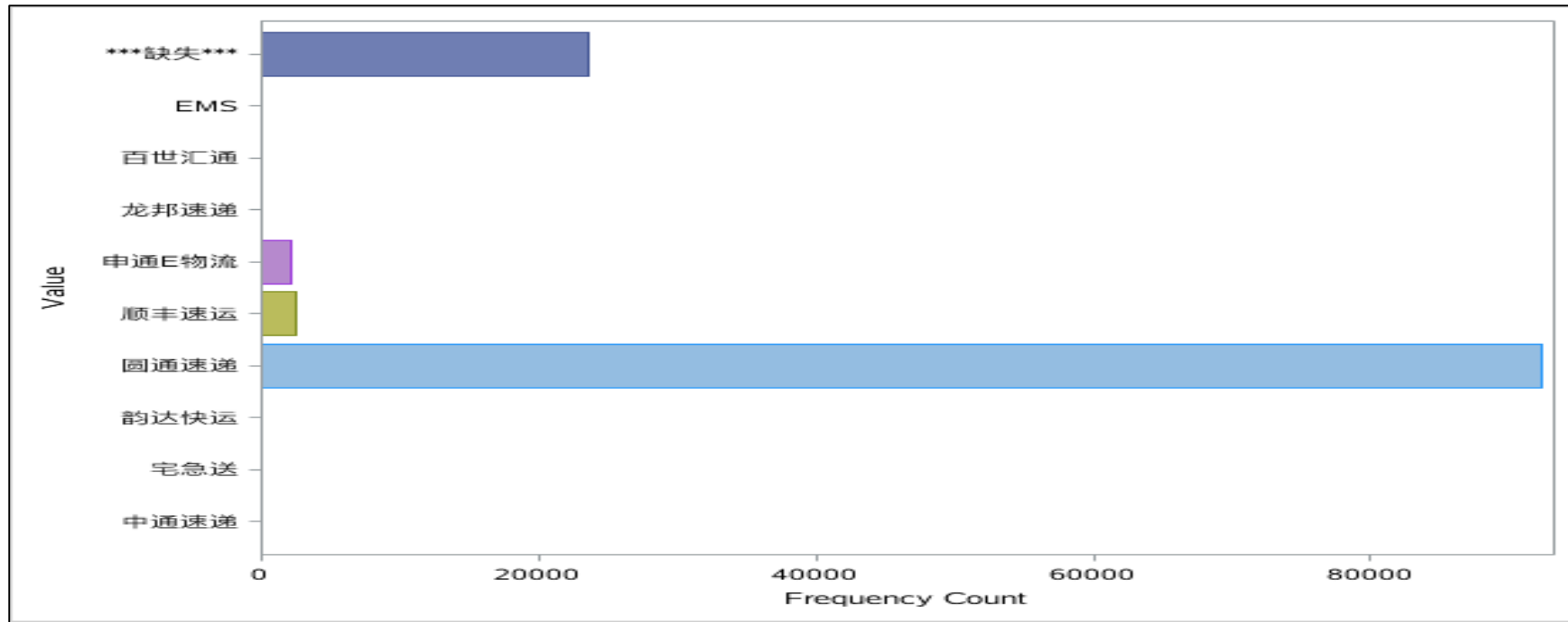


# 数据基本信息

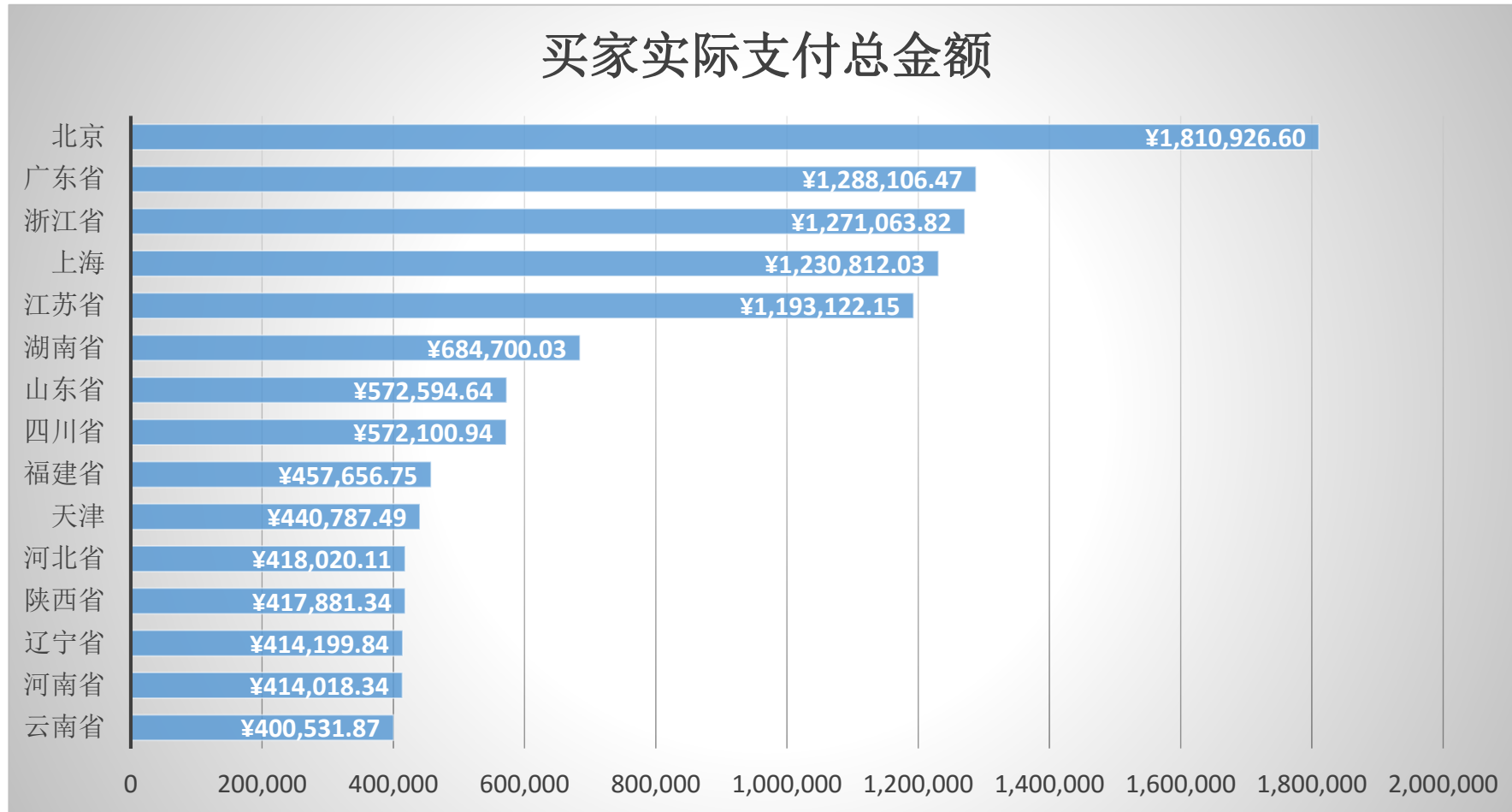
- 本次淘宝订单数据为2013年全年交易数据共**120,757**条，交易总金额为**RMB 19,662,721**元
- 共有**99,180**条订单有付款时间记录，而交易成功的订单数为**95,596**条，占全部的**79.16%**；交易成功总金额为**RMB 16,011,602元**，占全部的**81.43%**
- 通过对描述数据特征的分析，得出：**买家支付积分、买家实际支付积分、返点积分**都为0，对分析无参考意义，因此先筛选淘汰

# 物流基本信息

- 运送方式包括：快递、EMS、平邮、卖家承担运费，其中快递占约**99.90%**
- 物流公司存在约**19.60%**的缺失值，大部分由圆通快递承担，约占**76.45%**，顺丰、申通也仅占少数部分，分别为**2.14%**、**1.79%**

