

2026 MCM C题：与星共舞的数据 (Data With The Stars)

背景

《与星共舞》 (Dancing with the Stars, DWTS) 是根据英国节目《严格来跳舞》 (Strictly Come Dancing, 原名《来跳舞》) 改编的国际电视特许经营节目的美国版。该节目的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出现。本题关注的美国版已经完成了 34 季。名人与职业舞者搭档，每周表演舞蹈。专家评审团为每对选手的舞蹈打分，粉丝则（通过电话或网络）为他们那一周最喜欢的组合投票。粉丝可以投一票或多票，上限由每周公布。此外，粉丝投票是为了留住他们希望留下的明星，而不能投票淘汰明星。评委分数和粉丝投票将结合起来，以决定该周淘汰哪一对组合（总分最低者）。三对（在某些季度更多）组合进入决赛，在决赛周，粉丝和评委的综合得分用于将他们从第 1 名排到第 3 名（或第 4、第 5 名）。

结合粉丝投票和评委分数的方法有很多种。在美国版的前两季中，这种结合是基于**排名 (Ranks)** 的。**第 2 季的争议 (由于名人选手 Jerry Rice 尽管评委分数很低却成为了决赛选手) 导致了规则的修改，即使用百分比 (Percentages) 代替排名。**附录中提供了这两种方法的示例。

在第 27 季，当名人选手 Bobby Bones 尽管评委分数一直很低却获得冠军时，发生了另一次“争议”。作为回应，从第 28 季开始，淘汰流程做了一点微调。通过结合评委分数和粉丝投票确定排名垫底的两名选手，然后在直播节目中由评委投票决定淘汰这两名中的哪一位。大约在这个赛季，制片人也恢复了使用排名法来结合评委分数和粉丝投票，就像第一季和第二季那样。这一变化发生的确切赛季尚不清楚，但合理假设是第 28 季。

评委的分数旨在反映哪些舞者在技术上更好，尽管对于什么才算更好的舞蹈存在一些主观性。粉丝的投票可能主观性更强，既受舞蹈质量的影响，也受名人的知名度和个人魅力的影响。节目制作人实际上可能在某种程度上更喜欢意见和投票的冲突，因为这种情况会提升粉丝的兴趣和兴奋度。

问题陈述

提供了评委分数和选手信息的数据，如下所述。您可以自行决定是否包含额外信息或其他数据，但必须完整记录来源。使用这些数据来：

建立一个数学模型（或多个模型）来估算每位参赛者在参赛周期间的粉丝投票数（这是未知且严密保守的秘密）。

您的模型是否正确估算了导致与每周淘汰结果一致的粉丝票数？请提供一致性的衡量标准。

您生成的粉丝投票总数有多大的确定性？这种确定性对于每位选手/每周都是一样的吗？请提供您估算的确定性衡量标准。

将您的粉丝投票估算值与其余数据结合使用，以：

比较和对比节目用来结合评委和粉丝投票的两种方法（即排名和百分比）在各个赛季产生的结果（即对每个赛季应用这两种方法）。如果结果存在差异，是否有一种方法看起来比另一种更偏向粉丝投票？

检查应用于发生“争议”（即评委和粉丝之间存在分歧）的特定名人的这两种投票方法。结合评委分数和粉丝投票的方法选择是否会导致这些选手的每一位都得到相同的结果？如果加入额外的机制——即由评委选择每周垫底的两对组合中淘汰哪一对——会如何影响结果？您可以考虑以下一些例子（可能还有您发现的其他例子）：

第 2 季——Jerry Rice，尽管 5 周评委评分最低，却获得亚军。

第 4 季——Billy Ray Cyrus，尽管 6 周评委评分倒数第一，却获得第 5 名。

第 11 季——Bristol Palin，12 次评委评分最低，却获得第 3 名。

第 27 季——Bobby Bones，尽管评委评分一直很低，却赢得了冠军。

基于您的分析，您建议在未来的赛季中使用这两种方法中的哪一种？为什么？您是否建议包括评委从垫底两名中进行选择的额外方法？

使用包括您的粉丝投票估算值在内的数据开发一个模型，分析各种职业舞者的影响以及数据中可用的名人特征（年龄、行业等）。

这些因素对一位名人在比赛中的表现有多大影响？它们对评委评分和粉丝投票的影响方式相同吗？

提出另一种您认为更“公平”（或在其他方面“更好”，例如让粉丝感到更刺激）的每周结合粉丝投票和评委分数的系统。请提供支持理由，说明为何节目制作人应采纳您的方法。

撰写一份不超过 25 页的报告展示您的发现，并包含一份一到两页的备忘录，总结您的结果，就评委和

粉丝投票结合方式的影响向《与星共舞》制作人提供建议，并对未来赛季如何操作提出推荐意见。

提交要求

您的 PDF 解决方案总页数不超过 25 页，应包括：

一页摘要页。

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（如果使用了 AI，不计入 25 页限制）。

数据文件：2026_MCM_Problem_C_Data.csv - 第 1 至 34 季的选手信息、结果和每周评委评分。

表 1：2026_MCM_Problem_C_Data.csv 数据说明

变量 解释 示例

celebrity_name 名人选手（明星）姓名 Jerry Rice, Mark Cuban, ...

ballroom_partner 职业舞伴姓名 Cheryl Burke, Derek Hough, ...

celebrity_industry 明星职业类别 运动员 (Athlete), 模特 (Model), ...

celebrity_homestate 明星家乡州（如果是美国人）俄亥俄州 (Ohio), 缅因州 (Maine), ...

celebrity_homecountry/region 明星家乡国家/地区 美国, 英国, ...

celebrity_age_during_season 赛季期间明星的年龄 32, 29, ...

season 节目季数 1, 2, 3, ..., 32

results 该明星的赛季结果 第1名, 第2周被淘汰, ...

placement 赛季最终排名（1 为最好） 1, 2, 3, ...

weekX_judgeY_score Y 评委在第 X 周的评分 1, 2, 3, ...

数据说明：

每支舞蹈的评委评分从 1（低）到 10（高）。

a. 在某些周，报告的分数包含小数（例如 8.5），因为每位名人表演了不止一支舞蹈，分数是平均后的结果。

b. 在某些周，给予了加分（斗舞等）；它们均匀分布在评委/舞蹈分数中。

c. 团队舞蹈分数是与每个团队成员的个人分数平均计算的。

评委按打分顺序列出；因此，“评委 Y”在不同周或不同赛季可能不是同一位评委。

不同赛季的名人数量不同，节目播出的周数也不相同。

第 15 季是唯一一个由回归名人组成的全明星阵容的赛季。

偶尔会有几周没有名人被淘汰，而其他几周则有一名以上被淘汰。

数据集中出现 N/A 值的情况：

a. 第 4 位评委的分数，如果该周没有第 4 位评委（通常有 3 位）；以及

b. 节目在某季未播出的周数（例如，第 1 季持续 6 周，因此第 7 至 11 周记录为 N/A 值）。

被淘汰的名人记录为 0 分。例如，在第 1 季，第一位被淘汰的名人是 Trista Sutter，她在第 2 周节目结束时被淘汰。因此她在本季剩余时间（第 3 周到第 6 周）的分数为 0。

附录：投票方案示例

1. 按排名结合 (COMBINED BY RANK, 用于第 1、2 季和 28 - 34 季)

在第 1 和第 2 季，评委和粉丝投票是按排名结合的。例如，在第 1 季第 4 周，还有四名参赛者。

Rachel Hunter 被淘汰，意味着她获得了最低的综合排名。在表 2 中显示了评委分数和排名，我们创建了一组可能产生该正确结果的粉丝投票。还有许多可能的粉丝投票值也能得出相同的结果。您不应将这些作为实际值使用，因为这只是一个示例。由于 Rachel 被评委排在第 2 位，为了最终获得最低的综合得分，她的粉丝投票最低（第 4 名），总排名为 6。

表 2：按排名结合评委和粉丝投票的示例（第 1 季，第 4 周）

参赛者 评委总分 评委分数排名 粉丝票数* 粉丝排名* 排名总和

Rachel Hunter 25 2 110 万 4 6

Joey McIntyre 20 4 370 万 1 5

John O'Hurley 21 3 320 万 2 5

Kelly Monaco 26 1 200 万 3 4

- 粉丝票数/排名未知，为产生正确最终排名而选择的假设值

2. 按百分比结合 (COMBINED BY PERCENT, 用于第 3 季至第 27 季)

从第 3 季开始，分数改为使用百分比结合而不是排名。以第 5 季第 9 周为例。在那一周，Jennie Garth 被淘汰。同样，我们人为创建了粉丝投票，产生总百分比以正确导致该结果。评委百分比是通过将该参赛者的评委总分除以所有 4 名参赛者的评委总分之和计算得出的。根据评委百分比，Jennie 排名第 3。然而，加上我们要分配给评委百分比的 1000 万人创建的粉丝投票的百分比后，她是第 4 名。

表 3：按百分比结合评委和粉丝投票的示例（第 5 季，第 9 周）

参赛者 评委总分 评委分数百分比 粉丝票数* 粉丝百分比* 百分比总和

Jennie Garth 29 $29/117 = 24.8\%$ 110 万 $1.1/10 = 11\%$ 35.8

Marie Osmond 28 $28/117 = 23.9\%$ 370 万 $3.7/10 = 37\%$ 60.9

Mel B 30 $30/117 = 25.6\%$ 320 万 $3.2/10 = 32\%$ 57.8

Helio Castroneves 30 $30/117 = 25.6\%$ 200 万 $2/10 = 20\%$ 45.6

总计 117 1000 万

- 粉丝票数未知，为产生正确最终排名而选择的假设值

注：恢复排名法的确切年份尚不确定；第 28 季是一个合理的假设。