```
from google.colab import files
uploaded = files.upload()
```

Choose Files Social_Network_Ads.csv

- **Social_Network_Ads.csv**(text/csv) - 4903 bytes, last modified: 5/29/2023 - 100% done

```
Saving Social_Network_Ads.csv to Social_Network_Ads.csv
```

```
import numpy as np
import pandas as pd
```

```
df = pd.read_csv('Social_Network_Ads.csv')
```

```
df
```

	Age	EstimatedSalary	Purchased
0	19	19000	0
1	35	20000	0
2	26	43000	0
3	27	57000	0
4	19	76000	0
...	...	...	...
395	46	41000	1
396	51	23000	1
397	50	20000	1
398	36	33000	0
399	49	36000	1

400 rows × 3 columns

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

```
Age               0
EstimatedSalary   0
Purchased         0
dtype: int64
```

```
x = df.iloc[:,0:2]
```

```
x
```

	Age	EstimatedSalary
0	19	19000
1	35	20000
2	26	43000
3	27	57000
4	19	76000
...	...	...
395	46	41000
396	51	23000
397	50	20000
398	36	33000
399	49	36000

400 rows × 2 columns

```
y = df.iloc[:,-1]
```

```
y
```

```
0      0
1      0
```

```
2      0
3      0
4      0
      ..
395    1
396    1
397    1
398    0
399    1
Name: Purchased, Length: 400, dtype: int64
```

```
from sklearn.model_selection import train_test_split
```

```
x_train,x_test,y_train,y_test = train_test_split(x,y,test_size=0.2)
```

```
x_train
```

	Age	EstimatedSalary
57	28	79000
366	58	47000
345	41	63000
174	34	72000
147	41	30000
...	...	...
361	53	34000
395	46	41000
338	38	55000
384	57	33000
370	60	46000

320 rows × 2 columns

```
x_test
```

	Age	EstimatedSalary
10	26	80000
311	39	96000
108	26	86000
118	40	59000
82	20	49000
...	...	...
170	21	88000
16	47	25000
303	37	79000
51	18	44000
186	20	82000

80 rows × 2 columns

```
y_train
```

```
57     0
366    1
345    0
174    0
147    0
      ..
361    1
395    1
338    0
384    1
370    1
Name: Purchased, Length: 320, dtype: int64
```

```
y_test
```

```
10     0
311    1
108    0
118    0
82     0
      ..
170    0
16     1
303    1
51     0
186    0
Name: Purchased, Length: 80, dtype: int64
```

```
from sklearn.linear_model import LogisticRegression
lr = LogisticRegression()
```

```
model = lr.fit(x_train,y_train)
```

```
y_pred = model.predict(x_test)
```

```
y_test
```

```
10     0
311    1
108    0
118    0
82     0
      ..
170    0
16     1
303    1
51     0
186    0
Name: Purchased, Length: 80, dtype: int64
```

```
y_pred
```

```
array([0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0])
```

```
from sklearn.metrics import mean_squared_error
mse = mean_squared_error(y_test,y_pred)
```

```
mse
```

```
0.35
```

```
from sklearn.metrics import accuracy_score
accuracy = accuracy_score(y_test,y_pred)
```

```
accuracy
```

```
0.65
```

```
from sklearn.metrics import confusion_matrix
cm = confusion_matrix(y_test,y_pred)
```

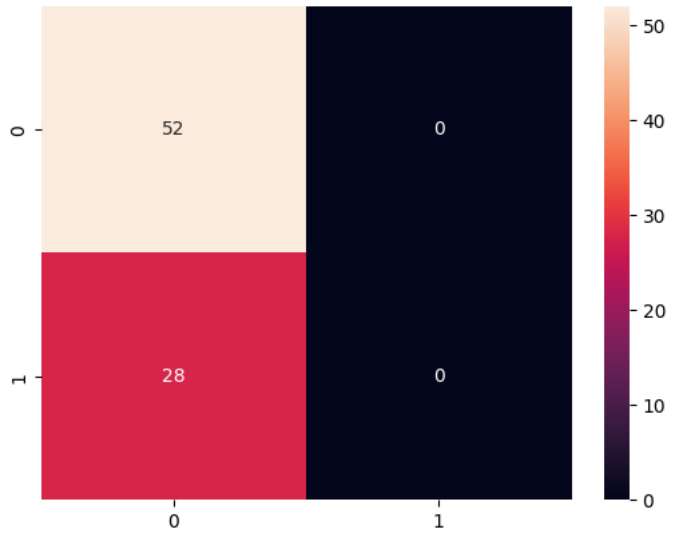
```
cm
```

```
array([[52,  0],
       [28,  0]])
```

```
import seaborn as sns
```

```
sns.heatmap(cm,annot=True)
```

<Axes: >



✓ 0s completed at 1:55 PM