

# 幸福指数影响因素探究

多元统计分析

于骏浩 2020012847

## Abstract

本文基于多元统计分析及因果推断的方法，对 2015-2022 年间 195 个国家的幸福指数和相关指标进行分析。运用因子分析和主成分分析进行降维和指标解读，利用层次聚类等方法对所获得因子进行聚类分析，利用 LDA、SVR 等方法对幸福指数进行判别分析和回归分析，得到判断某国家人民幸福程度的判别准则和预测方法。针对指标、因子得分和幸福指数之间的因果关系，在 Rubin 框架下使用 Neyman 方法和回归法对其进行分析。研究发现物质需求和精神需求的满足是提升人民幸福感的关键；在新冠疫情前后及不同发展程度的国家中，人民对这二者的看重程度有所不同。研究结果为各国政府部门提升本国人民幸福感提供了借鉴和参考。

## Contents

1	研究问题与背景描述	1
2	数据探索性分析	1
2.1	数据集描述	1
2.2	单变量分析	1
2.3	多变量相关性分析	2
2.4	PCA 分析与数据可视化	3
3	数据降维	4
3.1	因子分析	4
3.1.1	正交因子分析	4
3.1.2	斜交因子分析	4
4	聚类、判别与回归分析	5
4.1	聚类分析	5
4.2	判别分析	6
4.3	回归分析	7
5	因果推断	8
6	总结与反思	9
6.1	结论	9
6.2	政策建议	10
6.3	反思与展望	10
7	附录	10

# 1 研究问题与背景描述

从古至今，幸福感一直为人民群众所追求，进入新时代的中国，也愈发重视人民幸福感。2021 年刊发的《中共中央关于党的百年奋斗重大成就和历史经验的决议》就指出需要“不断实现好、维护好、发展好最广大人民的根本利益，使人民获得感、幸福感、安全感更加充实、更有保障、更可持续”[5]，因此对于人民幸福感的影响因素研究是非常具有现实意义的。根据文献调研，幸福感影响因素研究早在 20 世纪 50 年代就开始进入主流社会科学和经济学领域，但长久以来学界都将幸福感等同于国家以 GDP 为衡量标准的物质发展水平[6]，但近年来众多学者意识到这一评价标准的局限性。在此基础上，多国学者设计、参与并调研了世界上 195 个国家的幸福指数和相关指标[1]。本文即基于该报告及相关数据集，力求分析并回答如下问题：

1. 是否存在潜在的影响因素，影响幸福指数的相关指标？
2. 如何基于幸福指标，判断某国家提升人民幸福感的最有效措施？
3. 如何基于幸福指标，预测某国家的幸福指数？
4. 对提升人民幸福感具有显著帮助的潜在影响因素是什么？

## 2 数据探索性分析

### 2.1 数据集描述

原始数据来源于数据建模分析竞赛平台 Kaggle 上 World Happiness Report 2015-2022 项目[3]，其中包含变量如下所示：

变量	变量说明（不加说明的话，以下各变量均为各国间相对值）
Happiness Score	连续型变量；代表某国人民幸福程度
Economy (GDP per Capita)	连续型变量；代表某国人均国民生产总值
Family (Social Support)	连续型变量；代表某国人民人均获得的社会帮助或支持
Health (Life Expectancy)	连续型变量；代表某国人民人均健康水平或寿命期望
Freedom	连续型变量；代表某国人民自由程度
Trust (Government Corruption)	连续型变量；代表某国人民对政府的信任程度，以政府腐败程度为评判标准
Generosity	连续型变量；代表某国人民乐于助人、无私奉献的程度

### 2.2 单变量分析

首先我们绘制出各变量的箱型图以观察其整体分布特征。从中可以看出，前两个变量的尺度大于后四个变量的尺度，后三个变量的方差显著小于前三个变量，且最后两个变量存在明显的右偏性，这提示我们需要进行标准化及变量变换（如取对数等方法）。但整体来看，各变量的尺度大致相同，且具有相似的含义（即表示某国人民的某一方面的指标相对值），因此出于可解释性的考虑，本文并没有采用特殊变换，而仅仅在降维阶段对变量采用了标准化操作。

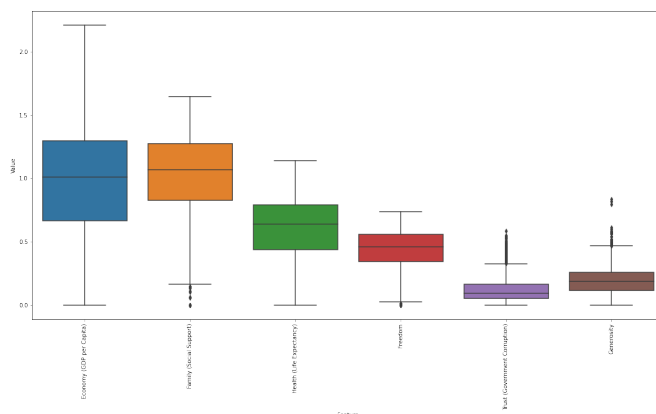


Figure 1: 各变量箱型图（受 Python 导出分辨率限制，pdf 不清晰，原 pdf 文件在代码中附上，后文不清晰 pdf 也均会附上）

接下来又根据国家所在的地区和指标调查的年份绘制了各变量的箱型图（图2和3），从中可以看出年份对变量的影响较小，而地区差异对变量有很大的影响，这也提示我们在之后的因果分析中需要使用分层的方法来控制变量，以期得到更可靠的结果。

