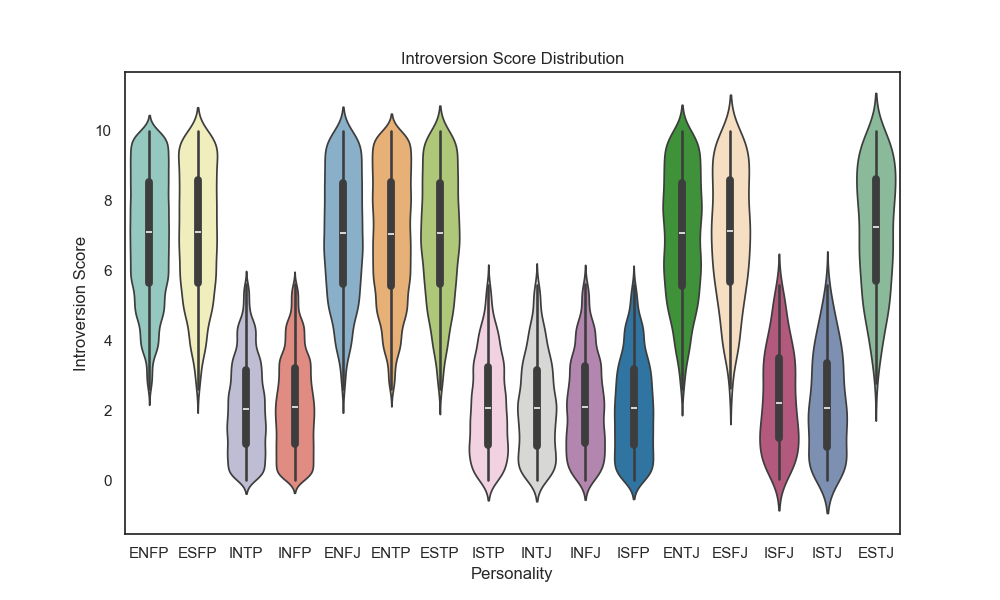
**一 数据集背景**：

这是一个来自Kaggle的大规模性格类型数据集，包含128,061个样本记录。研究变量涵盖人口统计学特征（年龄、性别）、教育程度（本科及以下/研究生及以上）、个人兴趣领域，以及四项性格维度的量化评分（0-10分制，包括内向-外向倾向、感知-直觉偏好、思维-情感取向、判断-知觉倾向）和MBTI性格类型（16种类型分类）。

本研究通过系统整合个人多维特征数据，分析性格特征与MBTI类型的关联模式、探索教育背景与兴趣爱好对性格形成的影响、研究人口特征在性格发展中的作用。研究对于优化个性化教学方案、提供科学的职业规划指导、改进人才筛选与管理策略等实践领域具有广泛的应用前景。

**二 图表分析（1~4为Violin Plot，5~6为Sunburst Diagram）**

**1. Introversion Score Distribution**



（1）数据解释：

一个范围从 0 到 10 的连续变量，代表个人的内向与外向倾向。分数越高表明外向倾向越大。

**Introversion Score:** A continuous variable ranging from 0 to 10, representing the individual's tendency toward introversion versus extraversion. Higher scores indicate a greater tendency toward extraversion.

（2）分布特征：

第一个字母为"E"的类型：表现出明显的外向特征，中位数大约为7分。

第一个字母为"I"的类型：表现出明显的内向特征，中位数大约为2分。

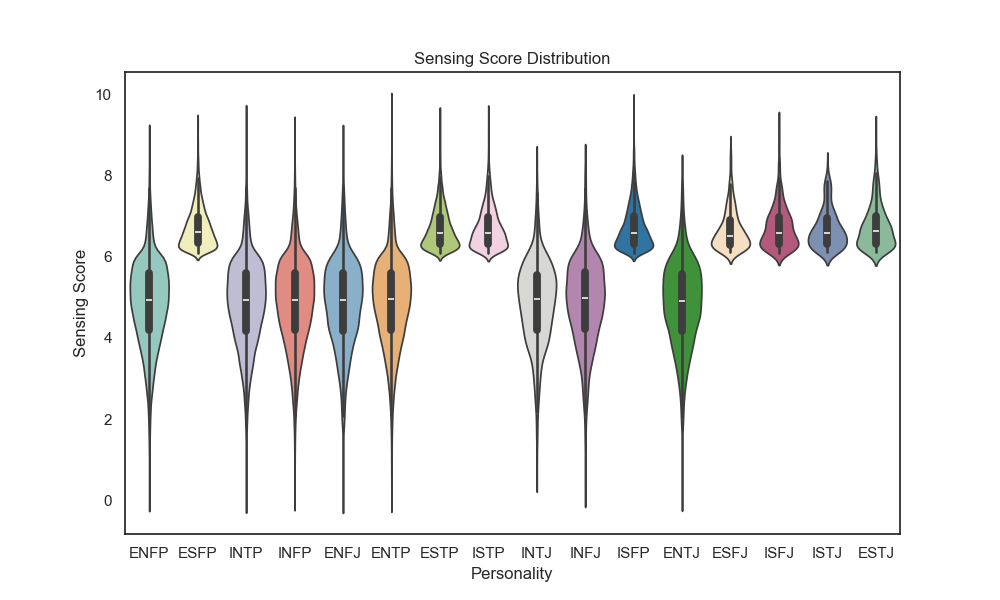
多数类型呈现多峰分布，表明内向分数可能受多个因素影响

（3）发现：

E类型和I类型的得分有明显的区分度，支持MBTI的分类准确性，也符合理论预期。

外向-内向特征呈现连续分布，而非绝对的二分类。

**2. Sensing Score Distribution**



（1）数据解释：

一个范围从 0 到 10 的连续变量，代表个人对感官与直觉的偏好。分数越高表示对感官的偏好。Sensing 类型的人倾向于关注具体的、现实的细节，依赖五感来感知世界。

**Sensing Score:** A continuous variable ranging from 0 to 10, representing the individual's preference for sensing versus intuition. Higher scores indicate a preference for sensing.

（2）分布特征：

第二个字母为"S"（感知），显示出较强的感官偏好，中位数都大于6。

第二个字母为"N"（直觉），以直觉为特征，显示出较低的感官倾向，中位数大约为5。

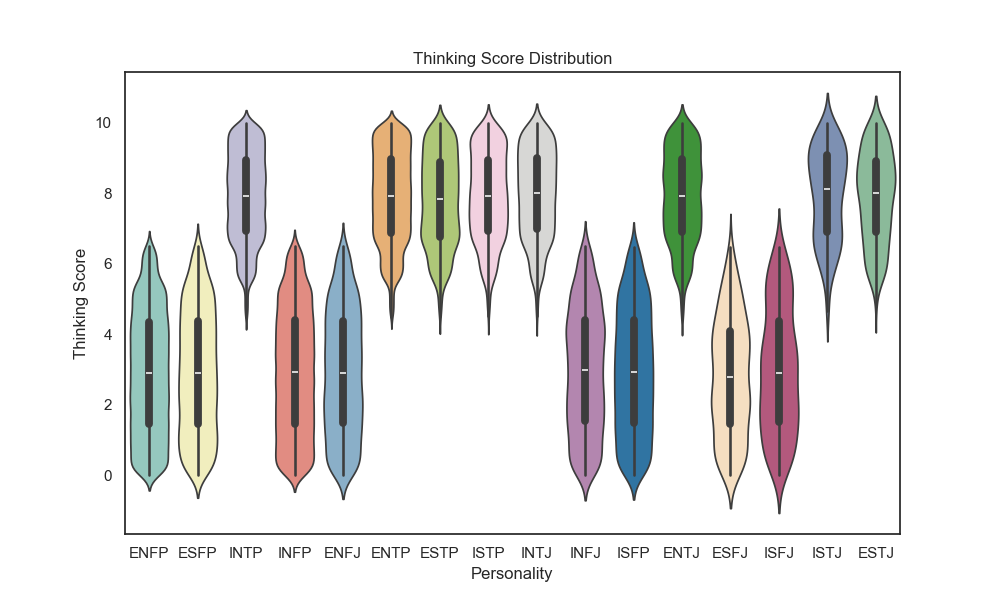
分布形态总体呈单峰或双峰分布。

（3）发现：

S类型（感知型）的得分普遍高于N类型（直觉型），符合MBTI理论。

两种类型分布的宽度各不相同，表明不同类型在感知特征上的稳定性存在差异。

**3. Thinking Score Distribution**

****

（1）数据解释：

一个范围从 0 到 10 的连续变量，表示个人对思考与情感的偏好。分数越高表明对思考的偏好。Feeling 类型的人倾向于通过考虑他人及自己的情感反应来做出判断。

**Thinking Score:** A continuous variable ranging from 0 to 10, indicating the individual's preference for thinking versus feeling. Higher scores indicate a preference for thinking.

（2）分布特征：

第三个字母为"T"（思考），显示出强烈的理性思考倾向，中位数大约为8。

第三个字母为"F"（情感），以情感为特征，表现出较低的思考倾向，中位数大约为3。

分布呈现明显的两极分化。

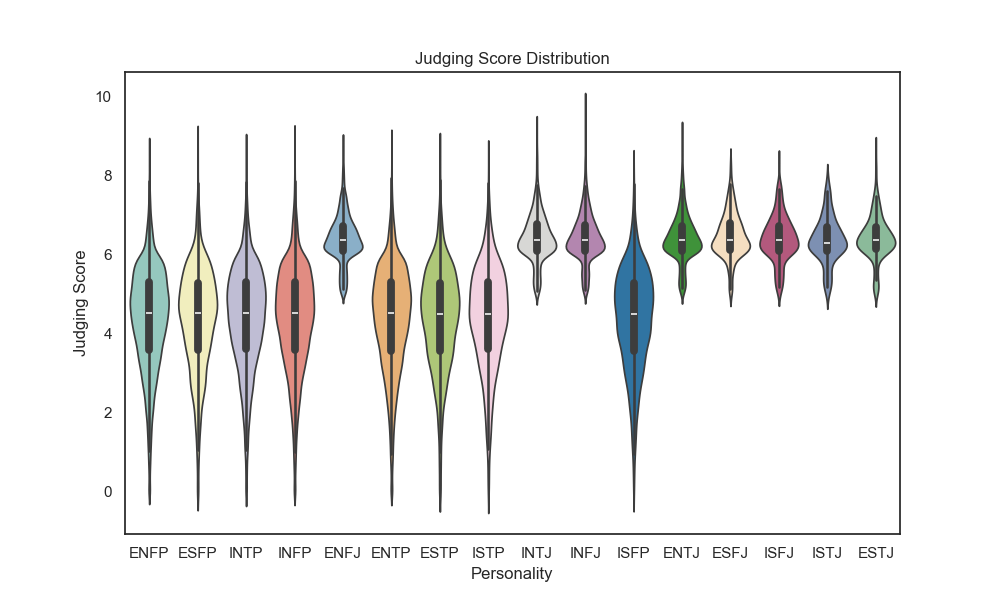
（3）发现：

MBTI中T-F维度的区分在数据中得到强有力的支持。

思考-情感特征表现出较强的二元性，而非连续渐变。

同类型（T或F）的性格表现出相似的得分模式。

**4. Judging Score Distribution**

****

（1）数据解释：

一个从 0 到 10 的连续变量，代表个人对判断与感知的偏好。分数越高表示对评判的偏好。Perceiving 代表的是一种开放、灵活和自发的生活态度。

**Judging Score:** A continuous variable ranging from 0 to 10, representing the individual's preference for judging versus perceiving. Higher scores indicate a preference for judging.

（2）分布特征：

第四分字母为"J"（判断），表现出较强的判断倾向，中位数大于6。

第四个字母为"P"（感知），显示出较低的判断倾向，中位数大约为5。

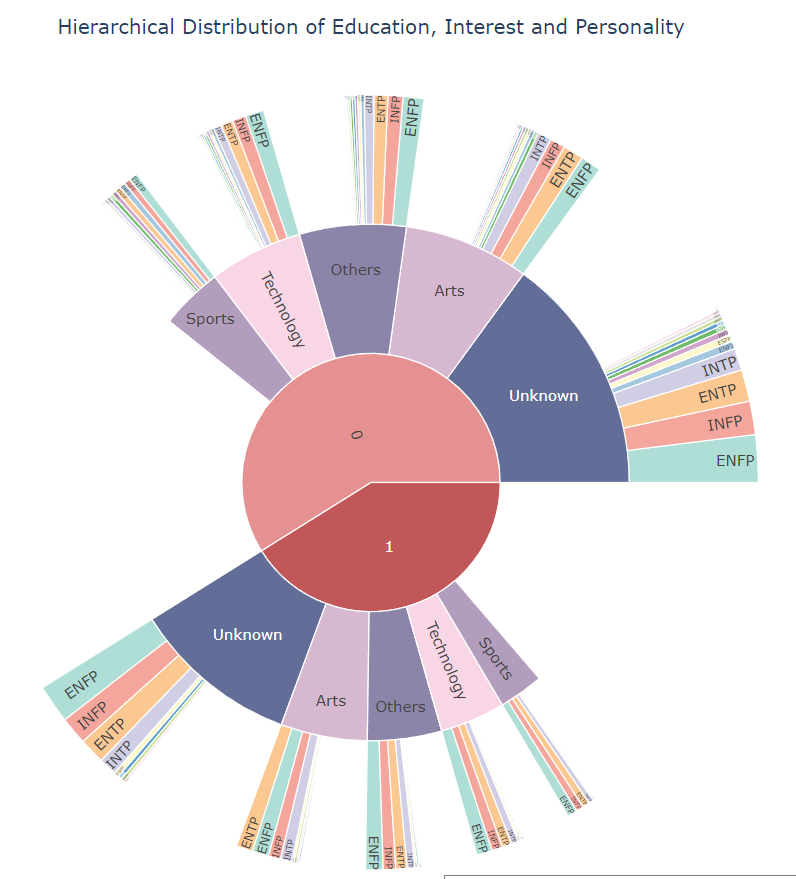
大多数得分集中在4-7分之间，分布形态相对均衡，波动较小。

（3）发现：

分布的离散程度相对较小，表明这一特征较为稳定。

各类型间的重叠程度较高，表明判断-感知特征可能更具灵活性。

**5. Hierarchical Distribution of Education, Interest and Personality**



（1）数据分布

样本中本科及以下（0）：75,408人，占总人数的58.9%。研究生及以上（1）：52,653人，占总人数的41.1%。

（2）发现

在MBTI性格类型与教育背景的交叉分析中，我们观察到一些有趣的分布模式。ENFP型人格展现出显著的兴趣多样性，他们在艺术、技术和运动等多个领域都表现出浓厚的兴趣，这种特征在本科及以下学历群体中表现得尤为明显，体现了ENFP型人格固有的探索精神和广泛的好奇心。

