Alcohol Consumption Insight

2015-2024

[ ]:
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
data=pd.read_csv('alcohol.csv',delimiter=",")
data
```

[39]:
```
data.head(5)
```

[39]:

	STATISTIC	Statistic Label	TLIST(A1)	\
0	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
1	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
2	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
3	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
4	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	

	Year	C02199V02655	Sex	C02076V03371	Age Group	UNIT	VALUE
0	2015	-	Both sexes	-	All ages	%	76.8
1	2015	-	Both sexes	300	15 - 19 years	%	52.2
2	2015	-	Both sexes	365	20 - 24 years	%	90.8
3	2015	-	Both sexes	410	25 - 29 years	%	84.8
4	2015	-	Both sexes	440	30 - 34 years	%	84.7

[40]:
```
data.shape
```

[40]:
```
(336, 10)
```

[41]:
```
data.isnull().sum()
```

[41]:
```
STATISTIC          0
Statistic Label    0
TLIST(A1)          0
Year               0
C02199V02655       0
Sex                0
C02076V03371       0
Age Group          0
UNIT               0
VALUE              0
```

```
dtype: int64
```

[ ]:

[42]:
```
data.notnull().sum()
```

[42]:
```
STATISTIC          336
Statistic Label    336
TLIST(A1)          336
Year               336
C02199V02655       336
Sex                336
