

# Build Your Data Science Portfolio Using GitHub

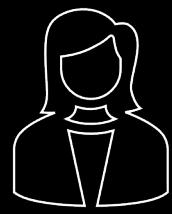




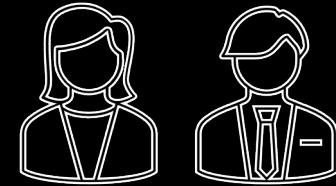
# Flow

- Why do we need to know about GitHub?
- How can it help in your career?
- What is GitHub?
- From project development to publicizing phase
- Project types that will be useful to demonstrate your skills

# Scenarios



Me



Recruiters



# Earn the Trust

Demonstrate hiring team that you have required skills

Build Portfolio of your Projects

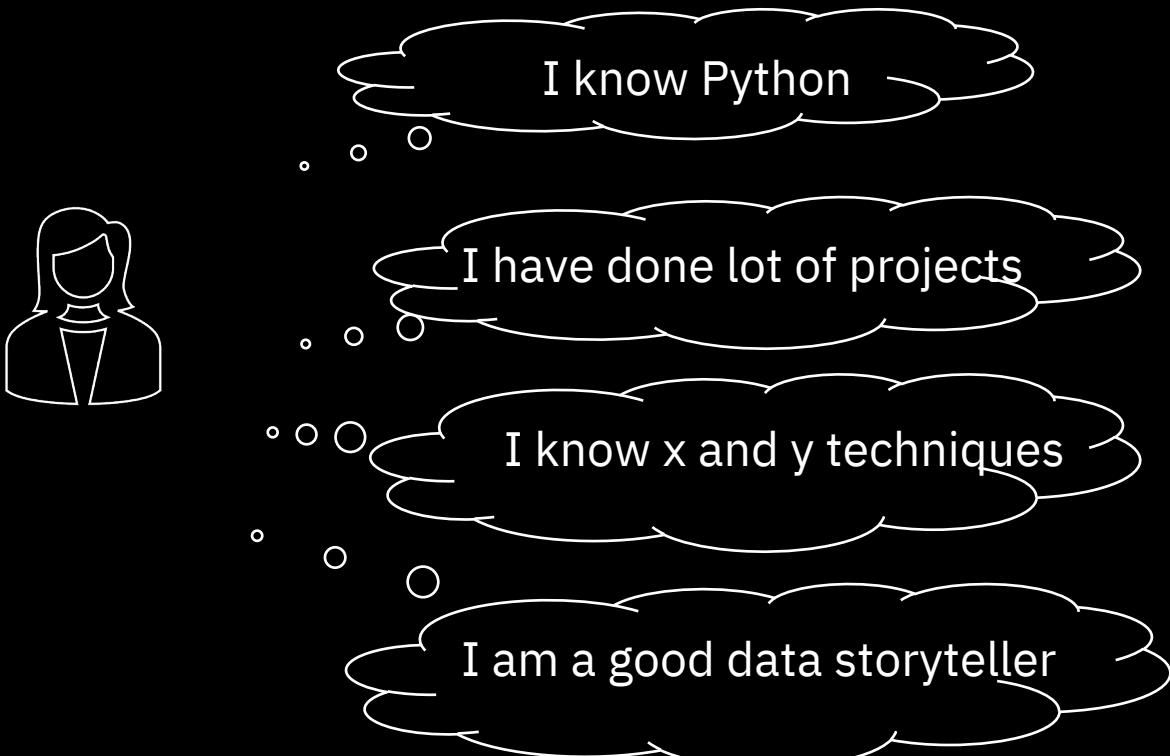




# Online Hosting Service

- Make your code visible to the community
- Document to reflect your skills
- Make it easy for hiring managers to find projects

# Which is better?



**Saishruthi**  
SSaishruthi

Edit profile

291 followers · 6 following · 5 repositories

Achievements

Organizations

**MAX-Text-Sentiment-Classifier**  
Forked from IBM/MAX-Text-Sentiment-Classifier  
Detect the sentiment captured in short pieces of text

Python 25 Apache License 2.0 Updated 3 hours ago

**Guvi\_AI\_For\_India**  
Materials used for Guvi Event

Jupyter Notebook Apache License 2.0 Updated 6 days ago

**pydax**  
Forked from CODAIT/pydax

Python 3 Apache License 2.0 Updated 24 days ago

**Kaggle\_Notebooks**  
This repo contains notebook related to kaggle dataset

Jupyter Notebook Apache License 2.0 Updated 25 days ago

**max-tutorial-app-python**  
Forked from IBM/max-tutorial-app-python  
A simple version of the MAX Object Detector Web App rewritten in python for use in the MAX tutorial

Python 9 Apache License 2.0 Updated on Mar 30

**VoiceAIM-website**  
Forked from ayeadaem/VoiceAIM-website  
Website for VoiceAIM

Python 3 Updated on Mar 30

**s2i**  
Test OpenShift S2I

Apache License 2.0 Updated on Mar 30

# Use Case – Predicting Bike Rentals

The screenshot shows the Dataquest platform interface. At the top, there's a logo featuring a rocket ship icon inside a circle next to the word "DATAQUEST". Below the logo, the title "Guided Project: Predicting Bike Rentals" is displayed, along with a "GET FREE LIFETIME ACCESS" button and some user icons.

The main area is divided into two sections: a "Learn" section on the left and a Jupyter notebook editor on the right.

**Learn Section:**

- Dashboard:** A link to return to the main Dataquest dashboard.
- Guided Project: Predicting Bike Rentals:** The current project being worked on.
- GET FREE LIFETIME ACCESS:** A button to sign up for free access.
- Basics.Ipynb:** The name of the Jupyter notebook currently open.
- Refresh:** A refresh button.
- Download:** A download button.
- Cell:** A cell toolbar with various options like Run, Stop, and Cell.
- Kernel:** A kernel toolbar with Python 3 selected.
- Help:** A help button.
- In [ ]:** An input cell placeholder.

**Project Description:**

Many U.S. cities have communal bike sharing stations where you can rent bicycles by the hour or day. Washington, D.C. is one of these cities. The District collects detailed data on the number of bicycles people rent by the hour and day.

Hadi Fanaee-T at the [University of Porto](#) compiled this data into a CSV file, which you'll work with in this project. The file contains 17380 rows, with each row representing the number of bike rentals for a single hour of a single day. You can download the data from the [University of California, Irvine's website](#). If you need help at any point, you can consult the solution notebook [in our GitHub repository](#).

Here's what the first five rows look like:

instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
1	2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0	3	13	16
2	2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.8	0	8	32	40
3	2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.8	0	5	27	32
4	2011-01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0	3	10	13
5	2011-01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0	0	1	1

Here are the descriptions for the relevant columns:

- **instant** - A unique sequential ID number for each row
- **dteday** - The date of the rentals
- **season** - The season in which the rentals occurred
- **yr** - The year the rentals occurred

**Instructions:**

1. Introduction to the Dataset

Back 1 / 7 Next

<https://app.dataquest.io/dashboard>

[Dashboard](#)

**Guided Project: Predicting Bike Rentals**

[GET FREE LIFETIME ACCESS](#)

### Learn

Many U.S. cities have communal bike sharing stations where you can rent bicycles by the hour or day. Washington, D.C. is one of these cities. The District collects detailed data on the number of bicycles people rent by the hour and day.

Hadi Fanaee-T at the [University of Porto](#) compiled this data into a CSV file, which you'll work with in this project. The file contains [17380](#) rows, with each row representing the number of bike rentals for a single hour of a single day. You can download the data from the [University of California, Irvine's website](#). If you need help at any point, you can consult the solution notebook [in our GitHub repository](#).

Here's what the first five rows look like:

instant	dteday	season	yr	mnth	hr	holiday	weekday	workingday	weathersit	temp	atemp	hum	windspeed	casual	registered	cnt
1	2011-01-01	1	0	1	0	0	6	0	1	0.24	0.2879	0.81	0	3	13	16
2	2011-01-01	1	0	1	1	0	6	0	1	0.22	0.2727	0.8	0	8	32	40
3	2011-01-01	1	0	1	2	0	6	0	1	0.22	0.2727	0.8	0	5	27	32
4	2011-01-01	1	0	1	3	0	6	0	1	0.24	0.2879	0.75	0	3	10	13
5	2011-01-01	1	0	1	4	0	6	0	1	0.24	0.2879	0.75	0	0	1	1

Here are the descriptions for the relevant columns:

- **instant** - A unique sequential ID number for each row
- **dteday** - The date of the rentals
- **season** - The season in which the rentals occurred
- **yr** - The year the rentals occurred

**Instructions**

1. Introduction to the Dataset

Basics.ipynb

Refresh Download

jupyter Basics (autosaved)

File Edit View Insert Cell Kernel Help Not Connected | Python 3

New Notebook Open... Make a Copy... Rename... Save and Checkpoint Revert to Checkpoint Print Preview Download as Trust Notebook Close and Halt

Notebook (.ipynb) Python (.py) HTML (.html) Markdown (.md) reST (.rst) PDF via LaTeX (.pdf)

## Create a new repository

A repository contains all project files, including the revision history. Already have a project repository elsewhere? [Import a repository](#).

### Repository template

Start your repository with a template repository's contents.

No template ▾

Owner \*

 SSAishruthi ▾

Repository name \*

Predicting\_Bike\_Rentals ✓

Great repository names are short and memorable. Need inspiration? How about [expert-system](#)?

### Description (optional)

This repo contains code that demonstrates predicting bike rentals using data science methods.

 Public

Anyone on the internet can see this repository. You choose who can commit.

 Private

You choose who can see and commit to this repository.

### Initialize this repository with:

Skip this step if you're importing an existing repository

Add a README file

This is where you can write a long description for your project. [Learn more](#).

Add .gitignore

Choose which files not to track from a list of templates. [Learn more](#).

.gitignore template: Python ▾

Choose a license

A license tells others what they can and can't do with your code. [Learn more](#).

License: Apache License 2.0 ▾

This will set  main as the default branch. Change the default name in your [settings](#).

### Grant your Marketplace apps access to this repository

You are subscribed to 1 Marketplace app

 Review Notebook App

Code Review and Collaboration for Jupyter Notebooks

**Create repository**

# GitHub Concepts

- Create a GitHub account
- Create a new repository
- Provide a repo name that reflects your work

# Upload the Code

The screenshot shows a GitHub repository page for "SSaishruthi / Predicting\_Bike\_Rentals". The "Code" tab is selected. A context menu is open over the repository summary, with the "Upload files" option highlighted by a red circle.

**Repository Summary:**

- Code: main branch, 1 branch, 0 tags
- Commits: 1 commit (by SSaishruthi Initial commit, 1 minute ago)
- Files: .gitignore, LICENSE, README.md
- README.md content: Predicting\_Bike\_Rentals
- Description: This repo contains code that demonstrates predicting bike rentals using data science methodologies

**Code Menu (highlighted):**

- Go to file
- Create new file
- Upload files** (highlighted with a red circle)

**About:**

This repo contains code that demonstrates predicting bike rentals using data science methodologies

**Readme:**

**Releases:**

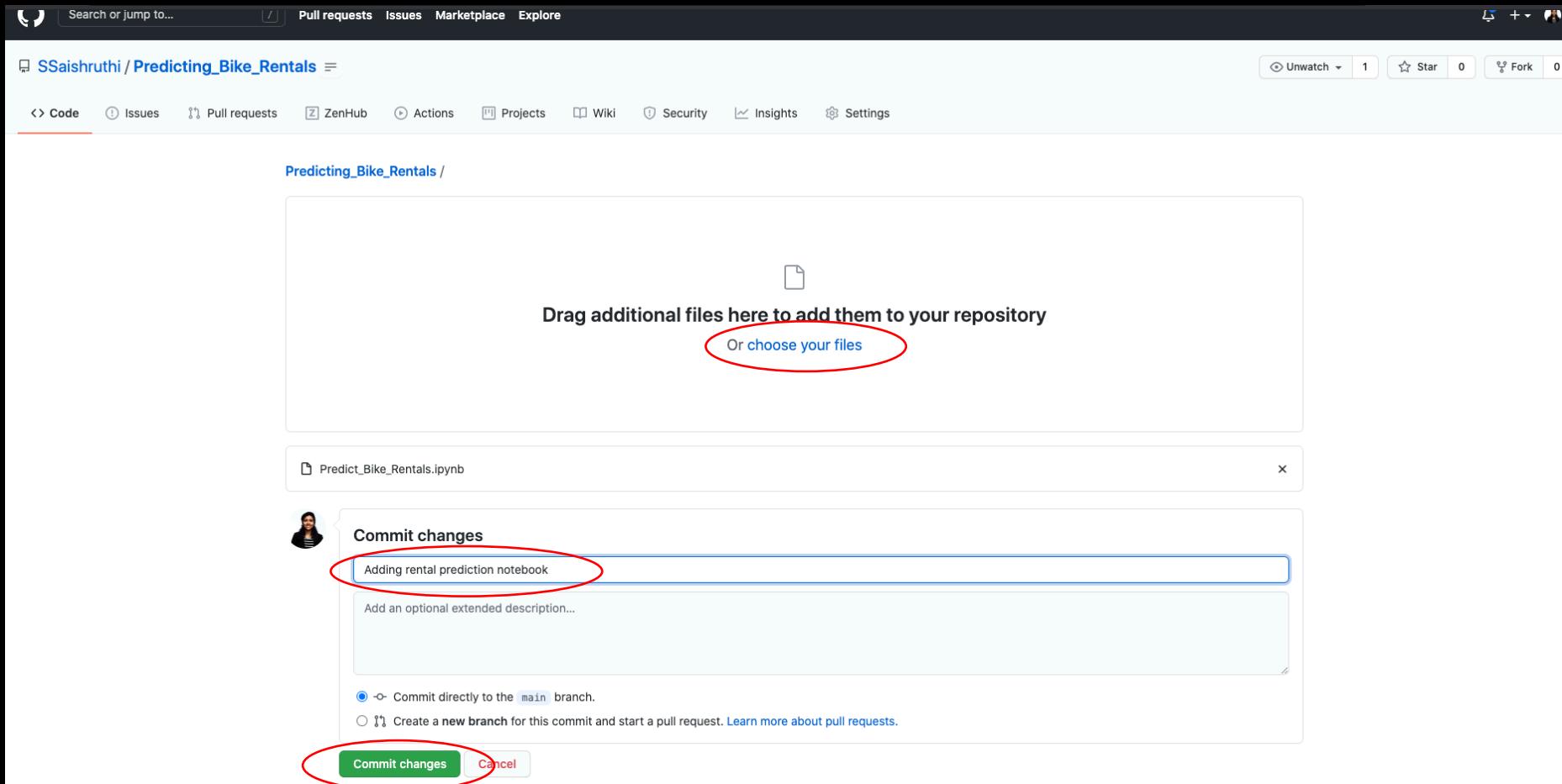
No releases published  
Create a new release

**Packages:**

No packages published  
Publish your first package

© 2021 GitHub, Inc. Terms Privacy Security Status Docs Contact GitHub Pricing API Training Blog About

# Upload the Code



# Does it explain your work completely?

The screenshot shows a GitHub repository page for the user 'SSaishruthi'. The repository name is 'Predicting\_Bike\_Rentals'. The page has a dark theme. At the top, there are navigation links: Pull requests, Issues, Marketplace, and Explore. On the right side, there are buttons for Unwatch (with 1), Star (0), Fork (0), and a profile icon. Below the header, there are tabs: Code (selected), Issues, Pull requests, ZenHub, Actions, Projects, Wiki, Security, Insights, and Settings. The main content area shows a list of files and commits. A commit by 'SSaishruthi' titled 'Adding rental prediction notebook' is highlighted. Below the commit list is a 'README.md' file, which contains the text 'Predicting\_Bike\_Rentals' and a descriptive paragraph about the repo's purpose. A red oval highlights the title 'Predicting\_Bike\_Rentals' and the descriptive text. A red square highlights the edit icon next to the 'README.md' file. To the right of the file list, there are sections for 'About', 'Readme', 'Apache-2.0 License', 'Releases' (which says 'No releases published. Create a new release'), and 'Packages' (which says 'No packages published. Publish your first package'). At the bottom, there are links for GitHub terms and conditions, privacy, security, status, and documentation, along with links for Contact GitHub, Pricing, API, Training, Blog, and About.

SSaishruthi / Predicting\_Bike\_Rentals

Code Issues Pull requests ZenHub Actions Projects Wiki Security Insights Settings

main · 1 branch · 0 tags Go to file Add file Code

SSaishruthi Adding rental prediction notebook 17c9984 now 2 commits

.gitignore Initial commit 11 minutes ago

LICENSE Initial commit 11 minutes ago

Predict\_Bike\_Rentals.ipynb Adding rental prediction notebook now

README.md Initial commit 11 minutes ago

**Predicting\_Bike\_Rentals**

This repo contains code that demonstrates predicting bike rentals using data science methodologies

About

This repo contains code that demonstrates predicting bike rentals using data science methodologies

Readme

Apache-2.0 License

Releases

No releases published Create a new release

Packages

No packages published Publish your first package

© 2021 GitHub, Inc. Terms Privacy Security Status Docs Contact GitHub Pricing API Training Blog About

# How about this?

The image shows two side-by-side screenshots of GitHub repository pages for 'Predicting\_Bike\_Rentals'.

