

主题：粮食短缺的经验

贫穷的农村家庭主要依靠粮食作物的生产来满足其大部分粮食需求，并靠赚取很少的收入来购买更多的粮食和其他必需品。

您将调查以下问题，调查这些家庭中的农作物总产量

- 这项研究中家庭的特征是什么？家庭贫困水平有哪些指标？
- 这些家庭赖以生存的农场的特点是什么？它们产生什么？他们有多少土地？
- 家庭是否正在面临食物短缺？多久一次？
- 种植更多农作物的家庭难免出现粮食短缺的情况？
- 较富裕的家庭（拥有更多土地，更好的住房）是否更不容易受到粮食短缺的影响？
- 家庭组成对粮食短缺经验的影响是什么？有更多的孩子意味着更多的短缺吗？或更年长的人？还是工作年龄人口？
- 您的调查结果对改善非洲贫困农村家庭的粮食安全计划中的优先事项有何影响？

您仅应将R-studio用于对该报告进行的任何分析。

下载数据的电子表格文件。将数据导入R-studio，并随机选择1,000个家庭作为样本进行分析。

为了获得一致的输出，请使用数字“1553”提取随机样本。

例子：

- `set.seed(1553)`
- `data <- as.data.frame(sample_n(excel数据名称, 1000))`
- 即使您重新运行代码，这也应生成相同的随机样本。
- 查看数据集中的变量列表，选择一些描述住户及其农场特征的变量。使用百分比(percentages)，摘要度量(summary measures)，直方图(histograms)等准备简短摘要。
- 根据您的主题，确定相关的“outcome”变量。描述该变量的样本特征。
- 针对可能与上述问题相关的其他一些重要变量，对结果变量执行适当的技术
- 估计适当的多元线性回归模型(multiple linear regression models)以解决一些问题。
- 解释结果后，请对development programs(开发程序)进行简要说明。

目标要求:

1.动机良好的导言，使用相关资源将给定的问题与一些背景研究明确联系在一起。

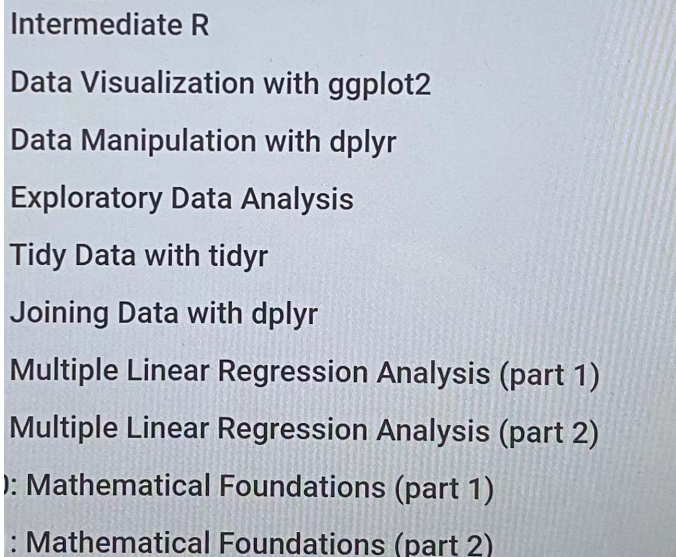
2.描述性分析

出色而适当的变量数字和图形摘要，对表格和图表有清晰的说明，表明您对变量和上下文有透彻的了解。

3.预测分析

回归分析的出色应用，并且可以很好地证明所选变量的合理性。通过提供有意义的见解清楚地解释分析的结果。

4.大致使用以下的知识点:



Intermediate R
Data Visualization with ggplot2
Data Manipulation with dplyr
Exploratory Data Analysis
Tidy Data with tidyr
Joining Data with dplyr
Multiple Linear Regression Analysis (part 1)
Multiple Linear Regression Analysis (part 2)
): Mathematical Foundations (part 1)
): Mathematical Foundations (part 2)

5.建议将相关策略与结果分析紧密关联。

6.结论总结清楚，对模型的改进有很好的建议（分析的局限性）。