

**《Python程序设计》**

**课程项目作业报告**



**序号及名称** 作业3：关于金融学专业数据分析

**学号与姓名** 3240103494 郑雨静

**年级与专业** 大一 工商管理

**学 院 (系)** 丹青学园

**日 期** 2025年4月28日

# 一、项目名称

人口超过100万的城市群中的人口（占总人口的百分比）

# 二、研究背景

研究动机：研究世界不同国家人口增长的趋势以及不同国家大城市人口在总人口中的占比。

数据来源：世界银行网站

# 三、数据获取与处理

#引入需要的库

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

# 读取数据

data = pd.read\_excel('数据.xls')

# 获取用户输入的多个国家名称（注意分隔符的中英文输入）

country\_names = input("请输入要分析的多个国家名称（用英文逗号隔开）: ")

countries = [country.strip() for country in country\_names.split(',')]

# 筛选出对应多个国家的数据

selected\_data = data[data['Country Name'].isin(countries)]

# 检查是否存在对应国家的数据

if selected\_data.empty:

print(f"未找到 {country\_names} 这些国家的数据，请检查输入是否正确。")

else:

# 获取用户输入的年份（注意分隔符的中英文输入）

year\_input = input(f"请输入要分析的年份（1960-2024）（用英文逗号隔开）: ")

years = [int(year.strip()) for year in year\_input.split(',')]

# 将年份列名从数字转换为字符串类型（确保能正确匹配列名）

years = [str(year) for year in years]

# 筛选出对应年份的数据

filtered\_data = selected\_data[selected\_data[years].notnull().any(axis=1)]

#检查是否存在对应年份的数据

if filtered\_data.empty:

print(f"{country\_names} 这些国家中未找到 {year\_input} 年份的数据，请检查输入是否正确。")

else:

# 描述性统计分析

print(filtered\_data[years].describe())

# 数据可视化 - 绘制直方图

sns.set(style="whitegrid")

plt.figure(figsize=(10, 6))

for year in years:

sns.histplot(filtered\_data[year].dropna(), kde=True, color='skyblue', alpha=0.6, label=year)

plt.title(f"{country\_names} 在 {year\_input} 年份人口超过100万的城市群中的人口（占总人口的百分比）分布")

plt.xlabel('百分比')

plt.ylabel('频率')

plt.legend()

plt.tight\_layout()

plt.show()

# 数据可视化 - 绘制折线图展示随时间变化趋势

plt.figure(figsize=(10, 6))

for country in countries:

country\_data = filtered\_data[filtered\_data['Country Name'] == country]

for year in years:

x = country\_data.index[country\_data[year].notnull()]

y = country\_data[year].dropna()

if not x.empty and not y.empty:

sns.lineplot(x=x, y=y, label=f"{country} - {year}")

plt.title(f"{country\_names} 在 {year\_input} 年份人口超过100万的城市群中的人口（占总人口的百分比）随时间变化")

plt.xlabel('年份')

