CONTINUED PROCESS VERIFICATION:
A PRACTICAL APPROACH

TABLE OF CONTENTS

INTRODUCTION

INTRODUCTION

To understand what to do in practice it is necessary to answer first the question:

What is

Continued (or Ongoing) Process Verification?

To answer this question, reference should be made to the relevant Regulatory documentation.

INTRODUCTION

- **Continued Process Verification:** Assuring that during routine production the process remains in a state of control.

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

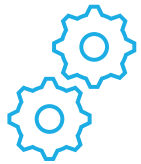
- **Ongoing Process Verification** (also known as continued process verification): Documented evidence that the process remains in a state of control during commercial manufacture.

EU Guidelines for GMP – Annex 15 : Qualification and Validation, Eudralex, Volume 4 (March 2015)

INTRODUCTION

- **Ongoing Process Verification** (also known as continued process verification): documented evidence that the process remains in a state of control during commercial manufacture.

PIC/S : Guide to Good Manufacturing Practice for Medicinal Product Annexes (February 2022)



- **Continuous Process Verification**: *an alternative approach to process validation* in which manufacturing *process performance is continuously monitored and evaluated*.

ICH guideline Q8 (R2) on pharmaceutical development (June 2017)

INTRODUCTION

- From the ICH Q8 document, the concept of Continuous Process Verification is introduced in the context of a Quality by Design (QbD) framework.
- The QbD framework is an approach to product development that begins with predefined objectives and emphasizes understanding of product and process, efficient and effective process control, and continual improvement through innovation.
- The Continuous Process Verification approach is based on thorough product and process understanding and process control. It requires a higher level of process understanding and control than the traditional process validation approach.

INTRODUCTION

Summarizing, in Europe, as described in Annex 15 of the GMPs, **three scenarios are practically possible**:

- ***Traditional Process Validation***
  - Manufacture of a number of batches of finished product under routine conditions to confirm its reproducibility.
  - It is generally considered acceptable to perform it with a *minimum of three consecutive batches*, although it is stated that this initial validation exercise involving three batches should be completed with additional data obtained from subsequent batches, as part of the on-going process verification program.

INTRODUCTION

- ***Continuous Process Verification***

  A Quality by Design (QbD) approach, where it has been scientifically proven during the development phase that the established control strategy provides a high degree of assurance of product quality.

- ***Hybrid approach***

  A hybrid of the traditional approach and continuous process verification.

INTRODUCTION

- The FDA does not carry out such a division and establishes that a *Process Performance Qualification* (PPQ) must be carried out, in which, among other aspects, ***the number of batches should be justified*** to demonstrate with a high degree of assurance that the process is able to consistently provide a quality product, considering an acceptable level of confidence in both intra-batch and inter-batch variability. In this regard, see, for instance: https://www.fda.gov/drugs/guidances-drugs/questions-and-answers-current-good-manufacturing-practice-regulations-production-and-process

- A possible approach based on CI is presented for drug products in: Ajay Pazhayattil *et al.*, *AAPS PharmSciTech*, Vol. 17, No. 4, August 2016

INTRODUCTION

Whichever method is used to carry out this phase

(*i.e.*, traditional, full QbD or hybrid),

THE THIRD PHASE IS OBLIGATORY

Continued Process Verification in FDA terminology (CPV) or

Ongoing Process Verification in EMA terminology (OPV)

Only a clarification:

Verification vs. Validation

INTRODUCTION

21 CFR Part 820 : Quality System Regulation:

- Sec. 820.3(1) **Process validation** means establishing by objective evidence that a process consistently produces a result or product meeting its predetermined specifications.

- Sec. 820.3(2) Design validation means establishing by objective evidence that device specifications conform with user needs and intended use(s).

- Sec. 820.3(aa) **Verification** means confirmation by examination and provision of objective evidence that specified requirements have been fulfilled.

https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-820/subpart-A/section-820.3

INTRODUCTION

21 CFR Part 820 : Quality System Regulation:

- 21 CFR Part 820 does not provide any specific definition of “process verification” ***however*** this general definition clearly intends that to verify that a process is working, you need to be able to provide some type of objective evidence (a measurement) that proves the outcome of the process meets your specified requirements.

https://www.ecfr.gov/current/title-21/chapter-I/subchapter-H/part-820/subpart-A/section-820.3

INTRODUCTION

Summarizing:

- **Validation** refers to the process of ensuring that a system, process, or equipment consistently produces results that meet predetermined specifications. It involves testing and documenting every step to ensure compliance with regulatory requirements.

- **Verification** focuses on confirming that a specific product, system, or component meets specified requirements. It involves reviewing documents, conducting inspections, and performing tests to ensure accuracy and completeness.

In a nutshell:

Validation ensures consistency and compliance, while **Verification** confirms accuracy and completeness.

Let's now go back to

Continued (or Ongoing) Process Verification

INTRODUCTION

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

- The **goal of Continued Process Verification** is continual assurance that the process remains in a state of control (the validated state) during commercial manufacture.
- A system or systems for detecting unplanned departures from the process as designed is essential to accomplish this goal.
- An ongoing program to collect and analyze product and process data that relate to product quality must be established and … the data should be statistically trended and reviewed by trained personnel.

INTRODUCTION

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

- The information collected should verify that **the quality attributes are being appropriately controlled throughout the process**.
- We recommend that a statistician or person with adequate training in statistical process control techniques develop the data collection plan and statistical methods and procedures used in **measuring and evaluating process stability and process capability**.
- Procedures should describe how trending and calculations are to be performed and should guard against overreaction to individual events as well as against failure to detect unintended process variability.

INTRODUCTION

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

- Production data should be collected to evaluate **process stability** and **capability**.
- The quality unit should review this information.
- Scrutiny of **intra-batch as well as inter-batch variation** is part of a comprehensive continued process verification program under § 211.180(e)
- We recommend continued monitoring and sampling of **process parameters** and **quality attributes** at the level established during the process qualification stage until sufficient data are available to generate significant variability estimates.

INTRODUCTION

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

- **Process variability should be periodically assessed**, and monitoring adjusted accordingly.
- Variation can also be detected by the timely assessment of defect complaints, out-of-specification findings, process deviation reports, process yield variations, batch records, incoming raw material records, and adverse event reports.
- Production line operators and quality unit staff should be encouraged to provide feedback on process performance.
- We recommend that the quality unit meet periodically with production staff to evaluate data, discuss possible trends or undesirable process variation, and coordinate any correction or follow-up actions by production.

INTRODUCTION

FDA Guidance for Industry – Process Validation: General Principles and Practices (January 2011)

- Data gathered during this stage might suggest ways to improve and/or optimize the process by altering some aspect of the process or product, such as the operating conditions (ranges and set-points), process controls, component, or in-process material characteristics.
- Maintenance of the facility, utilities, and equipment is another important aspect of ensuring that a process remains in control….. The equipment and facility qualification data should be assessed periodically to determine whether re-qualification should be performed and the extent of that re-qualification. Maintenance and calibration frequency should be adjusted based on feedback from these activities.

INTRODUCTION

EU Guidelines for GMP – Annex 15 : Qualification and Validation, Eudralex, Volume 4 (March 2015)

5.29. Manufacturers should monitor product quality to ensure that a state of control is maintained throughout the product lifecycle with the relevant process trends evaluated.

5.30. The extent and frequency of ongoing process verification should be reviewed periodically. At any point throughout the product lifecycle, it may be appropriate to modify the requirements taking into account the current level of process understanding and process performance.

INTRODUCTION

EU Guidelines for GMP – Annex 15 : Qualification and Validation, Eudralex, Volume 4 (March 2015)

5.31. Ongoing process verification should be conducted under an approved protocol or equivalent documents and a corresponding report should be prepared to document the results obtained. Statistical tools should be used, where appropriate, to support any conclusions with regard to the variability and capability of a given process and ensure a state of control.

5.32. Ongoing process verification should be used throughout the product lifecycle to support the validated status of the product as documented in the Product Quality Review. Incremental changes over time should also be considered and the need for any additional actions, *e.g.*, enhanced sampling, should be assessed.

Summarizing Regulatory prescriptions

To approach CPV in a practical way we need one or more documents which:

1. Define the methodological approach to be followed. The suitable document for this purpose is an SOP that indicates:
   - company department involved and their specific roles,
   - statistical tools useful for this purpose (this may be the subject of one or more specific SOPs),
   - flow of operations
   - “phases” descriptions,
   - *etc.*

INTRODUCTION

2. Provide the criteria for the identification of those critical process parameters (CPP) and critical quality attributes (CQA) that must be continuously monitored and sampled (*e.g.*, Risk Analysis and/or theoretical considerations)

3. Assess by historical data analysis:
   - process stability and variability also considering intra-batch as well as inter-batch variation
   - process capability (if applicable)
   - data trending

INTRODUCTION

4. Define the operational tools to extract information from the historical data referred to in point 3.

5. Define the criteria for evaluating the results with the aim of preventing overreactions to single events or the failure to detect unintended process variability.

6. Prescribe periodical meeting of the Quality Unit with Production Staff to evaluate data, discuss possible trends or undesirable process variation, and coordinate any correction or follow-up actions by production.

INTRODUCTION

7. Provide information on how often process variability should be evaluated and, if necessary, the monitoring appropriate accordingly.

8. Encourage, based on the data collected, the possibility of improvements and/or process optimizations or, in case of failures, escalate actions.

INTRODUCTION

What referred to in points 2 to 8 can be seen as elements of a **CPV Plan (or Protocol)** which should provide a rationale for :

- Parameters and attributes to be monitored
- CPV limits for each parameter and attribute combination
- Frequency of trend evaluations
- Statistical signals to be evaluated
- Default responses for each parameter-signal combination
- Best practices for information sharing, reporting (*i.e.*, **CPV Report**) and necessary actions to be undertaken

INTRODUCTION

Given these premises, let's enter even more deeply into the practical operation which, in the end, is what we have to do!

INTRODUCTION

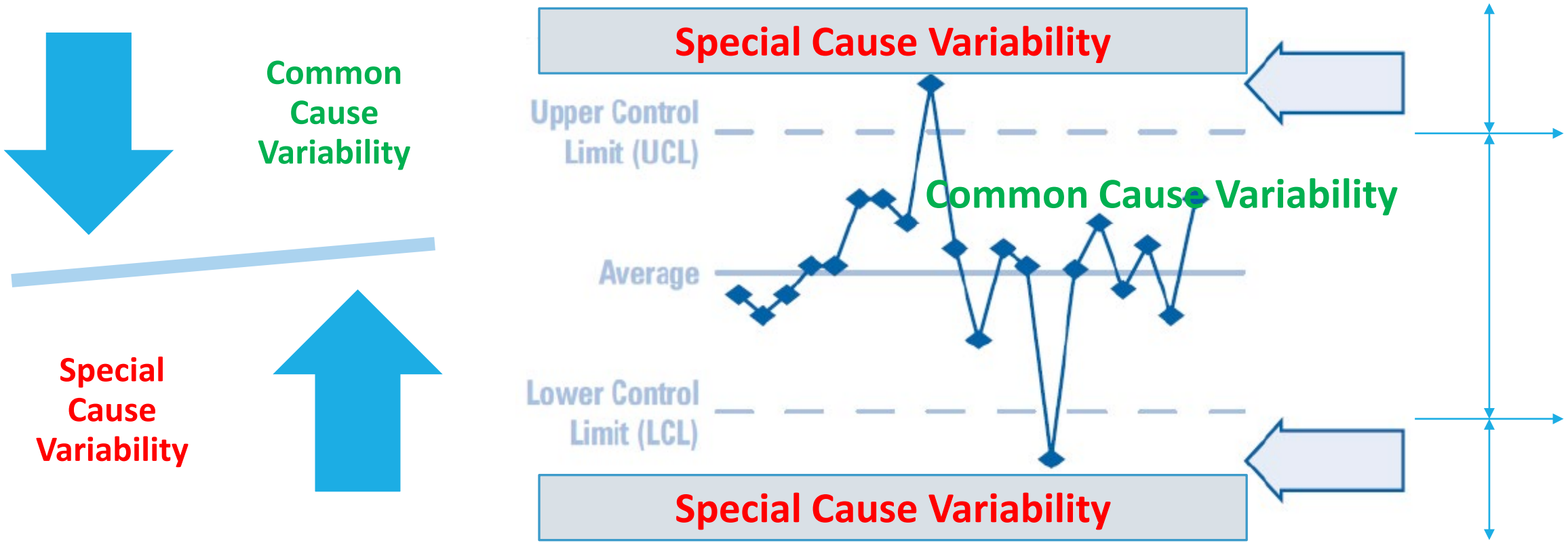
Given that, according Shewhart:

- the ***causes that contribute to the variability*** of a production process are essentially of two types: ***common causes*** and ***special (W.E. Deming) or assignable (W.A. Shewhart) causes.***
- a process is said to be ***under statistical control*** when ***its variability is due only to common causes.***
- **«** *… a phenomenon will be said to be controlled when, through the use of past experience, we can predict, at least within limits, how the phenomenon may be expected to vary in the future.* »

W.A. Shewhart, *Economic Control of Quality of Manufactured Product*, Van Nostrand (1931) p. 6

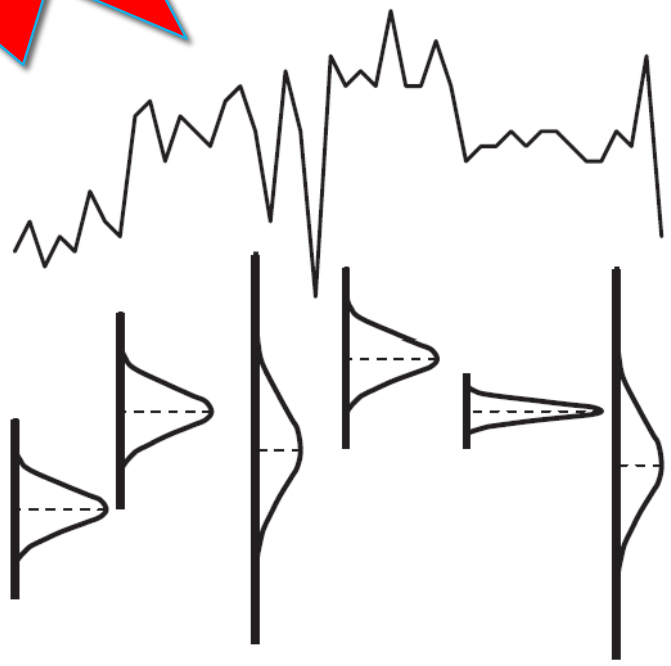
INTRODUCTION

The Shewhart Concept of Variation

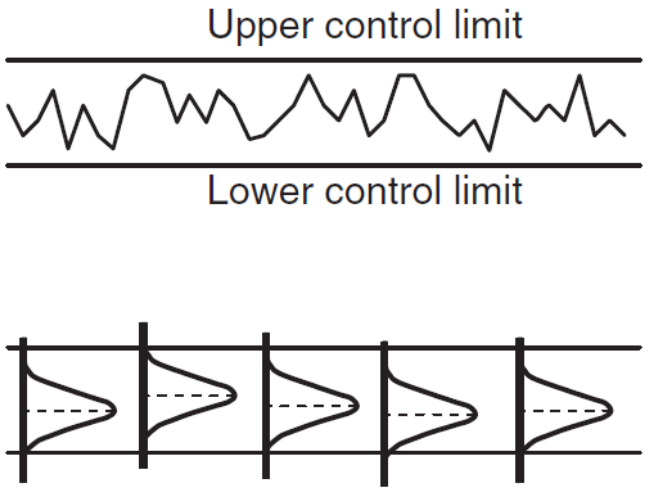


INTRODUCTION

Out of Control



Unstable process



Stable process

In Control

A process in (statistical) control is a predictable process !

INTRODUCTION

From the previous regulatory documents, **two phases** of the process aimed at defining a CPV strategy can be identified:

- **Phase 1:**
  - is the step in which a set of data is gathered and analyzed in a *retrospective* way to assess if the process has been in control over the period of time covered.
  - *At this stage **process variability is investigated** from **both** an **inter-batch** and an **intra-batch perspective.*** This aspect is **very important!**
  - If the process does not result to be stable, specific tools (*e.g.*, *control charts with trial control limits*) will be used to assist the following efforts intended to bring it into a state of statistical control.

INTRODUCTION

- **Phase 1 (*cont.*)**:
  - As a *rule-of- thumb*, at least 100 observations should be used to compute trial control limits, with approximately 300 observations used for "permanent" limits*. However, since processes change over time, parameters must eventually be re-estimated.

Please note that the above for Phase 1 also applies to APQR!

*C.P. Quesenberry, *The Effect of Sample Size on Estimated Limits for $\bar{X}$ and X Control Charts*, J. Quality Technology, 25 (1993)

INTRODUCTION

- **Phase 2**: is the phase that ***ideally*** begins only after the process has proven to be ***fully*** stable and operating under optimal conditions.

  It is aimed at *monitoring* it, trying to detect even small deviations as they may, perhaps, be the first signs of a progressive departure from the control conditions.

  Also, this phase is assisted by statistical tools that are specific for the intended purpose.

Ideally ? Fully ?

The meaning is rather simple: it is possible that once Phase 1 is completed, we conclude that for one or more parameters considered it is not possible or, better, it is not convenient (for economic reasons, impact of the change, *etc.*) to reduce variability. At that point we can still move on to Phase 2 but with the awareness of the choices made which, of course, will have to be justified and documented.

We will see later a practical example to clarify this point (i.e., Critical Process Temperature)

PHASE 1

PHASE 1

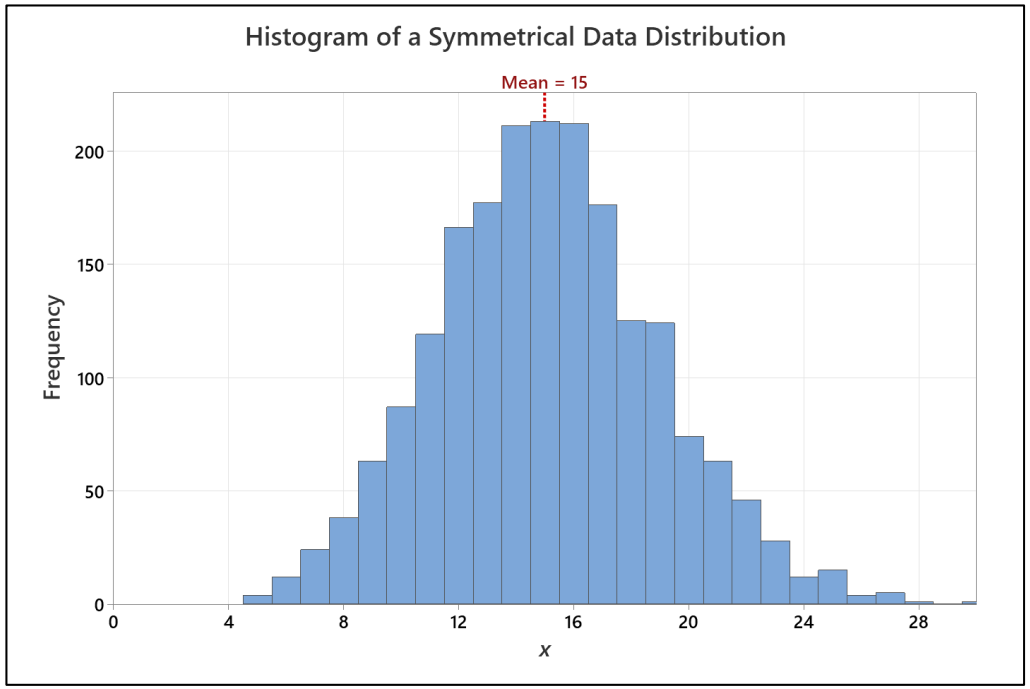
How process variability can be assessed in Phase 1?

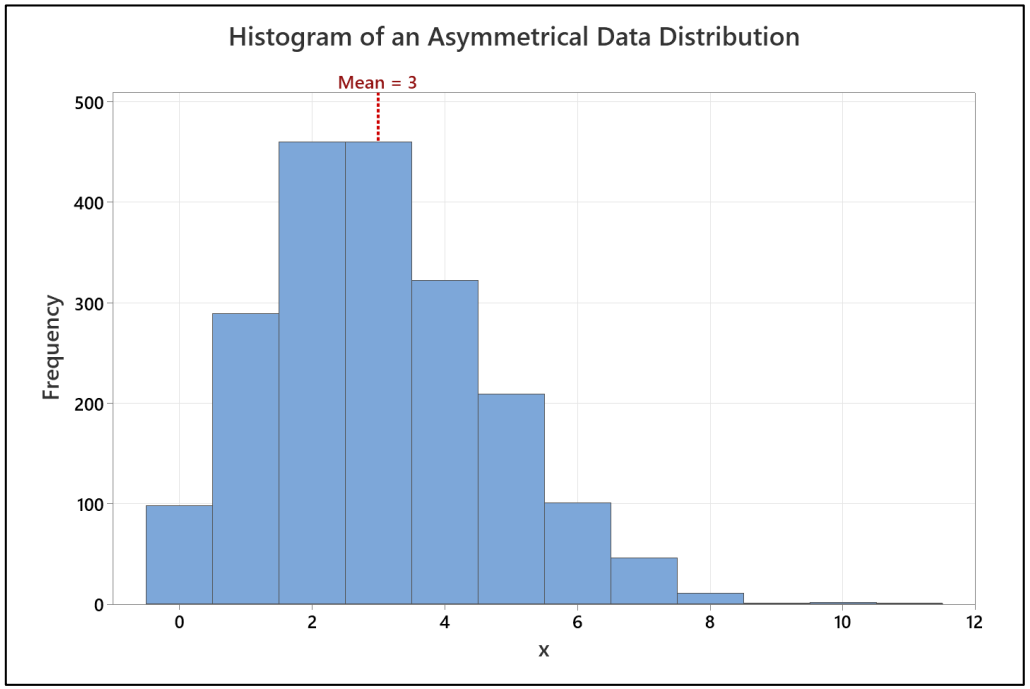
Since the variables we mainly deal with are ***numerical*** and ***continuous or discrete***, Statistics provides several types of tools, namely:

- **GRAPHICAL TOOLS**: histograms, boxplots and line plots
- **SUMMARY INDICES (OR *STATISTICS*)**: position, variability and shape indices
- **INFERENTIAL METHODS**: ANOVA, statistical intervals, modeling based on probability distributions, *etc.*
- **CONTROL CHARTS** (*individual* and *bar* charts) and **CAPABILITY INDICES**

PHASE 1

The graphs below show two data sets: the first corresponds to a symmetrical distribution with a mean of 15 while the second an asymmetrical distribution with a mean of 3. Here are visually illustrated the concepts of *histogram* and *shape* and *position (location) index* of a distribution (or set) of data.

