## 巧克力工艺、口味、评分数据分析

### 一、项目背景

自2007年以来，曼哈顿巧克力协会的成员一直聚会，唯一的目的是品尝和评判高品质巧克力，以教育消费者并激励巧克力制造商生产更高品质的巧克力。从那时起，他们已经编制了一个包含超过2200个巧克力条的数据集，由他们的成员根据以下等级评分：

4.0–5.00 = 杰出

3.5–3.99 = 强烈推荐

3.0–3.49 = 推荐

2.0–2.99 = 令人失望

1.0–1.99 = 不愉快

这些评分来自于一个考虑到香气、外观、质地、风味、余味和整体意见的评分标准，评分的巧克力条大多是较深的巧克力条，因为目的是欣赏可可的风味。除了评分，曼哈顿巧克力协会的数据集还包括许多特征，如可可豆种植的国家、巧克力条含有的成分数量、是否包含盐，以及用来描述它的词语。

### 二、问题描述

根据给出的数据，重点分析下什么样的巧克力可以拿到高分，高分巧克力是否有明显的地域性。

### 三、数据描述

company: 分类；巧克力条的制造商（超过500个不同的制造商）

company\_location: 分类；制造商的国家（66个不同的国家）

review\_date: 连续；巧克力条被评审的年份（从2006年到2020年）

country\_of\_bean\_origin: 分类；可可豆采集的国家（62个不同的国家）

cocoa\_percent: 分类；巧克力条中可可的百分比

rating: 连续；由曼哈顿巧克力协会给出的评分（可能的值：1-5）

counts\_of\_ingredients: 连续；巧克力条中的成分数量

cocoa\_butter: 二进制；是否使用可可脂制作

vanilla: 二进制；是否使用香草制作

lecithin: 二进制；是否使用卵磷脂制作

salt: 二进制；是否使用盐制作

sugar: 二进制；是否使用糖制作

sweetener\_without\_sugar: 二进制；是否使用无糖甜味剂制作

first\_taste: 文本；用来描述第一口味道的词（语）

second\_taste: 文本；用来描述第二口味道的（语）

third\_taste: 文本；用来描述第三口味道的词（语）

fourth\_taste: 文本；用来描述第四口味道的词（语）

### 四、数据来源

<http://flavorsofcacao.com/mcs_index.html>

### 五、分析思路

数据可视化分析。