nhehk7ow0

December 13, 2024

1 DATA Cleaning, Missing Value Treatment

```
[2]: #Name : Devesh J Arbat
     #Roll no. : 06
     #Section : A
[3]:
     #Aim: To perform Data Processing, Data Cleaning, Missing Value Treatment
     import pandas as pd
[5]:
     import os
     os.getcwd()
[6]: 'C:\\Users\\91876'
     os.chdir("C:\\Users\\91876\\OneDrive\\Desktop\\Data Science")
[8]: data=pd.read_csv("titanic.csv")
[9]:
     data
[9]:
           pclass
                   survived
                                                                           name
              1.0
     0
                         1.0
                                                 Allen, Miss. Elisabeth Walton
              1.0
                         1.0
     1
                                                Allison, Master. Hudson Trevor
     2
              1.0
                        0.0
                                                  Allison, Miss. Helen Loraine
     3
              1.0
                        0.0
                                         Allison, Mr. Hudson Joshua Creighton
     4
              1.0
                              Allison, Mrs. Hudson J C (Bessie Waldo Daniels)
                         0.0
              3.0
                                                         Zabour, Miss. Thamine
     1305
                        0.0
     1306
              3.0
                         0.0
                                                     Zakarian, Mr. Mapriededer
     1307
              3.0
                        0.0
                                                           Zakarian, Mr. Ortin
     1308
              3.0
                        0.0
                                                            Zimmerman, Mr. Leo
     1309
              NaN
                        NaN
                                                                            NaN
                             sibsp parch ticket
                                                                cabin embarked boat \
              sex
                       age
                                                        fare
                                      0.0
                                            24160
                                                  211.3375
                                                                   В5
     0
           female
                   29.0000
                               0.0
                                                                              S
```

1	male	0.9167	1.0	2.0	113781	151.5500	C22	C26	S	11
2	female	2.0000	1.0	2.0	113781	151.5500	C22	C26	S	NaN
3	male	30.0000	1.0	2.0	113781	151.5500	C22	C26	S	NaN
4	female	25.0000	1.0	2.0	113781	151.5500	C22	C26	S	${\tt NaN}$
•••	•••		•••	•••	•••		•••			
1305	female	NaN	1.0	0.0	2665	14.4542		${\tt NaN}$	С	NaN
1306	male	26.5000	0.0	0.0	2656	7.2250		${\tt NaN}$	C	NaN
1307	male	27.0000	0.0	0.0	2670	7.2250		${\tt NaN}$	C	NaN
1308	male	29.0000	0.0	0.0	315082	7.8750		${\tt NaN}$	S	NaN
1309	NaN	NaN	NaN	NaN	NaN	NaN		${\tt NaN}$	NaN	NaN
	body			h	ome.dest					
0	NaN			St L	ouis, MO					
1	NaN	Montreal,	PQ / C	hesterv	rille, ON					
2	NaN	Montreal,	PQ / C	hesterv	rille, ON					
3	135.0	Montreal,	PQ / C	hesterv	rille, ON					
4	NaN	Montreal,	PQ / C	hesterv	rille, ON					
•••	•••				•••					
1305	NaN				NaN					
1306	304.0				NaN					
1307	NaN				NaN					
1308	NaN				NaN					
1309	NaN				NaN					

[1310 rows x 14 columns]

[10]: data.head(40)

\	name	survived	pclass	[10]:
	Allen, Miss. Elisabeth Walton	1.0	1.0	0
	Allison, Master. Hudson Trevor	1.0	1.0	1
	Allison, Miss. Helen Loraine	0.0	1.0	2
	Allison, Mr. Hudson Joshua Creighton	0.0	1.0	3
	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	0.0	1.0	4
	Anderson, Mr. Harry	1.0	1.0	5
	Andrews, Miss. Kornelia Theodosia	1.0	1.0	6
	Andrews, Mr. Thomas Jr	0.0	1.0	7
	Appleton, Mrs. Edward Dale (Charlotte Lamson)	1.0	1.0	8
	Artagaveytia, Mr. Ramon	0.0	1.0	9
	Astor, Col. John Jacob	0.0	1.0	10
	Astor, Mrs. John Jacob (Madeleine Talmadge Force)	1.0	1.0	11
	Aubart, Mme. Leontine Pauline	1.0	1.0	12
	Barber, Miss. Ellen "Nellie"	1.0	1.0	13
	Barkworth, Mr. Algernon Henry Wilson	1.0	1.0	14
	Baumann, Mr. John D	0.0	1.0	15
	Baxter, Mr. Quigg Edmond	0.0	1.0	16
	Baxter, Mrs. James (Helene DeLaudeniere Chaput)	1.0	1.0	17

