**Scenario:**

You are a **Data Analyst** at a rising sports analytics startup called **OctaIQ**. Your company specializes in performance insights for fight promotions, athlete management, and sports betting.

The **Director of Strategic Development** has just called you into a high-stakes meeting:

“Team, we're about to sign a partnership deal with a new regional fight promotion. They want to use data to make smarter matchmaking decisions and boost audience engagement.

But here's the problem: Their marketing team wants flashy fights with knockout potential to grow popularity. Meanwhile, the coaches and managers want matchups that are statistically more likely to protect their fighters’ records and long-term careers.

We need your help. Dig into our historical UFC fight data and give us a strategy. I want real answers backed by real analysis. Don't just tell me who won — tell me **why** they won, and what it means for our matchmaking and promotion strategy."

**Your Mission:**

Use the UFC dataset to produce a strategic recommendation that answers the following:

**1. Risk vs. Reward:**

Do aggressive fighters who go for early KOs tend to win more, or burn out and lose?  
Do defensive, decision-style fighters win more consistently, even if fights are longer and less exciting?

**Goal**: Explore whether knockout-heavy fights (with short durations and high strikes per minute) are higher-risk but potentially higher-reward — or if consistent technical fighters are statistically safer bets.

**2. The Stats That Matter:**

Which specific stats (striking accuracy, takedown defense, submission averages) most strongly correlate with winning?

**Goal**: Help coaches identify which skills actually lead to victory — not just highlight reels — and shape training around data-driven priorities.

**3. The Finisher Formula:**

What’s the win rate by fight-ending method (KO/TKO, submission, decision)?  
Which divisions or fighter profiles have the highest finish rates?

**Goal**: Give promoters insight into which matchups are more likely to result in explosive finishes — which fans love — versus strategic, decision-heavy battles.

**4. The Weight of the Fight:**

Do title fights or main events significantly differ in performance patterns (e.g., longer duration, more decisions)?  
Are fighters more likely to lose when competing outside their home country?

**Goal**: Help fight organizers and bettors understand how context — not just stats — influences fight outcomes.

**5. The Matchmaking Sweet Spot:**

Based on divisions, styles, and outcomes, what’s the **optimal profile for matchmaking** that balances excitement and win probability?

**Goal**: Recommend the best division, fight style, and fight card structure to both **minimize risk** (fighter loss/injury) and **maximize promotional value** (audience excitement and finishes).

**Columns:**

 **event\_name**: Name of the UFC event.

 **date**: Date when the event took place.

 **location**: City and country where the event occurred.

 **division**: Weight class or division of the fight.

 **title\_fight**: Indicates if the fight was a title bout (1) or not (0).

 **method**: Method by which the fight was won (e.g., KO, Submission, Decision).

 **finish\_round**: Round in which the fight ended.

 **match\_time\_sec**: Total duration of the fight in seconds.

 **r\_kd**: Number of knockdowns scored by the red fighter.

 **r\_sig\_str**: Significant strikes landed by the red fighter.

 **r\_sig\_str\_att**: Significant strike attempts by the red fighter.

 **r\_sig\_str\_pct**: Percentage of significant strikes landed by the red fighter.

 **r\_total\_str**: Total strikes landed by the red fighter.

 **r\_td**: Takedowns landed by the red fighter.

 **r\_td\_att**: Takedown attempts by the red fighter.

 **r\_td\_pct**: Takedown success rate of the red fighter.

 **r\_sub\_att**: Submission attempts by the red fighter.

 **r\_pass**: Guard passes by the red fighter.

 **r\_rev**: Reversals by the red fighter.

 **r\_ctrl**: Control time for the red fighter.

 **r\_head**: Head strikes landed by the red fighter.

 **r\_body**: Body strikes landed by the red fighter.

 **r\_leg**: Leg strikes landed by the red fighter.

 **r\_distance**: Strikes landed at distance by the red fighter.

 **r\_clinch**: Strikes landed in the clinch by the red fighter.

 **r\_ground**: Strikes landed on the ground by the red fighter.

 **r\_win**: Whether the red fighter won (1) or not (0).

 **r\_draw**: Whether the red fighter drew (1) or not (0).

 **r\_splm**: Strikes per landed minute for the red fighter.

 **r\_str\_acc**: Striking accuracy percentage for the red fighter.

 **r\_sapm**: Strikes absorbed per minute by the red fighter.

 **r\_str\_def**: Striking defense percentage for the red fighter.

 **r\_td\_avg**: Average takedowns per 15 minutes for the red fighter.

 **r\_td\_avg\_acc**: Takedown accuracy percentage for the red fighter.

 **r\_td\_def**: Takedown defense percentage for the red fighter.

 **r\_sub\_avg**: Average submissions per 15 minutes by the red fighter.

 **b\_kd**: Number of knockdowns scored by the blue fighter.

 **b\_sig\_str**: Significant strikes landed by the blue fighter.

 **b\_sig\_str\_att**: Significant strike attempts by the blue fighter.

 **b\_sig\_str\_pct**: Percentage of significant strikes landed by the blue fighter.

 **b\_total\_str**: Total strikes landed by the blue fighter.

 **b\_td**: Takedowns landed by the blue fighter.

 **b\_td\_att**: Takedown attempts by the blue fighter.

 **b\_td\_pct**: Takedown success rate of the blue fighter.

 **b\_sub\_att**: Submission attempts by the blue fighter.

 **b\_pass**: Guard passes by the blue fighter.

 **b\_rev**: Reversals by the blue fighter.

 **b\_ctrl**: Control time for the blue fighter.

 **b\_head**: Head strikes landed by the blue fighter.

 **b\_body**: Body strikes landed by the blue fighter.

 **b\_leg**: Leg strikes landed by the blue fighter.

 **b\_distance**: Strikes landed at distance by the blue fighter.

 **b\_clinch**: Strikes landed in the clinch by the blue fighter.

 **b\_ground**: Strikes landed on the ground by the blue fighter.

 **b\_win**: Whether the blue fighter won (1) or not (0).

 **b\_draw**: Whether the blue fighter drew (1) or not (0).

 **b\_splm**: Strikes per landed minute for the blue fighter.

 **b\_str\_acc**: Striking accuracy percentage for the blue fighter.

 **b\_sapm**: Strikes absorbed per minute by the blue fighter.

 **b\_str\_def**: Striking defense percentage for the blue fighter.

 **b\_td\_avg**: Average takedowns per 15 minutes for the blue fighter.

 **b\_td\_avg\_acc**: Takedown accuracy percentage for the blue fighter.

 **b\_td\_def**: Takedown defense percentage for the blue fighter.

 **b\_sub\_avg**: Average submissions per 15 minutes by the blue fighter.

 **winner**: Name of the fighter who won the bout.

**Tasks:**

1. Delete unnecessary columns:

r\_nick\_name, b\_nick\_name

1. Solve null problems
2. Convert date column to datetime and create year column
3. Find top 5 winner
   1. Find top 5 winner by years
4. Find total count of method
   1. Find count of KO by years
5. Find count of finish round
6. Find total match time by years
7. Find average Fight Duration per Division
8. Find Top Referees by Number of Fight
9. Find top 10 location
10. Find different fighter performance
11. Find top 10 Fighter with submission
12. Find number fights per year
13. Find top 10 average takedown
14. Find corr matrix