**数据爬虫引擎需求文档**

1. **数据源**：

美团外卖，百度外卖，饿了么等三个平台移动端的数据（注意，这三个平台移动端和PC端的数据并不一致）

1. **数据爬取场景：**

系统提交一个位置信息，然后爬虫爬取该位置信息的任意数量的店铺数据，同时抓取这些店铺的相关数据，爬取需求是实时提交的，爬取数据要存到表里，处理后呈现在前端；

以美团平台为例：

1. 爬取店铺数据（红框和黄色框是需要爬取的数据）：



1. 爬取店铺活动（红框为需要爬取的数据）



1. 爬取店铺的排名前几的热销菜品（红框为所需数据）；



1. 爬取店铺排名前几的热销菜品的折扣价格（红框为所需数据）：



1. **还有用户评论数据(图省略)**
2. **系统性能要求要求：**
3. 因为这些平台有反爬虫策略，也有专门的反爬虫工程师，所以爬虫系统也要针对这些反爬虫方案有应对措施；
4. 由于用户对数据的拉取是实时的，所以需要系统有高度的稳定性，对于延迟的容忍度低，所以整套爬虫引擎需要高度的稳定，高效；