应用方向：

1、相关性分析：地区和国家使用趋势与能源类型特性（年度、月度）

2、相关性分析：可再生能源使用量的影响因素：地区发展状况、家庭收入情况、家庭规模、补贴情况与节约成本、能源类型的多样化程度

3、预测：各类型能源使用量及可再生能源使用总量趋势预测

分析适用性：

本方向与课程要求非常契合，特别是结合数据挖掘与场景分析。设想场景包括：

能源效率与经济发展：探讨地区发展水平和补贴政策如何影响可再生能源采纳。

家庭特征影响：分析家庭收入和规模对能源消费的约束。

复杂性较高，需要补充外部数据（如地区 GDP、补贴水平）。

优化建议：

数据补充：可以通过世界银行、联合国等统计数据补充外部特征信息（如 GDP、补贴）。

因果关系分析：除了简单相关性分析，尝试 Granger 检验或回归模型探讨因果性。

结果评价：将分析结果结合应用场景，如提出具体的补贴优化建议。

可用算法和工具：

算法：线性回归、多元回归、随机森林特征重要性。

工具：Scikit-learn（回归模型）、Statsmodels（因果分析）。

可视化：特征重要性条形图、因素变化趋势图。

代码步骤：

1.数据加载，数据处理（表格拆分，根据不同的分析目的），表格可视化

2.相关性分析：各因素之间的相关性分析，表格可视化（散点图、箱型图、折线图等），数据分布特性，变量之间的关系分析

3.算法与建模：回归分析、随机森林等，

参考资料：

[1] 曹军威, 袁仲达, 明阳阳, 等. 能源互联网大数据分析技术综述[J]. 南方电网技术, 2015, 9(11): 1-12.