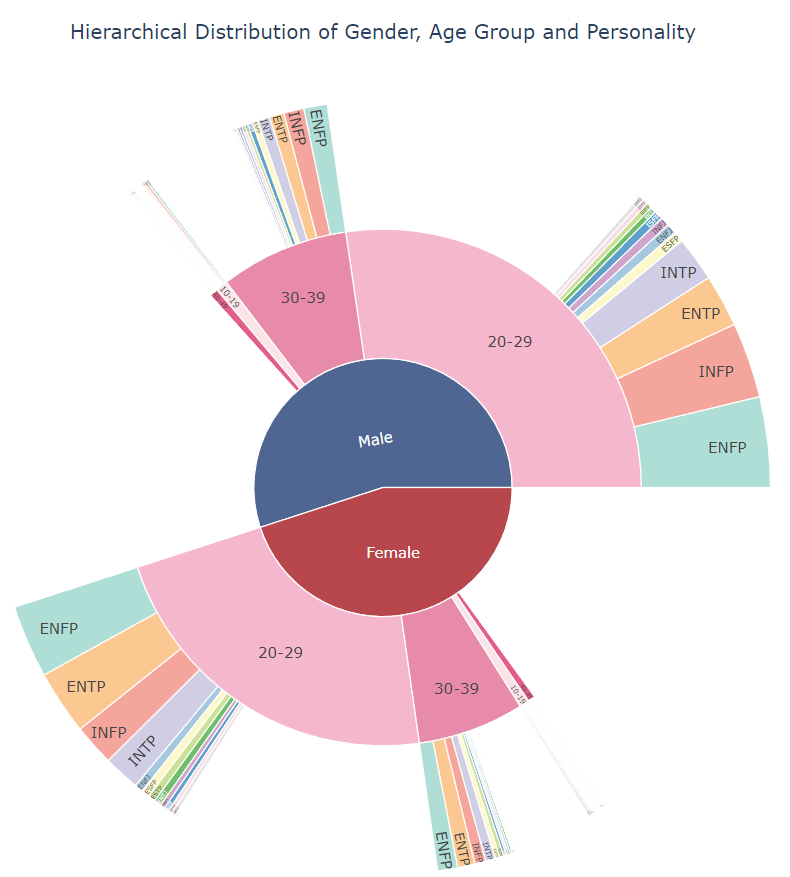
同时，研究发现ENTP、INFP和INTP类型对艺术领域具有普遍的强烈兴趣，这种倾向不受教育程度的影响，在所有学历层次中都保持着相当的一致性。这可能源于这些性格类型共同具备的创造力特质和对艺术表达的天然亲和力。

在职业定位方面，数据显示出明显的教育程度差异：本科及以下学历群体中，未确定兴趣方向的比例相对较高，这可能反映出他们在职业规划上的不确定性和探索阶段的特征。相比之下，研究生及以上群体表现出更加明确的职业兴趣取向，这或许与其接受更多专业训练和职业规划指导有关。

另一个显著特征是INTJ、ISFJ、ENTJ、ESFJ和ENFJ类型对运动领域普遍表现出较低的兴趣，这种特征在研究生及以上群体中尤为突出。这一现象可能与这些性格类型的内在特质以及他们对其他领域（如学术研究、管理领导等）的偏好有关。

这些发现不仅深化了我们对MBTI性格类型、教育背景与兴趣发展之间关系的理解，也为教育机构制定个性化培养方案和职业咨询机构提供职业规划建议提供了重要的参考依据。此外，这些观察结果也提示我们需要在职业生涯早期就为学生提供更多的职业探索机会和指导支持。

**6. Hierarchical Distribution of Gender, Age Group and Personality**



（1）数据分布

样本主要集中在20-29岁和30-39岁两个年龄段。性别构成上，男性占总样本的55%，女性占45%。

（2）发现

在性别和年龄的交叉分析中，我们观察到一些显著的性格类型分布特征。ENFP、ENTP、INFP和INTP这四种性格类型在各个性别和年龄层次中都呈现出明显的优势地位，它们共同的特点是都属于直觉(N)型人格，这反映了直觉型思维方式在当代人群中的普遍性。

特别引人关注的是20-29岁这一年轻群体，他们展现出最为丰富的MBTI类型分布特征，这种多样性可能源于年轻一代更开放的成长环境和更多元的教育背景。这一现象也可能反映了新生代在个性发展上的自我认知和表达更为自由。

整体来看，直觉(N)型性格在样本中占据主导地位，这一趋势跨越了性别和年龄的界限。这种现象可能与当代社会对创新思维、抽象推理和直觉判断能力的普遍重视有关，同时也可能反映了现代职场和生活方式对这类性格特质的培养和偏好。这些发现不仅深化了我们对人格类型分布的理解，也为人才发展策略和组织文化建设提供了有价值的参考依据。