xperiment-no-2-data-manipulation

October 23, 2023

1 Data Manipulation

```
[1]: #Exp no.: 2
     #Aim: Data Manipulation
     #Name: Krutika Nemade
     #Roll no.: 55
     #Sec: A
     #Subject: Data Science and Statistics (Lab 1)
[4]: import pandas as pd
     import os
[5]:
[6]:
     os.getcwd()
[6]: 'C:\\Users\\hp\\Downloads'
     os.chdir('C:\\Users\\hp\\Desktop')
[8]: data=pd.read_csv('diabetes.csv')
[9]: data.head(10)
[9]:
                                BloodPressure
                                                SkinThickness
        Pregnancies
                      Glucose
                                                                Insulin
                                                                           BMI
     0
                   6
                           148
                                            72
                                                            35
                                                                       0
                                                                          33.6
                           85
                                                            29
                                                                       0
                                                                          26.6
     1
                   1
                                            66
     2
                   8
                           183
                                            64
                                                             0
                                                                       0
                                                                          23.3
                                                            23
     3
                                                                          28.1
                   1
                           89
                                            66
                                                                      94
     4
                   0
                          137
                                            40
                                                            35
                                                                     168
                                                                          43.1
     5
                   5
                          116
                                            74
                                                             0
                                                                       0
                                                                          25.6
     6
                   3
                           78
                                            50
                                                            32
                                                                      88
                                                                          31.0
     7
                  10
                          115
                                                                          35.3
                                             0
                                                             0
                                                                       0
     8
                   2
                          197
                                            70
                                                            45
                                                                     543
                                                                          30.5
     9
                   8
                          125
                                                                       0
                                                                           0.0
                                            96
                                                             0
```

```
0
                             0.627
                                     50
                                                1
                             0.351
                                                0
      1
                                     31
                             0.672
      2
                                      32
                                                1
      3
                             0.167
                                     21
                                                0
      4
                             2.288
                                     33
                                                1
                             0.201
                                                0
      5
                                     30
      6
                             0.248
                                     26
                                                1
      7
                                                0
                             0.134
                                     29
      8
                             0.158
                                     53
                                                1
      9
                             0.232
                                     54
                                                1
[10]: data.tail()
           Pregnancies Glucose BloodPressure SkinThickness
                                                                 Insulin
[10]:
                                                                            BMI
      763
                     10
                             101
                                              76
                                                              48
                                                                      180 32.9
      764
                      2
                             122
                                              70
                                                              27
                                                                        0 36.8
      765
                      5
                                              72
                                                                      112 26.2
                             121
                                                              23
      766
                      1
                             126
                                              60
                                                               0
                                                                        0 30.1
      767
                      1
                              93
                                              70
                                                              31
                                                                        0 30.4
           DiabetesPedigreeFunction Age
                                           Outcome
      763
                               0.171
                                       63
                                                  0
      764
                               0.340
                                       27
                                                  0
      765
                               0.245
                                                  0
                                       30
      766
                               0.349
                                        47
                                                  1
      767
                               0.315
                                        23
                                                  0
[11]: data.shape
[11]: (768, 9)
[12]: data.size
[12]: 6912
[13]: data.ndim
[13]: 2
[14]: data.columns
[14]: Index(['Pregnancies', 'Glucose', 'BloodPressure', 'SkinThickness', 'Insulin',
             'BMI', 'DiabetesPedigreeFunction', 'Age', 'Outcome'],
            dtype='object')
[15]: data.head()
```

DiabetesPedigreeFunction

Age

Outcome

```
[15]:
         Pregnancies Glucose BloodPressure SkinThickness
                                                                 Insulin
                                                                            BMI \
                            148
                                                             35
                                                                       0
                                                                           33.6
      0
                    6
                                             72
                    1
                            85
                                                             29
                                                                           26.6
      1
                                             66
                                                                       0
      2
                    8
                            183
                                             64
                                                             0
                                                                       0
                                                                           23.3
      3
                    1
                                             66
                            89
                                                             23
                                                                      94
                                                                           28.1
      4
                    0
                            137
                                             40
                                                             35
                                                                     168
                                                                           43.1
         DiabetesPedigreeFunction
                                     Age
                                          Outcome
      0
                             0.627
                                      50
                                                 1
      1
                             0.351
                                      31
                                                 0
      2
                             0.672
                                      32
                                                 1
      3
                              0.167
                                      21
                                                 0
      4
                              2.288
                                      33
                                                 1
[16]: data.drop(labels="Age",axis=1)
[16]:
           Pregnancies
                         Glucose
                                   BloodPressure
                                                   SkinThickness
                                                                   Insulin
                                                                              BMI \
      0
                      6
                              148
                                               72
                                                               35
                                                                             33.6
      1
                               85
                                               66
                                                               29
                                                                             26.6
                      1
                                                                          0
      2
                      8
                                                                0
                              183
                                               64
                                                                         0
                                                                             23.3
      3
                      1
                                               66
                                                               23
                                                                         94 28.1
                               89
                      0
      4
                              137
                                               40
                                                               35
                                                                        168 43.1
      763
                                               76
                                                                        180 32.9
                     10
                              101
                                                               48
      764
                                               70
                                                               27
                                                                         0 36.8
                      2
                              122
                                                                       112 26.2
      765
                      5
                              121
                                               72
                                                               23
      766
                      1
                              126
                                               60
                                                                0
                                                                         0 30.1
      767
                      1
                               93
                                               70
                                                               31
                                                                          0 30.4
           DiabetesPedigreeFunction
                                       Outcome
      0
                                0.627
      1
                                0.351
                                              0
      2
                                0.672
                                              1
      3
                                0.167
                                              0
      4
                                2.288
                                              1
      . .
                                  •••
      763
                                0.171
                                              0
      764
                                0.340
                                              0
      765
                                0.245
                                              0
      766
                                0.349
                                              1
      767
                                0.315
                                              0
      [768 rows x 8 columns]
[17]: data.drop(labels=["Age", "Glucose"], axis=1)
```

[17]:	Pregnancies	BloodPressure	SkinThicl	kness	Insulin	BMI	\		
0	6	72		35	0	33.6			
1	1	66		29	0	26.6			
2	8	64		0	0	23.3			
3	1	66		23	94				
4	0	40		35	168	43.1			
	•••	•••	•••						
76	33 10	76		48	180	32.9			
76		70		27	0	36.8			
76		72		23					
					112				
76		60		0	0				
76	57 1	70		31	0	30.4			
	DiabetesPedi	greeFunction	Outcome						
0	Diabotobi oa	0.627							
			1						
1		0.351	0						
2		0.672	1						
3		0.167	0						
4		2.288	1						
76		0.171	0						
76	34	0.340	0						
76	35	0.245	0						
76	36	0.349	1						
76		0.315	0						
7 () (0.313	O						
_		_							
[7	768 rows x 7 col	lumns]							
[18] : da	ata.head(10)								
[18]:	Pregnancies (Glucose BloodP:	ressure Sl	kinThi	ckness	Insulin	BMI	\	
	6	148	72		35	0	33.6	`	
0						_			
1	1	85	66		29	0			
2	8	183	64		0	0	23.3		
3	1	89	66		23	94	28.1		
4	0	137	40		35	168	43.1		
	5	116	74		0	0	25.6		
5									
6	3	78	50		32	88	31.0		
7	10	115	0		0	0	35.3		
8	2	197	70		45	543	30.5		
9	8	125	96		0	0	0.0		
3	J	120	50		5	J	0.0		
	DiabetesPedig								
0		0.627 5							
1		0.351 3	1 0						
2		0.672 3							
2		0.072 0.							

0.167

```
4
                       2.288
                                33
                                          1
5
                       0.201
                                30
                                          0
                       0.248
6
                                          1
                                26
7
                       0.134
                                29
                                          0
                       0.158
8
                                          1
                                53
9
                       0.232
                                54
                                          1
```

[19]: data.drop(labels=[2,3],axis=0)

[19]:	Pregnancies	Glucose	BloodPressure	SkinThickness	Insulin	BMI	\
0	6	148	72	35	0	33.6	
1	1	85	66	29	0	26.6	
4	0	137	40	35	168	43.1	
5	5	116	74	0	0	25.6	
6	3	78	50	32	88	31.0	
		•••	•••		•••		
763	10	101	76	48	180	32.9	
764	2	122	70	27	0	36.8	
765	5	121	72	23	112	26.2	
766	1	126	60	0	0	30.1	
767	1	93	70	31	0	30.4	

	${\tt DiabetesPedigreeFunction}$		Outcome
0	0.62	27 50	1
1	0.35	31	0
4	2.28	33	1
5	0.20	1 30	0
6	0.24	18 26	1
		•••	•••
763	0.17	'1 63	0
764	0.34	10 27	0
765	0.24	£5 30	0
766	0.34	19 47	1
767	0.31	.5 23	0

[766 rows x 9 columns]