北京工商大学

本科实验报告­­­­

课程名称：机器学习

实验名称：空气质量监测数据的预处理和基本分析

专业班级： 学号：

学生姓名：

指导教师：高超

年 月 日

1. 实验目的

(1)掌握Python中Numpy和Pandas进行数据分析的方法。本案例基于北京市空气质量监测数据，聚焦数据建模中的数据预处理和基本分析环节，说明Numpy和Pandas的数据读取、数据分组、数据重编码、分类汇总等数据加工处理功能。

1. 实验内容

(1) 空气质量监测数据的预处理：根据空气质量监测数据的日期，生成对应的季度标志变量；对空气质量指数AQI分组，获得对应的空气质量等级。

(2) 空气质量监测数据的基本分析：计算各季度AQI和PM2.5的平均值等描述性统计量；找到空气质量较差的若干天数据，以及各季度中空气质量较差的若干天数据；计算季度和空气质量等级的交叉列联表；派生空气质量等级的虚拟变量。

1. 实验程序

(1) 空气质量监测数据的预处理程序代码：

(2) 空气质量监测数据的基本分析程序代码：

1. 实验结果与分析

(1) 空气质量监测数据预处理的实验结果：

(2) 空气质量监测数据基本分析的实验结果：