pandas library

```python
In [1]: import pandas as pd
```

```python
In [3]: df = pd.read_csv("https://raw.githubusercontent.com/datasciencedojo/datasets/ma
```

In [4]: `df`

Out[4]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 12 columns

In [5]: `df.head()`

Out[5]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cal
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	N
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	N

In [6]: `df.tail()`

Out[6]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.00	NaN
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.00	B42
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.45	NaN
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.00	C148
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.75	NaN

In [7]: `df`

Out[7]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 12 columns

In [8]: `df.dtypes`

Out[8]:
```
PassengerId      int64
Survived         int64
Pclass           int64
Name            object
Sex             object
Age            float64
SibSp            int64
Parch            int64
Ticket          object
Fare           float64
Cabin           object
Embarked        object
dtype: object
```

In [9]: `df.describe()`

Out[9]:

	PassengerId	Survived	Pclass	Age	SibSp	Parch	Fare
count	891.000000	891.000000	891.000000	714.000000	891.000000	891.000000	891.000000
mean	446.000000	0.383838	2.308642	29.699118	0.523008	0.381594	32.204208
std	257.353842	0.486592	0.836071	14.526497	1.102743	0.806057	49.693429
min	1.000000	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
25%	223.500000	0.000000	2.000000	20.125000	0.000000	0.000000	7.910400
50%	446.000000	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	668.500000	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
max	891.000000	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

In [10]: `df`

Out[10]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 12 columns

In [13]: `df[['Name','Sex','Ticket','Cabin','Embarked']]`

Out[13]:

	Name	Sex	Ticket	Cabin	Embarked
0	Braund, Mr. Owen Harris	male	A/5 21171	NaN	S
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	PC 17599	C85	C
2	Heikkinen, Miss. Laina	female	STON/O2. 3101282	NaN	S
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	113803	C123	S
4	Allen, Mr. William Henry	male	373450	NaN	S
...	...	...	...	...	...
886	Montvila, Rev. Juozas	male	211536	NaN	S
887	Graham, Miss. Margaret Edith	female	112053	B42	S
888	Johnston, Miss. Catherine Helen "Carrie"	female	W./C. 6607	NaN	S
889	Behr, Mr. Karl Howell	male	111369	C148	C
890	Dooley, Mr. Patrick	male	370376	NaN	Q

891 rows × 5 columns

In [14]: `df.dtypes == 'object'`

Out[14]:
```
PassengerId    False
Survived       False
Pclass         False
Name            True
Sex             True
Age            False
SibSp          False
Parch          False
Ticket          True
Fare           False
Cabin           True
Embarked        True
dtype: bool
```

In [19]: `df[df.dtypes[df.dtypes == 'float64'].index]`

Out[19]:

	Age	Fare
0	22.0	7.2500
1	38.0	71.2833
2	26.0	7.9250
3	35.0	53.1000
4	35.0	8.0500
...	...	...
886	27.0	13.0000
887	19.0	30.0000
888	NaN	23.4500
889	26.0	30.0000
890	32.0	7.7500

891 rows × 2 columns

In [20]: `df[df.dtypes[df.dtypes == 'int64'].index]`

Out[20]:

	PassengerId	Survived	Pclass	SibSp	Parch
0	1	0	3	1	0
1	2	1	1	1	0
2	3	1	3	0	0
3	4	1	1	1	0
4	5	0	3	0	0
...	...	...	...	...	...
886	887	0	2	0	0
887	888	1	1	0	0
888	889	0	3	1	2
889	890	1	1	0	0
890	891	0	3	0	0

891 rows × 5 columns

In [21]: `df`

Out[21]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	C
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 12 columns

In [25]: `df.head()`

Out[25]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cal
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	N
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	N
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	N

```
In [29]: df.tail(15)
```

Out[29]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Far
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.845
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.895
878	879	0	3	Laleff, Mr. Kristo	male	NaN	0	0	349217	7.895
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.158
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.895
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.516
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.500
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.050
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.125
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.000
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.000
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.450
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.000
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.750

```
In [30]: df.columns
```

```
Out[30]: Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age', 'SibSp',
                'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
               dtype='object')
```

```
In [31]: df[['Ticket']]
```

Out[31]:

	Ticket
0	A/5 21171
1	PC 17599
2	STON/O2. 3101282
3	113803
4	373450
...	...
886	211536
887	112053
888	W./C. 6607
889	111369
890	370376

891 rows × 1 columns

```
In [32]: df[['Ticket']][4:16]
```

Out[32]:

	Ticket
4	373450
5	330877
6	17463
7	349909
8	347742
9	237736
10	PP 9549
11	113783
12	A/5. 2151
13	347082
14	350406
15	248706

In [33]: `df[['Ticket']][4:16:2]`

Out[33]:

	Ticket
4	373450
6	17463
8	347742
10	PP 9549
12	A/5. 2151
14	350406

In [34]: `df[['Ticket','Cabin']]`

Out[34]:

	Ticket	Cabin
0	A/5 21171	NaN
1	PC 17599	C85
2	STON/O2. 3101282	NaN
3	113803	C123
4	373450	NaN
...	...	...
886	211536	NaN
887	112053	B42
888	W./C. 6607	NaN
889	111369	C148
890	370376	NaN

891 rows × 2 columns

In [35]: `df[['Ticket','Cabin']][4:17]`

Out[35]:

	Ticket	Cabin
4	373450	NaN
5	330877	NaN
6	17463	E46
7	349909	NaN
8	347742	NaN
9	237736	NaN
10	PP 9549	G6
11	113783	C103
12	A/5. 2151	NaN
13	347082	NaN
14	350406	NaN
15	248706	NaN
16	382652	NaN

In [36]: `df[['Ticket','Cabin']][4:17:3]`

Out[36]:

	Ticket	Cabin
4	373450	NaN
7	349909	NaN
10	PP 9549	G6
13	347082	NaN
16	382652	NaN

In [37]: `df[['Ticket','Cabin','Name']][4:17:3]`

Out[37]:

	Ticket	Cabin	Name
4	373450	NaN	Allen, Mr. William Henry
7	349909	NaN	Palsson, Master. Gosta Leonard
10	PP 9549	G6	Sandstrom, Miss. Marguerite Rut
13	347082	NaN	Andersson, Mr. Anders Johan
16	382652	NaN	Rice, Master. Eugene

In [38]: `df`

Out[38]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 12 columns

In [45]: `df['swati'] = 2.5`

In [46]: df

Out[46]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500	
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	

891 rows × 14 columns

In [47]: df

Out[47]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500
...	...	...	...	...	...	...	...	...	...	...
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4500
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500

891 rows × 14 columns

In [48]: df.insert(loc = 3, column = 'Food', value = 0)

In [49]: `df`

Out[49]:

	PassengerId	Survived	Pclass	Food	Name	Sex	Age	SibSp	Parch	Ticket	
0	1	0	3	0	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.:
1	2	1	1	0	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.:
2	3	1	3	0	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.!
3	4	1	1	0	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.
4	5	0	3	0	Allen, Mr. William Henry	male	35.0	0	0	373450	8.(
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	0	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.(
887	888	1	1	0	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.(
888	889	0	3	0	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W./C. 6607	23.4
889	890	1	1	0	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.(
890	891	0	3	0	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.:

891 rows × 15 columns

In [54]: `df.insert(loc = 6, column = 'equipments', value = "oxygen cylinder")`

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[54], line 1
----> 1 df.insert(loc = 6, column = 'equipments', value = "oxygen cylinder")

File C:\ProgramData\anaconda3\lib\site-packages\pandas\core\frame.py:4817, in
DataFrame.insert(self, loc, column, value, allow_duplicates)
   4811        raise ValueError(
   4812            "Cannot specify 'allow_duplicates=True' when "
   4813            "'self.flags.allows_duplicate_labels' is False."
   4814        )
   4815 if not allow_duplicates and column in self.columns:
   4816        # Should this be a different kind of error??
-> 4817        raise ValueError(f"cannot insert {column}, already exists")
   4818 if not isinstance(loc, int):
   4819        raise TypeError("loc must be int")

ValueError: cannot insert equipments, already exists
```

In [53]: `df`

Out[53]:

	PassengerId	Survived	Pclass	Food	Name	Sex	equipments	Age	SibSp	Parch	
0	1	0	3	0	Braund, Mr. Owen Harris	male	oxygenn cylinder	22.0	1	0	A
1	2	1	1	0	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	oxygenn cylinder	38.0	1	0	F
2	3	1	3	0	Heikkinen, Miss. Laina	female	oxygenn cylinder	26.0	0	0	S
3	4	1	1	0	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	oxygenn cylinder	35.0	1	0	
4	5	0	3	0	Allen, Mr. William Henry	male	oxygenn cylinder	35.0	0	0	
...	...	...	...	...	...	...	...	...	...	...	
886	887	0	2	0	Montvila, Rev. Juozas	male	oxygenn cylinder	27.0	0	0	
887	888	1	1	0	Graham, Miss. Margaret Edith	female	oxygenn cylinder	19.0	0	0	
888	889	0	3	0	Johnston, Miss. Catherine Helen "Carrie"	female	oxygenn cylinder	NaN	1	2	
889	890	1	1	0	Behr, Mr. Karl Howell	male	oxygenn cylinder	26.0	0	0	
890	891	0	3	0	Dooley, Mr. Patrick	male	oxygenn cylinder	32.0	0	0	

891 rows × 16 columns

In [ ]: