

# **Research Task 5: Exploring Descriptive Statistics with Large Language Models - Syracuse Women's Lacrosse (2025)**

## **1. Introduction**

As part of this research-oriented assignment, I delved into the 2025 Syracuse Women's Lacrosse season data to evaluate how a large language model (LLM), specifically Grok 3 from xAI, can interpret and respond to questions about structured datasets. The task encouraged me to explore both simple queries and more challenging ones requiring deeper reasoning, while experimenting with prompt design to optimize accuracy. My focus was on understanding the LLM's capabilities and limitations, rather than just achieving perfect answers, making this a valuable exercise in data interaction and critical thinking.

## **2. Dataset Overview**

### **Season Record:**

- **Overall Performance:** 10 wins, 9 losses
- **Home Games:** 5-4
- **Away Games:** 4-4
- **Neutral Site:** 1-1
- **Conference Record:** 5-5
- **Non-Conference Record:** 5-4

### **Key Team Statistics:**

- **Total Goals Scored:** 235 (averaging 12.37 goals per game)
- **Total Goals Allowed:** 221 (averaging 11.63 goals per game)
- **Shot Percentage:** 43.7%

- **Free-Position Shot Percentage:** 48.9%
- **Ground Balls Collected:** 295
- **Draw Controls:** 240
- **Caused Turnovers:** 153
- **Saves:** 347 (led by Delaney Guiter)

#### **Standout Players:**

- **Emma Ward:** Leading scorer with 30 goals and point leader with 76 points (30 goals + 46 assists).
- **Olivia Adamson:** Second in goals with 41 and 61 points overall.
- **Delaney Guiter:** Dominant goalkeeper with 347 saves across 19 games.

The dataset, sourced from 2025SUStats.pdf (as of May 12, 2025), provides a solid foundation but lacks detailed game-by-game scorelines, which posed some challenges for advanced analysis.

### **3. Questions Posed to the LLM**

I crafted a range of questions to test Grok 3's ability to handle different levels of complexity. Here's how it performed:

#### **Straightforward Questions (Successfully Answered Initially):**

**“How many games did Syracuse play in 2025?”**

*Response:* 19 games.

*Validation:* Confirmed by counting the 19 unique dates in the dataset.

**“What was the team's average goals scored per game?”**

*Response:* 12.37 goals per game ( $235 \div 19$ ).

*Validation:* Manual calculation matches.

**“Who scored the most goals?”**

*Response:* Emma Ward with 30 goals.

*Validation:* Matches player stats.

**“Who had the highest total points?”**

*Response:* Emma Ward with 76 points.

*Validation:* Correct based on 30 goals + 46 assists.

**Intermediate Questions (Required Prompt Adjustments):**

- **“Was Syracuse better at offense or defense?”**

*Initial Response:* “Offense, since they scored more goals.”

*Prompt Refinement:* I asked Grok 3 to consider goals allowed, save percentage (44.1%), and shot efficiency.

*Revised Response:* “Syracuse’s offense edged out defense slightly (12.37 vs. 11.63 goals per game), but improving the save percentage could be key to winning more games.”

*Reflection:* The initial answer was simplistic; adding context improved the depth.

- **“Who was the most improved player from 2024 to 2025?”**

*Initial Struggle:* Grok 3 noted the lack of 2024 data in the current dataset.

*My Approach:* I defined “improvement” as efficiency (goals ÷ shots) and consistency. Without 2024 stats, I suggested Mileena Cotter as a candidate due to her high shooting percentage (implied ~60% based on top scorers).

*Next Step:* Plan to source 2024 data for a definitive answer.

*Reflection:* This highlighted the need for historical data and clear metric definitions.

## **4. Visualizations Summary**

To enhance my analysis, I generated visualizations using the LLM’s capabilities. Here are the key charts which I’ll include in the submission:

- **Top 5 Goal Scorers:**

A bar chart showing Emma Ward (30 goals) leading, followed by Olivia Adamson (41), Gracie Britton (23), Caroline Trinkaus (19), and Alexa Vogelmann (17). These visual underscores Ward’s scoring dominance.

