In [44]:

```
print("Name :")
print("We will start learning about histogram and read, clean, and understand the Titanic d
print("We will also derive in which class people died the most.")
```

```
Name :
We will start learning about histogram and read, clean, and understand the T
itanic dataset, and plot a histogram for showing the age group who has the h
ighest death rate and who has the highest survival rate
We will also derive in which class people died the most.
```

In [45]:

```
#predefine code for image
from IPython.display import Image
Image(filename='titanic.jpg')
#predefine code end
```

Out[45]:

In [46]:

```
#import the required packages
import pandas as pd
import matplotlib.pyplot as plt

#Read the csv file.
df = pd.read_csv("Titanic.csv")
df
```

Out[46]:

	PassengerId	Survived	Pclass	Name	Gender	Age	SibSp	Parch	Ticket	Far
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.250
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.283
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.925
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.100
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.050
...	...	...	...	...	...	...	...	...	...	.
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

891 rows × 12 columns

Activity 1 - Plotting histogram showing the passenger

Activity 1: Plotting histogram showing the passenger of different age groups who survived

In [47]:

```
#Remove the rows from the data set where there are NaN values
df = df.dropna()

#Find out the passengers from the Titanic who survived in Titanic.
passengers_survived = df.loc[df['Survived'] == 1]
passengers_survived
```

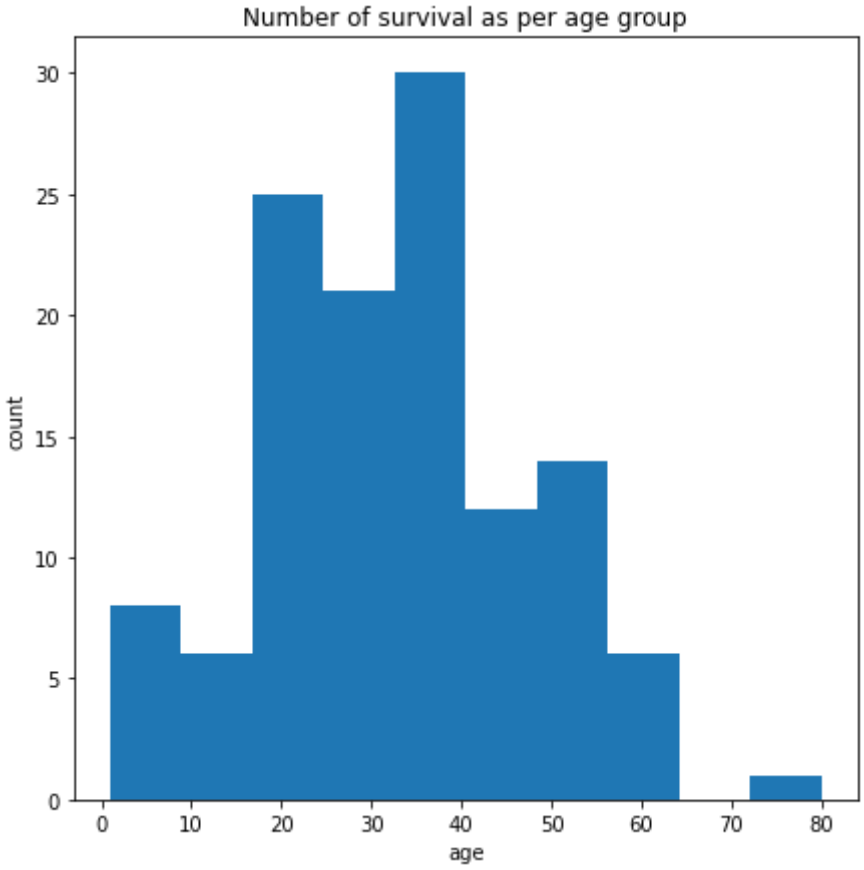
Out[47]:

	PassengerId	Survived	Pclass	Name	Gender	Age	SibSp	Parch	Ticket	Fare
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000
...	...	...	...	...	...	...	...	...	...	...
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000

123 rows × 12 columns

In [48]:

```
#Plot a histogram showing the passenger of titanic from different age groups who survived.
plt.figure(figsize=(7,7))
plt.hist(passengers_survived["Age"], bins=10)
plt.title("Number of survival as per age group")
plt.ylabel("count")
plt.xlabel("age")
plt.show()
```

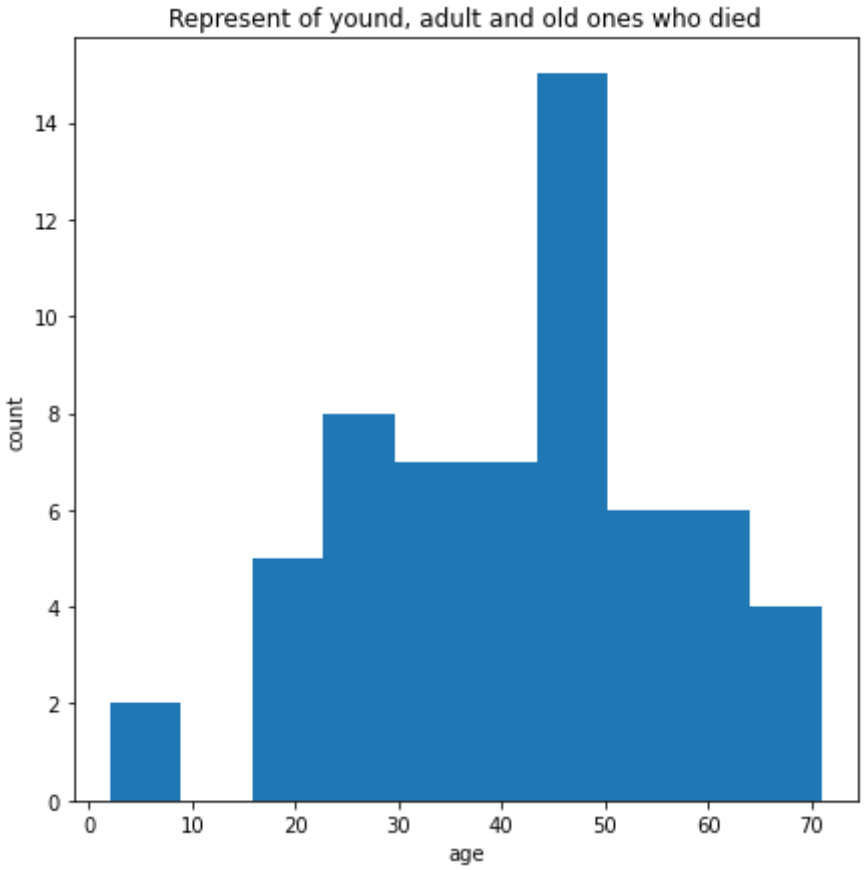


Conclusion: So from the Histogram we, can conclude that, most passengers who survived were in the age group of 20-40.

Activity 2 - Plotting histogram showing the passenger of different age groups who died.

In [49]:

```
from IPython.display import Image
Image(filename='titanic.jpg')
import pandas as pd
import matplotlib.pyplot as plt
df=pd.read_csv('Titanic.csv')
df
df=df.dropna()
passengers_died=df.loc[df['Survived']==0]
passengers_died
plt.figure(figsize=(7,7))
plt.hist(passengers_died['Age'],bins=10)
plt.title('Represent of yound, adult and old ones who died')
plt.xlabel('age')
plt.ylabel('count')
plt.show()
```



In [ ]:

```

```

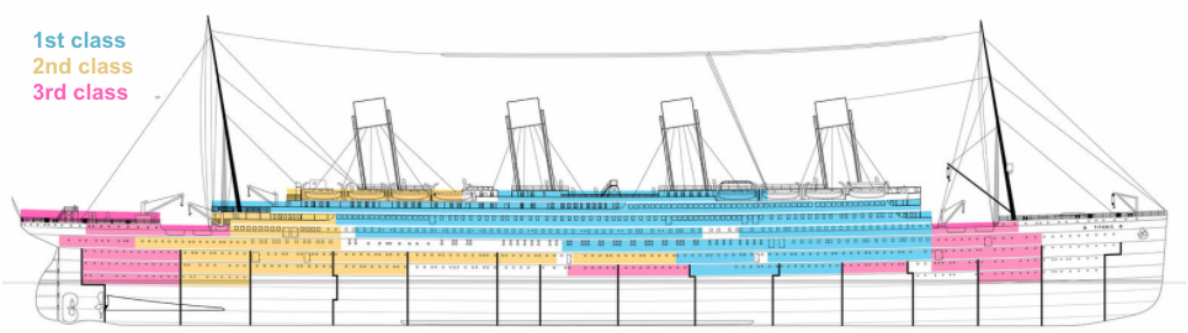
Conslusion:

Activity 3 - Find the total number of passengers who were travelling and total number of passengers who died from class1, class 2, and class 3.

In [50]:

```
#predefine code for image
Image(filename='titanic_classes.png')
#predefine code end
```

Out[50]:



In [51]:

```
survived1=df.loc[(df['Survived']==1)&(df['Pclass']==1)]

total_class1=df.loc[df['Pclass']==1]['Pclass'].count()
total_class1

survive_in_class1=survived1['Pclass'].count()
survive_in_class1
```

Out[51]:

```
106
```

In [52]:

```
survived2=df.loc[(df['Survived']==1)&(df['Pclass']==2)]

total_class2=df.loc[df['Pclass']==2]['Pclass'].count()
total_class2

survive_in_class2=survived2['Pclass'].count()
survive_in_class2
```

Out[52]:

```
12
```

In [53]:

```
survived3=df.loc[(df['Survived']==3)&(df['Pclass']==2)]

total_class3=df.loc[df['Pclass']==3]['Pclass'].count()
total_class3

survive_in_class3=survived3['Pclass'].count()
survive_in_class3
```

Out[53]:

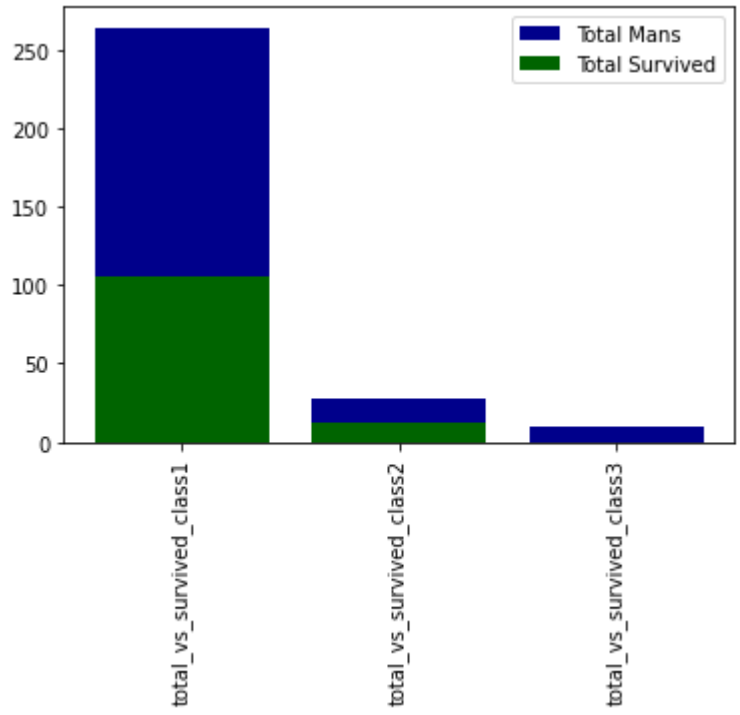
```
0
```

In [54]:

```
total_passengers=[total_class1,total_class2,total_class3]
total_survival=[survive_in_class1,survive_in_class2,survive_in_class3]
index=['total_vs_survived_class1','total_vs_survived_class2','total_vs_survived_class3']
```

In [55]:

```
plt.bar(index,total_passengers,bottom=total_survival,color='darkblue',label='Total Mans')
plt.bar(index,total_survival,color='darkgreen',label='Total Survived')
plt.xticks(rotation='vertical')
plt.legend()
plt.show()
```



In [ ]:

```
```

In [56]:

```
#Find the total number of passengers who survived from class 3
```

In [57]:

```
#Plot a stacked bar graph for showing highest numbers of deaths was in which class
```

Conclusion:

In [ ]:

```

```