

CSVファイルの生成

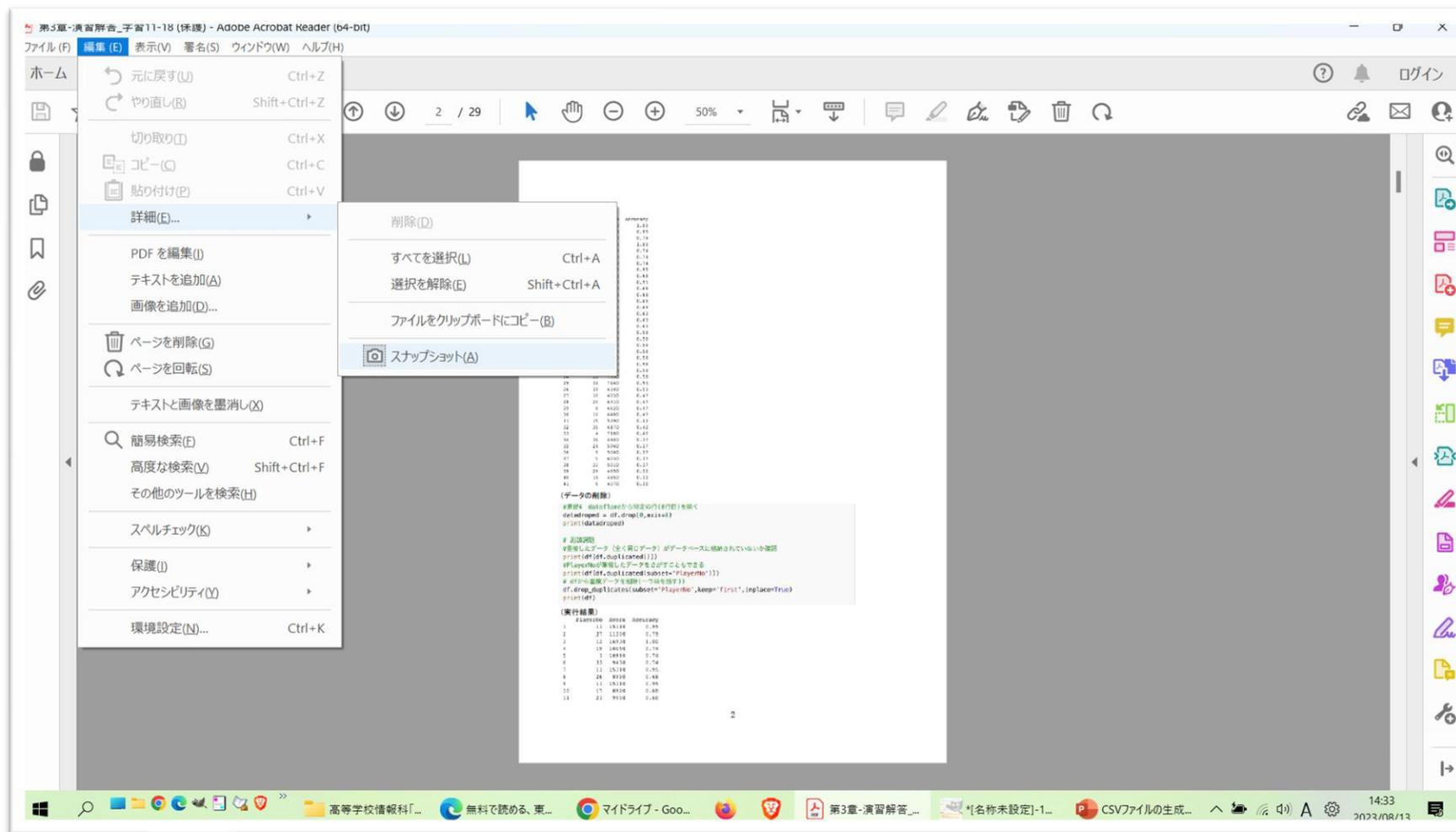
令和5(2023)年10月2日

CSVファイルの生成

- 方針
 1. 演習問題解答の画像ファイルから文字起こし
 - 演習問題解答のデータ部分をスナップショットにより画像化
 - Googleドライブに画像ファイルを置く
 - Googleドキュメントで読み込んで文字起こし
 2. テキスト形式ファイルを読み込んでPythonでCSV形式ファイルとして出力する

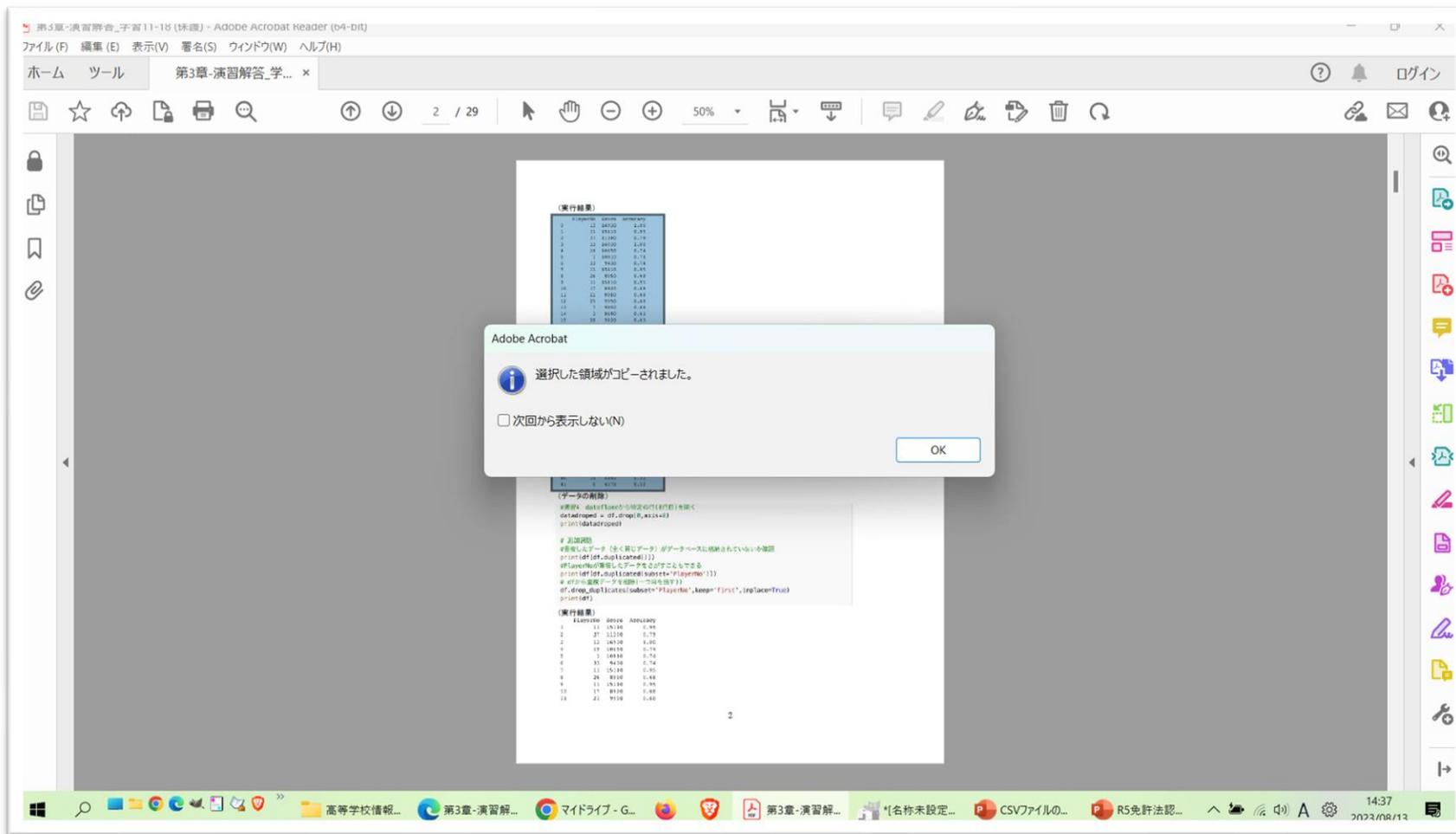
CSVファイルの生成

- 第3章-演習解答 学習11-18 (PDF:2.5MB)



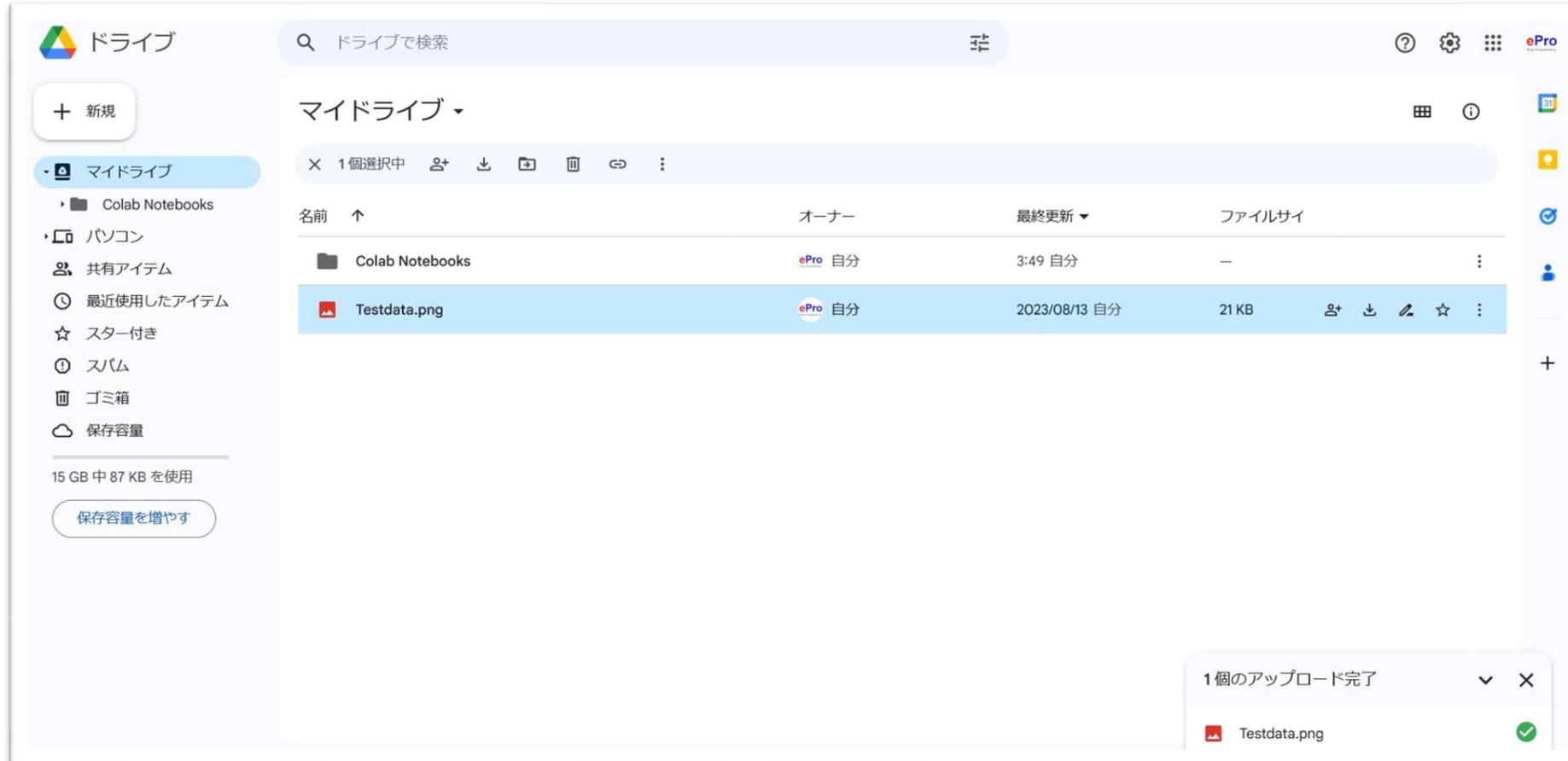
CSVファイルの生成

- 範囲を指定してコピー



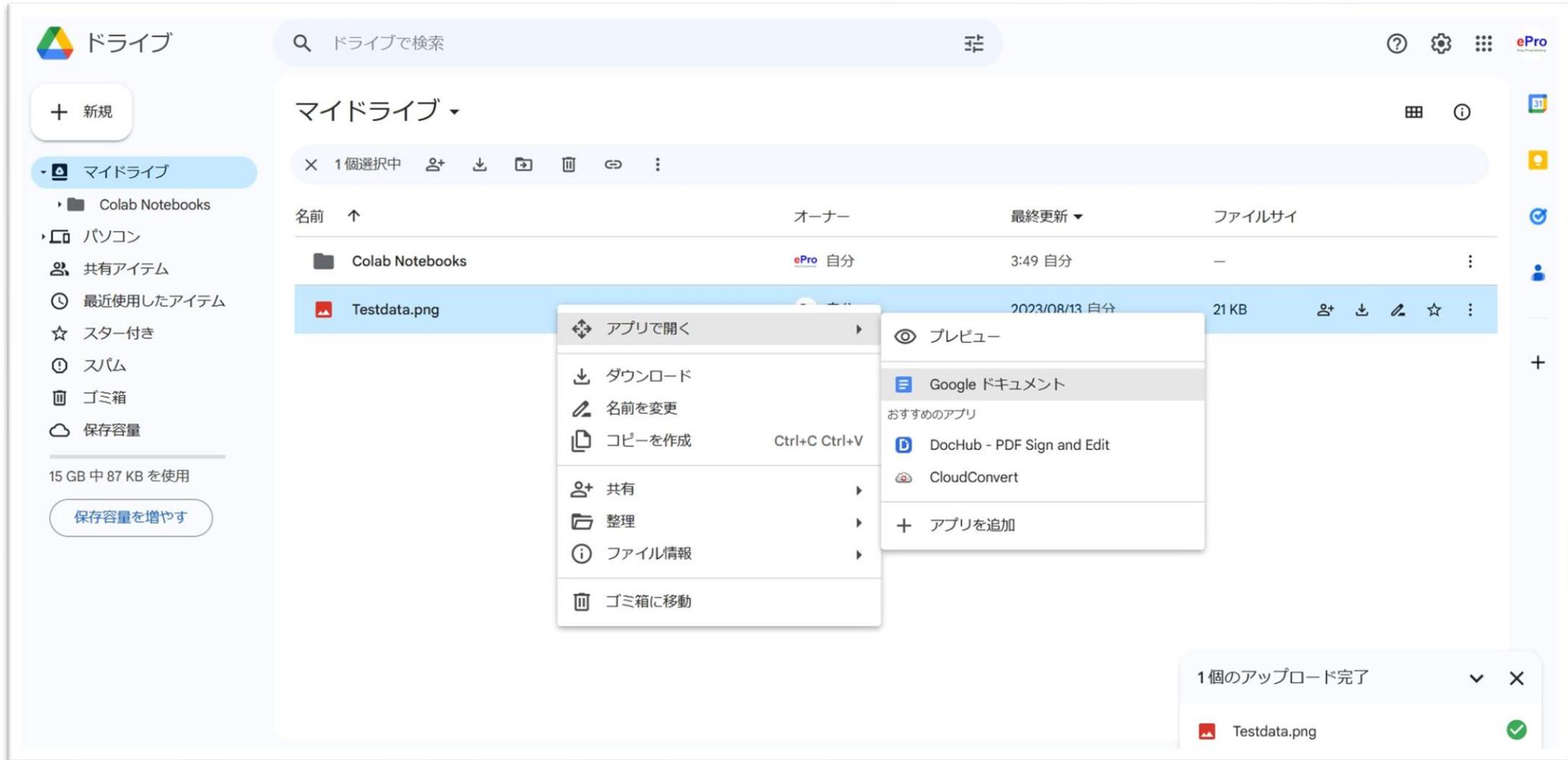
CSVファイルの生成

- Gimpに貼り付けてPNG形式で保存したものをGoogleドライブにアップロード



CSVファイルの生成

- Google ドキュメントで開く



CSVファイルの生成

- 下にスクロール

Testdata ☆ 📁 🌐
ファイル 編集 表示 挿入 表示形式 ツール 拡張機能 ヘルプ