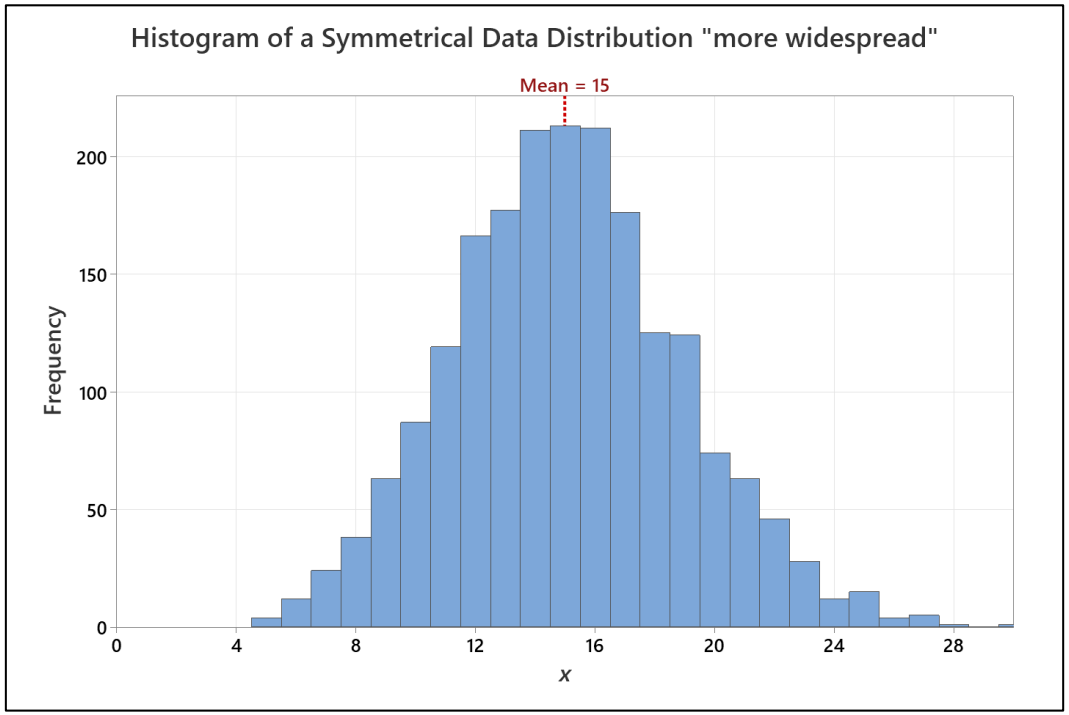


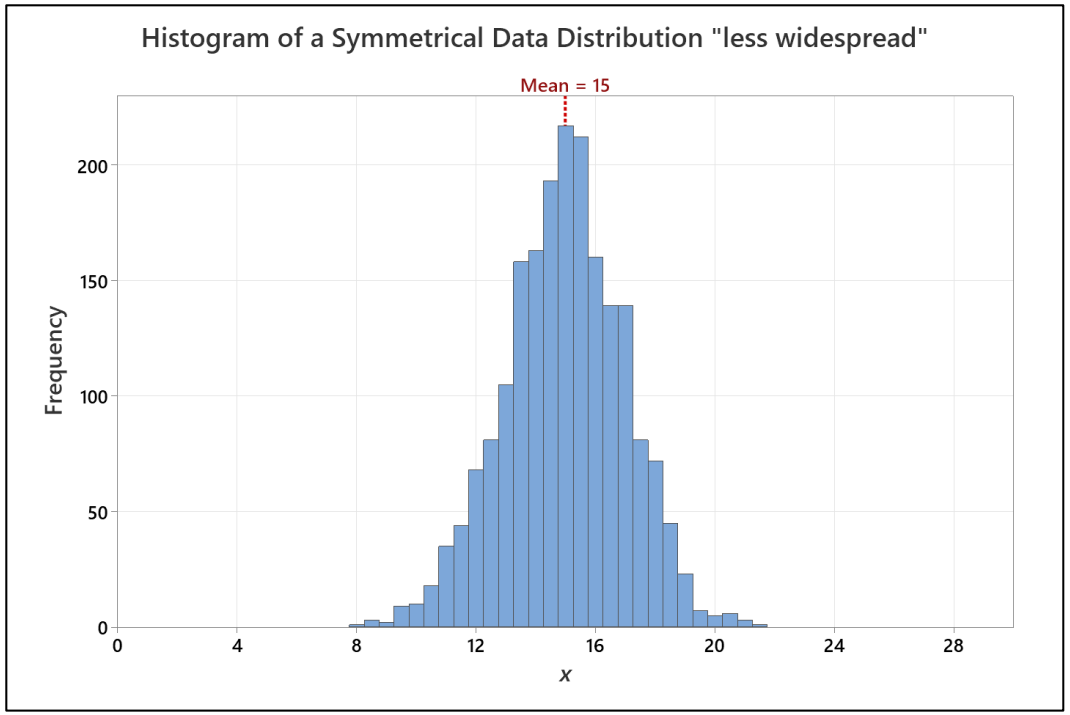


PHASE 1

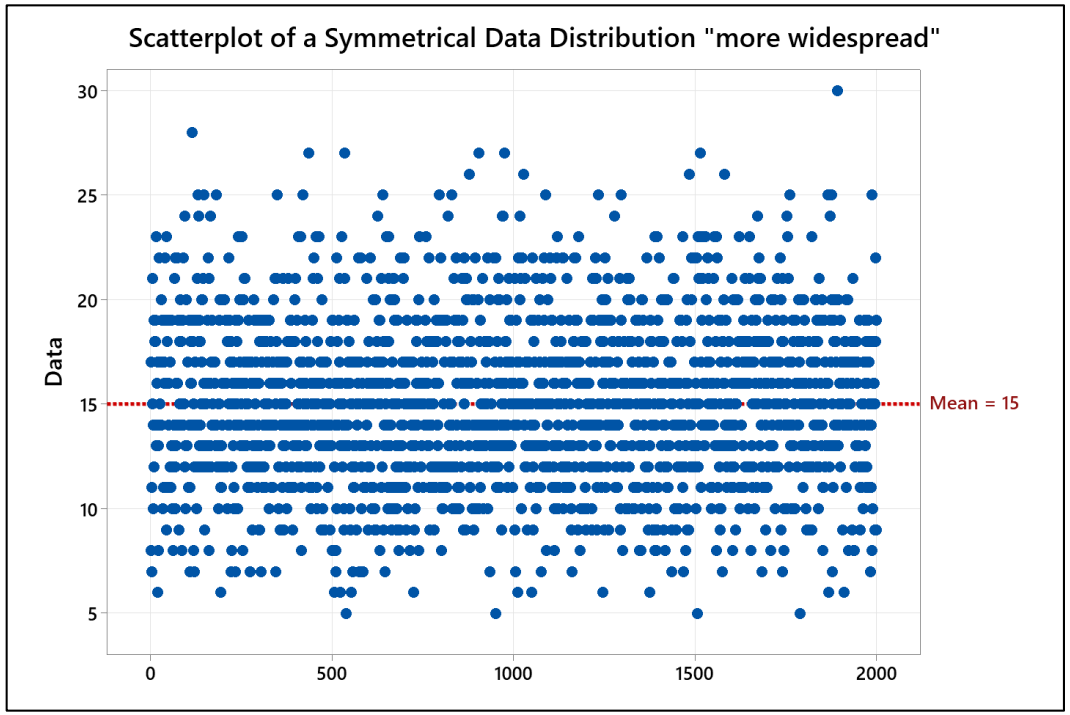
Alongside the shape (symmetrical or not) a data distribution can be more or less widespread. The histograms below show the concept of *variability index* or *dispersion with respect to a center* (*standard deviation*).

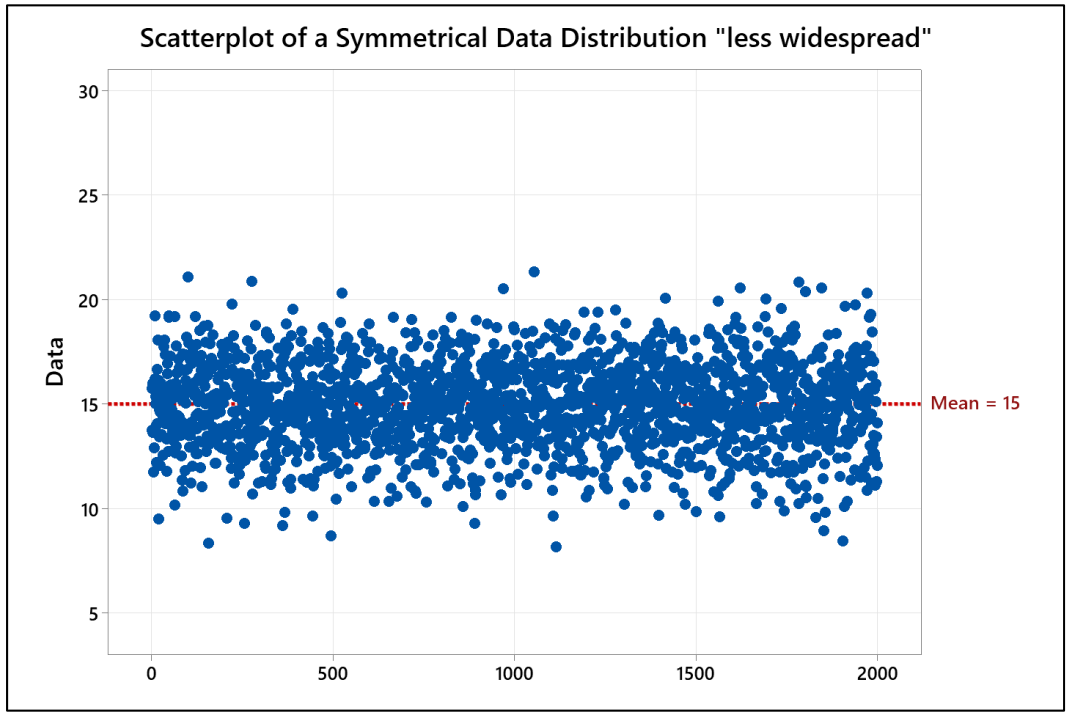
It is evident that, by itself, a position index is insufficient to describe a distribution of data !





PHASE 1





It is evident that, by itself, a position index is insufficient to describe a distribution of data !

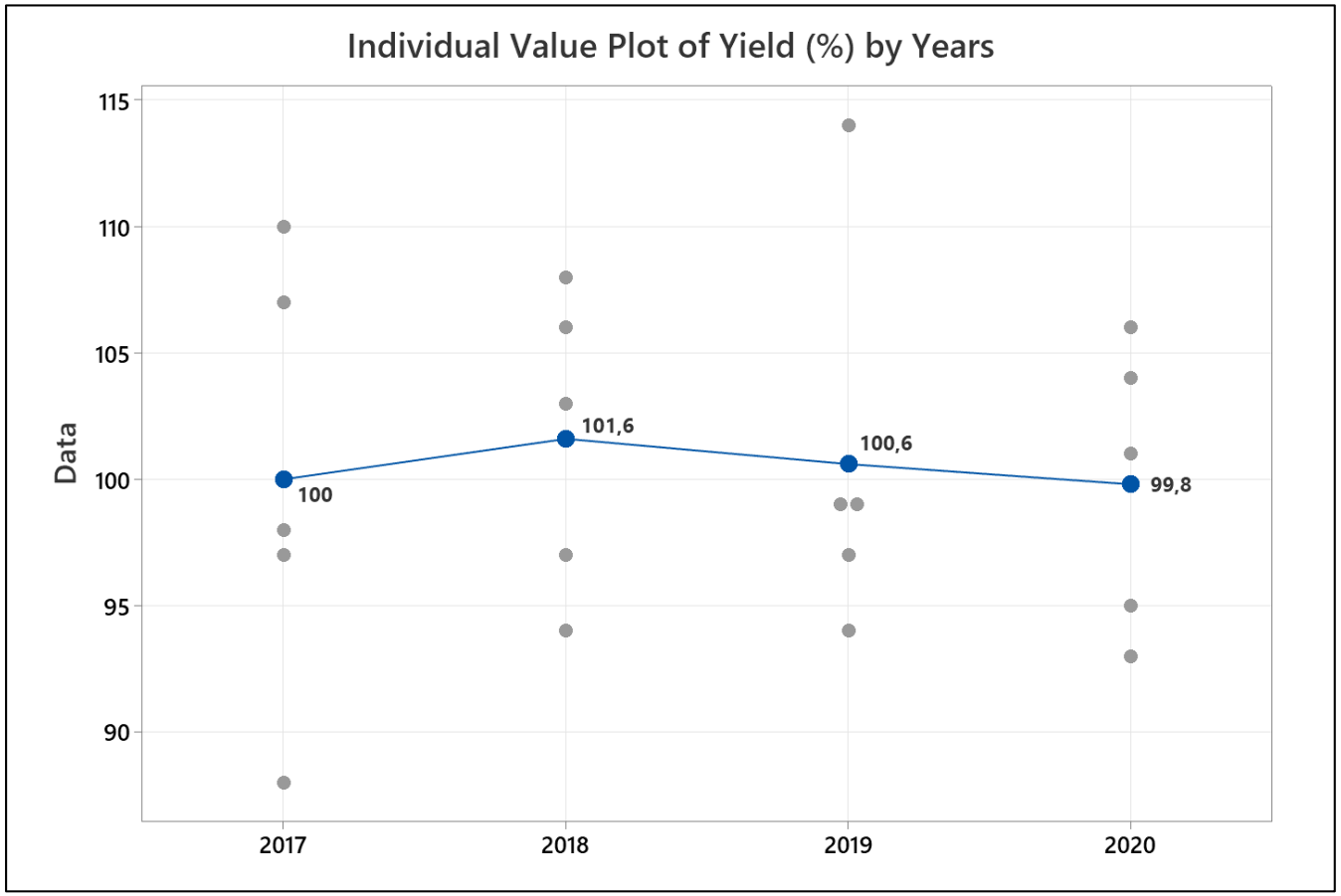
PHASE 1

The concepts of centrality and dispersion of a data distribution are also well illustrated using another graphical tool, the

Individual plot

which works in this case since it is a matter of few values.

pH1	pH2	pH3	pH4
5,05	5,68	5,39	4,85
4,88	4,92	4,92	5,34
5,16	5,20	5,88	5,34
5,14	4,59	4,68	4,81
5,07	4,56	5,45	6,18



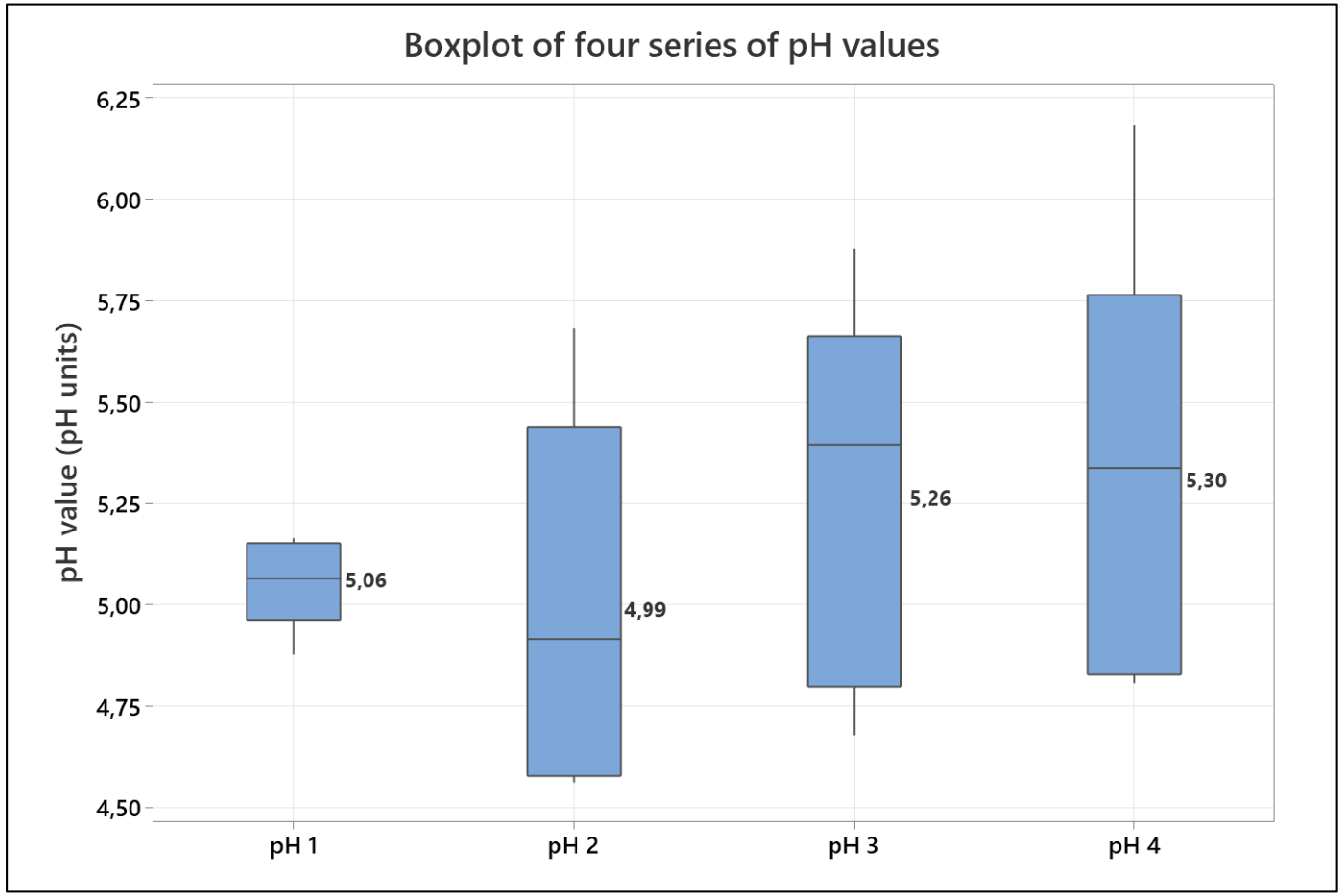
PHASE 1

In the case of multiple values, and in any case as a completely general approach, the

boxplot

is undoubtedly the best solution.

pH1	pH2	pH3	pH4
5,05	5,68	5,39	4,85
4,88	4,92	4,92	5,34
5,16	5,20	5,88	5,34
5,14	4,59	4,68	4,81
5,07	4,56	5,45	6,18



PHASE 1

1st Quartile, Q1: 25% of the data ≤ this value

Median, Q2: 50% of the data ≤ this value

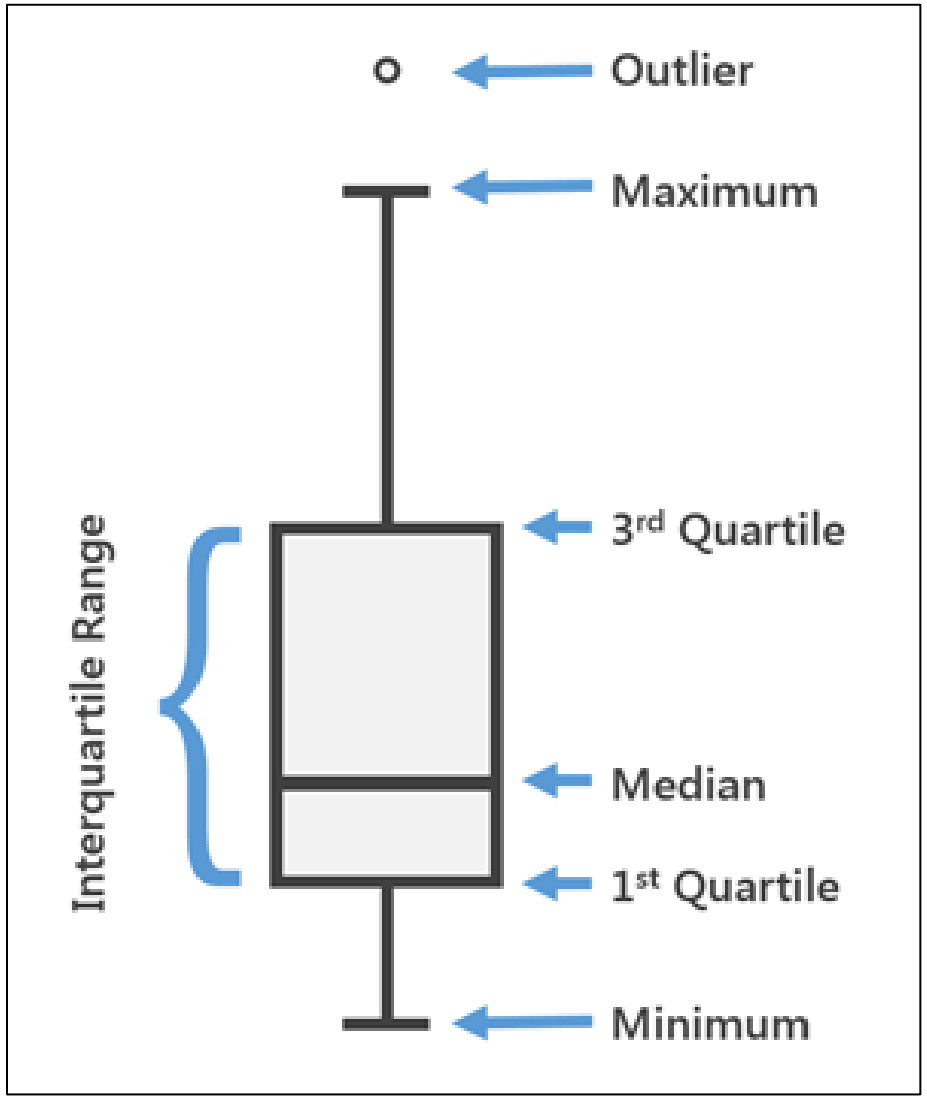
3rd Quartile, Q3: 75% of the data ≤ this value

Interquartile range: 50% of the data

Whiskers: extend to the minimum / maximum date point within 1.5 IQR from the bottom / top of the box

Outlier : observation beyond upper or lower whisker, *i.e.,* over 1.5IQR

J.W. Tukey, Exploratory Data Analysis, Addison Wesley, 1977



PHASE 1

POSITION INDICES: are summary indices that replace <u>all</u> values of a variable with a single value that can be considered "representative of all the others".

- **MODE:** is the value that appears most often in a data set.

  e.g.: 3, 3, 5, 6, **7, 7, 7**, 8, 8, 10 → *Mode = 7*

- **MEDIAN**: is the middle point in a dataset. ***Is a "robust" central trend indicator !***

  e.g.: 0, 0, 1, 1, **2**, 3, 3, 3, 4 → *Median = 2 (Mean = 1.89)*

- **ARITHMETIC MEAN** : can be considered the ***center of gravity*** of the dataset where the differences are balanced.

  e.g.: 3, 5, 10 → arithmetic mean: $\bar{x} = \frac{1}{3}(3 \times 1 + 5 \times 1 + 10 \times 1) = \frac{1}{3}(18) = 6$

PHASE 1

VARIABILITY INDICES :

The most commonly used are:

- ***Range*** : It is the simplest dispersion index. It is equal to the maximum value minus the minimum value
- ***Standard Deviation*** : measures the degree of dispersion of a dataset relative to the arithmetic mean
- ***Variance*** : is the square of standard deviation
- ***Coefficient of Variation*** : it allows you to compare the variability of two different data distributions

- **Range**
  - It is the simplest dispersion index.
  - It is equal to the maximum value minus the minimum value.

Range = **Maximum age — Minimum age** = 57 – 27 = **30**

- **Standard Deviation** – measures the degree of dispersion of a dataset relative to the arithmetic mean.

$$s = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}}$$

where: "n" is the number of elements forming the dataset

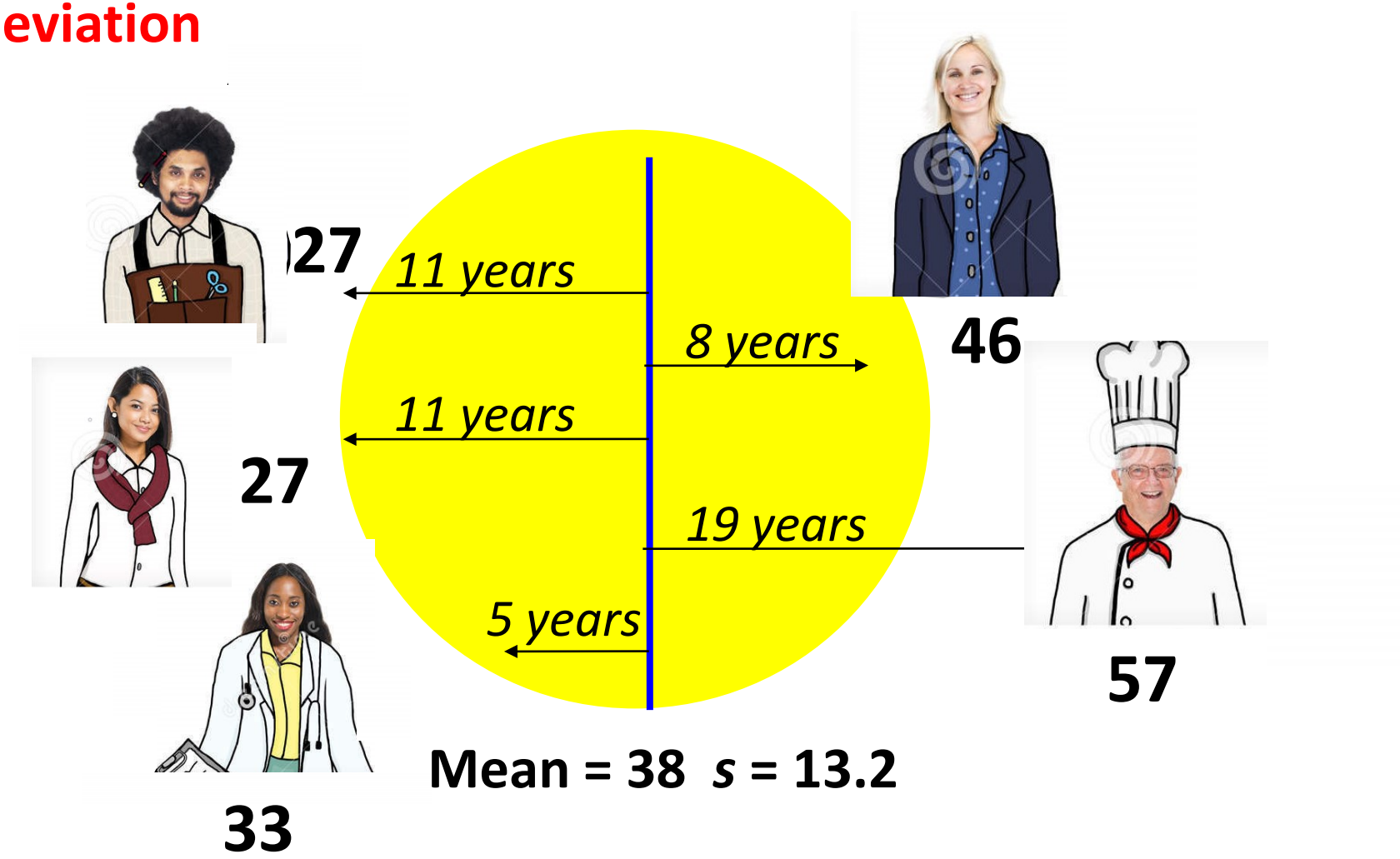
"X_i" is the value of each observation in the dataset

"$\bar{X}$" is the mean value of all observations forming the dataset

The standard deviation has the same units of measurement as the variable under study !

PHASE 1

Standard Deviation



PHASE 1

- While ***s*** refers to the sample, ***σ*** refers to the population.

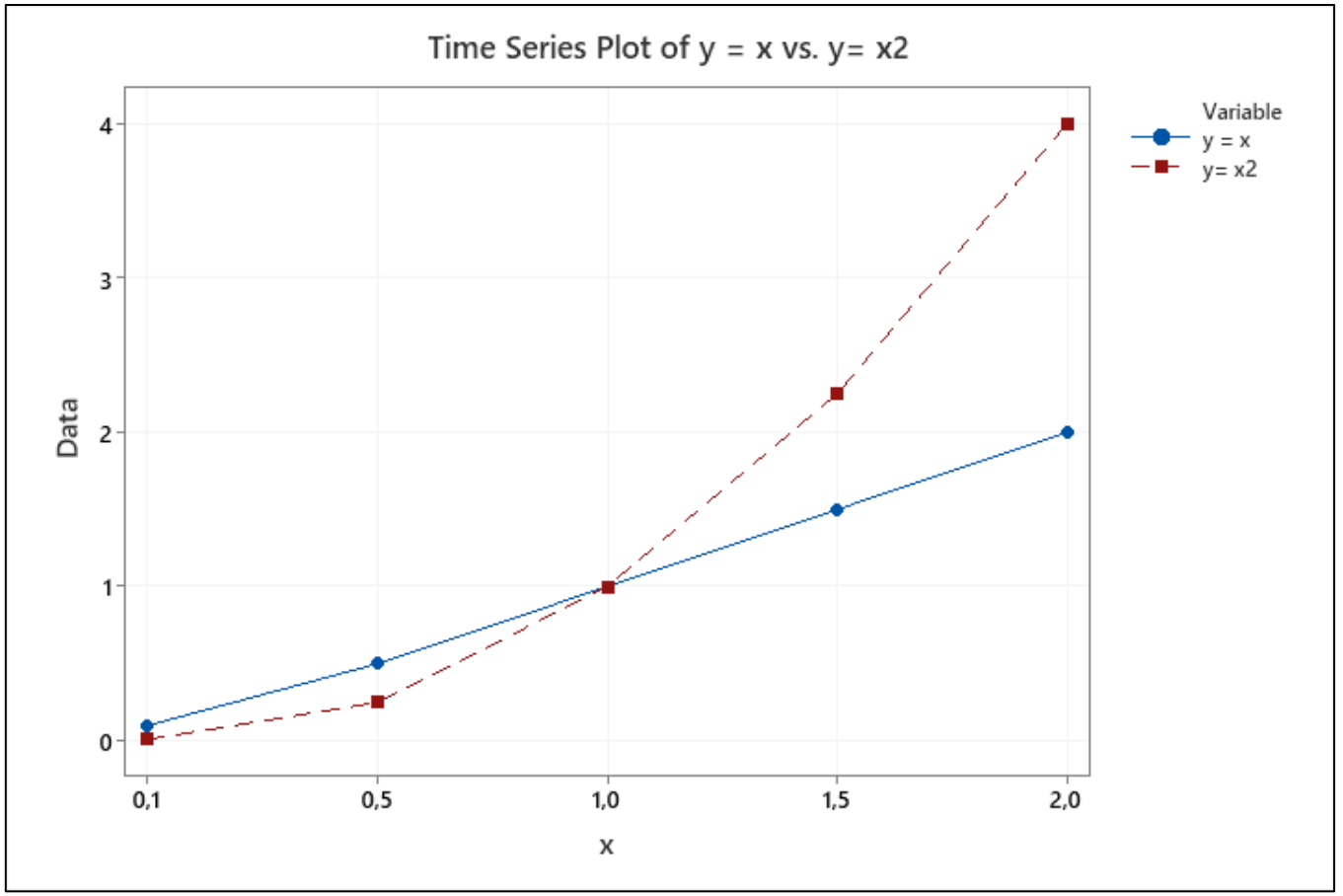
$$s = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n-1}} \qquad \sigma = \sqrt{\frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2}{n}}$$

- The reason for the difference between the two denominators is simply that if you divided by *n*, the standard deviation (or variance) of the sample would underestimate the standard deviation (or variance) of the population. That is, it would be a « *distorted statistic* ».

PHASE 1

IMPORTANT !!

In statistical indices such as: variance, standard deviation, etc. the differences (or deviations) are always squared not only because in this way they do not cancel each other out, but also because by squaring the small differences are "rewarded" and the large ones "penalized".



- **Variance** – is the square of standard deviation.

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

Where “n” is the number of the samples.

“X_i” is the value of each observation.

“$\bar{X}$” is the mean value of all the samples.

PHASE 1

The variance, unlike the standard deviation, has the ***property of additivity***. This means that if the elementary data form subgroups, then the total variance can be obtained as the sum of the variance "within groups" and the "variance between groups":

$$\sigma^2 = \sigma^2_{Within} + \sigma^2_{Between}$$

This « variance decomposition theorem » is the basis of the so-called

Analysis of Variance or ANOVA

PHASE 1

- The « between variance », $\sigma^2_{Between}$, or « variance of group means », measures how different the group means are from each other.

- The « within variance », σ^2_{Within} , or « mean of group variances », provides a summary of the level of variability present within each data group.

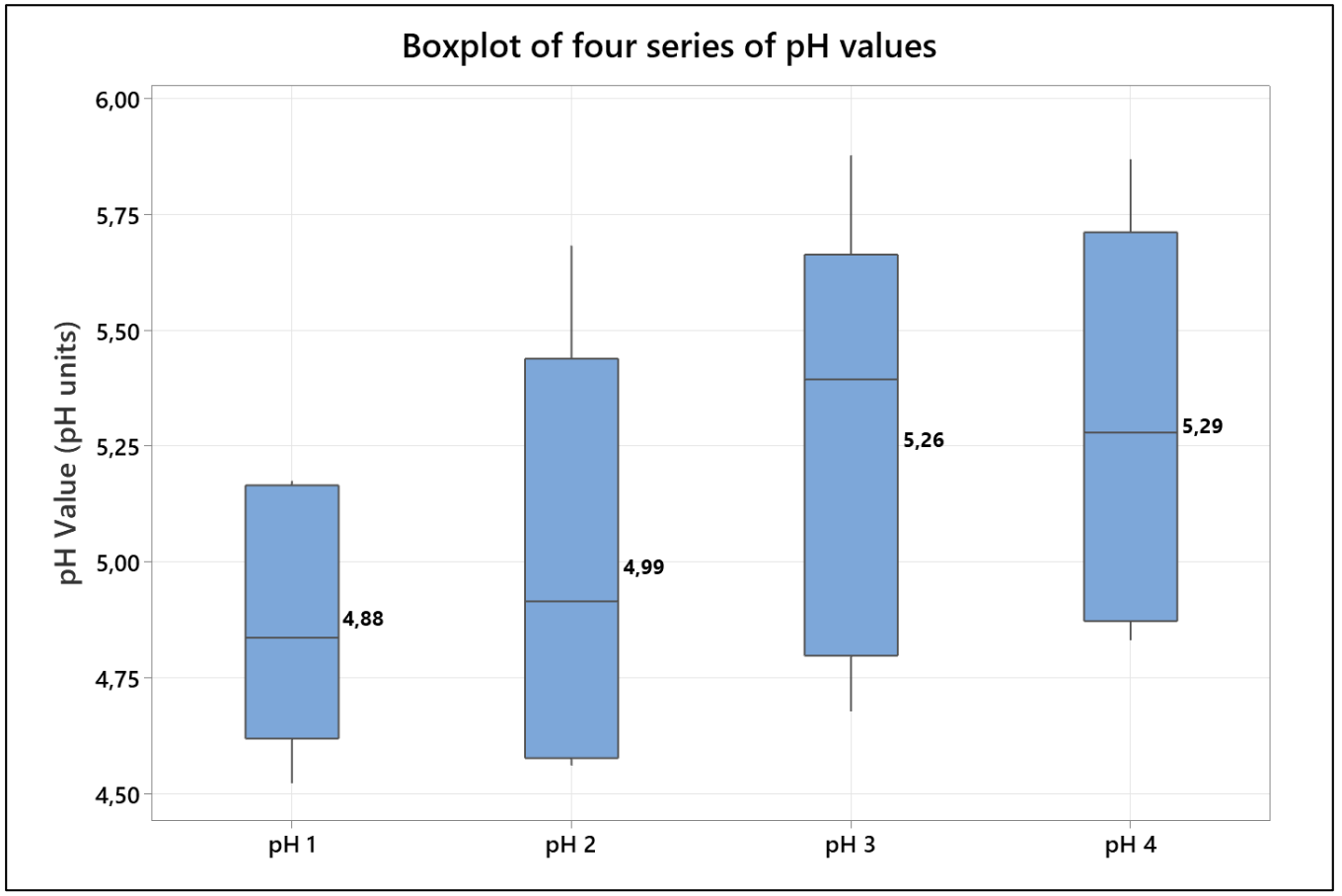
- In applying these criteria to regression analysis using the least squares method, the $\sigma^2_{Between}$ is called the ***explained variance*** while the σ^2_{Within} is called the ***residual variance***.

PHASE 1

Let us consider for example the four series of pH values and the corresponding boxplots seen previously…

are their means different or not?

pH1	pH2	pH3	pH4
5,18	5,68	5,39	5,87
5,16	4,92	4,92	4,91
4,72	5,20	5,88	5,28
4,52	4,59	4,68	4,83
4,84	4,56	5,45	5,55



PHASE 1

Let’s see ANOVA One-Way (or One factor) results:

SUMMARY						
Groups	**Count**	**Sum**	**Mean**	**Variance**	**Dev. Std.**	**CV%**
pH1	5	24,41	4,88	0,0799	0,2827	**5,79**
pH2	5	24,95	4,99	0,2171	0,4659	**9,34**
pH3	5	26,32	5,26	0,2225	0,4717	**8,96**
pH4	5	26,45	5,29	0,1890	0,4348	**8,22**

ANOVA						
Source of variation	**Sum of Squares**	***dof***	**Mean of Squares**	**F calc**	**P-value**	**F tab**
Between groups	0,6120	3	**0,20**	**1,15**	0,36	**3,24**
Within groups	2,8342	16	**0,18**			
Totale	3,4462	19				

What does ANOVA One-Way tell us?

In a few words:

- ***variability between data groups is practically comparable to that within them***
- ***average values of the data groups are not significantly different from each other***

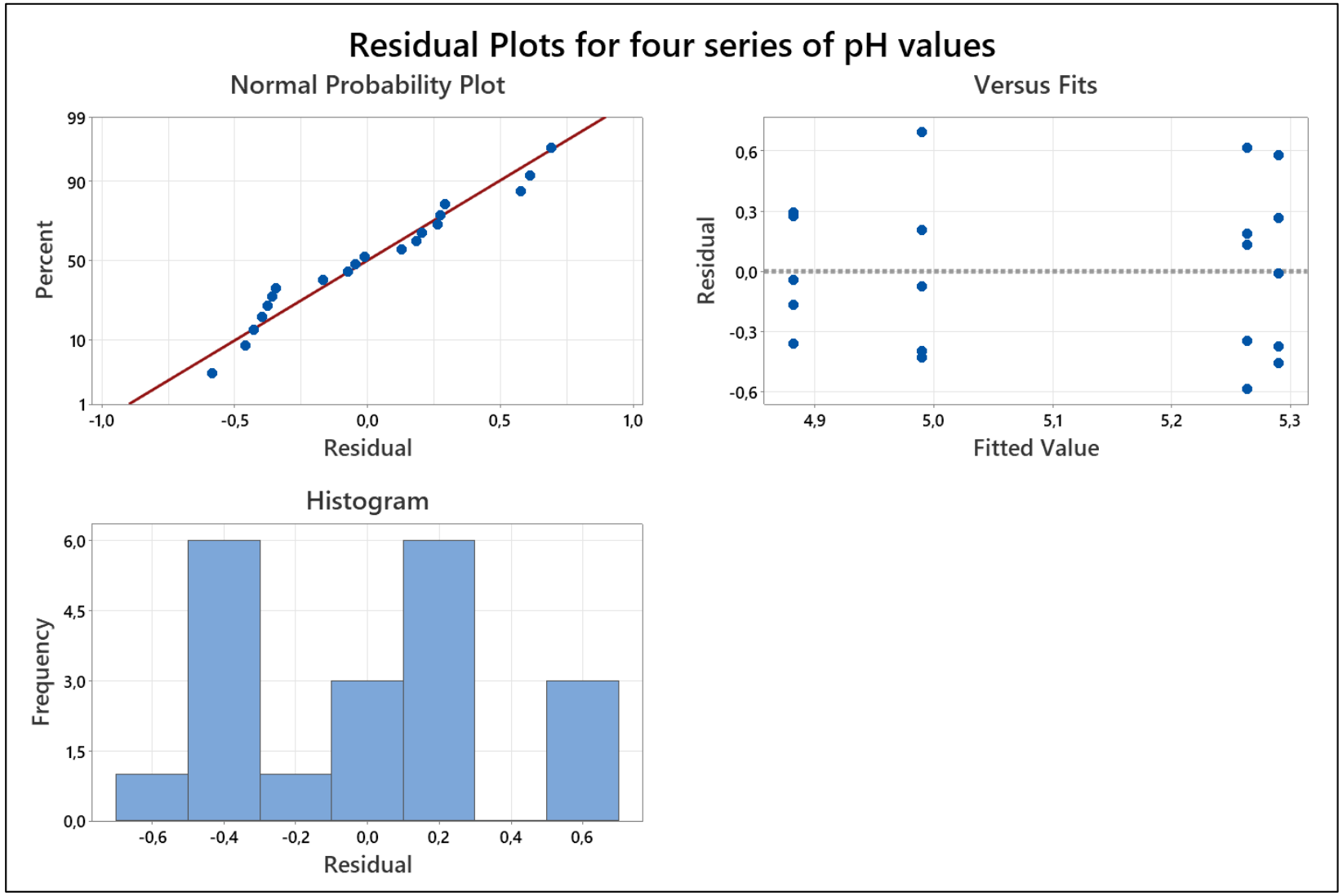
Remember that ANOVA requires that:

- ***Data are independently and randomly sampled from the reference population. This means that each observation must be independent of the others and that each observation must be chosen randomly from the population of interest***
- ***Variances of the data groups are similar (homoscedasticity)***
- ***Normality of residuals***

If these assumptions are heavily violated, it can lead to incorrect conclusions about the significance of the factor being tested !

PHASE 1

In the case under study, residuals are normally distributed and do not reveal any pattern



Remember that:

Even if ANOVA and Linear Regression are used for different purposes:

- ANOVA: comparison between the means of different groups
- Linear Regression: modeling the relationship between a dependent variable (y) and one or more independent variables (x_i)

both are based on linear models and share some underlying assumptions, including testing for residuals.

Types of Residual Plots

- **Histogram of Residuals**: used to test whether the residuals are normally distributed, a key assumption for ANOVA. A symmetrical, bell-shaped histogram indicates that the assumption of normality is reasonable.

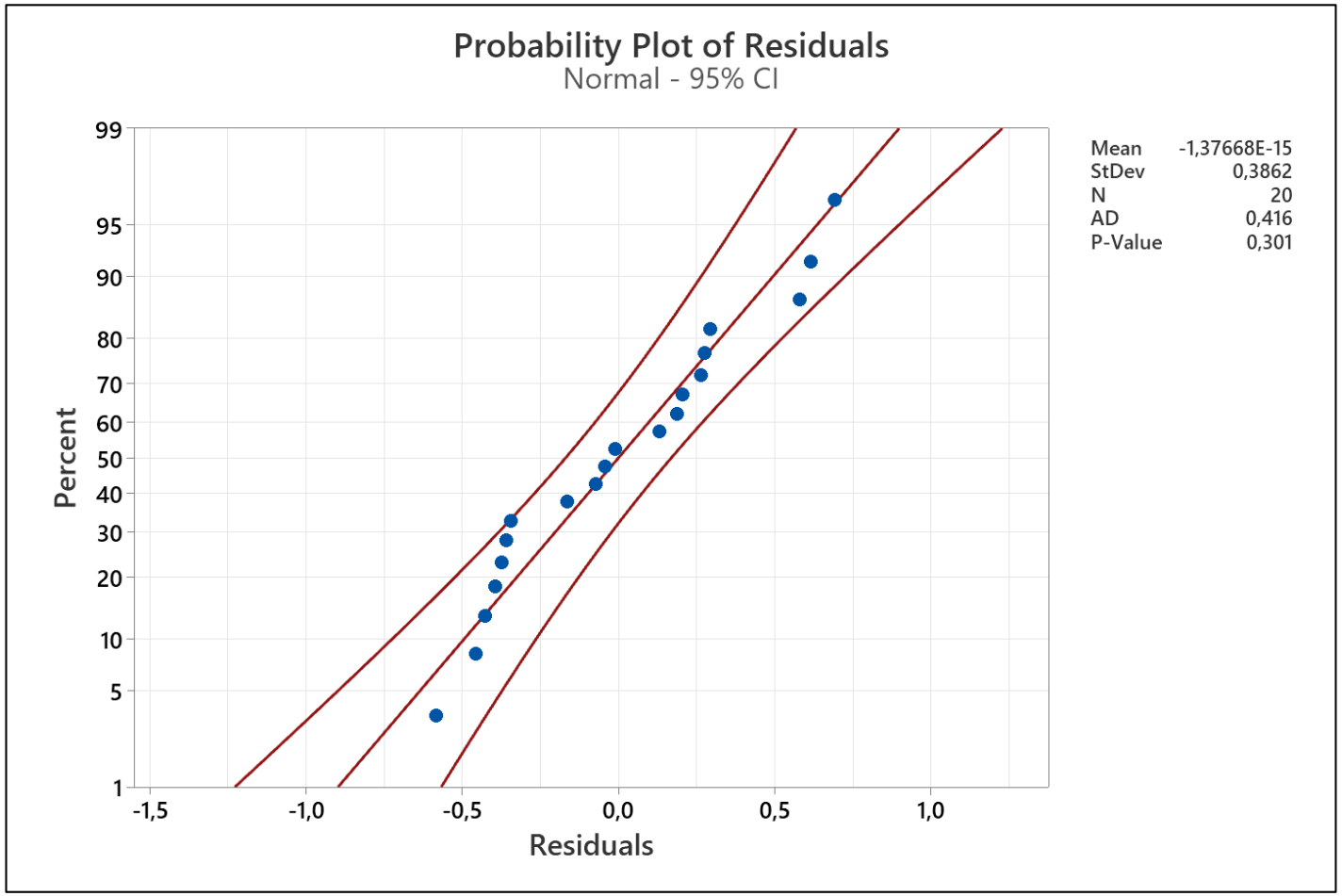
- **Scatter Plot of the Residuals *vs.* Fitted Values**: it is used to check homoscedasticity, *i.e.*, whether the variance of the residuals is constant for all levels of the factor. In a well-behaved graph, the points should form a random "cloud" around zero, with no obvious pattern.

- **Probability Plot:** it is also used to check the normality of the residuals. In a Probability Plot, the points should fall along a line inclined at 45 degrees if the data is normally distributed.

PHASE 1

The ***Probability Plot*** is a graphical technique for evaluating whether a set of data follows a given probability distribution such as, for example, the Normal.

The data is plotted against a theoretical distribution such that the points approximately form a straight line. Deviations from this line indicate deviations from the specified theoretical distribution.



J. Chambers, W. Cleveland ,B. Kleiner, P. Tukey, *Graphical Methods for Data Analysis*, Wadsworth (1983)

PHASE 1

Let’s go back to ANOVA: what could be the possible applications?

Comparison of multiple (> 2) data series such as:

- Yields of different lots obtained using the same process or different processes
- Assay, pH, LOD, *etc.* values relating to different production years for the same product (APQR)
- *etc.*

PHASE 1

Coefficient of Variation – is defined as:

$$\mathrm{CV} = \mathrm{RSD} = \frac{\sigma}{\mu} \quad or, alternatively \quad \mathrm{CV}\% = \mathrm{RSD}\% = \frac{\sigma}{\mu} \times 100$$

The usefulness of this index derives from the fact that it allows you to ***compare the variability*** of two different data distributions, *e.g.*, yields of two processes (or of the same process but conducted in different conditions / places), *etc.*

PHASE 1

Example:

Yield Process A (%): 99.8 100.1 100.0 100.7 99.7 100.0 100.2 100.7 98.8

Mean: 100.0 *s* = 0.52 CV% = 0.52%

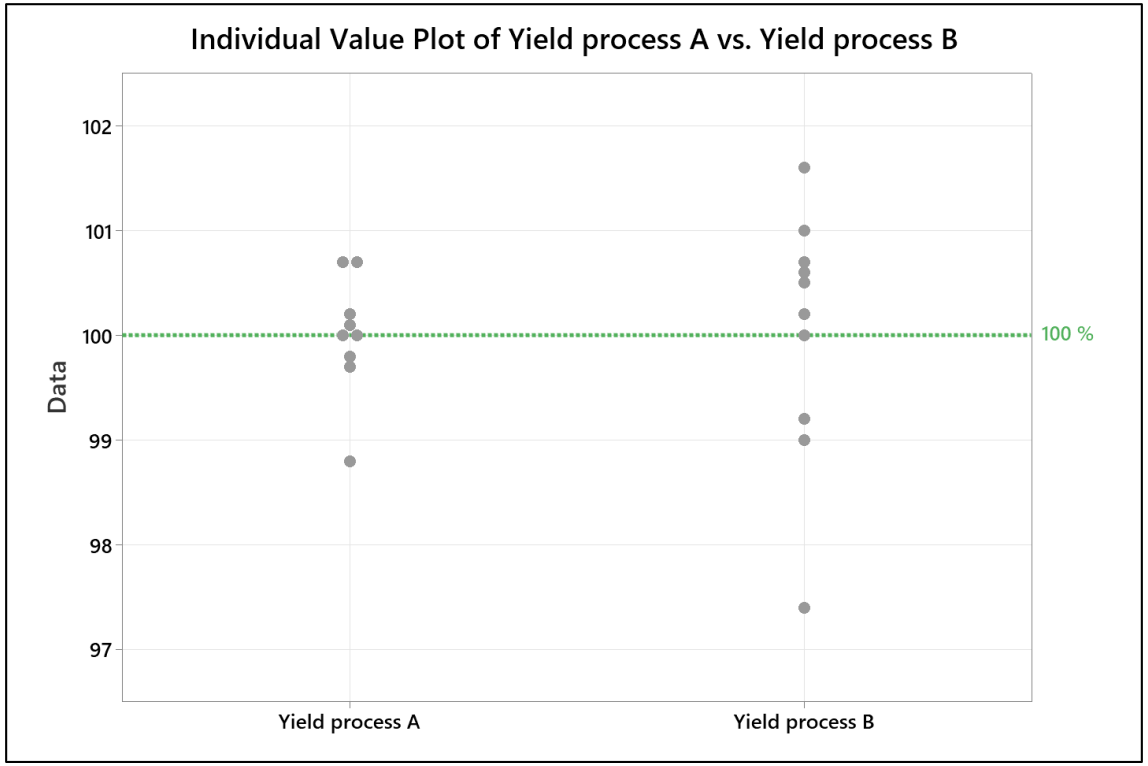
Yield Process B (%): 97.4 99.2 101.0 101.6 99.0 100.2 100.6 100.7 100.0 100.5

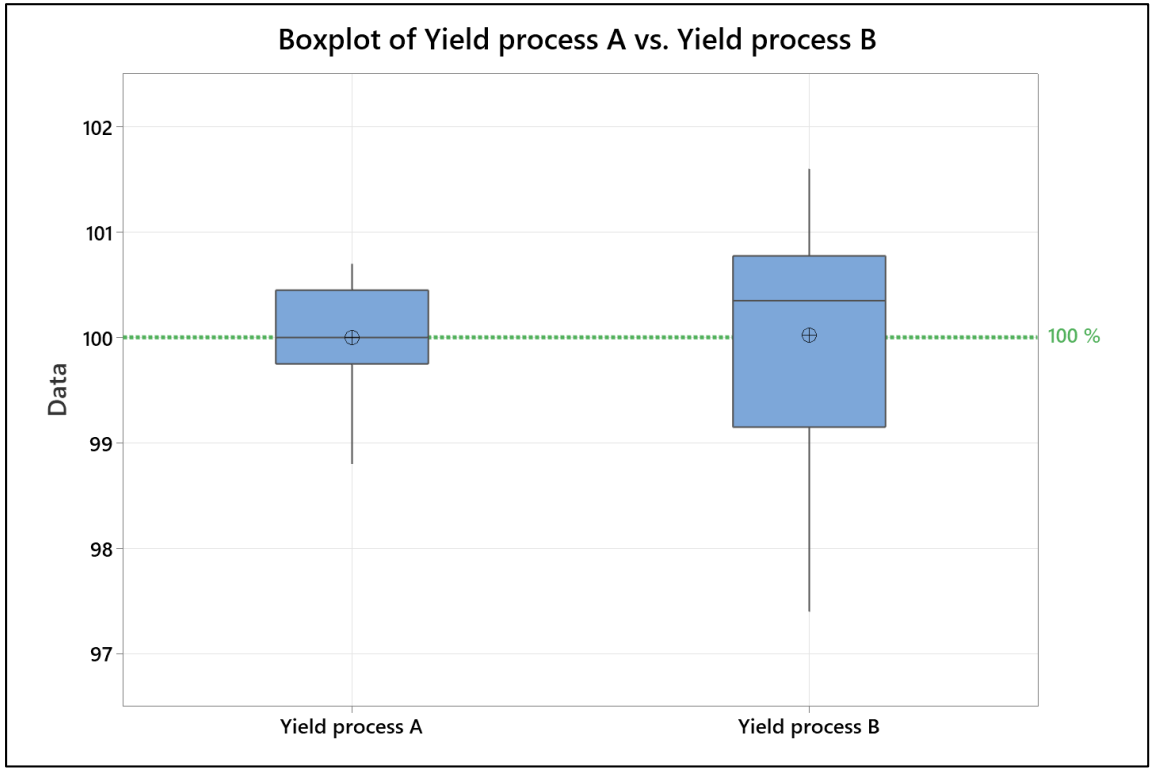
Mean: 100.0 *s* = 1.20 CV% = 1.20%

Conclusion: The yield of both processes is, on the average, equal to 100.0%, but process B is more variable.

ATTENTION: In this case, since the means are equal, the *s* value sufficed, but if the means are different from each other, the CV is needed.

PHASE 1





The difference in variability is also evident graphically !

