



BIMM 143

Course Summary & GitHub Portfolio

Class 19

Barry Grant
UC San Diego

<http://thegrantlab.org/bimm143>

Today's Menu

- Summary of major learning goals,
- Course discussion and feedback (https://board.net/p/bimm143_f23)
- 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://board.net/p/bimm143_f23)
- 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://board.net/p/bimm143_f23)
- 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_guidlines.pdf

The screenshot shows a web browser window with the URL https://bioboot.github.io/bimm143_S18/class-material/BIMM143_exam_guidlines.pdf in the address bar. The page content is a PDF document titled "BIMM-143: INTRODUCTION TO BIOINFORMATICS" with the URL <http://thegrantlab.org/bimm143>. Below the title, there is a section titled "Preparing for the Final Exam". The text discusses the final exam format, point distribution, and the exclusion of specific lecture material.

BIMM-143: INTRODUCTION TO BIOINFORMATICS
<http://thegrantlab.org/bimm143>

Preparing for the Final Exam

Overview: The final exam for BIMM-143 will be an open-book, open-notes 150-minute test consisting of 35 questions.

Questions will be predominantly short answer (typically worth 2 points) with a number of more involved longer answer questions (typically worth 5 points).

The number of points for each question is indicated at the beginning of each question. There are 80 total points on offer.

There will be no questions covering the material from lecture 10 (the git version control system). However, major points from all other lecture material are examinable

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 Version: https://board.net/p/bimm143_f23

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

Thank you very much!

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

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 \(NGS, RNA-Seq, and FastQC\)](#)

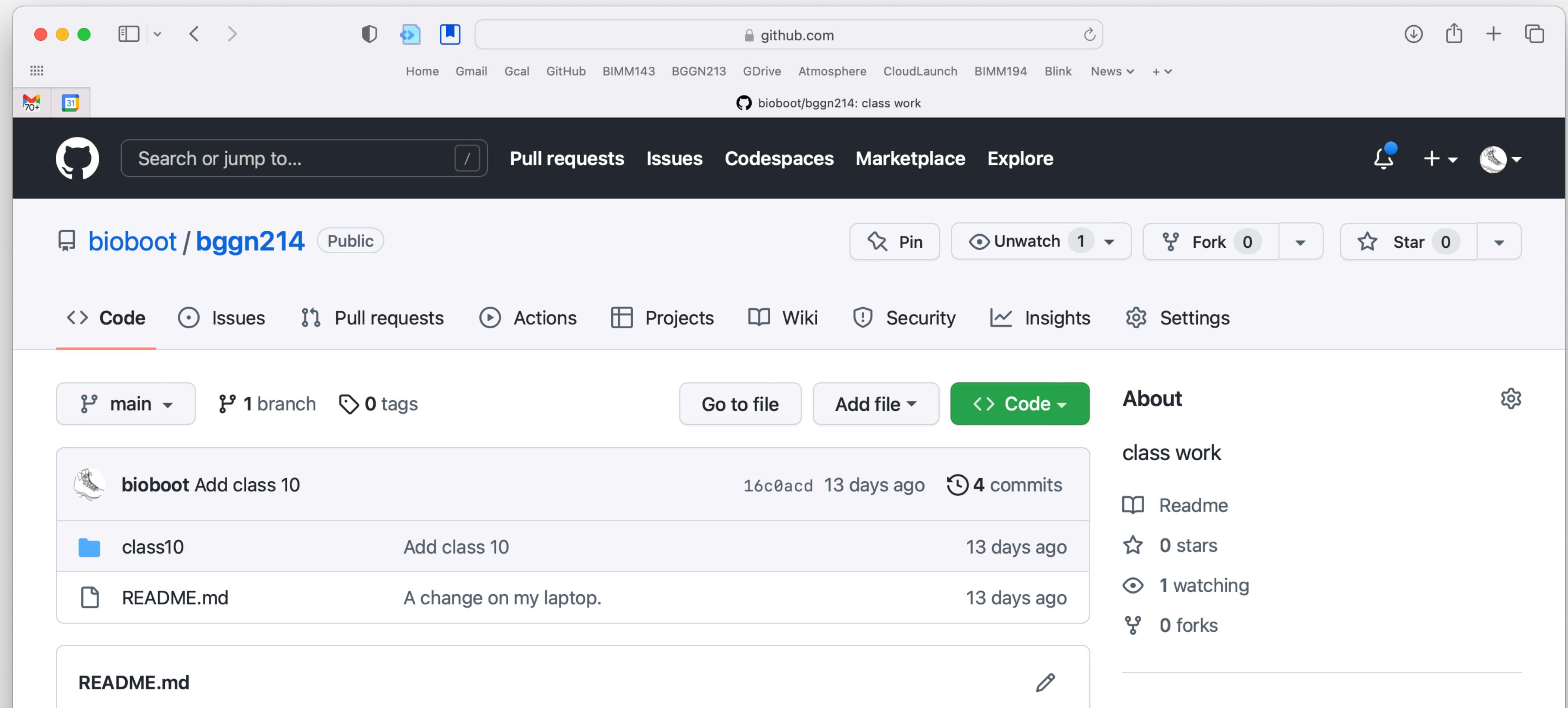
Class 14 - [Transcriptomics and RNA-Seq Analysis](#)

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

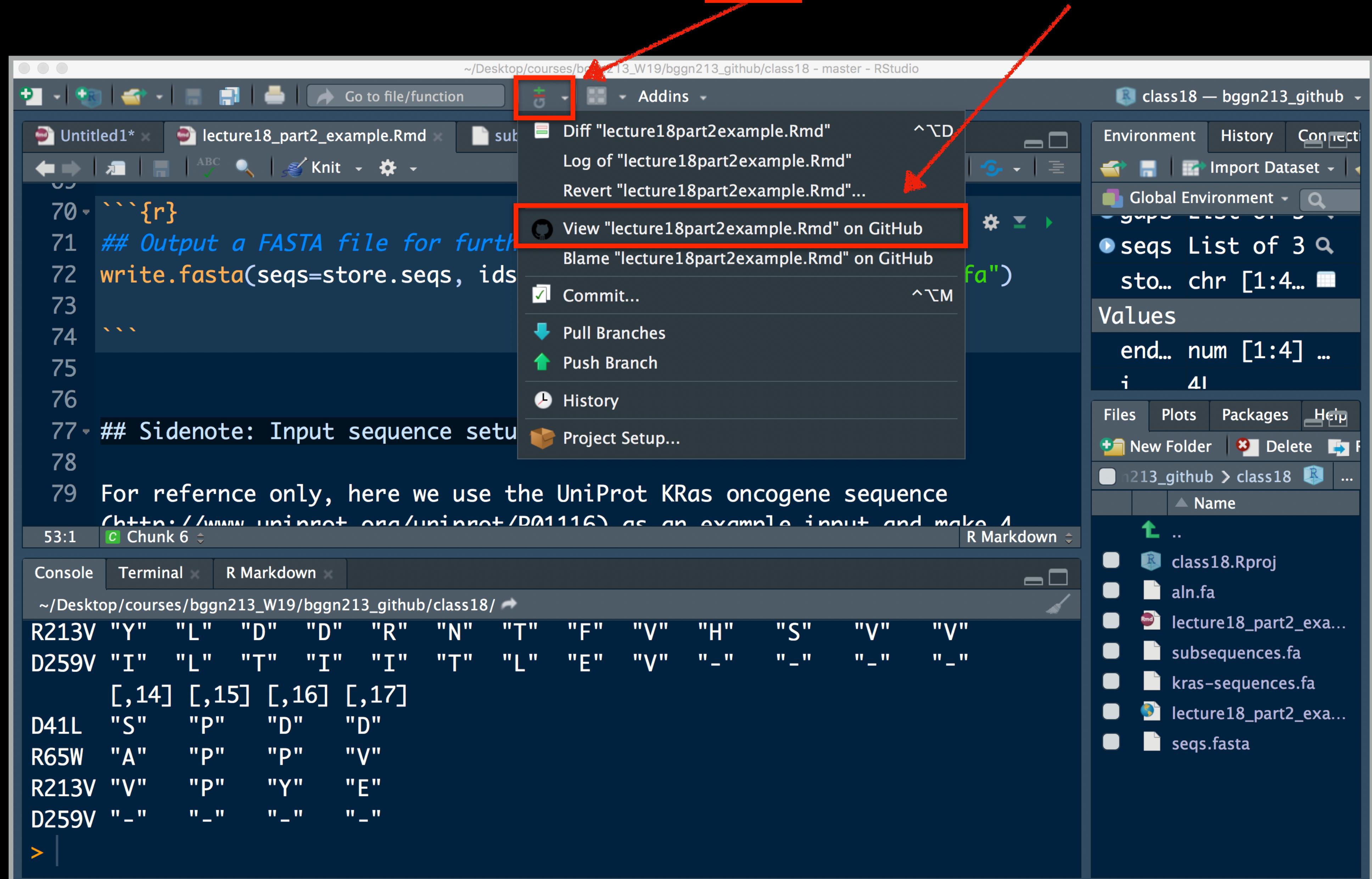
Side-note:

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



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

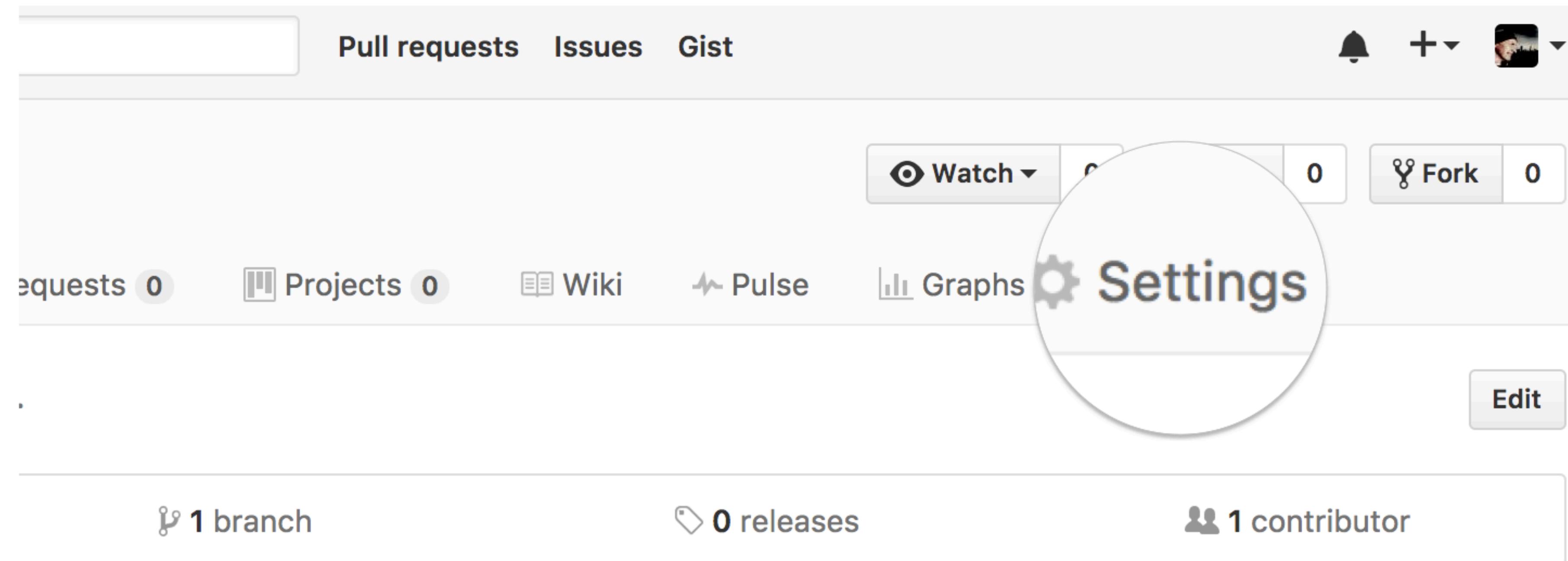


1

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”

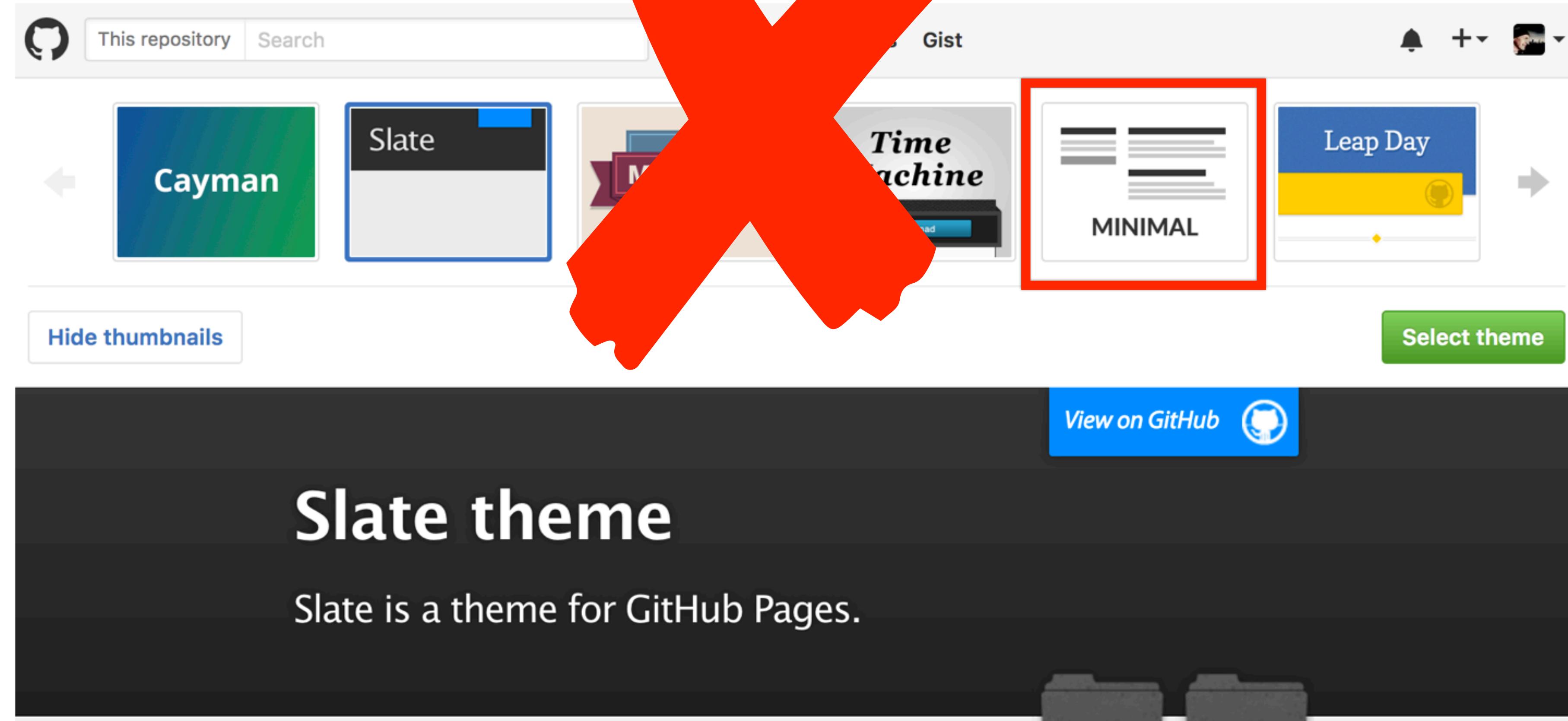
The screenshot shows the GitHub repository settings page for a repository named "luislafuente/learn-react". The left sidebar lists various settings categories: General, Access, Collaborators, Moderation options, Code and automation, Branches, Tags, Actions, Webhooks, Environments, Codespaces, and Pages. The "Pages" option is highlighted with a red box. The main content area is titled "GitHub Pages" and contains the following information:

- GitHub Pages**: A brief description stating "GitHub Pages is designed to host your personal, organization, or project pages from a GitHub repository."
- Build and deployment**: A section under "Code and automation" with a dropdown menu labeled "Deploy from a branch".
- Source**: A section where the "Branch" dropdown is set to "main" (which is highlighted with a red box). There is also a "Save" button.
- Branch**: A note stating "GitHub Pages is currently disabled. Select a source below to enable GitHub Pages for this repository." with a "Learn more" link.
- Private publishing**: A callout box with an icon showing a document and a lock, stating "Publish privately to people with read access to this repository". It includes links for "Try risk-free for 30 days" and "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.



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 repository interface for 'jldec / new-pages-site'. The top navigation bar includes the repository name, a 'Watch' button (0), a 'Star' button (0), and a 'Fork' button (0). Below the navigation bar are links for 'Issues 0', 'Pull requests 0', 'Projects 0', 'Wiki', 'Pulse', 'Graphs', and 'Settings'. The main content area shows the 'new-pages-site / README.md' file. The file content is as follows:

```
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 \[documentation\]\(https://help.github.com/categories/gi\)  
\(\) and we'll help you sort it out.  
38
```



Commit changes

Add content to new pages site

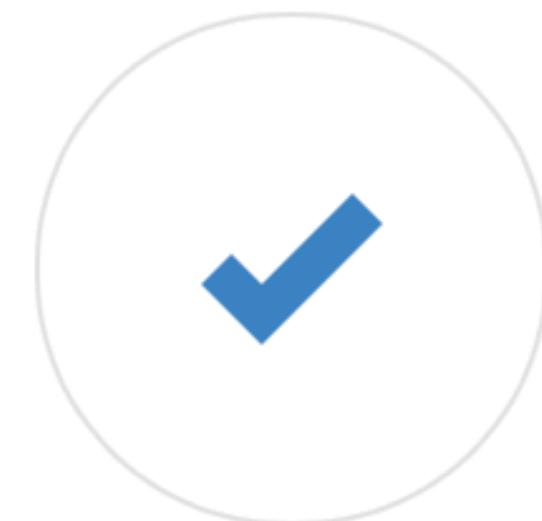
Add an optional extended description...

Commit directly to the `master` branch.

6

...and you're done!

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



The screenshot shows a web browser window with the URL bioboot.github.io/bimm143-1/ in the address bar. The page content is as follows:

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) forked Serina's Repo, (2) Chose the “minimal” theme, (3) Edited `_config.yml` (adding logo and title)

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) forked Serina's Repo, (2) Chose the “minimal” theme, (3) Edited `_config.yml` (adding logo and title)

bioboot / bimm143_serina_f18
forked from serinahuang/bimm143

Unwatch 1 Star 0 Fork 1

Code Pull requests 0 Projects 0 Wiki Insights Settings

Branch: master bimm143_serina_f18 / _config.yml Find file Copy path

bioboot Update _config.yml 3b72493 just now

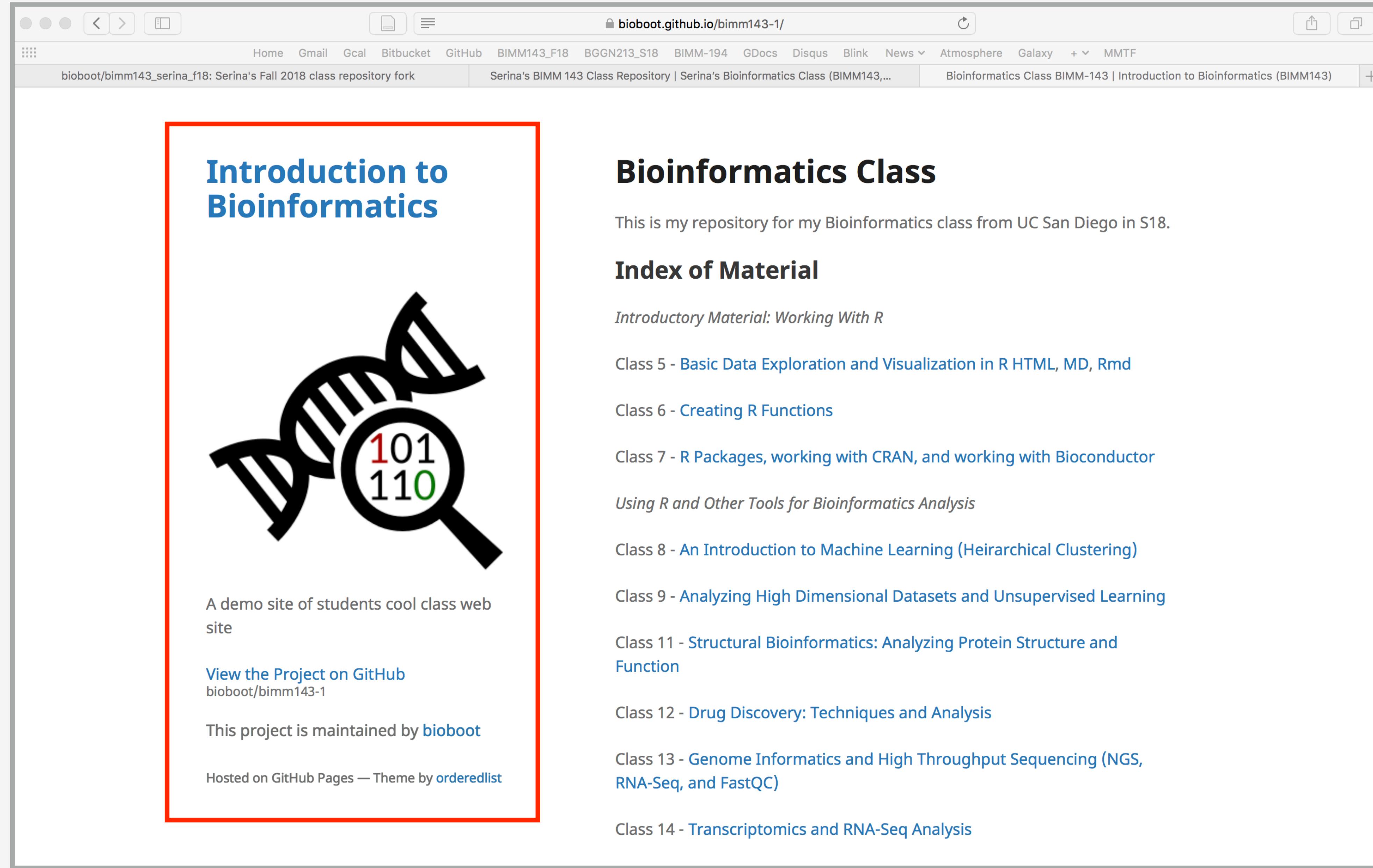
1 contributor

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

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



The screenshot shows a web browser window with the URL `bioboot.github.io/bimm143-1/` in the address bar. The page content is as follows:

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](#)

The screenshot shows a GitHub repository page for 'bioboot / tmp_test'. The repository has 1 unwatched star and 0 forks. The 'Code' tab is selected. The README.md file contains the following content:

```
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

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 for the DataCamp website (www.datacamp.com/tracks/custom-bioinformatics-extension). The page title is "CUSTOM TRACK Bioinformatics Extension". The main content area describes Barry's suggested DataCamp courses for delving deeper into R, Git, the Shell, data visualization, and Bioinformatics data analysis generally. It includes a yellow "Enroll" button, a DNA helix icon with a magnifying glass over binary code (101 110), and a summary of 111 hours and 26 Courses across R Language, Shell, Git, and Spreadsheets.

CUSTOM TRACK
Bioinformatics Extension

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!

Enroll

R Language | Shell | Git | Spreadsheets | 111 hours | 26 Courses

Introduction to Shell for Data Science

The Unix command line helps users combine existing programs in new ways, automate repetitive tasks, and run programs ...

INSTRUCTORS

 **Greg Wilson**
Co-founder of Software Carpentry

 **Jonathan Cornelissen**

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.

GitHub Copilot Hands-on Lab Session

Barry Grant
UC San Diego

<http://thegrantlab.org>

GitHub Copilot

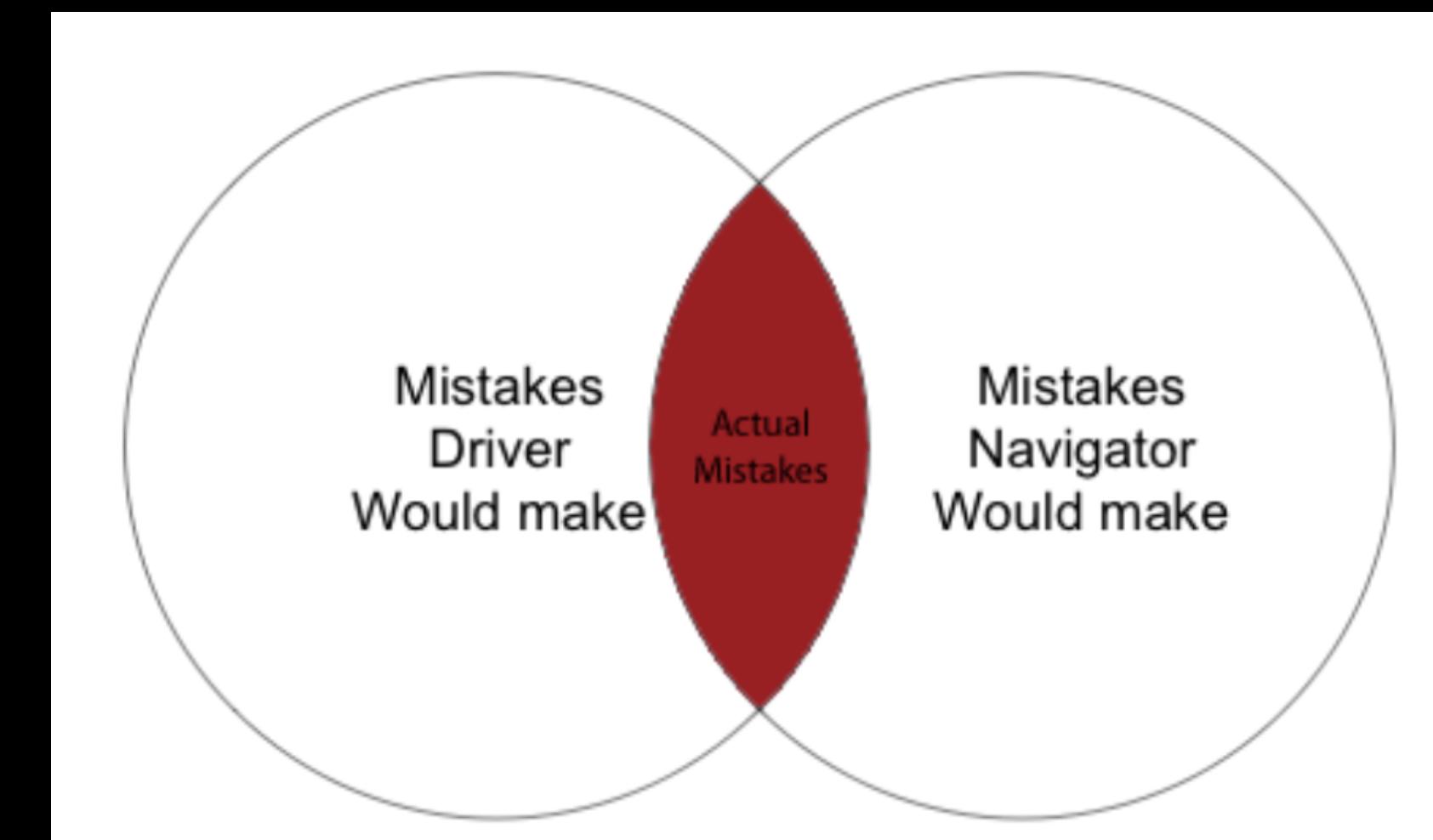
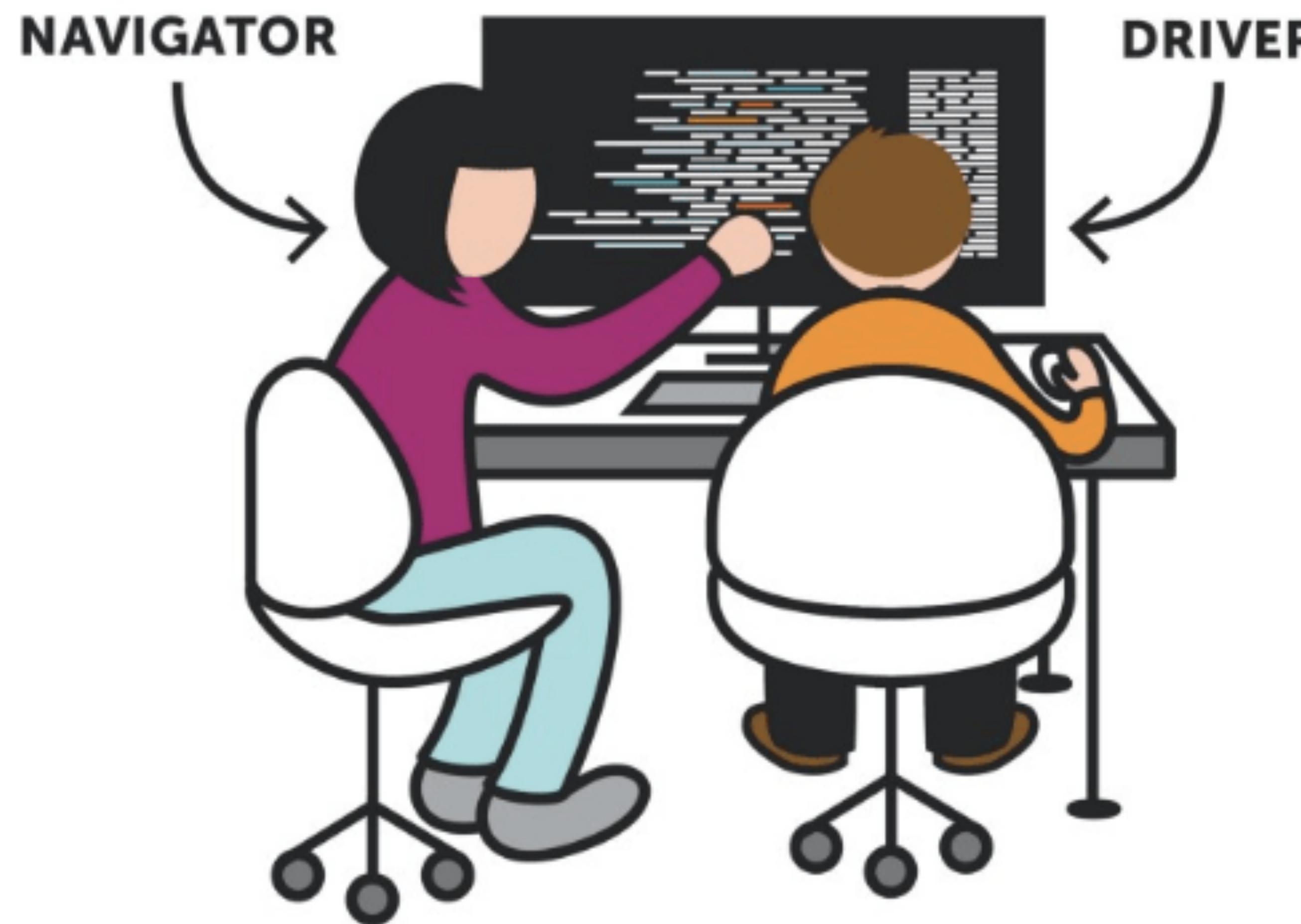


GitHub Copilot



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

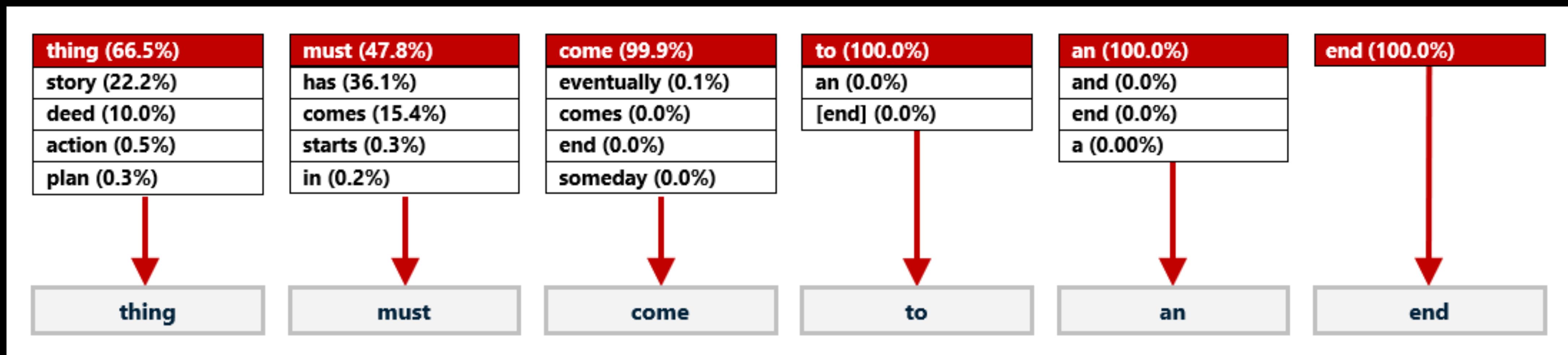
GitHub Copilot is an AI coding partner

- It provides autocomplete suggestions while you code.
- Get suggestions by typing code or describing it in natural language.
- Copilot analyzes your file (and related files), offering suggestions in your text editor.
- It uses OpenAI Codex to help derive context from written code and comments, and then suggests new lines or entire functions.

- 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>
 -

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.”



What is Copilot?

- 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
 - GitHub Copilot docs
- <https://colorado.posit.co/rsc/rstudio-copilot/#/TitleSlide>
-



Generative AI Models



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'
```

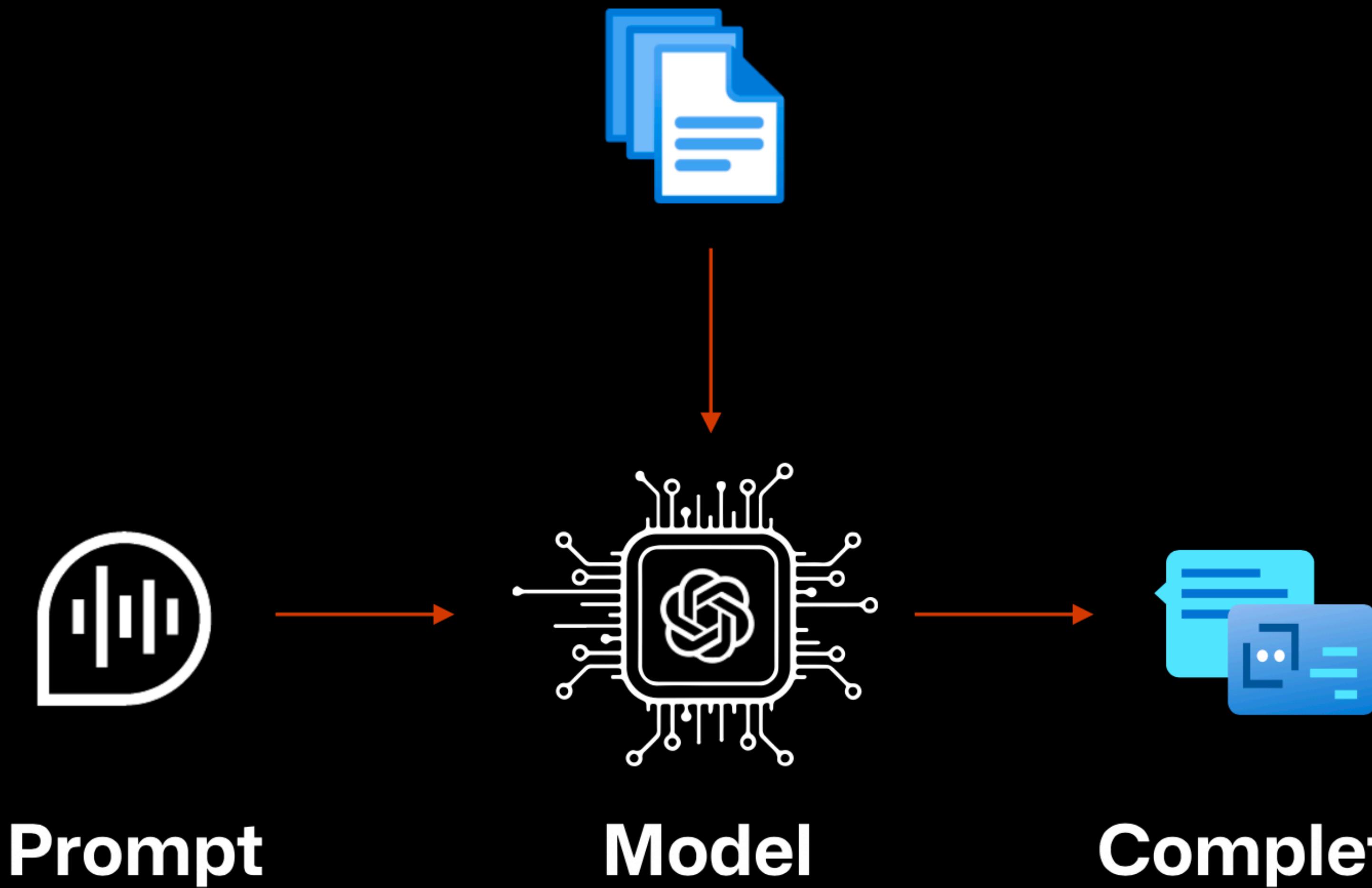
DALL·E

Prompt: A white Siamese cat

Response:

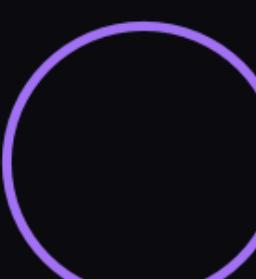


Training Data





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.

A screenshot of a web browser window showing the GitHub Student Developer Pack landing page. The page has a dark header with the GitHub logo and 'Education' link. Below the header is a large image of a person working at a desk with multiple monitors. The main title 'GitHub Student Developer Pack' is displayed prominently. A green button encourages users to 'Sign up for Student Developer Pack'. Below the button, social sharing links for Twitter and Facebook are present. A section titled 'Experiences' describes curated bundles of resources. The URL in the address bar is https://education.github.com/pack.

GitHub Student Developer Pack

GitHub.com

Education Students Teachers Schools Benefits Events Sign in

GitHub Student Developer Pack

Learn to ship software like a pro. There's no substitute for hands-on experience. But for most students, real world tools can be cost-prohibitive. That's why we created the GitHub Student Developer Pack with some of our partners and friends.

Sign up for Student Developer Pack

Love the pack? Spread the word

[Tweet](#) [Like 77K](#)

Experiences

Discover the best ways to use pack offers with Experiences. Experiences are curated bundles of pack partner products, GitHub tools, and other resources that are designed for you learn new skills and make the most out of the Student Developer Pack and your journey in Global Campus.

 GitHub Student Developer Pack X

← → C 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 

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

Tools

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 

[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 Edu X Virtual Event Kit - GitHub Edu X Intro to GitHub Flow - GitHub Edu X Data Science & Machine Learnin X Primer: Copilot - GitHub Edu X +

GitHub.com

GitHub.com

Education Students Teachers Schools Benefits Events Global 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.