Technological Basics II

Streamlit App with User Authentication

Reminder: Tech Basic Project demos and feedback will be on the 17th December, 9.45 AM.

If you can’t make the demos, please let me know asap.

Depending on what the weather/vibe is like, we could head to an Xmas market afterwards for those up for it :)

Reminder: TBII Exam Demoes

By now, it should be clear that for the TBII Exam project, you will be building a **Streamlit app** based on your TBI project. The assessment criteria for the exam will be shared within the next two classes.

For the pitches, you will have **a maximum of 5 minutes** to present your TBI project and share the features you plan to develop. This pitch accounts for **10% of your grade**.

The main purpose of this exercise is to celebrate your projects and gather constructive feedback from your peers.

Goals for today

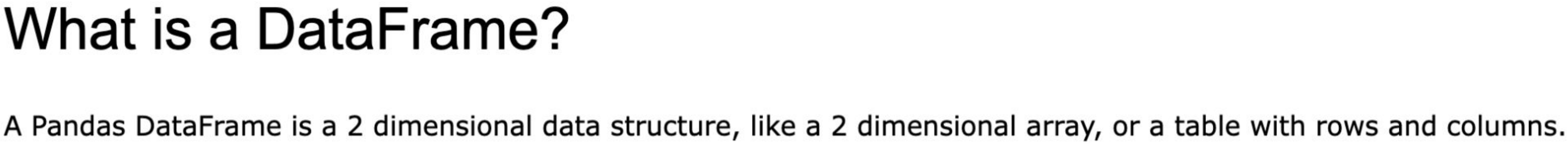
- Pandas to read and write data to MongoDB
- Developing and Deploying a Streamlit app with User Authentication

pandas - The Data Analytics Library

pandas

- Python library used for working with data sets
- It has **functions** for analysing, cleaning, exploring and manipulating data
- It is **fast**, **powerful**, **flexible** and **easy to use**
- The name pandas actually has nothing to do with pandas - and is in reference to “Panel Data” and “**Python Data Analysis**”
- Pandas gives you answers about the data. Like:
  - Is there a correlation between two or more columns?
  - What is average value? Max value? Min value?
  - Pandas are also able to delete rows that are not relevant, or contains wrong values, like empty or NULL values. This is called *cleaning* the data.

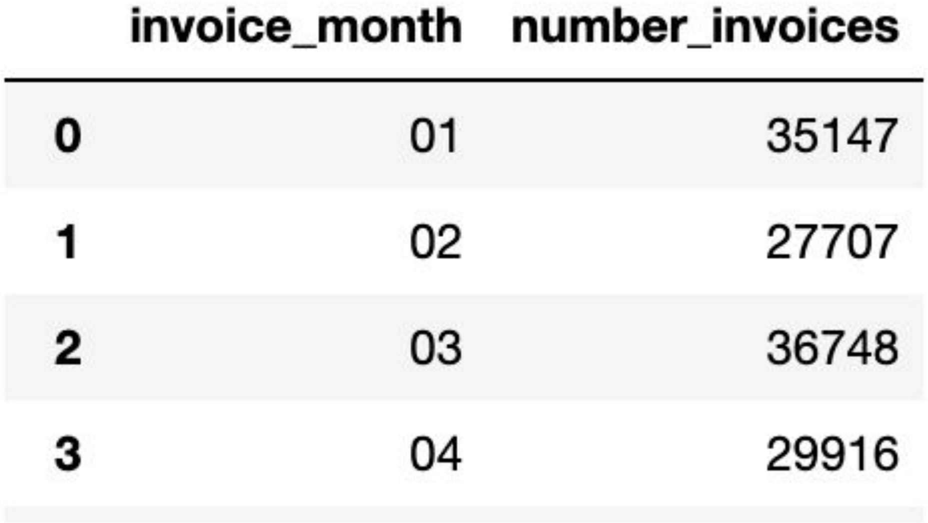
Pandas DataFrames



What is a DataFrame?

A Pandas DataFrame is a 2 dimensional data structure, like a 2 dimensional array, or a table with rows and columns.

0, 1, 2, 3 is the **index**. 0 is row 1, 1 is row 2 and 2 is row 3….



	invoice_month	number_invoices
0	01	35147
1	02	27707
2	03	36748
3	04	29916

Set up: Launch a Jupyter Notebook

- Open your **command line, locate your tech basics two folder, create a 08Lecture folder** and launch the jupyter notebook dashboard:

  `jupyter notebook` or

  `python -m notebook`

- Rename the notebook, call it reading-and-writing-data
- Can you download the data from here and upload to a data folder folder in your 08Lecture folder.

Activity: Do you know how to “read” the .csv flat file from your data folder and store as a variable?

Hint:

```
pd.read_csv("data/sample_data_cleaned.csv")
```

DATA
DOESN'T
LIE

For reference, more Pandas Functions

- Check out this resource for a complete list of pandas general functions and more:

https://pandas.pydata.org/docs/reference/general_functions.html

- If you want to work with other file extensions, this resource will be handy:

https://pandas.pydata.org/docs/reference/io.html#

Recap: Connect Streamlit to a Data Source

Reference:
https://docs.streamlit.io/develop/tutorials/databases

Step 1. Open MongoDB Atlas. If you were not here last week, create an account and run a cluster.

Create an Atlas Account here, and follow the instructions to create a “cluster”:

https://www.mongodb.com/cloud/atlas/register

Please do not provide any credit card information (at least not for the purpose of the course) and keep track of the username and password you created.

Step 2. Create a .streamlit/secrets.toml file in your 08Lecture Folder. You can copy across last weeks secrets file.

Add your username and password to this file.

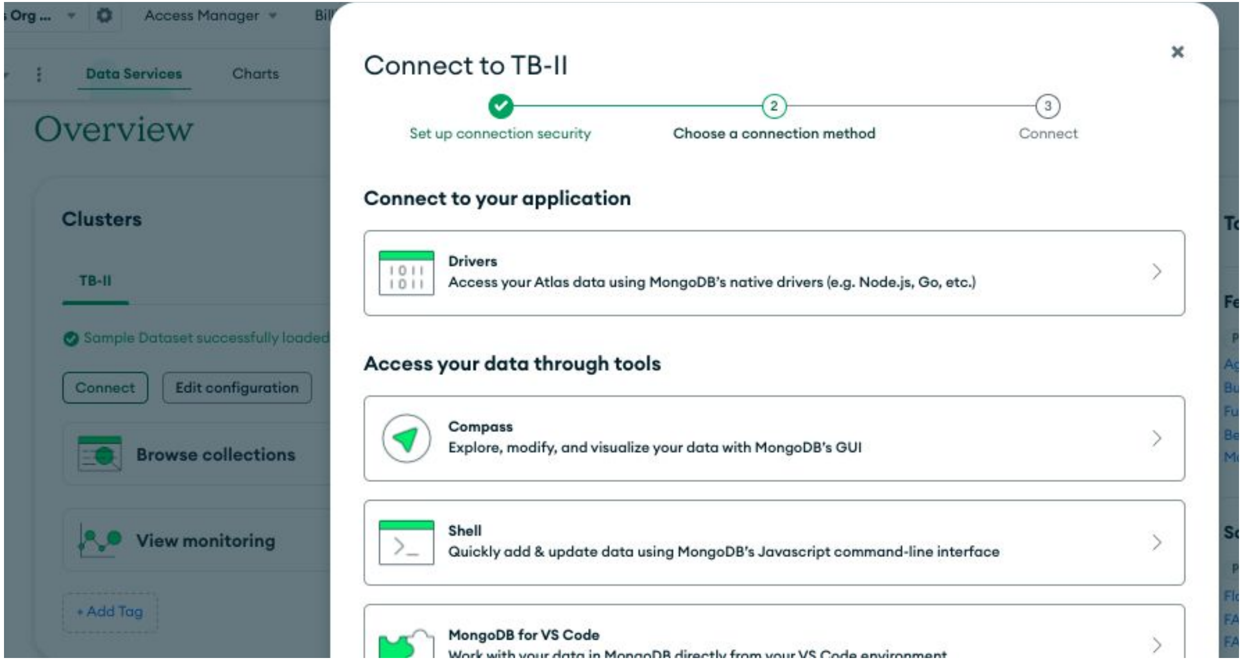
```
username = "xxx"

password = "xxx"
```

Step 3. Do you remember how to connect to your cluster?

You will need to install the dependency pymongo, if you were not here last week

```
pip3 install pymongo
```



Connect to TB-II

Set up connection security — 2 Choose a connection method — 3 Connect

Connect to your application

Drivers
Access your Atlas data using MongoDB's native drivers (e.g. Node.js, Go, etc.)

Access your data through tools

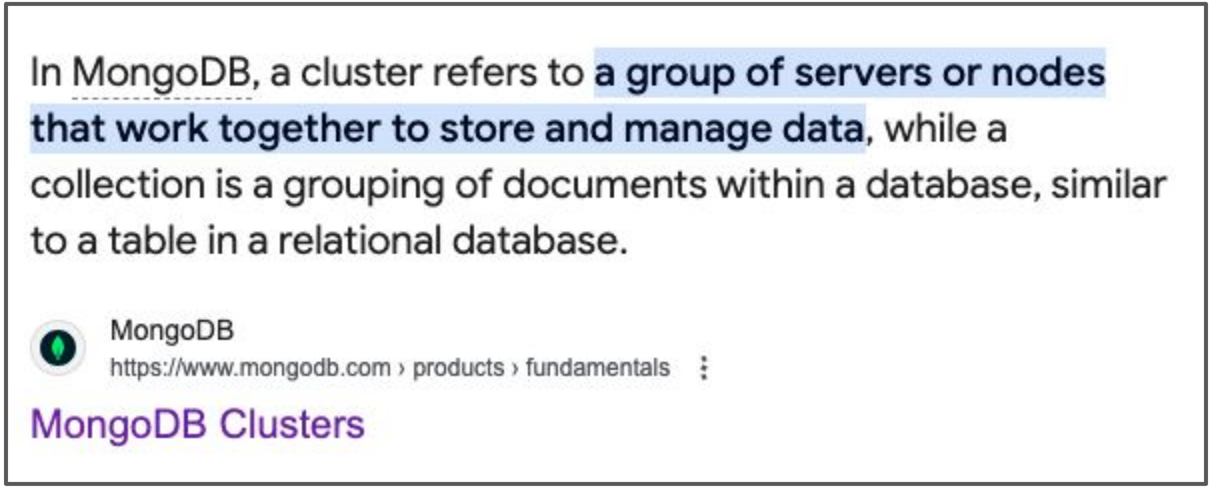
Compass
Explore, modify, and visualize your data with MongoDB's GUI

Shell
Quickly add & update data using MongoDB's Javascript command-line interface

MongoDB for VS Code
Work with your data in MongoDB directly from your VS Code environment

Or you can use my code, but double check your uri path:
https://github.com/shaq31415926/tech-basics/blob/main/tech_basics_two/08Lecture/src/helpers.py#L5

A cluster is a group of servers working together for data storage, high availability and scalability

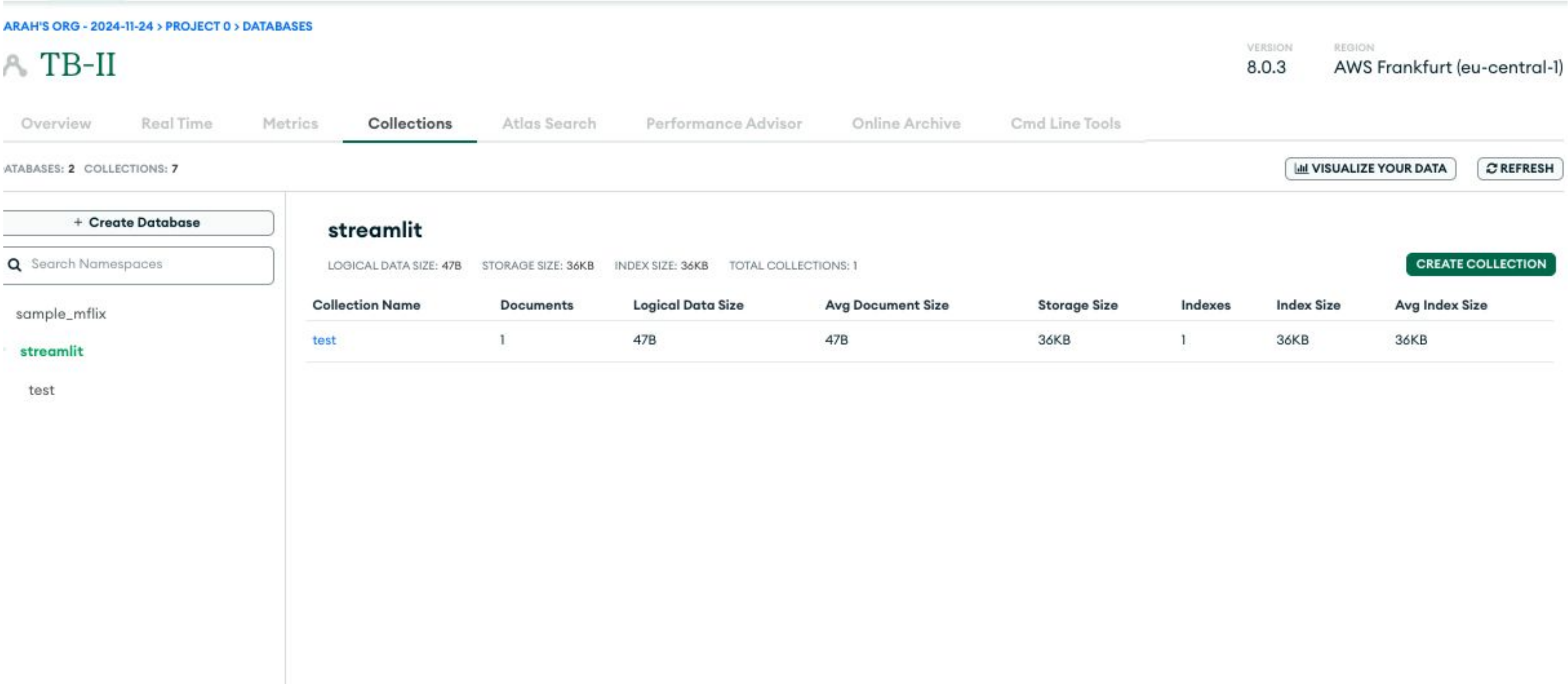


In MongoDB, a cluster refers to a group of servers or nodes that work together to store and manage data, while a collection is a grouping of documents within a database, similar to a table in a relational database.

MongoDB
https://www.mongodb.com › products › fundamentals

MongoDB Clusters

Step 4. If you were not here last week, create a Database and a Collection



ARAH'S ORG - 2024-11-24 > PROJECT 0 > DATABASES

TB-II

VERSION 8.0.3

REGION AWS Frankfurt (eu-central-1)

ATABASES: 2 COLLECTIONS: 7

sample_mflix

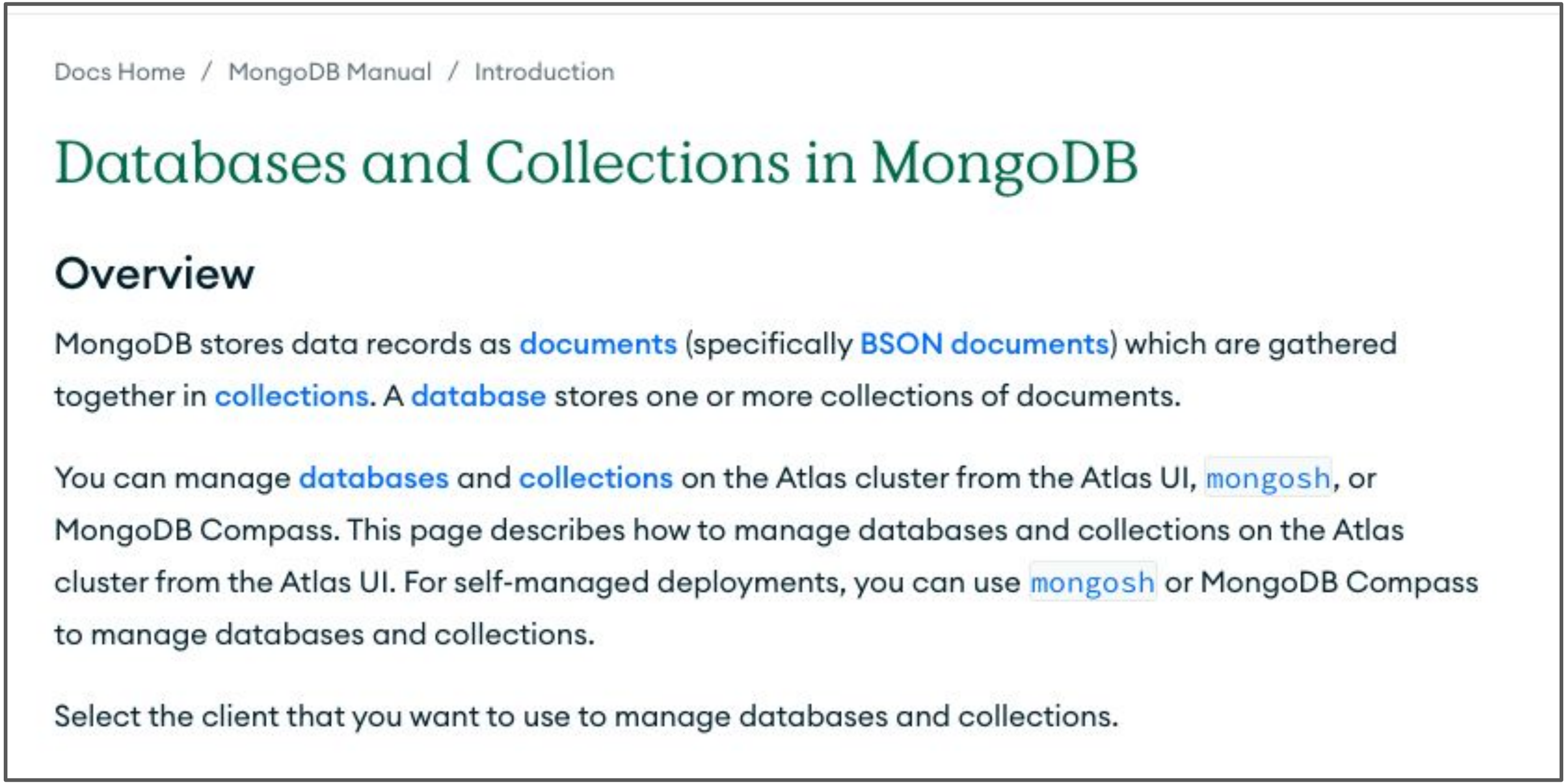
streamlit

test

streamlit

LOGICAL DATA SIZE: 47B STORAGE SIZE: 36KB INDEX SIZE: 36KB TOTAL COLLECTIONS: 1

Collection Name	Documents	Logical Data Size	Avg Document Size	Storage Size	Indexes	Index Size	Avg Index Size
test	1	47B	47B	36KB	1	36KB	36KB



Docs Home / MongoDB Manual / Introduction

Databases and Collections in MongoDB

Overview

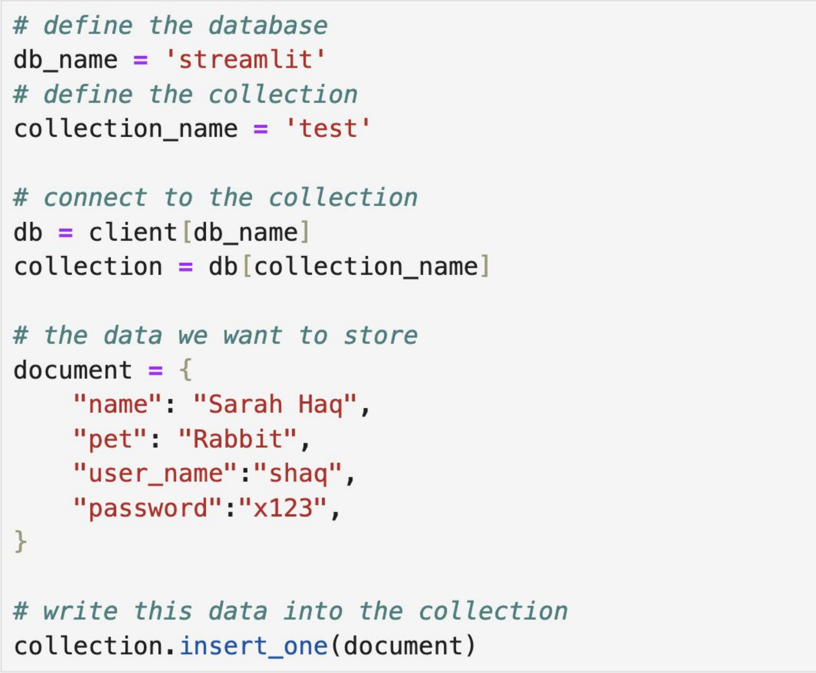
MongoDB stores data records as documents (specifically BSON documents) which are gathered together in collections. A database stores one or more collections of documents.

You can manage databases and collections on the Atlas cluster from the Atlas UI, `mongosh`, or MongoDB Compass. This page describes how to manage databases and collections on the Atlas cluster from the Atlas UI. For self-managed deployments, you can use `mongosh` or MongoDB Compass to manage databases and collections.

Select the client that you want to use to manage databases and collections.

Reference: https://www.mongodb.com/docs/manual/core/databases-and-collections/

Step 5. If you were not here last week, write some data to your collection and check if it worked



```
# define the database
db_name = 'streamlit'
# define the collection
collection_name = 'test'

# connect to the collection
db = client[db_name]
collection = db[collection_name]

# the data we want to store
document = {
    "name": "Sarah Haq",
    "pet": "Rabbit",
    "user_name":"shaq",
    "password":"x123",
}

# write this data into the collection
collection.insert_one(document)
```

Activity: Can you fetch the password for a specific username of your choice?

- Connect to your collection from last week. Look over the last week's notebook or you can use this function: https://github.com/shaq31415926/tech-basics/blob/c8285f4c6ee3eb923b240f39fa7991ab38aee839/tech_basics_two/08Lecture/src/helpers.py#L27
- Subset the Data Frame with a user choice of your choice, and manipulate the output to get the output.

Activity: Can you fetch the list of usernames?

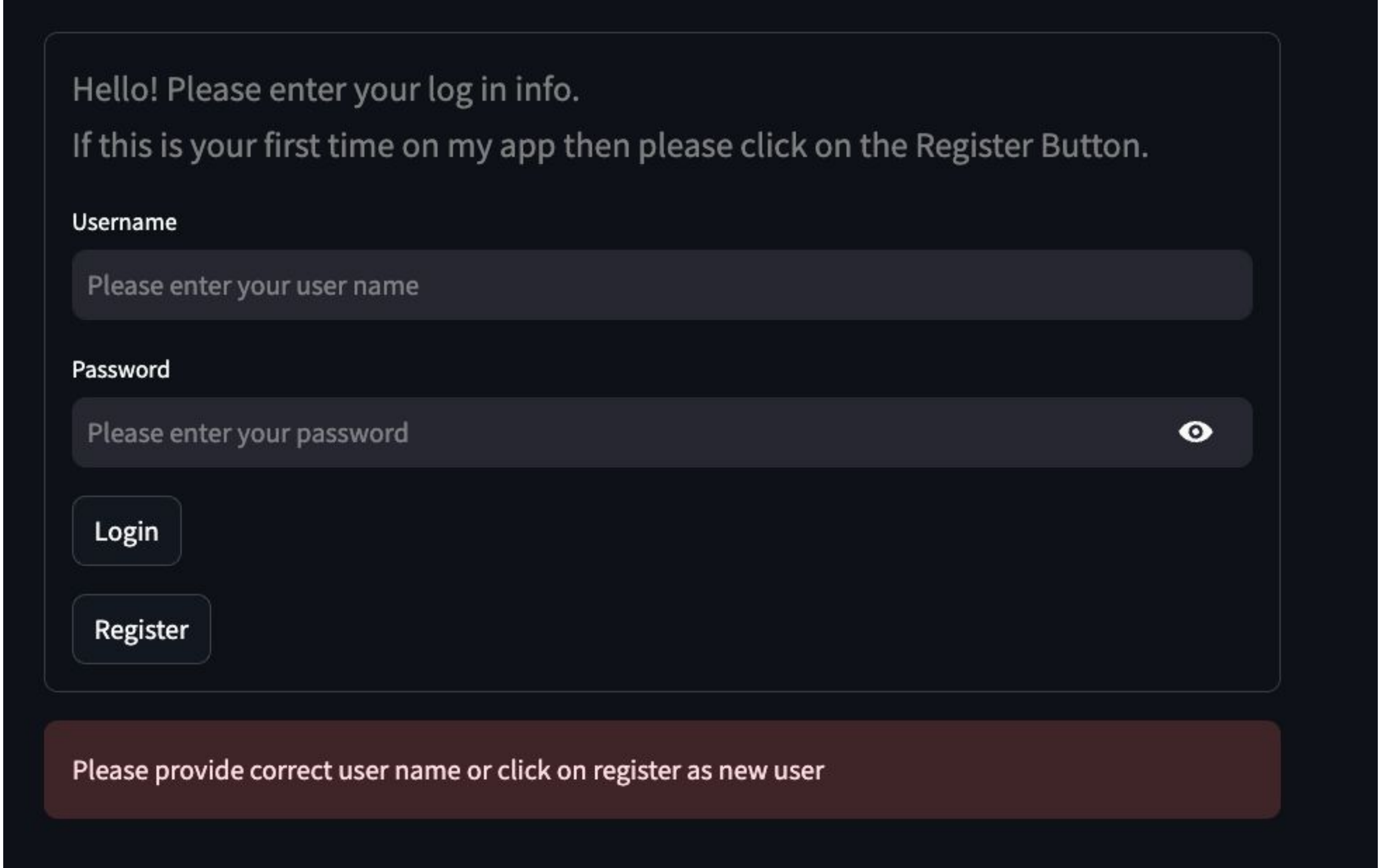
- Look over last week’s notebook if you are stuck.

Activity: Do you need to do any data cleaning?!

Streamlit App *with some very lightweight* User authentication

Activity: Create a Streamlit app with some lightweight user authentication

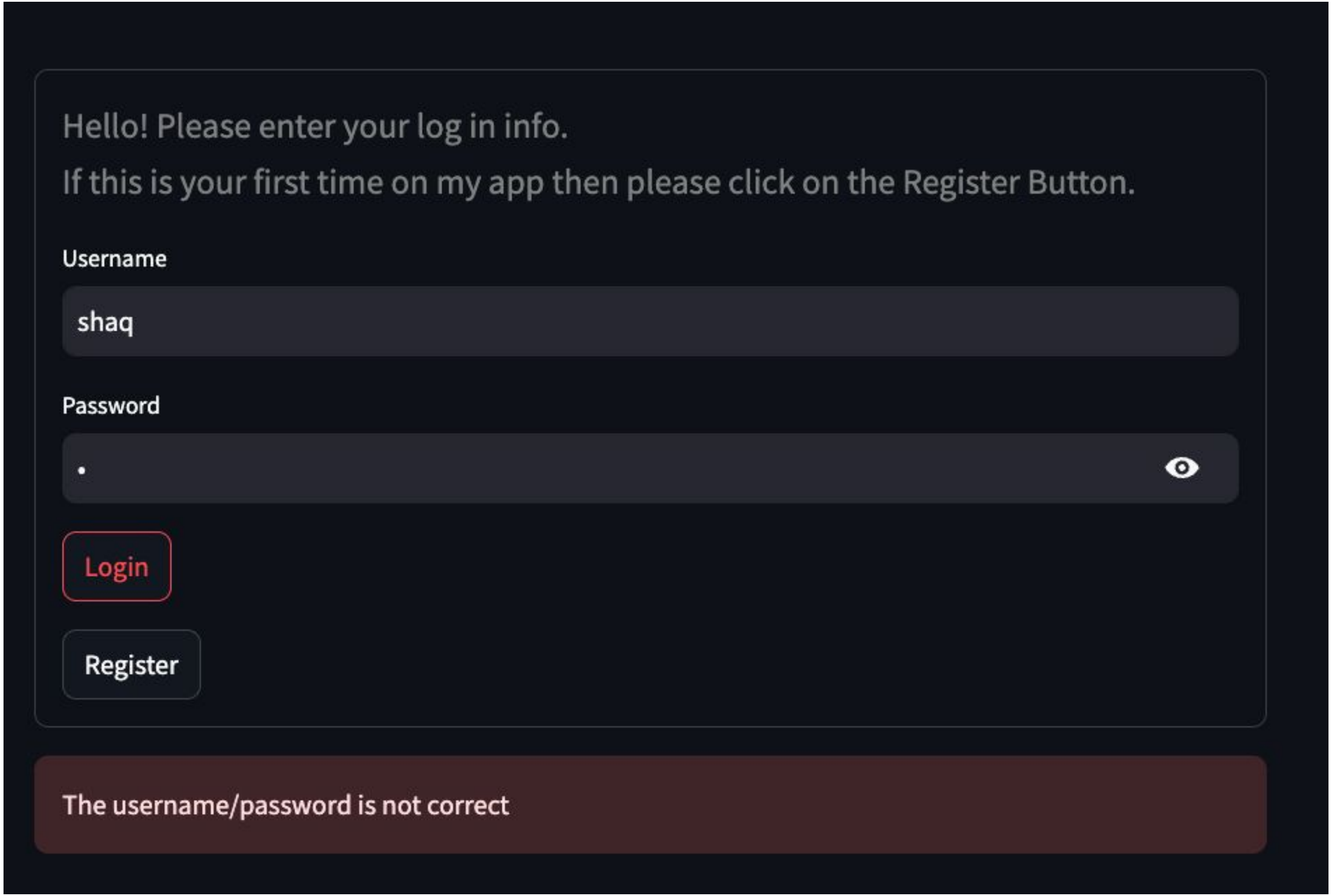
- Open PyCharm and create a new python file in your 08Lecture folder name it `login_page.py`
- Write the code for a Streamlit app that will ask the user for their username and password. We will use st.forms for this.
- On clicking login, the credentials should be checked.
- On clicking register, the user should go to the registration page from last week.



Hello! Please enter your log in info.

If this is your first time on my app then please click on the Register Button.

Please provide correct user name or click on register as new user



Hello! Please enter your log in info.

If this is your first time on my app then please click on the Register Button.

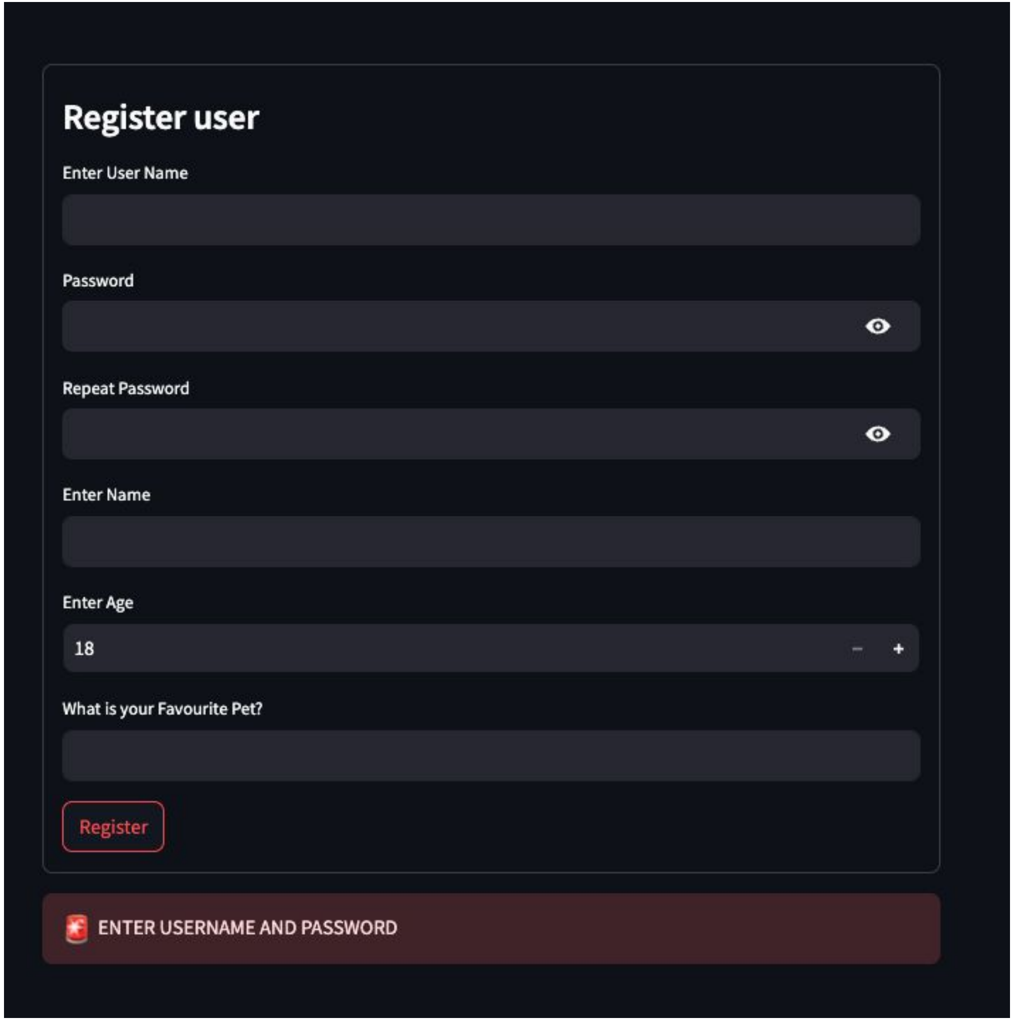
Username

shaq

Password

•

The username/password is not correct



Register user

Enter User Name

Password

Repeat Password

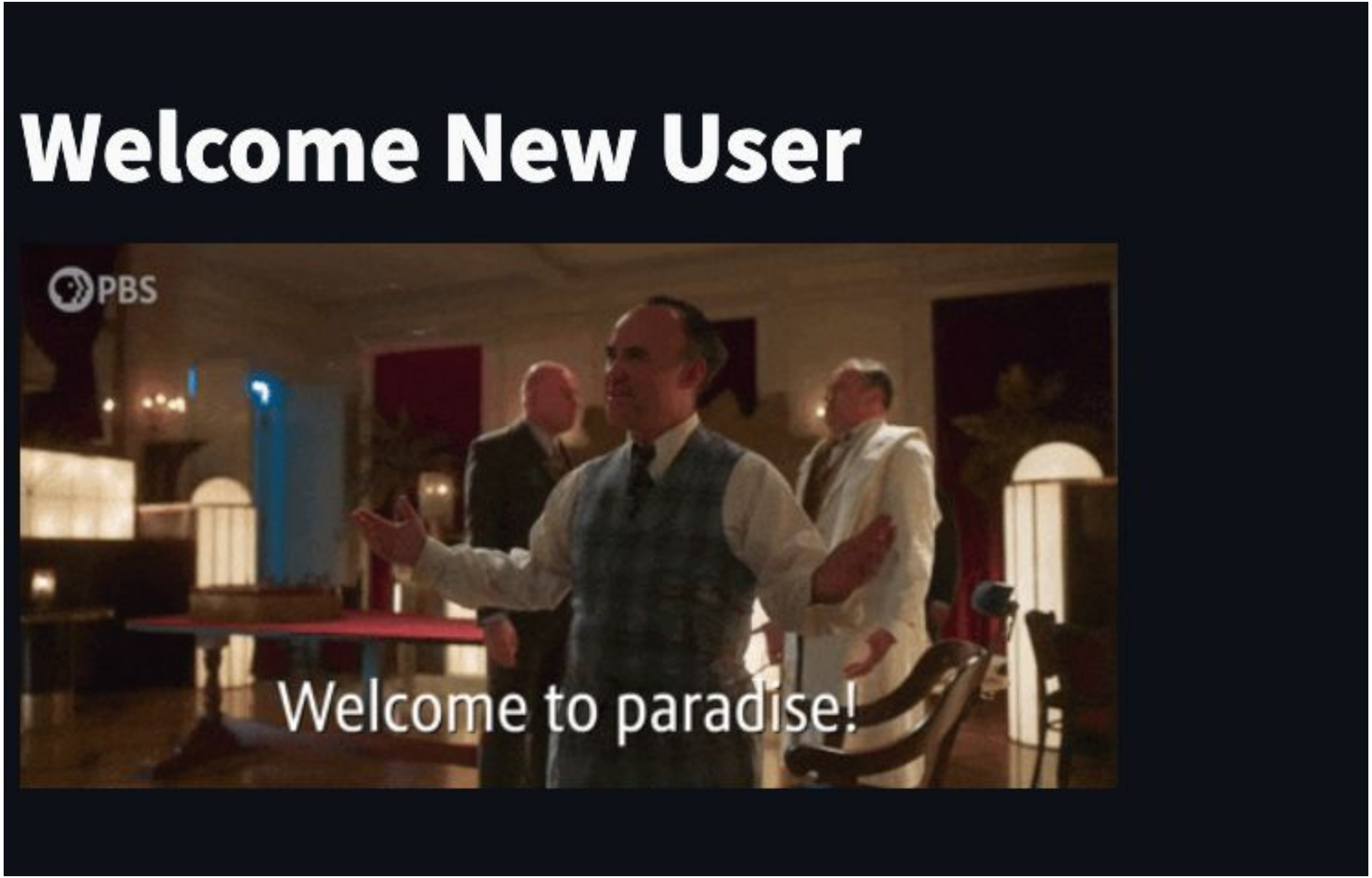
Enter Name

Enter Age

18

What is your Favourite Pet?

🚨 ENTER USERNAME AND PASSWORD

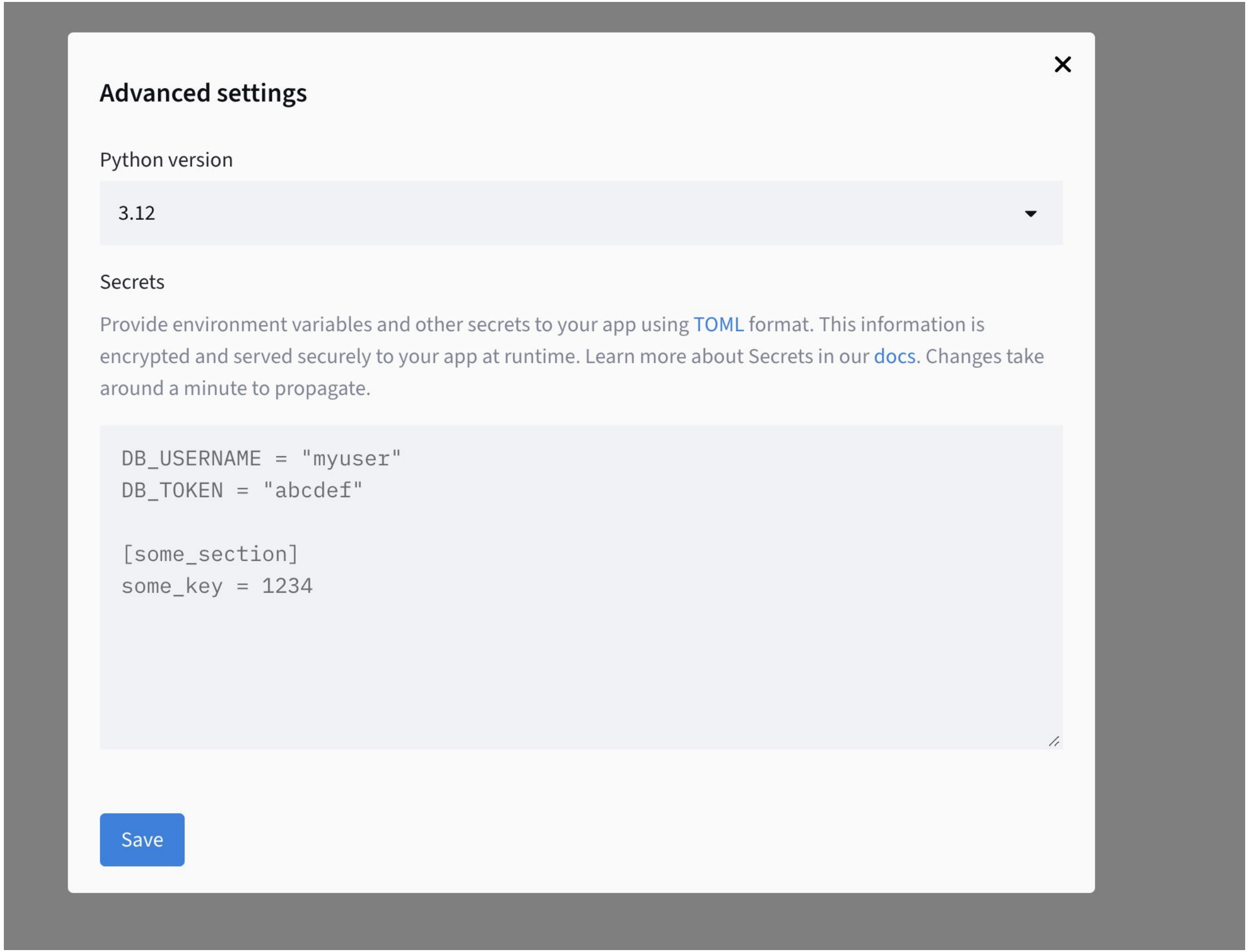


Welcome New User

Activity: Can you deploy your new app?

Do not upload your passwords to your github repo!

More on streamlit secrets: https://docs.streamlit.io/develop/api-reference/connections/secrets.toml



Advanced settings

Python version

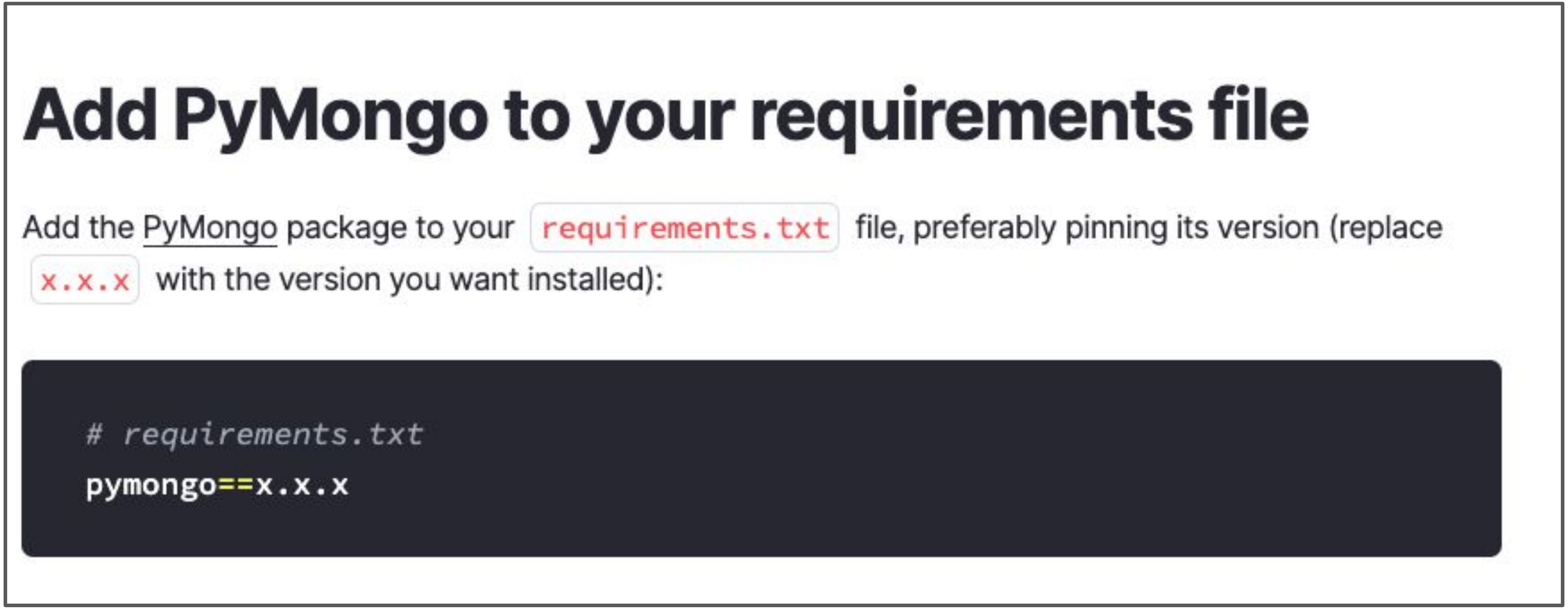
3.12

Secrets

Provide environment variables and other secrets to your app using TOML format. This information is encrypted and served securely to your app at runtime. Learn more about Secrets in our docs. Changes take around a minute to propagate.

```
DB_USERNAME = "myuser"
DB_TOKEN = "abcdef"

[some_section]
some_key = 1234
```



Add PyMongo to your requirements file

Add the PyMongo package to your `requirements.txt` file, preferably pinning its version (replace `x.x.x` with the version you want installed):

```
# requirements.txt
pymongo==x.x.x
```

All the code from class can be found here: https://github.com/shaq31415926/tech-basics/tree/main/tech_basics_two/08Lecture

I will share the hangman app code later on =)

Questions?!

References

- https://docs.streamlit.io/develop/api-reference/execution-flow/st.form
- https://www.mongodb.com/products/platform/atlas-database