**Batch:** T3

**Group 7:** 2019BTECS00001, 2019BTECS00014

**Date:** 15/2/2022

**Software Engineering Tools Lab**

**Assignment 2**

**Google colab:**

1. **Original author:** Google with the Jupyter Development Team

2**. Developers:** Google with the Jupyter Development Team

3. **Initial release:** October 2017

4**. Stable release:** April 2020

5. **Preview release:** 2017

6. **Repository (with cloud support ):** Google drive

7. **Written in (Languages):** IPython, ZeroMQ, Tornado, jQuery, Bootstrap, MathJax

8. **Operating System support:** Windows, Linux, macOS

9. **Platform portability:** It is entirely built on Google cloud so one just needs a browser to use colab.

10. **Available in (Total languages):** All the languages supported by chrome browser

11. **List of languages supported:** It supports over 40 programming languages. Only Python and Swift are officially supported. For other languages various packages are available.

12. **Type (Programming tool, integrated development environment etc.):** Integrated development environment (**a web IDE for python)**

13. **Website:** https://colab.research.google.com/

14. **Features:**

* Interactive tutorials to learn machine learning and neural networks.
* Write and execute Python 3 code without having a local setup.
* Execute terminal commands from the Notebook.
* Import datasets from external sources such as Kaggle.
* Save your Notebooks to Google Drive.
* Import Notebooks from Google Drive.
* Free cloud service, GPUs and TPUs.
* Integrate with PyTorch, Tensor Flow, Open CV.
* Import or publish directly from/to GitHub.

15. **Size (in MB, GB etc.):** We can run these notebooks for not more than 12 hours and we can use only **12 GB RAM**.

16. **Privacy and Security:** Google Colab is **safe**, at least as safe as our private Google Doc is. No one can access our own private Colab notebooks. And Google has the incentive to make it as safe as possible for their reputation.

17. **Type of software (Open source/License):** Google Colaboratory or Google Colab is **a free Jupyter Notebook environment**. It is a free cloud-based service by Google which means we don't have to pay anything. And **Jupyter** is the open source project on which Colab is based.

18. **If License- Provide details.:**   —

19. **Latest version:** Google Colab uses **Python 3.6.** **9**.

20. **Cloud support (Yes/No):** Yes

21. **Applicability:**

* Getting started with TensorFlow.
* Developing and training neural networks.
* Experimenting with TPUs.
* Disseminating AI research.
* Creating tutorials.

22. **Drawbacks (if any):**

**Closed-Environment:** [Anyone can use Google Colab](https://analyticsindiamag.com/explained-how-to-access-jupyterlab-on-google-colab/) to write and run arbitrary Python code in the browser. However, it is still a relatively closed environment, as machine learning practitioners can only run the python package already pre-added on the Colab. There is no way that one can add their own python package and start running the code. Hence, the platform can provide common tools but is not suitable for specialisation.

**Repetitive Tasks:** Imagine one has to repeat the same set of actions repeatedly to execute a task — not only will it be exhausting, but it will also consume a lot of time. Similarly, for every new session in the Google Colab, a programmer must install all of the specific libraries that aren’t included with the standard Python package.

**No Live-Editing:** [Writing a code](https://analyticsindiamag.com/how-to-run-a-development-server-for-flask-web-applications-using-google-colab/) and sharing the same with your partner or a team allows you to collaborate. However, the option for live editing is completely missing in Google Colab, which restricts two people to write, or edit codes at the same time. Hence, it further leads to a lot of back and forth re-sharing. Additionally, this feature is provided by its other competitors, including CoCalc.

**Saving & Storage Problems:** Uploaded files are removed when the session is restarted because Google Colab does not provide a persistent storage facility. So, if the device is turned off, the data can get lost, which can be a nightmare for many. Moreover, as one uses the current session in Google Storage, a downloaded file that is required to be used later needs to be saved before the session’s expiration. In addition to that, one must always be logged in to their Google account, considering all Colaboratory notebooks are stored in Google Drive.

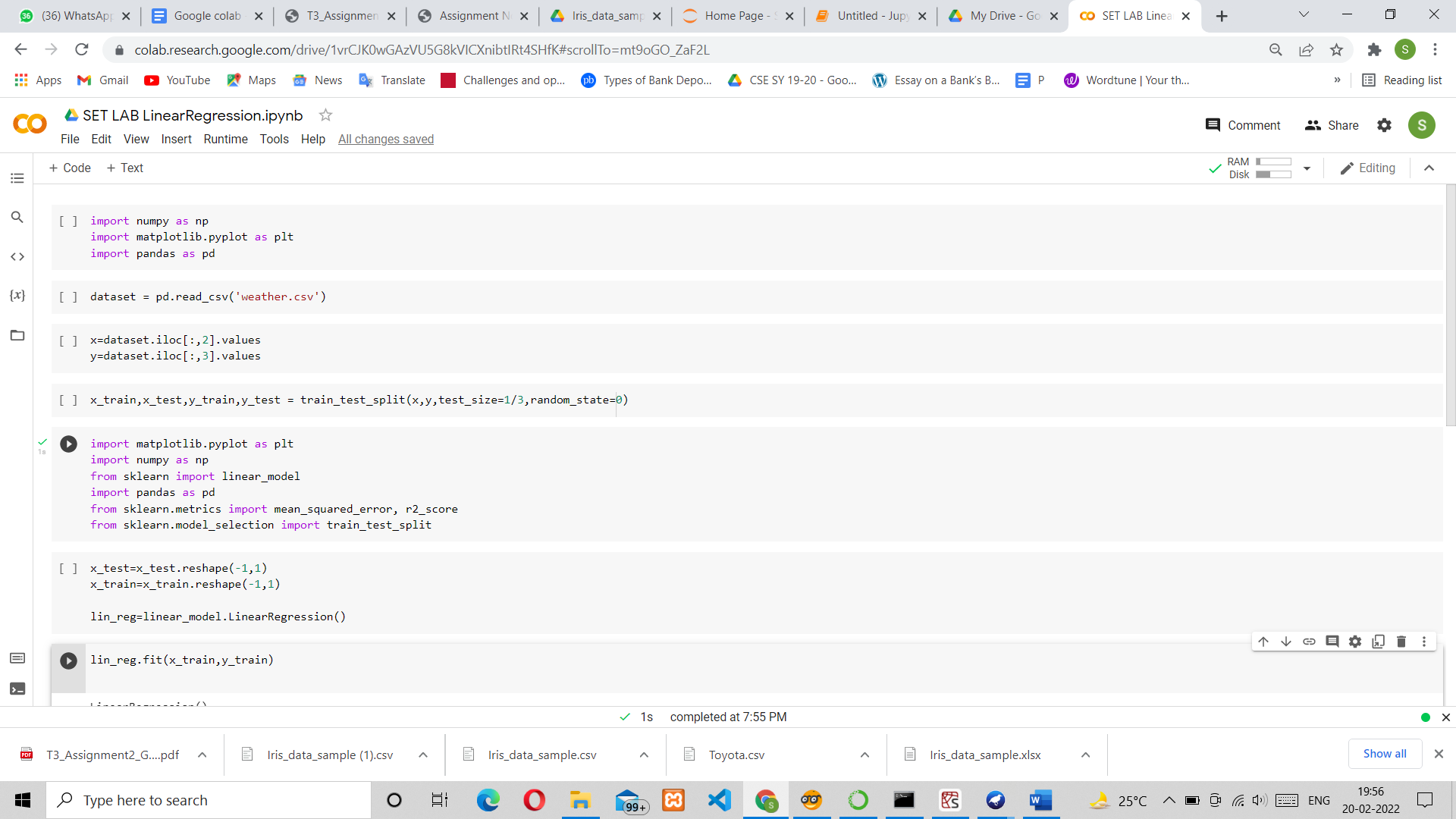
**Limited Space & Time:** The Google Colab platform stores files in Google Drive with a free space of 15GB; however, working on bigger datasets requires more space, making it difficult to execute. This, in turn, can hold most of the complex functions to execute.

Google Colab allows users to run their notebooks for at most 12 hours a day, but in order to work for a longer period of time, users need to access the paid version, i.e. Colab Pro which allows programmers to stay connected for 24 hours. Finally, the less talked about drawback of the platform is its inability to execute codes or run properly on a mobile device.

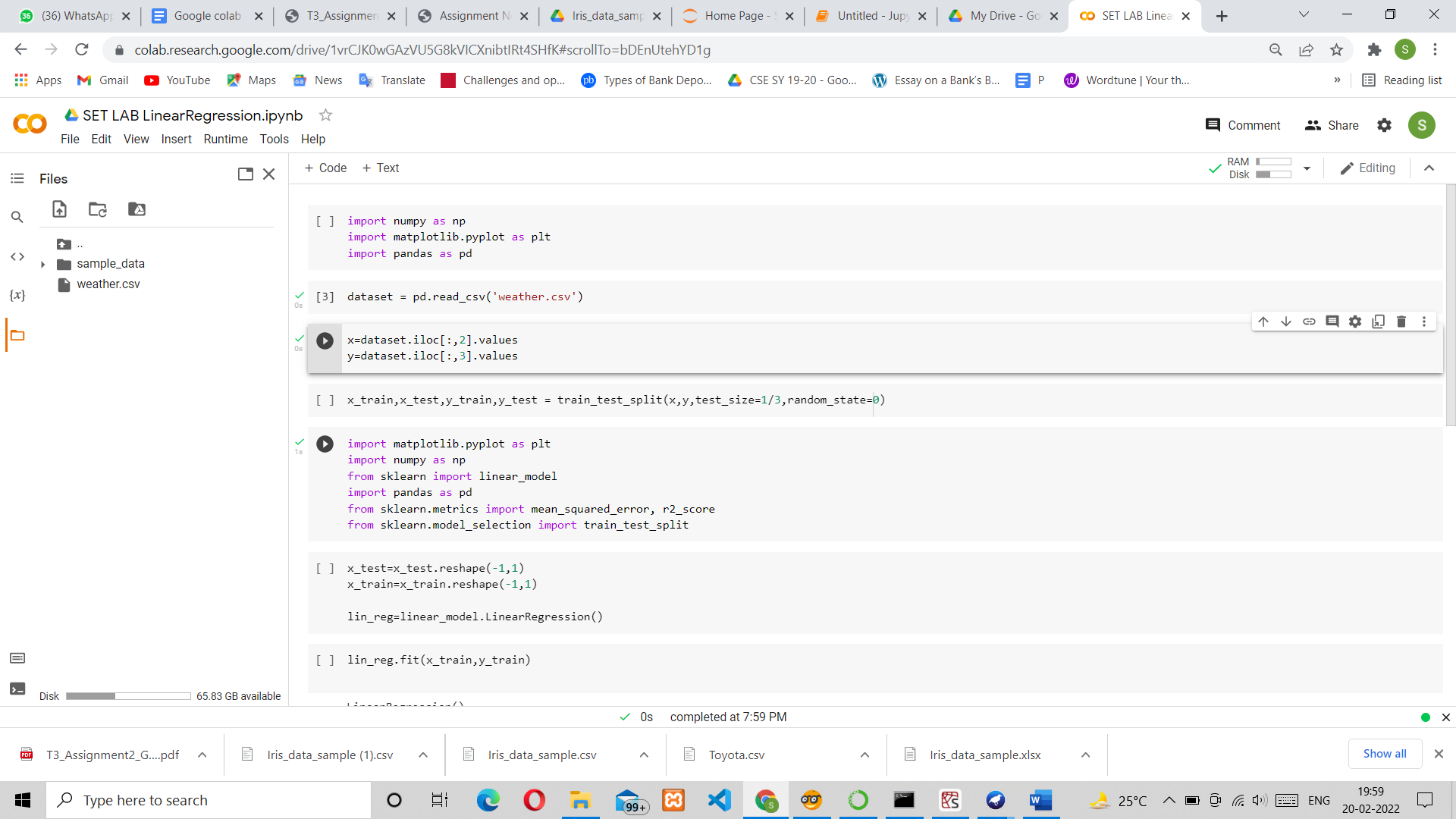
**Q.2 Implement linear regression problem using Google Colab (Perform preprocessing, training and testing)**

