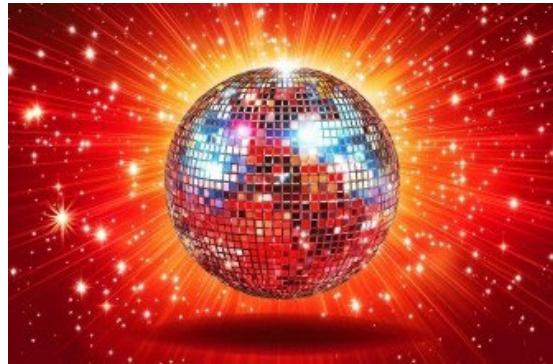


2026 MCM 问题 C:
数据与星星



与星共舞 (DWTS) 是根据英国节目“舞动奇迹”（原名“舞动起来”）改编的国际电视系列剧的美国版。该剧的多个版本已在阿尔巴尼亚、阿根廷、澳大利亚、中国、法国、印度和许多其他国家上映。本问题的焦点美国版已完结34季。

名人与专业舞者合作，然后每周表演舞蹈。专家评审团对每对舞者的舞蹈进行评分，粉丝们则（通过电话或在线）投票选出本周他们最喜欢的舞者。粉丝可以投票一次或多次，最多可达每周公布的限额。此外，粉丝可以投票选出他们希望保留的明星，但不能投票淘汰明星。评委和粉丝投票相结合，以确定本周淘汰哪对组合（综合得分最低）。三对（在某些赛季中更多）组合进入决赛，在决赛周，根据球迷和评委的综合评分，将他们从 1st 到 3rd（或 4th、5th）进行排名。

将粉丝投票和评委评分结合起来的可能方法有很多。在美国节目的前两季中，组合是根据排名进行的。对第二季的担忧（由于名人参赛者杰里·赖斯 (Jerry Rice) 尽管评判分数很低但仍进入决赛）导致修改为使用百分比而不是排名。附录中提供了这两种方法的示例。

第 27 季，另一场“争议”发生了，名人选手鲍比·博恩斯 (Bobby Bones) 尽管评委评分一直很低，但还是获胜。作为回应，从第 28 赛季开始，淘汰流程进行了轻微修改。根据评委评分和粉丝投票的综合结果确定排名垫底的两名参赛者，然后在现场表演期间，评委投票决定淘汰这两名选手中的哪一位。大约在同一季，制片人还像第一季和第二季一样，重新使用排名方法将评委评分与粉丝投票结合起来。这一变化发生的确切季节尚不清楚，但可以合理地假设是第 28 季。

评委的评分旨在反映哪些舞者在技术上更好，尽管舞蹈更好的因素存在一些主观性。粉丝投票可能更加主观，不仅受到舞蹈质量的影响，还受到名人的受欢迎程度和魅力的影响。在某种程度上，节目制作人实际上可能更喜欢意见和投票的冲突，因为此类事件会激发粉丝的兴趣和兴奋。

下面提供并描述了包含评委分数和参赛者信息的数据。您可以自行决定选择包含其他信息或其他数据，但您必须完整记录来源。使用数据来：

- 开发一个（或多个）数学模型，为每位参赛者在比赛的几周内产生估计的粉丝投票（这是未知的且严格保密的秘密）。
 - 您的模型是否正确估计了粉丝投票，从而得出与每周被淘汰的人一致的结果？提供一致性的度量。
 - 您产生的粉丝投票总数有多少确定性？每个参赛者/周的确定性是否始终相同？提供估算确定性的衡量标准。
- 将您的粉丝投票估计值与其余数据结合起来可以：
 - 比较和对比展览所使用的两种方法产生的结果
将各个赛季的评委和球迷投票（即排名和百分比）结合起来（即对每个赛季应用两种方法）。如果结果存在差异，是否一种方法似乎比另一种方法更有利粉丝投票？◦ 检查适用于存在“争议”的特定名人的两种投票方法，这意味着评委和粉丝之间存在差异。选择结合评委评分和粉丝投票的方法会给每位参赛者带来相同的结果吗？让评委每周选择淘汰最后两对组合中的哪对的额外方法会对结果产生怎样的影响？您可能会考虑一些示例（您可能还发现了其他示例）：第2季 - 杰瑞·赖斯 (Jerry Rice)，尽管评委评分为5周内最低，但仍获得亚军。第4季 - 比利·雷·赛勒斯 (Billy Ray Cyrus) 取得了5th的成绩，尽管评委在6周内得分垫底。第11季 - 布里斯托·佩林 (Bristol Palin) 12次以3rd获得最低评委分数。第27季 - 尽管评委评分一直较低，鲍比·博恩斯还是赢得了比赛
 - 根据您的分析，您建议在未来的季节中使用这两种方法中的哪一种？为什么？您是否建议加入评委从最后两对夫妇中选择的额外方法？
- 使用包括粉丝投票估计在内的数据来开发一个模型，该模型可以分析各种职业舞者的影响以及数据中可用的名人的特征（年龄、行业等）。这些事情对名人在比赛中的表现有多大影响？它们会以同样的方式影响评委评分和粉丝投票吗？
- 提出另一种系统，每周使用粉丝投票和评委评分，您认为该系统更“公平”（或其他方面“更好”，例如使节目对粉丝来说更令人兴奋）。提供支持，说明为何节目制作人应采用您的方法。
- 制作一份不超过25页的报告，其中包含您的发现，并包括一份一到两页的备忘录，总结您的结果，并为DWTS制作人提供关于如何将评委和粉丝投票结合起来的影响的建议，以及如何在未来赛季中这样做的建议。

