**唐山市职业学校技能大赛**

**大数据应用与服务赛项规程**

赛项名称： 大数据应用与服务

英文名称：Big data applications and services

赛项组别： 中等职业教育

1. 赛项信息

赛项名称：大数据应用与服务

赛项组别：中职组

英文名称：Big data applications and services

赛项归属：电子与信息大类

1. 竞赛目的

此次竞赛的举办旨在有效促进大数据教学模式的探索性改良，推进相关专业课程体系、教学内容和教学方法等教学资源的质量提升和丰富完善，进而推动大数据专业教育上层建筑体系质的飞跃。

通过大数据技能竞赛，能够激发学生的自主学习热情，树立正确积极的职业价值观和人生观。通过大赛，可以提高实践教学课时量，学生可在“大数据竞赛平台”中以实际大数据项目案例开展平台搭建数据采集、数据分析与挖掘等方面得到有效锻炼，提高学生的专业技能，并逐步实践“理实一体化”、“学教一体化”的教学模式。

1. 竞赛内容

结合实际情况，本次竞赛采用理论与实操结合方式，理论 20%，实操 80%权重。实操考核采用自动评分机制，实时公布竞赛动态及成绩排名。

比赛赛题由易到难，赛项设置了集群搭建、数据采集、数据分析及可视化、案例分析等环节。通过赛项的设置，考核锻炼选手如下能力：大数据环境搭建与运维水平、数据采集与预处理能力、数据分析软件使用水平、数据关联分析、数据挖掘能力。具体考核内容如下：

|  |  |
| --- | --- |
| 模块 | 模块名称及考核内容 |
| A | 理论知识模块 |
| 分析工具知识  数据库知识  数据分析知识  数据管理知识  数据安全知识  数据保密知识  相关法律知识 |
| B | 数据分析工具应用 |
| 运用Excel进行数据分析  运用SQL语句进行数据查询及分析  运用SPSS进行数据分析及建模  运用BI工具进行数据分析及展示  运用Python进行基本数据分析 |
| C | 大数据处理与应用模块 |
| 基础环境配置  数据库环境配置  分布式网络环境配置  大数据集群部署与管理 |
| D | 数据采集 |
| 网络数据采集  数据存储 |
| E | 数据案例分析 |
| 数据清洗  数据分析  算法实现  数据可视化 |

1. 竞赛方式

**（一）选手构成**

大数据应用与服务赛项为团体赛。每支参赛队由 3 名参赛选手组成，参赛选手不得跨校组队。

每队可配 1名指导教师，指导教师须为本校专兼职教师，参赛选手和指导教师报名获得确认后不得随意更换。

参赛选手须为中等职业学校全日制在籍学生或五年制高职一至三年级（含三年级）全日制在籍学生。凡在往届全国职业院校技能大赛中获得本赛项一等奖的选手，不能再参赛。参赛选手的资格审查工作按照《全国职业院校技能大赛制度汇编》要求执行。

**（二）竞赛方式**

本赛项设单一场次，竞赛时长为6小时（1天内进行），具体的时间安排以正式发布的竞赛指南为准。参赛选手在规定时间内，使用现场提供的工具完成模块任务，并按照要求提交作品。

竞赛平台：青椒课堂（<https://www.qingjiaoclass.com/），具体使用流程详见附件1>

**（三）竞赛评分**

本赛项实操考核采用自动评分机制，实时公布竞赛动态及成绩排名。

1. 竞赛流程

**（一）竞赛流程图**

大数据应用与服务赛项的竞赛流程如图1所示。



**图1 赛事流程**

**（二）竞赛时间表**

如下表，竞赛时间一览表作为参考，以实际安排为准。

**表1 竞赛时间一览表**

|  |  |  |
| --- | --- | --- |
| **7月13日** | **时间** | **事项** |
| 报到 | 7:00—8:00 | 各领队带队与学生一起到C楼阶梯教室报到 |
| 开幕式 | 8:10—8:40 | 地点：C楼阶梯教室，按指定位置就坐 |
| 赛场检录 | 8:40—9:20 | 一次加密：参赛队抽取参赛编号  二次加密：参赛队抽取赛位号 |
| 选手比赛 | 9:20—9:30 | 参赛队进入比赛赛位，进行赛前设备、材料检查 |
| 9:30—15:30 | 选手比赛 |
| 15:30—17:30 | 申诉受理 |
|  |  | 成绩核定和解密 |

1. 竞赛规则

1. 大数据应用与服务赛项为团体赛。每支参赛队由 3 名参赛选手组成，参赛选手不得跨校组队。每队可配1名指导教师，指导教师须为本校专兼职教师，参赛选手和指导教师报名获得确认后不得随意更换。参赛选手须为中等职业学校全日制在籍学生或五年制高职一至三年级（含三年级）全日制在籍学生。参赛选手年龄须不超过21周岁（年龄计算的截止时间以2024年6月1日为准）。凡在往届全国职业院校技能大赛中获得本赛项一等奖的选手，不能再参赛。参赛选手的资格审查工作按照《全国职业院校技能大赛制度汇编》要求执行。

2. 比赛赛位通过抽签决定，比赛期间参赛选手原则上不得离开比赛场地。

3. 比赛所需的硬件、软件和辅助工具统一提供，参赛队不得使用自带的任何具有存储和通信功能的设备，如硬盘、光盘、U盘、手机、随身听、智能手表、PDA等。

4. 参赛选手在赛前10分钟领取比赛任务，并进入比赛赛位。比赛正式开始方可进行相关操作。

5. 比赛过程中，参赛选手组内进行协作交流时，不得影响其他参赛成员，参赛选手如有疑问，应举手示意，现场裁判应按要求及时予以答疑。如遇设备或软件等故障，参赛选手应举手示意，现场裁判、技术人员等应及时予以解决。确因计算机软件或硬件故障，致使操作无法继续，经赛场裁判长确认，予以启用备用设备。

6. 比赛时间结束，选手应全体起立，结束操作。经工作人员查收清点所有文档后方可离开赛场，离开赛场时不得带走任何资料。

7. 赛项裁判应严格遵守赛项各项规章制度，确保比赛公平、公正、公开。

8. 比赛结束后，评分裁判方可入场进行成绩评判。最终竞赛成绩经复核无误、裁判长签字确认后，按要求公布。

9. 赛项中每个比赛环节裁判评分的原始材料和最终成绩等结果性材料，经裁判长签字后，装袋密封留档；由赛项承办院校封存，并委派专人妥善保管。

1. 技术规范

按照《全国职业院校技能大赛赛项规程编制要求》，结合企业职业岗位对人才培养需求，并参照表中相关国家职业标准制定。参赛代表队在实施竞赛项目中要求遵循表中的规范。

**表2 专业技术规范**

|  |  |  |
| --- | --- | --- |
| **序号** | **标准号** | **标准名称** |
| 1 | GB/T 11457-2006 | 信息技术、软件工程术语 |
| 2 | GB8566-88 | 计算机软件开发规范 |
| 3 | GB/T 12991-2008 | 信息技术数据库语言 SQL 第 1 部分：框架 |
| 4 | GB/T 21025-2007 | XML 使用指南 |
| 5 | GB/T 28821-1012 | 关系数据管理系统技术要求 |

**表 3 大数据相关标准**

|  |  |  |
| --- | --- | --- |
| 6 | GB/T 38672-2020 | 信息技术 大数据接口基本要求 |
| 7 | GB/T 38673-2020 | 信息技术 大数据大数据系统基本要求 |
| 8 | GB/T 38676-2020 | 信息技术 大数据存储与处理系统功能测试  要求 |
| 9 | GB/T 38643-2020 | 信息技术 大数据分析系统功能测试要求 |
| 10 | GB/T 38675-2020 | 信息技术 大数据计算系统通用要求 |
| 11 | GB/T 38633-2020 | 信息技术 大数据系统运维和管理功能要求 |
| 12 | GB/T 38672-2020 | 信息技术 大数据接口基本要求 |
| 13 | GB/T 38673-2020 | 信息技术 大数据大数据系统基本要求 |

**表4 软件开发标准**

|  |  |  |
| --- | --- | --- |
| 14 | GB/T 8566-2001 | 信息技术软件生存周期过程 |
| 15 | GB/T 15853-1995 | 软件支持环境 |
| 16 | GB/T 14079-1993 | 软件维护指南 |
| 17 | [GB/T 22032—2021](https://openstd.samr.gov.cn/bzgk/gb/javascript:void(0)) | [系统与软件工程 系统生存周期过程](https://openstd.samr.gov.cn/bzgk/gb/javascript:void(0)) |
| 18 | GB/T 17544-1998 | 信息技术软件包质量要求和测试 |

1. 技术环境

**（一）竞赛环境**

1.竞赛场地。现场保证良好的采光、照明和通风，提供稳定的水、电和供电应急设备，同时提供指导教师休息场所。

2.竞赛设备。所有竞赛设备由承办院校负责提供和保障，竞赛区按照参赛队数量准备比赛所需的软硬件平台，为参赛队提供标准竞赛设备。

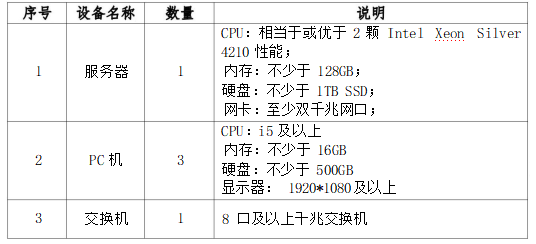
3.竞赛工位。每个工位上标明编号，工作台上面摆放3台计算机，并安装竞赛所需的相关软件；赛场全程视频监控。

4.裁判区。裁判休息的区域。

5.技术支持区。具备电脑、打印机、备用机器等，主要对选手的工位进行技术保障。

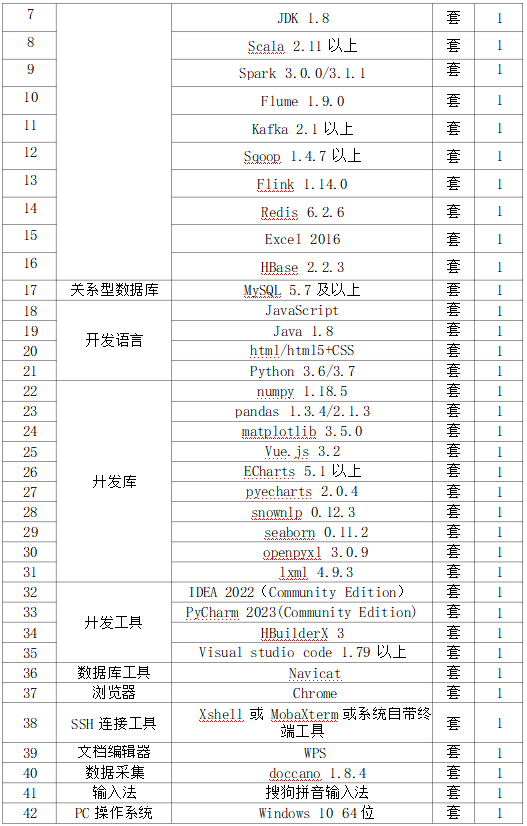
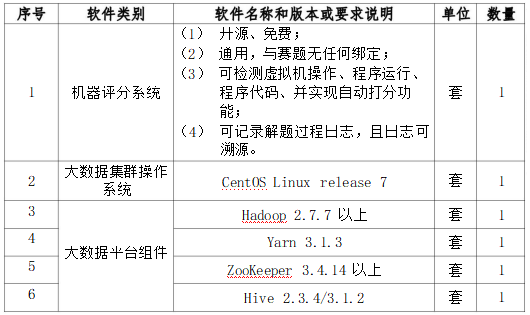
**（二）选手竞赛设备**

**1.硬件平台**

**表5 硬件平台**

**2.软件平台**

**表6软件平台**



1. 成绩评定

**（一）评分原则**

**1. 评分标准制定原则**

竞赛评分制定严格遵守公平、公正的原则，大数据应用与服务赛项评分采用赛项结果评分方法，始终贯彻落实竞赛一贯坚持的公平、公正和公开原则。

1.参与竞赛成绩管理的组织机构包括裁判组、监督仲裁组等，裁判组实行“裁判长负责制”，监督仲裁组组成结构、人数和相关要求参考《2023 全国职业院校技能大赛监督仲裁工作管理办法》中监督仲裁遴选条件和办法。

2.裁判评分方法，根据评分标准，各项目评分裁判根据选手操作过程和操作结果进行评分，独立评分。

3.成绩产生方法。为保证公开、公平、公正、透明地进行成绩评定，在裁判员的评分中，取两个评分裁判平均分作为选手技能得分。

4.成绩审核方法为各裁判员首先审核自身对选手的原始打分成绩，并签名，裁判长对所有裁判员的打分成绩进行审核，并签名，再由监督组对竞赛成绩抽检复核。

**2. 结果评分原则**

本赛项奖项设团体奖。设奖比例为：以赛项实际参赛队总数为基数，一、二、三等奖获奖比例分别为 10%、20%、30%（小数点后四舍五入）。

如出现参赛队总分相同情况，按照模块二、三、一顺序的得分高低排定名次顺序，即总成绩相同的情况下优先比较模块二的成绩，模块二成绩高的排名优先，其次按照模块三、一的成绩以此类推完成相同成绩的排序。如果各模块分值相同，则比较模块二各任务分值，按照任务编号从大到小的得分值排序。

**3. 二层加密原则**

比赛过程采取二层加密，通过抽取参赛编号和赛位号，屏蔽参赛队信息，每个环节设置独立裁判，每个环节结束后，数据立即封存于裁判长处，确保成绩评定公平、公正。

**4. 成绩抽查复核**

对赛项总成绩排名前60%的成绩进行复核，其中前15%的成绩进行全部复核，其余成绩进行抽检复核，抽检覆盖率为30%。

**（二）评分标准**

**表7 评分标准**

|  |  |  |  |
| --- | --- | --- | --- |
| **模块** | **任务** | **主要知识与技能点** | **分值** |
| 模块一:平台搭建与运  淮 | 任务一：大数据平台搭建 | Hadoop完全分布式下的 JDK 的解压安装、JDK 环竟变量配置、节点配置、Hadoop 配置文件修改、运行测试等。 | 10 |
| 任务二:数据库配置维护 | 使用SQL语句建立数据库和表 使用SQL语句对表进行增删改操作  使用 SQL 语句对表进行统计查询操作。  将 CSV 文件导入到数据库中。 | 20 |
| 模块二:数据获取与处理 | 任务一：数据获取与清洗 | 使用Python 程序读取 CSV 文件。  使用Python 程序处理空字段数据、异常字段数 | 10 |
| 任务二：数据标注 | 使用Python 对指定数据进行分类标注。使用Python 将标注后的数据保存到指定位置。 | 10 |
| 任务三：数据统计 | HDFS 上传 CVS 文件到指定目录下。  编写 MapReduce 程序对数据求解最大值、最小值、平均值和分区、分组操作等。  将计算结果保存到 HDFS 指定目录下。 | 15 |
| 模块三：业务数据分析与可视化 | 任务一：数据可视化 | 使用堆叠图展示相关统计数据  使用散点图展示相关统计数据  使用柱状图展示相关统计数据 | 20 |
| 任务二：业务分析 | 理解额业务场景，对数据报表进行分析研究，给出相应的建议和举措 | 15 |
| 总分 | | | 100 |

**（四）评分方法**

1. 竞赛满分为100分。

2. 本赛项配备裁判若干人，其中包括裁判长、评分裁判、加密裁判、现场裁判。

3. 裁判对参赛队伍提交的作品采取评价分和测量分结合评分。采取分步得分、累计总分的计分方式。各环节分别计算得分，错误不传递，按规定比例计入团队总分。根据赛题模块情况进行任务划分，分组评分裁判负责模块任务进行独立评分。

4. 裁判长正式提交全部模块评分结果并复核无误后，在督察员监督下进行解密。

5. 为保障成绩评判的准确性，督察员对赛项总成绩排名前60%的成绩进行复核，其中前15%的成绩进行全部复核，其余成绩进行抽检复核，抽检覆盖率为30%。

6. 在竞赛过程中，参赛选手如有不服从裁判裁决、扰乱赛场秩序、舞弊等行为的，由裁判长按照规定扣减相应分数，情节严重的将取消比赛资格，比赛成绩计0分。

7. 赛项成绩解密审核无误后，经裁判长、督察员和仲裁组签字后，向全体参赛队进行公布。比赛成绩以最终公布结果为准。

1. 赛项安全

赛事安全是技能竞赛一切工作顺利开展的先决条件，是赛事筹备和运行工作必须考虑的核心问题。赛项执委会采取切实有效措施保证大赛期间参赛选手、指导教师、裁判员、工作人员及观众的人身安全。

**（一）比赛环境**

执委会须在赛前组织专人对比赛现场、住宿场所和交通保障进行考察，并对安全工作提出明确要求。赛场的布置，赛场内的器材、设备，应符合国家有关安全规定。承办单位赛前须按照执委会要求排除安全隐患。

赛场周围要设立警戒线，防止无关人员进入发生意外事件。比赛现场内应参照相关职业岗位的要求为选手提供必要的劳动保护。在具有危险性的操作环节，裁判员要严防选手出现错误操作。

执委会须会同承办单位制定赛场和体验区的人员疏导方案。赛场环境中存在人员密集、车流人流交错的区域，除了设置齐全的指示标志外，须增加引导人员，并开辟备用通道。

大赛期间，承办单位须在赛场管理的关键岗位，增加力量，建立安全管理日志。

参赛选手进入赛位、赛事裁判工作人员进入工作场所，严禁携带通讯、照相摄录设备，禁止携带记录用具。如确有需要，由赛场统一配置、统一管理。赛项可根据需要配置安检设备对进入赛场重要部位的人员进行安检。

**（二）组队责任**

1.各学校组织代表队时，须安排为参赛选手购买大赛期间的人身意外伤害保险。

2.各学校代表队组成后，须制定相关管理制度，并对所有选手、指导教师进行安全教育。

3.各参赛队伍须加强对参与比赛人员的安全管理，实现与赛场安全管理的对接。

**（三）应急安全预案**

比赛期间发生意外事故，发现者应第一时间报告大赛执委会，同时采取措施避免事态扩大。大赛执委会应立即启动预案予以解决并报告大赛办。赛项出现重大安全问题可以停赛，是否停赛由大赛办决定。事后，大赛执委会应向大赛办报告详细情况。

相关应急预案如表8所示。

**表8 相关应急预案**

|  |  |  |
| --- | --- | --- |
| **突发事件** | **预防措施** | **事件发生后应对措施** |
| 参赛选手突发病或受伤 | 在各赛位张贴安全操作说明 | 医务人员应采取紧急救护措施，及时进行救治，如病情或伤势严重，应及时送往最近医院进行救治 |
| 人员发生食物中毒 | 比赛期间指定的住宿/餐饮场地符合国家相关资质要求。并协调地方卫生部门做好检查工作 | 立即组织对中毒人员进行救治，必要时送往最近医院进行检查治疗。同时对可疑的食品、饮水及其有关原料、工具设备和场所以及可能受污染的区域采取保留、控制措施，组织开展现场调查，迅速查明原因，并及时向大赛执委会报告 |
| 设备损坏（如不能启动、反复重启等） | 提前一天烤机，所有设备开机运行，现场放置备机 | 参赛选手举手示意后，监考人员计时，裁判确认后更换备机，并由主裁判确定应计入延时的时间 |
| 设备掉电 | 竞赛前技术人员及监考人员检查所有电源插头，确保牢固；电源线尽量绑扎在参赛选手碰不到的地方，如桌子后面等；竞赛前提醒参赛选手注意尽量不要碰到电源，配置文件要随时保存 | 参赛选手举手示意后，监考人员计时，裁判确认后重启机器，并由主裁判确定应计入延时的时间 |
| 现场网络线缆故障 | 现场走线要规范，尽量走暗槽或现场人员接触不到的地方；对主要线路要在走线槽内留有备线 | 启用备线 |

**（四）处罚措施**

1. 因参赛队伍原因造成重大安全事故的，取消其获奖资格。

2. 参赛队伍有发生重大安全事故隐患，经赛场工作人员提示、警告无效的，可取消其继续比赛的资格。

3. 赛事工作人员违规的，按照相应的制度追究责任。情节恶劣并造成重大安全事故的，由司法机关追究相应法律责任。

1. 竞赛须知

**（一）参赛队须知**

1. 参赛队组成。每支参赛队由3名符合参赛资格学生组成，性别不限。

2. 指导教师。每支参赛队最多可配指导教师1名，指导教师经报名并通过资格审查后确定。

3. 已提交的报名信息，超过截止日期，不允许更改。

**（二）指导教师须知**

1.各参赛代表队要发扬良好道德风尚，听从指挥，服从裁判，不弄虚作假。如发现弄虚作假者，取消参赛资格，名次无效。

2.各代表队领队要坚决执行竞赛的各项规定，加强对参赛人员的管理，做好赛前准备工作，督促选手带好证件等竞赛相关材料。

3.竞赛过程中，除参加当场次竞赛的选手、执行裁判员、现场工作人员和经批准的人员外，领队、指导教师及其他人员一律不得进入竞赛现场。

4.参赛代表队若对竞赛过程有异议，在规定的时间内由领队向赛项仲裁工作组提出书面报告。

5.对申诉的仲裁结果，领队要带头服从和执行，并做好选手工作。参赛选手不得因申诉或对处理意见不服而停止竞赛，否则以弃权处理。

6.指导老师应及时查看大赛专用网页有关赛项的通知和内容，认真研究和掌握本赛项竞赛的规程、技术规范和赛场要求，指导选手做好赛前的一切技术准备和竞赛准备。

**（三）参赛选手须知**

1. 竞赛选手严格遵守赛场规章、操作规程和工艺准则，保证人身及设备安全，接受裁判员的监督和警示，文明竞赛。

2. 参赛选手在检录时需将身份证、学生证、参赛证等身份证件交由检录人员统一保管，不得带入场内。

3. 参赛选手进入赛场，不允许携带任何书籍和其他纸质资料（相关技术资料的电子文档由赛项执委会提供），不允许携带任何通信工具和存储设备。竞赛统一提供计算机以及应用软件。

4. 各参赛队应在竞赛开始前一天规定的时间段进入赛场熟悉环境，但不得触碰任何比赛设备及材料。

5. 竞赛时，在收到开赛信号前不得启动操作，各参赛队自行决定分工、工作程序和时间安排，在指定赛位上完成竞赛项目，严禁作弊行为。

6. 竞赛过程中，因严重操作失误或安全事故不能进行比赛的（例如因综合布线发生短路导致赛场断电的、造成设备不能正常工作的），现场裁判员有权中止该队比赛。

7. 凡在竞赛期间提前离开的选手，当天不得返回赛场。

8. 为培养技能型人才的工作风格，在参赛期间，选手应当注意保持工作环境及设备摆放符合企业生产“6S”（即整理、整顿、清扫、清洁、素养和安全）的原则，如果过于脏乱，裁判员有权酌情扣分。

9. 在竞赛中如遇非人为因素造成的设备故障，经裁判确认后，可向裁判长申请补足排除故障的时间。

10. 参赛队欲提前结束比赛，应向现场裁判员举手示意，记录比赛终止时间。比赛终止后，不得再进行任何与比赛有关的操作。

11. 各竞赛队按照大赛要求和赛题要求提交竞赛成果，禁止在竞赛成果上做任何与竞赛无关的记号。

竞赛操作结束后，参赛队要确认成功提交竞赛要求的文件，裁判员在比赛结果的规定位置做标记，并与参赛队一起签字确认。

**（四）工作人员须知**

1. 熟悉竞赛规则，服从管理，严格按照工作程序和有关规定办事。

2. 树立服务观念，本着一切为参赛选手着想的原则，以高度负责的精神、严肃认真的态度和严谨细致的作风，积极完成大赛工作任务。

3. 按规定佩戴胸卡，文明礼貌，保持良好形象。

4. 坚守工作岗位，不迟到，不早退，不无故离岗，特殊情况向组长请假。

5. 遇安全突发事件，按照工作预案及时组织疏散，确保人员安全。

6. 未经同意不得擅自发布关于比赛的言论，不得私自接受采访。

1. 申诉与仲裁

本赛项在比赛过程中若出现有失公正或有关人员违规等现象，参赛队领队可在比赛结束后2小时之内向仲裁组提出书面申诉。

书面申诉应对申诉事件的现象、发生时间、涉及人员、申诉依据等进行充分、实事求是的叙述，并由领队亲笔签名。非书面申诉不予受理。

赛项仲裁工作组在接到申诉报告后的2小时内组织复议，并及时将复议结果以书面形式告知申诉方。

# 附件1：

**竞赛平台注册报名手册**

**一、报名准备**

团队成员均需在青椒课堂网站上注册个人账号，网址：<https://www.qingjiaoclass.com/。>

一定要记住账号和密码，登录竞赛平台使用，（在培训时已经注册的无需在注册，直接使用即可）

**二、组建团队**

由参赛队伍中的一名成员（选手/教师都可以）青椒课堂在竞赛专区板块内，

1、点击“队伍管理”；

2、进入页面后点击“创建或加入队伍”；

3、在创建队伍下输入信息后，队伍创建完成；

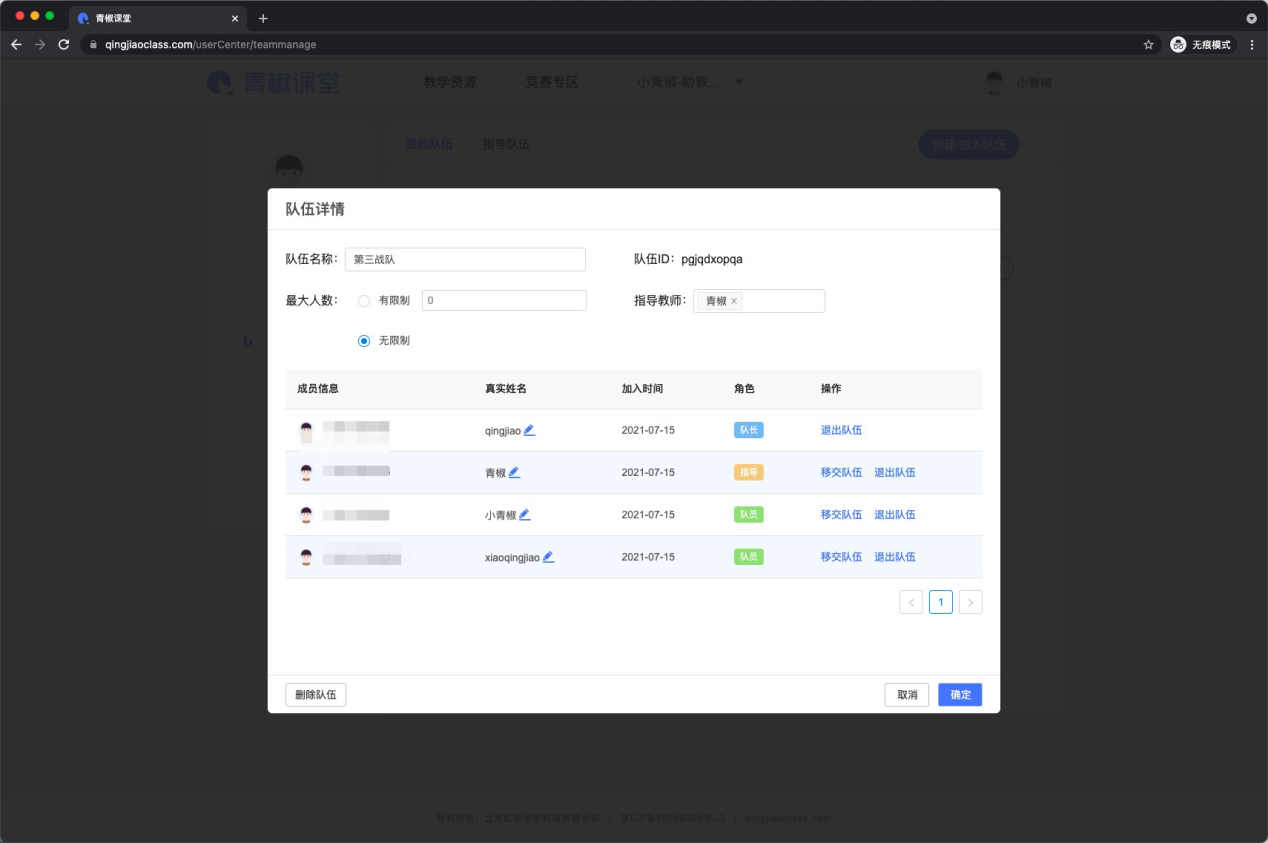
4、队伍创建成功后，队伍右上角出现邀请码，将邀请码发给同队人员即可;

5、同队人员加入队伍后，队长可以进行指定指导教师、转让队长身份等操作。（队长身份转让后需重新设置指导教师）

6、每支参赛队由3名选手和1名指导教师组成，（2支队伍可以重复指导教师）。





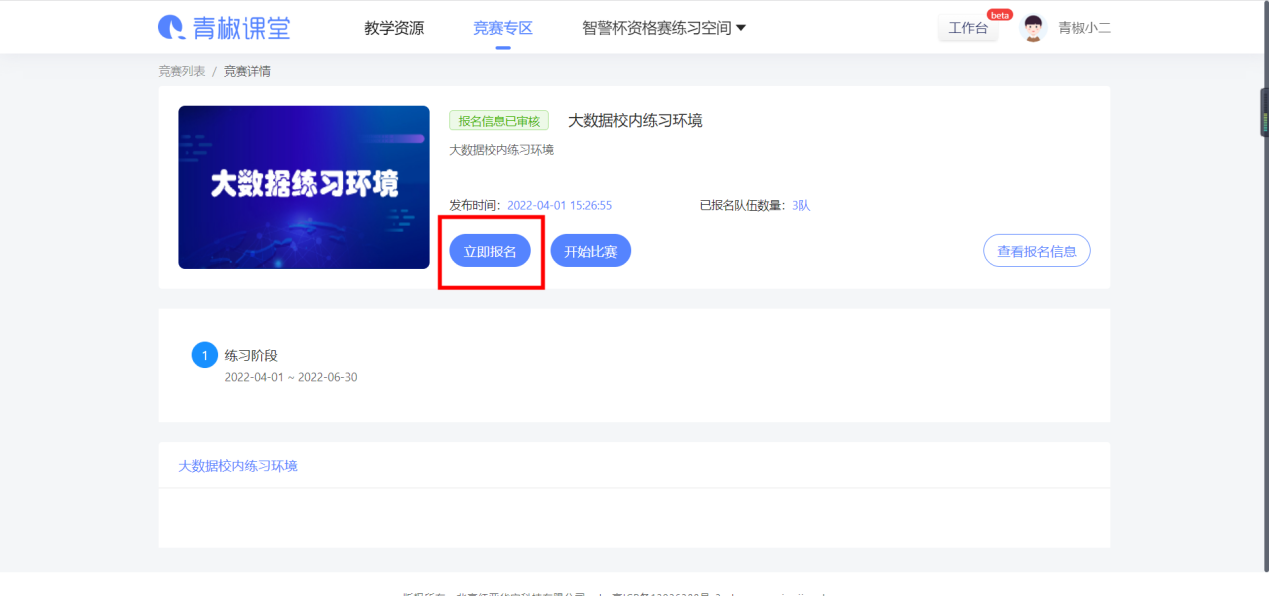


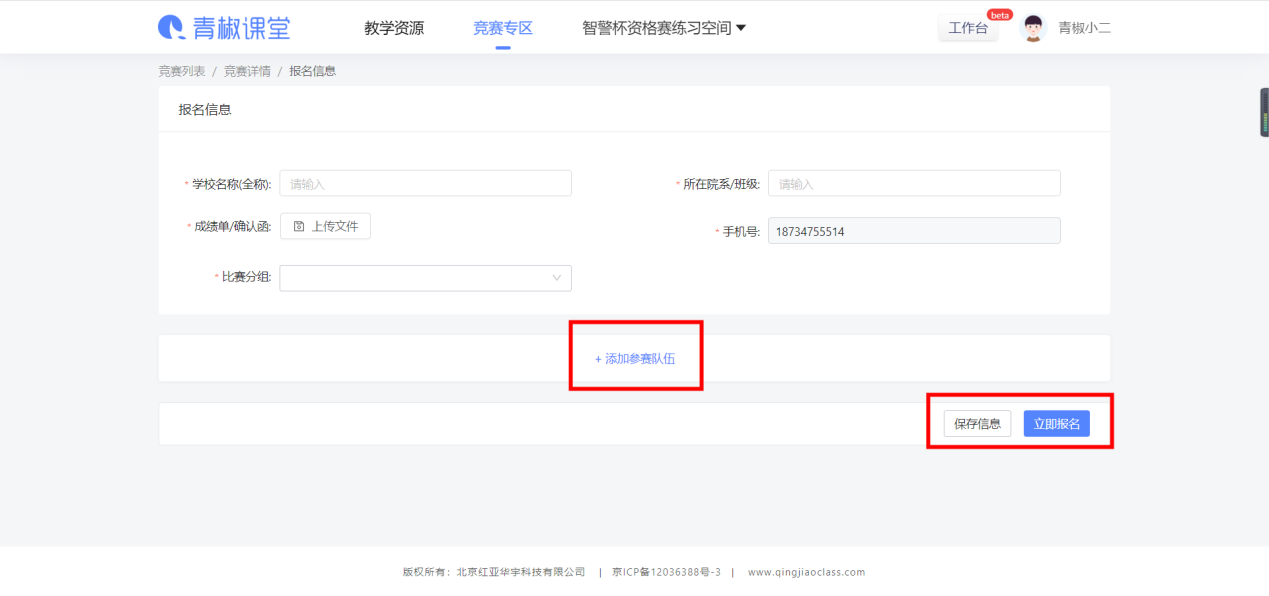
**三、团队-竞赛报名**

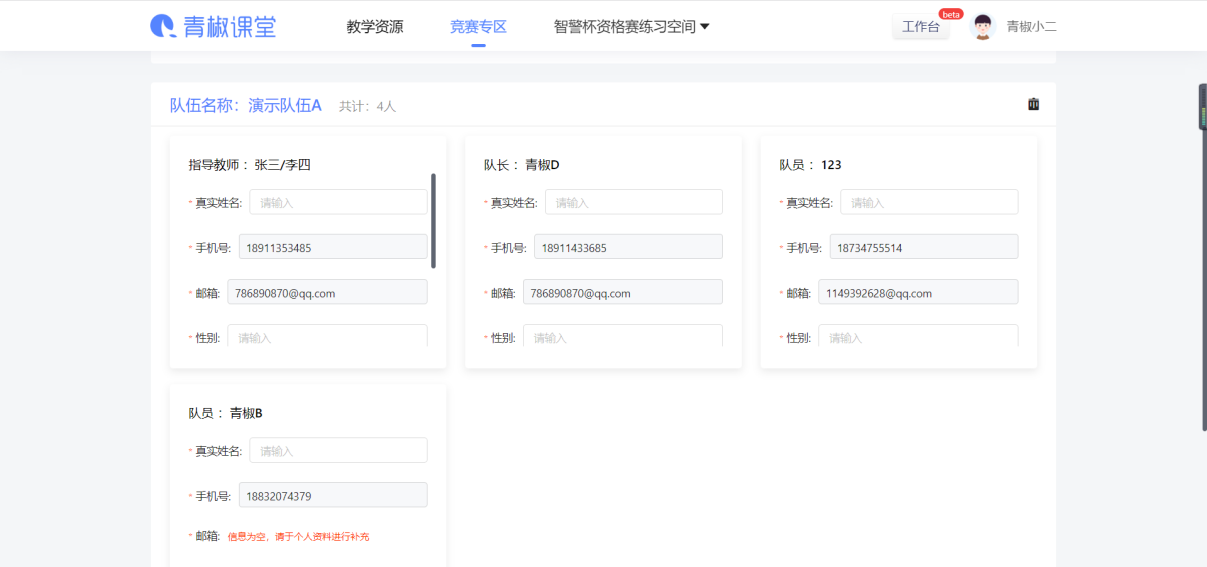
请团队中一位代表成员，由老师或竞赛负责人发放给大家的竞赛链接，登录青椒平台后，访问链接进行竞赛报名。

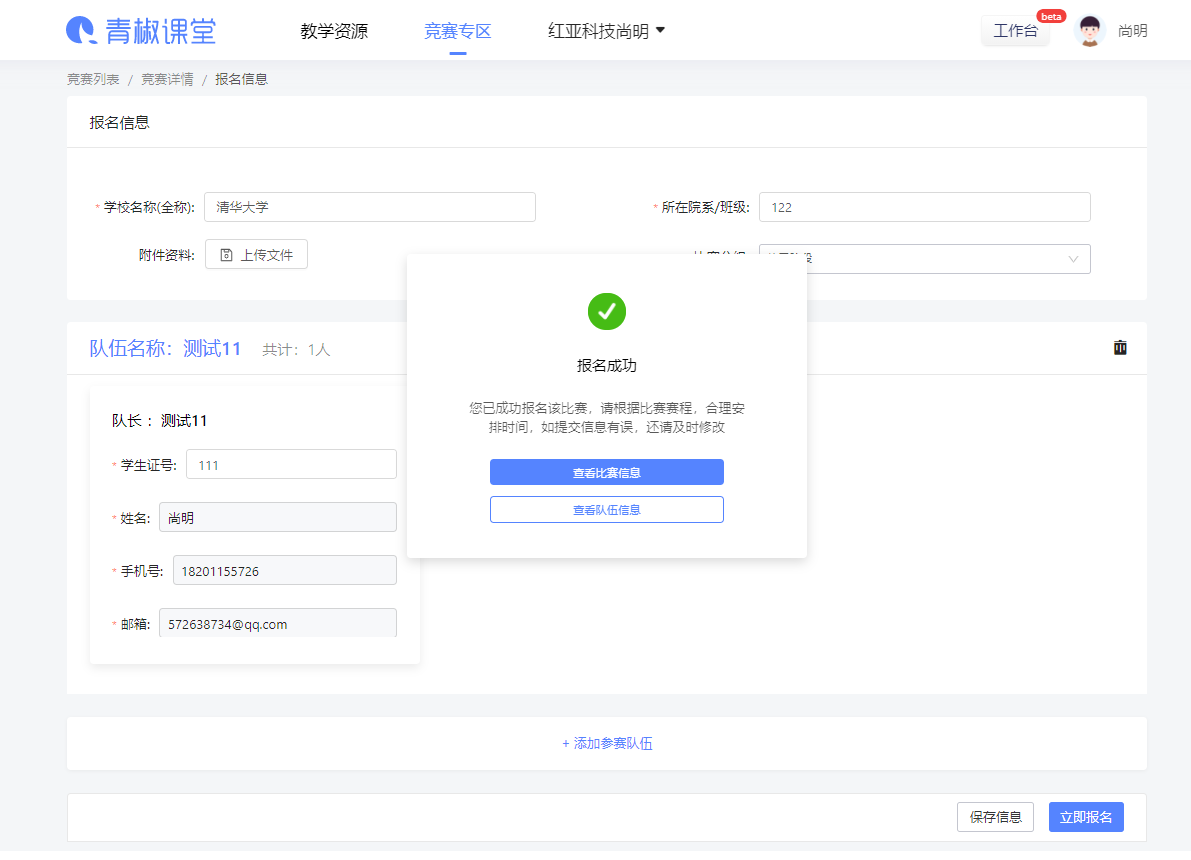
1、进入竞赛页面点击立即报名，进行信息填报、添加队伍、补充信息。

2、点击“立即报名”，即可完成竞赛报名。

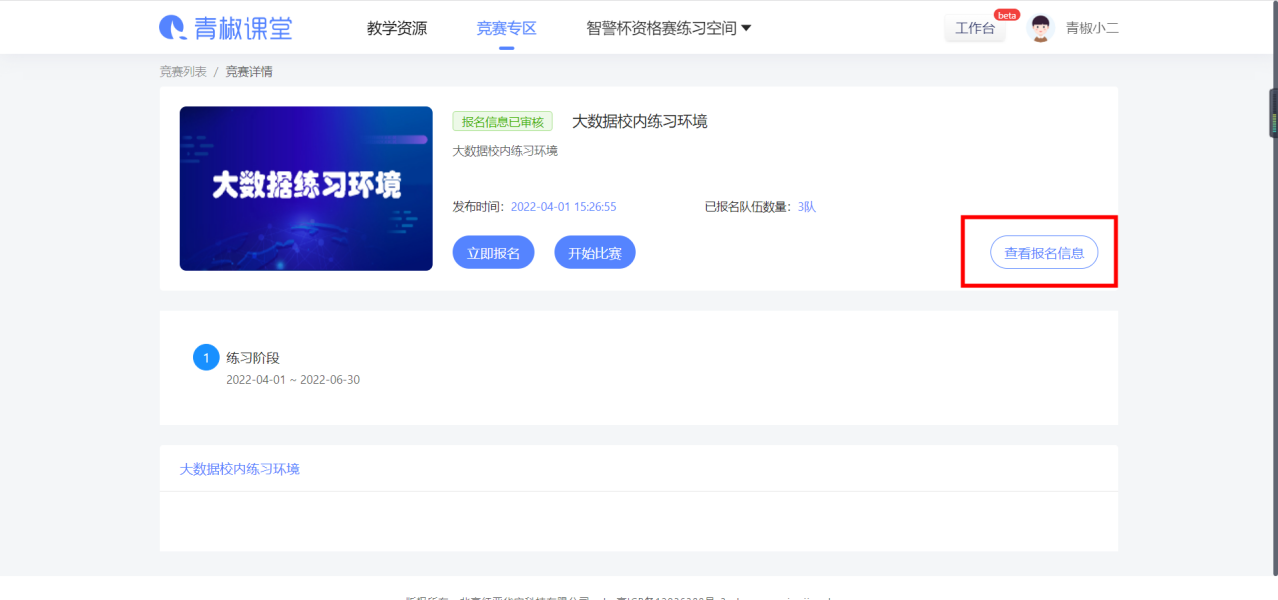


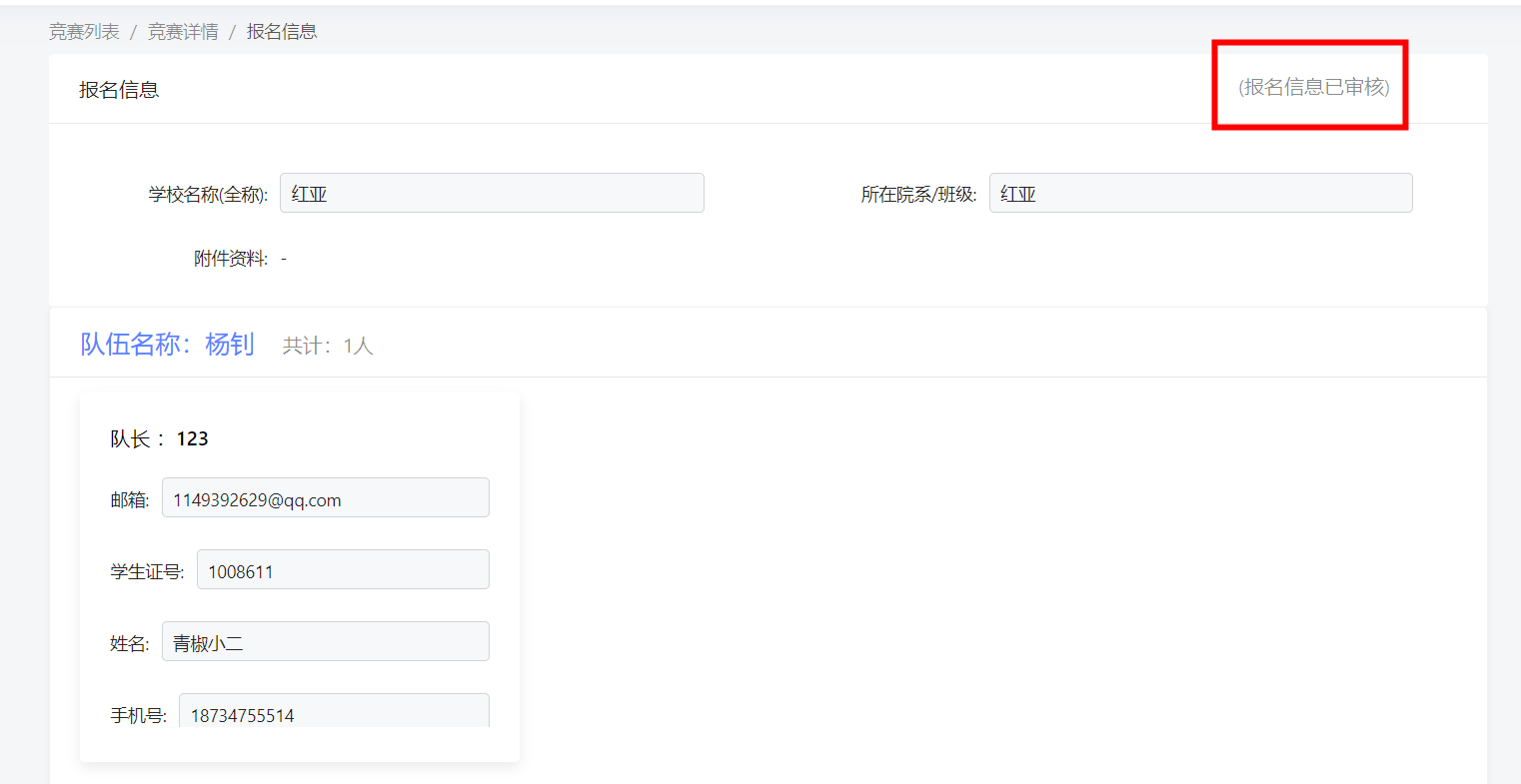






3、等待竞赛负责人会进行报名审核，缺少信息，将驳回报名，请保证报名信息的完整、正确。过程中可以通过查看队伍信息，了解审核状态。





4.审核将在报名截止后进行，报名截止前未审核状态属于正常状态。预祝参赛选手取得佳绩。