

```

In [1]: # Experiment No:6

In [2]: # Aim: Missing value Treatment

In [3]: # Name: Shravani Narendra Mahalle

In [4]: # Class: 3rd year(B)

In [5]: # Roll No: 17

In [6]: # Date: 23/08/24

In [7]: import pandas as pd

In [8]: import os

In [9]: os.getcwd()

Out[9]: 'C:\\Users\\hp'

In [10]: os.chdir('C:\\Users\\hp\\OneDrive\\Desktop')

In [11]: df=pd.read_csv("titanic.csv")

In [12]: df.head(12)

Out[12]:

```

|    | pclass | survived | name  | sex    | age     | sibsp | parch | ticket   | fare     | cabin   | embarked | boat | body  | home.dest                       |
|----|--------|----------|---|--------|---------|-------|-------|----------|----------|---------|----------|------|-------|---------------------------------|
| 0  | 1.0    | 1.0      | Allen, Miss. Elisabeth Walton                     | female | 29.0000 | 0.0   | 0.0   | 24160    | 211.3375 | B5      | S        | 2    | NaN   | St Louis, MO                    |
| 1  | 1.0    | 1.0      | Allison, Master. Hudson Trevor                    | male   | 0.9167  | 1.0   | 2.0   | 113781   | 151.5500 | C22 C26 | S        | 11   | NaN   | Montreal, PQ / Chesterville, ON |
| 2  | 1.0    | 0.0      | Allison, Miss. Helen Loraine                      | female | 2.0000  | 1.0   | 2.0   | 113781   | 151.5500 | C22 C26 | S        | NaN  | NaN   | Montreal, PQ / Chesterville, ON |
| 3  | 1.0    | 0.0      | Allison, Mr. Hudson Joshua Creighton              | male   | 30.0000 | 1.0   | 2.0   | 113781   | 151.5500 | C22 C26 | S        | NaN  | 135.0 | Montreal, PQ / Chesterville, ON |
| 4  | 1.0    | 0.0      | Allison, Mrs. Hudson J C (Bessie Waldo Daniels)   | female | 25.0000 | 1.0   | 2.0   | 113781   | 151.5500 | C22 C26 | S        | NaN  | NaN   | Montreal, PQ / Chesterville, ON |
| 5  | 1.0    | 1.0      | Anderson, Mr. Harry                               | male   | 48.0000 | 0.0   | 0.0   | 19952    | 26.5500  | E12     | S        | 3    | NaN   | New York, NY                    |
| 6  | 1.0    | 1.0      | Andrews, Miss. Kornelia Theodosia                 | female | 63.0000 | 1.0   | 0.0   | 13502    | 77.9583  | D7      | S        | 10   | NaN   | Hudson, NY                      |
| 7  | 1.0    | 0.0      | Andrews, Mr. Thomas Jr                            | male   | 39.0000 | 0.0   | 0.0   | 112050   | 0.0000   | A36     | S        | NaN  | NaN   | Belfast, NI                     |
| 8  | 1.0    | 1.0      | Appleton, Mrs. Edward Dale (Charlotte Lamson)     | female | 53.0000 | 2.0   | 0.0   | 11769    | 51.4792  | C101    | S        | D    | NaN   | Bayside, Queens, NY             |
| 9  | 1.0    | 0.0      | Artagaveytia, Mr. Ramon                           | male   | 71.0000 | 0.0   | 0.0   | PC 17609 | 49.5042  | NaN     | C        | NaN  | 22.0  | Montevideo, Uruguay             |
| 10 | 1.0    | 0.0      | Astor, Col. John Jacob                            | male   | 47.0000 | 1.0   | 0.0   | PC 17757 | 227.5250 | C62 C64 | C        | NaN  | 124.0 | New York, NY                    |
| 11 | 1.0    | 1.0      | Astor, Mrs. John Jacob (Madeleine Talmadge Force) | female | 18.0000 | 1.0   | 0.0   | PC 17757 | 227.5250 | C62 C64 | C        | 4    | NaN   | New York, NY                    |

```

In [13]: df.tail(45)

