#Assignment_2

```
import pandas as pd
import numpy as np
import plotly.express as px

df=pd.read_csv("/content/Churn_Modelling.csv")

df.isnull().any()
```

```
RowNumber          False
CustomerId         False
Surname            False
CreditScore        False
Geography          False
Gender             False
Age                False
Tenure             False
Balance            False
NumOfProducts      False
HasCrCard          False
IsActiveMember     False
EstimatedSalary    False
Exited             False
dtype: bool
```

```
df.isnull().sum()
```

```
RowNumber          0
CustomerId         0
Surname            0
CreditScore        0
Geography          0
Gender             0
Age                0
Tenure             0
Balance            0
NumOfProducts      0
HasCrCard          0
IsActiveMember     0
EstimatedSalary    0
Exited             0
dtype: int64
```

```
df.head(1)
```

```
   RowNumber  CustomerId   Surname  CreditScore Geography  Gender  Age
\
0          1    15634602  Hargrave          619    France  Female   42
```

```
   Tenure  Balance  NumOfProducts  HasCrCard  IsActiveMember
EstimatedSalary  \
0       2      0.0              1          1               1
101348.88

   Exited
0       1
```

```
df['RowNumber']
```

```
0           1
1           2
2           3
3           4
4           5
        ...
9995     9996
9996     9997
9997     9998
9998     9999
9999    10000
Name: RowNumber, Length: 10000, dtype: int64
```

```
df.shape
```

```
(10000, 14)
```

```
df[df.columns[1:3]]
```

```
      CustomerId    Surname
0       15634602   Hargrave
1       15647311       Hill
2       15619304       Onio
3       15701354       Boni
4       15737888   Mitchell
...          ...        ...
9995    15606229   Obijiaku
9996    15569892  Johnstone
9997    15584532        Liu
9998    15682355  Sabbatini
9999    15628319     Walker

[10000 rows x 2 columns]
```

```
df[df.columns[1:7]]
```

```
      CustomerId    Surname  CreditScore Geography  Gender  Age
0       15634602   Hargrave          619    France  Female   42
1       15647311       Hill          608     Spain  Female   41
```

```
2       15619304       Onio          502    France  Female   42
3       15701354       Boni          699    France  Female   39
4       15737888   Mitchell          850     Spain  Female   43
...          ...        ...          ...       ...     ...  ...
9995    15606229   Obijiaku          771    France    Male   39
9996    15569892  Johnstone          516    France    Male   35
9997    15584532        Liu          709    France  Female   36
9998    15682355  Sabbatini          772   Germany    Male   42
9999    15628319     Walker          792    France  Female   28

[10000 rows x 6 columns]
```

```
#we need not to perfrom descriptive statistic among the dataset
because we dont have any null values in the dataset
#we dont have any null values to perform handling missing values
```

```
px.histogram
```

```
<function plotly.express._chart_types.histogram(data_frame=None,
x=None, y=None, color=None, pattern_shape=None, facet_row=None,
facet_col=None, facet_col_wrap=0, facet_row_spacing=None,
facet_col_spacing=None, hover_name=None, hover_data=None,
animation_frame=None, animation_group=None, category_orders=None,
labels=None, color_discrete_sequence=None, color_discrete_map=None,
pattern_shape_sequence=None, pattern_shape_map=None, marginal=None,
opacity=None, orientation=None, barmode='relative', barnorm=None,
histnorm=None, log_x=False, log_y=False, range_x=None, range_y=None,
histfunc=None, cumulative=None, nbins=None, text_auto=False,
title=None, template=None, width=None, height=None)>
```

```
X = df.iloc[:, :-1].values
print(X)
```

```
[[1 15634602 'Hargrave' ... 1 1 101348.88]
 [2 15647311 'Hill' ... 0 1 112542.58]
 [3 15619304 'Onio' ... 1 0 113931.57]
 ...
 [9998 15584532 'Liu' ... 0 1 42085.58]
 [9999 15682355 'Sabbatini' ... 1 0 92888.52]
 [10000 15628319 'Walker' ... 1 0 38190.78]]
```

```
Y = df.iloc[:, -1].values
print(Y)
```

```
[1 0 1 ... 1 1 0]
```