【项目06】 多场景下的图表可视化表达

课程数据：

2016年奥运运动员数据，数据格式为xlsx，分3个sheet

作业要求：

1、分男女分别分析运动员的身高分布，并制作图表，数据为“奥运运动员数据.xlsx，sheet → 运动员信息”

要求：

① 制作分布密度图

② 计算出男女平均身高，并绘制辅助线表示

提示：

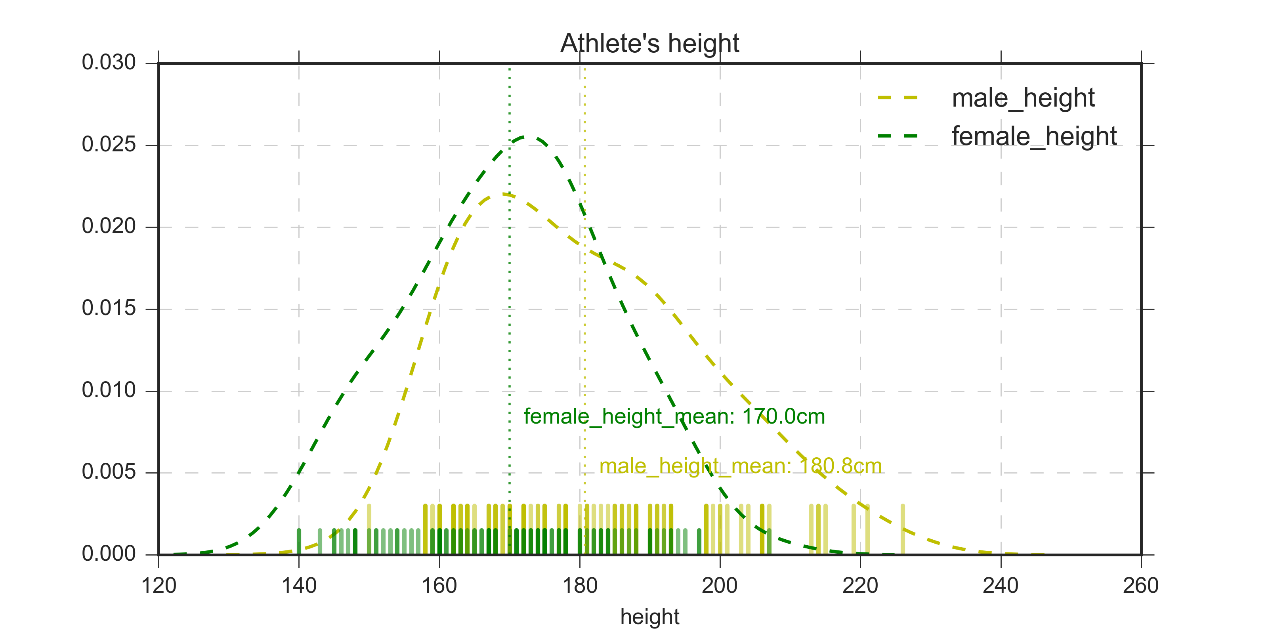
① 可视化制图方法 → sns.distplot()

② 辅助线制图方法 → plt.axvline()

③ 分男女分别筛选数据并制作图表

④ 不需要创建函数

身高分布密度图参考：

  
横坐标：运动员编号

纵坐标：男性身高，女性身高

2、综合指标判断运动员的身材，并找到TOP8的运动员，并制作图表，数据为“奥运运动员数据.xlsx，sheet → 运动员信息”

要求：

① 针对不同指标，绘制面积堆叠图

② TOP8的运动员，绘制雷达图表示

提示：

① 四个指标评判运动员身材，并加权平均

a. BMI 指数（BMI =体重kg ÷ 身高m\*\*2，详细信息可百度查询）→ 越接近22分数越高

b. 腿长/身高 指数 → 数据筛选，只选取小于0.7的数据，越大分数越高

c. 臂展/身高 指数 → 数据筛选，只选取大于0.7的数据，比值越接近1分数越高

d. 年龄 指数 → 年龄越小分数越高

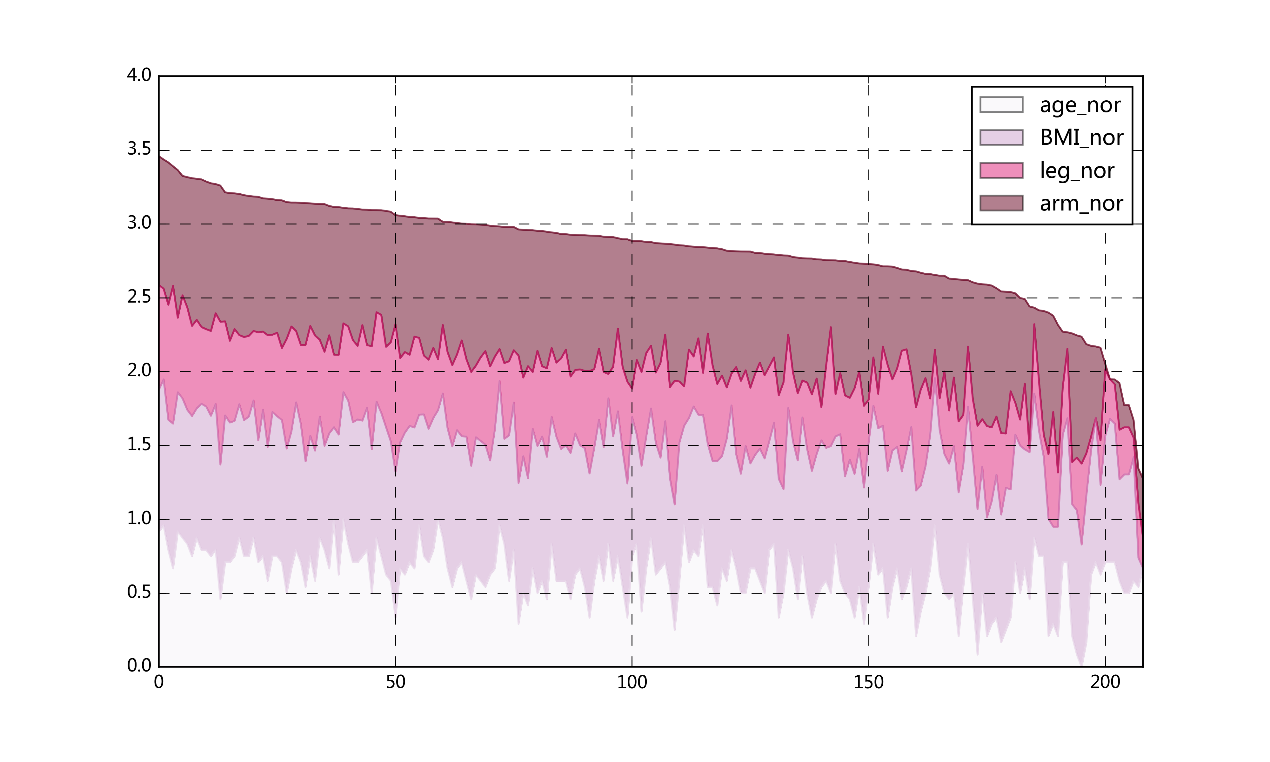
对上述abcd指标分别标准化得到n1,n2,n3,n4（划分到0-1的分值）

最后评分： finalscore = (n1 + n2 + n3 + n4)/4

② 制作堆叠面积图，data.plot.area()

③ 雷达图需要构建子图 + for循环遍历得到

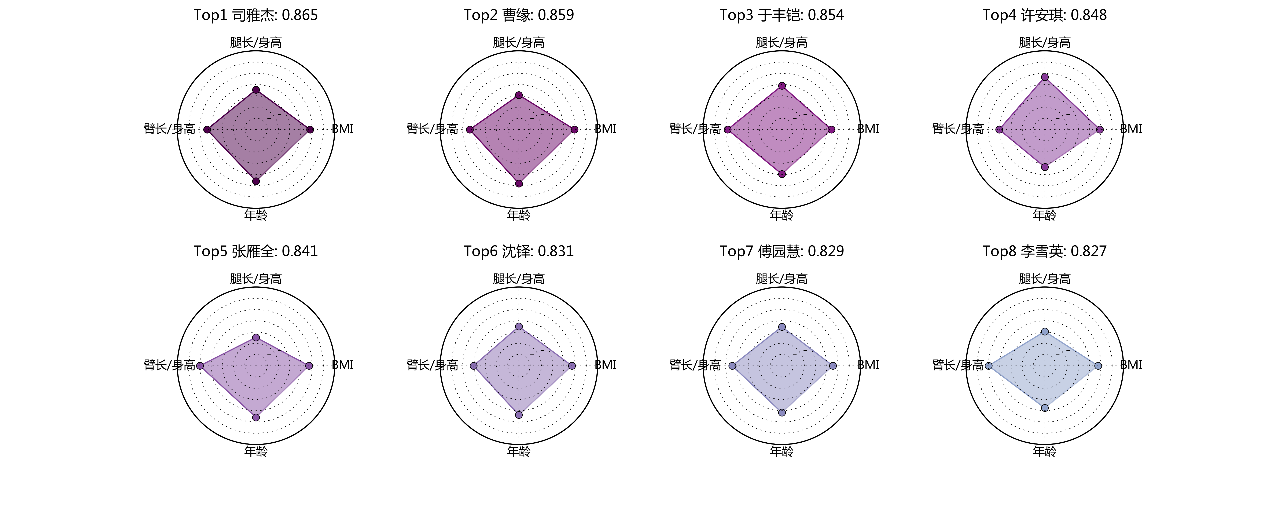
面积堆叠图参考：



横坐标：运动员编号

纵坐标：年龄指标，BMI指标，腿长指标，手长指标（注意四个指标均需要标准化）

雷达图参考：



3、根据运动员CP数据，分析出CP综合热度，通过python处理数据并导出，在Gephi中绘制图表，数据为“奥运运动员数据.xlsx，sheet → 运动员CP热度”

要求：

① 用python计算出综合热度指标

② 用Gephi绘制关系可视化图表

提示：

① 三个指标评判运动员CP综合热度，并加权平均

a. cp微博数量 → 数量越多分数越高

b. cp微博话题阅读量 → 阅读量越多分数越高

c. B站cp视频播放量 → 播放量越大分数越高

对上述abcd指标分别标准化得到n1,n2,n3,n4（划分到0-1的分值）

最后评分： finalscore = n1\*0.5 + n2\*0.3 + n3\*0.2

② Gephi中布局模式选择“ForceAtlas2”

③ Gephi中通过模块化计算，给关联结果做分组，并且以此分组设定点颜色

