**广州商学院**

**课程论文**

**题目： 关于上海市空气质量的分析报告**

|  |  |
| --- | --- |
| 课 程 名 称 | **商务分析报告设计** |
| 考 查 学 期 | **2018-2019学年第二学期** |
| 考 查 方 式 | **课程论文** |
| 姓 名 | **杨伟庭** |
| 学 号 | **201616060022** |
| 专 业 | **信息管理与信息系统** |
| 成 绩 |  |
| 指 导 教 师 | **吴晓玲** |

摘要

空气质量的好坏反映了空气污染程度，它是依据空气中污染物浓度的高低来判断的。空气污染是一个复杂的现象，在特定时间和地点空气污染物浓度受到许多因素影响。对人或物造成危害的现象叫做大气污染。我国大气污染指标AQI，作为新的污染程度衡量指标。AQI分为六级，分别为一级优，二级良，三级轻度污染，四级中度污染，五级重度污染，六级严重污染。参与评价的污染物有SO2、NO2、PM10、PM2.5、O3、CO六项。空气中对人、动植物以及人的财产产生影响的主要是SO2、NO2、PM10、PM2.5、O3、CO，因此目前大气污染研究的重点也在此。由于本次数据收集的限制和本人能力的限制，本文仅限于对AQI指数及其污染物做肤浅的研究。

通过Python获取相关的空气质量数据，运用R语言的可视化工具对数据进行可视化操作，主要分析AQI指数和其主要污染物的变化情况，分析近五年来上海空气质量的变化情况，尝试分析其原因与其解决方法。

**关键词：**上海；空气质量；特征分析；AQI；R;

Abstract

The quality of air reflects the degree of air pollution, which is judged by the concentration of pollutants in the air. Air pollution is a complex phenomenon, and air pollutant concentrations are affected by many factors at specific times and locations. The phenomenon of harm to people or things is called air pollution. China's air pollution index AQI, as a new indicator of pollution level. AQI is divided into six levels, which are first-grade, second-grade, third-class mild pollution, four-level moderate pollution, five-level heavy pollution, and six-level severe pollution. The pollutants involved in the evaluation include SO2, NO2, PM10, PM2.5, O3 and CO. The main influences on people, animals, plants and people's property in the air are SO2, NO2, PM10, PM2.5, O3, and CO. Therefore, the focus of air pollution research is also here. Due to the limitations of this data collection and the limitations of my ability, this article is limited to superficial research on the AQI index and its pollutants.

Obtain relevant air quality data through Python, use R language visualization tools to visualize the data, analyze the changes of AQI index and its main pollutants, analyze the changes of Shanghai air quality in the past five years, and try to analyze the reasons and Solution.

**Keywords**: city; air quality; characteristic analysis

目录

[摘要 I](#_Toc11596577)

[Abstract II](#_Toc11596578)

[目录 III](#_Toc11596579)

[1. 背景 1](#_Toc11596580)

[1.1. 数据分析目标 1](#_Toc11596581)

[1.2. 空气质量相关知识 1](#_Toc11596582)

[2. 数据收集及其分析 2](#_Toc11596583)

[2.1. 数据分析流程 2](#_Toc11596584)

[2.2. 数据采集 2](#_Toc11596585)

[2.3. AQI指数分析 5](#_Toc11596586)

[3. 分析结论 9](#_Toc11596587)

[参考文献 10](#_Toc11596588)

1. 背景

## 数据分析目标

改革开放以来，上海对如何建设国际经济中心进行了积极探索和实践，取得了丰硕的成果。肩负着面向世界、推动长三角地区一体化和长江经济带发展的重任，在全国经济建设和社会发展中具有十分重要的地位和作用。

经济的发展带来环境的污染是正常的现象，但是要依据污染程度来判断是否要继续以种种方式发展，如果污染的指数非常高且持续时间长的经济发展方式是不可取的。不仅是因为这种方式不可持续，而且是因为这种污染对人体健康造成威胁。

如今随着经济发展，环境问题日趋严重，更成为人们日常讨论关注的话题。可以说环境或者是空气质量问题成为人们普遍关注和防范的热点问题。人人都希望自己居住在环境好，空气质量好的地方。如此的社会环境迫使政府部门切实地保障空气质量。

报告根据已有的数据，运用R语言可视化工具，对环境空气质量进行科学合理的评价，预测与分析是一个很具有实用价值的问题。通过分析近五年上海空气的质量数据，分析空气质量不仅可以对污染物治理有一定帮助，对于预防和警示经济发展也有很大作用。。

## 空气质量相关知识

空气质量指数（英语：Air Quality Index, AQI）是定量描述空气质量状况的[非线性](https://zh.wikipedia.org/wiki/%E9%9D%9E%E7%BA%BF%E6%80%A7" \o "非线性)[无量纲](https://zh.wikipedia.org/wiki/%E6%97%A0%E9%87%8F%E7%BA%B2%E9%87%8F)指数。其数值越大、级别和类别越高、表征颜色越深，说明空气污染状况越严重，对人体的健康危害也就越大。AQI指数越大、级别和类别越高、表征颜色越深，说明空气污染状况越严重，对人体的健康危害也越大。看AQI信息时，不需要记住AQI的具体数值和级别，只需要注意优（绿色）、良（黄色）、轻度污染（橙色）、中度污染（红色）、重度污染（紫色）、严重污染（褐红色）等六种评价类别和表征颜色。当类别为优或良、颜色为绿色或黄色时，一般人群都可以正常活动；当类别为轻度污染以上，颜色为橙色、红色、紫色或褐红色时，各类人群就需要关注建议采取的措施，在安排自己的生活与出行时作为参考。

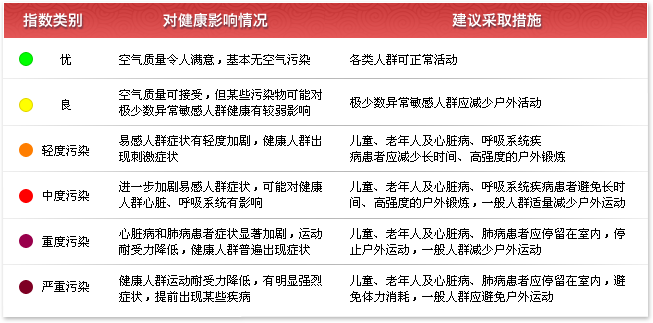


图1-1 AQI 等级及其对健康影响和采取措施

AQI计算与评价的过程大致可分为三个步骤：第一步是对照各项污染物的分级浓度限值，以细颗粒物（PM2.5）、可吸入颗粒物（PM10）、二氧化硫（SO2）、二氧化氮（NO2）、臭氧（O3）、一氧化碳（CO）等各项污染物的实测浓度值分别计算得出空气质量分指数；第二步是从各项污染物的IAQI中选择最大值确定为AQI，当AQI大于50时将IAQI最大的污染物确定为首要污染物；第三步是对照AQI分级标准，确定空气质量级别、类别及表示颜色、健康影响与建议采取的措施。

简言之，AQI就是各项污染物的空气质量分指数（IAQI）中的最大值，当AQI大于50时对应的污染物即为首要污染物。

# 数据收集及其分析

## 数据分析流程

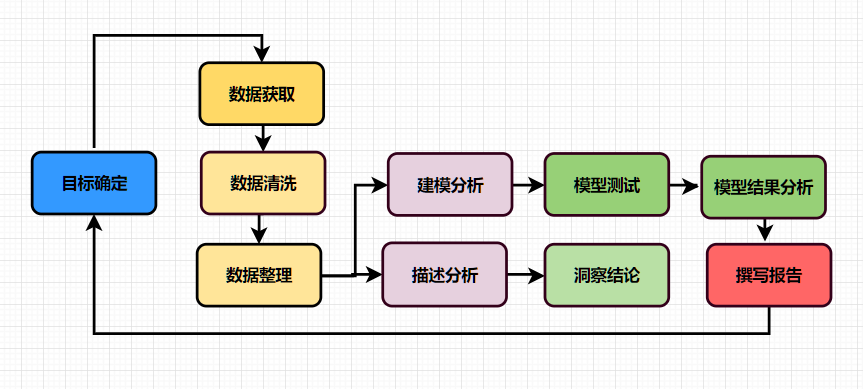


图2-1-1 数据分析流程图

总体的分析流程可分为三大部：

一、数据的设计。

二、数据的清洗建模分析。

三、撰写数据分析报告。

## 数据采集

本次数据的来源为 <https://www.aqistudy.cn> ，此网站的数据易于Python抓取且来源于环保部官方数据，从最早2013年12月到现在的记录，数据的准确性和权威性可以得到保障。

Python 爬虫说明：方法Python+Selenium

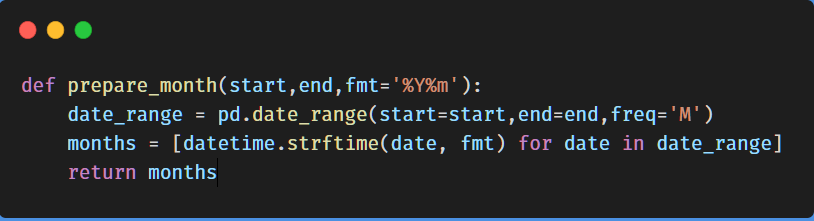


图表 3-1导入相关的包和构造存储数据字段



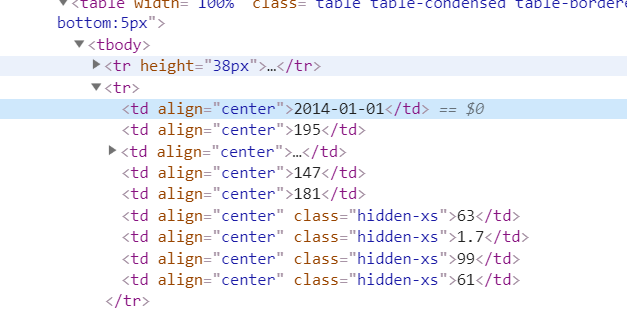
图表 3-1-1上海历年月份空气质量链接

图表3-1网站的上海空气质量指数月统计历史数据都在 table 下的td 的a标签内，且链接的格式为目标城市名和年份和日期。因此可以构造5年来的年月日期来构造完整的5年月份的链接。



图表 3-1-2构造年份和月份

构造一个函数，传入开始和结束的日期，从开始日期循环到结束日期，保存到months列表。

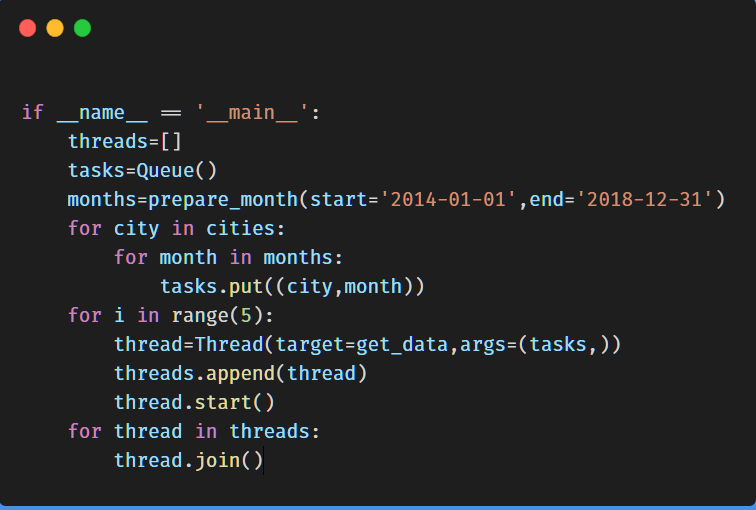


图表 3-1-3获取目标网页结构

每个月明天的数据都是在table 下的每个tr 里的td中。我们可以循环解析得到所需要的数据保存到csv文件中。如下图中使用 Selement 开启多线程，获取数据保存到csv中。



图表 3-1-4浏览器对象和数据解析

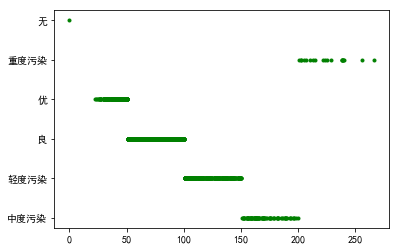


图表 3-3-5主函数

## AQI指数分析

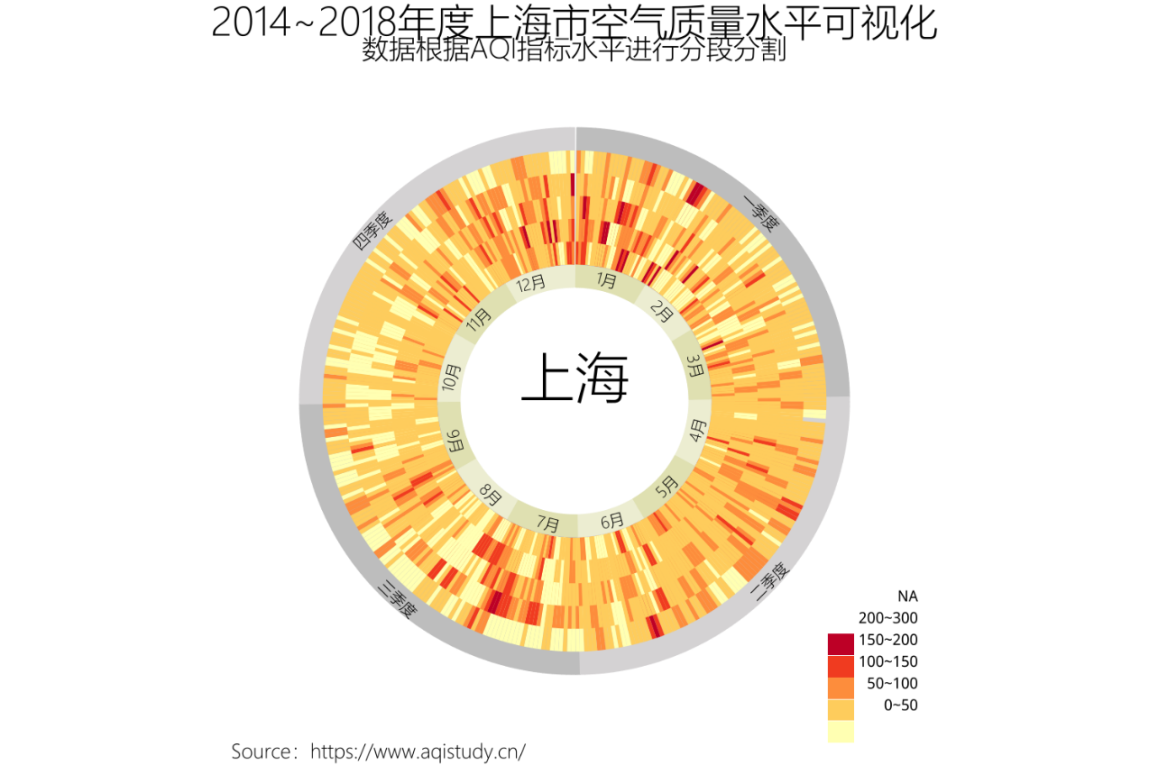
图表 4-1-1 上海2016-2018 AQI指数直方图

图表4-1-1可看出上海五年来AQI指数基本上时大于50，达到了对应的污染物即为首要污染物的指标。近五年上海空气质量数据中相当大的一部分的AQI值不容乐观，基本上所记录的相关污染物都达到了其标准的最低标准。部分数据更是达到轻度污染甚至重度污染的标准。这意味着这五年来上海的空气质量并未得到改善，警示者有关部门应积极采取相关措施把控好空气质量这一事关民生的问题。



图表 4-1上海五年空气质量散点图

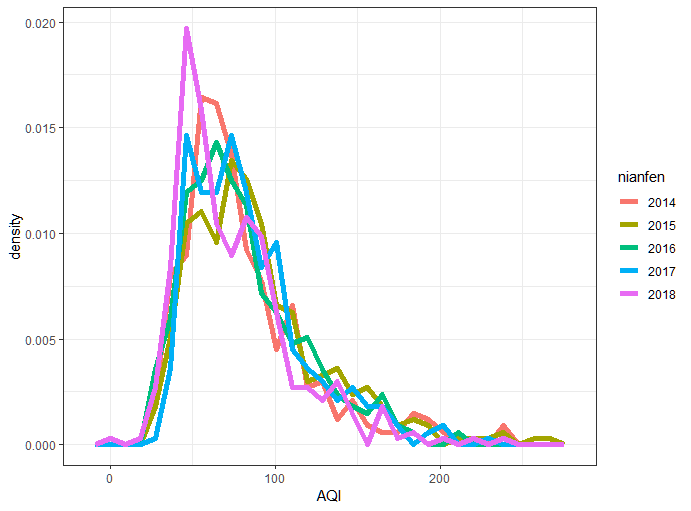
上图上海五年来空气质量指数AQI达到优或良要比空气质量达到轻度污染以上的情况要少得多。绝大多数都是达到了污染的标准。进一步说明近五年来上海的空气质量情况比较的悲观。



图表 4-1-2近五年上海市空气质量AQI指标图

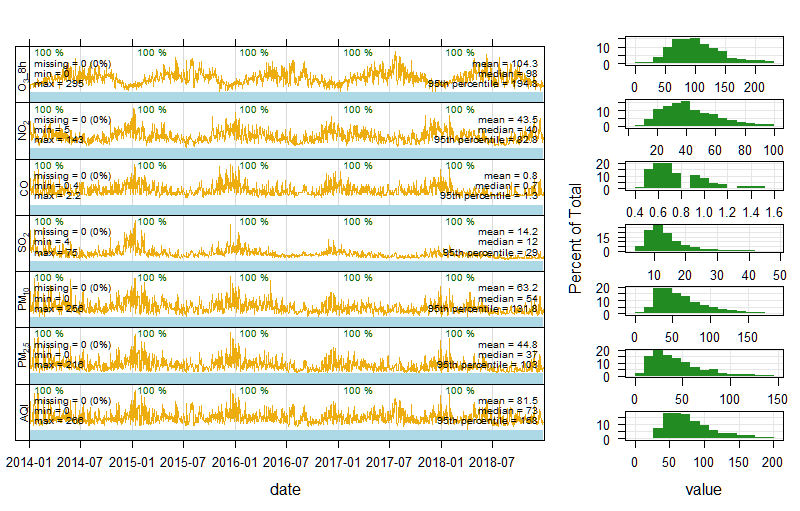
由图4-1-1可以看出上海近五年空气质量AQI呈现出明显的季节性特征。AQI空气质量指标在夏季和冬季。由此可推断在低湿度、高强度紫外线辐射等气象因素影响下，臭氧浓度易快速升高，形成污染。

冬季气团干燥，不利于形成降水，降雨量少且持续时间较短，风速和风力较小，对空气中污染物的冲刷效果不明显，也是导致AQI浓度居高不下的原因之一。上海临近北方，北方冬季集中供暖，供暖需要燃烧大量的煤炭，随着冷空气的南下，各种污染物也随这冷空气来到上海，加剧了上海的空气污染情况。



图表 4-1-2上海AQI年度曲线图

图4-1-2中AQI 指数年变化近五年来每年的情况都是比较类似的，也就是说每年AQI状况没有明年的改善，2018年还有上升的情况。可以说上海治理空气污染仍然需要加大力度，做到落实相关的环保措施，推动上海空气质量真正地得到改善，还市民能清新的空气。



图表 4-2-3近五年各污染物谁随时间的变化情况

图4-2-2各污染物和AQI指数近五年年随时间的变化的趋势时比较一致的，呈现出一定的线性关系。近五年上海市各污染物浓度最高出现的时间都是冬季和夏季。

国家气候中心监测数据显示，近年来年厄尔尼诺持续发展加强，达到1986年以来的最强值，成为极强厄尔尼诺事件。受其影响，冬季影响我国的冷空气明显偏弱，长江中下游等地区天气静稳，大气扩散条件很差。致使污染物在这两季节内无法扩散，导致空气质量在这两季节里相对来说比较差。

# 分析结论与建议

通过分析上海近五年来的空气质量数据，上海的空气质量原因是冬季北方比较严重的雾霾天气与燃煤供暖有关，南下的冷空气将污染物携带至上海，与不利的扩散条件一起加重了上海空气污染程度。尽管近年来上海市采取了一定的措施来治理空气质量问题，例如控制汽车排放，提倡绿色出行和绿色可持续发展经济产业，但是近五年的空气质量数据分析表明上海的空气质量或者说空气污染物是一直处于高水平的状态。这无疑要政府有关部门应高度重视。

为了更清新的空气与经济发展并肩齐区，我认为不是一个地方一个地区能够完全解决的，我认为治理空气污染上海应该**统筹考虑，控制大、中、小型污染源，把落后的重污染的企业进行技术改造或者依法取缔，控制污染还需要“最佳管理方案”，如小型污染源的控制。**并非所有控制措施都需要昂贵的技术或新技术，对小型污染源控制来说，常用方法还是最佳管理方案，这只需要常识和尽一切努力保护环境的意识。方案通常都不复杂，但日积月累，就对改善地区空气质量做出了很大贡献。

在空气质量比较差的季节，更应做好相关的预防工作。为市民提供相关的污染。平时加强环保相关的宣传工作，使人们潜移默化地养成环保的好习惯，绿色出行，多做好相关公共交通的建设，市民亦可以绿色出行，减少驾车从而达到改善空气质量。

参考文献

[1][刘洁](https://s.ixueshu.com/index.html?q=%E5%88%98%E6%B4%81),[张小玲](https://s.ixueshu.com/index.html?q=%E5%BC%A0%E5%B0%8F%E7%8E%B2),[徐晓峰](https://s.ixueshu.com/index.html?q=%E5%BE%90%E6%99%93%E5%B3%B0).[北京地区SO\_2、NO\_x、O\_3和PM\_(2.5)变化特征的城郊对比分析](http://www.ixueshu.com/document/69a3887b70652d6a.html)[J].[环境科学](https://s.ixueshu.com/index.html?q=%E7%8E%AF%E5%A2%83%E7%A7%91%E5%AD%A6),2008(04):214-220.

[2][王珩](https://s.ixueshu.com/index.html?q=%E7%8E%8B%E7%8F%A9),[于金莲](https://s.ixueshu.com/index.html?q=%E4%BA%8E%E9%87%91%E8%8E%B2).[大气中PM\_(10)浓度的影响因素及其污染变化特征分析](http://www.ixueshu.com/document/d202c392a0e1bfa4.html)[J][上海师范大学学报(自然科学版)](https://s.ixueshu.com/index.html?q=%E4%B8%8A%E6%B5%B7%E5%B8%88%E8%8C%83%E5%A4%A7%E5%AD%A6%E5%AD%A6%E6%8A%A5(%E8%87%AA%E7%84%B6%E7%A7%91%E5%AD%A6%E7%89%88)),2004(03):

[3][李琼](https://s.ixueshu.com/index.html?q=%E6%9D%8E%E7%90%BC),[李福娇](https://s.ixueshu.com/index.html?q=%E6%9D%8E%E7%A6%8F%E5%A8%87),[叶燕翔](https://s.ixueshu.com/index.html?q=%E5%8F%B6%E7%87%95%E7%BF%94).[珠江三角洲地区天气类型与污染潜势及污染浓度的关系](http://www.ixueshu.com/document/bd1c800c032a45d5318947a18e7f9386.html)[J].[热带气象学报](https://s.ixueshu.com/index.html?q=%E7%83%AD%E5%B8%A6%E6%B0%94%E8%B1%A1%E5%AD%A6%E6%8A%A5),1999(04):363-36

[4][任阵海](https://s.ixueshu.com/index.html?q=%E4%BB%BB%E9%98%B5%E6%B5%B7),[苏福庆](https://s.ixueshu.com/index.html?q=%E8%8B%8F%E7%A6%8F%E5%BA%86),[高庆先](https://s.ixueshu.com/index.html?q=%E9%AB%98%E5%BA%86%E5%85%88).[边界层内大气排放物形成重污染背景解析](http://www.ixueshu.com/document/35ad89cfc5c7fe1f.html)[J].[大气科学](https://s.ixueshu.com/index.html?q=%E5%A4%A7%E6%B0%94%E7%A7%91%E5%AD%A6),2005(01):63-69+175.

[5][杨德保](https://s.ixueshu.com/index.html?q=%E6%9D%A8%E5%BE%B7%E4%BF%9D),[王式功](https://s.ixueshu.com/index.html?q=%E7%8E%8B%E5%BC%8F%E5%8A%9F),[黄建国](https://s.ixueshu.com/index.html?q=%E9%BB%84%E5%BB%BA%E5%9B%BD).[兰州市区大气污染与气象条件的关系](javascript:;)[J].[兰州大学学报](https://s.ixueshu.com/index.html?q=%E5%85%B0%E5%B7%9E%E5%A4%A7%E5%AD%A6%E5%AD%A6%E6%8A%A5),1994(01):132-136.

**课程论文（设计）评分表**

|  |  |  |
| --- | --- | --- |
| 评分项目 | 分值 | 得分 |
| 对相关行业背景理解到位程度 | 10 |  |
| 需求分析和分析目标合理性与清晰度 | 10 |  |
| 数据分析框架逻辑性 | 15 |  |
| 数据挖掘模型采用情况 | 5 |  |
| 分析结论正确性 | 15 |  |
| 数据分析图表制作合理与专业水平 | 25 |  |
| 是否结合相关行业知识分析发现产生问题的原因 | 15 |  |
| 是否具有相对全面合理的解决措施与建议 | 5 |  |
| **总得分**： |  | |