

read_write_csv_excel (3)

January 12, 2020

##

Read/Write CSV and Excel Files in Pandas

0.0.1 Read CSV

```
[1]: import pandas as pd
```

```
df = pd.read_csv("stock_data.csv")
df
```

```
[1]:
```

	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

```
[2]: df = pd.read_csv("stock_data.csv", skiprows=3)
df
```

```
[2]:
```

	MSFT	-1	85	64	bill gates
0	RIL	not available	50	1023	mukesh ambani
1	TATA	5.6	-1	n.a.	ratan tata

```
[3]: df = pd.read_csv("stock_data.csv", header=2) # skiprows and header are kind of
      ↳ same
df
```

```
[3]:
```

	WMT	4.61	484	65	n.a.
0	MSFT	-1	85	64	bill gates
1	RIL	not available	50	1023	mukesh ambani
2	TATA	5.6	-1	n.a.	ratan tata

```
[4]: df = pd.read_csv("stock_data.csv", header=None, names =
      ↳ ["ticker", "eps", "revenue", "people"])
df
```

```
[4]:
```

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

```
[5]: df = pd.read_csv("stock_data.csv", header=None)
df
```

```
[5]:
```

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

```
[6]: df = pd.read_csv("stock_data.csv", nrows=2)
df
```

```
[6]:
```

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845	larry page
1	WMT	4.61	484	65	n.a.

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

```
[7]:
```

	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

```
[8]: df = pd.read_csv("stock_data.csv", na_values={
    'eps': ['not available'],
    'revenue': [-1],
    'people': ['not available', 'n.a.']
})
df
```

```
[8]:
```

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87.0	845	larry page
1	WMT	4.61	484.0	65	NaN
2	MSFT	-1.00	85.0	64	bill gates

3	RIL	NaN	50.0	1023	mukesh ambani
4	TATA	5.60	NaN	n.a.	ratan tata

0.0.2 Write to CSV

```
[9]: df.to_csv("new2.csv", index=True)
```

```
[10]: df.columns
```

```
[10]: Index(['tickers', 'eps', 'revenue', 'price', 'people'], dtype='object')
```

```
[11]: df.to_csv("new.csv", header=False)
```

```
[12]: df.to_csv("new.csv", columns=["tickers", "price"], index=False)
```

0.0.3 Read Excel

```
[13]: df = pd.read_excel("stock_data.xlsx", "Sheet1")
df
```

```
[13]:
```

	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

```
[14]: def convert_people_cell(cell):
        if cell=="n.a.":
            return 'Sam Walton'
        return cell

    def convert_price_cell(cell):
        if cell=="n.a.":
            return 50
        return cell

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

```
[14]:
```

	tickers	eps	revenue	price	people
0	GOOGL	27.82	87	845	larry page
1	WMT	4.61	484	65	Sam Walton

2	MSFT	-1	85	64	bill gates
3	RIL	not available	50	1023	mukesh ambani
4	TATA	5.6	-1	50	ratan tata

0.0.4 Write to Excel

```
[15]: df.to_excel("new.xlsx", sheet_name="stocks", index=False, startrow=2,
↳ startcol=1)
```

Write two dataframes to two separate sheets in excel

```
[16]: 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']
})
```

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