

```
# Import necessary libraries
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# URL to the white wine dataset
url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/wine-quality/winequality-white.csv'

# Load data into DataFrame, using semicolon as separator
try:
    wine_df = pd.read_csv(url, sep=';')
    print("Data loaded successfully.")
except Exception as e:
    print(f"Data loading error: {e}")
```

➡ Data loaded successfully.

```
# Display first 5 rows for familiarization
print("First 5 rows of data:")
print(wine_df.head())
```

```
# 1. Calculate descriptive statistics for all indicators
print("\nDescriptive statistics:")
# .T transposes the output, making it more readable
print(wine_df.describe().T)
```

➡ First 5 rows of data:

	fixed acidity	volatile acidity	citric acid	residual sugar	chlorides	\
0	7.0	0.27	0.36	20.7	0.045	
1	6.3	0.30	0.34	1.6	0.049	
2	8.1	0.28	0.40	6.9	0.050	
3	7.2	0.23	0.32	8.5	0.058	
4	7.2	0.23	0.32	8.5	0.058	

	free sulfur dioxide	total sulfur dioxide	density	pH	sulphates	\
0	45.0	170.0	1.0010	3.00	0.45	
1	14.0	132.0	0.9940	3.30	0.49	
2	30.0	97.0	0.9951	3.26	0.44	
3	47.0	186.0	0.9956	3.19	0.40	
4	47.0	186.0	0.9956	3.19	0.40	

	alcohol	quality
0	8.8	6
1	9.5	6
2	10.1	6
3	9.9	6
4	9.9	6

Descriptive statistics:

	count	mean	std	min	25%	\
fixed acidity	4898.0	6.854788	0.843868	3.80000	6.300000	
volatile acidity	4898.0	0.278241	0.100795	0.08000	0.210000	
citric acid	4898.0	0.334192	0.121020	0.00000	0.270000	
residual sugar	4898.0	6.391415	5.072058	0.60000	1.700000	
chlorides	4898.0	0.045772	0.021848	0.00900	0.036000	
free sulfur dioxide	4898.0	35.308085	17.007137	2.00000	23.000000	
total sulfur dioxide	4898.0	138.360657	42.498065	9.00000	108.000000	
density	4898.0	0.994027	0.002991	0.98711	0.991723	
pH	4898.0	3.188267	0.151001	2.72000	3.090000	
sulphates	4898.0	0.489847	0.114126	0.22000	0.410000	
alcohol	4898.0	10.514267	1.230621	8.00000	9.500000	
quality	4898.0	5.877909	0.885639	3.00000	5.000000	

	50%	75%	max
fixed acidity	6.80000	7.3000	14.20000
volatile acidity	0.26000	0.3200	1.10000
citric acid	0.32000	0.3900	1.66000
residual sugar	5.20000	9.9000	65.80000
chlorides	0.04300	0.0500	0.34600
free sulfur dioxide	34.00000	46.0000	289.00000
total sulfur dioxide	134.00000	167.0000	440.00000
density	0.99374	0.9961	1.03898
pH	3.18000	3.2800	3.82000
sulphates	0.47000	0.5500	1.08000
alcohol	10.40000	11.4000	14.20000
quality	6.00000	6.0000	9.00000

```
# 2. Calculate mode for quality indicator
quality_mode = wine_df['quality'].mode()[0]
print(f"\nMode for 'quality' indicator: {quality_mode}")
```

➡

Mode for 'quality' indicator: 6

```
# 3. Build visualization – histogram of wine quality distribution
plt.style.use('seaborn-v0_8-whitegrid') # Use modern style
fig, ax = plt.subplots(figsize=(10, 6))

sns.histplot(wine_df['quality'], bins=6, kde=False, ax=ax, color='skyblue', discrete=True)

ax.set_title('Distribution of White Wine Quality Ratings', fontsize=16)
ax.set_xlabel('Quality Rating (from 3 to 9)', fontsize=12)
ax.set_ylabel('Number of Samples', fontsize=12)
ax.set_xticks(range(wine_df['quality'].min(), wine_df['quality'].max() + 1)) # Clear marks on X axis

# Add labels with counts on each bar
for p in ax.patches:
    ax.annotate(f'{int(p.get_height())}', (p.get_x() + p.get_width() / 2., p.get_height()),
                ha='center', va='center', xytext=(0, 9), textcoords='offset points')

plt.show()
```