C02076V03371       336
Age Group          336
UNIT               336
VALUE              336
dtype: int64
```

[43]:
```
data['Sex'].unique()
```

[43]:
```
array(['Both sexes', 'Male', 'Female'], dtype=object)
```

[44]:
```
data['Year'].unique()
```

[44]:
```
array([2015, 2016, 2017, 2018, 2021, 2022, 2023, 2024], dtype=int64)
```

[45]:
```
filtered_data = data[data['Sex'] == 'Male']
filtered_data.head(5)
```

[45]:

	STATISTIC	Statistic Label	TLIST(A1)	\
14	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
15	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
16	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
17	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
18	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	

	Year	C02199V02655	Sex	C02076V03371	Age Group	UNIT	VALUE
14	2015	1	Male	-	All ages	%	79.7
15	2015	1	Male	300	15 - 19 years	%	56.5
16	2015	1	Male	365	20 - 24 years	%	90.9
17	2015	1	Male	410	25 - 29 years	%	87.8
18	2015	1	Male	440	30 - 34 years	%	88.5

[46]:
```
filtered_data = data[data['VALUE'] > 50]
filtered_data.head(5)
```

[46]:

	STATISTIC	Statistic Label	TLIST(A1)	\
0	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	

```
1  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
2  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
3  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
4  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
```

	Year	C02199V02655	Sex	C02076V03371	Age Group	UNIT	VALUE
0	2015	-	Both sexes	-	All ages	%	76.8
1	2015	-	Both sexes	300	15 - 19 years	%	52.2
2	2015	-	Both sexes	365	20 - 24 years	%	90.8
3	2015	-	Both sexes	410	25 - 29 years	%	84.8
4	2015	-	Both sexes	440	30 - 34 years	%	84.7

[47]:
```
filtered_data = data.loc[data['Age Group'] == '15 - 19 years', 'VALUE']
filtered_data
```

[47]:
```
1      52.2
15     56.5
29     48.0
43     51.7
57     56.0
71     46.8
85     54.1
99     54.3
113    53.9
127    45.7
141    50.0
155    41.1
169    59.2
183    67.3
197    50.7
211    58.3
225    55.4
239    61.7
253    67.1
267    71.6
281    62.3
295    63.5
309    62.6
323    64.6
Name: VALUE, dtype: float64
```

[48]:
```
filtered_data = data.loc[data['Age Group'] == '15 - 19 years', 'VALUE'].sum()
filtered_data
```

[48]:
```
1354.6
```

[49]:
```
filtered_data = data[data['Age Group'] >= '70 years']
filtered_data.head(5)
```

[49]:
```
   STATISTIC                                  Statistic Label  TLIST(A1)  \
0   HIS45C01  Persons who consumed alcohol in the last twelv…       2015
13  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
14  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
