In [63]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [64]:
```
df = pd.read_csv("titanic.csv")
```

In [65]:
```
df.head()
```

Out[65]:

	PassengerId	Survived	Pclass	Lname	Name	Sex	Age	SibSp	Parch	Ticket	
0	1	0	3	Braund	Mr. Owen Harris	male	22.0	1	0	A/5 21171	7
1	2	1	1	Cumings	Mrs. John Bradley (Florence Briggs Thayer)	female	38.0	1	0	PC 17599	71
2	3	1	3	Heikkinen	Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7
3	4	1	1	Futrelle	Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53
4	5	0	3	Allen	Mr. William Henry	male	35.0	0	0	373450	8

In [66]:
```
df.info()
```

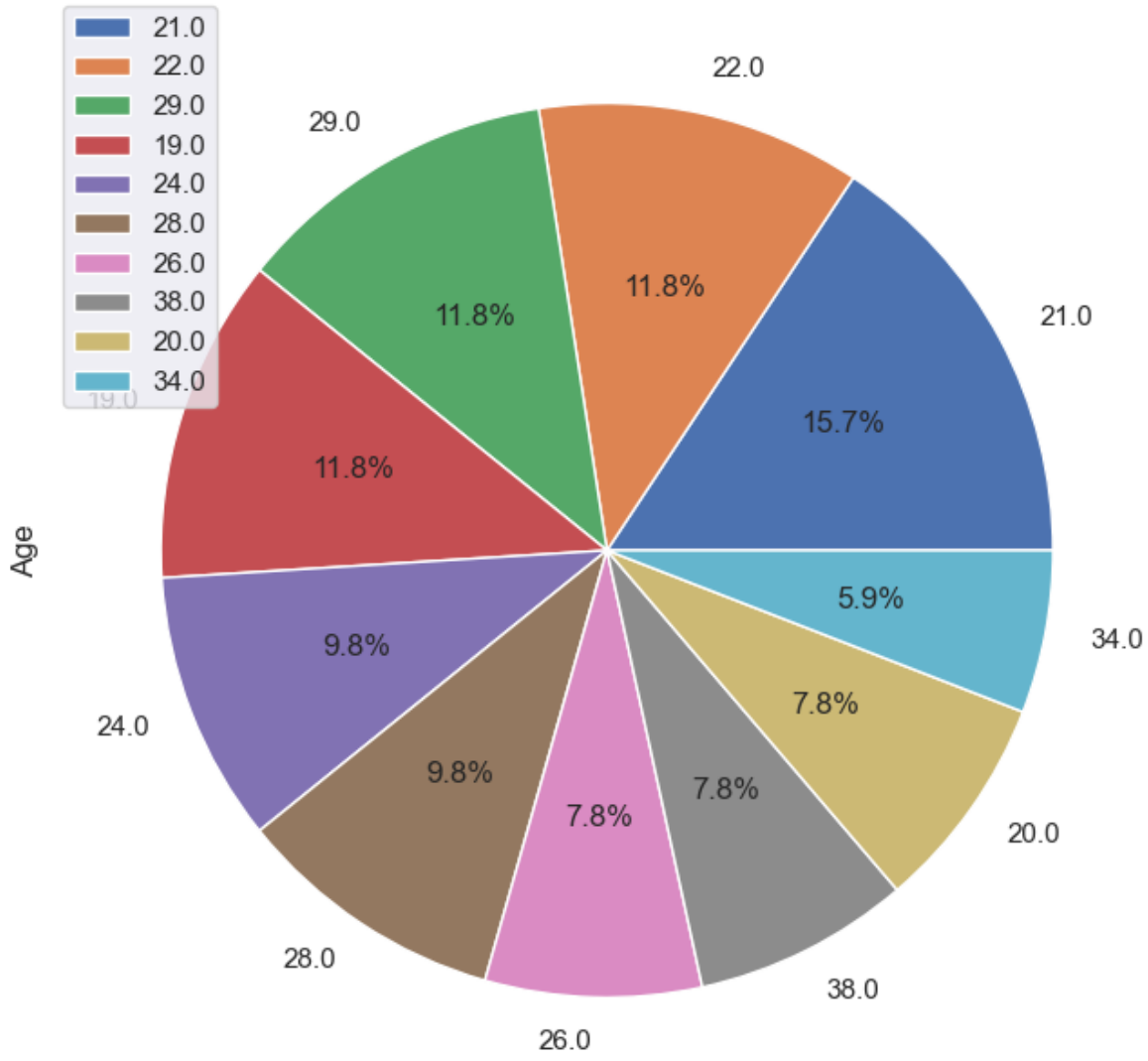
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 156 entries, 0 to 155
Data columns (total 13 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   PassengerId  156 non-null    int64  
 1   Survived     156 non-null    int64  
 2   Pclass       156 non-null    int64  
 3   Lname        156 non-null    object 
 4   Name         156 non-null    object 
 5   Sex          156 non-null    object 
 6   Age          126 non-null    float64
 7   SibSp        156 non-null    int64  
 8   Parch        156 non-null    int64  
 9   Ticket       156 non-null    object 
 10  Fare         156 non-null    float64
 11  Cabin        31 non-null     object 
 12  Embarked     155 non-null    object 
dtypes: float64(2), int64(5), object(6)
memory usage: 16.0+ KB
```

In [67]:
```
df.isnull().sum()
```

Out[67]:
```
PassengerId      0
Survived         0
Pclass           0
Lname            0
Name             0
Sex              0
Age             30
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          125
Embarked         1
dtype: int64
```

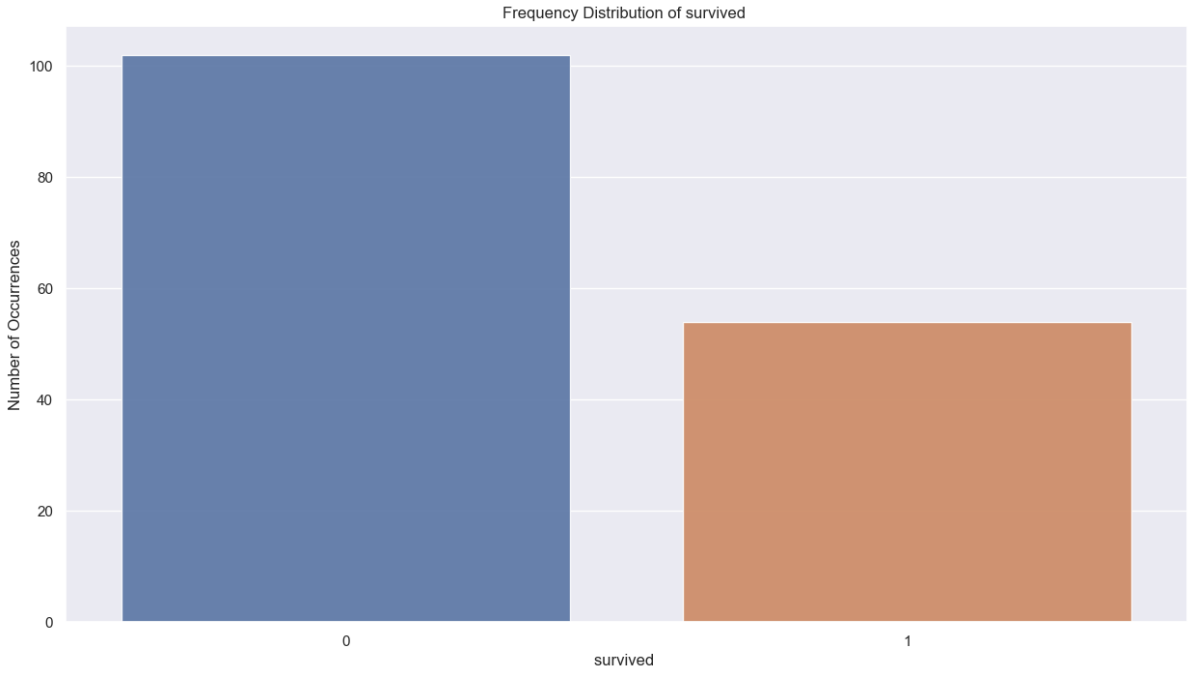
In [68]:
```
df["Age"].value_counts().head(10).plot(kind = 'pie', autopct='%1.1f%%', figsize
```

Out[68]:
```
<matplotlib.legend.Legend at 0x269e475ecd0>
```



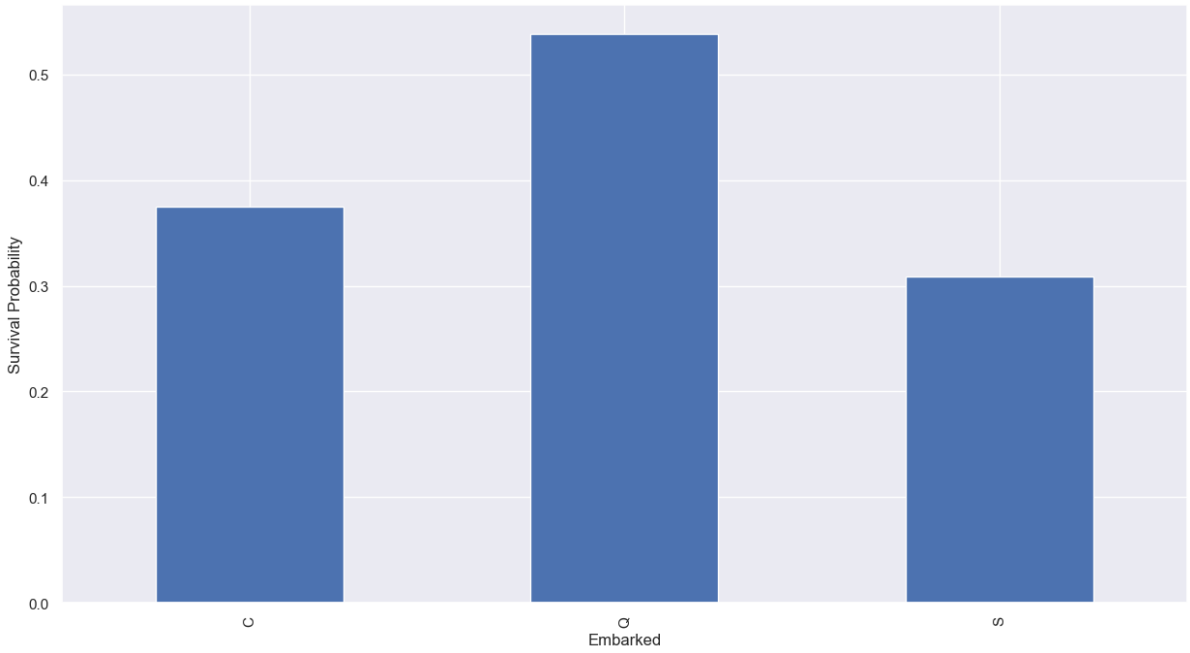
In [69]:

```
carrier_count = df["Survived"].value_counts()
sns.set(style="darkgrid")
data = {'carrier': carrier_count.index, 'count': carrier_count.values}
mf = pd.DataFrame(data)                          #v
sns.barplot(x='carrier', y='count', data=mf, alpha=0.9)
