

```
In [3]: import pandas as pd
import numpy as np
import seaborn as sns
import pylab as py

In [2]: %pylab inline
Populating the interactive namespace from numpy and matplotlib

In [3]: cdf=pd.read_csv('train (1).csv')

In [4]: cdf.head()
```

	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	31.01282	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	

```
In [6]: cdf.sample(10)
```

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
265	266	0	2	Reeves, Mr. David	male	36.0	0	0	C.A. 17248	10.5000	NaN	S
972	972	1	1	Appleton, Mrs. Edward Dale (Charlotte Lamson)	female	53.0	2	0	11769	51.4792	C101	S
225	226	0	3	Berglund, Mr. Karl Ivar Sven	male	22.0	0	0	PP 4348	9.3500	NaN	S
808	809	0	2	Meyer, Mr. August	male	39.0	0	0	248723	13.0000	NaN	S
17	18	1	2	Williams, Mr. Charles Eugene	male	NaN	0	0	244373	13.0000	NaN	S
153	154	0	3	van Billard, Mr. Austin Blyler	male	40.5	0	2	A/5. 851	14.5000	NaN	S
56	57	1	2	Rugg, Miss. Emily	female	21.0	0	0	C.A. 31026	10.5000	NaN	S
88	89	1	3	Andersson, Miss. Elna Alexandra	female	17.0	4	2	3101281	7.9250	NaN	S
585	586	1	1	Tausig, Miss. Ruth	female	18.0	0	2	110413	79.6500	E58	S
486	487	1	1	Hoyt, Mrs. Frederick Maxfield (Jane Anne Forby)	female	35.0	1	0	19943	90.0000	C93	S

```
In [7]: cdf.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
#   column      Non-Null Count  Dtype
---  --
0   PassengerId  891 non-null      int64
1   Survived    891 non-null      int64
2   Pclass      891 non-null      int64
3   Name        891 non-null      object
4   Sex         891 non-null      object
5   Age         714 non-null      float64
6   SibSp       891 non-null      int64
7   Parch       891 non-null      int64
8   Ticket      891 non-null      object
9   Fare        891 non-null      float64
10  Cabin       204 non-null      object
11  Embarked    889 non-null      object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
```

```
In [9]: cdf.isnull().sum()
```

```
PassengerId    0
Survived        0
Pclass         0
Name            0
Sex            0
Age          177
SibSp          0
Parch          0
Ticket         0
Fare           0
Cabin        687
Embarked       2
dtype: int64
```

```
In [11]: fig=py.figure(figsize=(35,15))
sns.heatmap(cdf.isnull(),yticklabels=False,)
```

```
Out[11]: <AxesSubplot>
```