

美赛问题

Problem A.旱灾对植物群落的打击

[2023 MCM Problem A](#)

背景.不同种类的植物对生存压力的反应不一样。例如，草原对干旱非常敏感。干旱发生的频率和严重程度各不相同。大量的观察表明，在植物群落适应持续几代时间的干旱时，不同物种的数量发挥着重要作用。只有一种植物的群落，后代对干旱的适应不如有四种植物的群落。这些观察数据产生了许多问题。例如，一个植物群落要从这种局部的生物多样性中获益，最少需要多少物种？这种现象如何随着物种数量的增加而扩大？这对植物群落的长期生存意味着什么？

要求.考虑到植物群落中干旱适应性与物种数量的关系，你的任务是探索和更好地理解这一现象。具体来说，你应该。

1. 建立一个数学模型来预测植物群落在经历各种不规则的天气周期时是如何随时间变化的。包括干旱时期，那时降水应该充足。该模型应考虑干旱循环期间不同物种之间的相互作用。
2. 探索你可以从你的模型中得出的关于植物群落和更大环境的长期相互作用的结论。考虑以下问题。
 1. 群落需要多少种不同的植物才能受益？随着物种数量的增加会发生什么？
 2. 群落中的物种类型如何影响你的研究结果？
 3. 在未来的天气周期中，更频繁和更广泛的干旱发生会产生什么影响？如果干旱不那么频繁，那么物种的数量是否会对整个群落产生同样的影响？
 4. 其他因素（如环境污染和栖息地减少）如何影响你的结论？
 5. 你的模型表明应该做什么来确保植物群落的长期生存能力，以及对更大的环境有什么影响？

Problem B.重构马赛马拉保护区

[2023 MCM Problem B](#)

背景.肯尼亚的野生动物保护区最初是为了保护野生动物和其他自然资源而建立的。肯尼亚议会于2013年通过了《野生动物保护和管理法》，以提供更公平的资源共享，并允许其他基于社区的管理工作。此后，肯尼亚增加了修正案，以弥补立法中的空白，为违反者提供更明确的行政、财政和惩罚。

要求.专注于一个大型的狩猎保护区。马赛马拉（Masai Mara），你的任务是确定备用的方法来管理当前公园边界内外的资源。具体来说，你应该。

1. 考虑并推荐当前保护区内不同地区的具体政策和管理策略，以保护野生动物和其他自然资源，同时平衡该地区居民的利益。这些政策和策略应该有助于减轻生活在保护区附近的人们受到的影响，并将动物和被吸引到保护区的人们之间的负面互动最小化。
2. 开发和描述一种方法，以确定哪些政策和管理策略将产生最好的结果。你的报告应该讨论如何根据你的方法对产生结果的好坏进行排序和比较。一定要包括用来预测动物和人之间相互作用的模型的描述和分析，以及由此对保护区内和周围地区的经济影响。
3. 根据你提出的计划，给出你的建议将导致的长期趋势的预测。分析并估计可能出现的长期结果的把握和影响。你还应该描述你的方法如何应用到其他野生动物管理领域。

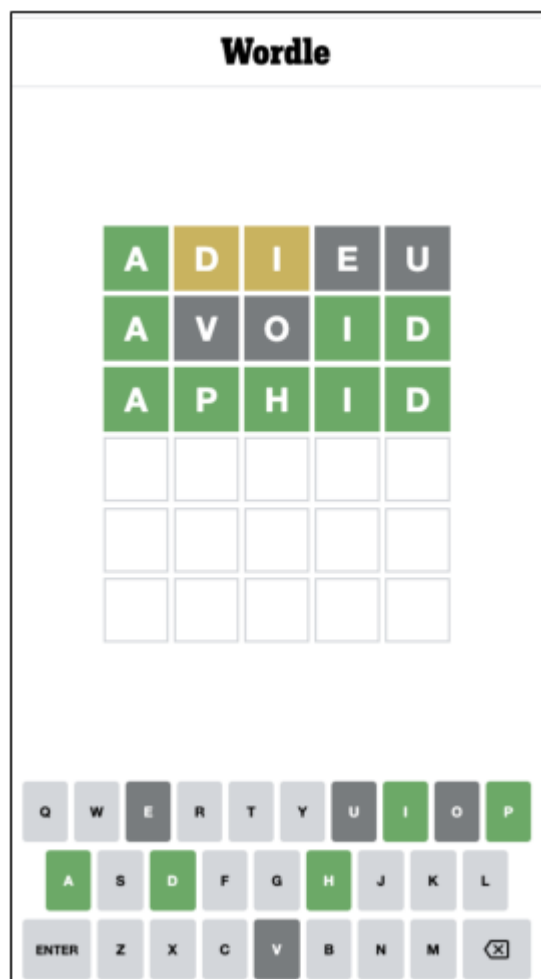
参考文献.

- [1] The Wildlife Conservation and Management Act, 2013. Republic of Kenya, Kenya Gazette Supplement No. 181 (Acts No. 47), 2013.
- [2] The Wildlife Conservation and Management Act (No. 47 of 2013). Legal Notice No. 155. Republic of Kenya, Kenya Gazette Supplement No 141 (Acts No. 47), 2020.

Problem C.Wordle

[2023 MCM Problem C](#)

背景.《纽约时报》目前每天都提供一款很受欢迎的字谜游戏“Wordle”。玩家尝试通过猜一个由 5 个字母组成的单词来解决谜题，每次猜都会得到反馈。对于这个版本，每个猜测必须是一个实际的英语单词。不允许猜中不被比赛识别的单词。Wordle 越来越受欢迎，游戏版本现在有 60 多种语言。《纽约时报》网站上关于 Wordle 的说明指出，在你提交你的文字后，瓷砖的颜色将会改变。黄色的 tile 表示该 tile 中的字母在单词中，但它在错误的位置。绿色瓦片表示该瓦片中的字母在单词中，并且位于正确的位置。灰色 tile 表示该 tile 中的字母完全不包含在单词中 (见附件 2)[2]。图 1 是一个示例解决方案，其中在三次尝试中找到了正确的结果。



玩家可以选择常规模式或“困难模式”。“《Wordle》的困难模式要求玩家一旦在一个单词中找到正确的字母 (方块是黄色或绿色)，就必须在随后的猜测中使用这些字母，从而增加了游戏的难度。图 1 中的例子是在困难模式下进行的。

许多 (但不是所有) 用户在 Twitter 上报告了他们的分数。对于这个问题, MCM 已经生成了 2022 年 1 月 7 日至 2022 年 12 月 31 日的每日结果文件 (见附件 1), 该文件包括日期、比赛编号、当天的单词、当天报告分数的人数、难度模式的玩家人数以及猜出单词的一次、两次、三次、四次、五次、六次或不能解决该谜题的比例 (用 X 表示), 例如图 2 中 7 月 20 日的单词, 2022 是“TRITE”, 结果是通过挖掘 Twitter 获得的。尽管图 2 中的百分比总和为 100%, 但在某些情况下, 由于四舍五入的关系, 这可能不是真的。

要求.《纽约时报》要求您对该文件中的结果进行分析, 并回答几个问题。

1. 报告的结果数量每天都在变化。开发一个模型来解释这种变化, 并使用您的模型来为 2023 年 3 月 1 日报告的结果数量创建一个预测区间。单词的特征是否会影响在困难模式中玩家各个得分的占比? 如果是, 怎么做? 如果没有, 为什么?
2. 对于未来日期的给定答案, 开发一个模型, 该模型允许您预测报告结果的分布。换句话说, 预测未来某一天 (1、2、3、4、5、6、X) 的相关百分比。你的模型和预测有哪些不确定性? 举一个你预测 2023 年 3 月 1 日单词 EERIE 的具体例子。你对你的模型的预测有多大信心?
3. 开发并总结一个按难度对解题词进行分类的模型。识别与每个分类相关联的给定单词的属性。使用你的模型, 单词 EERIE 有多难? 讨论你的分类模型的准确性。
4. 列出并描述该数据集的其他一些有趣的特性。

最后, 在给《纽约时报》的编辑的一到两页的信中总结你的结果。

附件. 给出的数据是唯一可用的信息。

分析

单词建模的特征可以有:

1. 单词长度: 单词长度可能会影响玩家猜测的难度。通常情况下, 单词长度越长, 玩家猜测的难度就越大, 因此可能会影响得分。
2. 单词中的元音和辅音数量: 元音和辅音的比例可能会影响单词的易读性和易发音性, 从而影响玩家的得分。
3. 单词的难度: 可以使用单词的普遍性或使用频率来衡量单词的难度。使用不太常见的单词可能会使得猜测的难度增加, 从而可能会降低玩家的得分。
4. 单词的语言学特征: 例如单词的词性、词根、前缀和后缀等语言学特征可能会影响单词的易懂度, 从而影响玩家的得分。
5. 单词的主题和背景: 单词的主题和背景可能会对玩家的得分产生影响。例如, 如果单词与某一主题有关, 那么熟悉该主题的玩家可能会更容易猜出单词, 从而获得更高的得分。

可以在这些地方找到单词词频之类的数据:

1. COCA (Corpus of Contemporary American English) : 是一个包含当代美国英语的语料库, 它是目前最大的英语语料库之一, 包括了约 20 亿个单词。COCA 提供了一个名为 Word and Phrase 的搜索引擎, 可以搜索单词及短语的词频信息, 并按照不同的子类别进行分类。提供了一些免费的公共资源, 包括一些已授权的、限制性的样本数据和语料库文件, 可以供科研用途使用。使用这些免费资源需要先进行注册, 并遵守使用条款和限制。
2. BNC (British National Corpus) : 是一个英国英语语料库, 包括了大约 4.2 亿个单词, 收录了来自不同领域和语言风格的文本。BNC 提供了一个名为 Sketch Engine 的工具, 可以搜索单词及短语的词频信

息，并按照不同的子类别进行分类。提供了一些免费的公共资源，包括一些样本数据和限制性的语料库文件，可以供科研用途使用。使用这些免费资源需要先进行注册，并遵守使用条款和限制。

3. Google Books Ngram Viewer: 是一个可以搜索和比较单词、短语在Google书籍中出现频率的工具。该工具基于Google书籍数字化计划中的数百万本书籍，可以提供跨越多个世纪的数据，并可以按照时间段、单词频率、语言等进行筛选。是一个免费的在线工具，可以搜索和比较单词、短语在Google书籍中出现频率的数据。这些数据可用于非商业科研用途，并且有一些限制和许可条款需要遵守。
4. WordNet: 是一个英语词汇数据库，包含了大量的单词及其词性、同义词、反义词、词根、搭配等信息。WordNet 提供了一个名为"word frequency lists"的工具，可以查询某个单词在语料库中的出现频率。是一个免费的英语词汇数据库，可以通过Princeton WordNet网站进行下载和使用。WordNet可用于个人和科研用途，但不允许将其用于商业目的。

可用的基于时间序列的预测模型：

1. Autoregressive Integrated Moving Average (ARIMA) 模型：ARIMA 模型是一种经典的时间序列预测模型，能够对非平稳时间序列进行建模和预测。它包含了自回归（AR）、差分（I）和移动平均（MA）三个部分，可以描述时间序列的长期趋势、季节性和随机性变化。
2. Exponential Smoothing (ES) 模型：ES模型是一种广泛使用的时间序列预测模型，通过加权平均预测未来的值。ES模型包括三种类型：简单指数平滑法、双指数平滑法和三指数平滑法。简单指数平滑法适用于没有趋势和季节性的时间序列，双指数平滑法适用于具有趋势但没有季节性的时间序列，三指数平滑法适用于具有趋势和季节性的时间序列。
3. Seasonal Autoregressive Integrated Moving Average (SARIMA) 模型：SARIMA模型是ARIMA模型的扩展，它考虑了时间序列的季节性变化。SARIMA模型包括季节性自回归（SAR）、差分（I）和季节性移动平均（SMA）三个部分，可以捕捉时间序列中的季节性模式。
4. Prophet 模型：Prophet模型是Facebook开发的一种时间序列预测模型，能够自适应地处理时间序列中的趋势、季节性和节假日等影响因素。Prophet模型使用可扩展的加法模型，能够处理多个季节性周期和非线性趋势变化。

Problem D.联合国的可持续性发展目标排序

背景.联合国制定了 17 项可持续发展目标 (SDGs)。实现这些目标最终将改善世界各地许多人的生活。这些目标不是相互独立的。因此，在某些目标中取得的积极成果通常会对其他目标产生影响 (积极的或消极的，有时两者兼有)。这种相互联系使实现所有目标成为一个流动的过程，供资限制和其他国家和国际优先事项可能优先于此。此外，技术进步、全球流行病、气候变化、区域战争和难民流动的影响对许多目标产生了严重影响。

要求.

1. 建立 17 个可持续发展目标之间的关系网络。
2. 利用各个可持续发展目标以及你们的网络结构，确定能够最有效地推动联合国工作向前发展的优先事项。你是如何评估每个优先级的有效性的？如果你开始优先考虑，那么在未来 10 年里，什么是合理的目标呢？
3. 如果实现了其中一个可持续发展目标 (例如，没有贫困或饥饿)，那么由此产生的网络结构将是什么？这一成就将如何影响你的团队的优先事项？是否有其他目标应该包括或建议联合国纳入？
4. 讨论技术进步、全球流行病、气候变化、区域战争、难民流动或其他国际危机对团队网络的影响，以及团队选择的优先事项。从网络的角度看，对联合国的进展有哪些重大影响？

5. 讨论你的网络方法如何帮助其他公司和组织设定他们目标的优先级。

联合国可持续发展目标

1. 无贫困目标
2. 零饥饿目标
3. 良好的健康和福祉目标
4. 优质教育目标
5. 性别平等目标
6. 清洁饮水和卫生目标
7. 可负担的清洁能源目标
8. 体面工作和经济增长目标
9. 工业、创新和基础设施目标
10. 减少不平等目标
11. 可持续城市和社区目标
12. 负责任的消费和生产目标
13. 气候行动目标
14. 保护水下生态系统目标
15. 保护陆地生态目标
16. 和平与正义
17. 为实现这一目标建立伙伴关系

其他

您的总页数不超过 25 页的 PDF 解决方案应包括.