plt.title('Frequency Distribution of survived ')
plt.ylabel('Number of Occurrences', fontsize=12)
plt.xlabel('survived ', fontsize=12)       #1
plt.show()
```



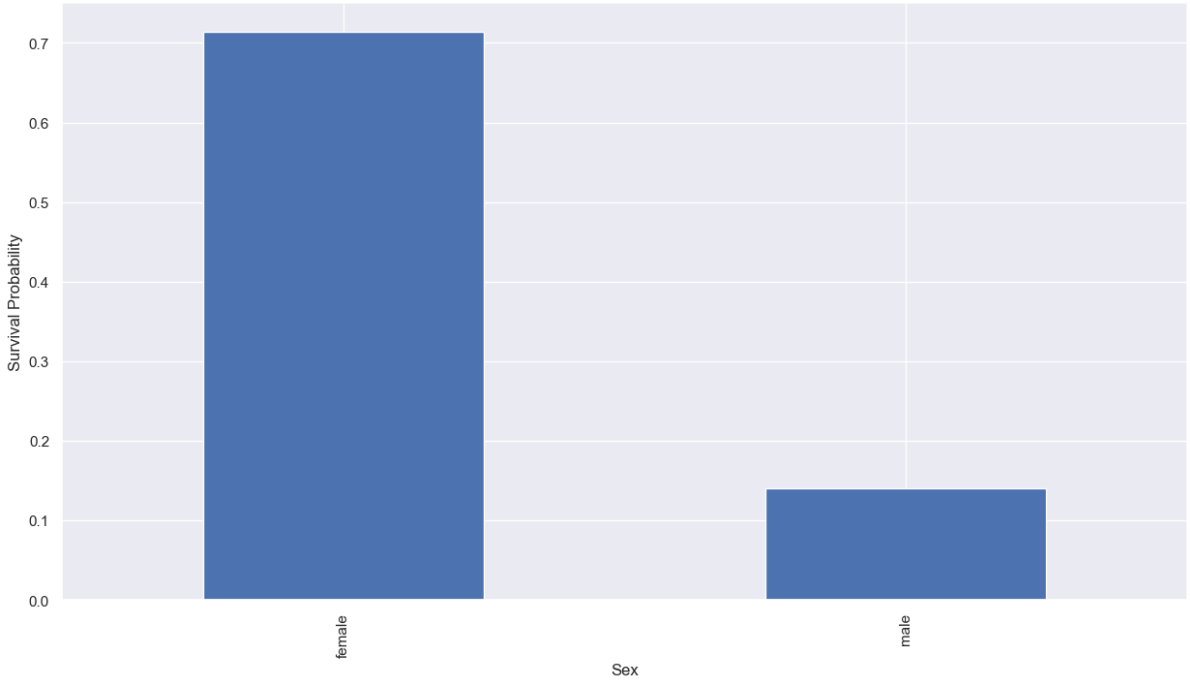
In [70]:
```
plt = df[['Embarked', 'Survived']].groupby('Embarked').mean().Survived.plot(kir
plt.set_xlabel('Embarked')                    #a
plt.set_ylabel('Survival Probability')
```

Out[70]: Text(0, 0.5, 'Survival Probability')



In [71]:
```
plt = df[['Sex', 'Survived']].groupby('Sex').mean().Survived.plot(kind='bar')
plt.set_xlabel('Sex')
plt.set_ylabel('Survival Probability')
```

Out[71]: Text(0, 0.5, 'Survival Probability')

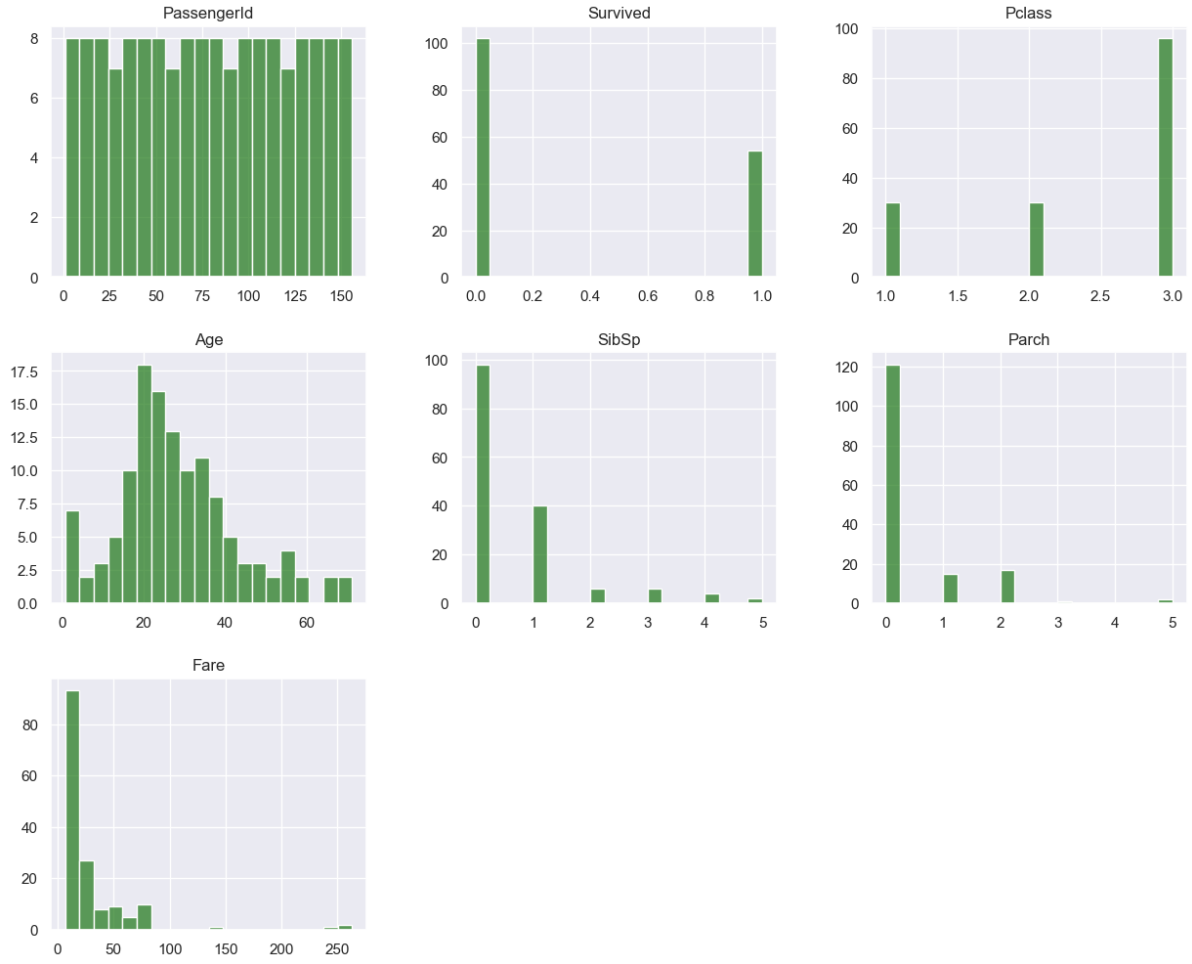


In [72]:
```
sns.pairplot(df,hue='Survived')
```

Out[72]:
```
<seaborn.axisgrid.PairGrid at 0x269e5106290>
```

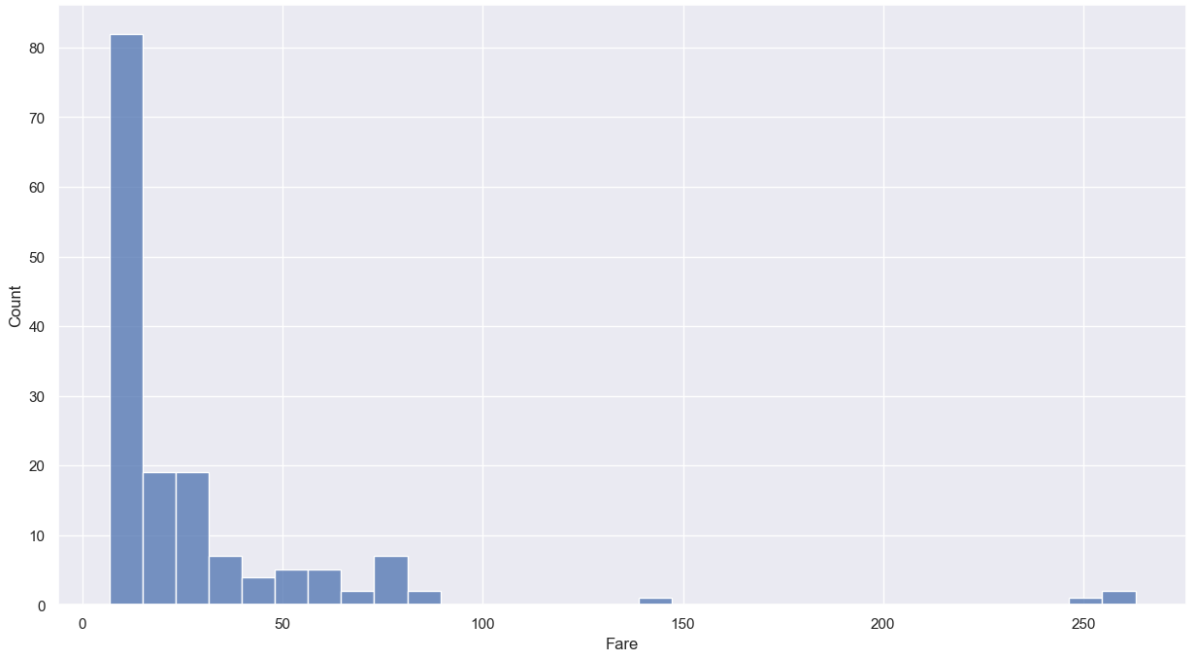
In [73]:
```
df.hist(figsize=(15,12),bins = 20, color="#107009AA")           #k
```

Out[73]:
```
array([[<Axes: title={'center': 'PassengerId'}>,
        <Axes: title={'center': 'Survived'}>,
        <Axes: title={'center': 'Pclass'}>],
       [<Axes: title={'center': 'Age'}>,
        <Axes: title={'center': 'SibSp'}>,
        <Axes: title={'center': 'Parch'}>],
       [<Axes: title={'center': 'Fare'}>, <Axes: >, <Axes: >]],
      dtype=object)
```



In [74]:
```
sns.set(rc = {'figure.figsize':(15,8)})
sns.histplot(data=df, x="Fare")
```

Out[74]:
```
<Axes: xlabel='Fare', ylabel='Count'>
```



In [ ]: