



OceanTeacher
GLOBAL ACADEMY



**Online Training Course
On**

Fundamental of Statistics for Oceanographers using Excel

Organized by

**International Training Centre for Operational Oceanography (ITCOO)
ESSO-INCOIS, MoES, Hyderabad, India**

Jan 04 - 07, 2021

ravikumar.jha@incois.gov.in

Terminology

- Data: Information + Noise
- Population: Any individual or items through a conclusion can be drawn
- Sample: Subset of a population
 - Small : < 30 quantity
 - Large: ≥ 30 quantity
- Parameter: Measurements describe characteristic of a population
 - Ex: Mean, SD,... etc
- Statistic: Numerical measure describing the characteristic of a sample

***Which is better, population or sample?

STATISTICS

```
graph TD;
  S[STATISTICS] --> D[Descriptive];
  S --> I[Inferential];
  D --> V[Variable];
  V --> Dis[Discrete];
  V --> Con[Continuous];
  I --> Est[Estimation];
  I --> Hyp[Hypothesis];
```

Descriptive

- Collect, summarize, & describe data
- Information of a small sample

Variable

Discrete

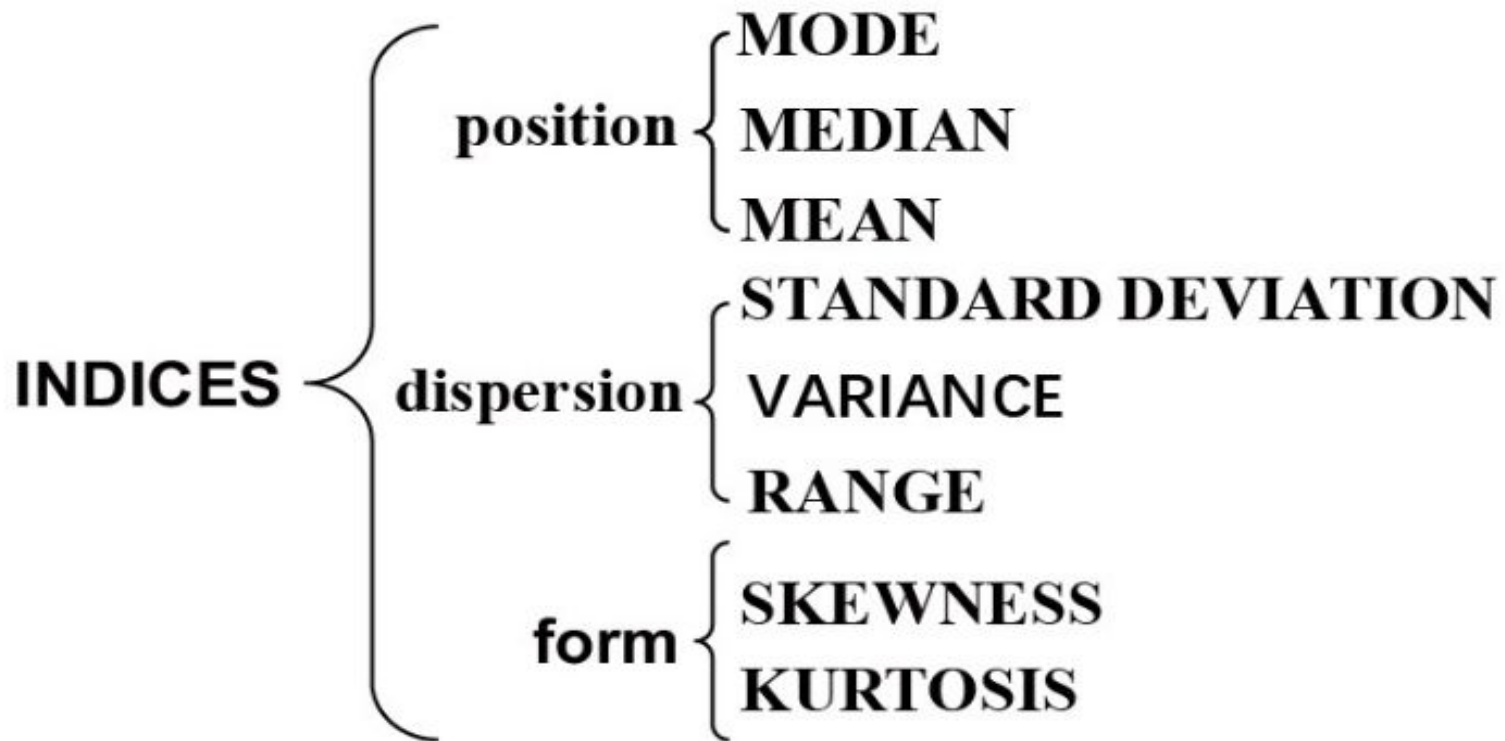
Continuous

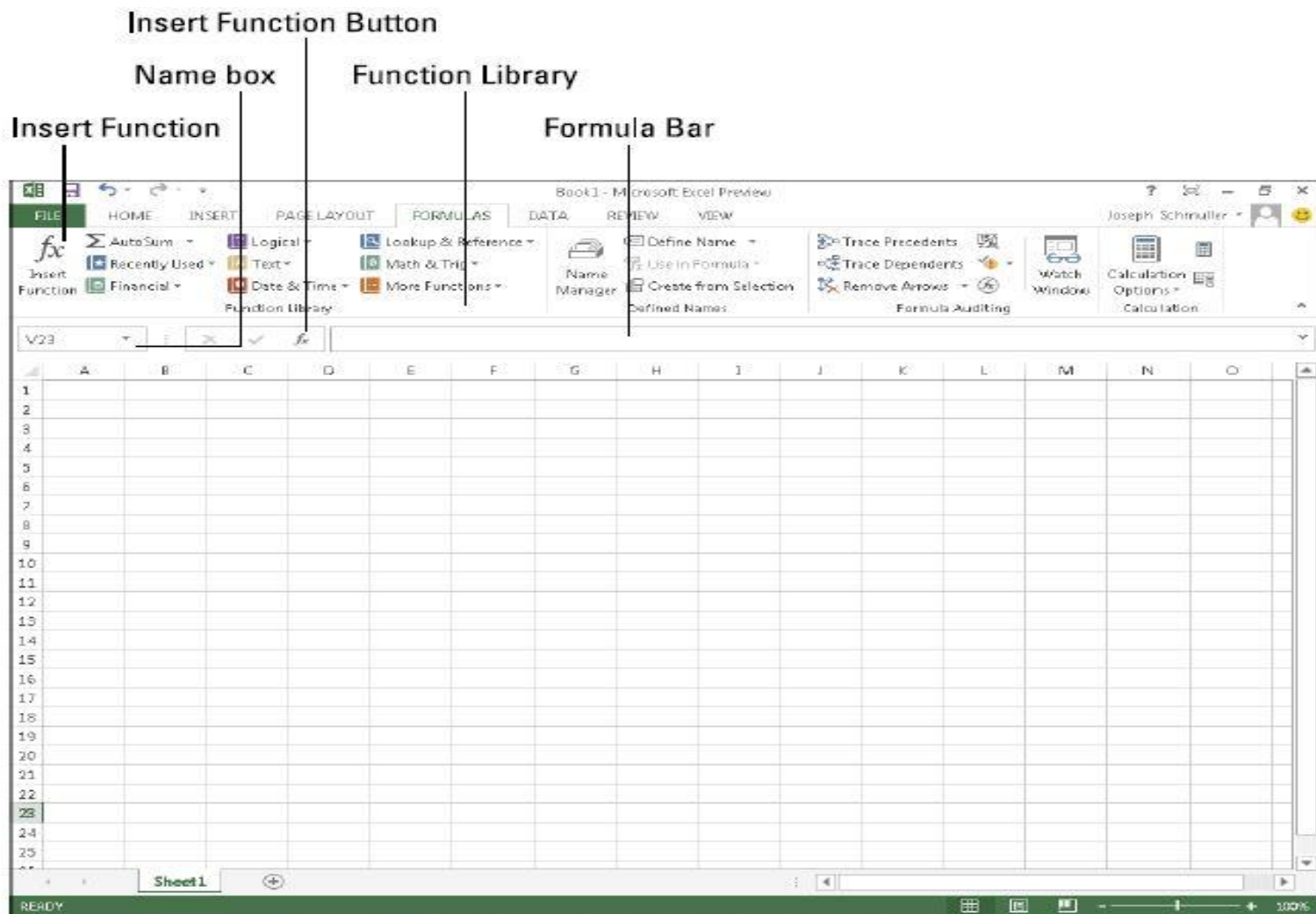
Inferential

- Get information from population through sample
- Conclusion, & decision making (model based)

Estimation

Hypothesis





Data in Excel

- Open data (.txt) in excel
- Sort & filter
- Data formatting & outliers
- Handling duplicates
- Random numbers
- Add-Ins
- Random sampling
- Functions

- Excel panel > open > data(.txt) > Delimited(tab/comma/space/other)
- Hide/Delete cells: Unwanted cells can be hidden or deleted (not recommend)
