LSSH
lean six sigma for healthcare

Basic Statistics

U4 – Measure

E2 – Statistics

The element ‘Statistics’ reviews the basics of statistics such as mean, deviation and probability. This element reviews a range of graphs that can be used to visualize data as well.

Lifelong Learning Programme

Yellow Belt

Types of Data

Scale types

Attributive data – Qualitative

- **Categorical / Nominal**
  - Objects have **no natural order** (example: gender man/ woman)
- **Ordinal**
  - Objects have a **natural order**, but the **differences cannot be quantified and compared** (example: seniority in a nursing team)

Variable data – Quantitative

- **Interval**
  - Objects have a **natural order** and **differences can be quantified and compared** (example: Temperature in an operating theatre)
- **Ratio**
  - Objects have a **natural order**, **differences can be quantified and compared** and **there is a zero** (example: one patient (140 kg) is 2 times heavier than another patient (69 kg))

Types of Data

Attributive data – Qualitative (categorical and ordinal)

Based on categories

Categorical variables

- Countable number of possible outcomes/ no natural order

(example: 'months in a year' has 12 possible outcomes

Ordinal variables

- Countable number of possible outcomes/ natural order

(example : 'I find this hospital service': 1-very bad; 2-bad; 3-moderate; 4-good; 5-excellent)

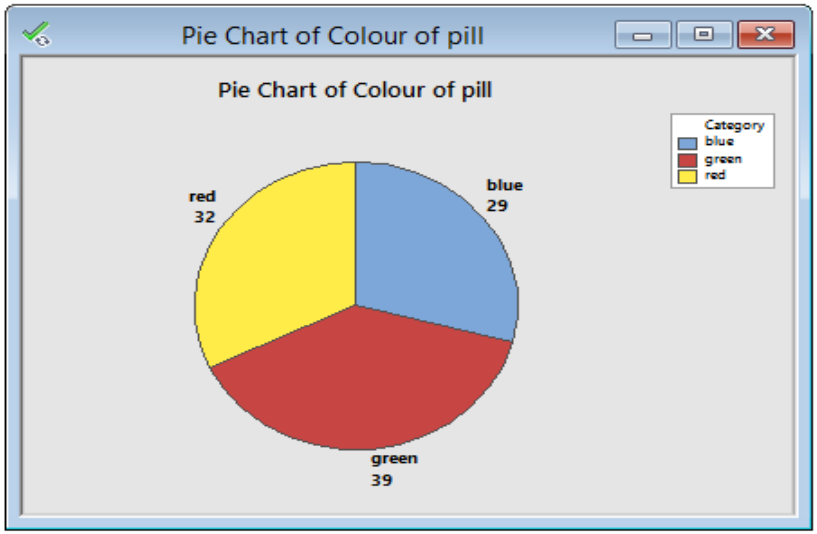
Types of Data

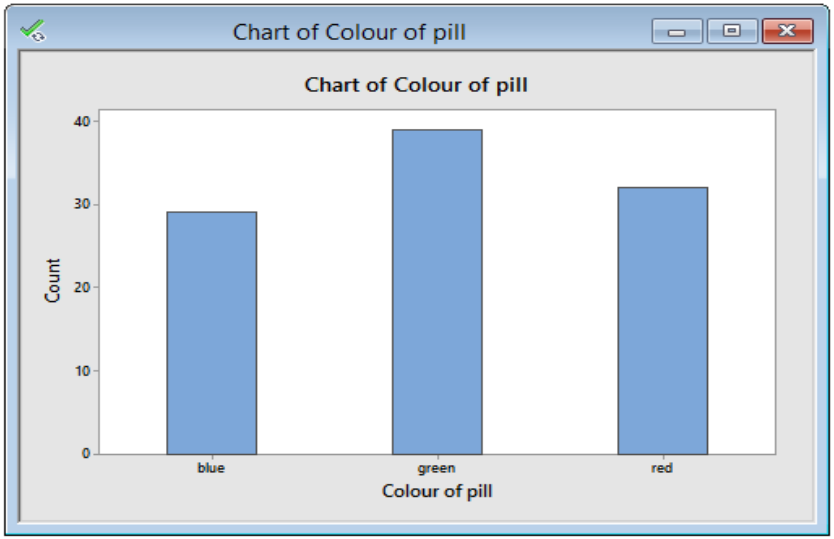
Example Categorical data

Nominal Data

Patients were given different pills during a clinical trial each pill was identified by its colour

red
green
blue
red
red
green
green
green
green
blue





Types of Data

Variable data – Quantitative (interval and ratio)