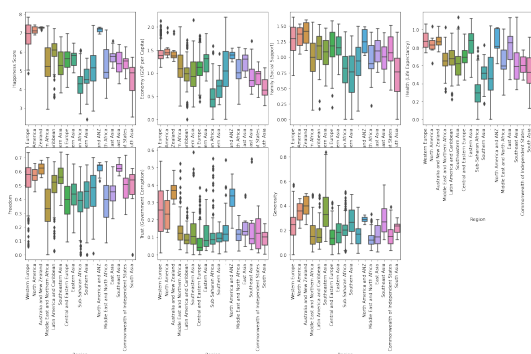


Figure 2: 各地区变量箱型图

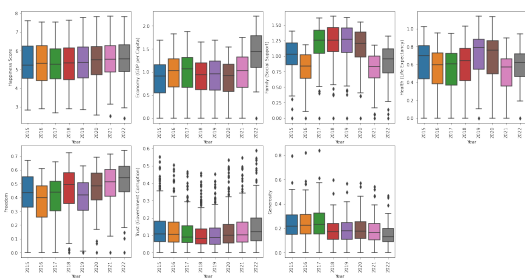


Figure 3: 各年份变量箱型图

## 2.3 多变量相关性分析

接下来我们又绘制了所有变量的相关系数图4。从单个变量的分布图可以看出，大部分变量并不服从正态分布，其中 Trust 和 Generosity 两个变量出现了明显的右偏，因此需要谨慎使用依赖正态性假设的模型。从散点图及相关系数矩阵中可以看出，多数变量之间存在明显的线性相关关系，这为我们后文继续进行降维分析提供了坚实的依据。

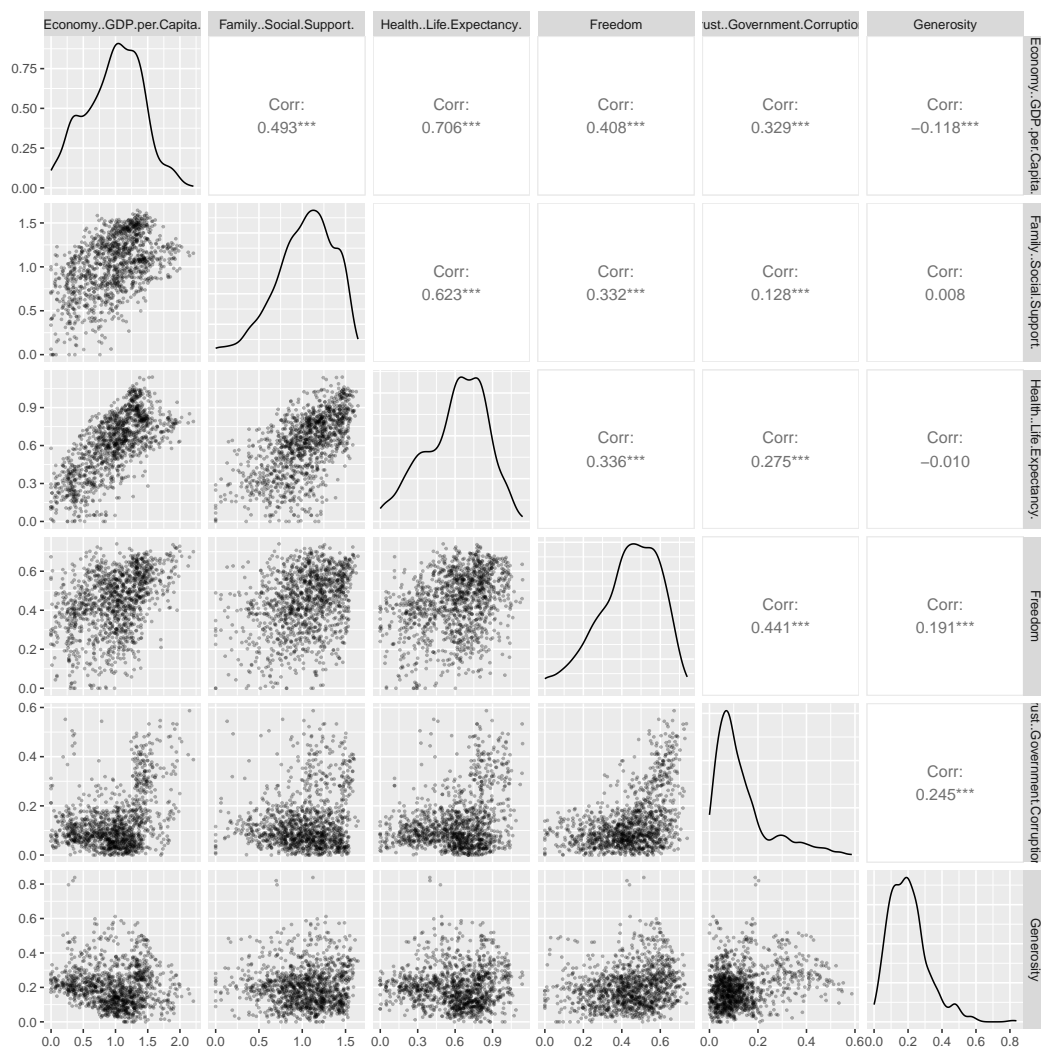


Figure 4: 变量相关系数图

2.4 PCA 分析与数据可视化

为了加深对数据的了解，寻找可能影响幸福指标的潜在因素，以及可视化多维数据，我们对 6 个幸福指标在标准化后进行主成分分析。为了为后续因子分析中因子个数选择提供依据，我们首先画出 scree plot (图5)。注意到在第二个主成分时出现了转折，同时参考数据也可以发现前两个主成分解释了 66% 的方差，而后面四个主成分每个的解释方差比例均在 10% 左右。为了可视化方便，综合考虑选择保留数据的前两个主成分得分，并将其前两个主成分得分可视化。

