

In [3]:

```
import pandas as pd
import numpy
import matplotlib.pyplot as plt
```

In [10]:

```
data = pd.read_excel('./data/mt.xls', names=['user', 'comment', 'commenttime', 'userid', 'star', 'menu'])
# data.head(n=10)
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1539 entries, 0 to 1538
Data columns (total 6 columns):
#   Column          Non-Null Count  Dtype  
---  --
0   user             1539 non-null   object 
1   comment          1539 non-null   object 
2   commenttime      1539 non-null   int64  
3   userid           1539 non-null   int64  
4   star             1539 non-null   int64  
5   menu             1398 non-null   object 
dtypes: int64(3), object(3)
memory usage: 72.3+ KB
```

## 项目思路

- 0 user 1539 non-null object 用户分析
- 1 comment 1539 non-null object 评价分析
- 2 commenttime 1539 non-null int64 销售量分析
- 3 userid 1539 non-null int64 用户分析
- 4 star 1539 non-null int64 评价分析
- 5 menu 1398 non-null object 套餐销售分析

## 用户分析的建议（饼状图）

1. 可以以用户的消费次数入手
2. 用户的月两次购买客户或n次（绘制饼状图）计算回购率与复购率等销售指标。

## 销售量分析的建议（折线图为主）

1. 统计销售量
2. 窗口期的问题 按天算 整体数据进行可视化反应的大的趋势
3. 需要将时间戳数据转换为时间字符串数据
4. 按时间轴划分

## 评价分析的建议（词云图）

1. 字符串处理去除无语义的词汇或者符号（字符串处理时难点）针对难点给定技术建议（1.去停用词、字符串处理）
2. 生成评价词云图表（总词云）消极词云（筛选评分25以下的用户进行评价词云展示）

# 套餐销售分析的建议（柱状图）

- 1. 数据统计
- 2. 柱状图可视化

In [14]:

```
data.head(n=3)
```

Out[14]:

	user	comment	commenttime	userid	star	menu
0	meiki8182	菜品味道从来没让我失望过，环境更是一流，领孩子也跟舒适！将军牛排，九折板，那绝对是必点菜式，...	1653720234	4448597	50	仅售93元！价值100元将军牛排韩国餐厅代金券1张
1	bzE766087788	一直想着再来吃一次，今天终于来了。上菜速度一如既往的快，黑椒牛肉饭真的超好吃，不油不腻，梨大...	1654359586	977700597	50	特色4人套餐，提供免费WiFi
2	花自飘零372	喜欢吃他家食物，跟朋友经常去，喜欢杂酱面，拌饭，牛肉，鸡爪都是不错的海鲜饼也好吃，辣口的菜也...	1652168718	47117506	50	NaN