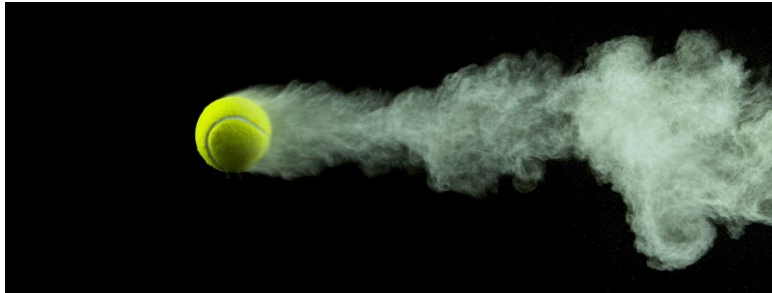


**2024 MCM**  
**Problem C: Momentum in Tennis**



In the 2023 Wimbledon Gentlemen's final, 20-year-old Spanish rising star Carlos Alcaraz defeated 36-year-old Novak Djokovic. The loss was Djokovic's first at Wimbledon since 2013 and ended a remarkable run for one of the all-time great players in **Grand Slams**.

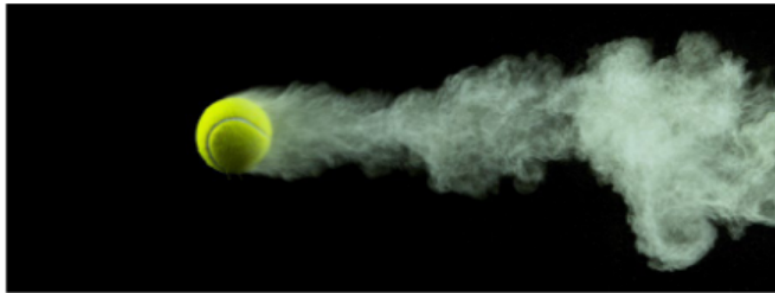
The match itself was a remarkable battle.<sup>[1]</sup> Djokovic seemed destined to win easily as he dominated the first set 6 – 1 (winning 6 of 7 games). The second set, however, was tense and finally won by Alcaraz in a tie-breaker 7 – 6. The third set was the reverse of the first, Alcaraz winning handily 6 – 1. The young Spaniard seemed in total control as the fourth set started, but somehow the match again changed course with Djokovic taking complete control to win the set 6 – 3. The fifth and final set started with Djokovic carrying the edge from the fourth set, but again a change of direction occurred and Alcaraz gained control and the victory 6 – 4. The data for this match is in the provided data set, “match\_id” of “2023-wimbledon-1701”. You can see all the points for the first set when Djokovic had the edge using the “set\_no” column equal to 1. The incredible swings, sometimes for many points or even games, that occurred in the player who seemed to have the advantage are often attributed to “momentum.”

One dictionary definition of momentum is “strength or force gained by motion or by a series of events.”<sup>[2]</sup> In sports, a team or player may feel they have the momentum, or “strength/force” during a match/game, but it is difficult to measure such a phenomenon. Further, it is not readily apparent how various events during the match act to create or change momentum if it exists.

Data is provided for every point from all Wimbledon 2023 men's matches after the first 2 rounds. You may choose to include additional player information or other data at your discretion, but you must completely document the sources. Use the data to:

- Develop a model that captures the flow of play as points occur and apply it to one or more of the matches. Your model should identify which player is performing better at a given time in the match, as well as how much better they are performing. Provide a visualization based on your model to depict the match flow. *Note: in tennis, the player serving has a much higher probability of winning the point/game. You may wish to factor this into your model in some way.*
- A tennis coach is skeptical that “momentum” plays any role in the match. Instead, he postulates that swings in play and runs of success by one player are random. Use your model/metric to assess this claim.

## 2024 MCM 问题 C: 网球的势头



在2023年温网绅士赛决赛中，20岁的西班牙新星卡洛斯·阿尔卡拉斯击败了36岁的诺瓦克·德约科维奇。这场失利是德约科维奇自2013年以来首次在温布尔登输球，并结束了这位大满贯历史上最伟大的球员之一的非凡表现。

这场比赛本身就是一场非凡的战斗。德约科维奇似乎注定要轻松获胜，因为他以 6 比 1 统治了第一盘（赢得了 7 场比赛中的 6 场）。然而，第二盘比赛很紧张，最终阿尔卡雷斯在决胜局中以7-6获胜。第三盘与第一盘相反，阿尔卡拉斯以6-1轻松获胜。在第四盘开始时，这位年轻的西班牙人似乎完全控制了比赛，但不知何故，比赛再次改变了方向，德约科维奇完全控制了比赛，以6-3赢得了这一盘。第五盘也是最后一盘开始时，德约科维奇从第四盘开始占据优势，但再次改变了方向，阿尔卡拉斯获得了控制权，并以6-4获胜。本场比赛的数据在提供的数据集“2023-wimbledon-1701”的“match\_id”中。当德约科维奇使用等于 1 的“set\_no”列获得优势时，您可以看到第一盘的所有分数。在看似有优势的球员身上发生的令人难以置信的波动，有时是许多点甚至比赛，通常归因于“动量”。

动量的一个字典定义是“通过运动或一系列事件获得的力量或力”。在体育运动中，一支球队或球员可能会觉得他们在比赛/比赛中有动力或“力量/力量”，但很难衡量这种现象。此外，比赛中的各种事件如何创造或改变势头（如果存在）并不明显。

2023年温布尔登网球公开赛前两轮比赛后，每场积分均提供数据。您可以自行决定是否包含其他玩家信息或其他数据，但您必须完整地记录来源。使用数据可以：

- 开发一个模型，在得分时捕捉比赛流程，并将其应用于一个或多个比赛。您的模型应该确定哪些球员在比赛中的给定时间表现更好，以及他们的表现有多好。提供基于模型的可视化效果，以描述匹配流程。注意：在网球比赛中，发球的球员赢得积分/比赛的概率要高得多。您可能希望以某种方式将其考虑到您的模型中。
- 一位网球教练对“动力”在比赛中发挥任何作用持怀疑态度。相反，他假设一个球员在比赛中的波动和成功是随机的。使用您的模型/指标来评估此声明。

- Coaches would love to know if there are indicators that can help determine when the flow of play is about to change from favoring one player to the other.
  - Using the data provided for at least one match, develop a model that predicts these swings in the match. What factors seem most related (if any)?
  - Given the differential in past match “momentum” swings how do you advise a player going into a new match against a different player?
- Test the model you developed on one or more of the other matches. How well do you predict the swings in the match? If the model performs poorly at times, can you identify any factors that might need to be included in future models? How generalizable is your model to other matches (such as Women’s matches), tournaments, court surfaces, and other sports such as table tennis.
- Produce a report of no more than 25 pages with your findings and include a one- to two-page memo summarizing your results with advice for coaches on the role of “momentum”, and how to prepare players to respond to events that impact the flow of play during a tennis match.

Your PDF solution of no more than 25 total pages should include:

- One-page Summary Sheet.
- Table of Contents.
- Your complete solution.
- One- to two-page memo.
- References list.
- [AI Use Report](#) (If used does not count toward the 25-page limit.)

**Note:** There is no specific required minimum page length for a complete MCM submission. You may use up to 25 total pages for all your solution work and any additional information you want to include (for example: drawings, diagrams, calculations, tables). Partial solutions are accepted. We permit the careful use of AI such as ChatGPT, although it is not necessary to create a solution to this problem. If you choose to utilize a generative AI, you must follow the [COMAP AI use policy](#). This will result in an additional AI use report that you must add to the end of your PDF solution file and does not count toward the 25 total page limit for your solution.

### Files provided:

- [Wimbledon\\_featured\\_matches.csv](#) – data set of Wimbledon 2023 Gentlemen’s singles matches after second round.
- [data\\_dictionary.csv](#) – description of the data set.
- [data\\_examples](#) – examples to help understand the provided data.

### Glossary

**Grand Slam:** The Grand Slam in tennis is the achievement of winning all four major championships in one discipline in a calendar year. The four Grand Slam tournaments are the Australian Open, the French Open, Wimbledon, and the US Open, with each played over two weeks.

- 教练们很想知道是否有指标可以帮助确定比赛的流程何时会从偏袒一名球员转变为偏袒另一名球员。
  - 使用为至少一场比赛提供的数据，开发一个模型来预测比赛中的这些波动。哪些因素似乎最相关（如果有的话）？
  - 考虑到过去比赛的“动量”波动差异，您如何建议球员与另一名球员进行新比赛？
- 在一个或多个其他匹配项上测试您开发的模型。你对比赛中的波动预测有多好？如果模型有时表现不佳，您能否确定未来模型中可能需要包含的任何因素？模型对其他比赛（如女子比赛）、锦标赛、球场表面和其他运动（如乒乓球）的泛化程度。
- 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，并为教练提供关于“动力”作用的建议，以及如何让球员做好准备以应对影响比赛流程的事件在网球比赛中。

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。
- 一到两页的备忘录。
- 引用列表。
- AI 使用报告（如果使用不计入 25 页限制）。

注意：完整的 MCM 提交没有具体的最小页面长度要求。您最多可以使用 25 页来完成所有解决方案工作以及要包含的任何其他信息（例如：图纸、图表、计算、表格）。接受部分解决方案。我们允许谨慎使用 ChatGPT 等 AI，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 使用策略。这将导致一个额外的 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。

提供的文件：

- Wimbledon\_featured\_matches.csv - 2023 年温布尔登网球公开赛第二轮后男子单打比赛的数据集。
- data\_dictionary.csv - 数据集的描述。
- data\_examples - 帮助理解所提供数据的示例。

词汇表

大满贯：网球大满贯是在一年内赢得一个项目的四个大满贯冠军的成就。四项大满贯赛事是澳大利亚网球公开赛、法国网球公开赛、温布尔登网球公开赛和美国网球公开赛，每场比赛都在两周内进行。

## Glossary of key terms/concepts:

- **Scoring:**<sup>[3]</sup>
  - o **Match:** best of five sets (for Gentlemen's matches at Wimbledon)
  - o **Set:** collection of games; 6 games win a set, but players must win by two games until the set is tied 6 – 6 when a tie-breaker is played (see below)
  - o **Game:** collection of points; a player wins when reaching 4 points but must win by two. See “scoring a game” below.
- **Scoring a game:**<sup>[3]</sup>
  - o 0 points = Love
  - o 1 point = 15
  - o 2 points = 30
  - o 3 points = 40
  - o Tied score = All (e.g., “30 all”)
  - o 40 – 40 = Deuce (players have won the same number of points, at least 3 points each)
  - o Server wins a deuce point = Ad-in (or “advantage in”)
  - o Receiver wins a deuce point = Ad-out
- **Serve:** players alternate games as the “server” (the player who hits the initial shot of a point) and “returner.” In professional tennis, the server tends to have a big advantage. A player is given two serves to put the ball in play (into the “service box”) on each point. Failure to hit a serve in play in two attempts is a “double fault” and the returning player is awarded the point.
  - o **Breaking serve** – when the returning player wins a game.
  - o **Break point** – a point in which if the returner wins, they would win the game.
  - o **Holding serve** – when the serving player wins the game.
- **Tie-breakers:** each set ends when a player has won 6 games, as long as they are ahead by at least two games (i.e., 6 – 4). If not, play continues until a tie at 6 – 6 is reached. At this point a tie-breaker is played. At Wimbledon tie-breakers are first to 7 points (must win by 2 points) except in the 5<sup>th</sup> set of a match when it is first to 10 points (must win by 2 points).
- **Rest breaks/sides of court:** players switch sides of the court after game 1 and then after every two games. 90 second rest breaks are allowed starting at the 3<sup>rd</sup> game at every change of sides. During tie-breakers, players change sides every six points. Players also rest for at least 2 minutes after the conclusion of each set. Medical timeouts and one bathroom break are permitted.

## 关键术语/概念词汇表:

- 得分:
  - o 比赛: 五局两胜制 (温布尔登绅士赛)
  - o 套装: 游戏集合; 6 场比赛赢得一盘, 但玩家必须赢两局, 直到决胜局时盘数 6 - 6 平局 (见下文)
  - o 游戏: 收集积分; 玩家在达到 4 分时获胜, 但必须以 2 分获胜。请参阅下面的“为比赛得分”。
  
- 为比赛打分:
  - o 0 分 = 爱
  - o 1 分 = 15
  - o 2 分 = 30
  - o 3 分 = 40
  - o 平局分数 = 全部 (例如, “全部 30 分”)
  - o 40 - 40 = 平分 (玩家赢得相同数量的积分, 每人至少 3 分)
  - o 服务器赢得平分点 = Ad-in (或 “优势”)
  - o 接收者赢得平分点 = Ad-out
  
- 发球: 玩家作为 “服务器” (击中点的初始击球的玩家) 和 “返回者” 交替进行游戏。在职业网球中, 服务器往往具有很大的优势。球员有两次发球机会, 在每一分上将球投入比赛 (进入 “发球框”)。两次尝试未能击中发球是 “双重错误”, 返回的球员将获得积分。
  - o 破发 - 当接发球的球员赢得比赛时。
  - o 破发点 - 如果返回者获胜, 他们将赢得比赛的点。
  - o 保住发球 - 当发球球员赢得比赛时。
  
- 决胜局: 每盘在玩家赢得 6 局时结束, 只要他们至少领先两局 (即 6-4)。如果没有, 比赛继续进行, 直到达到 6 - 6 的平局。此时将进行决胜局。在温布尔登, 决胜局首先获得 7 分 (必须以 2 分获胜), 但在第 5 盘比赛中, 当它首先获得 10 分 (必须以 2 分获胜) 时除外。
  
- 休息时间/球场两侧: 球员在第 1 场比赛后更换球场两侧, 然后每两场比赛后更换一次球场两侧。从 90 场比赛开始, 每次换边都允许 3 秒的休息时间。在决胜局中, 球员每六分换边。每盘结束后, 球员也要休息至少 2 分钟。允许医疗暂停和一次洗手间休息。

## References:

- [1] Braidwood, J. (2023), Novak Djokovic has created a unique rival – is Wimbledon defeat the beginning of the end, The Independent,  
<https://www.independent.co.uk/sport/tennis/novak-djokovic-wimbledon-final-carlos-alcaraz-b2376600.html>.
- [2] <https://www.merriam-webster.com/dictionary/momentum>
- [3] Rivera, J. (2023), Tennis scoring, explained: A guide to understanding the rules terms & point system at Wimbledon, The Sporting News,  
<https://www.sportingnews.com/us/tennis/news/tennis-scoring-explained-rules-system-points-terms/7uzp2evdhbd11obdd59p3p1cx>.

引用:

[1] 布雷德伍德, J. (2023), 诺瓦克·德约科维奇创造了一个独特的对手——温布尔登的失败是终结的开始, 《独立报》, <https://www.independent.co.uk/sport/tennis/novak-djokovic-wimbledon-final-carlos-alcarazb2376600.html>。

[2] <https://www.merriam-webster.com/dictionary/momentum>

[3] Rivera, J. (2023), 网球得分, 解释: 了解温布尔登规则条款和积分系统的指南, 《体育新闻》, <https://www.sportingnews.com/us/tennis/news/tennis-scoring-explained-rules-system-pointsterms/7uzp2evdhbd11obdd59p3plcx>。



## Examples to Help Understand the Data Set

### Example 1: row 5

Column(s)	Value(s)	Description
<i>match_id</i>	"2023-wimbledon-1301"	The 3 in "1301" indicates a round 3 match and the "01" indicates the first match listed from that round.
<i>elapsed_time</i>	"0:01:31"	The point begins with a serve 1 minute and thirty-one seconds after the start of the first point of the match.
<i>point_no, game_no, set_no</i> (“no” is an abbreviation for number)	4, 1, 1	The point played is the 4 <sup>th</sup> point of the 1 <sup>st</sup> game of the 1 <sup>st</sup> set of the match.
<i>p1_sets, p2_sets, p1_games, p2_games</i>	0, 0, 0, 0	Since this is the first game of the match neither player has won a game or set yet.
<i>p1_score, p2_score</i>	15, 30	The score when the point is played is 15 (player 1), to 30 (player 2). Thus, player 1 won one of the previous points and player 2 won two points.
<i>server</i>	1	Player 1 (Alcaraz) is serving on this point.
<i>serve_no</i>	1	The point was played on the first serve meaning Alcaraz hit his first serve in play.
<i>point_victor</i>	1	Alcaraz wins this point (player 1).
<i>p1_points_won, p2_points_won</i>	2, 2	Player 1 (Alcaraz) is the point victor so his total is now 2 for the match (it was previously 1). For player 2 the value remains 2 since player 2 lost the point.
<i>game_victor, set_victor</i>	0, 0	Alcaraz winning the point makes the score in the game 30 – 30 (2 points each) so neither a game or set was won by either player on this point (both = 0).
<b>Columns U – AC</b>		Allow us to determine how the point was won:
<i>p1_winner</i>	1	Alcaraz won the point by hitting an “untouchable” shot.
<i>p1_ace</i>	0	The shot was not a serve (since = 0).
<i>winner_shot_type</i>	F	The shot was a forehand (as opposed to a backhand).
<i>p2_net_pt</i>	1	Player 2 (Jarry) positioned himself near the net somewhere during the point.
<i>p2_net_pt_won</i>	0	Since Alcaraz won the point, although Jarry was at the net during the point this value is 0.
<b>Columns AH – AM</b>		Even had player 2 won the point, the game would not have been over so the point was not a “break point” and these are all 0.
<i>p1_distance_run, p2_distance_run</i>	51.108, 75.631	The distance each player ran (in meters) on this point.
<i>rally_count</i>	13	Number of shots hit during the point by both players combined.
<i>speed_mph, serve_width, serve_depth, return_depth</i>	130, BW, CTL, D	Alcaraz (the server) hit a 130 serve “Body/Wide” of the returner (we saw it was a first serve previously) and close to the line denoting in or out of play. Jarry (the returner) returned the ball “Deep” in the court (so near the other end of the court).

# 帮助理解数据集的示例

示例 1：第 5 行

列	价值	描述
match_id	“2023-温布尔登-1301”	“1301”中的 3 表示第 3 轮比赛，“01”表示该轮列出的第一场比赛。
elapsed_time	“0:01:31”	该点从比赛第一分开始后 1 分 31 秒的发球开始。
point_no、game_no、set_no (“no”是数字的缩写)	4, 1, 1	打出的点是1盘比赛的1局的4分。
p1_sets、p2_sets、 p1_games、p2_games	0, 0, 0, 0	由于这是比赛的第一场比赛，因此两位球员都没有赢得一局或一盘。
p1_score、p2_score	15, 30	打出该点时的分数是 15（玩家 1）到 30（玩家 2）。因此，玩家 1 赢得了前一个积分，玩家 2 赢得了两个积分。
服务器 serve_no	1 1	1号球员（阿尔卡拉斯）在这一点上发球。 这个分数是在第一个发球局打出的，这意味着阿尔卡拉斯在比赛中击中了他的第一个发球局。
point_victor p1_points_won、p2_points_won	1 2, 2	阿尔卡拉斯赢得这一分（玩家 1）。 1号玩家（阿尔卡拉斯）是得分胜者，所以他的比赛总分现在是2分（以前是1分）。对于玩家 2，该值仍为 2，因为玩家 2 失去了点数。
game_victor、set_victor	0, 0	阿尔卡拉斯赢得这一分使比赛的比分成为30-30（各得2分），因此在这点上，任何一方都没有赢得一场比赛或一盘（均=0）。
U - AC 柱 p1_winner	1	请允许我们确定该点是如何赢得的： 阿尔卡拉斯通过击中“不可触碰”的射门赢得了这一分。
p1_ace winner_shot_type	0 F	击球不是发球（因为 = 0）。 击球是正手（而不是反手）。
p2_net_pt	1	2号球员（Jarry）在得分时将自己定位在球网附近的某个地方。
p2_net_pt_won	0	由于阿尔卡拉斯赢得了这一分，尽管贾里在该点期间在网前，但该值为0。
色谱柱 AH - AM	全部 = 0	即使玩家 2 赢了一分，比赛也不会结束，所以这个点不是“破发点”，这些都是 0。
p1_distance_run、 p2_distance_run	51.108, 75.631	每个玩家在这点上跑的距离（以米为单位）。
rally_count	13	两名球员在得分期间的投篮次数加起来。
speed_mph、serve_width、 serve_depth、 return_depth	130、带宽、CTL、D	阿尔卡拉斯（发球手）击中了接发球手的130个发球 “身体/宽度”（我们之前看到这是第一次发球），并且靠近表示比赛或出局的线。贾里（回球手）在球场上“深”地回球（所以靠近球场的另一端）。

**Example 2: rows 8 – 12**

The final four points of the first game illustrate the concept of tied score (“deuce”) and advantage (“ad”). Each row is a subsequent point in time in the match.

Row	Column(s)	Value(s)	Description
Row 8	<i>p1_score,</i> <i>p2_score</i>	40, 40	The score is 40 – 40 meaning each player has won 3 previous points (this is also called “deuce”).
	<i>point_victor</i>	1	Alcaraz wins point 7 (in row 8).
Row 9	<i>p1_score,</i> <i>p2_score</i>	AD, 40	Since Alcaraz won the previous point (point 7) the score on point 8 is now “AD” for Alcaraz and “40” for Jarry meaning Alcaraz has won one more point and could win the game on the next point.
	<i>point_victor</i>	2	Jarry (player 2) wins point 8 (in row 9).
Row 10	<i>p1_score,</i> <i>p2_score</i>	40, 40	The score returns to 40 – 40 (“deuce”) meaning each player has won the same number of previous points although now it is 4 points each.
	<i>point_victor</i>	1	Alcaraz wins point 9 (in row 10).
Row 11	<i>p1_score,</i> <i>p2_score</i>	AD, 40	Alcaraz again has the advantage having won point 9.
	<i>point_victor</i>	1	Alcaraz wins point 10 (in row 11) which means he has won the game (has score 2 more points now).
Row 12	<i>game_no</i>	2	This is now the first point of game 2.
	<i>p1_games</i>	1	Alcaraz won game 1.

**Example 3: row 51**

The 51<sup>st</sup> point of the match illustrates “break points” – points where the player not serving (the player who is returning serve) has an opportunity to win the game.

Row	Column(s)	Value(s)	Description
Row 51	<i>p1_score,</i> <i>p2_score</i>	40, 30	The score is 40 – 30 meaning player 1 (Alcaraz) is ahead.
	<i>server</i>	2	Jarry (player 2) is serving.
	<i>p1_break_pt</i>	1	If Alcaraz wins the point he will win the game; since he is not serving this is a “break point.”
	<i>point_victor</i>	1	Alcaraz wins the point (and therefore the game).
	<i>p1_break_pt_won</i>	1	Alcaraz won the game and was not serving on the point.

示例 2：第 8 - 12 行

第一场比赛的最后四分说明了平局（“平分”）和优势（“ad”）的概念。  
每一行都是比赛中的后续时间点。

行列值			描述
第 8 行	p1_score, p2_score	40, 40	分数是 40 - 40，这意味着每个玩家都赢得了 3 个先前的积分（这也称为“平分”）。
第 9 行	point_victor p1_score, p2_score	1 公元40年	阿尔卡拉斯赢得第7分（第8排）。 由于阿尔卡拉斯赢得了上一分（第7分），因此第8点的分数现在阿尔卡拉斯为“AD”，贾里为“40”，这意味着阿尔卡拉斯又赢了一分，并可能在下一分赢得比赛。
第 10 行	point_victor p1_score, p2_score	2 40, 40	Jarry（玩家 2）赢得第 8 分（第 9 排）。 比分回到 40 - 40（“平分”），这意味着每个玩家都赢得了相同数量的先前积分，尽管现在是每人 4 分。
第 11 行	point_victor p1_score, p2_score	1 公元40年	阿尔卡拉斯赢得第9分（第10排）。 阿尔卡拉斯再次获得第9分的优势。
第 12 行	point_victor game_no p1_games	1 2 现在是第 2 场比赛的第一点。 1	阿尔卡拉斯赢得第10分（第11排），这意味着他赢得了比赛（现在又得了2分）。 阿尔卡拉斯赢得了第1场比赛。

示例 3：第 51 行

比赛的 51 分说明了“破发点”——不发球的球员（接发球的球员）有机会赢得比赛的点。

行列值			描述
第 51 行	p1_score, p2_score	40, 30	比分是40-30，这意味着1号球员（阿尔卡拉斯）领先。
	服务器	2	贾里（球员 2）正在发球。
	p1_break_pt	1	如果阿尔卡拉斯赢得一分，他将赢得比赛；由于他没有服务，这是一个“断点”。
	point_victor	1	阿尔卡拉斯赢得了一分（因此也赢得了比赛）。
	p1_break_pt_won	1	Alcaraz won the game and was not serving on the point.

## **Use of Large Language Models and Generative AI Tools in COMAP Contests**

This policy is motivated by the rise of large language models (LLMs) and generative AI assisted technologies. The policy aims to provide greater transparency and guidance to teams, advisors, and judges. This policy applies to all aspects of student work, from research and development of models (including code creation) to the written report. Since these emerging technologies are quickly evolving, COMAP will refine this policy as appropriate.

Teams must be open and honest about all their uses of AI tools. The more transparent a team and its submission are, the more likely it is that their work can be fully trusted, appreciated, and correctly used by others. These disclosures aid in understanding the development of intellectual work and in the proper acknowledgement of contributions. Without open and clear citations and references of the role of AI tools, it is more likely that questionable passages and work could be identified as plagiarism and disqualified.

Solving the problems does not require the use of AI tools, although their responsible use is permitted. COMAP recognizes the value of LLMs and generative AI as productivity tools that can help teams in preparing their submission; to generate initial ideas for a structure, for example, or when summarizing, paraphrasing, language polishing etc. There are many tasks in model development where human creativity and teamwork is essential, and where a reliance on AI tools introduces risks. Therefore, we advise caution when using these technologies for tasks such as model selection and building, assisting in the creation of code, interpreting data and results of models, and drawing scientific conclusions.

It is important to note that LLMs and generative AI have limitations and are unable to replace human creativity and critical thinking. COMAP advises teams to be aware of these risks if they choose to use LLMs:

- **Objectivity:** Previously published content containing racist, sexist, or other biases can arise in LLM-generated text, and some important viewpoints may not be represented.
- **Accuracy:** LLMs can ‘hallucinate’ i.e. generate false content, especially when used outside of their domain or when dealing with complex or ambiguous topics. They can generate content that is linguistically but not scientifically plausible, they can get facts wrong, and they have been shown to generate citations that don’t exist. Some LLMs are only trained on content published before a particular date and therefore present an incomplete picture.
- **Contextual understanding:** LLMs cannot apply human understanding to the context of a piece of text, especially when dealing with idiomatic expressions, sarcasm, humor, or metaphorical language. This can lead to errors or misinterpretations in the generated content.
- **Training data:** LLMs require a large amount of high-quality training data to achieve optimal performance. In some domains or languages, however, such data may not be readily available, thus limiting the usefulness of any output.

## Guidance for teams

Teams are required to:

1. **Clearly indicate the use of LLMs or other AI tools in their report**, including which model was used and for what purpose. Please use inline citations and the reference section. Also append the Report on Use of AI (described below) after your 25-page solution.
2. **Verify the accuracy, validity, and appropriateness** of the content and any citations generated by language models and correct any errors or inconsistencies.
3. **Provide citation and references, following guidance provided here.** Double-check citations to ensure they are accurate and are properly referenced.
4. **Be conscious of the potential for plagiarism** since LLMs may reproduce substantial text from other sources. Check the original sources to be sure you are not plagiarizing someone else's work.

<p><b>COMAP will take appropriate action when we identify submissions likely prepared with undisclosed use of such tools.</b></p>
---

## Citation and Referencing Directions

Think carefully about how to document and reference whatever tools the team may choose to use. A variety of style guides are beginning to incorporate policies for the citation and referencing of AI tools. Use inline citations and list all AI tools used in the reference section of your 25-page solution.

Whether or not a team chooses to use AI tools, the main solution report is still limited to 25 pages. If a team chooses to utilize AI, following the end of your report, add a new section titled Report on Use of AI. This new section has no page limit and will not be counted as part of the 25-page solution.

Examples (this is *not* exhaustive – adapt these examples to your situation):

### **Report on Use of AI**

1. OpenAI *ChatGPT* (Nov 5, 2023 version, ChatGPT-4)  
Query1: *<insert the exact wording you input into the AI tool>*  
Output: *<insert the complete output from the AI tool>*
2. OpenAI *Ernie* (Nov 5, 2023 version, Ernie 4.0)  
Query1: *<insert the exact wording of any subsequent input into the AI tool>*  
Output: *<insert the complete output from the second query>*
3. Github *CoPilot* (Feb 3, 2024 version)  
Query1: *<insert the exact wording you input into the AI tool>*  
Output: *<insert the complete output from the AI tool>*
4. Google *Bard* (Feb 2, 2024 version)  
Query: *<insert the exact wording of your query>*  
Output: *<insert the complete output from the AI tool>*