

2026 MCM 问题 C：星光数据 (Data With The Stars)

《与星共舞》(Dancing with the Stars, 简称 DWTS) 是一个国际电视节目的美国版本, 该节目源自英国节目《舞动奇迹》(Strictly Come Dancing)。该节目的不同版本已出现在阿尔巴尼亚、阿根廷、澳大利亚、中国、法国、印度等许多国家。作为本题焦点的美国版目前已完成 34 个赛季。

名人 (Celebrities) 与专业舞者组成搭档, 每周进行舞蹈表演。专家评审团为每对组合的舞蹈评分, 粉丝则 (通过电话或网络) 为他们当周喜爱的组合投票。粉丝可以投票一次或多次, 上限由每周公布的规则决定。此外, 粉丝投票是支持他们希望留下的明星, 不能投票淘汰某位明星。评委分数和粉丝投票将结合起来, 以决定当周淘汰哪对组合 (综合得分最低者)。在决赛周, 会有三对 (某些赛季更多) 组合进入决赛, 届时将根据评委和粉丝的综合得分对他们进行第 1 到第 3 (或第 4、第 5) 名的排名。

结合粉丝投票和评委分数的方法有很多种。在美国版的前两个赛季中, 结合方式是基于**排名 (Ranks) 的。由于第 2 季中名人选手 Jerry Rice 虽然评委分数极低但仍进入了决赛, 引发了人们的关注, 随后规则修改为使用百分比 (Percentages) **代替排名。附录中提供了这两种方法的示例。

在第 27 季中, 另一场“争议”发生了: 名人选手 Bobby Bones 虽然评委分数持续偏低, 但最终获得了冠军。作为回应, 从第 28 季开始, 淘汰流程进行了微调。通过评委分数和粉丝投票的综合结果确定最后两名 (Bottom Two), 然后在直播节目中由评委投票决定淘汰其中的哪一位。大约在同一时期, 制作方也将结合方式恢复为前两个赛季使用的排名法。具体恢复的赛季尚不明确, 但可以合理假设是从第 28 季开始的。

评委分数旨在反映选手的技术水平, 尽管舞蹈的优劣评价具有一定的主观性。粉丝投票可能更加主观, 受舞蹈质量影响, 也受名人的知名度和魅力的影响。在某种程度上, 节目制作方可能更倾向于观点和投票之间的冲突, 因为这种情况能提高粉丝的关注度和兴奋度。

任务要求

提供的数据包含评委分数和参赛者信息。您可以根据需要选择包含额外信息或其他数据, 但必须完整记录来源。请使用数据完成以下任务:

1. **开发一个或多个数学模型, 估算每位参赛者在参赛周内的粉丝投票数** (粉丝投票数据是未知的, 且属于商业机密)。
 - 您的模型估算的粉丝投票结果是否能与每周实际的淘汰结果保持一致? 请提供一致性的衡量标准。
 - 您生成的粉丝投票总量的确定性 (Certainty) 有多大? 这种确定性在不同选手/周次之间是否相同? 请为估算结果提供确定性的衡量指标。
2. **利用您的粉丝投票估算结果和其余数据完成以下分析:**
 - **对比两种方法:** 比较节目中使用的两种结合评委分数和粉丝投票的方法 (即排名法和百分比法) 在不同赛季的表现 (即将两种方法应用于每个赛季)。如果结果存在差异, 其中一种方法是否比另一种更偏向粉丝投票?

