

实例分析

餐饮客户价值分析：

- 部分餐饮客户的消费行为特征数据如下，根据数据将客户分成不同客户群，并评价这些客户群的价值

	A	B	C	D	E
1	Id	R	F	M	
2	1	27	6	232.61	
3	2	3	5	1507.11	
4	3	4	16	817.62	
5	4	3	11	232.81	
6	5	14	7	1913.05	
7	6	19	6	220.07	
8	7	5	2	615.83	
9	8	26	2	1059.66	
10	9	21	9	304.82	
11	10	2	21	1227.96	
12	11	15	2	521.02	
13	12	26	3	438.22	
14	13	17	11	1744.55	



餐饮客户价值分析：

□ Python代码

```
1  -*- coding: utf-8 -*-
2  #使用K-Means算法聚类消费行为特征数据
3  import pandas as pd
4  #参数初始化
5  inputfile = '../data/consumption_data.xls' #销量及其他属性数据
6  outputfile = '../tmp/data_type.xls' #保存结果的文件名
7  k = 3 #聚类的类别
8  iteration = 500 #聚类最大循环次数
9  data = pd.read_excel(inputfile, index_col = 'Id') #读取数据
10 data_zs = 1.0*(data - data.mean())/data.std() #数据标准化
11 from sklearn.cluster import KMeans
12 model = KMeans(n_clusters = k, n_jobs = 1, max_iter = iteration) #分为k类，并发数4
13 model.fit(data_zs) #开始聚类
14 #简单打印结果
15 r1 = pd.Series(model.labels_).value_counts() #统计各个类别的数目
16 r2 = pd.DataFrame(model.cluster_centers_) #找出聚类中心
17 r = pd.concat([r2, r1], axis = 1) #横向连接（0是纵向），得到聚类中心对应的类别下的数目
18 r.columns = list(data.columns) + [u'类别数目'] #重命名表头
19 #详细输出原始数据及其类别
20 r = pd.concat([data, pd.Series(model.labels_, index = data.index)], axis = 1) #详细输出每个样本对应的类别
21 r.columns = list(data.columns) + [u'聚类类别'] #重命名表头
22 r.to_excel(outputfile) #保存结果
23 def density_plot(data): #自定义作图函数
24     import matplotlib.pyplot as plt
25     plt.rcParams['font.sans-serif'] = ['SimHei'] #用来正常显示中文标签
26     plt.rcParams['axes.unicode_minus'] = False #用来正常显示负号
27     p = data.plot(kind='kde', linewidth = 2, subplots = True, sharex = False)
28     [p[i].set_ylabel(u'密度') for i in range(k)]
29     plt.legend()
30     return plt
31 pic_output = '../tmp/pd_' #概率密度图文件名前缀
32 for i in range(k):
33     density_plot(data[r[u'聚类类别']==i]).savefig(u'%s%s.png' %(pic_output, i))
34
```



