

# 南开大学

## C++程序设计大作业论文

### “天气衣指南”——基于python的生活指数网站构建和天气信息可视化



程序设计基础 0920 班第三小组

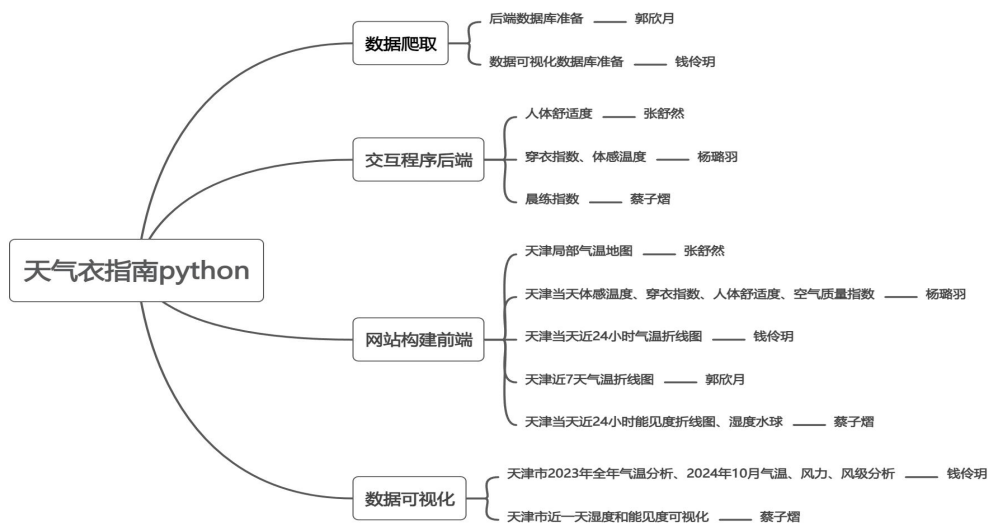
2024 年 12 月

# 目录

小组成员以及组内分工 .....	3
一、 摘要 .....	4
二、 绪论 .....	6
2.1 研究背景 .....	6
2.2 研究现状 .....	6
三、提出问题与分析问题 .....	7
3.1 选题原因 .....	7
3.2 选题意义 .....	7
3.3 分析问题 .....	7
四、解决问题 .....	9
4.1 数据爬取以及数据预处理 .....	9
4.1.1 数据爬取 .....	9
4.1.2 数据预处理 .....	9
4.1.3 数据存储 .....	9
4.2 后端公式选择与代码编写 .....	10
4.2.1 人体舒适度 .....	10
4.2.2 衣物厚度 .....	10
4.2.3 体感温度 .....	11
4.2.4 晨练指数 .....	12
4.3 前端网页程序编写 .....	13
4.4 数据分析可视化 .....	14
4.4.1 天津 2023 年全年度气温分析 .....	14
4.4.2 天津 2024 年 10 月气温、风力、风级数据分析 .....	15
4.4.3 天津湿度数据分析 .....	18
4.4.4 天津可见度数据分析 .....	18
五、效果评价以及总结 .....	20
六、 参考文献 .....	22

## 小组成员以及组内分工

小组成员	工作
钱伶俐（组长）	总体方向把控，数据爬取查找，数据整理，文献整体搜索，python 数据可视化分析（天津 2023 年气温分析、天津 2024 年 10 月气温、风力、风级可视化），网站天津市 24 小时气温折线图代码构建，报告展示以及答辩
郭欣月	数据爬取查找，数据整理，搜索 python 爬虫以及天气情况文献资料，网站天津市近 7 天气温折线图代码构建，报告展示以及答辩
杨璐羽	相关数据整理与量化计算，文献搜集与评价，穿衣指数相关代码设计，网站天津市体感温度、穿衣指数、人体舒适度、空气质量指数代码编写，报告展示以及答辩
张舒然	数据资源的整合和量化分析，搜索中英文文献，相关文献总结，人体舒适度相关代码设计，网站天津市局部气温地图代码编写、报告展示以及答辩
蔡子熠	文献检索与解读，相关数据爬取，晨练指数相关代码构建，python 可视化(天津 11 月 20 日逐小时湿度、能见度)，网站天津当天近 24 小时能见度折线图、湿度水球构建，报告以及展示答辩



## 一、 摘要

本文旨在帮助天津地区居民灵活应对天气变化, 并为其提供直观的生活指南, 本文主要介绍使用 Python 构建综合个性化天气网站的详细全过程。

首先通过 python 获取天津地区过去一年及特定月份的天气数据, 在此基础上运用准确、实时的数据获取技术、科学的计算处理技术以及多种数据可视化手段, 最后构建涵盖天气预报、人体舒适度、晨练指数、体感温度等多样、精确、科学的数据信息以及衣物厚度分析等具有实用性的功能的天气网站。

此网站为用户提供直观、详细、实用且仅针对于天津地区的天气指南服务, 从而进一步有效地为天津地区的用户提供合理规划日常生活与出行安排的建议。

**关键词:** 天津地区天气网站; 生活指数; Python; 数据可视化

# Abstract

This article aims to help residents in Tianjin area flexibly respond to weather changes and provide them with intuitive living guidelines. It mainly introduces the whole process of building a comprehensive personalized weather website by using Python.

Firstly, Python is used to obtain the weather data of Tianjin area in the past year and specific months. Based on these, accurate and real-time data acquisition techniques, scientific calculation and processing techniques as well as multiple data visualization methods are applied. Finally, a weather website with diverse, precise and scientific data information such as weather forecasts, human comfort levels, morning exercise indexes, perceived temperatures, etc., and practical functions like analysis of clothing thickness is constructed.

This website provides users with intuitive, detailed and practical weather guide services specifically for Tianjin area, thus further effectively offering suggestions for users in Tianjin area to reasonably plan their daily lives and travel arrangements.

**Key words:** Tianjin area weather website; living indexes; Python; data visualization

## 二、绪论

### 2.1 研究背景

随着互联网技术的飞速发展，数字化时代已经来临，信息获取的便捷性成为衡量社会进步的重要标志。在众多信息中，天气信息与人们的日常生活息息相关。无论是日常出行、健身锻炼，还是衣物选择，天气条件都起着关键的决定性作用。为了满足人们对天气信息的需求，各类天气服务网站应运而生。在此背景下，利用 Python 编程语言构建个性化、专业化的天气服务网站具有重要的现实意义。

Python 作为一种功能强大的编程语言，其在数据处理与网络开发领域的应用日益广泛。Python 的高效性、简洁性以及丰富的第三方库，为天气数据挖掘、处理与可视化呈现提供了坚实的技术支撑。在此基础上，本研究旨在深度整合天津地区的天气数据资源，打造一个功能完备、界面友好的天气服务网站，以满足用户对天气信息及其衍生生活指数的精准需求。

### 2.2 研究现状

目前，国内外天气服务网站的发展呈现出以下特点：

- (1) 数据丰富：天气服务网站普遍提供了丰富的天气数据，包括实时天气、历史天气、天气预报等。
- (2) 功能多样：部分天气服务网站已开始尝试提供与生活紧密关联的天气衍生信息，如人体舒适度、晨练指数等。
- (3) 可视化展示：天气数据可视化技术逐渐成熟，使得天气信息呈现更加直观、易懂。

然而，现有天气服务网站在以下几个方面仍存在不足：

- (1) 个性化服务不足：大部分网站未能针对用户需求提供个性化的天气信息服务。
- (2) 数据挖掘与处理能力有待提高：虽然天气数据丰富，但如何从中挖掘出有价值的信息仍需进一步研究。
- (3) 界面友好性不足：部分天气服务网站界面设计较为复杂，不利于用户快速获取所需信息。

## 三、提出问题与分析问题

### 3.1 选题原因

在日常生活中，天气对人们的出行和着装有着直接的影响。随着信息技术的飞速发展，大量的天气数据可供获取。Python 作为一种强大的编程语言，具备丰富的数据处理和分析库，能够高效地处理天气数据。通过 Python 处理天气数据并生成穿衣建议，用户能够提前知晓天气对应的穿衣搭配，无需在出门前匆忙决定，节省时间且确保穿着适宜。我们希望为用户提供一款能够根据实时天气数据，快速有效地给出建议穿着；同时可以为不同人群提供便捷、个性化生活服务的网站。

### 3.2 选题意义

从实用的视角出发，它能够把难以精准感知的气温数据转化为直观的穿衣指南，切实助力人们更妥善地应对天气变化，有效规避因着装不适而诱发的健康隐患。

在科技应用方面，此应用充分彰显了 Python 在数据处理范畴的实用价值，有力推动数据处理技术融入日常场景并得以广泛普及。

在商业领域，这类应用可被整合至智能穿戴设备或手机应用程序之中，从而为用户赋予增值服务体验，显著提升产品的魅力与竞争力。

在数据科学教育层面，该选题为学习者呈献了一个完备的实操案例，对深入领悟数据的获取、清理、分析以及应用流程大有裨益，利于培育解决实际问题的能力，推动数据科学知识的广泛传播与持续发展，使更多人真切体会到数据处理技术为生活品质优化所带来的积极效能。

### 3.3 分析问题

对于以上提到的问题，我们需要每一天的天气状况以及详细的数据。由于天气数据大多不会以直接的数据表的形式呈现给我们，因此，在数据上，我们选择 Python 中的库 BeautifulSoup 进行网页解析和数据爬虫，并使用 pandas 库以 csv 的方式进行数据存储。从实时数据而言，我们选择了历史天气数据网，虽然这个网站命名以历史为名，但是实际上该网站拥有丰富的实时天气资料，适合爬虫获取（图 1）；从历史天气数据而言，我们选择了天气网进行数据的获取。（图 2）

日期时间	天气	天气逐小时降水量	湿度	aqi	气压	风向	风力	风速	能见度
2024/12/6 16:00	晴	5.1℃	0mm	33%	23 1024hPa	北风	2级	8km/h	30km
2024/12/6 15:00	晴	6℃	0mm	31%	24 1023hPa	北风	3级	14km/h	30km
2024/12/6 14:00	阴	6.4℃	0mm	30%	27 1023hPa	北风	3级	16km/h	30km

图 1：动态天气数据爬取数据实例（12 月 6 日当天 17 时爬取的数据）

日期	天气情况	最高气温	最低气温	风力	风级
2024年10月1日	晴/晴	20	10	西北风	3-4级
2024年10月2日	多云/晴	19	10	北风	1-3级
2024年10月3日	晴/晴	22	12	西南风	1-3级
2024年10月4日	晴/晴	23	12	西南风	1-3级
2024年10月5日	多云/阴	24	16	西南风	1-3级
2024年10月6日	阴/阴	21	13	东北风	1-3级
2024年10月7日	晴/晴	20	12	北风	1-3级
2024年10月8日	晴/晴	23	15	南风	1-3级
2024年10月9日	阴/阴	22	14	西南风	1-3级
2024年10月10日	晴/晴	20	11	北风	3-4级
2024年10月11日	晴/晴	23	12	西南风	1-3级
2024年10月12日	晴/多云	26	17	东南风	1-3级
2024年10月13日	阴/阴	21	16	东风	1-3级
2024年10月14日	多云/小雨	25	16	西南风	1-3级
2024年10月15日	多云/多云	20	12	东北风	3-4级
2024年10月16日	多云/阴	21	14	东南风	1-3级
2024年10月17日	阴/小雨	23	16	东南风	1-3级
2024年10月18日	小雨/晴	18	6	东北风	3-4级
2024年10月19日	晴/多云	10	4	东北风	3-4级
2024年10月20日	阴/小雨	12	6	东北风	1-3级
2024年10月21日	多云/阴	14	8	北风	1-3级
2024年10月22日	阴/晴	14	5	北风	3-4级
2024年10月23日	晴/晴	18	9	西南风	1-3级
2024年10月24日	晴/晴	21	10	东南风	1-3级
2024年10月25日	晴/多云	23	11	东南风	1-3级
2024年10月26日	雾/阴	20	12	东南风	1-3级
2024年10月27日	多云/晴	17	8	东北风	1-3级
2024年10月28日	晴/多云	19	9	西南风	1-3级
2024年10月29日	多云/晴	19	9	西南风	1-3级
2024年10月30日	阴/阴	19	12	东风	1-3级
2024年10月31日	阴/阴	20	12	西风	1-3级

图 2：2024 年 10 月历史天气状况爬取数据实例

与此同时，由于天气状况的数据较为模糊，我们需要具体的量化参数对于不同的天气状况进行描述，于是我们选择了不同的日常生活指数进行计算，并将指数转化为相应的小建议和可视化的方式输出给用户。

更进一步而言，除了在当下的数据中获得及时的数据和指南以外，我们希望能从过去的天气数据中获得对于天气状况的大方向把控。为了实现这个目标，我们选择使用天津过去一年的气温状况绘制气温折线图，以直观地显示环境温度随时间的变化。10 月作为承接天气炎热的 8 月 9 月，迎接接下来的降温的 11 月，的“过渡月”，我们同时也希望通过分析 10 月的天气数据如气温、风力、风级，从而给出一份“过渡月”的参考。

为了能够呈现给用户一个具象化的结果，我们采用了制作网站的方式对于前端进行构建。在原先的指数的基础上，我们添加了天气地图、水球、仪表盘、折线图等图案，让数据更加可读。

总而言之，为了实现对于天津市天气状况的可视化，我们以“数据准备-指数量化-历史数据分析-网站前端构建”四个方面来分析解决问题的过程。对于以上的四个方面，我们选择了具备爬虫、函数计算、可视化以及网站构建功能的 Python 程序语言来实现上述过程。



## 四、解决问题

### 4.1 数据爬取以及数据预处理

#### 4.1.1 数据爬取

基于我们所需要的天津各区的实时和历史天气数据，并综合考虑数据的种类多样性、数据科学性、准确性以及时效性，我们选择天气网的天津天气预报和历史天气数据网作为目标网站。前者作为天津近 7 天天气的数据来源，而后者是天津各区逐小时天气数据的来源。

前端数据获取的过程主要涉及 python 的三个库，分别是：requests 库, BeautifulSoup 库, pandas 库。具体过程阐释如下：

首先我们利用 `getHTMLtext` 函数获取网页内容，用于向指定的 URL（也就是目标网站）发送 GET 请求，以获取网页的 HTML 内容。具体实现方式如下：在函数内部使用 requests 库向指定的 URL 发送 HTTP 请求，并设置了超时时间为 30 秒，若请求成功，则根据网页内容自动判断编码并设置编码格式（通过 `apparent_encoding` 属性），最后返回获取到的 HTML 文本；若请求过程出现异常（比如网络问题、URL 错误等），则打印“访问错误”信息，并返回空字符串。通过以上过程可以在程序运行过程中清晰地判断是否网站访问成功，这一步是后续用于解析的数据基础。

#### 4.1.2 数据预处理

下一步基于已经获取到的网页 HTML 文本，我们要深入解析其中的天气相关内容，精准提取出各项我们关注的天气数据，完成数据的初步整理工作。于是，我们选择使用了 `get_content` 函数帮助我们接收 HTML 文本以及一个表示要获取天数的参数 `num`（在完整使用场景中用于控制获取未来几天的天气数据）。在函数内部，首先使用 BeautifulSoup 解析（这里使用的解析器是 `html.parser`）传入的 HTML 内容，然后通过层层查找 HTML 标签元素的方式定位到存放天气信息的相关节点。对于每一天的天气数据，它会进一步提取日期、天气状况、最低气温、最高气温、风向以及风级等具体信息，并将这些信息整理成一个列表，最后将所有天的信息列表汇总到一个总的列表 `final` 中返回。

#### 4.1.3 数据存储

最后一步，我们使用 `get_weather_7` 作为一个整合函数，它将获取 HTML、解析提取数据以及整理成规范格式这一系列流程串起来，最终达到输出方便使用和处理的天气数据表格形式的效果。从功能上看，`get_weather_7` 函数不光包含了前面两个函数的功能，额外使用了 pandas 库将天气数据列表转换为带有指定

列名（['日期', '天气', '最低气温', '最高气温', '风向 1', '风向 2', '风级']）的 DataFrame 格式数据，返回这个结构化的天气数据 DataFrame。

这一步我们选择使用 DataFrame 而非 CSV 文件基于以下几点考虑：其一，从数据结构的本质上进行对比：DataFrame 类似于二维表格，有行索引和列索引。可以看作是一个带有标签（索引）的多维数组，而这些标签使得数据的选取、过滤和操作更加方便。而 CSV 则是一种纯文本的文件格式，用于存储表格数据。它本质上是一种简单的文本格式，数据按行存储，每行中的不同数据字段（列）之间通常用逗号隔开。所以它没有像 DataFrame 那样复杂的数据结构，只是简单的文本内容，没有内置的索引等信息。其二：从使用功能上比较：DataFrame：支持丰富的数据转换、计算、交互等功能。CSV 功能仅限于数据的存储和交换。其三：对于应用场景而言，DataFrame 是核心的数据存储和处理工具。而 CSV 则更适合用于数据的长期存储和数据共享。故此，基于以上几点，我们最终将储存 CSV 文件的程序修改为储存为 DataFrame。

## 4.2 后端公式选择与代码编写

### 4.2.1 人体舒适度

人体舒适度的计算综合考虑多个天气因素，采用以下常见公式：

$$ICHB = (1.8t + 32) - 0.55 * (1 - hu/100) * (1.8t - 26) - 3.2 * \sqrt{v}$$

其中，ICHB 表示人体舒适度指数，t 为气温（℃），hu 为相对湿度（%），v 为风速（m/s）。

根据采集到的气温，湿度与风速数据，代入公式计算得出人体舒适度指数。

将计算结果根据指数范围划分为不同等级，并标注感觉程度（寒冷，极不舒适；偏冷，大部分人感觉不舒适；偏凉，少部分人感觉不舒适；凉爽，大部分人感觉舒适；最舒适，绝大部分人感觉很舒适；暖和，大部分人感觉舒适；偏热，少部分人感觉不舒适；炎热，多数人感觉不舒适；酷热，感觉很不适）并在网站上以直观的文字描述或对应的表情图标形式呈现给用户。

通过这个方式，我们能够帮助用户快速判断当前天气条件下人体的实际舒适感受程度，以便用户更好地调整室内环境参数（如空调温度）或安排户外活动计划。同时，能够为天气、建筑、旅游等多个领域提供有价值的人体舒适度信息，有助于提升人们对环境的适应和应对能力，进一步改善人们的生活质量和工作效率。

### 4.2.2 衣物厚度

穿衣指数定义为在自然环境条件下，天空状况、气温、相对湿度及风等最主要的天气条件对人体感觉温度的影响，从而对人们适宜穿着的服装分级，以提醒人们根据天气变化适当着装。

在这里，我们引进依据前人实验数据结果再加以推导出的经验公式：

$$Y = \frac{0.61(25.8 - X)}{1 - 0.01165 V^2}$$

Y 为室外环境所需服装厚度预报值（单位：mm），x 为环境温度预报值，v 为环境风力预报值。从公式可以导出，当风力不变时，着装厚度随气温的升高而变薄。风力较小时，着装厚度随风力的变化很小；当风力变大后，着装增厚的幅度变大。

通过分析围场近 30 年的天气数据资料可知：围场地区的平均气温的全年穿衣指数在 1-5 月之间，最高气温的穿衣指数只有在 7 月份小于 0，而最低气温的穿衣指数在 1 月~3 月、11 月~12 月均在 5 以上。

通过统计、观察，围场最低气温的穿衣指数均在八级以上，天气最冷时需要穿棉衣、冬大衣、皮夹克、厚呢外套、呢帽、手套、羽绒服、裘皮大衣，夏季适合穿衬衫、薄长裙、T 恤的日子只有 7 月份。全年平均的穿衣指数在四~七级之间，冬季穿衣较厚，多在 8 级，夏季最热时也在 2-3 级之间。

围场地区近 30 年来，7 月份最高气温的穿衣指数，呈下降趋势。7 月份天气最热时，人们需要穿的衣服虽在逐年减少，但穿衣指数值均大于-1，只适合穿棉麻面料的衬衫、薄长裙、T 恤。

围场地区近 30 年来，1 月份最低气温的穿衣指数，呈上升趋势。1 月份天气最冷时，人们所需穿的衣服在逐年增厚，且均超过 8 级，只有穿棉衣、冬大衣等最厚实的衣服才能度过围场的漫漫冬季。围场冬季穿衣指数成提升的趋势，而夏季穿衣指数呈下降趋势，即日后冬季穿衣厚度在逐年提高，夏季穿衣厚度在逐年减少。

穿衣指数是天气部门根据自然环境对人体感觉温度影响最主要的天空状况、对人们适宜穿着的服装进行分级提示。了解穿衣指数可以帮助我们预防疾病、提高舒适度、适应不同的场合、便于安排活动。提高生活效率，例如减少选择时间、合理规划购物。

#### 4.2.3 体感温度

体感温度，是指人体感受空气的温度，与实际环境的温度可以有出入。这是因为其他与天气相关的状况，使空气和人体之间的热对流受到影响，从而使人感觉的度与实际有分别，因此又名表观气温。空气对热的吸收会受到相对湿度及其密度影响；而风速会影响到与人体表面可以接触到的空气的分量，当风速增加时，

与人体所接触的空气会增加，所以其所带走或带来的热量亦相应地增加。这里可以参考台湾中央气象局目前所采用的计算公式：

$$\text{体感温度} = 1.07 * T + 0.2 * e - 0.065 * v - 2.7$$

$$e = \frac{RH}{100} \times 6.105 \times \exp\left(\frac{17.27 \times T}{237.27 + T}\right)$$

其中，T 为气温（℃）、e 为水汽压（hPa）、V 为风速（m/sec）、RH 为相对湿度（%）。

计算水汽压通常需要知道空气的温度和相对湿度，而不是直接从气压中计算。不过，通过已知的气压、温度和相对湿度，我们可以间接计算出水汽压。具体步骤如下：

1. 饱和水汽压：首先，需要根据温度计算出饱和水汽压（ $e_0$ ），这是在特定温度下，空气能够容纳的最大水汽量对应的水汽压。这通常使用经验公式如 Goff-Gratch 公式来计算，该公式考虑了温度的影响。

2. 相对湿度：相对湿度 (RH) 是实际水汽压（ $e_a$ ）与饱和水汽压（ $e_0$ ）的比值， $RH = (e_a / e_0) * 100\%$ 。

3. 计算实际水汽压：知道了饱和水汽压和相对湿度后，可以通过以下关系计算实际水汽压： $e_a = RH / 100 * e_0$

体感温度是指人体所感受到的冷暖程度，它综合考虑了气温，湿度，风速等因素对人体的影响。了解体感温度对我们有以下优点：合理安排日常生活，例如穿着合适，安排适宜的活动；健康防护，例如预防疾病；提高生活效率，在办公或户外环境调整着装，优化工作和生活安排，选择合适的出行方式和出行装备等。

#### 4.2.4 晨练指数

晨练指数是一个用于指导公众进行早晨锻炼的健康指数。它主要考虑了气象因素对户外运动的影响，为市民提供早晨锻炼时的天气状况参考。晨练指数是根据早晨的天气条件（如温度、湿度、风力、空气质量等）和人体舒适度来综合判断早晨是否适宜进行户外锻炼的指标。通过晨练指数，人们可以根据实际情况调整自己的锻炼计划，选择最适宜的时间和方式进行身体锻炼，从而更好地保护身体健康。

晨练指数的计算综合考虑多个天气因素，采用以下常见公式：

$$MI = MIT + MIH + MIV + MIN$$

其中，MI 为晨练指数，无量纲；MIT 为气温范围取值，无量纲；MIRH 为相对湿度范围取值，无量纲；MIV 为最大风速范围取值，无量纲；MIN 为天气现象范围取值，无量纲。

1. MIT 值的确定：将 08 时的气温划分区间，对应不同的 MIT 值
2. MIN 值的确定：将爬取的天气数值正则化，因为是晨练指数，故只保留“转”前面的天气值，舍去“转”后面的天气值，从而分别对应不同的 MIN 值
3. MIRH 值的确定：将爬取的 08 时的相对湿度划分区间，分别对应不同的 MIRH 值
4. MIV 值的确定：将爬取的 08 时的风速划分区间，分别对应不同的 MIRH 值。

最后根据公式，将上述值加起来，可求得 MI 值，每个 MI 区间分别对应不同

适合晨练的程度，可划分为不同的晨练等级。

晨练指数是一个综合反映早晨气象条件对户外运动适宜程度的指标，它考虑了温度、湿度、风力、空气质量以及天气状况等因素，以不同四个等级来评价早晨是否适宜进行户外锻炼，旨在指导公众根据实时气象信息合理安排晨练活动，保障健康。

### 4.3 前端网页程序编写

为了让我们的后端能够更加直观地展现在用户的面前，我们采取了编写网页的方式。主要功能包括：天气地图，湿度、空气污染指数、体感温度、穿衣指数（衣物厚度）的指数显示，未来一周气温变化图像，24 小时气温变化图像，晨练指数和人体舒适度变化图像。

要实现让后端的代码能够应用到前端，我们首先建立了名为 `data_7` 的程序爬取 7 天以内的天气、气温、风向、风级数据，用作预备；而后建立了 `end_version` 的程序，用于爬取人体舒适度和晨练指数相关的数据以及其他指数数据如空气污染指数等，并定义了人体舒适度和晨练指数的函数公式。

为了能够创建有交互性的 web 可视化，我们选择了 `pyecharts` 渲染为 html 文件。在可视化 `py` 程序中，我们先整理了所需要的天气数据信息，并添加了穿衣指数、体感温度的函数计算，而后利用 `pyecharts` 绘图。通过 `options` 配置图标的各个属性，`themetype` 定义整体网页风格。

我们希望先给用户一个直观的天津整体气温的情况，于是我们首先选择在网站最上方构建“天气地图”并添加鼠标悬停显示数据的功能，让网页更加具有互动性。由于 `pyecharts` 内置的 `Map` 类别已经内置了天津地图，因此我们只需调整地图显示的参数即可。



若单个数据如湿度、空气污染指数、体感温度、穿衣指数直接以数值的方式存放在网页中，会减少整体美观度，因此我们使用水球图（Liquid）和仪表盘（Gauge）来让数值更加具有吸引力，将四个图用 Grid 模块进行网页聚合。

对于七天天气预报折线图、24 小时气温折线图、晨练指数、人体舒适度指数、能见度的折线图，我们采用了 Line 类别，读取 data\_7 和 end\_version 中的数据，绘制图像，并对于横纵轴以及呈现方式进行调整，最后进行网页聚合。

## 4.4 数据分析可视化

### 4.4.1 天津 2023 年全年气温分析

我们借助 Python 的数据可视化库 Matplotlib，绘制天津过去一年的气温变化可视化图表，希望能够对于新来天津的人们一些在气温上面的参考，从而及时做出调整以适应当地气温。

首先，我们通过 Python 爬虫爬取了 2023 年 1 月-2023 年 12 月前十天的气温数据（包括每日最高气温和最低气温），并把它存储在 csv 文件中，作为数据库的预备。

而后，我们从数据库中读取数据，以时间作为横坐标，气温数值作为纵坐标，运用 Matplotlib 的绘图函数绘制折线图。其中红色折线表示最高气温随时间的变化趋势，蓝色折线表示最低气温的变化情况。（如图 3 所示）通过设置图表标题、坐标轴标签、图例以及调整画布大小，以增强图表的可视化效果，让用户能够直观地感受天津全年气温的起伏变化规律、季节差异以及极端气温分布情况，为用户提供全面深入的气温变化概览信息。

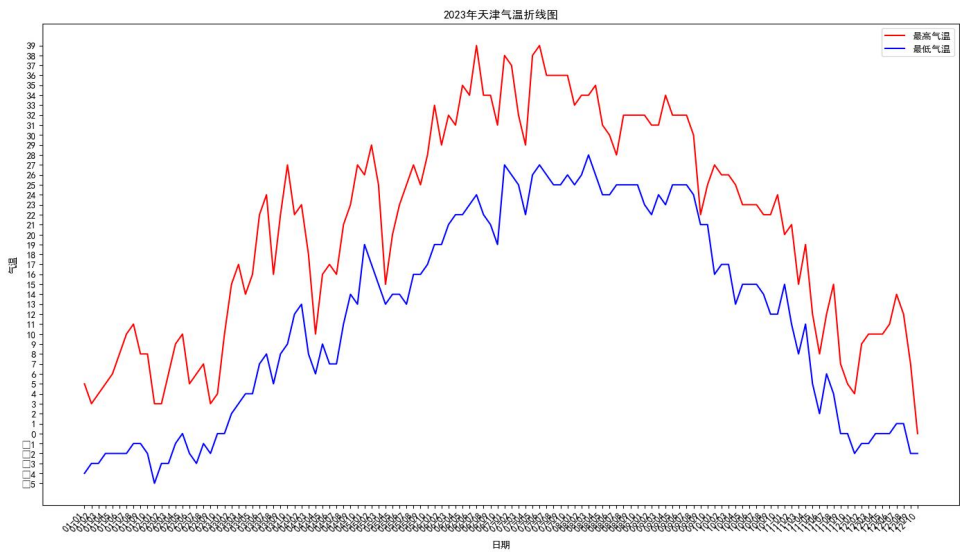


图 3：2023 年天津气温折线图

从图中，我们可以知道：

整体上：

最高气温和最低气温大概变化趋势相同；以 15 摄氏度为界限，1-3 月、11-12 月气温较低；7-8 月气温最高；4-5 月份迎来了一次特殊降温。

部分上：

最高气温范围在 0-40 摄氏度之间；在 2、4 月份迎来降温，从 3 月中旬的 27 摄氏度降温到了 4 月初的 11 摄氏度；与此同时，在 5 月份也有一次跨越将近 10 摄氏度的降温；自 5 月之后开始持续升温；从 8 月开始逐日气温下降；但总体上气温都在零上。

最低气温范围在-5-27 摄氏度之间；从 1 月到 7 月总体上呈现逐日上升的趋势；而后从 8 月起，气温逐日下降。

从以上图表的分析中，我们可以得出一个结论：在最高温度上，相较于长江以南地区的夏日 40 度高温，天津的夏天较为宜人；在最低温度上，会有零下温度出现，但是出现日期不多。在 2、4、5、8、9 月有温度骤变，人们在以上月份需要注意气温的变化。

#### 4.4.2 天津 2024 年 10 月气温、风力、风级数据分析

天津的 10 月是从 8、9 月的“秋老虎”向 11 月的寒冷的过渡月，我们希望通过分析天津 10 月的天气状况进行数据分析，从而得出合理的出行建议。

##### 4.4.2.1 天津 2024 年 10 月气温数据分析

同样的，通过 Python 的数据可视化库 Matplotlib，我们针对天津 10 月的气温数据，以横坐标为日期（十月的每一天），设置纵坐标为气温，绘制气温变化折线图。

首先，我们通过 Python 爬虫爬取了 2024 年 10 月 1 日-2024 年 10 月 31 日的气温数据（包括每日最高气温和最低气温），并把它存储在 csv 文件中，作为数据库的预备。

而后，我们从数据库中读取数据，绘制折线图。其中红色折线表示最高气温随时间的变化趋势，蓝色折线表示最低气温的变化情况。（如图 4 所示）同样地，这个可视化的方式能够直接地让用户知晓天津 10 月的气温变化规律，以供日常出行。

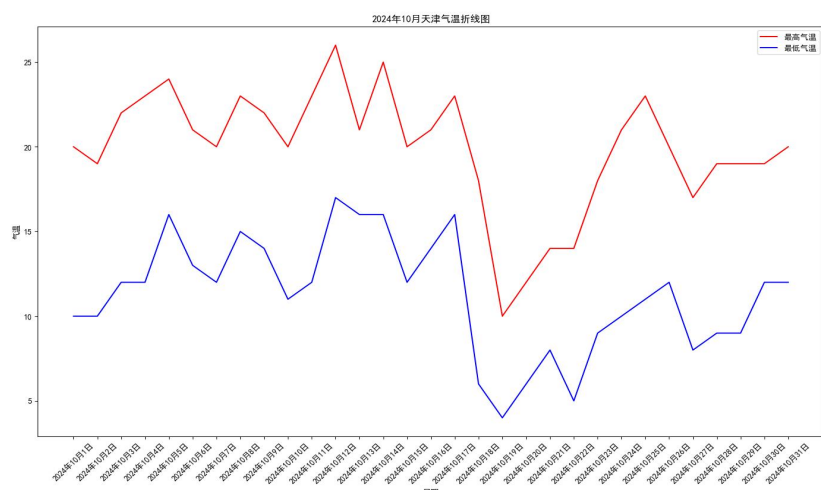


图 4：2024 年 10 月天津气温折线图

从图中，我们可以知道：

整体上：

最高气温和最低气温大概变化趋势相同，气温不断变化；以 10 月 17 日为分界线，前半个月气温较高，后半个月气温较低，但在 10 月 23 日有所回温。

部分上：

最高气温范围在 10-26 摄氏度之间；每 3-4 天气温具有单调性，而后单调性改变，变化较大。

最低气温范围在 4-17 摄氏度之间；每 3-4 天气温具有单调性，而后单调性改变，变化较大。

从以上图表的分析中，我们可以得出一个结论：天津的 10 月份气温总体上不稳定，气温跨度达到将近 20 摄氏度，需要人们关注天气情况进行调整。

#### 4.4.2.2 天津 2024 年 10 月风级数据分析

针对天津 10 月的风级数据，我们以横坐标为风级，纵坐标为频数，绘制柱状风级图。

首先，我们通过 Python 爬虫爬取了 2024 年 10 月 1 日-2024 年 10 月 31 日的风级数据，并把它存储在 csv 文件中，作为数据库的预备。

而后，我们从数据库中读取数据，绘制柱状图。（如图 5 所示）以供用户了解风级大小，从而安全出行。



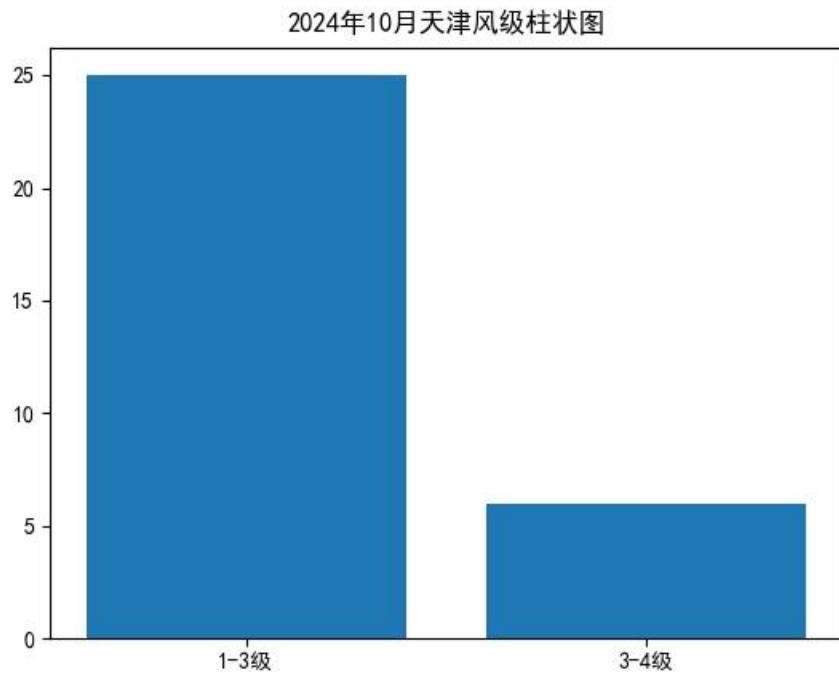


图 5：2024 年 10 月天津风级柱状图

对于风级而言，1-3 级风主要是软风、轻风、微风，在这 3 个风级下，人们能够感觉到有风，但是给人的感觉较为舒适。

3-4 级风主要是和风，在这个风级下，人会感到风的力量比较强劲，受到一定的阻力。

从天津的风级柱状统计图中可以知道：

天津 10 月份的风总体上给人以较大的舒适度，但是也要注意有将近一周的时间会有较为强劲的风，需要注意防风。

#### 4.4.2.3 天津 2024 年 10 月风力数据分析

针对天津 10 月的风力（风的方向）数据，我们绘制了风力饼图。

首先，我们通过 Python 爬虫爬取了 2024 年 10 月 1 日-2024 年 10 月 31 日的风力数据，并把它存储在 csv 文件中，预备数据库。

而后，我们从数据库中读取数据，绘制饼图。（如图 6 所示）以供用户了解风的来向，从而在骑车、运动的时候防止逆风而行导致的受伤。

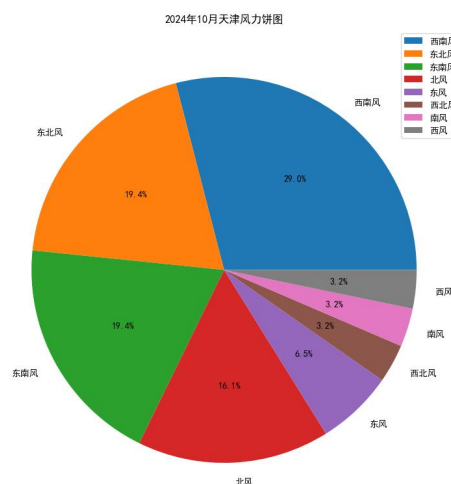


图 6：2024 年 10 月天津风力饼图

可以知道：在 10 月，天津经常刮西南风、东北风、东南风。由于我们的地理知识可以知道：

对于西南风：10 月处于季节转换期，副热带高压西侧的暖湿气流常向北方输送，可能被引导至天津；在气旋西侧，气流呈顺时针旋转流入中心，可能出现西南风。

对于东北风：10 月处于秋季，冷空气东侧地面常发展出气旋，冷空气是高压而气旋是低压，二者气压梯度大，导致空气从高压区流向低压区，使得天津容易刮东北风；天津北部有山脉等地形阻挡，冷空气南下时，气流方向发生改变，在天津形成东北风；虽然 10 月海陆热力性质差异不太明显，但仍然存在，影响风向。

因此，人们无论在日常生活中，还是在实验中，都要注意到这些风向的影响。

#### 4.4.3 天津湿度数据分析

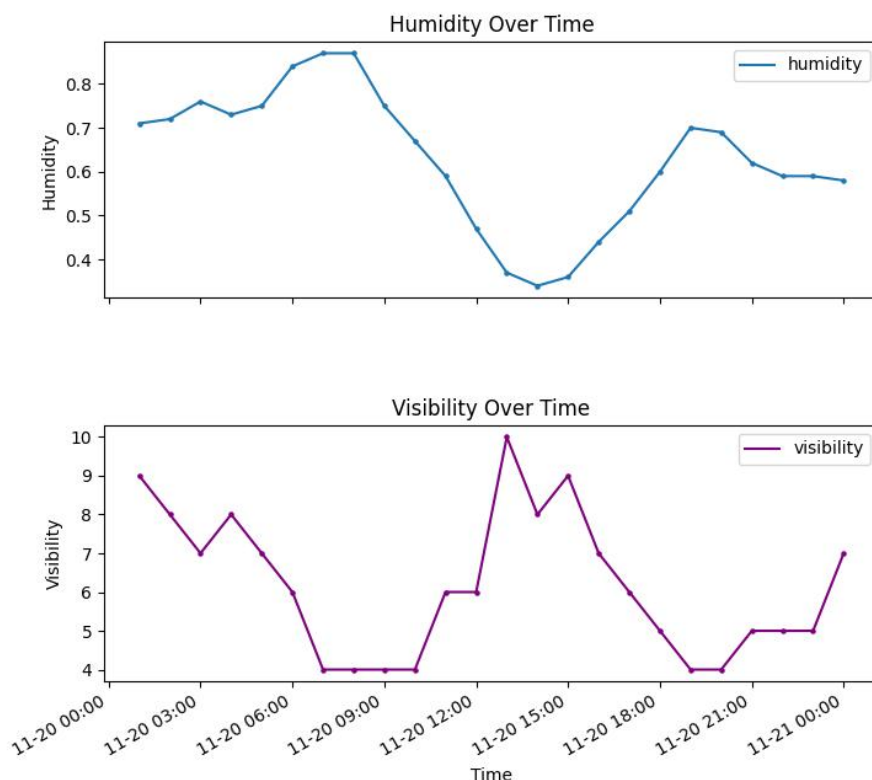


图 7：湿度、能见度折线图

以天津 11 月 20 日的湿度数据为例，我们绘制了湿度折线图。

整体上来看，出现两个峰值，分别在 6：00-9：00，以及 18：00-19：00 左右。分析可知，这两个时间段为温度变化较大的时间段，故由于液化与凝华的物理现象，这两个时段湿度较大。

在 12:00-15:00 时段，出现低谷，此时正值晌午，也就是一天中最热的时间段。小液滴得以最大程度地气化，从而使湿度降低。

由于一天内湿度变化较大，对人们的生活有一定影响。以下是一些建议，帮助大家更好地应对湿度变化，保持舒适的生活环境：

1. 早晨湿度较高时：保持室内通风，尤其是起床后，开窗换气，以降低室内湿度，避免潮湿引起的不适；晨练时，尽量避免在湿度较大的环境中长时间运动，以免出现身体不适。

2. 中午湿度较低时：注意补充水分，因为湿度较低时，空气较为干燥，容易导致皮肤和呼吸道干燥；使用加湿器，以保持适宜的室内湿度。

3. 傍晚湿度再次上升时：关闭门窗，防止湿气进入室内，可以使用除湿器降低室内湿度；晚餐适量食用一些利水渗湿的食物，如红豆、冬瓜、薏米等，有助于身体排湿。

4. 全天候建议：关注天气预报，了解当日湿度变化，提前做好应对措施；保持室内整洁，避免潮湿环境导致的细菌滋生。

遵循以上建议，可以有效应对一天内湿度变化对生活的影响，保持身体健康和舒适的生活环境。

#### 4.4.4 天津能见度数据分析

从上图可知，可见度和湿度二者明显地呈负相关互补的趋势。

整体上来看，出现两个低谷，分别在 6:00-9:00，以及 18:00-19:00 左右。分析可知，这两个时间段为温度变化较大的时间段，故由于液化与凝华的物理现象，这两个时段湿度较大，故能见度较低。

在 12:00-15:00 时段，出现峰值，此时正值晌午，也就是一天中最热的时段。

一天内能见度的变化可能会对出行、工作甚至健康产生影响。以下是一些建议，帮助大家应对能见度变化，确保生活安全与便利：

1. 早晨能见度较低时：出行前，关注天气预报和实时交通信息，了解当天的能见度情况；驾驶车辆时，开启雾灯、近光灯，减速慢行，保持安全距离；步行或骑行时，穿着反光衣物或佩戴反光配件，提高自身在低能见度环境中的可见性。

2. 中午能见度好转时：驾驶车辆时，根据能见度情况适时调整车速，但切勿超速行驶；利用能见度较好的时段进行户外活动，享受阳光，但要注意防晒。

3. 傍晚能见度再次降低时：尽量避免在能见度不佳的情况下进行长途驾驶或户外活动；家中应备有应急照明设备，如手电筒、蜡烛等，以防停电时使用。

4. 全天候建议：家中、办公室和车内应备有必要的防护用品，如口罩、护目镜等，以应对雾霾等低能见度天气；对于有呼吸道疾病或对空气质量敏感的人群，低能见度天气时应减少外出，必要时佩戴口罩。

遵循以上建议，可以帮助大家更好地应对一天内能见度的变化，确保生活秩序和人身安全。

## 五、效果评价以及总结

### 5.1 效果评价

#### 5.1.1 数据展示：

地图展示天气数据，直观地展示不同地区的天气状态，数据易于理解。湿度、

天气质量、衣物厚度和体感温度等以水球图、仪表盘合适的形式展示。折线图融合了多种数据，相比传统折线图展示效果更好。

天气数据仅限于温度和各种数值指标，未能完全反映复杂的天气状况（如天气、风向和风速等），而且没有文字性的信息，可以进一步丰富数据。

#### 5.1.2. 交互性：

启用了交互功能，用户可以通过点击或悬浮查看详细信息，增强了图表的互动性和用户体验；折线图也设置了时间滑块和各种辅助标志线，具有较强交互性。但是代码使用了基础的 Page 方法结合，结构较为简单，如果有能力可以进一步换用 Grid 方法并添加 Timeline，实现更高级展示效果。

#### 5.1.3. 美观性：

代码使用了深色主题，可以增强视觉冲击力，适合呈现数据。但是可能需要对颜色的使用进行优化，确保不同区域的数据对比足够明显；文字大小和字体等也可以添加艺术字使其更为美观。

#### 5.1.4. 易用性：

较大尺寸的图表适合展示更多数据，且地图展示方式直观，易于理解。但是区域数据过多，可能导致用户视觉疲劳，可以考虑添加二级界面等。与此同时，由于本身 Pyecharts 无法直接添加二维表格，导致在显示指数如晨练指数、人体舒适度指数的同时，无法显示对应的指数文字意义，会造成一定的认知模糊。

## 5.2 总结

#### 5.2.1 技术实现：

本作业使用了 Pyecharts 和 Python 实现了天气数据的可视化，充分发挥了 Pyecharts 的地图和动态动画效果优势，尤其适合天气这种具有空间分布特征的数据。

#### 5.2.2 创新性与独特性：

将天气数据与地理信息（如温度与区域名称）结合，提供了较好的展示方式。可以考虑加入更多天气数据（如降水量、风速等）以增强可视化效果。

#### 5.2.3 潜在改进方向：

##### 5.2.3.1 增加更多天气维度：

如紫外线指数、气压等，提供更加全面的天气信息。

##### 5.2.3.2 优化用户交互体验：

增加搜索框、缩放功能或详细信息窗口，以方便用户快速查找感兴趣的区域。添加指数对应的文字描述，更加具有可读性。

整体来说，该项目在展示天气数据上有很好的基础，通过进一步的优化和丰

富，可以更好地提高用户体验和可视化效果。

## 六、参考文献

- [1]贾硕,吴萍萍,李仪芳,等.基于天气数据的天气穿衣指数的应用研究[J].农村实用技术,2019,(03):89-90.
- [2]穆璐,李菁,李萌萌.基于天气数据的天气穿衣指数产品 模型设计与应用[J].天气科技进展,2017,7(06):194-196.
- [3]海南省市场监督管理局.常用生活天气指数海南省地方标准:DB46/ XXXXX—XXXX[S].
- [4]向怀菊,宋光勇,陈诚,蒋松林.黄平县近30年人体舒适度指数特征分析[J].天气变化研究快报,2024,13(4):823-829. DOI: 10.12677/ccr1.2024.134094.
- [5]霍寿喜.天气与穿着[J].百科知识,2008,40(9):1821.
- [6]田晓亮,章大海,王兆俊,孙晖.人体着装过程热湿舒适性仿真研究[J].青岛大学学报(工程技术版),2005,40(1):15.
- [7]国家质量监督检验检疫总局,中国国家标准化管理委员会.环境空气质量指数(AQI)技术规定:HJ 633—2012[S].2012-02-29发布.