100% 標準デキ... Arial 11 B I U A

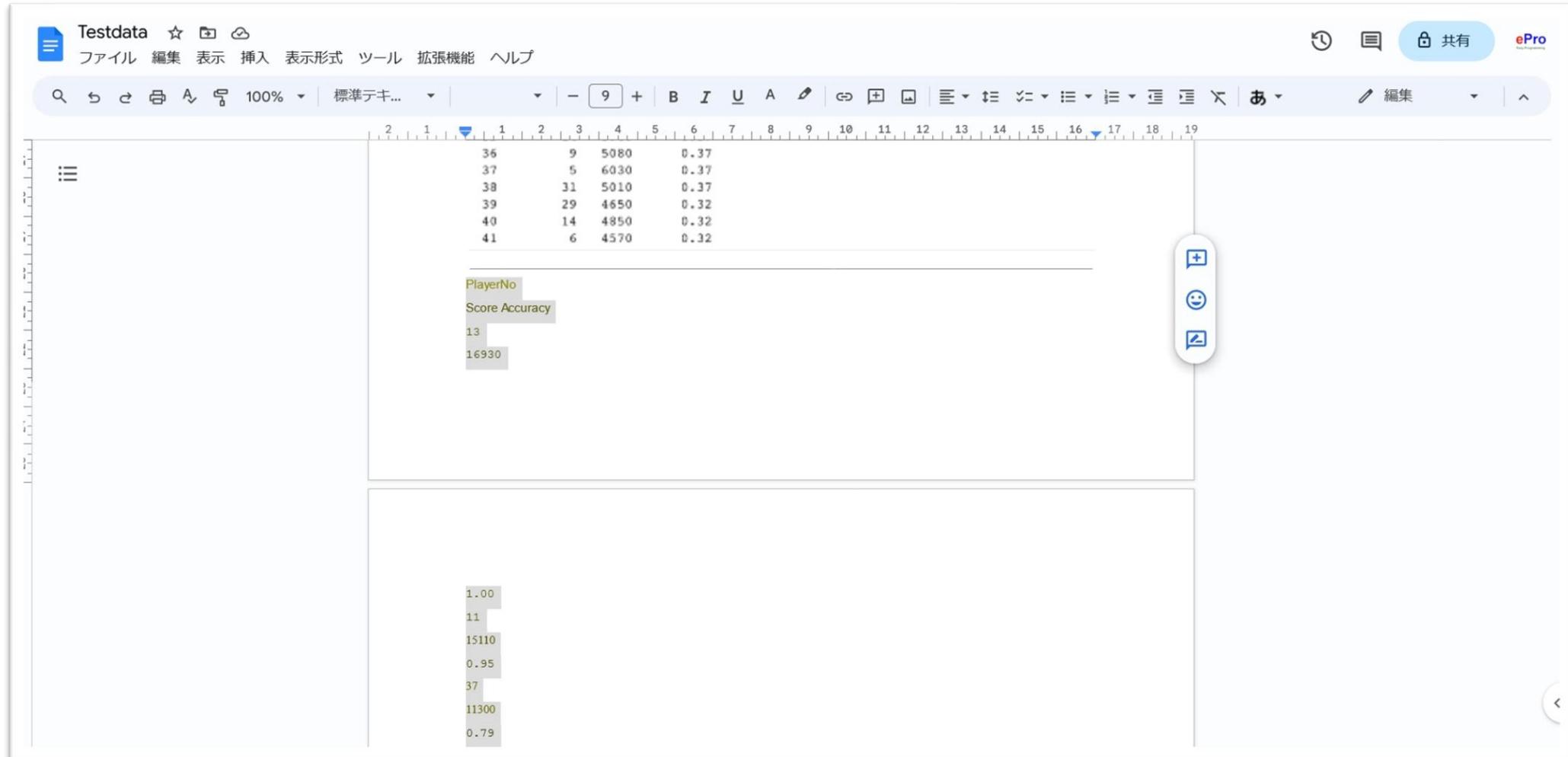
30	12	6480	0.47
31	15	5390	0.42
32	35	6170	0.42
33	4	7180	0.42
34	36	4940	0.37
35	24	5960	0.37
36	9	5080	0.37
37	5	6030	0.37
38	31	5010	0.37
39	29	4650	0.32
40	14	4850	0.32
41	6	4570	0.32

PlayerNo
Score Accuracy
13
16930

1.00
11
15110

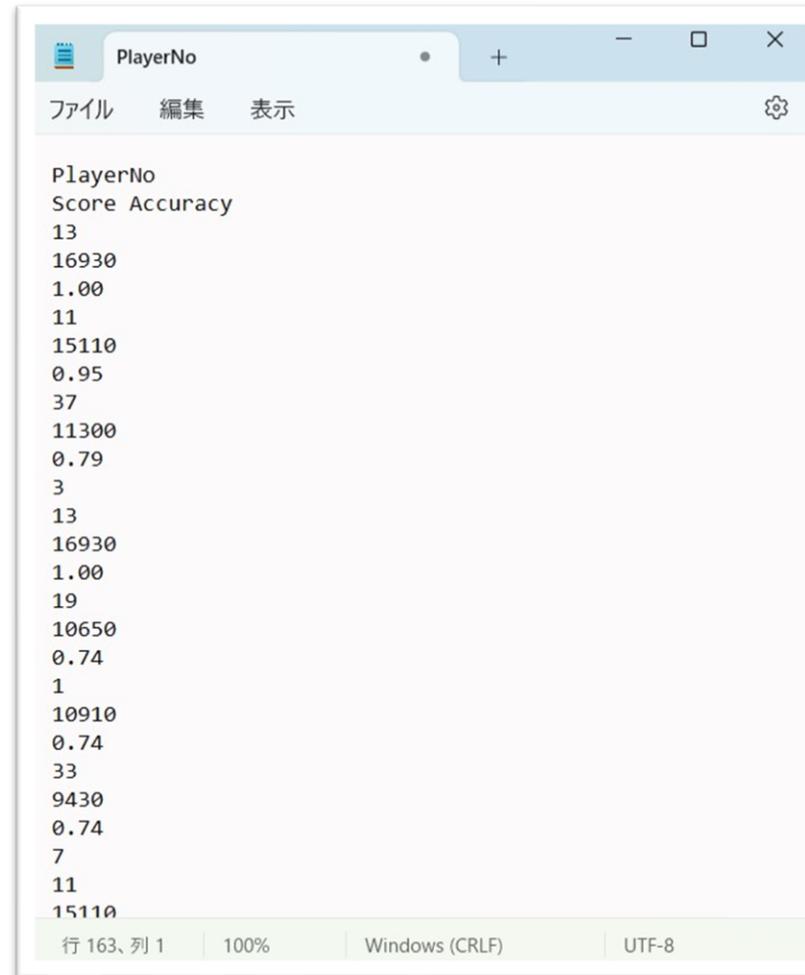
CSVファイルの生成

- テキスト部分を選択してコピー



CSVファイルの生成

- メモ帳を開いて貼り付ける



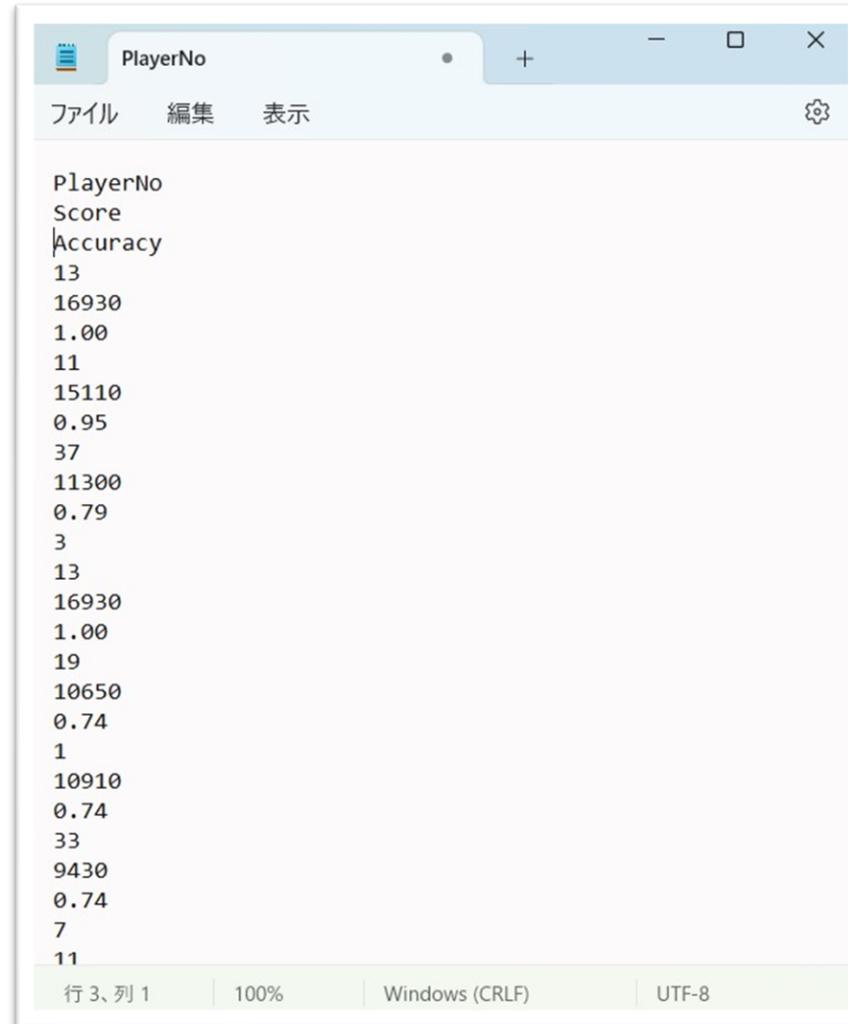
The screenshot shows a Notepad window titled "PlayerNo" with a menu bar containing "ファイル", "編集", and "表示". The text content is as follows:

```
PlayerNo
Score Accuracy
13
16930
1.00
11
15110
0.95
37
11300
0.79
3
13
16930
1.00
19
10650
0.74
1
10910
0.74
33
9430
0.74
7
11
15110
```

The status bar at the bottom indicates "行 163、列 1", "100%", "Windows (CRLF)", and "UTF-8".

CSVファイルの生成

- 「Score Accuracy」の行を2行に分割する



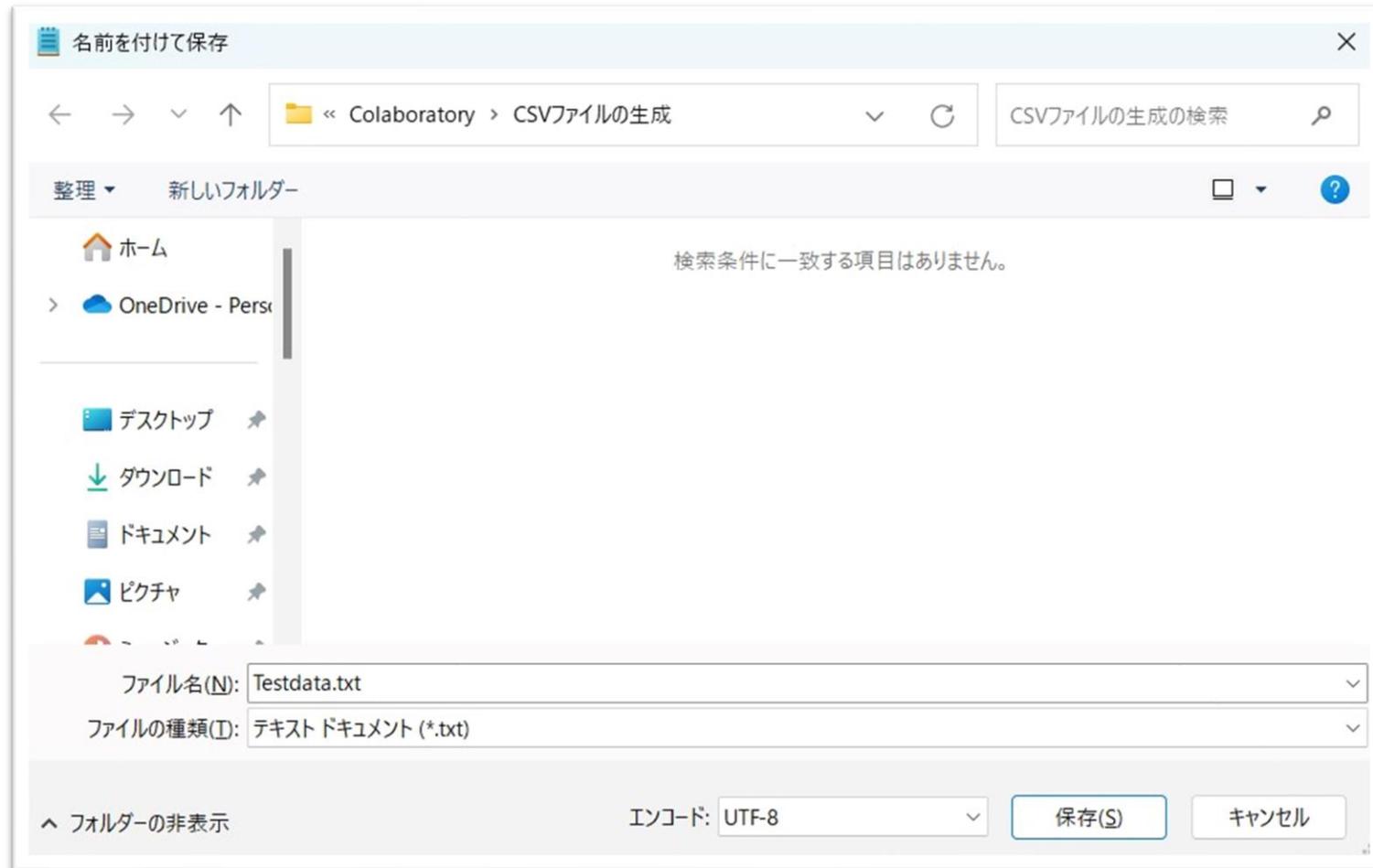
The screenshot shows a text editor window titled "PlayerNo" with a menu bar containing "ファイル", "編集", and "表示". The main text area contains the following CSV data:

```
PlayerNo
Score
Accuracy
13
16930
1.00
11
15110
0.95
37
11300
0.79
3
13
16930
1.00
19
10650
0.74
1
10910
0.74
33
9430
0.74
7
11
```

