**Glossary for Dataset**

* **game\_id** – id of each game
* **result\_id** – result of each game (1 if team1 wins, 0 if team1 loses)
* **season** – Years that the NCAA season spanned. The tournament takes place in the latter year
* **team1\_id** – Team 1 id
* **team2\_id** – Team 2 id
* **team1\_score** – Final score of team 1
* **team2\_score** – Final score of team2
* **numot** – Number of overtimes
* **team1\_seed** – Team 1 seed in the tournament. There are 4 brackets of teams seeded 1 through 16.
* **team2\_seed** – Team 2 seed in the tournament.
* **host\_site** – The location of game
* **host\_lat** – Latitude of location of game
* **host\_long** – Longitude of location of game

**The following stats are for both team 1 and team 2:**

* **lat** – Latitude of team’s home campus.
* **long –** Longitude of team’s home campus.
* **teamname** – Team name
* **fg2pct –** Shooting percentage on 2 point field goals.
* **fg3pct –** Shooting percentage on 3 point field goals.
* **ftpct –** Shooting percentage on free throws.
* **blockpct -** Blocked shots divided by opponents 2 point field goal attempts.
* **oppfg2pct –** Opponent’s shooting percentage on 2 point field goals.
* **oppfg3pct** – Opponent’s shooting percentage on 3 point field goals.
* **oppftpct** – Opponent’s shooting percentage on free throws.
* **oppblockpct** – Opponent’s blocked shots divided by 2 point field goal attempts.
* **f3grate** – Percentage of field goal attempts that are 3 point field goal attempts.
* **oppf3grate** – Percentage of opponent’s field goal attempts that are 3 point field goal attempts.
* **arate** – Assists divided by field goals made. Percentage of field goals that were preceded by an assist.
* **opparate** – Opponent’s Assists divided by opponent’s field goals made.
* **stlrate** – Steals divided by defensive possessions.
* **oppstlrate** – Opponent’s steals divided by opponent’s defensive possessions.
* **tempo** - Estimated possessions using this formula: FGA-OR+TO+0.475xFTA. For each team, possessions are counted for the team and their opponents, and then averaged. A team’s average tempo is total possessions divided by minutes.
* **adj\_tempo** – An estimate of the tempo (possessions per 40 minutes) a team would have against the team that wants to play at an average D-I tempo.
* **oe** – Points scored per 100 offensive possessions.
* **adjoe** – An estimate of the offensive efficiency (points scored per 100 possessions) a team would have against the average D-I defense.
* **de –** Points allowed per 100 defensive possessions.
* **adjde** – An estimate of the defensive efficiency (points allowed per 100 possessions) a team would have against the average D-I offense
* **ap\_preseason** – The preseason AP Poll ranking of each team (top 25 only)
* **ap\_final** – The final AP Poll ranking of each team (top 25 only)
* **coaches\_preseason** - The preseason Coaches Poll ranking of each team (top 25 only)
* **coaches\_before\_final** – The most recent Coaches Poll rankings before the final. The Coach’s poll final poll is released after the tournament (making it not very valuable for predictions)

**The following stats are coach data**

* **coach\_id –** coach’s id
* **coach\_name** – coach’s name
* **pt\_school\_ncaa** – Number of NCAA Tournament appearances at current school
* **pt\_overall\_ncaa** – Career overall number of NCAA Tournament appearances
* **pt\_school\_s16** – Number of NCAA Sweet Sixteen appearances at current school
* **pt\_overall\_s16** – Career overall number of NCAA Sweet Sixteen appearances
* **pt\_school\_ff** – Number of NCAA Final Four appearances at current school
* **pt\_overall\_ff** – Career overall number of NCAA Final Four appearances
* **pt\_career\_school\_wins** – Number of wins at current school
* **pt\_career\_school\_losses** – Number of losses at current school
* **pt\_career\_overall\_losses** – Career overall number of losses
* **pt\_team\_season\_wins** – Team’s number of wins in this season
* **pt\_team\_season\_losses** – Team’s number of losses in this season
* **pt\_coach\_season\_wins** – Coach’s number of wins in this season
* **pt\_coach\_season\_losses** – Coach’s number of losses in this season