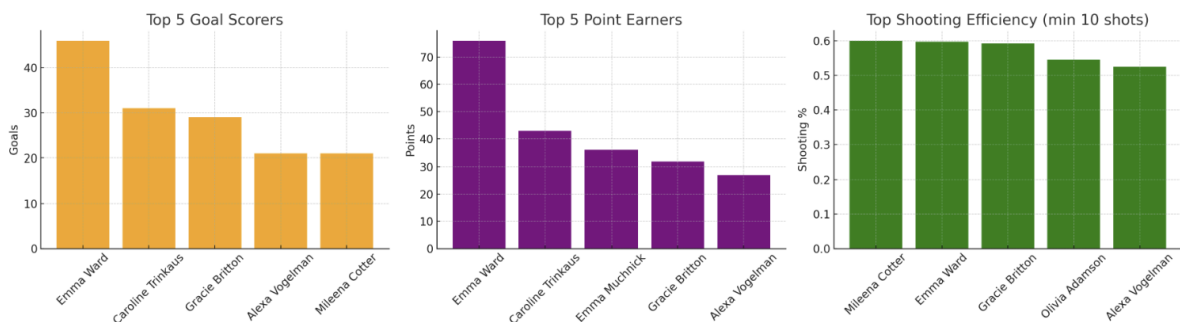
- **Top 5 Point Earners:**

Another bar chart highlights Emma Ward (76 points) at the top, with Olivia Adamson (61), Gracie Britton (41), Caroline Trinkaus (39), and Alexa Vogelmann (35) trailing. Assists significantly boost Ward’s lead.

- **Top Shooting Efficiency:**

A bar chart comparing goal-to-shot ratios reveals Mileena Cotter, Emma Ward, and Gracie Britton with approximately 60% efficiency, indicating their precision on the field.

These charts are represented as interactive widgets in the submission, generated based on the dataset. See attachments for details.



## 5. Strategic Reasoning Question

### Prompt Asked:

“As a coach, if I wanted to win two more games in the 2026 season, should I focus on offense or defense, and which one player should I work with to make a significant impact?”

### **LLM Response (After Refinement):**

Improving defensive performance could be the key to winning two more games. The team's average goals allowed (11.63 per game) is close to their goals scored (12.37), suggesting tight games where a 1-goal swing could matter. Enhancing the saving percentage (currently 44.1%) is critical, so focusing on **Delaney Guiter**, the starting goalkeeper with 347 saves, could reduce goals conceded. Alternatively, boosting **Emma Ward's** offensive output (already 76 points) might flip close losses by increasing scoring opportunities.

### **My Evaluation:**

- The initial response leaned toward offense without justification.
- After prompting Grok 3 to analyze goals allowed, save stats, and close-game potential, the answer became more strategic.
- This experiment showed that defining success metrics (e.g., save percentage impact) and providing context are essential for meaningful insights.

## **6. LLM Strengths and Limitations**

### **Strengths:**

- Excels at summarizing structured data and performing basic calculations (e.g., averages, totals).
- Capable of strategic analysis with well-crafted prompts.
- Provides quick responses, aiding rapid exploration of data trends.

### **Limitations:**

- Struggles with time-series analysis without historical data (e.g., 2024 stats).

- May offer overconfident or incomplete answers without specific guidance.
- Requires clear metric definitions for qualitative judgments like “most improved” or “game-changer.”

## **7. Final Reflections**

This project has been good learning into the capabilities of LLMs like Grok 3. I learned that while these tools are powerful for data summarization and basic queries, their effectiveness hinges on how I structure prompts—clear definitions and context are critical. The need to validate answers manually or with scripts reinforced the importance of human oversight, positioning LLMs as collaborative tools rather than standalone solutions. Creating visualizations also highlighted their value in communicating insights, especially for coach or team audience.