The cursor is positioned at the start of the "Accuracy" column. The status bar at the bottom indicates "行 3、列 1", "100%", "Windows (CRLF)", and "UTF-8".

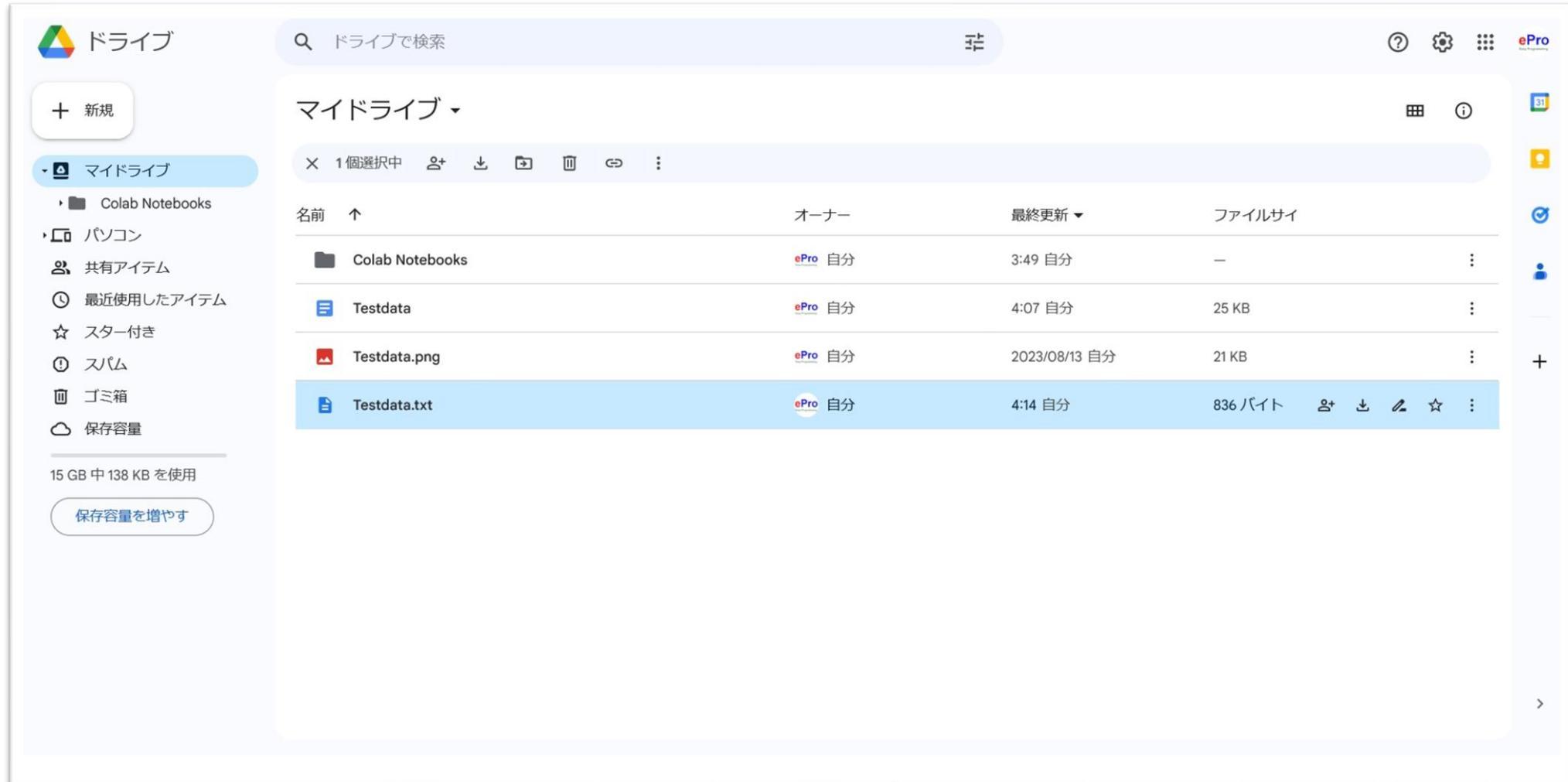
CSVファイルの生成

- 「Testdata.txt」という名前でテキストファイルとして保存



CSVファイルの生成

- Googleドライブにアップロード



CSVファイルの生成

- プログラムを入力し ドライブにマウント



```
import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

CSVファイルの生成

- プログラムを入力しドライブにマウント

CSVファイルの生成

- 別の方法：左のファイルボタンを押してもマウントできる



The screenshot shows a Google Colab notebook titled "CSVファイルの生成.ipynb". The notebook contains two code cells. The first cell mounts the Google Drive. The second cell imports the csv module, reads data from a text file in Google Drive, and writes it to a CSV file.

```
[4] # 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')

import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

CSVファイルの生成

- [ドライブをマウント]ボタンを押す



The screenshot shows a Jupyter Notebook titled "CSVファイルの生成.ipynb". The interface includes a file browser on the left, a code editor in the center, and a status bar at the bottom. The code in the notebook is as follows:

```
[ ] # 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')

import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

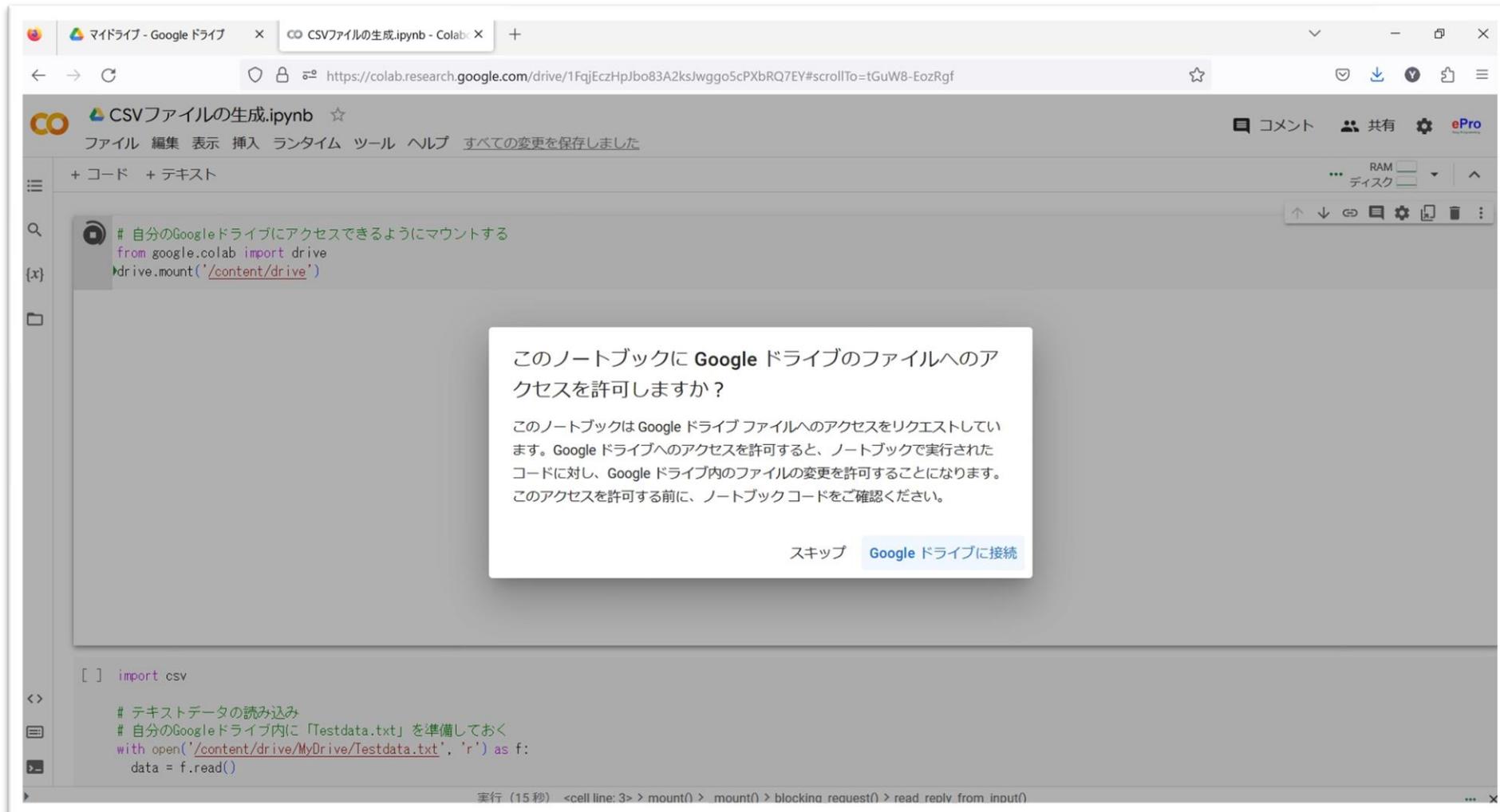
# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

The status bar at the bottom indicates "0秒 完了時間: 5:04".

CSVファイルの生成

- アクセスの許可



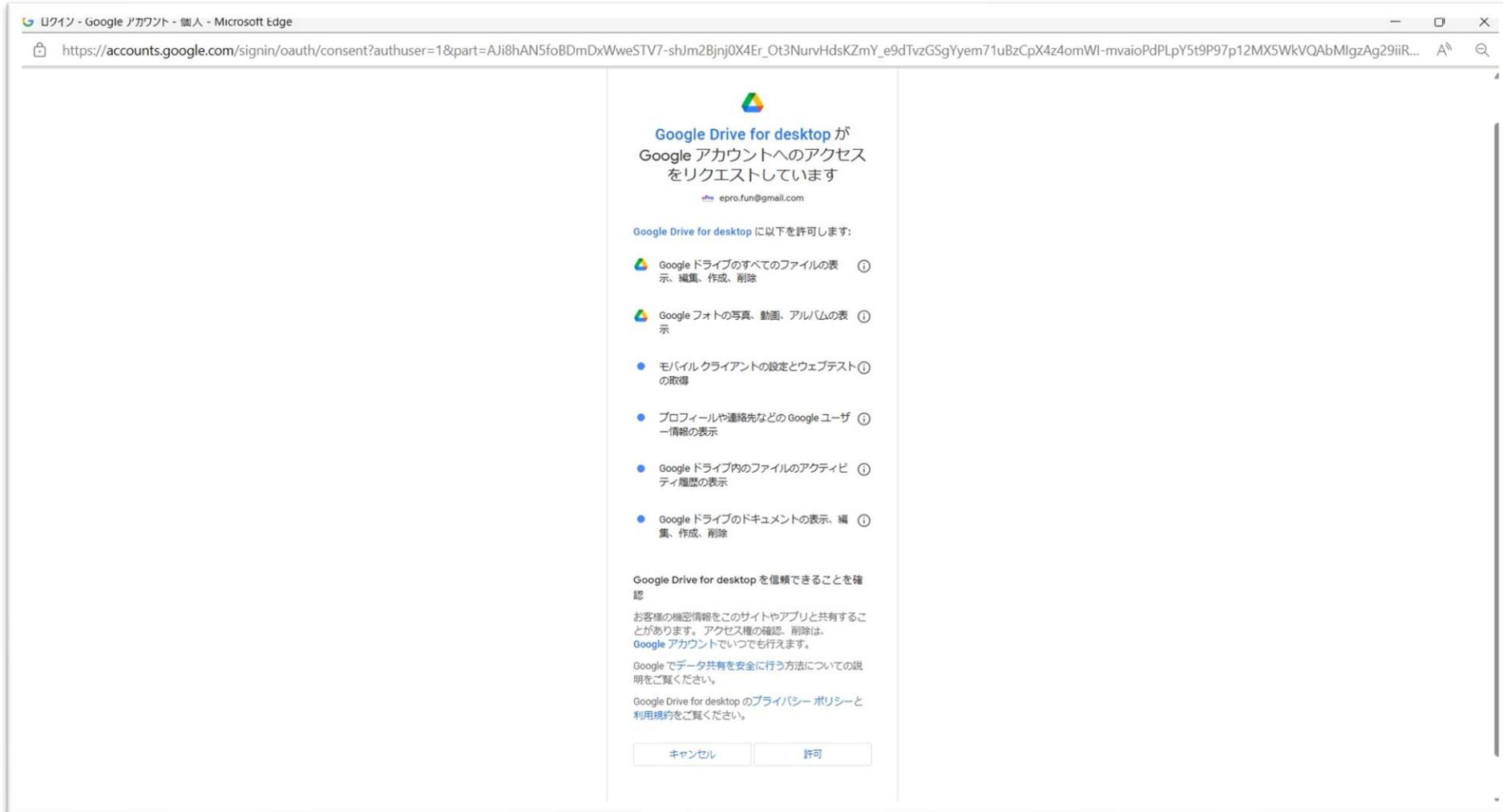
CSVファイルの生成

- アカウントの選択



CSVファイルの生成

- アカウントの認証



CSVファイルの生成

- マウント完了



The screenshot shows a Jupyter Notebook interface with the following content:

```
CSVファイルの生成.ipynb ☆
ファイル 編集 表示 挿入 ランタイム ツール ヘルプ
RAM
ディスク
+ コード + テキスト
# 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')
Mounted at /content/drive

[ ] import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

1分 17秒 完了時間: 502

CSVファイルの生成

- 左のファイルを開く



The screenshot shows a Google Colab notebook titled "CSVファイルの生成.ipynb". The left sidebar displays a file explorer with a folder named "sample_data". The main area shows a code cell that has been executed, with the following code and output:

```
+ コード + テキスト  
1分 ✓  
# 自分のGoogleドライブにアクセスできるようにマウントする  
from google.colab import drive  
drive.mount('/content/drive')  
  
Mounted at /content/drive  
  
[ ] import csv  
  
# テキストデータの読み込み  
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく  
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:  
    data = f.read()  
  
# print(data) #読み込んだデータの確認  
  
# リストに変換する  
data_list = data.split()  
# リストの内容を確認する  
print(data_list)  
  
with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:  
    writer = csv.writer(f)  
    for i in range(0, len(data_list), 3):  
        # リストの内容を3個ずつCSVファイルに書き出す  
        writer.writerow(data_list[i:i + 3])
```

At the bottom of the notebook, a status bar indicates "1分 17秒" and "完了時間: 5:02".

CSVファイルの生成

- フォルダを開いて（パスをたどって）「MyDrive」を表示する



The screenshot shows a Google Colab notebook titled "CSVファイルの生成.ipynb". The left sidebar displays a file explorer with the following structure:

- drive
 - MyDrive
 - Colab Notebooks
 - CSVファイルの生成.ipynb
 - Testdata.gdoc
 - Testdata.png
 - Testdata.txt
 - sample_data

The main area shows two code cells. The first cell contains the following code:

```
# 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')
```

The output of this cell is "Mounted at /content/drive".

The second cell contains the following code:

```
[ ] import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

CSVファイルの生成

- プログラムを実行する

CSVファイルの生成.ipynb ☆

ファイル 編集 表示 挿入 ランタイム ツール ヘルプ

RAM
ディスク

コメント 共有 ePro

ファイル

drive

MyDrive

Colab Notebooks

CSVファイルの生成.ipynb

Testdata.gdoc

Testdata.png

Testdata.txt

sample_data

```
[2] # 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

import csv

# テキストデータの読み込み
# 自分のGoogleドライブ内に「Testdata.txt」を準備しておく
with open('/content/drive/MyDrive/Testdata.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/Testdata.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(0, len(data_list), 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])
```

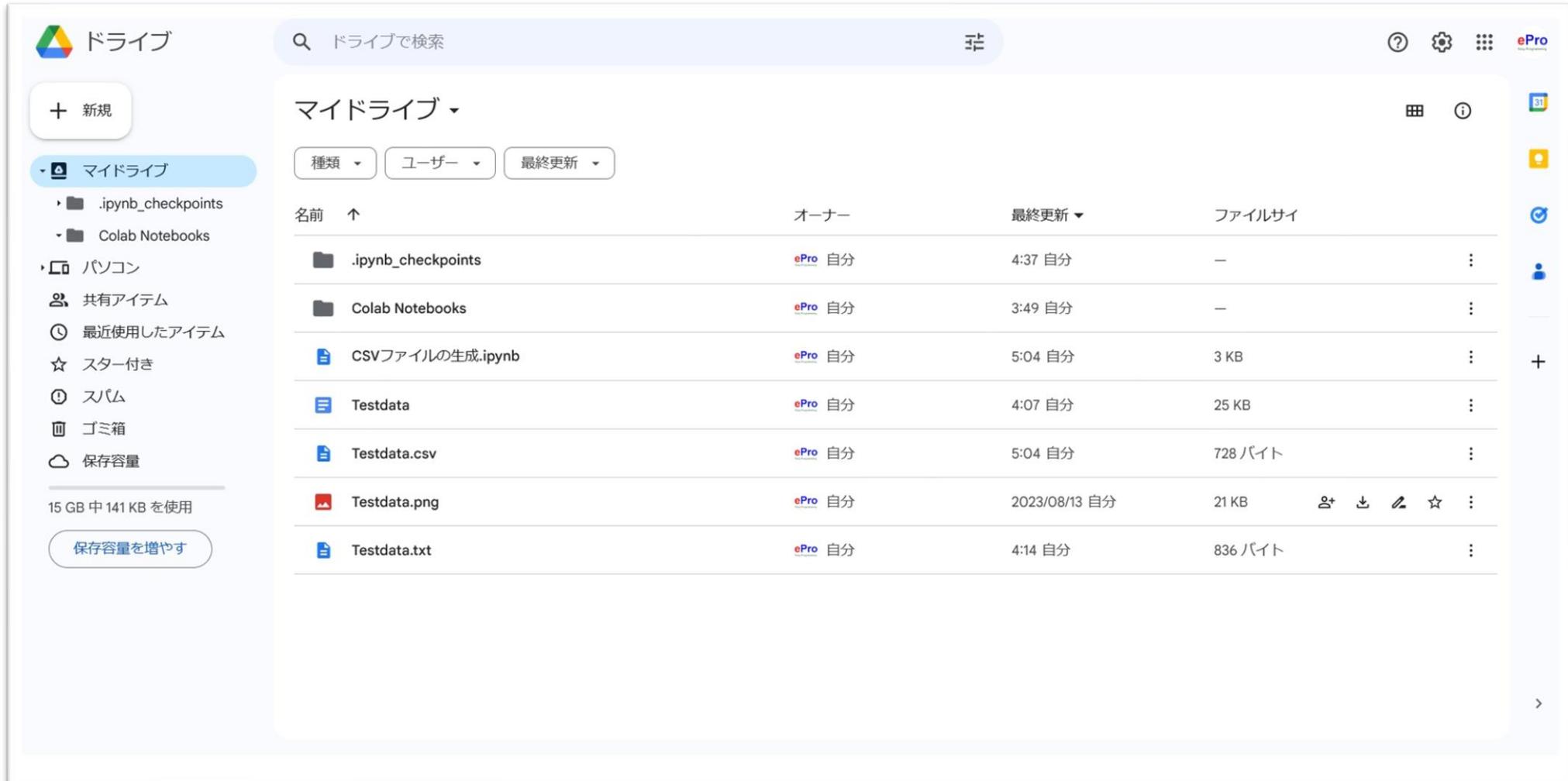
['PlayerNo', 'Score', 'Accuracy', '13', '16930', '1.00', '11', '15110', '0.95', '37', '11300', '0.79', '3', '13', '16930', '1.00', '19', '10650', '0.74', '1

0秒 完了時間: 5:04

ディスク 81.48 GBが利用可能

CSVファイルの生成

- 「Testdata.csv」ファイルが生成される



The screenshot shows the Google Drive interface. On the left, there is a sidebar with navigation options like 'マイドライブ' (My Drive) and 'パソコン' (Computer). The main area displays a list of files and folders. The file 'Testdata.csv' is highlighted in blue. Below is a table representing the file list data.

名前 ↑	オーナー	最終更新 ▼	ファイルサイ
📁 .ipynb_checkpoints	ePro 自分	4:37 自分	—
📁 Colab Notebooks	ePro 自分	3:49 自分	—
📄 CSVファイルの生成.ipynb	ePro 自分	5:04 自分	3 KB
📄 Testdata	ePro 自分	4:07 自分	25 KB
📄 Testdata.csv	ePro 自分	5:04 自分	728 バイト
🖼️ Testdata.png	ePro 自分	2023/08/13 自分	21 KB
📄 Testdata.txt	ePro 自分	4:14 自分	836 バイト

CSVファイルの生成

- 開いて確認する

	A	B	C
1	PlayerNo	Score	Accuracy
2	13	16930	1
3	11	15110	0.95
4	37	11300	0.79
5	3	13	16930
6	1	19	10650
7	0.74	1	10910
8	0.74	33	9430
9	0.74	7	11
10	15110	0.95	8
11	26	8950	0.68
12	9	11	15110
13	0.95	10	17
14	8920	0.68	11
15	21	9950	0.68
16	12	25	9950
17	0.68	13	7
18	9850	0.68	14
19	3	8600	0.63
20	15	28	9120
21	0.63	16	38
22	8460	0.63	17
23	2	7760	0.58
24	18	23	8260
25	0.58	19	16
26	7860	0.58	20
27	30	8430	0.58
28	21	27	7780
29	0.58	22	32
30	7740	0.58	23
31	18	8120	0.58

CSVファイルの生成

- 補足
 - CSVファイルはデータサイエンスの基本
 - ダウンロードしてExcelでCSVファイルを読み込むと表形式で表示される。

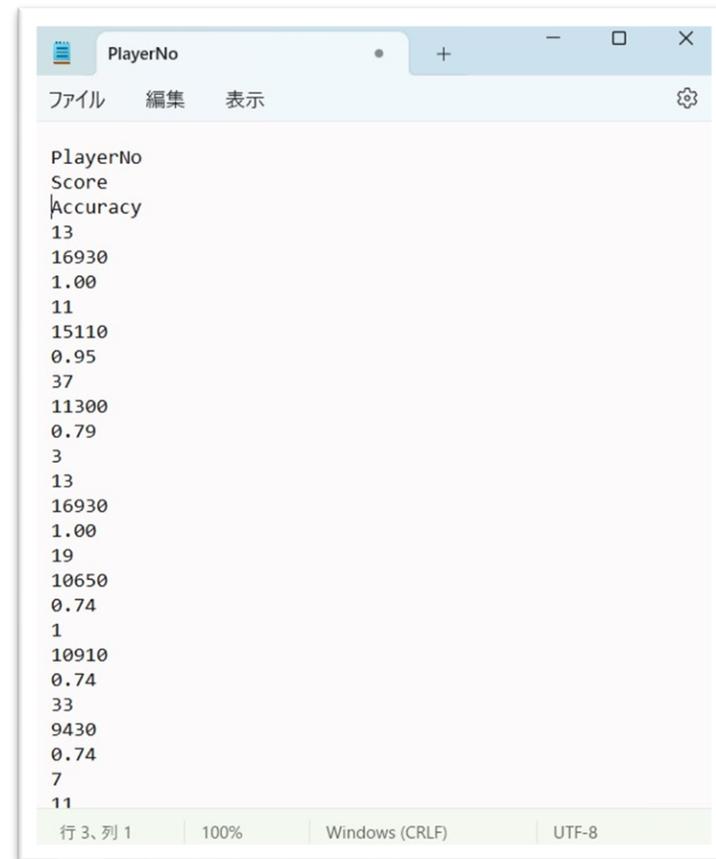
データ損失の可能性 このブックをコマンド区切り (.csv) 形式で保存すると、一部の機能が失われる可能性があります。機能が失われたいようにするには、Excel ファイル形式で保存してください。 次回から表示しない 名前を付けて保存...

PlayerNo	Score	Accuracy
Score	Accuracy	13
Accuracy	13	16930
13	16930	1
16930	1	11
1	11	15110
11	15110	0.95
15110	0.95	37
0.95	37	11300
37	11300	0.79
11300	0.79	3
0.79	3	13
3	13	16930
13	16930	1
16930	1	19
1	19	10650
19	10650	0.74
10650	0.74	1
0.74	1	10910
1	10910	0.74
10910	0.74	33

CSVファイルの生成

ようやく「演習4」の入り口にたどり着いた！？

- 補足
 - CSVファイルの本当の姿はメモ帳（あるいはテキストエディタ）での表示（形式）



The screenshot shows a text editor window titled "PlayerNo". The window contains the following text, which is a CSV file with three columns: PlayerNo, Score, and Accuracy. The data is as follows:

PlayerNo	Score	Accuracy
13	16930	1.00
11	15110	0.95
37	11300	0.79
3	16930	1.00
13	19	10650
1	0.74	1
10910	0.74	33
9430	0.74	7
11		

The status bar at the bottom of the window indicates the current position is "行 3、列 1", the zoom is "100%", the line ending is "Windows (CRLF)", and the encoding is "UTF-8".

Googleドライブとマウントとの対応関係

The image shows two overlapping windows. The left window is a JupyterLab notebook titled "CSVファイルの生成". The right window is a Google Drive interface.

JupyterLab Notebook:

```
# 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

[] import csv

# data = ['13 16930 1 11 15110 0.95 37 11300 0.79 3 13 16930']

# テキストデータの読み込み
# 自分のGoogleドライブ内に「TestData.txt」を準備しておく
with open('/content/drive/MyDrive/TestData.txt', 'r') as f:
    data = f.read()

# print(data) #読み込んだデータの確認

# リストに変換する
data_list = data.split()
# リストの内容を確認する
print(data_list)

with open('/content/drive/MyDrive/TestData.csv', 'w') as f:
    writer = csv.writer(f)
    for i in range(len(data_list) // 3):
        # リストの内容を3個ずつCSVファイルに書き出す
        writer.writerow(data_list[i:i + 3])

['13', '16930', '1', '11', '15110', '0.95', '37', '11300']
```

Google Drive Interface:

The Google Drive interface shows the "マイドライブ" (My Drive) section. A table lists the files and folders:

名前	オーナー	最終更新	ファイルサイズ
CSVファイルの生成	ePro 自分	2023/08/13 自分	4 KB
TestData	ePro 自分	2023/08/13 自分	25 KB
TestData.csv	ePro 自分	2023/08/13 自分	702 バイト
TestData.png	ePro 自分	2023/08/13 自分	21 KB
TestData.txt	ePro 自分	2023/08/13 自分	803 バイト

Red arrows indicate the mapping between the notebook and the Drive interface:

- From the notebook's `drive.mount('/content/drive')` line to the "drive" folder in the file browser.
- From the notebook's `with open('/content/drive/MyDrive/TestData.txt')` line to the "MyDrive" folder.
- From the notebook's `with open('/content/drive/MyDrive/TestData.csv')` line to the "TestData.csv" file in the Drive interface.
- From the notebook's `with open('/content/drive/MyDrive/TestData.txt')` line to the "TestData.txt" file in the Drive interface.

演習 4

The screenshot shows a Google Colab notebook titled "TestcsvimportD.ipynb". The notebook contains three code cells. The first cell mounts the local Google Drive. The second cell imports pandas and reads a CSV file. The third cell, titled "#演習4 dataframeから特定の行(0行目)を除く", drops the first row of the dataframe and prints the result. The output is a table with 17 rows and 4 columns: PlayerNo, Score, Accuracy, and an unlabeled column.

```
[1] # 自分のGoogleドライブにアクセスできるようにマウントする
from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

[2] import pandas as pd
# df = pd.read_csv('/content/drive/My Drive/Colab Notebooks/Testdata.csv')
df = pd.read_csv('/content/drive/MyDrive/Colab Notebooks/Testdata.csv')
print(df)

#演習4 dataframeから特定の行(0行目)を除く

datadropped = df.drop(0, axis=0)
print(datadropped)
```

	PlayerNo	Score	Accuracy	
1	16930.00	1.00	11.00	
2	1.00	11.00	15110.00	
3	11.00	15110.00	0.95	
4	15110.00	0.95	37.00	
5	0.95	37.00	11300.00	
6	37.00	11300.00	0.79	
7	11300.00	0.79	3.00	
8	0.79	3.00	13.00	
9	3.00	13.00	16930.00	
10	13.00	16930.00	1.00	
11	16930.00	1.00	19.00	
12	1.00	19.00	10650.00	
13	19.00	10650.00	0.74	
14	10650.00	0.74	1.00	
15	0.74	1.00	10910.00	
16	1.00	10910.00	0.74	
17	10910.00	0.74	33.00	