Excelを読み書きする

初めてのプログラミング 2019年度 只木進一(理工学部)

Pandasパッケージ

- ▶表や列を扱う目的のパッケージ
 - ▶行や列のラベル指定が可能
- DataFrame
 - ▶表形式のデータ
- Series
 - ▶列形式のデータ
- ▶入出力

Excelファイルの読み込み

■ pandas.read_excel()を利用して excelを読む

import pandas
with pandas.ExcelFile(filename) as f:
 data = pandas.read_excel(f)

- excelの内容はpandas.DataFrame クラス
- https://github.com/firstprogramming-saga/excelAndCSV

pandas.DataFrame

- ▶一行目がcolumnsとして識別される
 - DataFrame.columnsで参照
- ■一列目がindexとして認識される
 - DataFrame.indexで参照

columns English Math Science Social 80 90 95 Tim index 70 100 John 60 Kim 100 60 65 80 Sally 80 95 70 70 Tom 80 70 80 60 70 100 90 80 Bob

初めてのプログラミング©只木進一

Columnの指定

- ▶一行目がcolumnsとして識別される
 - ■data[列名]: 一行のデータ
 - ■pandas.Seriesクラス
 - ■data[列名][行名]: 指定位置のデータ

data['Math']

	English	Matk	Science	Social
Tim	80	90	95	70
John	80	60	70	100
Kim	100	60	65	80
Sally	70	80	95	70
Tom	80	70	80	60
Bob	70	100	90	80

初めてのプログラミング©只木進一

data['Social'] ['Sally']

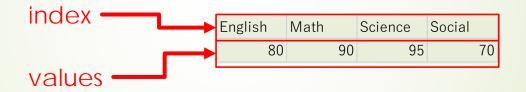
Indexの指定

- ■一列目はindexとして識別される
 - ■data.loc[行名]:一行のデータ
 - ■pandas.Seriesクラス
- data.loc[行名][列名]: 指定位置の

テータ		English	Math	Science	Social
	Tim	80	90	95	70
	John	80	60	70	100
data.loc['John']	Kim	100	60	65	80
	Sally	710	80	95	70
data.loc['Sally'] ['Math']	Tom	80	70	80	60
	Bob	70	100	90	80

pandas.Series

■ DataFrameから一行または一列取り 出すとSeriesとなる



for k in ser.index:
 v = ser[k]
 print(f'ser[{k}]:{v}')

CSVを読む

- csv (comma-separated values) ファイルはテキストファイル
- pandas.DataFrameとしても読むことができる

import pandas

data = pandas.read_csv(filename)

1カラムに名前がある場合

name	English	Math	Science	Social
Tim	80	90	95	70
John	80	60	70	100
Kim	100	60	65	80
Sally	70	80	95	70
Tom	80	70	80	60
Bob	70	100	90	80

▶番号のindexを付けられてしまう。

	name	English	Math	Science	Social
0	Tim	80	90	95	70
1	John	80	60	70	100
2	Kim	100	60	65	80
3	Sally	70	80	95	70
4	Tom	80	70	80	60
5	Bob	70	100	90	80

■set_index()で最初のカラムを指定

新規にDataFrameを作る

新しいデータの追加

■新しい列

```
label = '体育'
record = [90,70,80,100,90,70]
df[label]=record
```

▶新しい行

```
name = '古賀'
record = [90,90,86,70]
df.loc[name]=record
```

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)