- **争议案例研究：** 针对特定的“争议性”名人（即评委与粉丝之间存在显著意见分歧的案例），研究这两种投票方法。结合方法的选择是否会导致这些选手的最终结果发生改变？如果加入“由评委从最后两名中选择一人淘汰”的规则，结果会受到什么影响？您可以考虑以下示例（或您发现的其他案例）：
 - 第 2 季：Jerry Rice，尽管有 5 周评委分数垫底，仍获得亚军。
 - 第 4 季：Billy Ray Cyrus，尽管有 6 周评委分数最后一名，仍获得第 5 名。
 - 第 11 季：Bristol Palin，曾 12 次获得最低评委分，仍获得第 3 名。
 - 第 27 季：Bobby Bones，尽管评委分数持续低迷，最终夺冠。
 - **建议：** 根据您的分析，您建议未来赛季使用哪种方法？为什么？您是否建议保留“由评委从最后两名中选择”的做法？
3. **利用包含粉丝投票估算值在内的数据，开发一个模型分析专业舞者以及名人的特征（年龄、行业等）对比赛的影响。**
- 这些因素对名人在比赛中的表现影响有多大？
 - 它们对评委分数和粉丝投票的影响方式是否一致？
4. **提议一套新的系统：** 您认为更“公平”（或在其他方面更好，如更能激发粉丝热情）的结合每周粉丝投票和评委分数的新系统。提供支持理由，说明制作方为何应采用您的方案。
5. **编写一份报告：** 总页数不超过 25 页，包含您的研究发现，并附带一份 1 到 2 页的备忘录，总结您的结果，并就评委/粉丝投票结合方式的影响及未来建议为 DWTS 制作方提供咨询。

数据说明

文件 2026_MCM_Problem_C_Data.csv 包含第 1-34 季的参赛者信息、结果和每周评委分数。

变量	说明	示例
celebrity name	名人选手（明星）姓名	Jerry Rice, Mark Cuban...
ballroom partner	专业舞者搭档姓名	Cheryl Burke, Derek Hough...
celebrity industry	明星职业类别	Athlete (运动员), Model (模特)...
celebrity homestate	明星家乡（若是美国人）	Ohio, Maine...
celebrity homecountry/region	明星家乡国家/地区	United States, England...
celebrity age during season	明星在该赛季的年龄	32, 29...
season	节目赛季	1, 2, 3...
results	该明星的赛季结果	1st Place (第 1 名), Eliminated Week 2 (第 2 周淘汰)...

placement	赛季最终名次	1, 2, 3...
weekX_judgeY_score	第 X 周第 Y 位评委的分数	1, 2, 3...

数据说明：

- 1. 评委打分为 1（低）到 10（高）。
 - 某些周的分数包含小数（如 8.5），因为名人表演了多场舞蹈并取了平均分。
 - 某些周有奖金分（如对抗赛），这些分平均摊在评委/舞蹈分数中。
 - 团队舞蹈分数与各成员的个人分数取平均。
- 2. 评委按评分顺序排列，因此“Judge Y”在不同周次或赛季中未必是同一人。
- 3. 名人数量和节目播出周数在不同赛季间有所不同。
- 4. 第 15 季是唯一的“全明星”赛季，由往届名人回归参赛。
- 5. 偶尔有某周无人淘汰，或一周淘汰多人。
- 6. N/A 值出现的情况：第 4 位评委不存在（通常只有 3 位）；或者该赛季没有该周（例如第 1 季只有 6 周，则 7-11 周为 N/A）。
- 7. 已淘汰选手的得分为 0。

附录：投票方案示例

1. 排名结合法 (Rank) - 用于第 1, 2 季及第 28-34 季

评委分数和粉丝投票按排名结合。 示例（第 1 季第 4 周）： 有 4 对选手。 Rachel Hunter 被淘汰，意味着她的综合排名最低。

参赛者	评委总分	评委排名	粉丝投票*	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

*粉丝投票和排名为假设值，用于演示结果。

2. 百分比结合法 (Percent) - 用于第 3 至 27 季

示例（第 5 季第 9 周）： Jennie Garth 被淘汰。 评委百分比 = 该选手评委总分 / 所有人评委总分之和。

参赛者	评委总分	评委百分比	粉丝投票*	粉丝百分比*	百分比总和
Jennie Garth	29	24.8%	110 万	11%	35.8
Marie Osmond	28	23.9%	370 万	37%	60.9
Mel B	30	25.6%	320 万	32%	57.8
Helio Castroneves	30	25.6%	200 万	20%	45.6
总计	117	100%	1000 万	100%	

*粉丝投票为假设值。

©2026 by COMAP | www.comap.org