



BIMM 143

Course Summary & GitHub Portfolio

Class 20

Barry Grant

UC San Diego

<http://thegrantlab.org/bimm143>

Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bimm143_s24)
- CAPs evaluation ([Link](#))
- Final exam
 - ➡ Test structure, guidelines and rules
 - ➡ Topics and example questions
 - ➡ Exam preparation, discussion and open study
- Polish our GitHub content and publish your own website portfolios

Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bimm143_s24)
- CAPs evaluation ([Link](#))
- Final exam
 - Test structure, guidelines and rules
 - Topics and example questions
 - Exam preparation, discussion and open study
- Polish our GitHub content and publish your own website portfolios

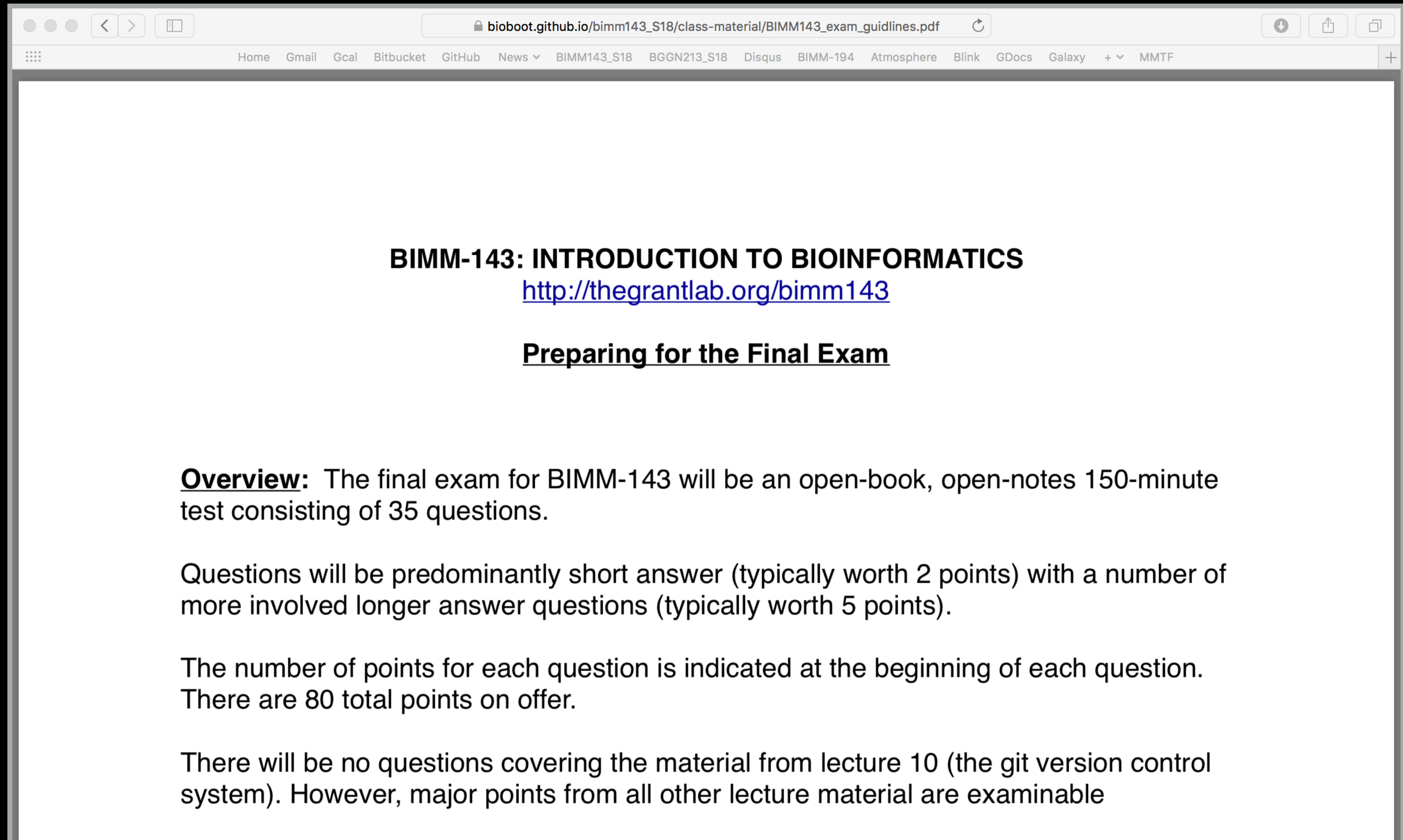


Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://etherpad.wikimedia.org/p/bimm143_s24)
- CAPs evaluation ([Link](#))
- Project troubleshooting
 - Test structure, guidelines and rules
 - Topics and example questions
 - Exam preparation, discussion and open study
- Polish our GitHub content and publish your own website portfolios



https://bioboot.github.io/bimm143_F22/class-material/BIMM143_exam_guidelines.pdf



- Q1.** Did you enjoy this course? (Rank in relation to others you have experienced at UCSD)?
- Q2.** Should this course be offered again?
- Q3.** If so what changes would you recommend for this course? (e.g. more/less DataCamp & Projects)
- Q4.** Was the course effectively organized (lecture and lab material online vs handout or TritonEd site)?
- Q5.** What advice would you give to another student who is considering taking this course?
- Q6.** Considering both the limitations and possibilities of the subject matter and the course, how would you rate the overall effectiveness of this course and instructor?
- Q7.** Do you agree or disagree - The course developed my abilities and skills for the subject?
- Q8.** On average, how many hours per week have you spent on this course, including attending classes, doing homework's and assignments?
- Q9.** Any other comments you would like to share?

EtherPad: https://etherpad.wikimedia.org/p/bimm143_s24

Form Version: <https://tinyurl.com/bimm143-end>

Thank you very much!

Announcement!

Bonus:

Bioinformatics & Genomics in Industry

Live Stream Video

Enjoy a set of short open ended guest lectures from leading genomic scientists at **Illumina Inc.**, **Synthetic Genomics Inc.**, and the **La Jolla Institute for Allergy and Immunology**. Feel free to contact these scientists for networking and to have your questions about industry careers in Bioinformatics and Genomics answered.

Bonus:

GitHub Spit & Polish

Introduction to Bioinformatics



A demo site of students cool class web site

[View the Project on GitHub](#)
bioboot/bimm143-1

This project is maintained by [bioboot](#)

Hosted on GitHub Pages — Theme by [orderedlist](#)

Bioinformatics Class

This is my repository for my Bioinformatics class from UC San Diego in S18.

Index of Material

Introductory Material: Working With R

Class 5 - Basic Data Exploration and Visualization in R HTML, MD, Rmd

Class 6 - Creating R Functions

Class 7 - R Packages, working with CRAN, and working with Bioconductor

Using R and Other Tools for Bioinformatics Analysis

Class 8 - An Introduction to Machine Learning (Heirarchical Clustering)

Class 9 - Analyzing High Dimensional Datasets and Unsupervised Learning

Class 11 - Structural Bioinformatics: Analyzing Protein Structure and Function

Class 12 - Drug Discovery: Techniques and Analysis

Class 13 - Genome Informatics and High Throughput Sequencing (NGS, RNA-Seq, and FastQC)

Class 14 - Transcriptomics and RNA-Seq Analysis

In your web browser navigate to your GitHub class repository < <https://github.com/> >

Side-note:

To find the link to your GitHub repository from RStudio, open one of your past class projects and in the terminal type:

```
git remote -v
```

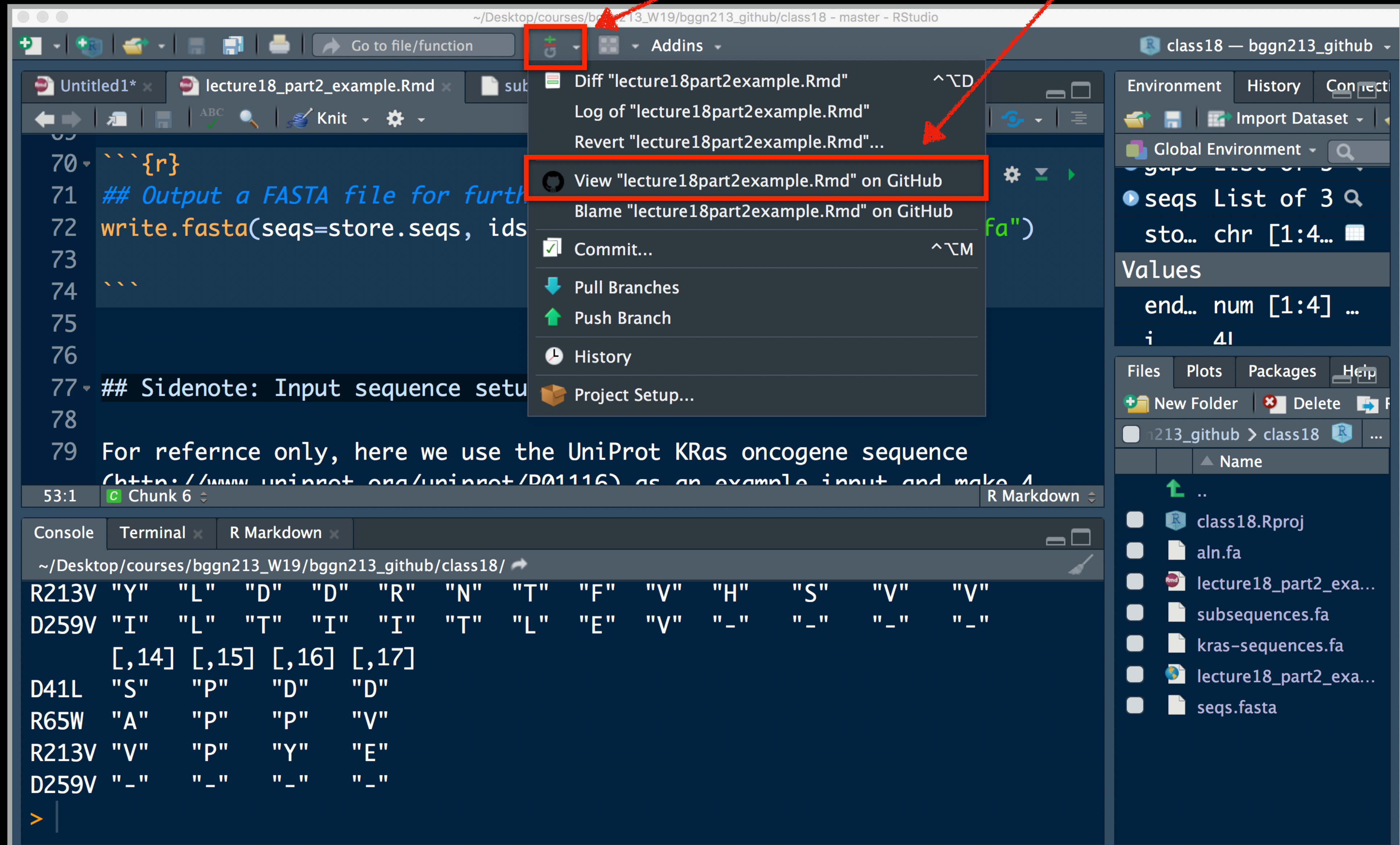
The screenshot shows a web browser window displaying the GitHub repository page for `bioboot/bggn214`. The browser's address bar shows `github.com`. The repository page includes a search bar, navigation links for Pull requests, Issues, Codespaces, Marketplace, and Explore, and a header for the repository `bioboot / bggn214` (Public). Below the header, there are tabs for Code, Issues, Pull requests, Actions, Projects, Wiki, Security, Insights, and Settings. The main content area shows the repository's file structure with a commit history table. The table lists the following commits:

Commit Hash	Commit Message	Time Ago	Commits
16c0acd	13 days ago	4 commits	
class10	Add class 10	13 days ago	
README.md	A change on my laptop.	13 days ago	

On the right side of the repository page, there is an 'About' section for `class work` with the following statistics:

- Readme
- 0 stars
- 1 watching
- 0 forks

Or for a given GitHub tracked file click **GIT** icon and **"VIEW on GITHUB"**



The screenshot shows the RStudio interface with the file `lecture18_part2_example.Rmd` open. The Git icon (a green 'G' with a plus sign) in the top toolbar is highlighted with a red box. A red arrow points from the text 'GIT icon' in the instruction above to this icon. Another red arrow points from the text '"VIEW on GITHUB"' in the instruction above to the 'View "lecture18part2example.Rmd" on GitHub' option in the Git menu. The menu is open, showing options like 'Diff', 'Log', 'Revert', 'View on GitHub', 'Blame', 'Commit', 'Pull Branches', 'Push Branch', 'History', and 'Project Setup...'. The 'View on GitHub' option is highlighted with a red box. The main editor shows R Markdown code, and the console at the bottom displays the output of the `write.fasta` function, showing a FASTA file format with sequence headers and nucleotide sequences.

```
70 ````{r}
71 ## Output a FASTA file for further analysis
72 write.fasta(seqs=store.seqs, ids=store.names, file="seqs.fasta")
73 
74 ````
75 
76 
77 ## Sidenote: Input sequence setup
78 
79 For reference only, here we use the UniProt KRas oncogene sequence
80 (http://www.uniprot.org/uniprot/P0116) as an example input and make 4
```

Console output:

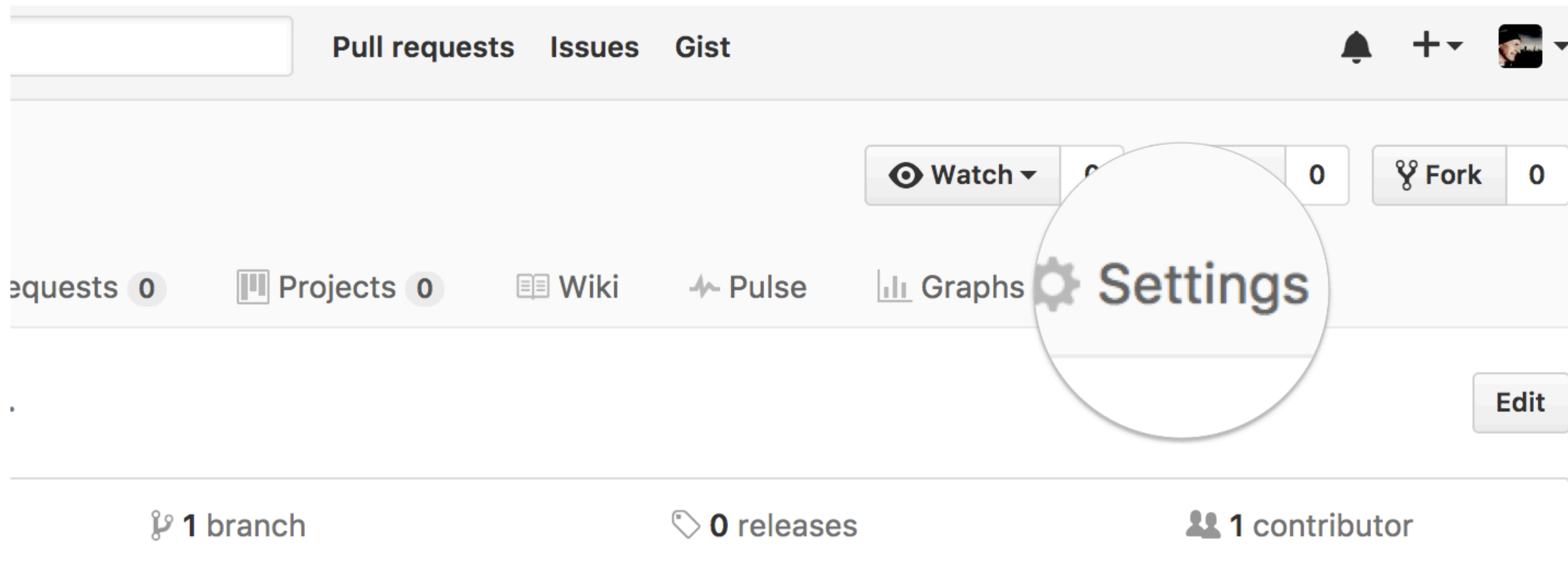
```
~/Desktop/courses/bgg213_W19/bgg213_github/class18/
R213V "Y" "L" "D" "D" "R" "N" "T" "F" "V" "H" "S" "V" "V"
D259V "I" "L" "T" "I" "I" "T" "L" "E" "V" "_" "_" "_" "_"
      [,14] [,15] [,16] [,17]
D41L  "S"  "P"  "D"  "D"
R65W  "A"  "P"  "P"  "V"
R213V "V"  "P"  "Y"  "E"
D259V "_"  "_"  "_"  "_"
>
```


1

Repository Settings

Head over to your [GitHub](#) class repository (where you are “pushing” all your class work)

Click on the **Settings** tab.



Theme chooser

Scroll down to the **GitHub Pages** section. ~~Press **Choose a theme.**~~
And set the **Source** to "main branch"

General

Access

Collaborators

Moderation options

Code and automation

Branches

Tags

Actions

Webhooks

Environments

Codespaces

Pages

Security

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Build and deployment

Source

Deploy from a branch


Branch

GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository. [Learn more.](#)

main

/ (root)

Save

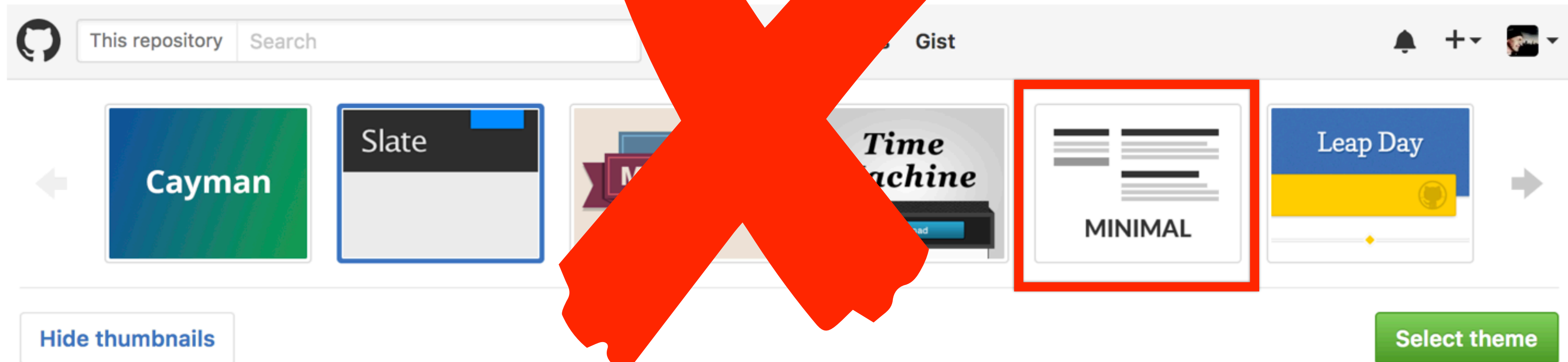


Publish privately to people with read access to this repository
Try risk-free for 30 days using a GitHub Enterprise organization, or [learn more about changing the visibility of your GitHub Pages site.](#) ×

Pick a theme

Choose one of the themes from the carousel at the top.

When you're done, click **Select theme** on the right.



Slate theme

Slate is a theme for GitHub Pages.

[View on GitHub](#)



Side-note:

Scroll down again to the **GitHub Pages** section to find the link to your new website.
Open this link in a **New Tab** of your browser:

GitHub Pages

GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository.

Your site is ready to be published at https://bioboot.github.io/bimm143_serina_f18/.

Source

Your GitHub Pages site is currently being built from the master branch. [Learn more](#).

master branch ▾

Save

Theme Chooser

Select a theme to publish your site with a Jekyll theme. [Learn more](#).

Your site is currently using the Minimal theme.

Change theme

Edit content

Back on the repository main page use the [GitHub](#) online editor to add content.
In particular, add links to each classes .MD file

The screenshot shows the GitHub interface for the repository 'jldec / new-pages-site'. At the top, there are buttons for 'Watch', 'Star', and 'Fork', each with a count of 0. Below these are tabs for 'Code', 'Issues', 'Pull requests', 'Projects', 'Wiki', 'Pulse', 'Graphs', and 'Settings'. The 'Code' tab is selected, and the file 'README.md' is shown. The editor interface includes a toolbar with 'Edit file' and 'Preview changes' tabs, and settings for 'Spaces' (2) and 'Soft wrap'. The main content area displays the README.md file with the following text:

```
1 ## Welcome to GitHub Pages
2
3 You can use the [editor on GitHub](https://github.com/jldec/new-pages-site/edit/master/README.md) to maintain and preview the content
  for your website in Markdown files.
4
5 Whenever you commit to this repository, GitHub Pages will run [Jekyll](https://jekyllrb.com/) to rebuild the pages in your site, from
  the content in your Markdown files.
6
7 ### Markdown
8
9 Markdown is a lightweight and easy-to-use syntax for styling your writing. It includes conventions for
10
11 ```markdown
12 Syntax highlighted code block
13
14 # Header 1
```

Commit

Enter a commit comment and click on **Commit changes** below the editor.

```
35   ### Support or Contact
36
37   Having trouble with Pages? Check out our \[documentation\]\(https://help.github.com/categories/gi
38    \(https://github.com/contact\) and we'll help you sort it out.
```



Commit changes

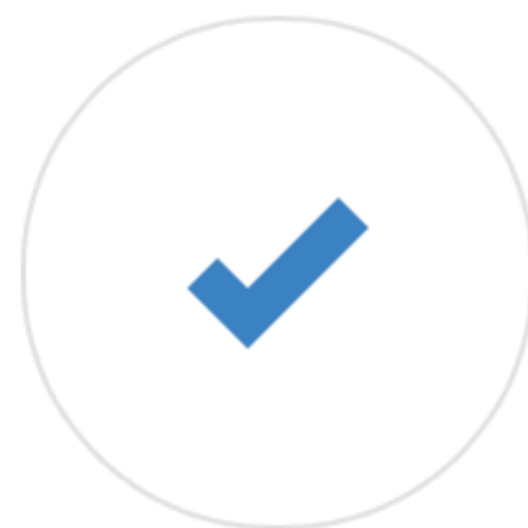
Add content to new pages site|

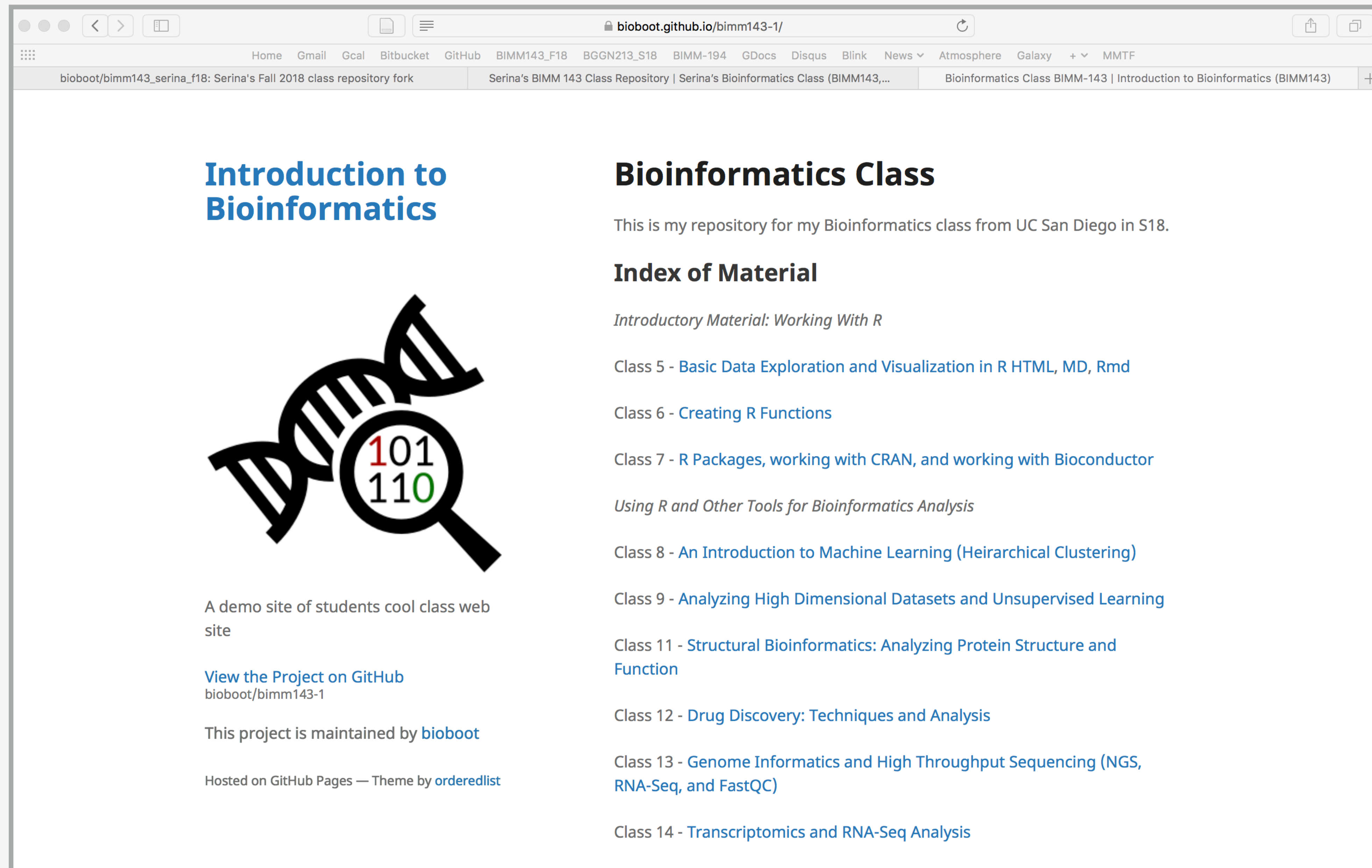
Add an optional extended description...

☒ Commit directly to the `master` branch.

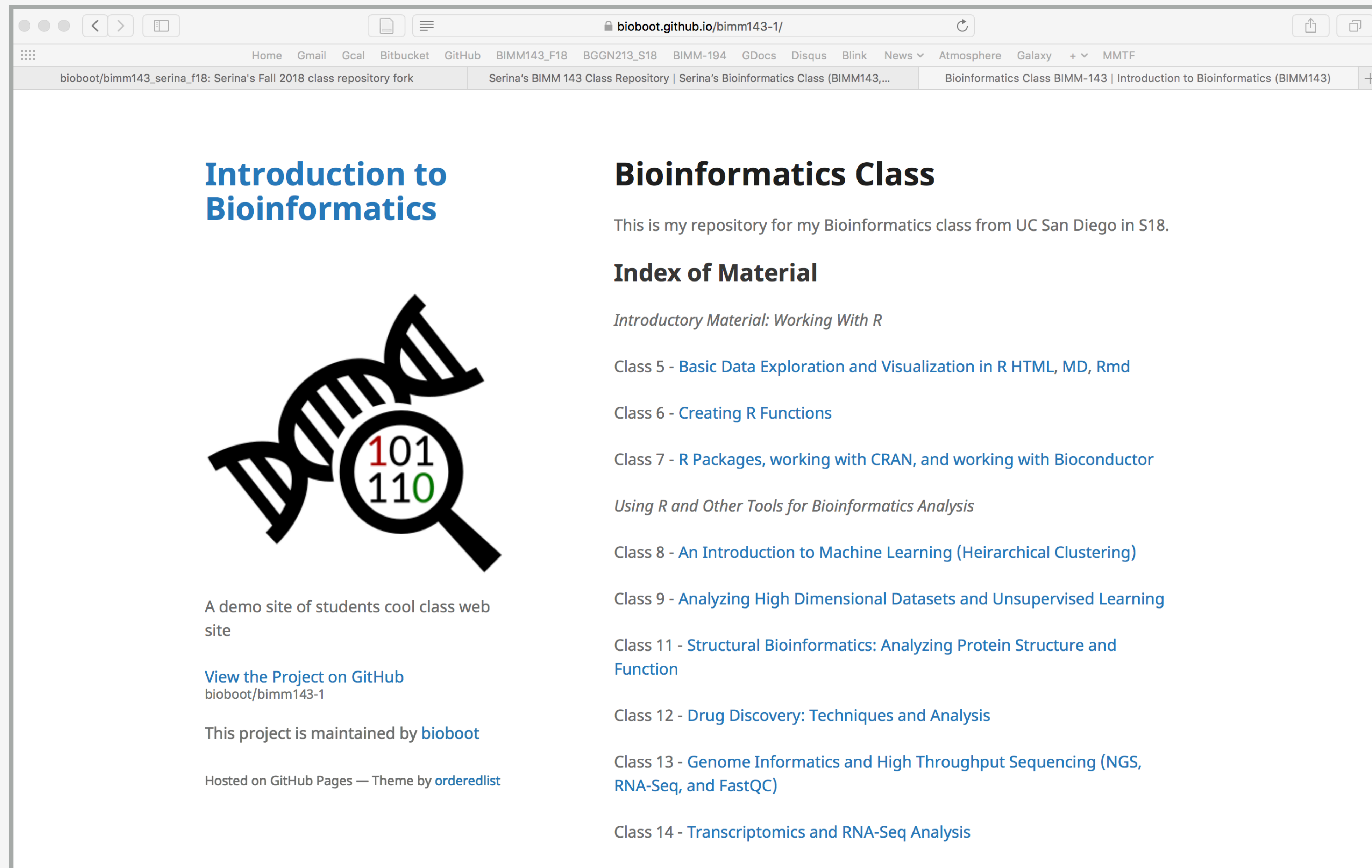
...and you're done!

Fire up a browser and go to **`http://username.github.io/repository`**.





Here I: (1) forked Serina's Repo, (2) Chose the "minimal" theme, (3) Edited _config.yml (adding logo and title)



Here I: **(1)** forked Serina's Repo, **(2)** Chose the "minimal" theme, **(3)** Edited _config.yml (adding logo and title)



bioboot Update _config.yml

3b72493 just now

1 contributor

4 lines (3 sloc) | 151 Bytes

Raw

Blame

History



```
1  theme: jekyll-theme-minimal
2  logo: https://bioboot.github.io/bimm143_F18/assets/img/logo.png
3  title: Serina's Bioinformatics Class (BIMM143, Fall 2018)
```


Here I: (1) Chose the “minimal” theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md

bioboot.github.io/bimm143-1/

Home Gmail Gcal Bitbucket GitHub BIMM143_F18 BGGN213_S18 BIMM-194 GDocs Disqus Blink News Atmosphere Galaxy + MMTF

bioboot/bimm143_serina_f18: Serina's Fall 2018 class repository fork Serina's BIMM 143 Class Repository | Serina's Bioinformatics Class (BIMM143,... Bioinformatics Class BIMM-143 | Introduction to Bioinformatics (BIMM143)

Introduction to Bioinformatics



A demo site of students cool class web site

[View the Project on GitHub](#)
bioboot/bimm143-1

This project is maintained by [bioboot](#)

Hosted on GitHub Pages — Theme by [orderedlist](#)

Bioinformatics Class

This is my repository for my Bioinformatics class from UC San Diego in S18.

Index of Material

Introductory Material: Working With R

Class 5 - [Basic Data Exploration and Visualization in R HTML, MD, Rmd](#)

Class 6 - [Creating R Functions](#)

Class 7 - [R Packages, working with CRAN, and working with Bioconductor](#)

Using R and Other Tools for Bioinformatics Analysis

Class 8 - [An Introduction to Machine Learning \(Heirarchical Clustering\)](#)

Class 9 - [Analyzing High Dimensional Datasets and Unsupervised Learning](#)

Class 11 - [Structural Bioinformatics: Analyzing Protein Structure and Function](#)

Class 12 - [Drug Discovery: Techniques and Analysis](#)

Class 13 - [Genome Informatics and High Throughput Sequencing \(NGS, RNA-Seq, and FastQC\)](#)

Class 14 - [Transcriptomics and RNA-Seq Analysis](#)

Here I: (1) Chose the “minimal” theme, (3) Edited `_config.yml` (adding logo and title), (4) Edited README.md

bioboot / tmp_test

Unwatch 1Star 0Fork 0

<> CodeIssues 0Pull requests 0Projects 0WikiInsightsSettings

tmp_test / README.md or cancel

<> Edit filePreview changesSpaces 2Soft wrap

```
1  # BIMM143 Classwork
2
3  This is a store of my class-work for [BIMM143 Winter 2019](https://bioboot.github.io/bimm143_W19/) at UC San Diego.
4
5  ## Content
6  - Class05: [R fundamentals](https://github.com/bioboot/tmp_test/blob/master/class05/class05.md)
7  - Class06: [R graphics]()
8  - Class07: R Functions
9  - Class08: R packages from CRAN, Bioconductor and GitHub
10 - Class09: Introduction to machine learning
11 - Class10: Some thing else
12 - Class11: (Structural Bioinformatics)(https://github.com/bioboot/tmp_test/blob/master/class11/class11.md)
13 - Class12: etc. etc.
14
```


Here I: (1) Chose the “minimal” theme, (3) Edited _config.yml (adding logo and title), (4) Edited README.md

bioboot.github.io/bimm143-1/

Home Gmail Gcal Bitbucket GitHub BIMM143_F18 BGGN213_S18 BIMM-194 GDocs Disqus Blink News Atmosphere Galaxy + MMTF

bioboot/bimm143_serina_f18: Serina's Fall 2018 class repository fork Serina's BIMM 143 Class Repository | Serina's Bioinformatics Class (BIMM143,... Bioinformatics Class BIMM-143 | Introduction to Bioinformatics (BIMM143)

Introduction to Bioinformatics



A demo site of students cool class web site

[View the Project on GitHub](#)
bioboot/bimm143-1

This project is maintained by [bioboot](#)

Hosted on GitHub Pages — Theme by [orderedlist](#)

Bioinformatics Class

This is my repository for my Bioinformatics class from UC San Diego in S18.

Index of Material

Introductory Material: Working With R

Class 5 - [Basic Data Exploration and Visualization in R HTML, MD, Rmd](#)

Class 6 - [Creating R Functions](#)

Class 7 - [R Packages, working with CRAN, and working with Bioconductor](#)

Using R and Other Tools for Bioinformatics Analysis

Class 8 - [An Introduction to Machine Learning \(Heirarchical Clustering\)](#)

Class 9 - [Analyzing High Dimensional Datasets and Unsupervised Learning](#)

Class 11 - [Structural Bioinformatics: Analyzing Protein Structure and Function](#)

Class 12 - [Drug Discovery: Techniques and Analysis](#)

Class 13 - [Genome Informatics and High Throughput Sequencing \(NGS, RNA-Seq, and FastQC\)](#)

Class 14 - [Transcriptomics and RNA-Seq Analysis](#)

Here I: (1) Chose the “minimal” theme, (3) Edited `_config.yml` (adding logo and title), (4) Edited README.md

Thank you very much!

Please do fill out your CAPs evaluation ([Link!](#)) if you get a chance.
It is important to the courses we offer in the future and how we teach them!

Thank you very much!

Please do fill out your CAPs evaluation ([Link!](#)) if you get a chance.
It is important to the courses we offer in the future and how we teach them!

Post to GradeScope your [GitHub Pages](#) portfolio [URL](#) to GradeScope!

Going Further With DataCamp

https://bioboot.github.io/bimm143_F19/class-material/datacamp_extras.pdf

The screenshot shows a web browser window with the URL `www.datacamp.com/tracks/custom-bioinformatics-extension`. The browser's address bar and tabs are visible at the top. The DataCamp logo is in the top left corner. A search bar contains the text "What would you like to learn t". Navigation links include "Learn", "Practice", "Projects", "Pricing", "My Classes", and a user profile icon showing "3,050 XP".

The main content area features the heading "CUSTOM TRACK" in orange, followed by "Bioinformatics Extension" in large white text. Below this, a paragraph reads: "Barry's suggested DataCamp courses for delving deeper into R, Git, the Shell, data visualization, and Bioinformatics data analysis generally. Please do reach out if you encounter problems. Happy DataCamping!". A yellow "Enroll" button is positioned below the text.

To the right of the text is a square icon with a black background, featuring a white DNA double helix and a magnifying glass with the binary code "101110" inside it.

At the bottom of the main section, a horizontal list of tags includes: "R Language", "Shell", "Git", "Spreadsheets", "111 hours", and "26 Courses".

Below the main section, two white boxes are visible. The first box has a blue terminal icon at the top, followed by the title "Introduction to Shell for Data Science". Below the title, it says: "The Unix command line helps users combine existing programs in new ways, automate repetitive tasks, and run programs ...". The second box is titled "INSTRUCTORS" and lists two people: "Greg Wilson" (Co-founder of Software Carpentry) and "Jonathan Cornelissen".

Announcement!

Bonus:

Bioinformatics & Genomics in Industry

Live Stream Video

Enjoy a set of short open ended guest lectures from leading genomic scientists at **Illumina Inc.**, **Synthetic Genomics Inc.**, and the **La Jolla Institute for Allergy and Immunology**. Feel free to contact these scientists for networking and to have your questions about industry careers in Bioinformatics and Genomics answered.

The background of the slide is an abstract composition of numerous overlapping, semi-transparent spheres in shades of purple, pink, orange, and blue. These spheres are arranged in a way that creates a sense of depth and movement, resembling a molecular structure or a cluster of cells. The overall color palette is vibrant and modern.

GitHub Copilot

Hands-on Lab Session

Barry Grant
UC San Diego

<http://thegrantlab.org>

GitHub Copilot

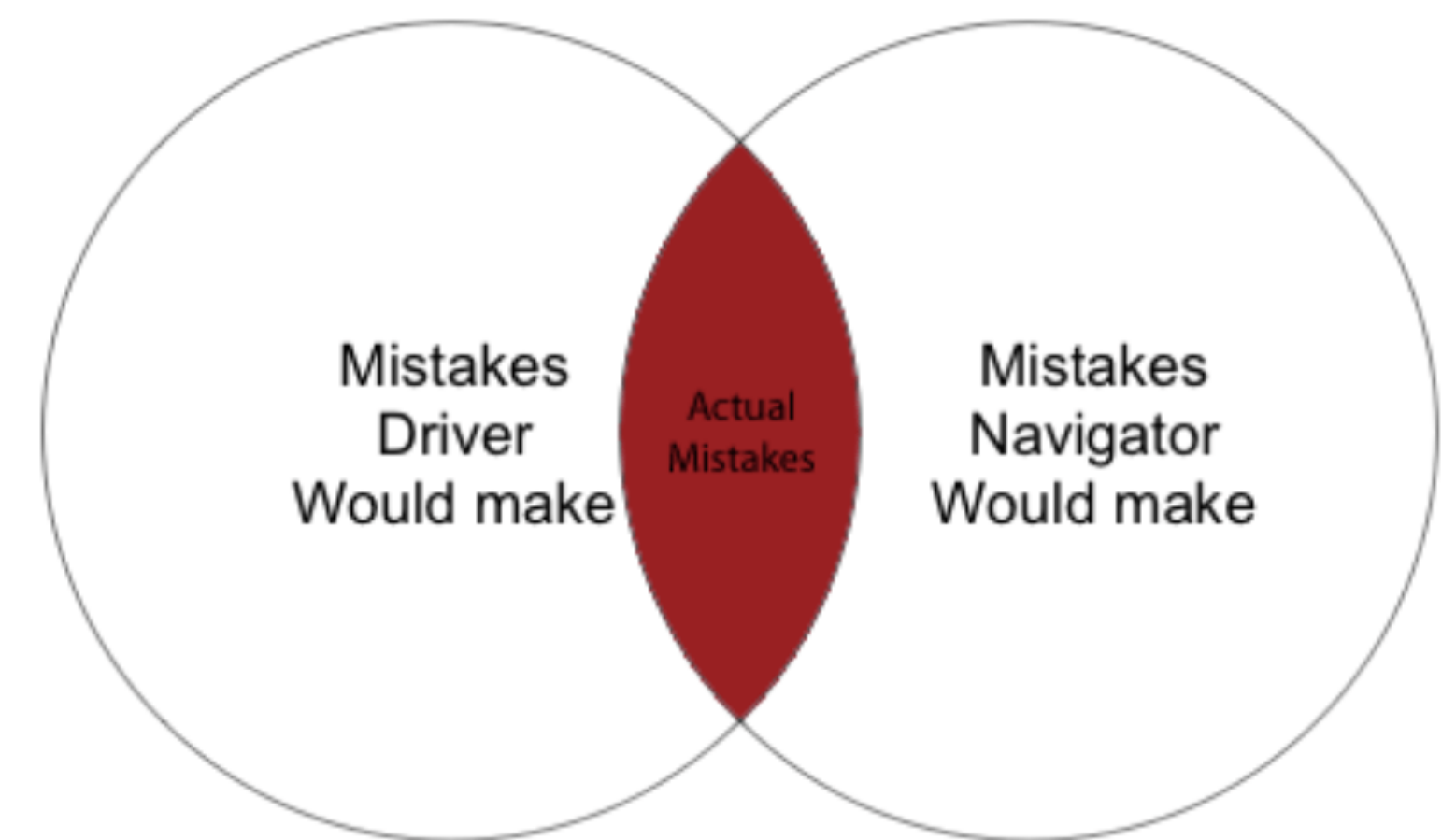
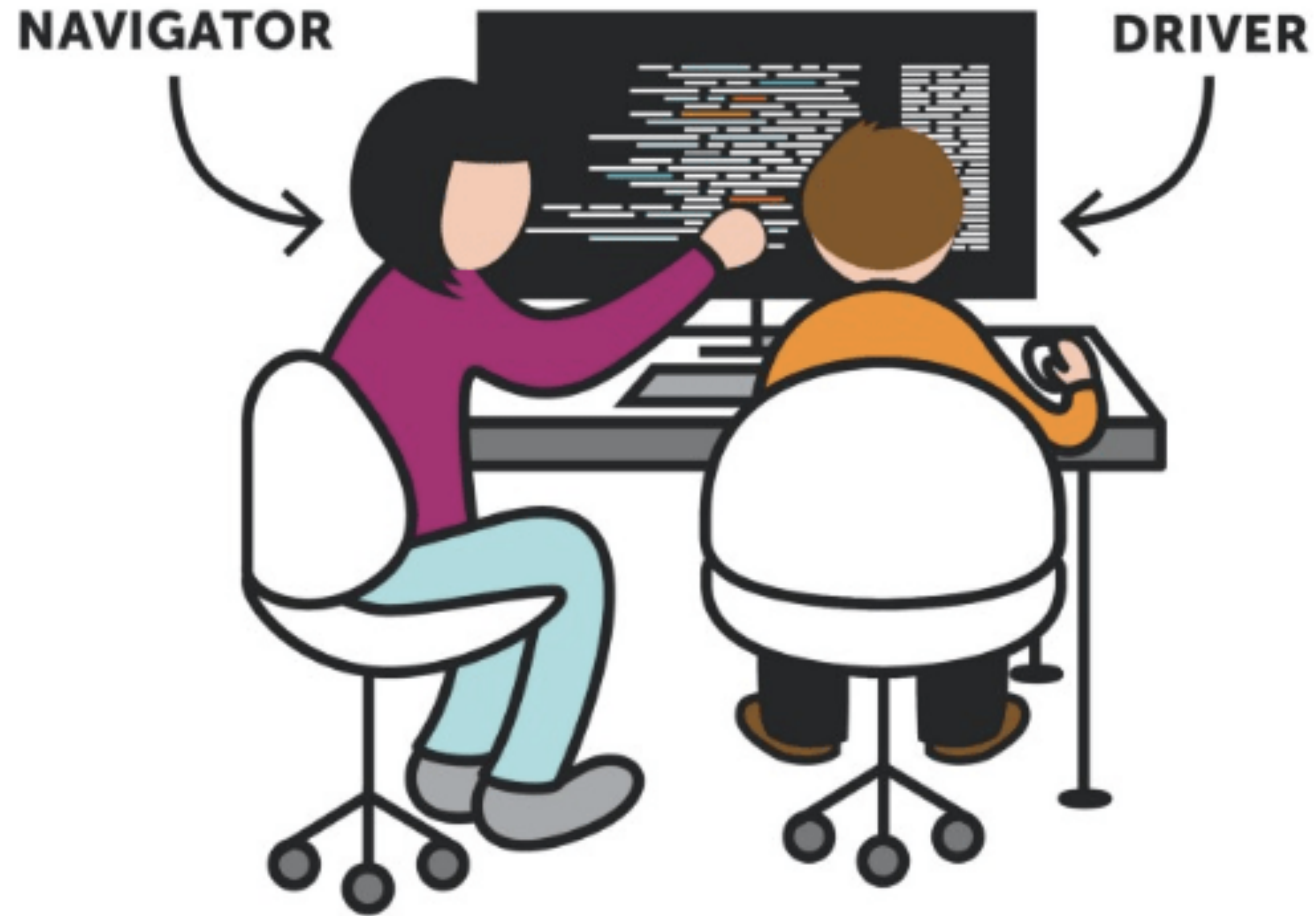


GitHub Copilot



Your AI “pair programmer”

PAIR PROGRAMMING



GitHub Copilot



AI powered add-on that aims to give helpful suggestions when writing code or documentation

AI pair programmer

GitHub Copilot is an AI pair programmer that offers autocomplete-style suggestions and real-time hints for the code you are writing by providing suggestions as “ghost text” based on the context of the surrounding code

What can GitHub Copilot do?

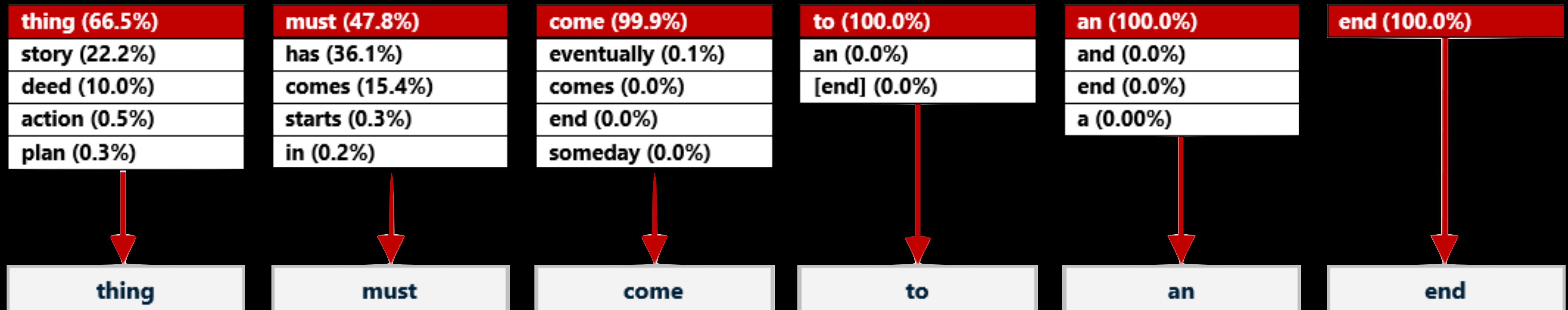
- It provides autocomplete suggestions while you code.
- You can get suggestions by typing code or describing it in natural language (i.e. code comments or markdown text).
- Copilot analyzes your file (and related files), offering suggestions in your editor.
- It uses [OpenAI Codex](#) to derive context from your code and comments, and then suggests new lines or entire functions.



Generative AI Models

Generative AI

- For text generation, Generative AI just wants to predict the next word/token/string!
- I might ask ChatGPT: > “Complete the sentence every good...”



OpenAI

Generative AI

GPT-3

Prompt:

Write a tagline for an ice cream shop.

Response:

We serve up smiles with every scoop!

Codex

Prompt:

```
Table customers, columns =  
[CustomerId, FirstName,  
LastName, Company, Address,  
City, State, Country,  
PostalCode]
```

Create a SQL query for all customers in Texas named Jane
query =

Response:

```
SELECT *  
FROM customers  
WHERE State = 'TX' AND FirstName  
= 'Jane'
```

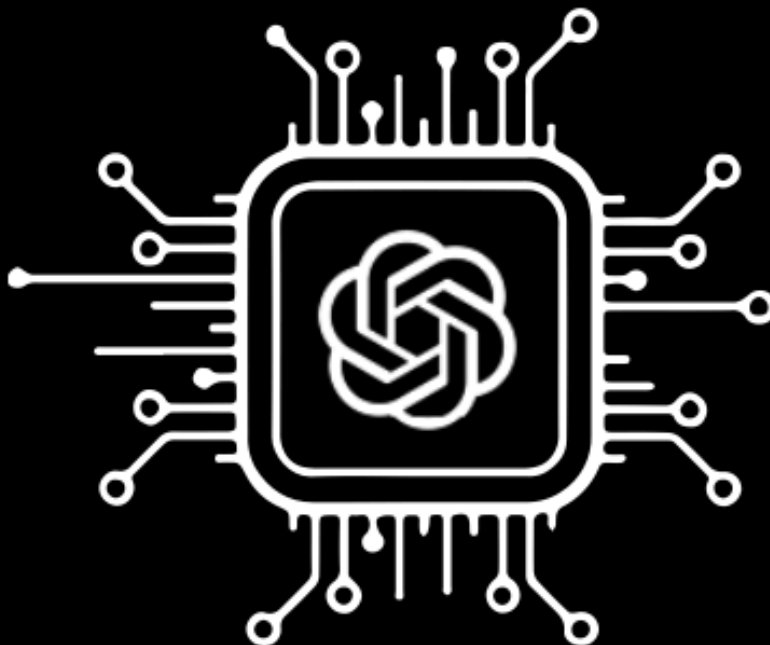
DALL·E

Prompt: A white Siamese cat

Response:



Training Data



Prompt

Model

Completion





Generative AI can:



Generate text, images and code

Different models are trained on different corpuses, depending on the application.



Generate “humanlike” output

What is a likely continuation of the prompt, given the training data?



Extract information

The continuation is likely to be similar to text frequently represented in the training data.



Create novel content

Text, images and code not contained in its training set. Translations. “Creative” works.



Generative AI is not:



Intelligent

It's just a predictive system, designed to give a likely continuation of the prompt given the training data.



Deterministic

Run the same prompt. Get back a different response (probably).



Trustworthy

It can “hallucinate” facts and confidently assert them to be true.

TODO



Generative AI does not:



Learn

The model is fixed at the time of its training.



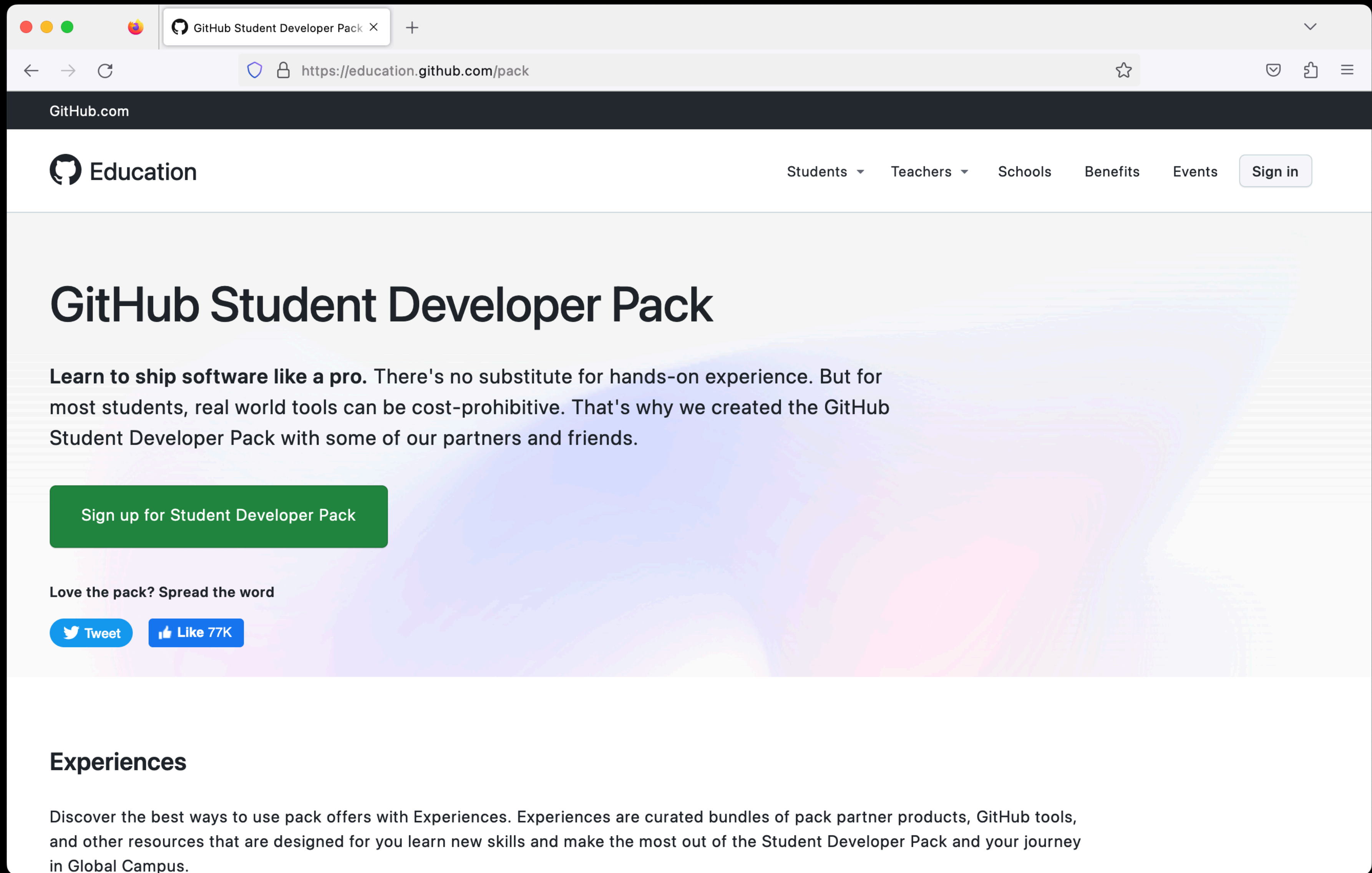
Contain all of the information of its training set

Think: a [blurry jpeg](#) of its training data.



Include verbatim copies of its training data

But it can generate stuff that looks like it.



GitHub Student Developer Pack

+

← → ↺


https://education.github.com/pack

☆

🔒

📄

☰





Tools


Security and Monitoring


Explore industry standard tools that keep your code safe, your permissions secure and your infrastructure monitored so that you can confidently scale your brilliant ideas.

Offers in this bundle

 GitHub


 Blackfire

 Honeybadger

 AstraSecurity

+3

[Learn more about Security and Monitoring >](#)





Primer


Primer: Copilot

What is Copilot and how do you get started using it? Use this Experience as a Primer on Copilot. We share tutorials, guides and templates to learn how to use AI in your development process.


Offers in this bundle

 Github Copilot

 GitHub Codespaces

 Visual Studio Code

[Learn more about Primer: Copilot >](#)





Tools


Career Readiness


Whether you're approaching graduation or just want to get a head start, it's never too early to prepare for your future in the industry.

Offers in this bundle

 Educative


 InterviewCake

 DailyBot

 GitHub Pages

+2


[Learn more about Career Readiness >](#)



Tools

Intro to Web Dev


Everything you need to build your next website.



Tools

Hackathon in the Cloud

Learn how to organize, promote, and communicate



Tools

Aspiring Creatives

Working on a creative project? Develop your design

Request a discount - GitHub EduXVirtual Event Kit - GitHub EducaXIntro to GitHub Flow - GitHub EdXData Science & Machine LearninXPrimer: Copilot - GitHub EducatiX+

https://education.github.com/discount_requests/application

GitHub.com

Education

StudentsTeachersSchoolsBenefitsEventsGlobal Campus

Home / Benefits / Benefits application

Get your GitHub benefits

Learn and teach using real-world developer tools

Select your academic status *

☐

Teacher

☒

Student

Benefits for Students

STUDENT

FREE GitHub Pro while you are a student

STUDENT

Valuable GitHub Student Developer Pack partner offers

STUDENT

GitHub Campus Expert training for qualified applicants

To qualify for student benefits, you must:

• Have a GitHub account.

• Be at least 13 years old.

Quick Start Guide

<https://docs.github.com/en/copilot/quickstart>

Additional Learning Resources

- **Video series from GitHub:**
 - <https://learn.microsoft.com/en-us/shows/introduction-to-github-copilot/>
- **Blog post series:**
 - <https://github.blog/2023-06-20-how-to-write-better-prompts-for-github-copilot/>
- **RStudio Docs:**
 - <https://docs.posit.co/ide/user/ide/guide/tools/copilot.html>



The End!

<http://thegrantlab.org>



Thank You!

<http://thegrantlab.org>