

Python

1. Read the csv or excel file.
2. Change the datatypes of the column and check the memory usage before and after the change in the data types.
3. Dump the data into the mysql database.
4. index the column after the data is inserted.

> Read the excel file.

```
In [1]: import pandas as pd
df= pd.read_excel("C:\\Users\\PC-chetan\\Desktop\\used_bikes.xlsx")
```

```
In [2]: df
```

```
Out[2]:
```

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000	Ahmedabad	17654	First Owner	3	110	TVS
1	Royal Enfield Classic 350cc	119900	Delhi	11000	First Owner	4	350	Royal Enfield
2	Triumph Daytona 675R	600000	Delhi	110	First Owner	8	675	Triumph
3	TVS Apache RTR 180cc	65000	Bangalore	16329	First Owner	4	180	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000	Bangalore	10000	First Owner	3	150	Yamaha
...
144	Royal Enfield Standard 350cc	115000	Hyderabad	14900	First Owner	3	350	Royal Enfield
145	Honda CB Shine 125cc Disc	65000	Mumbai	2000	First Owner	3	125	Honda
146	Honda CB Unicorn ABS 150cc	99000	Mumbai	4000	First Owner	2	150	Honda
147	Yamaha YZF-R15 2.0 150cc	68500	Delhi	68500	Second Owner	7	150	Yamaha
148	Yamaha YZF-R15 V3 150cc	140000	Delhi	2473	First Owner	2	150	Yamaha

149 rows × 8 columns

> Change the datatypes of the column and check the memory usage before and after the change in the data types.

```
In [87]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149 entries, 0 to 148
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype
---  -
0   bike_name       149 non-null    object
1   price           149 non-null    int64
2   city            149 non-null    object
3   kms_driven      149 non-null    int64
4   owner           149 non-null    object
5   age             149 non-null    int64
6   power           149 non-null    int64
7   brand           149 non-null    object
dtypes: int64(4), object(4)
memory usage: 9.4+ KB
```

```
In [88]: df_new = df.convert_dtypes()
```

```
In [89]: df_new
```

Out[89]:

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000	Ahmedabad	17654	First Owner	3	110	TVS
1	Royal Enfield Classic 350cc	119900	Delhi	11000	First Owner	4	350	Royal Enfield
2	Triumph Daytona 675R	600000	Delhi	110	First Owner	8	675	Triumph
3	TVS Apache RTR 180cc	65000	Bangalore	16329	First Owner	4	180	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000	Bangalore	10000	First Owner	3	150	Yamaha
...
144	Royal Enfield Standard 350cc	115000	Hyderabad	14900	First Owner	3	350	Royal Enfield
145	Honda CB Shine 125cc Disc	65000	Mumbai	2000	First Owner	3	125	Honda
146	Honda CB Unicorn ABS 150cc	99000	Mumbai	4000	First Owner	2	150	Honda
147	Yamaha YZF-R15 2.0 150cc	68500	Delhi	68500	Second Owner	7	150	Yamaha
148	Yamaha YZF-R15 V3 150cc	140000	Delhi	2473	First Owner	2	150	Yamaha

149 rows × 8 columns

```
In [90]: df_new.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 149 entries, 0 to 148
Data columns (total 8 columns):
#   Column          Non-Null Count  Dtype
---  -
0   bike_name       149 non-null    string
1   price           149 non-null    Int64
2   city            149 non-null    string
```

```

3   kms_driven    149 non-null    Int64
4   owner         149 non-null    string
5   age           149 non-null    Int64
6   power         149 non-null    Int64
7   brand         149 non-null    string
dtypes: Int64(4), string(4)
memory usage: 10.0 KB

```

```

In [91]: # memory usage before : 9.4+ KB
         # memory usage after : 10.0 KB

```

> Dump the data into the mysql database.

```

In [92]: df_new.reindex()

```

```

Out[92]:

```

	bike_name	price	city	kms_driven	owner	age	power	brand
0	TVS Star City Plus Dual Tone 110cc	35000	Ahmedabad	17654	First Owner	3	110	TVS
1	Royal Enfield Classic 350cc	119900	Delhi	11000	First Owner	4	350	Royal Enfield
2	Triumph Daytona 675R	600000	Delhi	110	First Owner	8	675	Triumph
3	TVS Apache RTR 180cc	65000	Bangalore	16329	First Owner	4	180	TVS
4	Yamaha FZ S V 2.0 150cc-Ltd. Edition	80000	Bangalore	10000	First Owner	3	150	Yamaha
...
144	Royal Enfield Standard 350cc	115000	Hyderabad	14900	First Owner	3	350	Royal Enfield
145	Honda CB Shine 125cc Disc	65000	Mumbai	2000	First Owner	3	125	Honda
146	Honda CB Unicorn ABS 150cc	99000	Mumbai	4000	First Owner	2	150	Honda
147	Yamaha YZF-R15 2.0 150cc	68500	Delhi	68500	Second Owner	7	150	Yamaha
148	Yamaha YZF-R15 V3 150cc	140000	Delhi	2473	First Owner	2	150	Yamaha

149 rows × 8 columns

```

In [93]: import mysql.connector
         import xlr

```

```

In [94]: mydb = mysql.connector.connect(
         host="localhost",
         user="root",
         password="Tar@un1234",
         database='mysql'
         )

```

```

In [95]: Loading [MathJax]/extensions/Safe.js ).cursor()# work as a pointer

```

```
In [ ]: my_cursor.execute("CREATE TABLE bikes_name (s_no int PRIMARY KEY, bike_name VARCHAR(50),  
mydb.close()
```

```
In [102... loc = ('C:\\Users\\PC-chetan\\Desktop\\used_bikes.xlsx')  
l= list()
```

```
In [ ]: a = xlrd.open_workbook(loc)  
sheet = a.sheet_by_index(0)  
sheet.cell_value(0,0)  
for i in range(1,150):  
    l.append(tuple(sheet.row_values(i)))
```

```
In [ ]: q = 'insert into bikes_name(bike_name,price, city, kms_drive, owner, age, power, brand)va'  
  
my_cursor.executemany(q,l)  
mydb.commit()  
mydb.close()
```