Dataset: <https://archive.ics.uci.edu/ml/datasets/KDD+Cup+1998+Data>

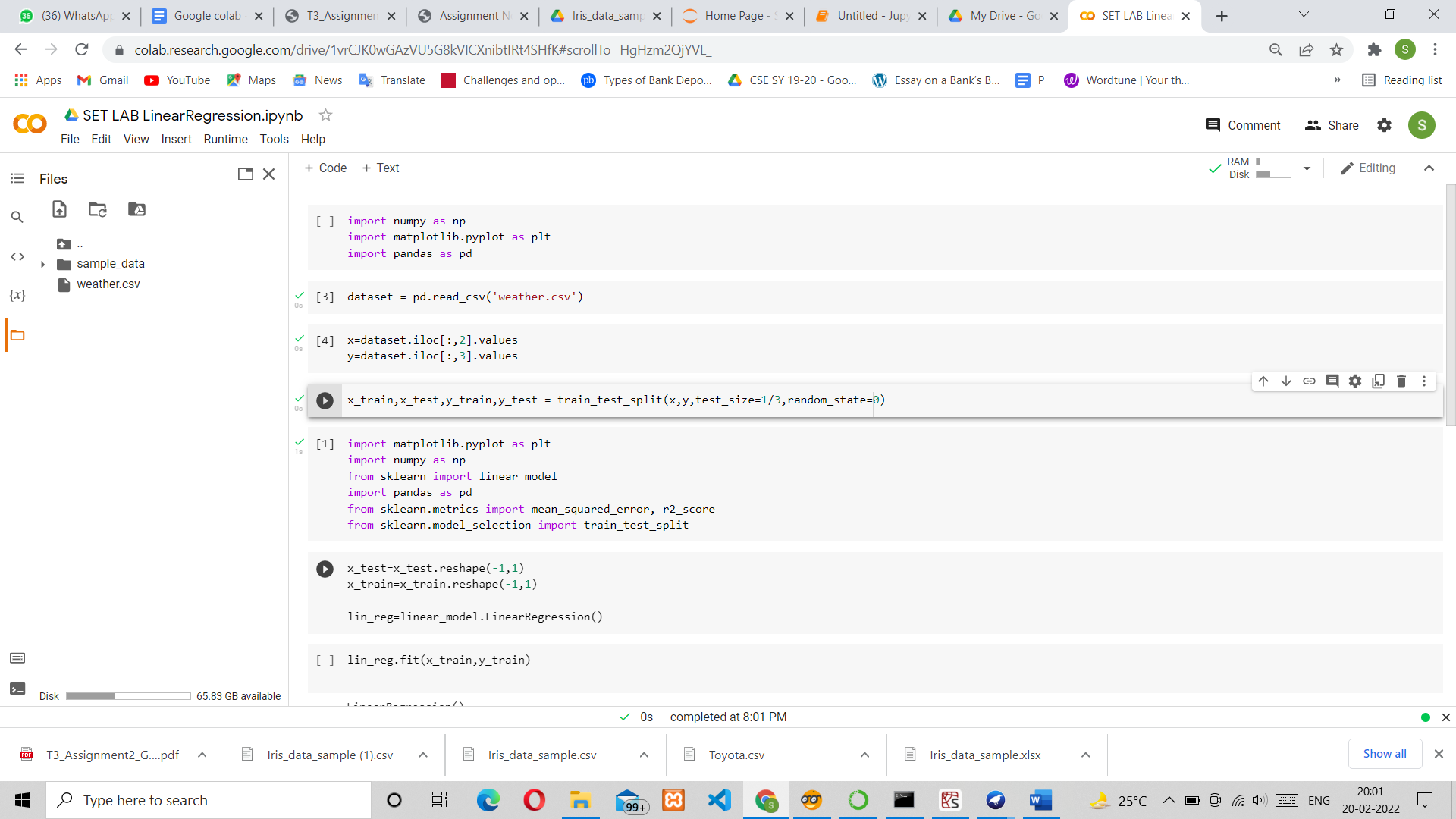
Libraries:



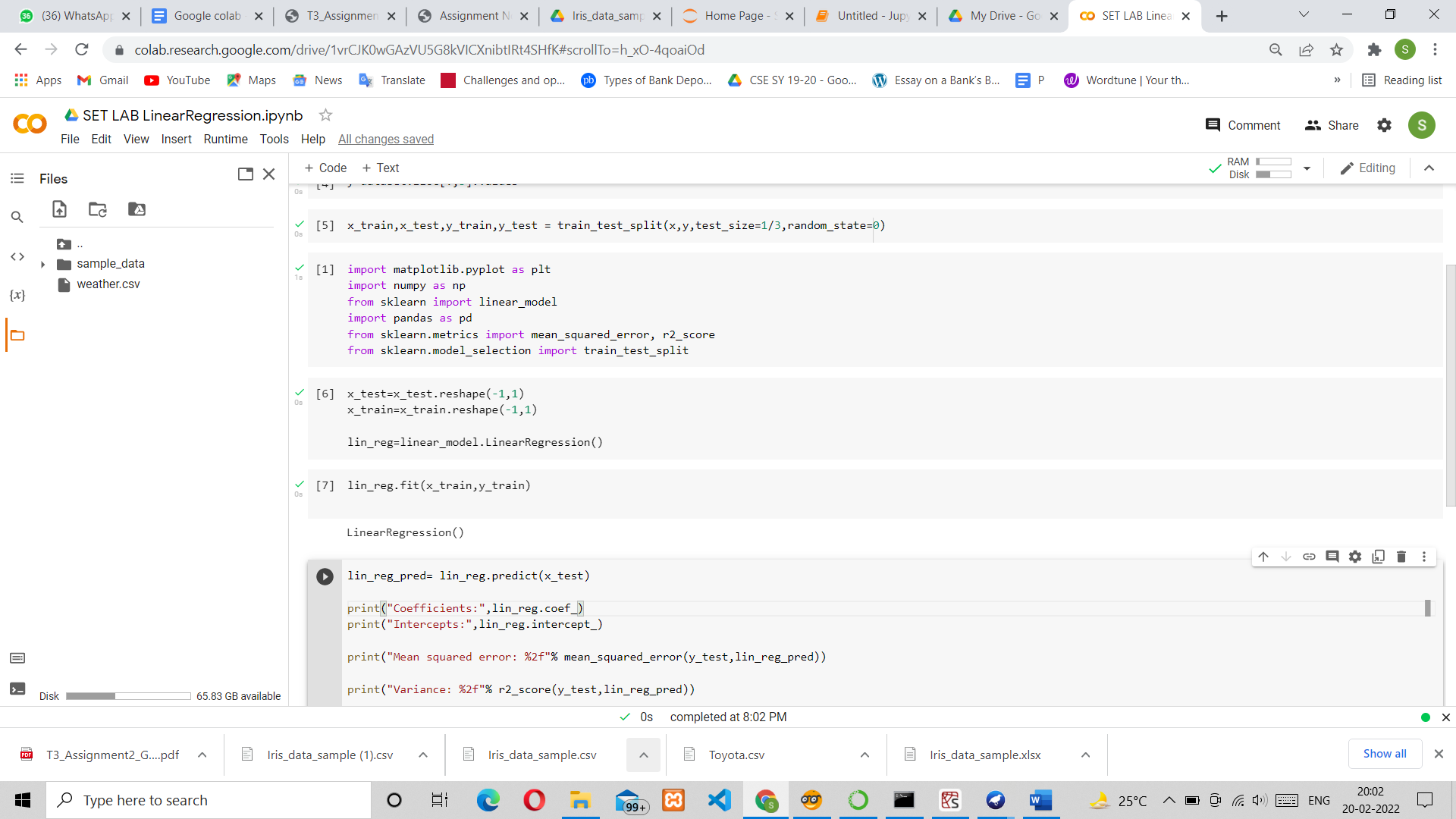
Dataset and setting Variables:



Splitting into Training and Testing:



Linear Regression Model:



Prediction:

