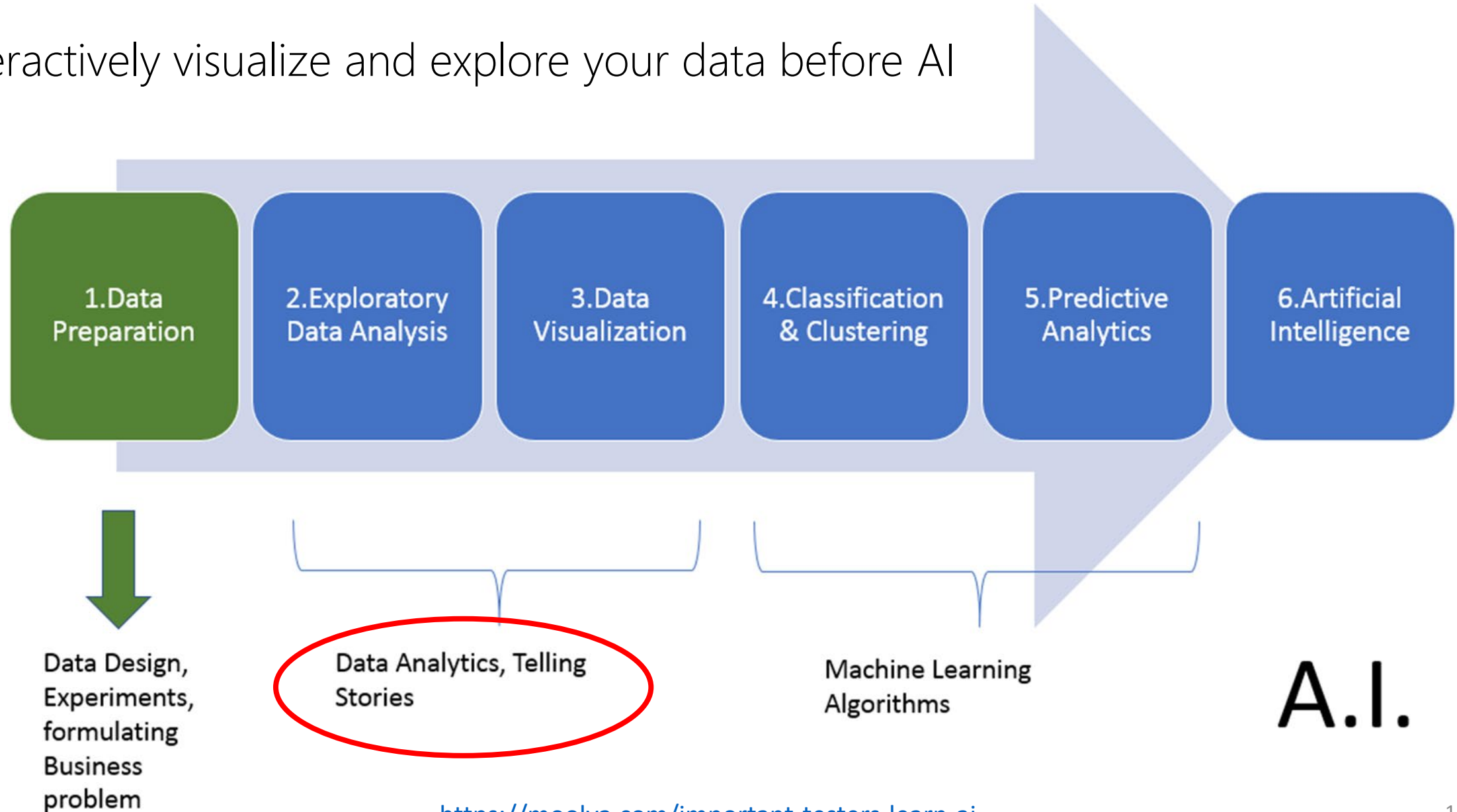
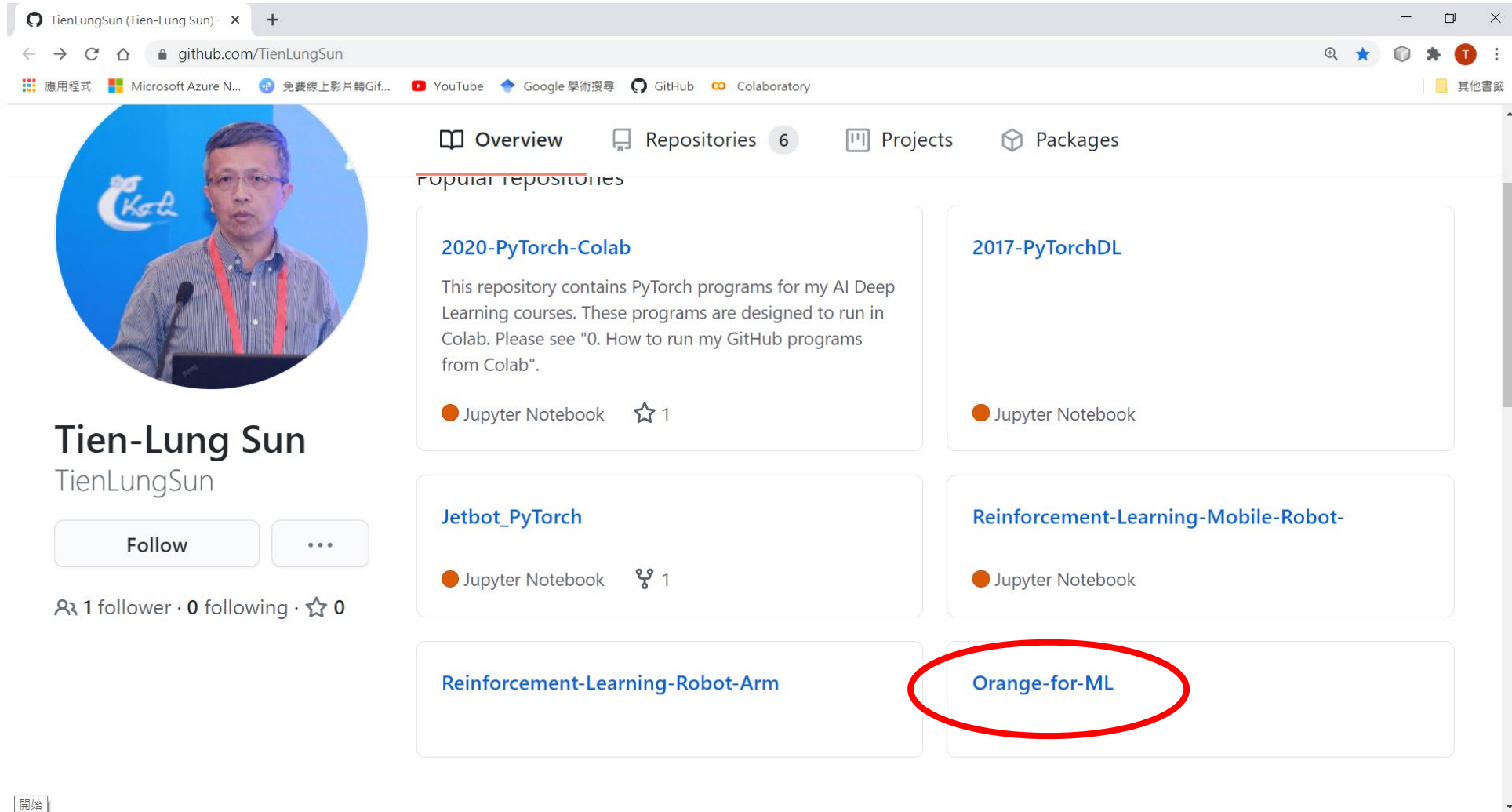


HI before AI

Interactively visualize and explore your data before AI



Download data files from my GitHub



The screenshot shows a web browser window displaying the GitHub profile of TienLungSun. The browser's address bar shows the URL `github.com/TienLungSun`. The profile page includes a circular profile picture of a man with glasses, the name 'Tien-Lung Sun', and the username 'TienLungSun'. Below the name is a 'Follow' button and a menu icon. The statistics show '1 follower · 0 following · 0 stars'. The 'Repositories' tab is selected, showing a list of repositories. The repository 'Orange-for-ML' is circled in red.

Repository Name	File Type	Stars
2020-PyTorch-Colab	Jupyter Notebook	1
2017-PyTorchDL	Jupyter Notebook	0
Jetbot_PyTorch	Jupyter Notebook	1
Reinforcement-Learning-Mobile-Robot-	Jupyter Notebook	0
Reinforcement-Learning-Robot-Arm	Jupyter Notebook	0
Orange-for-ML	Jupyter Notebook	0

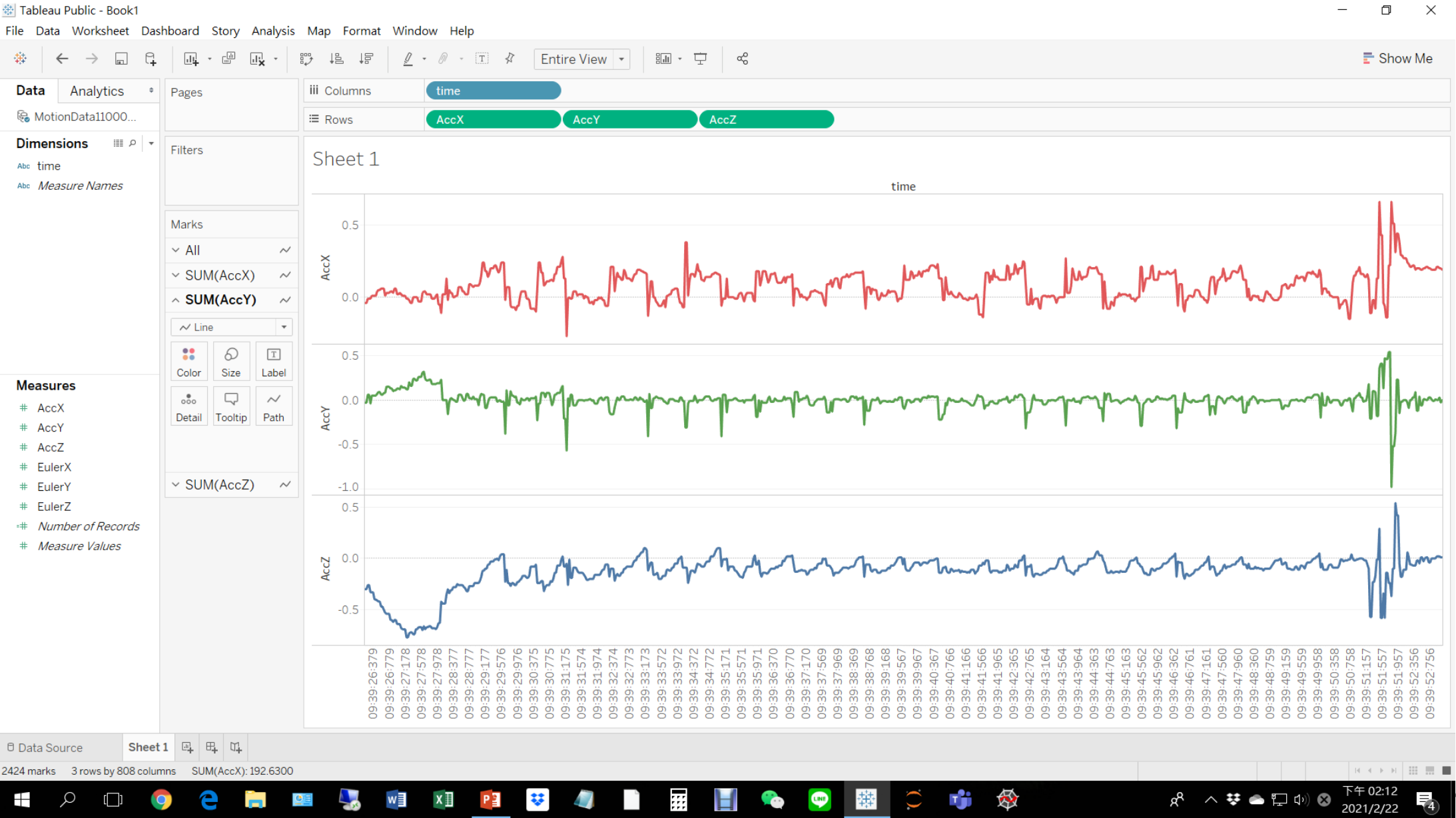
<https://github.com/TienLungSun>

Interactively visual exploration – (1) Tableau Public



<https://public.tableau.com/s/>

Tableau Public



Practice – Tableau public

1. Download and install Tableau Public
2. Visualize the motion data file
3. Search Kaggle (<https://www.kaggle.com/>) to find a sensor data file (see next slide)
4. Use Tableau public to visualize the data file

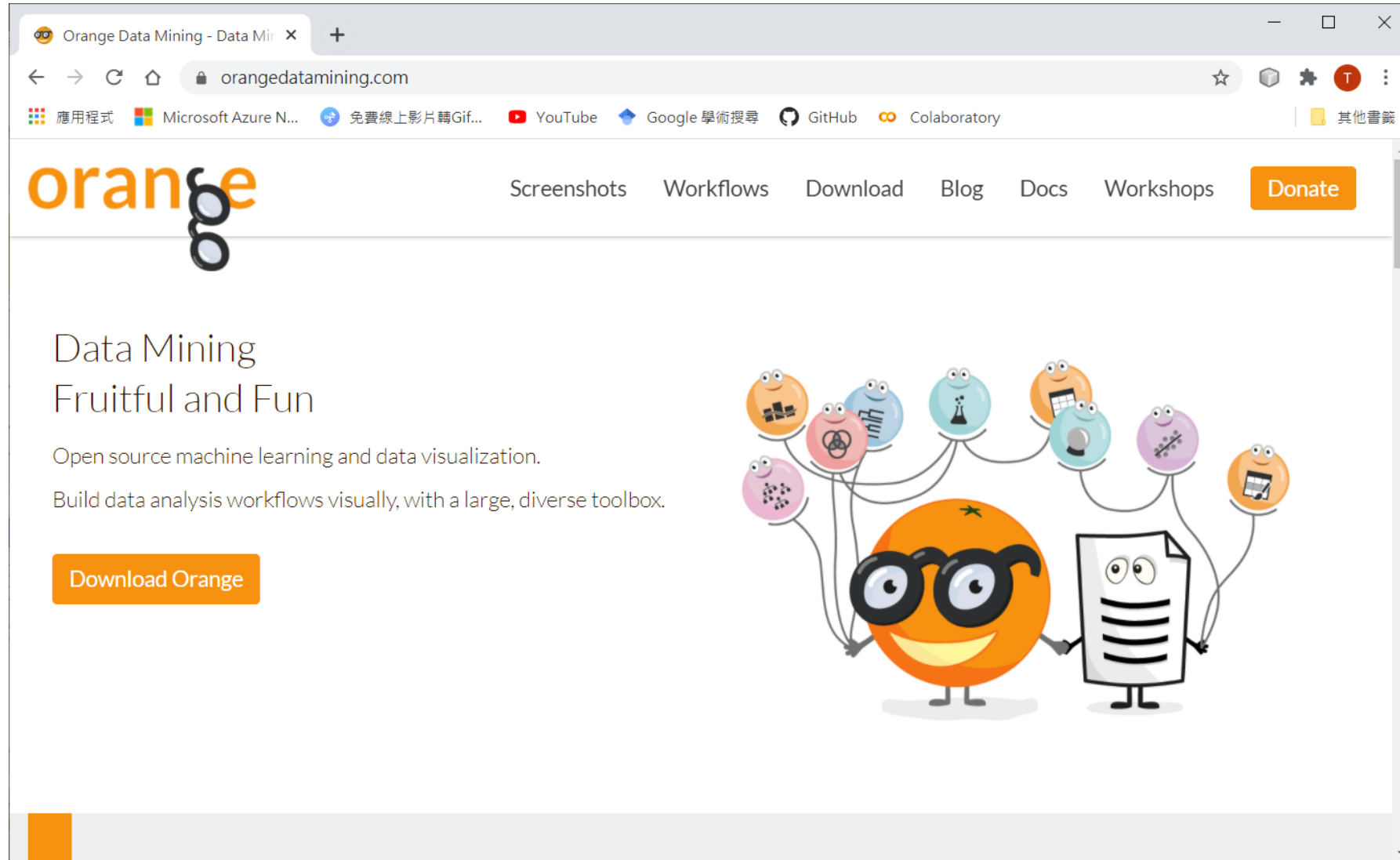
Kaggle

The screenshot shows the Kaggle website interface. In the left sidebar, the 'Data' link is circled in red. In the main content area, the search bar is also circled in red, containing the text 'motion sensor'. Below the search bar, there are tabs for 'Datasets', 'Tasks', 'Computer Science', 'Education', 'Classification', 'Computer Vision', 'NLP', and 'Data Visualization'. The 'Datasets' tab is selected, showing a list of 70 datasets. The top three datasets are:

- MotionSense Dataset : Smartphone Sensor Data - HAR**
Mohammad Malekzadeh · Updated 3 years ago
Usability 7.6 · 361 Files (CSV) · 72 MB
171 votes, Silver medal
- Emotions Sensor Data Set**
jon_bill · Updated 2 years ago
Usability 8.2 · 1 File (CSV) · 27 KB
122 votes, Bronze medal
- Hand Gesture Recognition Database**
GT1 · Updated 3 years ago
Usability 7.5 · 20000 Files (other) · 2 GB
397 votes, Silver medal

At the bottom of the page, there is a cookie consent banner and a Windows taskbar with various application icons.

Interactively visual exploration – (2) Orange



<https://orangedatamining.com/>

Orange

The screenshot displays the Orange3 data mining software interface. The main workflow window shows a sequence of three widgets: CSV File Import, Select Columns, and t-SNE, connected by 'Data' links. The Select Columns widget is currently active, showing a list of available variables on the left and a list of selected features on the right. The features list includes t1, t2, t3, t4, t5, t6, n1, and n2, which are circled in red. The target variable is set to y1, also circled in red. The t-SNE widget settings are visible on the right, showing a perplexity of 30, PCA components of 15, and a color mapping to y1. The resulting t-SNE plot shows two distinct clusters of data points, colored blue (0) and red (1).

Workflow: CSV File Import → Data → Select Columns → Data → t-SNE

Select Columns Widget:

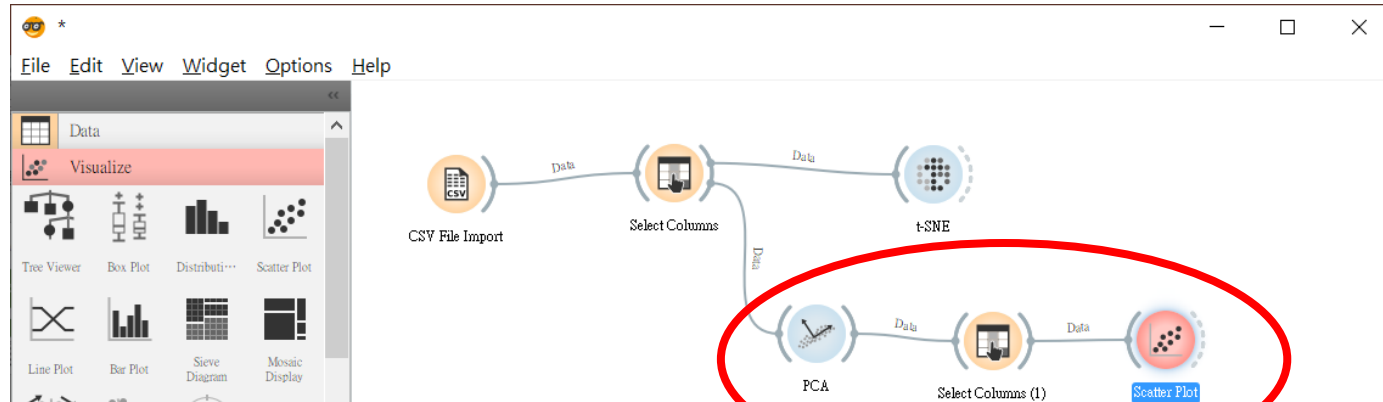
- Available Variables:** Gender, TUG, BBS, y2, No, age, time, Step frequency, Steps
- Features:** t1, t2, t3, t4, t5, t6, n1, n2
- Target Variable:** y1
- Meta Attributes:**

t-SNE Widget Settings:

- Perplexity: 30
- Exaggeration: 1
- PCA components: 15
- Normalize data: ☒
- Color: y1
- Shape: (Same shape)
- Size: (Same size)
- Label: (No labels)
- Symbol size: [Slider]
- Opacity: [Slider]
- Jittering: [Slider]
- Show color regions: ☐
- Show legend: ☒
- Zoom/Select: [Buttons]

t-SNE Plot: A scatter plot showing two clusters of data points. The legend indicates that blue points represent class 0 and red points represent class 1.

Orange



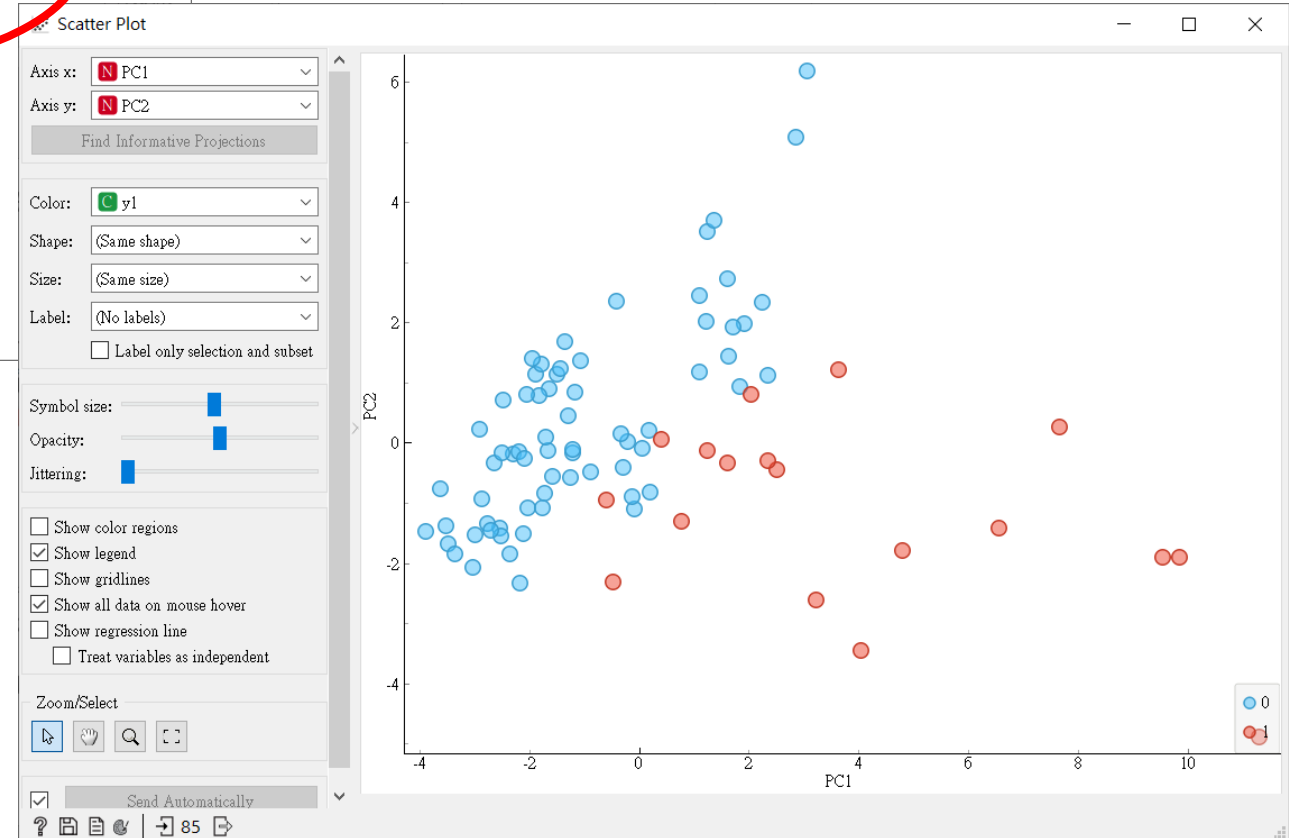
Scatter Plot

Interactive scatter plot visualization with intelligent data visualization enhancements.

[more...](#)



The 'Select Columns (1)' widget interface is shown. It features a list of 'Available Variables' on the left, including t1 through t6 and n1 through pz. A 'Filter' box is present above this list. On the right, the 'Features' section contains a list of selected variables: PC1 and PC2, which are circled in red. Below this, the 'Target Variable' is set to y1. The 'Meta Attributes' section is empty. At the bottom, there are 'Reset' and 'Send Automatically' buttons.



Practice – Orange

1. Download and install Orange
2. Visualize the 3M TUG data file
3. Search Kaggle to find a classification data file
4. Use Orange to visualize the high dimensional data

Kaggle

Find Open Datasets and Machi x +

kaggle.com/datasets?search=classification

應用程式 Google 學術搜尋 YouTube Colaboratory GitHub TienLungS... 李弘毅 ML

kaggle

Home

Compete

Data

Code

Communities

Courses

More

Search

Sign In Register

+ New Dataset Your Work

Search classification Filters

Datasets Tasks Computer Science Education Classification Computer Vision NLP Data Visualization

209 Tasks See All

Species Classification
Sachin Sharma · 48 Submissions

Classification Problem
Pratham Tripathi · 25 Submissions

Mushroom Classification using Random Forest Classifier
Sanchita Karmakar · 34 Submissions

Fetal Health Classification
Larxel · 26 Submissions

Iris Species
UCI Machine Learning · Usa...

Drug Classification
Pratham Tripathi · Usability 1...

Mushroom Classifica...
UCI Machine Learning · Usa...

Fetal Health Classific...
Larxel · Usability 10.0

3,010 Datasets Hottest

We use cookies on Kaggle to deliver our services, analyze web traffic, and improve your experience on the site. By using Kaggle, you agree to our use of cookies. Got it Learn more

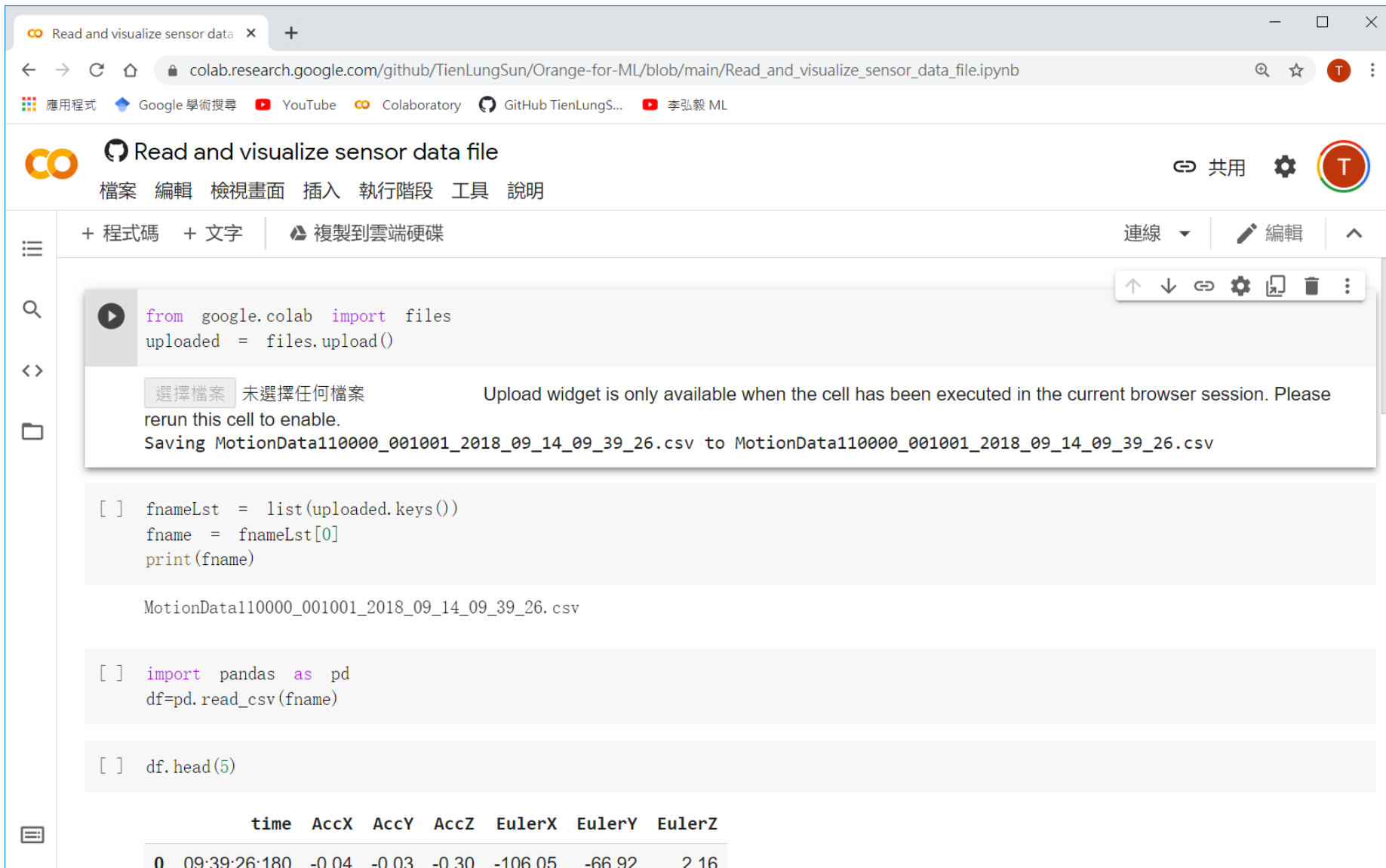
Windows taskbar: Chrome, Edge, File Explorer, Mail, Word, Excel, PowerPoint, Teams, OneDrive, Calculator, WeChat, Slack, VS Code, Docker, Zoom, and system tray with date/time (下午 03:26 2021/2/22).

