Basketball Analytics Tool

Illinois Men's Basketball Interview Assignment

Chen-Wei Hsu

Rationale: Addressing the challenge of opponent scouting and game planning in college basketball by creating a tool to give custom analyses and insights

- Coaches have limited time and resources during a long season
- It's tough to break down every opponent in detail, especially on short notice
- Currently, there isn't a simple way to quickly get data-driven insights for each specific matchup
- Thus, I wanted to create a tool that makes scouting faster, smarter, and easier

My Goals

- Help coaches prepare for games more effectively by solving the challenge of opponent analysis and strategy
- Pull data from KenPom and turn it into clear insights in order to give coaching staff the information they need without digging through multiple spreadsheets
- Support better decision-making and game planning

Methods

Why Python and Streamlit?

- Easy to work with for large amounts of data
- Fast to build and test ideas
- Very compatible with Streamlit to make simple, interactive tools in a working application with nice user interface
- Anyone on the coaching staff can use the tool without needing any coding or technical background

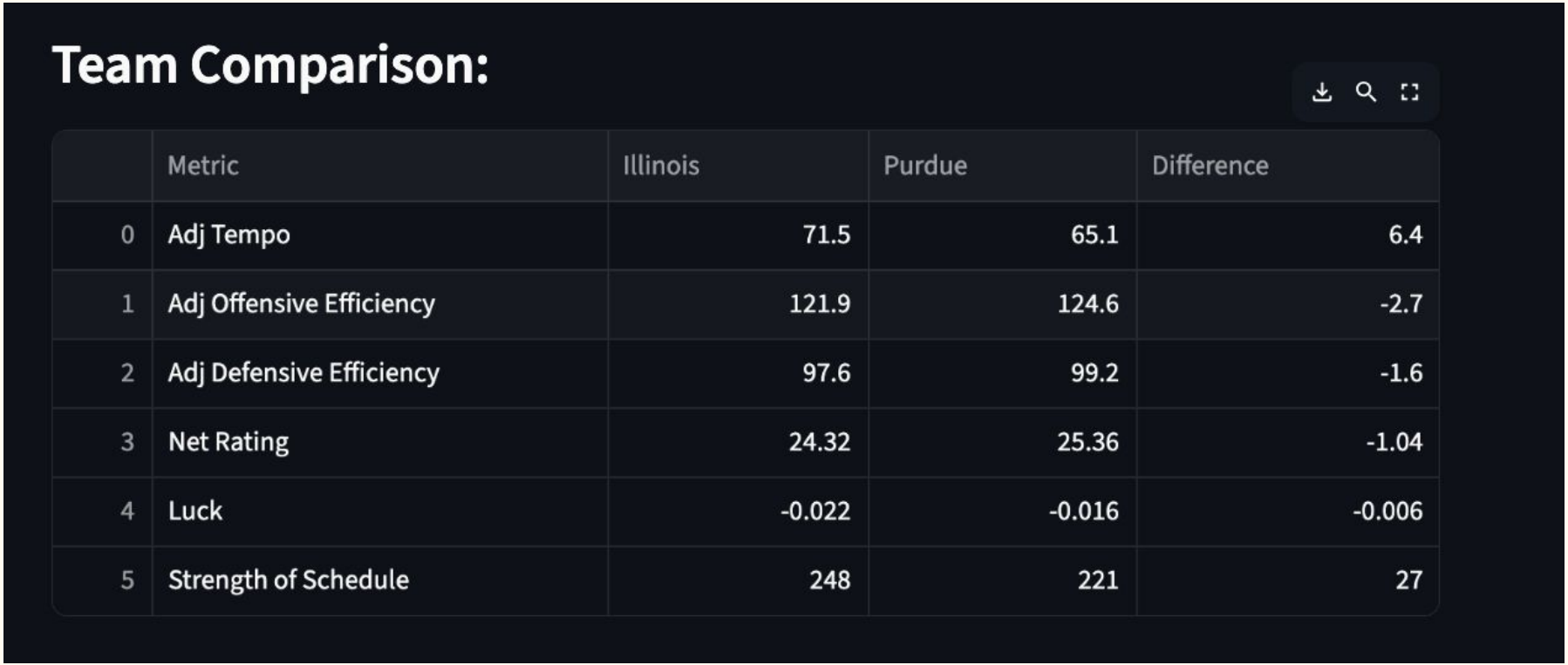
Tools in Python Used

- pandas – to load, organize, and compare KenPom team stats
- re (regular expressions) – to clean up messy column names
- difflib.get_close_matches – to fix team name typos and suggest the right ones

Results - Team Comparison

- Developed a tool that compares two teams' *Tempo*, *Offensive Efficiency*, *Defensive Efficiency*, *Net Rating*, *Luck*, *and Strength of Schedule* by extracting 2025 KenPom data
- Calculates the difference between teams for each metric
- Outputs results in a clear, side-by-side table to highlight matchup advantages and differences

