masai

Master Class on Github and Markdown

masai

Your Dream Offer

4 HR Round

3 Technical Interview

2 Technical Assignment

1 Resume, Portfolio and Github

Your Journey to secure the dream Job, 4 Step Process

masai

You will learn ...

What is Github

Why Github is important for Data Analysts

How to make an impactful Github Profile

What is Github

GitHub is a web-based platform that provides a range of services for version control and collaboration on different projects. It uses Git, a distributed version control system, to track changes in files and facilitate collaboration between multiple users.

Key Features :

- Portfolio building
- Reproducibility
- Documentation
- Collaboration
- Community

Overview of Github

- Web-based system for version control and collaborative development
- Used by developers and data analysts to track changes to code and data
- Powerful tools for organizing and contributing to open-source projects
- Offers an easy-to-use graphical user interface and command-line tools
- Provides access to vast libraries of reusable code and data

Why Data Analysts Need Github

Streamlining the Data Analysis Process

Why Data Analysts Need Github

- Track changes to data and scripts over time
- Collaborate with team members on data projects
- Easily share data visualizations and analyses with others
- Access powerful tools for data cleaning, manipulation, and analysis
- Contribute to open-source data projects and build a professional portfolio

Photo by Pexels

Collaboration Made Easy

- Track changes made to data and analysis scripts
- Collaborate seamlessly with team members
- Efficiently merge changes made to code
- Increase productivity through task distribution

Showcase Your Work

- Showcase code samples for job applications
- Provide public access to project portfolio
- Display personal data visualization projects
- Increase opportunities for professional networking

A Github Portfolio is a key way to impress potential employers and clients. After all, you've worked so hard on these projects, why not show them off?

Let’s start creating

Getting Started with Github

- Create a free Github account
- Create a repository to store data and scripts
- Clone a repository from Github to work on locally
- Commit changes to a repository and push changes to Github
- Collaborate with others using Github's issue tracking and pull request features

masai

How to make an impactful Github Repository

- Choose a Clear Repository Name
- Write a Compelling README
- Include Data Analysis Report
- Visualize Results
- Structure Your Repository

Choose a Clear Repository Name

Do-

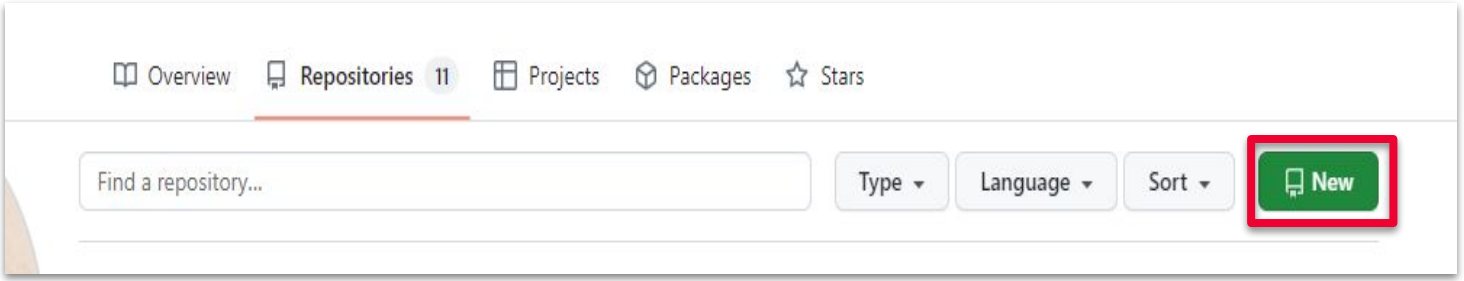
- Be Descriptive: DataAnalysisProjects
- Be Specific: PythonDataVisualizations
- Keep it Concise: DataAnalyticsPortfolio

Don’t -

- Avoid Ambiguity: ProjectX
- Don't Overuse Abbreviation: DAProj
- Avoid Special Characters: Data_Analysis_2023!

Steps to Create a Github Repository

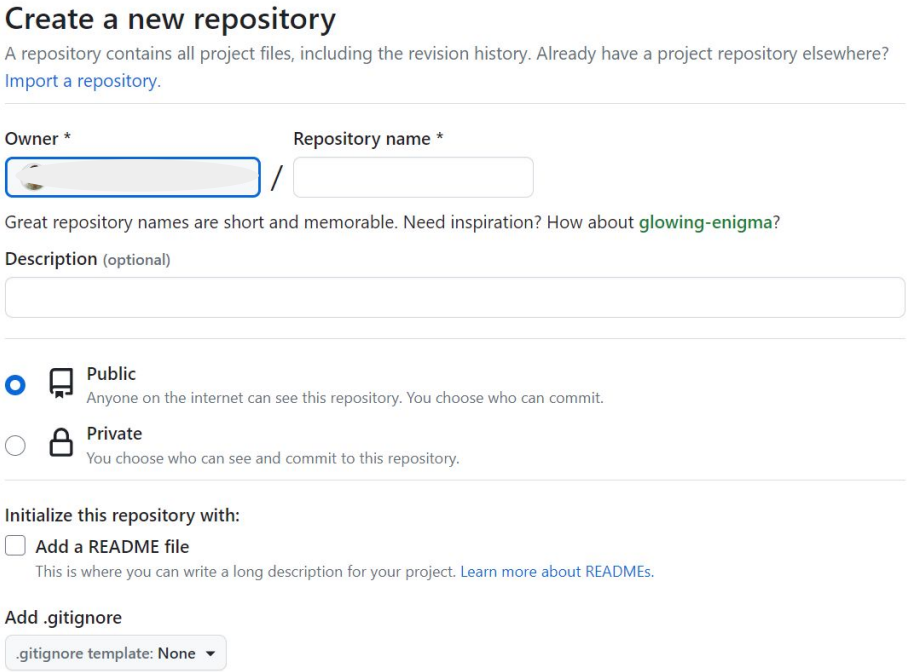
Step 1 : Click on New on the Repository Tab



masai

Steps to Create a Github Repository

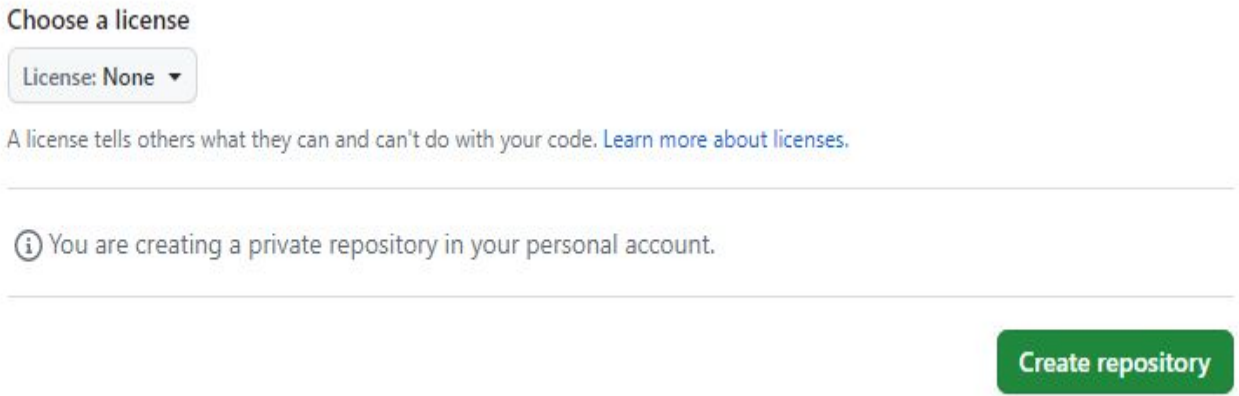
Step 2 : Insert the Repository Details



masai

Steps to Create a Github Repository

Step 3 : Click on Create Repository



masai

What is a README File

A README file is a text file that provides an introduction and overview of a project . Some common things to include in a README file are:

- A brief description of what the project does
- Steps involved in creating the project
- Analysis Generated
- Dashboard and Visualization
- Future Scope of the Project
- Conclusion

Markdown

Markdown is a lightweight markup language that's used to format text. It's used extensively on GitHub, where it's used to format README files, comments, and other types of text. Markdown is designed to be easy to read and write, even if it's not rendered as formatted text.

Markdown provides a simple syntax for formatting text, including:

- Headings and subheadings
- Lists (ordered and unordered)
- Links and images
- Bold and italic text
- Code blocks and inline code

Data Analysis Report

masai

Introduction:

Objective and key terms defined for analysis in report.

Data Description:

Data sources, variables, cleaning steps summarized with origin, methods, limitations.

Methodology:

Analytical methods explained with steps, transformations, modeling, assumptions justified.

Results / Insights:

Clear presentation of findings with visualizations, interpretation, and insightful explanations.

Future Scope: [Optional]

Main findings and conclusions summarized with implications, applications, and recommendations.

References and Citations: [Optional]

List of references and citations provided for external sources, datasets, software, libraries used.

masai

Checkpoints for GitHub

- Clear and Descriptive name for repository
- Folder structure for project files
- Make sure all uploaded files are accessible
- All codes are properly commented and documented
- README.md added to root of repository
- README starts with clear project description
- Outline of problem aimed to solve
- Description of data cleaning / pre processing
- Brief overview of data exploration and analysis
- Visualizations / Models / Algorithms
- Explanation of findings and conclusion from analysis at each step
- Summary of limitations / challenges faced

masai

Sample GitHub Portfolios

https://github.com/juniorcl/transaction-fraud-detection

https://github.com/s1dewalker/Airbnb-listings-NYC

masai

Thank You