

# GitHub for Academic Research

---

Mars Lee

Open Science Training



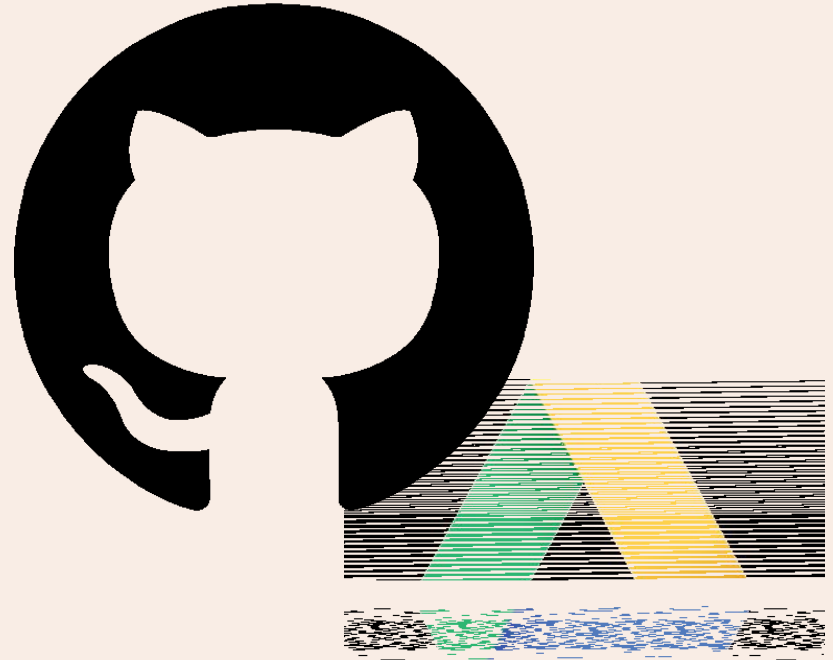
# GitHub

As a place to

- **Upload**
- **Store**
- **Comment**

on your research work.

**Advantages** over sharing over another tools, such as Google Docs.



# Student



# Advisor



# Student

# Advisor

Research Project Analysis

File Edit View Insert Format Tools Extensions Help

100%

Normal text

Arial

11

B I U A

Phenotypic Variations in Drosophila melanogaster

In this section of the notebook, we utilize Python's powerful data visualization libraries to analyze phenotypic variations observed in Drosophila melanogaster populations from different environments. Our goal is to provide a clear visual representation of the data collected during our experiment, highlighting differences in eye color and wing size across various environments.

### Libraries Used

- Pandas:** A library providing high-performance, easy-to-use data structures and data analysis tools for Python. In this context, we use Pandas to load our experimental data from a CSV file into a DataFrame. This allows us to manipulate and prepare the data efficiently before plotting.
- Matplotlib:** A comprehensive library for creating static, animated, and interactive visualizations in Python. Matplotlib is used here to create bar charts that visually compare the phenotypic traits (eye color and wing size) among fruit fly populations across different environments. It offers extensive customization options to make the visualizations informative and aesthetically pleasing.

### Plotting Code Explanation

- Loading Data:** First, we load our experimental data into a Pandas DataFrame. This step involves reading data from a CSV file that contains our observations of phenotypic variations across different environments.

Morphological examination of the fruit flies revealed distinct phenotypic variations across the three sampled populations. The source for this data can be found in .

Environment	Red-Eyed Flies (%)	Larger Wings (%)
Cafeteria	76%	49%
Chemistry Laboratory	58%	53%
Outdoor Green Spaces	63%	68%

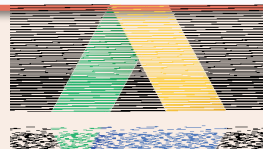
### Statistical Analysis of Phenotypic Differences

Math can be included using the `'''{math}'''` directive with a label, and



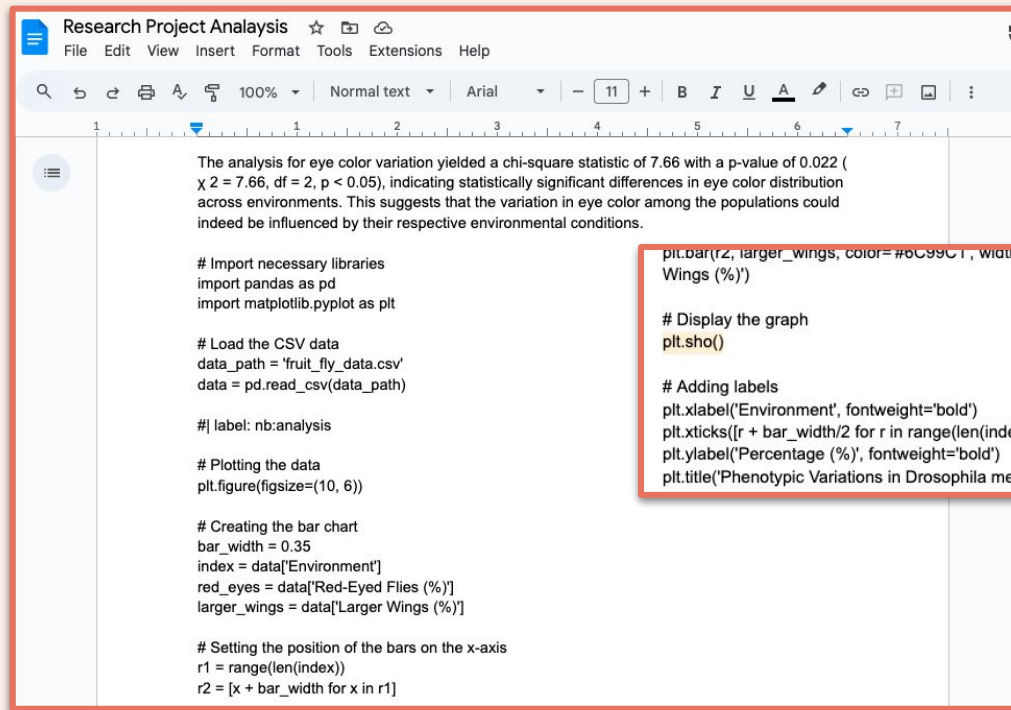
Mars Lee  
9:11 AM Today

Is this the right lab?