27  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
28  HIS45C01  Persons who consumed alcohol in the last twelv…       2015

    Year C02199V02655         Sex C02076V03371          Age Group UNIT  VALUE
0   2015            -  Both sexes            -           All ages    %   76.8
13  2015            -  Both sexes          605  75 years and over    %   53.1
14  2015            1        Male            -           All ages    %   79.7
27  2015            1        Male          605  75 years and over    %   61.7
28  2015            2      Female            -           All ages    %   74.0
```

[50]:
```
data['VALUE'].max()
```

[50]:
```
92.7
```

[51]:
```
data['VALUE'].min()
```

[51]:
```
33.3
```

[52]:
```
filtered_data= data.loc[data['VALUE'].idxmax()]
filtered_data
```

[52]:
```
STATISTIC                                                  HIS45C01
Statistic Label    Persons who consumed alcohol in the last twelv…
TLIST(A1)                                                     2018
Year                                                          2018
C02199V02655                                                     1
Sex                                                           Male
C02076V03371                                                  365
Age Group                                             20 - 24 years
UNIT                                                             %
VALUE                                                         92.7
Name: 142, dtype: object
```

[53]:
```
filtered_data= data.loc[data['VALUE'].idxmin()]
filtered_data
```

[53]:
```
STATISTIC                                                  HIS45C01
Statistic Label    Persons who consumed alcohol in the last twelv…
TLIST(A1)                                                     2022
Year                                                          2022
C02199V02655                                                     2
```

```
Sex                                                          Female
C02076V03371                                                    605
Age Group                                          75 years and over
UNIT                                                              %
VALUE                                                          33.3
Name: 251, dtype: object
```

[54]:
```
filtered_data.duplicated()
```

[54]:
```
STATISTIC          False
Statistic Label    False
TLIST(A1)          False
Year                True
C02199V02655       False
Sex                False
C02076V03371       False
Age Group          False
UNIT               False
VALUE              False
Name: 251, dtype: bool
```

[55]:
```
data_dropped = data.drop(columns=['C02199V02655',␣
↪'C02076V03371','UNIT','TLIST(A1)'])
data_dropped.head(5)
```

[55]:

	STATISTIC	Statistic Label	Year	Sex	Age Group	VALUE
0	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	Both sexes	All ages	76.8
1	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	Both sexes	15 - 19 years	52.2
2	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	Both sexes	20 - 24 years	90.8
3	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	Both sexes	25 - 29 years	84.8
4	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	Both sexes	30 - 34 years	84.7

[56]:
```
data_rename=data.rename(columns={'Statistic Label': 'STATISTIC LABEL','Year':
↪'YEAR','Sex':'SEX','Age Group':'AGE GROUP'})
data_rename.head(5)
```

[56]:

	STATISTIC	STATISTIC LABEL	TLIST(A1)	\
0	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
1	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	
2	HIS45C01	Persons who consumed alcohol in the last twelv…	2015	

```
3  HIS45C01  Persons who consumed alcohol in the last twelv…       2015
4  HIS45C01  Persons who consumed alcohol in the last twelv…       2015

   YEAR C02199V02655         SEX C02076V03371      AGE GROUP UNIT  VALUE
0  2015             -  Both sexes             -       All ages    %   76.8
1  2015             -  Both sexes           300  15 - 19 years    %   52.2
2  2015             -  Both sexes           365  20 - 24 years    %   90.8
3  2015             -  Both sexes           410  25 - 29 years    %   84.8
4  2015             -  Both sexes           440  30 - 34 years    %   84.7
```

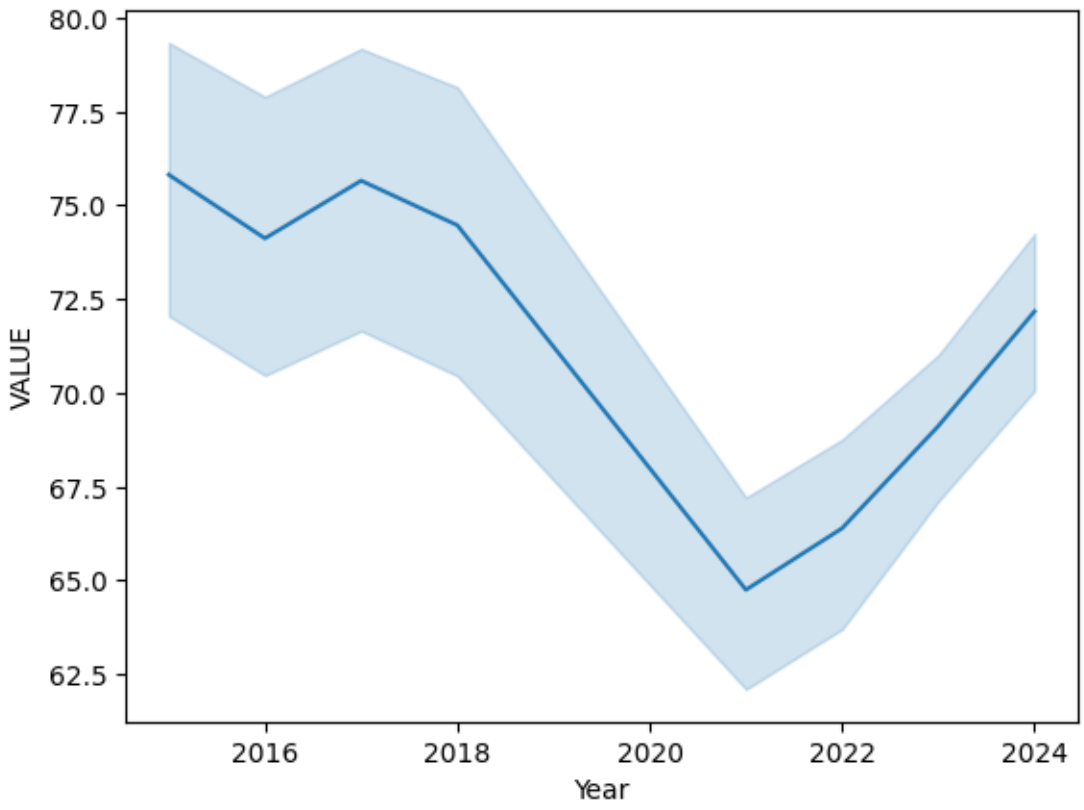
[57]:
```
filtered_data = data.groupby('Sex')['VALUE'].max()
filtered_data
```

[57]:
```
Sex
Both sexes    90.8
Female        90.7
Male          92.7
Name: VALUE, dtype: float64
```

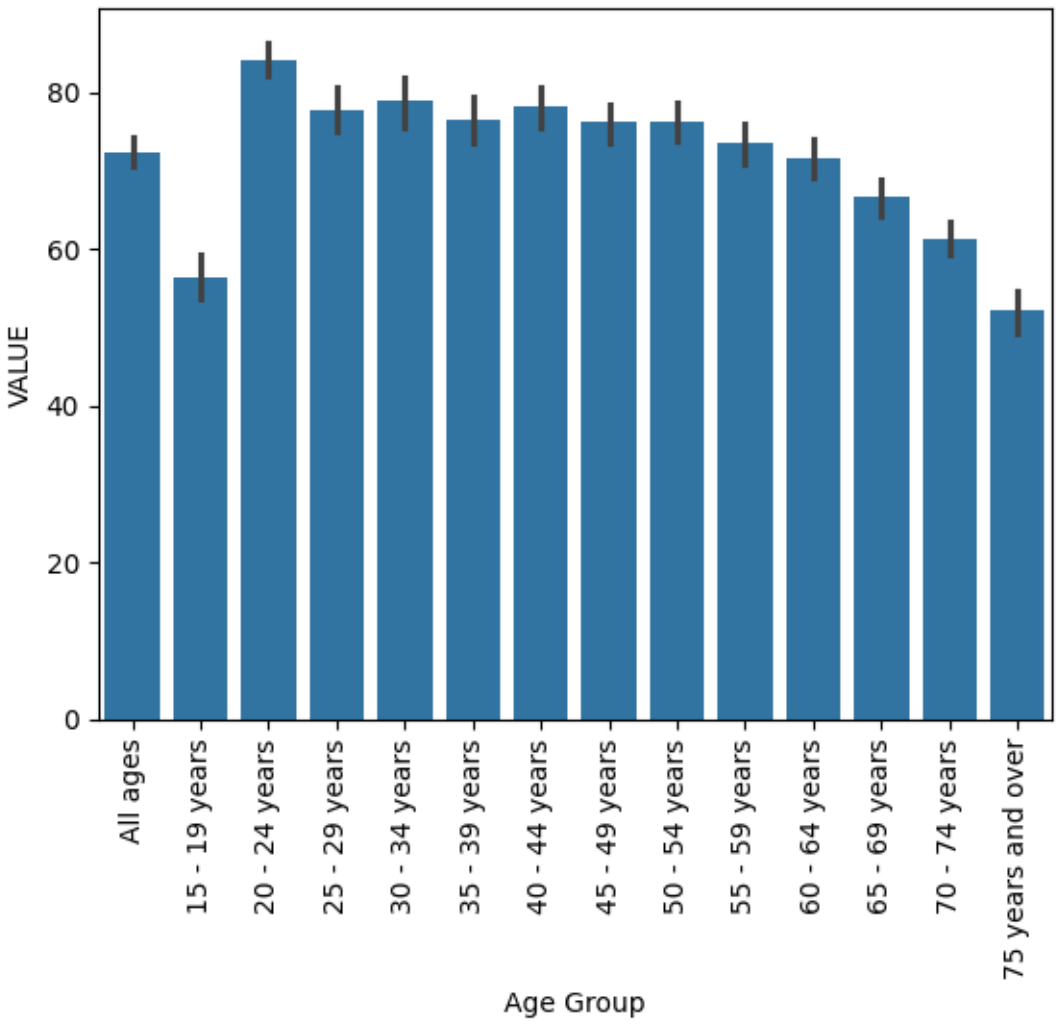
[58]:
```
filtered_data = data.groupby('Sex')['VALUE'].min()
filtered_data
```

[58]:
```
Sex
Both sexes    45.7
Female        33.3
Male          50.0
Name: VALUE, dtype: float64
```

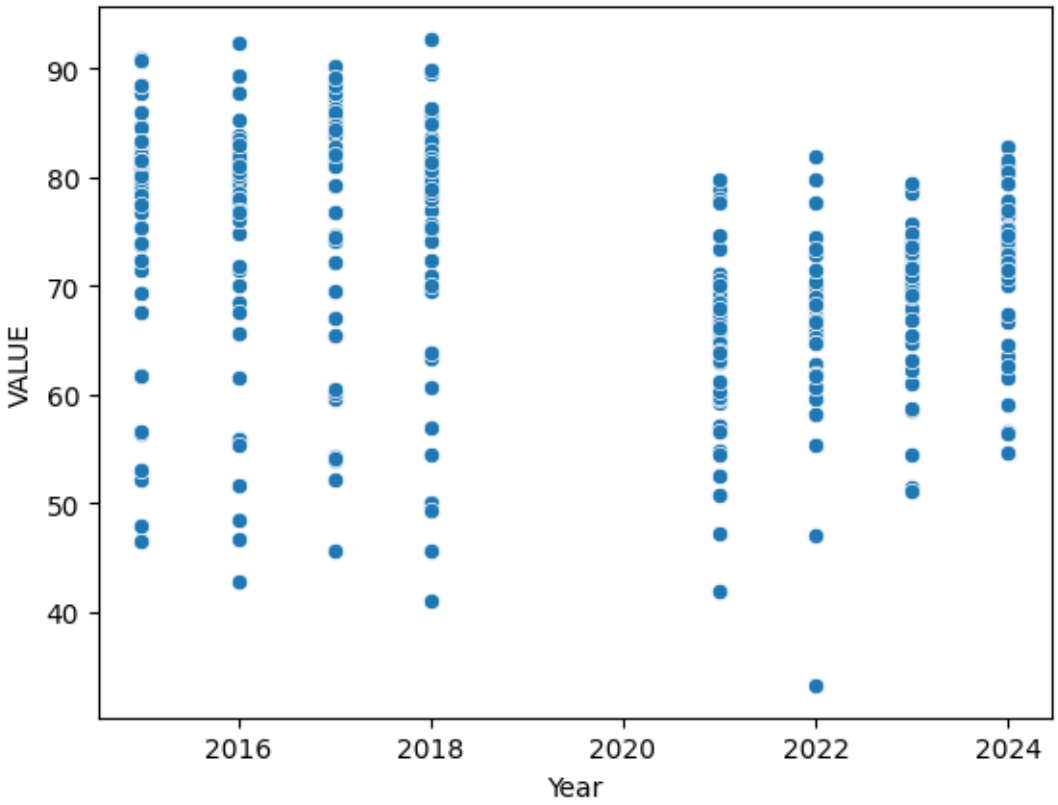
[59]:
```
sns.lineplot(data=data, x='Year', y='VALUE')
plt.show()
```



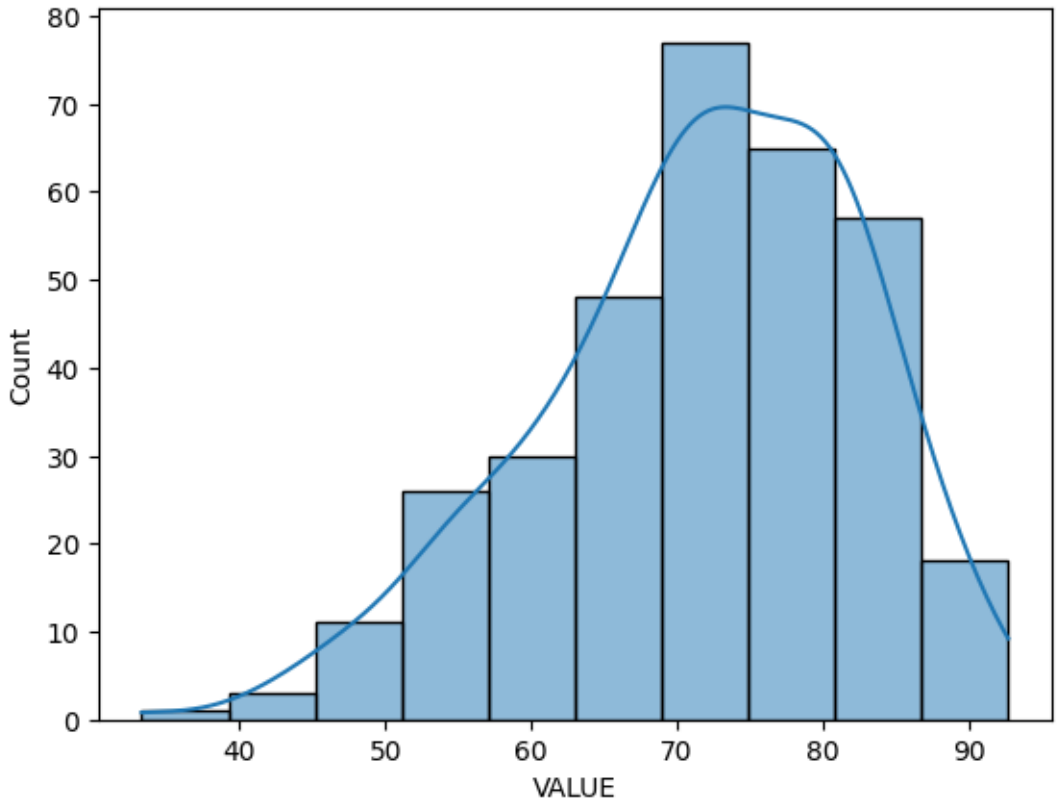
[60]:
```
sns.barplot(data=data, x='Age Group', y='VALUE')
plt.xticks(rotation=90)  # Rotate x-axis labels if needed
plt.show()
```



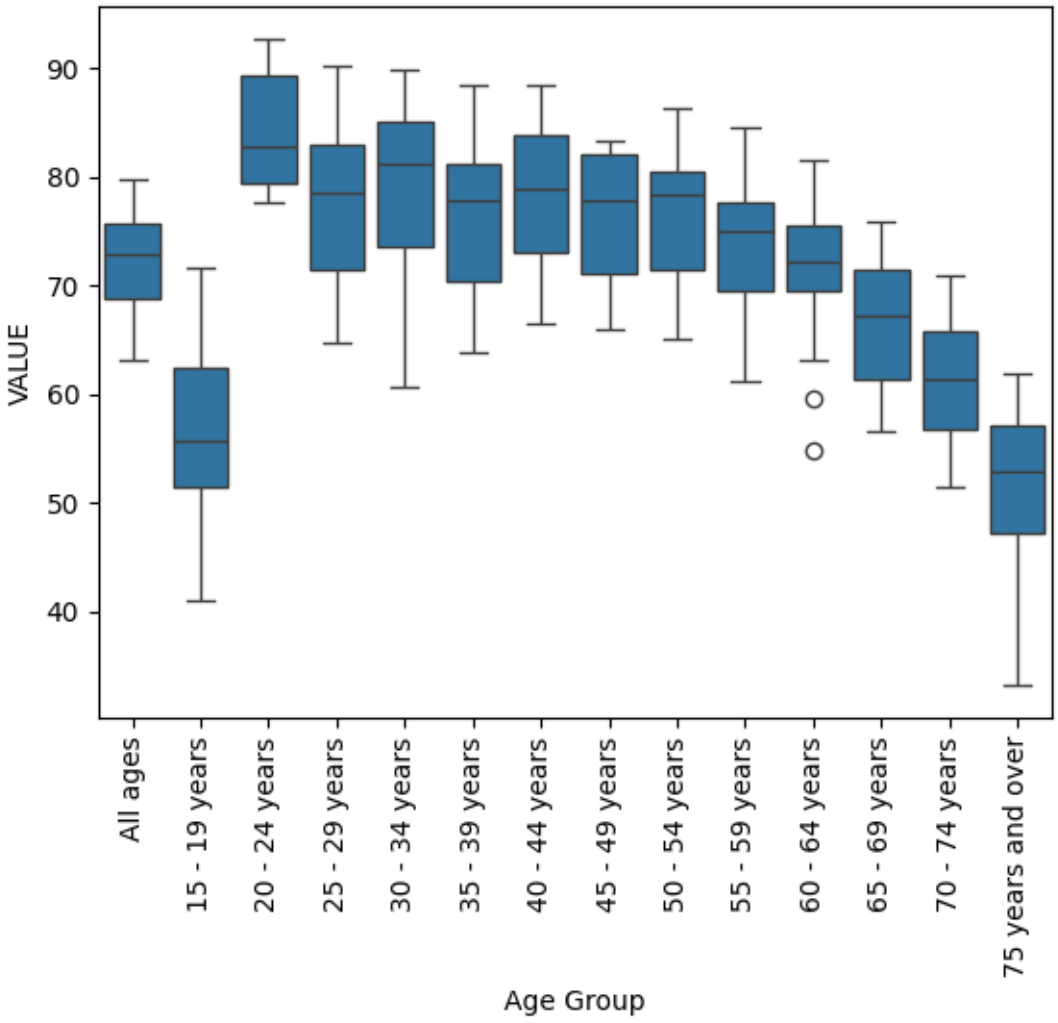
[61]:
```
sns.scatterplot(data=data, x='Year', y='VALUE')
plt.show()
```



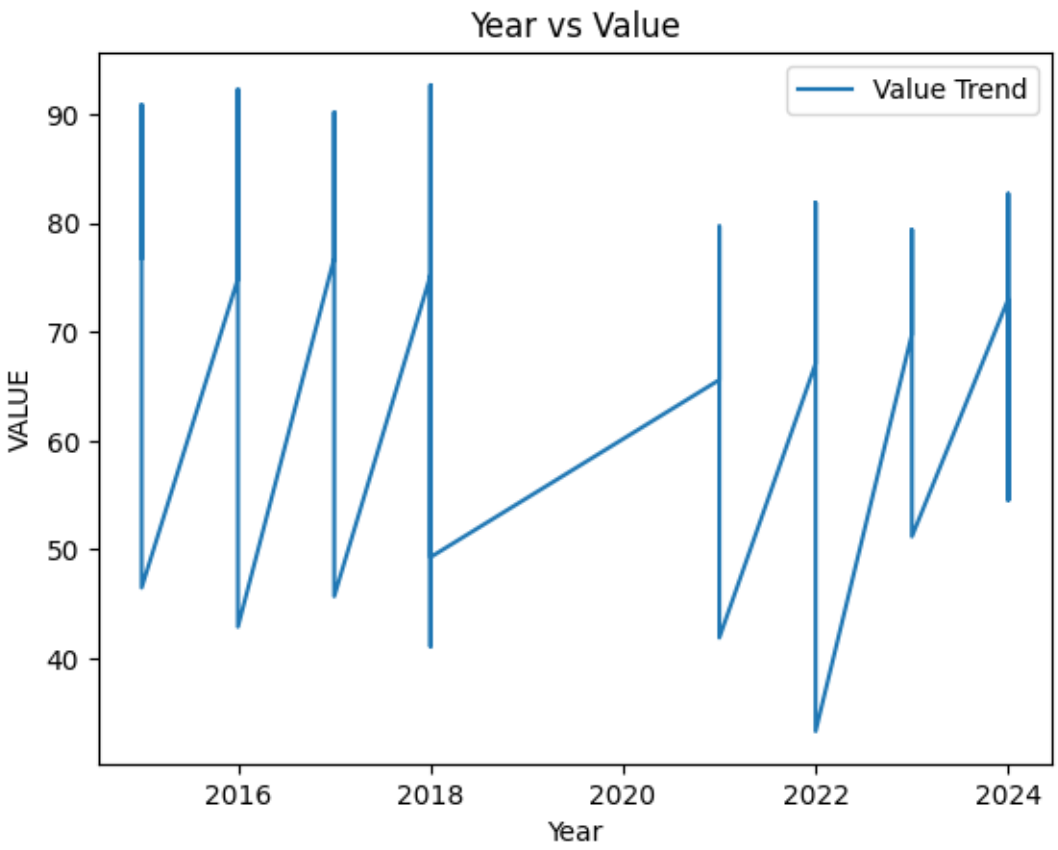
[62]:
```
sns.histplot(data=data, x='VALUE', bins=10, kde=True)
plt.show()
```



[63]:
```
sns.boxplot(data=data, x='Age Group', y='VALUE')
plt.xticks(rotation=90)
plt.show()
```

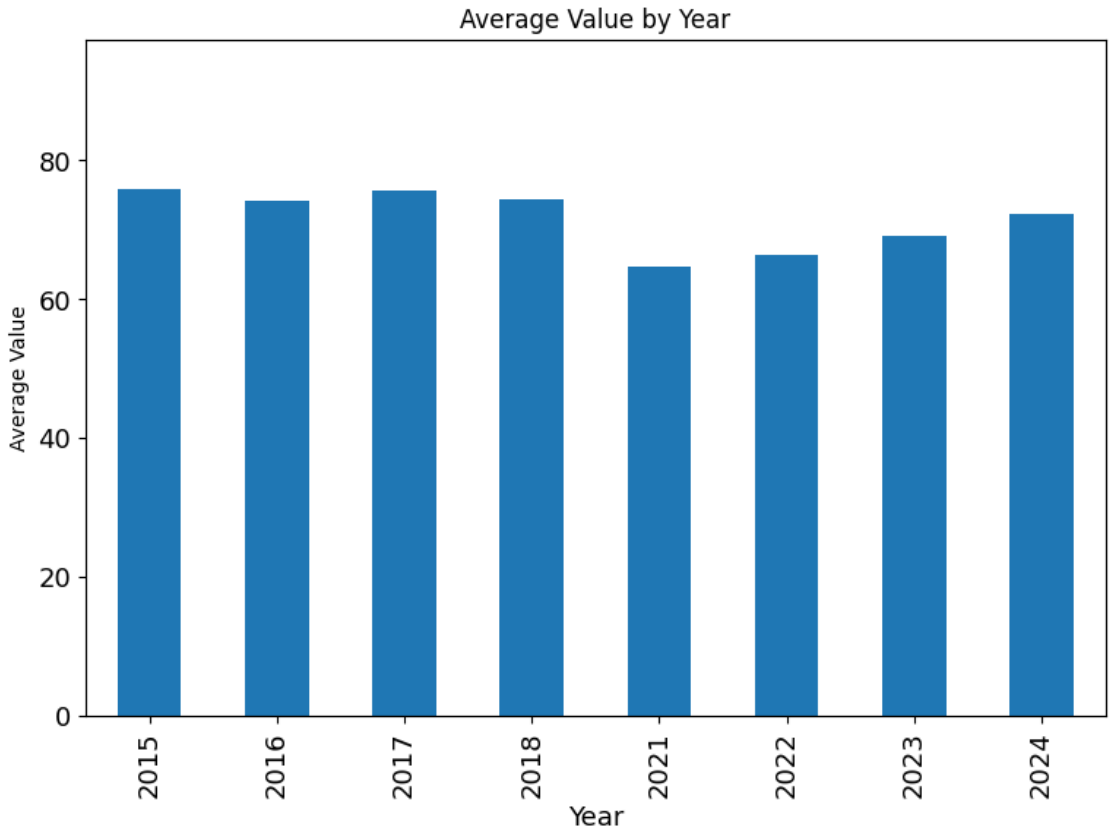


[64]:
```
plt.plot(data['Year'], data['VALUE'], label='Value Trend')
plt.xlabel('Year')
plt.ylabel('VALUE')
plt.title('Year vs Value')
plt.legend()
plt.show()
```



[ ]:

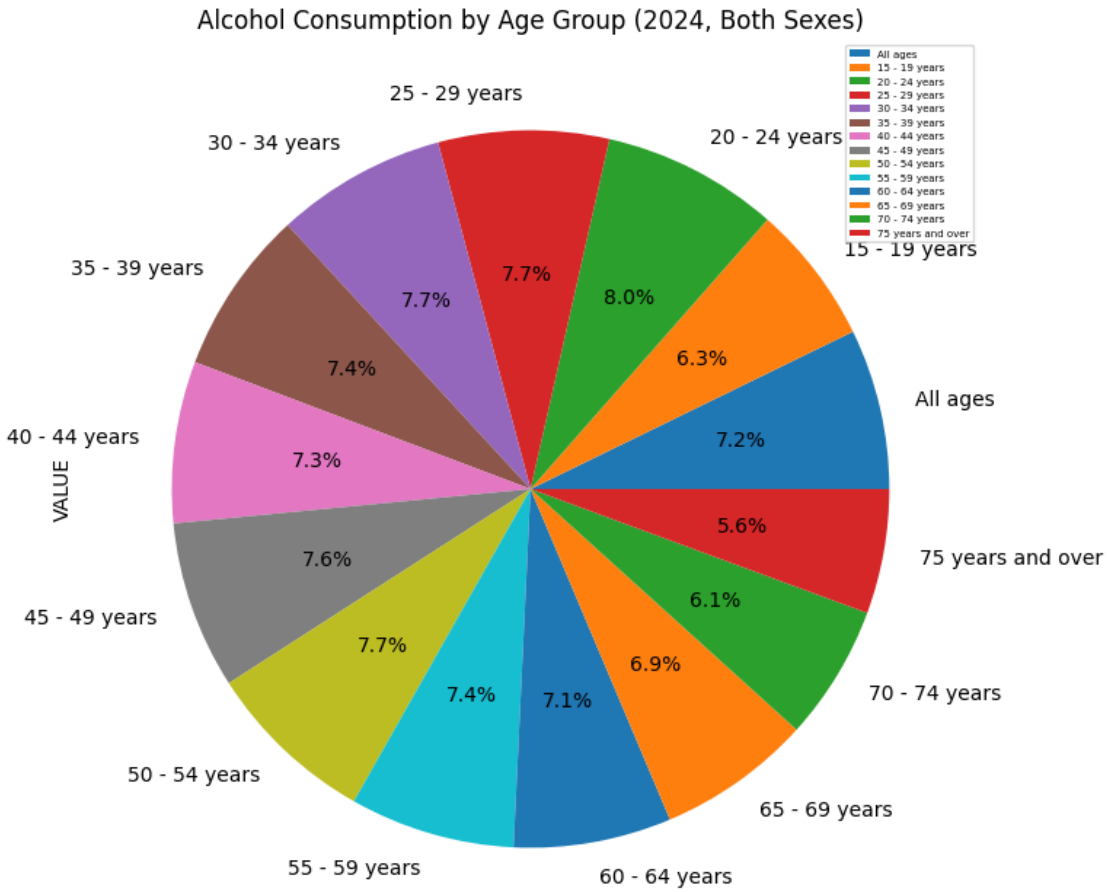
[65]:
```
plt.figure(figsize=(9,6))
plt.bar(x=data["Year"],height=data["VALUE"],color="purple")
data.groupby('Year')['VALUE'].mean().plot(kind='bar')
plt.xlabel('Year',fontsize=13)
plt.ylabel('Average Value')
plt.xticks(fontsize=13)
plt.yticks(fontsize=13)
plt.title('Average Value by Year')
plt.show()
```



[66]:
```
filtered_data = data[(data['Year'] == 2024) & (data['Sex'] == 'Both sexes')]

# Create a simple pie chart
filtered_data.plot.pie(
    y='VALUE',
    labels=filtered_data['Age Group'],
    autopct='%1.1f%%',
    figsize=(8, 8),
    title='Alcohol Consumption by Age Group (2024, Both Sexes)',

)
plt.legend(loc="upper right",fontsize="5")
plt.show()
```


Alcohol Consumption by Age Group (2024, Both Sexes)
All ages
15 - 19 years
20 - 24 years
25 - 29 years
30 - 34 years
35 - 39 years
40 - 44 years
45 - 49 years
50 - 54 years
55 - 59 years
60 - 64 years
65 - 69 years
70 - 74 years
75 years and over
VALUE
7.2%
6.3%
8.0%
7.7%
7.7%
7.4%
7.3%
7.6%
7.7%
7.4%
7.1%
6.9%
6.1%
5.6%

[ ]:

[ ]:

[ ]: