

《非参数统计》

期末报告

报告题目：经合（OECD）国家女性创业及劳动情况探究

学生姓名：曹梓上

第一部分：案例展示

在全球经济快速发展的背景下,女性在创业和劳动领域的参与度日益成为衡量一个国家经济活力与社会进步的重要指标。经济与合作发展组织（OECD）成员国作为全球经济体系的重要组成部分,其在推动性别平等、促进女性创业与就业方面的实践经验具有重要的研究价值。深入探究 OECD 国家女性创业及劳动情况,不仅有助于揭示不同国家在相关领域的发展特点与差异,还能为其他国家制定促进女性经济参与的政策提供参考,对推动全球性别平等和经济可持续发展具有深远意义。

本研究基于 Kaggle 官网提供的经合（OECD）国家女性创业及劳动情况数据集,该数据集源自 2015 年发布的女性创业指数和全球创业指数报告,全面记录了多个关键维度的数据信息,为研究提供了坚实的数据基础。

具体地,数据集由 51 条数据构成,包括 8 个变量,具体变量如表 1 所示:

表 1 案例变量简介

变量名称	变量含义	数据类型
Country	各个 OECD 国家名称	文本型
Level of development	各国家发展程度	名义（定性）变量
European Union Membership	各国家欧盟成员国身份标识	名义（定性）变量
Currency	各国家货币类型	文本型
Women Entrepreneurship Index	各国家女性创业指数	数值型
Entrepreneurship Index	各国家所有创业者创业指数	数值型
Inflation rate	各国家通货膨胀率	数值型
Female Labor Force Participation Rate	各国家女性劳动参与程度	数值型

完整数据集见附录

基于上述对数据集及其变量的清晰展示,后续将采用非参数分析方法进一步挖掘数据背后的规律与特征。

第二部分：案例分析

该部分主要对应用于案例的四种非参数方法适用性进行验证,主要涉及的 4 个数值型因素均为连续型变量。

(1) Wilcoxon 符号秩检验

Wilcoxon 符号秩检验（Wilcoxon Signed-Rank Test）是一种非参数统计方法,用于检验

配对样本或单个样本是否来自中位数特定的总体，主要通过分析差值的符号和大小（秩次）来推断总体分布是否存在差异。

➤ 问题描述

假定世界范围内女性创业指数的中位数大于 50 为基本合格标准，以经合国家为样本推断世界全体国家女性创业指数是否达到合格标准。

➤ 适用条件

为便于观察，对“女性创业指数”因素的分布进行可视化。

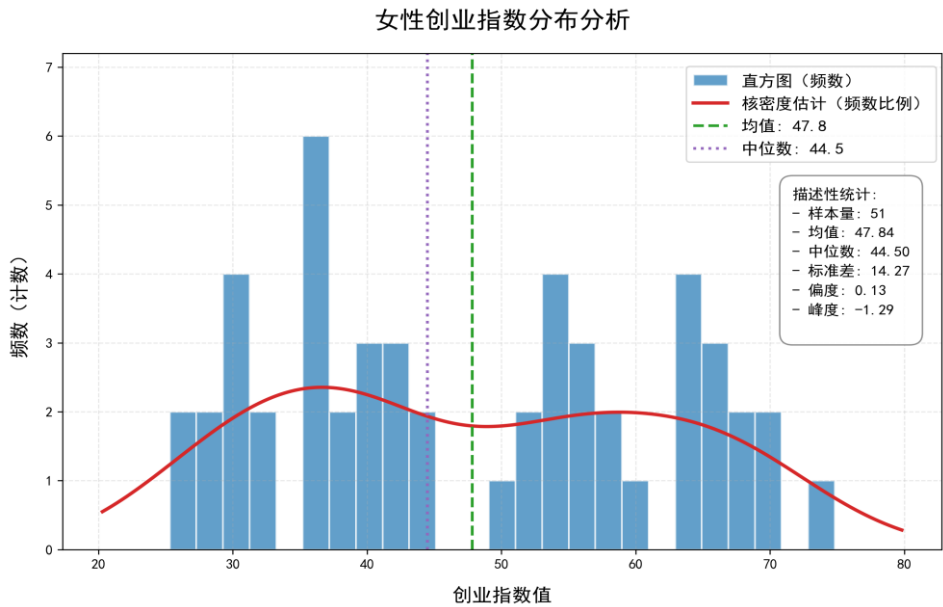


图 1 女性创业指数分布图

由图 1 易观察到，该因素满足 Wilcoxon 符号秩检验的适用条件：

1. 数据性质为单样本连续型数据（有序分类数据同样适用）
2. 观测值相互独立

样本中的每个观测值都应相互独立，即一个观测值的取值不会对其他观测值的取值产生影响。在研究女性创业指数时，各个国家的女性创业指数相互独立，一个国家的情况不会直接影响另一个国家的女性创业指数。

3. 不依赖正态分布

Wilcoxon 符号秩检验属于非参数检验，不要求数据服从正态分布。这是它相较于参数检验（如配对 t 检验）的一大优势。在实际研究中，数据往往不满足正态分布，尤其是样本量较小时，这时 Wilcoxon 符号秩检验就更适用。

4. 差值分布近似对称

在进行单样本或配对样本的 Wilcoxon 符号秩检验时，理论上要求数据与假设中位数的差值分布近似对称。由图 1 易看出，对于中位数 44.5，左右两侧分布近似对称，即与中位数的差值分布近似对称。

（2）Wilcoxon (Mann-Whitney) 秩和检验

Wilcoxon (Mann-Whitney) 秩和检验，也称为曼 - 惠特尼 U 检验，是一种非参数检验方法，用于推断两个独立样本所来自的总体分布是否存在显著差异。

➤ 问题描述

发达国家与发展中国家的女性创业指数的中位数是否相同

➤ 适用条件

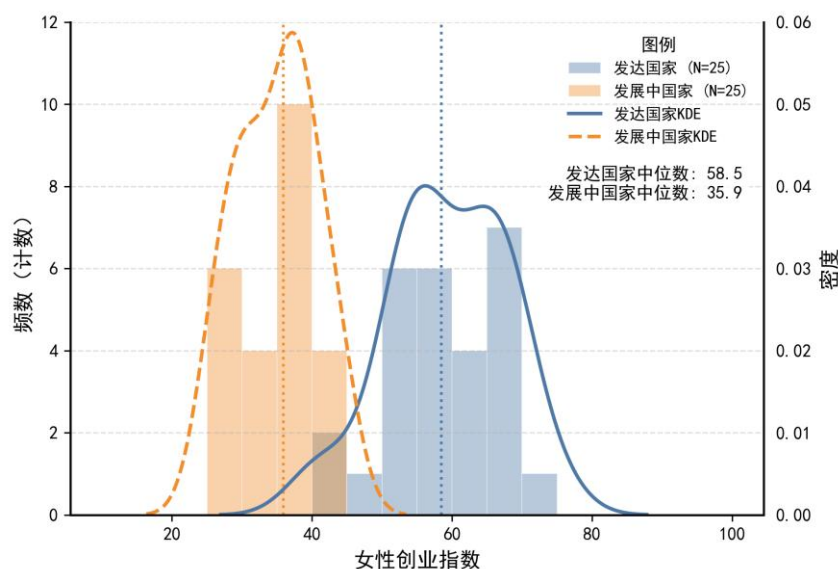


图 2 女性创业指数分组分布图

为便于观察，分别对发达国家与发展中国家的“女性创业指数”因素的分布进行分组可视化。

由图 2 易观察到，该因素满足 Wilcoxon（Mann--Whitney）秩和检验的适用条件：

1. 数据类型：两组连续型变量。
2. 独立性：两组数据独立（发达国家与发展中国家样本无重叠）。
3. 分布形状：数据量小（ $N = 25$ ）且不服从正态分布，且两组数据分布形状相似。

（3）三类分布检验方法

Kolmogorov-Smirnov 检验（简称 K-S 检验）主要面向连续型数据（有序离散数据需先转换），用于检验样本数据是否服从某个已知的理论分布（如正态分布、均匀分布等），或检验两个独立样本是否来自同一分布。它通过比较经验分布函数（ECDF）与理论分布函数的最大垂直距离来判断拟合优度。

Shapiro-Wilk 检验是一种专门用于判断样本数据是否服从正态分布的方法，属于显著性检验中的非参数检验。基于样本数据的排序信息，通过构造 W 统计量来衡量样本分布与正态分布的拟合优度。

Pearson 卡方拟合优度检验主要面向分类数据（连续型数据需先转换），是一种用于判断样本数据是否服从某一理论分布的非参数检验方法。其核心思想是通过比较样本观测频数（Observed Frequencies）与理论期望频数（Expected Frequencies）的差异，构建卡方统计量，衡量数据与理论分布的拟合程度。

➤ 问题描述

经合国家的女性劳动参与度是否服从正态分布。

➤ 适用条件

为便于观察，将女性劳动参与度因素的数据分布进行可视化：

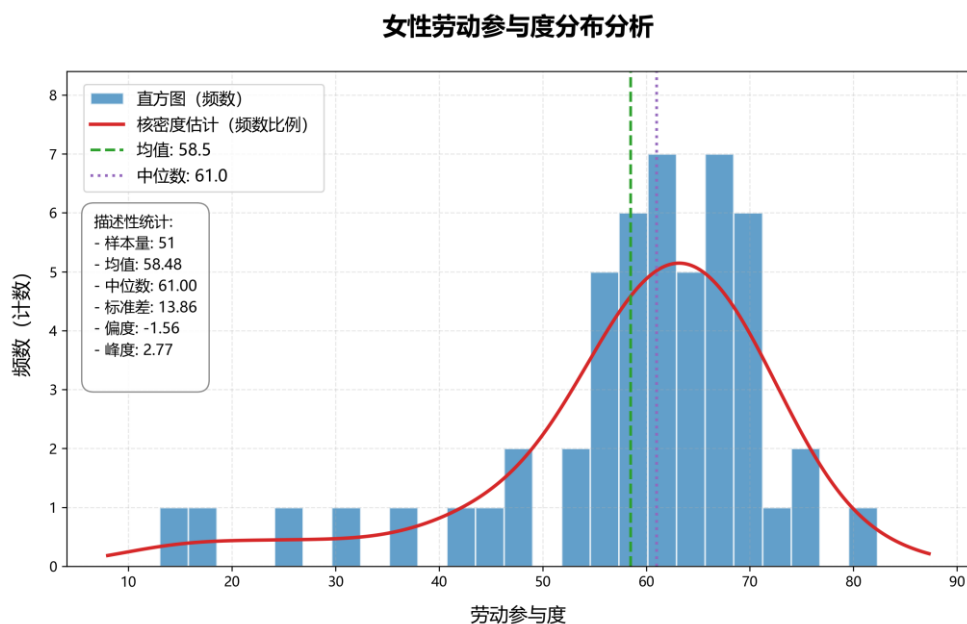


图 3 女性参与度分布图

由图 3 观察到，“女性劳动参与度”因素频数拟合曲线似乎服从正态分布，为了进一步检验这一假设，案例使用上述三类检验方法进行验证并加以比对，探究其中差异，该因素变量满足以下条件：

1. 连续型数据
2. 样本观测值相互独立
3. 理论分布的具体形式和参数已知

第三部分：非参数分析

(1) 世界女性创业情况

假定世界范围内以女性创业指数大于 50 为基本合格标准，使用 Wilcoxon 符号秩检验以经合国家为样本推断世界全体国家女性创业指数是否达到合格标准。

1. 假设

H_0 : 世界范围内以女性创业指数小于等于 50

H_1 : 世界范围内以女性创业指数大于 50

2. 计算 Wilcoxon 统计量（使用 python）

Wilcoxon 统计量: 541.5

p 值: 0.8726252171441391

3. 检验

在 $\alpha = 0.05$ 的显著性水平下，相应 p 值大于 0.05，不能拒绝原假设，没有足够证据表明总体女性劳动参与率大于 50，即不能认为总体女性创业指数大于 50。

(2) 国家发展情况的影响

将女性创业指数依据国家发展情况分为两类，使用 Wilcoxon (Mann-Whitney) 秩和检验方法检验发达国家与发展中国家的女性创业指数的中位数是否相同。

1. 假设

H0: 发达国家与发展中国家的女性创业指数的中位数相同

H1: 发达国家与发展中国家的女性创业指数的中位数不同

2. 数据整理与排序赋秩

表 2 国家分组排秩

发达国家组 (27 个)
54.9, 63.6, 55.4, 66.4, 68.8, 63.6, 43.0, 64.3, 51.4, 56.6, 58.5, 69.3, 54.8, 55.9, 52.5, 49.9, 69.7, 53.7, 57.7, 66.7, 74.8, 68.0, 40.0, 66.3, 59.8, 63.7, 53.4
发展中国家组 (24 个)
27.4, 35.7, 29.7, 31.6, 31.1, 38.3, 36.1, 32.3, 27.7, 29.9, 25.8, 25.3, 38.6, 41.2, 39.2, 42.8, 36.9, 43.6, 35.6, 37.0, 36.6, 30.7, 39.3, 44.5

3. 计算秩和

发达国家组秩和 (R_1) :

$$R_1 = 20 + 23 + 26 + 27 + 28 + 29 + 30 + 31 + 32 + 33 + 34 + 35 + 36 + 37 + 38 + 39.5 + 39.5 + 41 + 42 + 43 + 44 + 45 + 46 + 47 + 48 + 49 + 50 = 1165.5$$

发展中国家组秩和 (R_2) :

$$R_2 = \frac{N(N+1)}{2} - R_1 = \frac{51 \times 52}{2} - 1165.5 = 1326 - 1165.5 = 160.5$$

4. 计算 U 统计量

$$U_1 = R_1 - \frac{n_1(n_1+1)}{2} = 1165.5 - \frac{27 \times 28}{2} = 1165.5 - 378 = 787.5$$

$$U_2 = R_2 - \frac{n_2(n_2+1)}{2} = 160.5 - \frac{24 \times 25}{2} = 160.5 - 300 = -139.5$$

$$U = \min(U_1, U_2) = \min(787.5, 139.5) = 139.5$$

5. 计算 Z 值与 p 值

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} = \frac{139.5 - \frac{27 \times 24}{2}}{\sqrt{\frac{27 \times 24 \times 52}{12}}} = \frac{139.5 - 324}{\sqrt{2808}} = \frac{-184.5}{52.99} \approx -3.48$$

查标准正态分布表:

$$Z = -3.48 \Rightarrow \text{双尾 } p \text{ 值} < 0.001$$

6. 效应量计算

$$\delta = \frac{2U}{n_1 n_2} - 1 = \frac{2 \times 139.5}{648} - 1 = 0.43 - 1 = -0.57$$

效应量解释：Cliff's Delta = -0.57，表示大效应（ $|\delta| > 0.474$ ）。

表 3 Wilcoxon 秩和检验结果汇总表

统计量	值
发达国家样本量 (n_1)	27
发展中国家样本量 (n_2)	24
U 值	139.5
Z 值	-3.48
p 值	< 0.001
Cliff's Delta	-0.57

由表 3 易得结论：在显著性水平 $\alpha=0.05$ 下， $p < 0.001$ ，拒绝原假设，即发达国家与发展中国家的女性创业指数中位数存在显著差异，发达国家显著更高（效应量大）。

（3）女性劳动参与度的正态性检验

综合三类正态性检验方法检验经合国家的女性劳动参与度是否服从正态分布。

1. 假设

H0：经合国家的女性劳动参与度服从正态分布

H1：经合国家的女性劳动参与度不服从正态分布

2. 计算三类统计量

分别使用 Kolmogorov-Smirnov 检验、Shapiro-Wilk 检验及 Pearson 卡方拟合优度检验进行编程检验，其中较为特殊地，Pearson 卡方拟合优度检验需要额外进行分类变量转换操作，因此对其详细展开阐述。

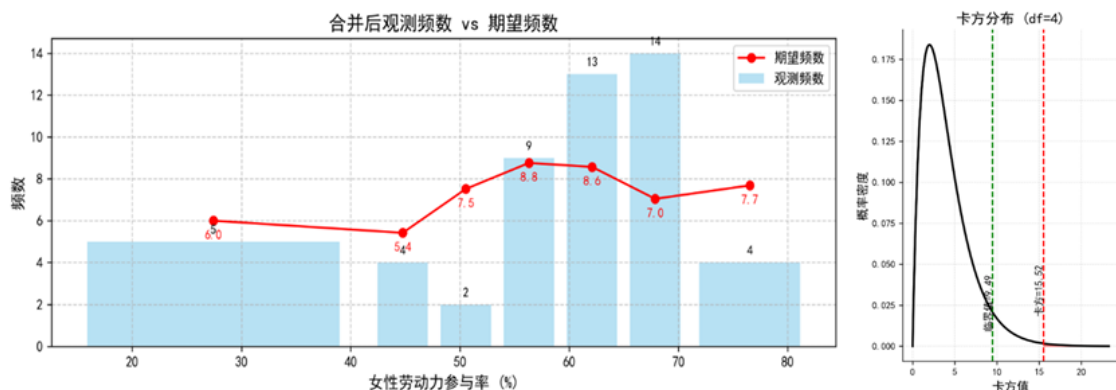


图 4 Pearson 卡方拟合优度检验分箱结果

首先对女性劳动力参与率数据分箱处理，将原始 12 箱合并为 7 箱，具体分箱情况图左子图。接着估计正态分布参数，算出各分箱期望频数，统计观测频数。之后依据公式：

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

其中 O_i 是第 i 个分箱的观测频数， E_i 是第 i 个分箱的期望频数。计算卡方统计量为 15.5184。随后确定自由度为 4（分箱数 7 减去估计参数个数 2 再减 1），并据此算出 p 值。

