

首都天气密码

气温作为线索的统计探究

小组成员：陈博远、杨淏天、付华钰



内容索引

1.1.数据来源与格式

1.2.数据处理思路

2.1.每日温度预期值的全年变化趋势

2.2.预测每个日期的气温

3.1.温度与空气质量的关系

3.2.风速对空气质量的影响 (续)

4. 温度和压强的关系 (?.随机分布的彩蛋

5. 总结



1.1.数据来源与格式

当地时间 北京市	T/°C	P/mmHg	相对湿度/%	风向/16向	风速/ms ⁻¹
16.01.2024 23:00	17.6	767.0	73	从东方吹来的风	1
16.01.2024 20:00	18.0	766.7	75	从东北偏东方向吹来的风	2
16.01.2024 17:00	21.1	765.5	65	从西南偏南方向吹来的风	2
16.01.2024 14:00	23.5	765.7	54	从东南偏东方向吹来的风	2
16.01.2024 11:00	20.2	767.8	62	从东北偏北方向吹来的风	3
16.01.2024 08:00	16.6	767.2	75	从东方吹来的风	2
16.01.2024 05:00	16.7	766.3	75	从东南偏东方向吹来的风	3
16.01.2024 02:00	17.4	766.3	75	从东方吹来的风	2
15.01.2024 23:00	18.4	766.4	77	从东方吹来的风	2
15.01.2024 20:00	19.7	766.0	72	从东方吹来的风	2
15.01.2024 17:00	23.2	764.8	46	从东南偏东方向吹来的风	2

数据特点：

- 1.每3小时一组数据
- 2.风向以文字记录 (需要数据处理)

数据来源^[1]

[1] Reliable Prognosis. 北京历史天气[EB/OL].(2024-1-13)[2024-1-17].
https://rp5.ru/%E5%97%E4%BA%AC%E5%8E%86%E5%8F%B2%E5%A4%A9%E6%B0%94_

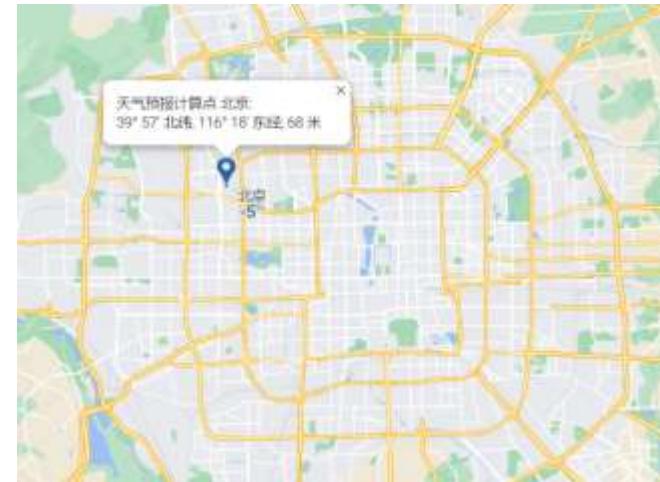


图1.1: 数据来源气象站位置



1.1.数据来源与格式 (续)

Date	AQI	PM2.5	PM10	CO	SO2	NO2	O3_8h
2015/11/1	98	72	106	1	7	78	39
2015/11/2	102	76	116	1.4	8	78	44
2015/11/3	199	149	181	1.9	16	90	63
2015/11/4	260	210	0	2	15	108	29
2015/11/5	144	110	0	1	6	53	26
2015/11/6	37	20	9	0.5	2	29	38
2015/11/7	37	19	0	0.5	2	29	41
2015/11/8	82	60	89	0.9	3	46	37
2015/11/9	163	124	0	1.6	4	56	7

数据特点：

- 1.每天一组数据，记录了多种污染物浓度
- 2.取AQI值衡量空气质量情况

数据来源^[2]

[2]中国气象局. 北京空气质量指数历史数据[EB/OL].(2024-1-1)[2024-1-17].

<https://www.aqistudy.cn/historydata/monthdata.php?city=%e5%8c%97%e4%ba%ac>



1.2.数据处理思路

一、温度:

$$T = T(y, d, t)$$

1. 温度平均值

1.1. 日平均气温:

$$T_t = E_t(T)(y, d)$$

1.2. 这个日期的温度的参考值:

$$T_{t,y} = E_y(E_t(T))(d)$$

2. 昼夜温差

2.1. 日温度极差 (昼夜温差) :

$$\Delta T(y, d)$$

2.2. 这个日期的昼夜温差的参考值:

$$E_y(\Delta T)(d)$$

二、压强:

$$P = P(y, d, t)$$

1. 日平均压强:

$$P_t = E_t(P)(y, d)$$

2. 这个日期的压强的参考值:

$$P_{t,y} = E_y(E_t(P))(d)$$

三、空气质量:

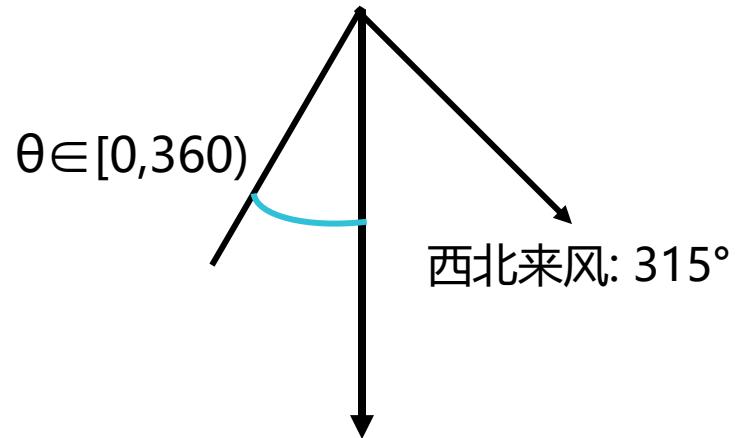
$$AQI = AQI(y, d)$$



1.2. 数据处理思路 (续)

四、风向D(Direction):

1. 建立坐标系如下, 转换成角度描述



五、风速V(Velocity):

$$V = V(y, d, t)$$

1. 每日风速的大小

1.1. 日平均风速 (方均根) :

$$V_t = RMS_t(V)(y, d)$$

1.2. 这个日期的风速的参考值:

$$V_{t,y} = E_y(RMS_t(P))(d)$$

2. 选定一个方向(θ_0), 计算风速的投影:

2.1. 日平均风速:

$$V_{//t} = RMS_t(V \cos(\theta - \theta_0))(y, d)$$

2.2. 这个日期的方向风速的参考值:

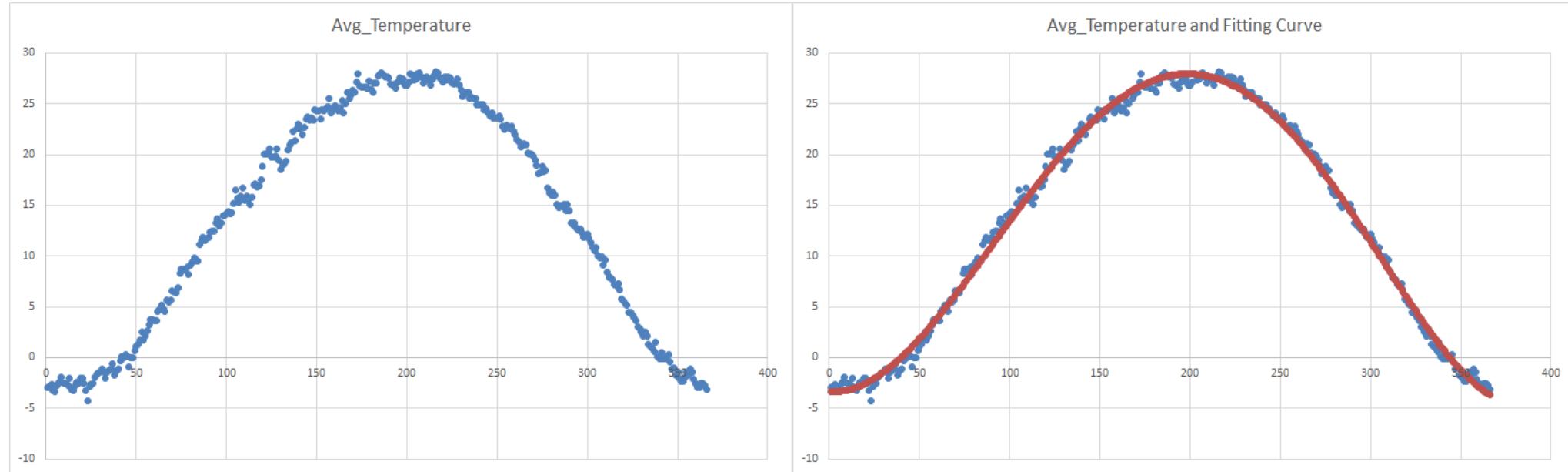
$$V_{//t,y} = E_y(RMS_t(V \cos(\theta - \theta_0)))(d)$$

2.3. 从不同方向吹来的风的数据点个数:

$$N = N_{(\theta)}$$



2.1. 每日温度预期值的全年变化趋势



每日温度的组成：随季节时令波动的预期值，和贡献随机性的波动值

$$\bar{T}_{(daily)} = T_{base(season)} + \varepsilon$$

若要研究每日平均气温的随机性(ε)，应消除这个系统误差 $T_{base(season)}$



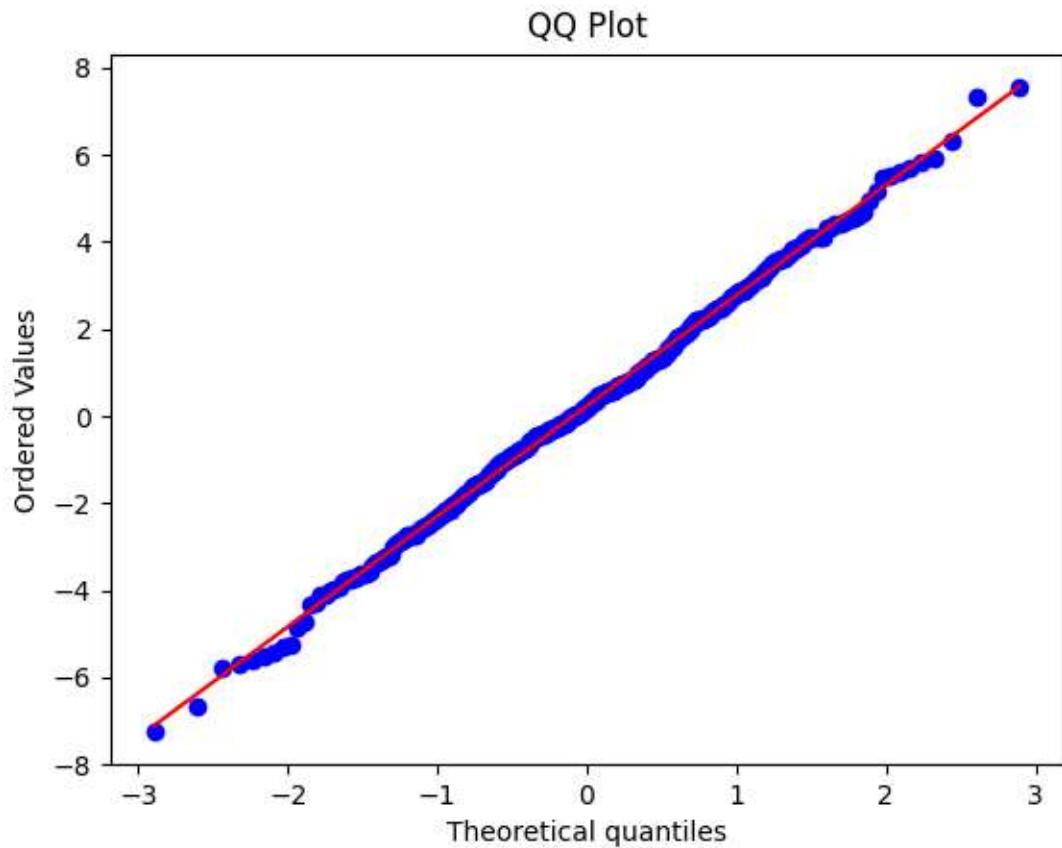
2.2. 预测每个日期的气温

(小付同学说, 他能预言明年今天的温度?)

假设不同年份同一天的日平均温度服从正态分布,
即 $T_t \sim N(\mu, \sigma)$:

1. 检验上述假设: 非参数检验对比总体
思想: 数据Q-Q图 \rightarrow CDF, 比较CDF
(Kolmogorov-Smirnov假设检验方法)

发现: P值为0.10, 未能拒绝



2.2. 预测每个日期的气温

(小付同学说, 他能预言明年今天的温度?)

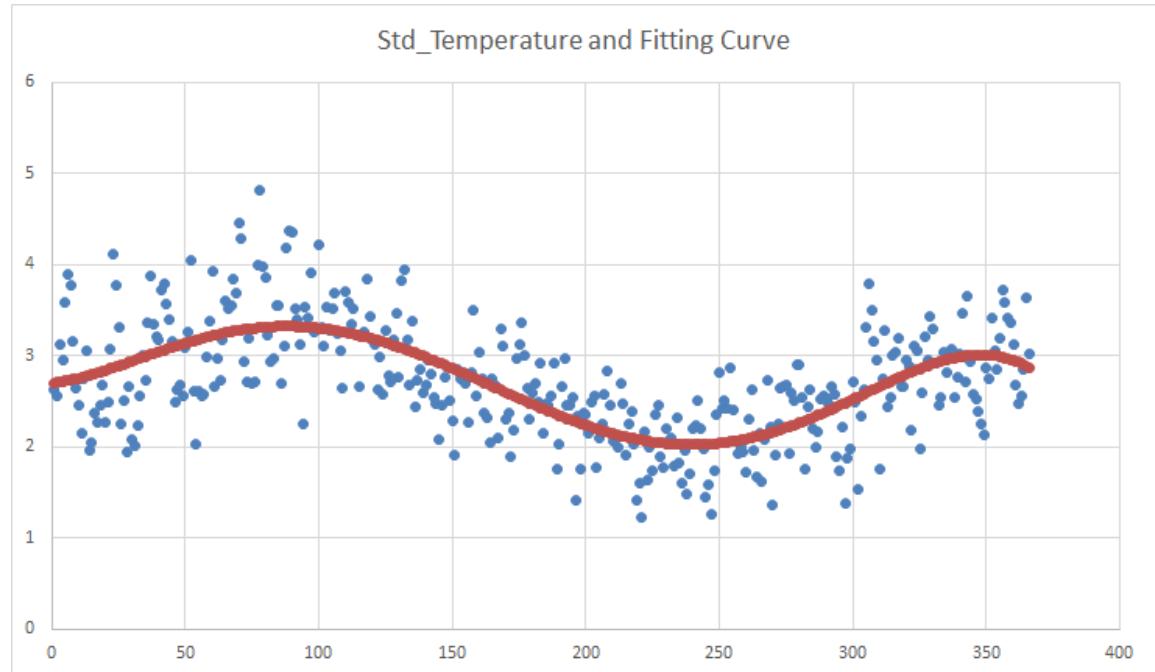
假设不同年份同一天的日平均温度服从正态分布,
即 $T_t \sim N(\mu, \sigma)$:

2. 分布参数的估计: (其中 μ 和 σ 是日期的函数)

$$\mu = \sum_{i=1}^{i=16} \frac{x_i}{16}, \quad \sigma_{16}^2 = \frac{1}{15} \sum_{i=1}^{16} (x_i - \bar{x})^2$$

发现: 方差随季节变化 (见右图)

(春天温度不稳定, 推荐春捂秋冻)



2.2. 预测每个日期的气温 (续)

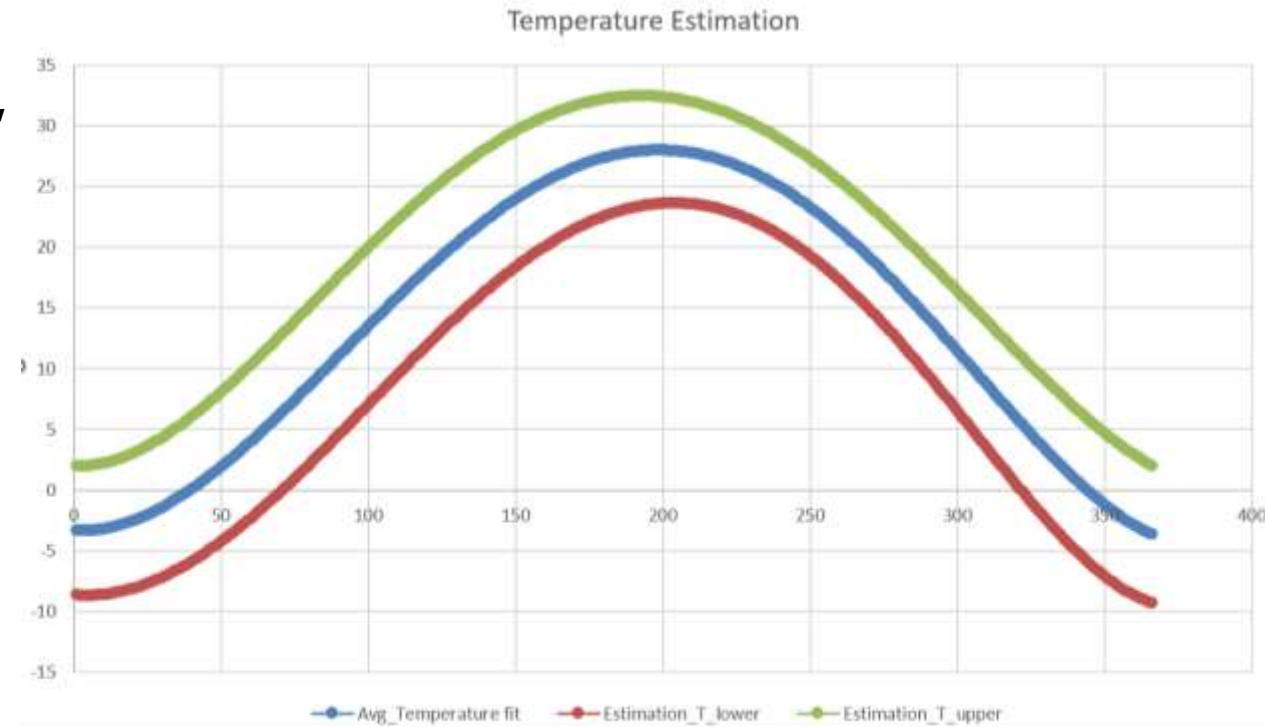
(小付同学说, 他能预言明年今天的温度?)

假设不同年份同一天的日平均温度服从正态分布,
即 $T_t \sim N(\mu, \sigma)$:

3. 利用分布估计一年中每天的日平均温度的95%
置信区间 (右图) :

$$[\mu - u_{0.025} \sigma, \mu + u_{0.025} \sigma]$$

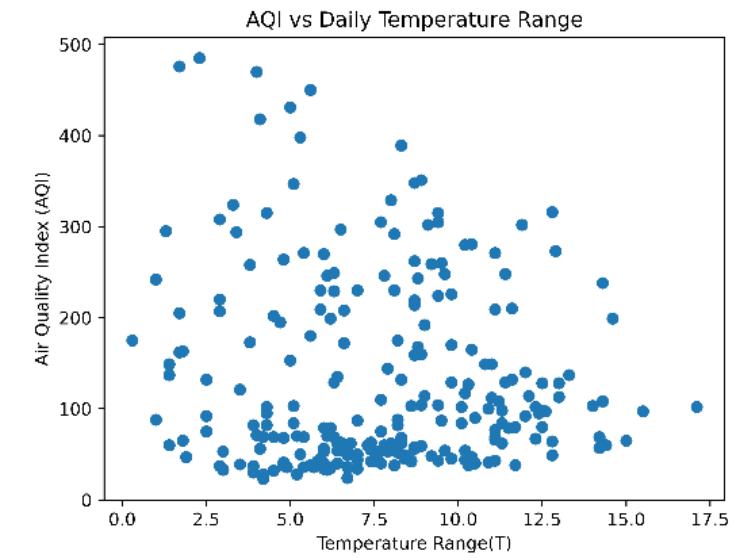
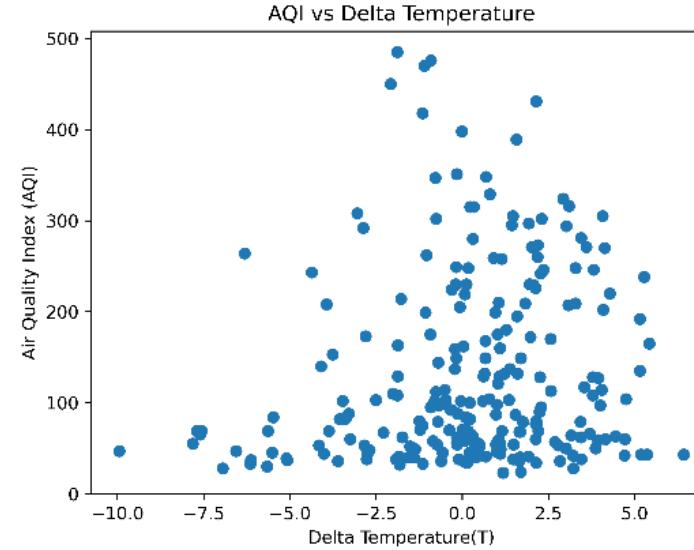
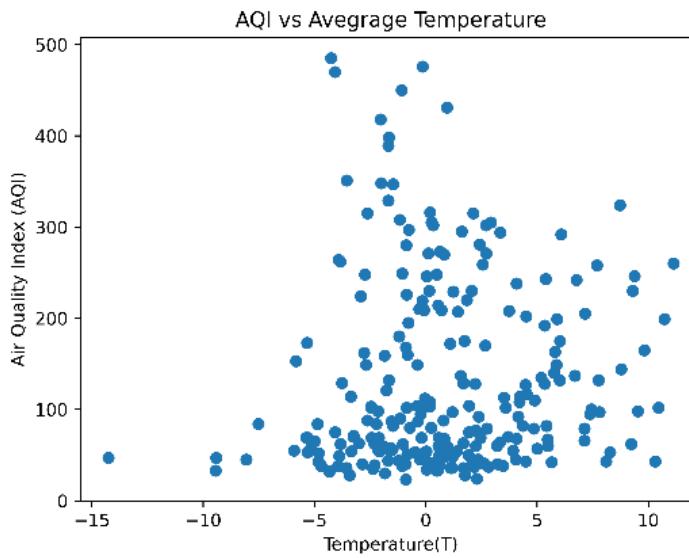
检验这个区间估计: 2022年的实际日平均温度
有336个点落在各自的置信区间内, 占比92%



3.1. 温度与空气质量的关系

(小杨同学：以前的北京，冬天真难受，起床一看，不是降温，就是雾霾-_-

1. 研究日均温度 T_t 和空气质量指数 (AQI) 的相关性 (左图)
2. 消除时令季节的影响，也研究 $T_t - T_{t,y}$ 和空气质量指数 (AQI) 的相关性 (中图)
3. 研究每日昼夜温差 ΔT 和空气质量指数 (AQI) 的相关性 (右图)



结果：均无明显相关性 (很喜欢小杨同学的一句话：啊?)



3.2. 风速对空气质量的影响

(小杨同学很不甘心，继续解析空气质量：着眼于风

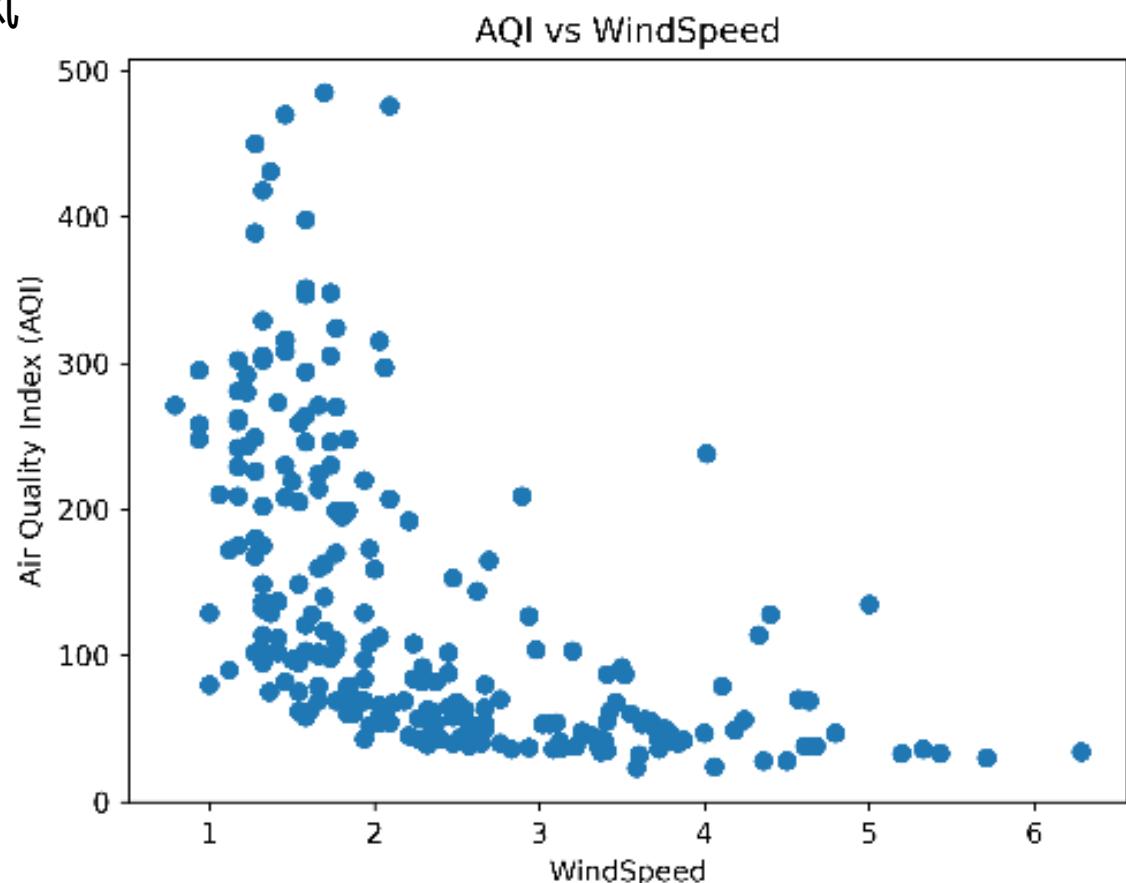
1. 研究每日风速大小 V 和空气质量指数 (AQI) 的相关性 (如图)

结果：

空气质量指数 (AQI) 和风速有一定的负相关性

(相关系数 $\text{Corr} = 0.49$)

(大风散雾霾，符合直观)

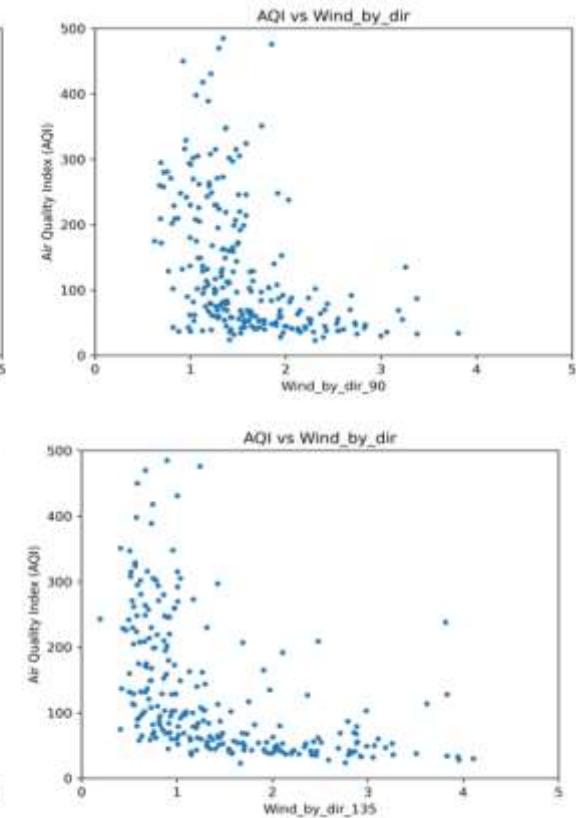
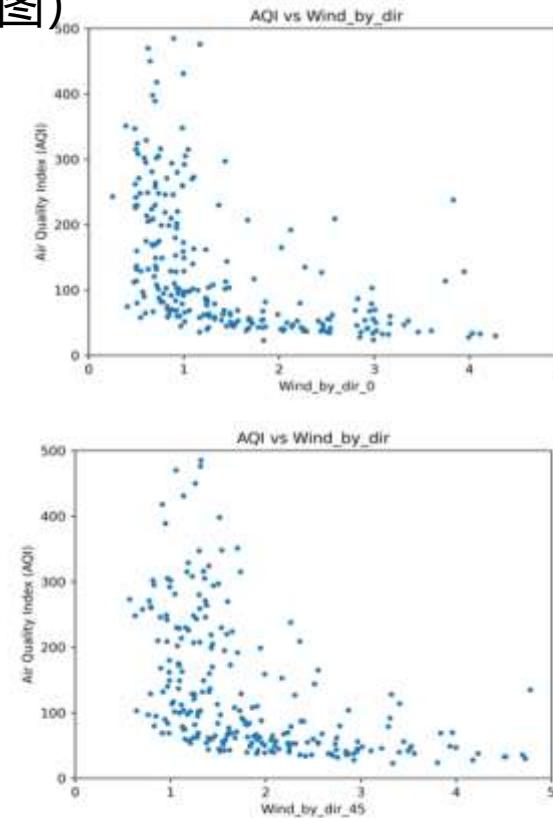
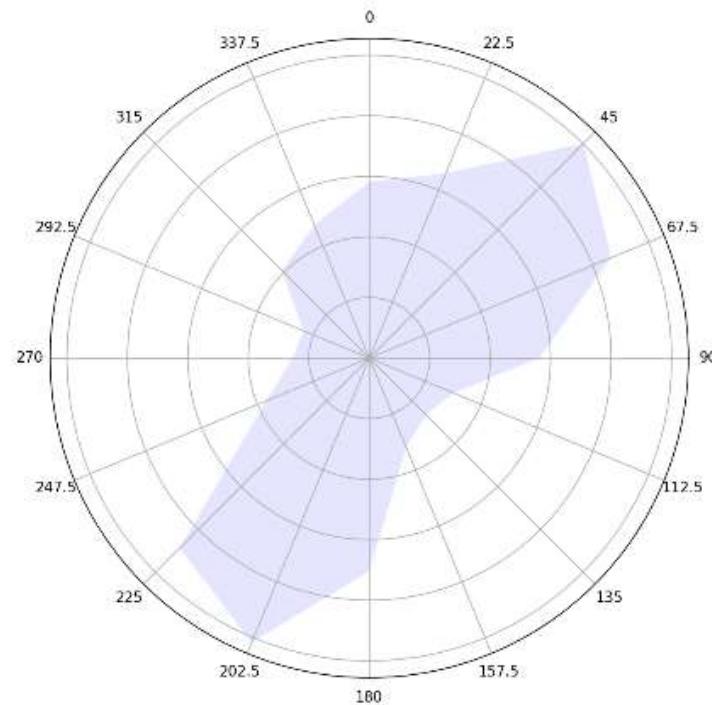


3.2. 风速对空气质量的影响 (续)

(小杨同学很不甘心, 继续解析空气质量: 着眼于风

2. 研究不同方向上的风速大小和空气质量指数 (AQI) 的相关性 (右图组)

背景信息: 北京西北面地势高, 此方向少风 (左图)



结果: 空气质量指数 (AQI) 和风速有一定的负相关性 (相关系数 $Corr = 0.49$)

风的方向对这个关系影响不大



4. 温度和压强的关系

(小陈同学想起了温度和压强的热力学关系：

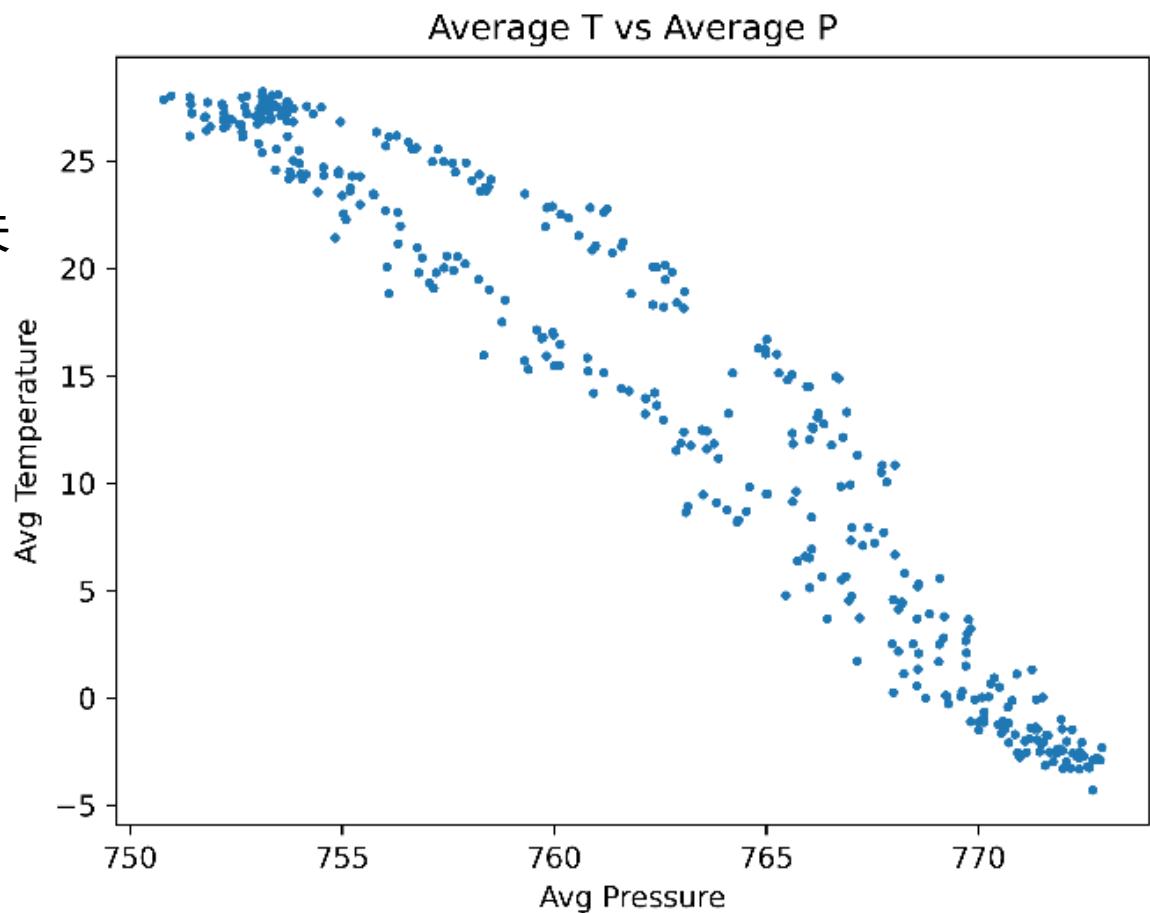
定性验证全年同一地点处，温度和压强的负相关关系 (如图)：

结果：

良好的线性负相关关系

但是分两支，像李萨如图形

反映了参量变化的相位不同



5.总结

- 1.温度预测模型性能良好
- 2.温度和污染的关系不符合预期
- 3.温度和压强的关系符合现实

(ps.只有小杨同学受伤的世界达成了TAT



谢谢！

陈博远、付华钰、杨淏天

