# 2023百万立方米

**问题C:预测Wordle结果**

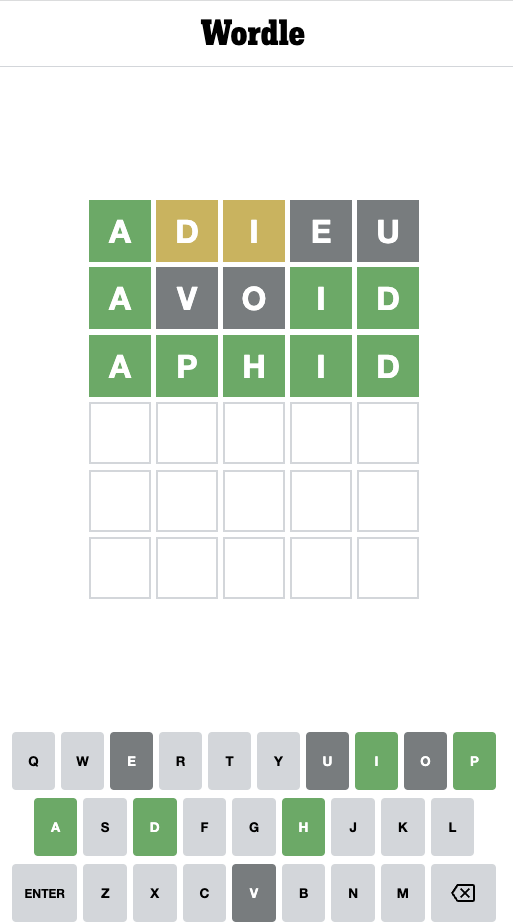


**图片来源:nytco.com[1]**

# 背景

Wordle是纽约时报每日提供的一个流行谜题。玩家尝试通过在六次或更少的尝试中猜测一个五个字母的单词来解决这个难题，每次猜测都会收到反馈。对于这个版本，每个猜测必须是一个实际的英语单词。不允许被竞赛识别为单词的猜测。Wordle越来越受欢迎，现在已经有超过60种语言的版本。

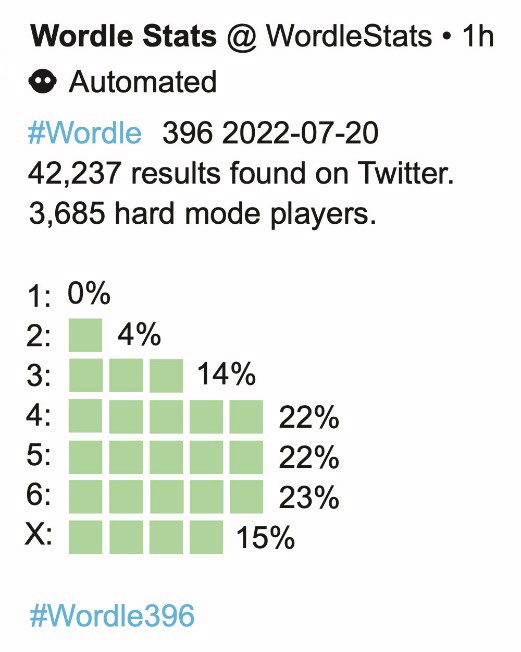
《纽约时报》网站Wordle的说明称，在你提交你的单词后，瓷砖的颜色会发生变化。黄色方块表示该方块中的字母在单词中，但位置错误。绿色方块表示该方块中的字母在单词中，并且位于正确的位置。灰色方块表示该方块中的字母根本不包含在单词中(参见[**Attachment 2**](#_bookmark1))[2].图1是一个示例解决方案，通过三次尝试找到了正确的结果。



# 图1:2022年7月21日的Wordle难题的示例解决方案[3]

玩家可以玩常规模式或“硬模式”Wordle的硬模式使游戏变得更加困难，因为它要求一旦玩家在单词中找到正确的字母(瓷砖是黄色或绿色)，这些字母必须在随后的猜测中使用。图1中的例子是在硬模式下播放的。

许多(但不是所有)用户在Twitter上报告他们的分数。对于这个问题，MCM已经生成了2022年1月7日到2022年12月31日的每日结果文件(参见[**Attachment 1**](#_bookmark0)).该文件包括日期、比赛编号、每日单词、当天报告分数的人数、处于困难模式的玩家人数，以及在一次尝试、两次尝试、三次尝试、四次尝试、五次尝试、六次尝试中猜出单词的百分比，或者无法解决谜题的百分比(用X表示)。例如，在图2中，单词2022年7月20日是“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，由于四舍五入的原因，这可能不是事实。



