In [2]:

```python
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

# Load the dataset
data = pd.read_csv('Data Cleaning and Preprocessing.csv')

# Display basic information about the dataset
print(data.info())
print(data.describe())

# Set visual aesthetics
plt.style.use('seaborn-darkgrid')

# Plotting distributions and boxplots for the main columns
fig, axes = plt.subplots(nrows=5, ncols=2, figsize=(15, 20))
numeric_columns = data.select_dtypes(include=['float64']).columns[:10]  #

for i, col in enumerate(numeric_columns):
    # Histogram for distribution
    sns.histplot(data[col].dropna(), bins=20, kde=True, ax=axes[i // 2, i
    axes[i // 2, i % 2].set_title(f'Distribution of {col}')
    axes[i // 2, i % 2].set_xlabel(col)
    axes[i // 2, i % 2].set_ylabel('Frequency')

plt.tight_layout()
plt.show()

# Plotting boxplots to check for outliers
fig, axes = plt.subplots(nrows=5, ncols=2, figsize=(15, 20))
for i, col in enumerate(numeric_columns):
    sns.boxplot(x=data[col], ax=axes[i // 2, i % 2])
    axes[i // 2, i % 2].set_title(f'Boxplot of {col}')
    axes[i // 2, i % 2].set_xlabel(col)

plt.tight_layout()
plt.show()

# Correlation heatmap
plt.figure(figsize=(12, 10))
corr_matrix = data.corr(numeric_only=True)
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f", linewidth
plt.title('Correlation Heatmap of Variables')
plt.show()
```
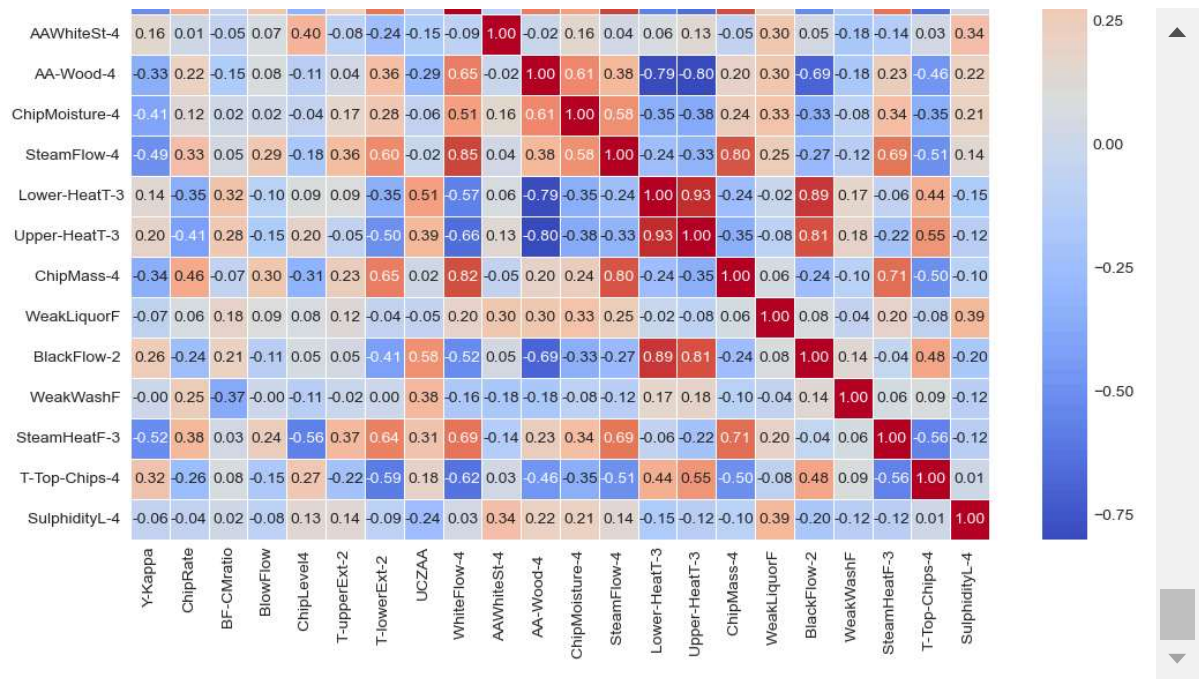
Explanation: Data Loading: This loads the CSV file and prints basic info about the dataset. Distributions: The sns.histplot function is used to show each variable's distribution. Outliers: Boxplots using sns.boxplot reveal potential outliers for each variable. Correlation Heatmap: The heatmap visualizes correlations between variables, highlighting strong relationships.

In [ ]: 1