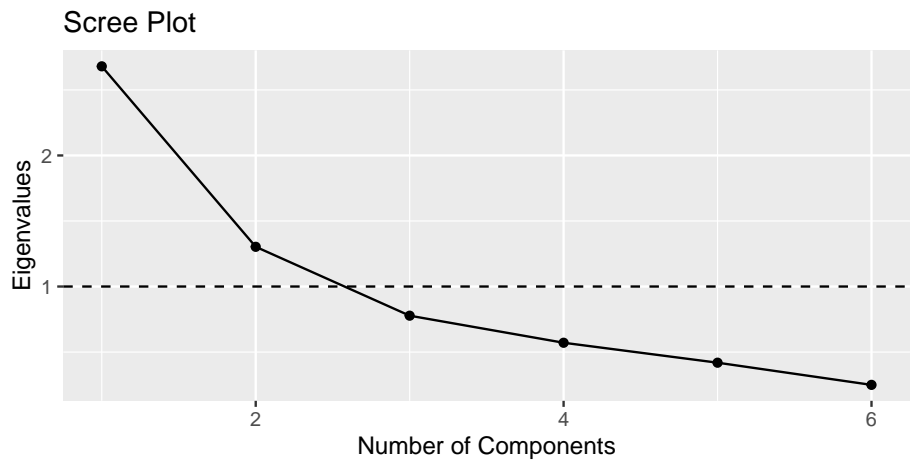


Figure 5: 崖底碎石图

其中，图6中的每个变量在横纵坐标上的投影就代表了其对相应主成分的贡献大小，从中也可以鲜明的看出代表 Economy、Health 和 Social Support 的三个变量对前两个主成分的贡献方式非常相似，且均位于第四象限；而剩余 3 个变量对于前两个主成分的贡献方式比较相似，均位于第一象限。这种现象说明我们可以将变量分为两组（根据其第二主成分贡献方式的不同），这为后续因子分析中的可解释方向提供了指导。

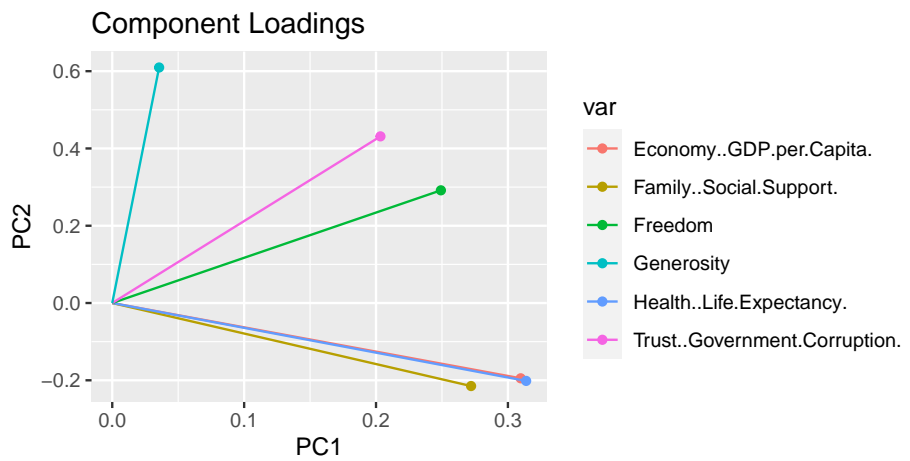


Figure 6: PCA 变量-主成分关系图

在样本主成分得分图7中，我们利用颜色的深浅表征了其幸福指数的大小（颜色越浅代表幸福指数越高），横纵坐标则为前两个主成分得分。其显示出我们可以将样本依据主成分得分大致分为两类，即幸福指数较高的和幸福指数较低的（可大致以  $PC1 = 0$  这条线作为区分）。再观察可以发现，幸福指数较高的组，其第一主成分明显较高；而幸福指数较低的组，其第一主成分明显较低。两组在第二主成分上的表现则难以区分。这提示我们后续进行因子分析时可以进行旋转来得到更好的解释性和分类能力。与此同时，在两类内部，我们还可以发现两个主成分间存在一定的相关关系（以幸福指数较高的组为例，可以发现随着第一主成分得分增加，第二主成分得分也在增加），这提示我们在后续因子分析中，可能存在因子间的相关性，单纯使用正交因子模型可能会出现由于假设违反带来的效用降低。

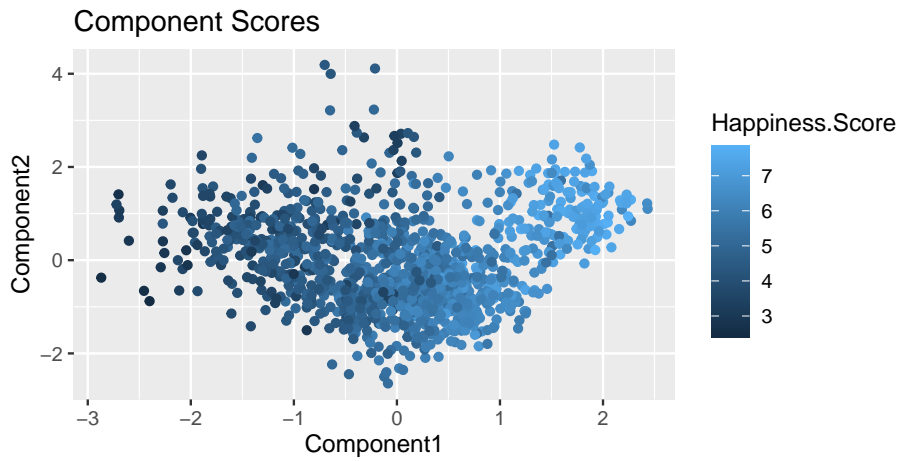


Figure 7: PCA 主成分得分图

### 3 数据降维

#### 3.1 因子分析

##### 3.1.1 正交因子分析

根据前文 PCA 分析得到的结果，我们首先利用最大似然估计法来求得可解释这 6 个变量的潜在影响因素，为了平衡数据尺度的问题，我们针对数据的相关系数矩阵进行因子分析。经过尝试和对比 2 因子模型和 3 因子模型的结果，我们最终选定了 2 因子模型，并通过旋转（旋转方式为 varimax）来获得更好的解释性，模型结果如下表所示。

