

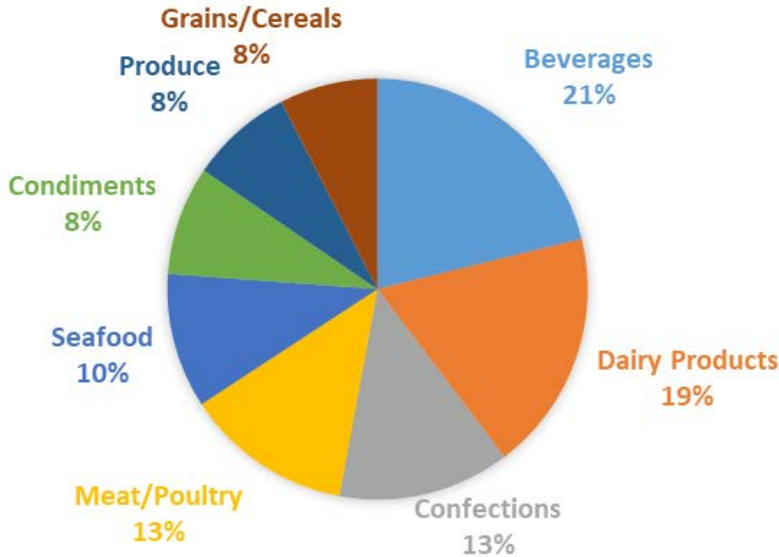
# 项目四、基于零售数据挖掘业务特征

XB

2019.2.15

# 公司产品中饮料销售额最高

各类商品销售额比例

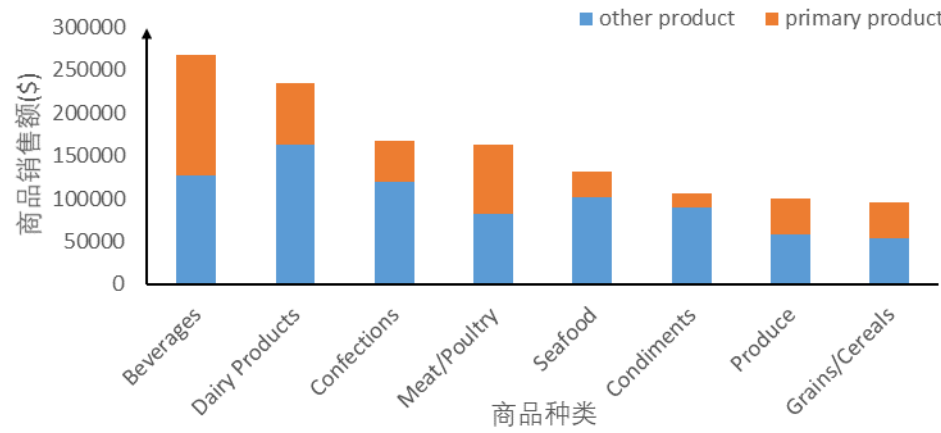


从图中可以看出饮料(Beverages)销售额比例最高，达到21%；乳制品(Dairy Products)次之，为19%；肉类(Meat/Poultry)比例13%；海产品(seafood)比例为10%，调味品(Condiments)比例为8%，农产品(Produce)比例为8%，谷物(Grain/Cereals)比例为8%。

categoryName	Beverages	Dairy Products	Confections	Meat/Poultry	Seafood	Condiments	Produce	Grains/Cereals	total
sum_money	267868.2	234507.3	167357.2	163022.4	131261.7	106047.1	99984.58	95744.59	1265793

# 公司各类商品首要商品销售额比重很大

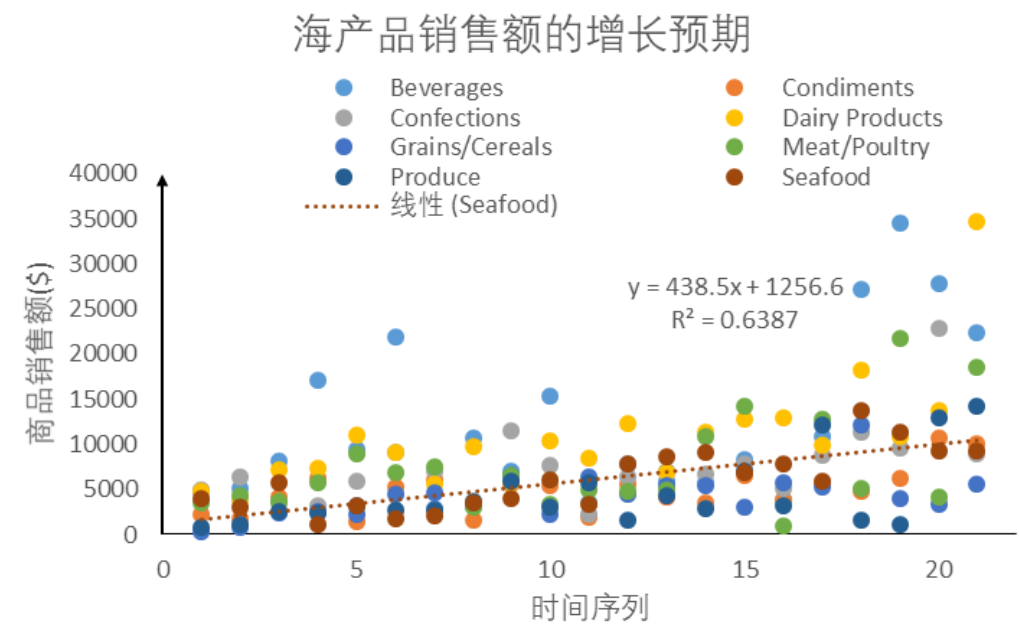
首要商品占各类商品的比例



从图中可以看出饮料(Beverages)的首要产品销售额比例最高，达到52.79%；乳制品(Dairy Products)、肉类(Meat/Poultry)、海产品(seafood)、农产品(Produce)和谷物(Grain/Cereals)销售额对于首要商品的依赖也大于20%，最低的调味品(Condiments)对首要商品的依赖也达到15.75%。这显示公司销售额对于明星产品极度依赖。

Category	sum money	product name	to money	percentage
Beverages	267868.2	银象 de C	141396.7	52.79
Dairy Prod	234507.3	Raclette Co	71155.7	30.34
Confection	167357.2	Tarte au su	47234.97	28.22
Meat/Poul	163022.4	银线 Th	80368.67	49.30
Seafood	131261.7	Carnarvon	29171.88	22.22
Condiment	106047.1	Vegie-spre	16701.1	15.75
Produce	99984.58	Manjimup	41819.65	41.83
Grains/Cer	95744.59	Gnocchi di	42593.06	44.49

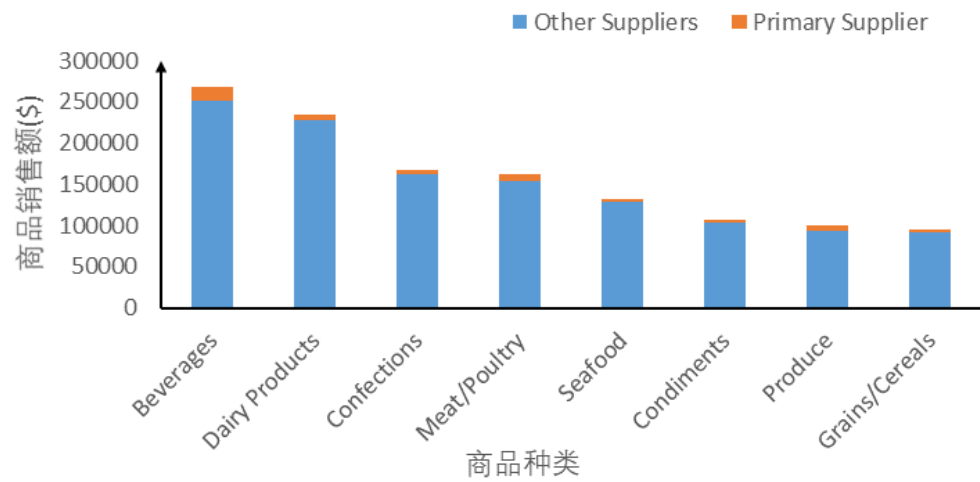
# 海产品销售额会稳步增长



将2014年8月至2016年4月各个月不同产品销售额统计做散点图，并计算回归方程和判定系数。只有海产品销售额与时间的回归方程拟合性较好( $R^2 > 0.5$ )，其他产品随着时间变化波动极大，所以认为海产品的销售额在未来会稳步增长

# 公司进货渠道繁杂

公司各类商品首要供应商占比重



统计各类商品总销售额和各类商品的单一供应商产品的最大销售额，可以看到公司进货渠道繁杂，各类产品首要供应商的商品总额比例不到该产品销售额的10%。这会对公司商品采购和质量控制带来挑战。

类型	商品总额(\$)	供应商	供应商总额(\$)	比例 (%)
Beverages	267868.18	Aux joyeux eccl 铝网 iastiques	15810	5.902157
Dairy Products	234507.29	Gai p 铝象 urage	6050	2.579877
Confection s	167357.23	Heli S 铝斤拷 waren GmbH & Co. KG	5268	3.147758
Meat/Poul try	163022.36	Plutzer Lebensmittel gro 铝网铝 鋸kte AG	9903.2	6.07475
Seafood	131261.74	Pavlova, Ltd.	3125	2.380739
Condimen ts	106047.09	For 铝象s d' 铝鋸ables	2565	2.418737
Produce	99984.58	G'day, Mate	6042	6.042932
Grains/Cer eals	95744.588	Plutzer Lebensmittel gro 铝网铝 鋸kte AG	4322.5	4.514616