

# 2021 年（第 14 届）中国大学生计算机设计大赛

## 大数据主题赛（直报赛区）赛题

### 一、 比赛题目

#### 题目描述：数据解读国家规划与中国经济发展

改革开放以来，中国经济和社会经历了一轮又一轮的发展和提升。我们从关注 GDP 增长到关注碳中和，从关注对外贸易逆差到关注国内大循环，从关注外汇储备到关注人民币国际化。政策的规划与转变的背后是基于数据的决策；政策的决策与执行又会产生新的数据，而这些新的数据所带来的洞见能验证这个政策规划的执行情况。从洞见中见微知著，从宏观规划上找寻数据依据。比如：

- 商业规划：有人蹲点观测瑞幸咖啡门店的客流量来推测瑞幸咖啡真正的业务收入；
- 生活需求：有人从中国人口的变化和年龄变化来预测房价的变化和购房的需求；
- 国际关系：有人从收集义乌给美国大选做订单的数据来预测美国大选的结果。

每一个政策的转变都会带来新的数据关注点，而每一个新的数据洞察也会带来新的政策变化。基于给出的数据集，自立题目，运用数据思维，发现与解释经济社会现象，探讨与解决社会问题。

#### 数据说明：

本次比赛提供以下数据集，参赛选手必须选取本次比赛提供的一个或多个数据集作为基础数据。同时，鼓励选手补充其他数据集以充实数据分析，但需要提供完整的数据集和描述。

- [国民经济和社会发展第十二五、十三五、十四五规划全文](#)
- 国民经济数据 [Part 1](#) & [Part 2](#)
- [全国各省份结婚登记和离婚数据](#)
- [2020 年世界卫生组织统计数据](#)

### 二、 作品提交要求

#### 1. 提交要求

参赛选手最终至少需提交两份材料：一份**主题分析报告**和一份**数据处理源程序**。

**主题分析报告：**围绕题目撰写分析报告，运用数据思维，发现与解释经济社会现象，探讨与解决社会问题。**以 PDF 形式提交。**

**数据处理源程序：**从技术应用的角度描述数据来源、数据处理、数据分析方法。另外，分析报告中的分析结果（包括并不限于图表、统计数字等）均应在**程序源代码注释或说明**中列出详细的分析步骤，否则相应的结果不计入作品评审中。采用 Python 语言的源程序，请在主题赛技术支持平台进行**在线分析和提交**；采用其他开发和分析工具的源代码或源文件，请以附件形式提交。

**数据集：**如引入其他支持数据集，则需提交使用的数据集。

### 三、 比赛日程

#### 比赛报名时间

报名开始时间：2021 年 4 月 25 日

报名截止时间：2021 年 5 月 20 日（23:59:59）

组队截止时间：2021 年 5 月 20 日（23:59:59）

#### 初赛阶段

2021 年 4 月 25 日（10:30:00）至 2021 年 5 月 31 日（23:59:59）

- 提交截止时间：2021 年 5 月 31 日（23:59:59）
- 初赛评审阶段：2021 年 6 月 1 日至 2021 年 6 月 7 日
- 公示：2021 年 6 月 8 日 - 2021 年 6 月 20 日

形式：线上赛

#### 决赛阶段

初步计划于 2021 年 7 月 17 日-7 月 21 日之间进行，以最终版本为准

- PPT 提交截止时间：7 月 10 日（20:00）

形式：现场答辩（如遇特殊情况，将改为线上）

地址：上海市松江区人民北路 2999 号 东华大学

### 四、 支持单位

#### 主办：

中国大学生计算机设计大赛组委会

#### 承办：

东华大学

#### 分析工具支持：

Heywhale 和鲸

#### 培训支持：

深圳市深科数据科技有限公司

### 五、 其他

大数据主题赛相关的更多细节信息，请持续关注：

- 赛事 QQ 交流群：878606734
- 分赛区承办方官方主页