注意到在 EDA 过程中，我们发现部分变量并不符合正态分布，而最大似然估计法恰恰使用了正态分布假设，因此在假设不成立的情况下其分析结果存在有偏性。为了解决这一问题，我们尝试使用最小残差法来代替最大似然估计法，其没有预先假设数据分布，因此可以一定程度缓解数据非正态分布的问题，模型结果同样如下表所示：

Variable	Maximum Likelihood			Minimum Residual		
	Factor 1	Factor 2	$\psi_i$	Factor 1	Factor 2	$\psi_i$
Economy (GDP per Capita)	0.772	0.185	0.370	0.79	0.14	0.35
Family (Social Support)	0.683	0	0.534	0.66	0.05	0.56
Health (Life Expectancy)	0.893	0	0.195	0.88	0.08	0.22
Freedom	0.369	0.492	0.621	0.39	0.54	0.55
Trust (Government Corruption)	0.233	0.746	0.390	0.24	0.66	0.51
Generosity	0	0.347	0.873	-0.10	0.40	0.83
Cumulative proportion of total standardized sample variance	0.343	0.503		0.34	0.50	

从上表可以看出，两种优化方式的结果大致相同，因此以 MLE 方法的结果为例进行分析。可以明确看出，第一个因子中代表 Economy、Health 和 Social Support 的三个变量的 loading 非常高，而在第二个因子中代表 Freedom、Trust 和 Generosity 的三个变量的 loading 非常高。因此我们可以将因子 1 解释为“人民物质需求满足的程度”（以下简称为“物质水平”），其主要表征某个国家可以量化的“硬实力”，如人均 GDP 等；将因子 2 解释为“人民精神需求满足的程度”（以下简称为“精神水平”），其主要表征某个国家难以被量化的“软实力”，如政府的信誉、人民的自由程度等等。其中因子 1 解释了 35% 的方差，而因子 2 仅解释了 15% 的方差，说明因子 1，即物质水平对人民幸福感影响更大。

##### 3.1.2 斜交因子分析

在刚刚的讨论过程中，特别是在提出两个因子的解释意义后，我们意识到两个因子正交的假设并不合理。首先从直觉上考虑，“仓廩实而知礼节”，一个物质水平高度发展的国家，往往其精神水平也会有长足的发展；相反，一个连基本的生存需求都无法保障的国家，也很难要求该国家的公民都很慷慨，因此这两个因子间应当存在一定的正相关性。其次在之前的 PCA 分析和 EDA 中，可以清晰地发现变量之间具有一定的相关性，所以引入斜交因子模型以捕捉这一信息。同样为了避免正态性违反带来的模型失效，我们使用 minres 方法优化模型，结果如下所示：



Variable	Minimum Residual		
	Factor 1	Factor 2	$\psi_i$
Economy (GDP per Capita)	0.80	0.02	0.35
Family (Social Support)	0.68	-0.05	0.56
Health (Life Expectancy)	0.90	-0.08	0.22
Freedom	0.31	0.50	0.55
Trust (Government Corruption)	0.13	0.65	0.51
Generosity	-0.17	0.43	0.83
Cumulative proportion of total standardized sample variance	0.35	0.50	

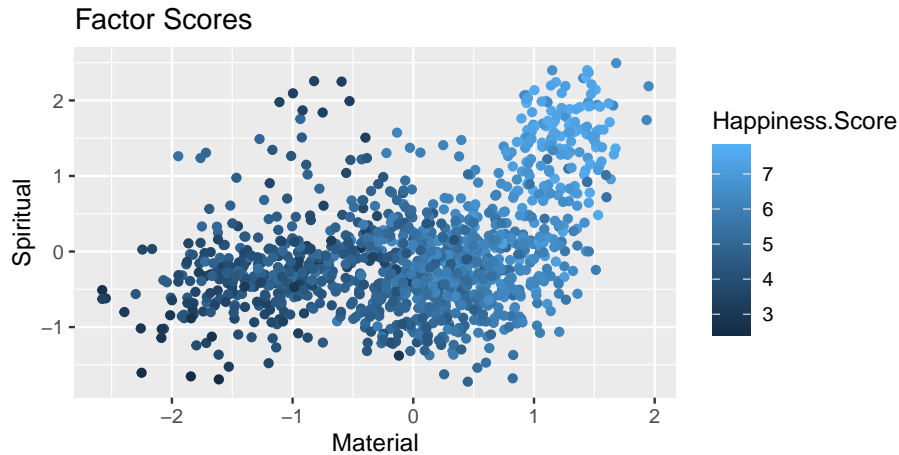


Figure 8: FA 因子得分图

其中，在 minres 方法下，两个因子之间的相关系数为 0.31，这个结果证明了引入斜交因子模型的必要性。同样绘制出因子得分图<sup>8</sup>，观察上述结果和得分图，我们可以注意到如下几个关键点：

1. 无论是基于哪种模型得到的结果，在物质水平因子中，均是 Health 变量载荷比较显著。这也符合直观认知，即人们只有在身体健康的情况下，才能享受到经济水平等其他指标带来的红利。
2. 无论是基于哪种模型得到的结果，在精神水平因子中，均是 Trust 变量载荷比较显著。这可能是因为在廉洁的政府（Trust 变量的评判标准即为政府腐败程度）往往意味着其受到及时有效的人民监督，即政府的权力不会无限制的扩张，因此其国民的自由程度等与精神水平相关的指标也会更高，从而说明精神水平的高低更容易反映在政府信誉这一指标上。
3. 在得分图的右上角（即物质水平和精神水平均高的国家），其人民幸福指数高；得分图的左下角（即物质水平和精神水平均低的国家），其人民幸福指数低。这也与我们的常识相符。
4. 在物质水平较低的情况下，单纯提升精神水平并不能显著提升人民的幸福指数；相反，在精神水平较低的情况下，提升物质水平可以显著提升人民的幸福指数。这说明在情况较恶劣时，人民更渴望得到物质需求的满足，而只有在物质水平达到一定水平后，才会开始追求精神需求的满足（之后我们会利用因果推断的方法证实并阐明这一结论）。

