



SKKU Biostats and Big data

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



Lecture 01

Course overview

Introduction

Lab webpage

<https://cocoanlab.github.io/>

COCOAN lab

Computational Cognitive Affective Neuroscience Laboratory



Cocoan lab is a neuroimaging research lab led by Dr. Choong-Wan Woo.

We are joining the [IBS Center for Neuroscience Imaging Research \(CNIR\)](#) at [Sungkyunkwan University \(Wikipedia\)](#) located in Suwon, South Korea (starting March 2017).

The **keywords** of our research include:

fMRI; Machine learning; Neuroimaging biomarkers; Data science; Translational research; Predictive modeling; Brain decoding; Encoding-decoding model; Pain; Emotions; Psychiatric and neurologic disorders; Mind-body interaction; Behavioral medicine; Network science; Psychological and social pain modulation; Emotion regulation, and more.

The mission of our lab is to understand **pain** and **emotions** in the perspective of data science, cognitive/affective/social neuroscience, and psychology. We also aim to develop clinically useful neuroimaging models and tools that can be used and shared across different research groups and clinical settings.

Our main research tools include functional Magnetic Resonance Imaging (fMRI; we're using 3T and 7T MRI), psychophysiology measures (skin conductance, pupillometry, electrocardiogram, respiration), electroencephalogram (EEG), and other behavioral measures such as facial expression, eye movement, etc. Most importantly, we use data science (computational) tools to model and understand affective, cognitive, and behavioral responses.



Survey

Please complete this before our first livestream class (2/28):

<https://forms.gle/GgvHDnYGD4JF6nqd6>

'Biostat and Big Data' course survey (Spring 2022)

This survey isn't meant to test your knowledge. Rather, it'll help us adjust the class just for you, and make it fun! (이 설문조사는 여러분을 평가하기 위한 것이 아닙니다. 수업이 여러분에게 더 도움이 되고, 즐겁게 만들기 위한 것이니 솔직하고 편하게 작성해주세요!)

 choongwan.woo@gmail.com (not shared) [Switch account](#) 

* Required

Name *

Your answer

Student ID (e.g., 2019123456) *

Your answer

Do you have any experience with data analysis? (데이터 분석을 해본 경험이 있나요?) *

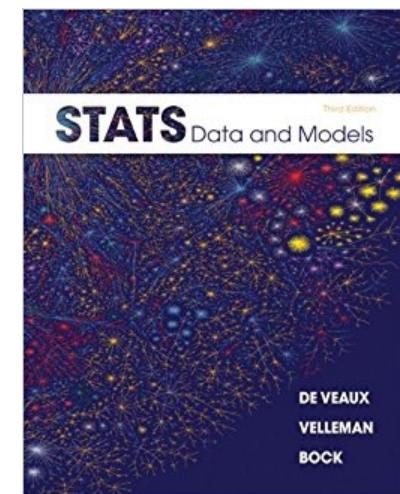
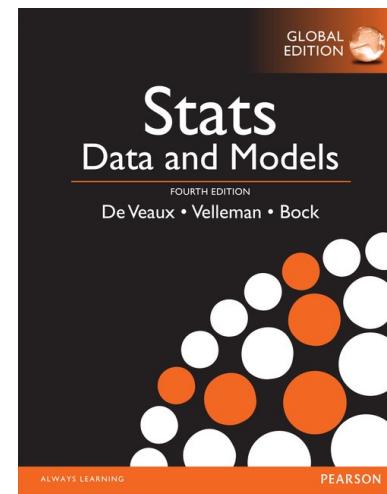
Yes

No

If you have experience with data analysis, please describe it here: (데이터 분석을 해

Course overview

- Biostats and Big data (previously GBME's Probability and Statistics)
- Class code: GBE 3064
- Intro Stats
- Textbook: "Stats: Data and Models" by De Veaux, Velleman, and Bock, Fourth Edition
 - Third edition also works!



Course overview

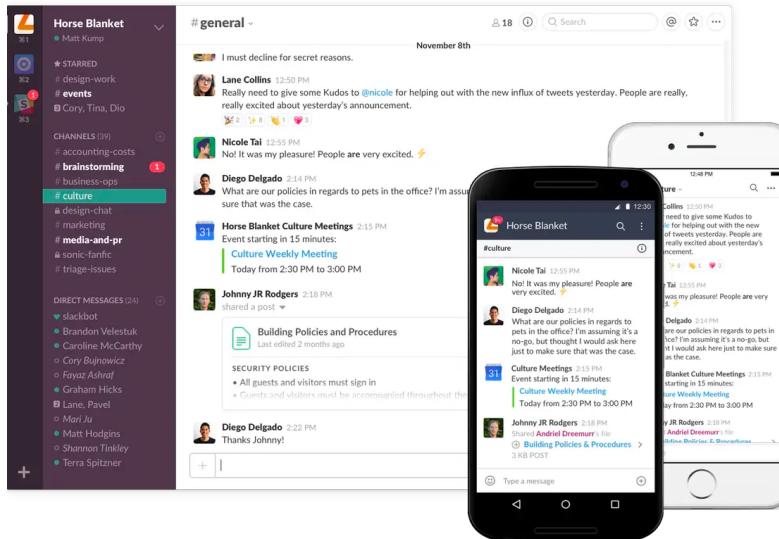
- Biostats and Big data (previously GBME's Probability and Statistics)
- Class code: GBE 3064
- Intro Stats
- Textbook: "Stats: Data and Models" by De Veaux, Velleman, and Bock, Fourth Edition
 - Third edition also works!
 - And also I'm writing a textbook in Korean.. Maybe coming soon.. I will share some of them.
- Two useful web-based platforms for the class:
 1. **Slack** for class-related discussion and questions
 2. **GitHub** for posting class-related materials, assignments, data, codes, etc





- A great platform for team communications
- I will share the link for the class during the first week of the class (you should join)

Works for both desktop and mobile phone



Real examples

The screenshots illustrate the use of Slack for academic collaboration. The '#general' channel shows typical team communication. The '#advanced-fmri-analysis' channel is used for more specialized discussions, including sharing research papers and tools.

#general

#advanced-fmri-analysis

Dec 13th, 2019

Andrew You 1001 PM

Hello people. Please download the following for tmrw's Network analysis tutorial.
SPM12 and matlab with statistic toolbox is required

GraphVar_2.02b.zip
40 MB Zip

Dec 12th, 2019

Wani 8:17 PM

@Andrew You if you could post your ppt, that would be helpful!

@channel It was a great class. I enjoyed it a lot. See many of you in the next semester. I hope we can cover some new methods like

- Non-linear dimension reduction methods (e.g., UMAP, isomap, etc.)
- Shared response model
- NLP or Topic modeling
- Different methods in network analysis (e.g., multilayer community detection)
- Multivariate pattern dependence (Anzellotti et al., 2017)
- Temporal communities by trajectory clustering (Thompson1 et al., 2019)
- Deep transfer learning link
- Other neuroimaging analysis platforms we didn't cover: Neuroscout, MRIQC, fMRIDenoise, Neurolearn

and more

Message #advanced-fmri-analysis

Dec 13th, 2019

Wani 8:17 PM

@Andrew You if you could post your ppt, that would be helpful!

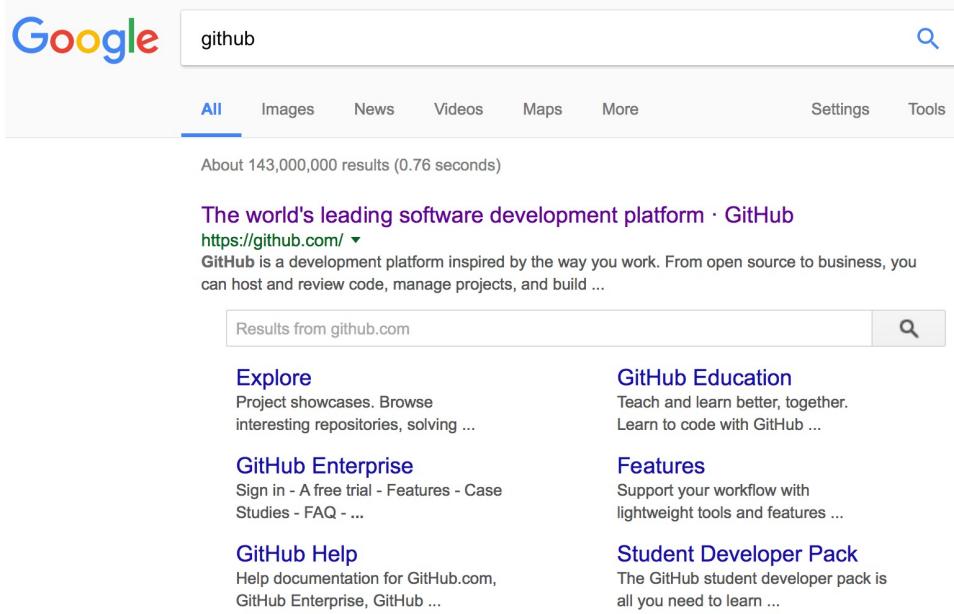
8:18 @channel It was a great class. I enjoyed it a lot. See many of you in the next semester. I hope we can cover some new methods like

- Non-linear dimension reduction methods (e.g., UMAP, isomap, etc.)
- Shared response model
- NLP or Topic modeling
- Different methods in network analysis (e.g., multilayer community detection)
- Multivariate pattern dependence (Anzellotti et al., 2017)
- Temporal communities by trajectory clustering (Thompson1 et al., 2019)
- Deep transfer learning link
- Other neuroimaging analysis platforms we didn't cover: Neuroscout, MRIQC, fMRIDenoise, Neurolearn

and more



- Most popular software development platform: version control, easy-sharing scripts and data, etc.



Google search results for "github":

About 143,000,000 results (0.76 seconds)

The world's leading software development platform · GitHub
<https://github.com/> ▾

GitHub is a development platform inspired by the way you work. From open source to business, you can host and review code, manage projects, and build ...

Results from github.com

Explore
 Project showcases. Browse interesting repositories, solving ...

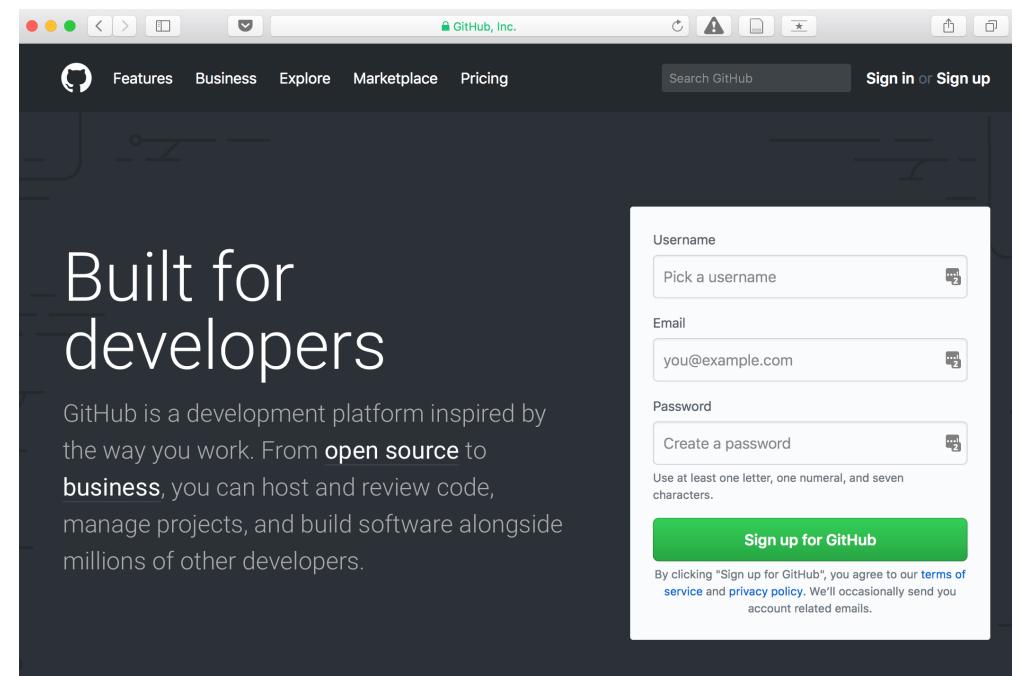
GitHub Enterprise
 Sign in - A free trial - Features - Case Studies - FAQ - ...

GitHub Help
 Help documentation for GitHub.com, GitHub Enterprise, GitHub ...

GitHub Education
 Teach and learn better, together. Learn to code with GitHub ...

Features
 Support your workflow with lightweight tools and features ...

Student Developer Pack
 The GitHub student developer pack is all you need to learn ...



Built for developers

GitHub is a development platform inspired by the way you work. From **open source** to **business**, you can host and review code, manage projects, and build software alongside millions of other developers.

Sign up for GitHub

By clicking "Sign up for GitHub", you agree to our [terms of service](#) and [privacy policy](#). We'll occasionally send you account related emails.



- Most popular software development platform: version control, easy-sharing scripts and data, etc.
- Everyone uses GitHub for the development, including Google, Facebook, etc.

Google <https://opensource.google.com/>

Repositories 1059 People 1,485

Search repositories... Type: All Language: All

oss-fuzz
OSS-Fuzz - continuous fuzzing of open source software
security fuzzing stability
Shell ★ 1,939 313 Updated a minute ago

syzkaller
syzkaller is an unsupervised, coverage-guided Linux system call fuzzer
testing linux security kernel fuzzing fuzz-testing
Go ★ 1,157 220 Updated 13 minutes ago

ci_edit
python console coding text-editor
Python ★ 21 13 Updated 25 minutes ago

Facebook <https://code.facebook.com/projects/>

Repositories 179 People 170

Search repositories... Type: All Language: All

buck
A fast build system that encourages the creation of small, reusable modules over a variety of platforms and languages.
android python java ios buck build-tool
Java ★ 5,025 776 Updated 22 minutes ago

PathPicker
PathPicker accepts a wide range of input -- output from git commands, grep results, searches -- pretty much anything. After parsing the input, PathPicker presents you with a nice UI to select which files you're interested in. After that you can open them in your favorite editor or execute arbitrary commands.
Python ★ 3,469 215 Updated an hour ago

chisel
Chisel is a collection of LLDB commands to assist debugging iOS apps.
Python ★ 6,159 529 Updated 3 hours ago



- Most popular software development platform: version control, easy-sharing scripts and data, etc.
- Everyone uses GitHub for the development, including Google, Facebook, etc.
- More and more people are creating and sharing a GitHub repository for their research and publications:



- Most popular software development platform: version control, easy-sharing scripts and data, etc.
- Everyone uses GitHub for the development, including Google, Facebook, etc.
- More and more people are creating and sharing a GitHub repository for their research and publications:
 - E.g., Lee et al., 2021, *Nature Medicine*

ARTICLES
<https://doi.org/10.1038/s41591-020-1142-7>

A neuroimaging biomarker for sustair experimental and clinical pain

Jae-Joong Lee^{1,2}, Hong Ji Kim^{1,2}, Marta Čeko^{3,4}, Bo-yong Park^{1,5}, Soo Ahn Mathieu Roy^{7,8}, Seong-Gi Kim^{1,2}, Tor D. Wager^{1,9} and Choong-Wan Woo^{1,2}

nature medicine

[Check for updates](#)

Data availability
The dynamic functional connectivity data of cocoanlab.github.io/tops and <https://github.com/cocoanlab/tops> tutorial. In addition, all the data to generate results in this study are publicly available at <http://www.openpain.org>. The data that were not used in this study can also be shared through <https://cocoanlab.github.io/tops>.

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

cocoanlab / tops Public

Code Issues Pull requests Actions Projects Wiki Security Insights

master 1 branch 0 tags

jaejoonglee92 update MCA form d27516a on 21 Oct 2021 34 commits

File	Description	Last Commit
functions	update PDF and HTML	15 months
model	update script	16 months
scripts	Update README.md	15 months
.gitignore	update	16 months
LICENSE	Create LICENSE	14 months
README.md	update MCA form	4 months

README.md

Tonic pain signature (ToPS)

Our class website: https://github.com/wanirepo/Stats_2022Spring

- README.md is the syllabus.
- Let's visit the website and read the syllabus together.

The screenshot shows a GitHub repository page for 'wanirepo / Stats_2022Spring'. The repository is public and contains one branch ('master') and no tags. The README.md file is displayed, containing the following text:

```
2022 Spring "Biostats and Big Data" at SKKU GBME

• Lecturer: Choong-Wan Woo, Ph.D. Assistant professor (GBME).
• Office: N-center, 86335
• Web: Cocoan lab
• E-mail: please send the message through direct message in slack. If you haven't joined the slack, send me the email to choongwan.woo at gmail dot com
• Class: Mon 1:30-2:45, Wed 12:00-1:15 (online; youtube livestream)
• Office hours: Mon 10:00-12:00

Download

You can download the class materials using the following command line.
```

https://github.com/wanirepo/Stats_2022Spring

Evaluation

Absolute evaluation will be used for this course.

1. Attendance (40%)
2. Participation (including pop-questions) (35%)
3. Final exam (25%)

For *attendance*,

- I will give you pop-quizes during the livestreaming class using google forms
- For each pop-quiz item, you will need to report your **student ID** and the google forms will record the **time** you answered
- Later, I will count the **student ID** for each item for a specific **period**

https://github.com/wanirepo/Stats_2022Spring

Textbooks

Main textbook:

"Stats: Data and Models" by De Veaux, Velleman, and Bock

Supplements:

"Statistical Thinking for the 21st Century" by Russ Poldrack [Link](#)

"Seeing theory" [Link](#)

"The Seven Pillars of Statistical Wisdom" by Stephen M. Stigler [Amazon Link](#)

"History of Statistics Reading Group" (CMU) [Link](#)

Softwares

Teaching how to use statistics packages or programming is not the main focus of this class. This class is more about statistical methods and theories. However, software packages and computer programming are actually essential in learning statistics. Therefore, I will provide two lectures about new software packages for statistical analysis, [JAMOVI](#) and [JASP](#). These are brand-new and open-source tools for statistical analysis. In addition, I might use Matlab, Python, or R sometimes. You can download Matlab through SKKU. Python and R are open-source programming languages. There will be "Biostats and Bigdata 2" next semester in which you can learn more about JAMOVI.

To-do before the next class

- Survey before the next class (2/28): <https://forms.gle/GgvHDnYGD4JF6nqd6>
- Watch the first lecture
 - [V01] Statistics and Data: <https://youtu.be/hkMrXhG0ot0>
- I will send the email or icampus message to ask you to sign up for slack
- Next class: 2/28 Monday 1:30 PM (Link: <https://youtu.be/6lrKqKfeaBc>)