```

|      | pclass | survived | name  | sex    | age  | sibsp | parch | ticket | fare    | cabin | embarked | boat     | body | home.dest                      |
|------|--------|----------|---|--------|------|-------|-------|--------|---------|-------|----------|----------|------|--------------------------------|
| 1265 | 3.0    | 0.0      | Van Impe, Miss. Catharina                         | female | 10.0 | 0.0   | 2.0   | 345773 | 24.1500 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1266 | 3.0    | 0.0      | Van Impe, Mr. Jean Baptiste                       | male   | 36.0 | 1.0   | 1.0   | 345773 | 24.1500 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1267 | 3.0    | 0.0      | Van Impe, Mrs. Jean Baptiste (Rosalie Paula Go... | female | 30.0 | 1.0   | 1.0   | 345773 | 24.1500 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1268 | 3.0    | 0.0      | van Melkebeke, Mr. Philemon                       | male   | NaN  | 0.0   | 0.0   | 345777 | 9.5000  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1269 | 3.0    | 0.0      | Vande Velde, Mr. Johannes Joseph                  | male   | 33.0 | 0.0   | 0.0   | 345780 | 9.5000  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1270 | 3.0    | 0.0      | Vande Walle, Mr. Nestor Cyriel                    | male   | 28.0 | 0.0   | 0.0   | 345770 | 9.5000  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1271 | 3.0    | 0.0      | Vanden Steen, Mr. Leo Peter                       | male   | 28.0 | 0.0   | 0.0   | 345783 | 9.5000  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1272 | 3.0    | 0.0      | Vander Cruyssen, Mr. Victor                       | male   | 47.0 | 0.0   | 0.0   | 345765 | 9.0000  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1273 | 3.0    | 0.0      | Vander Planke, Miss. Augusta Maria                | female | 18.0 | 2.0   | 0.0   | 345764 | 18.0000 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1274 | 3.0    | 0.0      | Vander Planke, Mr. Julius                         | male   | 31.0 | 3.0   | 0.0   | 345763 | 18.0000 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1275 | 3.0    | 0.0      | Vander Planke, Mr. Leo Edmondus                   | male   | 16.0 | 2.0   | 0.0   | 345764 | 18.0000 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1276 | 3.0    | 0.0      | Vander Planke, Mrs. Julius (Emelia Maria Vande... | female | 31.0 | 1.0   | 0.0   | 345763 | 18.0000 | NaN   | S        | NaN      | NaN  | NaN                            |
| 1277 | 3.0    | 1.0      | Vartanian, Mr. David                              | male   | 22.0 | 0.0   | 0.0   | 2658   | 7.2250  | NaN   | C        | 13<br>15 | NaN  | NaN                            |
| 1278 | 3.0    | 0.0      | Vendel, Mr. Olof Edvin                            | male   | 20.0 | 0.0   | 0.0   | 350416 | 7.8542  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1279 | 3.0    | 0.0      | Vestrom, Miss. Hulda Amanda Adolfina              | female | 14.0 | 0.0   | 0.0   | 350406 | 7.8542  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1280 | 3.0    | 0.0      | Vovk, Mr. Janko                                   | male   | 22.0 | 0.0   | 0.0   | 349252 | 7.8958  | NaN   | S        | NaN      | NaN  | NaN                            |
| 1281 | 3.0    | 0.0      | Waelens, Mr. Achille                              | male   | 22.0 | 0.0   | 0.0   | 345767 | 9.0000  | NaN   | S        | NaN      | NaN  | Antwerp, Belgium / Stanton, OH |
| 1282 | 3.0    | 0.0      | Ware, Mr. Frederick                               | male   | NaN  | 0.0   | 0.0   | 359309 | 8.0500  | NaN   | S        | NaN      | NaN  | NaN                            |

Warren Mr

|      |     |     |   |        |      |     |     |                  |         |     |   |     |       |     |
|------|-----|-----|---|--------|------|-----|-----|------------------|---------|-----|---|-----|-------|-----|
| 1283 | 3.0 | 0.0 | Waller, Mr. Charles William                     | male   | NaN  | 0.0 | 0.0 | C.A. 49867       | 7.5500  | NaN | S | NaN | NaN   | NaN |
| 1284 | 3.0 | 0.0 | Webber, Mr. James                               | male   | NaN  | 0.0 | 0.0 | SOTON/OQ 3101316 | 8.0500  | NaN | S | NaN | NaN   | NaN |
| 1285 | 3.0 | 0.0 | Wenzel, Mr. Linhart                             | male   | 32.5 | 0.0 | 0.0 | 345775           | 9.5000  | NaN | S | NaN | 298.0 | NaN |
| 1286 | 3.0 | 1.0 | Whabee, Mrs. George Joseph (Shawneene Abi-Saab) | female | 38.0 | 0.0 | 0.0 | 2688             | 7.2292  | NaN | C | C   | NaN   | NaN |
| 1287 | 3.0 | 0.0 | Widegren, Mr. Carl/Charles Peter                | male   | 51.0 | 0.0 | 0.0 | 347064           | 7.7500  | NaN | S | NaN | NaN   | NaN |
| 1288 | 3.0 | 0.0 | Wiklund, Mr. Jakob Alfred                       | male   | 18.0 | 1.0 | 0.0 | 3101267          | 6.4958  | NaN | S | NaN | 314.0 | NaN |
| 1289 | 3.0 | 0.0 | Wiklund, Mr. Karl Johan                         | male   | 21.0 | 1.0 | 0.0 | 3101266          | 6.4958  | NaN | S | NaN | NaN   | NaN |
| 1290 | 3.0 | 1.0 | Wilkes, Mrs. James (Ellen Needs)                | female | 47.0 | 1.0 | 0.0 | 363272           | 7.0000  | NaN | S | NaN | NaN   | NaN |
| 1291 | 3.0 | 0.0 | Willer, Mr. Aaron ("Abi Weller")                | male   | NaN  | 0.0 | 0.0 | 3410             | 8.7125  | NaN | S | NaN | NaN   | NaN |
| 1292 | 3.0 | 0.0 | Wiley, Mr. Edward                               | male   | NaN  | 0.0 | 0.0 | S.O./P.P. 751    | 7.5500  | NaN | S | NaN | NaN   | NaN |
| 1293 | 3.0 | 0.0 | Williams, Mr. Howard Hugh "Harry"               | male   | NaN  | 0.0 | 0.0 | A/5 2466         | 8.0500  | NaN | S | NaN | NaN   | NaN |
| 1294 | 3.0 | 0.0 | Williams, Mr. Leslie                            | male   | 28.5 | 0.0 | 0.0 | 54636            | 16.1000 | NaN | S | NaN | 14.0  | NaN |
| 1295 | 3.0 | 0.0 | Windelov, Mr. Einar                             | male   | 21.0 | 0.0 | 0.0 | SOTON/OQ 3101317 | 7.2500  | NaN | S | NaN | NaN   | NaN |
| 1296 | 3.0 | 0.0 | Wirz, Mr. Albert                                | male   | 27.0 | 0.0 | 0.0 | 315154           | 8.6625  | NaN | S | NaN | 131.0 | NaN |
| 1297 | 3.0 | 0.0 | Wiseman, Mr. Phillippe                          | male   | NaN  | 0.0 | 0.0 | A/4. 34244       | 7.2500  | NaN | S | NaN | NaN   | NaN |
| 1298 | 3.0 | 0.0 | Wittevrongel, Mr. Camille                       | male   | 36.0 | 0.0 | 0.0 | 345771           | 9.5000  | NaN | S | NaN | NaN   | NaN |
| 1299 | 3.0 | 0.0 | Yasbeck, Mr. Antoni                             | male   | 27.0 | 1.0 | 0.0 | 2659             | 14.4542 | NaN | C | C   | NaN   | NaN |
| 1300 | 3.0 | 1.0 | Yasbeck, Mrs. Antoni (Selini Alexander)         | female | 15.0 | 1.0 | 0.0 | 2659             | 14.4542 | NaN | C | NaN | NaN   | NaN |
| 1301 | 3.0 | 0.0 | Youseff, Mr. Gerious                            | male   | 45.5 | 0.0 | 0.0 | 2628             | 7.2250  | NaN | C | NaN | 312.0 | NaN |
| 1302 | 3.0 | 0.0 | Yousif, Mr. Wazli                               | male   | NaN  | 0.0 | 0.0 | 2647             | 7.2250  | NaN | C | NaN | NaN   | NaN |
| 1303 | 3.0 | 0.0 | Yousseff, Mr. Gerious                           | male   | NaN  | 0.0 | 0.0 | 2627             | 14.4583 | NaN | C | NaN | NaN   | NaN |
| 1304 | 3.0 | 0.0 | Zabour, Miss. Hileni                            | female | 14.5 | 1.0 | 0.0 | 2665             | 14.4542 | NaN | C | NaN | 328.0 | NaN |
| 1305 | 3.0 | 0.0 | Zabour, Miss. Thamine                           | female | NaN  | 1.0 | 0.0 | 2665             | 14.4542 | NaN | C | NaN | NaN   | NaN |

|      |     |     |                              |      |      |     |     |        |        |     |     |     |       |     |
|------|-----|-----|------------------------------|------|------|-----|-----|--------|--------|-----|-----|-----|-------|-----|
| 1306 | 3.0 | 0.0 | Zakarian,<br>Mr. Mapriededer | male | 26.5 | 0.0 | 0.0 | 2656   | 7.2250 | NaN | C   | NaN | 304.0 | NaN |
| 1307 | 3.0 | 0.0 | Zakarian,<br>Mr. Ortin       | male | 27.0 | 0.0 | 0.0 | 2670   | 7.2250 | NaN | C   | NaN | NaN   | NaN |
| 1308 | 3.0 | 0.0 | Zimmerman,<br>Mr. Leo        | male | 29.0 | 0.0 | 0.0 | 315082 | 7.8750 | NaN | S   | NaN | NaN   | NaN |
| 1309 | NaN | NaN | NaN                          | NaN  | NaN  | NaN | NaN | NaN    | NaN    | NaN | NaN | NaN | NaN   | NaN |

In [14]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1310 entries, 0 to 1309
Data columns (total 14 columns):
#   Column      Non-Null Count  Dtype
---  -
0    pclass      1309 non-null   float64
1    survived    1309 non-null   float64
2    name        1309 non-null   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
8    fare        1308 non-null   float64
9    cabin       295 non-null    object
10   embarked    1307 non-null   object
11   boat        486 non-null    object
12   body        121 non-null    float64
13   home.dest   745 non-null    object
dtypes: float64(7), object(7)
memory usage: 143.4+ KB
```

In [15]: `df.describe()`

Out[15]:

|       | pclass      | survived    | age         | sibsp       | parch       | fare        | body       |
|-------|-------------|-------------|-------------|-------------|-------------|-------------|------------|
| count | 1309.000000 | 1309.000000 | 1046.000000 | 1309.000000 | 1309.000000 | 1308.000000 | 121.000000 |
| mean  | 2.294882    | 0.381971    | 29.881135   | 0.498854    | 0.385027    | 33.295479   | 160.809917 |
| std   | 0.837836    | 0.486055    | 14.413500   | 1.041658    | 0.865560    | 51.758668   | 97.696922  |
| min   | 1.000000    | 0.000000    | 0.166700    | 0.000000    | 0.000000    | 0.000000    | 1.000000   |
| 25%   | 2.000000    | 0.000000    | 21.000000   | 0.000000    | 0.000000    | 7.895800    | 72.000000  |
| 50%   | 3.000000    | 0.000000    | 28.000000   | 0.000000    | 0.000000    | 14.454200   | 155.000000 |
| 75%   | 3.000000    | 1.000000    | 39.000000   | 1.000000    | 0.000000    | 31.275000   | 256.000000 |
| max   | 3.000000    | 1.000000    | 80.000000   | 8.000000    | 9.000000    | 512.329200  | 328.000000 |

In [16]: `df.shape`

Out[16]: (1310, 14)

In [17]: `df.size`

Out[17]: 18340

In [18]: `df.ndim`

Out[18]: 2

In [19]: `df.isnull()`

Out[19]:

|      | pclass | survived | name  | sex   | age   | sibsp | parch | ticket | fare  | cabin | embarked | boat  | body  | home.dest |
|------|--------|----------|-------|-------|-------|-------|-------|--------|-------|-------|----------|-------|-------|-----------|
| 0    | False  | False    | False | False | False | False | False | False  | False | False | False    | False | True  | False     |
| 1    | False  | False    | False | False | False | False | False | False  | False | False | False    | False | True  | False     |
| 2    | False  | False    | False | False | False | False | False | False  | False | False | False    | True  | True  | False     |
| 3    | False  | False    | False | False | False | False | False | False  | False | False | False    | True  | False | False     |
| 4    | False  | False    | False | False | False | False | False | False  | False | False | False    | True  | True  | False     |
| ...  | ...    | ...      | ...   | ...   | ...   | ...   | ...   | ...    | ...   | ...   | ...      | ...   | ...   | ...       |
| 1305 | False  | False    | False | False | True  | False | False | False  | False | True  | False    | True  | True  | True      |
| 1306 | False  | False    | False | False | False | False | False | False  | False | True  | False    | True  | False | True      |
| 1307 | False  | False    | False | False | False | False | False | False  | False | True  | False    | True  | True  | True      |
| 1308 | False  | False    | False | False | False | False | False | False  | False | True  | False    | True  | True  | True      |
| 1309 | True   | True     | True  | True  | True  | True  | True  | True   | True  | True  | True     | True  | True  | True      |

1310 rows × 14 columns

In [20]:

df.isnull().any()

Out[20]:

pclass

survived

name

sex

age

sibsp

parch

ticket

fare

cabin

embarked

boat

body

home.dest

dtype: bool

In [21]:

df.isnull().sum()

Out[21]:

pclass

survived

name

sex

age

sibsp

parch

ticket

fare

cabin

embarked

boat

body

home.dest

dtype: int64

In [24]:

```

-----
KeyError                                Traceback (most recent call last)
File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3805, in Index.get_loc(self, key)
    3804 try:
-> 3805     return self._engine.get_loc(casted_key)
    3806 except KeyError as err:

File index.pyx:167, in pandas._libs.index.IndexEngine.get_loc()

File index.pyx:196, in pandas._libs.index.IndexEngine.get_loc()

File pandas\_libs\hashtable_class_helper.pxi:7081, in pandas._libs.hashtable.PyObjectHashTable.get_item()

File pandas\_libs\hashtable_class_helper.pxi:7089, in pandas._libs.hashtable.PyObjectHashTable.get_item()

KeyError: 'Age'

The above exception was the direct cause of the following exception:

KeyError                                Traceback (most recent call last)
Cell In[24], line 1
----> 1 df["Age"].fillna(25.0000)

File ~\anaconda3\Lib\site-packages\pandas\core\frame.py:4102, in DataFrame.__getitem__(self, key)
    4100 if self.columns.nlevels > 1:
    4101     return self._getitem_multilevel(key)
-> 4102 indexer = self.columns.get_loc(key)
    4103 if is_integer(indexer):
    4104     indexer = [indexer]

File ~\anaconda3\Lib\site-packages\pandas\core\indexes\base.py:3812, in Index.get_loc(self, key)
    3807 if isinstance(casted_key, slice) or (
    3808     isinstance(casted_key, abc.Iterable)
    3809     and any(isinstance(x, slice) for x in casted_key)
    3810 ):
    3811     raise InvalidIndexError(key)
-> 3812     raise KeyError(key) from err
    3813 except TypeError:
    3814     # If we have a listlike key, _check_indexing_error will raise
    3815     # InvalidIndexError. Otherwise we fall through and re-raise
    3816     # the TypeError.
    3817     self._check_indexing_error(key)

KeyError: 'Age'

```

In [ ]:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js