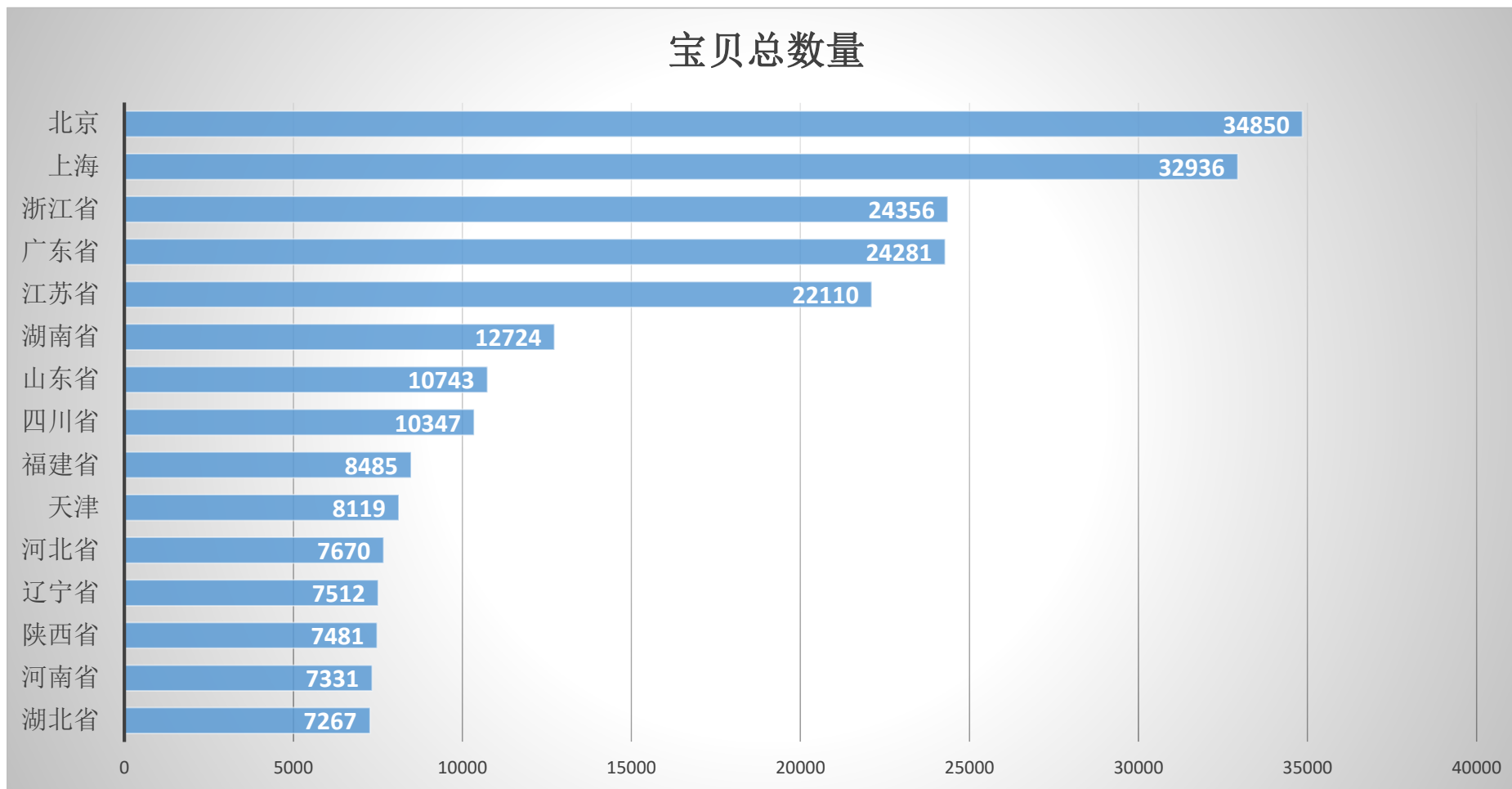


全年消费金额最高的15个省直辖市的条形图

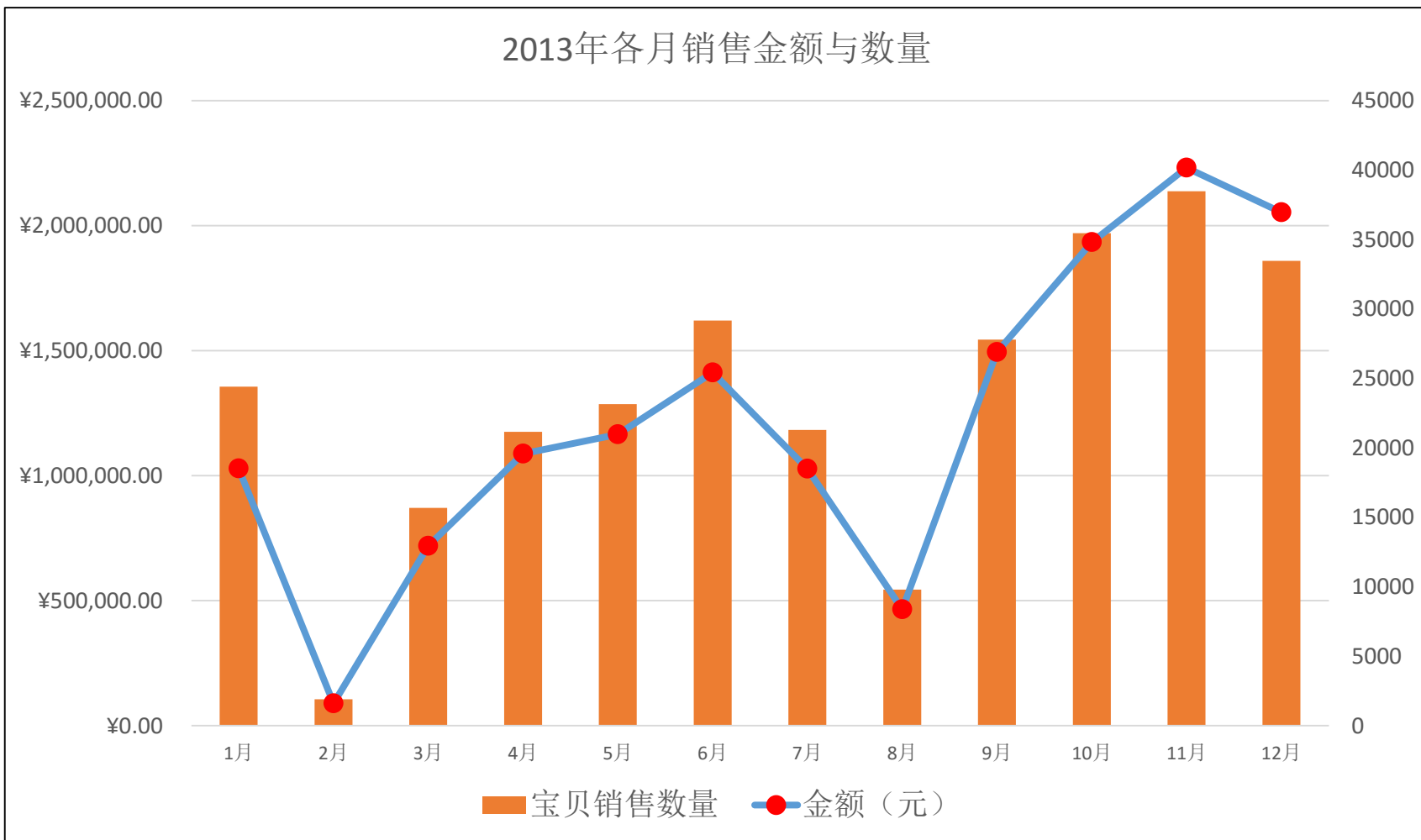


从上图中可看出，北京的消费金额是第1梯队要远高于其他地区，同时江苏，上海，浙江，广东4个地区的消费金额成为第2梯队，远高于除北京外的其他省和自治区。根据数据统计得出全国年的消费金额是¥14,715,880.96，消费金额前5的省总和为¥6,798,183.61，占总体的46.20%。

全年购买宝贝最高的15个省或自治区的条形图



# 2013年各月销售金额与数量



月份	金额 (元)	宝贝销售数量
1月	¥1,027,987.19	24412
2月	¥89,011.73	1899
3月	¥718,748.62	15677
4月	¥1,087,635.34	21156
5月	¥1,165,397.32	23156
6月	¥1,412,565.59	29160
7月	¥1,027,277.43	21281
8月	¥465,364.19	9796
9月	¥1,494,345.41	27785
10月	¥1,933,087.51	35451
11月	¥2,231,611.34	38463
12月	¥2,052,789.38	33454

## 对第4季度的交易成功商品的关联分析

SET_SIZE	置信度	支持度	提升度	COUNT	_LHAND	_RHAND
2	96.875	2.52169 2	1.22606 4	93	S特. 自留款 舒适超弹力 透气蕾丝提臀女士平角 内裤 -6608	S特. 冬季加厚款 纯棉 杂线翻边女袜短袜子 - 5953
2	92.9577 5	1.78958 8	1.17648 7	66	特. 2013秋冬装 韩版简 洁显瘦 针织女款包臀 裙子 -5933	S特. 冬季加厚款 纯棉 杂线翻边女袜短袜子 - 5953
2	91.0256 4	1.92516 3	1.15203 4	71	特. 韩版仿羊绒伞摆修 身显瘦圆领套头针织打 底毛衣 -6631	S特. 冬季加厚款 纯棉 杂线翻边女袜短袜子 - 5953

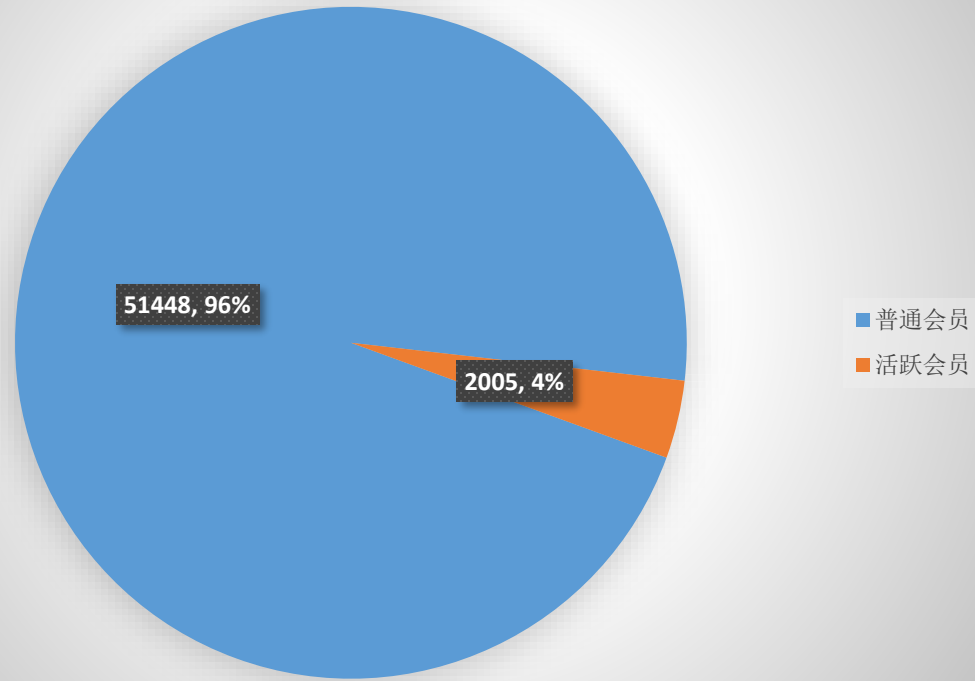
对第4季度交易成功的商品进行关联分析，发现许多商品都与商品**5953**存在较强的相关，可以把这些商品与商品**5953**进行捆绑销售，或者进行相关推荐。



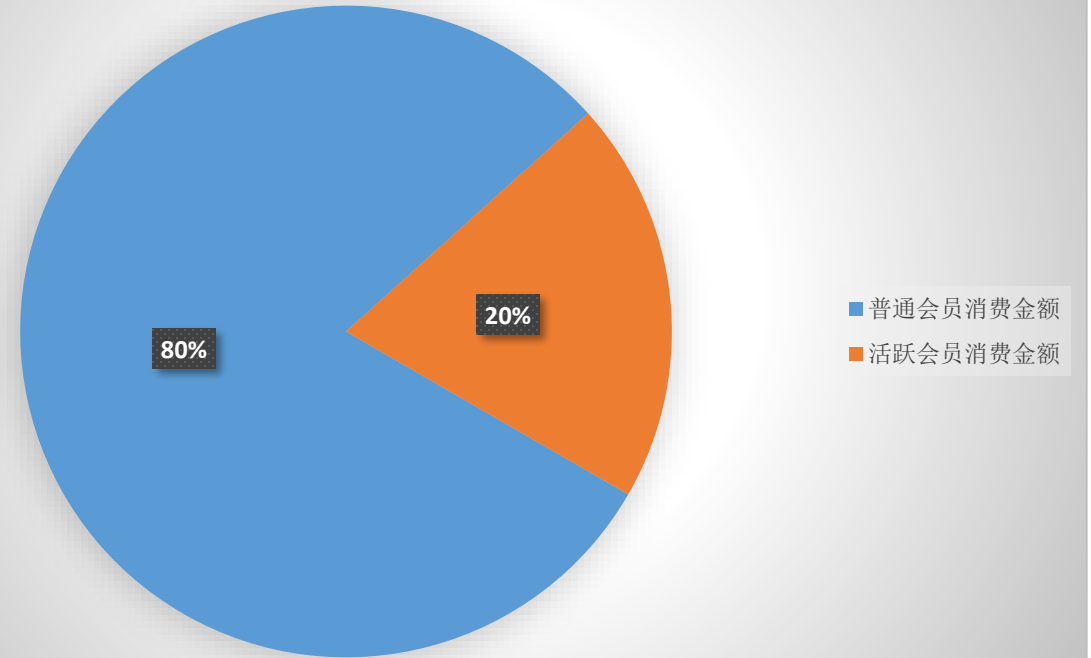
关联模型

# 会员情况

## 会员情况



## 会员消费金额情况



由左图可知，在2013年成功交易的会员中有4%（2005位）为活跃会员(即一年内购买次数大于10次)，即这部分会员在本店发生多次购买行为。其中活跃会员的消费金额占总体会员的20%，所占比例比较大。

## 对交易成功的订单客户建立RFM模型

买家会员名	订单付款时间	sum_jiaoyi	sum_cishu	time_diff	购买金额	购买频次	购买时间	rfm_score
c000000001	10/26/2013	2015.84	13	750	5	5	3	355
c000000002	11/15/2013	725.4	6	730	5	5	4	455
c000000003	05/02/2013	154.1	2	927	3	4	1	143
c000000004	09/26/2013	611.05	9	780	5	5	3	355
c000000005	12/23/2013	1638.9	7	692	5	5	5	555
c000000006	10/21/2013	1041.2	4	755	5	5	3	355
c000000007	05/02/2013	168	1	927	3	0	1	103
c000000008	11/25/2013	1427.2	14	720	5	5	4	455
c000000009	05/10/2013	253	3	919	4	5	1	154
c000000010	06/03/2013	353.38	2	895	4	4	1	144
c000000011	08/13/2013	467.8	2	824	5	4	2	245
c000000012	12/31/2013	74283.45	1246	684	5	5	5	555

对于2013年所有交易成功的客户建立RFM模型把全年购买总金额，总频数，最后一次购买时间至今时间间隔按大小分为0-5的6个等份（金额，频数从大到小按5-0打分，时间间隔按从大到小0-5打分）。最后将3列得分合并，数值越大的为越优质客户。在后续进行老客户的维系以及发展有价值的客户是能够更加准确。

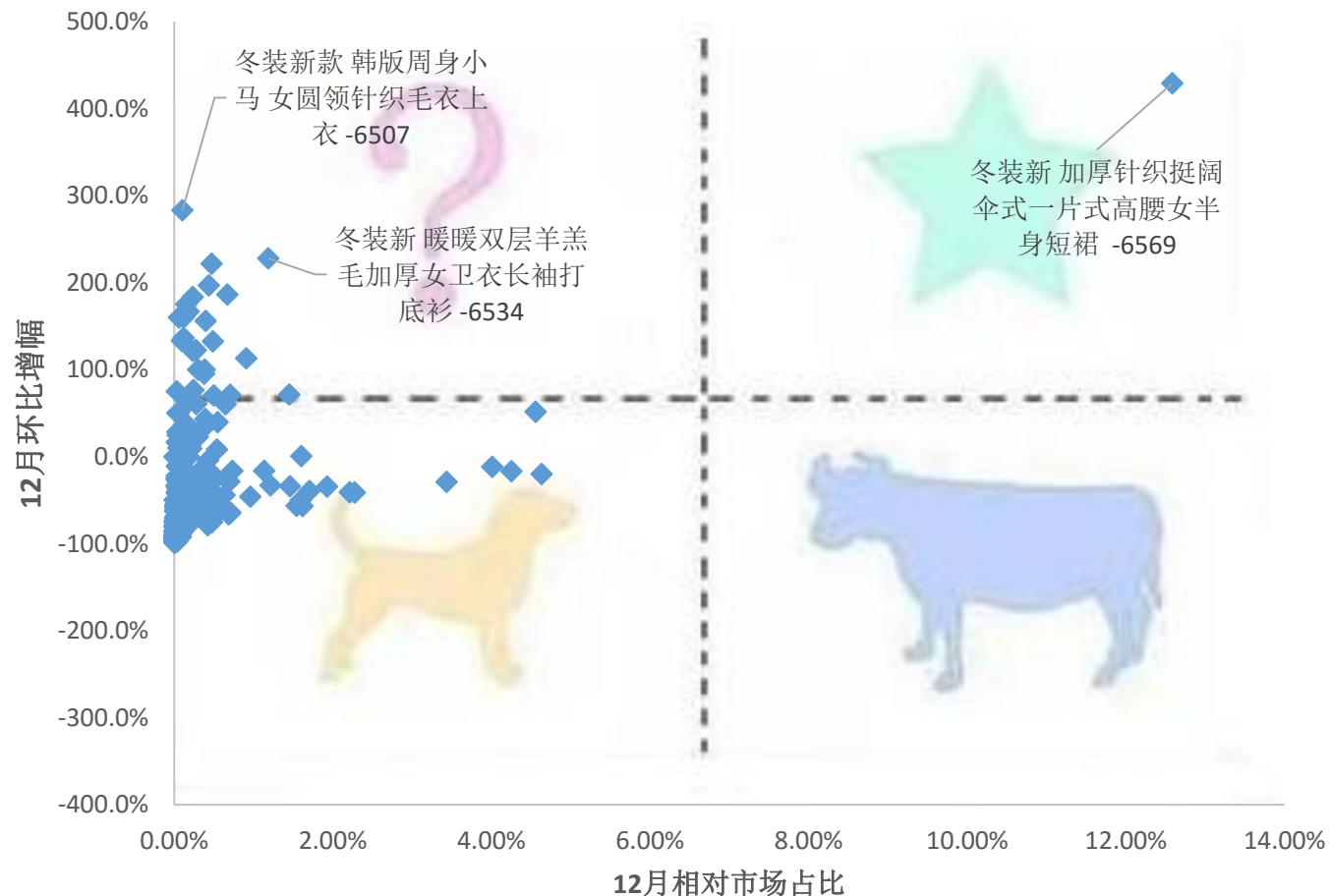


RFM模型交易成功得分



# 2013年12月对比11月产品—市场分析

波士顿矩阵



在12月中有**5%**的产品增长率大于**100%**;

其中增长率最高的产品为**6569**, 环比增长率为**429.1%**;

此产品也是本月的明星产品, 可以将其加大投资, 使其迅速发展, 增加市场占有率, 让其进入成熟期, 获得更大利润。

产品**6507**和**6534**它是处于高增长率、低市场占有率的产品群中, 利润率较低, 市场机会大, 前景好, 但是在市场营销上存在问题。

# 关于是否包邮对交易是否成功的影响分析

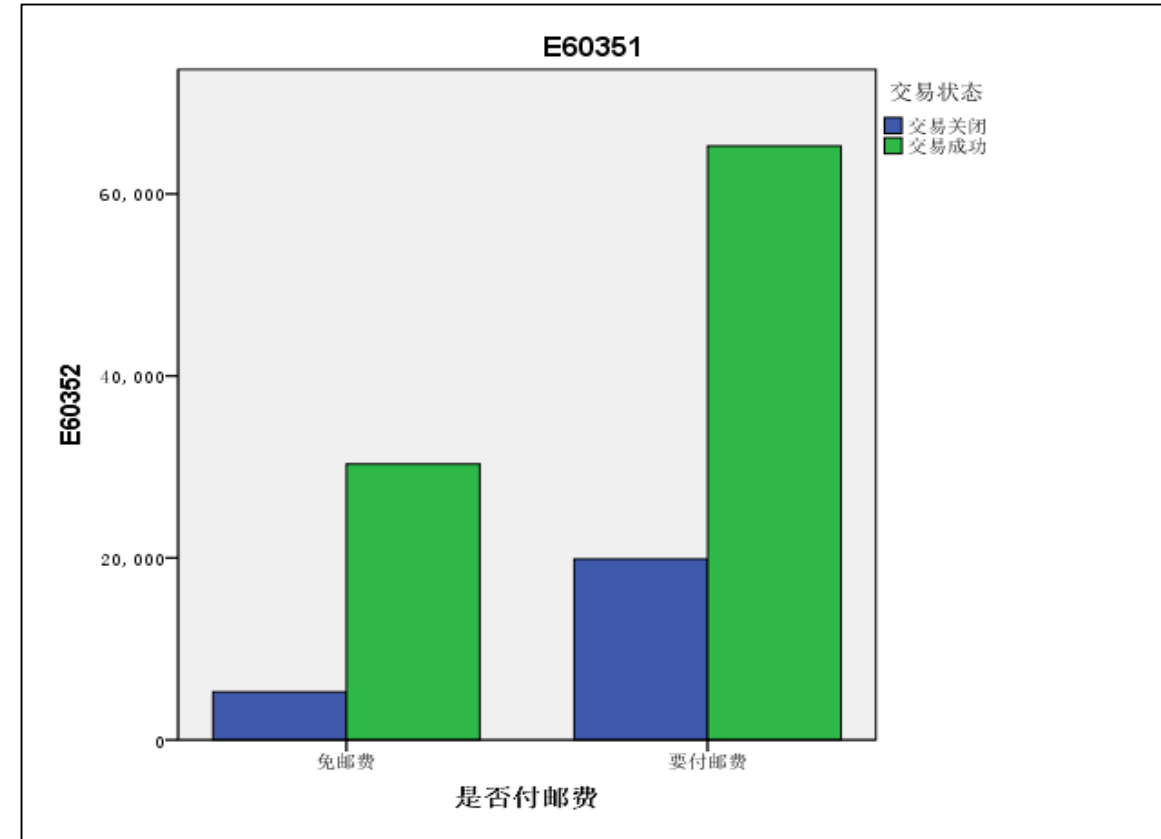
计数		交易状态		总计
		交易关闭	交易成功	
是否付邮费	免邮费	5284	30318	35602
	要付邮费	19874	65278	85152
总计		25158	95596	120754

	值	自由度	渐近显著性 (双向)	精确显著性 (双向)	精确显著性 (单向)
皮尔逊卡方	1099.120 <sup>a</sup>	1	.000		
连续校正 <sup>b</sup>	1098.605	1	.000		
似然比(L)	1155.310	1	.000		
Fisher 精确 检验				.000	.000
线性关联	1099.111	1	.000		
有效个案数	120754				

交易关闭的订单中有**5284**单免邮费，**19874**单要付邮费。对交易订单情况与包邮情况进行分析，由卡方分析得出，交易是否成功与是否包邮存在显著的相关。

# 关于是否包邮对交易是否成功的影响分析

风险评估			
	值	95% 置信区间	
		下限	上限
是否付邮费 (免邮费 / 要付邮费) 的比值比	.572	.554	.592
对于 cohort 交易状态 = 交易关闭	.636	.619	.654
对于 cohort 交易状态 = 交易成功	1.111	1.105	1.117
有效个案数	120754		



从上表可看出对于交易关闭的订单，免邮费的订单数是要付邮费的订单数的0.636倍。对于交易成功的订单，免邮费的订单数是要付邮费的订单数的1.111倍。所有可以通过包邮的形式来提升订单的交易成功率。

# 各省市不包邮的交易关闭详细情况

是否付邮 费	要付邮费				
交易状态	交易关闭				
省市名称	买家应付货款	买家应付邮费	总金额	邮费占总金额的百分比	宝贝总数量
广东省	230373.22	16651	247024.22	6.74%	3657
北京	235191.21	11386	246577.21	4.62%	3778
浙江省	219339.83	15661	235000.83	6.66%	3387
江苏省	201111.11	15046	216157.11	6.96%	3160
上海	160244.23	9788	170032.23	5.76%	2574
山东省	143108.99	11280	154388.99	7.31%	2162
湖南省	130015.91	8991	139006.91	6.47%	1980
河北省	94923.21	7823	102746.21	7.61%	1412
福建省	93649.89	6722	100371.89	6.70%	1465
四川省	89450.71	7394	96844.71	7.63%	1365
河南省	79574.77	6569	86143.77	7.63%	1224
陕西省	77347.89	6192	83539.89	7.41%	1167
湖北省	74875.66	5889	80764.66	7.29%	1174
安徽省	73849.04	6043	79892.04	7.56%	1122
天津	72665.68	5495	78160.68	7.03%	1161
辽宁省	71673.38	5527	77200.38	7.16%	1115

收货地址 省	值	自由度	渐近显著性 (双向)	精确显著性 (双向)	精确显著性 (单向)
北京	皮尔逊卡方	95.343 <sup>d</sup>	1	.000	
广东省	皮尔逊卡方	119.452 <sup>g</sup>	1	.000	
浙江省	皮尔逊卡方	128.740 <sup>ah</sup>	1	.000	
江苏省	皮尔逊卡方	93.813 <sup>q</sup>	1	.000	
上海	皮尔逊卡方	83.279 <sup>z</sup>	1	.000	

通过卡方检验北京，广东，浙江，江苏，上海5个省市的交易是否成功与是否包邮存在显著的相关。5个省市的成交额和成交量都是全国最高的而且邮费金额在总金额中占比较少，可以首先实施该5个省市的全面包邮措施。