- 一页摘要表。
- 目录。
- 您的完整解决方案。
- 参考文献列表。

注. MCM 竞赛有 25 页的限制。您提交的所有方面都计入 25 页的限制 (摘要页, 目录, 报告, 参考列表, 和任何附录)。你必须在你的报告中引用你的想法、图像和其他材料的来源。

Problem E. 光污染

光污染用于描述过度或不良使用人造光。我们称之为光污染的一些现象包括光侵入、过度照明和光杂波。在大城市, 太阳落山后, 这些现象最容易在天空中看到; 然而, 它们也可能发生在更偏远的地区。光污染会改变我们对夜空的想法, 对环境造成影响, 并影响我们的健康和安全。例如, 植物成熟可能会延迟或加速, 野生动物的迁徙模式也会受到影响。过度的人造光可能会扰乱我们的昼夜节律, 导致睡眠质量下降, 可能还会导致身心健康问题。人造灯光引起的眩光可能会导致一些机动车事故。社区官员或地方团体可以实施干预策略, 以减轻光污染的负面影响。然而, 人造光具有积极和消极的影响, 以不同的方式影响不同的位置。例如, 为了避免上述光污染的负面影响, 一些社区选择了光线较低的社区,

这反过来可能导致犯罪增加。光污染的影响可能取决于当地的发展水平、人口、生物多样性、地理和气候等因素。因此，评估任何干预策略的影响程度和潜在影响必须针对特定地点。

要求

中国商飞的照明控制任务（ICM）正在努力提高对光污染影响的认识，并制定干预策略以减轻这些影响。为了支持ICM的这项工作，您的任务是解决测量和减轻不同地点光污染的影响，同时考虑人类和非人类的问题。具体来说，您应该：

· 制定一个广泛适用的指标，以确定一个地点的光污染风险水平。

① 可用的系列判定指标：

1. 光度计测量指标：通过光度计测量，可以得到光照度、光线强度等指标，其中常用的指标包括：
 - 光照度：单位为勒克斯（lux），表示单位面积上落在该面积上的光通量。
 - 光线强度：单位为坎德拉（candela），表示光源在某一方向上单位立体角内的光通量。
 - 光密度：单位为坎德拉每平方米（candela/m²），表示在单位面积上单位立体角内的光通量。
2. 星空照片分析指标：通过分析星空照片，可以得到以下指标：
 - 星等（magnitude）：表示天体的亮度。较亮的天体星等较小。
 - 裸眼可见星等极限（naked eye limiting magnitude）：表示在该地点肉眼能看到的最暗的星等。通常与该地点的视场和背景光等因素有关。
 - Bortle光污染指数：由约翰·伯特尔（John Bortle）提出，将地面观测者的星空可见程度分为九个等级，从最暗（天空很黑，星星非常多）到最亮（天空灯光充斥，几乎看不到星星）。除此之外，还有其他一些可能的光污染指标，如噪声指标、光谱指标等，具体应根据实际测量情况和研究目的来选择。

② 可用于查找不同地区光污染风险程度的数据集网站

1. [Dark Sky Finder] (<https://darksitefinder.com/maps/world.html>)：该网站提供了世界各地的星空亮度地图，可以通过地图查找并比较不同地区的光污染水平，也可以通过搜索地名来查找该地区的星空亮度。
 2. [Light Pollution Map] (<https://www.lightpollutionmap.info>)：该网站提供了世界各地的光污染地图，以色块形式显示不同地区的光污染等级，可以通过地图查找并比较不同地区的光污染水平，也可以通过搜索地名来查找该地区的光污染等级。
 3. [Globe at Night] (<https://www.globeatnight.org>)：该网站是一个国际性的公众参与项目，旨在通过公众提交的数据来监测全球的光污染水平。在网站上，用户可以学习如何观测并报告星空亮度数据，也可以查看世界各地的数据分布和趋势。
- 在以下四种不同类型的地点应用度量并解释其结果：
 - 收保护的士地

- 农村地区
- 郊区地区
- 城镇地区
- 描述解决光污染的三种可能的干预策略。讨论实施每项战略的具体行动，以及这些行动对光污染总体影响的潜在影响。
- 选择您的两个地点，并使用您的度量标准来确定您的干预策略中哪一个对他们最有效。讨论所选择的干预策略如何影响所在地的风险水平。
- 最后，针对您确定的一个地点及其最有效的干预策略，制作一张1页传单，宣传该地点的策略。