# 第六组-毛弘宇-2020141461142-第六周周报

本周工作：

1、对部分数据进行了清洗处理；

2、爬取了热门歌曲的名称、播放量、评论数等数据；

3、建立了部分数据库；

4、创建了项目管理库GitHub。

下周计划：

1、完成数据清洗工作；

2、进一步完善数据库；

3、爬取歌曲评论、IP等数据。

存在的问题：

1、数据库设计存在一些问题；

2、动态数据难以抓取，需要转换为静态数据；

3、JSON格式数据抓取有困难，需要建立JSON库；

4、数据清洗逻辑有错误，正在修复中。