**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**КАФЕДРА ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ**

**Лабораторна робота №**2

з дисципліни «Інтелектуальний аналіз та візуалізація даних»

тема «ВІЗУАЛІЗАЦІЯ НАБОРІВ ДАНИХ. ЗАГАЛЬНА СТАТИСТИКА»

**Варіант №8**

Виконала студентка

групи КН-21

Іванова А.О.

Перевірив(-ла):

МІнаєва Ю.І.

**Київ – 2025**

Мета роботи: Вивчення можливостей і основних принципів роботи по виконанню статистичного аналізу даних. Графічний розвідувальний аналіз даних. Первинний статистичний аналіз даних.

# Імпорт бібліотек

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

from scipy.stats import skew, kurtosis

from sklearn.datasets import load\_wine

# Завантаження другого датасету (Climate Change Data) з Datahub.io

climate\_data\_url = "https://datahub.io/core/co2-ppm/r/co2-mm-mlo.csv"

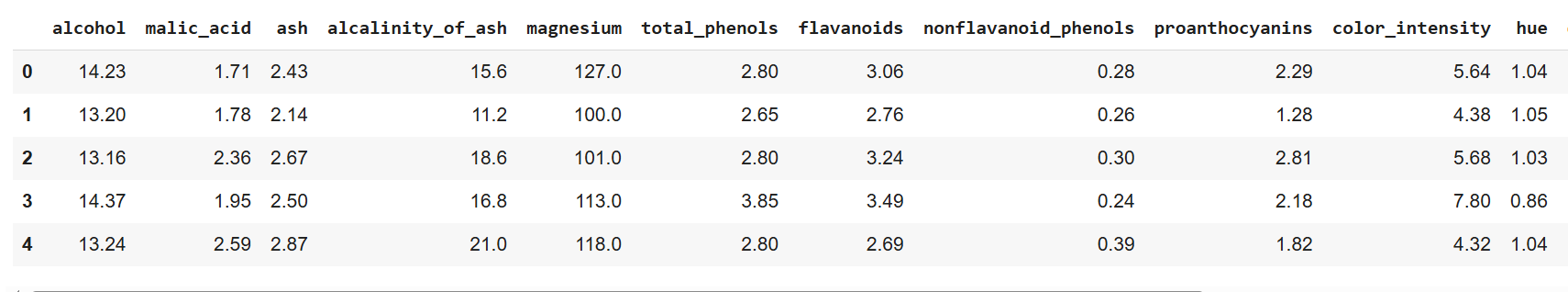
df\_climate = pd.read\_csv(climate\_data\_url)

wine = load\_wine()

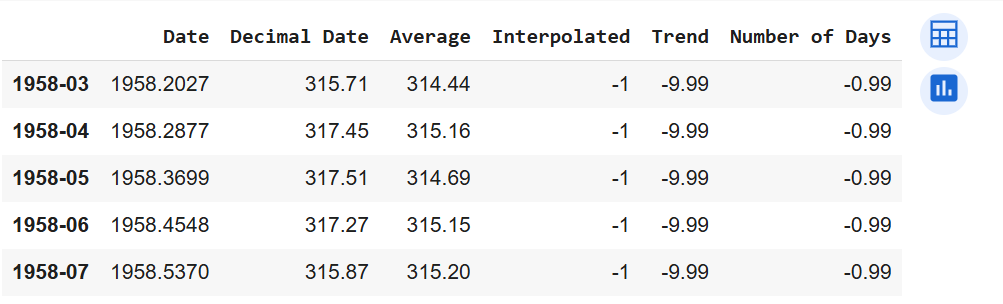
df\_wine = pd.DataFrame(data=wine.data, columns=wine.feature\_names)

df\_wine['target'] = wine.target

df\_wine.head()



df\_climate.head()



# Обираємо числові стовпці для аналізу

wine\_feature = 'alcohol'

climate\_feature = 'Interpolated'

# Обчислення описових статистик

wine\_stats = {

"Математичне сподівання": df\_wine[wine\_feature].mean(),

"Оцінка медіани": np.median(df\_wine[wine\_feature]),

"Дисперсія": np.var(df\_wine[wine\_feature]),

"Стандартне відхиленя": np.std(df\_wine[wine\_feature]),

"Асиметрія": skew(df\_wine[wine\_feature]),

"Ексцес": kurtosis(df\_wine[wine\_feature])

}

climate\_stats = {

"Математичне сподівання": df\_climate[climate\_feature].mean(),

"Оцінка медіани": np.median(df\_climate[climate\_feature]),

"Дисперсія": np.var(df\_climate[climate\_feature]),

"Стандартне відхиленя": np.std(df\_climate[climate\_feature]),

"Асиметрія": skew(df\_climate[climate\_feature]),

"Ексцес": kurtosis(df\_climate[climate\_feature])

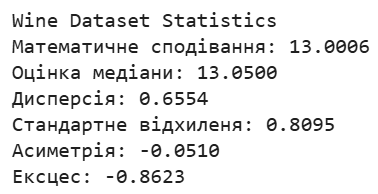
}

# Вивід статистичних показників

print("Wine Dataset Statistics")

for key, value in wine\_stats.items():

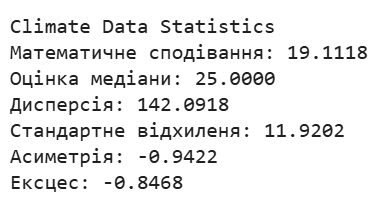
print(f"{key}: {value:.4f}")



print("\nClimate Data Statistics")

for key, value in climate\_stats.items():

print(f"{key}: {value:.4f}")



# Побудова гістограми для Wine Dataset

plt.figure(figsize=(10, 5))

sns.histplot(df\_wine[wine\_feature], bins=15, kde=True, color='blue', label='Wine Dataset')

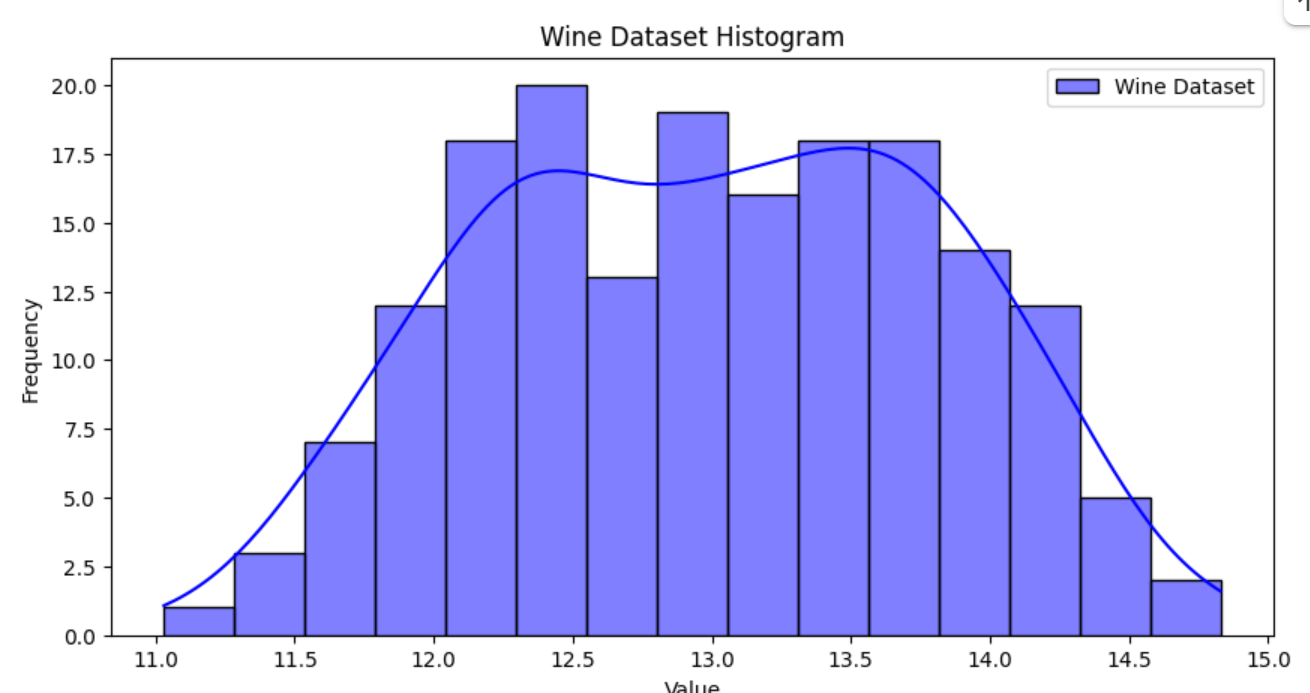
plt.legend()

plt.title("Wine Dataset Histogram")

plt.xlabel("Value")

plt.ylabel("Frequency")

plt.show()



# Побудова гістограми для Climate Change Dataset

plt.figure(figsize=(10, 5))

sns.histplot(df\_climate[climate\_feature], bins=15, kde=True, color='red', label='Climate Change Dataset', alpha=0.7)

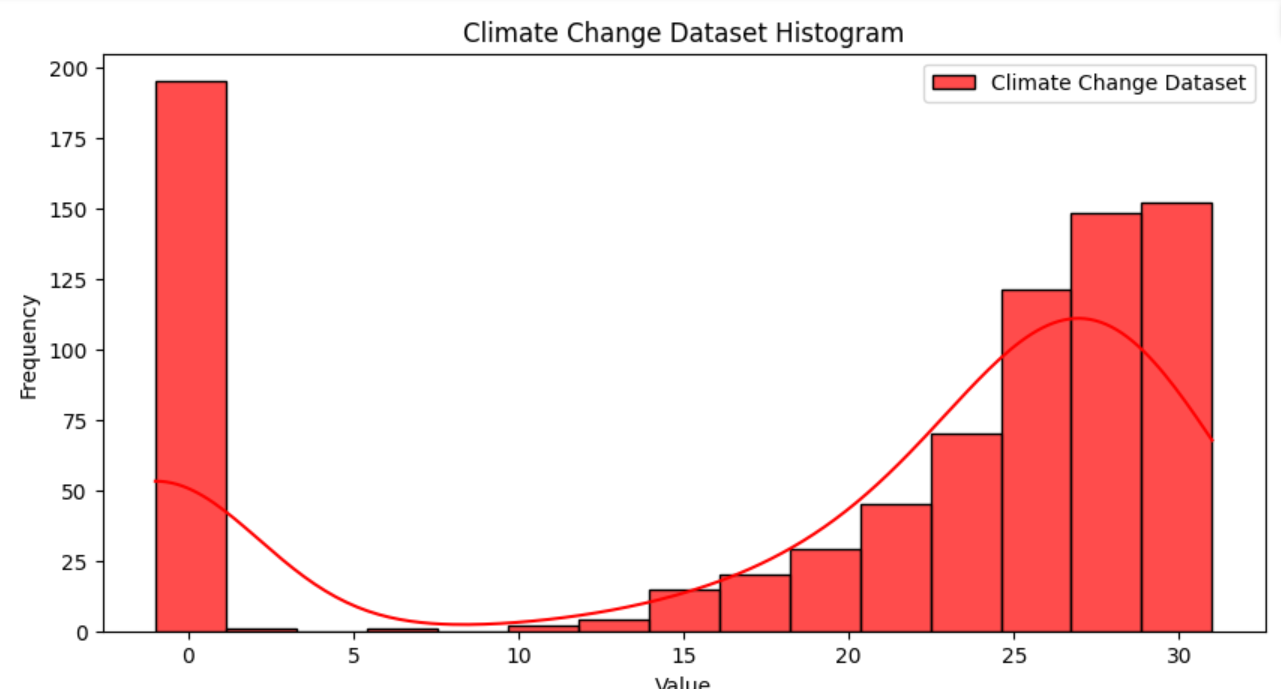
plt.legend()

plt.title("Climate Change Dataset Histogram")

plt.xlabel("Value")

plt.ylabel("Frequency")

plt.show()



# Побудова радіальної діаграми

def plot\_radial\_chart(data, feature\_name):

angles = np.linspace(0, 2 \* np.pi, len(data), endpoint=False)

stats\_values = list(data.values())

stats\_values.append(stats\_values[0]) # Замкнути діаграму

angles = np.concatenate((angles, [angles[0]]))

plt.figure(figsize=(6, 6))

ax = plt.subplot(111, polar=True)

ax.plot(angles, stats\_values, 'o-', linewidth=2)

ax.fill(angles, stats\_values, alpha=0.25)

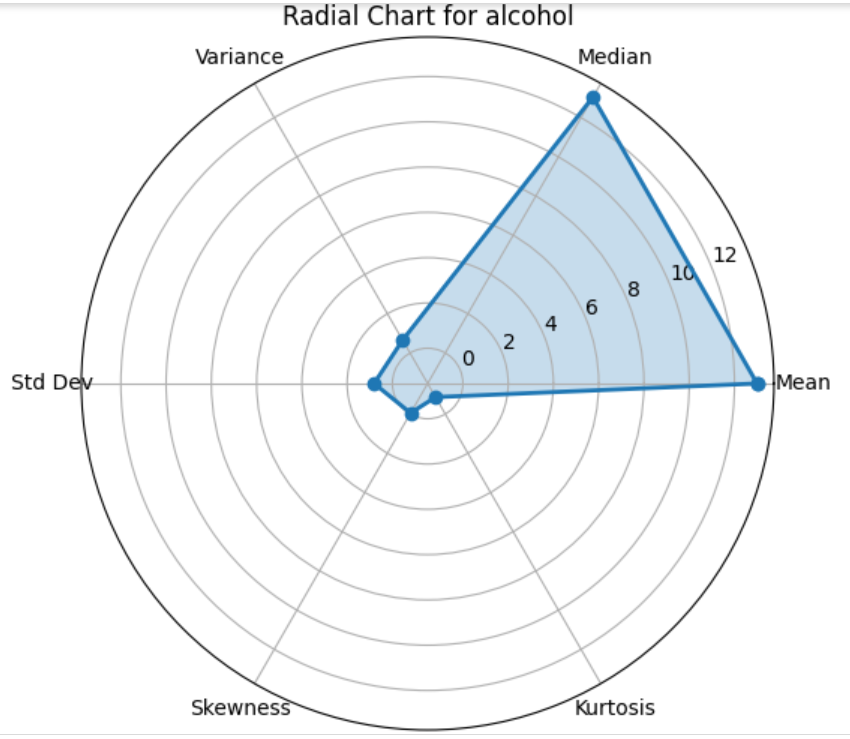
ax.set\_xticks(angles[:-1])

ax.set\_xticklabels(data.keys(), fontsize=10)

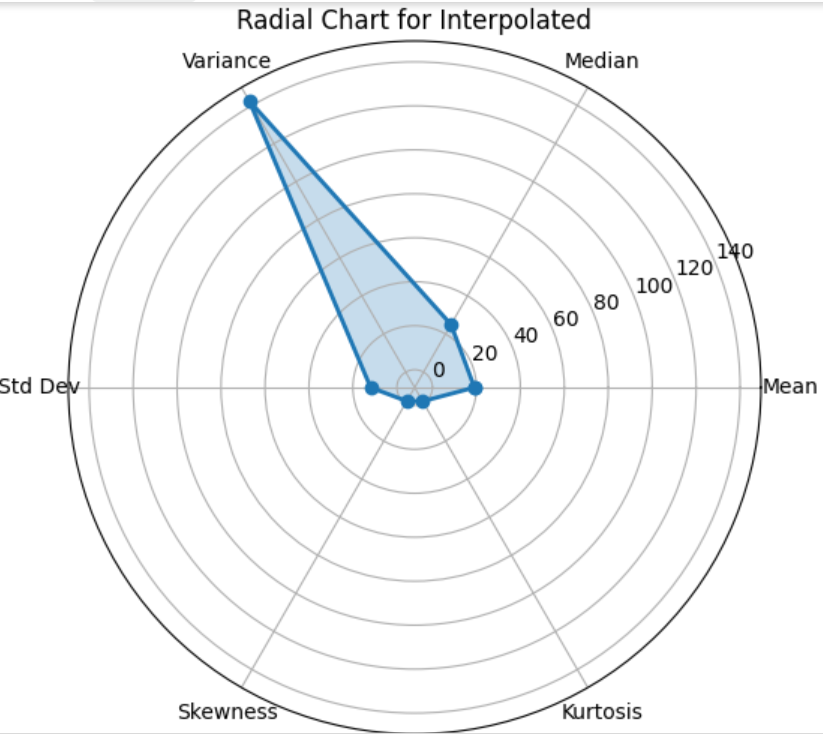
plt.title(f'Radial Chart for {feature\_name}')

plt.show()

plot\_radial\_chart(wine\_stats, wine\_feature)



plot\_radial\_chart(climate\_stats, climate\_feature)



**Висновок:** лабораторній роботі було виконано категоризованих та радіальних діаграм, проведено аналіз за допомогою описових статистик.