



# Excelを作る

初めてのプログラミング

2020年度

只木進一（理工学部）

# 今日のサンプルファイル

- ▶ <https://github.com/first-programming-saga/excelAndCSV2>

# 新規にDataFrameを作る

```
In [1]: 1 import pandas
```

DataFrameの作り方：その1

```
In [2]: 1 data={'数学':[90,70,100,60],  
2         '理科':[80,90,90,70],  
3         '英語':[80,100,80,90],  
4         '社会':[70,90,70,80]}  
5 index=['山田','佐藤','鈴木','渡辺']  
6 df = pandas.DataFrame(data,index)  
7 print(df)
```

	数学	理科	英語	社会
山田	90	80	80	70
佐藤	70	90	100	90
鈴木	100	90	80	70
渡辺	60	70	90	80

excelAndCSV2/createDataFrame.ipynb

# 新規にDataFrameを作る

## DataFrameの作り方：その2

```
In [4]: 1 data2={'名前':['山田','佐藤','鈴木','渡辺'],
2         '数学':[90,70,100,60],
3         '理科':[80,90,90,70],
4         '英語':[80,100,80,90],
5         '社会':[70,90,70,80]}
6 df2 = pandas.DataFrame(data2)
7 df2.set_index('名前', inplace=True)
8 print(df2)
```

	数学	理科	英語	社会
名前				
山田	90	80	80	70
佐藤	70	90	100	90
鈴木	100	90	80	70
渡辺	60	70	90	80

# 新しいデータの追加

## 新しい列の生成

In [12]:

```
1 label = '体育'  
2 record = [90,70,80,100]  
3 df[label] = record  
4 print(df)
```

	数学	理科	英語	社会	体育
山田	90	80	80	70	90
佐藤	70	90	100	90	70
鈴木	100	90	80	70	80
渡辺	60	70	90	80	100

## 1 新しい行の生成

In [13]:

```
1 name='古賀'  
2 recordKoga = [90,90,85,70,70]  
3 df.loc[name] = recordKoga  
4 print(df)
```

	数学	理科	英語	社会	体育
山田	90	80	80	70	90
佐藤	70	90	100	90	70
鈴木	100	90	80	70	80
渡辺	60	70	90	80	100
古賀	90	90	85	70	70

# ファイルへの出力

## ■ エクセル出力

```
df.to_excel(ファイル名)
```

## ■ CSV出力

```
df.to_csv(ファイル名)
```

# データ集計 入力ファイル

name	English	Math	Science	Social
Per	80	90	95	70
John	80	60	70	100
Kim	100	60	65	80
Sally	70	80	95	70
Tom	80	70	80	60
Maki	70	100	90	80
Joe	75	90	100	70
Jane	90	80	80	100
Ann	90	100	80	100
Kei	100	70	90	75

```
excelAndCSV2/StatFromExcel.ipynb
```

# データ集計 目標とする出力

name	English	Math	Science	Social
Per	80	90	95	70
John	80	60	70	100
Kim	100	60	65	80
Sally	70	80	95	70
Tom	80	70	80	60
Maki	70	100	90	80
Joe	75	90	100	70
Jane	90	80	80	100
Ann	90	100	80	100
Kei	100	70	90	75
SubjectAverage	83.5	80	84.5	80.5



# データ集計 手順の大枠

1. エクセルファイルを読み、  
dataFrameに変換
2. 各科目の平均を記入する欄を作成
3. 各科目の平均を計算し、記入

# 平均値を記入する欄を作る

- ▶ 各教科の平均値を記入する行の生成
  - ▶ 0をlen(df.columns)個記入

科目と個人の平均を記入するために、新規に列と行を追加する

```
In [ ]: 1 def addAverageRow(df):  
        2     df.loc[subjectAverageLabel]=[0.]*len(df.columns)
```

# データ集計 科目毎平均

```
1. forall (各列){
2.     sum = 0 #和
3.     count = 0 #人数
4.     ser = 一列のデータ
5.     forall (serのキー){
6.         if (キーは科目平均ではない){
7.             sum += データ
8.             count += 1
9.         }
10.    }
11.    average = sum/count
12.    average を記入
13. }
```

## 科目の成績

- 各列（各科目）を選ぶ（ただし、ラベルが'IndividualAverage'でないもの）。
  - 列のデータはSeries。キー（人）で一つずつ読む（ただし、ラベルが'SubjectAverage'でないもの）

In [4]:

```
1 def subjectRecord(data):
2     for subject in data.columns:#各列を選ぶ
3         print(f'----{subject}----')
4         ser = data[subject]#一列のデータ(series)
5         for k in ser.keys():
6             if k != subjectAverageLabel:
7                 v = ser[k]
8                 print(f'[k]:{v}')
```