Analysing and presenting data: **practical hints**

Daniele CEI, Giorgio MATTEI

Contacts:
daniele.cei@centropiaggio.unipi.it
giorgio.mattei@centropiaggio.unipi.it

Course: Ingegneria dei Tessuti Biologici
Date: 27 May 2014

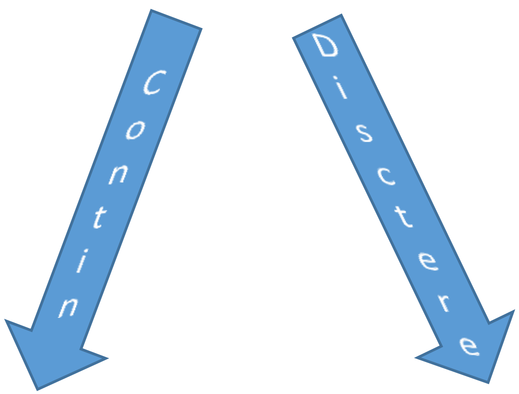
Outline

- ***Basic knowledge*** of statistic and statistical inference
- «***How to use***» common software for data analysis
- ***Practical*** examples
- ***Writing a paper***: literature research and reference management

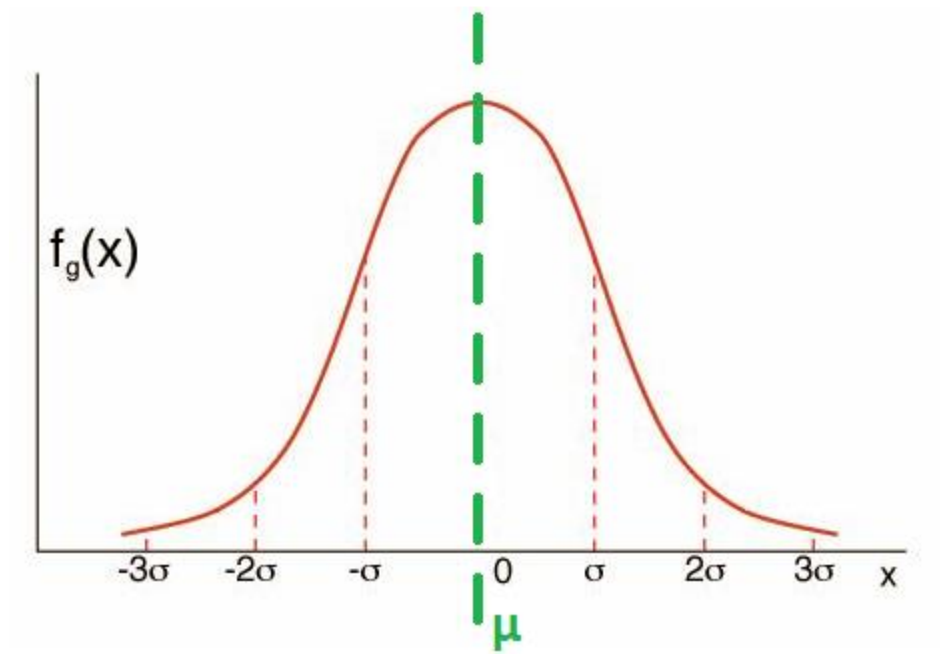
Statistical analysis of data
basic knowledge

- **Mean**: *expected value of a random variable*

$$\mathbb{E}(X) := \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$$



$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f(x)\, dx \qquad \mathbb{E}[X] = \sum_{i=1}^{\infty} x_i\, p_i$$



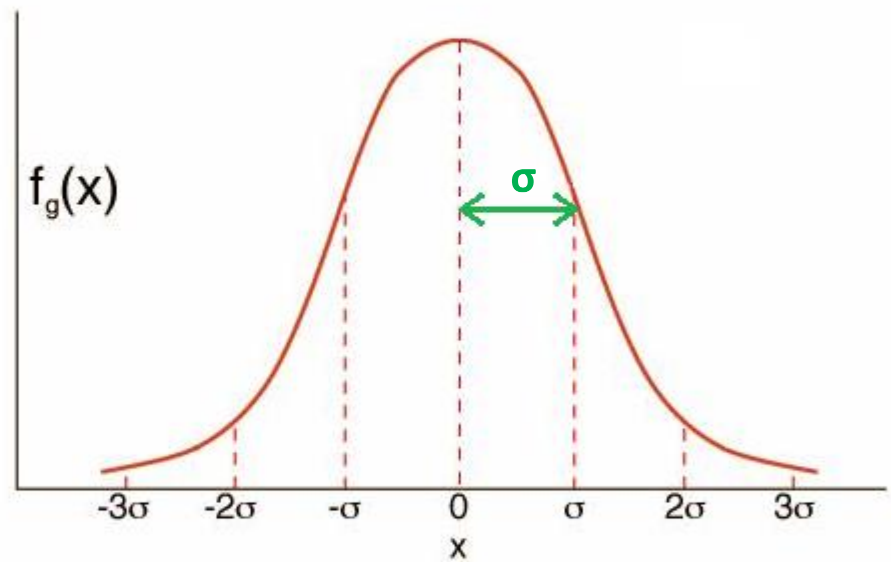
Statistical analysis of data
basic knowledge

- **Variance** **(σ^2)** and **standard deviation** **(σ)**: *measure of how far a set of numbers is spread out*

The variance of a random variable *X* is its ***second central moment***

$$\operatorname{Var}(X) = \operatorname{E}\left[(X-\mu)^2\right] = \operatorname{E}\left[X^2\right] - (\operatorname{E}[X])^2$$

$$\begin{aligned}\operatorname{Var}(X) = \sigma^2 &= \int (x-\mu)^2 f(x)\,dx \\ &= \int x^2 f(x)\,dx - \mu^2 \\ &= \sum_{i=1}^{n} p_i \cdot (x_i-\mu)^2 = \sum_{i=1}^{n} (p_i \cdot x_i^2) - \mu^2\end{aligned}$$



The *standard deviation* is the ***square root*** of the variance.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2} \qquad \sigma = \sqrt{\int_{\mathbf{X}} (x-\mu)^2\, p(x)\, dx},$$

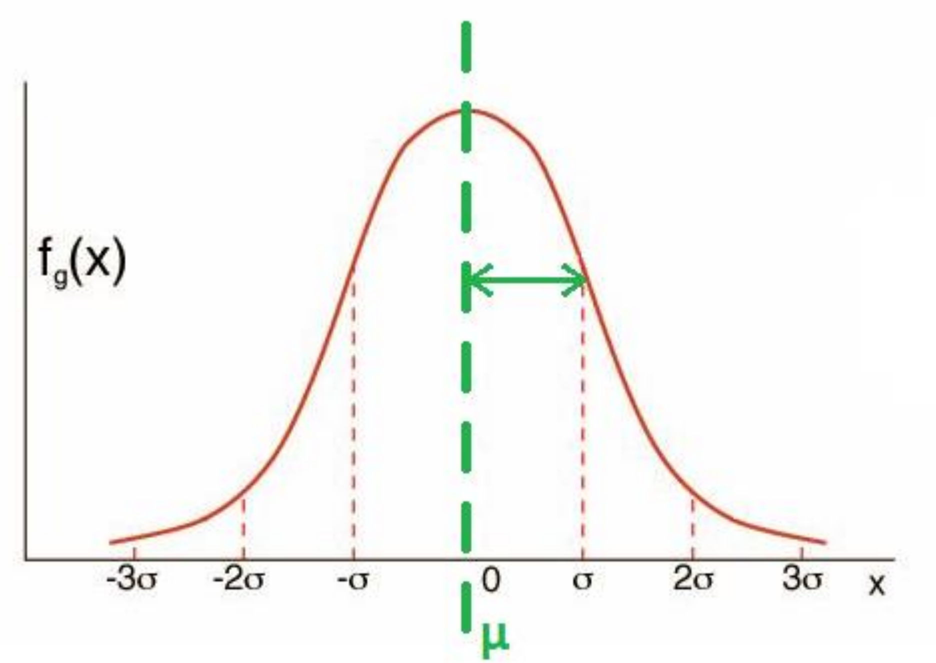
Statistical analysis of data
basic knowledge

- **Correction for sample**: we do not use population in ''real'' case

$$s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2.$$

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \overline{x})^2}.$$

$$\mu = \frac{1}{N}\sum_{i=1}^{N}x_i.$$



Data Analysis
how people really analyse data

Microsoft Excel
Useful shortcuts

- **Sum**: *sum of the selected cells*
- **Mean**: *mean of the selected cells*
- **Standard deviation (sample)**:

  *=ST.DEV.C(**number1**,number2,...)*
- **Use of $**: *freeze coordinates (useful when links cells and sheets)*
- **How to link cells and sheets**

MathWorks MatLab
Useful Shortcuts

- **Sum:**

```
S = sum(A,dim)
```

- **Mean:**

```
M = mean(A,dim)
```

- **Standard deviation (sample):**

```
s = std(X,flag,dim)
```

dim = 1 evaluation by **columns** (default)

dim = 2 evaluation by **rows**

Matlab evaluate corrected standard deviation by default

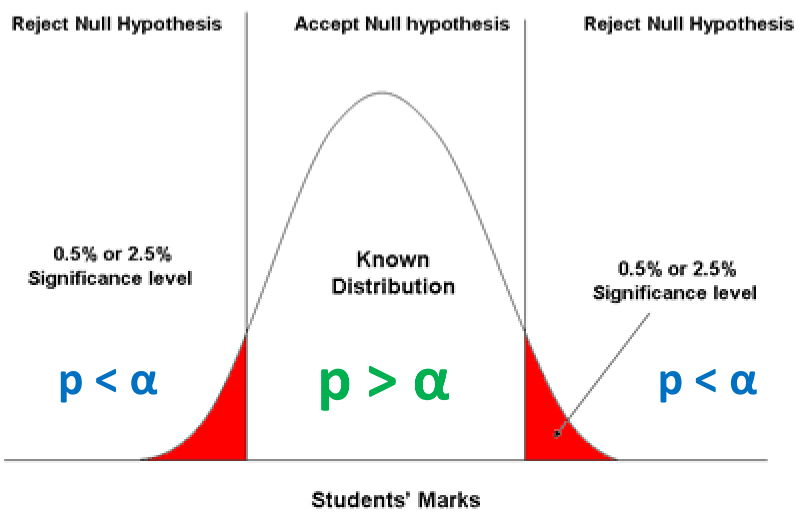
Statistical analysis of data
test of Hypothesis

A ***statistical hypothesis test*** is a method of making decisions using data from a scientific study. In statistics, a result is called statistically significant if it has been predicted as unlikely to have occurred by chance alone, according to a pre-determined threshold probability (called *significance level*, α)

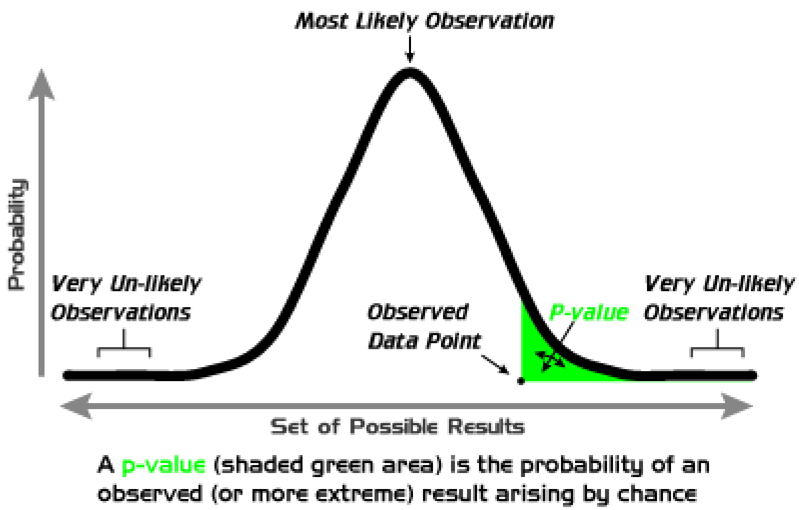
Null *and* ***alternative hypothesis*** (H_0 and H_1) formulation → ***Statistic Test*** selection → ***Significance level*** (α) selection → ***Compute*** from observation value t_{obs} and ***probability*** under the null hypothesis

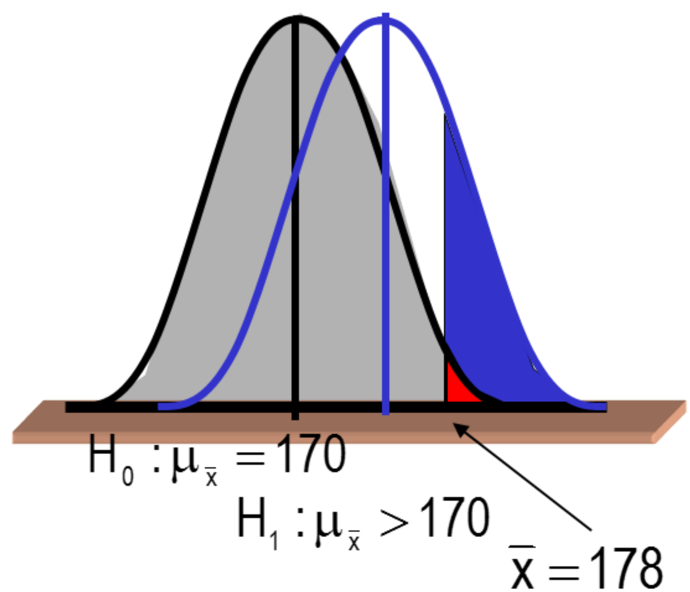
Reject the null hypothesis in favor of the alternative **p < α**

Accept the null hypothesis **p > α**



Interpreting the *p*-value





In conclusion, the **smaller** the ***p*-value** the **more statistical evidence** exists to **support** the **alternative hypothesis** (H_1)

Common Test *mean comparison*

- ***Z test***: one-sample location test comparing the mean of a set of measurements to a given constant
  - ***Independence*** (independent and identically distributed samples, i.i.d.)
  - ***Parameters variance*** should be known, or estimated with high accuracy
  - The test statistic should follow a ***normal distribution***.

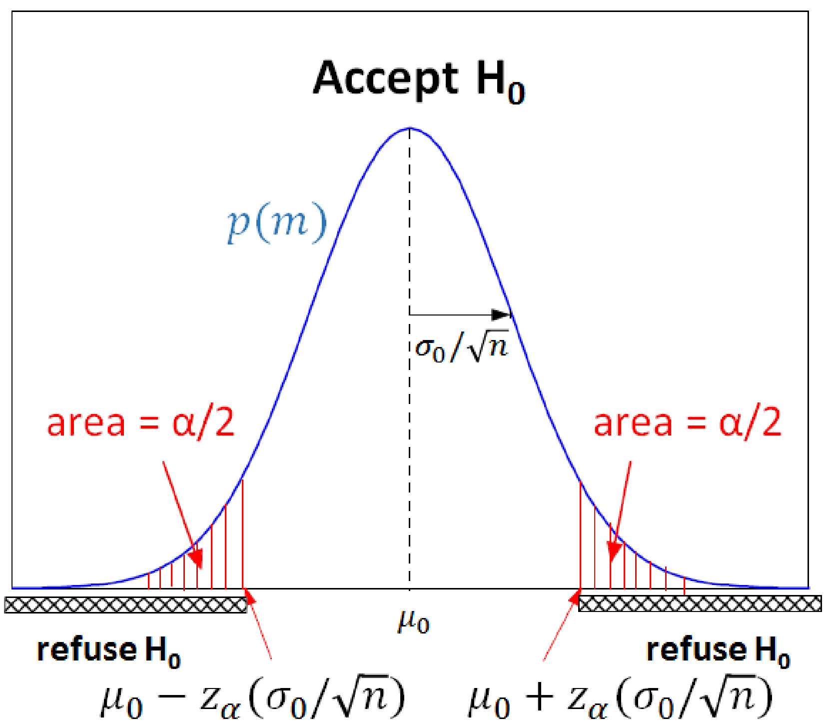
$$z = \frac{M - \mu}{\mathrm{SE}} \qquad \mathrm{SE} = \frac{\sigma}{\sqrt{n}}$$

- ***T test***: A ***one-sample location test*** of whether the mean of a population has a value specified in a null hypothesis; it is used instead of a z-test when the **sample variance is unknown** and is replaced by an estimate based on the data.
  - ***Normal Population*** Assumption
  - ***Parameters variance*** should be known, or estimated with high accuracy

$$t = \frac{\overline{x} - \mu_0}{s / \sqrt{n}}$$

Case A: unknown μ_0, known σ_0
$\bar{z}$ *statistic* (***z-test***)

- Mean survival time from the diagnosis of a given disease
  - **Population** = **38.3** ± **43.3** months ($\mu_0 \pm \sigma_0$)
  - **100** patients treated with a **new technique** = **46.9** months ($\bar{m}$)
- **H**$_0$ □ $\bar{m} = \mu_0$ and $\bar{s} = \sigma_0$ and **H**$_1$ □ $\bar{m} \neq \mu_0$



$$\bar{z} = \frac{\bar{m} - \mu_0}{\sigma_0/\sqrt{n}} = \frac{46.9 - 38.3}{43.3/\sqrt{100}} = \frac{8.6}{4.33} = 1.99$$

H$_0$ is refused with a **significance level α if $\bar{z}$ < - $z_{0.05}$ or $\bar{z}$ > $z_{0.05}$**

Since $z_{0.05}$ ***= 1.96*** and $z_{0.01}$ ***= 2.58*** **what can we say?**

Case b: unknown μ_0, unknown σ_0
$\bar{t}$ *statistic* (***t-test***)

- Rat uterine weight
  - **Population** = **24** mg (μ_0)
  - $\boldsymbol{n}$**=20** rats: [9, 14, 15, 15, 16, 18, 18, 19, 19, 20, 21, 22, 22, 24, 24, 26, 27, 29, 30, 32]
  - $\boldsymbol{\nu} = n - 1 = \mathbf{19}$

- **H$_0$** □ $\boldsymbol{\overline{m} = \mu_0}$ and $\cancel{\bar{s} = \sigma_0}$

$$\bar{t} = \frac{\overline{m} - \mu_0}{\overline{s}/\sqrt{n}} = \frac{21 - 24}{1.3219} = -2.27$$

Since $\boldsymbol{t_{19,\,0.05}}$ **= 2.093** and $\boldsymbol{t_{19,\,0.02}}$ **= 2.539** **what can we say?**

Sample and **population** are **significantly different** with a **significance level** comprised between **2 %** and **5 %** (**0.02 <** $\boldsymbol{p}$ **< 0.05**; calculated *p-value* for $\boldsymbol{t_{19,\,p}}$ **= 2.27** is $\boldsymbol{p}$ **= 0.035**)

MATLAB

z-test

H = 0, H_0 cannot be refused at **α**

H = 1, refuse H_0 at α

Confidence interval for the «true» value ***μ*** at a level **1 - α**

z-statistic value

significance level

[H,P,CI,ZVAL] = ZTEST(X,mean,sigma,alpha,tail)

p-value *(i.e. the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true)*

sample

population parameters

'both' □“$\overline{X}$ is not mean" (two-tailed test)

'right' □"$\overline{X}$ is greater than mean" (right-tailed test)

'left' □"$\overline{X}$ is less than mean" (left-tailed test)

MATLAB
z-test: example

```
>> X=[8.3 9.2 12.5 7.6 10.2 12.9 11.7 10.8 11.7 9.6];
>> sigma=2.1;
>> mean=12;
>> alpha=0.05;
>> [H,P,CI,ZVAL]=ztest(X,mean,sigma,alpha)
```

MATLAB
t-test

H = 0, H_0 cannot be refused at **α**
H = 1, refuse H_0 at α

Confidence interval for the «true» value ***μ*** at a level **1 - α**

Data **structure** containing **t-statistics value** and **number of DoF**

significance level

[H,P,CI,STATS] = TTEST(X,mean,alpha,tail)

***p-value** (i.e. the probability of obtaining a test statistic at least as extreme as the one that was actually observed, assuming that the null hypothesis is true)*

sample

population mean

'both' □“$\overline{X}$ is not mean" (two-tailed test)
'right' □"$\overline{X}$ is greater than mean" (right-tailed test)
'left' □"$\overline{X}$ is less than mean" (left-tailed test)

MATLAB
t-test: example

```
>> X=[22.3 25.1 27 23.4 24.7 26.5 25.7 24.1 23.9 22.8];
>> mean=23;
>> alpha=0.05;
>> [H,P,CI,STAT]=ttest(X,mean,alpha)
```

Microsoft Excel
Useful function

- **z.test**: *Returns the one-tailed probability-value of a z-test. For a given hypothesized population mean, x, **Z.TEST** returns the probability that the sample mean would be greater than the average of observations in the data set (**array**) — that is, the observed sample mean.*

=ZTEST (array, x, sigma)

array	The array or range of data against which to test *x*.
x	The value to test.
sigma	The population (known) standard deviation.

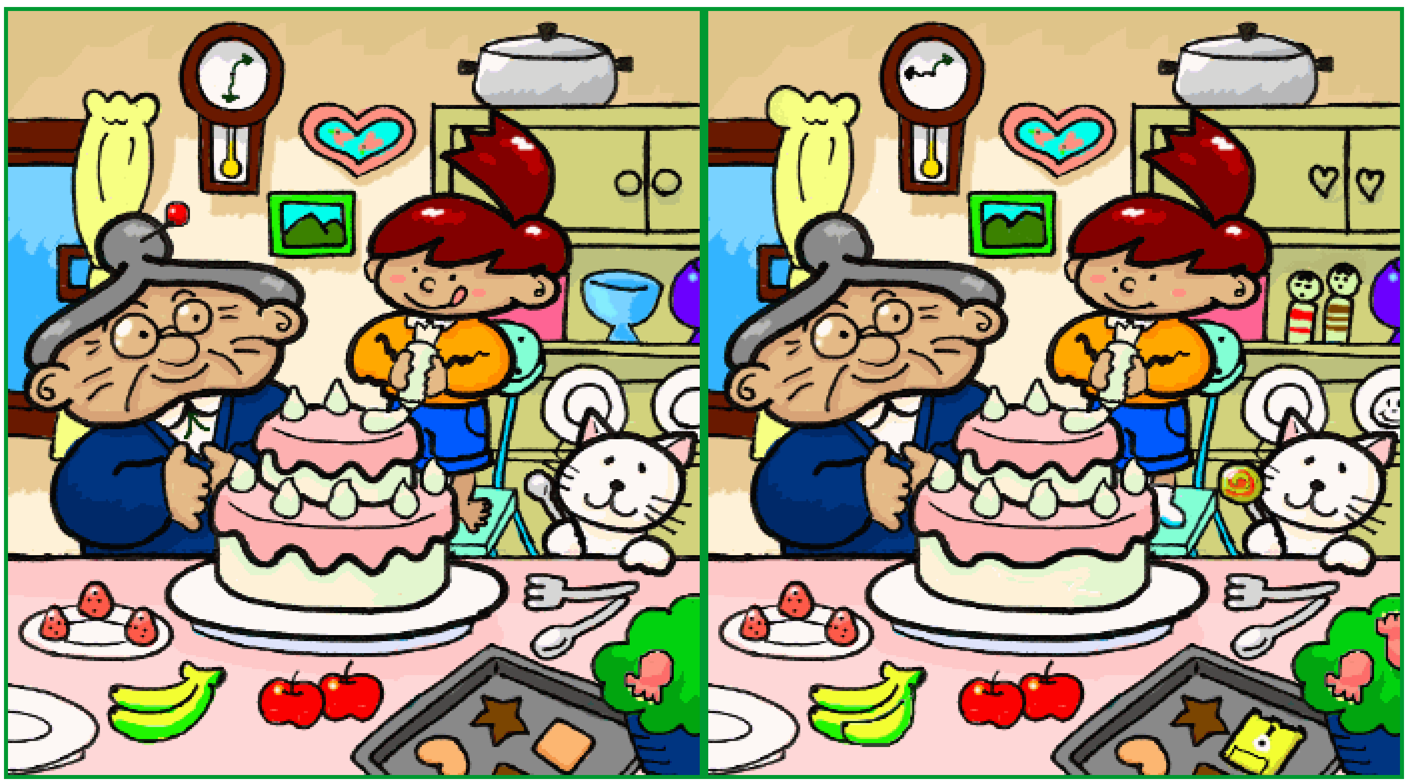
- **t.test**: *Returns the probability that is associated with a Student's t-Test. Use **T.TEST** to determine whether two samples are likely to have come from the same two underlying populations that have the same mean*

=TTEST (array1, array2, tails, type)

Array1, Array2	The array or range of data to test.
Tails	specifies the number of distribution tails: tails = 1 -> uses the one-tailed distribution tails = 2 -> uses the two-tailed distribution.
Type	is the kind of t-Test to perform.

Equal or different?

*The case of **two samples***



Independent two-sample *t-test*

***Equal** sample sizes (**n**), **equal** variances ($S_{X_1X_2}$)*

The ***t* statistic** to test whether the **means** of **group 1** ($\overline{X_1}$) and **group 2** ($\overline{X_2}$) **are different** can be calculated as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \cdot \sqrt{\frac{2}{n}}}$$

$$S_{X_1X_2} = \sqrt{\frac{1}{2}(S_{X_1}^2 + S_{X_2}^2)}$$

«pooled» standard deviation

t-test DoFs = 2n - 2

H_0 is refused with a **significance level α if**

$t < -t_{DoF,\alpha}$ **or** $t > t_{DoF,\alpha}$

Independent two-sample *t-test*

Unequal *sample sizes (**n_1** and **n_2**),* ***equal*** *variances (**$S_{X_1X_2}$**)*

The ***t* statistic** to test whether the **means** of **group 1** ($\overline{X_1}$) and **group 2** ($\overline{X_2}$) **are different** can be calculated as follows:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{X_1X_2} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$S_{X_1X_2} = \sqrt{\frac{(n_1 - 1)S_{X_1}^2 + (n_2 - 1)S_{X_2}^2}{n_1 + n_2 - 2}}$$

«pooled» standard deviation

t-test DoFs = $n_1 + n_2 - 2$

H_0 is refused with a **significance level α if**
$t < -t_{DoF,\alpha}$ or $t > t_{DoF,\alpha}$

Independent two-sample *t-test*

Unequal *sample sizes (**n_1** and **n_2**),* ***unequal*** *variances ($S_{X_1X_2}$)*

The ***t* statistic** to test whether the **means** of **group 1** ($\overline{X_1}$) and **group 2** ($\overline{X_2}$) **are different** can be calculated as follows:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_{\overline{X}_1 - \overline{X}_2}}$$

$$s_{\overline{X}_1 - \overline{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

«unpooled» standard deviation

$$t\text{-}test \text{ DoFs} = \frac{(s_1^2/n_1 + s_2^2/n_2)^2}{(s_1^2/n_1)^2/(n_1-1) + (s_2^2/n_2)^2/(n_2-1)}$$

Welch–Satterthwaite equation

H_0 is refused with a **significance level α if**

$t < -t_{DoF,\alpha}$ or $t > t_{DoF,\alpha}$

Independent two-sample *t-test* (*unequal sample sizes* and *equal variances*): an *example*

- Two groups of 10 *Dapnia magna* eggs, randomly extracted from the same clone, were reared in two different concentrations of hexavalent chromium
- After a month survived individuals were measured: 7 in group A and 8 in group B

A	B
2,7	2,2
2,8	2,1
2,9	2,2
2,5	2,3
2,6	2,1
2,7	2,2
2,8	2,3
	2,6
2.714	2.250

Mean 2.714 2.250

$$s_p^2 = \frac{0,10825 + 0,18000}{6 + 7} = 0,022173$$

«pooled» variance

$$t_{13} = \frac{2,714 - 2,250}{\sqrt{0,022173 \cdot \left(\frac{1}{7} + \frac{1}{8}\right)}} = 6,02$$

t with 13 DoF

Since $\boldsymbol{t_{13,\,0.05}}$ **= 2.160 what can we say?**

MATLAB
Independent two-sample t-test

H = 0, H_0 cannot be refused at **α**
H = 1, refuse H_0 at α

Confidence interval for the «true» **difference** of **population means**

Data **structure** containing **t-statistics value** and **number of DoF**

significance level

[H,P,CI,STATS] = TTEST2($\underbrace{\text{X,Y}}_{\textbf{samples}}$,alpha,tail,vartype)

p-value *(i.e. the probability of observing the given result, or one more extreme, by chance if the null hypothesis is true)*

'equal' or **'unequal'**

'both' □“means are not equal" (two-tailed test)
'right' □"$\overline{X}$ is greater than $\overline{Y}$" (right-tailed test)
'left' □"$\overline{X}$ is less than $\overline{Y}$" (left-tailed test)

MATLAB

Ind. 2-sample t-test: an example

```
>> X=[2.7 2.8 2.9 2.5 2.6 2.7 2.8]';
>> Y=[2.2 2.1 2.2 2.3 2.1 2.2 2.3 2.6]';
>> [H,P,CI,STATS] = ttest2(X,Y,0.05,'both','equal')
```

Dependent two-sample *t-test*

one sample tested twice *or* ***two "paired" samples***

$$t = \frac{\overline{X}_D - \mu_0}{s_D/\sqrt{n}}$$

- Calculate the differences between all ***n*** pairs (X_D), then substitute their average ($\overline{X_D}$) and standard deviation (s_D) in the equation above to test if the average of the differences is significantly different from $\boldsymbol{\mu_0}$ ($\boldsymbol{\mu_0}$ **= 0** under $\mathbf{H_0}$ **, DoFs =** ***n*** **- 1**)
- The "pairs" can be either one person's pre-test and post-test scores (repeated measures) or persons matched into meaningful groups (e.g. same age)

Example of repeated measures			
Number	**Name**	**Test 1**	**Test 2**
1	Mike	35%	67%
2	Melanie	50%	46%
3	Melissa	90%	86%
4	Mitchell	78%	91%

Example of matched pairs			
Pair	**Name**	**Age**	**Test**
1	John	35	250
1	Jane	36	340
2	Jimmy	22	460
2	Jessy	21	200

Dependent two-sample *t-test*: an example

Student	Pre-module score	Post-module score	Difference
1	18	22	+4
2	21	25	+4
3	16	17	+1
4	22	24	+2
5	19	16	-3
6	24	29	+5
7	17	20	+3
8	21	23	+2
9	23	19	-4
10	18	20	+2
11	14	15	+1
12	16	15	-1
13	16	18	+2
14	19	26	+7
15	18	18	0
16	20	24	+4
17	12	18	+6
18	22	25	+3
19	15	19	+4
20	17	16	-1

$$t = \frac{2.05}{0.634} = 3.231 \qquad \text{on 19 df}$$

Since $\boldsymbol{t_{19,\, 0.05}}$ **= 2.093 what can we say?**

MATLAB

Dependent two-sample t-test

H = 0, H_0 cannot be refused at **α**
H = 1, refuse H_0 at α

Confidence interval for the «true» **difference** of **population means**

Data **structure** containing **t-statistics value** and **number of DoF**

significance level

[H,P,CI,STATS] = TTEST($\underbrace{X,Y}_{\textbf{samples}}$,alpha,tail)

p-value *(i.e. the probability of observing the given result, or one more extreme, by chance if the null hypothesis is true)*

'both' □“means are not equal" (two-tailed test)
'right' □"$\overline{X}$ is greater than $\overline{Y}$" (right-tailed test)
'left' □"$\overline{X}$ is less than $\overline{Y}$" (left-tailed test)

MATLAB
Dep. 2-sample t-test: an example

```
>> X=[22 25 17 24 16 29 20 23 19 20 15 15 18 26 18 24 18 25 19 16]';
>> Y=[18 21 16 22 19 24 17 21 23 18 14 16 16 19 18 20 12 22 15 17]';
>> [H,P,CI,STATS] = ttest(X,Y,0.05,'both')
```

Equal or different?
more than ***two samples***

ANalysis Of VAriance (ANOVA)

- **More than 2** groups: **NO pairwise comparisons** (*t-test*)

  ↑ groups □↑ overall probability that at least one of them is significant
  (e.g. α=0.05 and n=20 □in average 1 group will be significantly different for the case, even if H_0 is true)

$$H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$$

H_1: not all means are equal

all means **are different**

...

one mean **is different** from the others, which are all equals

- **ANOVA**
  - uses **Fisher's distribution** (**F-**distribution)
  - the **sources of variations** on observed values of **two or more groups** can be **decomposed** and **accurately measured**
  - the **source of variation** is called **EXPERIMENTAL FACTOR** (or **TREATMENT**) and can be multi-levelled
  - each **unit or observation** of the experimental factor is called **REPLICATION**

1- and 2-way ANOVA

- ***ANOVA*** **(1way):** The ANOVA tests the null hypothesis that samples in *two or more groups* are drawn from populations with the *same* mean values. To do this, two assumption are made about the population variance.
  - Response variable are **normally distributed** (or approximately normally distributed), and **independent** and **identically** distributed
  - **Variances** of populations are **equal**.

- ***ANOVA*** **(2way):** is an extension of the one-way ANOVA test that examines the *influence of different categorical* independent variables on one dependent variable. The two-way ANOVA can determine the main effect of contributions of each independent variable and identifies if there is a significant interaction effect between the independent variables.
  - The populations from which the samples are obtained must be **normally distributed**, with a **correct** sampling. Observations for within and between groups must be **independent**.
  - The variances among populations must be equal (**homoscedastic**).

one-way ANOVA: an example

The problem

- Content of iron in air in 3 different zones (A, B, C) of a city (µg/N mc at 0 °C and 1013 mbar)

	EXPERIMENTAL FACTOR				
	A	B	C		
	2,71	1,75	2,22		
	2,06	2,19	2,38		
	2,84	2,09	2,56		
	2,97	2,75	2,60		
	2,55		2,72		
	2,78				
$\sum X_j$	15,91	8,78	12,48	$\sum X$	37,17
n_i	6	4	5	n	15
$\overline{X}_{\cdot j}$	2,652	2,195	2,496	$\overline{X}_{\cdot\cdot}$	2,478

MATLAB

one-way ANOVA

***p*-value for H_0**
(means of the groups are equal)

ANOVA table values

Structure of statistics useful **for performing a multiple comparison of means** with the **MULTCOMPARE** function

[P,ANOVATAB,STATS] = anova1(X,GROUP,DISPLAYOPT)

Matrix with **1 group per column**
(requires equal-sized samples)

Character array: one row per column of X, containing the **group names**

Vector of data

Vector: one group name for **each element of X**

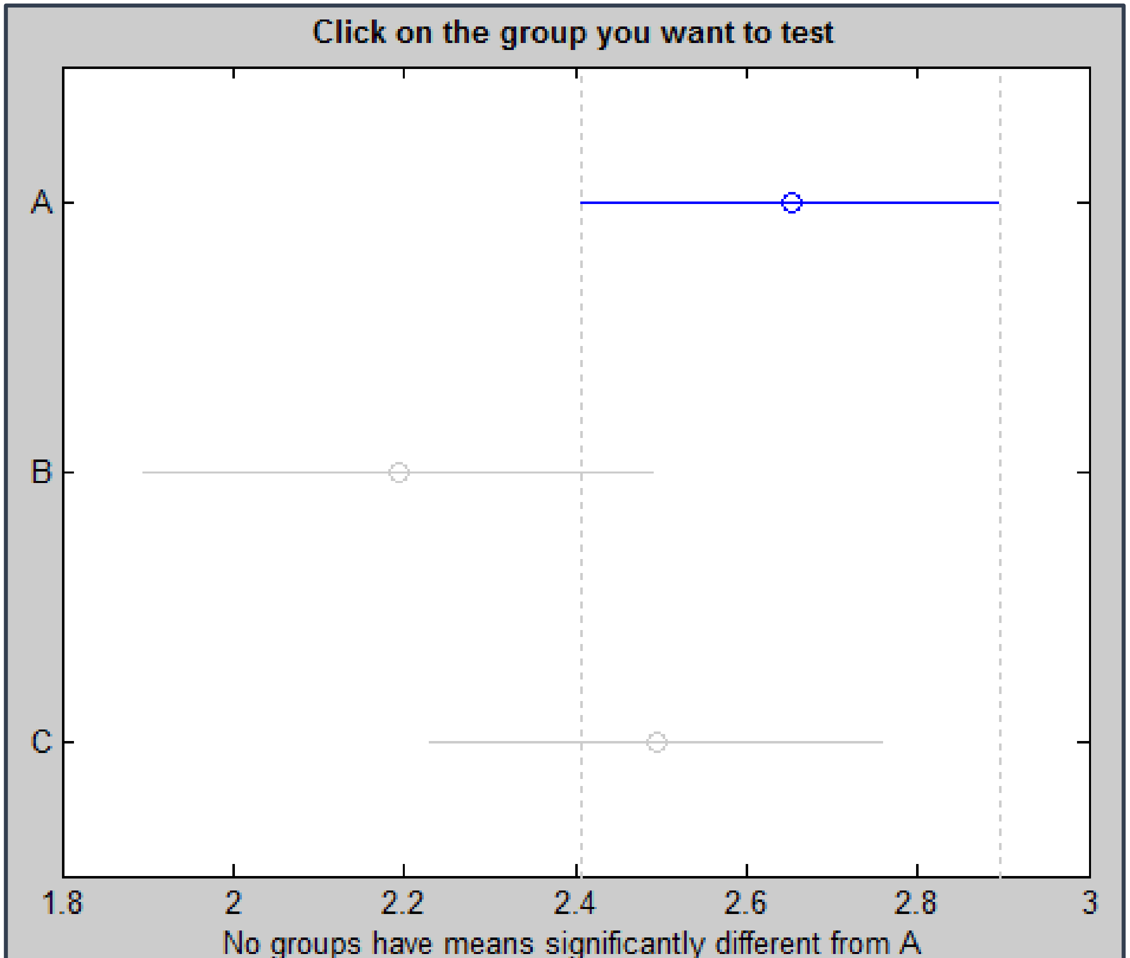
'on' (the default) to **display figures containing a standard one-way anova table and a boxplot**, or 'off' to omit these displays

MATLAB
one-way ANOVA: example

```
>> X=[2.71,2.06,2.84,2.97,2.55,2.78,1.75,2.19,2.09,2.75,2.22,2.38,2.56,2.6,2.72]';
>> GROUP=['A','A','A','A','A','A','B','B','B','B','C','C','C','C','C']';
>> [P,ANOVATAB,STATS] = anova1(X,GROUP)
```

MATLAB
one-way ANOVA: example

COMPARISON = multcompare(STATS)



two-way ANOVA: an example

The problem

- **Content** of **Pb** in air in **5 different urban zones** revealed every **6 hours** during the day

	TREATMENTS (urban zone)					X_{ij}	
BLOCKS (time)	**1**	**2**	**3**	**4**	**5**	**sums**	**means**
6 am	28	25	30	22	26	131	26,2
12 am	34	32	37	31	30	164	32,8
6 pm	22	21	24	20	19	106	21,2
12 pm	36	31	40	33	29	169	33,8
sums	120	109	131	106	104	570	
means	30,00	27,25	32,75	26,50	26,00		28,50

MATLAB
two-way ANOVA

***p*-value for H_0** (means of the groups are equal)

ANOVA table values

Structure of statistics useful **for performing a multiple comparison of means** with the **MULTCOMPARE** function

[P,ANOVATAB,STATS] = anova2(X,REPS,DISPLAYOPT)

Matrix of data (balanced ANOVA □equal number of repetitions)

Columns: 1st factor

Rows: 2nd factor

REPS indicates the **number** of **observations per "cell"**

A **"cell" contains REPS number** of **rows**

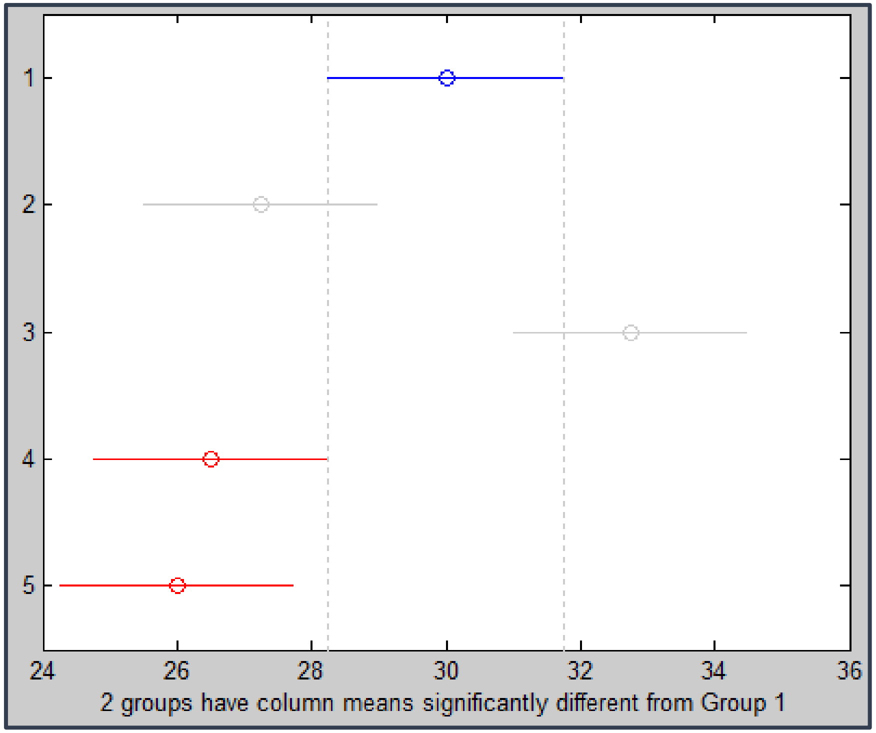
'on' (the default) to **display a standard two-way anova table**, or 'off' to skip the display

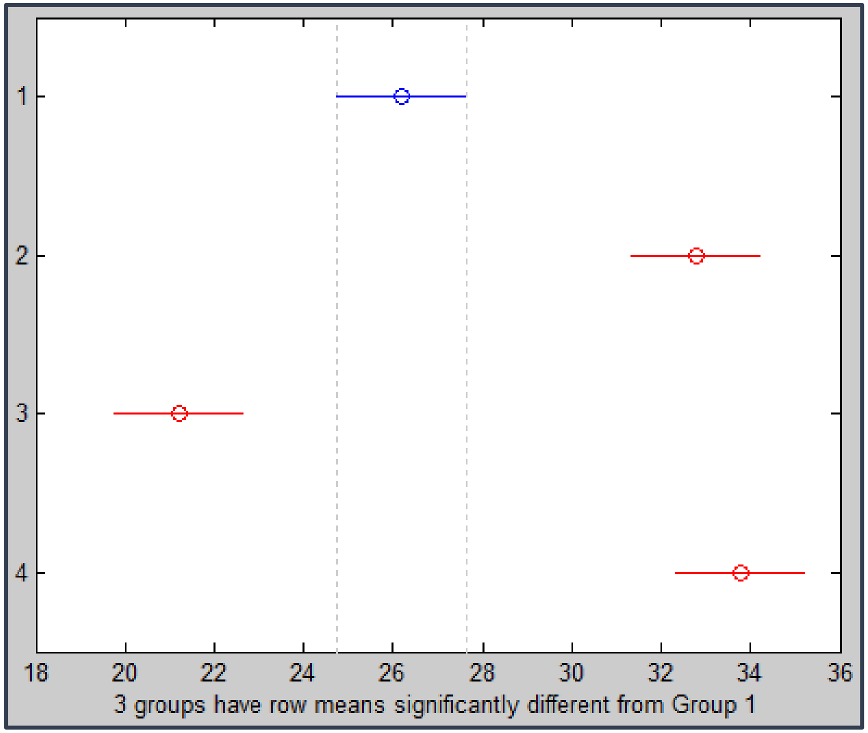
MATLAB
two-way ANOVA: example

```
>> X=[28 25 30 22 26;
34 32 37 31 30;
22 21 24 20 19;
36 31 40 33 29];
>> [P,ANOVATAB,STATS] = anova2(X)
```

MATLAB
two-way ANOVA: example

COMPARISON = multcompare(STATS, 'estimate', 'column' (default) or 'row')

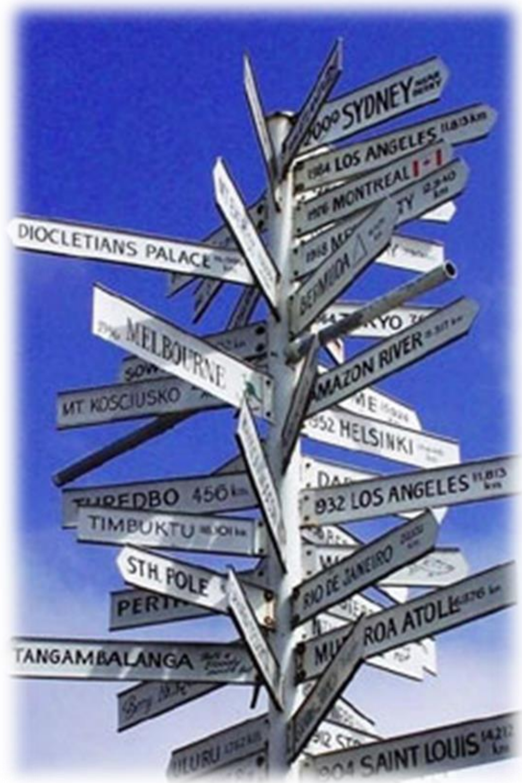




Columns (i.e. **urban zones**)

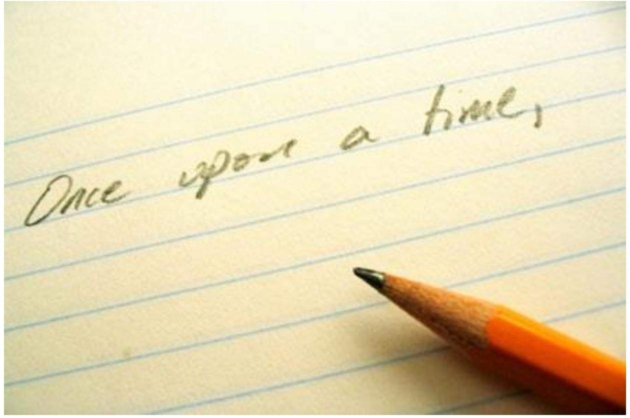
Rows (i.e. **times**)

MATLAB
anovan: N-way *analysis of variance*



How to write a paper

Literature research and reference management



The Sections of a Paper

- Most journal-style scientific papers are subdivided into sections which usually appear in the following prescribed order:

Experimental process	Section of Paper
What did I do in a nutshell?	Abstract
What is the problem?	Introduction
How did I solve the problem?	Materials and Methods
What did I find out?	Results
What does it mean?	Discussion
Who helped me out?	Acknowledgments (optional)
Whose work did I refer to?	Literature Cited
Extra Information	Appendices (optional)

Source: http://abacus.bates.edu/~ganderso/biology/resources/writing/HTWsections.html

Literature research

PubMed, Web of Science, or Google Scholar?

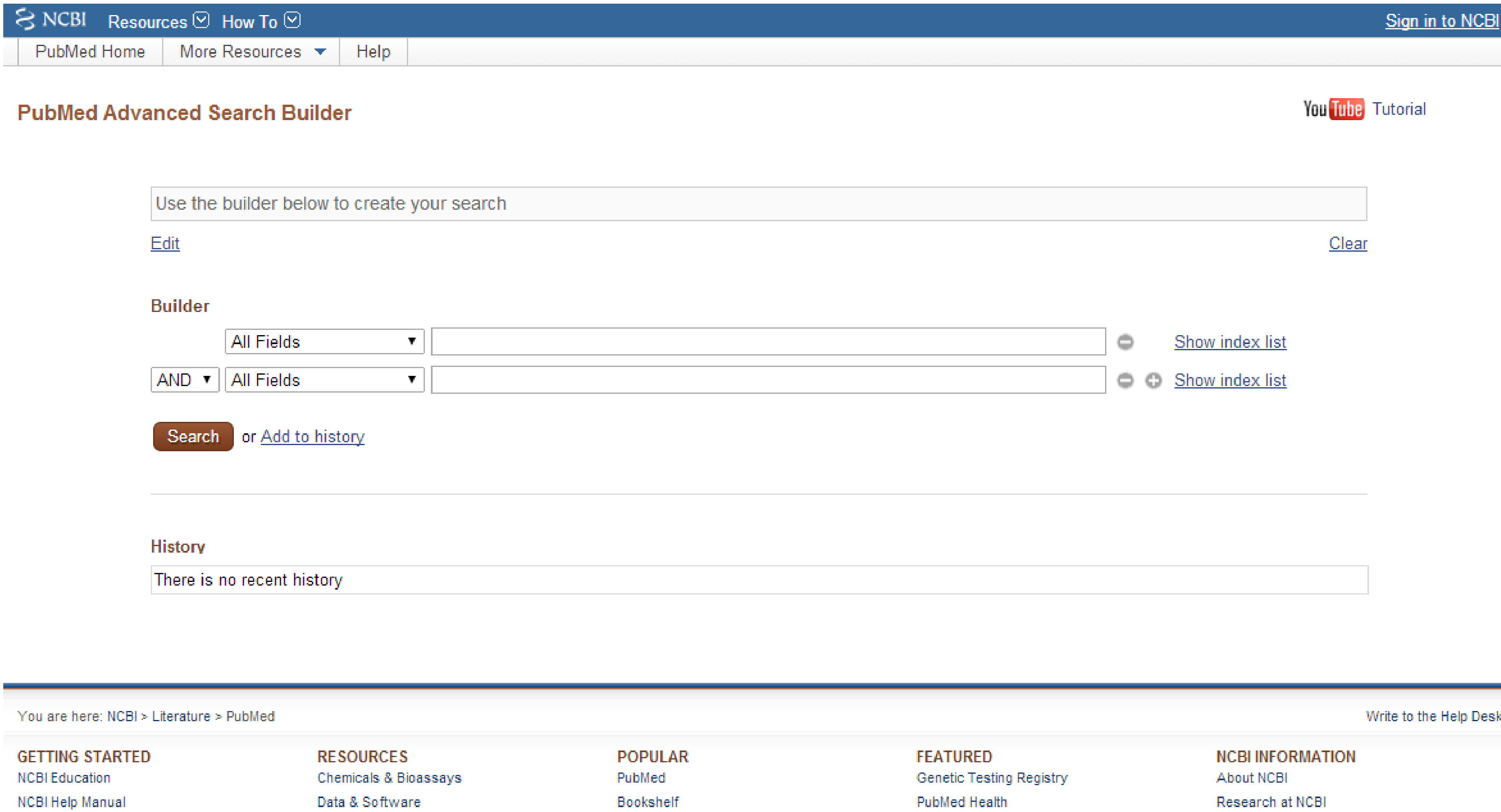
Major differences you must know

1. **Pubmed and Web of Science** are **human-curated** databases. **Google Scholar is not**. This is the key to most of the differences you will find in your search results.
2. **Web of Science** and **Google Scholar track citations,** but **PubMed does not.**
3. **Google Scholar searches full text of articles** but **PubMed and Web of Science search only the citation, abstract, and tagging information.**

Source: http://libguides.lib.msu.edu/pubmedvsgooglescholar

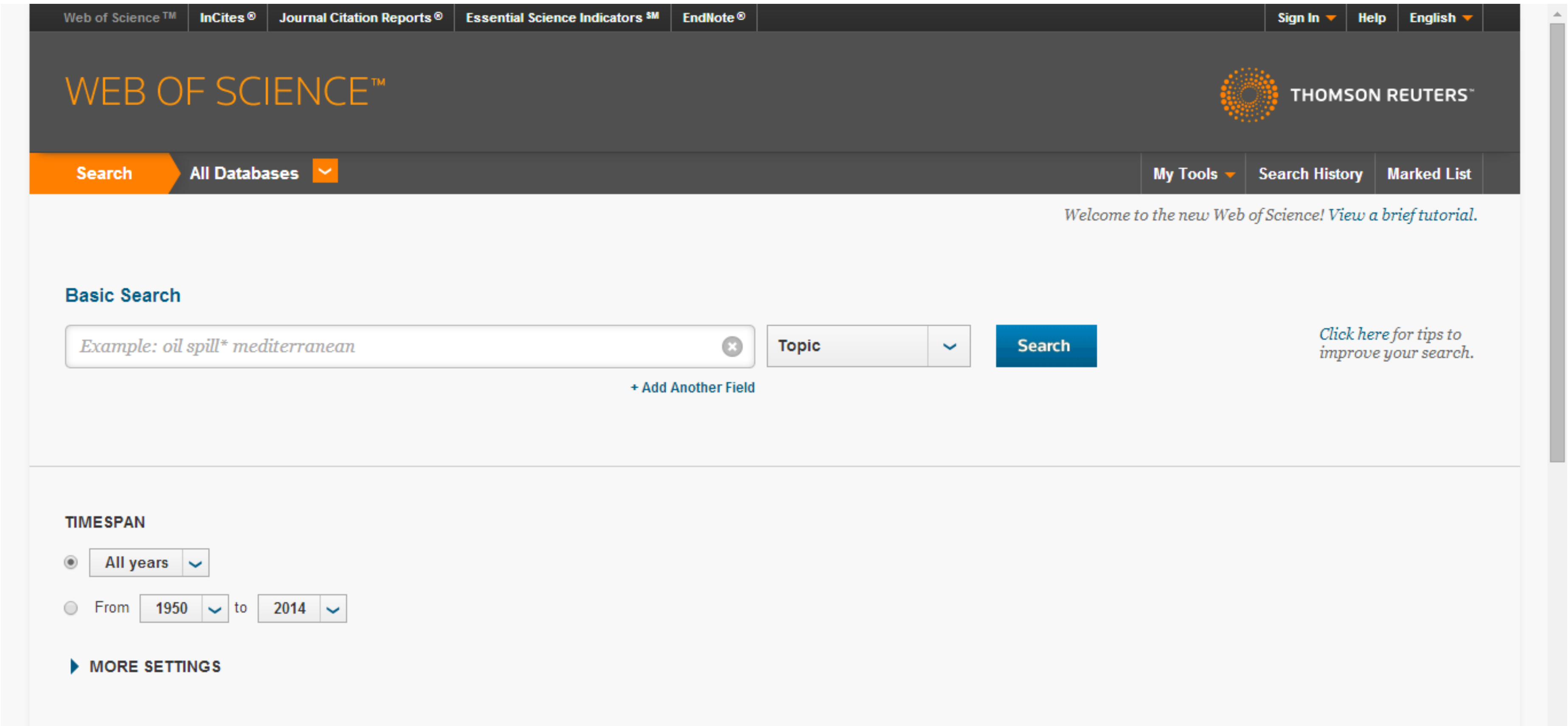
Pubmed

http://www.ncbi.nlm.nih.gov/pubmed



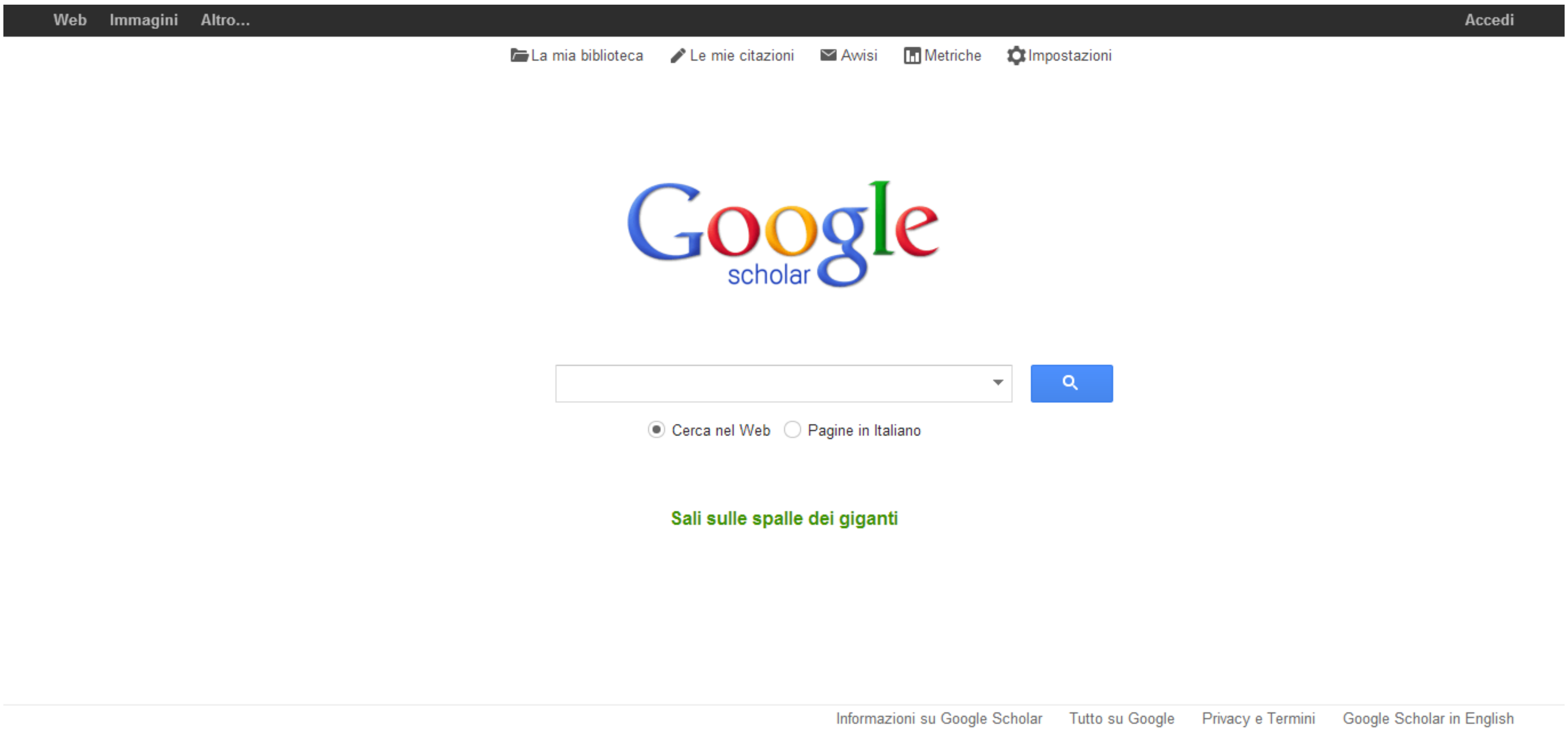
Web of Science

http://apps.webofknowledge.com/



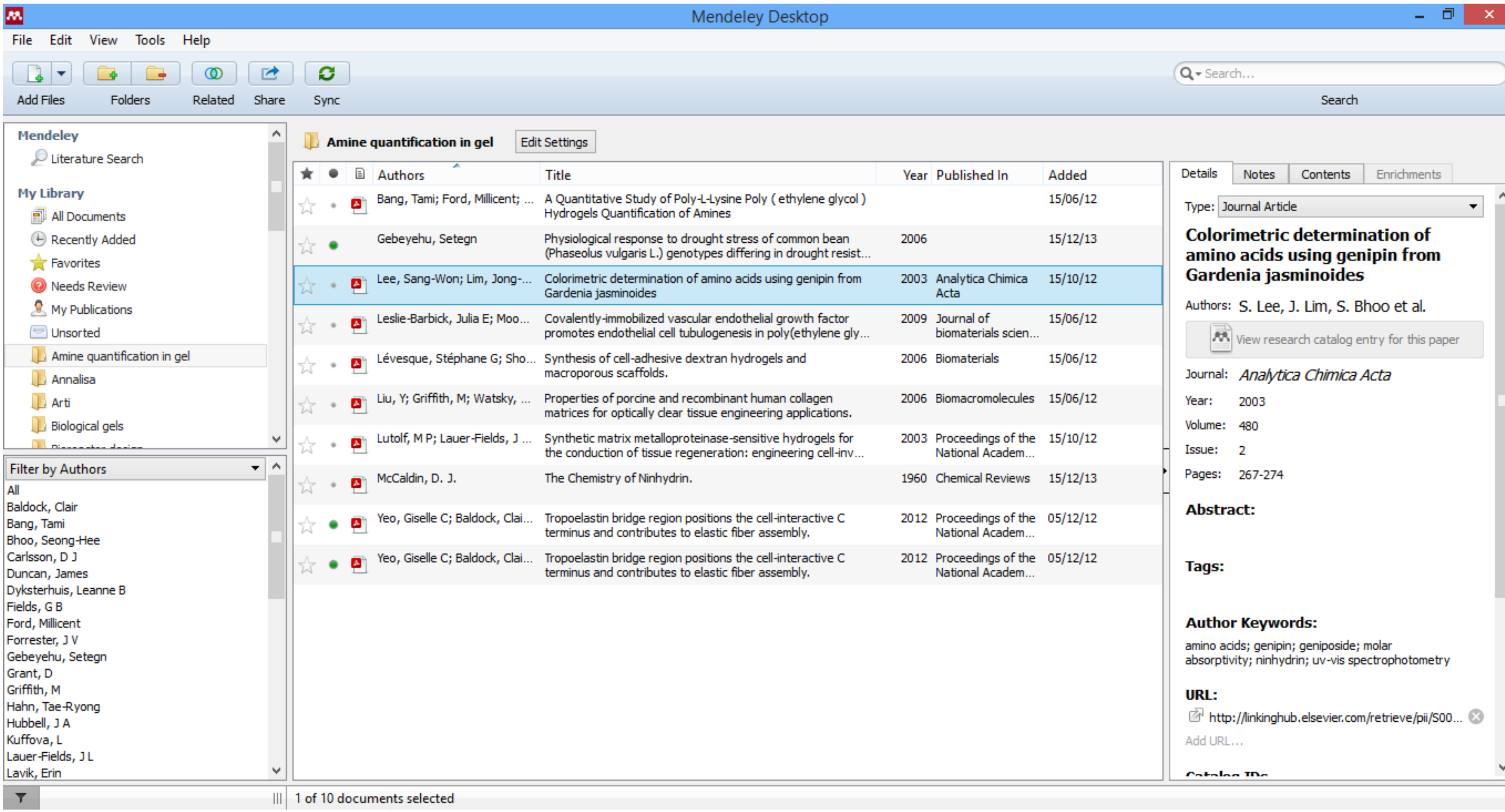
Google Scholar

http://scholar.google.it/



Reference management

Mendeley desktop: a **free** citation manager



Downloadable at http://www.mendeley.com

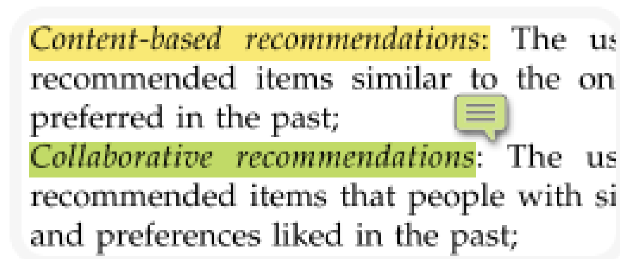
Mendeley’s features

Reference Manager

Generate citations and bibliographies in Microsoft Word, LibreOffice, and LaTeX.

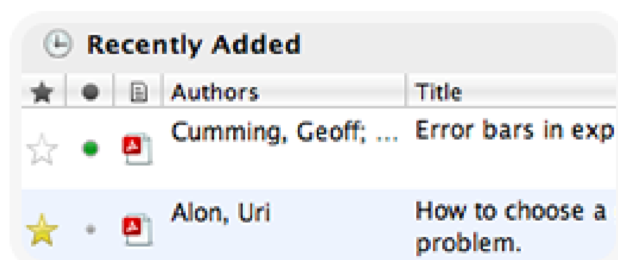
Read and Annotate

Open PDFs and capture your thoughts through sticky notes and highlights.



Add and Organize

Import and organize PDFs from your computer, EndNote™, Papers or Zotero.



Source: http://www.mendeley.com/features/

Mendeley’s features

Collaborate

Connect with colleagues and securely share your papers, notes and annotations.

Backup, Sync and Mobile

Access your papers on the web, iPhone or iPad.

Network and Discover

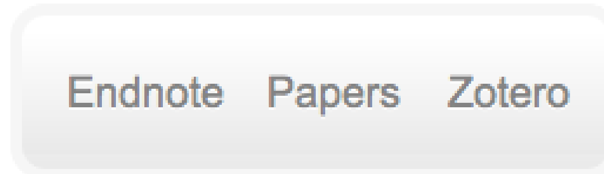
Discover papers, people and public groups.

Source: http://www.mendeley.com/features/

Add papers from anywhere

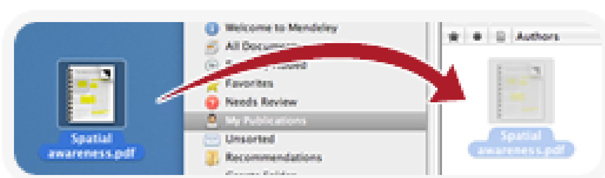
Import from websites

The web importer allows you to import papers, web pages and other documents into your library from search engines and academic databases.



Import from existing software

Mendeley can import BibTeX, RIS and EndNote™ XML files so you can easily transfer your library from EndNote™, Papers and Zotero.



Drag and drop

Drag any PDF directly into Mendeley from your desktop or folders which you can then instantly annotate, cite, and highlight.

Automatically watch folders

Choose a folder on your computer to "watch". When you next add papers to that folder, they are also automatically added to your Mendeley library.

Source: http://www.mendeley.com/features/add-and-organize/

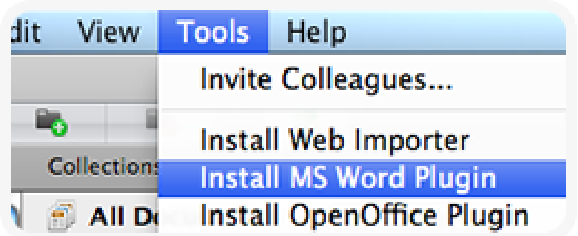
Reference manager

Free and fully compatible with…

- Windows Word 2003, 2007, 2010
- Mac Word 2008, 2011
- LibreOffice
- BibTeX

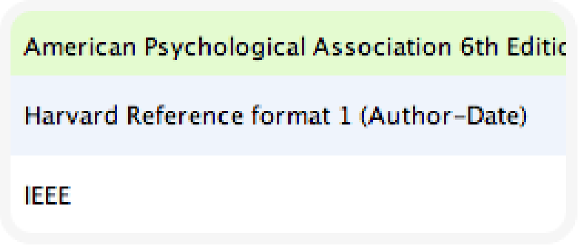
Quick and simple installation

Once you download the Mendeley reference manager, you can install the Word Plugin in 3 clicks. Hey presto – you are ready to create your bibliography.



Citation styles for thousands of journals

Quickly search and select your citation style from a rapidly growing community managed database, or create new styles with the new CSL Editor.



Source: http://www.mendeley.com/features/reference-manager/

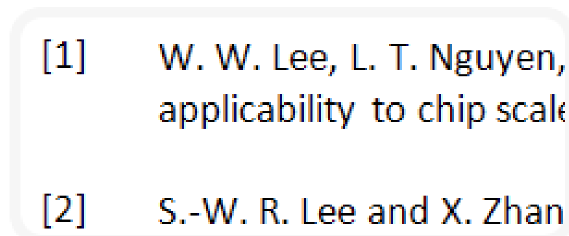
Reference manager

Create bibliographies instantly

Cite seamlessly without leaving Word. Format your citations and bibliography according to your chosen style.

Flexible formatting

Hanging indents, the use of "Ibid.", author disambiguation. Mendeley looks after the details of creating a bibliography so you can focus on writing.



Collaborate on bibliographies

Share bibliographies with your colleagues through a private group. Any member can add or edit any cited references.

Source: http://www.mendeley.com/features/reference-manager/