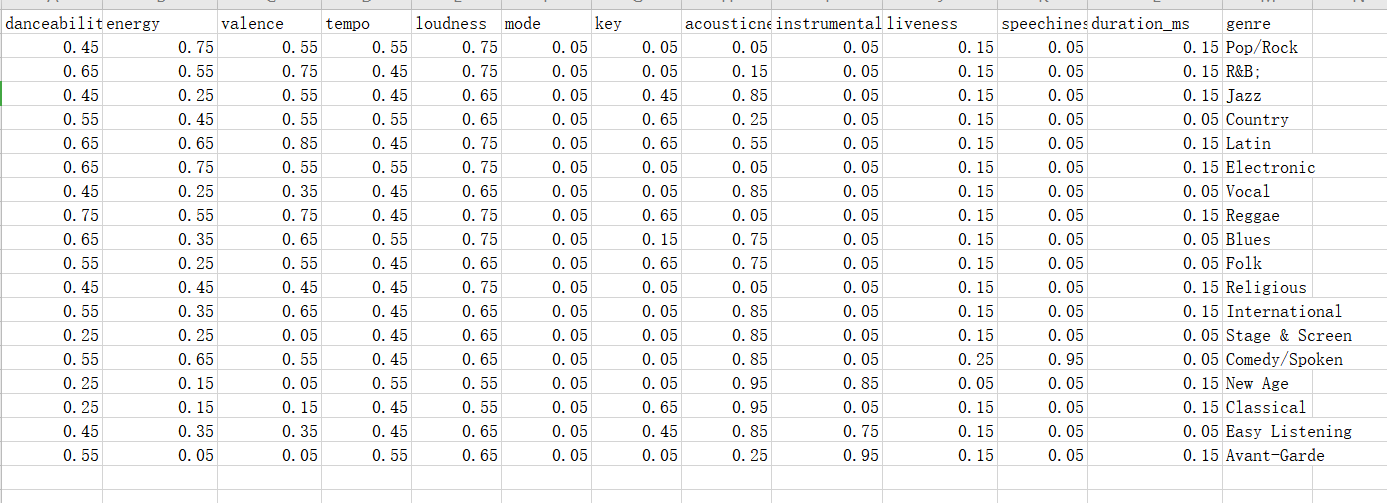
1. 十八个流派的十二个变量分析



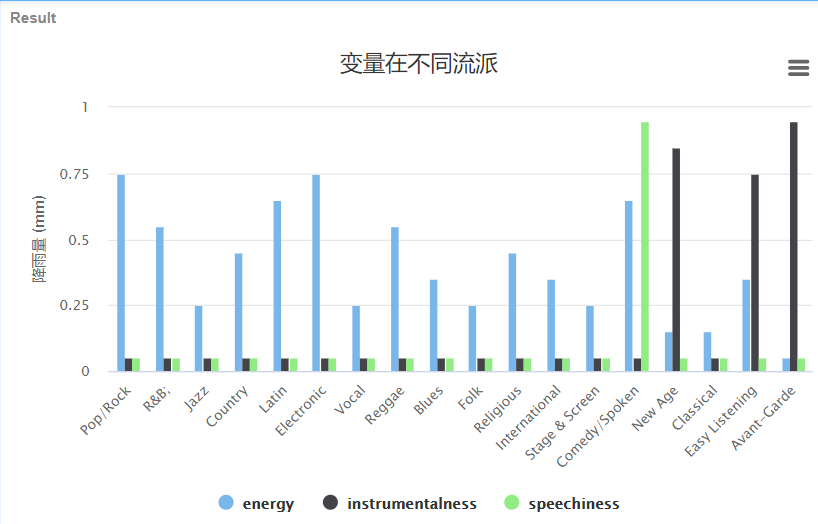
针对每个一流派的每一个变量（我们认为艺术家的popularity和cout与音乐流派的本身特征无关，因此把他们舍去），共十二个变量，我们将其标准化到[0,1]区间，以0.1为一个小区间，统计每个流派中所有 艺术家 的变量落在哪些区间，针对一个变量如 danceability , 某个流派落在0.4-0.5这个区间的艺术家比较多，我们便认为该流派 danceability 的主流为0.45（取了平均值），由此得到了18个流派的12个主流特征分布。

这12个主流特征作为流派的“标识符”，每列流派的12个特征序列都是独一无二的。为了验证其确实能够区分开其他的流派，我们选取The Rolling Stones，其特征为 [894465 ,0.50368237,0.719719881,0.659015688,0.528177341,0.778922974,1,0,0.294967335,0.181024725,0.267155588,0.030209567,0.115411008] 。其流派为Pop/Rock。我们这个艺术家的特征与18个流派的主流特征做 余弦相似度分析，得到了相似度列表为：

[0.83076436, 0.70348581, 0.74632624, 0.76742565, 0.82872646,0.7195927 , 0.7490484 , 0.77757226, 0.68521578, 0.82973572,0.7599794 , 0.64533123, 0.6807169 , 0.53960183, 0.5453823 ,0.66286441, 0.57595484]。与其相似度最高的流派为 Pop/Rock，恰当的将其分别了出来！

随机选取100个，准确率为26%，选取前四个最可能的，准确率提升为53%

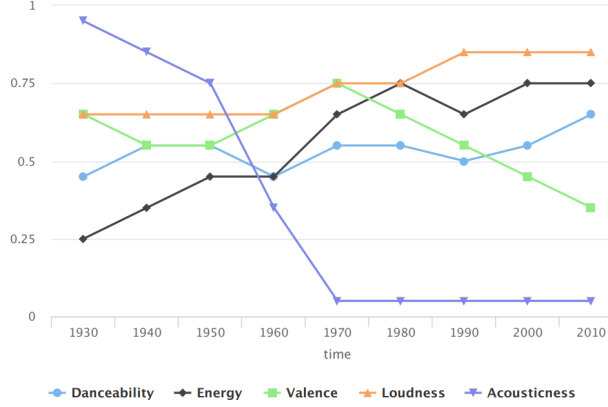
同时，我们发现某些流派在某些变量上面比其他流派更为独特（偏小或者偏大），而其他一些变量如danceability、tempo、loudness、liveness、duration\_ms、mode 等，所有的流派在其上取值都相同或者没有明显的区分，因此我们在可视化中将这些没有明显区分的变量舍去。



分析：由energy很高可以推断流派是POP/Rock 或者 Electronic，因为这两种流派的音乐往往频率比较高，energy比较大，如果energy较小则可以推测属于 NEW Age 或者 Classical。

Speechiness这个变量将 Comedy/Spocken这个流派明显同其他流派区别出来，而instrumealness这个变量有效的将New Age、Easy Listening , Avant-Garde这三个流派同其他流派明显的区分开来。

1. 我们将所有流派统一，不区分流派单纯区分年份，做了1930年到2010年世界音乐变量的变化图，基本没有变动的几个变量如tempo、mode、instrumentalness liveness speechiness duration\_ms等没有被放入图内。



结论：从图上来看，世界范围内的音乐的几个变量 danceability ,energy ,Loudness从1930年开始到2010年呈现上升趋势，而Valence 从1970年开始迅速下降，Acousticness 从1950年开始到1970年迅速下降，并维持一个很低的水平。

分析：随着近年来世界经济的上升，流行音乐开始向快节奏，嘈杂（高 energy）转变。

Energy 的解释可以参见 论文引用第1条，在task5分析中也分析了世界范围音乐各种变量随年份变化的原因。

1970年代电吉他和base开始流行于音乐中。Pop-Rock Music as Expressive Isomorphism: Blurring the National, the Exotic, and the Cosmopolitan in Popular Music Motti Regev

American Behavioral Scientist

Vol 55, Issue 5, pp. 558 - 573

1949 摇滚元年，摇滚起源于美国，其中里程碑作品已经无从考究。但鼻祖级别但作品还是有的，Rock Awile 和 Rock the Joint 这两首发布在1949年的歌提出了ROOOOOCK 的概念。