# Student

# Advisor



```
plt.bar(r2, larger_wings, color=#6C99C1, width=bar_width, edgecolor=gray, label=Larger  
Wings (%))
```

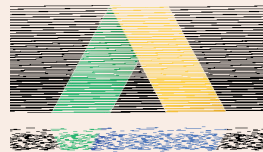
```
# Display the graph
plt.sho()
```

```
# Adding labels
plt.xlabel('Environment', fontweight='bold')
plt.xticks([r + bar_width/2 for r in range(len(index))], index)
plt.ylabel('Percentage (%)', fontweight='bold')
plt.title('Phenotypic Variations in Drosophila melanogaster Across Different Environments')
```



Mars Lee  
11:25 AM Today

It should be `plt.show()`



# Student

# Advisor

```
← Files simplified research-project / analysis.ipynb ↑ Top

Preview Code Blame Raw Copy Download Edit

In [1]:

# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt

In [2]:

# Load the CSV data
data_path = 'fruit_fly_data.csv'
data = pd.read_csv(data_path)

In [3]:

#| label: nb:analysis

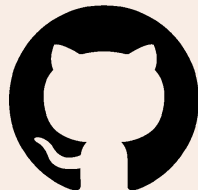
# Plotting the data
plt.figure(figsize=(10, 6))

# Creating the bar chart
bar_width = 0.35
index = data['Environment']
red_eyes = data['Red-Eyed Flies (%)']
larger_wings = data['Larger Wings (%)']

# Setting the position of the bars on the x-axis
r1 = range(len(index))
r2 = [x + bar_width for x in r1]

# Making the plot
plt.bar(r1, red_eyes, color='#D75D43', width=bar_width, edgecolor='gray',
plt.bar(r2, larger_wings, color='#6C99C1', width=bar_width, edgecolor='gra

# Adding labels
plt.xlabel('Environment', fontweight='bold')
plt.xticks([r + bar_width/2 for r in range(len(index))], index)
plt.ylabel('Percentage (%)', fontweight='bold')
plt.title('Phenotypic Variations in Drosophila melanogaster Across Differ
```



# Student

```
← Files simplified research-project / analysis.ipynb ↑ Top

Preview Code Blame Raw Copy Download Edit

In [1]:

# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt

In [2]:

# Load the CSV data
data_path = 'fruit_fly_data.csv'
data = pd.read_csv(data_path)

In [3]:

#| label: nb:analysis

# Plotting the data
plt.figure(figsize=(10, 6))

# Creating the bar chart
bar_width = 0.35
index = data['Environment']
red_eyes = data['Red-Eyed Flies (%)']
larger_wings = data['Larger Wings (%)']

# Setting the position of the bars on the x-axis
r1 = range(len(index))
r2 = [x + bar_width for x in r1]

# Making the plot
plt.bar(r1, red_eyes, color='#D75D43', width=bar_width, edgecolor='gray',
plt.bar(r2, larger_wings, color='#6C99C1', width=bar_width, edgecolor='gray')

# Adding labels
plt.xlabel('Environment', fontweight='bold')
plt.xticks([r + bar_width/2 for r in range(len(index))], index)
plt.ylabel('Percentage (%)', fontweight='bold')
plt.title('Phenotypic Variations in Drosophila melanogaster Across Differ
```

# Advisor

advisor-drosophila commented 3 days ago View reviewed changes

article.md

... @@ -95,7 +95,7 @@ The source for this data can be found in [](../fruit\_fly\_data.csv).

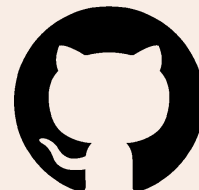
95	95	Environment	Red-Eyed Flies (%)	Larger Wings (%)	
96	96	-----	-----	-----	
97	97	Cafeteria	76%	49%	
98	-	Chemistry Laboratory	58%	53%	
98	+	Biology Laboratory	58%	53%	

advisor-drosophila 3 days ago Author ...

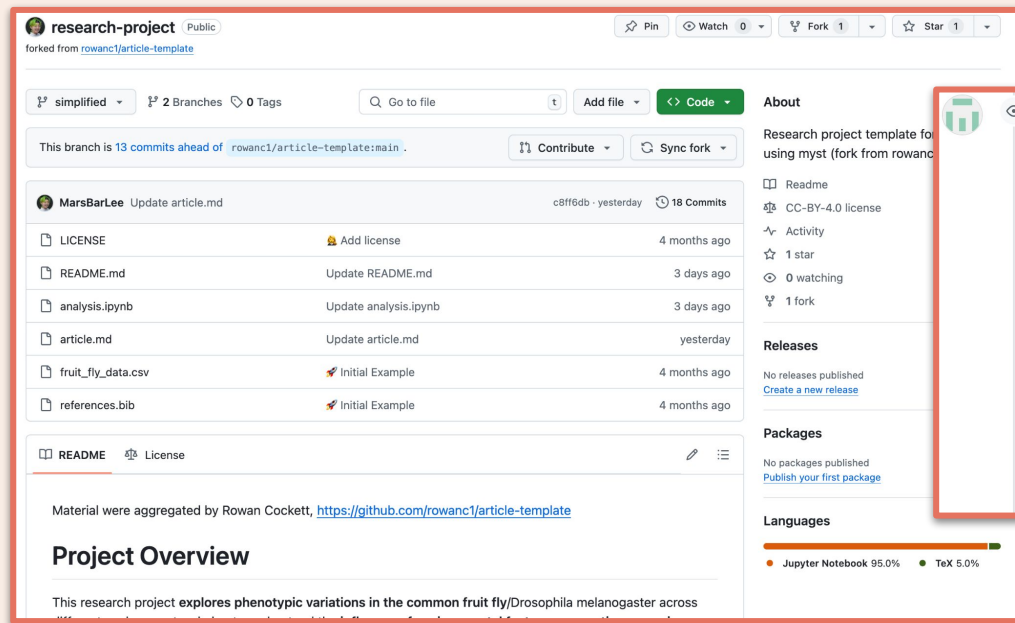
This sample was collected from the Biology Lab, not Chemistry Lab.

MarsBarLee 3 days ago Owner ...

Oh yes, I was at both labs that day and I mixed it up. Thanks for catching that!

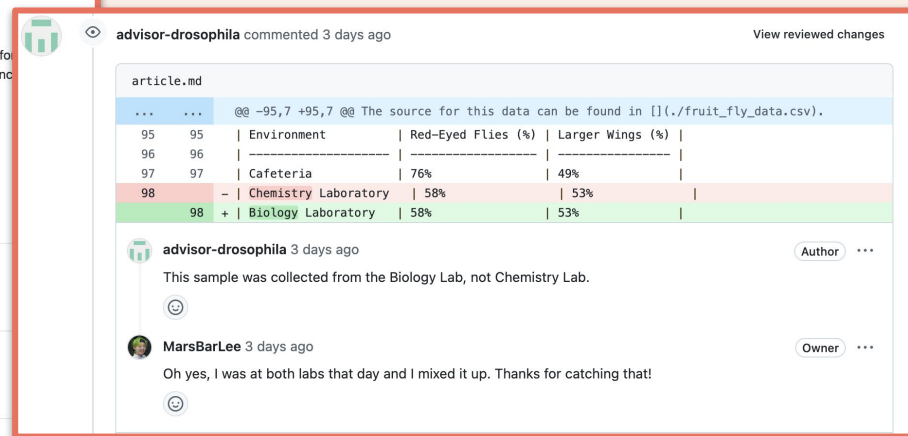


# Student



The screenshot shows a GitHub repository named 'research-project' (Public) forked from 'rowanc1/article-template'. The repository is 13 commits ahead of the upstream. The file list includes LICENSE, README.md, analysis.ipynb, article.md, fruit\_fly\_data.csv, and references.bib. The README section is visible, showing the project overview: 'Material were aggregated by Rowan Cockett, <https://github.com/rowanc1/article-template>'. The project overview states: 'This research project explores phenotypic variations in the common fruit fly/Drosophila melanogaster across'.

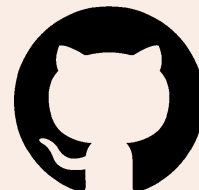
# Advisor



The screenshot shows a GitHub discussion thread. The first comment by 'advisor-drosophila' (commented 3 days ago) contains a code snippet from 'article.md' showing a table of data. The table has columns for Environment, Red-Eyed Flies (%), and Larger Wings (%). The data shows a transition from Chemistry Laboratory to Biology Laboratory.

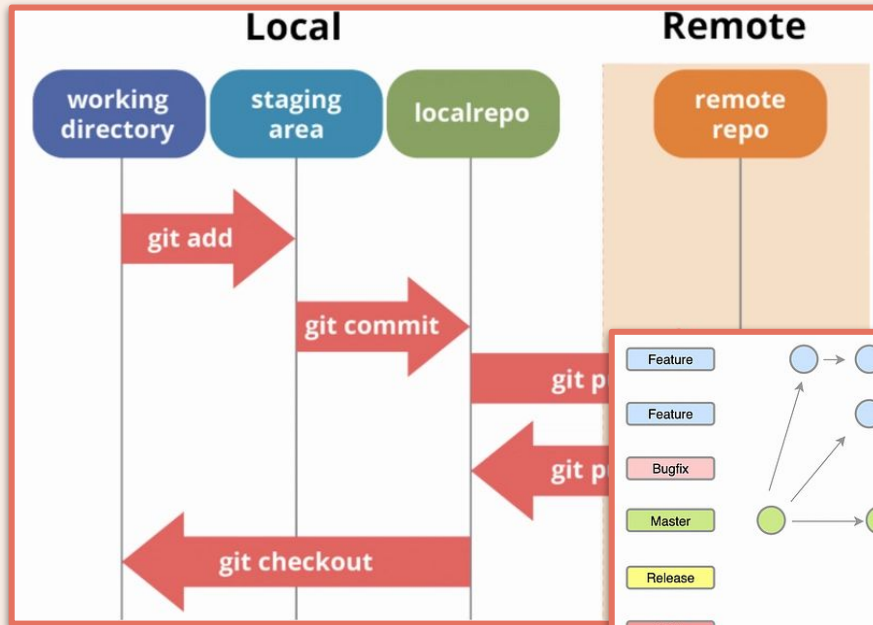
Environment	Red-Eyed Flies (%)	Larger Wings (%)
Chemistry Laboratory	58%	53%
Biology Laboratory	58%	53%

The second comment by 'MarsBarLee' (commented 3 days ago) responds: 'Oh yes, I was at both labs that day and I mixed it up. Thanks for catching that!'.

