

Analyzing Labor Discrepancies in Campus Recreation Operations

APPALACHIAN STATE UNIVERSITY RECREATION

Jonathan Taylor
MBA & MSADA Dual Degree Candidate | SUMMER 2025

Table of Contents

1. Executive Summary.....	2
2. Organization Background.....	3
3. Business Problem / Context.....	3
4. Project Goals / Problem Scope.....	3
5. Literature Review.....	4
6. Literature Review & Best Practices.....	4
7. Data Collection & Challenges.....	5
8. Data Cleaning.....	6
9. Data Storage.....	6
10. Data Analysis & Tools.....	6
11. Results & Visual Discussion.....	7
12. Limitations & Challenges.....	9
13. Implications.....	9
14. Recommendations & Next Steps.....	10
15. Conclusion.....	11
16. References.....	12
17. Appendices.....	13

Executive Summary

This project analyzed labor discrepancies between scheduled and actual work hours within University Recreation at Appalachian State. Using data from WhenToWork (scheduling) and TimeClock Plus (timeclock records), I calculated hour-by-hour differences across 2,583 shifts during the 2024–2025 academic year.

The majority of staff worked in close alignment with their assigned shifts. However, 6% of shifts had discrepancies over an hour, and overall actual hours were 111.4 hours under scheduled hours, saving the department an estimated \$1,000-\$1,365 in labor costs. While some employees consistently clocked out late or early, the average discrepancy was only two minutes, signaling reliability overall.

Key trends emerged around operational realities, such as weekend game cancellations and event-based early departures. Discrepancy spikes aligned with known busy periods and weather disruptions. The strongest discrepancies appeared on Saturdays, with early departures in club sports roles.

The findings suggest actionable improvements in shift planning, budget forecasting, and system integration. I recommend linking cancellation logs, assigning unique staff identifiers, and implementing automated discrepancy reporting. This framework can evolve into a repeatable decision-support tool for improving scheduling precision and operational efficiency.

Organization Background

University Recreation (UREC) at Appalachian State University provides facilities, programs, and employment opportunities that support student well-being, engagement, and development. As part of its student employment structure, UREC hires a large number of student workers to manage facility operations and run intramural events throughout the academic year.

The student staff experience is designed to foster leadership and accountability. However, the administrative responsibility of scheduling, tracking hours, and ensuring accurate pay is complex. Supervisors rely on a combination of systems, such as WhenToWork for scheduling and TimeClock Plus for timeclock management. These tools are not automatically integrated, which means analyzing how scheduled and actual hours compare has typically been a manual or surface-level process. This project sought to close that gap.

Business Problem / Context

UREC Sports Programs employs over 50 student staff members a year, and they often scheduled for specific shifts related to intramural events and facility operations. While most shifts are covered as intended, there is limited visibility into how well actual worked hours align with scheduled time. Discrepancies may arise from early departures, game cancellations, shift trades, or extended responsibilities. Without a systematic way to quantify and explore these discrepancies, it becomes difficult to identify inefficiencies, spot trends, or make data-informed staffing decisions. Before the present project, UREC has not had a joined analysis of TimeClock and WhenToWork data.

This analysis focuses on identifying those gaps and patterns. The goal is not to evaluate individual performance but to understand where UREC's operations may benefit from improved tracking, planning, or support. The insights gained from analyzing timeclock discrepancies could be utilized to optimize budgeting, payroll accuracy, fairness of shift distribution, and leadership decisions.

Project Goals / Problem Scope

This project aimed to answer the following business questions:

- To what extent do actual employee work hours differ from their scheduled shifts?
- Are there patterns in discrepancies across dates, employees, or shift types?
- What proportion of shifts result in employees working more than scheduled versus less?
- Which employees consistently go over or under their scheduled time?

- How can this analysis support staffing decisions, shift planning, and operational adjustments?

By answering these questions, the project provides a clearer view of labor efficiency within UREC's Sports programs and identifies potential areas for improvement or follow-up inquiry.

Literature Review & Best Practices in Labor and Scheduling Analytics

Effective workforce management relies on timely, accurate scheduling and time tracking. In large operational environments such as university recreation programs, discrepancies between scheduled and actual hours can create challenges in budgeting, staff reliability, and service delivery. This literature review synthesizes key findings from human resource analytics, labor discrepancy studies, and scheduling optimization best practices, while connecting them to the analytical techniques used in the current project.