3. 检验

表 4 正态检验结果汇总表

检验方法	统计量	p 值	结论
Kolmogorov-Smirnov 检验	0.1844	0.0544	不拒绝原假设（数据服从正态分布）
Shapiro-Wilk 检验	0.8595	0.0000	拒绝原假设（数据不服从正态分布）
Pearson 卡方拟合优度检验	15.5184	0.003738	拒绝原假设（数据不服从正态分布）

由表 4 易知：在显著性水平 $\alpha = 0.05$ 下，基于频率统计的 K-S 检验所得 p 值大于 0.05，接受原假设。但值得注意的是，所得 KS 统计量为 0.1844，查精确临界值表得临界值约为 0.190（ $n = 51$ ），分布差异并不显著。为保证检验的可靠性，结合后两种检验方法进一步分析，其中，专用于检验正态分布的 Shapiro-Wilk 检验结果显示 p 值接近于 0，拒绝原假设，将该因素转换为分类数据后（详见图 ），基于频数统计的 Pearson 卡方拟合优度检验结果同样显示拒绝原假设，即数据不服从正态分布。

同时，绘制正态 Q-Q 图方便进一步直观观察，以作辅证：

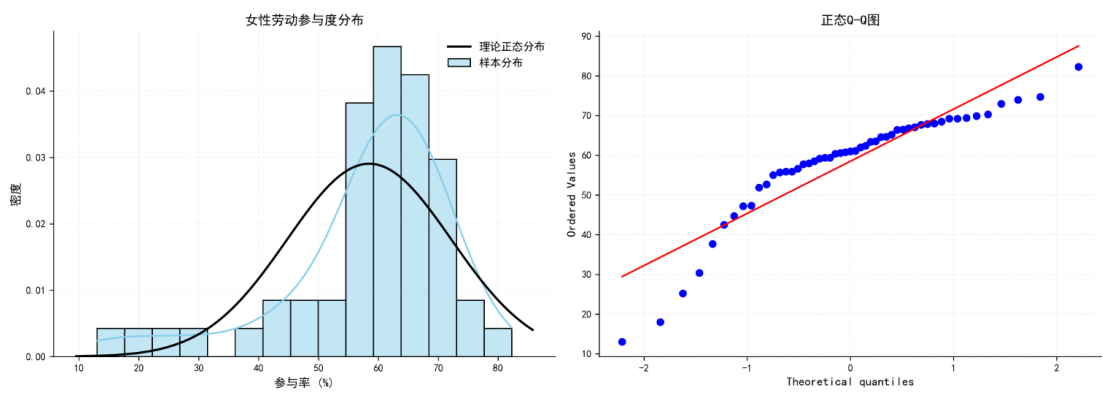


图 5 女性劳动参与度正态 Q-Q 图

由图 5 中右子图易观察到，正态分位数观测值明显偏离相应理论值，因此，得出最终结论：经合国家的女性劳动参与度不服从正态分布。

附录：

Country	Level of development	Membership	Currency	Women Entrepreneurship Index	Entrepreneurship Index	Inflation rate	Female Labor Force Participation Rate
Austria	Developed	Member	Euro	54.9	64.9	0.90	67.10
Belgium	Developed	Member	Euro	63.6	65.5	0.60	58.00
Estonia	Developed	Member	Euro	55.4	60.2	-0.88	68.50
Finland	Developed	Member	Euro	66.4	65.7	-0.20	67.70
France	Developed	Member	Euro	68.8	67.3	0.00	60.60
Germany	Developed	Member	Euro	63.6	67.4	0.50	69.90
Greece	Developed	Member	Euro	43.0	42.0	-1.70	42.50
Ireland	Developed	Member	Euro	64.3	65.3	-0.30	59.40
Italy	Developed	Member	Euro	51.4	41.3	0.00	47.20
Latvia	Developed	Member	Euro	56.6	54.5	0.20	66.40
Lithuania	Developed	Member	Euro	58.5	54.6	-0.90	66.50
Netherlands	Developed	Member	Euro	69.3	66.5	0.60	69.20
Slovakia	Developed	Member	Euro	54.8	45.4	-0.30	55.90
Slovenia	Developed	Member	Euro	55.9	53.1	-0.50	61.00
Spain	Developed	Member	Euro	52.5	49.6	-0.50	52.70
Croatia	Developed	Member	National Currency	49.9	40.6	-0.50	60.40
Denmark	Developed	Member	National Currency	69.7	71.4	0.50	70.30
Poland	Developed	Member	National Currency	57.7	47.4	-0.90	56.60
Sweden	Developed	Member	National Currency	66.7	71.8	0.00	74.00
Australia	Developed	Not Member	National Currency	74.8	77.6	1.50	66.80
Iceland	Developed	Not Member	National Currency	68.0	70.4	1.60	82.30
Japan	Developed	Not Member	National Currency	40.0	49.5	0.80	64.70
Norway	Developed	Not Member	National Currency	66.3	65.6	2.17	69.20
Singapore	Developed	Not Member	National Currency	59.8	68.1	-0.50	59.18