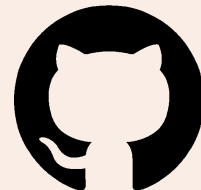
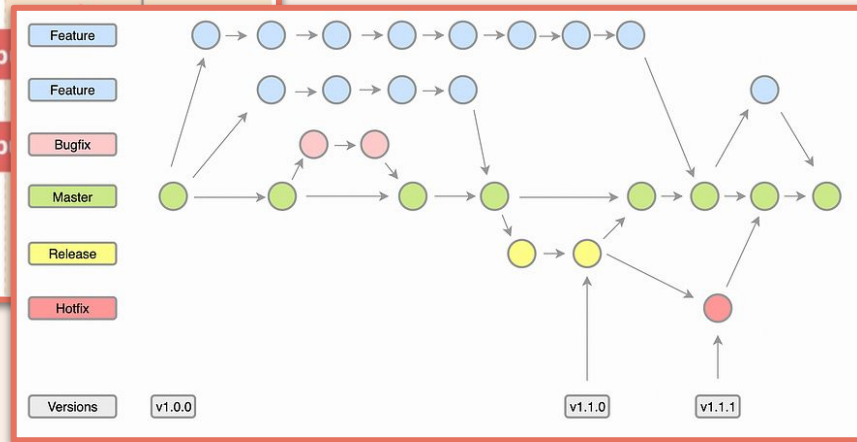




# Full Git workflow

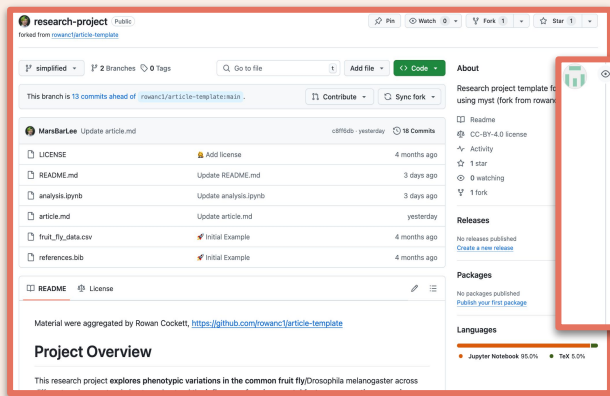


- Uses Terminal
- 6 people
- 6 branches
- Powerful but can be overwhelming

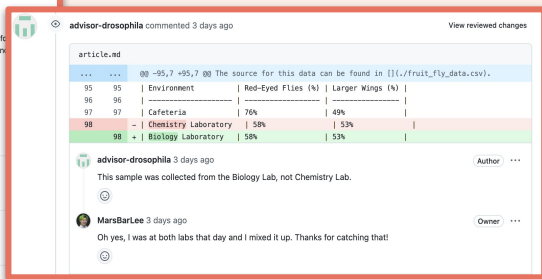


# Simplified Git workflow

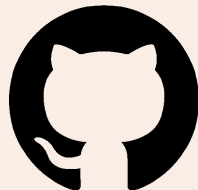
Student



Advisor



- No terminal OR CLI
- Uses github.com website
- 2 people
- 1 branch



# Student

The screenshot shows a GitHub repository page for 'research-project', which is a public fork of 'rowanc1/article-template'. The repository is managed by MarsBarLee and is 13 commits ahead of the original template. The file list includes LICENSE, README.md, analysis.ipynb, article.md, fruit\_fly\_data.csv, and references.bib. The README section is visible, stating that material was aggregated by Rowan Cockett from the original template. The project overview mentions exploring phenotypic variations in the common fruit fly (*Drosophila melanogaster*).

**research-project** Public  
forked from rowanc1/article-template

2 Branches 0 Tags

Go to file Add file <> Code

This branch is 13 commits ahead of rowanc1/article-template:main. Contribute Sync fork

**MarsBarLee** Update article.md c8ff6db · yesterday 18 Commits

File	Commit	Time
LICENSE	Add license	4 months ago
README.md	Update README.md	3 days ago
analysis.ipynb	Update analysis.ipynb	3 days ago
article.md	Update article.md	yesterday
fruit_fly_data.csv	Initial Example	4 months ago
references.bib	Initial Example	4 months ago

**README** License

Material were aggregated by Rowan Cockett, <https://github.com/rowanc1/article-template>

## Project Overview

This research project explores phenotypic variations in the common fruit fly/*Drosophila melanogaster* across

**About**

Research project template for an article, using myst (fork from rowanc1's repo)

- Readme
- CC-BY-4.0 license
- Activity
- 1 star
- 0 watching
- 1 fork

**Releases**

No releases published  
[Create a new release](#)

**Packages**

No packages published  
[Publish your first package](#)

**Languages**

Jupyter Notebook 95.0% TeX 5.0%

Link in chat:  
<https://github.com/MarsBarLee/research-project>

# Student

research-project Public  
forked from rowanc1/article-template

simplified 2 Branches 0 Tags  
Go to file Add file Cods About

This branch is 13 commits ahead of rowanc1/article-template:main. Contribute Sync fork

MarsBarLee Update article.md c8ff6db · yesterday 18 Commits

- LICENSE Add license 4 months ago
- README.md Update README.md 3 days ago
- analysis.ipynb Update analysis.ipynb 3 days ago
- article.md Update article.md yesterday
- fruit\_fly\_data.csv Initial Example 4 months ago
- references.bib Initial Example 4 months ago

README License

Material were aggregated by Rowan Cockett, <https://github.com/rowanc1/article-template>

### Project Overview

This research project explores phenotypic variations in the common fruit fly *Drosophila melanogaster* across

FORK

# Advisor

research-project Public  
forked from rowanc1/article-template

simplified 2 Branches 0 Tags  
Go to file Add file Cods About

This branch is 13 commits ahead of rowanc1/article-template:main. Contribute Sync fork

