Pandas Revision

Important Syntax in Panda's :

```
df=pd.DataFrame (data=value,index[],columns[])
df.set_index()
df.reset_index()
df.transpose()

## Reading Data from CSV and Excel :
df=pd.read_csv/excel("Path")

## Exploring Data sets :
df.info
df.describe
df.memory_usage(index=True/False,deep=True/False)

## Accessing Data set :
df[]

df.astype() .... for changing the data type

## Data Preparation :
df.isna()
df.dropna(how="all/any",inplace=True)
df.drop([values],axis=0/1)
df.rename(columns={},inplace=True)
df.fillna("_",inplace=True/False)
df[["prefix","sufix"]]=df["column_name"].str.split(pat="_",expand=True)

## Filter the Data set :
df.["_"].unique()
df.["_"].count()
df.["_"].nunique()
df.["_"].value_counts()

df["Fare"] > 100
df[df["Fare"] > 100]
df[(df["Fare"] > 100) & (df["class"]==3)]

df.iloc[rows,coulmns]...we hgave to type the indexes
df.loc[:,columns] ...we have to type the names

## Creating Custom Columns
df["revised_fare"] = df.apply(lambda x: x["Fare"]*100, axis=1)
df["revidef_fare"] = df["fare"] * 100
```

```
## Group by
df.groupby(by="_")

#Joining Two columns :
merge_columns=pd.merge(merge1,merge2,left_on="_",right_on="_",how="left/right")

#Saving the data file :
df.to_excel/csv("path/filename.xlsx or filename.csv", index = True/False)

## Some other functions
1. Lead and lag in the pandas
df["column_name"].shift(1/2/-1/-2 etc)

2. how to categories in panda's ?

cutoff = [0,10,20]
tags = ["bad","good","excellent"]

df["Category"]=pd.cut(df["column"],bins=cutoff,labels=tags)

3.sorting the values
df.sort_values(by="column_name",ascending=True/False)

4.Filtering the columns from the table
df[["column1","column2"]] or
df.filter ("column1","column2")

5.Ranking in the Panda's

df["Rank"]=df["column_to_rank"].rank(axis=0/1,method
="dense/first/min/max/average",ascending=True/False,na_option="keep/top/bottom")
```