

BikeBuyersSažimanje

August 1, 2025

```
[1]: import pandas as pd
```

```
[2]: bikes = pd.read_csv("../..../datasets/bike_buyers_clean.csv")
```

```
[3]: bikes.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 13 columns):
 #   Column                Non-Null Count  Dtype  
---  -
 0   ID                    1000 non-null   int64  
 1   Marital Status        1000 non-null   object  
 2   Gender                1000 non-null   object  
 3   Income                1000 non-null   int64  
 4   Children              1000 non-null   int64  
 5   Education             1000 non-null   object  
 6   Occupation            1000 non-null   object  
 7   Home Owner            1000 non-null   object  
 8   Cars                  1000 non-null   int64  
 9   Commute Distance      1000 non-null   object  
10   Region                1000 non-null   object  
11   Age                   1000 non-null   int64  
12   Purchased Bike        1000 non-null   object  
dtypes: int64(5), object(8)
memory usage: 101.7+ KB
```

```
[4]: bikes.head()
```

```
[4]:
```

	ID	Marital Status	Gender	Income	Children	Education \
0	12496	Married	Female	40000	1	Bachelors
1	24107	Married	Male	30000	3	Partial College
2	14177	Married	Male	80000	5	Partial College
3	24381	Single	Male	70000	0	Bachelors
4	25597	Single	Male	30000	0	Bachelors

	Occupation	Home Owner	Cars	Commute Distance	Region	Age \
0	Skilled Manual	Yes	0	0-1 Miles	Europe	42

1	Clerical	Yes	1	0-1 Miles	Europe	43
2	Professional	No	2	2-5 Miles	Europe	60
3	Professional	Yes	1	5-10 Miles	Pacific	41
4	Clerical	No	0	0-1 Miles	Europe	36

Purchased Bike	
0	No
1	No
2	No
3	Yes
4	Yes

```
[5]: icols = bikes.select_dtypes('integer').columns
```

```
[6]: icols
```

```
[6]: Index(['ID', 'Income', 'Children', 'Cars', 'Age'], dtype='object')
```

```
[7]: bikes[icols] = bikes[icols].apply(pd.to_numeric, downcast='integer')
```

```
[8]: bikes.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ID                    1000 non-null   int16
1   Marital Status        1000 non-null   object
2   Gender                1000 non-null   object
3   Income                1000 non-null   int32
4   Children              1000 non-null   int8
5   Education             1000 non-null   object
6   Occupation            1000 non-null   object
7   Home Owner            1000 non-null   object
8   Cars                  1000 non-null   int8
9   Commute Distance      1000 non-null   object
10  Region                1000 non-null   object
11  Age                   1000 non-null   int8
12  Purchased Bike        1000 non-null   object
dtypes: int16(1), int32(1), int8(3), object(8)
memory usage: 71.4+ KB
```

```
[9]: bikes.Education.value_counts()
```

```
[9]: Education
Bachelors          306
Partial College     265
```

```

High School          179
Graduate Degree      174
Partial High School   76
Name: count, dtype: int64

```

```
[10]: bikes["Education"] = bikes["Education"].astype("category")
```

```
[11]: bikes.info()
```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1000 entries, 0 to 999
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   ID                    1000 non-null  int16
1   Marital Status        1000 non-null  object
2   Gender                1000 non-null  object
3   Income                1000 non-null  int32
4   Children              1000 non-null  int8
5   Education              1000 non-null  category
6   Occupation            1000 non-null  object
7   Home Owner            1000 non-null  object
8   Cars                  1000 non-null  int8
9   Commute Distance      1000 non-null  object
10  Region                1000 non-null  object
11  Age                   1000 non-null  int8
12  Purchased Bike        1000 non-null  object
dtypes: category(1), int16(1), int32(1), int8(3), object(7)
memory usage: 64.8+ KB

```

```
[12]: bikes.head()
```

```

[12]:      ID Marital Status  Gender  Income  Children  Education \
0  12496      Married  Female   40000         1    Bachelors
1  24107      Married   Male   30000         3  Partial College
2  14177      Married   Male   80000         5  Partial College
3  24381       Single   Male   70000         0    Bachelors
4  25597       Single   Male   30000         0    Bachelors

      Occupation Home Owner  Cars  Commute Distance  Region  Age \
0  Skilled Manual      Yes    0      0-1 Miles  Europe  42
1      Clerical      Yes    1      0-1 Miles  Europe  43
2   Professional      No    2      2-5 Miles  Europe  60
3   Professional      Yes    1      5-10 Miles  Pacific  41
4      Clerical      No    0      0-1 Miles  Europe  36

      Purchased Bike
0                No

```

1	No
2	No
3	Yes
4	Yes