MarsBarLee Update article.md c8ff6db · yesterday 18 Commits

- LICENSE Add license 4 months ago
- README.md Update README.md 3 days ago
- analysis.ipynb Update analysis.ipynb 3 days ago
- article.md Update article.md yesterday
- fruit\_fly\_data.csv Initial Example 4 months ago
- references.bib Initial Example 4 months ago

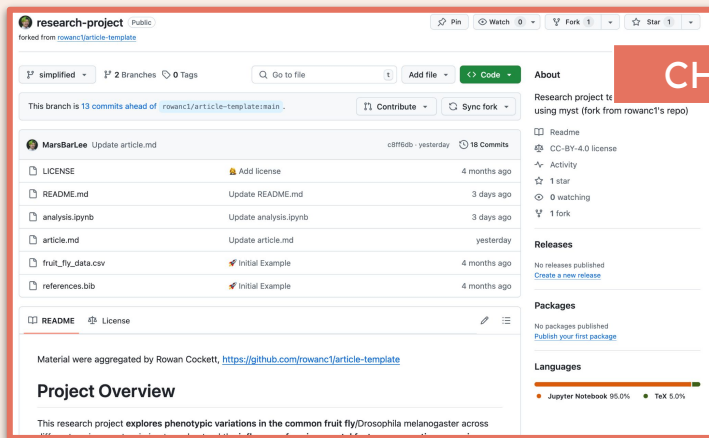
README License

Material were aggregated by Rowan Cockett, <https://github.com/rowanc1/article-template>

### Project Overview

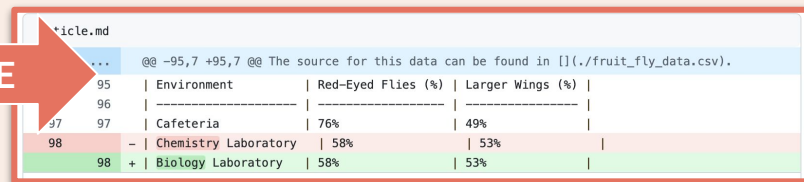
This research project explores phenotypic variations in the common fruit fly *Drosophila melanogaster* across

# Advisor



The screenshot shows a GitHub repository named 'research-project' which is a fork of 'rowanc1/article-template'. The repository is public and has 2 branches and 0 tags. The main branch is 'main' and it is 13 commits ahead of the upstream repository. The repository contains several files: LICENSE, README.md, analysis.ipynb, article.md, fruit\_fly\_data.csv, and references.bib. The 'article.md' file was updated yesterday. The repository also has a README section titled 'Project Overview' which states that the research project explores phenotypic variations in the common fruit fly *Drosophila melanogaster* across different environments.

CHANGE FILE



The screenshot shows a diff view for the file 'article.md'. The diff highlights changes to the source of data used in the project. The changes are as follows:

Line	Change	File	Environment	Red-Eyed Flies (%)	Larger Wings (%)
95	+	Environment	Red-Eyed Flies (%)	Larger Wings (%)	
96	-	Environment	Red-Eyed Flies (%)	Larger Wings (%)	
97	+	Cafeteria	76%	49%	
98	-	Chemistry Laboratory	58%	53%	
98	+	Biology Laboratory	58%	53%	

# Advisor

The screenshot shows a GitHub repository named 'research-project' which is a fork of 'rowanc1/article-template'. The repository has 2 branches and 0 tags. The file list on the left includes LICENSE, README.md, analysis.ipynb, article.md, fruit\_fly\_data.csv, and references.bib. The 'article.md' file is selected, showing its commit history and a diff view. The diff shows changes to the file's content, including a table of data. The repository description states: 'Research project template using myst (fork from rowanc1's repo)'. The 'About' section mentions 'MarsBarLee' updated 'article.md' yesterday with 18 commits. The 'Releases' section shows no releases published. The 'Packages' section shows no packages published. The 'Languages' section shows Jupyter Notebook at 95.0% and Text at 5.0%.

CHANGE FILE

The screenshot shows the content of the 'article.md' file. It includes a header section with a table of data. The table has 4 columns: 'Environment', 'Red-Eyed Flies (%)', 'Larger Wings (%)', and an unlabeled column. The table has 4 rows of data. The first row is highlighted in blue. The second row is highlighted in green. The third row is highlighted in red. The fourth row is highlighted in green.

Environment	Red-Eyed Flies (%)	Larger Wings (%)	
Environment	Red-Eyed Flies (%)	Larger Wings (%)	
Cafeteria	76%	49%	
Chemistry Laboratory	58%	53%	
Biology Laboratory	58%	53%	

MAKE PR

The screenshot shows a GitHub pull request discussion. The pull request is titled 'advisor-drosophila commented 3 days ago'. The discussion includes a comment from 'advisor-drosophila' and a response from 'MarsBarLee'. The comment from 'advisor-drosophila' states: 'This sample was collected from the Biology Lab, not Chemistry Lab.' The response from 'MarsBarLee' states: 'Oh yes, I was at both labs that day and I mixed it up. Thanks for catching that!'. The pull request is for the 'article.md' file, showing a diff of the file's content. The diff shows changes to the file's content, including a table of data. The table has 4 columns: 'Environment', 'Red-Eyed Flies (%)', 'Larger Wings (%)', and an unlabeled column. The table has 4 rows of data. The first row is highlighted in blue. The second row is highlighted in green. The third row is highlighted in red. The fourth row is highlighted in green.

advisor-drosophila commented 3 days ago

article.md

... @@ -95,7 +95,7 @@ The source for this data can be found in [](./fruit\_fly\_data.csv).

Environment	Red-Eyed Flies (%)	Larger Wings (%)	
Environment	Red-Eyed Flies (%)	Larger Wings (%)	
Cafeteria	76%	49%	
Chemistry Laboratory	58%	53%	
Biology Laboratory	58%	53%	

advisor-drosophila 3 days ago

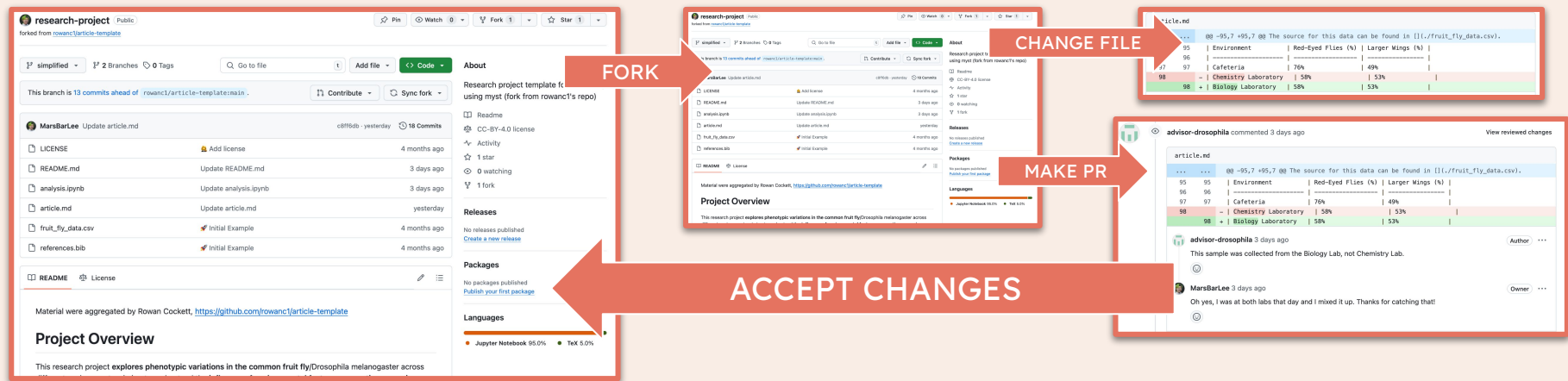
This sample was collected from the Biology Lab, not Chemistry Lab.

MarsBarLee 3 days ago

Oh yes, I was at both labs that day and I mixed it up. Thanks for catching that!

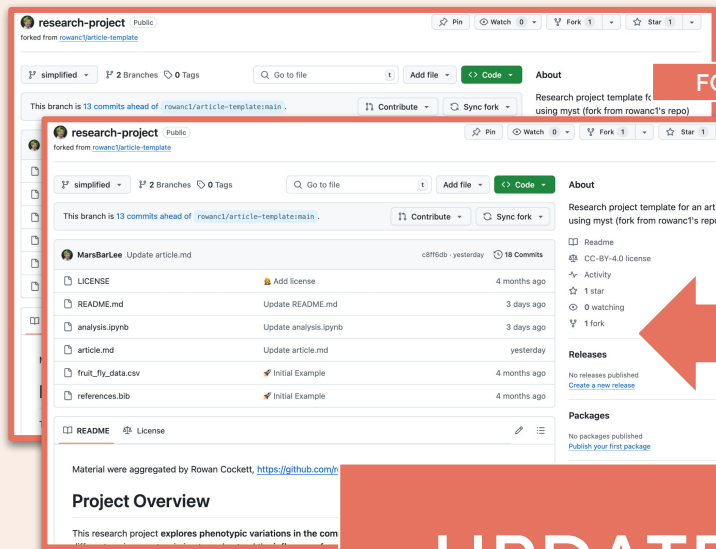
# Student

# Advisor

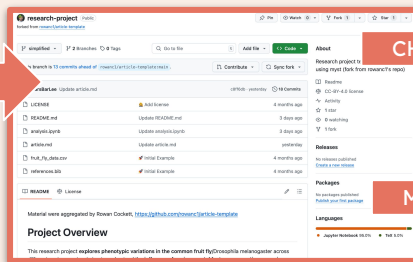


# Student

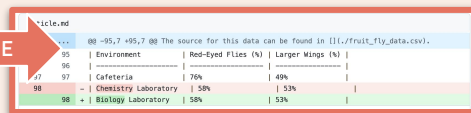
# Advisor



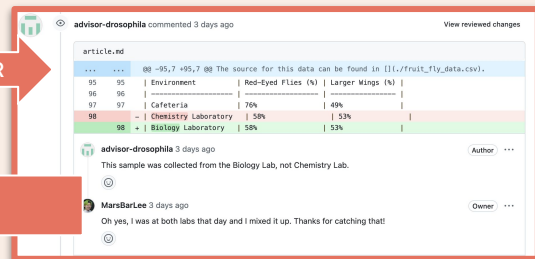
FORK



CHANGE FILE



MAKE PR



ACCEPT CHANGES

## UPDATED FILES!



# Advisor's fork

The screenshot shows a GitHub repository named 'research-project' which is a fork of 'rowanc1/article-template'. The repository is public and has 2 branches and 0 tags. It is 13 commits ahead of the parent repository. The repository was updated by MarsBarLee yesterday with 18 commits. The file list includes LICENSE, README.md, analysis.ipynb, article.md, fruit\_fly\_data.csv, and references.bib. The README section is visible, stating that material was aggregated by Rowan Cockett from the parent repository. The project overview mentions exploring phenotypic variations in the common fruit fly/Drosophila melanogaster across different conditions.

research-project Public  
forked from rowanc1/article-template

2 Branches 0 Tags  
Go to file  
Add file  
Code

This branch is 13 commits ahead of rowanc1/article-template:main  
Contribute Sync fork

MarsBarLee Update article.md c8ff6db · yesterday 18 Commits

LICENSE	Add license	4 months ago
README.md	Update README.md	3 days ago
analysis.ipynb	Update analysis.ipynb	3 days ago
article.md	Update article.md	yesterday
fruit_fly_data.csv	Initial Example	4 months ago
references.bib	Initial Example	4 months ago

README License

Material were aggregated by Rowan Cockett, <https://github.com/rowanc1/article-template>

## Project Overview

This research project explores phenotypic variations in the common fruit fly/Drosophila melanogaster across

About  
Research project template for an article, using myst (fork from rowanc1's repo)

- Readme
- CC-BY-4.0 license
- Activity
- 1 star
- 0 watching
- 1 fork

Releases  
No releases published  
[Create a new release](#)

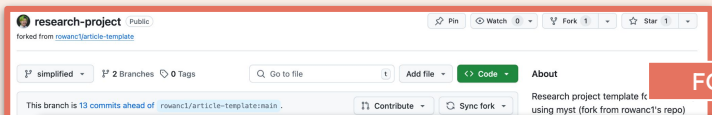
Packages  
No packages published  
[Publish your first package](#)

Languages  
Jupyter Notebook 95.0% TeX 5.0%

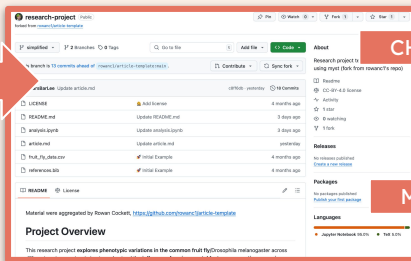
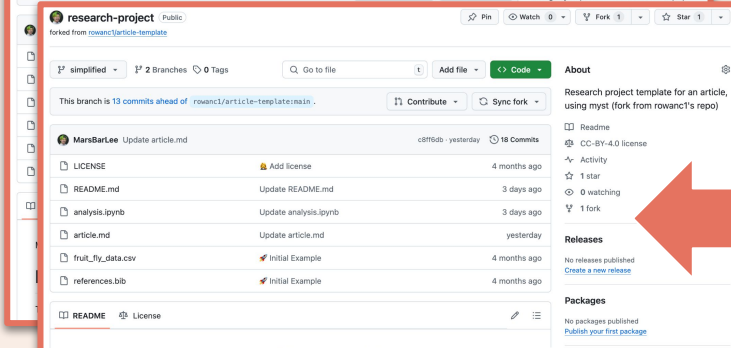
Link in chat:  
<https://github.com/MarsBarLee/research-project>

# Student

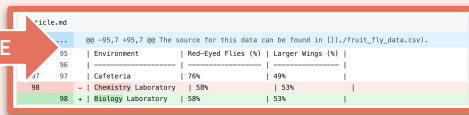
# Advisor



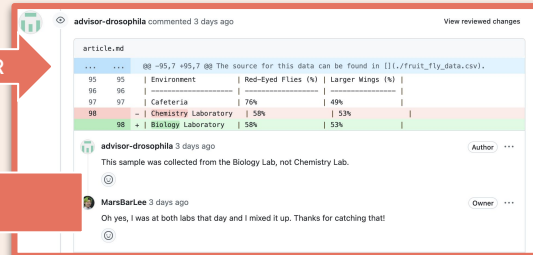
FORK



CHANGE FILE



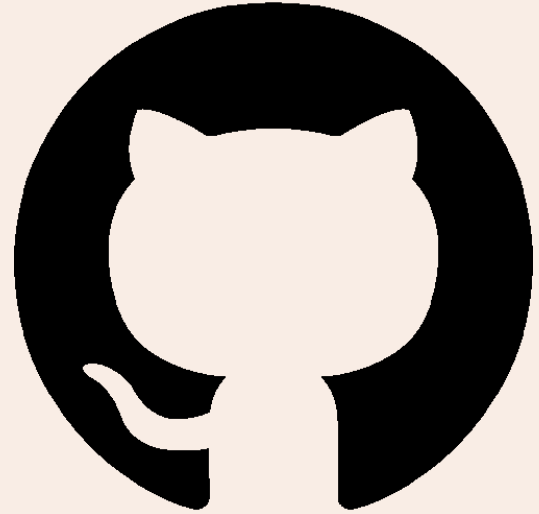
MAKE PR



ACCEPT CHANGES

UPDATED FILES!

# Why learn Git?



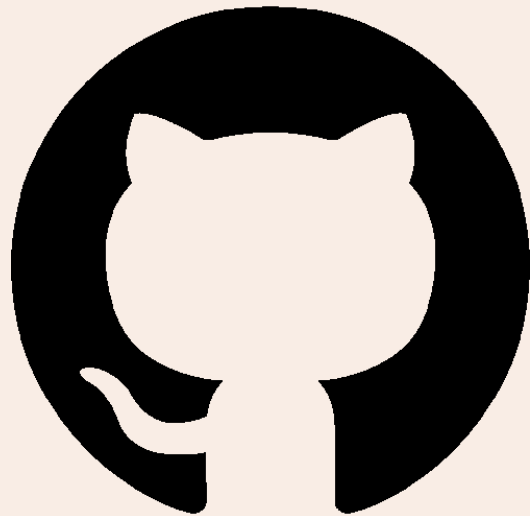
# Why learn Git?

