

关于中国人口，你需要关心的问题（二）

原创 闲欢 Python技术 3月12日

中国的人口总数已经突破14亿了，你知道吗？中国人口的年龄结构你了解吗？中国的城市化进程你清楚吗？平均每个成年人要抚养多少个老人孩子你有谱吗？我们接着上篇文章，继续来了解这些人口问题。

获取人口数据

我们的目标是：

获取新中国成立后70年的总人口数据，以及人口年龄结构和抚养比数据。

怎样查看请求以及请求参数的含义上篇文章已经做了详细的介绍，不明白的可以参考上篇文章。我们这里直接上代码：

```
1 # 爬取人口数据
2 def spider_population():
3     # 请求参数 sj（时间），zb（指标）
4     # 总人口
5     dfwds1 = '[{"wdcode": "sj", "valuecode": "LAST70"}, {"wdcode": "zb", "valuecode": "A0301"}]'
6     # 人口年龄结构和抚养比
7     dfwds2 = '[{"wdcode": "sj", "valuecode": "LAST70"}, {"wdcode": "zb", "valuecode": "A0303"}]'
8     url = 'http://data.stats.gov.cn/easyquery.htm?m=QueryData&dbcode=hgnd&rowcode=sj&colcode=zb&wds='
9     # 将所有数据放这里，年份为key，值为各个指标值组成的list
10    # 数据顺序为历年数据
11    population_dict = {
12
13    }
14
15    response1 = requests.get(url.format(dfwds1))
16    get_population_info(population_dict, response1.json())
17
18    response2 = requests.get(url.format(dfwds2))
19    get_population_info(population_dict, response2.json())
20
21    save_excel(population_dict)
22
23    return population_dict
24
25 # 提取人口数量信息
26 def get_population_info(population_dict, json_obj):
27     datanodes = json_obj['returndata']['datanodes']
28     for node in datanodes:
29         # 获取年份
```

```

30     year = node['code'][-4:]
31     # 数据数值
32     data = node['data']['data']
33     if year in population_dict.keys():
34         population_dict[year].append(data)
35     else:
36         population_dict[year] = [int(year), data]
37     return population_dict

```

同样的，我们将获取到的数据存储到 Excel 表格中，将两份数据合成一张表。

```

1  # 人口数据生成excel文件
2  def save_excel(population_dict):
3      # .T 是行列转换
4      df = pd.DataFrame(population_dict).T[:-1]
5      df.columns = ['年份', '年末总人口(万人)', '男性人口(万人)', '女性人口(万人)', '城镇人口(万人)', '乡村人
6                  '15-64岁人口(万人)', '65岁及以上人口(万人)', '总抚养比(%)', '少儿抚养比(%)', '老年抚养比(
7      writer = pd.ExcelWriter(POPULATION_EXCEL_PATH)
8      # columns参数用于指定生成的excel中列的顺序
9      df.to_excel(excel_writer=writer, index=False, encoding='utf-8', sheet_name='中国70年人口数据')
10     writer.save()
11     writer.close()

```

这样，我们就获得了我们所需要的数据，我们打开生成的 Excel 看看：

我们可以看到，及人口年龄结构和抚养比数据在1990年之前基本上是没有统计的，所以我们选取1990那年至2019年的数据来做分析。另外2019年数据也不完整，但是我们可以通过其他项计算出缺失的数据项。

分析人口数据

我准备通过三个方面来分析数据，分别是人口结构分析、抚养比例分析和城镇化进程分析。

人口结构分析

我们的统计数据中，人口结构分为三类：0-14岁人口、15-64岁人口、65岁及以上人口，对应的分别是少儿、成年人和老年人三个年龄段。我们将1990年至今的三个年龄段人口数量放在一个折线图里面，代码如下：

```

1  # 分析人口结构
2  def analysis_struct():
3      # 处理数据

```

```

4     x_data = pdata['年份'].map(lambda x: "%d" % x).tolist()
5     y_data1 = pdata['0-14岁人口(万人)'].map(lambda x: "%.2f" % x).tolist()
6     y_data2 = pdata['15-64岁人口(万人)'].map(lambda x: "%.2f" % x).tolist()
7     y_data3 = pdata['65岁及以上人口(万人)'].map(lambda x: "%.2f" % x).tolist()
8
9     # 人口结构折线图
10    line = Line()
11    line.add_xaxis(x_data)
12    line.add_yaxis('0-14岁人口', y_data1, label_opts=opts.LabelOpts(is_show=False))
13    line.add_yaxis('15-64岁人口', y_data2, label_opts=opts.LabelOpts(is_show=False))
14    line.add_yaxis('65岁及以上人口', y_data3, label_opts=opts.LabelOpts(is_show=False))
15    line.set_global_opts(
16        title_opts=opts.TitleOpts(title="人口结构", pos_bottom="bottom", pos_left="center"),
17        xaxis_opts=opts.AxisOpts(
18            name='年份',
19            name_location='end',
20            type_="category",
21            # axislabel_opts=opts.LabelOpts(is_show=True, color="#000", interval=0, rotate=90),
22            axistick_opts=opts.AxisTickOpts(is_show=True, is_align_with_label=True),
23            axispointer_opts=opts.AxisPointerOpts(type_="shadow", label=opts.LabelOpts(is_show=True)
24        ),
25        # y轴相关选项设置
26        yaxis_opts=opts.AxisOpts(
27            name='人口数（万人）',
28            type_="value",
29            position="left",
30            axislabel_opts=opts.LabelOpts(is_show=True)
31        ),
32        legend_opts=opts.LegendOpts(is_show=True)
33    )
34
35    # 渲染图像，将多个图像显示在一个html中
36    # DraggablePageLayout表示可拖拽
37    page = Page(layout=Page.DraggablePageLayout)
38    page.add(line)
39    page.render('population_struct.html')

```

运行代码，我们看到的图片为：

从图中我们可以得出以下结论：

- 我国的劳动人口（15-64岁）人数从2010年后开始逐步降低。
- 我国的少儿（0-14岁）人数从2010年之后呈现缓慢增长态势，但是增长率不明显。
- 我国的老年（65岁及以上）人数一直上升，并且近些年增长率有明显扩大趋势。

综合来看，我国劳动力人口在减少，老年人口在增加，说明我国正在向老龄化社会迈进。同时劳动力人口减少意味着我

们的人口红利快到尽头了，接下来需要我们向着高质量高效率的方向发展。

抚养比例分析

抚养比例是指非劳动力人口占劳动力人口的比例。通常用百分比表示。说明每100名劳动年龄人口大致要负担多少名非劳动年龄人口。用于从人口角度反映人口与经济发展的基本关系。

抚养比例一般从三个数据衡量，分别是少儿抚养比、老年抚养比和总抚养比。其中总抚养比等于少儿抚养比和老年抚养比之和。

我们我折线图来展现这三个数据趋势：

```
1  # 分析抚养比例
2  def analysis_raise():
3      # 处理数据
4      x_data = pdata['年份'].map(lambda x: "%d" % x).tolist()
5      y_data1 = pdata['总抚养比(%)'].map(lambda x: "%.2f" % x).tolist()
6      y_data2 = pdata['少儿抚养比(%)'].map(lambda x: "%.2f" % x).tolist()
7      y_data3 = pdata['老年抚养比(%)'].map(lambda x: "%.2f" % x).tolist()
8
9      line = Line()
10     line.add_xaxis(x_data)
11     line.add_yaxis('总抚养比(%)', y_data1, label_opts=opts.LabelOpts(is_show=False))
12     line.add_yaxis('少儿抚养比(%)', y_data2, label_opts=opts.LabelOpts(is_show=False))
13     line.add_yaxis('老年抚养比(%)', y_data3, label_opts=opts.LabelOpts(is_show=False))
14     line.set_global_opts(
15         title_opts=opts.TitleOpts(title="人口抚养比例", pos_bottom="bottom", pos_left="center"),
16         xaxis_opts=opts.AxisOpts(
17             name='年份',
18             name_location='end',
19             type_="category",
20             # axislabel_opts=opts.LabelOpts(is_show=True, color="#000", interval=0, rotate=90),
21             axistick_opts=opts.AxisTickOpts(is_show=True, is_align_with_label=True),
22             axispointer_opts=opts.AxisPointerOpts(type_="shadow", label=opts.LabelOpts(is_show=True)
23         ),
24         # y轴相关选项设置
25         yaxis_opts=opts.AxisOpts(
26             name='抚养比例(%)',
27             type_="value",
28             position="left",
29             axislabel_opts=opts.LabelOpts(is_show=True)
30         ),
31         legend_opts=opts.LegendOpts(is_show=True)
32     )
33
34     # 渲染图像，将多个图像显示在一个html中
35     # DraggablePageLayout表示可拖拽
36     page = Page(layout=Page.DraggablePageLayout)
37     page.add(line)
```

```
38 page.render('population_raise.html')
```

运行代码，我们可以看到图像如下：

从图中我们可以得出以下结论：

- 最近10年少儿抚养比例在缓慢增加。
- 最近10年老年抚养比例在显著上升，并且增加幅度逐年上升。
- 总的抚养比例已经达到40%多了。

综上所述，随着我国逐渐进入老龄化社会，我国的劳动力人口在下降，而老年人口在上升，所以平均每个劳动力人口负担的抚养比例在上升。后续我们的成年人家庭负担会越来越重，看到这里突然有点鸭梨山大有没有？

城镇化进程分析

城镇化进程是指一个农业人口转化为非农业人口、农业地域转化为非农业地域、农业活动转化为非农业活动的过程。简单理解就是城镇人口占总人口的比例，叫做城镇化（也叫城市化）比例。

我们将城镇人口和农村人口用叠加柱状图来表示，代码如下：

```
1 # 分析城镇化比例
2 def analysis_urban():
3     x_data = pdata['年份'].map(lambda x: "%d" % x).tolist()
4     # total = pdata['年末总人口(万人)'].map(lambda x: "%.2f" % (x / 1000)).tolist()
5     y_data1 = pdata['城镇人口(万人)'].map(lambda x: "%.2f" % (x / 1000)).tolist()
6     y_data2 = pdata['乡村人口(万人)'].map(lambda x: "%.2f" % (x / 1000)).tolist()
7
8     # 城镇化比例
9     # y_data_rate = pdata['城镇人口(万人)'] * 100 / pdata['年末总人口(万人)']
10
11     bar = Bar()
12     bar.add_xaxis(x_data)
13     bar.add_yaxis("城镇人口", y_data1, stack="stack1", category_gap="10%")
14     bar.add_yaxis("乡村人口", y_data2, stack="stack1", category_gap="10%")
15     bar.set_series_opts(label_opts=opts.LabelOpts(is_show=True, position="inside", rotate=90))
16     bar.set_global_opts(
17         title_opts=opts.TitleOpts(title="中国城镇化进程"),
18         xaxis_opts=opts.AxisOpts(
19             name='年份',
20             name_location='end',
21             type_="category",
22             # axislabel_opts=opts.LabelOpts(is_show=True, color="#000", interval=0, rotate=90),
23             axistick_opts=opts.AxisTickOpts(is_show=True, is_align_with_label=True),
24             axispointer_opts=opts.AxisPointerOpts(type_="shadow", label=opts.LabelOpts(is_show=True)
25         ),
```

```

26     # y轴相关选项设置
27     yaxis_opts=opts.AxisOpts(
28         name='人口数(千万人)',
29         type_="value",
30         position="left",
31         axislabel_opts=opts.LabelOpts(is_show=True)
32     ),
33     legend_opts=opts.LegendOpts(is_show=True)
34 )
35
36 # 渲染图像，将多个图像显示在一个html中
37 page = Page(layout=Page.DraggablePageLayout)
38 page.add(bar)
39 page.render('population_urban.html')

```

运行上面代码，我们得到的图像如下：

从图中我们可以看出，伴随着我国的城市化进程的推进，农村人口不断向城市迁移，截至2019年，我国城镇人口达到8.8亿，已经占总人口的62.8%了。而从发达国家的城市化比例来看，普遍达到70%之后才出现逆城市化，所以我国的城市化还有空间，城镇人口还有增长空间。另一方面，人口迁移过程中，也会进一步促进城市经济发展。而对于房价来说，能够吸引劳动力的城市会得到支撑，因为人口和经济双增。

总结

本文从国家数据官方网站获取到中国人口相关的数据，然后通过图标将人口结构、人口抚养比例以及中国城镇化进程等几个方面展现出来，从而可以直观地看到我国人口的发展情况。

文中示例代码：[python-100-days](#)

PS： 公号内回复「Python」即可进入 Python 新手学习交流群，一起 **100天计划**！

-END-

Python 技术
关于 Python [都在这里](#)