# 第四问前半部分

### 根据similarity 来验证influencers是否影响了followers

在本节中，为了确定DATA\_influence 数据集中的相似性是否表明识别出的影响者实际上影响了各自的艺术家，我们对music\_influence与similari相关性进行了分析。

**music\_influence与similari的相关性**

在这一节中，我们通过实际的实验结果来验证理论结果，即用influencers和followers之间的similarity来验证前者对后者influence程度。若followers和influencers之间的相似性可以表明influencers影响了followers，则我们可以通过similarity与influence之间的相关性来进行研究，若相关性高则说明influencers确实影响到了followers。

由1.2的模型，我们假设所构建的influence network，直接相连的influencer对follower影响最大；间接相连的influencer对follwer的影响程度随着跳步数，时间等因素增大而衰减；对于不相连的2个artist，认为他们对彼此没有影响。

在这里，我们依据跳步数1、2、3…来进行分组，使用皮尔逊相关系数分别计算符合此跳步数的”influencer对follower的influence”与“这二者的similarity”之间的相关性。

相关性表示为：



（图片+分析结果）