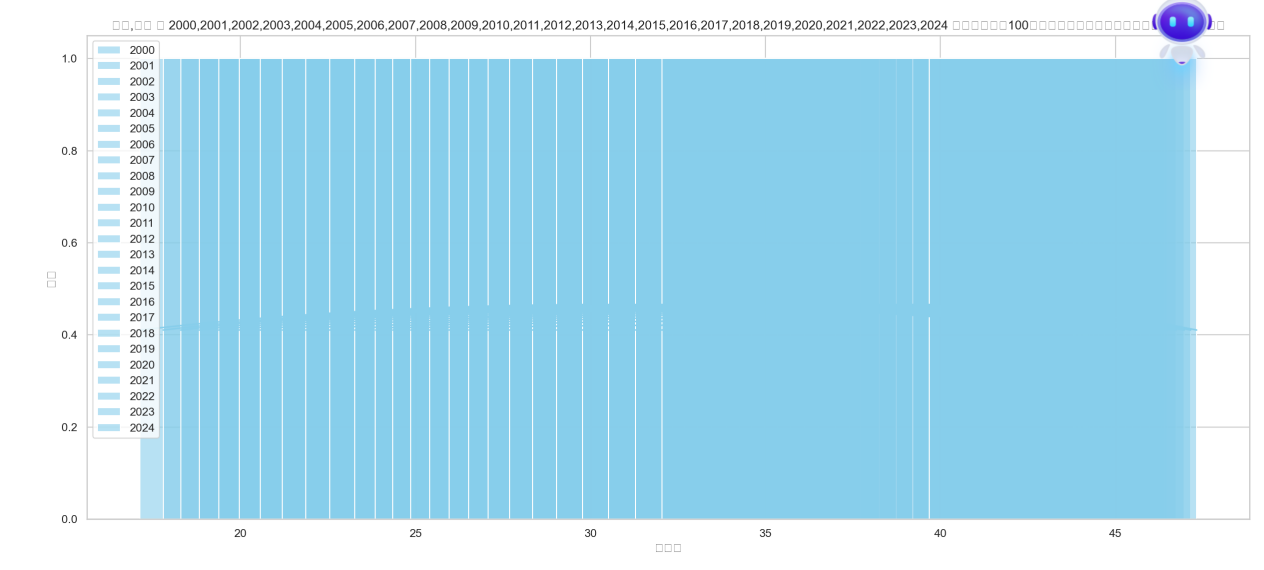
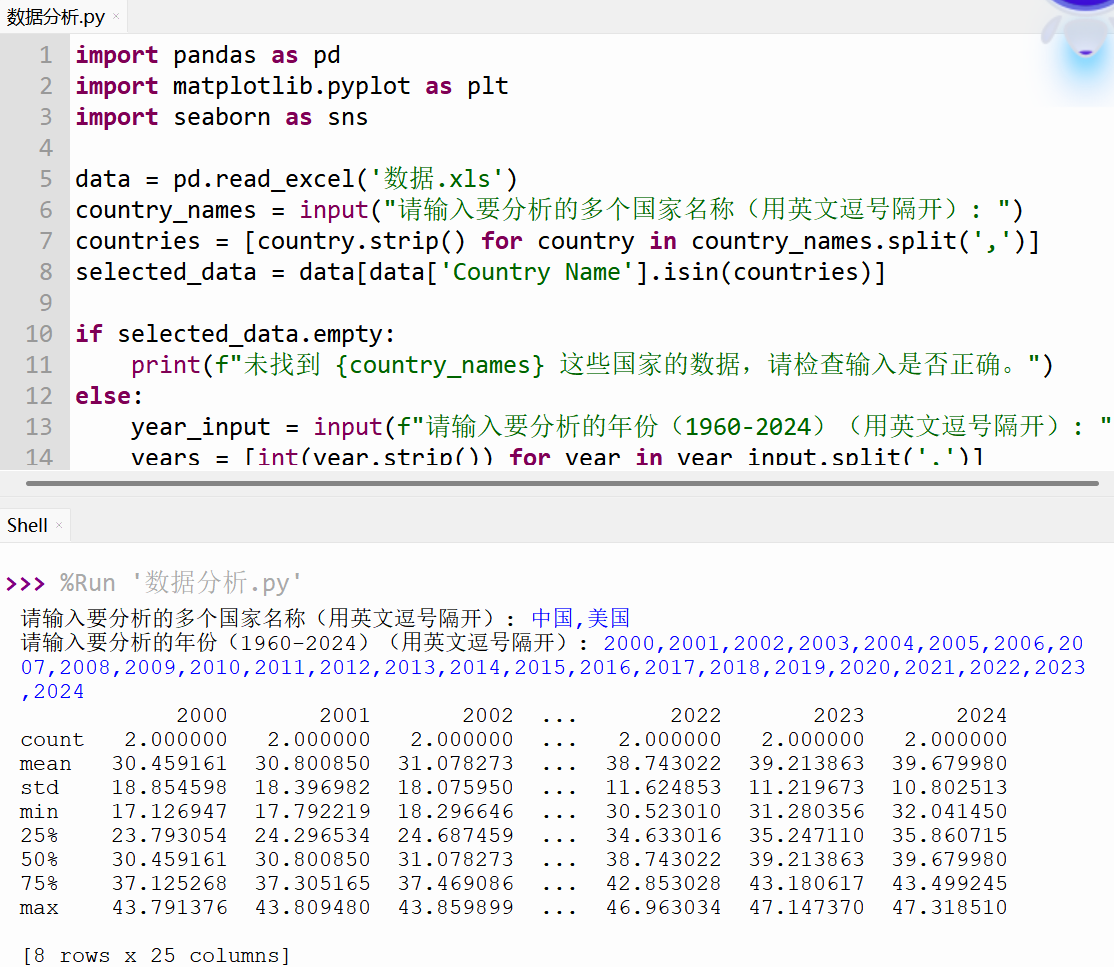
plt.ylabel('百分比')

