**数据结构实验报告14（数字）**

**学号：** 117060400222 **姓名**： 魏紫嫣 **班级：** 17应用统计学2班  **指导老师：** 林卫中

**实验名称**： 学校排名

**实验要求：** 爬出前100名学费高的学校

**实验题目：**

**算法实现：import requests**

**import re**

**from bs4 import BeautifulSoup**

**allUniv=[]**

**def getHTMLText(url):**

**send\_headers = {**

**"User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/61.0.3163.100 Safari/537.36",**

**"Connection": "keep-alive",**

**"Accept": "text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,\*/\*;q=0.8",**

**"Accept-Language": "zh-CN,zh;q=0.8"}**

**try:**

**r = requests.get(url, headers=send\_headers)**

**r.raise\_for\_status()**

**print(r.status\_code)**

**r.encoding = 'utf-8'**

**return r.text**

**except:**

**return ""**

**def fillUnivList(soup):**

**data = soup.find\_all('div',{'class':re.compile('shadow-dark')})**

**for div in data:**

**singleUniv = []**

**div1 = div.find('div',{'style':'margin-left: 2.5rem;'})**

**rank = div1.get\_text().strip()**

**singleUniv.append(rank.split(' ')[0])**

**univName = div.find('h3')**

**singleUniv.append(univName.get\_text().strip())**

**ldiv = div.find\_all('div',{'style':'padding-right: 0.5rem;'})**

**singleUniv.append(ldiv[0].strong.string)**

**singleUniv.append(ldiv[1].strong.string)**

**allUniv.append(singleUniv)**

**def printUnivList():**

**print("{:<6}{:<20}{:<6}{:<10}".format("排名","学校名称","学费","培养规模"))**

**for u in allUniv:**

**print("{:<6}{:<20}{:<10}{:<10}".format(u[0],u[1],u[2],u[3]))**

**def main(n):**

**url = 'https://www.usnews.com/best-colleges/rankings/national-universities?\_page='**

**for i in range(1,n):**

**ri = url + str(i)**

**html = getHTMLText(ri)**

**soup = BeautifulSoup(html,'html.parser')**

**fillUnivList(soup)**

**printUnivList()**

**main(11)**

**实验结果：**