```
# 1. Import all the required Python Libraries
import pandas as pd
import numpy as np

# 2. Locate an open source data from the web.
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"

# 3. Load the Dataset into pandas data frame
column_names = ['sepal_length', 'sepal_width', 'petal_length', 'petal_width', 'class']
iris_df = pd.read_csv(url, names=column_names)

# Display the first few rows of the dataset to verify the import
print("First few rows of the Iris dataset:")
print(iris_df.head())

# 4. Data Preprocessing:
# Check for missing values using pandas info(), describe() functions.
print("\nInformation about the dataset:")
print(iris_df.info())

print("\nDescriptive statistics of the dataset:")
print(iris_df.describe())

print(iris_df.isnull())

# Variable Descriptions:
# - Sepal Length, Sepal Width, Petal Length, Petal Width: Numeric variables.
# - Class: Categorical variable representing the species of iris flowers.

# Check the dimensions of the data frame.
print("\nDimensions of the dataset (rows, columns):", iris_df.shape)
```

```
# 5. Data Formatting and Normalization:
# Summarize the types of variables by checking data types.
print("\nData Types of Variables:")
print(iris_df.dtypes)

# 6. Turn categorical variables into quantitative variables.
# The 'class' variable is categorical; we can use one-hot encoding to convert it to quantitative.
iris_df = pd.get_dummies(iris_df, columns=['class'], drop_first=True)

# Display the updated dataframe.
print("\nUpdated DataFrame after one-hot encoding:")
print(iris_df.head())
```