- Home > Sort & Filter (make your choice)
- Data > Select data column(s) > Filter > Select through cursor > Deselect
- Home > Conditional formatting (used to highlight the odd values)
- Duplicates: Data > select a column > Remove duplicates > expand the selection > OK (remove duplicates from all columns)

Duplicates: Data > select a column > Remove duplicates > with the current selection > OK (remove duplicates from the selected column)

- Add-Ins : Data Analysis Package

Office > Excel options > Add-Ins > Manage (excel add-ins) > Go > Select Analysis tool > Ok (Location: Data > Data Analysis)

- Random number

Go to formula bar: = rand () > enter (between 0 and 1)

Go to formula bar : = randbetween (1,50)

- Random Sampling

Select input > Go to Data > Data Analysis > Sampling > select input range > no of sample > output range > enter

Functions

- Average= Sum of observation/count
- Mode= Frequently observe value
- Median= Middle value in a data set when the numbers are listed in either ascending/descending order
- IF: =if(logical test,[true value],[false value])
E.g. IF(J29<40,"FAIL","PASS") (if column J29 is less than 40, give FAIL)
- IF_AND : =IF(AND(A29>=30,B29>=20),"PASS","FAIL")
- IF_IF: =IF(M29="fail","FAIL",IF(J29>=75,"A+",IF(J29>=60,"B+")))
- SUMIF: =sumif(range,criteria,[sum_range])
- COUNTIF: =countif(range,criteria)
- AVERAGEIF: =range,criteria,[average_range])
- SUMIFS: =sumifs(sum_range, criteria range 1, [criteria1], criteria range 2, [criteria2])

THANK YOU!!

