Australian Open Tracking Data Dictionary

The dataset includes shot-level summaries of points played in men’s and women’s singles matches from the 2017 and 2018 Australian Open. Each point includes the final shot and information about the shot just before (if there were 2 shots or more). The variables include information include position of the players, position of the ball, and speed/height of the ball.

These data are based off of a multi-camera tracking system. It is important to note that the camera system provides estimates of the player and ball position, and measurement error in the system is possible (though anticipated to be small in most cases).

The only manually collected variable is the ‘outcome’. This is a subjective classification of the result of the point, which assigned 1 of 3 mutually exclusive possibilities: winner, unforced error and forced error. This labelling is performed by courtside annotators.

One note on shots that do not have a bounce (ie volleys). In cases where no bounce was recorded for a shot, we estimate the landing location by projecting the tracked arc of the shot (a 3D polynomial) to where the height of the arc would reach the ground.

|  |  |  |
| --- | --- | --- |
| Variable | Description | Value Range |
| event | Character identifier of event | AO2017 or AO2018 |
| year | Year match took place | 2017 or 2018 |
| matchid | Character identifier of match | Men’s matches begin with MS and women’s with WS |
| pointid | Character identifier of point | Syntax follows <year> <matchid>\_<set>\_<game>\_<point> |
| rally | The number of shots in the point counting serves and shots ending in errors | An integer from 1, 2, 3... |
| shotinrally | The number of the current shot in the rally (1 = serve) | An integer from 1, 2, 3... |
| serve | A number indicating whether the point was played on a first or second serve. | 1 = First  2 = Second |
| hitpoint | Shot category for point-ending shot | F = Forehand  B = Backhand  S = Serve  U = Unknown   * Note that unknown categories can occur when the tracking system encounters a shot type that is not recognized or it misses the identification of a true F or B shot |
| speed | Speed of point-ending shot | Continuous (m/s) |
| net.clearance | Distance above the net as point-ending shot passed the net | Continuous (cm) distance above net. Can be negative if shot did not pass above the net. NOTE: Di assumes missings are negative, and replaces with -0.5 |
| distance.from.sideline | Lateral distance of the point-ending shot bounce from the nearest singles sideline. | Perpendicular distance in meters. This can be negative if the ball passed the sideline. |
| depth | Distance of the point-ending shot bounce from the baseline | Perpendicular distance from the net in meters  (always positive even if out) |
| outside.sideline | Logical indicator of whether point-ending shot landed outside of the in-play singles sideline (4.115 meters from centre) | TRUE, FALSE |
| outside.baseline | Logical indicator of whether point-ending shot landed beyond the in-play baseline (11.89 m from net) | TRUE, FALSE |
| player.impact.distance.travelled | Distance travelled by player who made the current shot from start to the end of the shot | Euclidean distance in meters |
| player.impact.depth | Depth of position of player who made the current shot | Straight line distance from net in meters |
| player.impact.distance.from.center | Lateral position of player who made the current shot | Straight line distance from the center line in meters |
| opponent.distance.travelled | Distance travelled by player who received the current shot from start to the end of the shot | Euclidean distance in meters |
| opponent.depth | Depth of position of player who received the current shot | Straight line distance from net in meters |
| opponent.distance.from.center | Lateral position of player who received the current shot | Straight line distance from the center line in meters |
| player.depth | Distance of player who made point-ending shot from the net at the time the penultimate shot was made | Perpendicular distance along the length of court from net in meters |
| same.side | Logical indicator if both player and opponent were positioned on the same side of the center line (ad or deuce court) for the current shot | TRUE, FALSE |
| previous.speed | Speed of penultimate shot | Continuous (m/s)  *Missing Cases* = Will be missing if shot ended on serve NOTE: Di replaced missing with 0, lower than provided values |
| previous.net.clearance | Distance above the net as penultimate shot passed the net | Continuous (cm) distance above net. Can be negative if shot did not pass above the net.  *Missing Cases* = Will be missing if shot ended on serve NOTE: Di replaced missing with median 0.476, lower than provided values |
| previous.distance.from.sideline | Lateral distance of the penultimate shot bounce from the nearest singles sideline. | Perpendicular distance in meters (always positive even if out)  *Missing Cases* = Will be missing if shot ended on serve NOTE: Di replaced missing with 5, a large recognizable value |
| previous.depth | Distance of the penultimate shot bounce from the baseline | Perpendicular distance in meters  (always positive even if out)  *Missing Cases* = Will be missing if shot ended on serve NOTE: Di replaced missing with 15, a large recognizable value |
| previous.hitpoint | Shot category for penultimate shot | F = Forehand  B = Backhand  S = Serve  U = Unknown  *N = No previous point* |
| previous.time.to.net | Time for penultimate shot to be hit and pass the net | Continuous number in seconds  *Missing Cases* = Will be missing if shot ended on serve NOTE: Di replaced missing with 5, a large recognizable value |
| server.is.impact.player | Logical if player who made point-ending shot was the server of the point | TRUE, FALSE |
| serverwon | Logical if server won the point | TRUE, FALSE |
| outcome | Target variable, character with three categories indicating the type of shot that ended the point | W = Winner  F = Forced Error  U = Unforced Error |