Introduction to Pandas and NumPy: Part 1

```
import pandas as pd
import numpy as np
import numpy
print('numpy:{}'.format(numpy. version ))
import pandas
print('pandas:{}'.format(pandas.__version__))
  numpy:1.23.5
  pandas:1.5.3
titanic = pd.read csv("titanic.csv")
titanic.head().style.set_properties(**{'background-color': 'Black',
                        'color': 'white'.
                        'border-color': 'darkblack'})
   Passengerld Survived Pclass
                                                          Age SibSp Parch
                                                                                      Fare Cabin Embarked
                                   Braund, Mr. Owen Harris male 22.000000
                                                                            A/5 21171 7.250000
                        Cumings, Mrs. John Bradley (Florence Briggs
Thayer) female 38.000000
                                                                            PC 17599 71.283300
                                                                            STON/O2
2
         3
                                    Heikkinen, Miss. Laina female 26.000000
                                                                                    7.925000
                                                                                           nan
                        Futrelle, Mrs. Jacques Heath (Lily May Peel) female 35.000000
                                                                              113803 53 100000 C123
                                    Allen, Mr. William Henry male 35.000000
                                                                             373450
                                                                                   8.050000
wine = pd.read_csv("winequality-red.csv")
wine.head().style.background_gradient(cmap='Dark2')
   fixed acidity volatile acidity citric acid residual sugar chlorides free sulfur dioxide total sulfur dioxide density
                                                                               pH sulphates alcohol quality
 0
               0.700000
                                                                                   0.650000 9,800000
                               2.300000 0.092000
                                                              54.000000 0.997000 3.260000
                                                              60.000000 0.998000 3.160000
                                                                                   0.580000 9.800000
                                1.900000 0.075000
                                                  11 000000
                                                              34.000000 0.997800 3.510000 0.560000 9.400000
               0.700000 0.000000
#Drop Columns
titanic = titanic.drop(['PassengerId','Name','Ticket'],axis=1)
#Drop Rows
titanic = titanic.drop(labels=[0,3,6],axis=0)
titanic.head().style.background_gradient(cmap='copper')
                                                 SibSp
      Survived
                Pelass
                                                         Parch
                              Sex
                                          Age
                                                                               Cabin
                          female
                                                              0
  2
                           female
                                                      0
                                                                   7.925000
                                                                                                 s
                                                                                 nan
              O
                                                      O
                                                              O
                                                                   8.050000
  4
                       3
                            male
                                    35 000000
                                                                                 nan
                                                                                                 S
  5
              0
                                                                   8.458300
                                                                                                 Q
```

2.000000

male

1 21.075000

nan

titanic[['Age','Fare','Pclass']].agg(['sum','max','mean','std','skew','kur t']).style.background_gradient(cmap='copper')

	Age	Fare	Pclass
sum	21205.170000	28693.949300	2057.000000
max	80.000000	512.329200	3.000000
mean	29.699118	32.204208	2.308642
std	14.526497	49.693429	0.836071
skew	0.389108	4.787317	-0.630548
kurt	0.178274	33.398141	-1.280015

titanic[titanic['Survived']==0].describe().T.style.background_gradien t(subset=['mean','std','50%','count'], cmap='RdPu')

	count	mean	std	min	25%	50%	75%	max	
Survived	547.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
Pclass	547.000000	2.533821	0.733955	1.000000	2.000000	3.000000	3.000000	3.000000	
Age	422.000000	30.591232	14.153698	1.000000	21.000000	28.000000	39.000000	74.000000	
SibSp	547.000000	0.553931	1.290398	0.000000	0.000000	0.000000	1.000000	8.000000	
Parch	547.000000	0.330896	0.824430	0.000000	0.000000	0.000000	0.000000	6.000000	
Fare	547.000000	22.090690	31.413411	0.000000	7.854200	10.500000	26.000000	263.000000	

titanic.describe(percentiles=[0.05,0.25,0.35,0.5,0.75,0.85,0.95,0.995, 0.999]).style.background_gradient(cmap='copper')

	Survived	Pclass	Age	SibSp	Parch	Fare
count	888.000000	888.000000	711.000000	888.000000	888.000000	888.000000
mean	0.384009	2.310811	29.668312	0.522523	0.382883	32.186641
std	0.486634	0.834850	14.524290	1.104235	0.807113	49.761015
min	0.000000	1.000000	0.420000	0.000000	0.000000	0.000000
5%	0.000000	1.000000	4.000000	0.000000	0.000000	7.225000
25%	0.000000	2.000000	20.000000	0.000000	0.000000	7.917700
35%	0.000000	2.000000	24.000000	0.000000	0.000000	9.000000
50%	0.000000	3.000000	28.000000	0.000000	0.000000	14.454200
75%	1.000000	3.000000	38.000000	1.000000	0.000000	31.000000
85%	1.000000	3.000000	45.000000	1.000000	1.000000	56.495800
95%	1.000000	3.000000	56.000000	3.000000	2.000000	112.437905
99.5%	1.000000	3.000000	70.725000	8.000000	5.000000	263.000000
99.9%	1.000000	3.000000	75.740000	8.000000	5.113000	512.329200
max	1.000000	3.000000	80.000000	8.000000	6.000000	512.329200

titanic[['Age','Embarked','Fare']].count().to_frame().style.background _gradient(cmap='copper')



titanic['Pclass'].value_counts().tolist()

```
[490, 214, 184]
```

titanic['Embarked'].value_counts(normalize= True, sort = True, ascending= False).to_frame().style.background_gradient(cmap='copper')

Embarked

s 0.723476c 0.189616Q 0.086907

titanic['Embarked'][titanic['Sex']=='female'].value_counts(normalize=True)*100

S 64.951768 C 23.472669 Q 11.575563

Name: Embarked, dtype: float64

(titanic['Survived'].value_counts()/len(titanic['Survived'])).to_frame().style.background_gradient(cmap='copper')

Survived 0 0.615991 1 0.384009

corr = titanic.groupby(["Embarked"])[["Fare" , "Age"]].corr()
corr.head().style.background_gradient(cmap='copper')

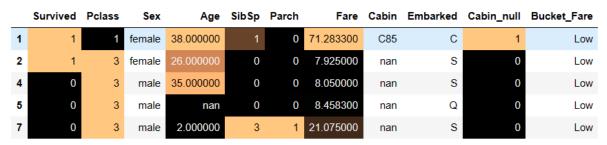
	Fare	Age
Embarked		

С	Fare	1.000000	0.160451
C	Age	0.160451	1.000000
	Fare	1.000000	0.027276
Q	Age	0.027276	1.000000
s	Fare	1.000000	0.049505

titanic['Cabin_null'] = np.where(titanic['Cabin'].isnull(),0,1) titanic.head().style.background_gradient(cmap='copper')

	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked	Cabin_null
1	1	1	female	38.000000	1	0	71.283300	C85	С	1
2	1	3	female	26.000000	0	0	7.925000	nan	S	0
4	0	3	male	35.000000	0	0	8.050000	nan	S	0
5	0	3	male	29.668312	0	0	8.458300	nan	Q	0
7	0	3	male	2.000000	3	1	21.075000	nan	S	0

titanic["Bucket_Fare"] = np.where(titanic["Fare"] < 250, "Low", "High") titanic.head().style.background_gradient(cmap='copper')



titanic_room= titanic.groupby(['Embarked','Sex'])['Age'].mean().reset_index() titanic_room.head().style.background_gradient(cmap='copper')

	Embarked	Sex	Age
0	С	female	28.344262
1	С	male	32.998841
2	Q	female	24.291667
3	Q	male	30.937500
4	S	female	27.732432

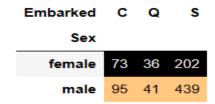
titanic.groupby("Embarked").agg({"Fare": np.mean, "Sex": np.size}).style.background_gradient(cmap='copper')



titanic.pivot_table(index="Embarked").style.background_gradient(cmap='c opper')

	Age	Cabin_null	Fare	Parch	Pclass	SibSp	Survived
Embarked	I						
C	30.814769	0.410714	59.954144	0.363095	1.886905	0.386905	0.553571
G	28.089286	0.051948	13.276030	0.168831	2.909091	0.428571	0.389610
S	29.404265	0.198128	27.031492	0.414977	2.354134	0.570983	0.336973

pd.crosstab(titanic['Sex'],titanic['Embarked']).style.background_gradient(cmap='copper')



titanic['sex Titanic map']=titanic['Sex'].map({'male':1,'female':0}) titanic.head().style.background_gradient(cmap='copper')

	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked	Cabin_null	Bucket_Fare	sex Titanic map
1	1	1	female	38.000000		0	71.283300	C85	С	1	Low	0
2	1	3	female	26.000000	0	0	7.925000	nan	S	0	Low	0
4	0	3	male	35.000000	0	0	8.050000	nan	S	0	Low	1
5	0	3	male	nan	0	0	8.458300	nan	Q	0	Low	1
7	0	3	male	2.000000	3	1	21.075000	nan	S	0	Low	1

titanic["age_bins"]= pd.cut(titanic["Age"] ,bins=[1,18,29 , 40 , 50 , 60 , 80] ,labels=["child","teen","adult" , "fortieth" , "old" , "ancient"]) titanic.head().style.background_gradient(cmap='copper')



titanic["Survived"].replace({0:"Died", 1:"Saved"}, inplace=True) titanic.head().style.background_gradient(cmap='copper')



titanic['Pclass'][titanic['Pclass'] == 1] = 'Rich' titanic['Pclass'][titanic['Pclass'] == 2] = 'Middel Class' titanic['Pclass'][titanic['Pclass'] == 3] = 'Poor'

titanic.head().style.background_gradient(cmap='copper')

	Survived	Pclass	Sex	Age	SibSp	Parch	Fare	Cabin	Embarked	Cabin_null	Bucket_Fare	sex Titanic map	age_bins
1	Saved	Rich	female	38.000000		0	71.283300	C85	С	1	Low	0	adult
2	Saved	Poor	female	26.000000	0	0	7.925000	nan	S	0	Low	0	teen
4	Died	Poor	male	35.000000	0	0	8.050000	nan	s	0	Low	1	adult
5	Died	Poor	male	nan	0	0	8.458300	nan	Q	0	Low	1	nan
7	Died	Poor	male	2.000000	3	1	21.075000	nan	S	0	Low	1	child