1950s这时候的摇滚还有一股舞厅气，起源于Jazz 的摇滚在这时还没完全脱离Jazz的影响。

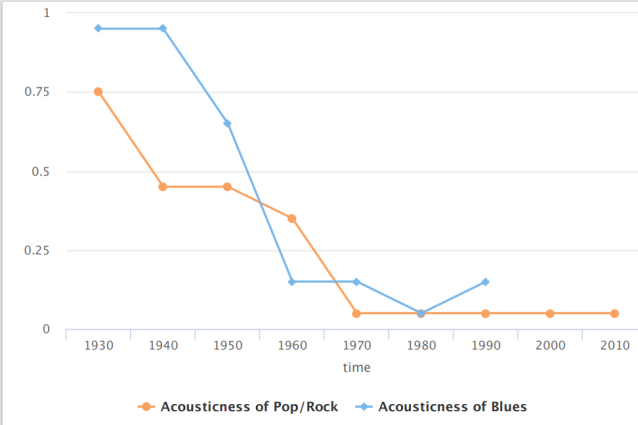
1960s 出口转内销，英国摇滚入侵，Beatles 在这十年席卷全球

1970s 是属于 Queen 的摇滚元年，整个摇滚乐队或许为人所知作品最多的时期恐怕也要集中在这10年

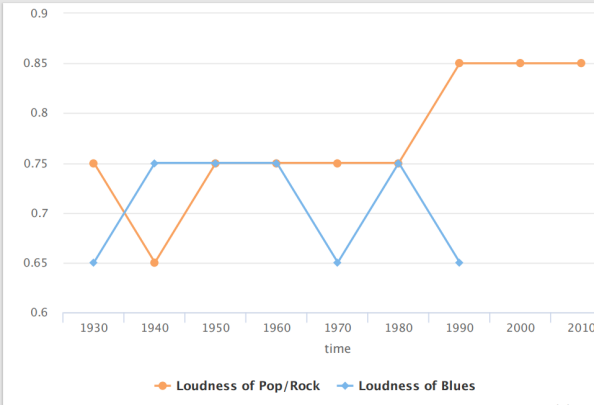
分析Acousticness 的下降：

磁带转化唱片业,到1950年代末,绝大多数商业录音被掌握在磁带上。电子革命之后,晶体管的发明带来了其他激进的变化,其中最重要的是引入了世界上第一个“个人音乐设备”

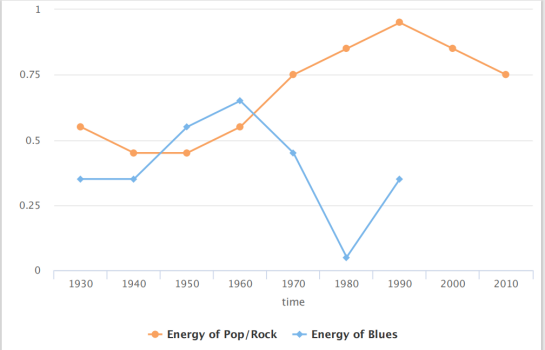
1. 我们针对性的选择了两个流派：Pop/Rock和Blues，并对他们流派的变量以时间为变量进行主流趋势的分析。



分析：两个变量的 Acoustiness 都符合世界音乐中急剧下降的趋势，这意味着这个趋势是普遍存在的，可能由技术革命引发？ 由电子技术革命引发。



分析：Pop/Rock 的Loudness有逐渐上升的趋势，意味着近年来 Pop/Rock的艺术家们更喜欢创作大音量的震撼作品。而Blues 的 Loudness在一个固定区间内波动，在这几十年中基本维持着自己的本色。



分析：Pop/Rock 的 Energy 近年来变得比较高，有一个上升趋势，而有趣的是，Blues在1960-1980年的Energy 出现了急剧的下降，直到1990年代才发生好转。

分析发现是由于当时有一位 叫 Chris Thomas King 的人物，他的energy为0.01，而且当时的bules艺术家非常少，导致他引领了潮流。

更新：

对于每一个流派，采用它每一项特征值（一共12个）作为单项的中心，那么十二个特征的平均值可以理解为簇的中心。对于判断一个歌手根据其风格，与簇的中心越接近，就越有可能被判断为属于该流派。

我们随机选取了100位艺术家，采用余弦相似度进行预测，若只选取最相似的流派作为预测流派，准确率达到了31%，若选取前四个最相似的流派作为预测流派，准确率达到了56.7%。

因此，我们可以说拿十二个特征的中心位置去判断一个艺术家是否属于某个流派是合理的。

如果已经知道和其有所关联的艺术家的流派，那么就有极大的概率预测出该艺术家处于哪个流派。