*Why learn another thing?!*



# Why learn Git?

- Academia is already **competitive** and we're **overworked**
- It's a **nice to have** not a must-have
- **Fear of being scooped up.**



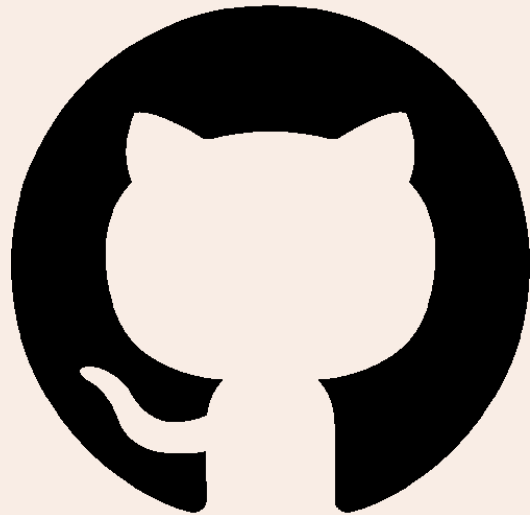
# Why learn Git?

- “I came to **research my passion**-not learn to code.”
- But we want to **learn Open Science**- and one useful tool is Git and Github



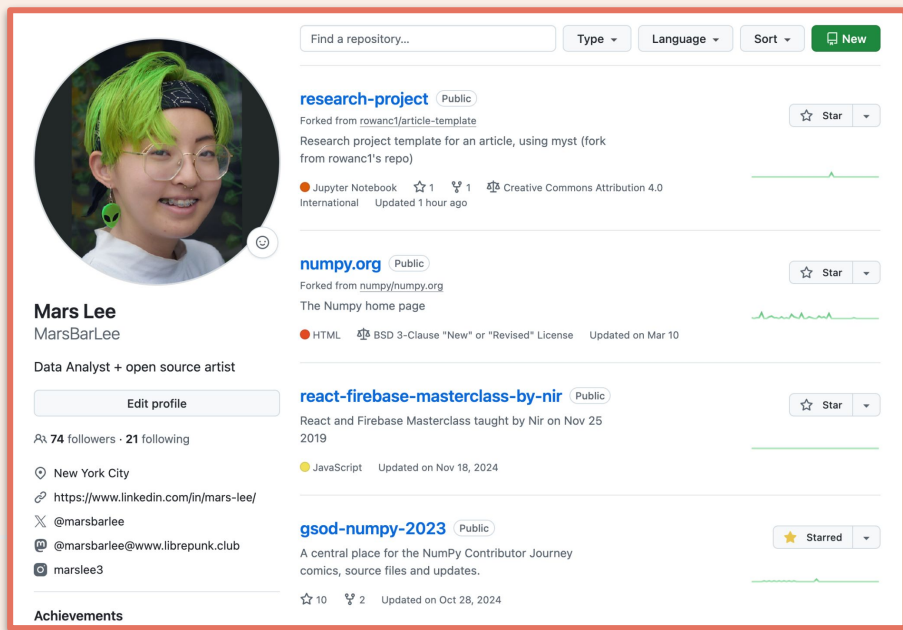
# Why learn Git?

- Get **credit for your code** work, not just articles
- **Visibility** of your research work beyond a few academic journals
- Help solve the **scientific reproducibility crisis**.



# Why learn Git?

- Learn a **new skill set!**



The screenshot displays the GitHub profile of Mars Lee (MarsBarLee), a Data Analyst and open source artist. The profile includes a circular profile picture of a person with green hair and glasses. Below the picture, the name 'Mars Lee' and handle 'MarsBarLee' are shown, along with the bio 'Data Analyst + open source artist' and an 'Edit profile' button. The profile statistics show 74 followers and 21 following. Social links for New York City, LinkedIn, Twitter, and other platforms are listed. The repository list on the right features four public repositories: 'research-project' (forked from rowanc1/article-template), 'numpy.org' (forked from numpy/numpy.org), 'react-firebase-masterclass-by-nir' (React and Firebase Masterclass taught by Nir on Nov 25 2019), and 'gsod-numpy-2023' (a central place for the NumPy Contributor Journey comics, source files and updates). Each repository entry includes its description, language, license, and update date.

Find a repository... Type Language Sort New

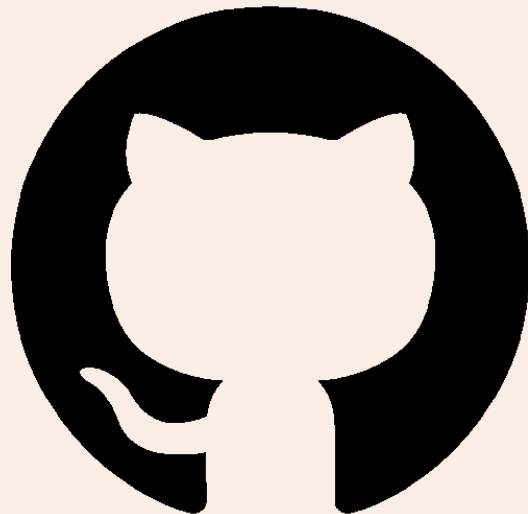
**research-project** (Public)  
Forked from rowanc1/article-template  
Research project template for an article, using myst (fork from rowanc1's repo)  
Jupyter Notebook 1 1 Creative Commons Attribution 4.0 International Updated 1 hour ago

**numpy.org** (Public)  
Forked from numpy/numpy.org  
The Numpy home page  
HTML BSD 3-Clause "New" or "Revised" License Updated on Mar 10

**react-firebase-masterclass-by-nir** (Public)  
React and Firebase Masterclass taught by Nir on Nov 25 2019  
JavaScript Updated on Nov 18, 2024

**gsod-numpy-2023** (Public)  
A central place for the NumPy Contributor Journey comics, source files and updates.  
10 2 Updated on Oct 28, 2024

**Achievements**





# Thank you!

Let's connect on LinkedIn!

---

Mars Lee

Open Science Training

