# 一背景介绍

## 1项目背景

餐饮行业作为我国第三产业中的一个传统服务性行业，始终保持着旺盛的增长势头，取得了突飞猛进的发展，展现出繁荣兴旺的新局面。与此同时，我国餐饮业发展的质量和内涵也发生了重大变化。根据国家统计局数据显示，餐饮行业餐费收入从2006到2015年都处于增长的趋势，但是同比增长率却有很大的波动。

目前，餐饮企业正面临着同行竞争压力大、对消费者的需求认识不到位、客户忠诚度低、菜品和服务缺乏个性化的几个问题。

## 2项目数据与目标

某餐饮企业的系统数据库中积累了大量的与客户用餐相关的数据，包含菜品详情表、订单表、订单详情表等。本项目对该餐饮企业的订单数据进行分析，利用协同过滤算法对就餐的客户进行菜品智能推荐。

## 3项目实操分析

1.1读取订单数据，并进行探索

使用read\_csv读取的订单详情表的长度为： 10037

订单详情表的元素个数为： 190703

订单详情表的维度数为： 2

订单详情表的形状为： (10037, 19)

1.2处理菜品名称字符数据

将菜品名称中的其他字符删除，保留菜品名称

1.3构建热销度评分指标

经Min-Max标准化后，最高分为1分，最低分为0分。

1.4热销top10

订单详情表dishes\_name频数统计结果前10为：

白饭/大碗 323

凉拌菠菜 269

谷稻小庄 239

麻辣小龙虾 216

辣炒鱿鱼 189

芝士烩波士顿龙虾 188

五色糯米饭(七色) 187

白饭/小碗 186

香酥两吃大虾 178

焖猪手 173

Name: dishes\_name, dtype: int64

2.1统计每个订单状态的占比

2.2选取有效订单数据

2.3选取主要特征

3.1将数据划分为训练集和测试集

3.2构建训练集数据客户-菜品二元矩阵

4.1计算相似度矩阵

4.2生成推荐列表

5.1构建测试集客户IP-菜品字典

5.2构建评价指标，并分析推荐结果