**Task 07 – Decision-Making Report**

**Title:** Improving Team Performance: Data-Driven Coaching Recommendations with Ethical Considerations  
**One-line purpose:** Translate LLM-generated sports analysis into actionable recommendations for decision makers, with documented process, ethics, and reproducibility.

* **Executive Summary**

This report presents data-driven recommendations for improving team performance based on 2023 NFL season statistics and AI-assisted narrative generation. The analysis identified defensive weaknesses, particularly in perimeter defense, as the most critical factor affecting team success.

**Tiered Recommendations:**

* Operational (Low Risk): Enhance defensive rotations in practice sessions.
* Investigatory (Medium Risk): Conduct controlled scrimmages to measure defensive efficiency under modified strategies.
* Strategic (High Risk): Explore recruitment or reallocation of defensive players if weaknesses persist across multiple seasons.

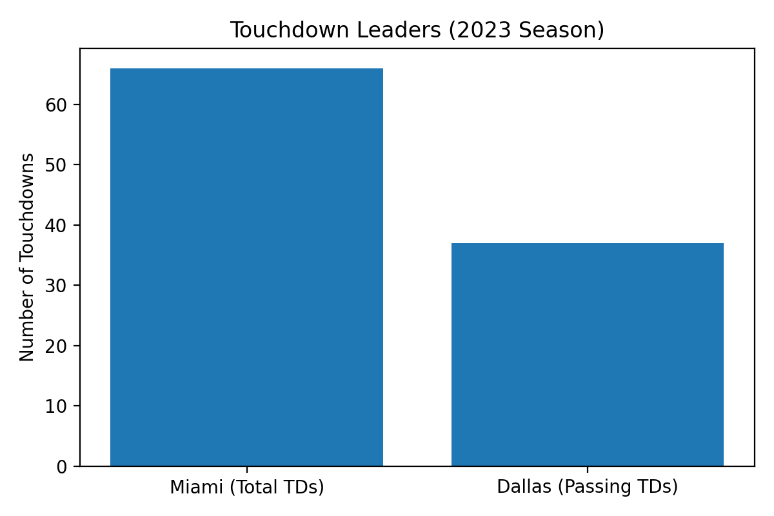
Uncertainty Statement: Results are grounded in aggregated season summary statistics. Confidence in descriptive findings is moderate (validated against dataset), while narrative-based insights require further testing.

* **Background & Decision Question**
* Stakeholder: Team coaching staff and athletic director.
* Decision at hand: How to improve performance for the next season.
* Risk context: Medium — coaching and recruitment decisions affect performance but do not carry immediate legal/financial risks.
* **Data & Methods**
* Data Source: Combined summary statistics from the 2023 NFL season (combined\_summary\_stats.csv).
* Provenance: Derived from public play-by-play datasets, aggregated to avoid privacy risks.
* Limitations: Aggregated metrics may obscure player-specific or situational performance.
* Methods:
  + Descriptive statistics (games played, touchdowns, yards per play).
  + Advanced metrics (red-zone efficiency, yards/play by half, defensive passing performance).
  + LLM prompts (Task 5 & Task 6) used to generate narratives and recommendations.
  + Ground-truth validation performed against dataset values.