实例分析

餐饮客户价值分析：

□ 运行结果

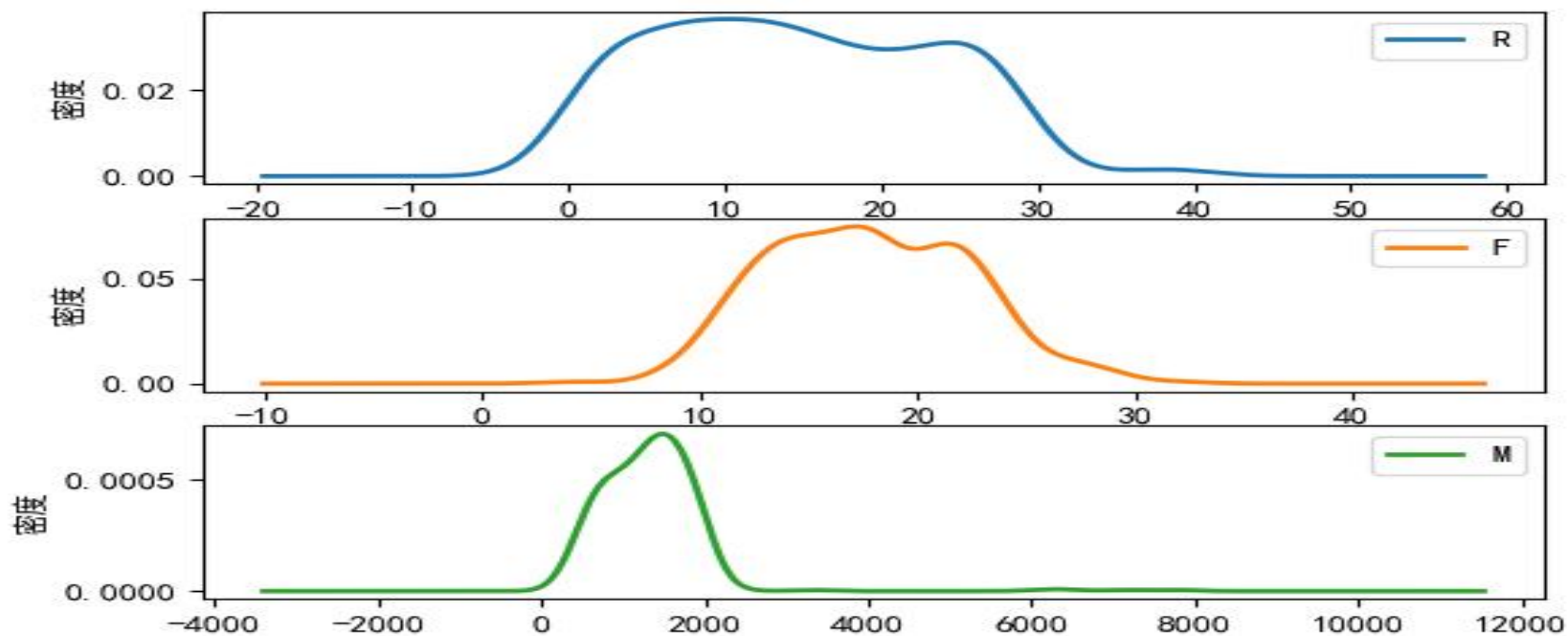
	A	B	C	D	E	F	
1	Id	R	F	M	聚类类别		
2	1	27	6	232.61	1		
3	2	3	5	1507.11	1		
4	3	4	16	817.62	0		
5	4	3	11	232.81	1		
6	5	14	7	1913.05	1		
7	6	19	6	220.07	1		
8	7	5	2	615.83	1		
9	8	26	2	1059.66	1		
10	9	21	9	304.82	1		
11	10	2	21	1227.96	0		
12	11	15	2	521.02	1		
13	12	26	3	438.22	1		
14	13	17	11	1744.55	0		
15	14	30	16	1957.44	0		
16	15	5	7	1713.79	1		
17	16	4	21	1768.11	0		
18	17	93	2	1016.34	2		



实例分析

餐饮客户价值分析：

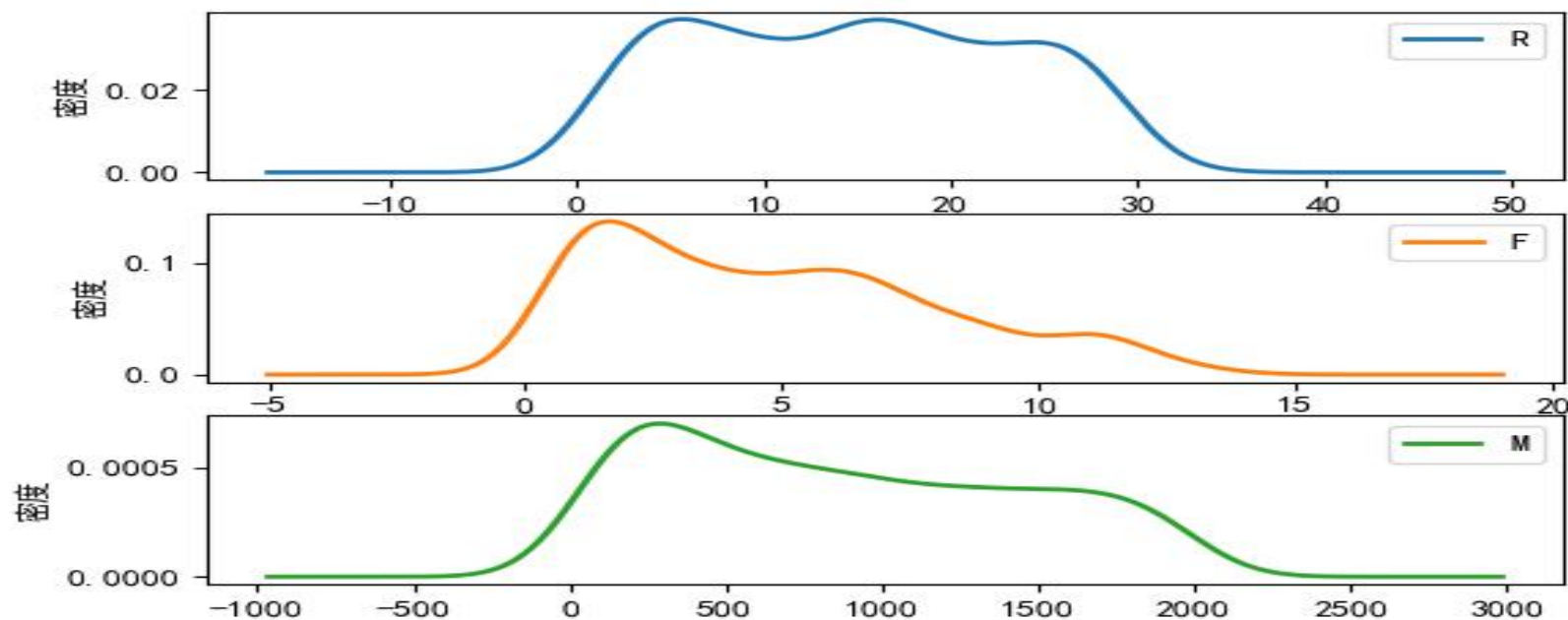
- 分群一结果分析：分群一的R间隔相对较小，主要集中在0~30天，消费次数集中在10~25次，消费金额在500~2000



实例分析

餐饮客户价值分析：

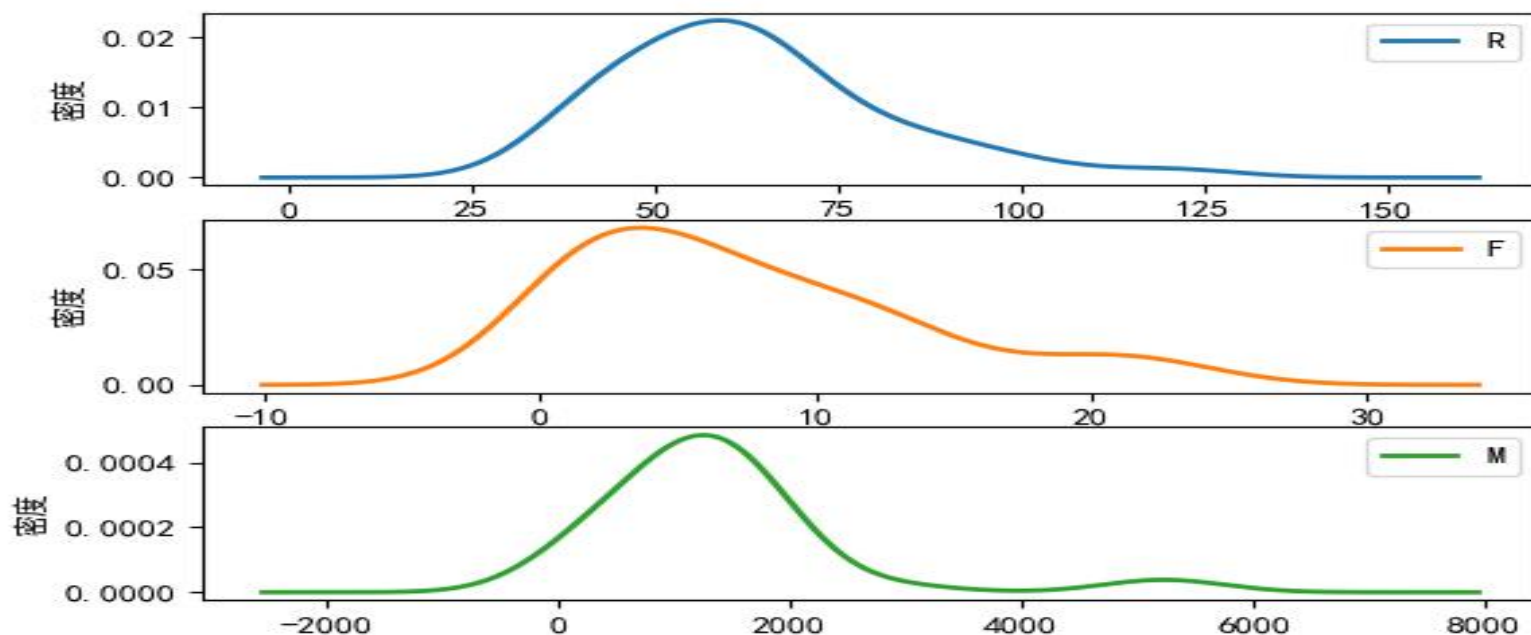
- 分群二结果分析：分群二的R间隔分布在0~30天，消费次数集中在0~12次，消费金额在0~1800



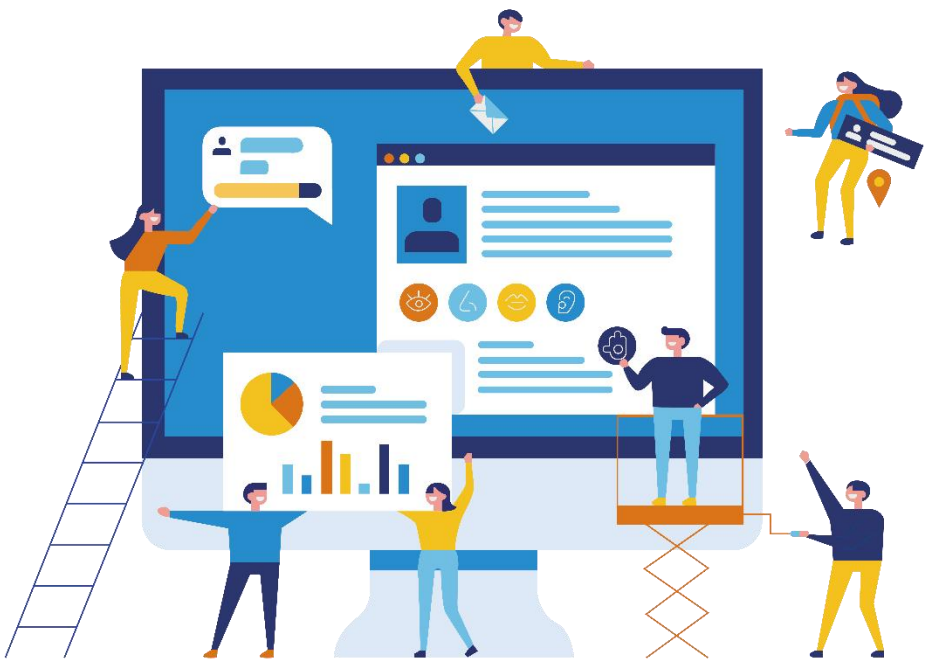
实例分析

餐饮客户价值分析：

- 分群三结果分析：分群三的R间隔较大，间隔分布在30~80天，消费次数集中在0~15次，消费金额在0~2000



实例分析



餐饮客户价值分析：

□ 对比分析

- 分群1的时间间隔较短，消费次数多，而且消费金额较大，时高消费、高价值人群。
- 分群二的时间间隔、消费次数和消费金额处于中等水平，代表着一般客户。
- 分群三的时间间隔较长，消费次数较少，消费金额也不是特别高，是价值较低的客户群体。