Country	Level of development	Membership	Currency	Women Entrepreneurship Index	Entrepreneurship Index	Inflation rate	Female Labor Force Participation Rate
Switzerland	Developed	Not Member	National Currency	63.7	68.6	-1.10	74.70
Taiwan	Developed	Not Member	National Currency	53.4	69.1	-0.61	55.00
Algeria	Developing	Not Member	National Currency	27.4	30.2	4.80	18.00
Argentina	Developing	Not Member	National Currency	35.7	37.2	26.50	47.30
Bolivia	Developing	Not Member	National Currency	29.7	28.0	4.10	69.40
Bosnia and Herzegovina	Developing	Not Member	National Currency	31.6	28.9	-1.00	51.90
Brazil	Developing	Not Member	National Currency	31.1	25.8	10.67	55.90
China	Developing	Not Member	National Currency	38.3	36.4	1.40	62.40
Costa Rica	Developing	Not Member	National Currency	36.1	37.7	0.80	59.40
Ecuador	Developing	Not Member	National Currency	32.3	28.2	-0.50	63.50
Egypt	Developing	Not Member	National Currency	27.7	28.1	11.00	64.60
El Salvador	Developing	Not Member	National Currency	29.9	29.6	-2.25	55.70
Ghana	Developing	Not Member	National Currency	25.8	24.8	17.20	60.80
India	Developing	Not Member	National Currency	25.3	25.3	5.90	61.10
Jamaica	Developing	Not Member	National Currency	38.6	27.2	3.70	37.70
Macedonia	Developing	Not Member	National Currency	41.2	37.1	3.70	73.00
Malaysia	Developing	Not Member	National Currency	39.2	40.0	2.30	58.50
Mexico	Developing	Not Member	National Currency	42.8	30.7	2.70	44.70

Country	Level of development	Membership	Currency	Women Entrepreneurship Index	Entrepreneurship Index	Inflation rate	Female Labor Force Participation Rate
Panama	Developing	Not Member	National Currency	36.9	32.2	0.10	67.90
Peru	Developing	Not Member	National Currency	43.6	30.9	3.50	63.40
Russia	Developing	Not Member	National Currency	35.6	31.7	15.50	65.20
Saudi Arabia	Developing	Not Member	National Currency	37.0	49.6	1.20	13.00
Thailand	Developing	Not Member	National Currency	36.6	32.1	-0.90	62.00
Tunisia	Developing	Not Member	National Currency	30.7	35.5	4.80	25.19
Turkey	Developing	Not Member	National Currency	39.3	54.6	7.70	30.40
Uruguay		Not Member	National Currency	44.5	41.4	8.67	68.00

部分代码:

```

01 # 1. Kolmogorov-Smirnov 检验 (正态性)
02 def ks_test(sample_data):
03     """执行 Kolmogorov-Smirnov 检验"""
04     mu, sigma = np.mean(sample_data), np.std(sample_data)
05     stat, p_value = stats.kstest(sample_data, 'norm', args=(mu, sigma))
06     conclusion = (
07         f"p 值={p_value:.4f}, 拒绝原假设 (数据不服从正态分布)"
08         if p_value < 0.05
09         else f"p 值={p_value:.4f}, 不拒绝原假设 (数据服从正态分布)"
10     )
11     return {
12         'test_type': 'Kolmogorov-Smirnov 检验',
13         'statistic': stat,
14         'p_value': p_value,
15         'conclusion': conclusion
16     }

```