Based on measurements and/ or counts

Discrete - Countable

- Can be counted: a discrete (whole) number of events
- Scale is a ratio scale
- Number of patients

Continuous - Measurable

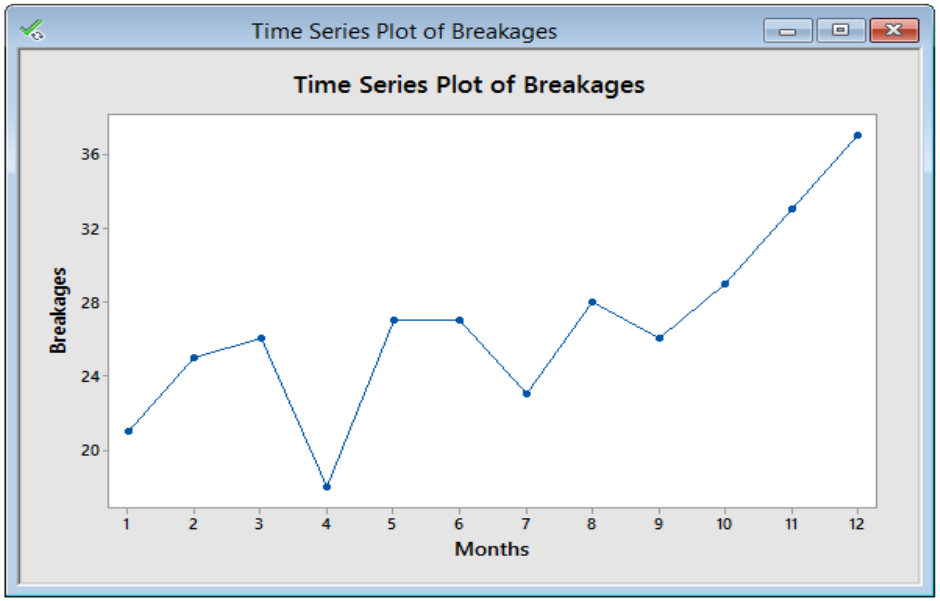
- Can be measured on a continuous (interval or ratio) scale
- Length, time, weight

Types of Data

Example Discrete data

Number of breakages in an orthopaedic clinic

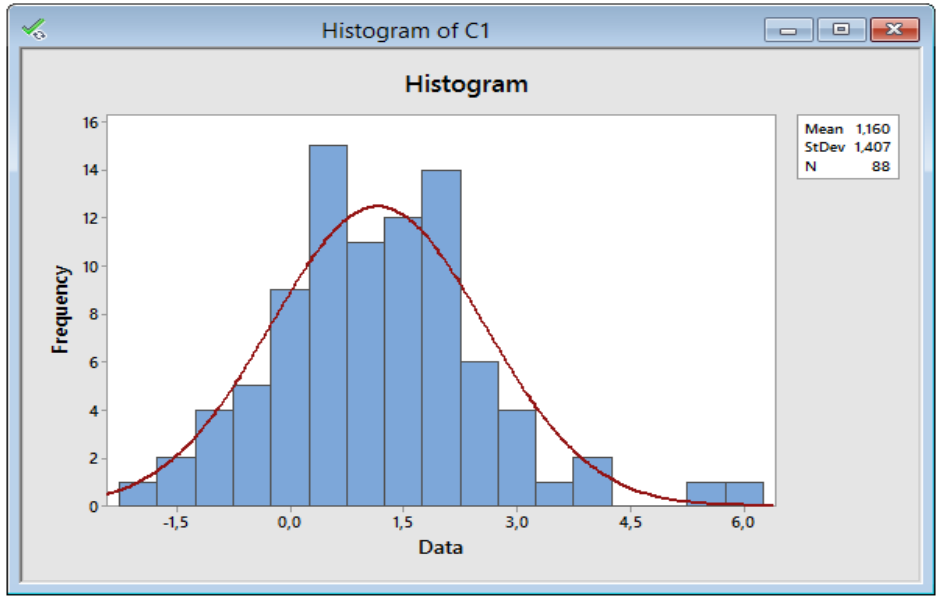
21
25
26
18
27
27
23
28
26
29
33
37



Types of Data

Example Continuous data

0.73	2.46	2.15	1.47
1.96	1.33	1.45	0.49
2.05	-0.02	-0.08	0.87
4.15	0.29	0.91	0.59
1.78	-0.12	0.58	1.83
-0.08	-0.78	2.22	-0.49
1.21	0.10	-2.00	1.80
1.48	-1.25	3.09	1.19
0.45	0.95	1.14	0.72
2.91	2.29	1.14	-1.39
2.85	-0.34	1.52	0.80
-0.11	5.48	2.06	0.07
2.25	-0.05	0.57	1.33
-0.45	1.47	-1.04	5.91
2.69	2.06	-0.18	1.84
1.69	1.32	0.50	0.60
-0.42	2.62	3.40	1.93
1.06	-0.88	2.47	0.97
0.93	1.69	-0.32	0.38
2.00	1.30	0.30	1.40
0.42	1.79	1.80	0.46
-1.49	0.68	3.16	4.04



Sampling methods

Introduction into sampling

If you want to say something about a large group, you first have to do some sort of measurements

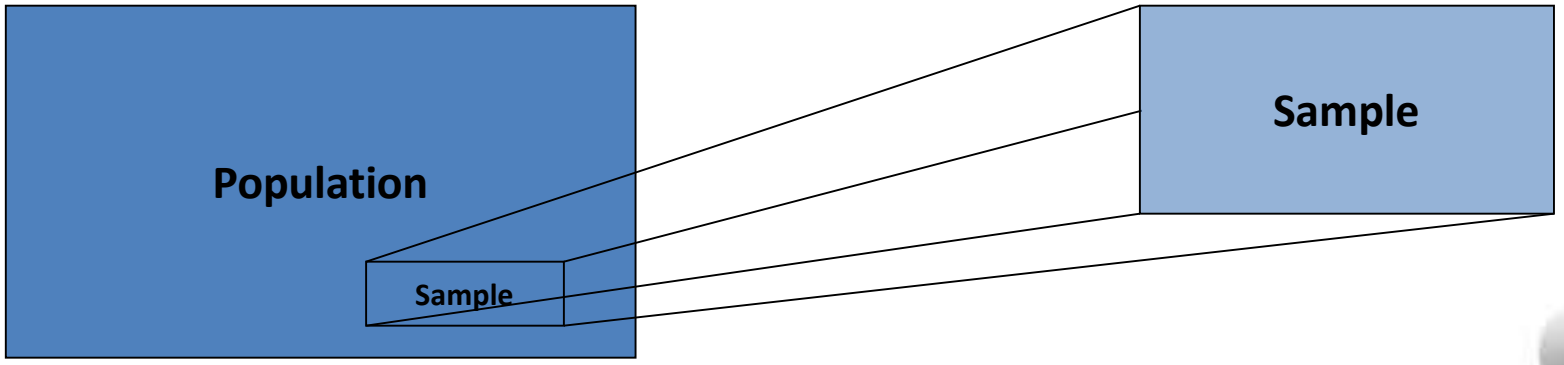
For example If we would like to say something about the length of people, basically we should measure everybody, and average that data

Obviously that would be a lot of work…

Sampling methods

Introduction into sampling

If we choose the right sample the average of the group will be pretty accurate



Sampling methods

Important definitions

Sampling

- **The process of selecting a subset of items from a population**
- **This subset should give information about the population**

Observation

- **Measures one or more properties of an item selected from a population**

Population

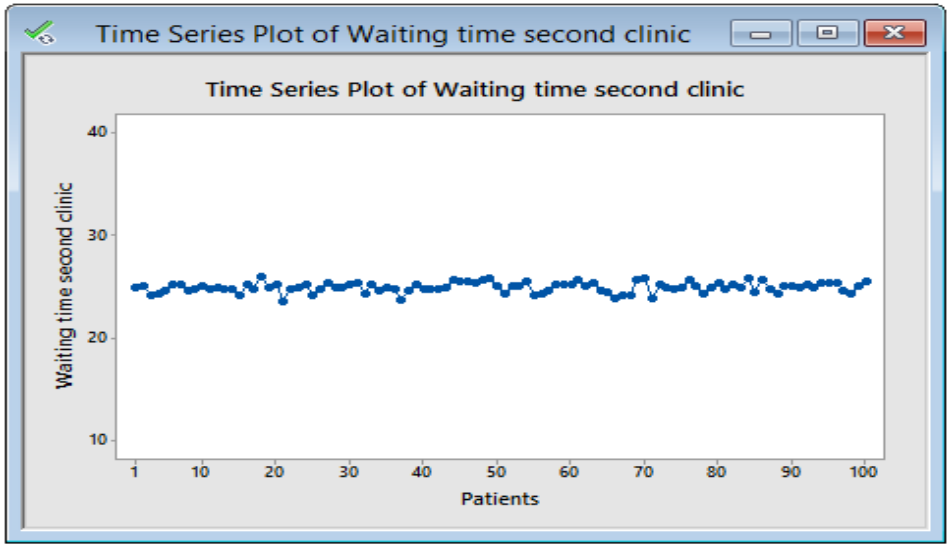
- **The complete set of items whose characteristics one wishes to investigate**

Variation

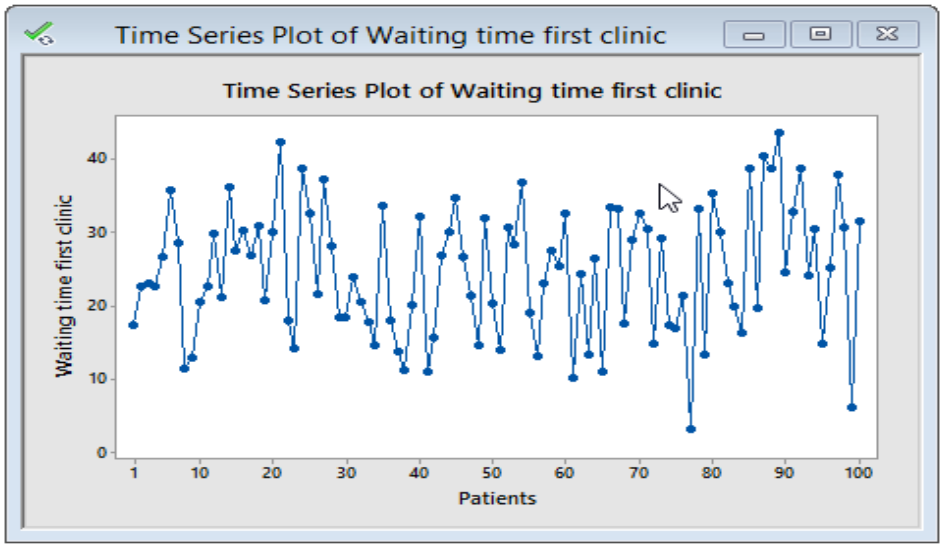
Variation is everywhere

Process A – the waiting time for one hospital clinic has relatively little variation compared to:

Process B – the waiting time for another hospital clinic



A



B

Variation

Is variation acceptable?

Variation is present in every process

We may accept variation when:

- **The process mean is "on target" e.g. the average number of patients seen in the Emergency Unit**
- **The total variation is relatively small compared to the specification (expectation)**
- **The process is stable over time**

Process Characteristics

Understand the process to make improvements

Process variation has several components

- **Centre:** **Is the process on target?**
- **Spread:** **How much variation is in the data?**

In all cases

- **Use graphs and statistical parameters to visualize the process**

Process Characteristics

Two measures of central tendency

Arithmetic average: Mean

- **Arithmetic mean of data**

$$\overline{X} = \frac{(3+5+4+7+5)}{5} = 4.8$$

Median Me: Median

- **Middle value of sorted data**

$$Median(\ 3;4;5;5;7\ ) = 5$$

Data: {3, 5, 4, 7, 5}

Process Characteristics

Exercise: determine mean and median

Arithmetic average: Mean

- **Arithmetic mean of data**

Median Me: Median

- **Middle value of sorted data**

Data : {2, 3, 4, 5, 3, 10}

Process Characteristics

Measures for spread of a sample

Range (R)

- **Absolute difference between maximum and minimum value from the dataset**

$$R = x_{\max} - x_{\min}$$

Variance (s^2)

- **Average squared distance of X**

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2$$

Standard deviation (s)

- **Square root of the variance s^2**

$$s = \sqrt{s^2}$$

Process Characteristics

Exercise: determine R, s^2 and s

Range (R)

- **Absolute difference between maximum and minimum value from the dataset**

Variance (s^2)

- **Average squared distance of X**

Standard deviation (s)

- **Square root of the variance s^2**

Data: {2, 3, 4, 5, 3, 7}

Normal distribution

Normal distribution or Gauss curve

- **This bell-shaped, continuous distribution is common and is characterized by the mean μ and standard deviation σ**