Output - Team Comparison



Team Comparison:

	Metric	Illinois	Purdue	Difference
0	Adj Tempo	71.5	65.1	6.4
1	Adj Offensive Efficiency	121.9	124.6	-2.7
2	Adj Defensive Efficiency	97.6	99.2	-1.6
3	Net Rating	24.32	25.36	-1.04
4	Luck	-0.022	-0.016	-0.006
5	Strength of Schedule	248	221	27

Results - Opponent Analysis & Strategy Recommendations

- Developed a tool that identifies an opponent's team and automatically pulls their *adjusted tempo* data from KenPom
- Generates custom strategy suggestions across three areas:
  - Offense: ideas for tempo flexibility and matchup-based sets
  - Defense: ways to disrupt opponent rhythm and handle pace shifts
  - Rotation: substitution and personnel strategies to match tempo
- Outputs tailored coaching notes designed to support quick prep and game planning

Output - Opponent Analysis & Strategy Recommendations

Recommended Coaching Notes:

Tempo Profile: ~65.1 possessions/game (Balanced classification).

Offensive Ideas:

- Be flexible between pace options.
- Mix early offense with patient sets.
- Use matchup reads to adjust on the fly.

Defensive Ideas:

- Prepare for pace shifts mid-game.
- Emphasize communication to handle adjustments.
- Force them to settle into uncomfortable rhythm.

Rotation/Personnel Notes:

- Favor versatile players who can handle different tempos.
- Adjust lineups to exploit matchups dynamically.
- Plan substitutions to match opponent's pace.

Results - Game Risk & Volatility Estimator

- Created a tool that compares two teams using 2025 KenPom data to assess game volatility and upset potential
- Uses three key metrics to evaluate game dynamics:
  - Net Rating gap – measures overall team strength difference
  - Luck differential – captures consistency or randomness in performance
  - Strength of Schedule – identifies potential differences in opponent difficulty
- Generates a combined written analysis based on these stats
- Outputs a final prediction with:
  - Expected winner
  - Confidence score
  - Upset risk level

How the Game Risk & Volatility Estimator Works

Score the Upset Risk (0–3 scale)

- Add 1 point for each:
- Net Rating gap < 5
- Luck difference > 0.05
- Strength of Schedule difference > 1.5
- Score determines Upset Risk Level:
  - $0 =$ Low, $1 =$ Moderate, $2–3 =$ High

Adjust Score to Better Measure Team Strength

- Adjusted Score $=$ Net Rating $\times 0.85 +$ Schedule $\times 0.1 —$ Luck $\times 0.05$

Predict Winner & Confidence

- The team with the higher adjusted score is picked to win
- Confidence Score (50%–99%) reflects how big that gap is
  - Higher = more certain; lower = less certain

Output - Game Risk & Volatility Estimator

Game Risk Assessment

Net Rating gap: -1.0 → This game could possibly be close and upset-prone.

Luck differential: -0.006 → Fairly stable performance from both teams.

Strength of Schedule difference: 27.0 → One team may have faced tougher competition.

Final Prediction:

- Predicted Winner: Illinois
- Confidence Score: 58.2%
- Upset Risk Level: High

Possible Limitations and Improvements

Confidence score logic: The current formula based on Net Rating generally works well, but it can give misleading results when comparing a low-major team to a high-major team.

- How to improve: Adjust the score calculation for extreme mismatches by using thresholds or scaling factors.

Lack of game context: The model doesn't consider important factors like home-court advantage, travel fatigue, or how recently a team played.

- How to improve: Add inputs for location, rest days, and game timing to make the prediction more game-specific.

Possible Limitations and Improvements

KenPom data limitations: KenPom provides strong team-level stats, but it doesn't capture player-level details or in-game situations.

- How to improve: If possible, integrate richer data like shot quality, player usage, or lineup efficiency.

Only looks at current season: The tool gives a snapshot based on this season alone, without looking at trends over time.

- How to improve: Add recent game momentum or multi-year data to give a fuller picture.

Applications and Examples of Specific Use-Cases

- *Scouting Opponents* – Quickly compare teams' strengths, weaknesses, and tempo styles to tailor game plans
- *Preparing for Upcoming Games* – Generate custom strategy suggestions based on how fast or slow the opponent plays
- *Making In-Season Adjustments* – Use updated KenPom stats to adapt to trends and refine team strategy week to week
- *Tournament Readiness* – Assess volatility and upset risk to stay prepared for unpredictable matchups
- *Supporting Analysts & Broadcasters* – Provide clear, data-driven talking points and matchup summaries for game coverage

Summary and Conclusion

- This project developed a practical tool to support basketball coaches with data-driven insights.
- It includes three key features:
  - Team Comparison
  - Opponent Analysis & Strategy Recommendations
  - Game Risk & Volatility Estimator
- The tool helps coaches make quicker, more informed decisions during scouting and prep.
- While it's currently based on KenPom data, the framework is flexible and could be expanded with more detailed inputs in the future.