

```

# Practical - 2
# Bendre Anushka A.
# TCO074
import pandas as pd
import numpy as np
index = pd.date_range("1/1/2000", periods=8)

s = pd.Series(np.random.randn(5), index=["a", "b", "c", "d", "e"])

s

a    -0.086528
b    -0.711500
c    -1.342835
d    -1.297275
e     2.132302
dtype: float64

df = pd.DataFrame(np.random.randn(8, 3), index=index, columns=["A",
"B", "C"])
df

```

	A	B	C
2000-01-01	-0.648552	0.659313	-0.099236
2000-01-02	-0.887142	0.667063	1.466192
2000-01-03	0.445046	-1.231343	-0.377259
2000-01-04	0.187401	2.075141	-1.254141
2000-01-05	-0.742831	0.669048	0.672633
2000-01-06	-0.488106	1.133684	-1.105367
2000-01-07	-2.042008	2.546807	-1.239859
2000-01-08	0.360245	-0.380727	0.112358

```
df.head()
```

	A	B	C
2000-01-01	-0.648552	0.659313	-0.099236
2000-01-02	-0.887142	0.667063	1.466192
2000-01-03	0.445046	-1.231343	-0.377259
2000-01-04	0.187401	2.075141	-1.254141
2000-01-05	-0.742831	0.669048	0.672633

```
df.head(2)
```

	A	B	C
2000-01-01	-0.648552	0.659313	-0.099236
2000-01-02	-0.887142	0.667063	1.466192

```
df.tail()
```

	A	B	C
2000-01-04	0.187401	2.075141	-1.254141
2000-01-05	-0.742831	0.669048	0.672633

```
2000-01-06 -0.488106  1.133684 -1.105367
2000-01-07 -2.042008  2.546807 -1.239859
2000-01-08  0.360245 -0.380727  0.112358
```

```
df.tail(2)
```

```
          A          B          C
2000-01-07 -2.042008  2.546807 -1.239859
2000-01-08  0.360245 -0.380727  0.112358
```

```
df = pd.DataFrame(
    {
        "one": pd.Series(np.random.randn(3), index=["a", "b", "c"]),
        "two": pd.Series(np.random.randn(4), index=["a", "b", "c",
"d"]),
        "three": pd.Series(np.random.randn(3), index=["b", "c", "d"]),
    }
)
df
```

```
      one      two      three
a -0.944549 -0.240045      NaN
b  0.517087 -1.841126  0.313439
c -0.396527  2.780265 -0.447179
d      NaN -1.313342 -0.506332
```

```
df.sub(df.iloc[2], axis="columns")
```

```
      one      two      three
a -0.548022 -3.020310      NaN
b  0.913614 -4.621391  0.760618
c  0.000000  0.000000  0.000000
d      NaN -4.093607 -0.059153
```

```
df.sub(df.iloc[2], axis=1)
```

```
      one      two      three
a -0.548022 -3.020310      NaN
b  0.913614 -4.621391  0.760618
c  0.000000  0.000000  0.000000
d      NaN -4.093607 -0.059153
```

```
df.sub(df["one"], axis=0)
```

```
      one      two      three
a  0.0  0.704504      NaN
b  0.0 -2.358213 -0.203648
c  0.0  3.176792 -0.050652
d  NaN      NaN      NaN
```

```
df2 = df.copy()
df2
```

	one	two	three
a	-0.944549	-0.240045	NaN
b	0.517087	-1.841126	0.313439
c	-0.396527	2.780265	-0.447179
d	NaN	-1.313342	-0.506332

```
df2.loc["d", "one"] = 0.08
```

```
df2
```

	one	two	three
a	-0.944549	-0.240045	NaN
b	0.517087	-1.841126	0.313439
c	-0.396527	2.780265	-0.447179
d	0.080000	-1.313342	-0.506332

```
df
```

	one	two	three
a	-0.944549	-0.240045	NaN
b	0.517087	-1.841126	0.313439
c	-0.396527	2.780265	-0.447179
d	NaN	-1.313342	-0.506332

```
df.gt(df2)
```

	one	two	three
a	False	False	False
b	False	False	False
c	False	False	False
d	False	False	False

```
df.lt(df2)
```

