

In [468]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

In [469]:

```
trainset = pd.read_csv('/home/lavahawk0123/Documents/Tasks-MRM/Tasks-ADMIN/Task 2 Logistic Regression/titanic_train.csv', sep= ",")
```

In [470]:

```
trainset.head(n=892)
```

Out[470]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	male	NaN	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	NaN	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	G6	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C103	S
12	13	0	3	Saunderscock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.0500	NaN	S
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	NaN	S
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14.0	0	0	350406	7.8542	NaN	S
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	NaN	S
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	NaN	Q
17	18	1	2	Williams, Mr. Charles Eugene	male	NaN	0	0	244373	13.0000	NaN	S
18	19	0	3	Vander Planke, Mrs. Julius (Emelia Maria Vande...	female	31.0	1	0	345763	18.0000	NaN	S
19	20	1	3	Masselmani, Mrs. Fatima	female	NaN	0	0	2649	7.2250	NaN	C
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	239865	26.0000	NaN	S
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000	D56	S
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	330923	8.0292	NaN	Q
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	A6	S
24	25	0	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	349909	21.0750	NaN	S
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	347077	31.3875	NaN	S
26	27	0	3	Emir, Mr. Farred Chehab	male	NaN	0	0	2631	7.2250	NaN	C
27	28	0	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	19950	263.0000	C23 C25 C27	S
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	NaN	0	0	330959	7.8792	NaN	Q
29	30	0	3	Todoroff, Mr. Lallo	male	NaN	0	0	349216	7.8958	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...
861	862	0	2	Giles, Mr. Frederick Edward	male	21.0	1	0	28134	11.5000	NaN	S
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292	D17	S
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	NaN	8	2	CA. 2343	69.5500	NaN	S
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	233866	13.0000	NaN	S
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	236852	13.0000	NaN	S
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	SC/PARIS 2149	13.8583	NaN	C
867	868	0	1	Roebling, Mr. Washington Augustus II	male	31.0	0	0	PC 17590	50.4958	A24	S
868	869	0	3	van Melkebeke, Mr. Philemon	male	NaN	0	0	345777	9.5000	NaN	S
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	347742	11.1333	NaN	S
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	349248	7.8958	NaN	S
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monyperry)	female	47.0	1	1	11751	52.5542	D35	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	B51 B53 B55	S
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	345765	9.0000	NaN	S
874	875	1	2	Abelson, Mrs. Samuel (Hannah Witosky)	female	28.0	1	0	P/PP 3381	24.0000	NaN	C
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	NaN	C
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.8458	NaN	S
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.8958	NaN	S
878	879	0	3	Laleff, Mr. Kristo	male	NaN	0	0	349217	7.8958	NaN	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C50	C
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	NaN	S
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	NaN	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	NaN	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	NaN	S
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTONVOQ 392076	7.0500	NaN	S
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	NaN	Q
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	NaN	1	2	W.C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows x 12 columns

In [471]:

```
#Function to replace null ages with average value of age
def add_age(cols):
    Age = cols[0]
    Pclass = cols[1]
    if pd.isnull(Age):
        return int(trainset[trainset["Pclass"] == Pclass]["Age"].mean())
    else:
        return Age
```

In [472]:

```
#Calling the function to replace null age values
trainset["Age"] = trainset[["Age", "Pclass"]].apply(add_age,axis=1)
```

In [473]:

trainset

Out[473]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S
5	6	0	3	Moran, Mr. James	male	25.0	0	0	330877	8.4583	NaN	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	E46	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	NaN	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	NaN	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	NaN	C
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	G6	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C103	S
12	13	0	3	Saunderscock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.0500	NaN	S
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	NaN	S
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14.0	0	0	350406	7.8542	NaN	S
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	NaN	S
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	NaN	Q
17	18	1	2	Williams, Mr. Charles Eugene	male	29.0	0	0	244373	13.0000	NaN	S
18	19	0	3	Vander Plank, Mrs. Julius (Emelia Maria Vande...	female	31.0	1	0	345763	18.0000	NaN	S
19	20	1	3	Massemi, Mrs. Fatma	female	25.0	0	0	2649	7.2250	NaN	C
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	239865	26.0000	NaN	S
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000	D56	S
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	330923	8.0292	NaN	Q
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	A6	S
24	25	0	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	349909	21.0750	NaN	S
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	347077	31.3875	NaN	S
26	27	0	3	Emir, Mr. Farred Chehab	male	25.0	0	0	2631	7.2250	NaN	C
27	28	0	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	19950	263.0000	C23 C25 C27	S
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	25.0	0	0	330959	7.8792	NaN	Q
29	30	0	3	Todoroff, Mr. Lallo	male	25.0	0	0	349216	7.8958	NaN	S
...	...	...	...	...	...	...	...	...	...	...	...	...
861	862	0	2	Giles, Mr. Frederick Edward	male	21.0	1	0	28134	11.5000	NaN	S
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292	D17	S
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	25.0	8	2	CA. 2343	69.5500	NaN	S
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	233866	13.0000	NaN	S
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	236852	13.0000	NaN	S
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	SC/PARIS 2149	13.8583	NaN	C
867	868	0	1	Roebling, Mr. Washington Augustus II	male	31.0	0	0	PC 17590	50.4958	A24	S
868	869	0	3	van Melkebeke, Mr. Philemon	male	25.0	0	0	345777	9.5000	NaN	S
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	347742	11.1333	NaN	S
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	349248	7.8958	NaN	S
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542	D35	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	B51 B53 B55	S
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	345765	9.0000	NaN	S
874	875	1	2	Abelson, Mrs. Samuel (Hannah Wozosky)	female	28.0	1	0	P/PP 3381	24.0000	NaN	C
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	NaN	C
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.8458	NaN	S
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.8958	NaN	S
878	879	0	3	Laleff, Mr. Kristo	male	25.0	0	0	349217	7.8958	NaN	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C50	C
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	NaN	S
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	NaN	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	NaN	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	NaN	S
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	NaN	S
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	NaN	Q
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	NaN	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	B42	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	25.0	1	2	W./C. 6607	23.4500	NaN	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C148	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	NaN	Q

891 rows × 12 columns

```
In [474]: #Cleaning the dataset
trainset.drop("Cabin",inplace=True,axis=1)#Removing all values in cabin column
trainset.dropna())#Removing rows where all the values are null
trainset.head(n=891)
```

Out[474]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	S
1	2	1	1	Cummings, 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	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	S
5	6	0	3	Moran, Mr. James	male	25.0	0	0	330877	8.4583	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	C
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	S
12	13	0	3	Saunderscock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.0500	S
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	S
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14.0	0	0	350406	7.8542	S
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	S
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	Q
17	18	1	2	Williams, Mr. Charles Eugene	male	29.0	0	0	244373	13.0000	S
18	19	0	3	Vander Planke, Mrs. Julius (Emelia Maria Vande...	female	31.0	1	0	345763	18.0000	S
19	20	1	3	Maselli, Mrs. Fatima	female	25.0	0	0	2649	7.2250	C
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	239865	26.0000	S
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000	S
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	330923	8.0292	Q
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	S
24	25	0	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	349909	21.0750	S
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	347077	31.3875	S
26	27	0	3	Emir, Mr. Farred Chehab	male	25.0	0	0	2631	7.2250	C
27	28	0	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	19950	263.0000	S
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	25.0	0	0	330959	7.8792	Q
29	30	0	3	Todoroff, Mr. Lalio	male	25.0	0	0	349216	7.8958	S
...	...	...	...	...	...	...	...	...	...	...	...
861	862	0	2	Giles, Mr. Frederick Edward	male	21.0	1	0	28134	11.5000	S
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292	S
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	25.0	8	2	CA. 2343	69.5500	S
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	233866	13.0000	S
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	236852	13.0000	S
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	SC/PARIS 2149	13.8583	C
867	868	0	1	Roebling, Mr. Washington Augustus II	male	31.0	0	0	PC 17590	50.4958	S
868	869	0	3	van Melkebeke, Mr. Philemon	male	25.0	0	0	345777	9.5000	S
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	347742	11.1333	S
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	349248	7.8958	S
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	S
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	345765	9.0000	S
874	875	1	2	Abelson, Mrs. Samuel (Hannah Wozosky)	female	28.0	1	0	P/PP 3381	24.0000	C
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	C
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.8458	S
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.8958	S
878	879	0	3	Laleff, Mr. Kristo	male	25.0	0	0	349217	7.8958	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	S
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	S
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	S
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	Q
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	25.0	1	2	W./C. 6607	23.4500	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	Q

891 rows x 11 columns

```
In [475]: #Converting the categorical data into numerical data
Gender = pd.get_dummies(trainset["Sex"])
Embarked = pd.get_dummies(trainset['Embarked'], drop_first=True)
Class_travelled = pd.get_dummies(trainset['Pclass'], drop_first=True)
```

```
In [476]: trainset
```

Out[476]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	S
1	2	1	1	Cummings, 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	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	S
5	6	0	3	Moran, Mr. James	male	25.0	0	0	330877	8.4583	Q
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	S
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	S
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	S
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	C
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	S
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	S
12	13	0	3	Saunderscock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.0500	S
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	S
14	15	0	3	Vestrom, Miss. Hilda Amanda Adolfin	female	14.0	0	0	350406	7.8542	S
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	S
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	Q
17	18	1	2	Williams, Mr. Charles Eugene	male	29.0	0	0	244373	13.0000	S
18	19	0	3	Vander Plank, Mrs. Julius (Emelia Maria Vande...	female	31.0	1	0	345763	18.0000	S
19	20	1	3	Massemi, Mrs. Fatma	female	25.0	0	0	2649	7.2250	C
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	239865	26.0000	S
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000	S
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	330923	8.0292	Q
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	S
24	25	0	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	349909	21.0750	S
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	347077	31.3875	S
26	27	0	3	Emir, Mr. Farred Chehab	male	25.0	0	0	2631	7.2250	C
27	28	0	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	19950	263.0000	S
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	25.0	0	0	330959	7.8792	Q
29	30	0	3	Todoroff, Mr. Lallo	male	25.0	0	0	349216	7.8958	S
...	...	...	...	...	...	...	...	...	...	...	...
861	862	0	2	Giles, Mr. Frederick Edward	male	21.0	1	0	28134	11.5000	S
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292	S
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	25.0	8	2	CA. 2343	69.5500	S
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	233866	13.0000	S
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	236852	13.0000	S
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	SC/PARIS 2149	13.8583	C
867	868	0	1	Roebbing, Mr. Washington Augustus II	male	31.0	0	0	PC 17590	50.4958	S
868	869	0	3	van Melkebeke, Mr. Philemon	male	25.0	0	0	345777	9.5000	S
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	347742	11.1333	S
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	349248	7.8958	S
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542	S
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	S
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	345765	9.0000	S
874	875	1	2	Abelson, Mrs. Samuel (Hannah Wozosky)	female	28.0	1	0	P/PP 3381	24.0000	C
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	C
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.8458	S
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.8958	S
878	879	0	3	Laleff, Mr. Kristo	male	25.0	0	0	349217	7.8958	S
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	S
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	S
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	S
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	S
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	S
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	Q
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	S
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	S
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	25.0	1	2	W./C. 6607	23.4500	S
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	Q

891 rows × 11 columns

```
In [477]: #We add the numerical data i.e concat it to the existing dataframe
trainset = pd.concat([trainset,Class_travelled,Gender,Embarked],axis=1)
```

```
In [478]: trainset
```

Out[478]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked	2	3	female	male	Q	S
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	S	0	1	0	1	0	1
1	2	1	1	Cummings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C	0	0	1	0	0	0
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	S	0	1	1	0	0	1
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	S	0	0	1	0	0	1
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	S	0	1	0	1	0	1
5	6	0	3	Moran, Mr. James	male	25.0	0	0	330877	8.4583	Q	0	1	0	1	1	0
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	S	0	0	0	1	0	1
7	8	0	3	Palsson, Master. Gosta Leonard	male	2.0	3	1	349909	21.0750	S	0	1	0	1	0	1
8	9	1	3	Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)	female	27.0	0	2	347742	11.1333	S	0	1	1	0	0	1
9	10	1	2	Nasser, Mrs. Nicholas (Adele Achem)	female	14.0	1	0	237736	30.0708	C	1	0	1	0	0	0
10	11	1	3	Sandstrom, Miss. Marguerite Rut	female	4.0	1	1	PP 9549	16.7000	S	0	1	1	0	0	1
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	S	0	0	1	0	0	1
12	13	0	3	Saunderscock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.0500	S	0	1	0	1	0	1
13	14	0	3	Andersson, Mr. Anders Johan	male	39.0	1	5	347082	31.2750	S	0	1	0	1	0	1
14	15	0	3	Vestrom, Miss. Hulda Amanda Adolfina	female	14.0	0	0	350406	7.8542	S	0	1	1	0	0	1
15	16	1	2	Hewlett, Mrs. (Mary D Kingcome)	female	55.0	0	0	248706	16.0000	S	1	0	1	0	0	1
16	17	0	3	Rice, Master. Eugene	male	2.0	4	1	382652	29.1250	Q	0	1	0	1	1	0
17	18	1	2	Williams, Mr. Charles Eugene	male	29.0	0	0	244373	13.0000	S	1	0	0	1	0	1
18	19	0	3	Vander Plank, Mrs. Julius (Emelia Maria Vand...	female	31.0	1	0	345763	18.0000	S	0	1	1	0	0	1
19	20	1	3	Massei, Mrs. Fatima	female	25.0	0	0	2649	7.2250	C	0	1	1	0	0	0
20	21	0	2	Fynney, Mr. Joseph J	male	35.0	0	0	239865	26.0000	S	1	0	0	1	0	1
21	22	1	2	Beesley, Mr. Lawrence	male	34.0	0	0	248698	13.0000	S	1	0	0	1	0	1
22	23	1	3	McGowan, Miss. Anna "Annie"	female	15.0	0	0	330923	8.0292	Q	0	1	1	0	1	0
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	S	0	0	0	1	0	1
24	25	0	3	Palsson, Miss. Torborg Danira	female	8.0	3	1	349909	21.0750	S	0	1	1	0	0	1
25	26	1	3	Asplund, Mrs. Carl Oscar (Selma Augusta Emilia...	female	38.0	1	5	347077	31.3875	S	0	1	1	0	0	1
26	27	0	3	Emir, Mr. Farred Chehab	male	25.0	0	0	2631	7.2250	C	0	1	0	1	0	0
27	28	0	1	Fortune, Mr. Charles Alexander	male	19.0	3	2	19950	263.0000	S	0	0	0	1	0	1
28	29	1	3	O'Dwyer, Miss. Ellen "Nellie"	female	25.0	0	0	330959	7.8792	Q	0	1	1	0	1	0
29	30	0	3	Todoroff, Mr. Lallo	male	25.0	0	0	349216	7.8958	S	0	1	0	1	0	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
861	862	0	2	Giles, Mr. Frederick Edward	male	21.0	1	0	28134	11.5000	S	1	0	0	1	0	1
862	863	1	1	Swift, Mrs. Frederick Joel (Margaret Welles Ba...	female	48.0	0	0	17466	25.9292	S	0	0	1	0	0	1
863	864	0	3	Sage, Miss. Dorothy Edith "Dolly"	female	25.0	8	2	CA. 2343	69.5500	S	0	1	1	0	0	1
864	865	0	2	Gill, Mr. John William	male	24.0	0	0	233866	13.0000	S	1	0	0	1	0	1
865	866	1	2	Bystrom, Mrs. (Karolina)	female	42.0	0	0	236852	13.0000	S	1	0	1	0	0	1
866	867	1	2	Duran y More, Miss. Asuncion	female	27.0	1	0	SC/PARIS 2149	13.8583	C	1	0	1	0	0	0
867	868	0	1	Roebeling, Mr. Washington Augustus II	male	31.0	0	0	PC 17590	50.4958	S	0	0	0	1	0	1
868	869	0	3	van Melkebeke, Mr. Philemon	male	25.0	0	0	345777	9.5000	S	0	1	0	1	0	1
869	870	1	3	Johnson, Master. Harold Theodor	male	4.0	1	1	347742	11.1333	S	0	1	0	1	0	1
870	871	0	3	Balkic, Mr. Cerin	male	26.0	0	0	349248	7.8958	S	0	1	0	1	0	1
871	872	1	1	Beckwith, Mrs. Richard Leonard (Sallie Monypeny)	female	47.0	1	1	11751	52.5542	S	0	0	1	0	0	1
872	873	0	1	Carlsson, Mr. Frans Olof	male	33.0	0	0	695	5.0000	S	0	0	0	1	0	1
873	874	0	3	Vander Cruyssen, Mr. Victor	male	47.0	0	0	345765	9.0000	S	0	1	0	1	0	1
874	875	1	2	Abelson, Mrs. Samuel (Hannah Wozosky)	female	28.0	1	0	P/PP 3381	24.0000	C	1	0	1	0	0	0
875	876	1	3	Najib, Miss. Adele Kiamie "Jane"	female	15.0	0	0	2667	7.2250	C	0	1	1	0	0	0
876	877	0	3	Gustafsson, Mr. Alfred Ossian	male	20.0	0	0	7534	9.8458	S	0	1	0	1	0	1
877	878	0	3	Petroff, Mr. Nedelio	male	19.0	0	0	349212	7.8958	S	0	1	0	1	0	1
878	879	0	3	Laleff, Mr. Kristo	male	25.0	0	0	349217	7.8958	S	0	1	0	1	0	1
879	880	1	1	Potter, Mrs. Thomas Jr (Lily Alexenia Wilson)	female	56.0	0	1	11767	83.1583	C	0	0	1	0	0	0
880	881	1	2	Shelley, Mrs. William (Imanita Parrish Hall)	female	25.0	0	1	230433	26.0000	S	1	0	1	0	0	1
881	882	0	3	Markun, Mr. Johann	male	33.0	0	0	349257	7.8958	S	0	1	0	1	0	1
882	883	0	3	Dahlberg, Miss. Gerda Ulrika	female	22.0	0	0	7552	10.5167	S	0	1	1	0	0	1
883	884	0	2	Banfield, Mr. Frederick James	male	28.0	0	0	C.A./SOTON 34068	10.5000	S	1	0	0	1	0	1
884	885	0	3	Sutehall, Mr. Henry Jr	male	25.0	0	0	SOTON/OQ 392076	7.0500	S	0	1	0	1	0	1
885	886	0	3	Rice, Mrs. William (Margaret Norton)	female	39.0	0	5	382652	29.1250	Q	0	1	1	0	1	0
886	887	0	2	Montvila, Rev. Juozas	male	27.0	0	0	211536	13.0000	S	1	0	0	1	0	1
887	888	1	1	Graham, Miss. Margaret Edith	female	19.0	0	0	112053	30.0000	S	0	0	1	0	0	1
888	889	0	3	Johnston, Miss. Catherine Helen "Carrie"	female	25.0	1	2	W./C. 6607	23.4500	S	0	1	1	0	0	1
889	890	1	1	Behr, Mr. Karl Howell	male	26.0	0	0	111369	30.0000	C	0	0	0	1	0	0
890	891	0	3	Dooley, Mr. Patrick	male	32.0	0	0	370376	7.7500	Q	0	1	0	1	1	0

891 rows × 17 columns

```
In [479]: #We drop/Remove the columns from the dataframe that we don't need
trainset.drop(["PassengerId", "Pclass", "Name", "Sex", "Ticket", "Embarked", "female"],axis=1,inplace=True)
```

Out[480]:

891 rows x 10 columns

```
In [516]: #Converting the variables dataset x and y to NumPy Arrays
```

```

0 1 1 1 0 0 0 0 1 1 1 0 0 0 1 0 1 0 1 1 1 1 0 1 0 0 1 0 0 1 1 0 0 0 1
0 0 1 0 0 0 1 1 1 0 0 1 0 0 0 0 0 1 1 0 1 1 0 1 0 0 1 0 0 1 0 0 1 0 0 0 0 1
1 0 0 0 1 0 1 0 1 1 1 0 1 1 0 0 1 0 0 0 0 0 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 1 0 0 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 1
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 1 0 0 0 1 0 1 1 1 1 0 0 0 1 0 0 0 0 0 0 0 0 0 1 1 1 0 1 0 0 0 0 1 1 0 1 0 1 0
0 0 1 0 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0 1 1 1
1 0 1 0 0 0 0 0 0 1 1 1 0 1 1 1 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 1 1 1 1 0 0 0 0 0
0 0 1 1 1 1 0 1 0 1 1 1 0 1 1 1 0 0 0 1 1 1 0 1 0 0 1 1 0 1 0 1 1 1 1 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 0 0 0 0 1 1 0 0 0 1 1 0 1 0 0 0 0 1 0 1 1 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0 0
1 0 0 0 0 1 0 1 0 1 1 0 0 0 0 0 0 0 0 0 0 1 1 0 1 1 1 0 0 1 0 0 1 0 0 1 0 0 1
1 1 1 1 1 1 0 0 0 1 0 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0 1
0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 1 0 0 0 0 0 0 1 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 1 0 1 1 1 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 0 1 1 0 1 1 1 0 0 0 0 1 1 0 0 0 1 1
0 0 0 0 1 0 0 1 1 0 0 0 0 0 0 0 1 1 1 0 0 0 1 0 0 1 1 1 0 0 0 0 1 1 0 0 0 0 1
0 1 0 1 0 0 0 1 0 0 0 0 0 0 1 0 1 1 1 0 1 0 1 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0
0 0 0 1 1 1 1 0 1 0 0 0 0 1 1 0 0 1 0 0 0 1 0 1 1 1 0 0 1 0 0 0 0 0 0 0 0 0 0 0
0 0 1 1 0 0 0 1 0 0 1 1 0 0 0 1 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 0 0 0 0 0 0 0 1 0
1 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 0 0 0 0 0 0 0 1 1 1 0 0 0 0 1 1 1 0 0 0 0 1 0 0 0
0 0 0 0 0 0 1 1 0 1 0 0 0 1 1 1 1 1 0 0 0 0 0 0 0 1 1 0 0 1 1 1 0 0 0 0 0 0 0 1 0
0 0 1 1 1 1 0 0 0 0 1 0 0 0 1 1 0 0 1 0 0 0 1 0 0 0 0 1 1 0 0 0 1 1 0 0 0 0 0 0
0 1 1

```

```
In [484]: #Function to define cost Function
def cost_function(x, y, theta):
    m = float(len(y))
    h = sigmoid(np.dot(x, theta))
    cost = (1 / m) * (np.dot(-y.T, np.log(h)) - np.dot((1 - y).T, np.log(1 - h)))
    return cost.sum()
```



In [491]:

testset

Out[491]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	NaN	S
5	897	3	Svensson, Mr. Johan Cervin	male	14.0	0	0	7538	9.2250	NaN	S
6	898	3	Connolly, Miss. Kate	female	30.0	0	0	330972	7.6292	NaN	Q
7	899	2	Caldwell, Mr. Albert Francis	male	26.0	1	1	248738	29.0000	NaN	S
8	900	3	Abraham, Mrs. Joseph (Sophie Halaut Easu)	female	18.0	0	0	2657	7.2292	NaN	C
9	901	3	Davies, Mr. John Samuel	male	21.0	2	0	A/4 48871	24.1500	NaN	S
10	902	3	Ilieff, Mr. Ylio	male	NaN	0	0	349220	7.8958	NaN	S
11	903	1	Jones, Mr. Charles Cresson	male	46.0	0	0	694	26.0000	NaN	S
12	904	1	Snyder, Mrs. John Pillsbury (Nelle Stevenson)	female	23.0	1	0	21228	82.2667	B45	S
13	905	2	Howard, Mr. Benjamin	male	63.0	1	0	24065	26.0000	NaN	S
14	906	1	Chaffee, Mrs. Herbert Fuller (Carrie Constance...	female	47.0	1	0	W.E.P. 5734	61.1750	E31	S
15	907	2	del Carlo, Mrs. Sebastiano (Argenia Genovesi)	female	24.0	1	0	SC/PARIS 2167	27.7208	NaN	C
16	908	2	Keane, Mr. Daniel	male	35.0	0	0	233734	12.3500	NaN	Q
17	909	3	Assaf, Mr. Gerios	male	21.0	0	0	2692	7.2250	NaN	C
18	910	3	Ilmakangas, Miss. Ida Livija	female	27.0	1	0	STON/O2. 3101270	7.9250	NaN	S
19	911	3	Assaf Khalil, Mrs. Mariana (Miriam)"	female	45.0	0	0	2696	7.2250	NaN	C
20	912	1	Rothschild, Mr. Martin	male	55.0	1	0	PC 17603	59.4000	NaN	C
21	913	3	Olson, Master. Artur Karl	male	9.0	0	1	C 17368	3.1708	NaN	S
22	914	1	Fliegenheim, Mrs. Alfred (Antoinette)	female	NaN	0	0	PC 17598	31.6833	NaN	S
23	915	1	Williams, Mr. Richard Norris II	male	21.0	0	1	PC 17597	61.3792	NaN	C
24	916	1	Ryerson, Mrs. Arthur Larned (Emily Maria Borie)	female	48.0	1	3	PC 17608	262.3750	B57 B59 B63 B66	C
25	917	3	Robins, Mr. Alexander A	male	50.0	1	0	A/5. 3337	14.5000	NaN	S
26	918	1	Ostby, Miss. Helene Ragnhild	female	22.0	0	1	113509	61.9792	B36	C
27	919	3	Daher, Mr. Shedid	male	22.5	0	0	2698	7.2250	NaN	C
28	920	1	Brady, Mr. John Bertram	male	41.0	0	0	113054	30.5000	A21	S
29	921	3	Samaan, Mr. Elias	male	NaN	2	0	2662	21.6792	NaN	C
...	...	...	...	...	...	...	...	...	...	...	...
805	1280	3	Canavan, Mr. Patrick	male	21.0	0	0	364858	7.7500	NaN	Q
806	1281	3	Paisson, Master. Paul Folke	male	6.0	3	1	349909	21.0750	NaN	S
807	1282	1	Payne, Mr. Vivian Ponsonby	male	23.0	0	0	12749	93.5000	B24	S
808	1283	1	Lines, Mrs. Ernest H (Elizabeth Lindsey James)	female	51.0	0	1	PC 17592	39.4000	D28	S
809	1284	3	Abbott, Master. Eugene Joseph	male	13.0	0	2	C.A. 2673	20.2500	NaN	S
810	1285	2	Gilbert, Mr. William	male	47.0	0	0	C.A. 30769	10.5000	NaN	S
811	1286	3	Kink-Heilmann, Mr. Anton	male	29.0	3	1	315153	22.0250	NaN	S
812	1287	1	Smith, Mrs. Lucien Philip (Mary Eloise Hughes)	female	18.0	1	0	13695	60.0000	C31	S
813	1288	3	Colbert, Mr. Patrick	male	24.0	0	0	371109	7.2500	NaN	Q
814	1289	1	Frolicher-Stehli, Mrs. Maxmillian (Margaretha ...	female	48.0	1	1	13567	79.2000	B41	C
815	1290	3	Larsson-Rondberg, Mr. Edvard A	male	22.0	0	0	347065	7.7750	NaN	S
816	1291	3	Conlon, Mr. Thomas Henry	male	31.0	0	0	21332	7.7333	NaN	Q
817	1292	1	Bonnett, Miss. Caroline	female	30.0	0	0	36928	164.8667	C7	S
818	1293	2	Gale, Mr. Harry	male	38.0	1	0	28664	21.0000	NaN	S
819	1294	1	Gibson, Miss. Dorothy Winifred	female	22.0	0	1	112378	59.4000	NaN	C
820	1295	1	Carrau, Mr. Jose Pedro	male	17.0	0	0	113059	47.1000	NaN	S
821	1296	1	Frauenthal, Mr. Isaac Gerald	male	43.0	1	0	17765	27.7208	D40	C
822	1297	2	Nourney, Mr. Alfred (Baron von Drachstedt)"	male	20.0	0	0	SC/PARIS 2166	13.8625	D38	C
823	1298	2	Ware, Mr. William Jeffery	male	23.0	1	0	28666	10.5000	NaN	S
824	1299	1	Widener, Mr. George Dunton	male	50.0	1	1	113503	211.5000	C80	C
825	1300	3	Riordan, Miss. Johanna Hannah""	female	NaN	0	0	334915	7.7208	NaN	Q
826	1301	3	Peacock, Miss. Treasteall	female	3.0	1	1	SOTON/O.Q. 3101315	13.7750	NaN	S
827	1302	3	Naughton, Miss. Hannah	female	NaN	0	0	365237	7.7500	NaN	Q
828	1303	1	Minahan, Mrs. William Edward (Lillian E Thorpe)	female	37.0	1	0	19928	90.0000	C78	Q
829	1304	3	Henriksson, Miss. Jenny Lovisa	female	28.0	0	0	347086	7.7750	NaN	S
830	1305	3	Spector, Mr. Woolf	male	NaN	0	0	A.5. 3236	8.0500	NaN	S
831	1306	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758	108.9000	C105	C
832	1307	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	NaN	S
833	1308	3	Ware, Mr. Frederick	male	NaN	0	0	359309	8.0500	NaN	S
834	1309	3	Peter, Master. Michael J	male	NaN	1	1	2668	22.3583	NaN	C

835 rows × 11 columns

In [492]:

```
#Function to replace null ages with average value of age
def add_age(cols):
    Age = cols[0]
    Pclass = cols[1]
    if pd.isnull(Age):
        return int(testset[testset["Pclass"] == Pclass]["Age"].mean())
    else:
        return Age
```

In [493]:

```
#Calling the function to replace null age values
testset["Age"] = testset[["Age", "Pclass"]].apply(add_age,axis=1)
```



```
In [494]: #Cleaning the dataset
testset.drop("Cabin",inplace=True,axis=1)#Removing all values in cabin column
testset.dropna()#Removing rows where all the values are null
testset.head(n=891)
```

Out[494]:

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	S
4	896	3	Hirvonen, Mrs. Alexander (Helga E Lindqvist)	female	22.0	1	1	3101298	12.2875	S
5	897	3	Svensson, Mr. Johan Cervin	male	14.0	0	0	7538	9.2250	S
6	898	3	Connolly, Miss. Kate	female	30.0	0	0	330972	7.6292	Q
7	899	2	Caldwell, Mr. Albert Francis	male	26.0	1	1	248738	29.0000	S
8	900	3	Abraham, Mrs. Joseph (Sophie Halaut Easu)	female	18.0	0	0	2657	7.2292	C
9	901	3	Davies, Mr. John Samuel	male	21.0	2	0	A/4 48871	24.1500	S
10	902	3	Ilieff, Mr. Ylio	male	24.0	0	0	349220	7.8958	S
11	903	1	Jones, Mr. Charles Cresson	male	46.0	0	0	694	26.0000	S
12	904	1	Snyder, Mrs. John Pillsbury (Nelle Stevenson)	female	23.0	1	0	21228	82.2667	S
13	905	2	Howard, Mr. Benjamin	male	63.0	1	0	24065	26.0000	S
14	906	1	Chaffee, Mrs. Herbert Fuller (Carrie Constance...	female	47.0	1	0	W.E.P. 5734	61.1750	S
15	907	2	del Carlo, Mrs. Sebastiano (Argentina Genovesi)	female	24.0	1	0	SC/PARIS 2167	27.7208	C
16	908	2	Keane, Mr. Daniel	male	35.0	0	0	233734	12.3500	Q
17	909	3	Assaf, Mr. Gerios	male	21.0	0	0	2692	7.2250	C
18	910	3	Ilmakangas, Miss. Ida Livija	female	27.0	1	0	STON/O2. 3101270	7.9250	S
19	911	3	Assaf Khalil, Mrs. Mariana (Miriam)"	female	45.0	0	0	2696	7.2250	C
20	912	1	Rothschild, Mr. Martin	male	55.0	1	0	PC 17603	59.4000	C
21	913	3	Olsen, Master. Artur Karl	male	9.0	0	1	C 17368	3.1708	S
22	914	1	Flegenheim, Mrs. Alfred (Antoinette)	female	40.0	0	0	PC 17598	31.6833	S
23	915	1	Williams, Mr. Richard Norris II	male	21.0	0	1	PC 17597	61.3792	C
24	916	1	Ryerson, Mrs. Arthur Larned (Emily Maria Borie)	female	48.0	1	3	PC 17608	262.3750	C
25	917	3	Robins, Mr. Alexander A	male	50.0	1	0	A/5. 3337	14.5000	S
26	918	1	Ostby, Miss. Helene Ragnhild	female	22.0	0	1	113509	61.9792	C
27	919	3	Daher, Mr. Shedid	male	22.5	0	0	2698	7.2250	C
28	920	1	Brady, Mr. John Bertram	male	41.0	0	0	113054	30.5000	S
29	921	3	Samaan, Mr. Elias	male	24.0	2	0	2662	21.6792	C
...	...	...	...	...	...	...	...	...	...	...
805	1280	3	Canavan, Mr. Patrick	male	21.0	0	0	364858	7.7500	Q
806	1281	3	Palsson, Master. Paul Folke	male	6.0	3	1	349909	21.0750	S
807	1282	1	Payne, Mr. Vivian Ponsonby	male	23.0	0	0	12749	93.5000	S
808	1283	1	Lines, Mrs. Ernest H (Elizabeth Lindsey James)	female	51.0	0	1	PC 17592	39.4000	S
809	1284	3	Abbott, Master. Eugene Joseph	male	13.0	0	2	C.A. 2673	20.2500	S
810	1285	2	Gilbert, Mr. William	male	47.0	0	0	C.A. 30769	10.5000	S
811	1286	3	Kirk-Heilmann, Mr. Anton	male	29.0	3	1	315153	22.0250	S
812	1287	1	Smith, Mrs. Lucien Philip (Mary Eloise Hughes)	female	18.0	1	0	13695	60.0000	S
813	1288	3	Colbert, Mr. Patrick	male	24.0	0	0	371109	7.2500	Q
814	1289	1	Frolicher-Stehli, Mrs. Maximilian (Margaretha ...	female	48.0	1	1	13567	79.2000	C
815	1290	3	Larsson-Rondberg, Mr. Edvard A	male	22.0	0	0	347065	7.7750	S
816	1291	3	Conlon, Mr. Thomas Henry	male	31.0	0	0	21332	7.7333	Q
817	1292	1	Bonnell, Miss. Caroline	female	30.0	0	0	36828	164.8667	S
818	1293	2	Gale, Mr. Harry	male	38.0	1	0	28664	21.0000	S
819	1294	1	Gibson, Miss. Dorothy Winifred	female	22.0	0	1	112378	59.4000	C
820	1295	1	Carrau, Mr. Jose Pedro	male	17.0	0	0	113059	47.1000	S
821	1296	1	Frauenthal, Mr. Isaac Gerald	male	43.0	1	0	17765	27.7208	C
822	1297	2	Nourney, Mr. Alfred (Baron von Drachstedt)"	male	20.0	0	0	SC/PARIS 2166	13.8625	C
823	1298	2	Ware, Mr. William Jeffery	male	23.0	1	0	28666	10.5000	S
824	1299	1	Widener, Mr. George Dunton	male	50.0	1	1	113503	211.5000	C
825	1300	3	Riordan, Miss. Johanna Hannah""	female	24.0	0	0	334915	7.7208	Q
826	1301	3	Peacock, Miss. Treasteall	female	3.0	1	1	SOTON/O.Q. 3101315	13.7750	S
827	1302	3	Naughton, Miss. Hannah	female	24.0	0	0	365237	7.7500	Q
828	1303	1	Minahan, Mrs. William Edward (Lillian E Thorpe)	female	37.0	1	0	19928	90.0000	Q
829	1304	3	Henriksson, Miss. Jenny Lovisa	female	28.0	0	0	347086	7.7750	S
830	1305	3	Spector, Mr. Woolf	male	24.0	0	0	A.5. 3236	8.0500	S
831	1306	1	Oliva y Ocana, Dona. Fermina	female	39.0	0	0	PC 17758	108.9000	C
832	1307	3	Saether, Mr. Simon Sivertsen	male	38.5	0	0	SOTON/O.Q. 3101262	7.2500	S
833	1308	3	Ware, Mr. Frederick	male	24.0	0	0	359309	8.0500	S
834	1309	3	Peter, Master. Michael J	male	24.0	1	1	2668	22.3583	C

835 rows x 10 columns

```
In [495]: #Converting the categorical data into numerical data
sex = pd.get_dummies(testset['Sex'], drop_first=True)
embarked = pd.get_dummies(testset['Embarked'], drop_first=True)
pclass = pd.get_dummies(testset['Pclass'], drop_first=True)
testset = pd.concat([testset, sex, embarked, pclass], axis=1)

In [496]: #We drop/Remove the columns from the dataframe that we don't need
testset.drop(['PassengerId', 'Pclass', 'Name', 'Sex', 'Ticket', 'Embarked'], axis=1, inplace=True)
```

```
In [497]: testset = testset.dropna() #Deletes all rows with all null values
testset
```

Out[497]:

	Age	SibSp	Parch	Fare	male	Q	S	2	3
0	34.5	0	0	7.8292	1	1	0	0	1
1	47.0	1	0	7.0000	0	0	1	0	1
2	62.0	0	0	9.6875	1	1	0	1	0
3	27.0	0	0	8.6625	1	0	1	0	1
4	22.0	1	1	12.2875	0	0	1	0	1
5	14.0	0	0	9.2250	1	0	1	0	1
6	30.0	0	0	7.6292	0	1	0	0	1
7	26.0	1	1	29.0000	1	0	1	1	0
8	18.0	0	0	7.2292	0	0	0	0	1
9	21.0	2	0	24.1500	1	0	1	0	1
10	24.0	0	0	7.8958	1	0	1	0	1
11	46.0	0	0	26.0000	1	0	1	0	0
12	23.0	1	0	82.2667	0	0	1	0	0
13	63.0	1	0	26.0000	1	0	1	1	0
14	47.0	1	0	61.1750	0	0	1	0	0
15	24.0	1	0	27.7208	0	0	0	1	0
16	35.0	0	0	12.3500	1	1	0	1	0
17	21.0	0	0	7.2250	1	0	0	0	1
18	27.0	1	0	7.9250	0	0	1	0	1
19	45.0	0	0	7.2250	0	0	0	0	1
20	55.0	1	0	59.4000	1	0	0	0	0
21	9.0	0	1	3.1708	1	0	1	0	1
22	40.0	0	0	31.6833	0	0	1	0	0
23	21.0	0	1	61.3792	1	0	0	0	0
24	48.0	1	3	262.3750	0	0	0	0	0
25	50.0	1	0	14.5000	1	0	1	0	1
26	22.0	0	1	61.9792	0	0	0	0	0
27	22.5	0	0	7.2250	1	0	0	0	1
28	41.0	0	0	30.5000	1	0	1	0	0
29	24.0	2	0	21.6792	1	0	0	0	1
...	...	...	...	...	...	...	...	...	...
805	21.0	0	0	7.7500	1	1	0	0	1
806	6.0	3	1	21.0750	1	0	1	0	1
807	23.0	0	0	93.5000	1	0	1	0	0
808	51.0	0	1	39.4000	0	0	1	0	0
809	13.0	0	2	20.2500	1	0	1	0	1
810	47.0	0	0	10.5000	1	0	1	1	0
811	29.0	3	1	22.0250	1	0	1	0	1
812	18.0	1	0	60.0000	0	0	1	0	0
813	24.0	0	0	7.2500	1	1	0	0	1
814	48.0	1	1	79.2000	0	0	0	0	0
815	22.0	0	0	7.7750	1	0	1	0	1
816	31.0	0	0	7.7333	1	1	0	0	1
817	30.0	0	0	164.8667	0	0	1	0	0
818	38.0	1	0	21.0000	1	0	1	1	0
819	22.0	0	1	59.4000	0	0	0	0	0
820	17.0	0	0	47.1000	1	0	1	0	0
821	43.0	1	0	27.7208	1	0	0	0	0
822	20.0	0	0	13.8625	1	0	0	1	0
823	23.0	1	0	10.5000	1	0	1	1	0
824	50.0	1	1	211.5000	1	0	0	0	0
825	24.0	0	0	7.7208	0	1	0	0	1
826	3.0	1	1	13.7750	0	0	1	0	1
827	24.0	0	0	7.7500	0	1	0	0	1
828	37.0	1	0	90.0000	0	1	0	0	0
829	28.0	0	0	7.7750	0	0	1	0	1
830	24.0	0	0	8.0500	1	0	1	0	1
831	39.0	0	0	108.9000	0	0	0	0	0
832	38.5	0	0	7.2500	1	0	1	0	1
833	24.0	0	0	8.0500	1	0	1	0	1
834	24.0	1	1	22.3583	1	0	0	0	1

833 rows x 9 columns

```
In [498]: x_test = np.array(testset) #Converting the dataset to an array
print(x_test)
```

```
[[ 34.5  0.  0. ...,  0.  0.  1. ]
 [ 47.   1.  0. ...,  1.  0.  1. ]
 [ 62.   0.  0. ...,  0.  1.  0. ]
 ...,
 [ 38.5  0.  0. ...,  1.  0.  1. ]
 [ 24.   0.  0. ...,  1.  0.  1. ]
 [ 24.   1.  1. ...,  0.  0.  1. ]]
```

```
test_pred = predict(x_test, theta)
for t in range(len(test_pred)):
    if test_pred[t] >= 0.5:
        test_pred[t] = 1
    else:
        test_pred[t] = 0
print(test_pred) # Prediction where 0 :- Dead 1 :- Alive
```