In [18]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
train_data='C:/Users/ramya/Desktop/New folder/traink.csv'
full_data=pd.read_csv(train_data)
full_data.describe()
test_data='C:/Users/ramya/Desktop/New folder/testk.csv'
full_test=pd.read_csv(test_data)
full_test.describe()
```

Out[18]:

	PassengerId	Pclass	Age	SibSp	Parch	Fare
count	418.000000	418.000000	332.000000	418.000000	418.000000	417.000000
mean	1100.500000	2.265550	30.272590	0.447368	0.392344	35.627188
std	120.810458	0.841838	14.181209	0.896760	0.981429	55.907576
min	892.000000	1.000000	0.170000	0.000000	0.000000	0.000000
25%	996.250000	1.000000	21.000000	0.000000	0.000000	7.895800
50%	1100.500000	3.000000	27.000000	0.000000	0.000000	14.454200
75%	1204.750000	3.000000	39.000000	1.000000	0.000000	31.500000
max	1309.000000	3.000000	76.000000	8.000000	9.000000	512.329200

In [5]:

```
a=full_data[['Age', 'Survived']].groupby(['Age'], as_index=False).mean().sort_values(by='Survived'
, ascending=True)
a.head(10)
```

Out[5]:

	Age	Survived
54	40.5	0.0
77	61.0	0.0
14	10.0	0.0
86	74.0	0.0
40	30.5	0.0
46	34.5	0.0
26	20.5	0.0
81	65.0	0.0
71	55.5	0.0
82	66.0	0.0

In [6]:

```
full_data[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean().sort_values(by='Survived',
ascending=True)
```

Out[6]:

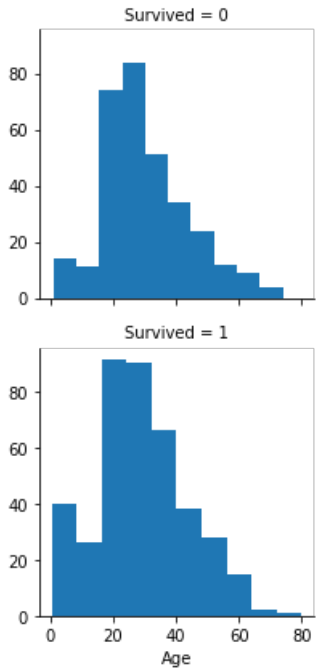
	Sex	Survived
1	male	0.403813
0	female	0.805732

In [7]:

```
g=sns.FacetGrid(full_data,'Survived')
g.map(plt.hist,'Age',bins=10)
```

Out[7]:

```
<seaborn.axisgrid.FacetGrid at 0x26aad335358>
```



In [8]:

```
full_data.drop('Ticket',axis=1)
```

Out[8]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	7.9250	NaN	S
3	4	0	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	male	NaN	0	0	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	21.0750	NaN	S
8	9	0	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	30.0708	NaN	C
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	16.7000	G6	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	26.5500	C103	S
12	13	0	3	Saundercock, Mr. William Henry	male	20.0	0	0	8.0500	NaN	S
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	31.2750	NaN	S
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14.0	0	0	7.8542	NaN	S
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	16.0000	NaN	S
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	29.1250	NaN	Q
17	18	1	2	Williams, Mr. Charles Eugene	male	NaN	0	0	13.0000	NaN	S
18	19	0	3	Vander Planke, Mrs. Julius (Emelia Maria Vande...	female	31.0	1	0	18.0000	NaN	S

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked
19	20	1	3	Masselmani, Mrs. Fatima	female	NaN	0	0	7.2250	NaN	C
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	26.0000	NaN	S
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	13.0000	D56	S
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	8.0292	NaN	Q
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	35.5000	A6	S
24	25	1	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	21.0750	NaN	S
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	31.3875	NaN	S
26	27	1	3	Emir, Mr. Farred Chehab	male	NaN	0	0	7.2250	NaN	C
27	28	1	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	263.0000	C23 C25 C27	S
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	NaN	0	0	7.8792	NaN	Q
29	30	1	3	Todoroff, Mr. Lalio	male	NaN	0	0	7.8958	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...
861	862	1	2	Giles, Mr. Frederick Edward	male	21.0	1	0	11.5000	NaN	S
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	25.9292	D17	S
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	NaN	8	2	69.5500	NaN	S
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	13.0000	NaN	S
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	13.0000	NaN	S
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	13.8583	NaN	C
867	868	0	1	Roebling, Mr. Washington Augustus II	male	31.0	0	0	50.4958	A24	S
868	869	0	3	van Melkebeke, Mr. Philemon	male	NaN	0	0	9.5000	NaN	S
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	11.1333	NaN	S
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	7.8958	NaN	S
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	52.5542	D35	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	5.0000	B51 B53 B55	S
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	9.0000	NaN	S
874	875	1	2	Abelson, Mrs. Samuel (Hannah Wizosky)	female	28.0	1	0	24.0000	NaN	C
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	7.2250	NaN	C
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	9.8458	NaN	S
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	7.8958	NaN	S
878	879	0	3	Laleff, Mr. Kristo	male	NaN	0	0	7.8958	NaN	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	83.1583	C50	C
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	26.0000	NaN	S
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	7.8958	NaN	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	10.5167	NaN	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	10.5000	NaN	S
884	885	1	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	7.0500	NaN	S
885	886	1	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	29.1250	NaN	Q
886	887	1	2	Montvila, Rev. Juozas	male	27.0	0	0	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	7.7500	NaN	Q

891 rows × 11 columns

In [9]:

```
full_data[['Parch', 'Survived']].groupby(['Parch'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

```
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[9]:

	Parch	Survived
6	6	0.000000
0	0	0.511799
3	3	0.600000
5	5	0.600000
2	2	0.637500
1	1	0.669492
4	4	0.750000

In [10]:

