SPSS

WEBINAR

THE T-TEST

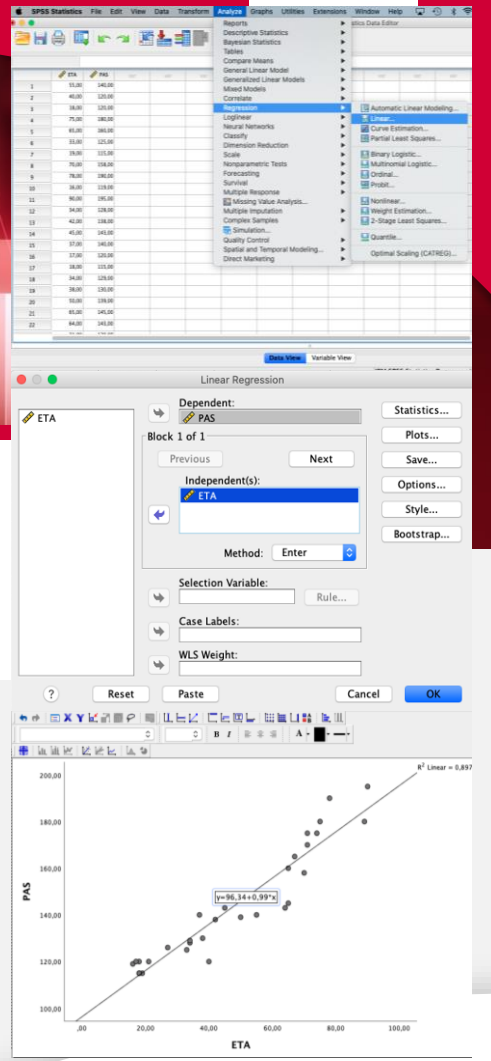
with Statistics for Data Analysis

June 13th

2.30 PM

Statistics for Data Analysis

WEBINAR AGENDA



About Us

25+ years of history… in brief

Statistics for Data Analysis

Services included in the solution (LaunchBox, Add-On, APP, webinar, Support, …)

T-Test with Statistics for Data Analysis

- Notes on the **Theory** of Estimation and Statistical **Inference**
- The main differences between **parametric** and **non-parametric** tests
- T-test **prerequisites**
- Practical **example**
- Q&A