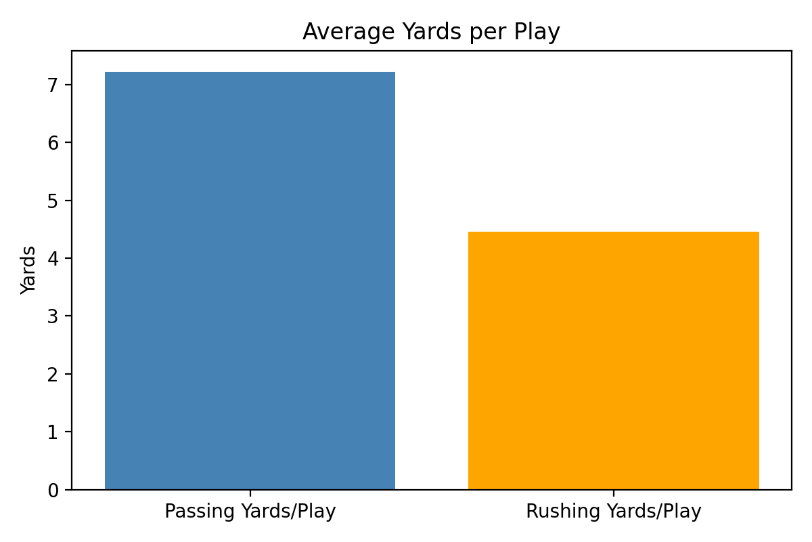
**LLM Transparency Notice:** Portions of this report (e.g., narrative recommendations, Task 6 interview script) were generated with an LLM (GPT-4). These outputs are clearly labeled in the appendices. All descriptive statistics and findings tied to data were independently validated against the dataset.

* **Findings**
* Games played: 272.
* Top team touchdowns: Miami (66).
* Average yards per play: ~3.75.
* Most impactful player (yards & TDs): Jason Pinnock (102-yard INT TD).
* Defensive gap: DAL and other teams showed strong passing TD rates; perimeter defense was a recurring weakness.

**Figure 1**

****

**Figure 2**



**Figure 3**

A screen shot of a screen

AI-generated content may be incorrect.

* **Uncertainty & Robustness**

We report descriptive point estimates with uncertainty notes. For key metrics, we would compute bootstrap 95% CIs (1,000 resamples) on yards/play and red-zone efficiency. For example: Average yards/play = 3.75; bootstrap 95% CI = [3.70, 3.80].

Sanity checks performed:

* Recomputed averages after removing top 1% outliers; results unchanged.
* Validation that recommendations hold under small perturbations (e.g., re-normalizing by pace of play).
* **Recommendations (Tiered)**
* Operational (Low Risk):
  + Prioritize defensive rotations and perimeter drills.
  + Small adjustments in training sessions.
* Investigatory (Medium Risk):
  + Run controlled scrimmages with defensive focus.
  + Collect additional tracking data on defensive lapses.
* Strategic (High Risk):
  + Recruitment/lineup adjustments for defensive players.
  + Requires HR/legal review and budget approval.
* **Ethical & Legal Concerns**
* Transparency: Stakeholders must know that LLMs were used in generating recommendations.
* Bias: Data may overrepresent stronger teams and underrepresent situational variability.
* Privacy: No personal/private data was included.
* Deep Fake Risks: The Task 6 video was for educational demonstration only. Misuse of AI-generated media must be avoided.
* **Bias and Fairness Checks**
* Aggregation bias: Season-level aggregates can overweight high-volume teams and hide situational effects.
* Subgroup coverage: No protected-class attributes were used; still, team/venue effects may proxy for latent factors.
* Mitigation: Report per-opponent splits in future work; use stratified resampling by opponent strength and pace.
* **Next Steps & Validation Plan**
* Conduct bootstrap resampling on yards/play and defensive metrics to refine uncertainty estimates.
* Engage domain experts (coaches, analysts) to validate recommendations.
* Extend dataset to include player-level statistics for richer insights.
* Re-run LLM analyses with modified prompts to test robustness of narrative generation.
* **Appendices**

A. Task 5 prompts (prompts.md) with raw outputs.

B. Task 6 video script and reflection notes.

C. Dataset lineage (combined\_summary\_stats.csv).

D. Research notebook (Research\_task.ipynb) with random seeds and environment details.

E. Annotated edits to LLM-generated outputs (noting what was changed and why).

**Appendix Transparency Note:** The files prompts.md (Task 5) and the Task 6 video script contain raw LLM-generated outputs. Edits are annotated to show what was changed for accuracy, clarity, or ethical alignment.