Module 5: Pandas

Assignment - 1 Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python for Data Science Certification Course Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to analyze the customer_churn dataset with various functions. Tasks To Be Performed:

1. Start off by importing the customer_churn.csv file in the jupyter notebook and store that in churn DataFrame.
2. From the churn DataFrame, select only 3rd, 7th, 9th, and 20th columns and all the rows and store that in a new DataFrame named newCols.
3. From the original DataFrame, select only the rows from the 200th index till the 1000th index(inclusive) column.
4. Now select the rows from 20th index till 200th index(exclusive),and columns from 2nd index till 15th index value.
5. Display the top 100 records from the original DataFrame.
6. Display the last 10 records from the DataFrame.
7. Display the last record from the DataFrame.
8. Now from the churn DataFrame, try to sort the data by the tenure column according to the descending order.
9. Fetch all the records that are satisfying the following condition: a. Tenure>50 and the gender as 'Female' b. Gender as 'Male' and SeniorCitizen as 0 c. TechSupport as 'Yes' and Churn as 'No' d. Contract type as 'Month-to-month' and Churn as 'Yes'
10. Use a for loop to calculate the number of customers that are getting the tech support and are male senior citizens

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

In [3]:
```
##1,
churn=pd.read_csv('C:/Users/Vikas/Desktop/customer_churn-1.csv')
churn
```

Out[3]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

7043 rows × 21 columns

In [4]:
```
##2
newCols = churn.iloc[:, [2, 6, 8, 19]]
newCols
```

Out[4]:

	SeniorCitizen	PhoneService	InternetService	TotalCharges
0	0	No	DSL	29.85
1	0	Yes	DSL	1889.5
2	0	Yes	DSL	108.15
3	0	No	DSL	1840.75
4	0	Yes	Fiber optic	151.65
...	...	...	...	...
7038	0	Yes	DSL	1990.5
7039	0	Yes	Fiber optic	7362.9
7040	0	No	DSL	346.45
7041	1	Yes	Fiber optic	306.6
7042	0	Yes	Fiber optic	6844.5

7043 rows × 4 columns

In [5]:
```
##3,
cols=churn.iloc[100:1001,]
cols
```

Out[5]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
100	6380-ARCEH	Male	0	No	No	1	Yes	No	No	No internet service	...	
101	3679-XASPY	Female	0	Yes	Yes	1	Yes	No	No	No internet service	...	
102	7123-WQUHX	Male	0	No	No	38	Yes	Yes	Fiber optic	No	...	
103	5386-THSLQ	Female	1	Yes	No	66	No	No phone service	DSL	No	...	
104	3192-NQECA	Male	0	Yes	No	68	Yes	Yes	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	...
996	6641-XRPSU	Female	0	No	No	34	Yes	No	Fiber optic	No	...	
997	1374-DMZUI	Female	1	No	No	4	Yes	Yes	Fiber optic	No	...	
998	2545-LXYVJ	Male	0	Yes	No	72	Yes	No	No	No internet service	...	
999	3234-VKACU	Male	0	No	No	2	Yes	No	DSL	No	...	
1000	8357-EQXFO	Female	0	No	No	7	Yes	No	Fiber optic	No	...	

901 rows × 21 columns

In [6]:
```
##4,

index_rows=churn.iloc[20:200,2:15]
index_rows
```

Out[6]:

	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection
20	1	No	No	1	No	No phone service	DSL	No	No	Yes
21	0	Yes	No	12	Yes	No	No	No internet service	No internet service	No internet service
22	0	No	No	1	Yes	No	No	No internet service	No internet service	No internet service
23	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No
24	0	Yes	Yes	49	Yes	No	DSL	Yes	Yes	No
...	...	...	...	...	...	...	...	...	...	...
195	0	Yes	No	20	Yes	No	Fiber optic	Yes	Yes	No
196	0	Yes	Yes	24	Yes	Yes	No	No internet service	No internet service	No internet service
197	0	No	No	59	Yes	Yes	Fiber optic	No	Yes	Yes
198	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	Yes
199	0	No	Yes	1	Yes	No	No	No internet service	No internet service	No internet service

180 rows × 13 columns

Module 5: Pandas

Assignment - 2 Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python for Data Science Certification Course Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to analyze the customer_churn dataset with various functions. Tasks To Be Performed:

1. Now select the rows from 20th index till 200th index (exclusive) and columns from 2nd index till 15th index value.

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

In [7]:
```
##1,
churn=pd.read_csv('C:/Users/Vikas/Desktop/customer_churn-1.csv')
churn
```

Out[7]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	...
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

7043 rows × 21 columns

In [10]:
```
##1,
extract=churn.iloc[20:200,2:16]
extract
```

Out[10]:

	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	OnlineBackup	DeviceProtection
20	1	No	No	1	No	No phone service	DSL	No	No	Yes
21	0	Yes	No	12	Yes	No	No	No internet service	No internet service	No internet service
22	0	No	No	1	Yes	No	No	No internet service	No internet service	No internet service
23	0	Yes	No	58	Yes	Yes	DSL	No	Yes	No
24	0	Yes	Yes	49	Yes	No	DSL	Yes	Yes	No
...	...	...	...	...	...	...	...	...	...	...
195	0	Yes	No	20	Yes	No	Fiber optic	Yes	Yes	No
196	0	Yes	Yes	24	Yes	Yes	No	No internet service	No internet service	No internet service
197	0	No	No	59	Yes	Yes	Fiber optic	No	Yes	Yes
198	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	Yes	Yes
199	0	No	Yes	1	Yes	No	No	No internet service	No internet service	No internet service

180 rows × 14 columns

Assignment - 3

 Intel iPaat Python for Data Science Certification Course Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to analyze the customer_churn dataset with various functions. Tasks To Be Performed:

1. Display the top 100 records from the original data frame.
2. Display the last 10 records from the data frame.
3. Display the last record from the data frame.

In [17]:
```
churn=pd.read_csv('C:/Users/Vikas/Desktop/customer_churn-1.csv')
churn
```

Out[17]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	...
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

7043 rows × 21 columns

In [18]:
```
top_record=churn.head(101)
top_record
```

Out[18]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Device
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...	
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...	
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	
96	9803-FTJCG	Male	0	Yes	Yes	71	Yes	Yes	DSL	Yes	...	
97	0278-YXOOG	Male	0	No	No	5	Yes	No	No	No internet service	...	
98	3212-KXOCR	Male	0	No	No	52	Yes	No	No	No internet service	...	
99	4598-XLKNJ	Female	1	Yes	No	25	Yes	No	Fiber optic	No	...	
100	6380-ARCEH	Male	0	No	No	1	Yes	No	No	No internet service	...	

101 rows × 21 columns

In [14]:
```
##2,
last_record=churn.tail(10)
last_record
```

Out[14]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
7033	9767-FFLEM	Male	0	No	No	38	Yes	No	Fiber optic	No	...	
7034	0639-TSIQW	Female	0	No	No	67	Yes	Yes	Fiber optic	Yes	...	
7035	8456-QDAVC	Male	0	No	No	19	Yes	No	Fiber optic	No	...	
7036	7750-EYXWZ	Female	0	No	No	12	No	No phone service	DSL	No	...	
7037	2569-WGERO	Female	0	No	No	72	Yes	No	No	No internet service	...	
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

10 rows × 21 columns

In [15]:
```
##3,
last_record=churn.tail()
last_record
```

Out[15]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes	...	
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No	...	
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	Yes	...	

5 rows × 21 columns

Module 5: Pandas

Assignment - 4 Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python for Data Science Certification Course Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to analyze the customer_churn dataset with various functions. Tasks To Be Performed:

1. Now from the churn DataFrame, try to sort the data by the tenure column according to the descending order.
2. Fetch all the records that are satisfying the following condition: a. Tenure>50 and the gender as 'Female' b. Gender as 'Male' and SeniorCitizen as 0 c. TechSupport as 'Yes' and Churn as 'No' d. Contract type as 'Month-to-month,churn as yes,

In [19]:
```
##1,
churn.sort_values(['tenure'],ascending=False)
```

Out[19]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
1672	4737-AQCPU	Male	0	Yes	Yes	72	Yes	Yes	DSL	Yes	...	
193	9680-NIAUV	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	Yes	...	
4553	5914-XRFQB	Male	0	Yes	No	72	Yes	Yes	Fiber optic	Yes	...	
483	5168-MQQCA	Female	0	Yes	No	72	Yes	Yes	Fiber optic	No	...	
3266	0464-WJTKO	Female	0	Yes	Yes	72	Yes	No	No	No internet service	...	
...	...	...	...	...	...	...	...	...	...	...	...	
1082	4367-NUYAO	Male	0	Yes	Yes	0	Yes	Yes	No	No internet service	...	
3826	3213-VVOLG	Male	0	Yes	Yes	0	Yes	Yes	No	No internet service	...	
936	5709-LVOEQ	Female	0	Yes	Yes	0	Yes	No	DSL	Yes	...	
6754	2775-SEFEE	Male	0	No	Yes	0	Yes	Yes	DSL	Yes	...	
1340	1371-DWPAZ	Female	0	Yes	Yes	0	No	No phone service	DSL	Yes	...	

7043 rows × 21 columns

In [21]:
```
##2,
##Fetch all the records that are satisfying the following condition:
##a. Tenure>50 and the gender as 'Female'
##b. Gender as 'Male' and SeniorCitizen as 0
##c. TechSupport as 'Yes' and Churn as 'No' d. Contract type as 'Month-to-month and churn as 'yes'

##a. Tenure>50 and the gender as 'Female'
churn[(churn['tenure']>50) &(churn['gender']=='Female')]

##b. Gender as 'Male' and SeniorCitizen as 0
churn[(churn['gender']=='Male') &(churn['SeniorCitizen']== 0)]


##c. TechSupport as 'Yes' and Churn as 'No'
churn[(churn['TechSupport']=='Yes') &(churn['Churn']== 'No')]

##d. Contract type as 'Month-to-month and churn as 'yes'
churn[(churn['Contract']=='Month-to-month') &(churn['Churn']== 'Yes')]
```

Out[21]:

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...	Devic
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...	
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...	
5	9305-CDSKC	Female	0	No	No	8	Yes	Yes	Fiber optic	No	...	
8	7892-POOKP	Female	0	Yes	No	28	Yes	Yes	Fiber optic	No	...	
13	0280-XJGEX	Male	0	No	No	49	Yes	Yes	Fiber optic	No	...	
...	...	...	...	...	...	...	...	...	...	...	...	
7018	1122-JWTJW	Male	0	Yes	Yes	1	Yes	No	Fiber optic	No	...	
7026	8775-CEBBJ	Female	0	No	No	9	Yes	No	DSL	No	...	
7032	6894-LFHLY	Male	1	No	No	1	Yes	Yes	Fiber optic	No	...	
7034	0639-TSIQW	Female	0	No	No	67	Yes	Yes	Fiber optic	Yes	...	
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No	...	

1655 rows × 21 columns

Module 5: Pandas

Assignment - 5 Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python for Data Science Certification Course Problem Statement: You work in XYZ Corporation as a Data Analyst. Your corporation has told you to analyze the customer_churn dataset with various functions. Tasks To Be Performed:

1. Use a for loop to calculate the number of customers that are getting the tech support and are male senior citizens.

In [23]:
```
##1,
count = 0
for i in range(len(churn)):
    if churn.loc[i, 'gender'] == 'Male' and churn.loc[i, 'SeniorCitizen'] >= 60 and churn.loc[i, 'TechSupport']
        count += 1
print(count)
```

```
0
```

Module 5: Pandas Case Study

Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python Certification Course Problem Statement: You work in XYZ Company as a Python developer. The company officials want you to build a Python program. Tasks To Be Performed:

1. Write a function that takes start and end of a range returns a pandas series object containing numbers within that range. In case the user does not pass start or end or both they should default to 1 and 10 respectively. E.g: -> range_series() -> Should Return a pandas series from 1 to 10 range_series(5) -> Should Return a pandas series from 5 to 10 range_series(5, 10) -> Should Return a pandas series from 5 to 15 Create a method that takes n NumPy arrays of the same dimensions, sums them and returns the answer.
2. Create a function that takes in two lists named keys and values as arguments Keys would be strings and contain n string values Values would be a list containing n lists The methods should return a new pandas DataFrame with keys as column names and values as their corresponding values, e.g: ->create_dataframe(["One", "Two"], [["X", "Y"], ["A", "B"]]) -> should return a data frame One Two 0 X A 1 Y B
3. Create a function that concatenates two DataFrames. Use a previously created function to create two DataFrames and pass them as parameters Make sure that the indexes are reset before returning. Contact us: support@intellipaat.com / © Copyright Intellipaat / All rights reserved Intel iPaat Python Certification Course
4. Write code to load data from cars.csv into a dataframe and print its details. Details like: 'count', 'mean', 'std', 'min', '25%', '50%', '75%', 'max'.
5. Write a method that will take a column name as argument and return the name of the column with which the given column has the highest correlation. The data to be used is the cars dataset. The returned value should not the column named that was passed as the parameters, e.g. : get_max_correlated_column('mpg') -> should return 'drat

In [30]:
```
import numpy as np

import pandas as pd
```

In [25]:
```
##1,
squares = [i**2 for i in range(1,11)]
print(squares)
```

```
[1, 4, 9, 16, 25, 36, 49, 64, 81, 100]
```

In [27]:
```
##2,
def is_leap(year):
    if year % 4 == 0:
        if year % 100 == 0:
            if year % 400 == 0:
                return True
            else:
                return False
        else:
            return True
    else:
        return False
```

In [28]:
```
##3,
my_array = np.zeros(3)
my_variable = 0.

def my_function():
    my_array = np.array([1.,2.,3.])
    my_variable = 99.
    return my_array,my_variable

my_function()
```

Out[28]:
```
(array([1., 2., 3.]), 99.0)
```

In [37]:
```
##4
cars=pd.read_csv('C:/Users/Vikas/Desktop/cars-1.csv')
```

In [39]:
```
cars.head(10)
```

Out[39]:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
5	Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
6	Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
7	Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
8	Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
9	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4

In [40]:
```
cars.describe()
```

Out[40]:

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
count	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.000000	32.0000
mean	20.090625	6.187500	230.721875	146.687500	3.596563	3.217250	17.848750	0.437500	0.406250	3.687500	2.8125
std	6.026948	1.785922	123.938694	68.562868	0.534679	0.978457	1.786943	0.504016	0.498991	0.737804	1.6152
min	10.400000	4.000000	71.100000	52.000000	2.760000	1.513000	14.500000	0.000000	0.000000	3.000000	1.0000
25%	15.425000	4.000000	120.825000	96.500000	3.080000	2.581250	16.892500	0.000000	0.000000	3.000000	2.0000
50%	19.200000	6.000000	196.300000	123.000000	3.695000	3.325000	17.710000	0.000000	0.000000	4.000000	2.0000
75%	22.800000	8.000000	326.000000	180.000000	3.920000	3.610000	18.900000	1.000000	1.000000	4.000000	4.0000
max	33.900000	8.000000	472.000000	335.000000	4.930000	5.424000	22.900000	1.000000	1.000000	5.000000	8.0000

In [31]:
```
##5,
cars=pd.read_csv('C:/Users/Vikas/Desktop/cars-1.csv')
```

In [33]:
```
cars.head(10)
```

Out[33]:

	model	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
0	Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
1	Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
2	Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
3	Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
4	Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
5	Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
6	Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
7	Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
8	Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
9	Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4

In [36]:

```
def get_max_correlated_column('mpg'):
    cars = pd.read_csv('C:/Users/Vikas/Desktop/cars-1.csv')
    corr_matrix = cars.corr(method='mpg')
    corr_values = corr_matrix['mpg'].sort_values(ascending=False)
    return corr_values.index[5]
```

```
  Cell In[36], line 1
    def get_max_correlated_column('mpg'):
                                  ^

SyntaxError: invalid syntax
```

In [ ]: