

Dummyvars

November 20, 2016

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

```
In [3]: train = pd.read_csv('http://bit.ly/kaggletrain')
```

```
In [4]: train.head(2)
```

```
Out[4]:
```

	PassengerId	Survived	Pclass	\	Name	Sex	Age	SibSp
0	1	0	3		Braund, Mr. Owen Harris	male	22.0	1
1	2	1	1		Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C

```
In [5]: train['Sex_male'] = train.Sex.map({'female' : 0, 'male' : 1})
```

```
In [6]: train.head(2)
```

```
Out[6]:
```

	PassengerId	Survived	Pclass	\	Name	Sex	Age	SibSp
0	1	0	3		Braund, Mr. Owen Harris	male	22.0	1
1	2	1	1		Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1

	Parch	Ticket	Fare	Cabin	Embarked	Sex_male
0	0	A/5 21171	7.2500	NaN	S	1
1	0	PC 17599	71.2833	C85	C	0

```
In [7]: pd.get_dummies(train.Sex)
```

```
Out[7]:
```

	female	male
0	0.0	1.0

1	1.0	0.0
2	1.0	0.0
3	1.0	0.0
4	0.0	1.0
5	0.0	1.0
6	0.0	1.0
7	0.0	1.0
8	1.0	0.0
9	1.0	0.0
10	1.0	0.0
11	1.0	0.0
12	0.0	1.0
13	0.0	1.0
14	1.0	0.0
15	1.0	0.0
16	0.0	1.0
17	0.0	1.0
18	1.0	0.0
19	1.0	0.0
20	0.0	1.0
21	0.0	1.0
22	1.0	0.0
23	0.0	1.0
24	1.0	0.0
25	1.0	0.0
26	0.0	1.0
27	0.0	1.0
28	1.0	0.0
29	0.0	1.0
..
861	0.0	1.0
862	1.0	0.0
863	1.0	0.0
864	0.0	1.0
865	1.0	0.0
866	1.0	0.0
867	0.0	1.0
868	0.0	1.0
869	0.0	1.0
870	0.0	1.0
871	1.0	0.0
872	0.0	1.0
873	0.0	1.0
874	1.0	0.0
875	1.0	0.0
876	0.0	1.0
877	0.0	1.0
878	0.0	1.0

879	1.0	0.0
880	1.0	0.0
881	0.0	1.0
882	1.0	0.0
883	0.0	1.0
884	0.0	1.0
885	1.0	0.0
886	0.0	1.0
887	1.0	0.0
888	1.0	0.0
889	0.0	1.0
890	0.0	1.0

[891 rows x 2 columns]

```
In [9]: # prefix argument => prefix = Sex
pd.get_dummies(train.Sex, prefix='Sex').iloc[:, 1:]
```

```
Out[9]:
```

	Sex_male
0	1.0
1	0.0
2	0.0
3	0.0
4	1.0
5	1.0
6	1.0
7	1.0
8	0.0
9	0.0
10	0.0
11	0.0
12	1.0
13	1.0
14	0.0
15	0.0
16	1.0
17	1.0
18	0.0
19	0.0
20	1.0
21	1.0
22	0.0
23	1.0
24	0.0
25	0.0
26	1.0
27	1.0
28	0.0

```

29      1.0
..      ...
861     1.0
862     0.0
863     0.0
864     1.0
865     0.0
866     0.0
867     1.0
868     1.0
869     1.0
870     1.0
871     0.0
872     1.0
873     1.0
874     0.0
875     0.0
876     1.0
877     1.0
878     1.0
879     0.0
880     0.0
881     1.0
882     0.0
883     1.0
884     1.0
885     0.0
886     1.0
887     0.0
888     0.0
889     1.0
890     1.0

```

```
[891 rows x 1 columns]
```

```
In [10]: train.Embarked.value_counts()
```

```

Out[10]: S      644
         C      168
         Q       77
         Name: Embarked, dtype: int64

```

```
In [11]: pd.get_dummies(train.Embarked, prefix = 'Embarked')
```

```

Out[11]:   Embarked_C  Embarked_Q  Embarked_S
0           0.0           0.0           1.0
1           1.0           0.0           0.0
2           0.0           0.0           1.0
3           0.0           0.0           1.0