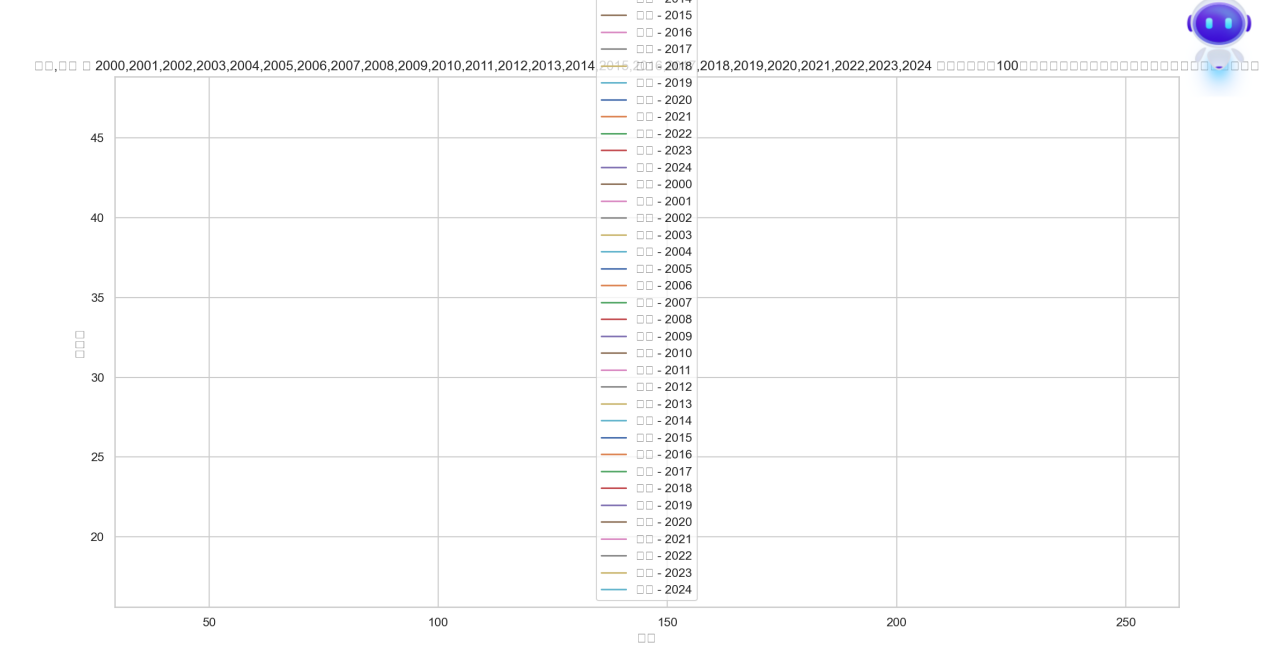
plt.legend()

plt.tight\_layout()

plt.show()

# 代码试运行（以中国和美国为例）





# 五、数据分析结论与建议

1. 从全球范围看，人口在百万以上城市群的占比呈整体增长趋势。但同时，这种增长的地区差异也很显著显著。在高收入国家和地区，像美国、澳大利亚等，该比例起始值较高且增长平稳（2024年美国大城市人口占比达47.32% ，体现其高度城市化）。中低收入国家，如中国、印度等，起始占比低，但增长迅速（中国从1960年的8.12%增长到2024年32.04% ，表明这些国家正处于快速城市化阶段）。非洲部分国家，如布隆迪、尼日尔，占比在研究期内较低，说明城市化发展较为滞后。而部分国家波动变化。
2. 发展建议：
   1. 高收入国家和地区：优化城市资源配置，加强基础设施维护与升级，如纽约、伦敦，要平衡各城区发展，避免过度集中带来的“大城市病”，如交通拥堵、环境污染等。提升城市的可持续发展能力，加大环保投入，发展绿色建筑和公共交通，如新加坡在城市规划中融入环保理念，打造花园城市。
   2. 中低收入国家：加大对城市基础设施建设投入，满足新增城市人口需求，例如印度应改善交通、水电供应等基础设施。制定合理的城市发展规划，引导人口有序聚集，防止城市无序扩张。如中国可以进一步发展新质生产力，促进城市化与工业化协同发展。