### 作业要求

### 1、作业内容Assignment content

请使用提供的数据集https://github.com/awesomedata/awesome-public-datasets或从网络上爬取自己感兴趣的数据完成数据挖掘任务。

### 作业基本要求Assignment requirements

1. 必须包括数据预处理。根据分析内容设计数据预处理的策略和方法，完成数据预处理。

2)必须包括探索性分析并实现统计可视化。

3)明确数据挖掘任务，选择或设计合适的数据挖掘算法完成数据挖掘任务并可视化数据挖掘结果。该部分需要介绍使用的平台或工具。可以调用相关机器学习库也可以直接编写算法程序。**若有创新的算法程序实现请注明**。

4)设计评估方案，对数据挖掘结果评估并给出评估结果。

5）可以使用相关AI工具辅助数据挖掘分析工作。

### 三、作业提交形式Assignment submission

1. 数据挖掘报告。明确数据挖掘任务、需要使用的数据、工具，设计技术路线和人员分工。阐述数据收集、数据预处理、探索性分析、数据挖掘、结果评估各阶段的工作成果。此外报告中需要明确小组各成员的工作量。
2. 汇报PPT及视频。制作数据挖掘大作业的汇报ppt并录制讲解视频
3. 程序代码。
4. 1-5人一组，最多不超过5人。

### 四、作业提交截止日期2025.11.15日