PHASE 1

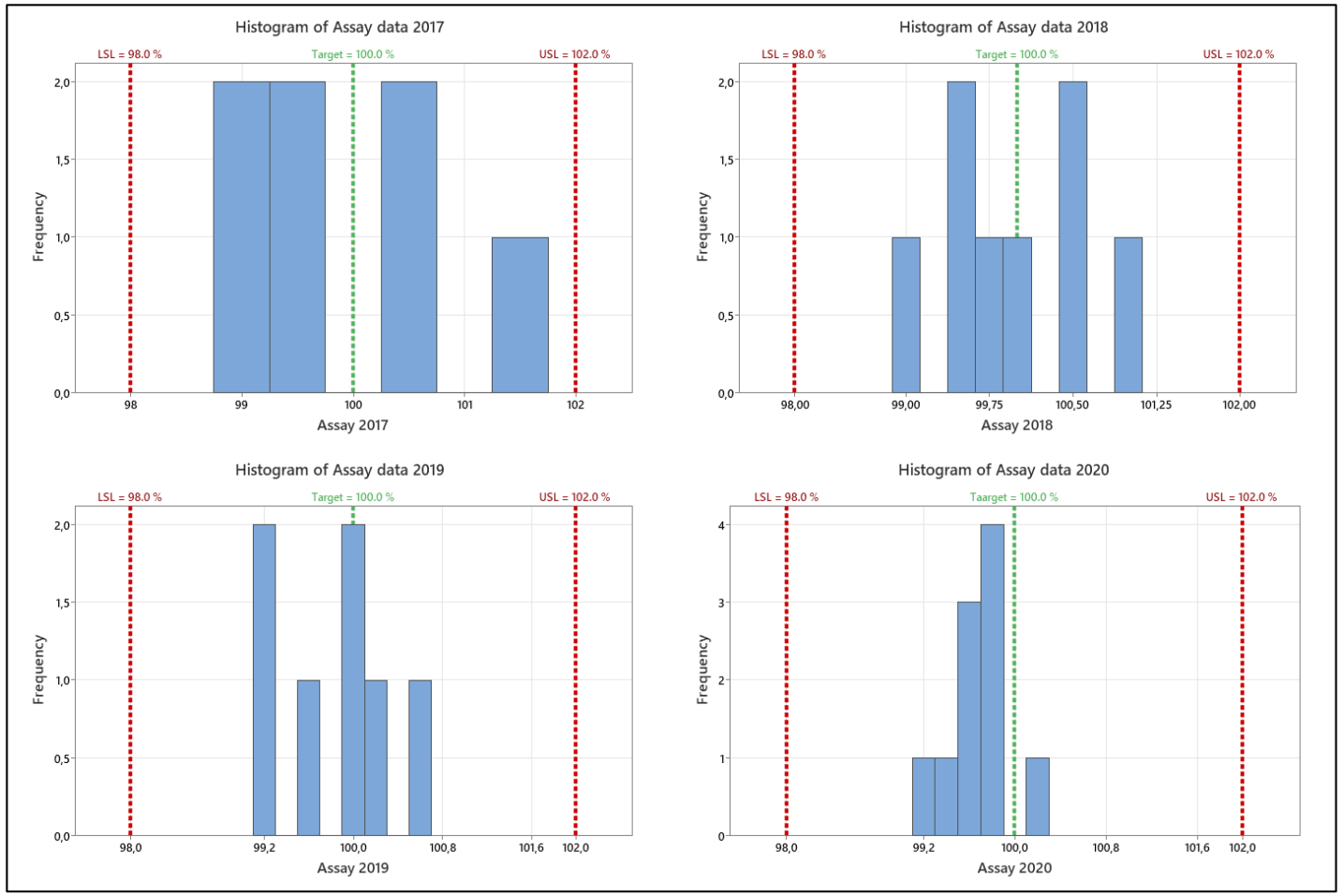
Now that we have seen a few statistical tools useful for investigating the variability of a process, let's see the case (theoretical) of the assay percent values of a manufacturing process over four years : *a typical APQR comparison !*

Assay 2017	Assay 2018	Assay 2019	Assay 2020
99,48	99,86	99,22	99,8
100,61	98,95	99,28	99,87
101,29	99,42	99,58	100,24
99,19	100,47	100,17	99,86
100,54	99,53	99,96	99,66
99,08	99,99	99,97	99,84
99,35	101,02	100,59	99,53
	100,58		99,27
			99,67
			99,35

PHASE 1

It is clear from the histograms on the side that this graphical tool is not so useful for the purposes of a practical study of the variability of the process over time.

What we can deduce from these graphs is only that the experimental data was within the specification limits which is obvious since we are dealing with released lots.



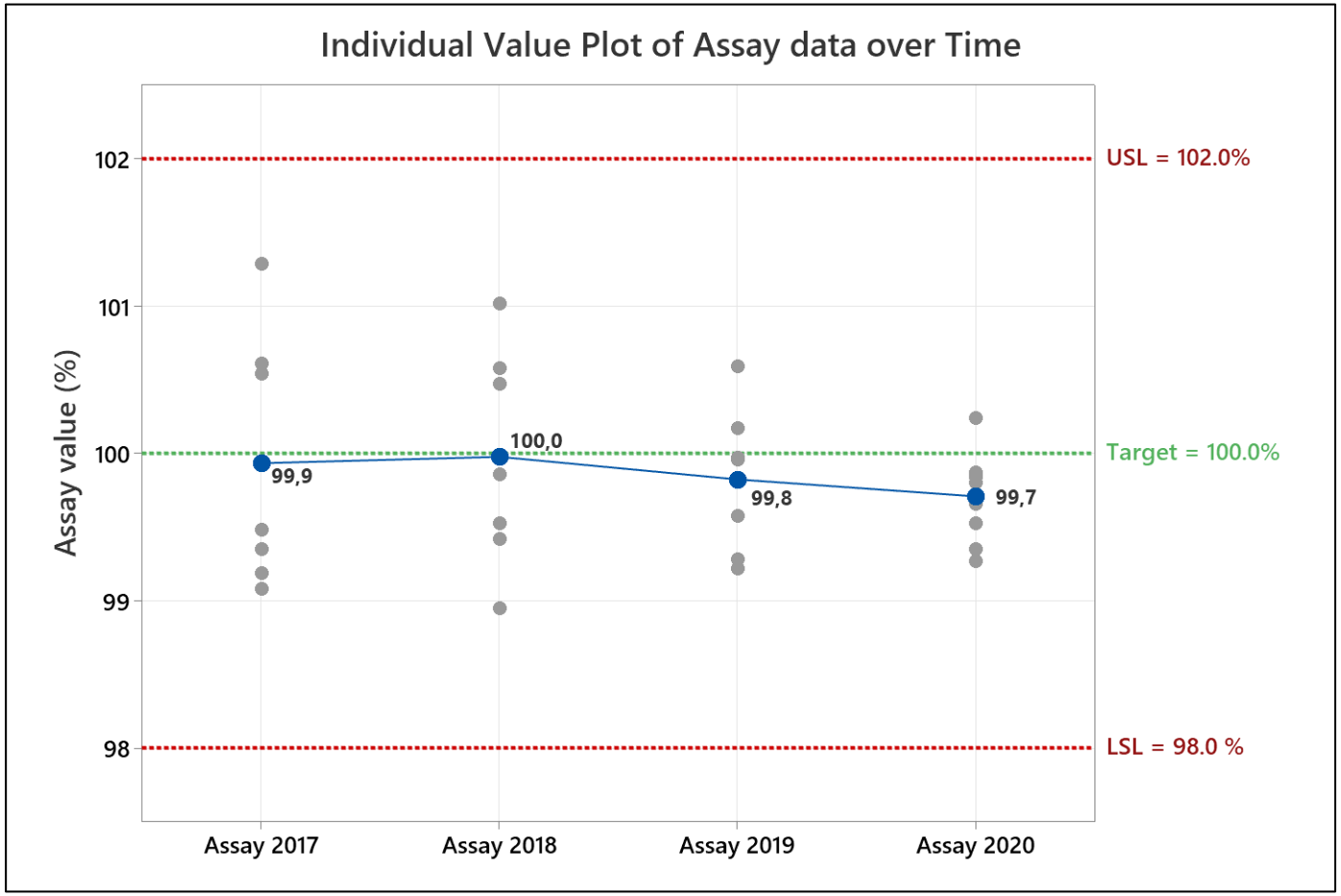
PHASE 1

The feedback improves using the

individual plot

From the graph to the side, it is evident that the data relating to the four years are characterized by different variability and that it decreases over time.

The average assay value for the different datasets remains close to 100%.



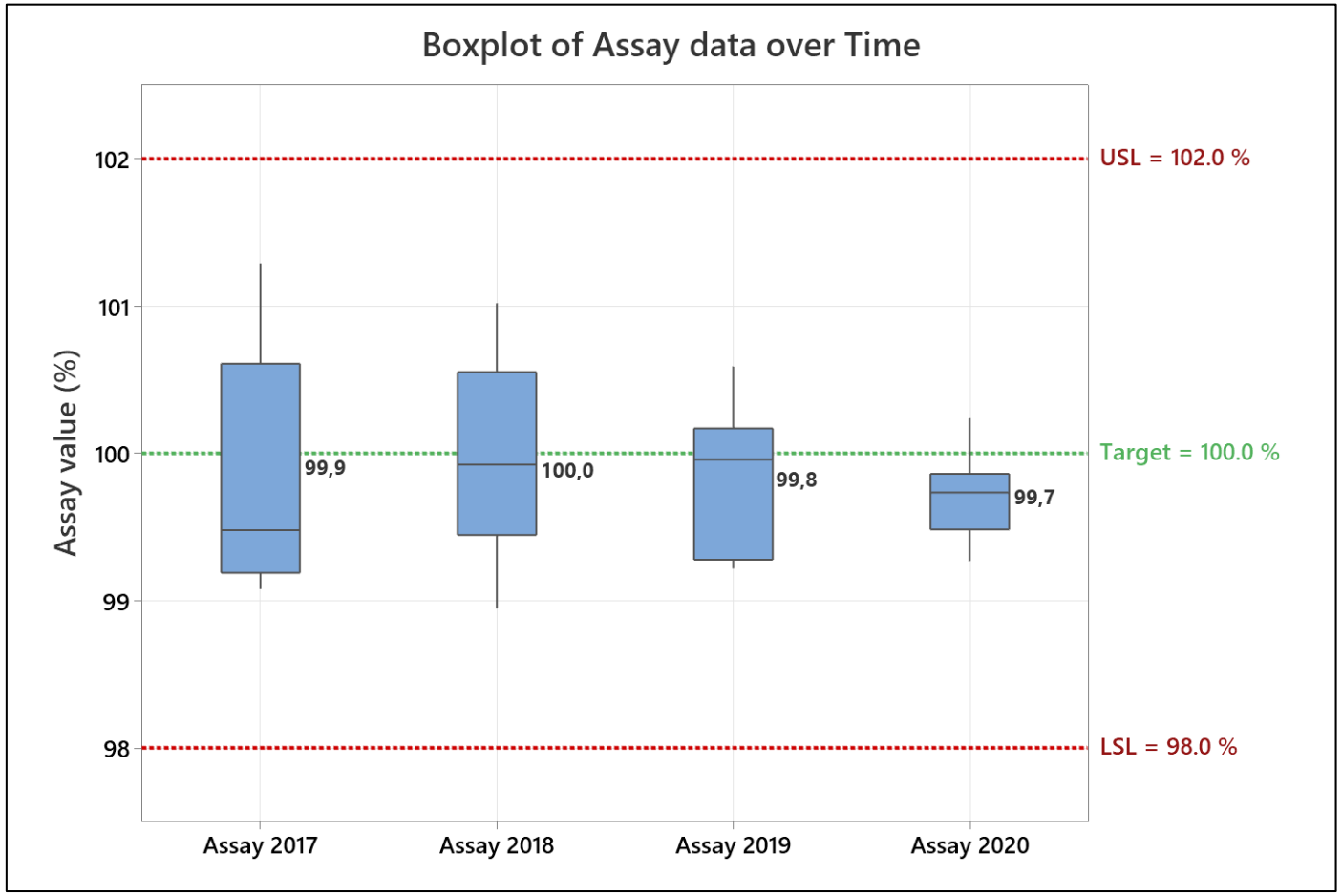
PHASE 1

The feedback is even clearer with

boxplots

The decrease over time of the variability (size of the boxes) is evident, as is the slight mean decentralization with respect to the ideal target of 100%, but

are these differences significant or just random?



PHASE 1

ANOVA One-Way (or One factor) can answer the question !

SUMMARY						
Groups	Count	Sum	Average	Variance	Std. Dev.	***CV%***
Assay 2017	7	699,54	99,9343	0,7488	0,8653	**0,87**
Assay 2018	8	799,82	99,9775	0,4676	0,6838	**0,68**
Assay 2019	7	698,77	99,8243	0,2444	0,4944	**0,50**
Assay 2020	10	997,09	99,7090	0,0793	0,2816	**0,28**

ANOVA						
Source of Variation	**Sum of Squares**	**dof**	**Mean of Squares**	**F calc**	**P-value**	**F crit**
Between Groups	0,3832	3	**0,13**	**0,36**	0,78	**2,95**
Within Groups	9,9464	28	**0,36**			
Total	10,3296	31				

What does ANOVA One-Way tell us?

In a few words:

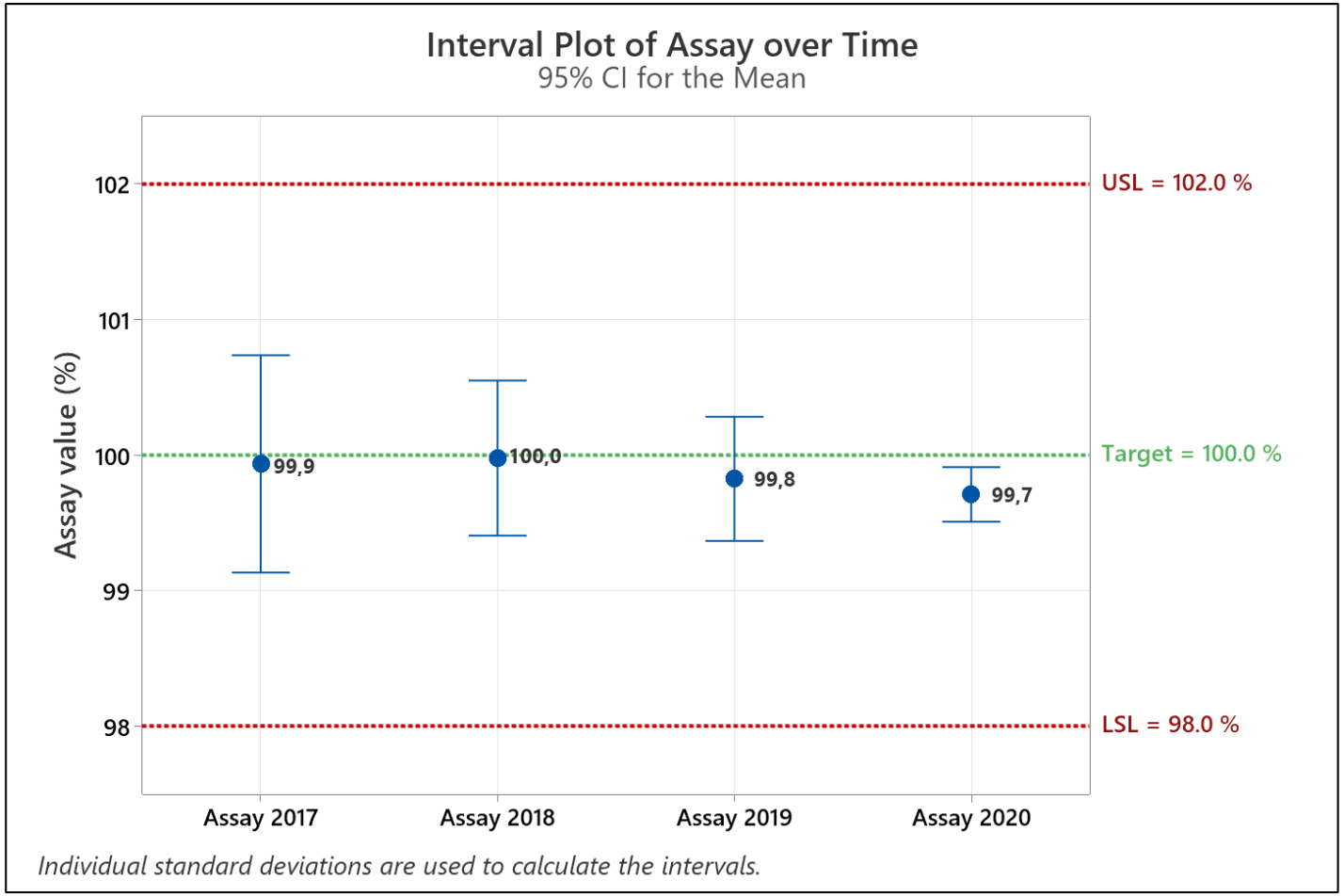
- ***variability is smaller between data groups than within them.***
- ***average assay values of the data groups are not significantly different from each other***

PHASE 1

Another useful graphical tool is the

Interval Plot

which is shown here.



What does the INTERVAL PLOT tell us?

- An interval plot typically presents the Confidence Intervals for the means of different data groups.
- The width of the interval reflects the uncertainty in the estimate of the mean, and it is influenced by the sample size and variability within each data group.

PHASE 1

Since:

- the interval plot shows a 95% Confidence Interval for the mean of each data group, and
- in this case, all interval plots well overlaps with each other and are within specification

this indicates that:

  - the group means are consistent with the specifications
  - the differences between the means are not significant as also emerged from the ANOVA
  - based on the data under study, the ***true mean of the process***, which is unknown by definition, is expected to be within specifications.

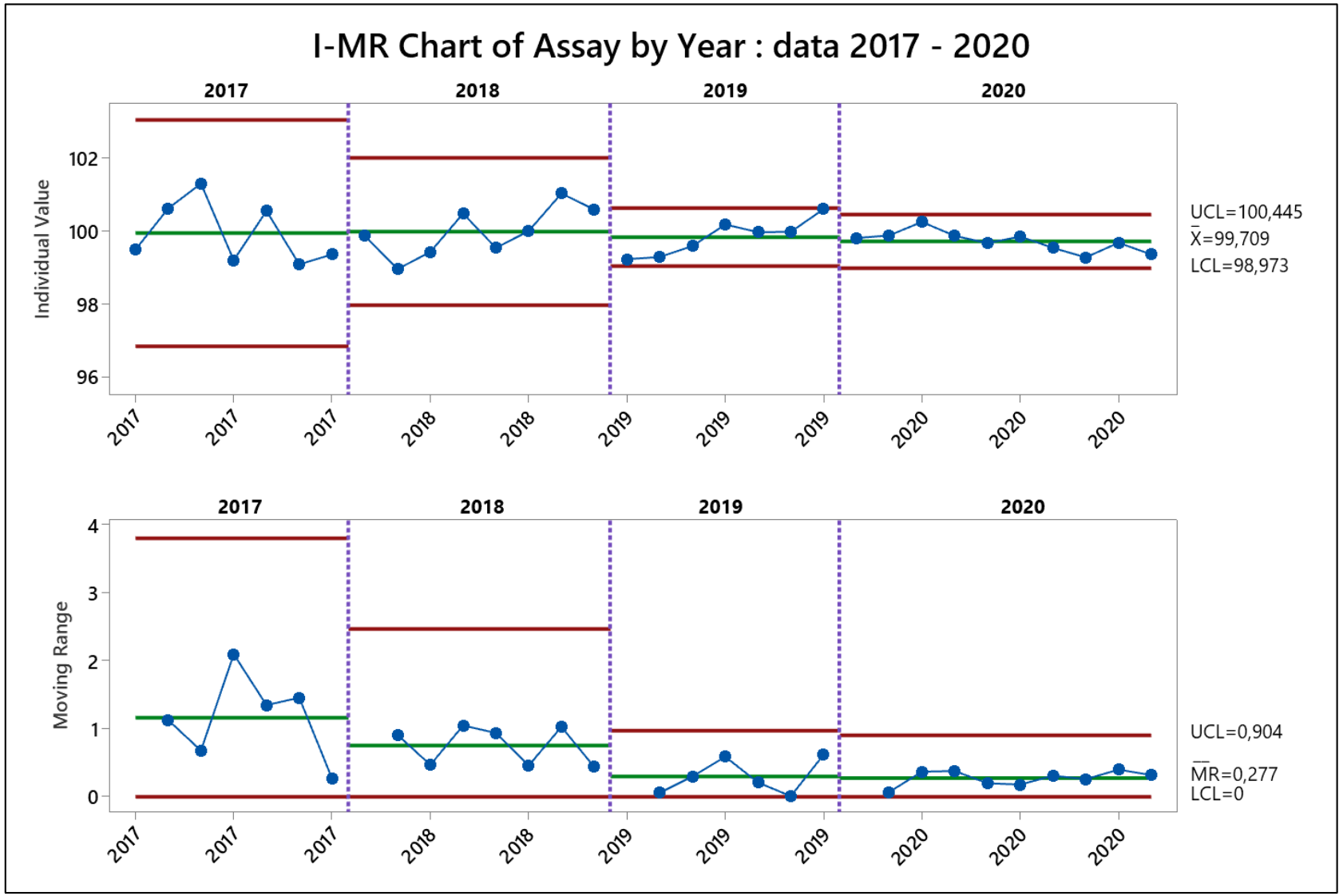
PHASE 1

- Alongside those just seen, another powerful statistical tool for studying process variability is represented by the ***Shewhart control charts*** which have nothing to do with the "famous average $\pm 3\sigma$ graph".

- Given the low overall number of batches produced, quite frequent in the chemical-pharmaceutical field, ***individual control charts*** are generally the most suitable in Phase 1.

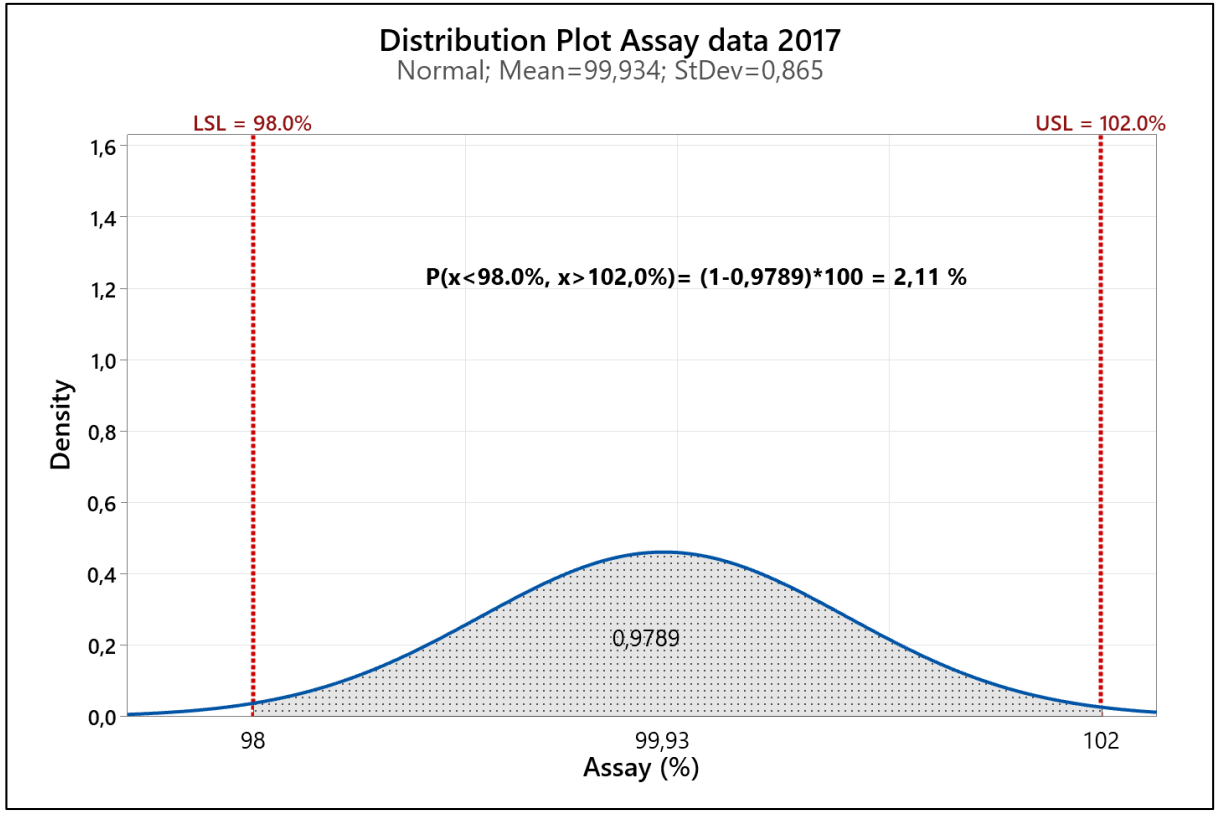
- In the following slide is the *I-MR chart* (Individual-Moving Range chart) relating to the study data divided by year (or *stage control charts*).

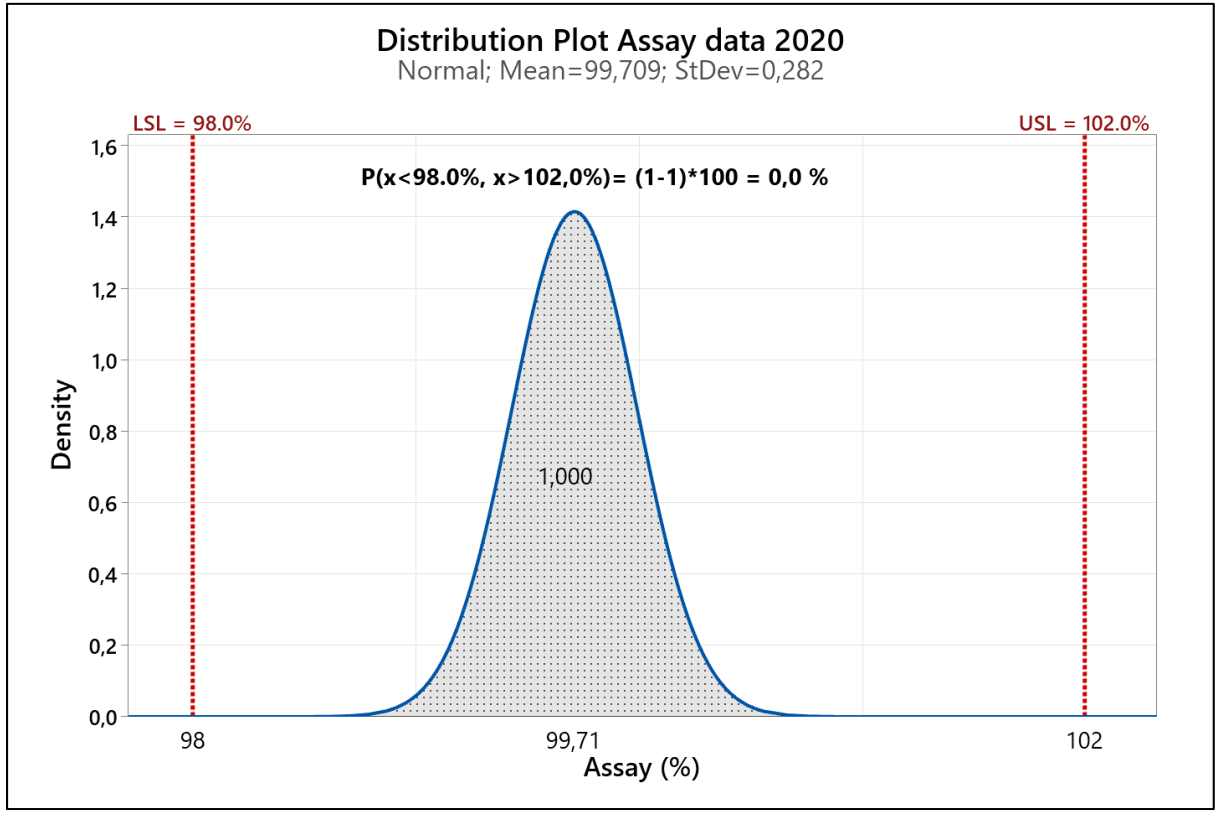
PHASE 1


I-MR Chart of Assay by Year : data 2017 - 2020
2017
2018
2019
2020
Individual Value
102
100
98
96
UCL=100,445
X̄=99,709
LCL=98,973
Moving Range
4
3
2
1
0
UCL=0,904
MR̄=0,277
LCL=0

PHASE 1

The process improvement from 2017 to 2020 seen from a *probabilistic point of view* is even more relevant !



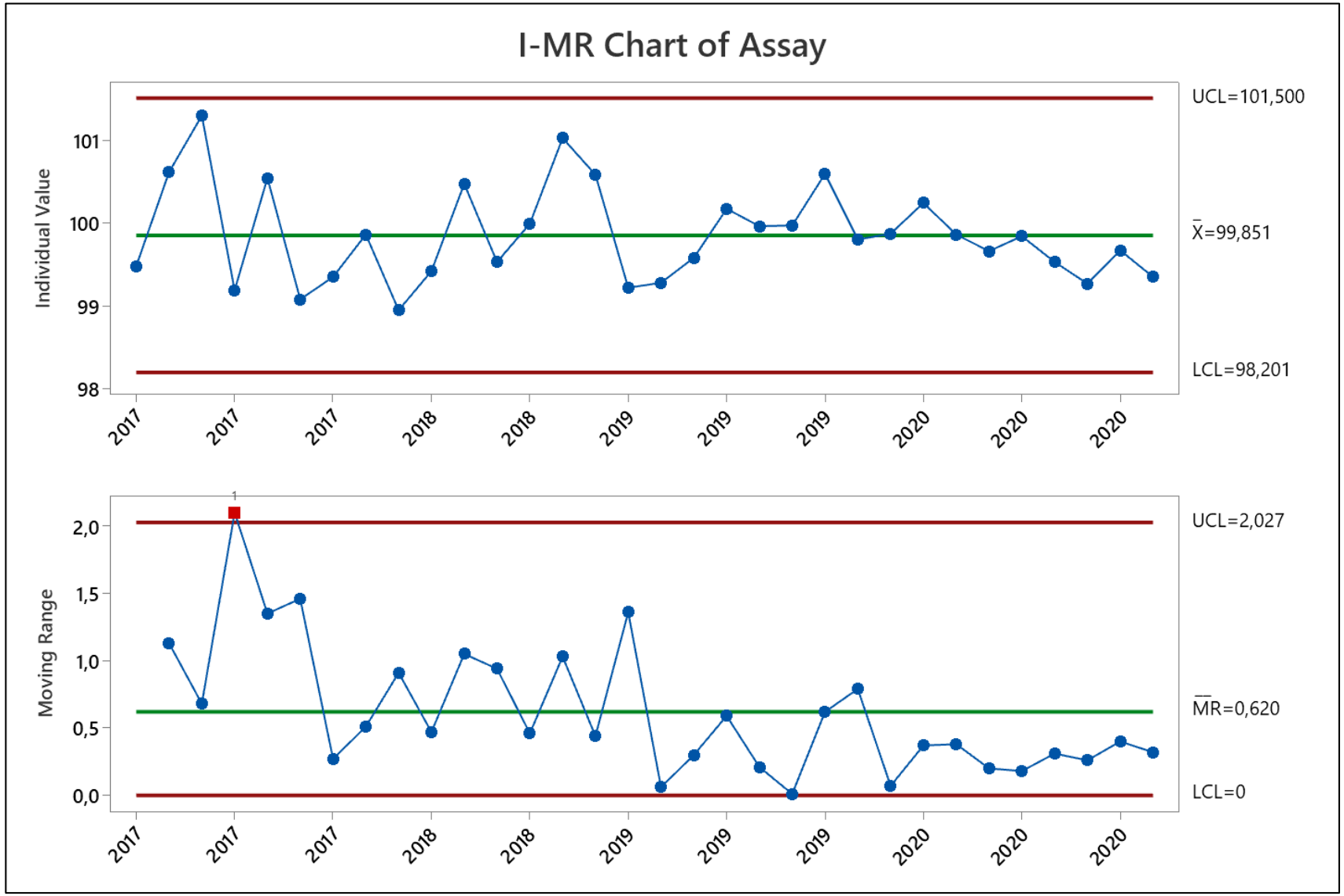


What does the *I-MR Chart* tell us?

- ***the process is under control in the time span considered*** as the limits of the chart (which are not specification limits, but control limits calculated taking into account the variability in the data) are never exceeded.
- ***the variability of the process has been decreasing over the years***.
- ***based on these results, it will therefore be possible to establish control limits to be adopted for future monitoring of the process.***
- ***it is evident that if all the historical data were considered together, the interval would be wider, and close to the "famous mean $\pm 3\sigma$" (see next slide).***

PHASE 1

It is also evident that all the considerations just made considering each year would have been impossible if the limits had been calculated on the whole dataset as shown here alongside !

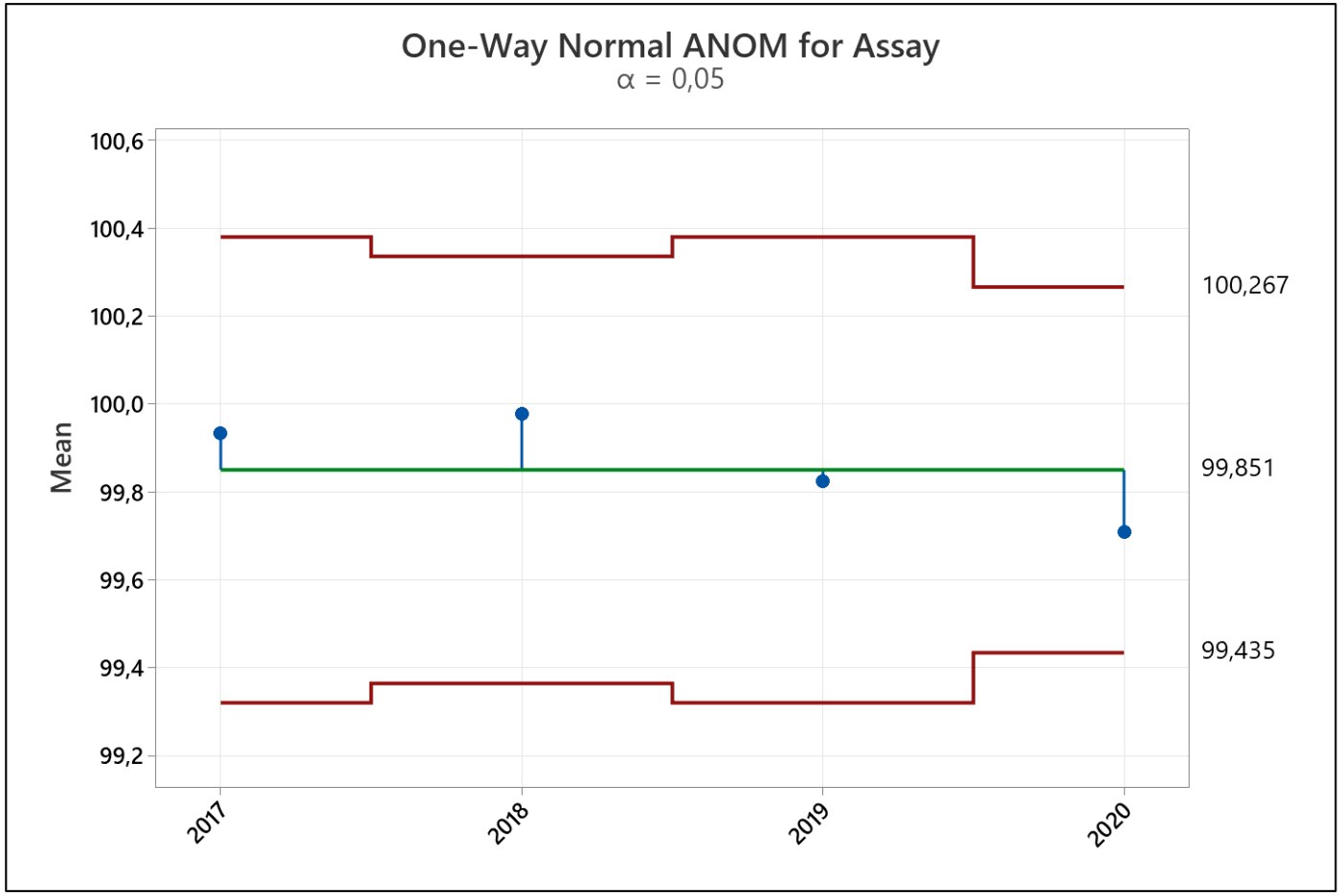


PHASE 1

Another useful graphical tool is the

Analysis of Means (ANOM)

a graphical procedure for comparing a collection of means, which is shown here.



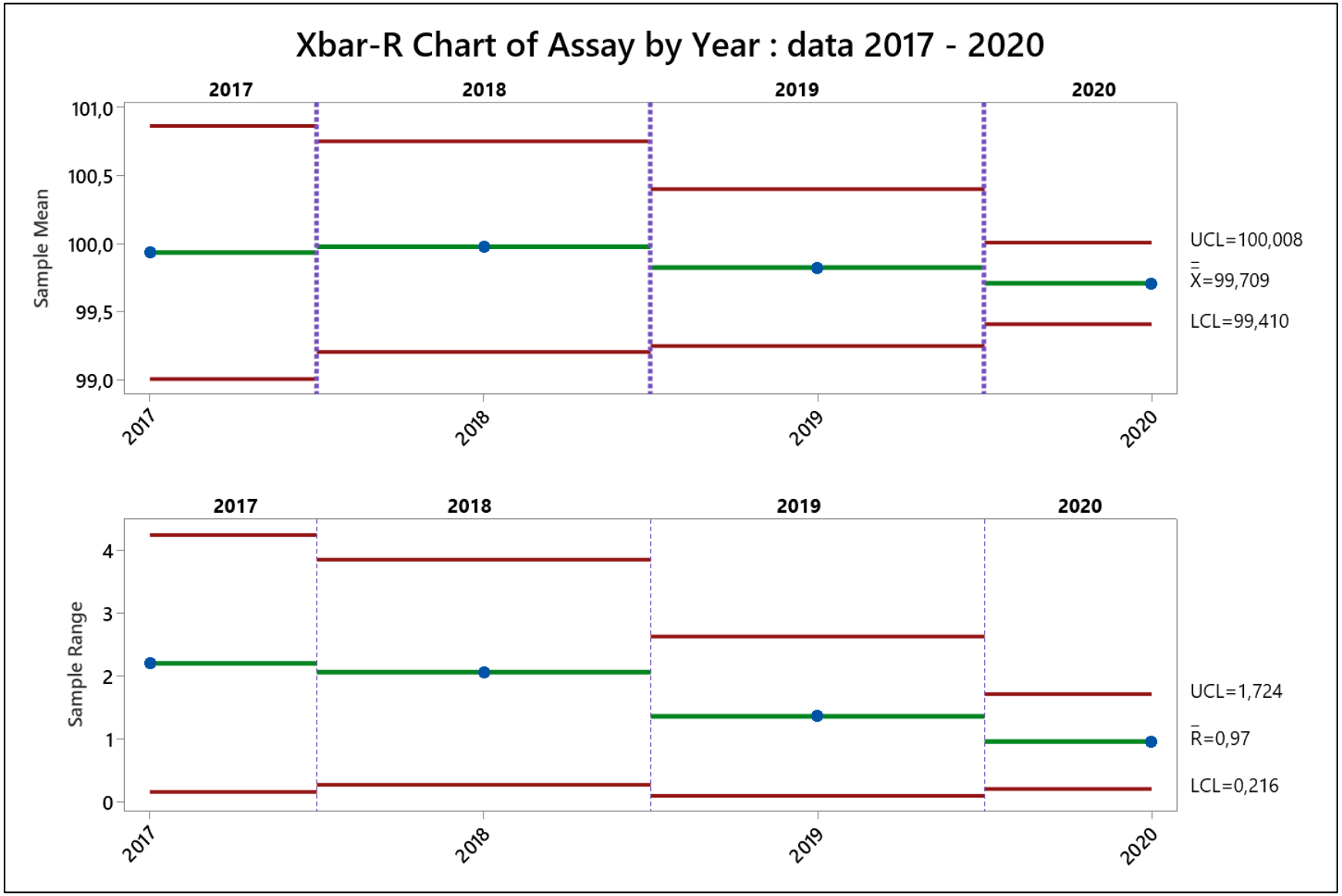
What does ANOM tell us?

- **ANOM** is a ***decision chart*** similar in appearance to the control charts seen so far
- It has in fact a centerline located at the overall mean and upper and lower decision limits
- The group means are plotted and since, in this case, none of them fall beyond the decision limits they are said to not differ significantly from the overall value.
- This result is in line with the comparison of means resulting from ANOVA and seen previously.

PHASE 1

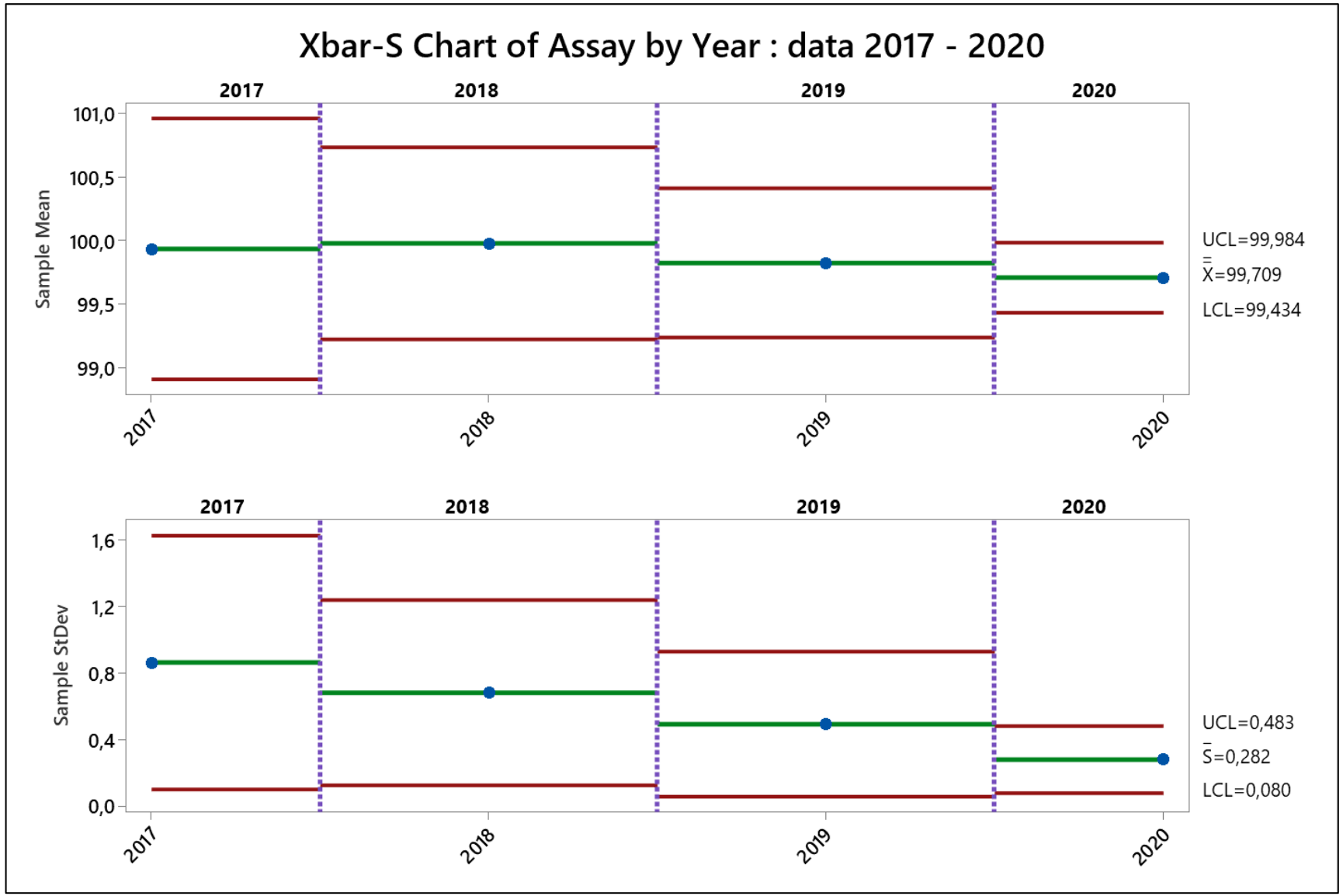
Alongside the *I-MR chart* seen previously, there are other Shewhart control charts, such as the *Xbar-R chart* shown alongside.

In this case, however, it does not provide additional information.



PHASE 1

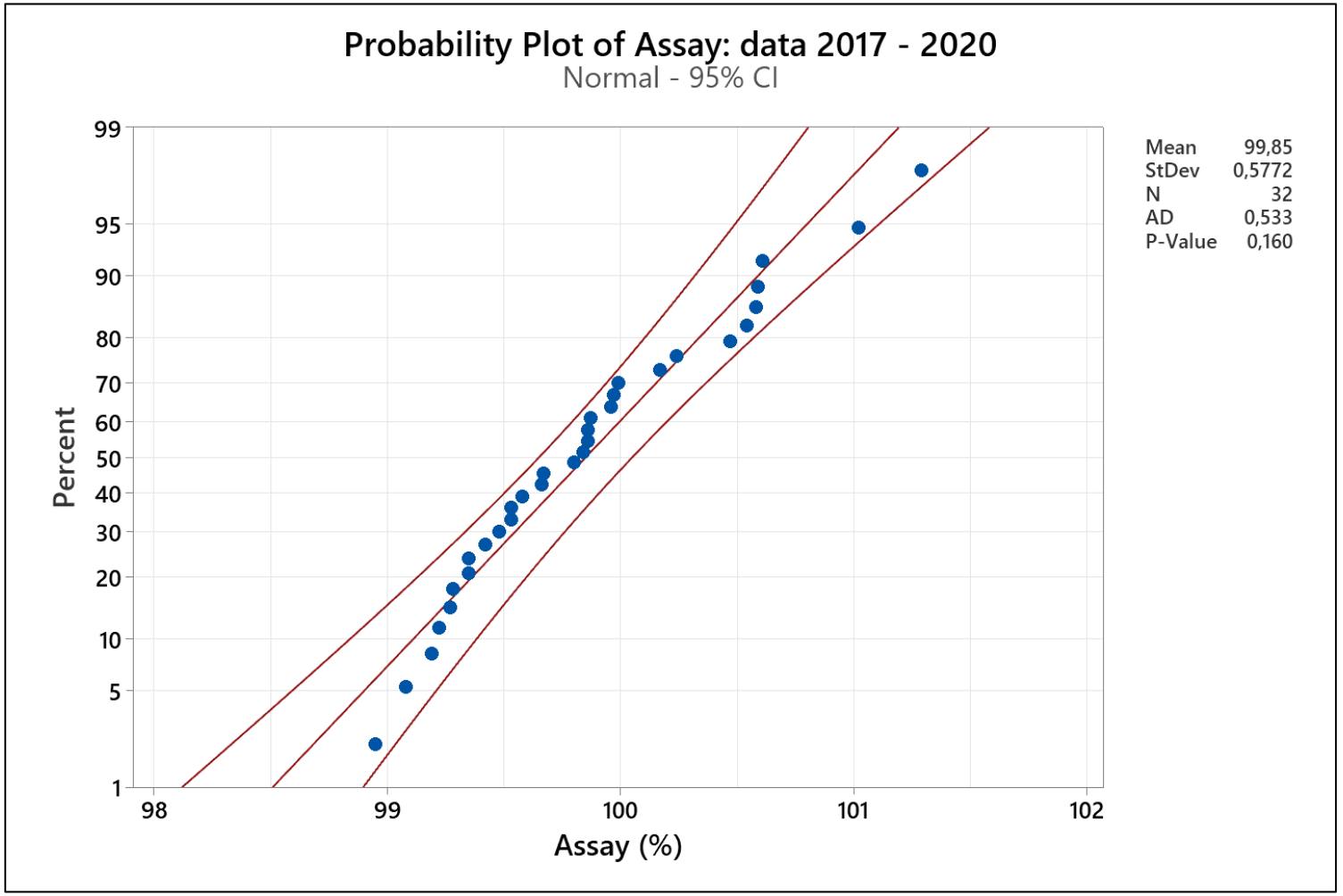
The same observation made in the previous slide for the *Xbar-R* chart also applies to the *Xbar-S chart* shown alongside.



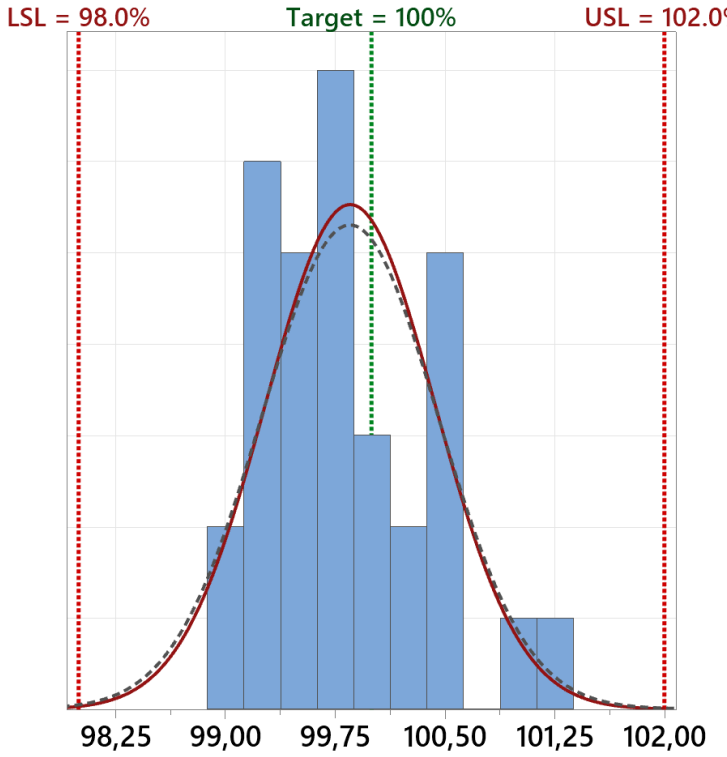
PHASE 1

Since the data examined as a whole return the image of a stable process, *i.e.*, “under control” or "predictable“, and moreover they are normally distributed (see the Normal Probability Plot opposite), it makes sense to conduct a

Process Capability Analysis



Process Capability Report for Assay : data 2017 - 2020



Process Data	
LSL	98
Target	100
USL	102
Sample Mean	99,8506
Sample N	32
StDev(Overall)	0,577246
StDev(Within)	0,601354

Overall
Within

Overall Capability	
Pp	1,15
PPL	1,07
PPU	1,24
Ppk	1,07
Cpm	1,13

Potential (Within) Capability	
Cp	1,11
CPL	1,03
CPU	1,19
Cpk	1,03

Performance	Observed	Expected Overall	Expected Within
% < LSL	0,00	0,07	0,10
% > USL	0,00	0,01	0,02
% Total	0,00	0,08	0,12

The actual process spread is represented by 6 sigma.

What does the *Process Capability Analysis* tell us?

In a few words:

- The ***Capability Ratio***, ***Cp***, is equal to **1.11** which can be considered adequate for batch chemical-pharmaceutical processes.
- The ***Centered Capability Ratio***, ***Cpk***, is equal to **1.03** since the process is slightly off center, but it can be considered acceptable.
- The ***Performance Ratio***, ***Pp***, is equal to **1.15** which is close to Cp = 1.11 and indicates that the process is *operated predictably*.
- The ***Centered Performance Ratio***, ***Ppk***, is equal to **1.07** as the process is slightly off center, but, again, it can be considered acceptable.