19	1.0	1.0				Bazz	zani, Miss	s. Albina		
10	1.0	0.0	Beattie, Mr. Thomson							
20	1.0	1.0	Beckwith, Mr. Richard Leonard							
21	1.0	1.0	Beck	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)						
22	1.0	1.0		Behr, Mr. Karl Howell						
23	1.0	1.0				Bido	ois, Miss	. Rosalie		
24	1.0	1.0					Bird, Mis	ss. Ellen		
25	1.0	0.0				Bi	rnbaum, N	Mr. Jakob		
26	1.0	1.0				Bishop	o, Mr. Dio	ckinson H		
27	1.0	1.0		Bi	shop, Mrs.	Dickinson	ı H (Heler	n Walton)		
28	1.0	1.0				Bisse	ette, Miss	s. Amelia		
29	1.0	1.0		Bjor	nstrom-Ste	ffansson,	Mr. Maur	itz Hakan		
30	1.0	0.0			В	lackwell,	Mr. Steph	nen Weart		
31	1.0	1.0					Blank, N	Mr. Henry		
32	1.0	1.0				Bonnel	l, Miss.	Caroline		
33	1.0	1.0				Bonnell	, Miss. H	Elizabeth		
34	1.0	0.0				Boreban	ık, Mr. Jo	ohn James		
35	1.0	1.0				Bowen,	Miss. Gra	ace Scott		
36	1.0	1.0				Bowerman,	Miss. Els	sie Edith		
37	1.0	1.0	В	radley,	Mr. Georg	e ("George	Arthur I	Brayton")		
38	1.0	0.0				Brady,	Mr. John	n Bertram		
39	1.0	0.0				E	Brandeis,	Mr. Emil		
	sex	age	sibsp	parch	ticket	fare	cabin	${\tt embarked}$	boat	\
0	female	29.0000	0.0	0.0	24160	211.3375	B5	S	2	
1	male	0.9167	1.0	2.0	113781	151.5500	C22 C26	S	11	
2	female	2.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	
3	male	30.0000	1.0	2.0	113781	151.5500	COO COC		3.T 3.T	
4	female						C22 C26	S	NaN	
		25.0000	1.0	2.0	113781	151.5500	C22 C26	S	NaN	
5	male	48.0000	0.0	0.0	19952	151.5500 26.5500	C22 C26 E12	S S	NaN 3	
6	female	48.0000 63.0000	0.0 1.0		19952 13502	151.5500 26.5500 77.9583	C22 C26 E12 D7	\$ \$ \$	NaN 3 10	
	female male	48.0000 63.0000 39.0000	0.0 1.0 0.0	0.0 0.0 0.0	19952 13502 112050	151.5500 26.5500 77.9583 0.0000	C22 C26 E12 D7 A36	\$ \$ \$ \$	NaN 3 10 NaN	
6 7 8	female male female	48.0000 63.0000 39.0000 53.0000	0.0 1.0 0.0 2.0	0.0 0.0 0.0	19952 13502 112050 11769	151.5500 26.5500 77.9583 0.0000 51.4792	C22 C26 E12 D7 A36 C101	\$ \$ \$ \$	NaN 3 10 NaN D	
6 7 8 9	female male female male	48.0000 63.0000 39.0000 53.0000 71.0000	0.0 1.0 0.0 2.0 0.0	0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042	C22 C26 E12 D7 A36 C101 NaN	\$ \$ \$ \$ \$ C	NaN 3 10 NaN D NaN	
6 7 8 9 10	female male female male	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000	0.0 1.0 0.0 2.0 0.0 1.0	0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250	C22 C26 E12 D7 A36 C101 NaN C62 C64	\$ \$ \$ \$ C	NaN 3 10 NaN D NaN NaN	
6 7 8 9 10 11	female male female male female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000	0.0 1.0 0.0 2.0 0.0 1.0	0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250	C22 C26 E12 D7 A36 C101 NaN C62 C64	\$ \$ \$ \$ C C	NaN 3 10 NaN D NaN NaN	
6 7 8 9 10 11 12	female male male male male female female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0	0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64	\$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$	NaN 3 10 NaN D NaN NaN 4	
6 7 8 9 10 11 12 13	female male female male female female female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 26.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN	\$ \$ \$ \$ \$ C C C \$	NaN 3 10 NaN D NaN NaN 4 9	
6 7 8 9 10 11 12 13	female male male male male female female female male	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500 30.0000	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23	\$ \$ \$ \$ \$ C C C C \$ \$	NaN 3 10 NaN D NaN NaN 4 9 6	
6 7 8 9 10 11 12 13 14 15	female male female male female female female male male	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 26.0000 80.0000	0.0 1.0 0.0 2.0 0.0 1.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500 30.0000 25.9250	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23	\$ \$ \$ \$ \$ C C C C \$ \$ \$ \$	NaN 3 10 NaN D NaN NaN 4 9 6 B	
6 7 8 9 10 11 12 13 14 15 16	female male female male female female female male male male	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23 NaN B58 B60	s s s s c c c c s s s c	NaN 3 10 NaN D NaN 4 9 6 B NaN	
6 7 8 9 10 11 12 13 14 15 16	female male female male female female female male male male male	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000 50.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558 PC 17558	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208 247.5208	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23 NaN B58 B60 B58 B60	s s s s c c c c s s s c c	NaN 3 10 NaN D NaN NaN 4 9 6 B NaN NaN NaN 6	
6 7 8 9 10 11 12 13 14 15 16 17 18	female male male male female female female male male male female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000 50.0000 32.0000	0.0 1.0 0.0 2.0 0.0 1.0 0.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558 PC 17558	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208 247.5208 76.2917	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23 NaN B58 B60 B58 B60 D15	s s s s c c c c s s s c c c c	NaN 3 10 NaN D NaN 4 9 6 B NaN NaN 6	
6 7 8 9 10 11 12 13 14 15 16 17 18	female male female male female female female male male male male female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000 50.0000 32.0000 36.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0 1	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558 PC 17558 PC 17558 11813 13050	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208 247.5208 76.2917 75.2417	C22 C26 E12 D7 A36 C101 NaN C62 C64 B35 NaN A23 NaN B58 B60 B58 B60 D15 C6	s s s s c c c c s s s c c c c	NaN 3 10 NaN D NaN NaN 4 9 6 B NaN NaN 6 8	
6 7 8 9 10 11 12 13 14 15 16 17 18 19 20	female male female male female female female male male male male female female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000 50.0000 32.0000 37.0000	0.0 1.0 0.0 2.0 0.0 1.0 0.0 0.0 0.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558 PC 17558 11813 13050 11751	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208 247.5208 76.2917 75.2417 52.5542	C22 C26 E12 D7 A36 C101 NaN C62 C64 C62 C64 B35 NaN A23 NaN B58 B60 B58 B60 D15 C6 D35	s s s s c c c c s s s s c c c c s	NaN 3 10 NaN D NaN NaN 4 9 6 B NaN NaN 6 8 A 5	
6 7 8 9 10 11 12 13 14 15 16 17 18	female male female male female female female male male male male female	48.0000 63.0000 39.0000 53.0000 71.0000 47.0000 18.0000 24.0000 80.0000 NaN 24.0000 50.0000 32.0000 36.0000	0.0 1.0 0.0 2.0 0.0 1.0 1.0 0.0 0.0 0.0 0.0	0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 1.0 1	19952 13502 112050 11769 PC 17609 PC 17757 PC 17757 PC 17477 19877 27042 PC 17318 PC 17558 PC 17558 PC 17558 11813 13050	151.5500 26.5500 77.9583 0.0000 51.4792 49.5042 227.5250 69.3000 78.8500 30.0000 25.9250 247.5208 247.5208 76.2917 75.2417	C22 C26 E12 D7 A36 C101 NaN C62 C64 B35 NaN A23 NaN B58 B60 B58 B60 D15 C6	s s s s c c c c s s s c c c c	NaN 3 10 NaN D NaN NaN 4 9 6 B NaN NaN 6 8	

```
23
    female 42.0000
                         0.0
                                0.0 PC 17757
                                                 227.5250
                                                                NaN
                                                                             С
                                                                                  4
                                                                C97
                                                                             S
                                                                                  8
24
    female
             29.0000
                         0.0
                                0.0
                                      PC 17483
                                                 221.7792
25
      male
             25.0000
                         0.0
                                0.0
                                         13905
                                                  26.0000
                                                                NaN
                                                                             С
                                                                                NaN
                                0.0
                                                                             С
26
      male
             25.0000
                         1.0
                                         11967
                                                  91.0792
                                                                B49
                                                                                  7
27
    female
            19.0000
                         1.0
                                0.0
                                         11967
                                                  91.0792
                                                                B49
                                                                             С
                                                                                  7
                                                                             S
28
    female
             35.0000
                         0.0
                                0.0 PC 17760
                                                 135.6333
                                                                C99
                                                                                  8
29
      male
             28.0000
                         0.0
                                0.0
                                                                C52
                                                                             S
                                                                                  D
                                        110564
                                                  26.5500
                         0.0
                                                                  Т
                                                                             S
30
      male
             45.0000
                                0.0
                                        113784
                                                  35.5000
                                                                                {\tt NaN}
                                                                             С
                         0.0
                                0.0
                                                                                  7
31
      male
             40.0000
                                        112277
                                                  31.0000
                                                                A31
32
    female
             30.0000
                         0.0
                                0.0
                                         36928
                                                 164.8667
                                                                 C7
                                                                             S
                                                                                  8
33
    female
                         0.0
                                0.0
                                                                             S
                                                                                  8
             58.0000
                                        113783
                                                  26.5500
                                                               C103
34
      male
            42.0000
                         0.0
                                0.0
                                        110489
                                                  26.5500
                                                                D22
                                                                             S
                                                                                NaN
35
   female
             45.0000
                         0.0
                                0.0 PC 17608
                                                 262.3750
                                                                NaN
                                                                             С
                                                                                  4
36
    female
             22.0000
                         0.0
                                1.0
                                        113505
                                                  55.0000
                                                                E33
                                                                             S
                                                                                  6
37
                         0.0
                                0.0
                                                                             S
                                                                                  9
      male
                 NaN
                                        111427
                                                  26.5500
                                                                NaN
                                                                             S
38
      male
            41.0000
                         0.0
                                0.0
                                        113054
                                                  30.5000
                                                                A21
                                                                                {\tt NaN}
39
                                                                             С
      male
            48.0000
                         0.0
                                0.0 PC 17591
                                                  50.4958
                                                                B10
                                                                                NaN
                                                 )
```

0NaNSt Louis, MO1NaNMontreal, PQ / Chesterville, ON2NaNMontreal, PQ / Chesterville, ON3135.0Montreal, PQ / Chesterville, ON4NaNMontreal, PQ / Chesterville, ON5NaNNew York, NY6NaNHudson, NY7NaNBelfast, NI8NaNBayside, Queens, NY922.0Montevideo, Uruguay10124.0New York, NY11NaNNew York, NY12NaNParis, France13NaNNaN14NaNHessle, Yorks15NaNNew York, NY16NaNMontreal, PQ17NaNMontreal, PQ18NaNNaN19NaNNew York, NY21NaNNew York, NY22NaNNew York, NY23NaNNew York, NY24NaNNaN25148.0San Francisco, CA26NaNDowagiac, MI27NaNDowagiac, MI		body	home.dest
Montreal, PQ / Chesterville, ON Montreal, PQ Montevideo, NY Montreal, PQ Montevideo, Uruguay Montredeo, Uruguay Montredeo, Uruguay Montreal, NY Montreal, PQ M	0	NaN	St Louis, MO
Montreal, PQ / Chesterville, ON Montreal, PQ / Chesterville, ON New York, NY Montreal, PQ / Chesterville, ON New York, NY Mew York, NY Man Belfast, NI Bayside, Queens, NY Montevideo, Uruguay Montevideo, Uruguay New York, NY Mew York, NY Mew York, NY Mew York, NY Man Paris, France Man Man Man Man Messle, Yorks Montreal, PQ Montreal, PA Montreal, PQ Montreal, PA Montre	1	NaN	Montreal, PQ / Chesterville, ON
Montreal, PQ / Chesterville, ON New York, NY Nan Hudson, NY Nan Belfast, NI New York, NY Montevideo, Uruguay New York, NY Nan Paris, France Nan Nan Hessle, Yorks Nan New York, NY Nan Montreal, PQ Nan Montreal, PQ Nan	2	NaN	Montreal, PQ / Chesterville, ON
5 NaN New York, NY 6 NaN Hudson, NY 7 NaN Belfast, NI 8 NaN Bayside, Queens, NY 9 22.0 Montevideo, Uruguay 10 124.0 New York, NY 11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	3	135.0	Montreal, PQ / Chesterville, ON
6 NaN Hudson, NY 7 NaN Belfast, NI 8 NaN Bayside, Queens, NY 9 22.0 Montevideo, Uruguay 10 124.0 New York, NY 11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	4	NaN	Montreal, PQ / Chesterville, ON
7NaNBelfast, NI8NaNBayside, Queens, NY922.0Montevideo, Uruguay10124.0New York, NY11NaNNew York, NY12NaNParis, France13NaNNaN14NaNHessle, Yorks15NaNNew York, NY16NaNMontreal, PQ17NaNMontreal, PQ18NaNNaN19NaNWinnipeg, MN20NaNNew York, NY21NaNNew York, NY22NaNNew York, NY23NaNNaN24NaNNaN25148.0San Francisco, CA26NaNDowagiac, MI	5	NaN	New York, NY
8 NaN Bayside, Queens, NY 9 22.0 Montevideo, Uruguay 10 124.0 New York, NY 11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 20 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	6	NaN	Hudson, NY
9 22.0 Montevideo, Uruguay 10 124.0 New York, NY 11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	7	NaN	Belfast, NI
10 124.0 New York, NY 11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	8	NaN	Bayside, Queens, NY
11 NaN New York, NY 12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	9	22.0	Montevideo, Uruguay
12 NaN Paris, France 13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN NaN 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	10	124.0	New York, NY
13 NaN NaN 14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	11	NaN	New York, NY
14 NaN Hessle, Yorks 15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	12	NaN	Paris, France
15 NaN New York, NY 16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	13	NaN	NaN
16 NaN Montreal, PQ 17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	14	NaN	Hessle, Yorks
17 NaN Montreal, PQ 18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	15	NaN	New York, NY
18 NaN NaN 19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	16	NaN	Montreal, PQ
19 NaN Winnipeg, MN 20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	17	NaN	Montreal, PQ
20 NaN New York, NY 21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	18	NaN	NaN
21 NaN New York, NY 22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	19	NaN	Winnipeg, MN
22 NaN New York, NY 23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	20	NaN	New York, NY
23 NaN NaN 24 NaN NaN 25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	21	NaN	New York, NY
24NaNNaN25148.0San Francisco, CA26NaNDowagiac, MI	22	NaN	New York, NY
25 148.0 San Francisco, CA 26 NaN Dowagiac, MI	23	NaN	NaN
26 NaN Dowagiac, MI	24	NaN	NaN
G .	25	148.0	San Francisco, CA
27 NaN Dowagiac, MI	26	NaN	Dowagiac, MI
8 4 4 7	27	NaN	Dowagiac, MI

```
29
             NaN
                  Stockholm, Sweden / Washington, DC
                                           Trenton, NJ
      30
             NaN
      31
             NaN
                                        Glen Ridge, NJ
      32
             NaN
                                        Youngstown, OH
      33
             NaN
                   Birkdale, England Cleveland, Ohio
      34
             NaN
                                London / Winnipeg, MB
      35
             NaN
                                       Cooperstown, NY
      36
             NaN
                    St Leonards-on-Sea, England Ohio
      37
             NaN
                                       Los Angeles, CA
             NaN
      38
                                           Pomeroy, WA
      39
          208.0
                                              Omaha, NE
[11]: data.tail(10)
[11]:
             pclass
                     survived
                                                                       name
                                                                                sex
                                                                                       age
      1300
                3.0
                           1.0
                                Yasbeck, Mrs. Antoni (Selini Alexander)
                                                                             female
                                                                                      15.0
      1301
                3.0
                           0.0
                                                     Youseff, Mr. Gerious
                                                                               male
                                                                                      45.5
                           0.0
      1302
                3.0
                                                        Yousif, Mr. Wazli
                                                                                       NaN
                                                                               male
                3.0
                           0.0
                                                    Yousseff, Mr. Gerious
      1303
                                                                               male
                                                                                       NaN
      1304
                3.0
                           0.0
                                                     Zabour, Miss. Hileni
                                                                             female
                                                                                      14.5
      1305
                3.0
                           0.0
                                                    Zabour, Miss. Thamine
                                                                             female
                                                                                       NaN
      1306
                3.0
                           0.0
                                                Zakarian, Mr. Mapriededer
                                                                               male
                                                                                      26.5
      1307
                3.0
                           0.0
                                                      Zakarian, Mr. Ortin
                                                                                      27.0
                                                                               male
      1308
                3.0
                           0.0
                                                       Zimmerman, Mr. Leo
                                                                               male
                                                                                      29.0
      1309
                NaN
                           NaN
                                                                        NaN
                                                                                NaN
                                                                                       NaN
             sibsp
                    parch
                           ticket
                                        fare cabin embarked boat
                                                                      body home.dest
      1300
               1.0
                       0.0
                              2659
                                     14.4542
                                                NaN
                                                            C
                                                               NaN
                                                                       NaN
                                                                                  NaN
      1301
                      0.0
                              2628
                                      7.2250
               0.0
                                                NaN
                                                            C
                                                               NaN
                                                                    312.0
                                                                                  NaN
      1302
               0.0
                       0.0
                              2647
                                      7.2250
                                                NaN
                                                            C
                                                               NaN
                                                                       NaN
                                                                                  NaN
                                                            С
      1303
               0.0
                      0.0
                              2627
                                     14.4583
                                                               NaN
                                                                       NaN
                                                                                  NaN
                                                NaN
                                     14.4542
                                                            C
      1304
               1.0
                      0.0
                              2665
                                                NaN
                                                               NaN
                                                                    328.0
                                                                                  NaN
                                                            С
      1305
                       0.0
                              2665
                                     14.4542
                                                               NaN
                                                                       NaN
                                                                                  NaN
               1.0
                                                NaN
                       0.0
                                                            C
                                                               NaN
      1306
               0.0
                              2656
                                      7.2250
                                                NaN
                                                                    304.0
                                                                                  NaN
      1307
               0.0
                       0.0
                              2670
                                      7.2250
                                                NaN
                                                            C
                                                               NaN
                                                                       NaN
                                                                                  NaN
      1308
               0.0
                       0.0
                            315082
                                      7.8750
                                                            S
                                                               NaN
                                                NaN
                                                                       NaN
                                                                                  NaN
      1309
               NaN
                       NaN
                               NaN
                                         NaN
                                                NaN
                                                          NaN
                                                               NaN
                                                                       NaN
                                                                                  NaN
[12]: data.info()
      <class 'pandas.core.frame.DataFrame'>
```

NaN

28

NaN

float64

RangeIndex: 1310 entries, 0 to 1309 Data columns (total 14 columns):

Non-Null Count Dtype

-----1309 non-null

Column

pclass

```
1309 non-null
      1
           survived
                                        float64
      2
                      1309 non-null
           name
                                        object
      3
           sex
                      1309 non-null
                                        object
      4
           age
                      1046 non-null
                                        float64
      5
           sibsp
                      1309 non-null
                                        float64
      6
           parch
                      1309 non-null
                                        float64
      7
           ticket
                      1309 non-null
                                        object
                      1308 non-null
      8
           fare
                                        float64
           cabin
                      295 non-null
                                        object
      10
           embarked
                      1307 non-null
                                        object
                      486 non-null
      11
          boat
                                        object
      12
                      121 non-null
                                        float64
          body
      13
          home.dest
                      745 non-null
                                        object
     dtypes: float64(7), object(7)
     memory usage: 143.4+ KB
[13]: data.describe()
                   pclass
                               survived
                                                              sibsp
                                                                            parch \
                                                  age
      count
             1309.000000
                           1309.000000
                                         1046.000000
                                                       1309.000000
                                                                     1309.000000
                 2.294882
      mean
                               0.381971
                                            29.881135
                                                           0.498854
                                                                         0.385027
      std
                 0.837836
                               0.486055
                                            14.413500
                                                           1.041658
                                                                         0.865560
      min
                               0.000000
                 1.000000
                                             0.166700
                                                           0.000000
                                                                         0.000000
      25%
                 2.000000
                               0.000000
                                            21.000000
                                                           0.000000
                                                                         0.000000
      50%
                 3.000000
                               0.000000
                                            28.000000
                                                          0.000000
                                                                         0.000000
      75%
                 3.000000
                               1.000000
                                            39.000000
                                                           1.000000
                                                                         0.000000
      max
                 3.000000
                               1.000000
                                            80.000000
                                                           8.000000
                                                                         9.000000
                     fare
                                  body
             1308.000000
                            121.000000
      count
                            160.809917
      mean
                33.295479
      std
                51.758668
                            97.696922
      min
                 0.000000
                              1.000000
      25%
                 7.895800
                            72.000000
      50%
                14.454200
                           155.000000
      75%
                31.275000
                            256.000000
      max
               512.329200
                           328.000000
      data.shape
[14]: (1310, 14)
      data.size
[15]: 18340
```

[13]:

[14]:

[15]:

[16]:

data.ndim

[16]: 2

[17]: data.isna() [17]: fare \ pclass survived sibsp parch ticket namesex age 0 False 1 False False False False False False False False 2 False False False False False False False False False 3 False False False False False False False False False 4 False False False False False False False False False ••• 1305 False False False False True False False False False 1306 False False False False False False False False False 1307 False False False False False False False False False 1308 False False False False False False False False False 1309 True True True True True True True True True body cabin embarked boat home.dest 0 False False False True False 1 True False False False False 2 False False True True False 3 False False True False False 4 False False False True True 1305 True False True True True 1306 True False True False True 1307 True True True False True 1308 False True True True True 1309 True True True True True

[1310 rows x 14 columns]

[18]: data.isna().any()

[18]:	pclass	True
	survived	True
	name	True
	sex	True
	age	True
	sibsp	True
	parch	True
	ticket	True
	fare	True
	cabin	True
	${\tt embarked}$	True
	boat	True
	body	True

```
dtype: bool
[19]: data.isna().sum()
[19]: pclass
                       1
      survived
                       1
      name
                       1
      sex
                       1
                     264
      age
                       1
      sibsp
      parch
                       1
      ticket
                       1
      fare
                       2
      cabin
                    1015
      embarked
                       3
      boat
                     824
      body
                    1189
      home.dest
                     565
      dtype: int64
[20]: data["age"].fillna(29.699118)
[20]: 0
              29.000000
      1
               0.916700
      2
               2.000000
      3
              30.000000
      4
              25.000000
      1305
              29.699118
      1306
              26.500000
      1307
              27.000000
      1308
              29.000000
      1309
              29.699118
      Name: age, Length: 1310, dtype: float64
[21]: data.isna().sum()
[21]: pclass
                       1
      survived
                       1
                       1
      name
      sex
                       1
      age
                     264
      sibsp
                       1
      parch
                       1
      ticket
                       1
                       2
      fare
```

home.dest

True

```
boat
                     824
      body
                    1189
      home.dest
                     565
      dtype: int64
[22]: data.any()
[22]: pclass
                    True
      survived
                    True
      name
                    True
                    True
      sex
                    True
      age
      sibsp
                    True
      parch
                    True
      ticket
                    True
      fare
                    True
      cabin
                    True
      embarked
                    True
      boat
                    True
      body
                    True
      home.dest
                    True
      dtype: bool
[23]: data=data.dropna()
[24]: data.any()
[24]: pclass
                    False
      survived
                    False
      name
                    False
      sex
                    False
                    False
      age
                    False
      sibsp
      parch
                    False
      ticket
                    False
      fare
                    False
      cabin
                    False
                    False
      embarked
      boat
                    False
                    False
      body
      home.dest
                    False
      dtype: bool
[25]: data.isna().sum()
```

1015

3

cabin embarked

[25]:	pclass	0.0
	survived	0.0
	name	0.0
	sex	0.0
	age	0.0
	sibsp	0.0
	parch	0.0
	ticket	0.0
	fare	0.0
	cabin	0.0
	embarked	0.0
	boat	0.0
	body	0.0
	home.dest	0.0
	dtvpe: float6	34