# 图2:2022年7月20日向Twitter[4]要求报告的结果分布

《纽约时报》要求你对该文件中的结果进行分析，以回答几个问题。

* 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，并使用您的模型为2023年3月1日报告的结果数量创建一个预测区间。这个单词的任何属性会影响硬模式下的分数百分比吗？如果有，如何实现？如果没有，为什么没有？
* 对于未来某个日期的给定未来解决方案单词，开发一个允许您预测报告结果的分布的模型。换句话说，预测未来日期的(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？请给出一个你对2023年3月1日这个词的预测的具体例子。你对你的模型的预测有多少信心？
* 开发并总结一个模型，根据难度对解决方案单词进行分类。识别与每个分类相关联的给定单词的属性。使用你的模型，怪诞这个词有多难？讨论你的分类模型的准确性。
* 列出并描述这个数据集的一些其他有趣的特征。

最后，给《纽约时报》的字谜编辑写一封一到两页的信，总结你的结果。

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。
* 一到两页的信。
* 参考列表。

注意:MCM竞赛有25页的限制。你提交的所有方面都算在25页的限制之内(摘要表、目录、报告、参考书目和任何附录)。你必须在你的报告中引用你的想法、图片和任何其他材料的来源。

# 附件

1. **数据文件。问题C数据Wordle.xlsx**

附加的数据文件包含您应该用于此目的的唯一数据

问题。该问题所需的所有信息都在问题陈述和数据文件中给出。你不需要访问纽约时报网站或推特网站。在这些网站上找不到其他信息。

# 数据文件条目描述

日期:以mm-dd-yyyy(月-日-年)格式表示的给定字谜的日期。

竞赛编号:沃尔多谜题的索引，从2022年1月7日的202开始。

单词:单词玩家试图猜测相关的日期和比赛号码。

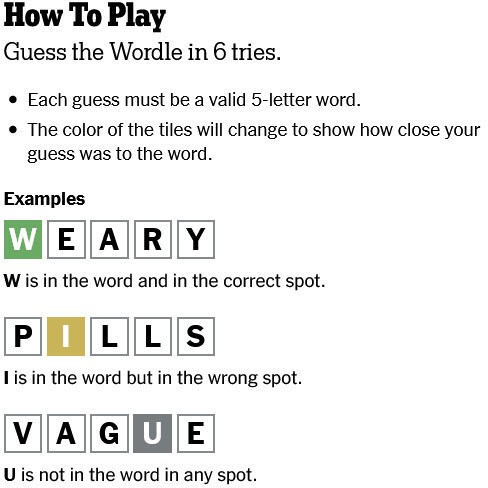
报告的结果数:当天在Twitter上记录的分数总数。

硬模式下的人数:当天在Twitter上记录的硬模式下的得分数。

* 1. 尝试:玩家在一次猜测中解决难题的百分比。
  2. 尝试次数:玩家在两次猜测中解出谜题的百分比。三次尝试:玩家在三次猜测中解决谜题的百分比。四次尝试:玩家在四次猜测中解决谜题的百分比。五次尝试:玩家在五次猜测中解决谜题的百分比。6次尝试:玩家在6次猜测中解决谜题的百分比。

7次或更多次尝试(X):玩家在6次或更少次尝试中不能解决谜题的百分比。注:由于四舍五入，百分比总和不一定是100%。

# 《纽约时报》网站发布了沃尔多的指示。[2]



**词汇表**

纽约时报:总部位于美国纽约州纽约市的日报，以印刷版和网络版发行。

Twitter:一个社交网站，允许用户发布不超过280个字符(从最初的140个字符增加)的简短帖子。

解答(单词难题):按照正确的顺序输入正确的字母，组成每日单词。

# 参考

注意:我们提供以下引用来支持问题陈述。我们从这些资源中提取了重要的想法。这些网站上没有解决这个MCM问题所需的附加信息。解决这个问题不需要访问纽约时报或Twitter网站。

1. 纽约时报网站上的Wordle标志。于2022年12月13日访问https://nytco-assets . nytimes . com/2022/08/cropped-Screen-Shot-2022-08-24-at-8 . 49 . 39-am . png。
2. "沃尔多-纽约时报"2022年纽约时报。于2022年12月13日通过https访问[://w](http://www.nytimes.com/games/wordle/index.html)ww。[nytimes.com/games/wordle/index.html.](http://www.nytimes.com/games/wordle/index.html)
3. "沃尔多-纽约时报"《纽约时报》，2022年7月21日。
4. "沃尔得统计数据"推特，2022年7月20日。