Reading Writing csv, excel Files

1. Read CSV

In [14]:
```
import pandas as pd
df = pd.read_csv("stock_data.csv")
df
```

Out[14]:

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845	larry page
1	WMT	4.61	484	65	n.a.
2	MSFT	-1	85	64	bill gates
3	RIL	not available	50	1023	mukesh ambani
4	TATA	5.6	-1	n.a.	ratan tata

Note: 1. if we want to skip certain rows, we need to give argument to read_csv as: skiprows=1 (how many) | and header =1 argument also suggest to make row 1 as header. 2. header = None (to keep no header), custom header titile => names = ["title1", "title2", ...] 3. to display custom rows: nrows = 3 (displays first 3 rows)

In [17]:
```
df = pd.read_csv("stock_data.csv", na_values=["not available",
"n.a."])
df
```

Out[17]:

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845.0	larry page
1	WMT	4.61	484	65.0	NaN
2	MSFT	-1.00	85	64.0	bill gates
3	RIL	NaN	50	1023.0	mukesh ambani
4	TATA	5.60	-1	NaN	ratan tata

In [27]:
```
#instead of supplying list we can supply dictionary to reach each col
umn values.
df = pd.read_csv("stock_data.csv", na_values={
    'eps': ["not available", "n.a."],
    'revenue' : ["not available", "n.a.", -1],
    'price' : ["not available", "n.a."],
    'people' : ["not available", "n.a."]
})
df
```

Out[27]:

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87.0	845.0	larry page
1	WMT	4.61	484.0	65.0	NaN
2	MSFT	-1.00	85.0	64.0	bill gates
3	RIL	NaN	50.0	1023.0	mukesh ambani
4	TATA	5.60	NaN	NaN	ratan tata

2. Write CSV

In [28]:
```
df.to_csv("mynew.csv")
```

The above code will generate/create mynew.csv file with the above data. It will also have index values. If we want to remove those index values, we can send argument:: index=False . Also, to write only limited columns :

df.to_csv ("mynew.csv", columns=["tickers", "eps"]

To escape header:: header = False

3. Read Excel

In [29]:
```
df = pd.read_excel("stock_data.xlsx") # df = pd.read_excel("stock_data.xlsx", "sheet1")
df
```

Out[29]:

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845	larry page
1	WMT	4.61	484	65	n.a.
2	MSFT	-1	85	64	bill gates
3	RIL	not available	50	1023	mukesh ambani
4	TATA	5.6	-1	n.a.	ratan tata

Using Converters

In [40]:
```
def convert_people_cell(cell):
    if cell == "n.a.":
        return 'Raju'
    return cell

def convert_nan_cell(cell):
    if cell == "not available" or cell == "n.a.":
        return None
    return cell


df = pd.read_excel("stock_data.xlsx", "Sheet1", converters = {
    'people' : convert_people_cell,
    'eps' : convert_nan_cell,
    'price' : convert_nan_cell
})
df
```

Out[40]:

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845.0	larry page
1	WMT	4.61	484	65.0	Raju
2	MSFT	-1.00	85	64.0	bill gates
3	RIL	NaN	50	1023.0	mukesh ambani
4	TATA	5.60	-1	NaN	ratan tata

4. Write Excel

In [9]:
```
df.to_excel("newexcel_data.xlsx")
```

Arguments:: sheet_name = "...." , startrow=1, startcol=2, index=false

To write 2 dataframes to different sheets of excel

In [10]:
```
df_stocks = pd.DataFrame({
    'tickers': ['GOOGL', 'WMT', 'MSFT'],
    'price': [845, 65, 64 ],
    'pe': [30.37, 14.26, 30.97],
    'eps': [27.82, 4.61, 2.12]
})

df_weather =  pd.DataFrame({
    'day': ['1/1/2017','1/2/2017','1/3/2017'],
    'temperature': [32,35,28],
    'event': ['Rain', 'Sunny', 'Snow']
})
```

In [12]:
```
with pd.ExcelWriter('stocks_weather.xlsx') as writer:
    df_stocks.to_excel(writer, sheet_name="stocks")
    df_weather.to_excel(writer, sheet_name="weather")
```