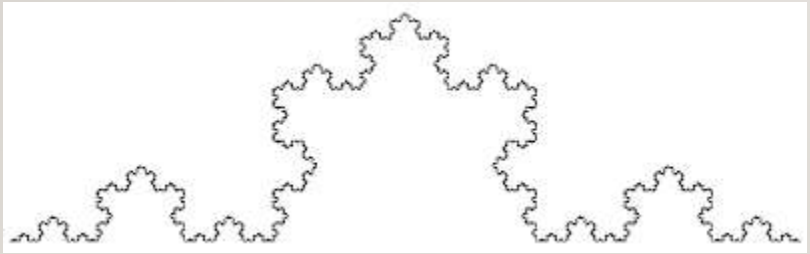
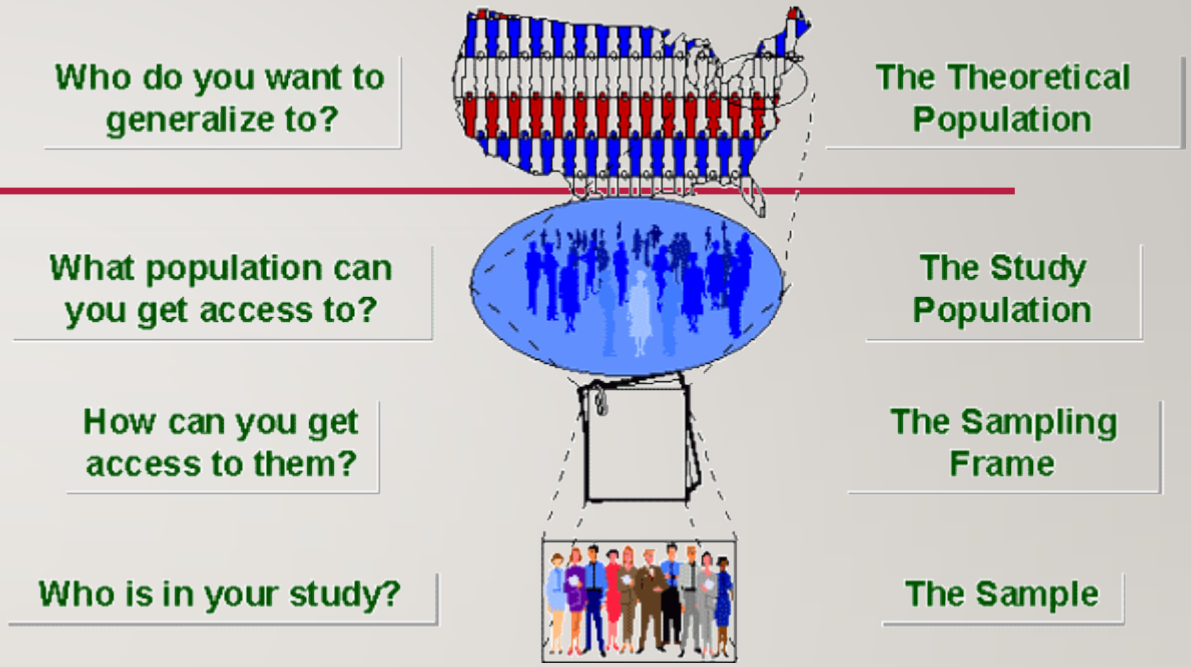
BIOSTATISTICS – FINAL CONSIDERATIONS

PLEASE EVALUATE THE ACTIVITY OF THIS COURSE!

STATISTICAL SETS

- Population
- Sample
  - Representative sample





VARIABLES

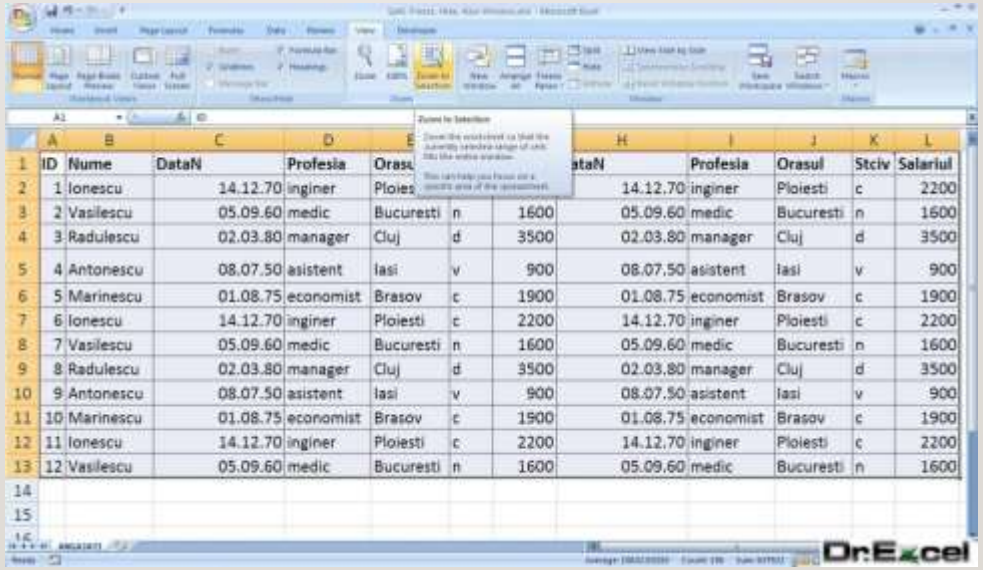
- The common element of the statistical analysis

- Categories:
  - Quantitative
    - Continuous
    - Discrete
  - Qualitative
    - Nominal
    - Ordinal

- Scales
    - Nominal
    - Ordinal
    - Ratio
    - Interval

STATISTICAL SERIES

- The result of the data collection process in a statistical set



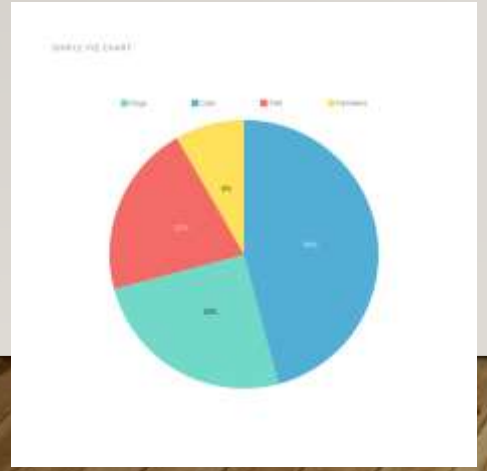
	A	B	C	D	E	F	G	H	I	J	K	L
1	ID	Nume	DataN	Profesia	Orasu			ataN	Profesia	Orasul	Stciv	Salariul
2	1	Ionescu	14.12.70	inginer	Ploies			14.12.70	inginer	Ploiesti	c	2200
3	2	Vasilescu	05.09.60	medic	Bucuresti	n	1600	05.09.60	medic	Bucuresti	n	1600
4	3	Radulescu	02.03.80	manager	Cluj	d	3500	02.03.80	manager	Cluj	d	3500
5	4	Antonescu	08.07.50	asistent	Iasi	v	900	08.07.50	asistent	Iasi	v	900
6	5	Marinescu	01.08.75	economist	Brasov	c	1900	01.08.75	economist	Brasov	c	1900
7	6	Ionescu	14.12.70	inginer	Ploiesti	c	2200	14.12.70	inginer	Ploiesti	c	2200
8	7	Vasilescu	05.09.60	medic	Bucuresti	n	1600	05.09.60	medic	Bucuresti	n	1600
9	8	Radulescu	02.03.80	manager	Cluj	d	3500	02.03.80	manager	Cluj	d	3500
10	9	Antonescu	08.07.50	asistent	Iasi	v	900	08.07.50	asistent	Iasi	v	900
11	10	Marinescu	01.08.75	economist	Brasov	c	1900	01.08.75	economist	Brasov	c	1900
12	11	Ionescu	14.12.70	inginer	Ploiesti	c	2200	14.12.70	inginer	Ploiesti	c	2200
13	12	Vasilescu	05.09.60	medic	Bucuresti	n	1600	05.09.60	medic	Bucuresti	n	1600
14												
15												

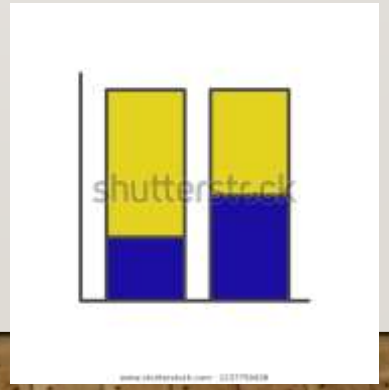
STATISTICS

- Descriptive statistics
  - Collecting data
  - Presenting data
    - Using tables
    - Using charts
    - Using synthetic parameters
- Inferential statistics
  - Estimating parameters in population
    - Confidence intervals
  - Testing statistical hypotheses
    - Statistical tests

QUALITATIVE DATA – DESCRIPTIVE STATISTICS

- Frequency
- Tables of frequencies / table of contingency
- Different charts

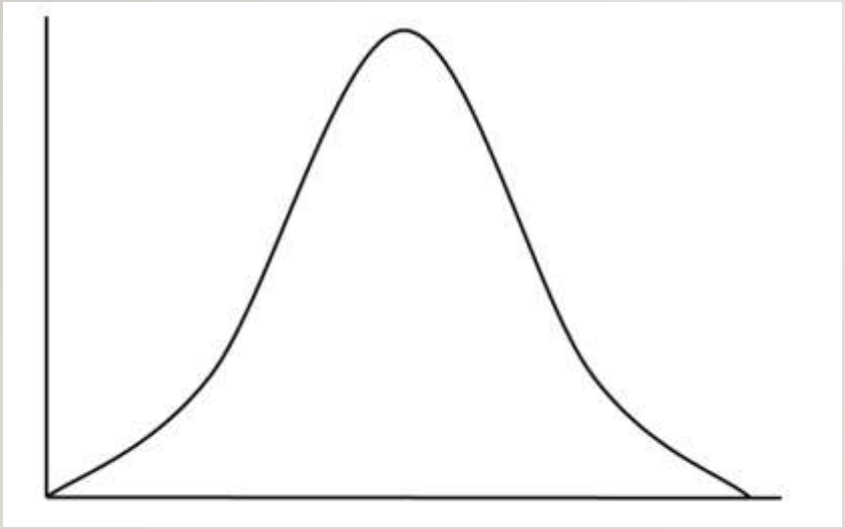




QUANTITATIVE DATA

- Histograms, Box&Wiskers

- Descriptive parameters
  - Centrality measures
  - Spreading measures
  - Localization measures
  - Normality measures



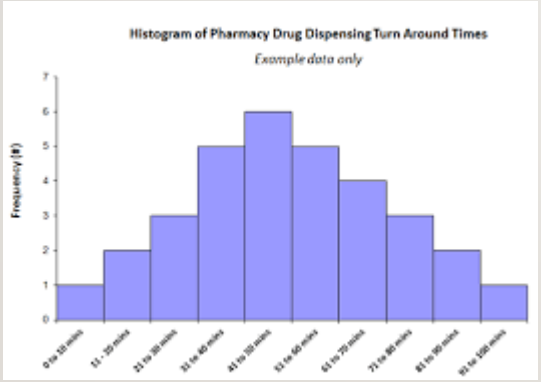
NORMAL DISTRIBUTION

- Theoretical distribution
- Are we close to normal distribution?

SUMMARY OF THE NUMERICAL SERIES

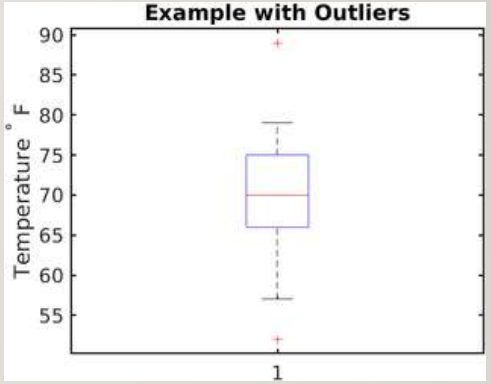
NORMAL DISTRIBUTED DATA

- Mean and standard deviation
  - 50 (10)
- Histogram



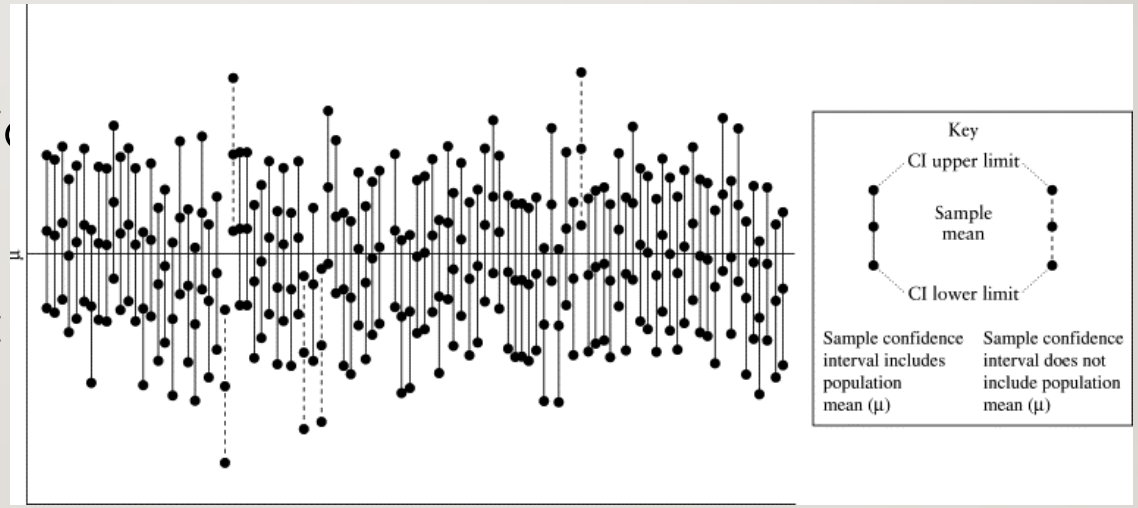
NOT NORMAL DISTRIBUTED DATA

- Median and Interquartile Range
  - 70 (65, 75)
- Boxplot



CONFIDENCE INTERVALS

- Are used to estimate the value of a statistical parameter in population
- Confidence and precisi
- Confidence interval for:
  - Means
  - Frequencies
  - etc.



STATISTICAL TESTS

- Use statistical tests whenever you need to make statements based on samples
- Level of signification -> Level of confidence
- Null hypothesis, Alternative hypothesis

Decision:

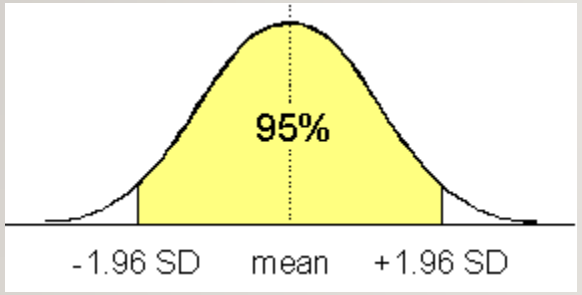
- can reject the null hypothesis and can accept the alternative hypothesis with high confidence

Or

- Cannot reject the null hypothesis and cannot demonstrate the alternative hypothesis

DECISION

- Using critical region



- Using p

ERRORS IN STATISTICAL TESTS

		Conclusion about null hypothesis from statistical test	
		Accept Null	Reject Null
Truth about null hypothesis in population	True	Correct	Type I error Observe difference when none exists
	False	Type II error Fail to observe difference when one exists	Correct

TEST THE DIFFERENCE – QUANTITATIVE VARIABLES

NORMAL DISTRIBUTED DATA

- Parametrical tests
  - Anova
  - Z
  - T

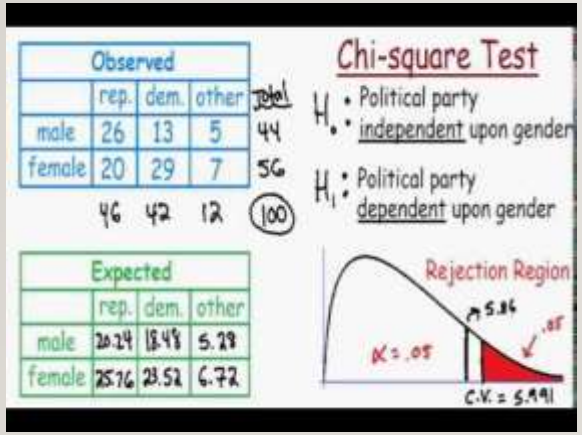
NOT NORMAL DISTRIBUTED DATA

- Non parametrical tests
  - Kruskal-Wallis
  - Man-Whitney-U
  - Wilcoxon sign rank,
  - etc.

TEST THE ASSOCIATION (RELATIONSHIP)

QUALITATIVE DATA

- Chi Square test (with corrections)



QUANTITATIVE DATA

- Correlation
  - Pearson's coefficient for normal distributed data
  - Spearman's rank coefficient for not normal distributed data
  - Direction, power, significance