```
full_data[['Pclass', 'Survived']].groupby(['Pclass'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[10]:

	Pclass	Survived
2	3	0.435845
1	2	0.630435
0	1	0.722222

In [11]:

```
full_data[['Cabin', 'Survived']].groupby(['Cabin'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[11]:

	Cabin	Survived
0	A10	0.0
38	B71	0.0
45	B86	0.0
46	B94	0.0
52	C110	0.0
54	C118	0.0
59	C128	0.0
64	C30	0.0
79	C82	0.0
82	C86	0.0
85	C91	0.0
26	B38	0.0
88	C95	0.0
106	D46	0.0
108	D48	0.0
110	D50	0.0
112	D6	0.0
122	E31	0.0
126	E38	0.0
127	E40	0.0
129	E46	0.0
132	E58	0.0

	Cabin	Survived
139	F G63	0.0
100	D30	0.0
25	B37	0.0
146	T	0.0
23	B30	0.0
18	B19	0.0
9	A32	0.0
6	A24	0.0
...	...	...
19	B20	1.0
53	C111	1.0
57	C125	1.0
58	C126	1.0
89	C99	1.0
12	A5	1.0
87	C93	1.0
24	B35	1.0
13	A6	1.0
84	C90	1.0
83	C87	1.0
14	A7	1.0
81	C85	1.0
15	B101	1.0
78	C78	1.0
92	D11	1.0
77	C70	1.0
72	C54	1.0
71	C52	1.0
70	C50	1.0
69	C49	1.0
68	C47	1.0
67	C46	1.0
66	C45	1.0
65	C32	1.0
63	C23 C25 C27	1.0
60	C148	1.0
17	B18	1.0
76	C7	1.0
73	C62 C64	1.0

147 rows × 2 columns

In [12]:

```
full_data[['SibSp', 'Survived']].groupby(['SibSp'],
as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[12]:

	SibSp	Survived
6	8	0.285714
4	4	0.388889

	SibSp	Survived
0	0	0.541543
3	3	0.562500
5	5	0.600000
1	1	0.641148
2	2	0.714286

In [13]:

```
full_data[['Embarked', 'Survived']].groupby(['Embarked'], as_index=False).mean().sort_values(by='Survived', ascending=True)
```

Out[13]:

	Embarked	Survived
2	S	0.504655
1	Q	0.507576
0	C	0.665158

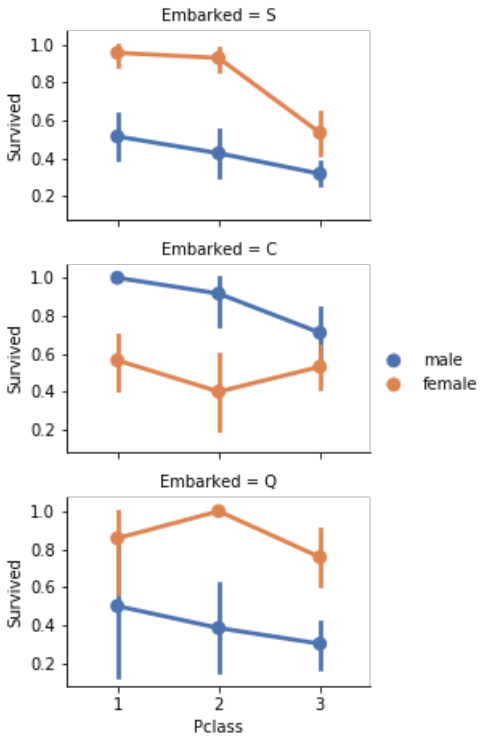
In [83]:

```
grid = sns.FacetGrid(full_data, row='Embarked', size=2.2, aspect=1.6)
grid.map(sns.pointplot, 'Pclass', 'Survived', 'Sex', palette='deep')
grid.add_legend()
```

```
C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:230: UserWarning: The `size`
paramter has been renamed to `height`; please update your code.
  warnings.warn(msg, UserWarning)
C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:715: UserWarning: Using the
pointplot function without specifying `order` is likely to produce an incorrect plot.
  warnings.warn(warning)
C:\Users\ramya\Anaconda3\lib\site-packages\seaborn\axisgrid.py:720: UserWarning: Using the
pointplot function without specifying `hue_order` is likely to produce an incorrect plot.
  warnings.warn(warning)
```

Out[83]:

```
<seaborn.axisgrid.FacetGrid at 0x1503cf765c0>
```



In [14]:

```
full_data.Sex[full_data.Sex=='male']=0
full_data.Sex[full_data.Sex=='female']=1
full_test.Sex[full_test.Sex=='male']=0
full_test.Sex[full_test.Sex=='female']=1
full_data.head()
```

```
C:\Users\ramya\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/indexing.html#indexing-view-versus-copy
  """Entry point for launching an IPython kernel.
C:\Users\ramya\Anaconda3\lib\site-packages\ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/indexing.html#indexing-view-versus-copy

C:\Users\ramya\Anaconda3\lib\site-packages\ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/indexing.html#indexing-view-versus-copy
  This is separate from the ipykernel package so we can avoid doing imports until
C:\Users\ramya\Anaconda3\lib\site-packages\ipykernel_launcher.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-
docs/stable/indexing.html#indexing-view-versus-copy
  after removing the cwd from sys.path.
```

Out[14]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	0	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	1	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	1	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	0	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	1	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	0	35.0	0	0	373450	8.0500	NaN	S

In [15]:

```
full_data[['Sex', 'Survived']].groupby(['Sex'], as_index=False).mean().sort_values(by='Survived',
ascending=True)
```

Out[15]:

	Sex	Survived
0	0	0.403813
1	1	0.805732