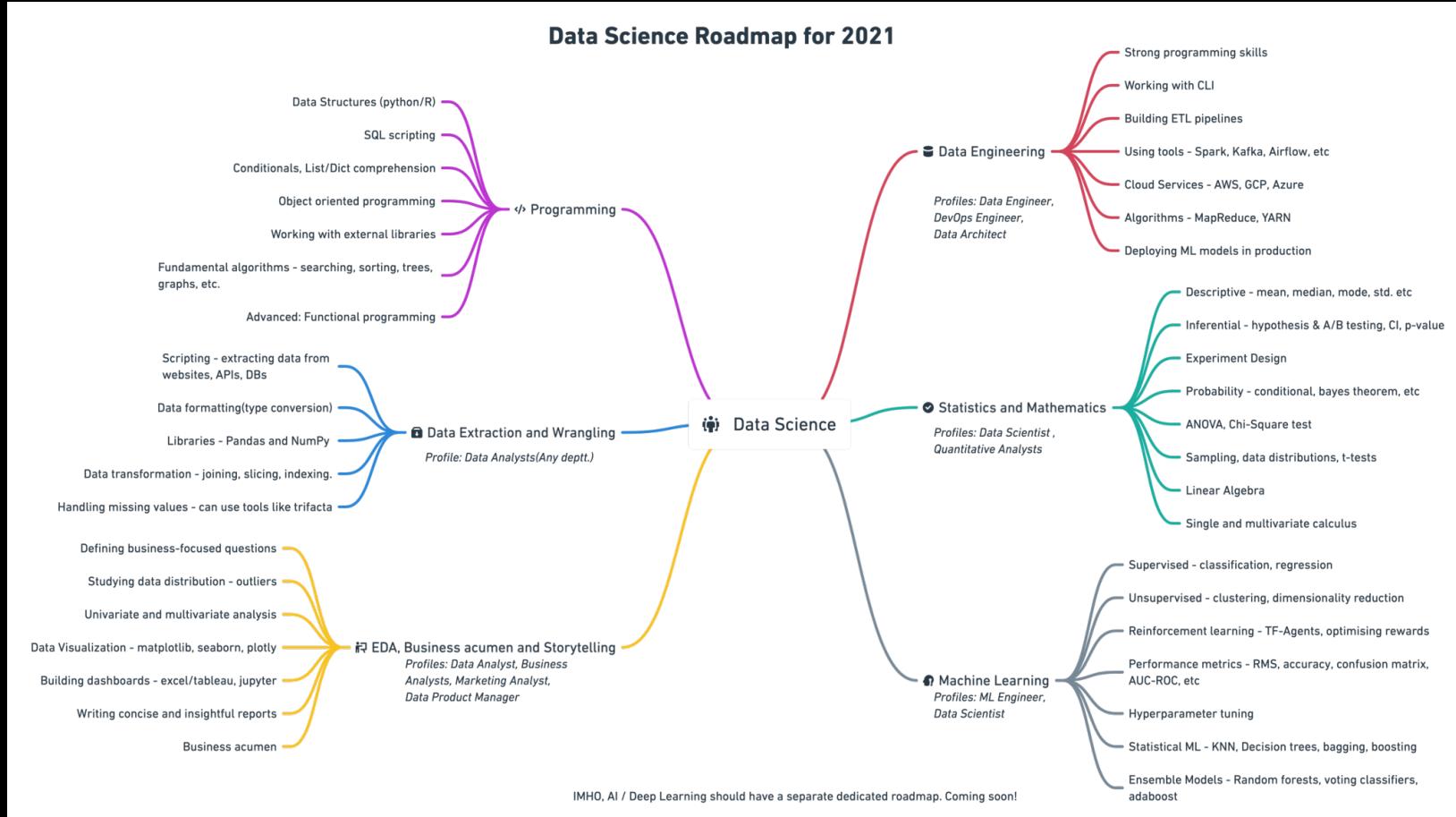
**Left Repository (Older Version):**

- Code:** main (1 branch), 0 tags.
- Commits:** SSaishruthi Adding rental prediction notebook (17c9984 now), .gitignore (Initial commit 11 minutes ago), LICENSE (Initial commit 11 minutes ago), Predict\_Bike\_Rentals.ipynb (Adding rental prediction notebook now), README.md (Initial commit 11 minutes ago).
- About:** This repo contains code that demonstrates predicting bike rentals using data science methodologies.
- Readme:** Predicting\_Bike\_Rentals
- Packages:** No packages published. Publish your first package.

**Right Repository (Updated Version):**

- Code:** main (2 branches), 0 tags.
- Commits:** SSaishruthi Update to readme (79b278a now), .gitignore (Initial commit 4 hours ago), LICENSE (Initial commit 4 hours ago), Predict\_Bike\_Rentals.ipynb (Adding rental prediction notebook 3 hours ago), README.md (Update to readme now).
- About:** This repo contains code that demonstrates predicting bike rentals using data science methodologies.
- Readme:** Predicting\_Bike\_Rentals
- Business Problem:** Peak hours have less bikes for rentals which impacts business profit.
- Dataset Used:** To address this business problem, Bike Sharing dataset from UCI repository.
- Tools Used:**
  - NumPy
  - Pandas
  - Matplotlib
  - Scikit-Learn
- Languages:** Jupyter Notebook 100.0%

# Good to Cover



<https://towardsdatascience.com/data-science-learning-roadmap-for-2021-84f2ba09a44f>

# My Story

towards  
data science

Following 587K Followers Editors' Picks Features Deep Dives Grow Contribute About

## Logistic Regression — Detailed Overview

Saishruthi Swaminathan Mar 15, 2018 · 5 min read

### Logistic Regression Model

Inputs:  $X_1, X_2, X_3$  || Weights:  $\theta_1, \theta_2, \theta_3$  || Outputs: Happy or Sad  
@dataaspirant.com

Figure 1: Logistic Regression Model (Source: <http://dataaspirant.com/2017/03/02/how-logistic-regression-model-works/>)

Logistic Regression was used in the biological sciences in early twentieth century. It was then used in many social science applications. Logistic Regression is used when the dependent variable(target) is categorical.

Saishruthi Swaminathan  
Passionate about transforming data into useful products. Happy sharing my knowledge in data science to all!

SAISHRUTHI SWAMINATHAN FOLLOWS

- Towards Data Science
- Arseny Kravchenko
- Code Like A Girl
- Jupyter Blog
- Tirthayoti Sarkar

See all (36)

Blogrolls help your readers discover writers you follow. Writers who have published most recently show up at the top.

Turn on Not now

Cost reduction over time

Figure 13: Cost Reduction

Train and test accuracy of the system is 100 %

This implementation is for binary logistic regression. For data with more than 2 classes, softmax regression has to be used.

This is an educational post and inspired from Prof. Andrew Ng's deep learning course.

Full code :  
[https://github.com/SSaishruthi/LogisticRegression\\_Vectorized\\_Implementation/blob/master/Logistic\\_Regression.ipynb](https://github.com/SSaishruthi/LogisticRegression_Vectorized_Implementation/blob/master/Logistic_Regression.ipynb)

6.3K Q 35

# GitHub Pages

- GitHub Pages is a static site hosting service
- You can host your site on GitHub's `github.io` domain

The screenshot shows a GitHub Pages website for Saishruthi Swaminathan. At the top, there is a navigation bar with links for "Publications", "Talks", "Blog Posts", "CV", and "Volunteer". Below the navigation bar, there is a profile section featuring a circular profile picture of Saishruthi, her name, and her title as a "Data Scientist / Developer Advocate in IBM". It also lists her location as "California" and provides links to her Twitter, LinkedIn, and GitHub profiles. To the right of the profile section, there is a heading "Blog posts" followed by a list of three blog entries. Each entry includes the title, a "less than 1 minute read" note, the publication date, and a brief description.

Saishruthi Swaminathan

Publications   Talks   Blog Posts   CV   Volunteer

**Blog posts**

2018

[Learn What Lies Beneath Our Ready-to-Run Deep Learning Models](#)  
less than 1 minute read  
Published: November 05, 2018

Model Asset Exchange is the one stop shop for open source deep learning models, covering common application domains, such as audio, image, text or video processing. When you pick a model from MAX, you:

- can be sure that the model code and intellectual property (IP) have been vetted and tested
- can deploy the model as a microservice in minutes
- can consume the model using a simple REST API Looks fascinating, right? Now, I am going to take you through what's happening behind the scenes.

[Logistic Regression—Detailed Overview](#)  
less than 1 minute read  
Published: March 15, 2018

This blog gives a detailed overview of logistic regression model

[Is R-Square value always between 0 and 1](#)  
less than 1 minute read  
Published: March 01, 2018

R-Square metric analysis

# Courses

Browse > Data Science > Data Analysis

## Data Visualization with Python

★★★★★ 4.5 9,079 ratings

 Saishruthi Swaminathan +1 more instructor

[Go To Course](#)   Already enrolled   Financial aid available

114,735 already enrolled

Offered By 

Browse > Data Science > Data Analysis

## Data Analysis with R

 Yiwen Li +1 more instructor

[Enroll for Free](#)   Starts Apr 30   Financial aid available

Try for Free: Enroll to start your 7-day full access free trial

Offered By 

Browse > Data Science > Data Analysis

## Data Visualization with R

 Saishruthi Swaminathan +1 more instructor

Starting May 2021   Already enrolled   Financial aid available

Offered By 