```

4	0.0	0.0	1.0
5	0.0	1.0	0.0
6	0.0	0.0	1.0
7	0.0	0.0	1.0
8	0.0	0.0	1.0
9	1.0	0.0	0.0
10	0.0	0.0	1.0
11	0.0	0.0	1.0
12	0.0	0.0	1.0
13	0.0	0.0	1.0
14	0.0	0.0	1.0
15	0.0	0.0	1.0
16	0.0	1.0	0.0
17	0.0	0.0	1.0
18	0.0	0.0	1.0
19	1.0	0.0	0.0
20	0.0	0.0	1.0
21	0.0	0.0	1.0
22	0.0	1.0	0.0
23	0.0	0.0	1.0
24	0.0	0.0	1.0
25	0.0	0.0	1.0
26	1.0	0.0	0.0
27	0.0	0.0	1.0
28	0.0	1.0	0.0
29	0.0	0.0	1.0
..
861	0.0	0.0	1.0
862	0.0	0.0	1.0
863	0.0	0.0	1.0
864	0.0	0.0	1.0
865	0.0	0.0	1.0
866	1.0	0.0	0.0
867	0.0	0.0	1.0
868	0.0	0.0	1.0
869	0.0	0.0	1.0
870	0.0	0.0	1.0
871	0.0	0.0	1.0
872	0.0	0.0	1.0
873	0.0	0.0	1.0
874	1.0	0.0	0.0
875	1.0	0.0	0.0
876	0.0	0.0	1.0
877	0.0	0.0	1.0
878	0.0	0.0	1.0
879	1.0	0.0	0.0
880	0.0	0.0	1.0
881	0.0	0.0	1.0

882	0.0	0.0	1.0
883	0.0	0.0	1.0
884	0.0	0.0	1.0
885	0.0	1.0	0.0
886	0.0	0.0	1.0
887	0.0	0.0	1.0
888	0.0	0.0	1.0
889	1.0	0.0	0.0
890	0.0	1.0	0.0

[891 rows x 3 columns]

```
In [12]: pd.get_dummies(train.Embarked, prefix = 'Embarked').iloc[:, 1:]
```

```
Out[12]:
```

	Embarked_Q	Embarked_S
0	0.0	1.0
1	0.0	0.0
2	0.0	1.0
3	0.0	1.0
4	0.0	1.0
5	1.0	0.0
6	0.0	1.0
7	0.0	1.0
8	0.0	1.0
9	0.0	0.0
10	0.0	1.0
11	0.0	1.0
12	0.0	1.0
13	0.0	1.0
14	0.0	1.0
15	0.0	1.0
16	1.0	0.0
17	0.0	1.0
18	0.0	1.0
19	0.0	0.0
20	0.0	1.0
21	0.0	1.0
22	1.0	0.0
23	0.0	1.0
24	0.0	1.0
25	0.0	1.0
26	0.0	0.0
27	0.0	1.0
28	1.0	0.0
29	0.0	1.0
...
861	0.0	1.0
862	0.0	1.0

863	0.0	1.0
864	0.0	1.0
865	0.0	1.0
866	0.0	0.0
867	0.0	1.0
868	0.0	1.0
869	0.0	1.0
870	0.0	1.0
871	0.0	1.0
872	0.0	1.0
873	0.0	1.0
874	0.0	0.0
875	0.0	0.0
876	0.0	1.0
877	0.0	1.0
878	0.0	1.0
879	0.0	0.0
880	0.0	1.0
881	0.0	1.0
882	0.0	1.0
883	0.0	1.0
884	0.0	1.0
885	1.0	0.0
886	0.0	1.0
887	0.0	1.0
888	0.0	1.0
889	0.0	0.0
890	1.0	0.0

[891 rows x 2 columns]

In [13]: embarked_dummies = pd.get_dummies(train.Embarked, prefix = 'Embarked').iloc

In [14]: train = pd.concat([train, embarked_dummies], axis =1)

In [15]: train.head(2)

Out[15]:

	PassengerId	Survived	Pclass	\	Name	Sex	Age	SibSp	
0	1	0	3		Braund, Mr. Owen Harris	male	22.0	1	
1	2	1	1		Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	

	Parch	Ticket	Fare	Cabin	Embarked	Sex_male	Embarked_Q	Embarke
0	0	A/5 21171	7.2500	NaN	S	1	0.0	
1	0	PC 17599	71.2833	C85	C	0	0.0	

```
In [16]: train1 = pd.read_csv('http://bit.ly/kaggletrain')
```

```
In [17]: train1.head(2)
```

```
Out[17]:
```

	PassengerId	Survived	Pclass	\	Name	Sex	Age	SibSp
0	1	0	3					
1	2	1	1					

	Parch	Ticket	Fare	Cabin	Embarked
0	0	A/5 21171	7.2500	NaN	S
1	0	PC 17599	71.2833	C85	C

```
In [18]: # dropped Sex and Embarked and replace with Dummy columns
pd.get_dummies(train, columns = ['Sex', 'Embarked'])
```

```
Out[18]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
5	6	0	3	
6	7	0	1	
7	8	0	3	
8	9	1	3	
9	10	1	2	
10	11	1	3	
11	12	1	1	
12	13	0	3	
13	14	0	3	
14	15	0	3	
15	16	1	2	
16	17	0	3	
17	18	1	2	
18	19	0	3	
19	20	1	3	
20	21	0	2	
21	22	1	2	
22	23	1	3	
23	24	1	1	
24	25	0	3	
25	26	1	3	
26	27	0	3	
27	28	0	1	
28	29	1	3	

29	30	0	3
..
861	862	0	2
862	863	1	1
863	864	0	3
864	865	0	2
865	866	1	2
866	867	1	2
867	868	0	1
868	869	0	3
869	870	1	3
870	871	0	3
871	872	1	1
872	873	0	1
873	874	0	3
874	875	1	2
875	876	1	3
876	877	0	3
877	878	0	3
878	879	0	3
879	880	1	1
880	881	1	2
881	882	0	3
882	883	0	3
883	884	0	2
884	885	0	3
885	886	0	3
886	887	0	2
887	888	1	1
888	889	0	3
889	890	1	1
890	891	0	3

	Name	Age	SibSp	Parch
0	Braund, Mr. Owen Harris	22.0	1	0
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	38.0	1	0
2	Heikkinen, Miss. Laina	26.0	0	0
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	35.0	1	0
4	Allen, Mr. William Henry	35.0	0	0
5	Moran, Mr. James	NaN	0	0
6	McCarthy, Mr. Timothy J	54.0	0	0
7	Palsson, Master. Gosta Leonard	2.0	3	1
8	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	27.0	0	2
9	Nasser, Mrs. Nicholas (Adele Achem)	14.0	1	0
10	Sandstrom, Miss. Marguerite Rut	4.0	1	1
11	Bonnell, Miss. Elizabeth	58.0	0	0
12	Saunderscock, Mr. William Henry	20.0	0	0
13	Andersson, Mr. Anders Johan	39.0	1	5

14	Vestrom, Miss. Hulda Amanda Adolfina	14.0	0	0
15	Hewlett, Mrs. (Mary D Kingcome)	55.0	0	0
16	Rice, Master. Eugene	2.0	4	1
17	Williams, Mr. Charles Eugene	NaN	0	0
18	Vander Planke, Mrs. Julius (Emelia Maria Vande...	31.0	1	0
19	Masselmani, Mrs. Fatima	NaN	0	0
20	Fynney, Mr. Joseph J	35.0	0	0
21	Beesley, Mr. Lawrence	34.0	0	0
22	McGowan, Miss. Anna "Annie"	15.0	0	0
23	Sloper, Mr. William Thompson	28.0	0	0
24	Palsson, Miss. Torborg Danira	8.0	3	1
25	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	38.0	1	5
26	Emir, Mr. Farred Chehab	NaN	0	0
27	Fortune, Mr. Charles Alexander	19.0	3	2
28	O'Dwyer, Miss. Ellen "Nellie"	NaN	0	0
29	Todoroff, Mr. Lalio	NaN	0	0
..
861	Giles, Mr. Frederick Edward	21.0	1	0
862	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	48.0	0	0
863	Sage, Miss. Dorothy Edith "Dolly"	NaN	8	2
864	Gill, Mr. John William	24.0	0	0
865	Bystrom, Mrs. (Karolina)	42.0	0	0
866	Duran y More, Miss. Asuncion	27.0	1	0
867	Roebling, Mr. Washington Augustus II	31.0	0	0
868	van Melkebeke, Mr. Philemon	NaN	0	0
869	Johnson, Master. Harold Theodor	4.0	1	1
870	Balkic, Mr. Cerin	26.0	0	0
871	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	47.0	1	1
872	Carlsson, Mr. Frans Olof	33.0	0	0
873	Vander Cruyssen, Mr. Victor	47.0	0	0
874	Abelson, Mrs. Samuel (Hannah Wizosky)	28.0	1	0
875	Najib, Miss. Adele Kiamie "Jane"	15.0	0	0
876	Gustafsson, Mr. Alfred Ossian	20.0	0	0
877	Petroff, Mr. Nedelio	19.0	0	0
878	Laleff, Mr. Kristo	NaN	0	0
879	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	56.0	0	1
880	Shelley, Mrs. William (Imanita Parrish Hall)	25.0	0	1
881	Markun, Mr. Johann	33.0	0	0
882	Dahlberg, Miss. Gerda Ulrika	22.0	0	0
883	Banfield, Mr. Frederick James	28.0	0	0
884	Sutehall, Mr. Henry Jr	25.0	0	0
885	Rice, Mrs. William (Margaret Norton)	39.0	0	5
886	Montvila, Rev. Juozas	27.0	0	0
887	Graham, Miss. Margaret Edith	19.0	0	0
888	Johnston, Miss. Catherine Helen "Carrie"	NaN	1	2
889	Behr, Mr. Karl Howell	26.0	0	0
890	Dooley, Mr. Patrick	32.0	0	0

	Ticket	Fare	Cabin	Sex_male	Embarked_Q	\
0	A/5 21171	7.2500	NaN	1	0.0	
1	PC 17599	71.2833	C85	0	0.0	
2	STON/O2. 3101282	7.9250	NaN	0	0.0	
3	113803	53.1000	C123	0	0.0	
4	373450	8.0500	NaN	1	0.0	
5	330877	8.4583	NaN	1	1.0	
6	17463	51.8625	E46	1	0.0	
7	349909	21.0750	NaN	1	0.0	
8	347742	11.1333	NaN	0	0.0	
9	237736	30.0708	NaN	0	0.0	
10	PP 9549	16.7000	G6	0	0.0	
11	113783	26.5500	C103	0	0.0	
12	A/5. 2151	8.0500	NaN	1	0.0	
13	347082	31.2750	NaN	1	0.0	
14	350406	7.8542	NaN	0	0.0	
15	248706	16.0000	NaN	0	0.0	
16	382652	29.1250	NaN	1	1.0	
17	244373	13.0000	NaN	1	0.0	
18	345763	18.0000	NaN	0	0.0	
19	2649	7.2250	NaN	0	0.0	
20	239865	26.0000	NaN	1	0.0	
21	248698	13.0000	D56	1	0.0	
22	330923	8.0292	NaN	0	1.0	
23	113788	35.5000	A6	1	0.0	
24	349909	21.0750	NaN	0	0.0	
25	347077	31.3875	NaN	0	0.0	
26	2631	7.2250	NaN	1	0.0	
27	19950	263.0000	C23 C25 C27	1	0.0	
28	330959	7.8792	NaN	0	1.0	
29	349216	7.8958	NaN	1	0.0	
..	
861	28134	11.5000	NaN	1	0.0	
862	17466	25.9292	D17	0	0.0	
863	CA. 2343	69.5500	NaN	0	0.0	
864	233866	13.0000	NaN	1	0.0	
865	236852	13.0000	NaN	0	0.0	
866	SC/PARIS 2149	13.8583	NaN	0	0.0	
867	PC 17590	50.4958	A24	1	0.0	
868	345777	9.5000	NaN	1	0.0	
869	347742	11.1333	NaN	1	0.0	
870	349248	7.8958	NaN	1	0.0	
871	11751	52.5542	D35	0	0.0	
872	695	5.0000	B51 B53 B55	1	0.0	
873	345765	9.0000	NaN	1	0.0	
874	P/PP 3381	24.0000	NaN	0	0.0	
875	2667	7.2250	NaN	0	0.0	
876	7534	9.8458	NaN	1	0.0	

877		349212	7.8958	NaN	1	0.0
878		349217	7.8958	NaN	1	0.0
879		11767	83.1583	C50	0	0.0
880		230433	26.0000	NaN	0	0.0
881		349257	7.8958	NaN	1	0.0
882		7552	10.5167	NaN	0	0.0
883	C.A./SOTON	34068	10.5000	NaN	1	0.0
884	SOTON/OQ	392076	7.0500	NaN	1	0.0
885		382652	29.1250	NaN	0	1.0
886		211536	13.0000	NaN	1	0.0
887		112053	30.0000	B42	0	0.0
888	W./C.	6607	23.4500	NaN	0	0.0
889		111369	30.0000	C148	1	0.0
890		370376	7.7500	NaN	1	1.0

	Embarked_S	Sex_female	Sex_male	Embarked_C	Embarked_Q	Embarked_S
0	1.0	0.0	1.0	0.0	0.0	1.0
1	0.0	1.0	0.0	1.0	0.0	0.0
2	1.0	1.0	0.0	0.0	0.0	1.0
3	1.0	1.0	0.0	0.0	0.0	1.0
4	1.0	0.0	1.0	0.0	0.0	1.0
5	0.0	0.0	1.0	0.0	1.0	0.0
6	1.0	0.0	1.0	0.0	0.0	1.0
7	1.0	0.0	1.0	0.0	0.0	1.0
8	1.0	1.0	0.0	0.0	0.0	1.0
9	0.0	1.0	0.0	1.0	0.0	0.0
10	1.0	1.0	0.0	0.0	0.0	1.0
11	1.0	1.0	0.0	0.0	0.0	1.0
12	1.0	0.0	1.0	0.0	0.0	1.0
13	1.0	0.0	1.0	0.0	0.0	1.0
14	1.0	1.0	0.0	0.0	0.0	1.0
15	1.0	1.0	0.0	0.0	0.0	1.0
16	0.0	0.0	1.0	0.0	1.0	0.0
17	1.0	0.0	1.0	0.0	0.0	1.0
18	1.0	1.0	0.0	0.0	0.0	1.0
19	0.0	1.0	0.0	1.0	0.0	0.0
20	1.0	0.0	1.0	0.0	0.0	1.0
21	1.0	0.0	1.0	0.0	0.0	1.0
22	0.0	1.0	0.0	0.0	1.0	0.0
23	1.0	0.0	1.0	0.0	0.0	1.0
24	1.0	1.0	0.0	0.0	0.0	1.0
25	1.0	1.0	0.0	0.0	0.0	1.0
26	0.0	0.0	1.0	1.0	0.0	0.0
27	1.0	0.0	1.0	0.0	0.0	1.0
28	0.0	1.0	0.0	0.0	1.0	0.0
29	1.0	0.0	1.0	0.0	0.0	1.0
..
861	1.0	0.0	1.0	0.0	0.0	1.0

862	1.0	1.0	0.0	0.0	0.0	1.0
863	1.0	1.0	0.0	0.0	0.0	1.0
864	1.0	0.0	1.0	0.0	0.0	1.0
865	1.0	1.0	0.0	0.0	0.0	1.0
866	0.0	1.0	0.0	1.0	0.0	0.0
867	1.0	0.0	1.0	0.0	0.0	1.0
868	1.0	0.0	1.0	0.0	0.0	1.0
869	1.0	0.0	1.0	0.0	0.0	1.0
870	1.0	0.0	1.0	0.0	0.0	1.0
871	1.0	1.0	0.0	0.0	0.0	1.0
872	1.0	0.0	1.0	0.0	0.0	1.0
873	1.0	0.0	1.0	0.0	0.0	1.0
874	0.0	1.0	0.0	1.0	0.0	0.0
875	0.0	1.0	0.0	1.0	0.0	0.0
876	1.0	0.0	1.0	0.0	0.0	1.0
877	1.0	0.0	1.0	0.0	0.0	1.0
878	1.0	0.0	1.0	0.0	0.0	1.0
879	0.0	1.0	0.0	1.0	0.0	0.0
880	1.0	1.0	0.0	0.0	0.0	1.0
881	1.0	0.0	1.0	0.0	0.0	1.0
882	1.0	1.0	0.0	0.0	0.0	1.0
883	1.0	0.0	1.0	0.0	0.0	1.0
884	1.0	0.0	1.0	0.0	0.0	1.0
885	0.0	1.0	0.0	0.0	1.0	0.0
886	1.0	0.0	1.0	0.0	0.0	1.0
887	1.0	1.0	0.0	0.0	0.0	1.0
888	1.0	1.0	0.0	0.0	0.0	1.0
889	0.0	0.0	1.0	1.0	0.0	0.0
890	0.0	0.0	1.0	0.0	1.0	0.0

[891 rows x 18 columns]

```
In [19]: # drops Embarked C and Sex
pd.get_dummies(train, columns = ['Sex', 'Embarked'], drop_first = True)
```

```
Out[19]:
```

	PassengerId	Survived	Pclass	\
0	1	0	3	
1	2	1	1	
2	3	1	3	
3	4	1	1	
4	5	0	3	
5	6	0	3	
6	7	0	1	
7	8	0	3	
8	9	1	3	
9	10	1	2	
10	11	1	3	
11	12	1	1	

12	13	0	3
13	14	0	3
14	15	0	3
15	16	1	2
16	17	0	3
17	18	1	2
18	19	0	3
19	20	1	3
20	21	0	2
21	22	1	2
22	23	1	3
23	24	1	1
24	25	0	3
25	26	1	3
26	27	0	3
27	28	0	1
28	29	1	3
29	30	0	3
..
861	862	0	2
862	863	1	1
863	864	0	3
864	865	0	2
865	866	1	2
866	867	1	2
867	868	0	1
868	869	0	3
869	870	1	3
870	871	0	3
871	872	1	1
872	873	0	1
873	874	0	3
874	875	1	2
875	876	1	3
876	877	0	3
877	878	0	3
878	879	0	3
879	880	1	1
880	881	1	2
881	882	0	3
882	883	0	3
883	884	0	2
884	885	0	3
885	886	0	3
886	887	0	2
887	888	1	1
888	889	0	3
889	890	1	1

890

891

0

3

	Name	Age	SibSp	Parch
0	Braund, Mr. Owen Harris	22.0	1	0
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	38.0	1	0
2	Heikkinen, Miss. Laina	26.0	0	0
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	35.0	1	0
4	Allen, Mr. William Henry	35.0	0	0
5	Moran, Mr. James	NaN	0	0
6	McCarthy, Mr. Timothy J	54.0	0	0
7	Palsson, Master. Gosta Leonard	2.0	3	1
8	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	27.0	0	2
9	Nasser, Mrs. Nicholas (Adele Achem)	14.0	1	0
10	Sandstrom, Miss. Marguerite Rut	4.0	1	1
11	Bonnell, Miss. Elizabeth	58.0	0	0
12	Saunderscock, Mr. William Henry	20.0	0	0
13	Andersson, Mr. Anders Johan	39.0	1	5
14	Vestrom, Miss. Hulda Amanda Adolfina	14.0	0	0
15	Hewlett, Mrs. (Mary D Kingcome)	55.0	0	0
16	Rice, Master. Eugene	2.0	4	1
17	Williams, Mr. Charles Eugene	NaN	0	0
18	Vander Planke, Mrs. Julius (Emelia Maria Vande...	31.0	1	0
19	Masselmani, Mrs. Fatima	NaN	0	0
20	Fynney, Mr. Joseph J	35.0	0	0
21	Beesley, Mr. Lawrence	34.0	0	0
22	McGowan, Miss. Anna "Annie"	15.0	0	0
23	Sloper, Mr. William Thompson	28.0	0	0
24	Palsson, Miss. Torborg Danira	8.0	3	1
25	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	38.0	1	5
26	Emir, Mr. Farred Chehab	NaN	0	0
27	Fortune, Mr. Charles Alexander	19.0	3	2
28	O'Dwyer, Miss. Ellen "Nellie"	NaN	0	0
29	Todoroff, Mr. Lalio	NaN	0	0
..
861	Giles, Mr. Frederick Edward	21.0	1	0
862	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	48.0	0	0
863	Sage, Miss. Dorothy Edith "Dolly"	NaN	8	2
864	Gill, Mr. John William	24.0	0	0
865	Bystrom, Mrs. (Karolina)	42.0	0	0
866	Duran y More, Miss. Asuncion	27.0	1	0
867	Roebling, Mr. Washington Augustus II	31.0	0	0
868	van Melkebeke, Mr. Philemon	NaN	0	0
869	Johnson, Master. Harold Theodor	4.0	1	1
870	Balkic, Mr. Cerin	26.0	0	0
871	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	47.0	1	1
872	Carlsson, Mr. Frans Olof	33.0	0	0
873	Vander Cruyssen, Mr. Victor	47.0	0	0
874	Abelson, Mrs. Samuel (Hannah Wizosky)	28.0	1	0

875	Najib, Miss. Adele Kiamie "Jane"	15.0	0	0
876	Gustafsson, Mr. Alfred Ossian	20.0	0	0
877	Petroff, Mr. Nedelio	19.0	0	0
878	Laleff, Mr. Kristo	NaN	0	0
879	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	56.0	0	1
880	Shelley, Mrs. William (Imanita Parrish Hall)	25.0	0	1
881	Markun, Mr. Johann	33.0	0	0
882	Dahlberg, Miss. Gerda Ulrika	22.0	0	0
883	Banfield, Mr. Frederick James	28.0	0	0
884	Sutehall, Mr. Henry Jr	25.0	0	0
885	Rice, Mrs. William (Margaret Norton)	39.0	0	5
886	Montvila, Rev. Juozas	27.0	0	0
887	Graham, Miss. Margaret Edith	19.0	0	0
888	Johnston, Miss. Catherine Helen "Carrie"	NaN	1	2
889	Behr, Mr. Karl Howell	26.0	0	0
890	Dooley, Mr. Patrick	32.0	0	0

	Ticket	Fare	Cabin	Sex_male	Embarked_Q	\
0	A/5 21171	7.2500	NaN	1	0.0	
1	PC 17599	71.2833	C85	0	0.0	
2	STON/O2. 3101282	7.9250	NaN	0	0.0	
3	113803	53.1000	C123	0	0.0	
4	373450	8.0500	NaN	1	0.0	
5	330877	8.4583	NaN	1	1.0	
6	17463	51.8625	E46	1	0.0	
7	349909	21.0750	NaN	1	0.0	
8	347742	11.1333	NaN	0	0.0	
9	237736	30.0708	NaN	0	0.0	
10	PP 9549	16.7000	G6	0	0.0	
11	113783	26.5500	C103	0	0.0	
12	A/5. 2151	8.0500	NaN	1	0.0	
13	347082	31.2750	NaN	1	0.0	
14	350406	7.8542	NaN	0	0.0	
15	248706	16.0000	NaN	0	0.0	
16	382652	29.1250	NaN	1	1.0	
17	244373	13.0000	NaN	1	0.0	
18	345763	18.0000	NaN	0	0.0	
19	2649	7.2250	NaN	0	0.0	
20	239865	26.0000	NaN	1	0.0	
21	248698	13.0000	D56	1	0.0	
22	330923	8.0292	NaN	0	1.0	
23	113788	35.5000	A6	1	0.0	
24	349909	21.0750	NaN	0	0.0	
25	347077	31.3875	NaN	0	0.0	
26	2631	7.2250	NaN	1	0.0	
27	19950	263.0000	C23 C25 C27	1	0.0	
28	330959	7.8792	NaN	0	1.0	
29	349216	7.8958	NaN	1	0.0	

...
861	28134	11.5000	NaN	1 0.0
862	17466	25.9292	D17	0 0.0
863	CA. 2343	69.5500	NaN	0 0.0
864	233866	13.0000	NaN	1 0.0
865	236852	13.0000	NaN	0 0.0
866	SC/PARIS 2149	13.8583	NaN	0 0.0
867	PC 17590	50.4958	A24	1 0.0
868	345777	9.5000	NaN	1 0.0
869	347742	11.1333	NaN	1 0.0
870	349248	7.8958	NaN	1 0.0
871	11751	52.5542	D35	0 0.0
872	695	5.0000	B51 B53 B55	1 0.0
873	345765	9.0000	NaN	1 0.0
874	P/PP 3381	24.0000	NaN	0 0.0
875	2667	7.2250	NaN	0 0.0
876	7534	9.8458	NaN	1 0.0
877	349212	7.8958	NaN	1 0.0
878	349217	7.8958	NaN	1 0.0
879	11767	83.1583	C50	0 0.0
880	230433	26.0000	NaN	0 0.0
881	349257	7.8958	NaN	1 0.0
882	7552	10.5167	NaN	0 0.0
883	C.A./SOTON 34068	10.5000	NaN	1 0.0
884	SOTON/OQ 392076	7.0500	NaN	1 0.0
885	382652	29.1250	NaN	0 1.0
886	211536	13.0000	NaN	1 0.0
887	112053	30.0000	B42	0 0.0
888	W./C. 6607	23.4500	NaN	0 0.0
889	111369	30.0000	C148	1 0.0
890	370376	7.7500	NaN	1 1.0

	Embarked_S	Sex_male	Embarked_Q	Embarked_S
0	1.0	1.0	0.0	1.0
1	0.0	0.0	0.0	0.0
2	1.0	0.0	0.0	1.0
3	1.0	0.0	0.0	1.0
4	1.0	1.0	0.0	1.0
5	0.0	1.0	1.0	0.0
6	1.0	1.0	0.0	1.0
7	1.0	1.0	0.0	1.0
8	1.0	0.0	0.0	1.0
9	0.0	0.0	0.0	0.0
10	1.0	0.0	0.0	1.0
11	1.0	0.0	0.0	1.0
12	1.0	1.0	0.0	1.0
13	1.0	1.0	0.0	1.0
14	1.0	0.0	0.0	1.0

15	1.0	0.0	0.0	1.0
16	0.0	1.0	1.0	0.0
17	1.0	1.0	0.0	1.0
18	1.0	0.0	0.0	1.0
19	0.0	0.0	0.0	0.0
20	1.0	1.0	0.0	1.0
21	1.0	1.0	0.0	1.0
22	0.0	0.0	1.0	0.0
23	1.0	1.0	0.0	1.0
24	1.0	0.0	0.0	1.0
25	1.0	0.0	0.0	1.0
26	0.0	1.0	0.0	0.0
27	1.0	1.0	0.0	1.0
28	0.0	0.0	1.0	0.0
29	1.0	1.0	0.0	1.0
..
861	1.0	1.0	0.0	1.0
862	1.0	0.0	0.0	1.0
863	1.0	0.0	0.0	1.0
864	1.0	1.0	0.0	1.0
865	1.0	0.0	0.0	1.0
866	0.0	0.0	0.0	0.0
867	1.0	1.0	0.0	1.0
868	1.0	1.0	0.0	1.0
869	1.0	1.0	0.0	1.0
870	1.0	1.0	0.0	1.0
871	1.0	0.0	0.0	1.0
872	1.0	1.0	0.0	1.0
873	1.0	1.0	0.0	1.0
874	0.0	0.0	0.0	0.0
875	0.0	0.0	0.0	0.0
876	1.0	1.0	0.0	1.0
877	1.0	1.0	0.0	1.0
878	1.0	1.0	0.0	1.0
879	0.0	0.0	0.0	0.0
880	1.0	0.0	0.0	1.0
881	1.0	1.0	0.0	1.0
882	1.0	0.0	0.0	1.0
883	1.0	1.0	0.0	1.0
884	1.0	1.0	0.0	1.0
885	0.0	0.0	1.0	0.0
886	1.0	1.0	0.0	1.0
887	1.0	0.0	0.0	1.0
888	1.0	0.0	0.0	1.0
889	0.0	1.0	0.0	0.0
890	0.0	1.0	1.0	0.0

[891 rows x 16 columns]

```
In [ ]:
```