【考核项目06】 婚恋配对实验

婚恋配对模拟规则：

① 按照一定规则生成了1万男性+1万女性样本：

\*\* 在配对实验中，这2万个样本具有各自不同的个人属性（财富、内涵、外貌），每项属性都有一个得分

\*\* 财富值符合指数分布，内涵和颜值符合正态分布

\*\* 三项的平均值都为60分，标准差都为15分

② 模拟实验。基于现实世界的提炼及适度简化，我们概括了三个最主流的择偶策略：

\*\* 择偶策略1：门当户对，要求双方三项指标加和的总分接近，差值不超过20分；

\*\* 择偶策略2：男才女貌，男性要求女性的外貌分比自己高出至少10分，女性要求男性的财富分比自己高出至少10分；

\*\* 择偶策略3：志趣相投、适度引领，要求对方的内涵得分在比自己低10分~高10分的区间内，且外貌和财富两项与自己的得分差值都在5分以内

③ 每一轮实验中，我们将三种策略随机平分给所有样本（即采用每种策略的男性有3333个样本）

④ 我们为每位单身男女随机选择一个对象，若双方互相符合要求就算配对成功，配对失败的男女则进入下一轮配对。

1、样本数据处理

\*\* 按照一定规则生成了1万男性+1万女性样本：

\*\* 在配对实验中，这2万个样本具有各自不同的个人属性（财富、内涵、外貌），每项属性都有一个得分

\*\* 财富值符合指数分布，内涵和颜值符合正态分布

\*\* 三项的平均值都为60分，标准差都为15分

要求：

① 构建函数实现样本数据生成模型，函数参数之一为“样本数量”，并用该模型生成1万男性+1万女性数据样本

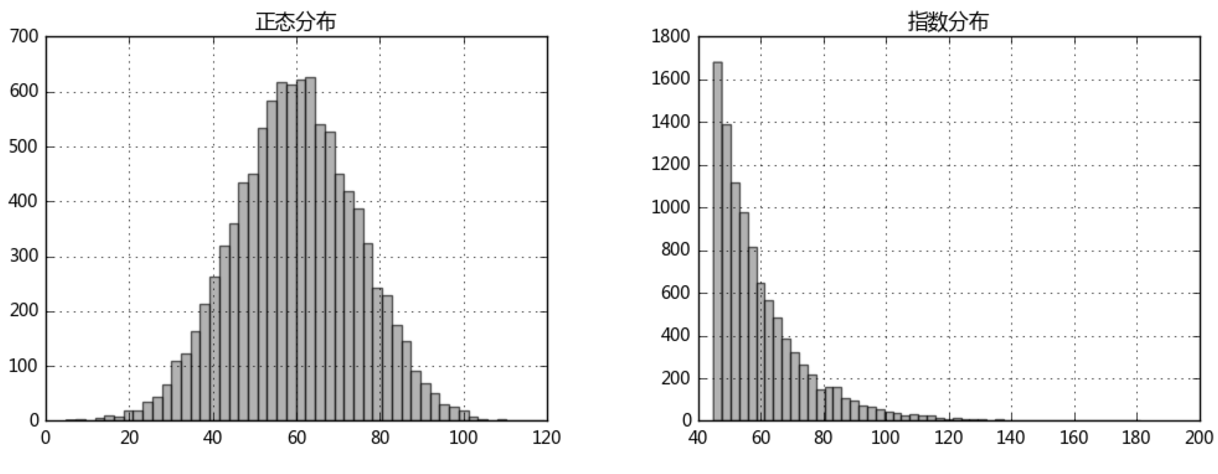
\*\* 包括三个指标：财富、内涵、外貌

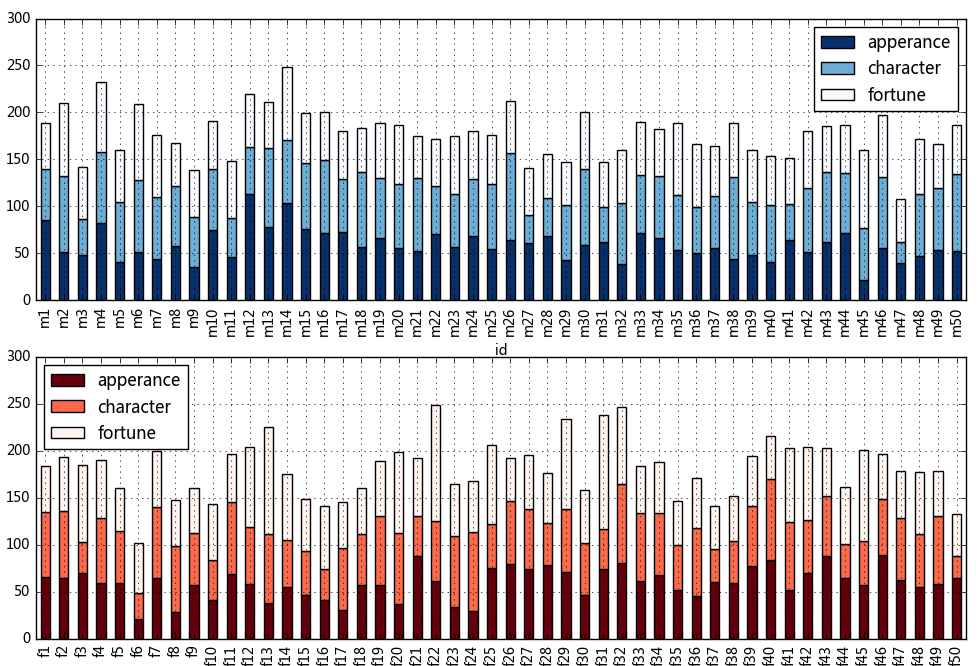
1. 绘制柱状图来查看每个人的属性分值情况

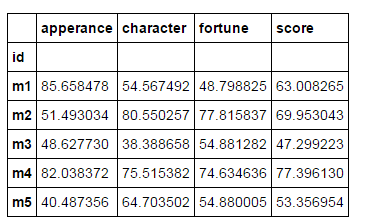
提示：

① 正态分布：np.random.normal(loc=60, scale=15, size=n)

② 指数分布：np.random.exponential(scale=15, size=n) + 45







数据生成后格式

2、生成99个男性、99个女性样本数据，分别针对三种策略构建算法函数

\*\* 择偶策略1：门当户对，要求双方三项指标加和的总分接近，差值不超过20分；

\*\* 择偶策略2：男才女貌，男性要求女性的外貌分比自己高出至少10分，女性要求男性的财富分比自己高出至少10分；

\*\* 择偶策略3：志趣相投、适度引领，要求对方的内涵得分在比自己低10分~高10分的区间内，且外貌和财富两项与自己的得分差值都在5分以内

\*\* 每一轮实验中，我们将三种策略随机平分给所有样本，这里则是三种策略分别33人

\*\* 这里不同策略匹配结果可能重合，所以为了简化模型

→ 先进行策略1模拟，

→ 模拟完成后去掉该轮成功匹配的女性数据，再进行策略2模拟，

→ 模拟完成后去掉该轮成功匹配的女性数据，再进行策略3模拟

要求：

① 生成样本数据

② 给男性样本数据，随机分配策略选择 → 这里以男性为出发作为策略选择方

③ 尝试做第一轮匹配，记录成功的匹配对象，并筛选出失败的男女性进入下一轮匹配

④ 构建模型，并模拟1万男性+1万女性的配对实验

⑤ 通过数据分析，回答几个问题：

\*\* 百分之多少的样本数据成功匹配到了对象？

\*\* 采取不同择偶策略的匹配成功率分别是多少？

\*\* 采取不同择偶策略的男性各项平均分是多少？

提示：

① 择偶策略评判标准：

\*\* 若匹配成功，则该男性与被匹配女性在这一轮都算成功，并退出游戏

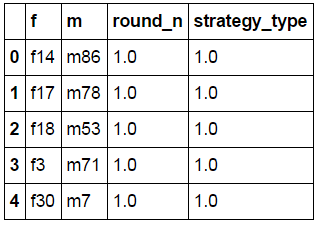
\*\* 若匹配失败，则该男性与被匹配女性再则一轮都算失败，并进入下一轮

\*\* 若同时多个男性选择了同一个女性，且满足成功配对要求，则综合评分高的男性算为匹配成功

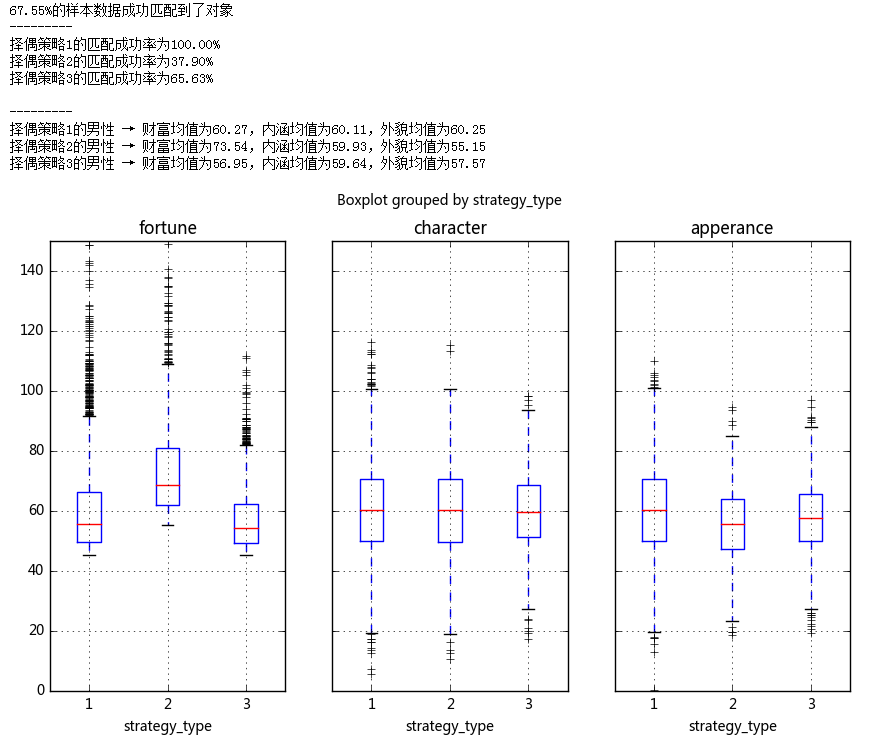
② 构建空的数据集，用于存储匹配成功的数据

③ 每一轮匹配之后，删除成功匹配的数据之后，进入下一轮，这里删除数据用df.drop()

④ 这里建议用while去做迭代 → 当该轮没有任何配对成功，则停止实验



匹配成功的数据存储格式





3、以99男+99女的样本数据，绘制匹配折线图

要求：

① 生成样本数据，模拟匹配实验

② 生成绘制数据表格

③ bokhe制图

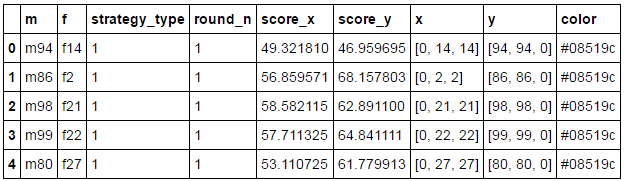
\*\* 这里设置图例，并且可交互（消隐模式）

提示：

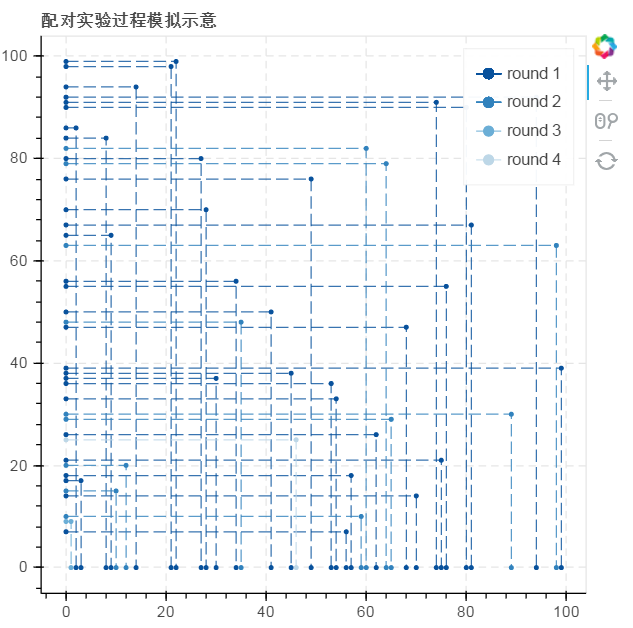
① bokeh制图时，y轴为男性，x轴为女性

② 绘制数据表格中，需要把男女性的数字编号提取出来，这样图表横纵轴好识别

③ bokhe绘制折线图示意：p.line([0,女性数字编号，女性数字编号],[男性数字编号，男性数字编号，0])



Bokeh绘图的数据格式



4、生成“不同类型男女配对成功率”矩阵图

要求：

① 以之前1万男+1万女实验的结果为数据

② 按照财富值、内涵值、外貌值分别给三个区间，以区间来评判“男女类型”

\*\* 高分（70-100分），中分（50-70分），低分（0-50分）

\*\* 按照此类分布，男性女性都可以分为27中类型：财高品高颜高、财高品中颜高、财高品低颜高、... （财→财富，品→内涵，颜→外貌）

③ bokhe制图

\*\* 散点图

\*\* 27行\*27列，散点的颜色深浅代表匹配成功率

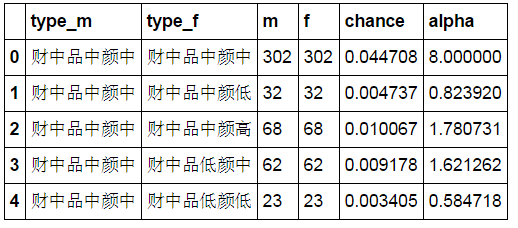
提示：

① 注意绘图的数据结构

② 这里散点图通过xy轴定位数据，然后通过设置颜色的透明度来表示匹配成功率

1. alpha字段为每种类型匹配成功率标准化之后的结果，再乘以一个参数

→ data['alpha'] = (data['chance'] - data['chance'].min())/(data['chance'].max() - data['chance'].min())\*8



Bokeh绘图数据结构

