

## Agenda

- Importing pandas
  - Importing the dataset
  - DataFrame/Series
- Basic ops on a DataFrame
  - df.info()
  - df.head()
  - df.tail()
  - df.shape
- Basic ops on columns
  - Different ways of accessing coils
  - Check for unique values
  - Rename column
  - Deleting col
  - Creating new coils
- Basic ops on rows
  - Implicit/explicit index
  - df.index
    - Indexing in series
    - Slicing in series
  - loc/iloc
  - Indexing/Slicing in dataframe
  - Adding a row
  - Deleting a row
  - Working for duplicates
- Working with both rows and columns

## Installation of pandas

### Importing pandas

```
In [1]: !pip install pandas
```

Requirement already satisfied: pandas in /Users/nikhilishanghi/opt/anaconda3/lib/python3.9/site-packages (1.3.4)  
Requirement already satisfied: python-dateutil<2.8,>=2.7.3 in /Users/nikhilishanghi/opt/anaconda3/lib/python3.9/site-packages (from pandas) (2021.3)  
Requirement already satisfied: numpy>=1.17.3 in /Users/nikhilishanghi/opt/anaconda3/lib/python3.9/site-packages (from pandas) (1.20.3)  
Requirement already satisfied: six>=1.5 in /Users/nikhilishanghi/opt/anaconda3/lib/python3.9/site-packages (from python-dateutil<2.8,>=2.7.3->pandas) (1.16.0)

```
In [2]: import numpy as np
import pandas as pd
```

### Importing the dataset

```
In [47]: df=pd.read_csv("Users/nikhilishanghi/Downloads/81_owl-course-main-live/batches/2_Sept_Beg_Tue_Oct_Beg_Tue/97_Pandas_1/mckinsey.csv")
df
```

```
Out[47]:
```

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1967	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1962	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1957	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11920563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows x 6 columns

```
In [4]: type(df)
```

```
Out[4]: pandas.core.frame.DataFrame
```

### Dataframe/Series

```
In [5]: # 1. row oriented approach

In [7]: df1=pd.DataFrame([[{"Afghanistan",1952,8425333,"Asia",28.801,779.445314},
                        [{"Afghanistan",1957,9240934,"Asia",30.332,820.853030}]],
                        columns=["country","year","population","continent","life_exp","gdp_cap"])

In [8]:
```

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030

```
In [9]:
```

	country	year	population	continent	life_exp	gdp_cap
0	...	...	...	...	...	...
1	...	...	...	...	...	...
2	...	...	...	...	...	...
3	...	...	...	...	...	...
4	...	...	...	...	...	...
5	...	...	...	...	...	...
6	...	...	...	...	...	...
7	...	...	...	...	...	...
8	...	...	...	...	...	...
9	...	...	...	...	...	...
10	...	...	...	...	...	...
11	...	...	...	...	...	...
12	...	...	...	...	...	...
13	...	...	...	...	...	...
14	...	...	...	...	...	...
15	...	...	...	...	...	...
16	...	...	...	...	...	...
17	...	...	...	...	...	...
18	...	...	...	...	...	...
19	...	...	...	...	...	...
20	...	...	...	...	...	...
21	...	...	...	...	...	...
22	...	...	...	...	...	...
23	...	...	...	...	...	...
24	...	...	...	...	...	...
25	...	...	...	...	...	...
26	...	...	...	...	...	...
27	...	...	...	...	...	...
28	...	...	...	...	...	...
29	...	...	...	...	...	...
30	...	...	...	...	...	...
31	...	...	...	...	...	...
32	...	...	...	...	...	...
33	...	...	...	...	...	...
34	...	...	...	...	...	...
35	...	...	...	...	...	...
36	...	...	...	...	...	...
37	...	...	...	...	...	...
38	...	...	...	...	...	...
39	...	...	...	...	...	...
40	...	...	...	...	...	...
41	...	...	...	...	...	...
42	...	...	...	...	...	...
43	...	...	...	...	...	...
44	...	...	...	...	...	...
45	...	...	...	...	...	...
46	...	...	...	...	...	...
47	...	...	...	...	...	...
48	...	...	...	...	...	...
49	...	...	...	...	...	...
50	...	...	...	...	...	...
51	...	...	...	...	...	...
52	...	...	...	...	...	...
53	...	...	...	...	...	...
54	...	...	...	...	...	...
55	...	...	...	...	...	...
56	...	...	...	...	...	...
57	...	...	...	...	...	...
58	...	...	...	...	...	...
59	...	...	...	...	...	...
60	...	...	...	...	...	...
61	...	...	...	...	...	...
62	...	...	...	...	...	...
63	...	...	...	...	...	...
64	...	...	...	...	...	...
65	...	...	...	...	...	...
66	...	...	...	...	...	...
67	...	...	...	...	...	...
68	...	...	...	...	...	...
69	...	...	...	...	...	...
70	...	...	...	...	...	...
71	...	...	...	...	...	...
72	...	...	...	...	...	...
73	...	...	...	...	...	...
74	...	...	...	...	...	...
75	...	...	...	...	...	...
76	...	...	...	...	...	...
77	...	...	...	...	...	...
78	...	...	...	...	...	...
79	...	...	...	...	...	...
80	...	...	...	...	...	...
81	...	...	...	...	...	...
82	...	...	...	...	...	...
83	...	...	...	...	...	...
84	...	...	...	...	...	...
85	...	...	...	...	...	...
86	...	...	...	...	...	...
87	...	...	...	...	...	...
88	...	...	...	...	...	...
89	...	...	...	...	...	...
90	...	...	...	...	...	...
91	...	...	...	...	...	...
92	...	...	...	...	...	...
93	...	...	...	...	...	...
94	...	...	...	...	...	...
95	...	...	...	...	...	...
96	...	...	...	...	...	...
97	...	...	...	...	...	...
98	...	...	...	...	...	...
99	...	...	...	...	...	...
100	...	...	...	...	...	...
101	...	...	...	...	...	...
102	...	...	...	...	...	...
103	...	...	...	...	...	...
104	...	...	...	...	...	...
105	...	...	...	...	...	...
106	...	...	...	...	...	...
107	...	...	...	...	...	...
108	...	...	...	...	...	...
109	...	...	...	...	...	...
110	...	...	...	...	...	...
111	...	...	...	...	...	...
112	...	...	...	...	...	...
113	...	...	...	...	...	...
114	...	...	...	...	...	...
115	...	...	...	...	...	...
116	...	...	...	...	...	...
117	...	...	...	...	...	...
118	...	...	...	...	...	...
119	...	...	...	...	...	...
120	...	...	...	...	...	...
121	...	...	...	...	...	...
122	...	...	...	...	...	...
123	...	...	...	...	...	...
124	...	...	...	...	...	...
125	...	...	...	...	...	...
126	...	...	...	...	...	...
127	...	...	...	...	...	...
128	...	...	...	...	...	...
129	...	...	...	...	...	...
130	...	...	...	...	...	...
131	...	...	...	...	...	...
132	...	...	...	...	...	...
133	...	...	...	...	...	...
134	...	...	...	...	...	...
135	...	...	...	...	...	...
136	...	...	...	...	...	...
137	...	...	...	...	...	...
138	...	...	...	...	...	...
139	...	...	...	...	...	...
140	...	...	...	...	...	...
141	...	...	...	...	...	...
142	...	...	...	...	...	...
143	...	...	...	...	...	...
144	...	...	...	...	...	...
145	...	...	...	...	...	...
146	...	...	...	...	...	...
147	...	...	...	...	...	...
148	...	...	...	...	...	...
149	...	...	...	...	...	...
150	...	...	...	...	...	...
151	...	...	...	...	...	...
152	...	...	...	...	...	...
153	...	...	...	...	...	...
154	...	...	...	...	...	...
155	...	...	...	...	...	...
156	...	...	...	...	...	...
157	...	...	...	...	...	...
158	...	...	...	...	...	...
159	...	...	...	...	...	...
160	...	...	...	...	...	...
161	...	...	...	...	...	...
162	...	...	...	...	...	...
163	...	...	...	...	...	...
164	...	...	...	...	...	...
165	...	...	...	...	...	...
166	...	...	...	...	...	...
167	...	...	...	...	...	...
168	...	...	...	...	...	...
169	...	...	...	...	...	...
170	...	...	...	...	...	...
171	...	...	...	...	...	...
172	...	...	...	...	...	...
173	...	...	...	...	...	...
174	...	...	...	...	...	...
175	...	...	...	...	...	...
176	...	...	...	...	...	...
177	...	...	...	...	...	...
178	...	...	...	...	...	...
179	...	...	...	...	...	...
180	...	...	...	...	...	...
181	...	...	...	...	...	...
182	...	...	...	...	...	...
183	...	...	...	...	...	...
184	...	...	...	...	...	...
185	...	...	...	...	...	...
186	...	...	...	...	...	...
187	...	...	...	...	...	...
188	...	...	...	...	...	...
189	...	...	...	...	...	...
190	...	...	...	...	...	...
191	...	...	...	...	...	...
192	...	...	...	...	...	...
193	...	...	...	...	...	...
194	...	...	...	...	...	...
195	...	...	...	...	...	...
196	...	...	...	...	...	...
197	...	...	...	...	...	...
198	...	...	...	...	...	...
199	...	...	...	...	...	...

```
In [10]:
```

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030

```
In [11]:
```

	country	year	population	continent	life_exp	gdp_cap
0	Afghanistan	1952	8425333	Asia	28.801	779.445314
1	Afghanistan	1957	9240934	Asia	30.332	820.853030
2	Afghanistan	1962	10267083	Asia	31.997	853.100710
3	Afghanistan	1967	11537966	Asia	34.020	836.197138
4	Afghanistan	1972	13079460	Asia	36.088	739.981106
...	...	...	...	...	...	...
1699	Zimbabwe	1967	9216418	Africa	62.351	706.157306
1700	Zimbabwe	1962	10704340	Africa	60.377	693.420786
1701	Zimbabwe	1957	11404948	Africa	46.809	792.449960
1702	Zimbabwe	2002	11920563	Africa	39.989	672.038623
1703	Zimbabwe	2007	12311143	Africa	43.487	469.709298

1704 rows x 6 columns

## Basic ops on a DataFrame

### df.info()

```
In [15]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1704 entries, 0 to 1703
Data columns (total 6 columns):
 #   Column      Non-Null Count  Dtype  ---
 0   country     1784 non-null    object
 1   year        1784 non-null    int64
 2   population  1784 non-null    float64
 3   continent   1784 non-null    object
 4   life_exp    1784 non-null    float64
 5   gdp_cap     1784 non-null    float64
 6   Europe      float64(2), int64(2), object(2)
memory usage: 88.6+ KB
```

### df.describe()

```
In [16]: df.describe()

Out[16]:
```

	country	year	population	life_exp	gdp_cap
count	1704.000000	1.704000e+03	1704.000000	1704.000000	1704.000000
mean	1979.500000	2.960121e+07	99.474439	72.15327081	7215.327081
std	17.265313	1.061170e+08	6.217107	9.957454643	9.957454643
min	1952.000000	6.001100e+04	23.999000	24.1168976	24.1168976
25%	1965.750000	2.79			