## Accessing Google Drive

```
from google.colab import drive
drive.mount('/content/drive')
```

```
Mounted at /content/drive
```

## Importing pandas library and getting the whole table by mentioning its path on drive.

```
import pandas as pd
h=pd.read_csv('/content/drive/MyDrive/test.csv')
h.head()
```

| | id | f0 | f1 | f2 | f3 | f4 | f5 | f6 | f7 |
|---|---|---|---|---|---|---|---|---|---|
| **0** | 600000 | 0.003229 | 4.838660 | 585.529 | 2.282910 | 0.713180 | 3.907830 | 0.480696 | 1.482270 |
| **1** | 600001 | 0.008602 | 0.505536 | -100.099 | 3.012670 | 0.027199 | 1.194610 | 5.036620 | 2.517440 |
| **2** | 600002 | 1.461000 | 2.437260 | -112.964 | 3.541230 | 0.752338 | 4.338310 | 1.648080 | 4.699910 |
| **3** | 600003 | 0.140556 | 3.085610 | 179.451 | 0.573945 | 0.057342 | 2.216790 | 1.623480 | 0.526174 |
| **4** | 600004 | 0.128876 | 5.199760 | 107.466 | -0.497149 | 0.080220 | 0.458121 | 0.629839 | 5.240460 |

5 rows × 101 columns

## Importing numpy and random libraries to get random values.

```
import numpy as np
import random

df·=·pd.read_csv(r'/content/drive/MyDrive/test.csv')

df['RN']·=·np.random.uniform(0,1,·df.shape[0])
print(df)
```

```
          id       f0       f1  ...      f98      f99       RN
0     600000  0.003229  4.838660  ...  0.131533  0.012047  0.755117
1     600001  0.008602  0.505536  ...  0.047418  0.120015  0.356781
2     600002  1.461000  2.437260  ...  0.108843  0.064687  0.105593
3     600003  0.140556  3.085610  ...  0.052524  0.011058  0.897167
4     600004  0.128876  5.199760  ...  0.145319 -0.050393  0.707918
...      ...       ...       ...  ...       ...       ...       ...
539995  1139995  0.431599  1.507560  ...  0.059733  0.029033  0.591295
539996  1139996  0.069713  2.355480  ...  0.016573  0.079498  0.608555
539997  1139997  0.385075  2.528890  ...  0.776967  0.123728  0.513979
539998  1139998  1.846240  3.415350  ...  0.054561  0.082622  0.685126
539999  1139999  0.475802  2.670740  ...  0.072163  0.102005  0.487208

[540000 rows x 102 columns]
```

## Applying Condition on RN(Random Number Column) to get the target value

```
df['target'] = [1 if x >= 0.5 else 0 for x in df['RN']]
print(df)
```

```
            id       f0       f1  ...      f99       RN  target
0       600000  0.003229  4.838660  ...  0.012047  0.755117       1
1       600001  0.008602  0.505536  ...  0.120015  0.356781       0
2       600002  1.461000  2.437260  ...  0.064687  0.105593       0
3       600003  0.140556  3.085610  ...  0.011058  0.897167       1
4       600004  0.128876  5.199760  ... -0.050393  0.707918       1
...        ...       ...       ...  ...       ...       ...     ...
539995  1139995  0.431599  1.507560  ...  0.029033  0.591295       1
539996  1139996  0.069713  2.355480  ...  0.079498  0.608555       1
539997  1139997  0.385075  2.528890  ...  0.123728  0.513979       1
539998  1139998  1.846240  3.415350  ...  0.082622  0.685126       1
539999  1139999  0.475802  2.670740  ...  0.102005  0.487208       0

[540000 rows x 103 columns]
```

## Returning a new list by copy function to get the desired column

```
ndf = df[['id','target']].copy()
print(ndf)
```

```
            id  target
0       600000       1
1       600001       0
2       600002       0
3       600003       1
4       600004       1
...        ...     ...
539995  1139995       1
539996  1139996       1
539997  1139997       1
539998  1139998       1
539999  1139999       0

[540000 rows x 2 columns]
```

## Finally exporting it to .csv file.

```
ndf.to_csv('sample.csv',index=False)
```

✓ 0s completed at 1:04 AM