## 4 聚类、判别与回归分析

从因子得分图中可以看出，样本点可以大致分为 3 类，即最左侧的低幸福指数类、中间的中幸福指数类以及右上角的高幸福指数类。在本部分，我们首先使用聚类分析的方法验证这一观察的正确性。接下来，我们针对幸福指数程度（按照幸福指数排名将其分为“高”、“中”、“低”三类）进行判别分析，以期得到某个国家提升人民幸福感的关键政策建议（即找到不同类间的关键分类变量）。针对幸福指数这一连续型相应变量，利用回归分析方法建立预测模型，以期在得到幸福相关指标的情况下，可以估计该国家人民的幸福指数。

### 4.1 聚类分析

注意到因子得分图中每一类别的样本大致呈椭圆形分布，而且存在离群点（如图左上角的点），因此在进行层次聚类时使用了 Ward's Linkage 来保证聚类的形状大致为椭圆形，且尽可能减少离群点对聚类的影响。

基于同样的原因，相比使用 Kmeans 方法，本文主要依赖 Kmedoids 方法以尽可能减轻离群点对聚类的影响。通过观察可以发现，两种方法的结果相仿，一定程度上说明了上述样本分类的正确性。

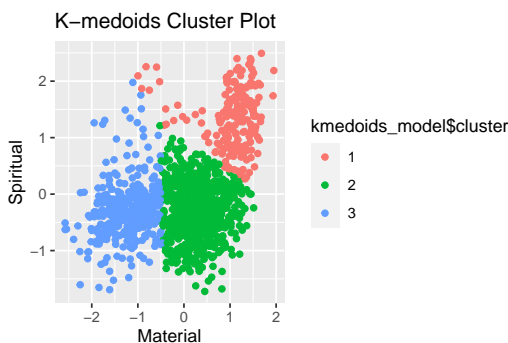


Figure 9: K-medoids 聚类结果

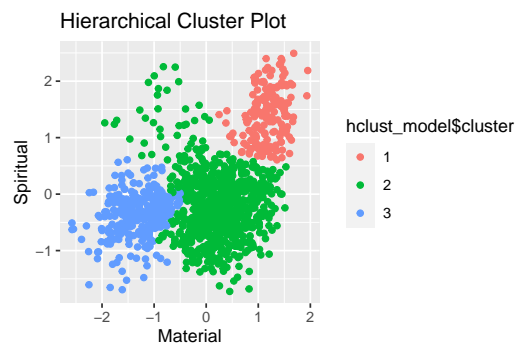


Figure 10: Hierarchical 聚类结果

## 4.2 判别分析

根据上文的分析结果，将样本按照幸福指数排名分为 3 类（1-50 名为高幸福度国家，51-100 名为中幸福度国家，101 名后为低幸福度国家），之后将数据集按照 7:3 的比例随机分为训练集和测试集，然后分别利用 LDA 等方法进行判别分析，训练集和测试集准确率如下表所示，测试集分类结果如下图所示（由于三种方法分类边界大致相同，因此只展示解释意义最明确的 LDA 方法的结果，全部结果可见代码）：

	LDA	QDA	SVM
Training Accuracy	0.7012	0.6907	0.7279
Testing Accuracy	0.729	0.6965	0.7561

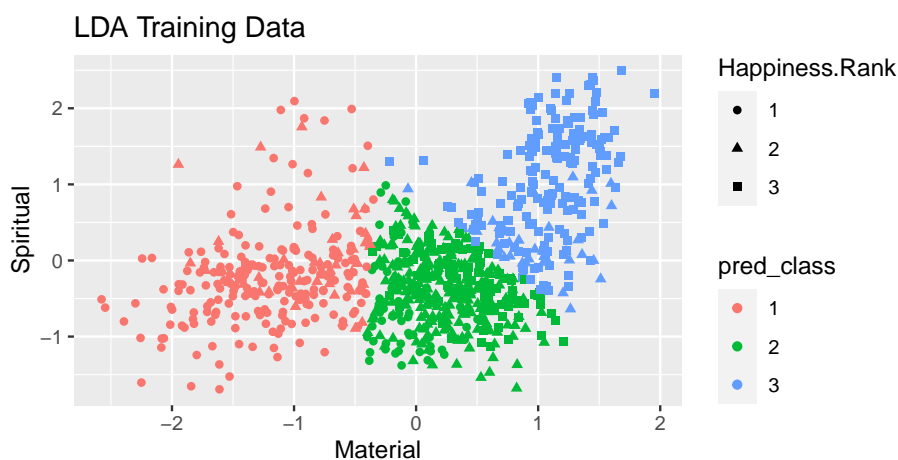


Figure 11: LDA 训练集分类结果

从如上的表格和结果图可以看到几个关键点：

1. LDA 表现比 QDA 表现好。这恰好呼应了 EDA 中部分变量并不符合正态性分布的性质，因此依赖于正态性假设的 QDA 方法效果并不好。正常来说，LDA 同样依赖于正态性假设，但本数据集存在一定特殊性：其三个类别较为均衡，即先验概率基本一致，且错判的 cost 一致，因此根据理论推导可知此时 LDA 得到的分类边界与 Fisher 方法得到的分类边界一致。注意到 Fisher 方法事实上不需要正态假设，因此在本数据集中，LDA 方法得到的分类边界由于不依赖于正态性假设，因此表现更好。
2. SVM 分类结果最好。这是因为 SVM 方法引入了核函数，将现有的特征拓展到高维空间，即丰富了特征的表示方式，因此允许更为灵活的分类型边界，相比 LDA 的线性分类边界有了很大的提升，因此分类效果更好。但是其一个典型问题是解释性较差，因此在准确率相差不大的情况下，本文选择了解释性更强的线性判别法为政府提升幸福感提供政策建议。
3. 主要分类错误的地方集中于误将第 2 类分为第 1 类或第 3 类。以误分为第 1 类为例，这可能是因为这类国家处于提升幸福感的过渡期，可能 GDP、反腐败等指标在政策层面已有实质性提升，因此统计时指

标较高，但由于政策落地需要时间，因此人民短时间内无法迅速感受到政策带来的幸福感提升，导致数据中的幸福指数相比指标来说偏低，因此导致错误分类。

4. 从图中可以看出，针对经济水平较差的国家（即图中粉色标识的类别），其首要任务是提升国家的物质水平而非精神水平，因为提升精神水平相当于平行于类别 1 和类别 2 的分类边界在移动，并不会为人民带来实质的幸福感提升。
5. 针对经济水平中等的国家（即图中绿色标识的类别），其应该精神水平和物质水平并举。注意到此时类别 2 和类别 3 的分类边界为略低于 45 度的斜向右下直线，因此这类国家还应将注意力稍多地分配在提升精神水平上，这样才能更有效率地提升人民幸福感。

### 4.3 回归分析

利用前文所述的因子分析和判别分析等方法，可以得到关于幸福指数的深度解读，但是此时没有建立预测的模型，因此利用回归模型来弥补这一缺憾。本部分同样按照 7:3 的比例，将数据集随机切分为训练集和测试集。由于时间限制和篇幅限制，本文仅仅尝试了线性回归分析和 SVR（其原理与 SVM 类似）方法，结果如下：

	Linear Regression	SVR
Training MSE	0.3072275	0.2342125
Testing MSE	0.3260827	0.2146241

由于线性回归分析中 6 个变量均非常重要，因此全部保留，给出如下所示的预测模型：

$$\begin{aligned} \text{Happiness.score} = & 2.1 + 0.96 \times \text{Economy} + 0.72 \times \text{Family} + 1.17 \times \text{Health} \\ & + 1.47 \times \text{Freedom} + 0.75 \times \text{Trust} + 0.90 \times \text{Generosity} \end{aligned}$$

我们可以绘制出残差分析图来寻找未来模型可能的改进方向：

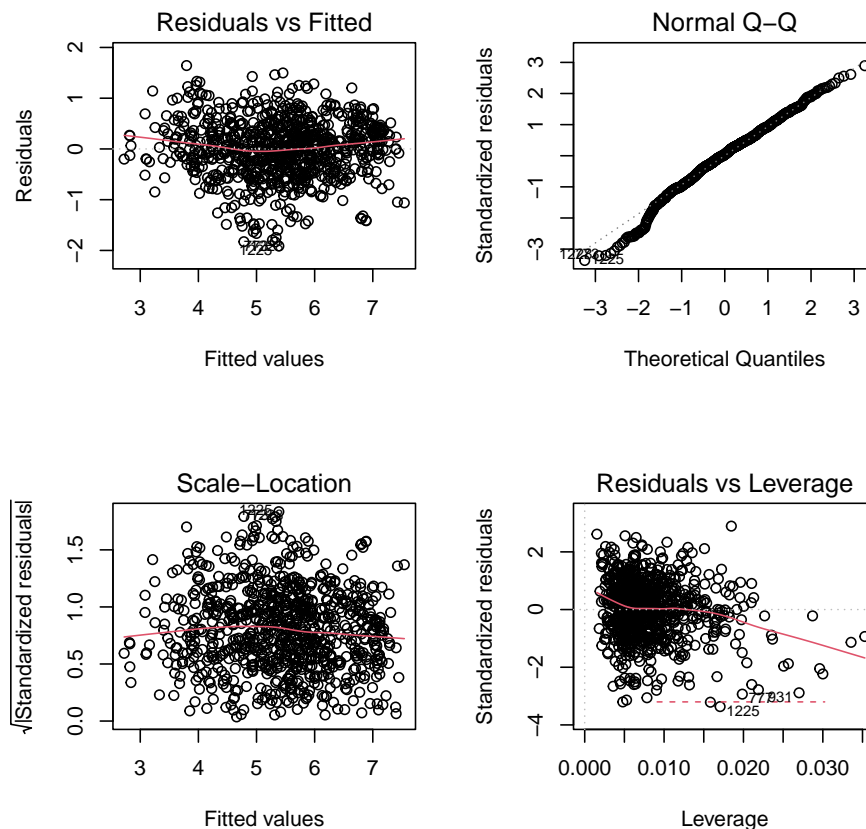


Figure 12: 线性回归残差分析图

从上面的训练结果和残差分析图中可以发现几个值得注意的关键点：



1. 线性回归分析中，全部 6 个变量均非常重要，因此没有明确理论支撑的情况下，可以全部保留。但根据残差分析结果，可以发现存在一定的异方差性，也存在一定数量的离群点，因此后期可以尝试使用加权线性回归的方式来消减异方差性，还可以通过去除离群点等方法使模型的泛化能力增强，而非拟合训练集中的噪声。
2. 线性回归分析中，仍可从 QQ 图中观察到其残差分布并不严格满足正态分布，这提示我们应当采取一定的数据变换，如 Box-Cox 等方法来使残差尽可能满足正态性。其中需要特别注意的是，形如 Box-Cox 这类灵活性很强的数据变换方法，可能由于数据集的随机性，导致其参数选择拟合了数据集中的噪声，从而损失了线性回归模型的可解释性。因此，在进行数据变换时，也应注意到保证模型的可解释性和泛化能力。
3. 注意到 SVR 的训练集的 MSE 明显高于测试集的 MSE，这意味着该模型存在一定的过拟合性，未来可以通过为模型添加正则化限制，降低其模型复杂度以更好地提升其泛化能力。
4. 正则化参数的选择方式（例如选择高斯核函数中的标准差  $\sigma$ ），可以利用交叉验证的方法来确定。具体来说，可以将训练集随机平均划分为 10 份，每次仅利用其中 9 份来训练模型，在余下的一份中检验模型表现，最终平均 10 个模型的表现效果。通过对比不同参数对应模型的表现，即可确定正则化参数的选择。之后再利用全部训练集数据，在选定的正则化参数下重新训练模型，以更全面的利用训练集信息。

## 5 因果推断

注意到上述方法均存在两个缺陷：

1. 上述方法均在寻找变量间的相关关系而非因果关系，因此我们无法确认某一因素对提升人民幸福感的真正作用（一个可能的例子是，GDP 对人民幸福感的提升作用事实上非常强，但由于其同时会显著影响人民的健康等因素，因此模型参数中反而 Health 的参数比 GDP 的参数更大）。
2. 上述方法均没有考虑到年份和地区带来的影响，但事实上在 EDA 中我们可以发现地区的差异对人民幸福指数和各项指标的影响是很大的，而上述模型并没有体现这一影响。

上文通过因子分析和判别分析，我们可以看出物质水平的提升无论在低幸福度国家还是中幸福度国家，都能带来人民幸福度的提升。因此我们基于因子分析的结果，力求通过因果推断的方法回答如下两个问题：

1. 物质水平的提升和人民幸福感的提升是否确实存在因果关系？
2. 不同国家或年份，物质水平提升对人民幸福感提升的因果效用是否存在差异？

为了回答上述两个问题，我们首先进行建模：

1. 根据 EDA 的结果，我们将 Region 分为两类：发达国家地区和发展中国家地区。一方面是因为 EDA 中发现这两类内部不同地区间各项指标的差异较小，而类间的差异较大，因此在尽可能保证类内数据共性的基础上，体现了类间数据的显著差异性，进而在后续的分析中能体现出地区差异对结果的影响。另一方面是因为这样划分解释性较强（按照常识，发展中国家人民对于物质水平的提升需求更强，而发达国家的人民可能更倾向于精神水平的提升），且避免了某一地区内数据量过小，导致结果可信度下降。
2. 根据文献调研和常识推断，我们将年份以 2019 年为界划分成前后两段。这样划分有两个原因，其一是因为在 EDA 过程中，2020 年-2022 年的数据中部分变量表现出与前面不同的分布情况，因此这样划分能够体现类间数据的差异性；其二是因为 2019 年底爆发了新冠疫情，将 2020 年及以后的数据单独划分出来能一定程度上反应新冠疫情对幸福指数和相关指标的影响，因此更好的提升了模型的可解释性。
3. 按照上述分类方式，我们可以将数据划分为四类。在每一类内，我们选择物质水平因子得分的中位数作为划分的阈值，将大于该阈值的样本视为实验组（即对其使用了 treatment），将小于该阈值的样本视为对照组（即对其使用了 control）。进一步，我们保留精神水平因子得分作为协变量。
4. 由于数据被切分为 4 类，我们可以将其视为 Rubin 框架下的 SRE (Stratified Random Experiment)。进一步，我们分别使用 Neyman's Approach<sup>[4]</sup> 和回归方法对物质水平的因果效应进行研究。Neyman 方法的核心是利用实验组（对照组）的样本均值来表征所有样本在（不）接受 treatment 后的样本均值，从而解决了缺失数据的问题（即同一样本不可能既接受 treatment 又接受 control，因此无法获得某一特定个体的因果效应）。

根据上述建模，我们可以得到如下结果（其中  $\hat{\tau}$  即为估计的因果效应）：

	Developed Region		Developing Region	
	Before Covid-19	After Covid-19	Before Covid-19	After Covid-19
$\hat{\tau}$	0.6789	0.3410	1.0457	1.2104
$\hat{\sigma}(\hat{\tau})$	0.1130	0.1315	0.1230	0.1799
CI	(0.4574, 0.9003)	(0.0833, 0.5986)	(0.7949, 1.2966)	(0.8579, 1.5630)

Table 1: Neyman's Approach without regression for SRE

	Developed Region		Developing Region	
	Before Covid-19	After Covid-19	Before Covid-19	After Covid-19
$\hat{\tau}$	0.22378	-0.04958	0.72993	0.85592
$\hat{\sigma}(\hat{\tau})$	0.06677	0.07621	0.23560	0.26672
CI	(0.09291, 0.3546)	(-0.1990, 0.09979)	(0.2682, 1.1917)	(0.3331, 1.3787)

Table 2: Neyman's Approach with regression for SRE

从表中可以发现如下几个值得注意的关键点：

1. 在同一张表内进行对比，可以发现发展中国家地区的因果效应更为显著，这也与前文聚类分析和判别分析的结果相符，即发展中国家人民最渴望得到物质水平的提升，这也与我们的常识认识相符。值得注意的是，对比新冠前后的结果，可以发现发达国家地区和发展中国家地区的因果效应变化趋势相反。这可能是因为对发达国家地区来说，人们从严重的疫情恢复过来后，由于经历过生死考验，因此更注重及时行乐带来的精神愉悦（即 *Freedom* 这一变量），更珍惜和身边人的亲情和友情（即 *Generosity* 这一变量），而且其物质水平也支持这些地区的人民这样做。相反，对于发展中国家地区来说，由于疫情对其经济造成了重创，使其本就较低的物质水平再次降低，因此人们更追求物质水平的提升（一个典型的例子可参考图 11 中第 2 类别的国家，其如果物质水平下滑导致其滑落至第 1 类别，那么精神水平的提升并不能为本国人民幸福感带来多少提升）。
2. 通过对比有无引入协变量精神水平的结果，首先可以发现精神水平对国民幸福感的提升是至关重要的（表现在物质水平的因果效应均有明显下滑），这也侧面证明了引入协变量的重要性。进一步分析可以发现引入精神水平后，发达国家地区的物质水平和人民幸福度之间基本可以认为不存在显著的因果效应，再次证明了发达国家要想提升人民的幸福度，其关键任务是满足本国人民的精神需要；而发展中国家的因果效应仍然非常显著，也说明了发展中国家要想提升人民的幸福度，关键任务是提升本国人民的物质水平。

