2023 年（第 16 届）中国大学生计算机设计大赛
大数据主题赛——“信创大数据”赛题说明

一、比赛题目

题目：上海旅游数据分析

信息技术应用创新发展是目前的一项国家战略，也是当今形势下国家经济发展的新动能， 信创产业发展已经成为经济数字化转型、提升产业链发展的关键，信创技术也经成为学生创新能力培养的重要方面。

在赛题指定的信创环境下，以给定的上海旅游统计数据集为基础，利用开源大数据组件搭建大数据系统，实现大数据 Web 应用的设计和开发，可包含数据可视化分析、决策辅助支持、离线和实时推荐等功能。

1.信创环境要求

主办方提供系统软件供参赛选手下载安装（提供下载包和镜像，镜像含必选软件）。

1）必选软件：

操作系统：统信桌面操作系统 V20（教育版），统信服务器操作系统 V20

中间件：TongWeb 应用服务器，TongHttpServer 负载均衡

数据库：达梦数据库 DM8

2）可选组件：

TongRDS 分布式数据缓存，TongLINK/Q 消息中间件，TongETL 数据处理

2.数据集说明

本次比赛提供以下数据集，参赛选手必必须选取其中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大数据应用开发，但需要提供完整的数据集和描述。

1) 上海旅游统计数据-本市接待入境旅游者人数（人次）
2) 上海旅游统计数据-旅游饭店客房平均出租率和平均房价
3) 上海旅游统计数据-入境外国人按主要客源国分（人次）
4) 上海旅游统计数据-旅行社国内旅游组团人数和接待人数（人次）
5) 上海旅游统计数据-旅行社外联、接待入境旅游人数（人次）

数据集下载地址：

https://pan.baidu.com/s/1DxxDcnavFfU3GhOII9oc9A?pwd=rthz 提取码：rthz

数据来源：上海市公共数据开放平台

3.信创软件下载及技术资料

统信操作系统 V20 和达梦数据库软件 DM8 通过企业官网链接下载，东方通中间件从百度网盘下载。三类软件的相关技术资料从百度网盘下载。为了减化软件安装过程，百度网

盘还提供了一个虚拟机镜像文件，已安装部署赛题要求的必选软件，可以直接使用。

百度网盘地址如下：

https://pan.baidu.com/s/1DxxDcnavFfU3GhOII9oc9A?pwd=rthz 提取码：rthz

各类软件下载方式详细说明如下：

① 统信操作系统软件：

https://cdimage-download.chinauos.com/uniontechos-server-20-1050a-amd64.iso?token=eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJleHAiOjE2NjUxOTg4OTgsImZpbGVfbmFtZSI6InVuaW9udGVjaG9zLXNlcnZlciOyMCOxMDUwYS1hbWQ2NC5pc28ifQ.OwEMc8vy2Ukm23HZErjEOrz7o9jDcpAyKo3YlWQ9ojk

统信操作系统技术资料：百度网盘下载。

②达梦数据库及技术文档：https://eco.dameng.com/download/

选择 DM8 安装包：开发版（信创平台），CPU：选择对应的 CPU（如果是 X86 可选兆芯），操作系统：统信 Uos20

达梦数据库部署规范：百度网盘下载。

③东方通中间件及技术资料：百度网盘下载。

④赛题信创环境系统虚拟机镜像文件：百度网盘下载。

二、作品提交要求

参赛选手需提交的材料：大数据系统设计方案、源码及部署包、大数据应用展示文件、自增的数据集。

- **作品设计报告：**参考国赛大数据/人工智能赛作品报告模板。文档中要包含已部署可访问的 Web 网站链接。

- **Web 网站源码及部署包：**系统源码及部署包。
- **作品展示视频：**初赛选手采用视频（mp4）演示大数据应用系统运行和功能；决赛选手需在参赛现场做汇报和演示（如遇特殊情况或改为线上演示）。
- **数据集：**如引入其他支持数据集，需提交使用的数据集或部分样本。

三、比赛日程

省赛阶段

- 选手下载主办方提供的信创环境软件或镜像，进行信创环境的搭建；
- 在自己搭建的信创环境上，根据题目要求，进行大数据应用开发；
- 按校赛或省赛规定的报名要求、时间和提交方式提交作品，由校赛或省赛负责竞赛的组织和评审。

国赛阶段

- 国赛决赛于 2023 年 7 月 17 日-7 月 21 日进行，以国赛公告为准。
- 决赛形式：答辩

作品评审：

作品评审参考大数据实践赛评审标准，主要从主题创意、功能效果、技术实现和作品展示四个方面进行评价。注意作品开发必须符合本赛题规定的信创环境和基础数据要求。

四、竞赛培训

为使参赛选手尽快熟悉信创环境，赛题方将为参赛选手提供竞赛培训。培训内容包括：统信操作系统、达梦数据库、东方通中间件，介绍相关产品的安装及应用部署方法、信创环境下大数据开发的注意事项，大数据应用设计要点等。

五、大赛组织

主办：中国大学生计算机设计大赛组委会

承办：东华大学

协办：上海计算机软件技术开发中心

武汉达梦数据库股份有限公司

统信软件技术有限公司

北京东方通科技股份有限公司

联系方式：

- 赛题咨询： 乐老师 邮箱：lqh@sscenter.sh.cn

- 技术支持 QQ 群：645675570（申请加入请注明：学校_姓名）