

附件 2

2023 年“领航杯”江苏省青少年大数据分析 应用能力竞赛规则

一、参赛对象

我省具有正式学籍的全日制高校学生，省教育厅负责对参赛学生进行资格审查。

二、组队要求

本次比赛为团体赛，每支参赛队伍限报 3 名参赛队员、1 名指导老师，每名指导老师限指导 1—2 支队伍，不得跨校组队。每所高校参加线上初赛的队伍不超过 6 支，每校仅限 1 支队伍入围线下决赛。高校在不同城市的校区视为不同高校，可分别组队参赛并具有入围决赛的资格。

三、竞赛流程

1. 竞赛模式

线上初赛和线下决赛竞赛时长均为 3 小时。赛题由理论和实操两部分组成，理论部分成绩占 20%、实操部分成绩占 80%，合并计算总成绩。每个参赛队共用同一套题目环境，每位队员均可操作题目环境。参赛队自由分工协作，共同完成全部题目。选手通过竞赛系统完成答题，可多次提交验证。平台将根据赛题标准结果进行实时评判，根据比赛进度和解题情况实时公布比赛成绩。

2.竞赛内容

线上初赛的理论知识包括数据库基础、数据分析基本流程、常见数据分析技术、数据安全与管理、操作系统基础知识等，实操技能包括 Linux 系统操作、基础环境管理、数据库配置与管理、SQL 语句应用、Python 数据分析、Hadoop 平台及组件部署与管理等。线下决赛的理论知识包括爬虫框架与反爬技术、数据清洗基本流程、数仓分层设计与规范、大数据生态圈应用技术、数据可视化理论基础等，实操技能包括网络数据采集、离线数据分析、实时数据分析、数据可视化等等。

3.计分规则

竞赛总成绩 1000 分，其中理论部分 100 道题，每题 2 分，计 200 分，实操部分 800 分。

竞赛实操部分设不同任务，不同分数和实操题型的题目。比赛中由评分系统自动验证结果，评分方式为代码或命令是否正确、结果是否正确。两者均正确情况下，系统自动给出相应题目得分。每个任务先对的先得分，后提交的会根据提交顺序按 1% 进行分值递减，递减至 20% 后不再递减。

4.竞赛环境

所有参赛选手须自带笔记本电脑进行答题，过程中所使用的各类技术工具均由参赛选手自行准备。比赛过程中开放内部局域网、互联网，选手必须使用组委会提供的网络环境进行答题。竞赛过程中全程采用“防作弊”技术实时监测，一经发现，取消参赛队参赛资格。