AI 程序设计@NJU

实验 6 可视化与数据探索参考答案

1.

```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt

wine = pd.read_csv("winequality-red.csv")
print(wine.corr().quality)

# 之前的参考图是 4 个图分别截取的，这样的效果不佳且很麻烦，后面遇到的数据探索一般是所有的特征列与标签列（这里是 quality）比较，因此图的数量会比较大，是否想过或尝试过把所有的图作为子图放在一张图中？试一试如下代码。

fig, axes = plt.subplots(2, 2, dpi = 100)
sns.barplot(x = 'quality', y = 'volatile acidity', data = wine, ax = axes[0,0], estimator = np.median)
sns.barplot(x = 'quality', y = 'citric acid', data = wine, ax = axes[0,1], estimator = np.median)
sns.barplot(x = 'quality', y = 'sulphates', data = wine, ax = axes[1,0], estimator = np.median)
sns.barplot(x = 'quality', y = 'alcohol', data = wine, ax = axes[1,1], estimator = np.median)
plt.show()
```

2.

```
# -*- coding: utf-8 -*-
from sklearn import datasets
import matplotlib.pyplot as plt

iris = datasets.load_iris()       # 载入数据
print(iris.data)                  # 输出数据
print(iris.data.shape)            # 输出数据形状
print(iris.target)                # 输出数据标签
X = [item[0] for item in iris.data]   # 获取萼片长度
Y = [item[2] for item in iris.data]   # 获取花瓣长度
# 前 50 个山鸢尾样本
```

```
plt.scatter(X[:50], Y[:50], color = 'red', marker = 'o', label = 'setosa')
# 中间 50 个变色鸢尾样本
plt.scatter(X[50:100], Y[50:100], color = 'green', marker = '*', label =
'versicolor')
# 后 50 个弗吉尼亚鸢尾样本
plt.scatter(X[100:], Y[100:], color = 'blue', marker = 'D', label = 'virginica')
plt.legend(loc = 'best')
plt.show()
```

3.

```
from pyecharts import Radar

# 3 组数据，每组有 5 个维度
data1 = [[1, 0.99, 0.95, 0.95, 0.8]]
data2 = [[0.9, 0.6, 0.95, 0.8, 0.99]]
data3 = [[0.7, 0.7, 0.8, 0.85, 0.01]]
# 调整雷达图各维度的范围大小，可用如"max": 1, "min": -1 这样的完整范围，默认 min
为 0
c_schema= [{"name": "热血", "max": 1},
           {"name": "吃货", "max": 1},
           {"name": "勇气", "max": 1},
           {"name": "友谊", "max": 1},
           {"name": "路痴", "max": 1}]

radar = Radar()
radar.config(c_schema = c_schema, shape = "circle")
radar.add("Luffy", data1, item_color = 'green', area_color = 'green',
area_opacity = 0.35, legend_top = 'bottom', line_width = 2)
radar.add("Zoro", data2, item_color = 'blue', area_color = 'blue', area_opacity
= 0.35, legend_top = 'bottom', line_width = 2)
radar.add("Nami", data3, item_color = 'yellow', area_color = 'yellow',
area_opacity = 0.35, legend_top = 'bottom', line_width = 2)
radar.render("One Piece.html")
```