您的 PDF 解决方案总页数不超过 25 页，应包括：

- 一页摘要表。
- 目录。
- 您的完整解决方案。
- 一到两页的备忘录。
- 参考文献列表。
- AI 使用报告（如果使用，则不计入 25 页限制。）

注意：完整的 MCM 提交没有具体要求的最小页面长度。您最多可以使用 25 页来完成您的所有解决方案工作以及您想要包含的任何附加信息（例如：绘图、图表、计算、表格）。部分解决方案被接受。我们允许谨慎使用 ChatGPT 等人工智能，尽管没有必要为此问题创建解决方案。如果您选择使用生成式 AI，则必须遵循 COMAP AI 使用政策。这将导致额外的 AI 使用报告，您必须将其添加到 PDF 解决方案文件的末尾，并且不计入解决方案的 25 页总页数限制。

数据文件: 2026_MCM_Problem_C_Data.csv – 第 1 – 34 季按周划分的参赛者信息、结果和评委分数。表 1 中提供了数据描述。

表 1: 2026_MCM_Problem_C_Data.csv 的数据描述

Variables	Explanation	Example
celebrity_name	Name of celebrity contestant (Star)	Jerry Rice, Mark Cuban, ...
ballroom_partner	Name of professional dancer partner	Cheryl Burke, Derek Hough, ...
celebrity_industry	Star profession category	Athlete, Model, ...
celebrity_homestate	Star home state (if from U.S.)	Ohio, Maine, ...
celebrity_homecountry/region	Star home country/region	United States, England, ...
celebrity_age_during_season	Age of the star in the season	32, 29, ...
season	Season of the show	1, 2, 3, ..., 32
results	Season results for the start	1st Place, Eliminated Week 2, ...
placement	Final place for the season (1 best)	1, 2, 3, ...
weekX_judgeY_score	Score from judge Y in week X	1, 2, 3, ...

数据注释：

1. 评委对每支舞蹈的评分为 1 (低) 至 10 (高)。

一个。在某些星期，报告的分数会包含小数（例如 8.5），因为每个名人都会表演不止一支舞蹈，并且每个名人的分数都会被平均。b. 在某些周内，会授予奖励积分（舞蹈比赛等）；它们均匀分布在评委/舞蹈乐谱中。c. 团队舞蹈分数是根据每个团队成员的分数来平均的。

2. 评委按照给舞蹈打分的顺序排列；因此，“Y 法官”可能每周、每季都不是同一位法官。

3. 不同季节的名人数量并不相同，节目播出的周数也不同。
4. 第 15 季是唯一由全明星演员回归的一季。
5. 偶尔有几周没有名人被淘汰，而其他几周则不止一名名人被淘汰。
6. 数据集中出现 N/A 值

一个。如果该周没有 4th 裁判（通常有 3 个），则 4th 裁判得分；b. 节目在一个季度中未播放的周数（例如，第一季持续了 6 周，因此记录了第 7 周到第 11 周的 N/A 值）。

7. 被淘汰的名人记为0分。例如，在第一季中，第一个被淘汰的名人是在第二周节目结束时的特里斯塔·萨特。因此，她在本赛季剩余时间（第 3 周到第 6 周）的得分为 0。

附录：投票方案示例

1. 按排名组合（用于第 1、2 和 28^a - 34 季）

在第一季和第二季中，评委和粉丝投票按排名合并。例如，在第一季第四周，还剩下四名参赛者。雷切尔·亨特（Rachel Hunter）被淘汰，这意味着她的综合排名最低。表 2 显示了评委的分数和排名，我们创建了一组可能产生正确结果的粉丝投票。粉丝投票有许多可能的值，也会给出相同的结果。您不应将这些值用作实际值，因为这只是一个示例。由于 Rachel 被评委评为 2nd，为了以最低的综合分数结束比赛，她的粉丝投票最低（4th 位），总排名为 6

◦

表 2：按排名合并评委和粉丝投票的示例（第 1 季第 4 周）

Contestant	Total Judges Score	Judges Score Rank	Fan Vote*	Fan Rank*	Sum of ranks
Rachel Hunter	25	2	1.1 million	4	6
Joey McIntyre	20	4	3.7 million	1	5
John O'Hurley	21	3	3.2 million	2	5
Kelly Monaco	26	1	2 million	3	4

* 粉丝投票/排名未知，选择假设值来产生正确的最终排名

2. 按百分比组合（用于第 3 季至第 27 季^a）

从第三季开始，分数是用百分比而不是排名来合并的。以第 5 季第 9 周为例。在那一周，Jennie Garth 被淘汰。同样，我们人为地创建了粉丝投票，产生了正确得出该结果的总百分比。评委百分比的计算方法是将参赛者的评委总分除以所有 4 名参赛者的评委总分之和。根据评委的百分比，Jennie 是 3rd。然而，加上我们分配给评委的 1000 万人工投票的百分比，她的百分比是 4th。

表 3：按百分比合并评委和粉丝投票的示例（第 5 季第 9 周）

Contestant	Total Judges Score	Judges Score Percent	<i>Fan Vote*</i>	<i>Fan Percent*</i>	Sum of Percents
Jennie Garth	29	$29/117 = 24.8\%$	<i>1.1 million</i>	$1.1/10 = 11\%$	35.8
Marie Osmond	28	$28/117 = 23.9\%$	<i>3.7 million</i>	$3.7/10 = 37\%$	60.9
Mel B	30	$30/117 = 25.6\%$	<i>3.2 million</i>	$3.2/10 = 32\%$	57.8
Helio Castroneves	30	$30/117 = 25.6\%$	<i>2 million</i>	$2/10 = 20\%$	45.6
Total	117		10 million		

* 粉丝投票未知，重视假设以产生正确的最终排名

a 恢复基于排名的方法的年份尚不清楚；第28季是一个合理的假设。