Data visualization – (3) Python coding

The screenshot shows the Google Colaboratory web interface. At the top, the browser address bar displays `colab.research.google.com/notebooks/intro.ipynb#recent=true`. The main interface has a sidebar on the left with a '目錄' (Table of Contents) section containing links like '開始使用' (Get started), '數據資料學' (Data science), '機器學習' (Machine learning), and '其他資源' (Other resources). The main area features a modal window for connecting to a GitHub repository. This modal has a tabbed interface with '範例' (Examples), '最近' (Recent), 'Google 雲端硬碟' (Google Drive), 'GitHub' (highlighted with a red circle), and '上傳' (Upload). The 'GitHub' tab contains a search bar with the text '輸入 GitHub 網址或依機構或使用者搜尋' (Enter GitHub URL or search by organization or user). The search results show 'TienLungSun' (highlighted with a red circle) as the selected user. Below the search bar, there are dropdown menus for '存放區' (Repository) showing 'TienLungSun/Orange-for-ML' and '分支版本' (Branch version) showing 'main'. A list of notebooks is displayed below, including 'Read_and_visualize_high_dimensional_data.ipynb' and 'Read_and_visualize_sensor_data_file.ipynb'. At the bottom of the modal, there are buttons for '新增筆記本' (Add notebook) and '取消' (Cancel). The bottom of the screen shows a Windows taskbar with various application icons and a system clock indicating '下午 03:11 2021/2/22'.

<https://colab.research.google.com/>

Data visualization with Python



Read and visualize sensor data file

檔案 編輯 檢視畫面 插入 執行階段 工具 說明

+ 程式碼 + 文字 複製到雲端硬碟

連線 編輯

```
from google.colab import files
uploaded = files.upload()
```

選擇檔案 未選擇任何檔案

Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.

Saving MotionData110000_001001_2018_09_14_09_39_26.csv to MotionData110000_001001_2018_09_14_09_39_26.csv

```
[ ] fnameLst = list(uploaded.keys())
    fname = fnameLst[0]
    print(fname)

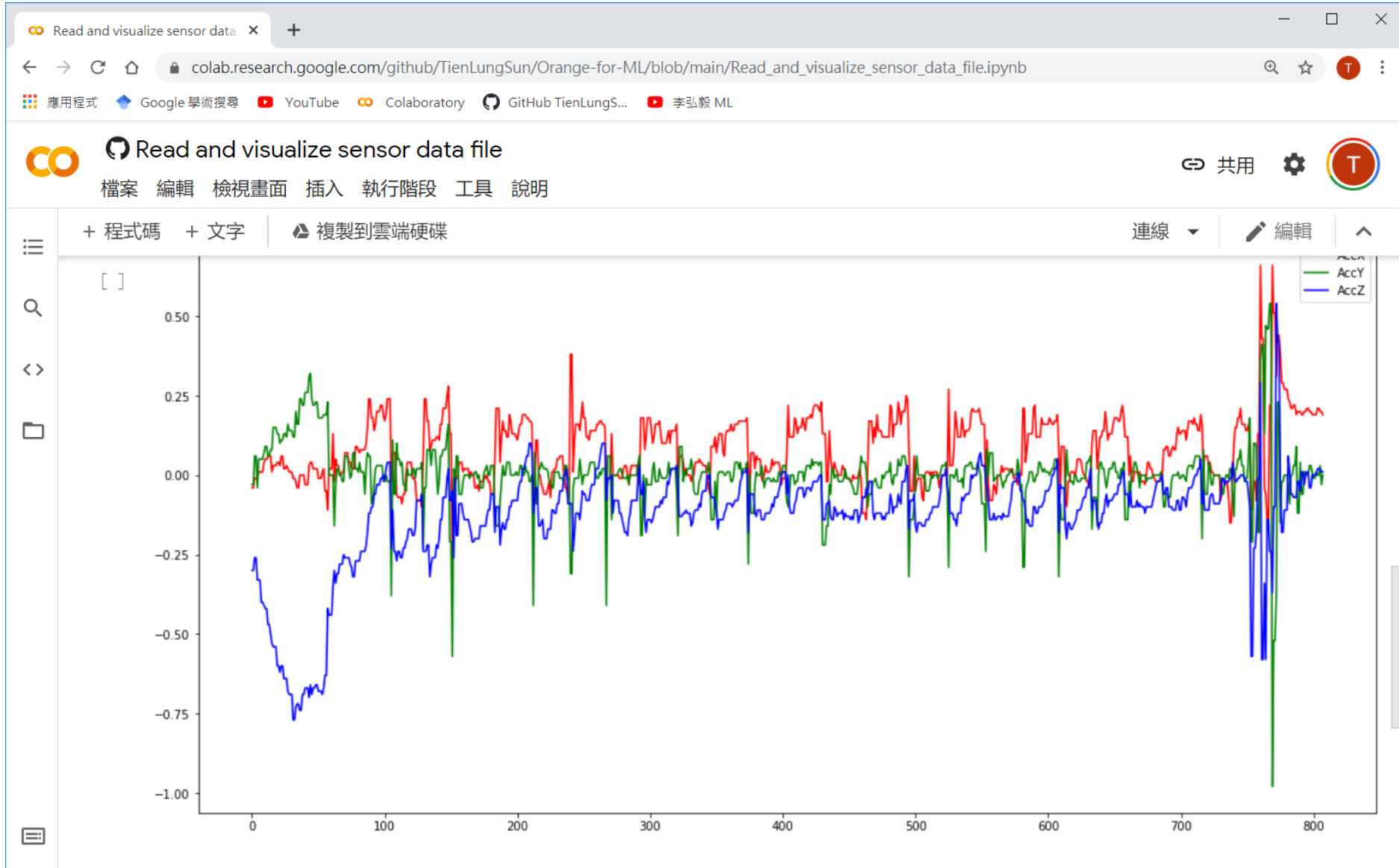
MotionData110000_001001_2018_09_14_09_39_26.csv
```

```
[ ] import pandas as pd
    df=pd.read_csv(fname)
```

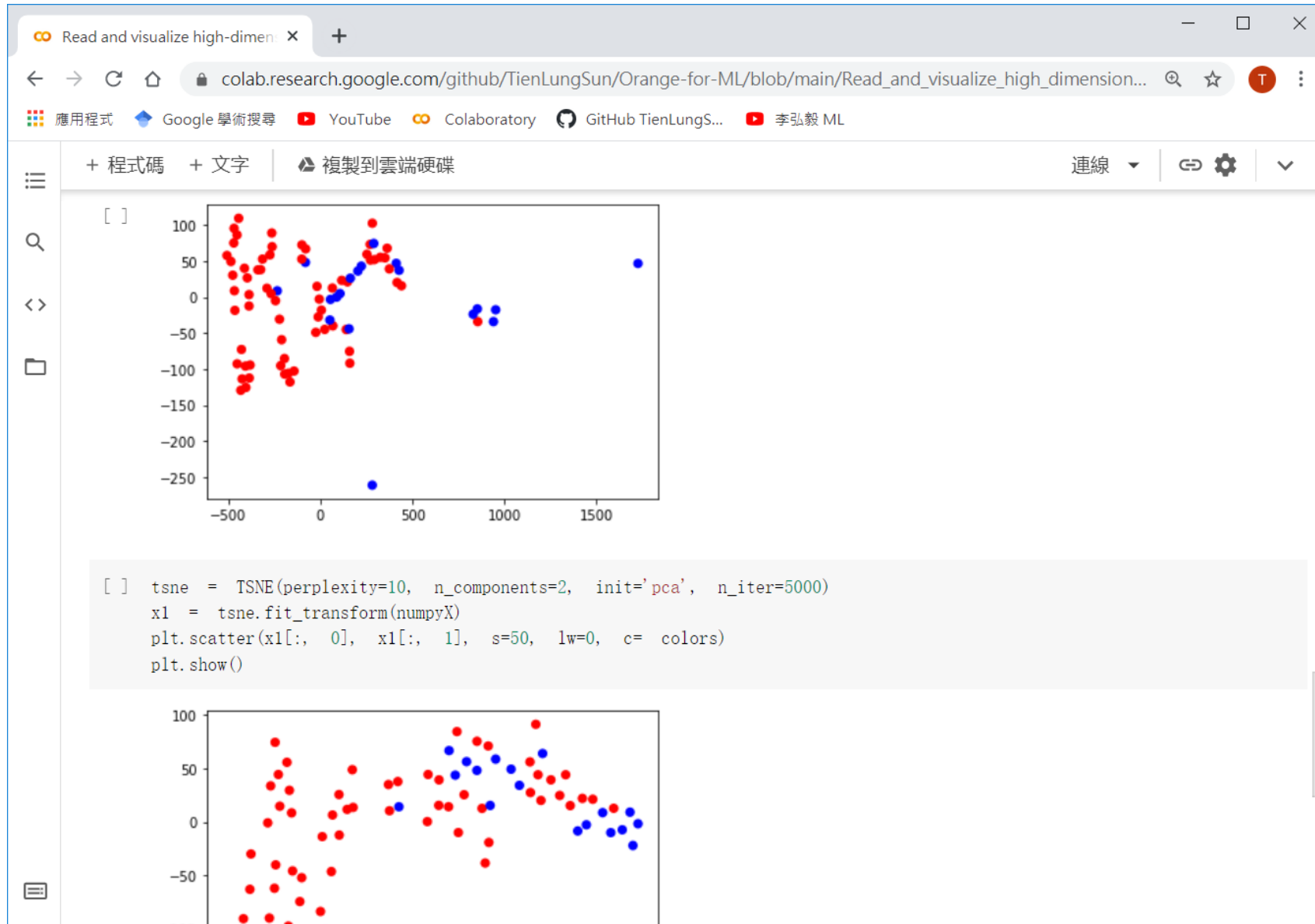
```
[ ] df.head(5)
```

	time	AccX	AccY	AccZ	EulerX	EulerY	EulerZ
0	09:39:26:180	-0.04	-0.03	-0.30	-106.05	-66.92	2.16

Data visualization with Python



Data visualization with Python



Practice – Python coding

1. Log in to Colab
2. Run python code to visualize the motion sensor data file
3. Run python code to visualize the 3M TUG data file
4. Run python code to visualize the two data files you download from Kaggle