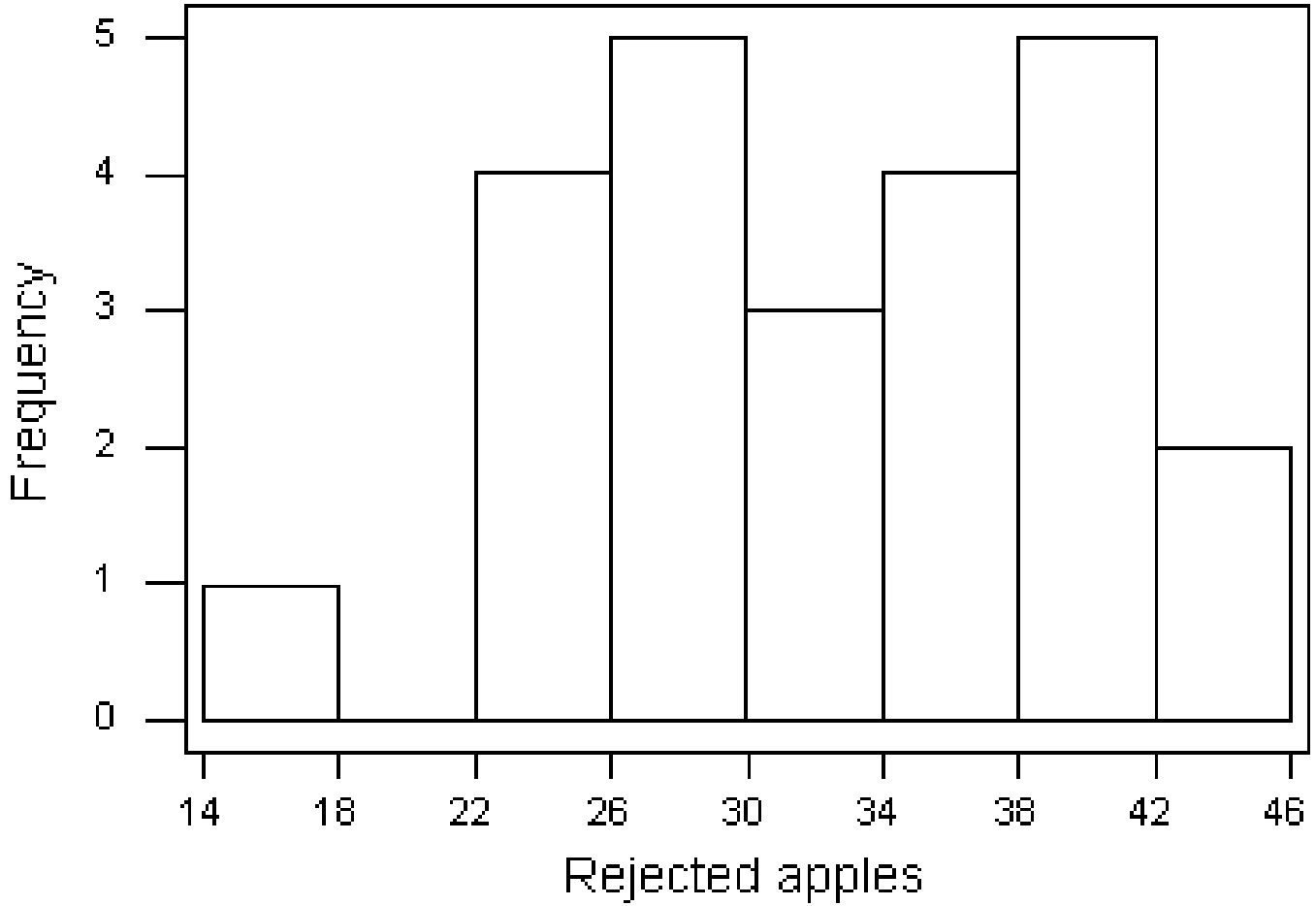
Histogram Example II

- The intervals do not include the right hand value so that each observation can fall in only one category.
- We denote this [14,18) to indicate the interval is from 14 to less than 18.
- Next we produce a table with the interval (and sometimes the midpoint), the frequency, the relative frequency and the cumulative relative frequency.

Histogram Example II

Interval	***Freq***	***Rel freq***	***Cum rel freq***
[14,18)	1	1÷24=0.0417	0.0417
[18,22)	0	0÷24=0	0.0417+0=0.0417
[22,26)	4	4÷24=0.1667	0.0417+0.1667=0.2084
[26,30)	5	5÷24=0.2083	0.2084+0.2083=0.4167
[30,34)	3	0.1250	0.5417
[34,38)	4	0.1667	0.7084
[38,42)	5	0.2083	0.9167
[42,46)	2	0.0833	1.0
Total	*24*	*1.0*	

Histogram Example II Chart



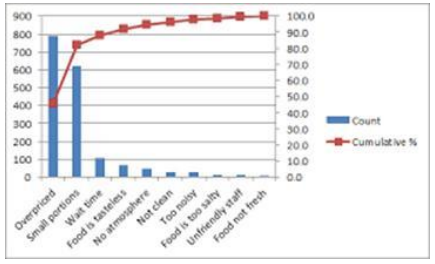
Pareto Analysis

- A Pareto Chart is a graphical tool to detect and prioritize multiple problems in a process.

- It is “a series of bars whose heights reflect the frequency or impact of problems. The bars are arranged in descending order of height from left to right. This means the categories represented by the tall bars on the left are relatively more significant than those on the right”.

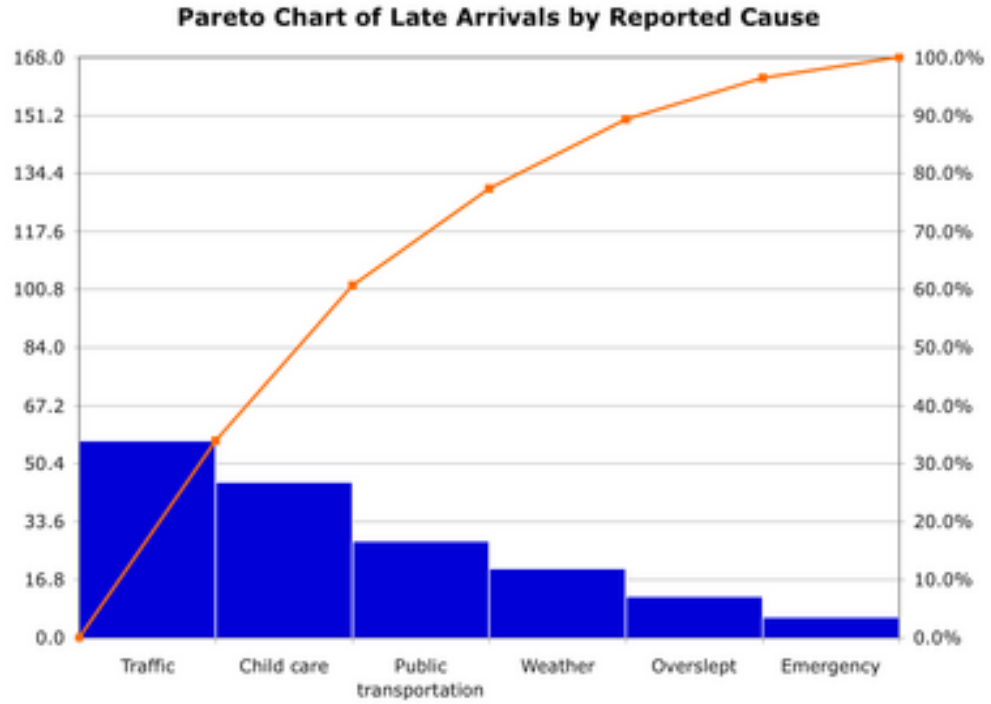
- They help you identify which of your problems are most significant so you can focus improvement efforts on areas where the largest gains can be made.

- We can use the Pareto Chart to focus on the area where we can have the greatest financial impact in the least amount of time, or with the fewest resources.



Pareto Analysis

Pareto Analysis is a systematic method of distinguishing between the ‘trivial many’ and the ‘vital few’ and identifying those ‘vital few’ that have the greatest influence.



Pareto Principle

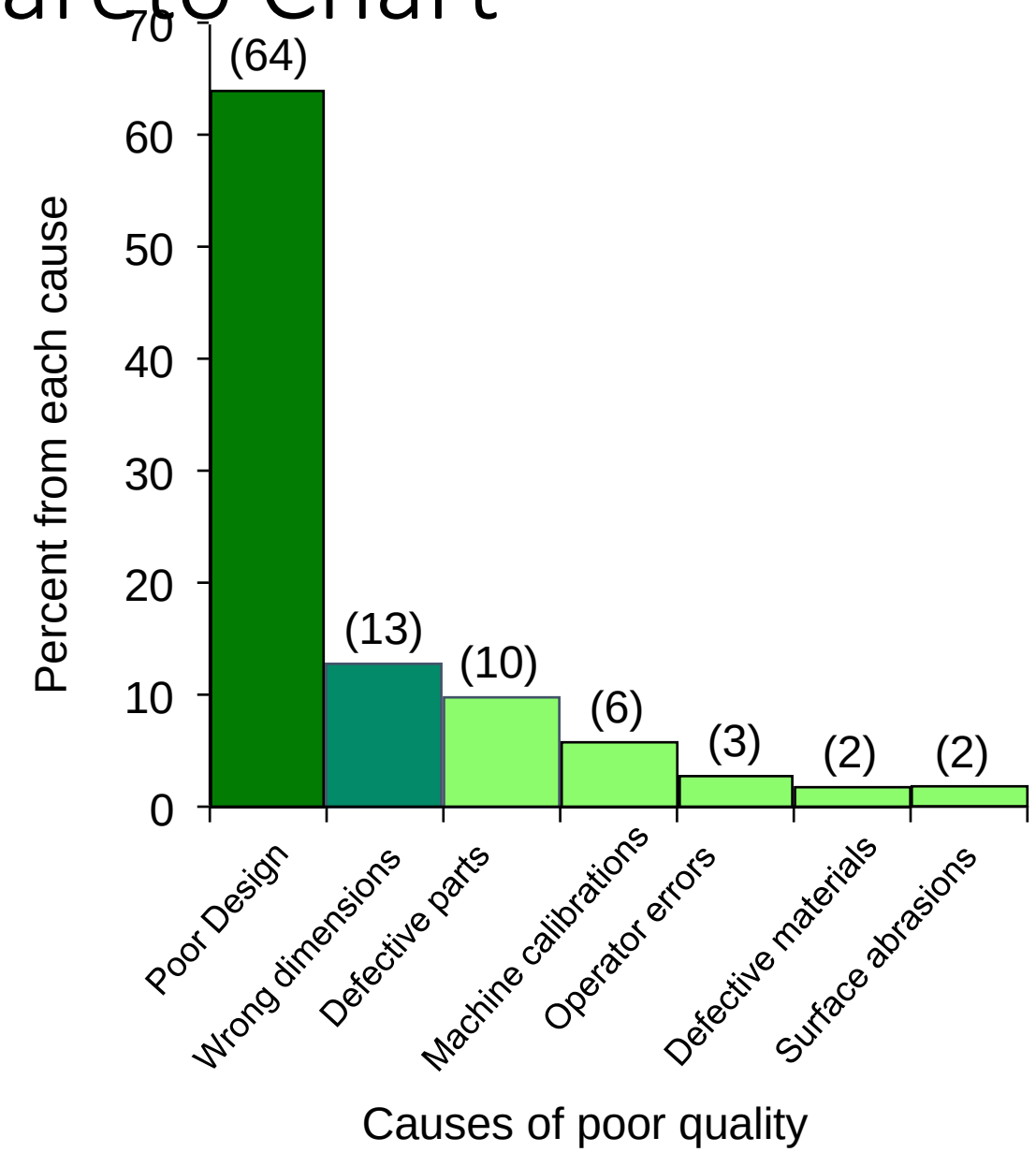
- Vilfredo Pareto (1848-1923) Italian economist
  - 20% of the population has 80% of the wealth

- Juran used the term “vital few, trivial many”. He noted that 20% of the quality problems caused 80% of the monetary loss.

Overview

- The causes of rejected parts, defects, failures etc. are classified according to their importance and the order of their importance is studied.
- A Histogram and Cumulative % line graph is then developed.
- Those causes which have the greatest influence are then identified and tackled.
- In this way, it separates the ‘trivial many’ from the ‘vital few’.
- Often called the ‘80/20 Rule’.

Pareto Chart



Developing Pareto Charts 1

- Identify the problem and a time period required.
- Define the type of data to be analysed (e.g. defects cause, defect type, or defect locations etc.)
- Collect representative data and categorise.
- Count and arrange the data in descending order
- The most frequent cause is presented on the left hand side of the graph.
- Calculate cumulative percentages.
- Group categories that collectively contain $\approx\leq$ 6% into an “other” category, plotted at extreme RHS.
- Plot the defect types (or locations) the x-axis.

Developing Pareto Charts 2

- Plot the frequency of occurrence on one y-axis.
- Plot the cumulative percentage on the other y-axis.
- A second Pareto diagram is usually presented with the monetary loss associated with each cause plotted on the y-axis in place of frequency.
- Cumulative percentages are also calculated.
- The ‘Vital few’ causes should now be easily distinguishable from the ‘trivial many’.
- These ‘vital few’ should now be tackled and corrected.

Pareto Chart Example I

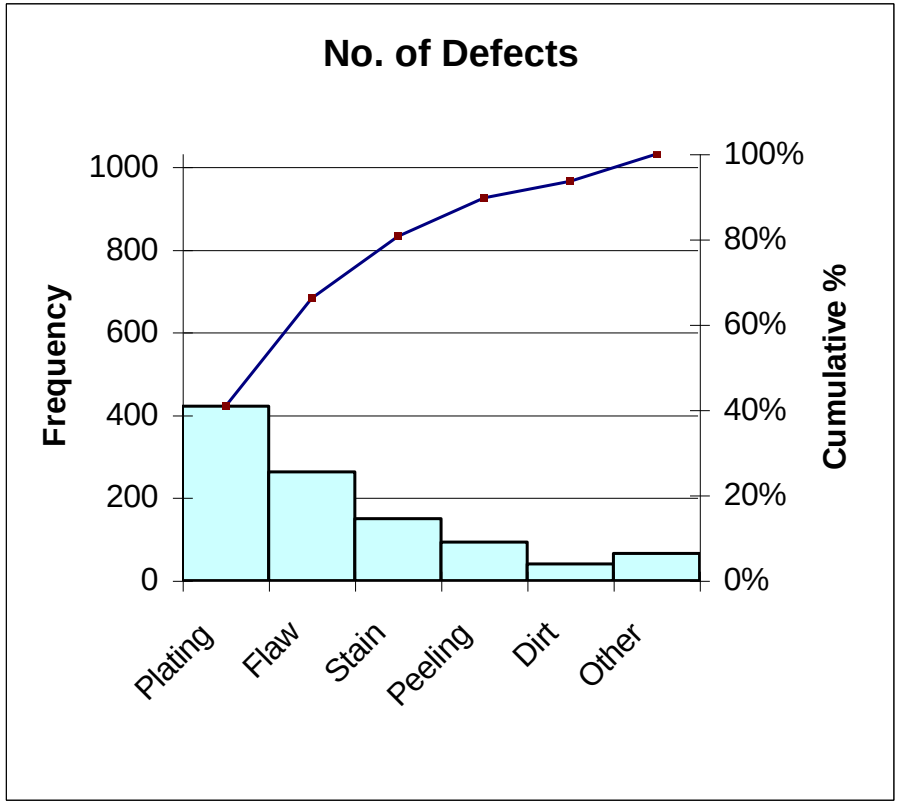
- An analysis of Surface Finish Defects reveals the following:
-

Defect type	Frequency	Cost
Flaw	262	€17,000
Stain	150	€14,000
Plating	421	€10,000
Dirt	40	€5,000
Peeling	93	€1,000
Other	65	€4,000

- Draw a Pareto Chart for the Frequency and Cost.

Pareto Chart Example I - Calculations for Frequency

Defect type	Frequency	% Cumulative	% of Total
Plating	421	40.83%	40.8%
Flaw	262	66.25%	25.4%
Stain	150	80.80%	14.5%
Peeling	93	89.82%	9.0%
Dirt	40	93.70%	3.9%
Other	65	100.00%	6.3%
	1031		

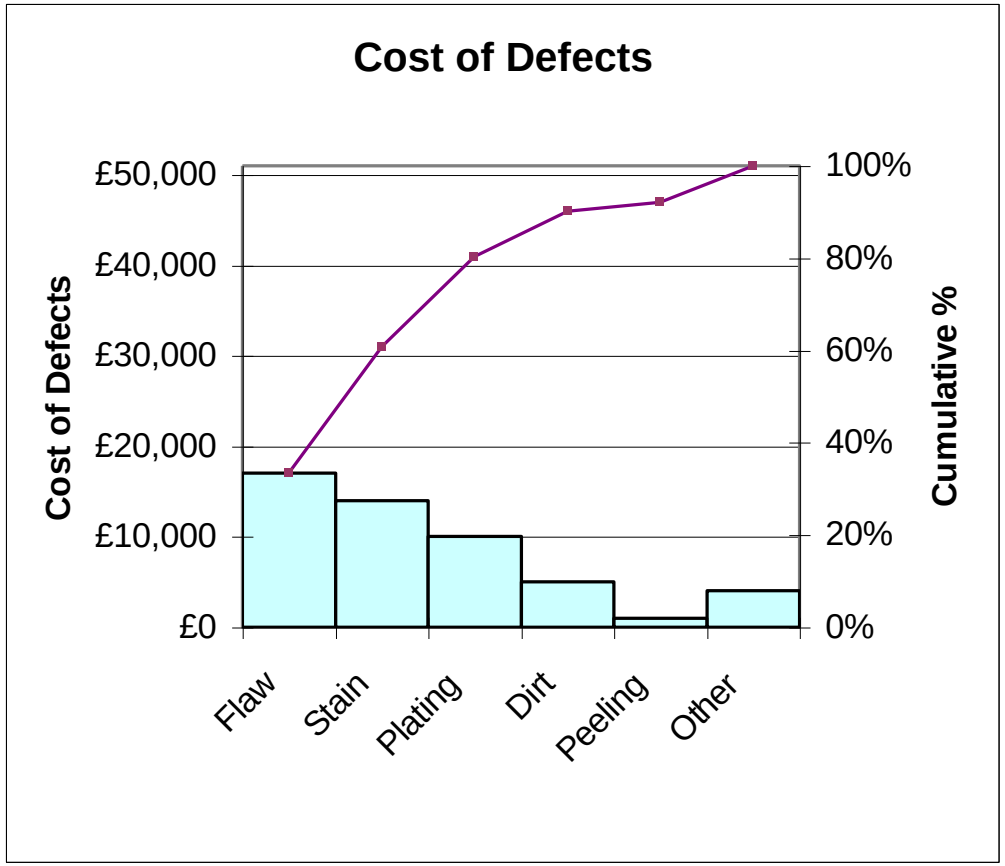


We should consider focusing our improvement efforts on **Plating**

- The Pareto Chart bars are divided into causes of defective product.
- The vertical axis shows the count of each error type.
- The blue line is the cumulative percentage.
- Plating is the most frequently occurring problem representing 41% of the total errors.
- Often is it used to re-pareto data using the cost per defect rather than the number of defects

Defect type	Cost	% Cumulative	% of Cost
Flaw	£17,000	33.33%	33.33%
Stain	£14,000	60.78%	27.45%
Plating	£10,000	80.39%	19.61%
Dirt	£5,000	90.20%	9.80%
Peeling	£1,000	92.16%	1.96%
Other	£4,000	100.00%	7.84%
	£51,000		

Pareto Chart Example I - Calculations for Cost



Pareto Problem Solving Method

- Pareto analysis is used to rank order the reasons for problems so that corrective action can be taken on the major causes of the problem.
- Pareto diagrams for a problem often lend themselves to further investigation and Pareto analysis. E.g.
  - Investigation of the most vital defect type may require further Pareto analysis to determine the root cause of the defect.
- Once the causes for the tallest bar have been resolved, proceed to the next tallest.
- Continue this process until the root causes have been eliminated, or reduced to a satisfactory level.

Why use Pareto Analysis?

- To prioritise actions needed to solve complex problems.
- To sort out the vital few from the trivial many.
- To separate important from unimportant causes contributing to a problem.
- To measure improvement after changes have been made.

Strengths and Weaknesses

Strengths:

- Progress can be measured.
- Efforts and resources are focused.
- The chart is easy to use and understand.
- It is an effective communication tool.

Weaknesses:

- The “vital few” could be misleading if only the number of occurrences are analysed without regard to costs.

Pareto Charts

1. Order categories from highest to lowest frequency.
2. Calculate cumulative frequencies.
3. Calculate relative frequencies.
4. Calculate relative cumulative frequencies.
5. Construct axes with frequency on the left-hand side and relative cumulative frequency on the right-hand.
6. Graph the histogram of frequencies and the line for relative cumulative frequencies.

Pareto - Example II

(Reasons for not learning material in course)

Book and Minitab difficult to understand	11
Failure to keep up with readings and homework	9
Poor teaching assistants	5
Computer software for statistical analyses	5
Uncomfortable lecture room	3
Difficult nature of subject	3
No interest in the course	2
Too few examples	2
Instructor generally not prepared	2
Instructor not accessible outside of class	0

Pareto - Example II

(Reasons for not learning course material)

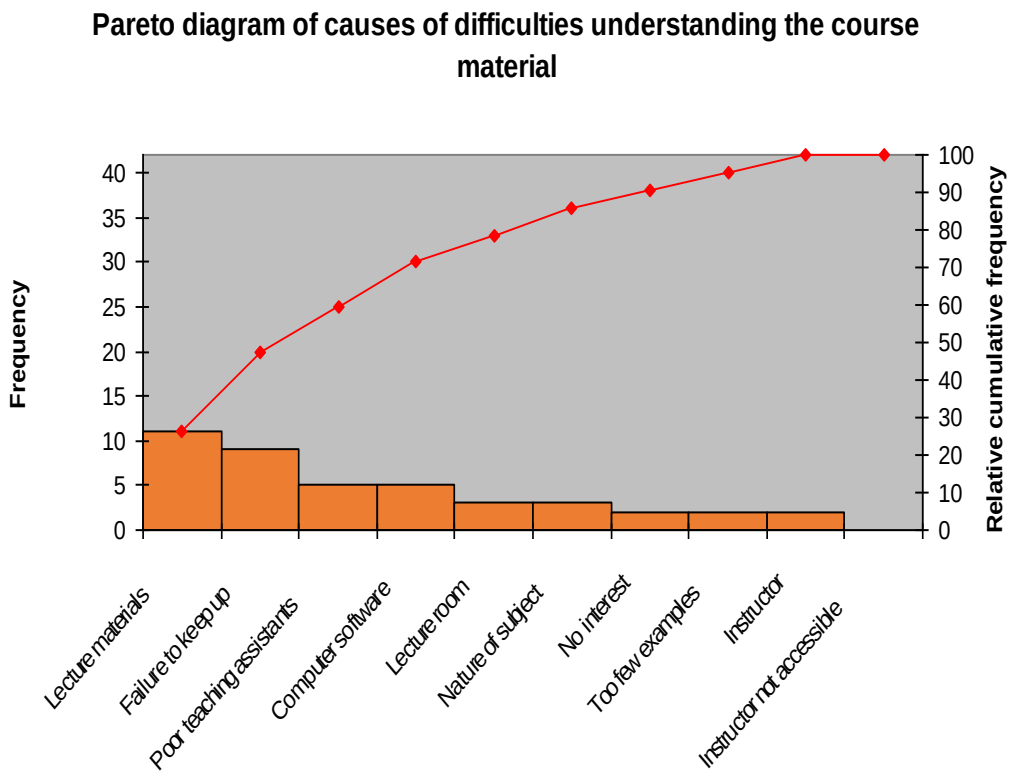
	f	*cf*
Book and Minitab difficult to understand	11	11
Failure to keep up with readings and homework	9	20
Poor teaching assistants	5	25
Computer software for statistical analyses	5	30
Uncomfortable lecture room	3	33
Difficult nature of subject	3	36
No interest in the course	2	38
Too few examples	2	40
Instructor generally not prepared	2	42
Instructor not accessible outside of class	0	42

Pareto - Example II

(Reasons for not learning course material)

	f	*cf*	*rf*	*rcf*
Book and Minitab difficult …	11	11	11÷42×100=26.2	26.2
Failure to keep up with …	9	20	9 ÷ 42 ×100=21.4	47.6
Poor teaching assistants	5	25	11.9	59.5
Computer software for …	5	30	11.9	71.4
Uncomfortable lecture room	3	33	7.1	78.5
Difficult nature of subject	3	36	7.1	85.6
No interest in the course	2	38	4.8	90.4
Too few examples	2	40	4.8	95.2
Instructor generally not …	2	42	4.8	100.0
Instructor not accessible …	0	42	0.0	100.0

Pareto Example II



Pareto Summary

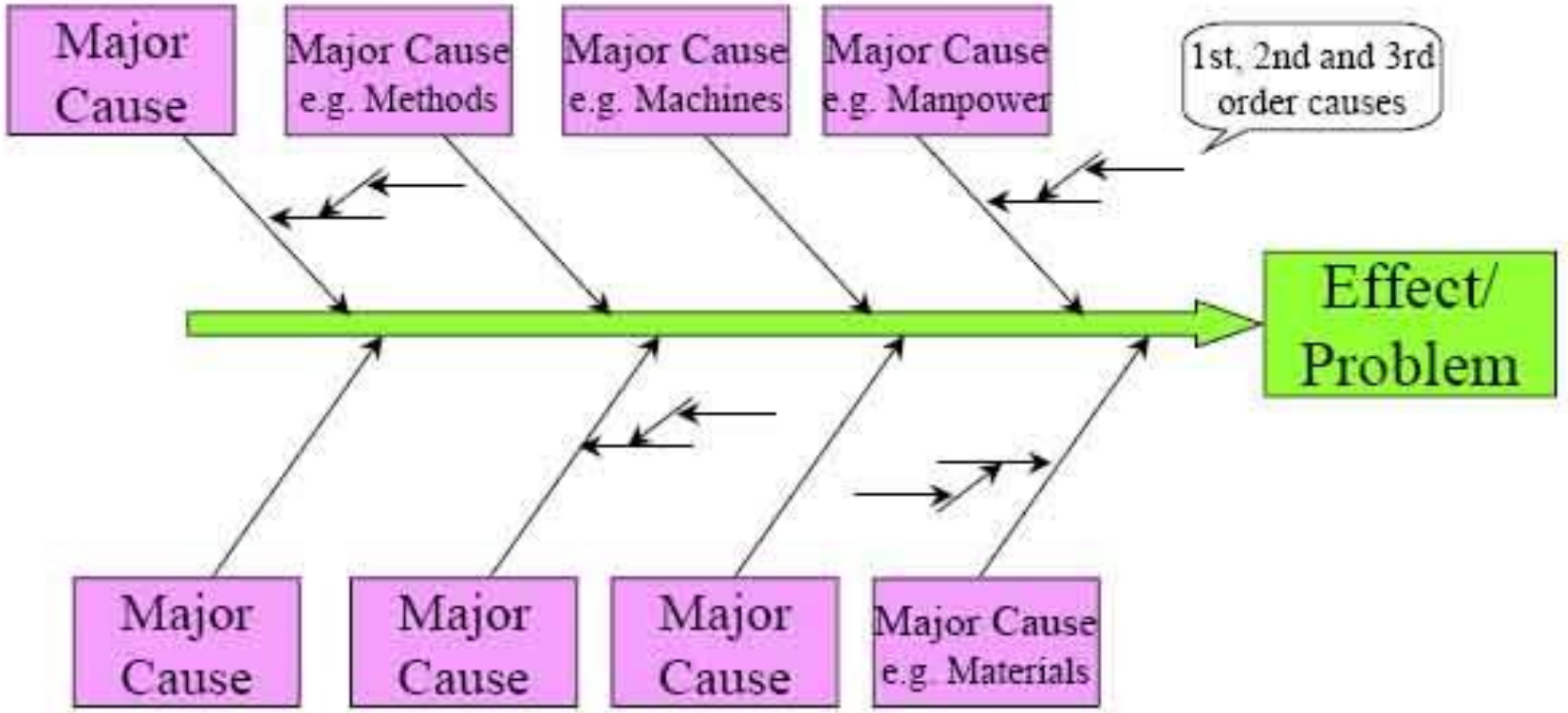
- A Pareto diagram will highlight frequent problems.
- It separates the “vital few” from the “trivial many”.
- The 80/20 rule says that 80% of quality problems can be attributed to 20% of the categories.
- Pareto diagrams are a vital tool for quality improvement – they identify the most common problems so these can be eliminated.

Cause & Effect Diagrams

- Show the relationships between a problem and its possible causes.
- Developed by Kaoru Ishikawa (1953)

- Also known as …
  - Fishbone diagrams
  - Ishikawa diagrams

Cause & Effect Diagrams

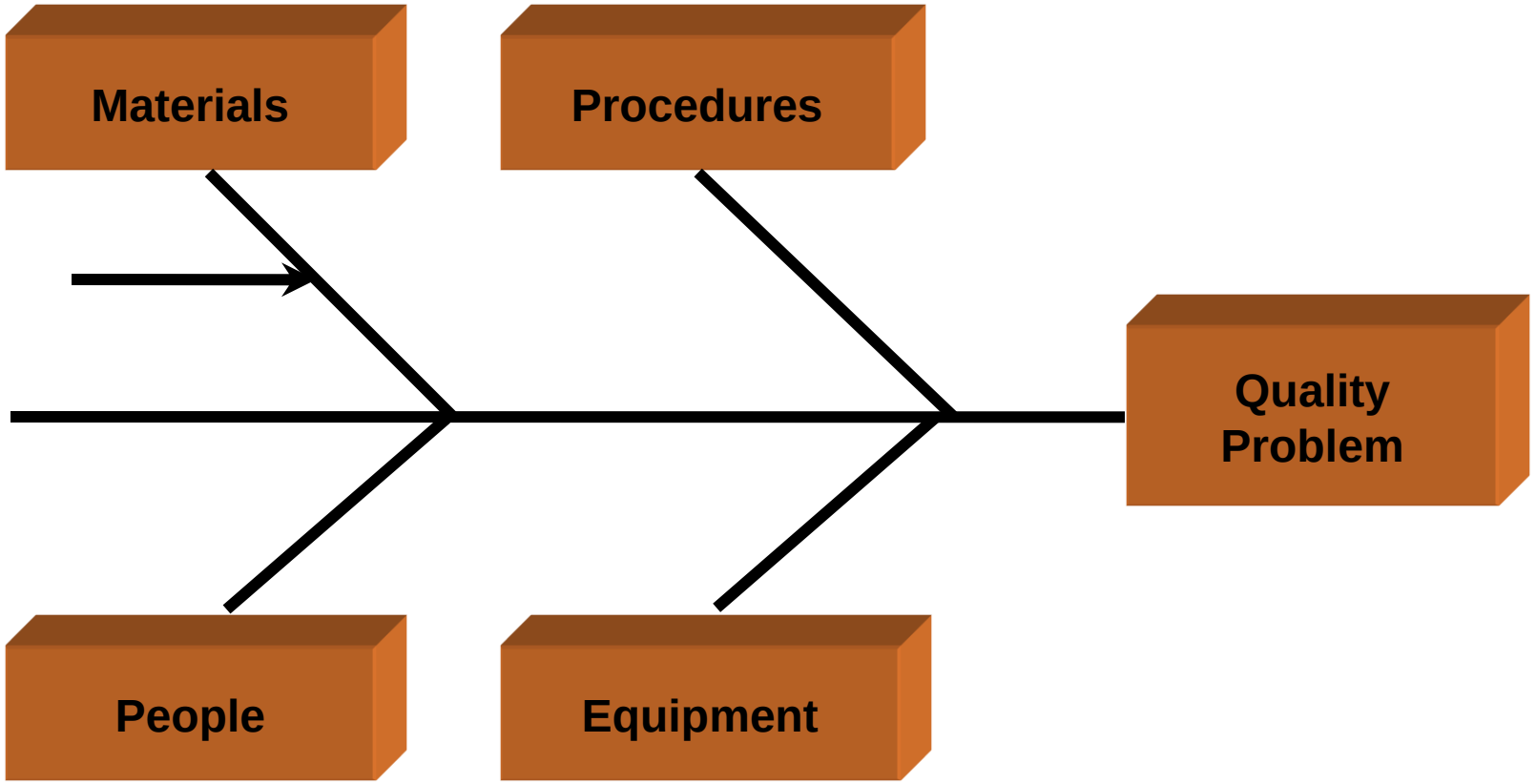


Cause and effect diagrams

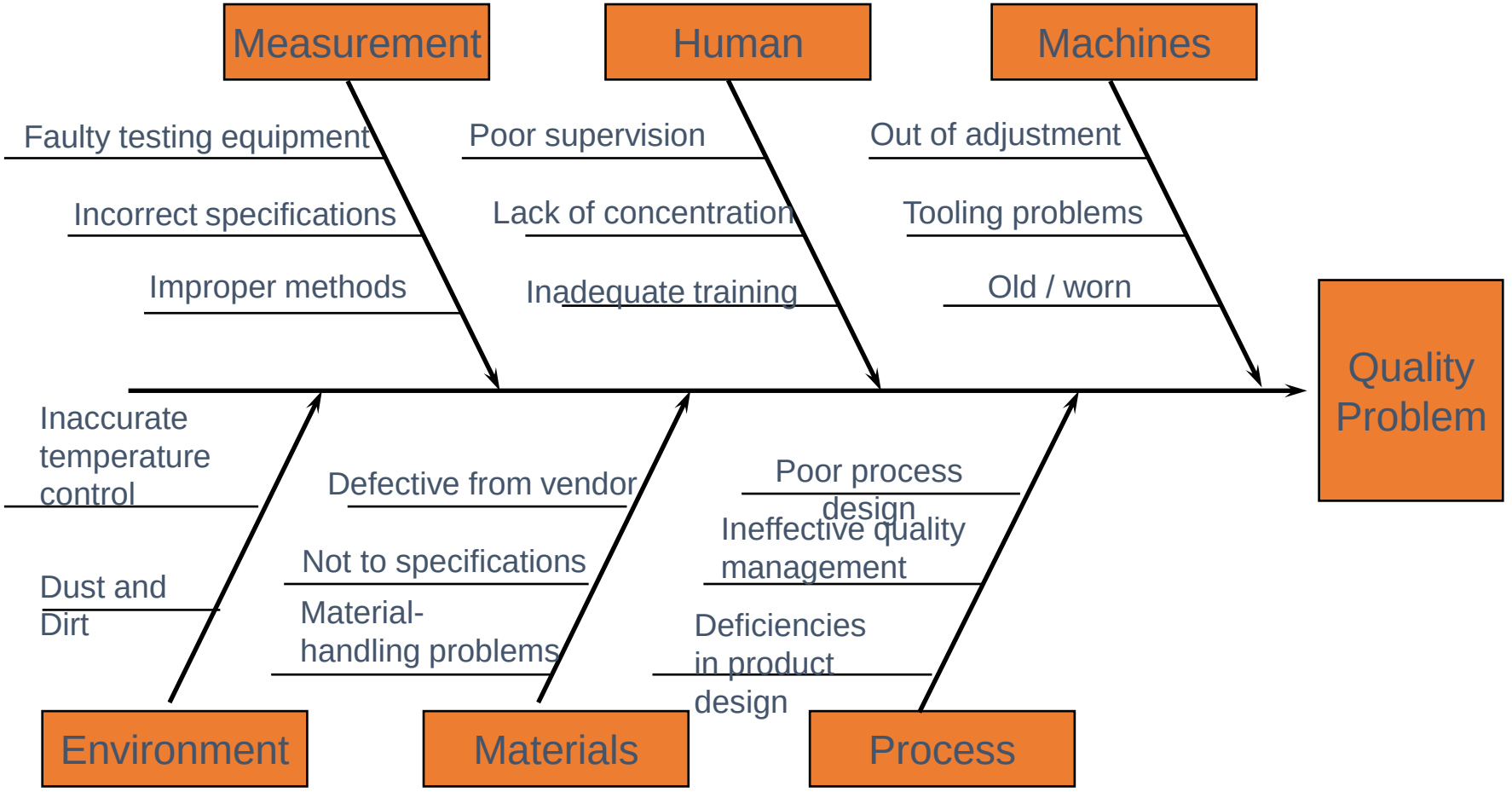
To construct the skeleton, remember:

- For manufacturing - the 6 M's
  - ✓man, method, measurement, machine, material, Mother Nature
- For service applications
  - ✓equipment, policies, procedures, people

Cause and Effect “Skeleton”



Fishbone Diagram
Product Quality Problems



Cause and effect diagrams

- We must always clearly define the effect or problem ensuring that it is a single effect, not a group of related effects.
- We should generate as many ideas as possible about the possible cause of an event (using a brainstorming session).
- After an idea is added, find other ideas by asking “what caused this?” but don’t evaluate ideas yet.

Cause and effect diagrams

Top-down generation of causes – start with the effect and work towards the cause by asking why (often referred to as why-why analysis). Problem – only causes identified in first step are explored further.

Bottom-up generation of causes – ideas are generated without any restrictions and by whatever means available (brainstorming or nominal group technique are used).

Process phases generation of causes – effects are often associated with processes, so this method examines each phase of a process one at a time.

Cause and effect diagrams

- Once a list of possible causes has been reached, the diagram can be constructed.
- It is not unusual to expect 100+ ideas from the brainstorming session – more ideas may be thought of and added as the diagram is constructed.
- The first step in constructing the diagram is to write the effect in a box on the right-hand side of the page.

Cause and effect diagrams

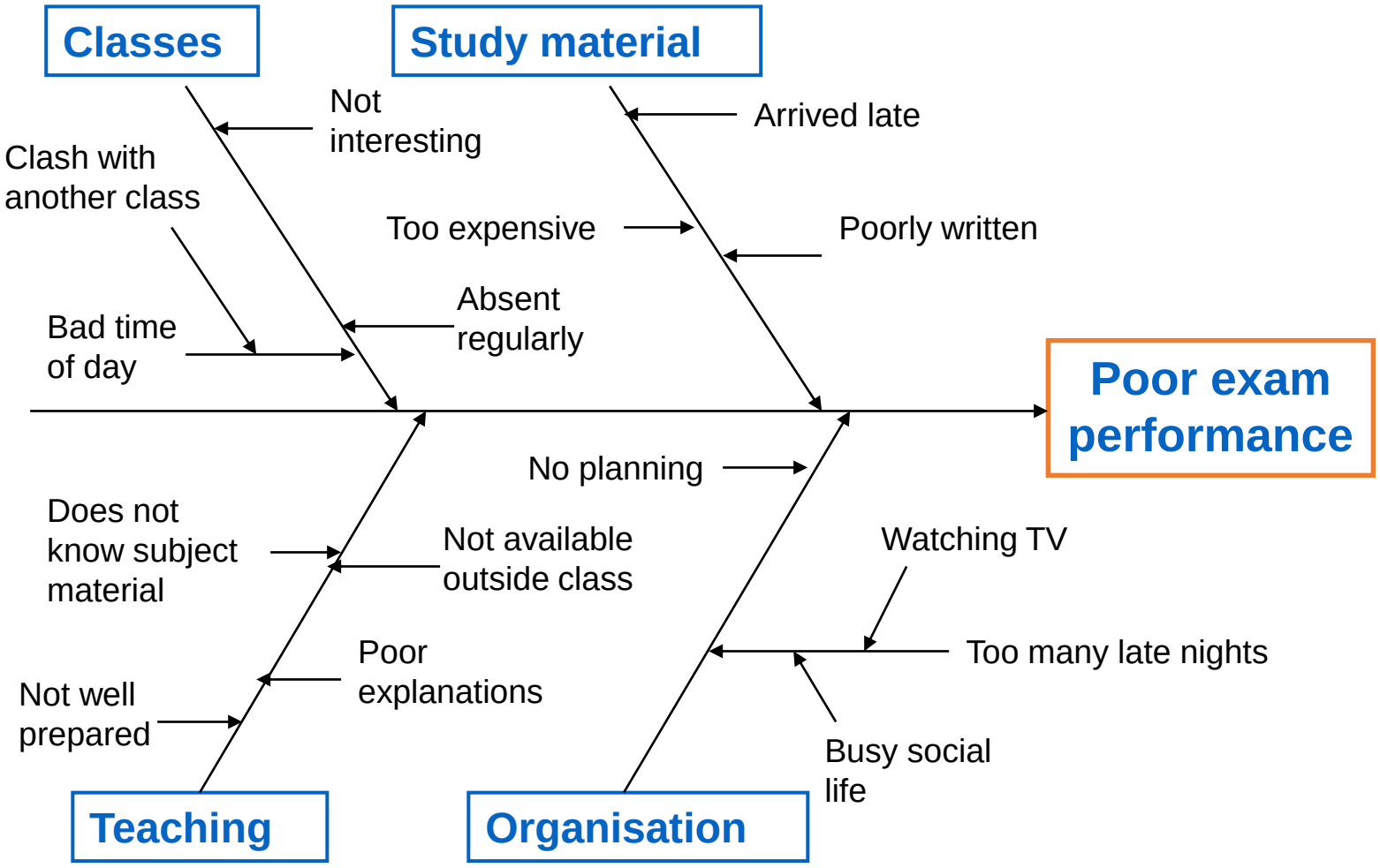
- Next, the possible causes should be grouped into categories, for example “work method”, “workers”, “equipment” and “measurement” are common headings for an industrial process.
- Possible causes may be placed under more than one heading if necessary.
- Now within each heading, identify different levels of causality.

Cause and effect diagrams

- The headings are usually called the first level causes, factors that contribute directly to these are second level causes, then third level causes etc.

- The (simplified) fishbone diagram on the following page illustrates this layering of causes.

Simplified Cause and Effect Diagram Example



Cause and effect diagrams

- Once a cause and effect diagram has been produced, we should now use it to determine the most likely cause(s) of the effect.

- Sometimes this may mean choosing one factor as most significant and leaving others for examination later in another revolution of the PDCA cycle.

- We search for the most likely cause by eliminating all possible causes except for one.

Cause and effect diagrams

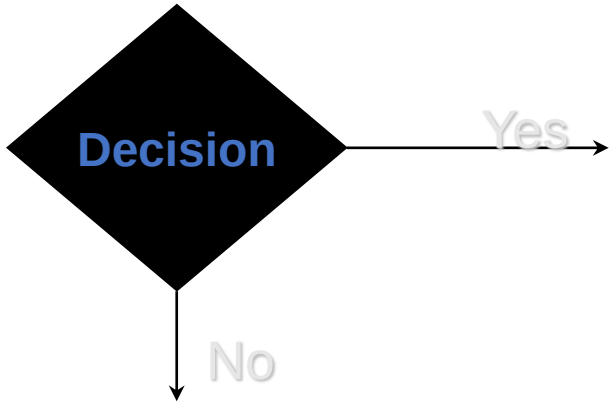
- Advantages
  - making the diagram is educational in itself
  - diagram demonstrates knowledge of problem solving team
  - diagram results in active searches for causes
  - diagram is a guide for data collection

Process Flowcharts

- A flowchart is a pictorial display of a sequence of events that make up a process.
- It is useful for describing the order of events and any interactions in the process.
- It can be used in the design of a new process or the documentation of an existing process.
- It can also be helpful in identifying problem areas in a process.

Symbols Used to Create A Process Map

Symbol	Name	Function
	Start/end	An oval represents a start or end point.
	Arrows	A line is a connector that shows relationships between the representative shapes.
	Input/Output	A parallelogram represents input or ouptut.
	Process	A rectangle represents a process.
	Decision	A diamond indicates a decision.

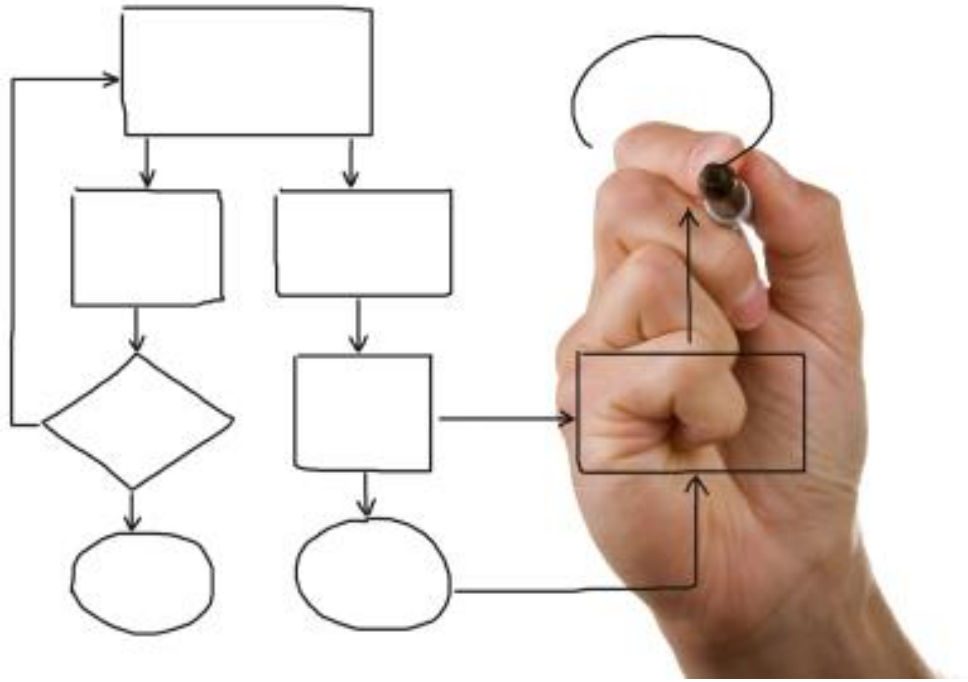


Diamonds represent a place where a decision must be made in the process (often has a yes/no response).

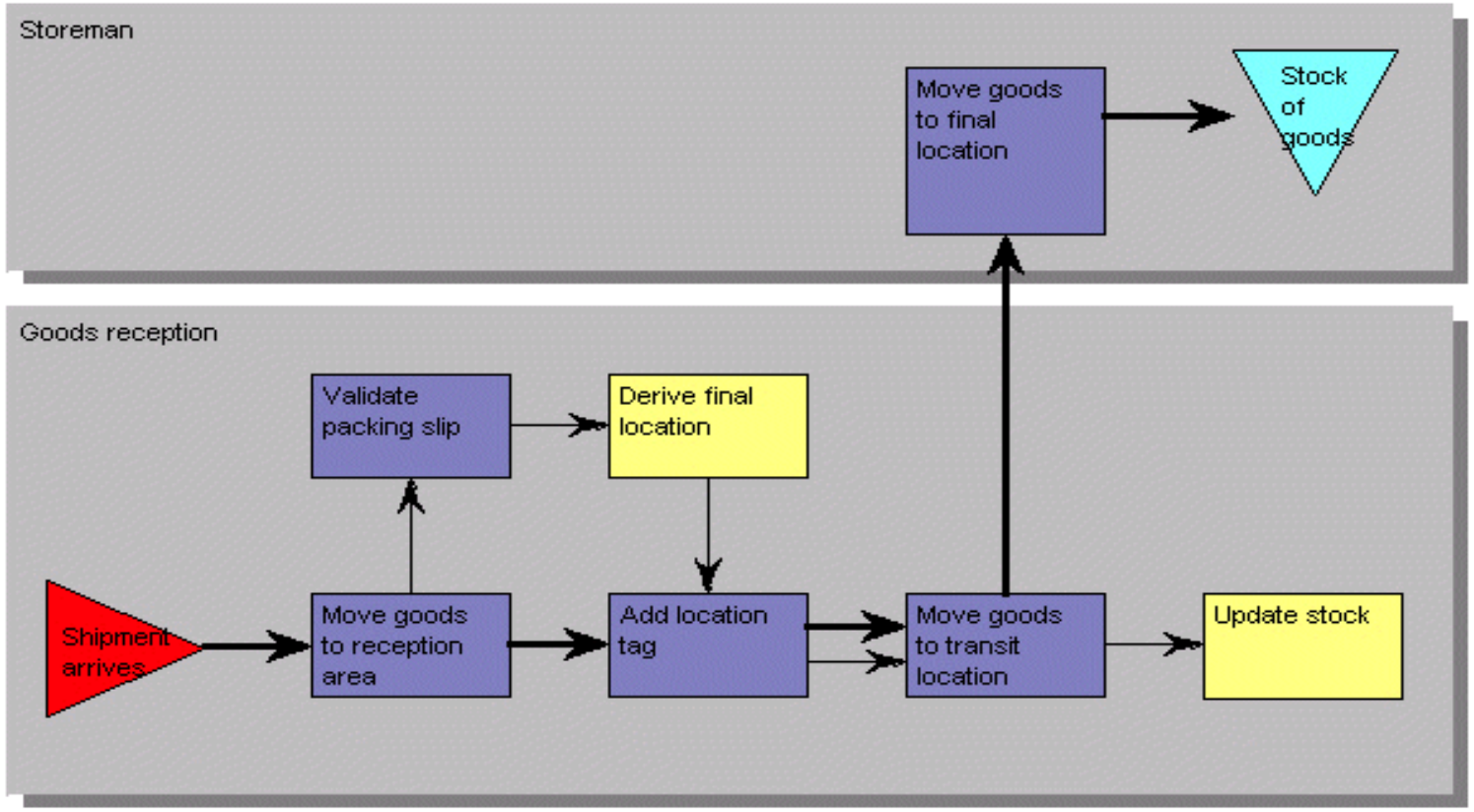
Flowchart

" Draw a flowchart for whatever you do. Until you do, you do not know what you are doing,
you just have a job.”

-- Dr. W. Edwards Deming.

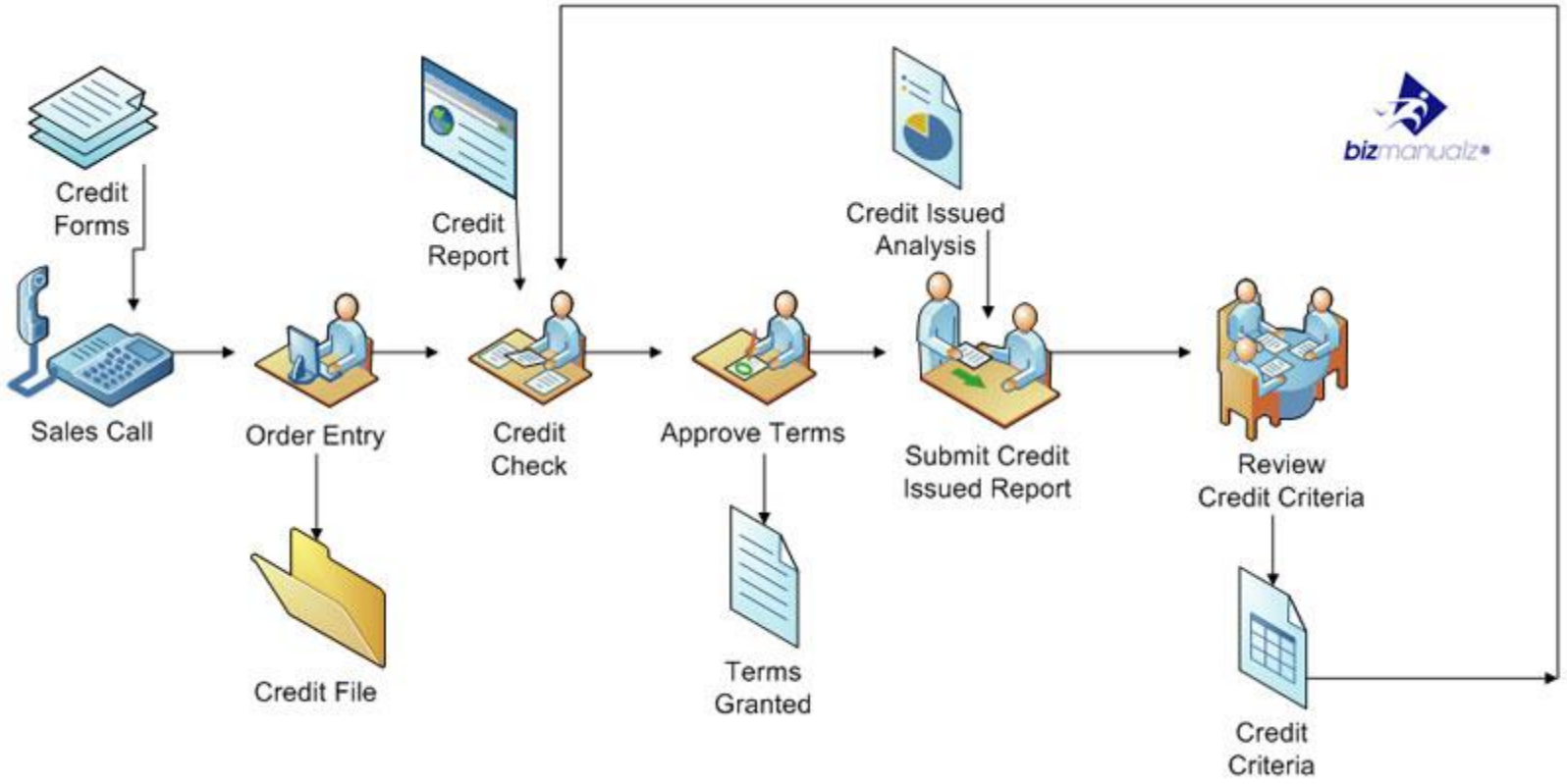


Flow Diagrams
Swim Lane

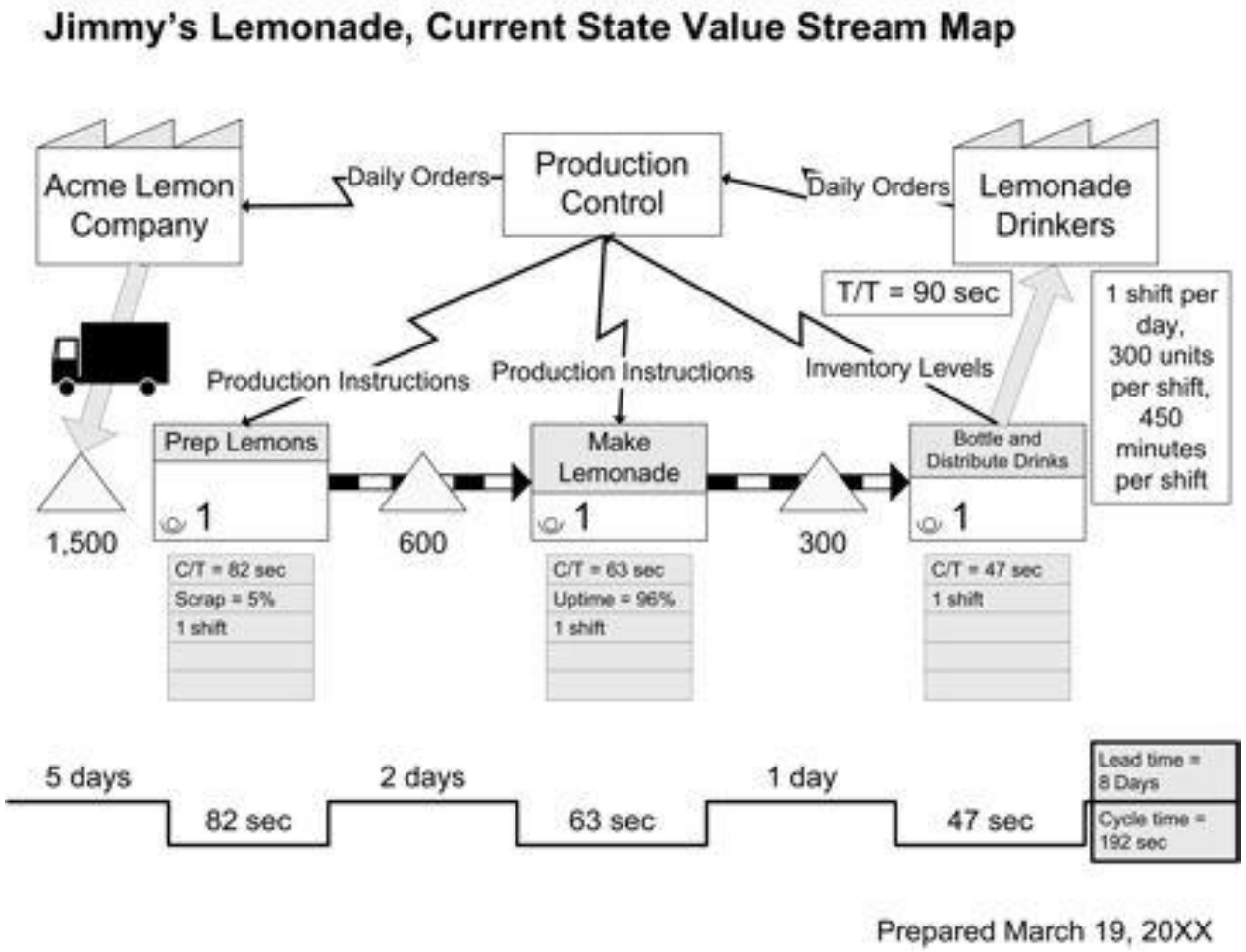


The thicker flow is the flow of material.
The thinner flow is the flow of operation/information.
Activities can be color coded: yellow means IT support, blue means manual activities.

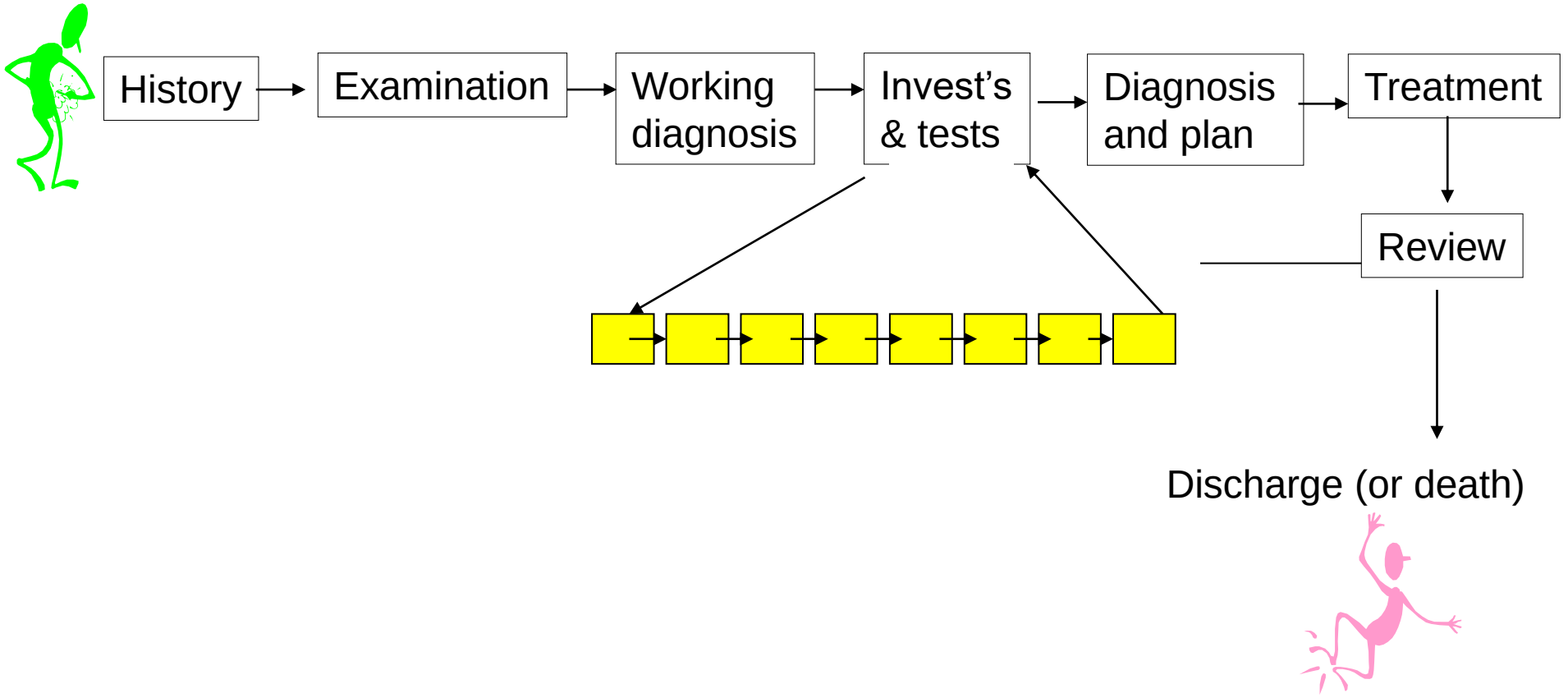
Flow Diagrams - Visio



Lean Value Stream Map

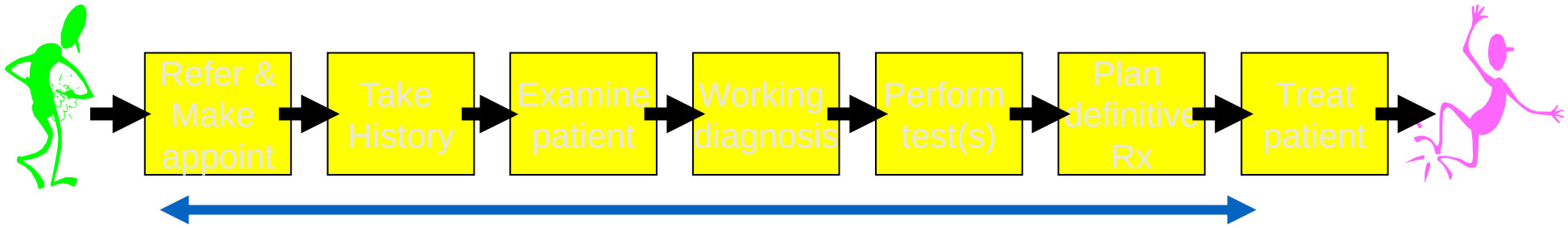


The Hospital's View:



Ref: Richard S Steyn

The Patient's view:

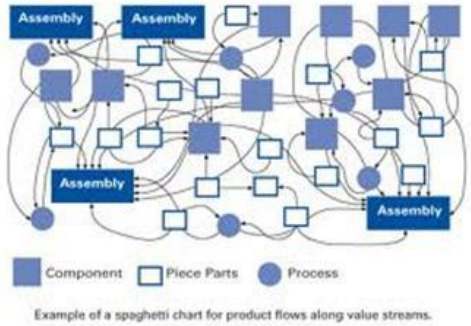


- Doctor requests 'test'
- Porter picks up request
- Porter delivers request to clerk
- Clerk logs request
- Clerks puts request for prioritisation
- Consultant prioritises request
- Consultant returns request
- Clerk files request in priority order
- Clerk draws request from file
- Clerk makes appointment
- Clerk sends appointment by post (>6 steps)
- Patient receives appointment
- Patient travels to hospital
- Patient finds car parking slot
- Patient finds X-ray department
- Patient checks in at reception

Ref: Richard S Steyn

Press Process Flow Chart

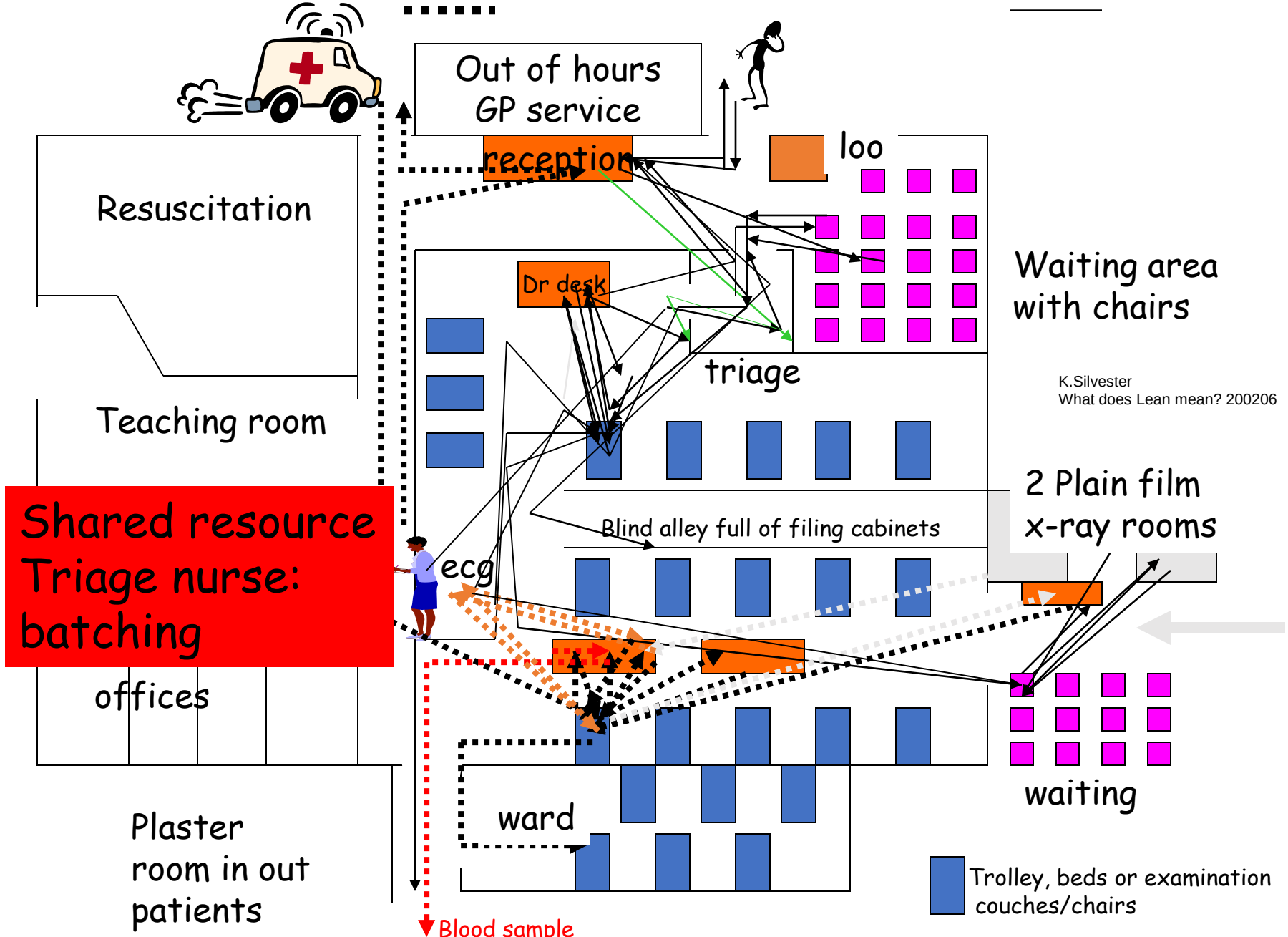
Spaghetti Diagrams



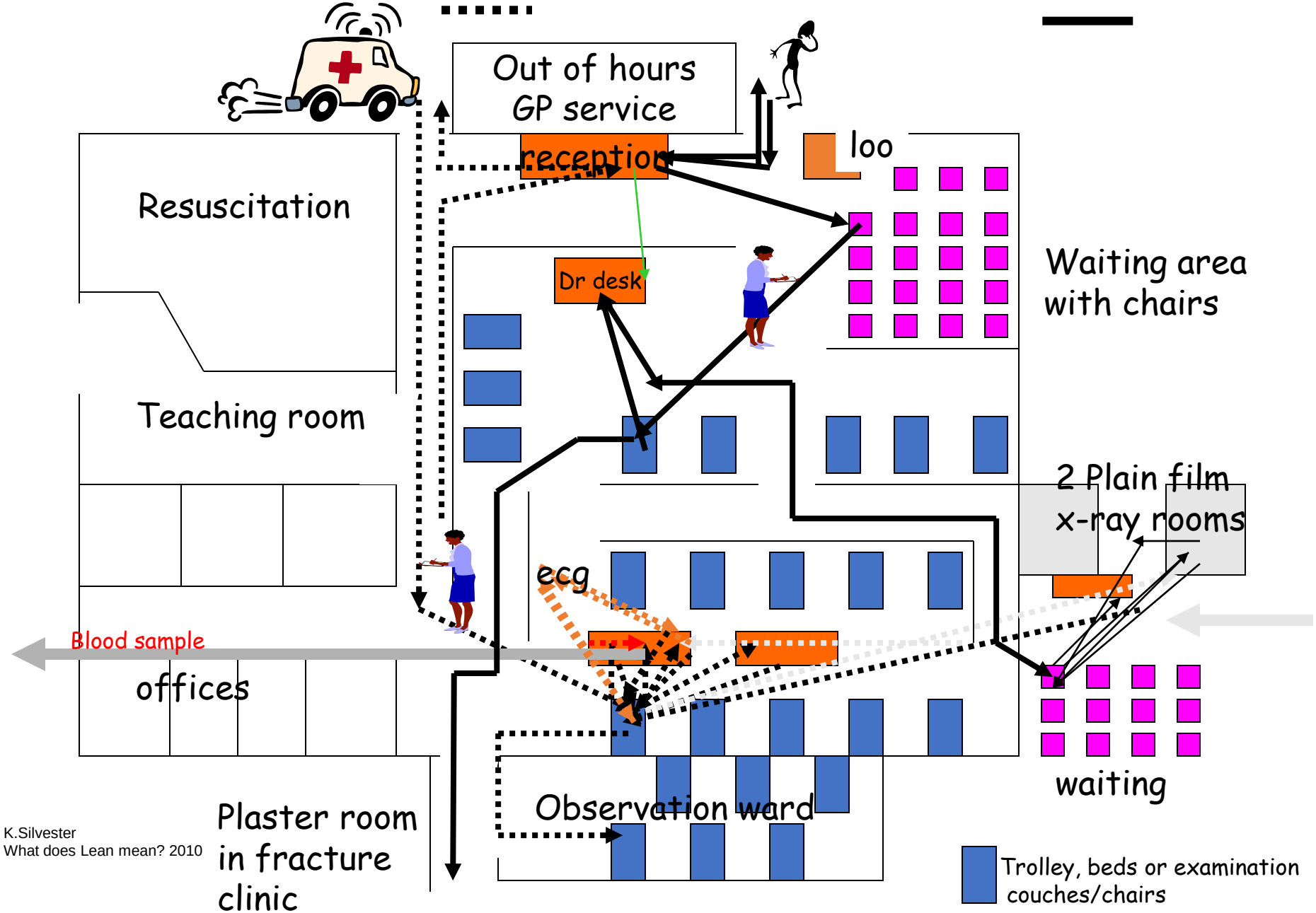
Example of a spaghetti chart for product flows along value streams.

- Diagrams that depict the physical flow of work or material in a process
- Used to improve the physical layout of a workspace e.g. office, factory, warehouse
- They are useful to depict the flow of information, material, or people
- Handoffs add significant delays and queue times.
- Therefore, if you see a lot of crisscrossing lines, investigate ways to reduce handoffs and simplify the layout

BEFORE Spaghetti Diagram for A&E department showing processes for patient with chest pain and patient with injured wrist.



AFTER Spaghetti Diagram for A&E department showing processes for patient with chest pain and patient with injured wrist.



K.Silvester
What does Lean mean? 2010