Scarica la nostra APP sugli store

GET IT ON Google Play

Download on the App Store

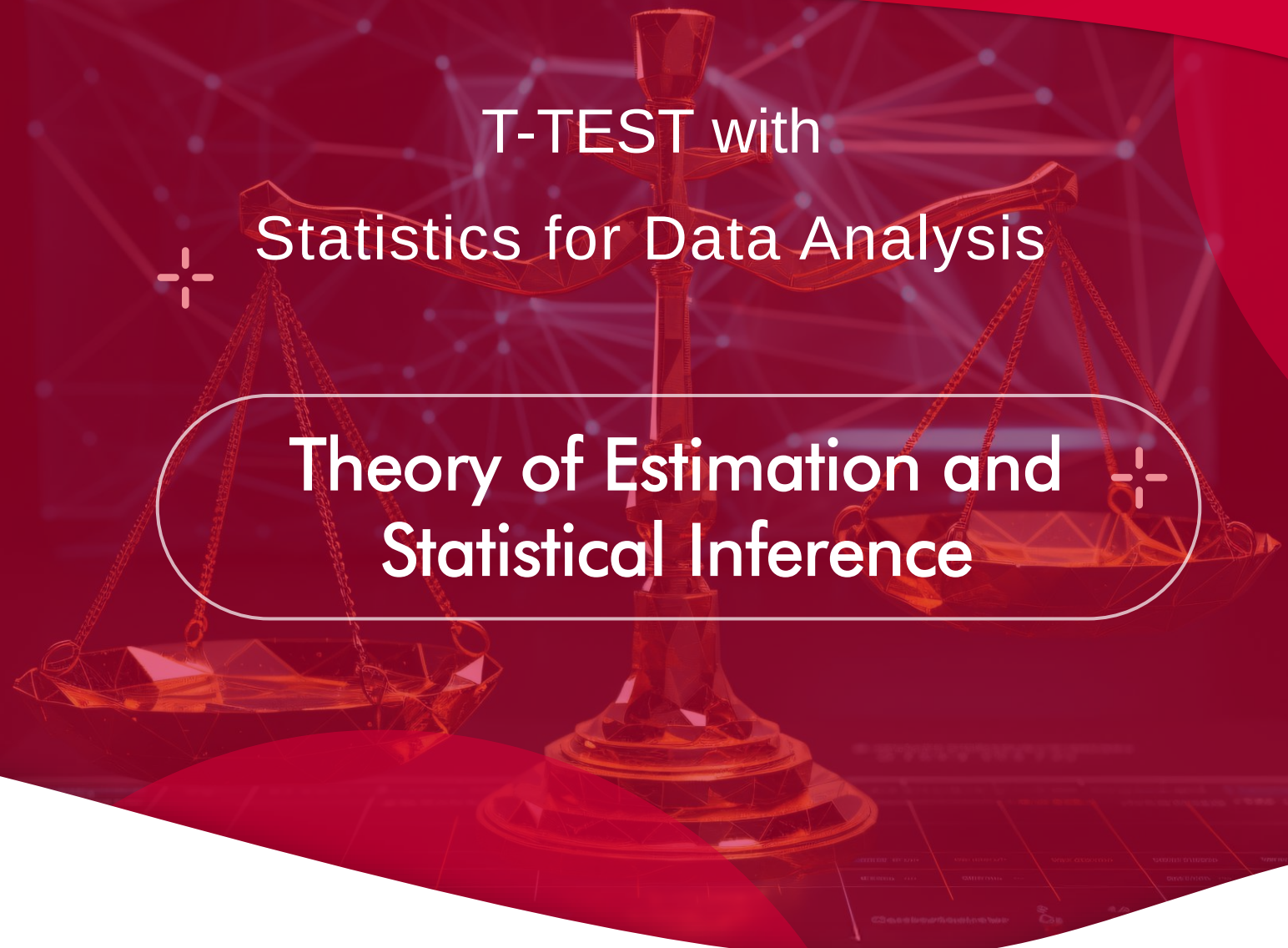
Statistics for Data Analysis

SPSS

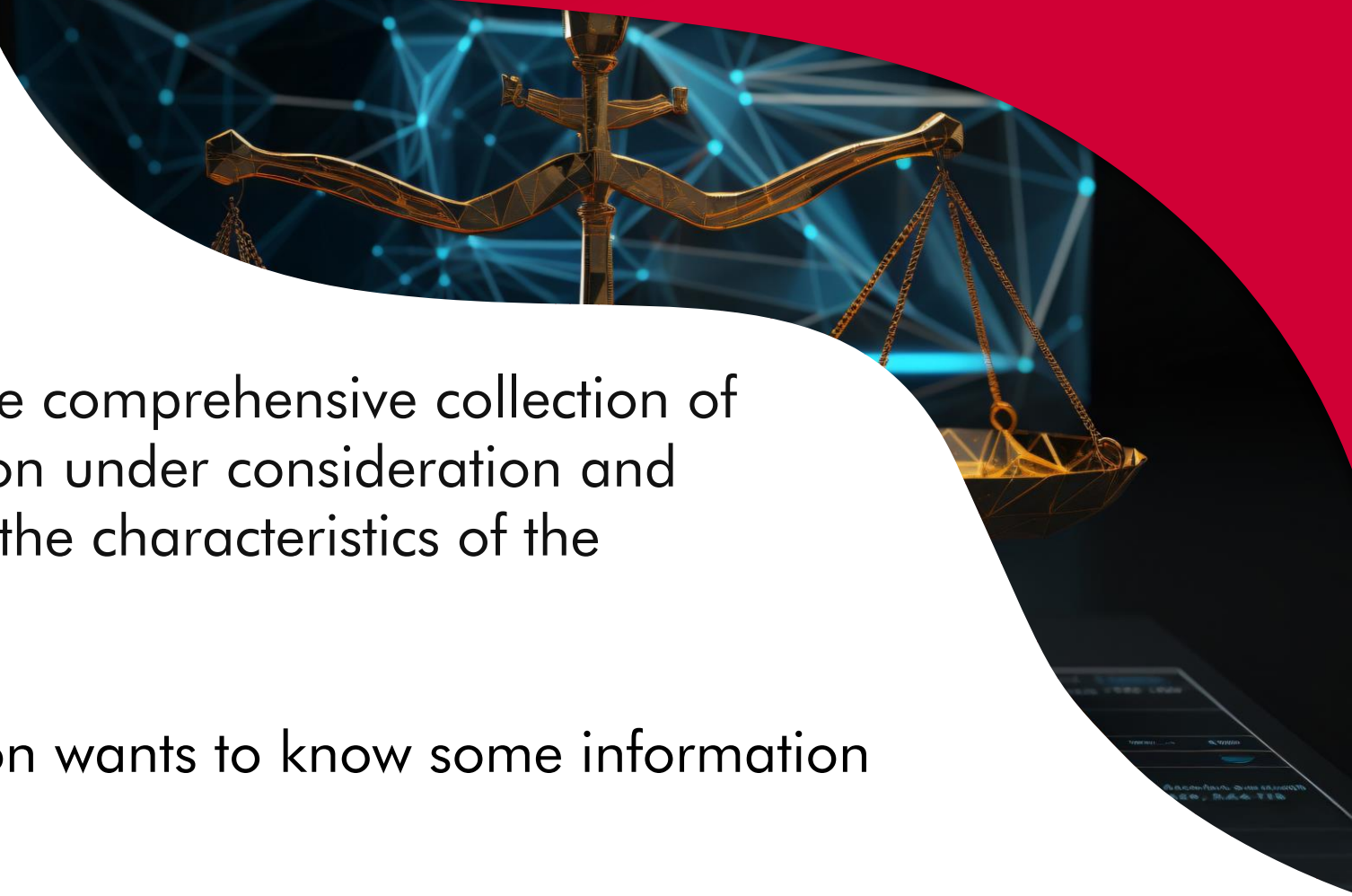
T-TEST with

Statistics for Data Analysis

Theory of Estimation and Statistical Inference



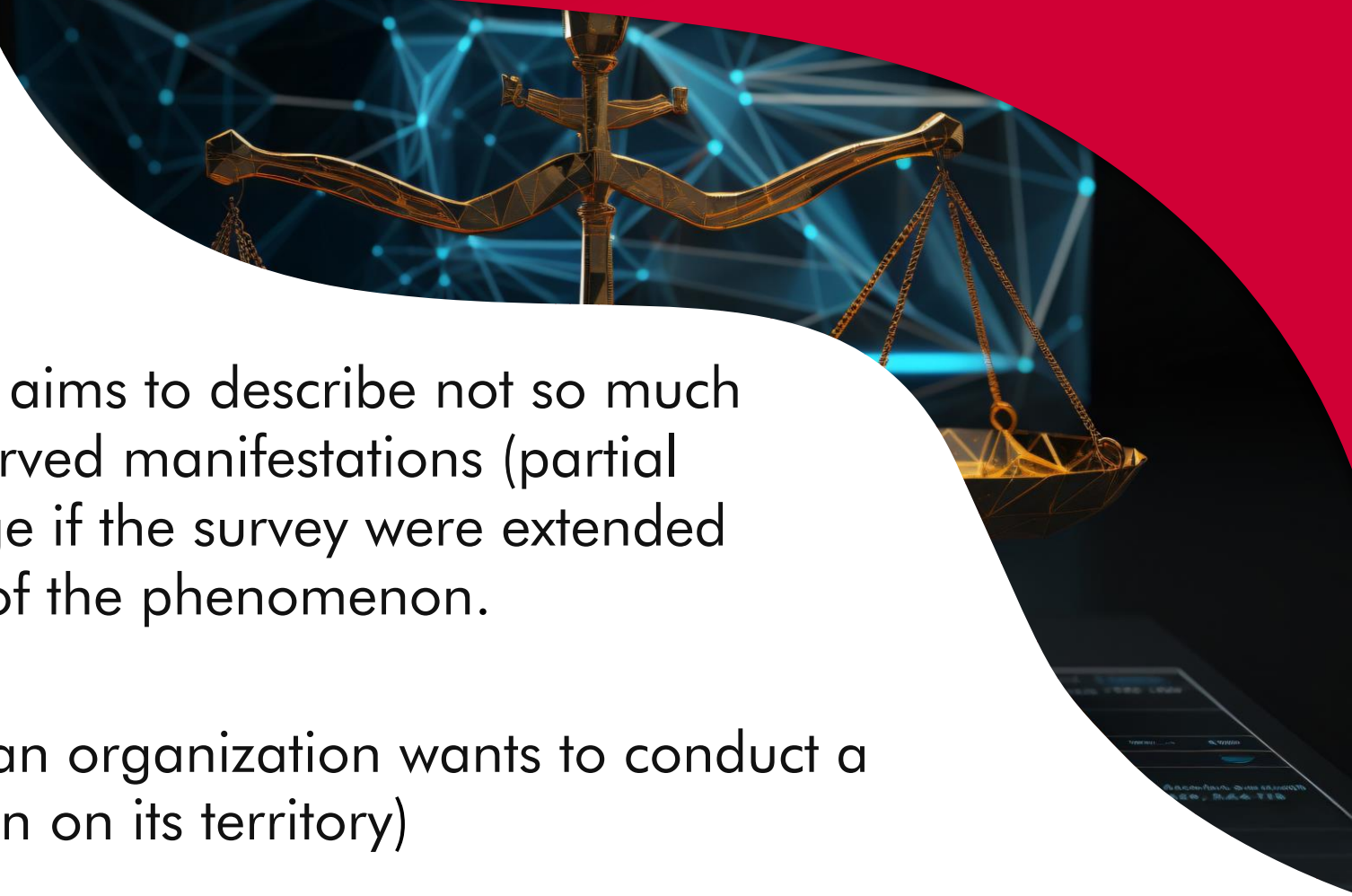
Descriptive statistics



Descriptive statistics deals with the comprehensive collection of manifestations of the phenomenon under consideration and provides a kind of description of the characteristics of the phenomenon itself.

For example, ... (an organisation wants to know some information about its territory)

Inferential Statistics



Inductive or inferential statistics aims to describe not so much what is apparent from the observed manifestations (partial surveys), but what would emerge if the survey were extended to the set of all manifestations of the phenomenon.

For example, ... (For example, an organization wants to conduct a survey on the impact of pollution on its territory)

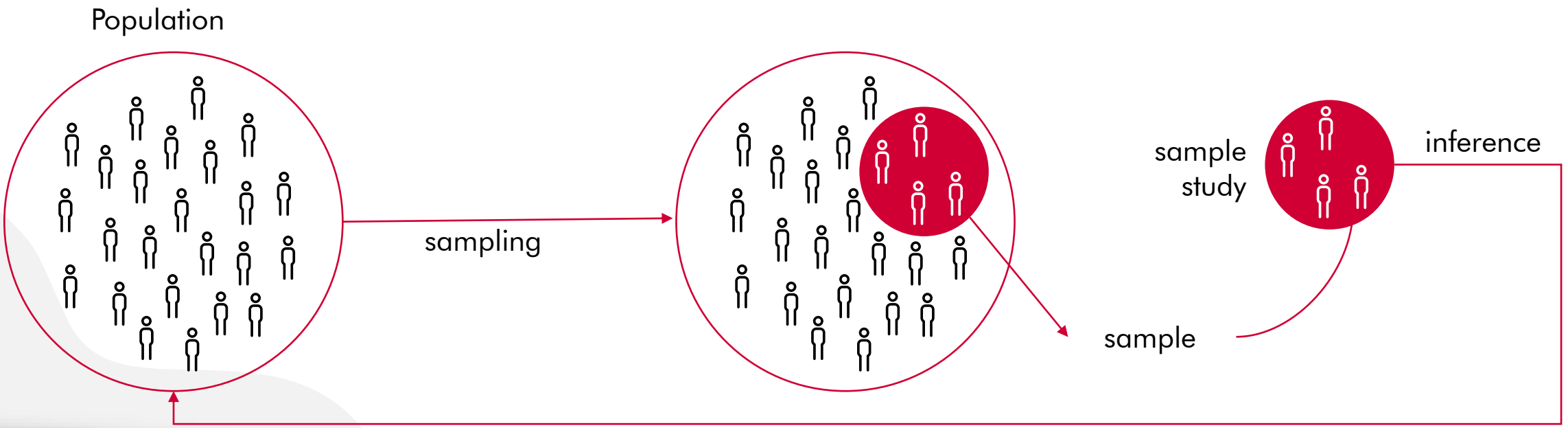
Inferential Statistics

The uncertainty that arises from the partiality of the survey is dominated by the **Theory of Probability**.

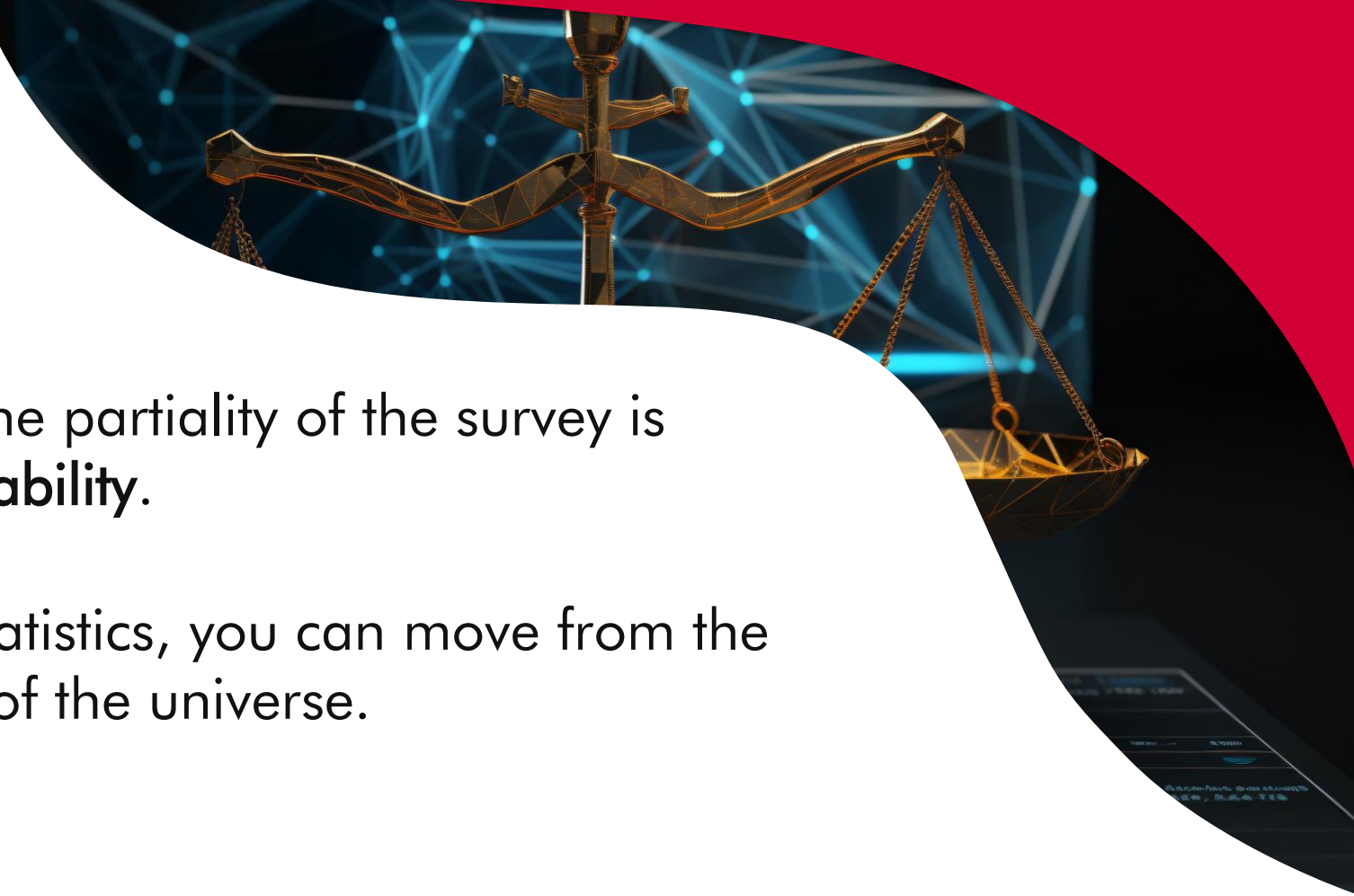
With the methods of inferential statistics, you can move from the knowledge of the sample to that of the universe.

Inferential Statistics

- Rarely ALL the units that make up a population can be studied.
- Therefore, often only a part (SAMPLE) of the population is studied; in order to generalize the results (this process of generalization is called 'inference').



Statistical assumption



The uncertainty that arises from the partiality of the survey is dominated by the **Theory of Probability**.

With the methods of inferential statistics, you can move from the knowledge of the sample to that of the universe.

Logic

Premise

↓

Logical arguments

↓

Contradiction
(reductio ad absurdum)

↓

Conclusions
Falsification of the premise

Statistics

null hypothesis

↓

Application of a test

↓

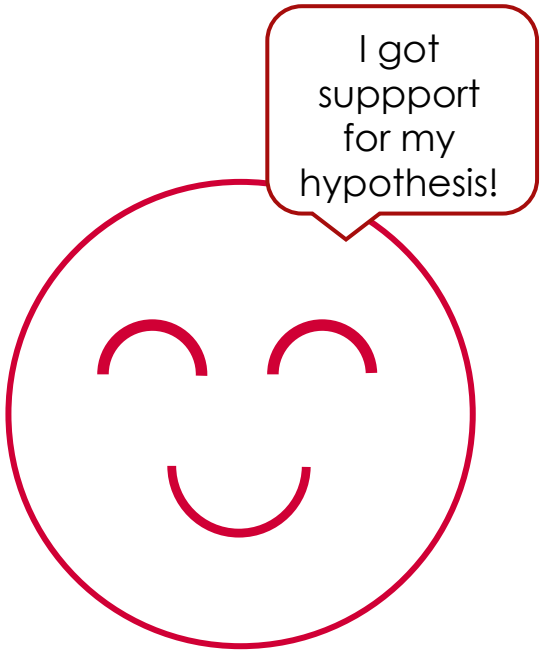
Unlikely outcome ($p<0,05$)

↓

Conclusions

- The hypothesis is not rejected (a not highly improbable result occurred)
- The hypothesis is rejected (the observed result is inconsistent with what is specified in the hypothesis)

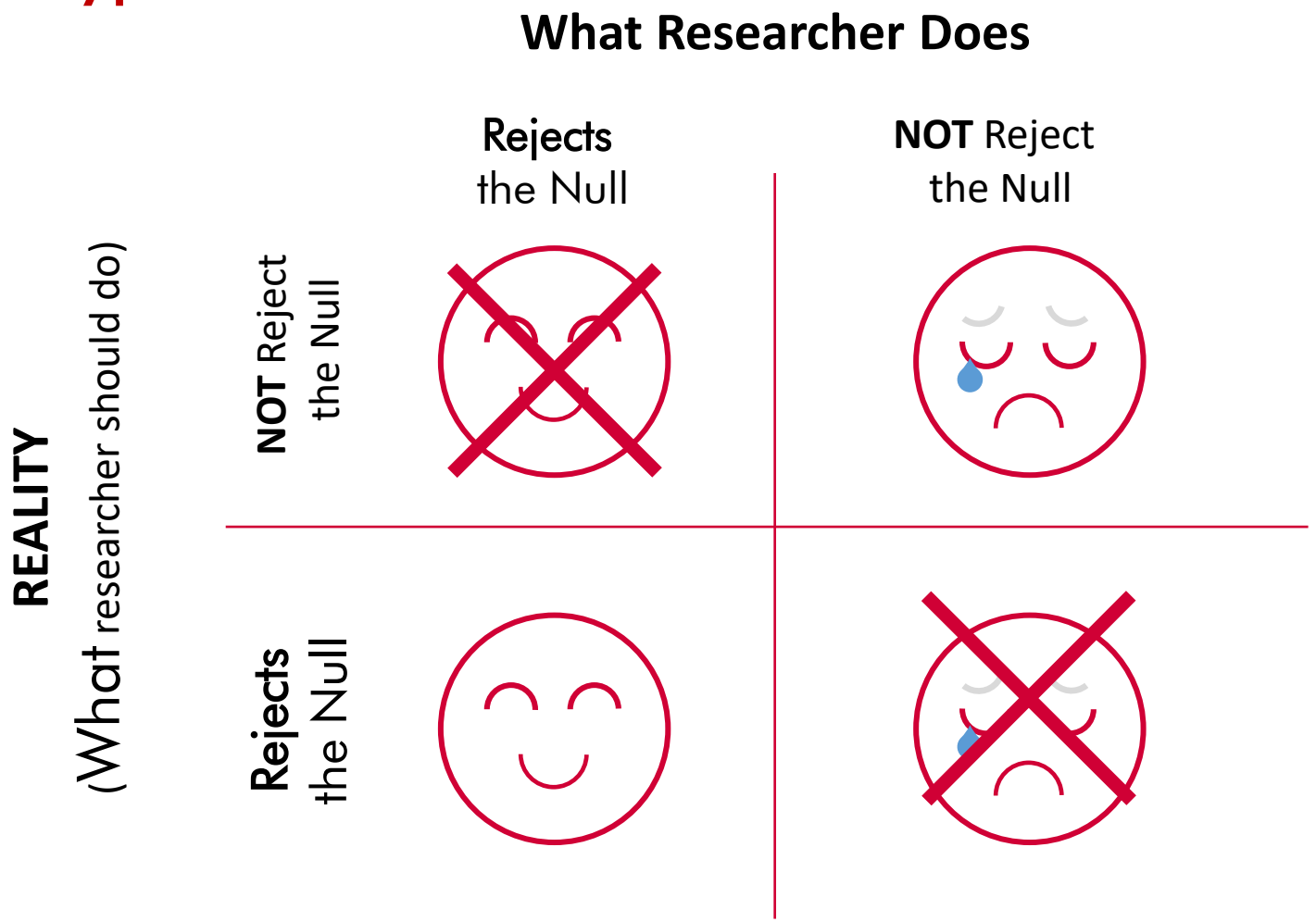
Verification of a Statistical Hypothesis



1. Correct
2. Incorrect (Data Reflects Chance)

1. Correct
2. Incorrect (Didn't Detect Effects)

Verification of a Statistical Hypothesis



Assumptions for parametric tests

- Sample observations must be independent, meaning the sample must be random.
- The observations must belong to populations that are normally distributed.
- The populations must have the same variance (homoscedasticity)
- The observed variables must be measurable on an interval scale

Many parametric tests are robust enough to withstand slight deviations from some of these postulates, especially when the sample size is sufficiently large.

Assumptions for parametric tests

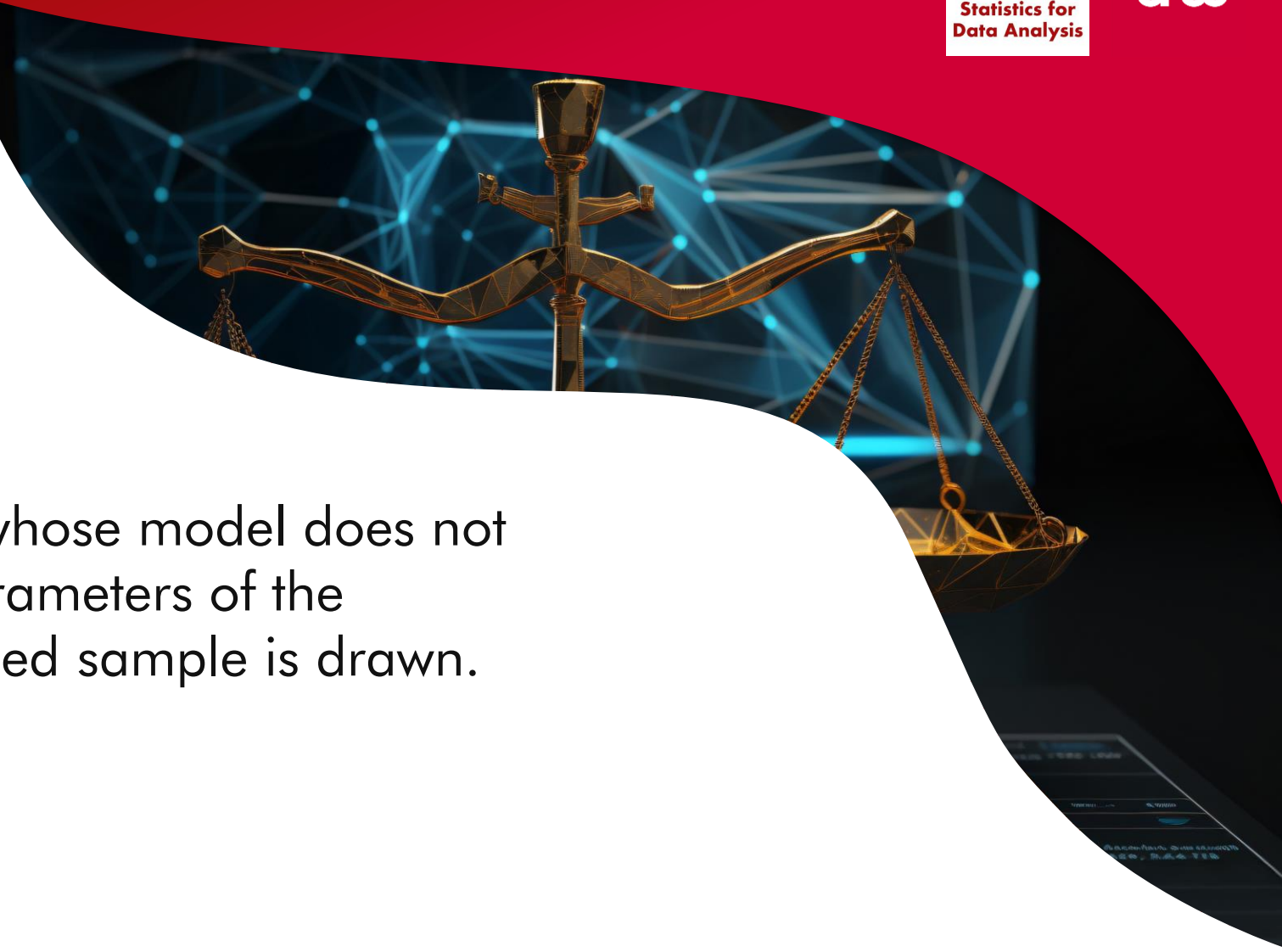
Difference between sample averages

(H_0 takes on this difference = 0)

$$t_{obs} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\hat{s}_{\bar{x}_1 - \bar{x}_2}}$$

Standard Error of the Difference between Sample Averages

(IF both group variances are equal)



NON-parametric tests

A non-parametric test is a test whose model does not specify conditions about the parameters of the population from which the studied sample is drawn.

NON-parametric tests

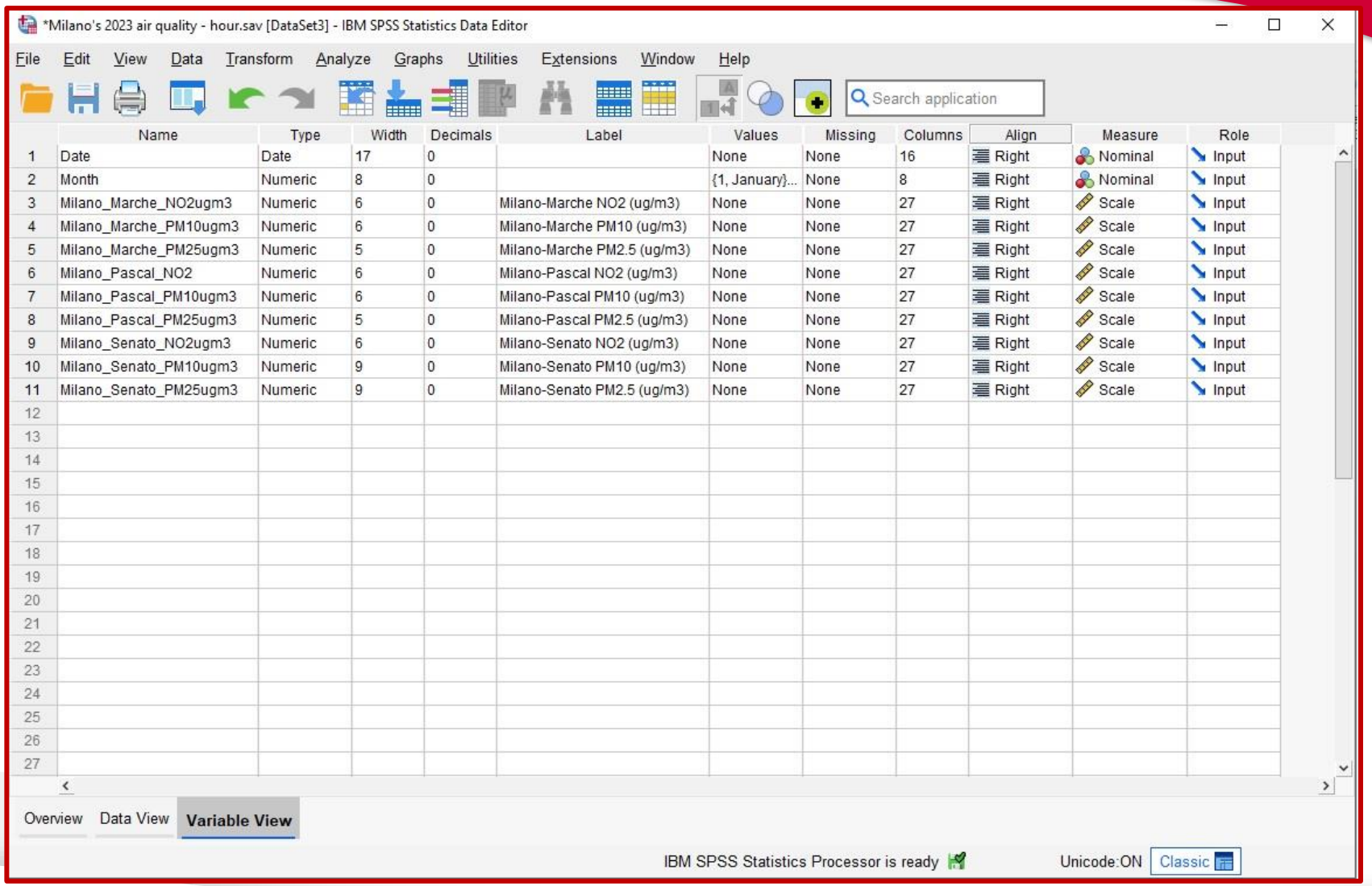
Advantages of non-parametric tests:

- Less restrictive assumptions
- Feasibility of use even with small samples
- Analysis of different populations
- Analysis of ranks or nominal data

Disadvantages of non-parametric tests:

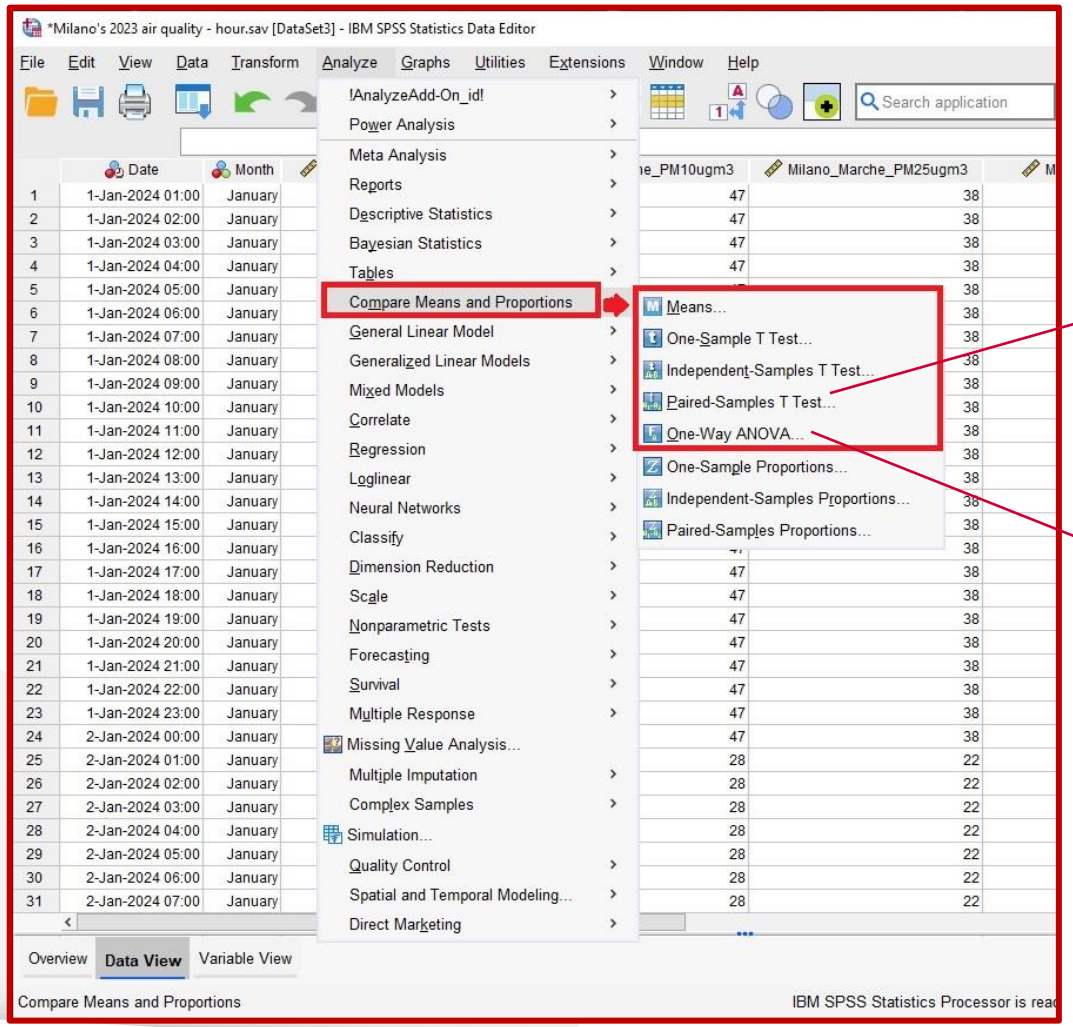
- Less precise with the same amount of information (if the assumptions were verified)

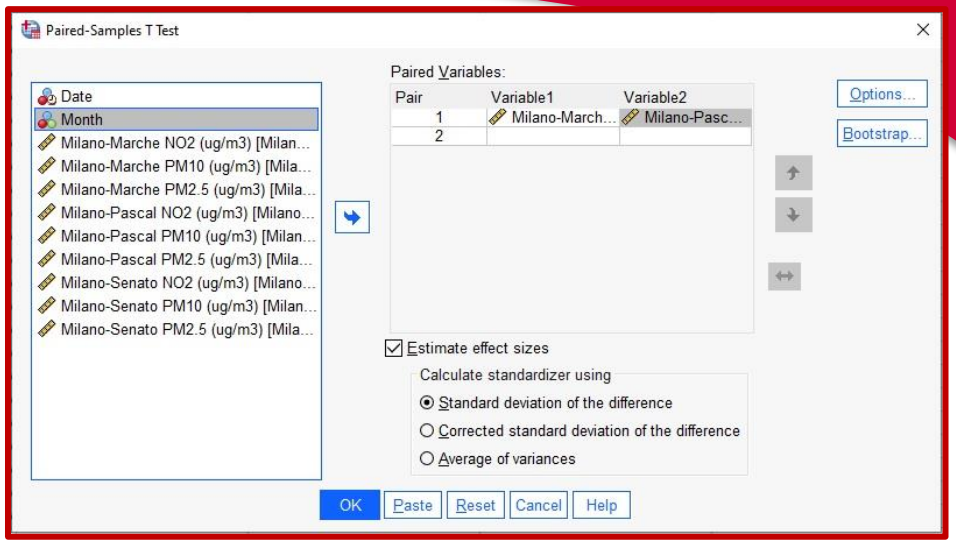
Examples

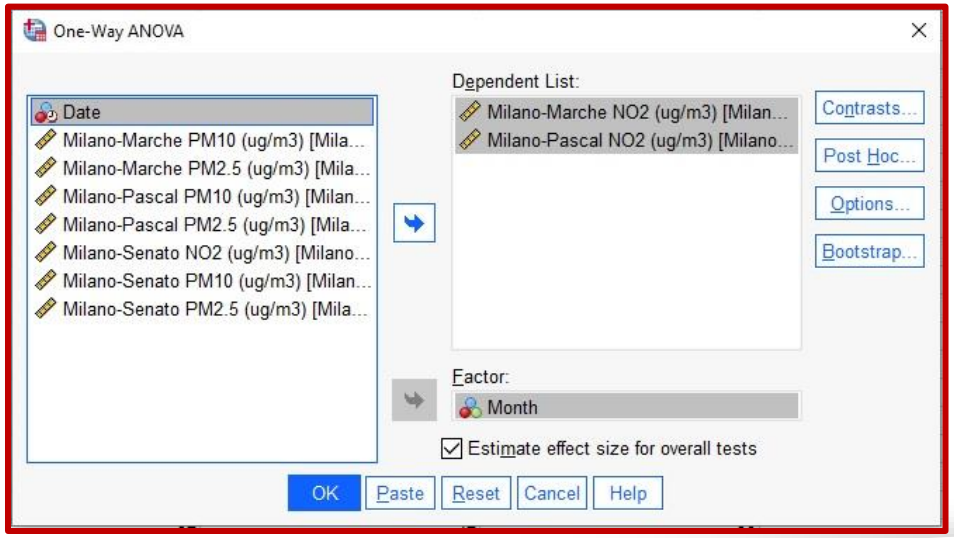


	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Date	Date	17	0		None	None	16	Right	Nominal	Input
2	Month	Numeric	8	0		{1, January}...	None	8	Right	Nominal	Input
3	Milano_Marche_NO2ugm3	Numeric	6	0	Milano-Marche NO2 (ug/m3)	None	None	27	Right	Scale	Input
4	Milano_Marche_PM10ugm3	Numeric	6	0	Milano-Marche PM10 (ug/m3)	None	None	27	Right	Scale	Input
5	Milano_Marche_PM25ugm3	Numeric	5	0	Milano-Marche PM2.5 (ug/m3)	None	None	27	Right	Scale	Input
6	Milano_Pascal_NO2	Numeric	6	0	Milano-Pascal NO2 (ug/m3)	None	None	27	Right	Scale	Input
7	Milano_Pascal_PM10ugm3	Numeric	6	0	Milano-Pascal PM10 (ug/m3)	None	None	27	Right	Scale	Input
8	Milano_Pascal_PM25ugm3	Numeric	5	0	Milano-Pascal PM2.5 (ug/m3)	None	None	27	Right	Scale	Input
9	Milano_Senato_NO2ugm3	Numeric	6	0	Milano-Senato NO2 (ug/m3)	None	None	27	Right	Scale	Input
10	Milano_Senato_PM10ugm3	Numeric	9	0	Milano-Senato PM10 (ug/m3)	None	None	27	Right	Scale	Input
11	Milano_Senato_PM25ugm3	Numeric	9	0	Milano-Senato PM2.5 (ug/m3)	None	None	27	Right	Scale	Input
12											
13											
14											
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											
27											

Examples


*Milano's 2023 air quality - hour.sav [DataSet3] - IBM SPSS Statistics Data Editor
File Edit View Data Transform Analyze Graphs Utilities Extensions Window Help
Search application
!AnalyzeAdd-On_id!
Power Analysis
Meta Analysis
Reports
Descriptive Statistics
Bayesian Statistics
Tables
Compare Means and Proportions
General Linear Model
Generalized Linear Models
Mixed Models
Correlate
Regression
Loglinear
Neural Networks
Classify
Dimension Reduction
Scale
Nonparametric Tests
Forecasting
Survival
Multiple Response
Missing Value Analysis...
Multiple Imputation
Complex Samples
Simulation...
Quality Control
Spatial and Temporal Modeling...
Direct Marketing
Means...
One-Sample T Test...
Independent-Samples T Test...
Paired-Samples T Test...
One-Way ANOVA...
One-Sample Proportions...
Independent-Samples Proportions...
Paired-Samples Proportions...
Date
Month
Milano_Marche_PM25ugm3
Overview
Data View
Variable View
Compare Means and Proportions
IBM SPSS Statistics Processor is rea


Paired-Samples T Test
Date
Month
Milano-Marche NO2 (ug/m3) [Milan...
Milano-Marche PM10 (ug/m3) [Mila...
Milano-Marche PM2.5 (ug/m3) [Mila...
Milano-Pascal NO2 (ug/m3) [Milano...
Milano-Pascal PM10 (ug/m3) [Milan...
Milano-Pascal PM2.5 (ug/m3) [Mila...
Milano-Senato NO2 (ug/m3) [Milano...
Milano-Senato PM10 (ug/m3) [Milan...
Milano-Senato PM2.5 (ug/m3) [Mila...
Paired Variables:
Pair
Variable1
Variable2
1
Milano-March...
Milano-Pasc...
2
Options...
Bootstrap...
Estimate effect sizes
Calculate standardizer using
Standard deviation of the difference
Corrected standard deviation of the difference
Average of variances
OK
Paste
Reset
Cancel
Help


One-Way ANOVA
Date
Milano-Marche PM10 (ug/m3) [Mila...
Milano-Marche PM2.5 (ug/m3) [Mila...
Milano-Pascal PM10 (ug/m3) [Milan...
Milano-Pascal PM2.5 (ug/m3) [Mila...
Milano-Senato NO2 (ug/m3) [Milano...
Milano-Senato PM10 (ug/m3) [Milan...
Milano-Senato PM2.5 (ug/m3) [Mila...
Dependent List:
Milano-Marche NO2 (ug/m3) [Milan...
Milano-Pascal NO2 (ug/m3) [Milano...
Contrasts...
Post Hoc...
Options...
Bootstrap...
Factor:
Month
Estimate effect size for overall tests
OK
Paste
Reset
Cancel
Help