cw1-e-dataFrame-reading-writing

March 15, 2022

[1]:
```
import pandas as pd
!cat examples/ex1.csv
```

```
alfa,beta,gamma,delta,word
5,6,7,8,ma
1,2,3,4,ala
9,10,11,12,kota
```

[2]:
```
df = pd.read_csv('examples/ex1.csv')
df
```

[2]:

	alfa	beta	gamma	delta	word
0	5	6	7	8	ma
1	1	2	3	4	ala
2	9	10	11	12	kota

[3]:
```
!cat examples/ex2.csv
```

```
5,6,7,8,ala
1,2,3,4,ma
9,10,11,12,kota
```

[4]:
```
pd.read_csv('examples/ex2.csv')
```

[4]:

	5	6	7	8	ala
0	1	2	3	4	ma
1	9	10	11	12	kota

[5]:
```
pd.read_csv('examples/ex2.csv', header = None)
```

[5]:

	0	1	2	3	4
0	5	6	7	8	ala
1	1	2	3	4	ma
2	9	10	11	12	kota

[6]:
```
pd.read_csv('examples/ex2.csv', names=['a','b','c','d','text'])
```

[6]:
```
   a   b   c   d  text
0  5   6   7   8   ala
1  1   2   3   4   ma
2  9  10  11  12  kota
```

[8]:
```
names = ['I','II','III','IV','word']
pd.read_csv('examples/ex2.csv', names = names, index_col='word')
```

[8]:
```
      I  II  III  IV
word
ala   5   6    7   8
ma    1   2    3   4
kota  9  10   11  12
```

[11]:
```
# opening a text file separated by variable amounts of whitespace symbols
list(open('examples/ex3.txt'))
```

[11]:
```
['            alfa         beta         gamma\n',
 'x -0.264438 -1.026059 -0.619500\n',
 'y  0.927272  0.302904 -0.032399\n',
 'z -0.264273 -0.386314 -0.217601\n',
 'v -0.871858 -0.348382  1.100491\n']
```

[13]:
```
# read it as df
pd.read_csv('examples/ex3.txt',sep = '\s+')
# index iferred automatically since the number of colnames is 1 less than␣
 ↪number of columns
```

[13]:
```
       alfa      beta     gamma
x -0.264438 -1.026059 -0.619500
y  0.927272  0.302904 -0.032399
z -0.264273 -0.386314 -0.217601
v -0.871858 -0.348382  1.100491
```

[14]:
```
# skipping rows
!cat examples/ex4.csv
```

```
# niepotrzebna pierwsza linia
alfa,beta,gamma,delta,word
# komentarz1, który trzeba pominąć w danych
# komentarz2
5,6,7,8,ala
1,2,3,4,ma
9,10,11,12,kota
```

[15]:
```
pd.read_csv('examples/ex4.csv',skiprows = [0,3,2])
```

[15]:

	alfa	beta	gamma	delta	word
0	5	6	7	8	ala
1	1	2	3	4	ma
2	9	10	11	12	kota

[19]:
```
# missing values (default: empty field or some standard name)
!cat examples/ex5.csv
```

```
name,a,b,c,d,word
ala,1,2,3,4,NA
ola,5,6,,8,ma
tosia,?,10,11,12,kota
```

[20]:
```
pd.read_csv('examples/ex5.csv')
```

[20]:

	name	a	b	c	d	word
0	ala	1	2	3.0	4	NaN
1	ola	5	6	NaN	8	ma
2	tosia	?	10	11.0	12	kota

[21]:
```
pd.read_csv('examples/ex5.csv',na_values='?')
```

[21]:

	name	a	b	c	d	word
0	ala	1.0	2	3.0	4	NaN
1	ola	5.0	6	NaN	8	ma
2	tosia	NaN	10	11.0	12	kota

[29]:
```
pd.read_csv('examples/ex5.csv',na_values={'a':['?'],'word':['?']})
```

[29]:

	name	a	b	c	d	word
0	ala	1.0	2	3.0	4	NaN
1	ola	5.0	6	NaN	8	ma
2	tosia	NaN	10	11.0	12	kota

[32]:
```
# writing data to text format
df.to_csv('examples/written.csv', na_rep='NaN')
```