

# Let's import Library

```
In [ ]: import pandas as pand
```

```
In [188... rw = pand.read_excel(r'C:\Users\hp\Downloads\27th\TASK\taskshikha.xlsx') ##Let's re  
rw
```

```
Out[188... 
```

|   | ADDR   | DT         | NAME | time    |
|---|--|------------|------|---------|
| 0 | 45 rd, kenith street, btm, bangalore, 500038 | 2022-03-04 | abc  | 8:30:21 |
| 1 | 45 rd, street, hitech, Hd, 500038            | 2022-04-02 | dec  | 5:30:21 |

```
In [189... rw.shape ##dimention of Dataframe
```

```
Out[189... (2, 4)
```

```
In [190... rw.columns ##columns of raw data
```

```
Out[190... Index(['ADDR', 'DT', 'NAME', 'time'], dtype='object')
```

```
In [196... rw[['House add', 'street', 'city', 'state', 'zip']] = (rw['ADDR'].str.split(',', expand
```

```
In [197... rw ##calling of Dataframe to view the output
```

```
Out[197... 
```

|   | ADDR   | DT         | NAME | time    | House<br>add | street        | city   | state     | zip    |
|---|--|------------|------|---------|--------------|---------------|--------|-----------|--------|
| 0 | 45 rd, kenith street, btm, bangalore, 500038 | 2022-03-04 | abc  | 8:30:21 | 45 rd        | kenith street | btm    | bangalore | 500038 |
| 1 | 45 rd, street, hitech, Hd, 500038            | 2022-04-02 | dec  | 5:30:21 | 45 rd        | street        | hitech | Hd        | 500038 |

```
In [203... from datetime import datetime ## Let's import one more library to split date & time
```

```
In [200... rw['DT'] = pand.to_datetime(rw['DT']) ##introducing new funtion (to_datetime) becau
```

```
In [201... rw['year'] = rw['DT'].dt.year  
rw['month'] = rw['DT'].dt.month  
rw['day'] = rw['DT'].dt.day  
rw
```

Out[201...

|   | ADDR  | DT             | NAME | time    | House<br>add | street           | city   | state     | zip    | year | month |
|---|---|----------------|------|---------|--------------|------------------|--------|-----------|--------|------|-------|
| 0 | 45 rd,<br>kenith<br>street,<br>btm,<br>bangalore,<br>500038 | 2022-<br>03-04 | abc  | 8:30:21 | 45 rd        | kenith<br>street | btm    | bangalore | 500038 | 2022 |       |
| 1 | 45 rd,<br>street,<br>hitech,<br>Hd,<br>500038               | 2022-<br>04-02 | dec  | 5:30:21 | 45 rd        | street           | hitech | Hd        | 500038 | 2022 |       |

In [204...

```
rw['time'] = pand.to_datetime(rw['time'])## same as above
```

In [205...

```
rw['hour'] = rw['time'].dt.hour
rw['minute'] = rw['time'].dt.minute
rw['second'] = rw['time'].dt.second
rw
```

Out[205...

|   | ADDR  | DT             | NAME | time                       | House<br>add | street           | city   | state     | zip    | year | month |
|---|---|----------------|------|----------------------------|--------------|------------------|--------|-----------|--------|------|-------|
| 0 | 45 rd,<br>kenith<br>street,<br>btm,<br>bangalore,<br>500038 | 2022-<br>03-04 | abc  | 2024-<br>03-28<br>08:30:21 | 45 rd        | kenith<br>street | btm    | bangalore | 500038 | 2022 |       |
| 1 | 45 rd,<br>street,<br>hitech,<br>Hd,<br>500038               | 2022-<br>04-02 | dec  | 2024-<br>03-28<br>05:30:21 | 45 rd        | street           | hitech | Hd        | 500038 | 2022 |       |

In [206...

```
rw.columns ## getting columns
```

Out[206...

```
Index(['ADDR', 'DT', 'NAME', 'time', 'House add', 'street', 'city', 'state',  
      'zip', 'year', 'month', 'day', 'hour', 'minute', 'second'],  
      dtype='object')
```

In [207...

```
columns_to_remove = ['ADDR','DT','time'] ## In order to print desired output we need
```

In [208...

```
clean_data = rw.drop(columns=columns_to_remove) ## Using drop functions to clean mu
```

In [209...

```
print(clean_data) ## printing clean data.
```

|   | NAME | House | add    | street | city   | state  | zip       | year   | month | \ |
|---|------|-------|--------|--------|--------|--------|-----------|--------|-------|---|
| 0 | abc  | 45    | rd     | kenith | street | btm    | bangalore | 500038 | 2022  | 3 |
| 1 | dec  | 45    | rd     |        | street | hitech | Hd        | 500038 | 2022  | 4 |
|   | day  | hour  | minute | second |        |        |           |        |       |   |
| 0 | 4    | 8     | 30     | 21     |        |        |           |        |       |   |
| 1 | 2    | 5     | 30     | 21     |        |        |           |        |       |   |

In [212...

Out[212]:



Desired Output:-----As per requirement

|                    |           |        |      |       |       |       |      |      |      |    |     |     |  |
|--------------------|-----------|--------|------|-------|-------|-------|------|------|------|----|-----|-----|--|
| THIS IS CLEAN DATA |           |        |      |       |       |       |      |      |      |    |     |     |  |
|                    | House add | street | city | state | zip   | month | data | year | name | hr | min | sec |  |
|                    | 45 rd     | kenith | btm  | banga | 50038 | 3     | 4    | 2022 | abc  | 8  | 30  | 21  |  |