Data-Driven Human Resources

Workforce analytics is defined as the practice of utilizing data in HR decision-making and business outcomes. Madhani (2023) suggested that utilizing HR data to make decisions can create value for the organization, driving performance and effectiveness. Furthermore, Margherita (2022) stated that "human resource analytics is crucial to enhance people-driven performance within organizations." These frameworks support the significance of tracking HR data, such as labor data, in order to gain insights into optimization and performance.

In this project, a labor discrepancy analysis was conducted by joining scheduling data with timeclock records. This reflects the best practices in workforce analytics, where the alignment of assigned responsibilities and actual labor input is analyzed to gain insights and create operational improvements.

Labor Discrepancy Studies

In operations, discrepancies between scheduled and actual shift times have been found to affect payroll accuracy and productivity. Furthermore, labor overages can result in cost overruns, while consistent early clock-outs may reflect disengagement from employees or inaccurate scheduling practices. Studies suggest that these kinds of discrepancies can result in compounding inefficiencies over time (Cohen, 2024; Lu, et al. 2022). Furthermore, the research indicates that being able to accurately schedule according to need can help to drive profitability (Perdikaki et al., 2012).

Applying this to the current study on University Recreation, the observed discrepancies in actual and scheduled shift times reveal trends not just unique to individuals but to the system.

Understanding these patterns allows managers to respond proactively and adjust scheduling

practices, plan for frequent cancellations, and identify employees who routinely leave early or stay late.

Scheduling Optimization and Automation

Research has shown that manual scheduling often leads to inefficiencies and inconsistencies and that stable scheduling practices directly increase productivity outcomes (Williams et al., 2018). Data-driven reporting, alerts, and models can support scheduling accuracy as well as reduce the workload of department leadership. Moreover, using historical data to create predictive scheduling models can inform and improve scheduling practices (Alabi, et al., 2024).

Incorporating automated workflows that are driven by historical data aligns directly with these findings, and this project recommends that kind of implementation. Dashboards or reports that flag discrepancies can indicate to supervisors that there are issues in real time, enabling them to stay ahead of issues and make decisions based on data. This is much better than relying on anecdotal feedback or making reactive corrections (IndeedFlex, 2025). In addition, linking contextual data, such as game cancellation logs or weather patterns, can help to indicate which departures are legitimate as well as highlight inefficiencies.

This project applied foundational practices from the literature by integrating historical shift data, creating visualizations and making actionable recommendations. If this work were to be replicated, it would benefit from contextual information (e.g., by role or location) and being implemented into an automated system. Doing so would provide foresight into scheduling practices as well as strengthen the UREC Sports Program's ability to manage and monitor student labor efficiently and equally.

Data Collection & Challenges

The data for this project came from two primary sources: the WhenToWork scheduling system, and TimeClock Plus reports. Both sources served a distinct purpose in mapping the full picture of employee labor.

- **WhenToWork Data:** Exported datasets from the scheduling platform provided detailed records of each employee's assigned shift times, including their name, position, and hours scheduled per day.
- **Timeclock Plus Data:** Timeclock exports offered timestamped records of clock-ins and clock-outs, along with calculated shift durations by employee and position.

While each dataset was accessible independently, a major challenge involved aligning records across systems that do not share a unique identifier like a Banner ID. Because both systems used employee names, the join process depended on name matching, which introduced some uncertainty due to inconsistent formatting (e.g., full names vs first names only).

Additionally, the datasets were stored in different formats and sometimes contained redundant or irrelevant information. Other challenges included shifts with missing clock-ins or very short durations, which required additional cleaning and interpretation.

Data Cleaning

Data cleaning was done using a combination of SQL and Python. After importing each dataset into SQL Server, columns were renamed to match across systems, and relevant fields were selected for analysis. Records with missing names or hours were excluded, and shift durations were calculated where necessary. Dates were standardized across datasets to allow for reliable joins.

Python was used later in the process to conduct additional filtering and remove records with null values in calculated discrepancy fields. This ensured that statistical summaries and visualizations were based only on valid comparisons between scheduled and actual hours.

Throughout this process, care was taken to avoid over-cleaning. Rather than excluding negative discrepancies, the data was kept intact to ensure that early departures and potential cancellations could be included in the analysis.

Data Storage

All datasets were imported into a centralized database named “UREC_Analytics” using SQL Server Management Studio (SSMS). Within this environment, individual tables were created for each data source, and transformed versions were added as needed to support analysis.

Views and Common Table Expressions (CTEs) were used to create modular queries that joined scheduled and actual hours by employee and date. This structure made it easy to explore the data iteratively while maintaining a clean data model.

Once the final merged dataset was prepared, it was exported as a CSV and brought into Python for analysis and visualization. This hybrid setup allowed SQL to handle the heavy lifting of joins and aggregation, while Python provided flexibility for visuals and summary statistics.

Data Analysis & Tools

The analysis focused primarily on descriptive methods, with the goal of quantifying and visualizing the difference between scheduled and actual hours for each employee on each shift day.

Key calculations included:

- Total scheduled hours per person per day
- Total actual hours per person per day

- The difference between those two totals (hour discrepancy)

From there, I created additional views of the data to support comparisons over time, averages by employee, and frequency distributions of discrepancies.

The tools used were:

- **SQL Server** for data manipulation, joins, and summarization
- **Python (Pandas, Matplotlib)** for statistical analysis and visualizations
- **Excel** for early-stage filtering and data exploration

The decision to use SQL was driven by the structure of the original datasets and the need to merge multiple tables cleanly. Python was chosen for its flexibility in generating visuals and exploring statistical relationships that would be difficult to produce in SQL alone.

Results & Visual Discussion

The analysis produced a joined dataset containing scheduled and actual hours for each employee on each day they were assigned to work. By calculating the difference between these two values, I was able to examine both individual and group-level trends in how closely employees worked in alignment with their schedules.

Overall Patterns

2583 shifts from the 2024-2025 school year were analyzed. 1356 shifts ended with employees clocking out late, while 835 ended early (Table 1). Most of these early or late clock outs were within 5-10 minutes of the shifts scheduled ending time, but 6% of discrepancies were over an hour (Table 2).

Across the full dataset, the majority of discrepancies were relatively small. A histogram of hourly differences showed that most employees either matched their scheduled time or deviated by less than one hour. The distribution was slightly skewed to the right, indicating that overages (working more than scheduled) were more common than leaving early (Figure 1).

The overall actual hours worked were 111.4 hours less than the scheduled hours, which means that sports programs spent at the minimum \$1058.30 less than the originally budgeted amount based on scheduled hours and around \$1365 less at the maximum, depending on the positions that employees were working (Table 2). In terms of budgeting for labor, this means that Sports Programs may want to budget a built-in labor variance buffer of approximately \$1000 due to routine cancellations. Games being canceled can be viewed as some sort of inevitability, given that there is commonly bad weather or forfeits during scheduled seasons.

According to the scatterplot shown in figure 3, actual hours and scheduled hours are highly positively correlated, which indicates that most shifts had little to no discrepancy. Most shifts were not far off of their scheduled time, and if they were off, they weren't off by a large amount. However, there are a decent number of shifts that moderately stray from the distribution, and some major outliers that may be skewing the data. Overall, employees tend to stick to their shift times, but there may be a small number of employees that consistently stray by 10 minutes under or over. The averages straying at least 15 minutes, both positive and negative, from 0 can be attributed to the larger discrepancies, especially since the average discrepancy as a whole is only about 2 minutes.

Time-Based Trends

Time-based analysis revealed that discrepancies followed seasonal and operational patterns. Negative discrepancies often occurred during weekends or known game cancellation periods, while positive spikes aligned with busy programming windows. This supports the idea that many early clock-outs were likely driven by operational realities rather than employee disengagement.

As seen in figure 2, the cumulative discrepancies reached their lowest point in February, likely due to weather-related cancellations during the Helene tournament and winter season transitions. Throughout the season, the trend started to go upwards and hit a peak of only 40 hours under, just to trend down to a low of around 120 under again by the end of the year.

As shown in Figure 4, multiple sharp negative spikes occurred, suggesting mass early clock-outs, likely caused by event cancellations or shortened shifts. This could include days where a large number of games were canceled, training ended early, or shifts were entered incorrectly.

Shift discrepancies tend to vary by day of the week. Overall, no day of the week has a major discrepancy, as the biggest discrepancy is leaving 10 minutes early on Saturdays. On Saturdays, there are only club sports events shifts, so it can be inferred that events end early or employees tend to leave those events a little early consistently. Every other day only has slight discrepancies, which indicates that there was a bit more of a uniform trend and that the day of the week did not have a huge effect on whether or not employees were leaving late or early (figure 5).

Employee-Level Differences

Averages were calculated for each employee to identify who tended to work more or less than scheduled. 20 employees averaged a positive discrepancy, and the highest maximum overage was over by around 20 minutes. 29 employees averaged a negative discrepancy, with the lowest average underage being around 20 minutes under. No employee surpassed over a 30-minute average discrepancy throughout the year, indicating that on average students were reliable. The graph displaying the data by employee cannot be shared due to protecting their personal

information. While this graph is omitted to protect employee confidentiality, it provided key insights into consistency and variation in staff time use.

As a whole, the staff's average overall discrepancy was about 2 minutes under. The average for those who stayed later was around 15 minutes. The average discrepancy for those who left early was around 36 minutes (Table 1).

Some employees consistently staying later may reflect dedication, inefficient closing policies, leadership roles, or longer event durations. Others consistently clocking out early potentially signals shift gaps, consistent game cancellations, personal availability issues, or a need for better scheduling alignment.

Together, these results show that while the UREC staff as a whole performs reliably, there are opportunities to refine staffing processes and better understand where time is gained or lost.

Limitations & Challenges

Several limitations were encountered in this project. First, the lack of a unique identifier across scheduling and timeclock systems made joins more error-prone and required careful name-based matching. This introduces a risk of mismatches or exclusions when names were formatted inconsistently. To reduce these limitations, future projects could benefit from integrating a unique staff ID across scheduling and timeclock systems.

Second, the analysis did not incorporate contextual information about individual shifts. For example, it was not always clear when a discrepancy resulted from a canceled game, an extended match, or a shift handoff. As a result, the reasons behind time differences had to be inferred rather than directly verified.

Additionally, the dataset excluded employees with missing data or invalid shift records, which may have slightly skewed the results. There was also limited metadata about positions or facilities worked, so deeper segmentation by job type was not possible within the scope of this project.

Implications

The results of this analysis carry both operational and strategic implications for UREC:

- **Budgeting and Forecasting:** A consistent discrepancy margin of ~\$1000 suggests UREC should factor a small buffer into future labor budgets, especially in seasons prone to cancellations or early event conclusions.
- **Shift Planning and Reliability:** Identifying which days or roles (e.g., Saturdays, club sports shifts) tend to underperform allows managers to proactively adjust shift lengths or allocate backup coverage.

- **Staff Performance and Recognition:** Employees with consistently positive discrepancies (staying longer) may be strong candidates for additional responsibility, promotions, or recognition — while others may benefit from check-ins or schedule adjustments.
- **Data Infrastructure:** The successful integration of WhenToWork and TimeClock Plus highlights the value of joining datasets to uncover patterns. This approach can be expanded to track job type, facility usage, or event attendance.
- **Cultural Insight:** The analysis supports the idea that students are generally reliable, with only a small number of outliers. This helps reinforce a culture of trust while still identifying areas to improve.

Recommendations & Next Steps

To build on this work and improve operational accuracy:

1. Automate Monthly Reporting

Develop a monthly SQL or Python-based workflow to flag shifts with large discrepancies, summarize department-wide trends, and email reports to leadership. This will reduce manual tracking and support proactive planning.

2. Incorporate Contextual Metadata

Integrate game cancellation logs, weather data, or shift notes to explain recurring discrepancies and reduce noise in the analysis.

3. Assign Unique Identifiers Across Systems

Where possible, attach Banner IDs or another consistent field across both systems to allow cleaner joins and better long-term tracking.

4. Segment by Role or Facility

With additional data, segment shifts by role or location to assess whether certain facilities or job types have more variance in attendance.

5. Train Successors and Institutionalize Process

Package this workflow into a training resource for future GAs or professional staff. This could include example dashboards, SQL queries, and documentation to institutionalize labor analytics within UREC.

6. Explore Predictive Scheduling

Once multiple seasons of data are available, explore basic predictive models to forecast labor needs, especially during tournaments or seasonal changes.

Conclusion

Completing this analysis was a meaningful learning experience and has real-world potential to improve how student labor is managed. This analysis demonstrates how even basic data integration can uncover meaningful insights. By pairing scheduled and actual work data, UREC can better support its staff, refine its resource use, and set a precedent for data-driven management. With continued refinement, this project can serve as a foundational model for ongoing decision-making and continuous improvement.

References

- Alabi, O. A., Ajayi, F. A., Udeh, C. A., & Efunniyi, F. P. (2024). Predictive analytics in human resources: Enhancing workforce planning and customer experience. *International Journal of Research and Scientific Innovation*, 11(9), 149-158.
- Cohen, T. (2024). 10 employee scheduling issues + expert solutions. Connectteam.
<https://connectteam.com/common-employee-scheduling-problems-solutions/>
- IndeedFlex. (2025, March 10). Workforce planning with data: Predict future staffing needs and optimize resources.
- Lu, G., Du, R. Y., & Peng, X. (2022). The impact of schedule consistency on shift worker productivity: An empirical investigation. *Manufacturing & Service Operations Management*, 24(5), 2780-2796.
- Madhani, P. M. (2023). Human resources analytics: leveraging human resources for enhancing business performance. *Compensation & Benefits Review*, 55(1), 31-45.
- Perdikaki, O., Kesavan, S., & Swaminathan, J. M. (2012). Effect of traffic on sales and conversion rates of retail stores. *Manufacturing & Service Operations Management*, 14(1), 145-162.
- Williams, J. C., Lambert, S. J., Kesavan, S., Fugiel, P. J., Ospina, L. A., Rapoport, E. D., ... & Adler-Milstein, S. (2018). Stable scheduling increases productivity and sales: The stable scheduling study. *Report, Worklife Law, University of California Hastings College of the Law, San Francisco*.

Appendix

Table 1

	Avg_Discrepancy	Min_Discrepancy	Max_Discrepancy	Total_Records
1	-0.0431281155710113	-24.0999999046326	5.75	2583

	Avg_Overtime	Smallest_Overage	Largest_Overage	Overages
1	0.288542096879454	2.38418579101563E-07	5.75	1356

	Avg_Underworked	Largest_Under	Smallest_Under	Underworked_Records
1	-0.609288491986015	-24.0999999046326	-2.38418579101563E-07	825

Table 2

# Total Shifts	...	# % Discrepancies > 1 Hour	# Total Discrepancy Hours	# Estimated Labor Cost Im...
2583		6.08	-111.4	-1281.1

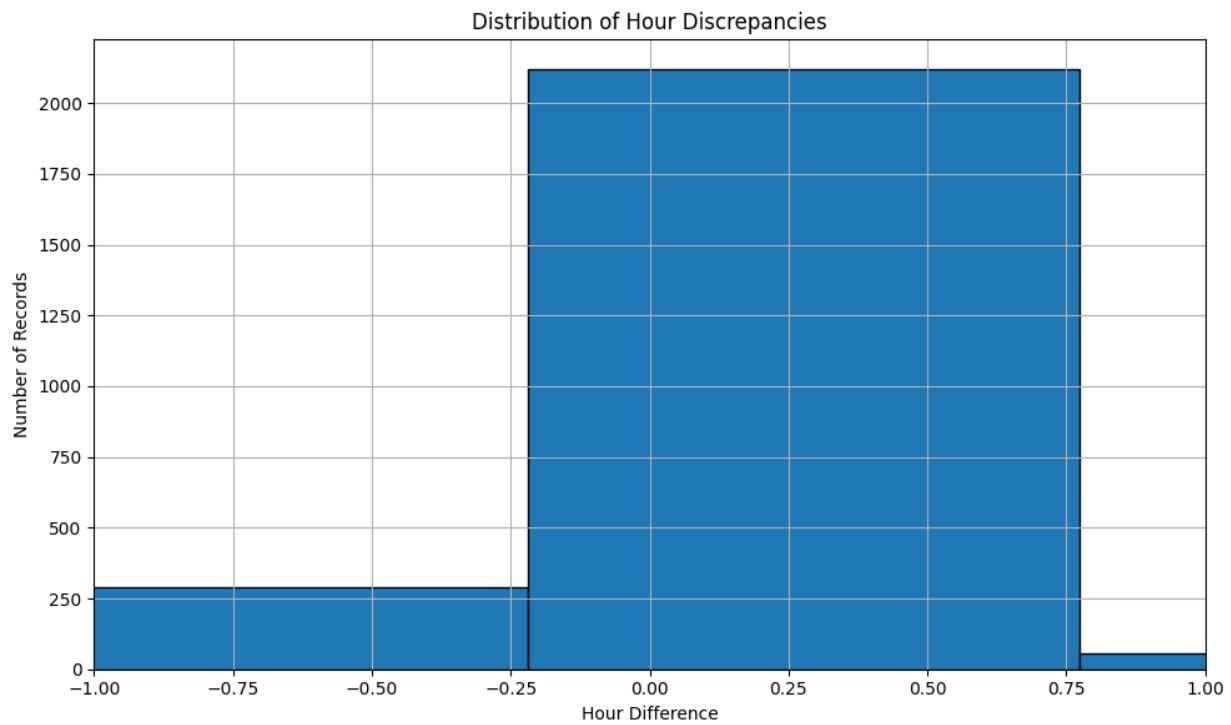


Figure 1

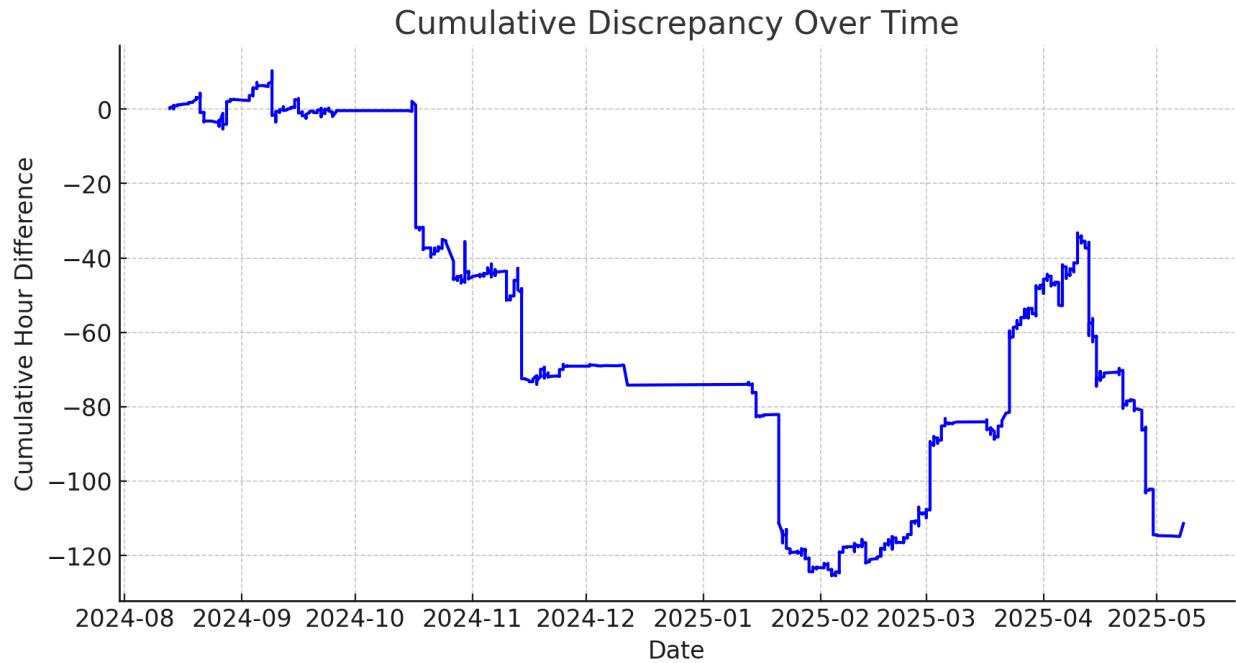