## 6 总结与反思

### 6.1 结论

根据如上的分析，我们可以回答在本文最初提出的问题：

1. 按照斜交因子模型的解释，存在两个潜在影响因素影响幸福指标，其显示意义分别是一个国家可以量化分析的“硬实力”——人民物质需求满足情况和一个国家难以量化分析的“软实力”——人民精神需求满足情况；且两者具有一定的相关性。若观测样本的两因子得分均较高，则该国国民大概率幸福指数较高；若观测样本的两因子得分均较低，则该国国民大概率幸福指数较低，这也体现了这两个因子对幸福指数的解释能力。
2. 按照聚类分析、判别分析和因果推断的结果，我们可以得出结论：对于物质水平较低的国家，无论其精神水平高低，其若想提升本国人民的幸福感，关键任务是大力发展生产力，提升本国的物质水平；而对于物质水平已经达到较高水平的国家，要想提升本国人民的幸福感，不仅要提升本国的物质水平，更要着眼于提升本国的精神水平；对于物质水平高度发达的国家来说，其首要任务就是提升本国的精神水平，单独提升物质水平对于提升本国国民的幸福感的边际作用已经很小了。
3. 根据回归分析的结果，要想根据幸福指标预测某国的幸福指数，应着眼于 *Health* 和 *Freedom* 这两个变量，因为其对幸福指数的影响比较显著。这也侧面说明，某国想要迅速提升本国人民的幸福度，首先应该从提升人民的平均预期寿命和人民自由程度入手。
4. 不同国家和不同时间提升国民幸福感的显著影响因素不尽相同。但总的来说，发展中国家的显著影响因素是物质水平和发达国家的显著影响因素是精神水平。疫情后发达国家人民更看重精神水平的提升，而发展中国家更看重物质水平的提升。

## 6.2 政策建议

根据上述分析结论，我们可以因地制宜地对我国国民幸福感提升提出政策建议：

1. 由于我国仍处于并将长期处于社会主义初级阶段，因此政府应首先大力发展生产力，提升国民的物质水平，这样可以迅速提升国民幸福度。具体来说，可以完善社会保障体系（即优化回归模型中的 **Social Support** 变量）；大力招商引资、提升科技水平，进而提升 GDP；完善医疗保障体系和养老保障体系，让人民老有所养、病有所医（即优化模型中的 **Health** 变量）。
2. 由于我国已经处于物质水平较高的阶段，因此也应将一部分注意力放在满足国民的精神需要上。具体来说，可以实现反腐倡廉制度化、常态化（即优化模型中的 **Trust** 变量）；可以在社会上宣传鼓励或物质激励见义勇为等行为，鼓励慈善事业发展，营造出良好的互帮互助的社会氛围（即优化模型中的 **Generosity** 变量）。

## 6.3 反思与展望

经过分析，我认为本文还存在如下缺陷及可能存在的改进方向：

1. 注意到本数据集有两个变量存在明显的右偏现象，并不满足正态性假设，因此后文中依赖正态性假设的模型均不同程度的出现了问题，因此在后续的工作中可以尝试对变量进行变换以保证其分布基本满足正态性假设，也可以通过调研文献等方式对数据的分布等信息进行更严格的推断，文章的逻辑性会更强，结论也会有更高的可信度。
2. 在进行判别分析时，仅采用了 3 种方法，且仅针对 LDA 方法进行了详尽的论述，而没有尝试其他同样具有高解释力的模型，例如梯度提升树等模型。未来可以尝试这类模型以期对幸福度的影响因素有更深刻的认识。还可以尝试 ECM 等方法更充分的利用信息进行判别（例如收集关于高中低三种幸福度国家概率分布的先验等）。
3. 在进行回归分析时，由于篇幅和时间所限，并没有尝试更丰富的回归方法，也并没有针对模型失败的地方做出详细的优化，未来可以针对我在回归分析 4.3 这一节中提出的问题和优化方法，使得模型的拟合能力、预测能力和解释能力得到进一步的提升。
4. 在因果推断中，我判断一个国家是否接受了 **treatment** 的方法是比较粗糙的，事实上也一定程度上违反了 SUTVA 假设 [4]，因为不同样本的物质水平因子得分事实上是不同的。一个可能的后果是某些国家的物质水平因子得分过高，甚至弭平了其精神因子得分低的缺陷，导致该国的幸福指数很高，进而导致对照组的因果效应被错误地高估了（注意到采用回归方法进行因果推断时，精神水平因子得分和幸福度呈正相关，因此低精神得分意味着高因果效应）。因此之后可以尝试纳入这一信息来实现更完善更有效的因果分析推断。
5. 在因果推断中，精神水平因子得分事实上存在一定的不平衡问题（详情可见代码），这也会削弱结果的有效性。因此未来可以使用 **Blocking** 等方法恢复数据平衡性，以完善因果推断过程及提升结果有效性。
6. 影响幸福指数的因素应该有更多，事实上在 2023 年的全球幸福指数报告 [2] 中就指出了一个新变量 **State Effectiveness**。因此未来也可以收集更多的相关数据，挖掘更多潜在的影响因素。

## 7 附录

由于本人代码习惯，将部分数据预处理的代码以 **python** 形式呈现，后面建模分析的代码则使用 **R** 呈现。其中，**R** 语言部分代码以 **R** 文件呈现，**python** 部分代码以 **jupyter notebook** 文件呈现。

## References

- [1] <https://worldhappiness.report/>.
- [2] <https://worldhappiness.report/ed/2023/executive-summary>.
- [3] <https://www.kaggle.com/datasets/shivkumarganesh/world-happiness-report-20152022>.
- [4] Guido W Imbens and Donald B Rubin. Causal inference in statistics, social, and biomedical sciences (2015). *Google Scholar Google Scholar Digital Library Digital Library*.
- [5] 人民日报. 中共中央关于党的百年奋斗重大成就和历史经验的决议, 2021.
- [6] 佩德罗·孔塞桑; 罗米娜·班德罗; 卢艳华; 主观幸福感研究文献综述, 2013.