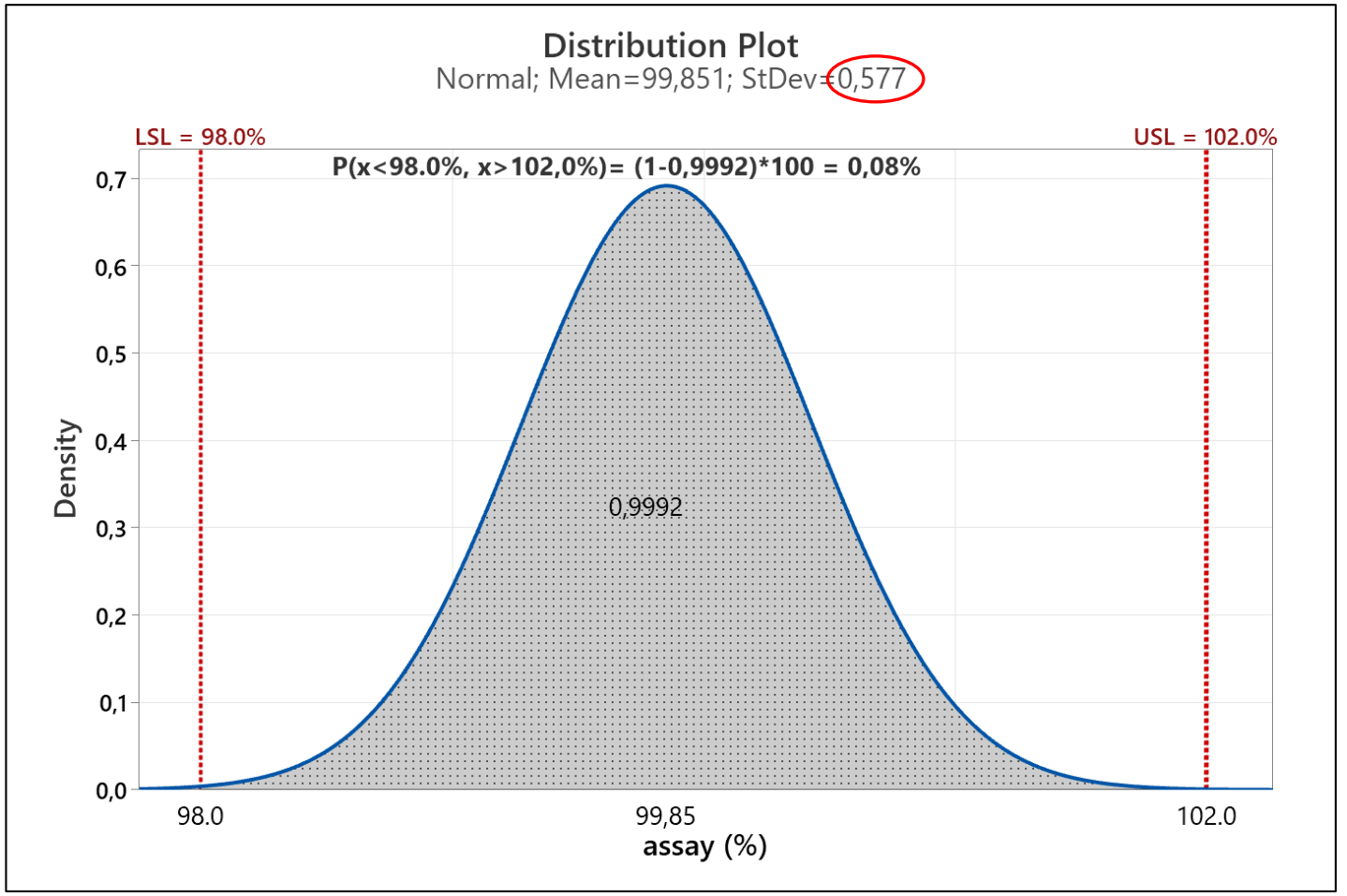
What does the *Process Capability Analysis* tell us?

- The ***Process Capability Index Cpm****, which evaluates the overall capability of the process relative to both the specification spread and the target, is equal to **1.13** which can be considered acceptable. Data, in fact, fall within the specification limits and the process is just slightly off target.

- ***% Total for Expected Overall Performance*** , which estimates the percentage of nonconforming items in the process, based on process overall variation, is equal to **0.08%**. This means that we expect 99.92% of conforming items ☺

DK.L. Chan *et al.*, *A New Measure of Process Capability:* C_{PM}, J. Quality Technology, Vol. 20, No. 3, (1988), pp. 162-175

What does the *Process Capability Analysis* tell us?



PHASE 1

- We can consider Phase 1 complete, when the historical data evaluation process just shown has been completed for all the CPPs and CQAs initially identified as relevant.
- In the light of the results of this analysis, if feasible, it is also possible to undertake actions that can only be for improvement.

ATTENTION : It is not said that "an improvement" can always be obtained or that it is convenient !

The following example is significant in this regard.

PHASE 1

Consider for example the case of a critical process temperature (CPP) which must remain within the safe range 75°C – 85°C, with a target of 80°C.

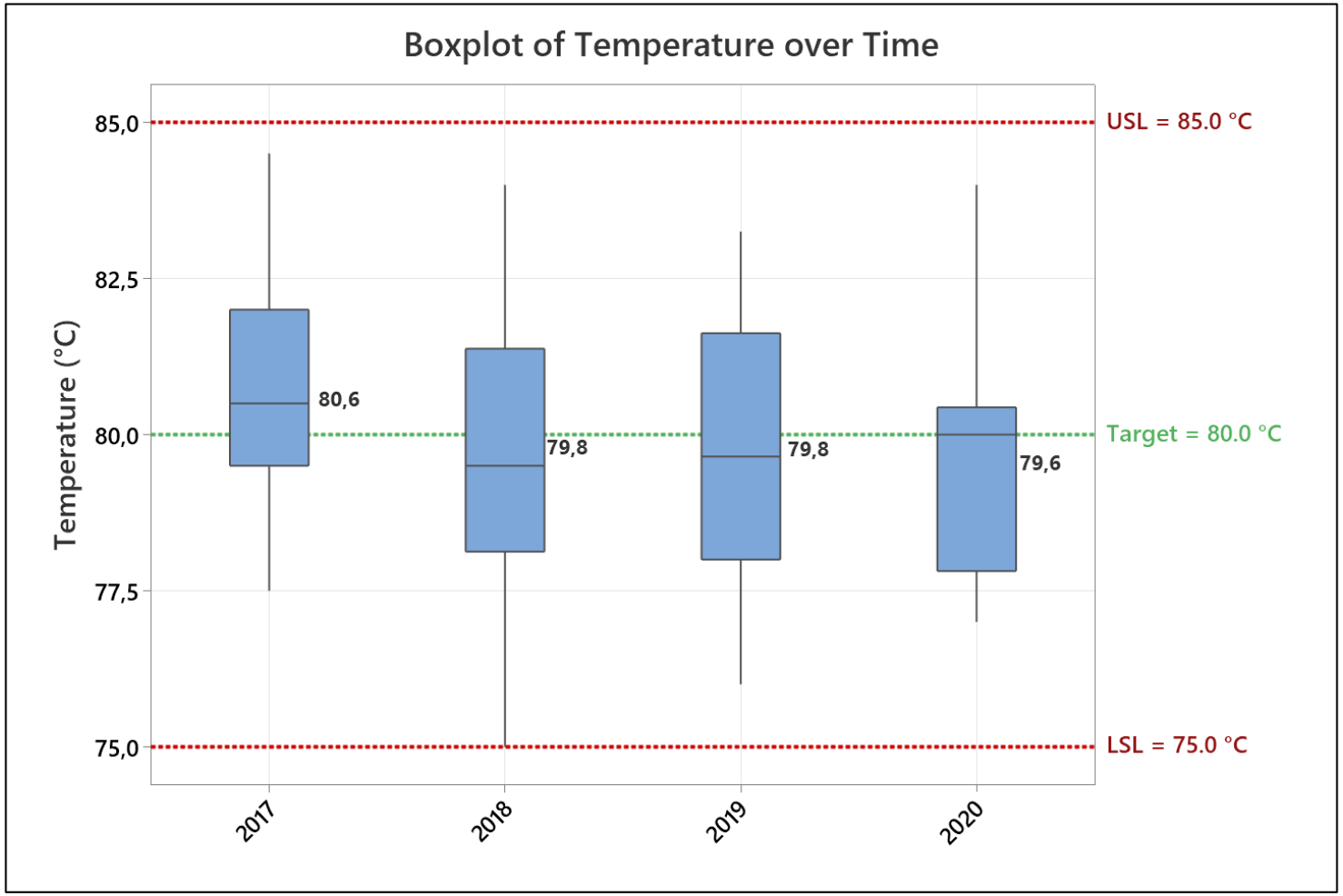
The measured values for each manufactured batch from 2017 to 2020 are shown alongside.

2017	2018	2019	2020
81,0	78,5	83,3	80,0
79,0	77,5	79,5	80,3
81,5	79,0	79,0	80,0
84,5	81,5	83,0	80,5
82,5	75,0	76,5	77,5
80,0	79,5	78,0	77,5
79,5	79,5	78,8	77,8
79,5	84,0	81,0	81,0
79,5	81,0	79,0	81,0
80,5	83,0	79,8	77,0
78,0	83,0	82,5	78,0
81,5	81,0	81,5	79,0
83,0	79,0	81,5	81,5
80,5	78,0	80,3	80,0
83,5	80,0	78,0	78,0
80,5	78,0	76,0	80,3
80,5		76,5	77,5
79,5		82,0	80,0
81,0			79,5
82,0			83,0
79,5			77,5
83,0			78,5
78,0			84,0
79,5			80,0
83,0			
78,0			
77,5			

PHASE 1

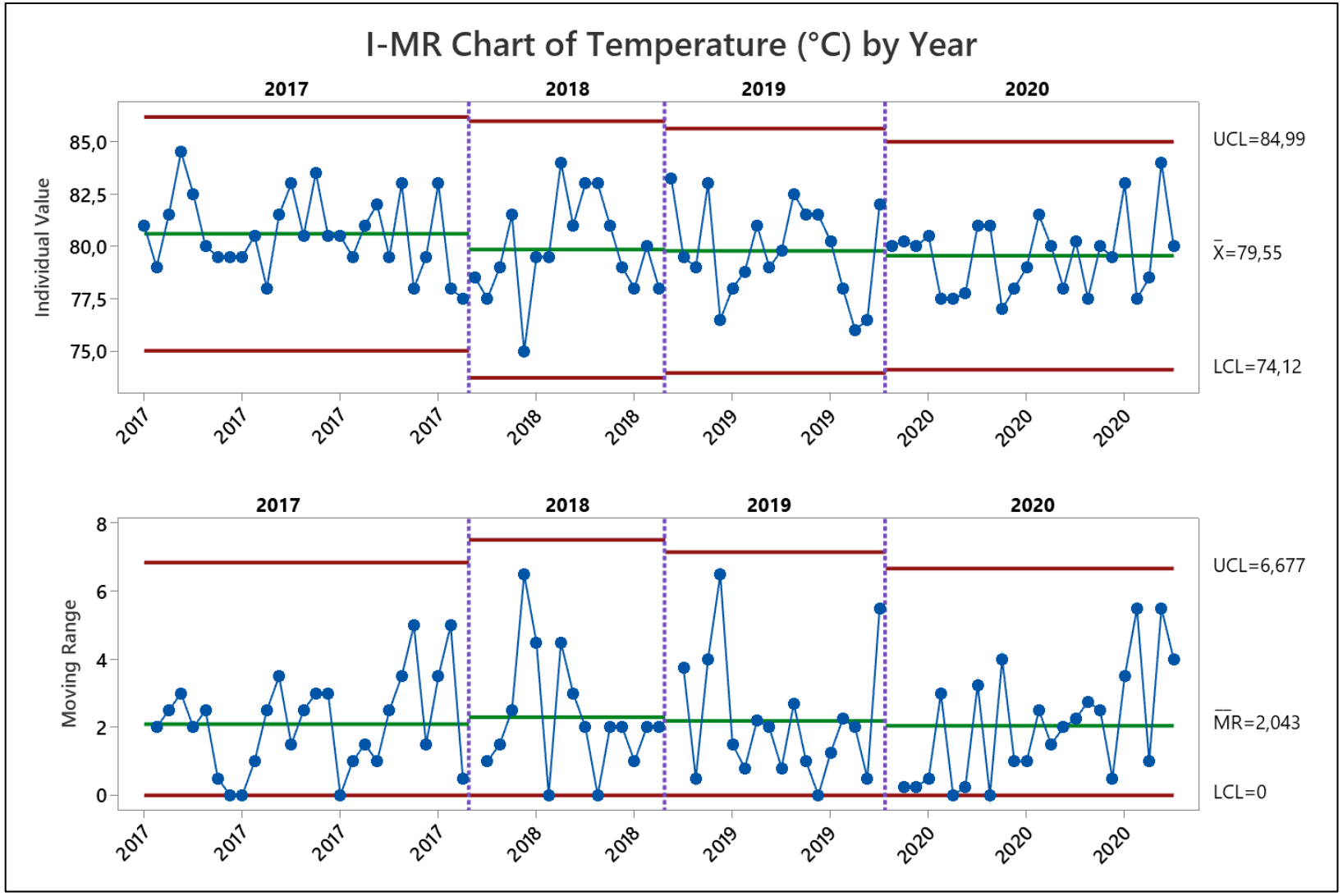
These boxplots show at least 3 interesting aspects of the datasets under study:

- the annual average values are all close to the target
- all values, although within the specification limits, sometimes come close to them
- the "boxes" are comparable in size.



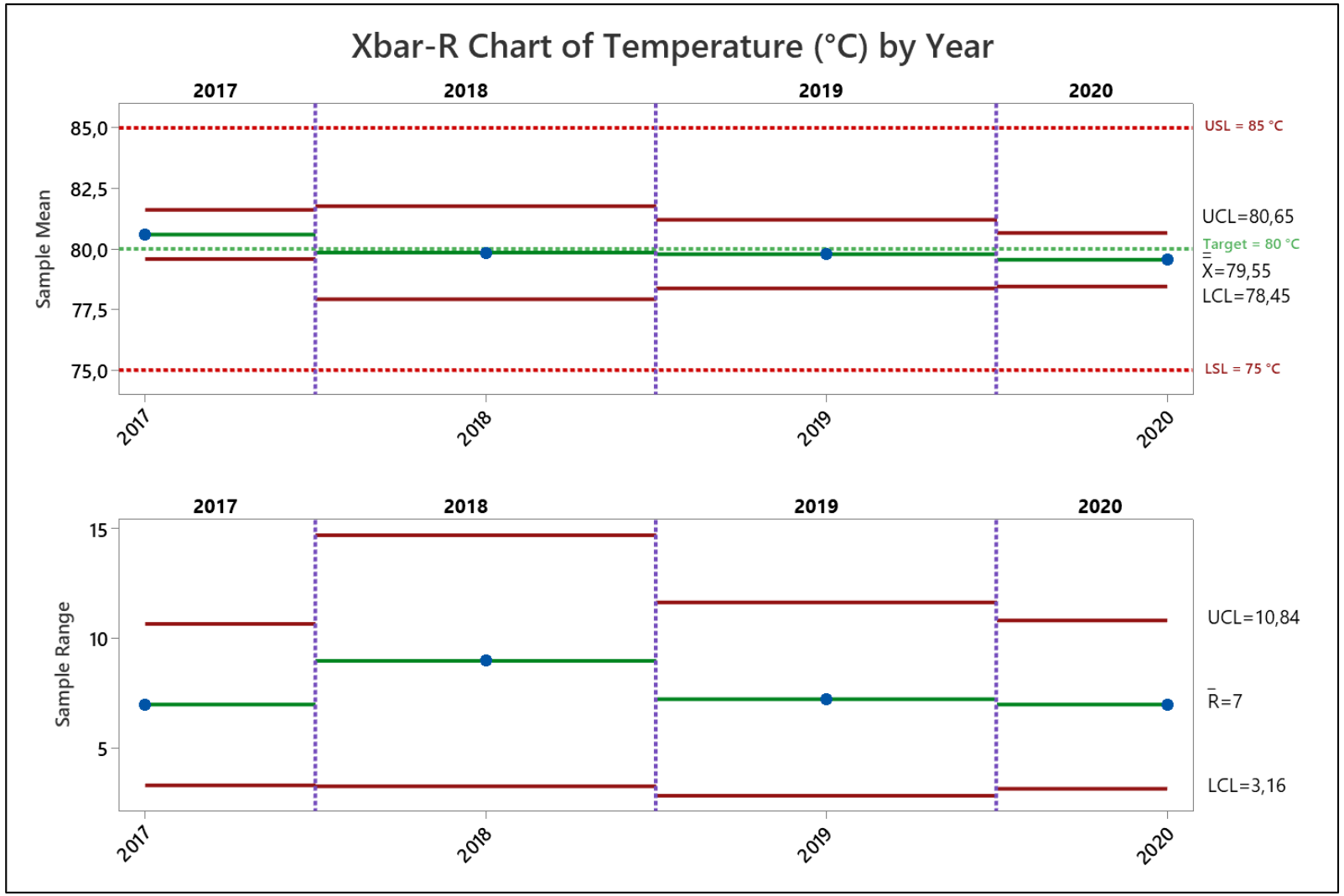
PHASE 1

The evidences shown by the boxplots are also evident by examining the *I-MR chart* shown here.



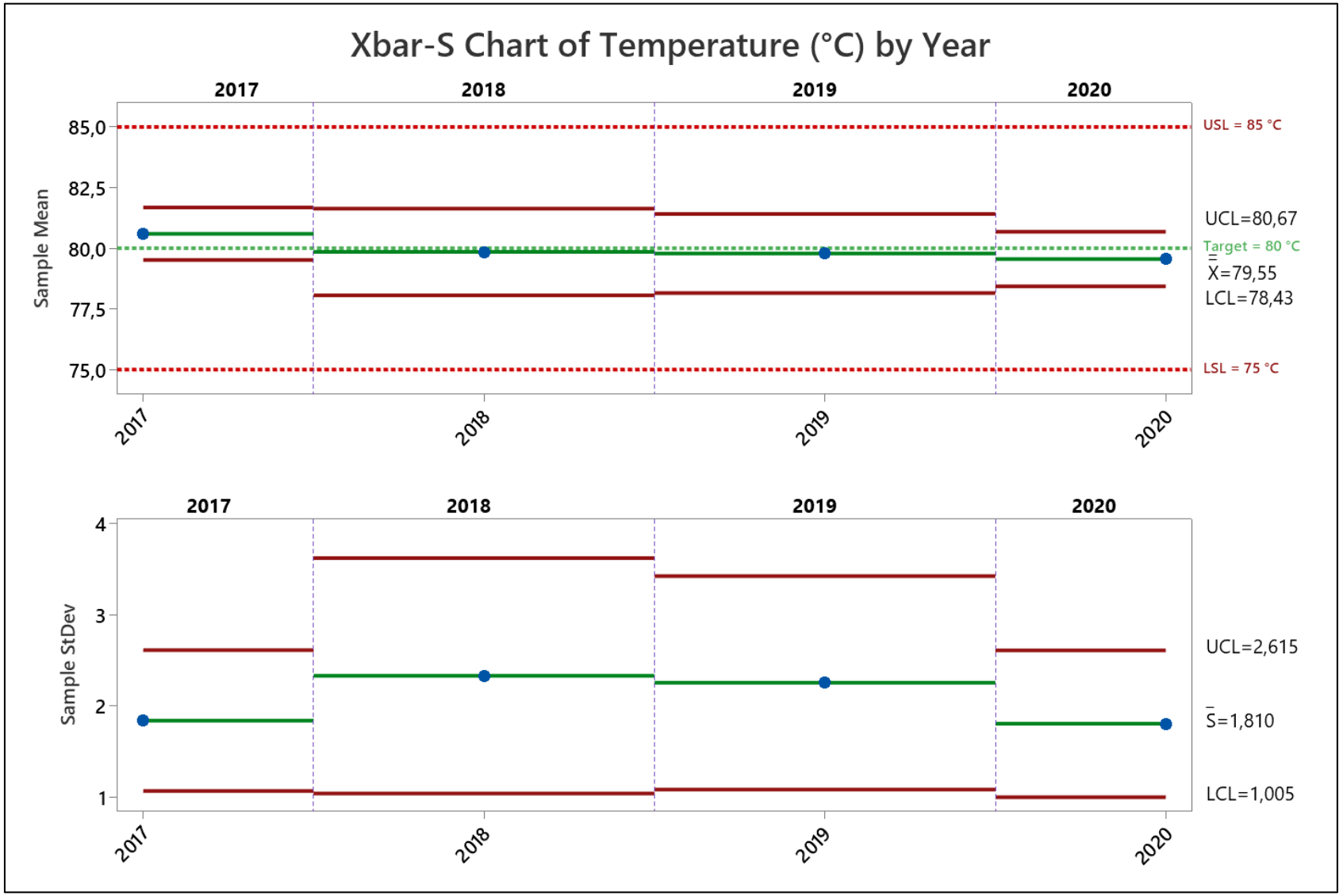
PHASE 1

The *Xbar-R chart* in this slide and the *Xbar-S chart* in the next confirm what above.



PHASE 1

In the next slide the four data groups are evaluated using ANOVA



PHASE 1

SUMMARY						
Groups	Count	Sum	Average	Variance	Std. Dev.	**CV%**
Year 2017	27	2176	80,6	3,4046	1,8451	**2,29**
Year 2018	16	1277,5	79,8	5,4573	2,3361	**2,93**
Year 2019	18	1436,1	79,8	5,1035	2,2591	**2,83**
Year2020	24	1909,25	79,6	3,2771	1,8103	**2,28**

ANOVA						
Source of Variation	**Sum of Squares**	**dof**	**Mean of Squares**	**F calc**	**P-value**	**F crit**
Between Groups	15,5167	3	**5,17**	**1,26**	0,29	**2,72**
Within Groups	332,5103	81	**4,11**			
Total	348,0269	84				

PHASE 1

ANOVA returns important information and confirmations:

- ***variability is smaller within groups than among them***
- ***there are no significant differences between the average temperature values over the four years covered by the study***
- ***the four datasets show similar variability (see comparable CV% values) as shown by the "test for Equal variances" in the next slide.***

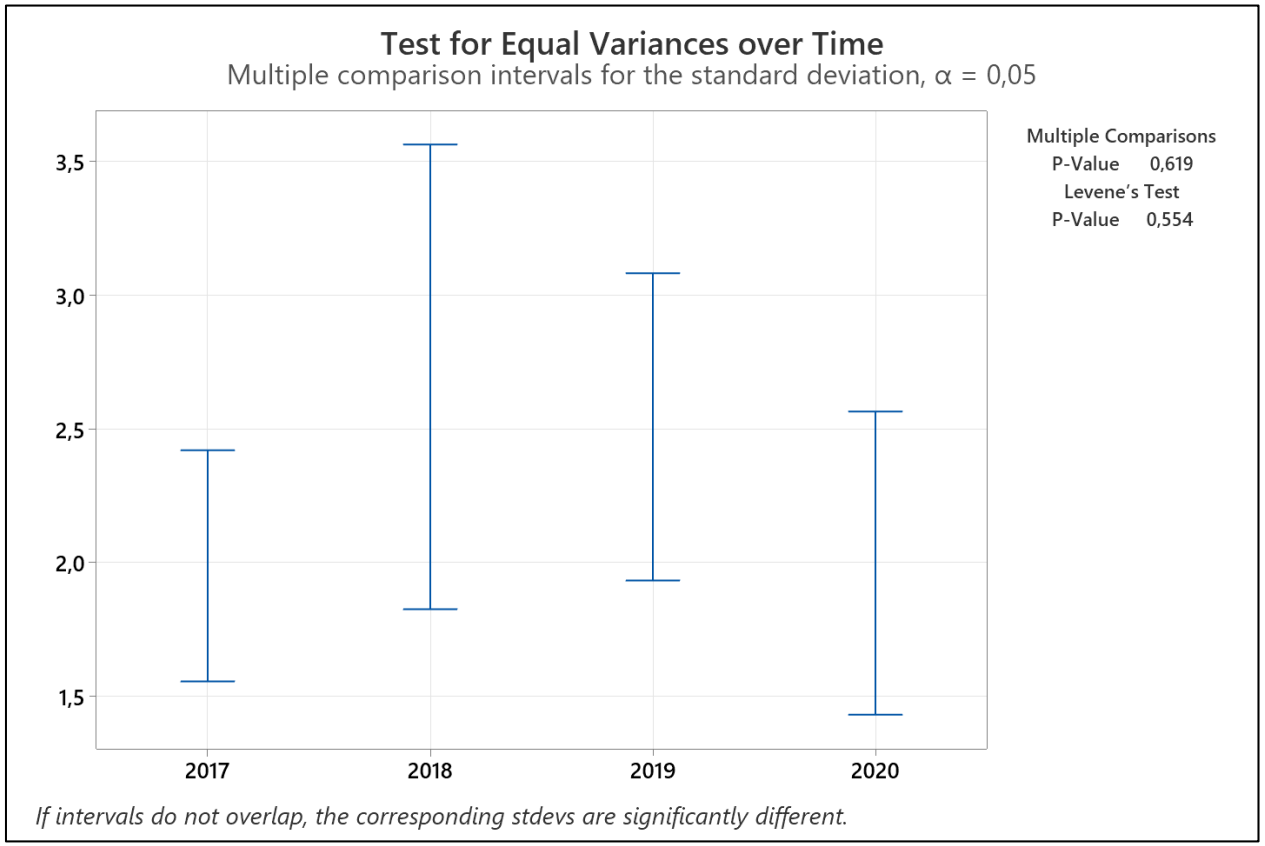
PHASE 1

Method

Null hypothesis	All variances are equal
Alternative hypothesis	At least one variance is different
Significance level	α = 0,05

Tests

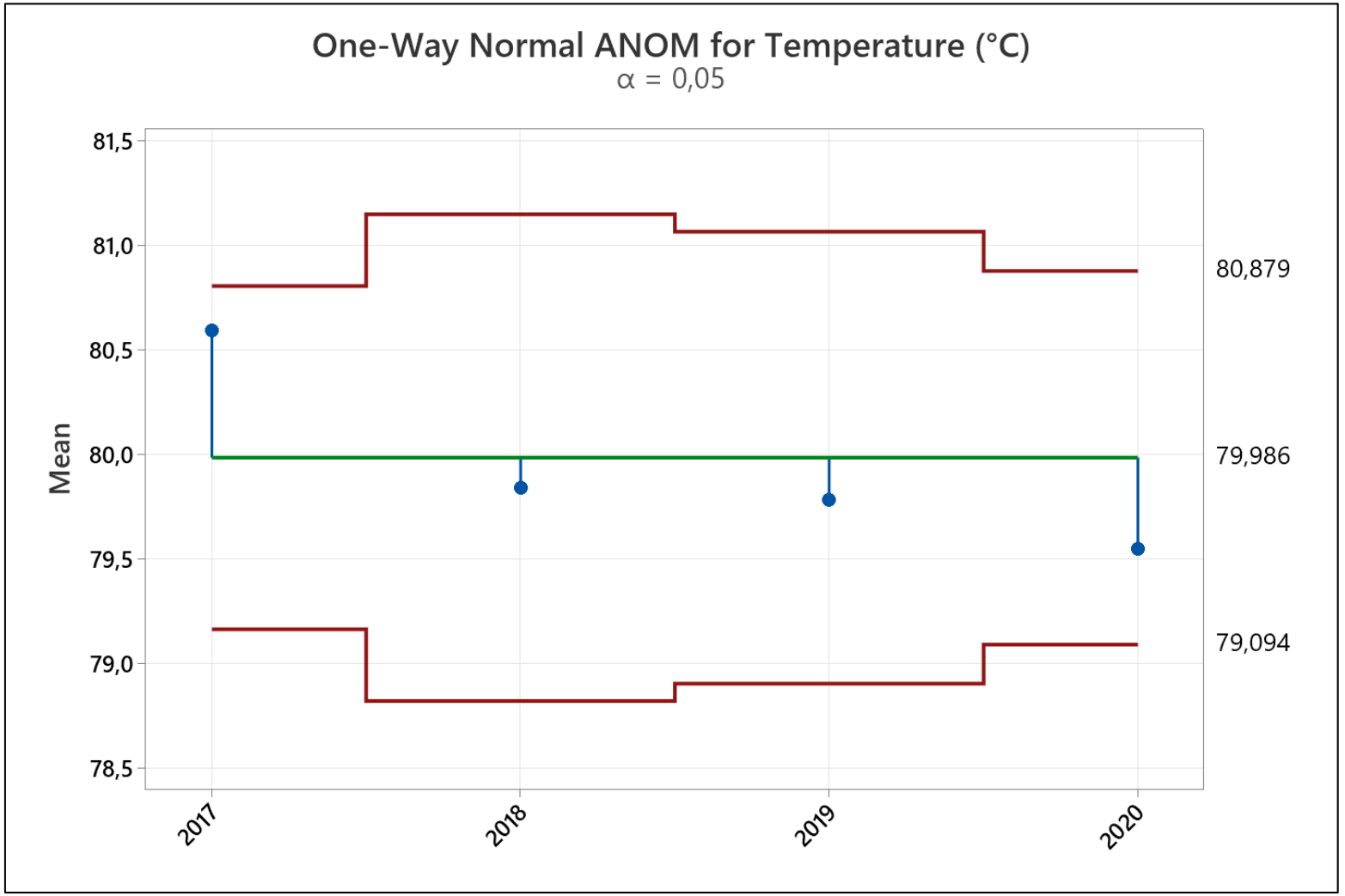
Method	Test Statistic	P-Value
Multiple comparisons	—	0,619
Levene	0,70	0,554



The results shown above are extremely important because they tells us not only that the process always behaved in the same way on average, but also that the variability does not show statistically significant differences over time and therefore reducing it is not easy!

PHASE 1

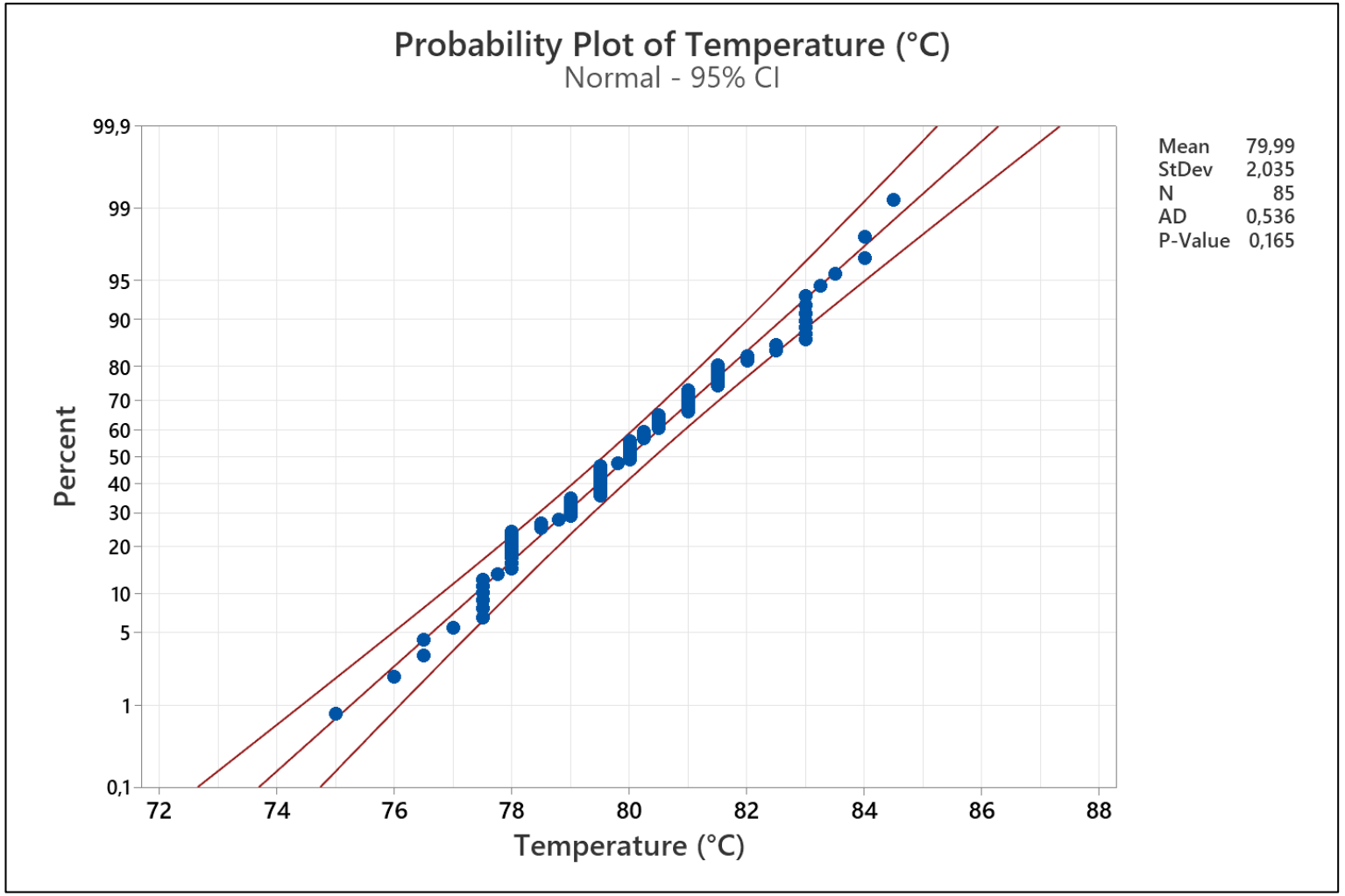
The Analysis of Means (ANOM) shown here returns the picture of average group data within the limits. The overall mean value is very close to the target of 80°C.



PHASE 1

Since the data examined as a whole return the image of a stable process, *i.e.*, “under control” or "predictable“, and moreover they are normally distributed (see the Normal Probability Plot opposite), it makes sense to conduct a

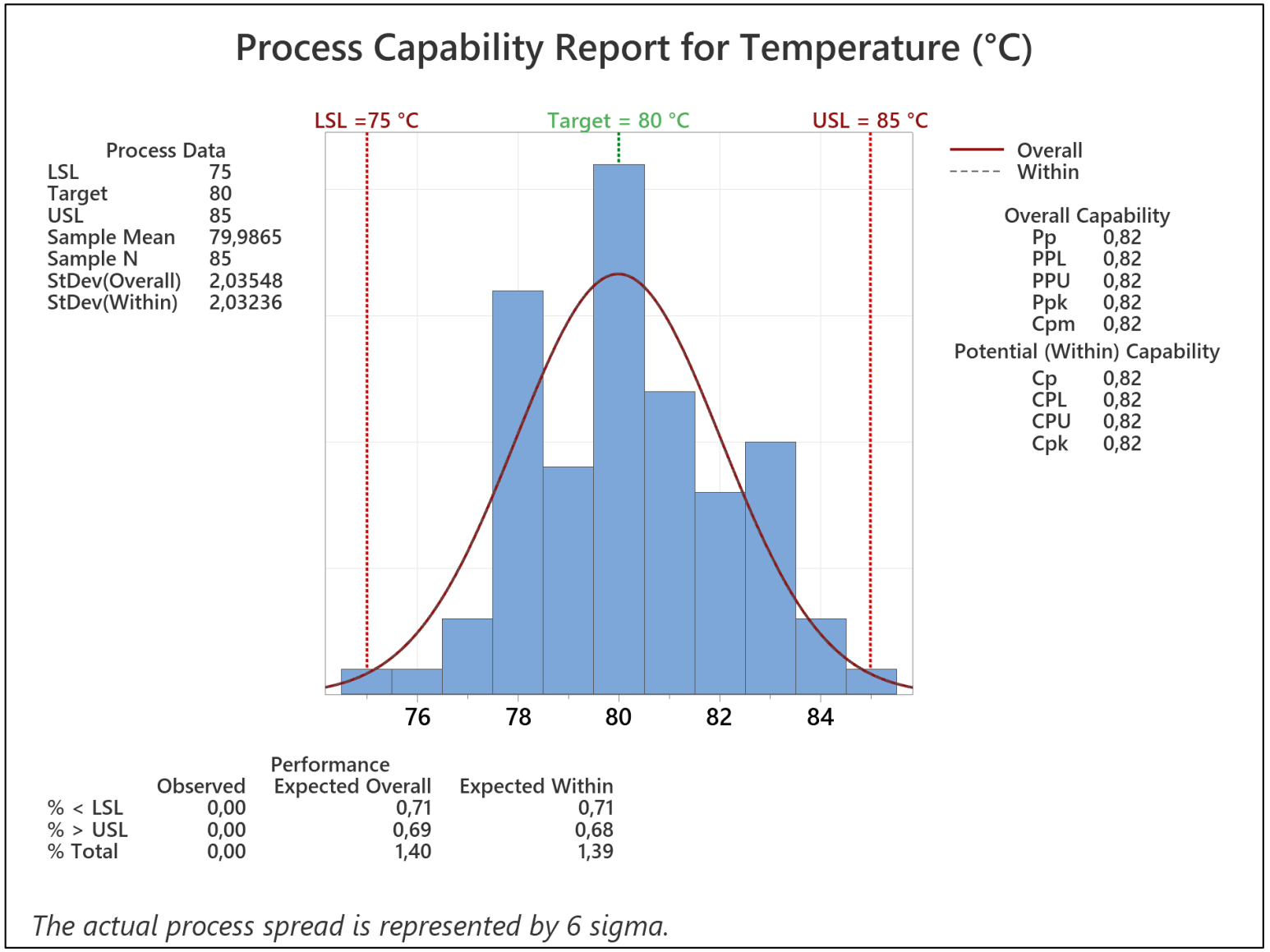
Process Capability Analysis



PHASE 1

The Capability Analysis returns, as expected, the image of a process:

- centered
- symmetrical with respect to the target
- with equal indices values which are < 1 due to the presence of experimental data close to the specification limits.



PHASE 1

In a case like this, precisely because of data constant variability over time, it is reasonable to expect that things will remain the same in the future as well.

That is, we are in the presence of a process affected only by "***common causes***" and therefore improving its Capability is much more difficult than in the presence of "***special causes***". The latter, in fact, if present, are easy to identify and remove. Common causes involve acting on the process in a much heavier way !

Will it be possible to act on the process? or, better yet,

Could the product pay for the investment ?

Summarizing

- Phase 1 can be considered completed only when analyzes of the type shown up to now have been conducted on each of the initially selected CPP and CQA.
- If CPPs or CQAs were found to be visibly out of control due to "special causes" from this analysis, those "special causes" must be identified and removed before proceeding.
- ***Only after the completion of all the activities belonging to Phase 1 can one pass to Phase 2.***

PHASE 2

PHASE 2

- Phase 2 consists in monitoring the CPPs and CQAs initially identified, trying to prevent any deviations from the stable starting conditions identified or established after the completion of Phase 1.
- In Phase 2, Shewhart control charts of the type used in Phase 1 are much less likely to be effective as they are not sensitive to small to moderate process changes.
- In Phase 2 Cumulative Sum (CUSUM) and Exponentially Weighed Moving Average (EWMA) control charts are much more likely to be effective.

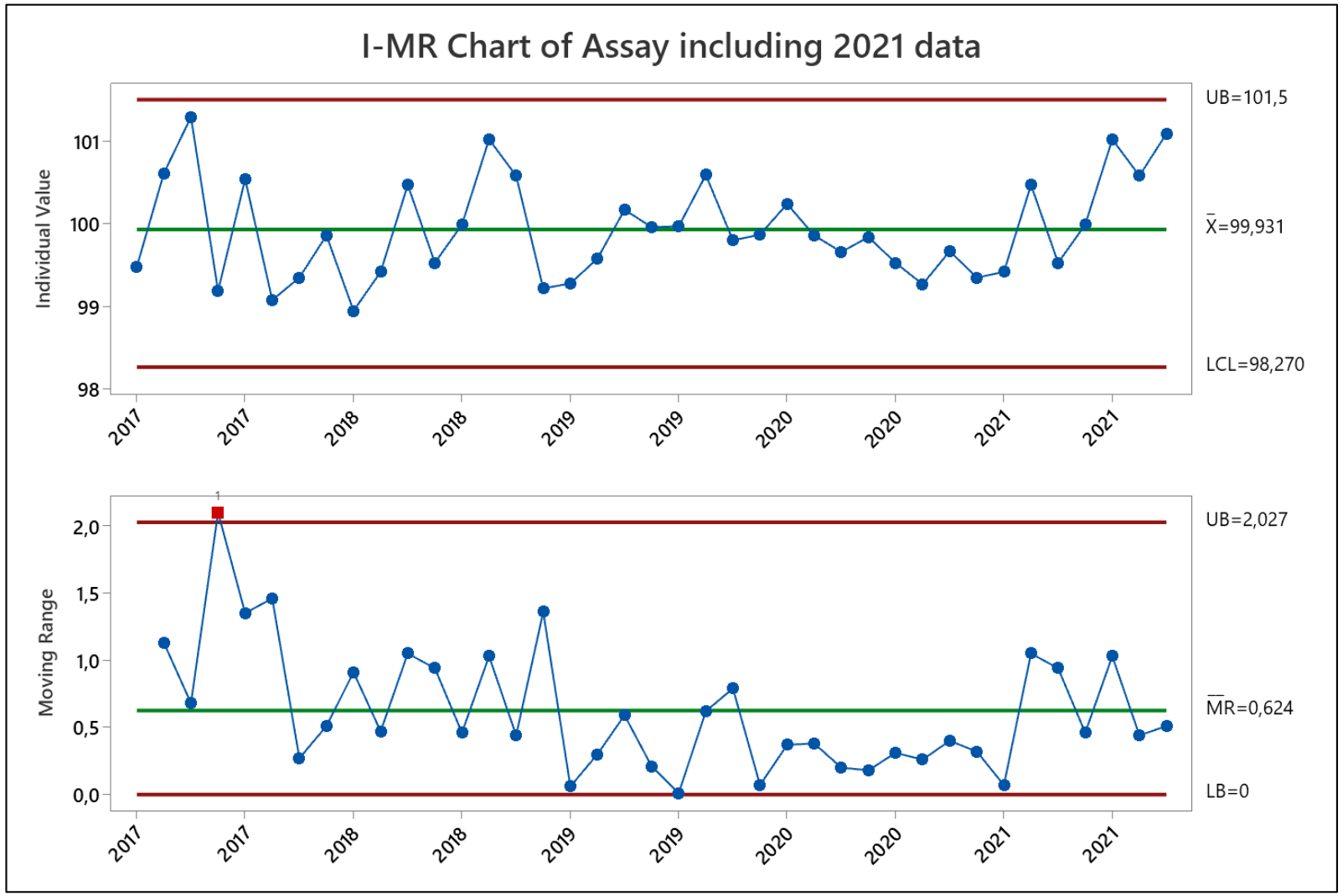
PHASE 2

As an example, let's take the (theoretical) case of the Assay percentage values seen previously and add some possible values for 2021.

Assay 2017	Assay 2018	Assay 2019	Assay 2020	Assay 2021
99,48	99,86	99,22	99,8	99,42
100,61	98,95	99,28	99,87	100,47
101,29	99,42	99,58	100,24	99,53
99,19	100,47	100,17	99,86	99,99
100,54	99,53	99,96	99,66	101,02
99,08	99,99	99,97	99,84	100,58
99,35	101,02	100,59	99,53	101,09
	100,58		99,27	
			99,67	
			99,35	

PHASE 2

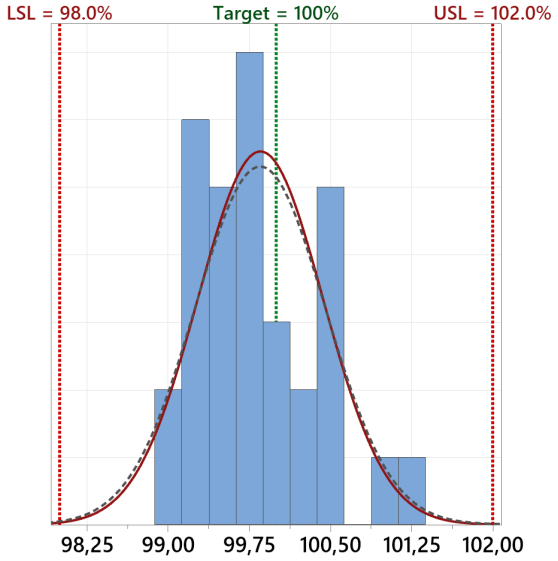
The addition of the 2021 data, if carried out using an I-MR graph with the limits calculated on the entire database of the previous four years, reveals nothing if not a slight growth like that seen in the past and in any case accompanied by a smaller variation of the average mobile.



PHASE 2

Even a comparison between overall Capability Analysis would not reveal significant changes.

Process Capability Report for Assay : data 2017 - 2020



Process Data	
LSL	98
Target	100
USL	102
Sample Mean	99,8506
Sample N	32
StDev(Overall)	0,577246
StDev(Within)	0,601354

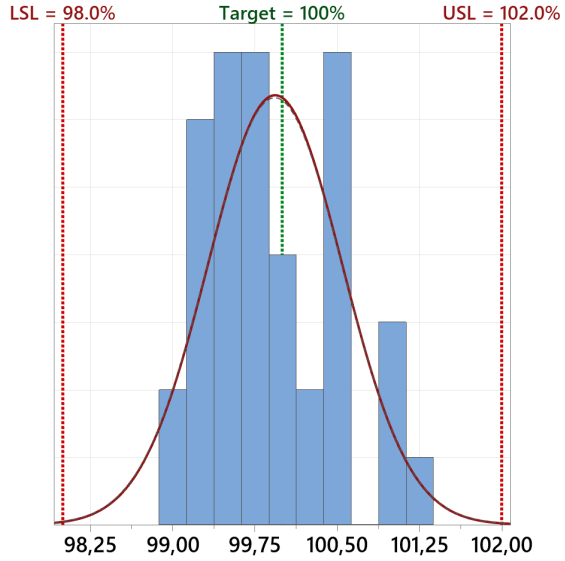
Overall Capability	
Pp	1,15
PPL	1,07
PPU	1,24
Ppk	1,07
Cpm	1,13

Potential (Within) Capability	
Cp	1,11
CPL	1,03
CPU	1,19
Cpk	1,03

Performance	Observed	Expected Overall	Expected Within
% < LSL	0,00	0,07	0,10
% > USL	0,00	0,01	0,02
% Total	0,00	0,08	0,12

The actual process spread is represented by 6 sigma.

Process Capability Report for Assay : data 2017 - 2021



Process Data	
LSL	98
Target	100
USL	102
Sample Mean	99,9313
Sample N	39
StDev(Overall)	0,611354
StDev(Within)	0,6147

Overall Capability	
Pp	1,09
PPL	1,05
PPU	1,13
Ppk	1,05
Cpm	1,10

Potential (Within) Capability	
Cp	1,08
CPL	1,05
CPU	1,12
Cpk	1,05

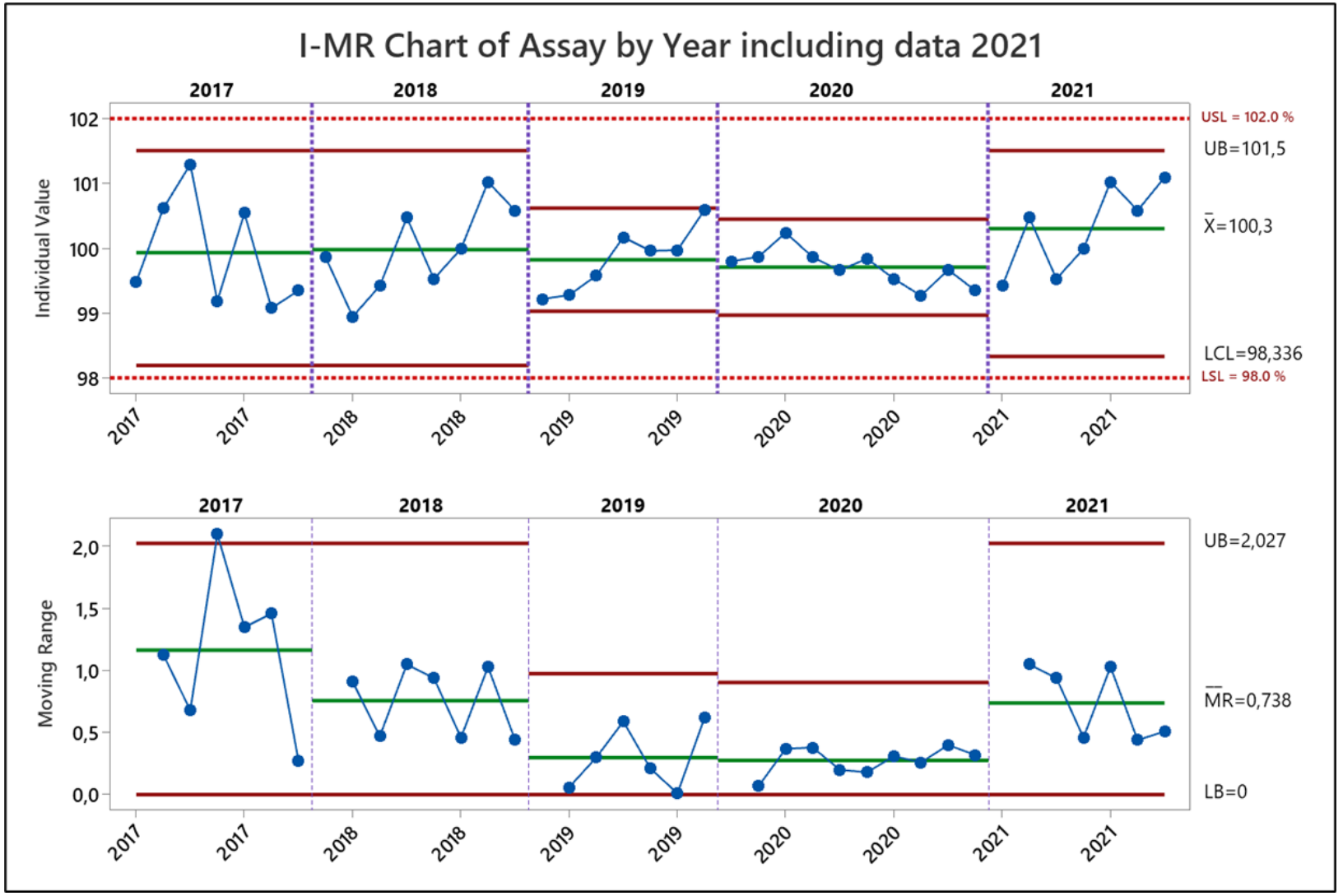
Performance	Observed	Expected Overall	Expected Within
% < LSL	0,00	0,08	0,08
% > USL	0,00	0,04	0,04
% Total	0,00	0,11	0,12

The actual process spread is represented by 6 sigma.

PHASE 2

Considering the data by year, even if everything remains within the specification limits, the following are evident:

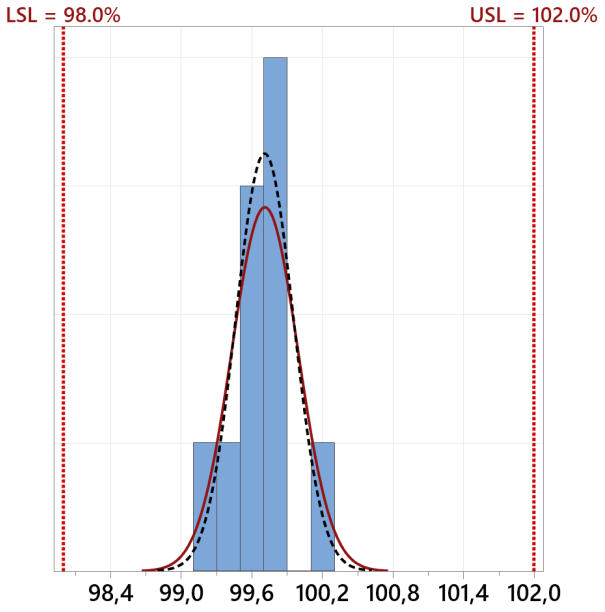
- a growth trend in the data and
- an increase in the value of the group average (99.7% in 2020 *vs.* 100.3% in 2021)



PHASE 2

Considering the data by year, even Capability Analysis is more informative.

Process Capability Report for Assay: data 2020



Process Data	
LSL	98
Target	*
USL	102
Sample Mean	99,709
Sample N	10
StDev(Overall)	0,281601
StDev(Within)	0,245272

Overall
Within

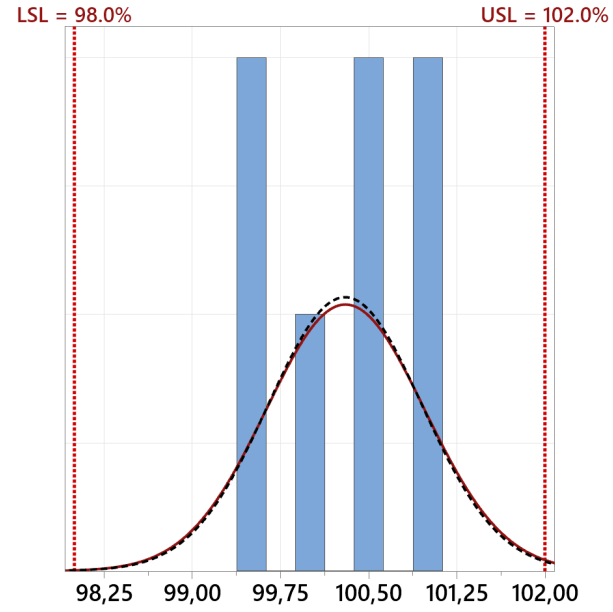
Overall Capability	
Pp	2,37
PPL	2,02
PPU	2,71
Ppk	2,02
Cpm	*

Potential (Within) Capability	
Cp	2,72
CPL	2,32
CPU	3,11
Cpk	2,32

Performance	Observed	Expected Overall	Expected Within
% < LSL	0,00	0,00	0,00
% > USL	0,00	0,00	0,00
% Total	0,00	0,00	0,00

The actual process spread is represented by 6 sigma.

Process Capability Report for Assay: data 2021



Process Data	
LSL	98
Target	*
USL	102
Sample Mean	100,3
Sample N	7
StDev(Overall)	0,672458
StDev(Within)	0,654551

Overall
Within

Overall Capability	
Pp	0,99
PPL	1,14
PPU	0,84
Ppk	0,84
Cpm	*

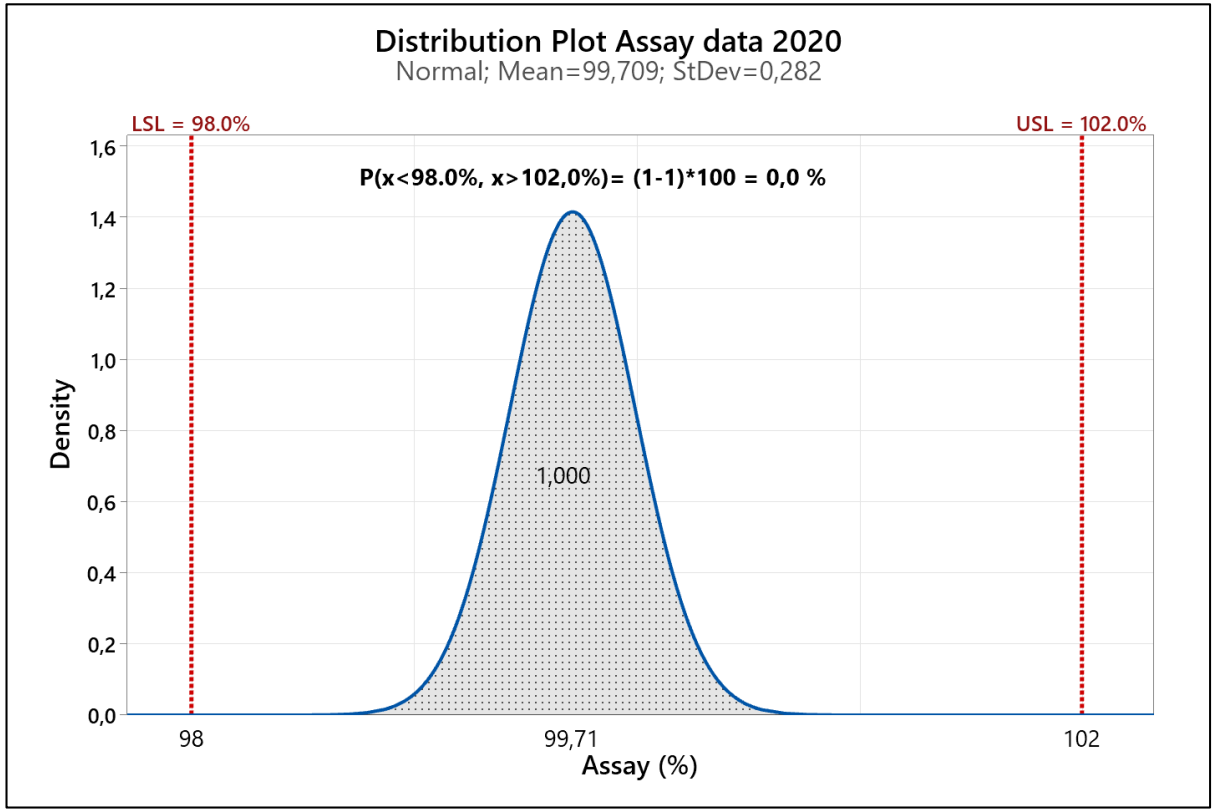
Potential (Within) Capability	
Cp	1,02
CPL	1,17
CPU	0,87
Cpk	0,87

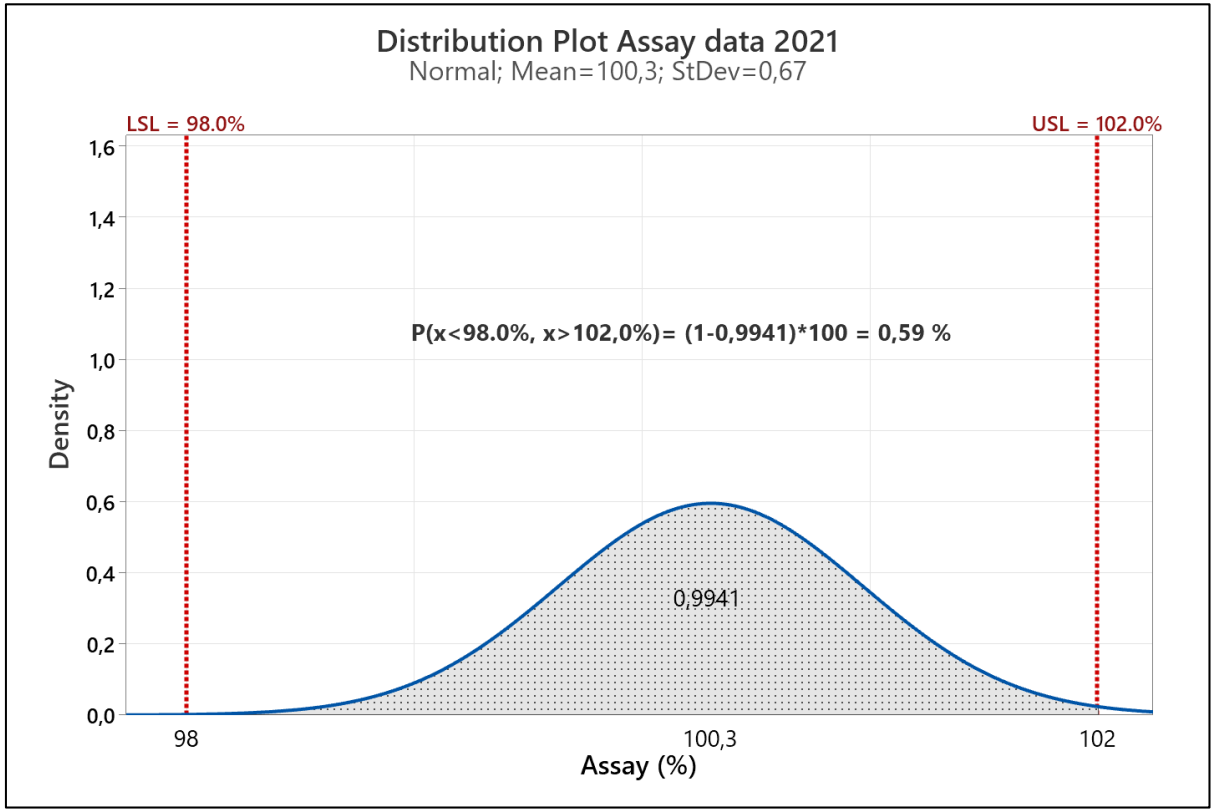
Performance	Observed	Expected Overall	Expected Within
% < LSL	0,00	0,03	0,02
% > USL	0,00	0,57	0,47
% Total	0,00	0,60	0,49

The actual process spread is represented by 6 sigma.

PHASE 2

From a probabilistic point of view, the process situation 2020 *vs.* 2021 is as follows:





PHASE 2

These results should come as no surprise!

Since Shewhart Control Charts use only the information about the process contained in the last sample observation and ignore any information given by the entire sequence of points, this makes them quite insensitive to small process shifts, *i.e.*, in the order of 1.5 standard deviation or less.

Shewhart's control charts are perfect for bringing an uncontrollable process under statistical control, but then they have to give way to other, much more sensitive types of charts.

PHASE 2

What does ANOVA tell us now?

Data 2017 - 2020

SUMMARY						
Groups	Count	Sum	Average	Variance	Std. Dev.	***CV%***
Assay 2017	7	699,54	99,9343	0,7488	0,8653	**0,87**
Assay 2018	8	799,82	99,9775	0,4676	0,6838	**0,68**
Assay 2019	7	698,77	99,8243	0,2444	0,4944	**0,50**
Assay 2020	10	997,09	99,7090	0,0793	0,2816	**0,28**

ANOVA						
Source of Variation	**Sum of Squares**	**dof**	**Mean of Squares**	**F calc**	**P-value**	**F crit**
Between Groups	0,3832	3	**0,13**	0,36	0,78	2,95
Within Groups	9,9464	28	**0,36**			
Total	10,3296	31				

Data 2017 - 2021

SUMMARY						
Groups	Count	Sum	Average	Variance	Std. Dev.	**CV%**
Assay 2017	7	699,54	99,9343	0,7488	0,8653	**0,87**
Assay 2018	8	799,82	99,9775	0,4676	0,6838	**0,68**
Assay 2019	7	698,77	99,8243	0,2444	0,4944	**0,50**
Assay 2020	10	997,09	99,7090	0,0793	0,2816	**0,28**
Assay 2021	7	702,10	100,3000	0,4522	0,6725	**0,67**

ANOVA						
Source of Variation	**Sum of Squares**	**dof**	**Mean of Squares**	**F calc**	**P-value**	**F crit**
Between Groups	1,5431	4	**0,39**	1,04	0,40	2,65
Within Groups	12,6596	34	**0,37**			
Total	14,2026	38				

PHASE 2

ANOVA One-Way tell us that considering 2021 data:

- ***variability between groups is comparable with that within them***
- ***average values of the data groups are not significantly different from each other***

However, these findings do not reveal much !

Let's now have a look to other Control Charts !!

PHASE 2

- ***Cumulative Sum (CUSUM) Control Chart*** :
  - It is typically designed with the focus of detecting (at least) a one standard deviation shift
  - it is a type of control chart which *considers all the information in the process data* by plotting the cumulative sums of the deviations of each sample value from a target (*e.g.*, 100% for the assay).
  - The CUSUM scheme employs two cumulative sums:
    - the first cumulative sum is for detecting mean increases
    - the second is for detecting mean decreases

PHASE 2

- ***Cumulative Sum (CUSUM) Control Chart (cont.)***:

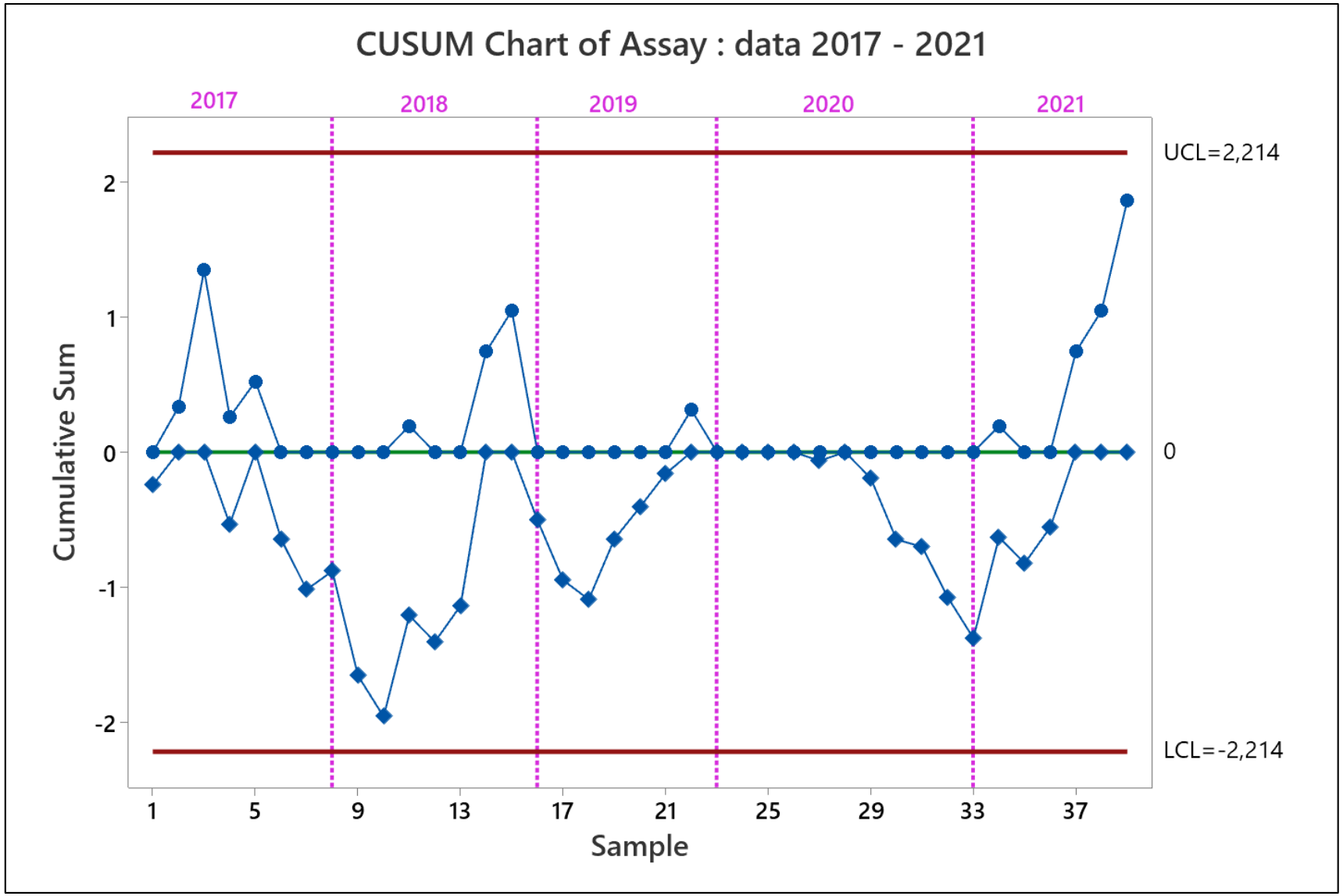
  - if the process remains in control at the target value, we should observe values which are randomly placed around zero.

    If, on the other hand, we see a significant shift up or down in the data points, this indicates that the process mean is shifting, and this is likely due to an "assignable cause“ which must be promptly identified and eliminated.

PHASE 2

Both the upper and lower CUSUM curves detect a strong upward shift in the level of the process.

The shift takes place from the start of production 2021.



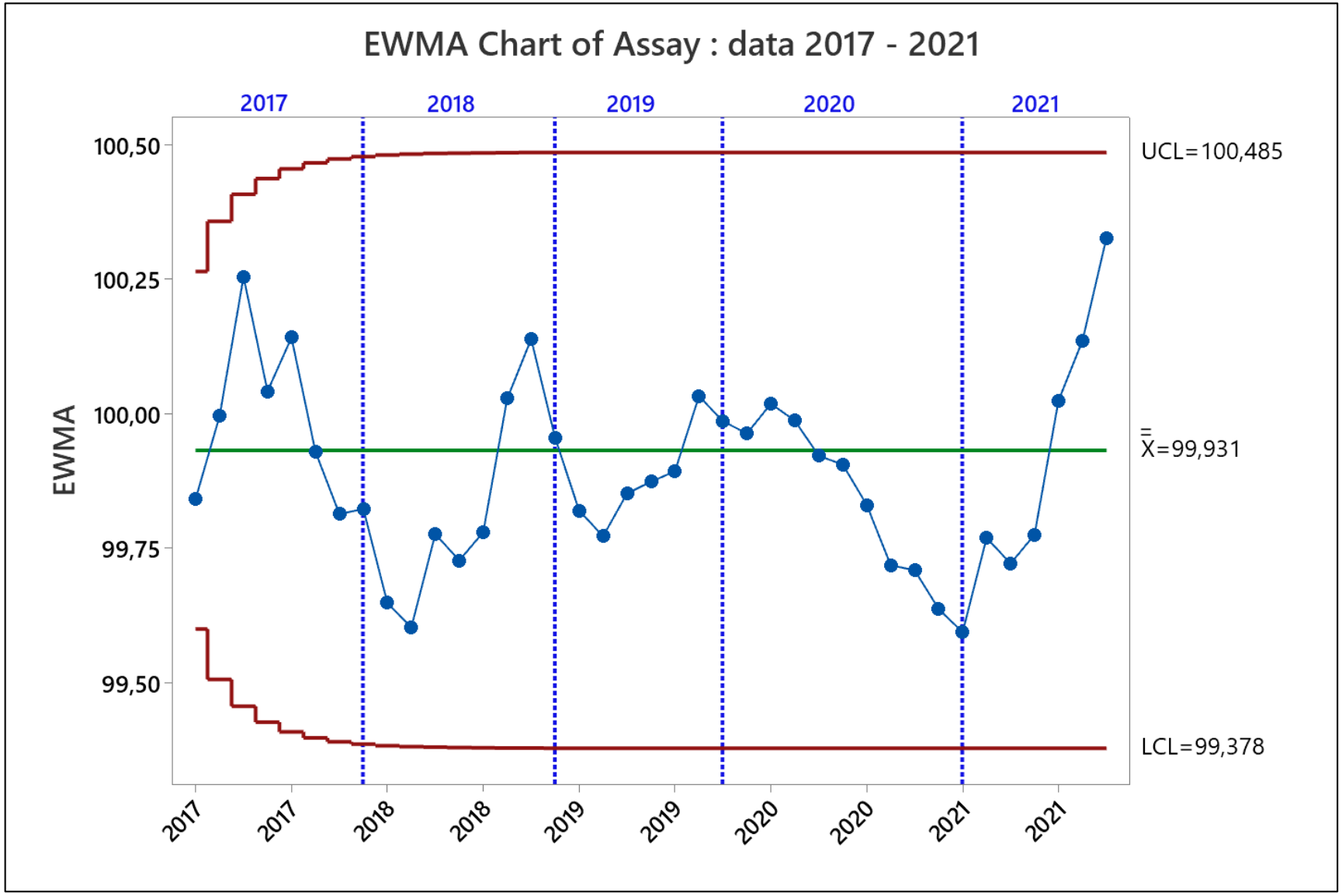
PHASE 2

- ***Exponentially Weighed Moving Average (EWMA) Control Chart*** :
  - it is a type of control chart which *considers all the information in the process data* by plotting a weighted average of all past and current observations.
  - since this type of chart is highly insensitive to data normality assumption, it is the ideal control chart to use with individual observations.

Like the CUSUM chart, the EWMA performs well on small shifts but does not react to large shifts as quickly as the Shewhart charts. This is why it is generally suggested to use both.

PHASE 2

The EWMA chart, like the CUSUM before it, also shows a strong upward shift since the start of production in 2021.



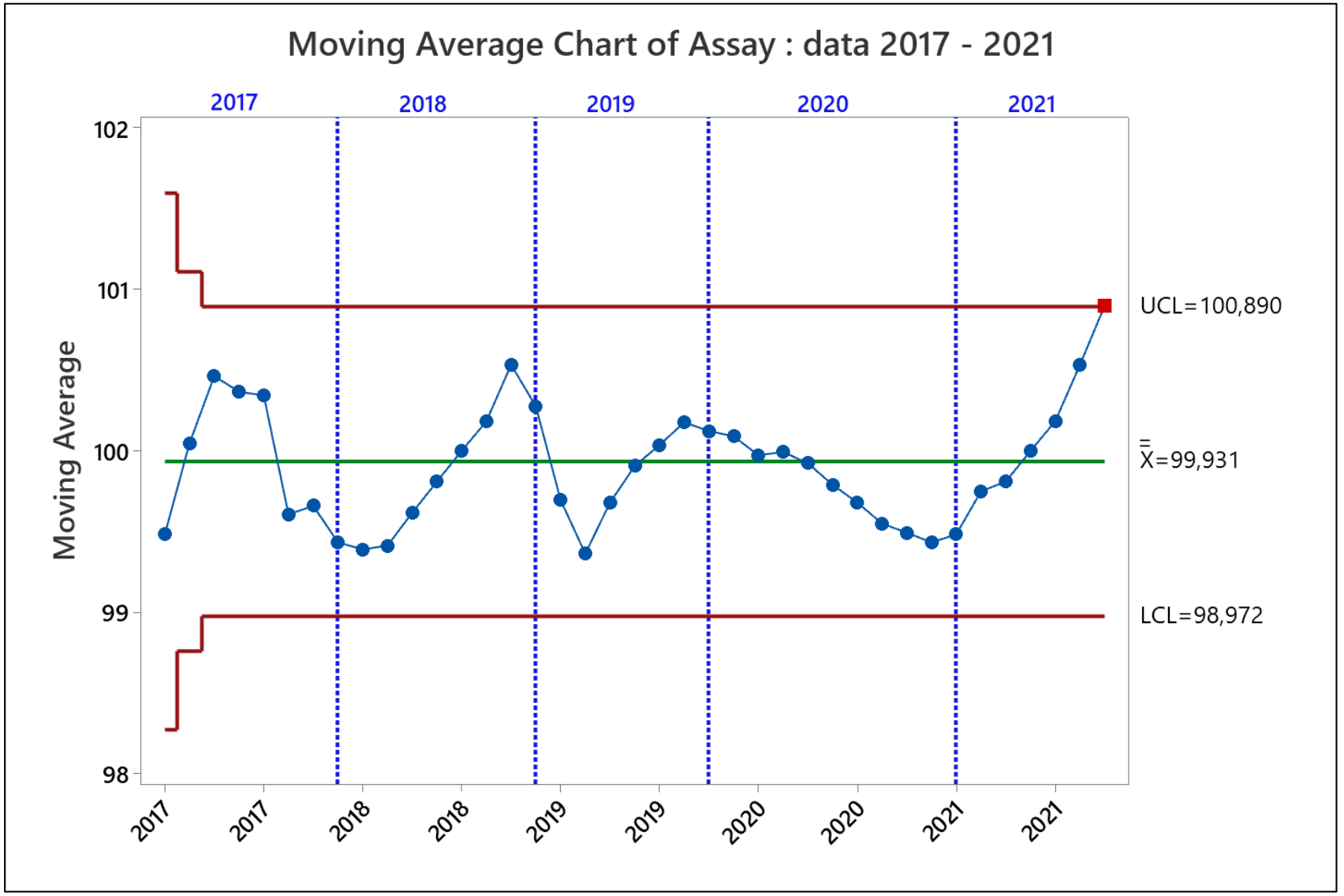
PHASE 2

- ***Moving Average (MA) Control Chart*** : it is a type of control chart which *considers all the information in the process data* by simply plotting an unweighted moving average of all individual observations.

  It is widely accepted that the MA chart is more effective than the Shewhart control charts in detecting small process shifts, but many Authors judge it to be less efficient, for example than the EWMA.
  However, in the chosen case study, the result is completely comparable.

PHASE 2

The shift upward takes place from the start of production 2021.



CONCLUSIONS

CONCLUSIONS

From all that has been discussed so far it is clear that to implement Continued Process Verification the following are required:

- a SOP that provides a description of the different stages of the process and how to approach them from a methodological point of view, for example:

  . identification of CPP and CQA of interest → Risk Analysis

  . retrospective data analysis → statistical tools

  . definition of monitoring criteria for CPV → ...

  . *etc.*

CONCLUSIONS

- One or more SOPs related to the different Statistical Methods available (Descriptive, Inferential, Statistical Process Control, *etc.*), each completed by practical examples. These procedures provide the "toolboxes“.

- An accurate historical analysis of the available data. This analysis, which represents Phase 1, will be conducted using the statistical methods described in the specific SOPs and which best apply to the case under study.

ATTENTION ! There is no one-size-fits-all approach!

CONCLUSIONS

- The historical analysis will be documented in an *ad hoc* report in the conclusions of which will be the methods and criteria to be used for the actual operational phase or Phase 2.
- Phase 2 will be carried out using specific control charts (*e.g.*, *EWMA*, *CUSUM*, *MA*) combined with Shewhart charts such as the *I-MR* chart. As usual in Statistics, the "golden rule" is to use multiple tools, ideally all of them, and not just one!
- The data will be statistically evaluated and reviewed by the Quality Unit which will share them with the Production staff.

alongside all this, please also consider that ...

CONCLUSIONS

While implementation is becoming a regulatory expectation, CPV can provide benefits beyond compliance by identifying opportunities to improve production processes and ultimately, the reliability of drug quality and supply.

ISPE Technical Document – *Continued Process Verification (CPV) Signal Responses in Biopharma* (February 2017)