Figure 2

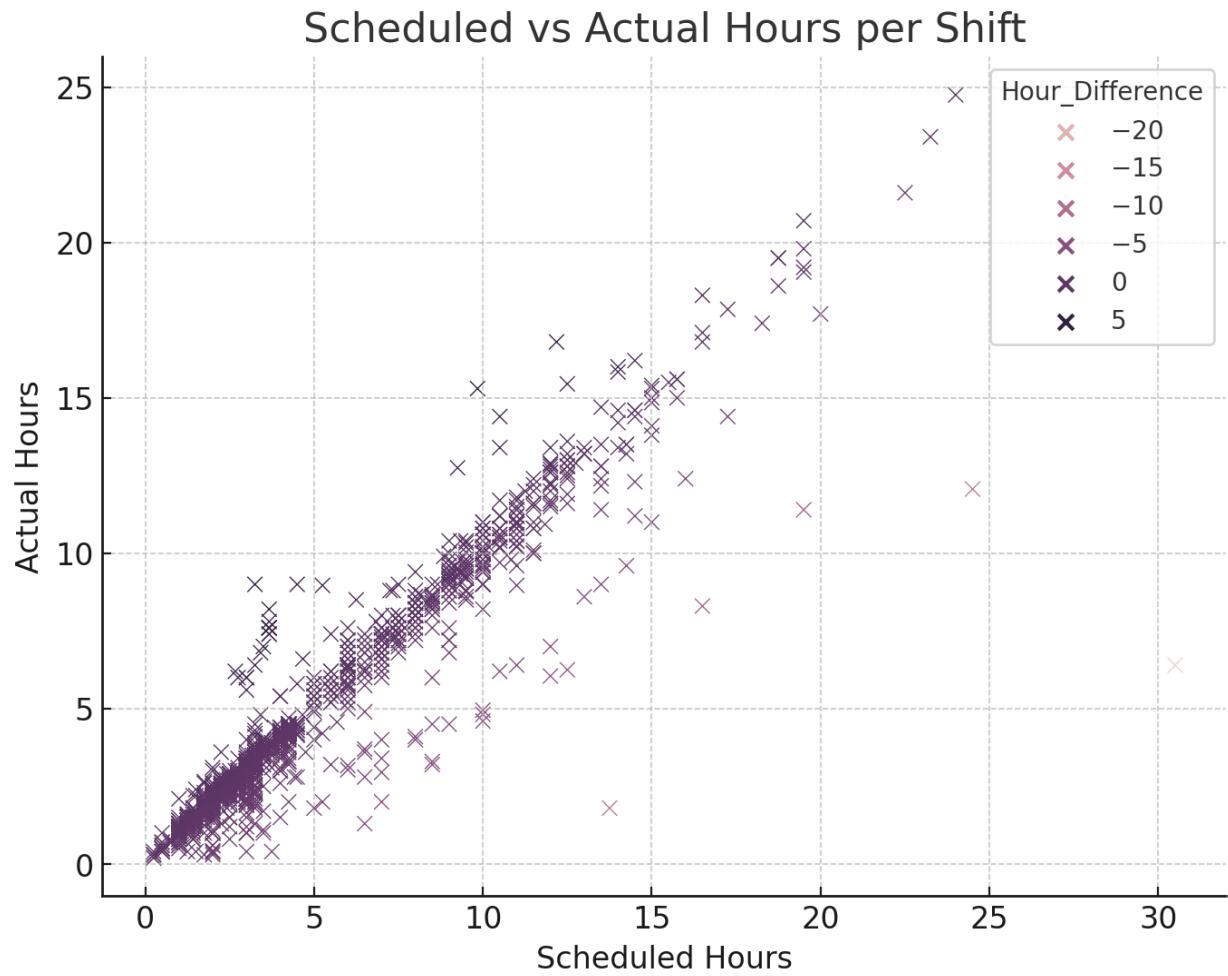


Figure 3

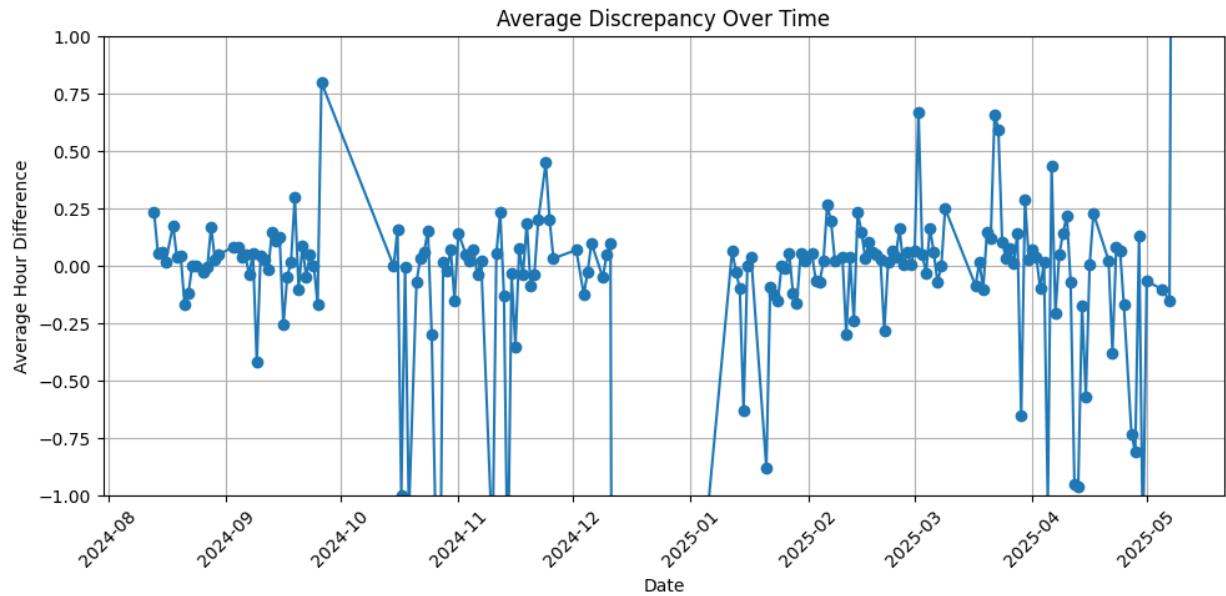


Figure 4

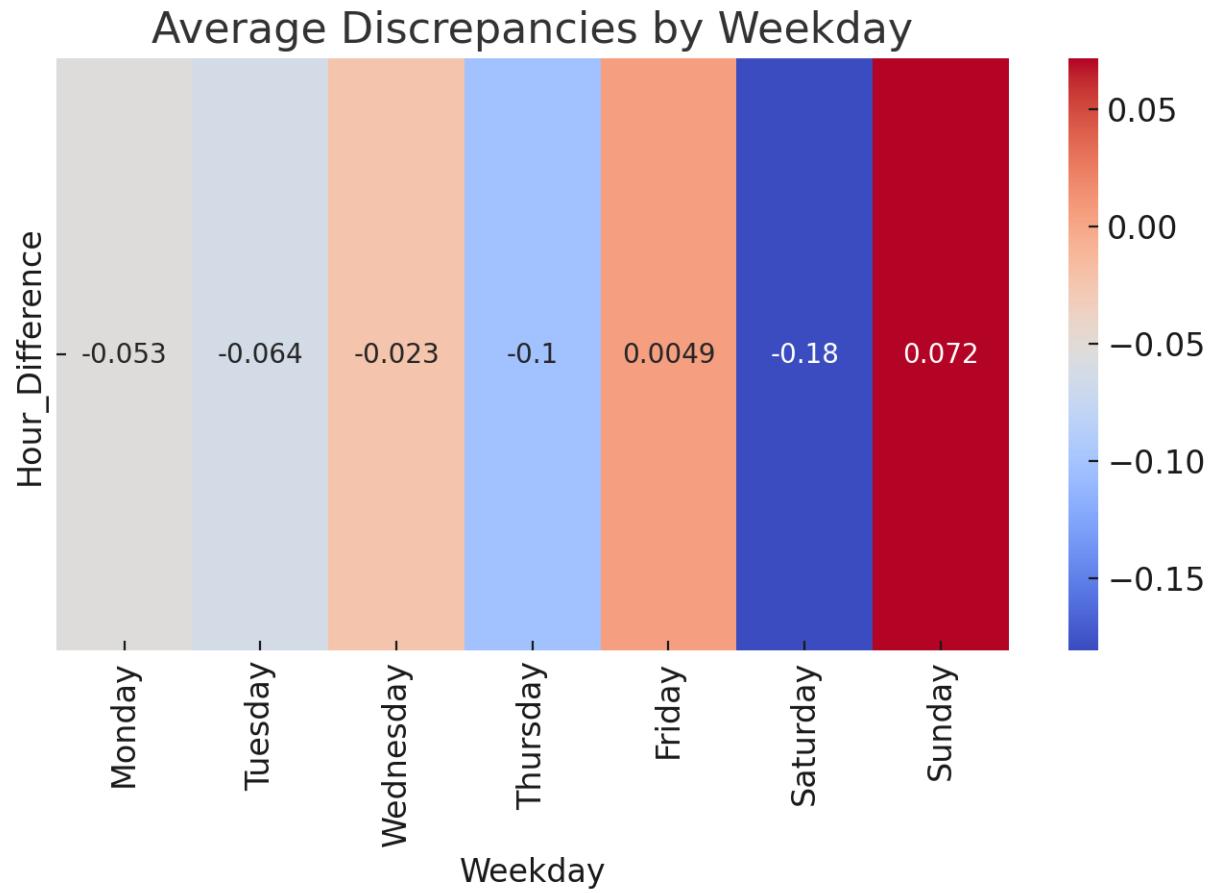


Figure 5