	one	two	three
a	False	False	False
b	False	False	False
c	False	False	False
d	False	False	False

```
df.ge(df2)
```

	one	two	three
a	True	True	False
b	True	True	True
c	True	True	True
d	False	True	True

```
df.ne(df2)
```

	one	two	three
a	False	False	True

```
b False False False
c False False False
d  True False False
```

```
(df > 0).any()
```

```
one      True
two      True
three    True
dtype: bool
```

```
df
```

```
      one      two      three
a -0.944549 -0.240045      NaN
b  0.517087 -1.841126  0.313439
c -0.396527  2.780265 -0.447179
d      NaN -1.313342 -0.506332
```

```
df.loc["d"] = df.loc["a"].copy()
```

```
df
```

```
      one      two      three
a -0.944549 -0.240045      NaN
b  0.517087 -1.841126  0.313439
c -0.396527  2.780265 -0.447179
d -0.944549 -0.240045      NaN
```

```
s = pd.Series(["a", "b", "c"], dtype=object)
```

```
s
```

```
0    a
1    b
2    c
dtype: object
```

```
s[0]=None
```

```
s
```

```
0    None
1      b
2      c
dtype: object
```

```
s.iloc[0] = np.nan
```

```
s
```

```
0    NaN
1      b
```

```

2      c
dtype: object

ser1 = pd.Series([np.nan, np.nan, 2, 3])
ser2=pd.Series([np.nan, 1, np.nan, 4])

ser1
0      NaN
1      NaN
2      2.0
3      3.0
dtype: float64

ser2
0      NaN
1      1.0
2      NaN
3      4.0
dtype: float64

ser1+ser2
0      NaN
1      NaN
2      NaN
3      7.0
dtype: float64

pd.Series([1,2,3,4,5], dtype="float64").sum()
15.0

ser1
0      NaN
1      NaN
2      2.0
3      3.0
dtype: float64

ser1.dropna() #remove null val
2      2.0
3      3.0
dtype: float64

ser2
0      NaN
1      1.0
2      NaN

```

```
3      4.0
dtype: float64
```

```
ser2.fillna(999) # replace NA by 999
```

```
0      999.0
1        1.0
2      999.0
3        4.0
dtype: float64
```

```
data=pd.read_csv("/home/ubuntu/Desktop/DATA.csv")
```

```
data
```

	gender	NationalITY	PlaceofBirth	StageID	GradeID	SectionID	\
0	M	KW	KuwaIT	lowerlevel	G-04	A	
1	M	KW	KuwaIT	lowerlevel	G-04	A	
2	M	KW	KuwaIT	lowerlevel	G-04	A	
3	M	KW	KuwaIT	lowerlevel	G-04	A	
4	M	KW	KuwaIT	lowerlevel	G-04	A	
..	
475	F	Jordan	Jordan	MiddleSchool	G-08	A	
476	F	Jordan	Jordan	MiddleSchool	G-08	A	
477	F	Jordan	Jordan	MiddleSchool	G-08	A	
478	F	Jordan	Jordan	MiddleSchool	G-08	A	
479	F	Jordan	Jordan	MiddleSchool	G-08	A	

	Topic	Semester	Relation	raisedhands	VisITedResources	\
0	IT	F	Father	15	16.0	
1	IT	F	Father	20	20.0	
2	IT	F	Father	10	7.0	
3	IT	F	Father	30	25.0	
4	IT	F	Father	40	50.0	
..	
475	Chemistry	S	Father	5	NaN	
476	Geology	F	Father	50	NaN	
477	Geology	S	Father	55	NaN	
478	History	F	Father	30	NaN	
479	NaN	S	Father	35	NaN	

	AnnouncementsView	Discussion	ParentAnsweringSurvey	\
0	2	20	Yes	
1	3	25	Yes	
2	0	30	No	
3	5	35	No	
4	12	50	No	
..	
475	5	8	No	
476	14	28	No	
477	25	29	No	

478	14	57	No
479	23	62	No

	ParentschoolSatisfaction	StudentAbsenceDays	Class
0	Good	Under-7	M
1	Good	Under-7	M
2	Bad	Above-7	L
3	Bad	Above-7	L
4	Bad	Above-7	M
..
475	Bad	Above-7	L
476	Bad	Under-7	M
477	Bad	Under-7	M
478	Bad	Above-7	L
479	Bad	Above-7	L

[480 rows x 17 columns]

```
df2=pd.DataFrame(data)
df2
```

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	\
0	M	KW	KuwaIT	lowerlevel	G-04	A	
1	M	KW	KuwaIT	lowerlevel	G-04	A	
2	M	KW	KuwaIT	lowerlevel	G-04	A	
3	M	KW	KuwaIT	lowerlevel	G-04	A	
4	M	KW	KuwaIT	lowerlevel	G-04	A	
..	
475	F	Jordan	Jordan	MiddleSchool	G-08	A	
476	F	Jordan	Jordan	MiddleSchool	G-08	A	
477	F	Jordan	Jordan	MiddleSchool	G-08	A	
478	F	Jordan	Jordan	MiddleSchool	G-08	A	
479	F	Jordan	Jordan	MiddleSchool	G-08	A	

	Topic	Semester	Relation	raisedhands	VisITedResources	\
0	IT	F	Father	15	16.0	
1	IT	F	Father	20	20.0	
2	IT	F	Father	10	7.0	
3	IT	F	Father	30	25.0	
4	IT	F	Father	40	50.0	
..	
475	Chemistry	S	Father	5	NaN	
476	Geology	F	Father	50	NaN	
477	Geology	S	Father	55	NaN	
478	History	F	Father	30	NaN	
479	NaN	S	Father	35	NaN	

	AnnouncementsView	Discussion	ParentAnsweringSurvey	\
0	2	20	Yes	
1	3	25	Yes	

2	0	30	No
3	5	35	No
4	12	50	No
..
475	5	8	No
476	14	28	No
477	25	29	No
478	14	57	No
479	23	62	No

	ParentschoolSatisfaction	StudentAbsenceDays	Class
0	Good	Under-7	M
1	Good	Under-7	M
2	Bad	Above-7	L
3	Bad	Above-7	L
4	Bad	Above-7	M
..
475	Bad	Above-7	L
476	Bad	Under-7	M
477	Bad	Under-7	M
478	Bad	Above-7	L
479	Bad	Above-7	L

[480 rows x 17 columns]

df2.head()

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	Topic
\							
0	M	KW	KuwaIT	lowerlevel	G-04	A	IT
1	M	KW	KuwaIT	lowerlevel	G-04	A	IT
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT
3	M	KW	KuwaIT	lowerlevel	G-04	A	IT
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT

	Semester	Relation	raisedhands	VisITedResources	AnnouncementsView
\					
0	F	Father	15	16.0	2
1	F	Father	20	20.0	3
2	F	Father	10	7.0	0
3	F	Father	30	25.0	5
4	F	Father	40	50.0	12

	Discussion	ParentAnsweringSurvey	ParentschoolSatisfaction	\
0	20	Yes	Good	
1	25	Yes	Good	
2	30	No	Bad	
3	35	No	Bad	
4	50	No	Bad	

	StudentAbsenceDays	Class
0	Under-7	M
1	Under-7	M
2	Above-7	L
3	Above-7	L
4	Above-7	M

df2.tail()

	gender	NationalITy	PlaceofBirth	StageID	GradeID	SectionID	\
475	F	Jordan	Jordan	MiddleSchool	G-08	A	
476	F	Jordan	Jordan	MiddleSchool	G-08	A	
477	F	Jordan	Jordan	MiddleSchool	G-08	A	
478	F	Jordan	Jordan	MiddleSchool	G-08	A	
479	F	Jordan	Jordan	MiddleSchool	G-08	A	

	Topic	Semester	Relation	raisedhands	VisITedResources	\
475	Chemistry	S	Father	5		NaN
476	Geology	F	Father	50		NaN
477	Geology	S	Father	55		NaN
478	History	F	Father	30		NaN
479	NaN	S	Father	35		NaN

	AnnouncementsView	Discussion	ParentAnsweringSurvey	\
475	5	8	No	
476	14	28	No	
477	25	29	No	
478	14	57	No	
479	23	62	No	

	ParentschoolSatisfaction	StudentAbsenceDays	Class
475	Bad	Above-7	L
476	Bad	Under-7	M
477	Bad	Under-7	M
478	Bad	Above-7	L
479	Bad	Above-7	L