

DS 09/21 Statistics Session-1

DS Statistics Session-1

Training Clarusway

Pear Deck - September 20, 2021 at 1:39PM

Part 1 - Summary

Use this space to summarize your thoughts on the lesson

Part 2 - Responses

Slide 1



Use this space to take notes:

Slide 2	Your Response
<p>Did you finish Statistics (Introduction, Types of Data, Level of Measurements) pre-class activity?</p>  <p><small>Students choose an option</small></p> <p>Pear Deck Interactive Slide <small>Do not remove this bar</small></p>	<p>You Chose</p> <ul style="list-style-type: none"> • I finished completely. <p>Other Choices</p> <ul style="list-style-type: none"> • I finished partially. • No, I didn't finish.

Use this space to take notes:

Slide 3



SUCCESS NEEDS
PREPARATION

CLARUSWAY®
www.clarusway.com

3

Use this space to take notes:

Slide 4



CLARUSWAY®
way to motivate yourself

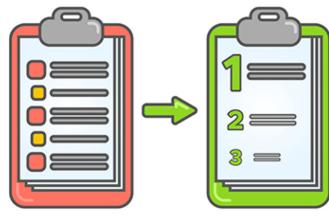
Use this space to take notes:

Slide 5



Use this space to take notes:

Slide 6



Prioritize

CLARUSWAY®
way to success yourself

8

Use this space to take notes:

Slide 7



CLARUSWAY®
way to success yourself

9

Use this space to take notes:

Slide 8

Scope of the Course

- ▶ Fundamentals of Statistics-1
- ▶ Fundamentals of Statistics-2
- ▶ Probability
- ▶ Probability Distributions
- ▶ Central Limit Theorem and Confidence Intervals
- ▶ Hypothesis Testing

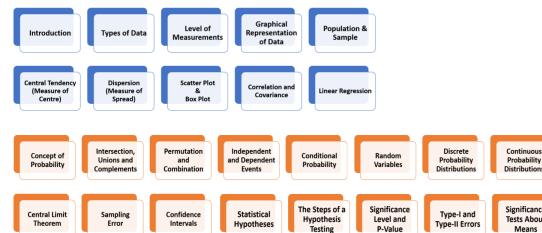


8

Use this space to take notes:

Slide 9

Course Content



9

Use this space to take notes:

Slide 10

► Sources

CLARUSWAY LMS (Pre-class Activities)

- Content
- Lets Practice, Check Yourself
- Assignments



Textbooks

- Brownlee J., Statistical Methods for Machine Learning.
- Rumsey, D. J. (2010). Statistics essentials for dummies. John Wiley & Sons.
- Wackerly, D., Mendenhall, W., & Scheaffer, R. L. (2014). Mathematical statistics with applications. Cengage Learning



CLARUSWAY®
WE TRY TO REINVENT YOURSELF



https://www.clarusway.com

10

Link(s) on this slide:

- <http://onlinestatbook.com/>
- <https://www.youtube.com/c/joshstarmer>
- <https://www.khanacademy.org/math/statistics-probability>

Use this space to take notes:

Slide 11

Table of Contents



- ▶ What is "Statistics"?
- ▶ Why Should You Learn Statistics?
- ▶ Probability vs. Statistics
- ▶ Stats with Python
- ▶ Types of Data
- ▶ Level of Measurements

CLARUSWAY®

11

Use this space to take notes:

Slide 12



1 ▶ What is “Statistics”?

CLARUSWAY®
WAY TO REINVENT YOURSELF

Use this space to take notes:

Slide 13	Your Response
<p>▶ What is “Statistics”?</p>  <p>Statistics is the grammar of science.</p> <p>-- Karl Pearson</p> <p> Students, write your response!</p> <p>Pear Deck Interactive Slide</p>	

Use this space to take notes:

Slide 14

► What is "Statistics"?



Most fundamentally, **statistics** is all about data.

- COLLECT
- CHARACTERIZE
- ANALYZE
- PRESENT

CLARUSWAY®
data to informed decisions



14

Use this space to take notes:

Slide 15

► What is "Statistics"?



WIKIPEDIA



CLARUSWAY®
data to informed decisions

Statistics is the discipline that concerns the collection, organization, analysis, interpretation and presentation of data.

A branch of mathematics dealing with the collection, analysis, interpretation, and presentation of masses of numerical data.

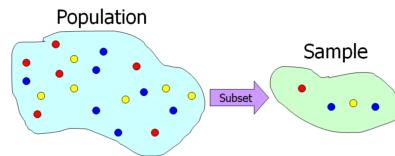
All the authors imply that statistics is a theory of information, with inference making as its objective.

15

Use this space to take notes:

Slide 16

► Parameters & Statistics



- Populations have Parameters (like μ , σ^2 , θ , p)
- Samples have Statistics, functions of observed data, like \bar{x} , \tilde{x} , s^2 , $\hat{\theta}$, \hat{p}

CLARUSWAY®
data to empower business

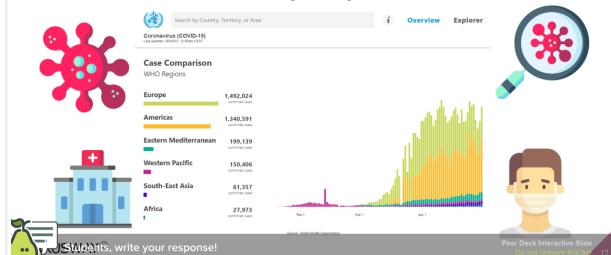
18

Use this space to take notes:

Slide 17

Your Response

► What are some examples of statistics in everyday life?



Use this space to take notes:

Slide 18

► What are some examples of statistics in everyday life?



CLARUSWAY®
www.clarusway.com

Weather Forecasts

Stock Market

Predicting Disease

Medical Studies

Insurance

Consumer Goods



18

Use this space to take notes:

Slide 19

Your Response

► Relation of Statistics with other Sciences ➤

Economy
Psychology
Medicine
Sociology
History



Statistics

?
Psychometrics
?
Sociometry
Cliometrics



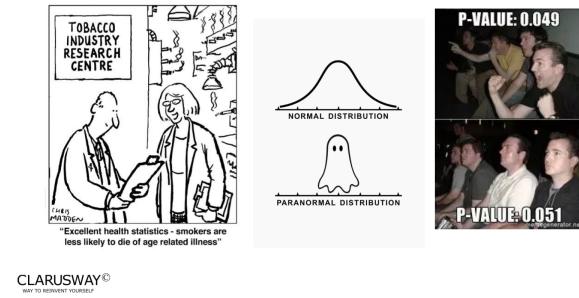
Students, write your response!

Pear Deck Interactive Slide

Use this space to take notes:

Slide 20

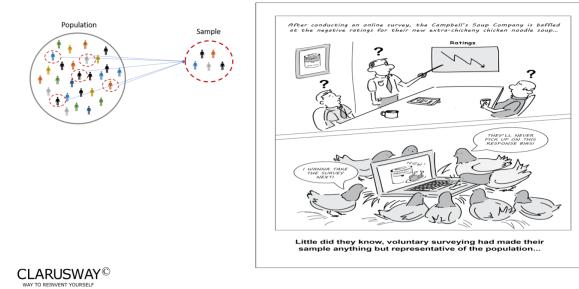
▶ Funny Statistics



Use this space to take notes:

Slide 21

▶ Sampling Bias



Use this space to take notes:

Slide 22



Why Should You Learn Statistics?

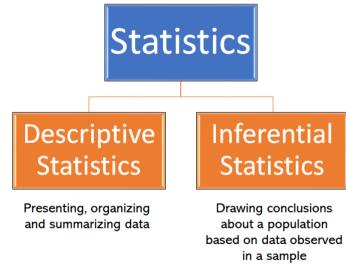


CLARUSWAY®
WAY TO REINVENT YOURSELF

Use this space to take notes:

Slide 23

► Descriptive vs Inferential Statistics ➤



CLARUSWAY®
WAY TO REINVENT YOURSELF

23

Use this space to take notes:

Slide 24

► Why "Statistics"?

- Should I run for the bus?
- Which stock should I buy?
- Which products should I list in Website?
- Should I take this medication?
- Should I have my children vaccinated?

CLARUSWAY®
data to empower decisions



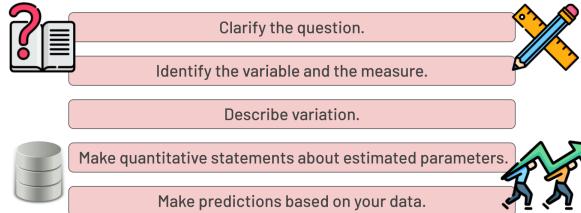
If you don't take this
medication, there is a
95% chance that you
will die.

24

Use this space to take notes:

Slide 25

► Statistical Process



CLARUSWAY®
data to empower decisions

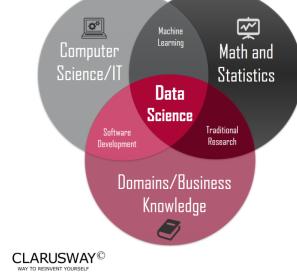
25

Use this space to take notes:

Slide 26

► Data Science vs. Statistics

“ A Data Scientist is one who knows more statistics than a programmer and more programming than a statistician ”



CLARUSWAY®
data to empower business

28

Use this space to take notes:

Slide 27

► Data Science vs. Statistics

- A data scientist makes hundreds of decisions every day.
- Many of these decisions require a strong foundation in math and statistics.
- Data science requires descriptive statistics and probability theory, at a minimum.

CLARUSWAY®
data to empower business

27

Use this space to take notes:

Slide 28

Data Science vs. Statistics

Data science	Statistics
<ul style="list-style-type: none">▸ multidisciplinary field▸ uses scientific methods, processes, and systems▸ extract knowledge from data▸ making predictions▸ optimizing search of large databases.▸ engineering background	<ul style="list-style-type: none">▸ mathematically-based field▸ collect and interpret quantitative data▸ focused on drawing conclusions about the world▸ math departments

CLARUSWAY®
www.clarusway.com

28

Use this space to take notes:

Slide 29	Your Response
<p>What is Data Science?</p> <p>Why are we learning Statistics?</p> <p>Pear Deck Pear Deck Interactive Slide Click & Answer this! ▶</p> <p> Students, write your response!</p>	

Use this space to take notes:

Slide 30



3

Probability vs. Statistics

CLARUSWAY®
WAY TO REINVENT YOURSELF

Use this space to take notes:

Slide 31

▶ Probability vs. Statistics

- ▶ In probability, the randomization mechanism is assumed or known, and calculations follow
 - ▷ E.g., Given a fair coin, what's the chance of observing 7 heads out of 10 coin flips?
- ▶ In statistics, the outcome is known and used to infer something about the mechanism
 - ▷ E.g., Given that I've just observed 7 heads out of 10 coin flips, is the coin fair?

CLARUSWAY®
WAY TO REINVENT YOURSELF



31

Use this space to take notes:

Slide 32



4

Stats with Python

CLARUSWAY®
WAY TO REINVENT YOURSELF

Use this space to take notes:

Slide 33

▶ Stats with Python

CLARUSWAY®
WAY TO REINVENT YOURSELF

33

Use this space to take notes:

Slide 34



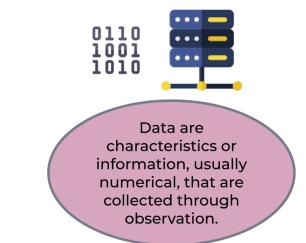
5 Types of Data

CLARUSWAY®
WAY TO REINVENT YOURSELF

Use this space to take notes:

Slide 35

► What is Data?



CLARUSWAY®
WAY TO REINVENT YOURSELF

35

Use this space to take notes:

Slide 36

► What is Data?

The screenshot shows a Microsoft Excel spreadsheet titled "Tutorial 18 - Nested IF V3 Dates.xlsx". The table has columns for Date, Movie, Tickets Sold, % of Capacity Sold, Tickets Remaining, % of Tickets to Sell, Status, and Description. The data includes showtimes for Grease, Jaws, Citizen Kane, and The Wizard of Oz at Fred's Movie Emporium.

A	B	C	D	E	F	G
1	Fred's Classic Movie Events					2
2	Venue	Fred's Movie Emporium		Today	13/12/2012	
3	Capacity	150				
4	Date	Movie	Tickets Sold	% of Capacity Sold	Tickets Remaining	% of Tickets to Sell
5	Wednesday 11 May 2011	Grease	50	33%	100	67%
6	Sunday 15 May 2011	Jaws	150	100%	0	0%
7	Monday 23 May 2011	Citizen Kane	105	70%	45	30%
8	Wednesday 01 Jun 2011	The Wizard of Oz	150	100%	0	0%
9	Friday 10 Jun 2011	Singin' In the Rain	85	57%	65	43%
10						Promote
11						Last Few Seats
12						Last Few Seats
13						

CLARUSWAY®
www.clarusway.com



38

Use this space to take notes:

Slide 37

► Statistical Data Analysis

- ▶ Statistics is basically a science that involves data collection, data interpretation and finally, data validation.
- ▶ Statistical data analysis is a procedure of performing various statistical operations.
- ▶ It is a kind of quantitative research, which seeks to quantify the data, and typically, applies some form of statistical analysis.

CLARUSWAY®
www.clarusway.com



37

Use this space to take notes:

Slide 38

► Statistical Data Analysis



CLARUSWAY®
data to drive your business

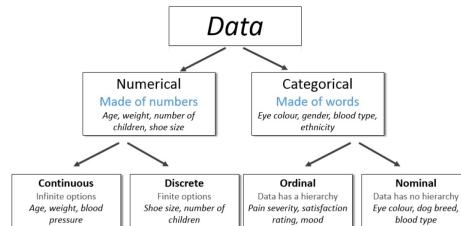
- ▶ Data in statistical data analysis consists of variable(s).
- ▶ Sometimes the data is univariate or multivariate.
- ▶ Depending upon the number and types of variables, the researcher performs different statistical techniques.
- ▶ If the data in statistical data analysis is multiple in numbers, then several multivariates can be performed.

38

Use this space to take notes:

Slide 39

► Types of Data



CLARUSWAY®
data to drive your business

39

Use this space to take notes:

Slide 40

► Numerical Data

Continuous Data

- ▶ Continuous data can have an infinite continuum of possible values.
 - ▷ height
 - ▷ weight
 - ▷ age
 - ▷ the amount of time it takes to complete an assignment

CLARUSWAY®
www.clarusway.com

Discrete Data

- ▶ Any variable with a finite number of possible values is discrete.
 - ▷ the number of pets in a household
 - ▷ the number of children in a family
 - ▷ the number of foreign languages in which a person is fluent

40

Use this space to take notes:

Slide 41

► Categorical Data

Ordinal Data

- ▶ Ordinal data requires an order
 - ▷ small, medium, large
 - ▷ good, average, poor
 - ▷ strongly agree, agree, disagree
- ▶ The distance between ordered categories is not measurable.
- ▶ No arithmetic can be done with the ordinal data as they show sequence only.

CLARUSWAY®
www.clarusway.com

Nominal Data

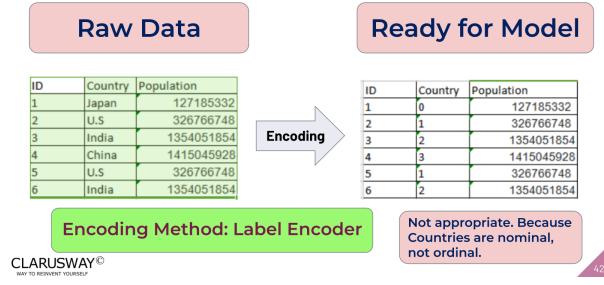
- ▶ Nominal data simply names something without an order being given.
 - ▷ employee's status
 - ▷ color
 - ▷ race
- ▶ Data obtained on nominal scale is in terms of frequency.

41

Use this space to take notes:

Slide 42

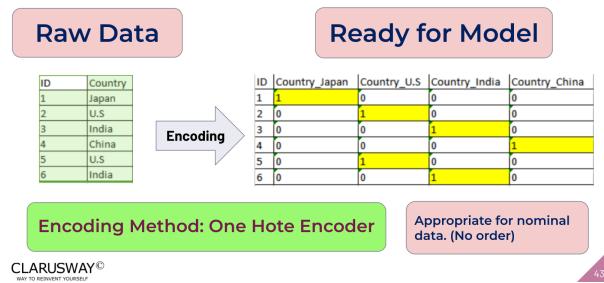
► Categorical Data in ML Models ➤



Use this space to take notes:

Slide 43

► Categorical Data in ML Models ➤



Use this space to take notes:

Slide 44

Your Response

You Chose

- **False**

Other Choices

Which variable is categorical?

Height

Race

Students choose an option

Pear Deck Interactive Slide
Do not remove this bar

- True

Use this space to take notes:

Slide 45



6

Level of Measurements

CLARUSWAY®
BORN TO INNOVATE YOURSELF

Use this space to take notes:

Slide 46

► Level of Measurement

Level of Measurement

A classification that describes the nature of information within the values assigned to variables.

Type	Measure property	Mathematical operators
Nominal	Classification, membership	=, ≠
Ordinal	Comparison, level	>, <
Interval	Difference, affinity	+, -
Ratio	Magnitude, amount	×, /

CLARUSWAY®
www.clarusway.com.br

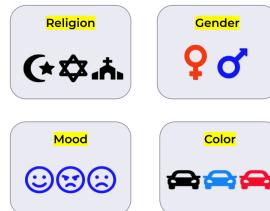
48

Use this space to take notes:

Slide 47

► Nominal Level of Measurement

- No order
- Classification
- Membership



CLARUSWAY®
www.clarusway.com.br

47

Use this space to take notes:

Slide 48

► Ordinal Level of Measurement

- ▶ Ordered
- ▶ Not actual value
- ▶ Comparison
- ▶ Level



CLARUSWAY®
data to empower learning

48

Use this space to take notes:

Slide 49

Your Response

► Let's Practice



What is the level of measurement for the variable *type of housing*?

- ▶ Apartment
- ▶ Condo
- ▶ Town home
- ▶ Single family home



You Chose

- **Nominal**

Other Choices

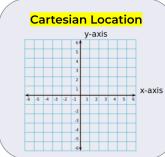
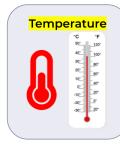
- Ordinal

Use this space to take notes:

Slide 50

► Interval Level of Measurement

- ▶ Ranked
- ▶ Measured
- ▶ Arbitrary zero
- ▶ Difference



CLARUSWAY®
used to remember knowledge

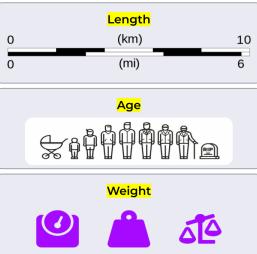
50

Use this space to take notes:

Slide 51

► Ratio Level of Measurement

- ▶ Ranked
- ▶ Measured
- ▶ True zero
- ▶ Magnitude



CLARUSWAY®
used to remember knowledge

51

Use this space to take notes:

Slide 52

Your Response

You Chose
• **Ratio**

Other Choices

▶ Let's Practice



- Interval

What is the level of measurement for the variable type of *Gross Domestic Product per capita*?



Peer Deck Interactive Slide
Do not remove this bar

Use this space to take notes:

Slide 53

▶ Level of Measurement



Variables can be conceptualized and operationalized in any number of ways.

Did you find the course challenging?



How challenging was the course?

NOT AT ALL SOMEWHAT VERY

In a range from 0 to 100,
how challenging was this course?



53

Use this space to take notes:

Slide 54

Kahoot!

CLARUSWAY®
way to innovative learning

54

Link(s) on this slide:

- <https://create.kahoot.it/details/jira-software/926c1447-fbc1-4be8-88e0-9038e2574f3e>

Use this space to take notes:

Slide 55

Your Response

How well did you like this lesson?

Students, drag the icon!

Pear Deck

Do not remove this bar

How well did you like this lesson?

Students, drag the icon!

Pear Deck

Do not remove this bar

Use this space to take notes:



Recap!

CLARUSWAY®
WAY TO REINVENT YOURSELF

58

Use this space to take notes:



THANKS!

Any questions?

You can find us at:

- ▶ mike_a@clarusway.com
- ▶ jason@clarusway.com



57

Use this space to take notes:



CLARUSWAY®
www.clarusway.com

58

Use this space to take notes: