基础案例 | 天子脚下，天气何如？

**背景介绍**

天气和人们一直都息息相关，天气的变化不仅影响的是温度，对人们的生活也是有很大的影响，比如这天气适不适合外出、适不适合运动..等等。天气究竟是什么呢？其实天气是指某一个地区距离地表较近的大气层在短时间内的具体状态。而天气现象则是指发生在大气中的各种自然现象，即某瞬时内大气中各种气象要素（如气温、气压、湿度、风、云、雾、雨、闪、雪、霜、雷、雹、霾等）空间分布的综合表现。

天气过程就是一定地区的天气现象随时间的变化过程。各种天气系统都具有一定的空间尺度和时间尺度，而且各种尺度系统间相互交织、相互作用。许多天气系统的组合，构成大范围的天气形势，构成半球甚至全球的大气环流。天气系统总是处在不断新生、发展和消亡过程中，在不同发展阶段有着其相对应的天气现象分布。那么我们就来看看北京近年来的天气情况吧。我们从某网站抓取了713条数据，来分析影响北京空气质量指数的因素。

**因变量**

**空气质量指数**：连续变量，取值范围[15，450]。

**自变量**

**时间**：时间数据，[2016/1/1，2017/12/13]；

**星期**：分类变量，包括星期一、星期二等；

**最低气温**：连续变量(℃)，取值范围[-13，29]；

**最高气温**：连续变量(℃)，取值范围[-2，41]；

**天气**：分类变量，包括晴天、多云转晴、阴转多云、阴转小雨、多云转雨夹雪、小雨转多云等；

**风向**：分类变量，包括东风、西风、南风、北风、东南风、东北风、西南风、西北风等；

**风力**：连续变量，取值范围[0，6]；

**空气污染程度**：分类变量，包括优、良、轻度污染、中度污染、重度污染和严重污染。

**样本解释**

本案例以北京2017年12月1日为例，2017年12月1日星期五，这一天的最高气温是5℃，最低气温是-5℃，天气为晴，风向为西南风，风力大小在1-2级之间，空气质量指数为91，空气污染程度为良。

**回归分析**

建立回归模型，分析影响空气质量指数的主要因素，为广大劳动者提供好的外出时间，并且为空气检测分析部门提供好的建议。