```
import pandas as pd
import numpy as np

np.random.seed(0)
df = pd.DataFrame(np.random.randn(10, 4), columns=list('ABCD'))

df.iloc[1, 2] = np.nan
df.iloc[4, 0] = np.nan
df.iloc[7, 3] = np.nan
df.iloc[9, 1] = np.nan

def highlight_na(s):
    return ['background-color: yellow' if pd.isna(v) else '' for v in s]

styled_df = df.style.apply(highlight_na)
styled_df
```

output

	A	B	C	D
0	1.764052	0.400157	0.978738	2.240893
1	1.867558	-0.977278	nan	-0.151357
2	-0.103219	0.410599	0.144044	1.454274
3	0.761038	0.121675	0.443863	0.333674
4	nan	-0.205158	0.313068	-0.854096
5	-2.552990	0.653619	0.864436	-0.742165
6	2.269755	-1.454366	0.045759	-0.187184
7	1.532779	1.469359	0.154947	nan
8	-0.887786	-1.980796	-0.347912	0.156349
9	1.230291	nan	-0.387327	-0.302303

```
import pandas as pd
import numpy as np

np.random.seed(0)
df = pd.DataFrame(np.random.randn(10, 4), columns=list('ABCD'))

styled_df = df.style.set_properties(**{
    'background-color': 'black',
    'color': 'yellow'
})
styled_df
```

output

	A	B	C	D
0	1.764052	0.400157	0.978738	2.240893
1	1.867558	-0.977278	0.950088	-0.151357
2	-0.103219	0.410599	0.144044	1.454274
3	0.761038	0.121675	0.443863	0.333674
4	1.494079	-0.205158	0.313068	-0.854096
5	-2.552990	0.653619	0.864436	-0.742165
6	2.269755	-1.454366	0.045759	-0.187184
7	1.532779	1.469359	0.154947	0.378163
8	-0.887786	-1.980796	-0.347912	0.156349
9	1.230291	1.202380	-0.387327	-0.302303

```
import pandas as pd
import numpy as np
df = pd.DataFrame({
    'A': [1, 2, np.nan, 4],
    'B': [np.nan, 2, 3, 4],
    'C': [1, 2, 3, np.nan],
    'D': [1, np.nan, np.nan, 4]
})

missing_values = df.isna()
print(missing_values)
```

output

```
       A      B      C      D
0  False   True  False  False
1  False  False  False   True
2   True  False  False   True
3  False  False   True  False
       A      B      C      D
0  False   True  False  False
1  False  False  False   True
2   True  False  False   True
3  False  False   True  False
```

```
import pandas as pd
import numpy as np

df = pd.DataFrame({
    'A': [1, 2, np.nan, 4],
    'B': [np.nan, 2, 3, 4],
    'C': [1, 2, 3, np.nan],
    'D': [1, np.nan, np.nan, 4]
})

df_filled = df.fillna(0)
print(df_filled)
```

output

```
     A    B    C    D
0  1.0  0.0  1.0  1.0
1  2.0  2.0  2.0  0.0
2  0.0  3.0  3.0  0.0
3  4.0  4.0  0.0  4.0
     A    B    C    D
0  1.0  0.0  1.0  1.0
1  2.0  2.0  2.0  0.0
2  0.0  3.0  3.0  0.0
3  4.0  4.0  0.0  4.0
```

```
import pandas as pd
import numpy as np

df = pd.DataFrame({
    'A': [1, 2, np.nan, 4],
    'B': [np.nan, 2, 3, 4],
    'C': [1, 2, 3, np.nan],
    'D': [1, np.nan, np.nan, 4]
})

df_with_nans = df[df.isna().sum(axis=1) >= 2]
print(df_with_nans)
```

output

```
     A    B    C   D
2  NaN  3.0  3.0 NaN
     A    B    C   D
2  NaN  3.0  3.0 NaN
```

```
import pandas as pd

df = pd.DataFrame({
    'school_code': ['s1', 's2', 's1', 's2', 's1', 's3'],
    'class': ['V', 'V', 'VI', 'VI', 'V', 'VI'],
    'name': ['Alberto', 'Gino', 'Ryan', 'Eddy', 'Steven', 'Smith'],
    'age': [12, 13, 11, 14, 12, 13]
})

grouped = df.groupby('school_code')

print(type(grouped))
```

output

```
<class 'pandas.core.groupby.generic.DataFrameGroupBy'>
<class 'pandas.core.groupby.generic.DataFrameGroupBy'>
```

```
import pandas as pd

df = pd.DataFrame({
    'school_code': ['s1', 's2', 's1', 's2', 's1', 's3'],
    'class': ['V', 'V', 'VI', 'VI', 'V', 'VI'],
    'name': ['Alberto', 'Gino', 'Ryan', 'Eddy', 'Steven', 'Smith'],
    'age': [12, 13, 11, 14, 12, 13]
})

grouped = df.groupby('school_code')['age'].agg(['mean', 'min', 'max'])
print(grouped)
```

output

```
                mean  min  max
school_code
s1          11.666667   11   12
s2          13.500000   13   14
s3          13.000000   13   13
```

```
import pandas as pd

df = pd.DataFrame({
    'school_code': ['s1', 's2', 's1', 's2', 's1', 's3'],
    'class': ['V', 'V', 'VI', 'VI', 'V', 'VI'],
    'name': ['Alberto', 'Gino', 'Ryan', 'Eddy', 'Steven', 'Smith'],
    'age': [12, 13, 11, 14, 12, 13]
})

grouped = df.groupby(['school_code',
'class']).size().reset_index(name='counts')
print(grouped)
```

output

	school_code	class	counts
0	s1	V	2
1	s1	VI	1
2	s2	V	1
3	s2	VI	1
4	s3	VI	1

```
import pandas as pd
file_path = '/content/world_alcohol.csv'
df = pd.read_csv(file_path)

print(df.shape)
print(df.columns)
)output
```

```
(3257, 5)
Index(['Year', 'WHO region', 'Country', 'Beverage Types', 'Display Value'], dtype='object')
```

```
import pandas as pd

df = pd.DataFrame({
    'name': ['Alberto', 'Gino', 'Ryan', 'Eddy', 'Steven', 'Smith'],
    'age': [12, 13, 11, 14, 12, 13]
})

substring = 'ry'
index = df['name'].str.find(substring)
print(index)
```

output

```
0   -1
1   -1
2   -1
3   -1
4   -1
5   -1
Name: name, dtype: int64
```