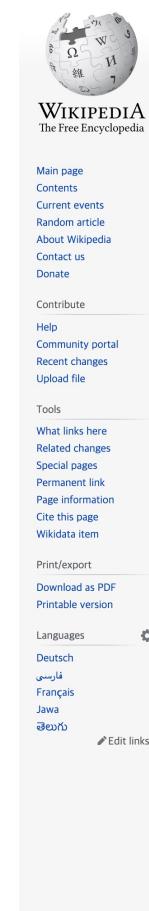


Lecture 13

Softwares and programming languages

Software packages for statistics

- There are so many software packages for statistics



https://en.wikipedia.org/wiki/List_of_statistical_software

The ones that I have used before... and now...

OLDIES (that I am not using any more)

Basic statistics

- SAS (during my undergrad)
- SPSS (during my masters)
- STATA (for my wife)

Multivariate

- LISREL
- Mplus
- AMOS

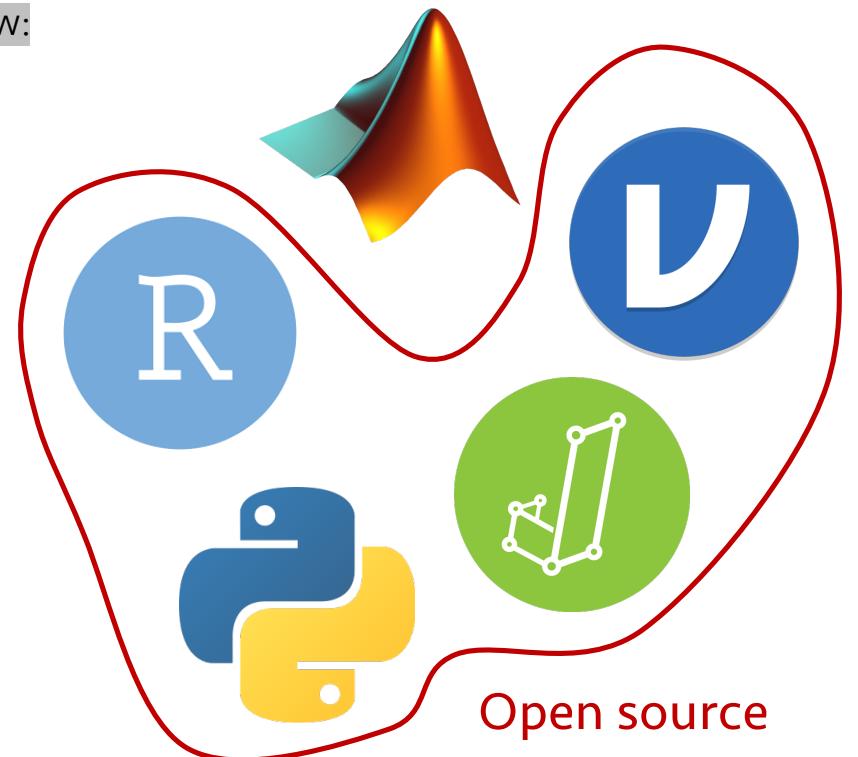
What I am using now:

Basic statistics

- Matlab
- R
- JASP
- JAMOVI
- Python

Multivariate

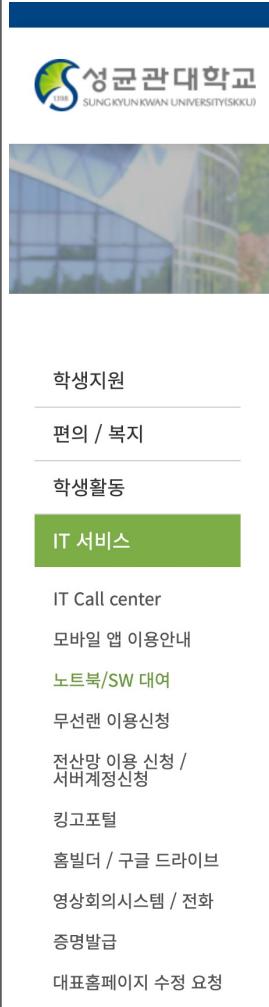
- Matlab
- Python
- R



Cultural changes for the last 20 years

The ones that I have used before... and now...

* 학내 공용 SW 목록 at SKKU

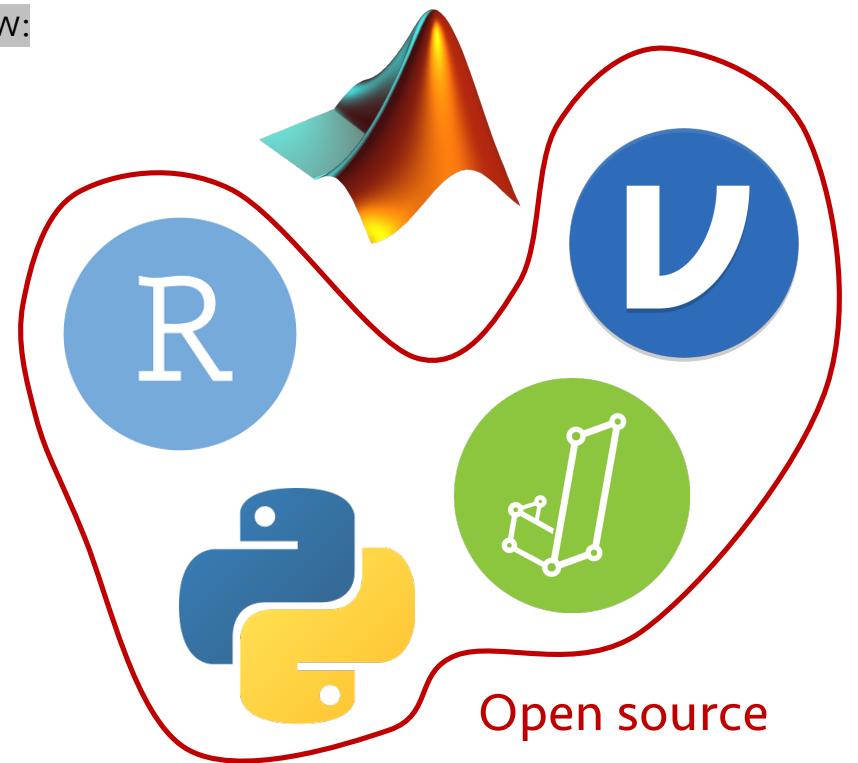


SW 종류	사용 지역		사용자 구분		
	인사캠	자과캠	교원	직원	학생
MS Window OS	○	○	○	○	×
MS Office	○	○	○	○	×
MS O365 Service	○	○	×	×	○
Adobe	○	○	○	○	×
아래한글	○	○	○	○	×
MATLAB	×	○	○	○	○
SAS	○	○	○	○	○
SPSS	○	○	○	○	×
STATA	○	○	○	○	×
V3	○	○	○	○	×
ALTOOLS	○	○	○	○	×
IT(OA)강좌	○	○	○	○	○

학내 지원
학생활동
IT 서비스
IT Call center
모바일 앱 이용안내
노트북/SW 대여
무선랜 이용신청
전산망 이용 신청 / 서버계정신청
킹고포털
홈빌더 / 구글 드라이브
영상회의시스템 / 전화
증명발급
대표홈페이지 수정 요청

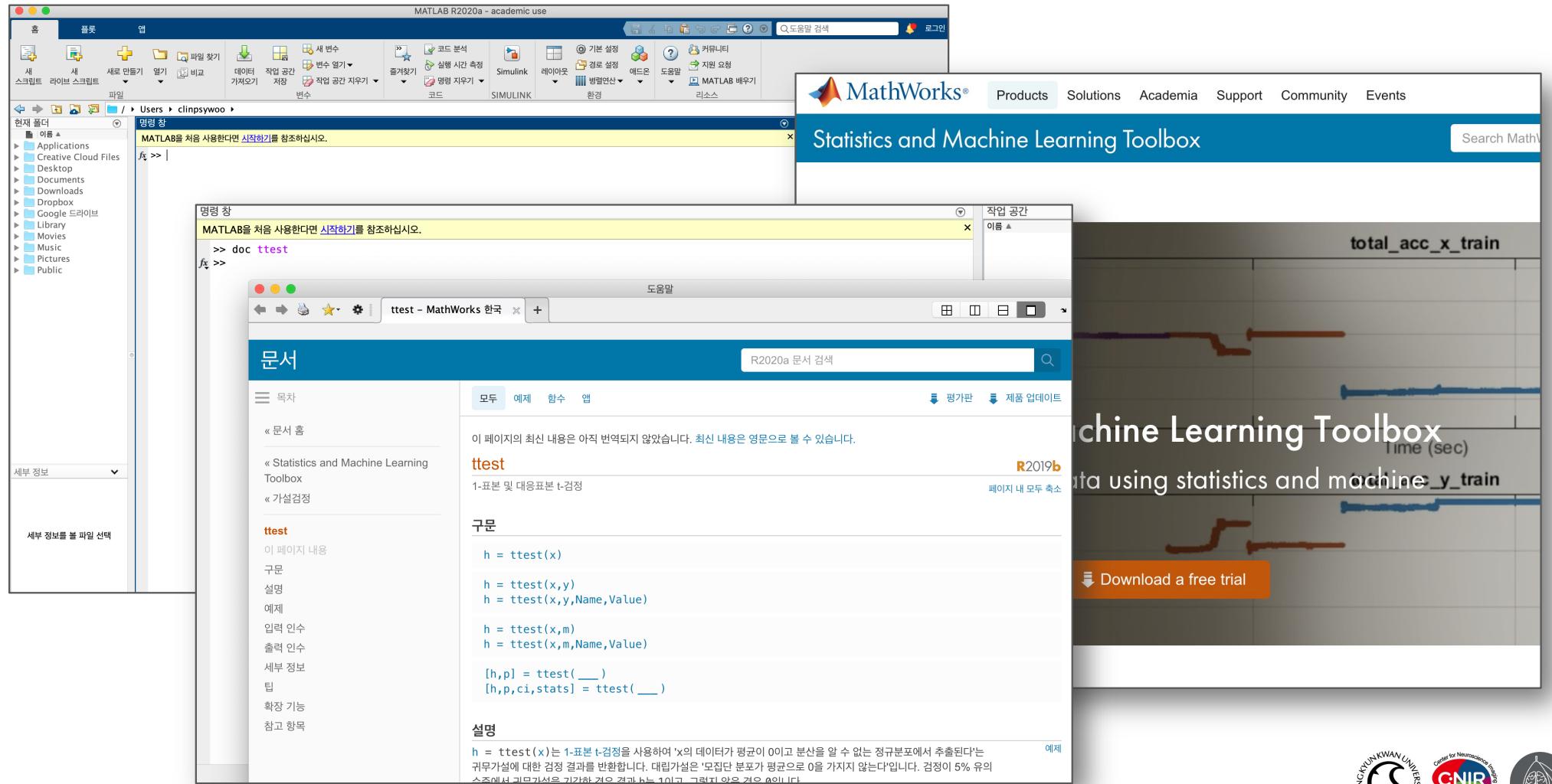
CHOONG-WAN WOO | COCOAN lab | <http://cocoanlab.github.io>

now:



Cultural changes for the last 20 years

Matlab



Matlab

Strengths:

- Command line coding
- Simple and fast
- Well-developed toolboxes (e.g., statistics and machine learning toolbox)
- Open source tools for neuroimaging and others (thick user base)

Weaknesses:

- Expensive (what if you graduate the school?)
- The popularity is decreasing (the user base is getting weaker)
- Deep learning?

R

The screenshot shows the RStudio interface with three main panes:

- Console (Left Pane):** Displays the R startup message and a command prompt (> |) where the user has typed `?t.test`.
- Global Environment (Middle Pane):** Shows the environment is empty.
- Help (Right Pane):** Provides documentation for the `t.test` function, including the description, usage, arguments, and source code.

```
R version 4.0.2 (2020-06-22) -- "Taking Off Again"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin19.5.0 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |
```

Environment **History** **Connections** **Tutorial**

Project: (None)

Global **File** **Edit** **View** **Plots** **Packages** **Help** **Viewer**

Go to file/function Addins

Console Terminal Jobs

~ /

> ?t.test

>

Environment is empty

Files Plots Packages Help Viewer

R: Student's t-Test Find in Topic

t.test {stats}

Student's t-Test

Description

Performs one and two sample t-tests on vectors of data.

Usage

```
t.test(x, ...)
```

```
## Default S3 method:
t.test(x, y = NULL,
      alternative = c("two.sided", "less", "greater"),
      mu = 0, paired = FALSE, var.equal = FALSE,
      conf.level = 0.95, ...)
```

```
## S3 method for class 'formula'
t.test(formula, data, subset, na.action, ...)
```

Arguments

x	a (non-empty) numeric vector of data values.
y	an optional (non-empty) numeric vector of data values.
alternative	a character string specifying the alternative hypothesis, must be one

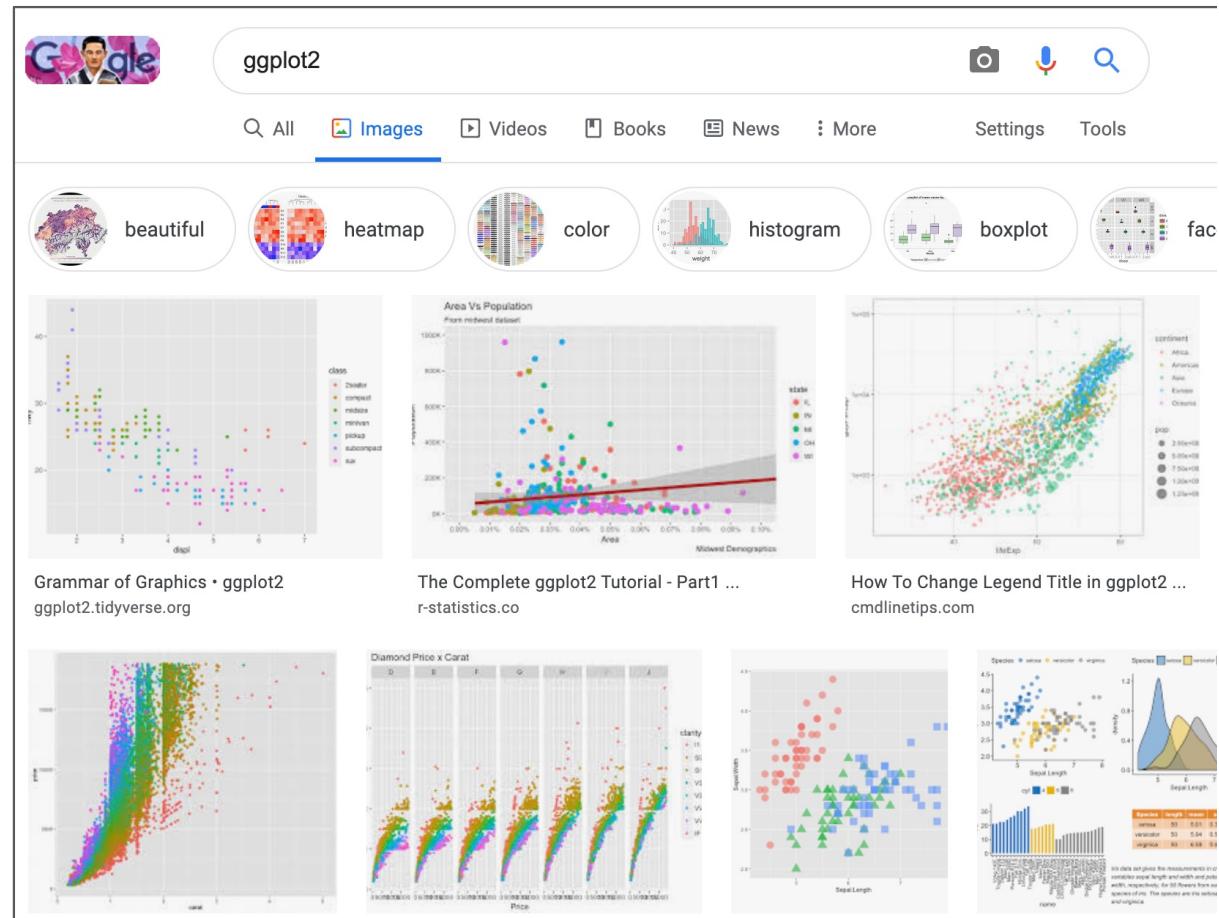
R

Strengths:

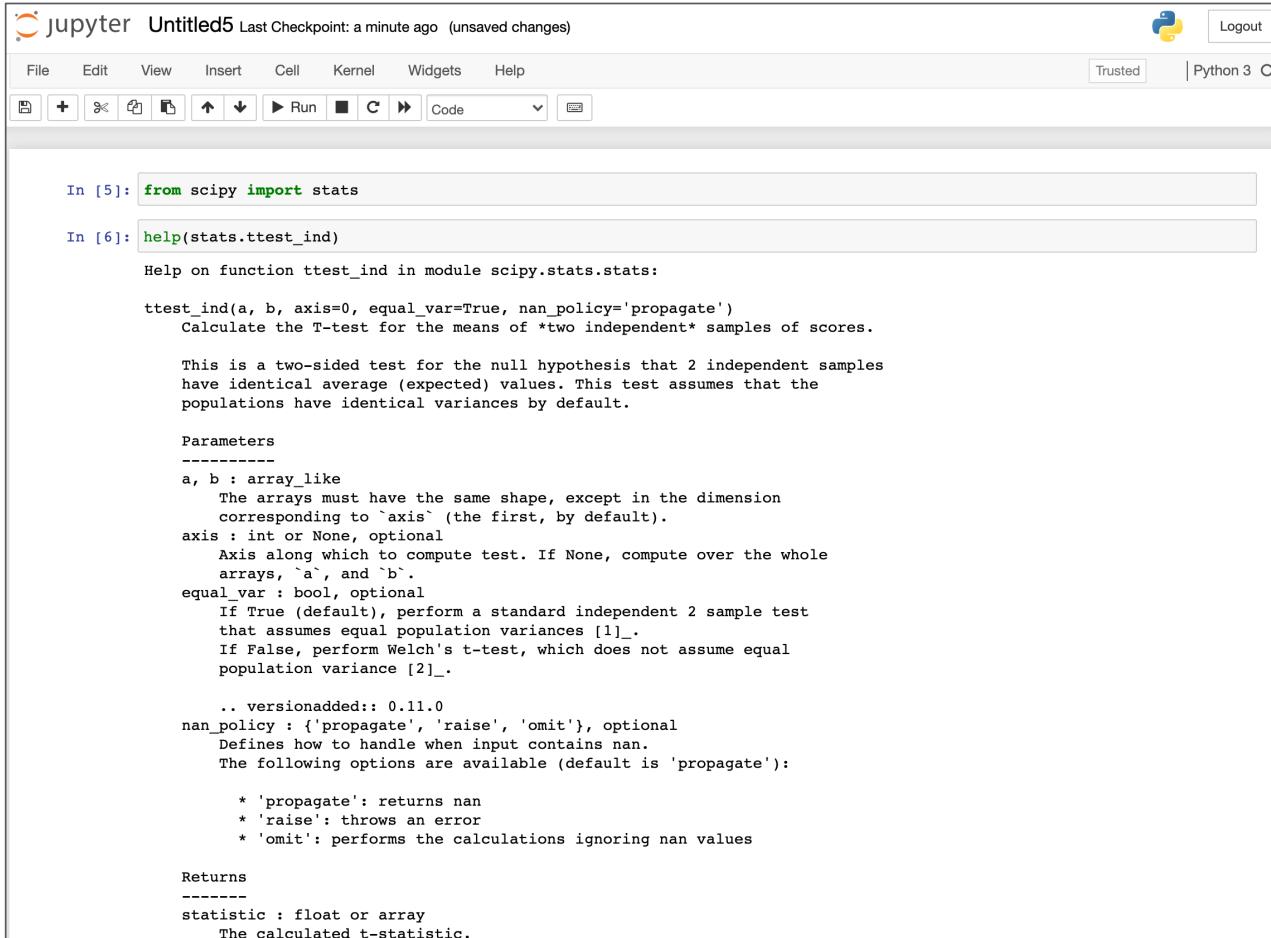
- Command line coding
- Simple and fast
- Thick user base!!! Esp. in statistics;
so many useful packages for advanced stats
- Good visualization tools (e.g., ggplot2)
- The popularity is increasing

Weaknesses:

- No good connection with neuroimaging
- The popularity is lower than python
- Deep learning?



Python



The screenshot shows a Jupyter Notebook interface with the following content:

```
In [5]: from scipy import stats
```

```
In [6]: help(stats.ttest_ind)
```

Help on function `ttest_ind` in module `scipy.stats.stats`:

```
ttest_ind(a, b, axis=0, equal_var=True, nan_policy='propagate')
Calculate the T-test for the means of *two independent* samples of scores.

This is a two-sided test for the null hypothesis that 2 independent samples
have identical average (expected) values. This test assumes that the
populations have identical variances by default.

Parameters
-----
a, b : array_like
    The arrays must have the same shape, except in the dimension
    corresponding to `axis` (the first, by default).
axis : int or None, optional
    Axis along which to compute test. If None, compute over the whole
    arrays, `a`, and `b`.
equal_var : bool, optional
    If True (default), perform a standard independent 2 sample test
    that assumes equal population variances [1].
    If False, perform Welch's t-test, which does not assume equal
    population variance [2].
    .. versionadded:: 0.11.0
nan_policy : {'propagate', 'raise', 'omit'}, optional
    Defines how to handle when input contains nan.
    The following options are available (default is 'propagate'):
        * 'propagate': returns nan
        * 'raise': throws an error
        * 'omit': performs the calculations ignoring nan values

Returns
-----
statistic : float or array
    The calculated t-statistic.
```

Python

Strengths:

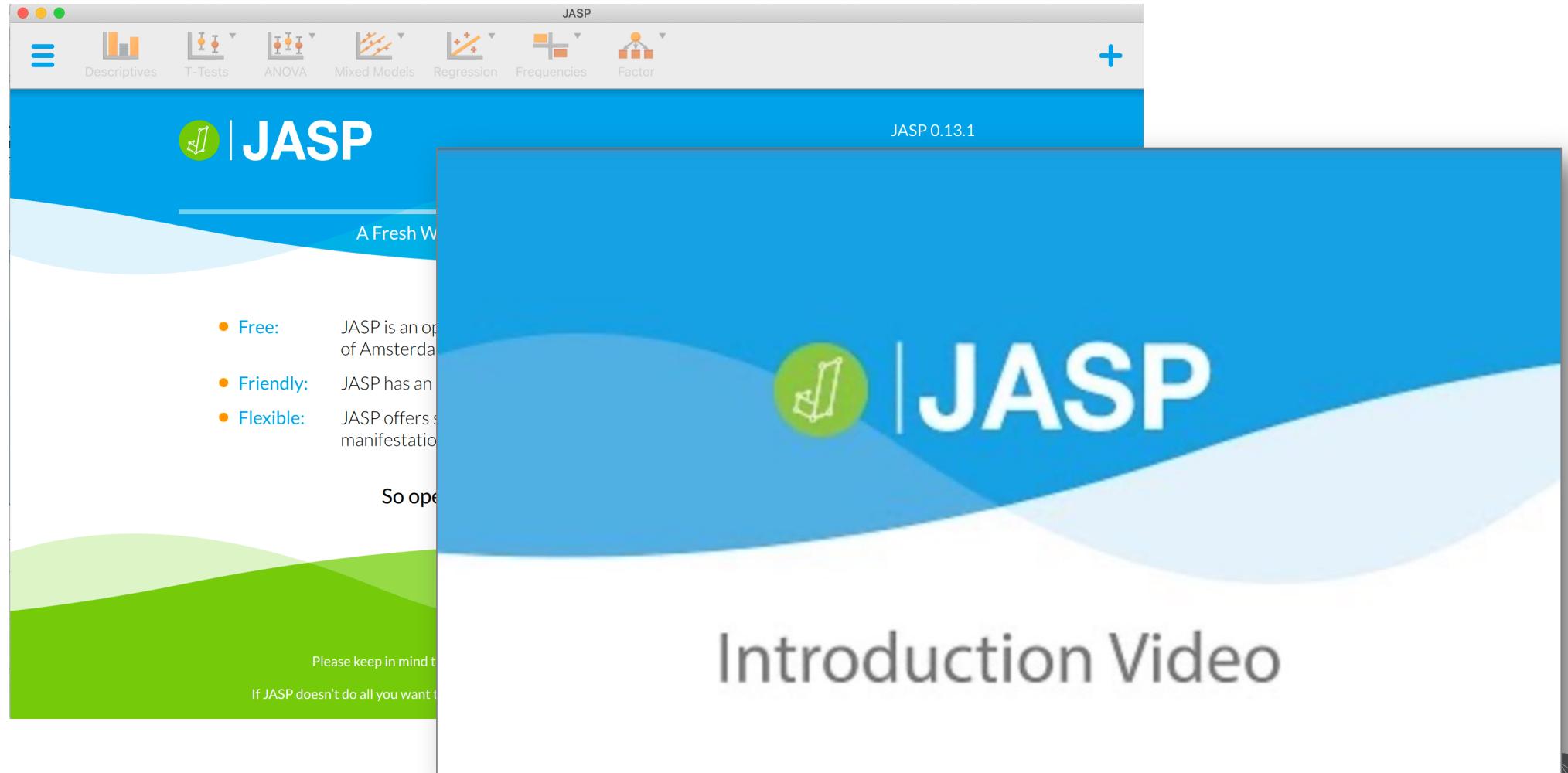
- Command line coding
- Simple and fast
- Thick user base! Esp. in machine learning and others;
so many useful packages for everything
- Deep learning!
- The popularity is rapidly increasing

Weaknesses:

- Not really good for statistics (maybe compared to R or matlab), but still okay

JASP

<https://jasp-stats.org/2017/06/27/new-5-minute-video-introducing-jasp/>



JASP

Strengths:

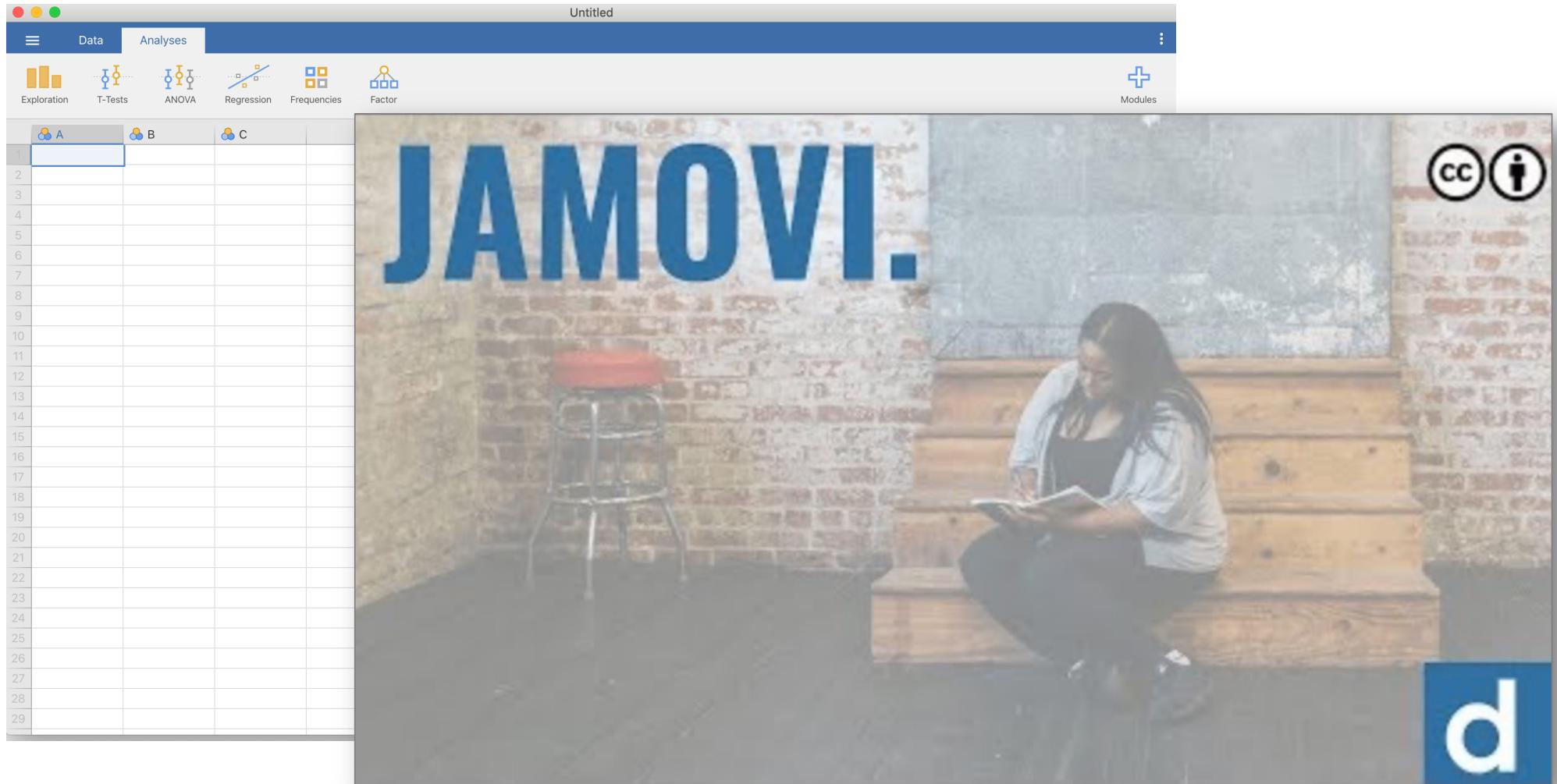
- Free!
- Graphical user interface
- Simple and easy
- Bayesian analysis

Weaknesses:

- Not command line
- No flexibility
- No development

JAMOVI

https://www.youtube.com/watch?v=Ej9e8lzaeDE&feature=emb_logo



JAMOVI (similar to JASP)

Strengths:

- Free!
- Graphical user interface
- Simple and easy
- Bayesian analysis

Interestingly, JAMOVI was developed by the previously JASP developers.

Weaknesses:

- Not command line
- No flexibility
- No development

<http://blog.efpsa.org/2017/03/23/introducing-jamovi-free-and-open-statistical-software-combining-ease-of-use-with-the-power-of-r/>

Thus, our choice for the class is...

JAMOVI

(JASP)

Matlab

R