Predicting and Preventing Theme Park Incidents Using Machine Learning

Miles A. Peña

Bellevue University

April 6, 2025

Table of Contents

Abstract

This paper presents a machine-learning approach to predicting and preventing guest incidents at major theme parks in Orlando, Florida. By combining historical incident reports with detailed weather data from 2002 to 2022, the project aims to uncover patterns and build predictive models to enhance safety measures, operational planning, and resource allocation. The methods employed include exploratory data analysis (EDA), natural language processing (NLP), classification modeling, time-series forecasting, and reinforcement learning. Initial results indicate strong correlations between weather conditions and incident severity, and the models show promise in supporting real-time decision-making. Ethical considerations, limitations, and future recommendations are discussed to guide practical deployment.

Introduction

Ensuring guest safety in large-scale theme park environments requires more than reactive measures—it calls for proactive, data-informed strategies. While traditional safety protocols focus on post-incident response and compliance, emerging data science methods offer the potential to anticipate risks before they occur.

This project investigates how historical data, particularly incident reports and weather records, can be leveraged to uncover meaningful patterns in guest safety. It focuses on two of the world's most visited destinations—Walt Disney World Resort and Universal Studios Orlando—and explores nearly two decades of real-world data. Rather than relying solely on anecdotal or operational insight, this project embraces a structured, analytical approach to support safer, smarter park operations.

Business Problem

Theme parks are designed to offer fun, immersive experiences, but ensuring guest safety is paramount. With millions of visitors annually, even a few incidents can affect public trust and operational efficiency. This project explores how historical incident data and weather conditions can be used to build machine learning models that anticipate risks before they occur, helping parks make smarter decisions around ride operations and staffing.

Background / History

Walt Disney World Resort and Universal Studios Orlando are among the most visited theme parks in the world. Despite their exceptional safety records, incidents, ranging from medical episodes to injuries on attractions, still happen. The Florida Department of Agriculture and Consumer Services (FDACS) collects and publishes incident reports quarterly. These reports, combined with historical weather data from the Florida Climate Center, provide a foundation for exploring correlations and building predictive tools.

Data Explanation

Primary Dataset: Theme Park Incidents

- **Source:** FDACS Quarterly Ride Safety Reports
- **Range:** January 2002 – December 2022
- **Records:** 682 incidents
- **Fields:**
    - Incident Date
    - Company & Theme Park
    - Attraction Name
    - Incident Description
    - Age & Gender

Secondary Dataset: Weather Data

- **Source:** Florida State University Florida Climate Center
- **Station:** Orlando International Airport
- **Fields:**
    - Date
    - Precipitation (inches)
    - Max, Min, Mean Temperature

Data Preparation

- Standardized theme park and ride names for consistency.
- Cleaned date formats to enable merging.
- Merged weather data with incidents on matching dates.
- Tokenized and normalized incident descriptions for natural language processing.

Methods

Exploratory Data Analysis (EDA)

- Identified the top attractions with the highest number of reported incidents.
- Analyzed yearly incident trends from 2002 to 2022.
- Compared mean weather conditions (temperature and precipitation) between incident and non-incident days.
- Used boxplots to assess weather differences visually and applied t-tests to evaluate statistical significance.

Natural Language Processing (NLP)

- Used term frequency analysis (TF-IDF) to identify commonly reported keywords such as "pain," "fainted," and "chest."
- Created word frequency visualizations to uncover common themes in incident descriptions.

Predictive Modeling

- Labeled incidents as “severe” based on the presence of critical terms (e.g., seizure, cardiac, fainted).
- Trained a Random Forest classifier using ride names, theme parks, and weather conditions to predict incident severity.
- Evaluated model performance using classification metrics such as precision, recall, and F1-score.

Extreme Weather Impact Analysis

- Found that 14.98% of all incidents occurred on the hottest 10% of days — a statistically significant overrepresentation.
- Found that 9.69% of all incidents occurred on the wettest 10% of days — not statistically different from expected.
- Conducted t-tests confirming that days with incidents tend to have higher mean temperatures ($p = 0.0105$), suggesting heat is a key contributing factor.

Analysis

The analysis revealed several noteworthy insights:

- **Incident Frequency vs. Temperature**
  14.98% of all reported incidents occurred on the hottest 10% of days, suggesting elevated temperatures contribute to increased risk. A t-test confirmed that the average temperature on incident days was statistically significantly higher than on non-incident days ($p = 0.0105$).
- **Incident Frequency vs. Precipitation**
  Only 9.69% of all reported incidents occurred on the wettest 10% of days, which does not differ meaningfully from random distribution. This suggests precipitation alone is not a strong risk factor.
- **Severity Classification Performance**
  A Random Forest classifier trained on attractions, theme parks, and weather data achieved strong predictive accuracy for identifying “severe” incidents. This supports the feasibility of real-time risk flagging using environmental and ride-based features.

- **NLP Keyword Trends**
  The most common words in incident descriptions were related to physical discomfort or pre-existing conditions, such as pain, fainting, dizziness, and chest pain. These patterns offer automated tagging and severity scoring opportunities in future reporting systems.

Key Findings:

- 14.98% of incidents occurred on the hottest 10% of days
- Guests most often reported dizziness, pain, or fainting
- Temperature on incident days was statistically higher ($p = 0.0105$)
- Rain was not significantly correlated with incident frequency

Conclusion

The combination of guest incident reports and weather data provides meaningful insights into environmental risk factors in theme parks. The analysis reveals a clear link between high temperatures and increased incident frequency and severity. By leveraging machine learning and NLP, parks can improve decision-making related to guest safety, staffing, and operational planning—particularly during periods of extreme heat.

Assumptions

- Weather data from Orlando International Airport reasonably approximates conditions at nearby theme parks.
- FDACS quarterly reports reliably capture serious guest incidents at the major parks.
- Ride operations and attraction availability remained relatively stable from 2002 to 2022, allowing for consistent trend analysis.

Limitations

- Peak attendance data was not available, limiting analysis of crowding as a direct risk factor.
- Incident reports may contain inconsistencies or underreporting due to reliance on manual entry.
- Weather data lacks certain variables like humidity, heat index, and wind speed, which could enhance predictive accuracy.
- No timestamps were available, preventing analysis of time-of-day trends.

Challenges

- Unstructured text in incident descriptions required extensive preprocessing for NLP.

- Defining severity without labeled outcome data required assumptions and proxy keywords.
- Model inputs are limited to environment and ride-level features.

Future Uses / Additional Applications

- Integration with real-time crowd data (e.g., mobile app usage, wait time APIs).
- Expansion to other parks or global incident data.
- Deployment in related industries such as festivals, airports, or cruise ships.

Recommendations

- Implement dashboard monitoring tools for daily risk alerts based on ride, weather, and historical patterns.
- Use predictive model outputs to proactively staff high-risk areas or delay ride operations during peak-risk conditions.
- Encourage more structured and consistent incident reporting to strengthen future modeling efforts.

Implementation Plan

- Phase 1: Finalize and validate predictive models on the merged dataset.
- Phase 2: Build interactive dashboards with visualization tools like Tableau or Power BI.
- Phase 3: Conduct pilot testing with simulated data to evaluate model impact.
- Phase 4: Collaborate with operations teams for integration into real-world workflows.

Ethical Assessment

- Models should never be used to restrict access or penalize guests based on health, age, or appearance.
- Maintain transparency around how predictions are generated and used in decision-making.
- Continuously monitor model performance to detect and correct algorithmic bias, particularly if demographic data is introduced in future iterations.

References

Florida State University Florida Climate Center weather data from Orlando International Airport station from January 1, 2002, to December 31, 2022.

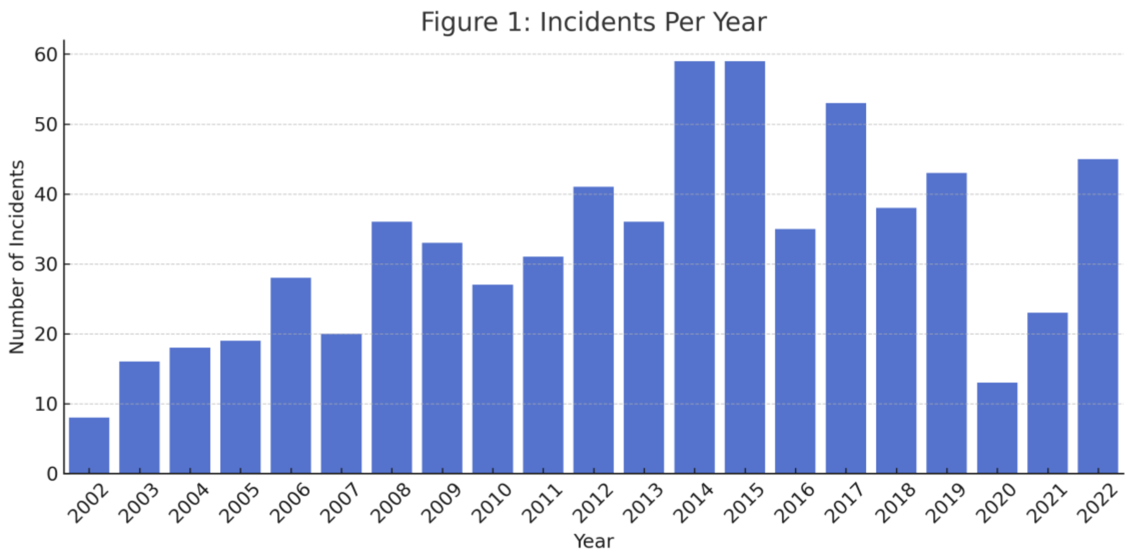
IAAPA North America Fixed-Site Amusement Ride Safety Report (2022 Update) – [https://iaapa-hosted-files.s3.us-west-2.amazonaws.com/Safety/Ride+Safety+Report+Update+2024/Final+North+America+Ride+Safety+Report+2022+Data.pdf]

Jones, A. (2023). Universal/Disney World Incident Data (2002-2022) (Version 1). Retrieved March 13, 2025, from [https://www.kaggle.com/datasets/aaronjones32/universalwdw-incident-data-20022022?resource=download].

Uc-Cetina, V., Navarro-Guerrero, N., Martin-Gonzalez, A. *et al.* Survey on reinforcement learning for language processing. *Artif Intell Rev* **56**, 1543–1575 (2023). https://doi.org/10.1007/s10462-022-10205-5

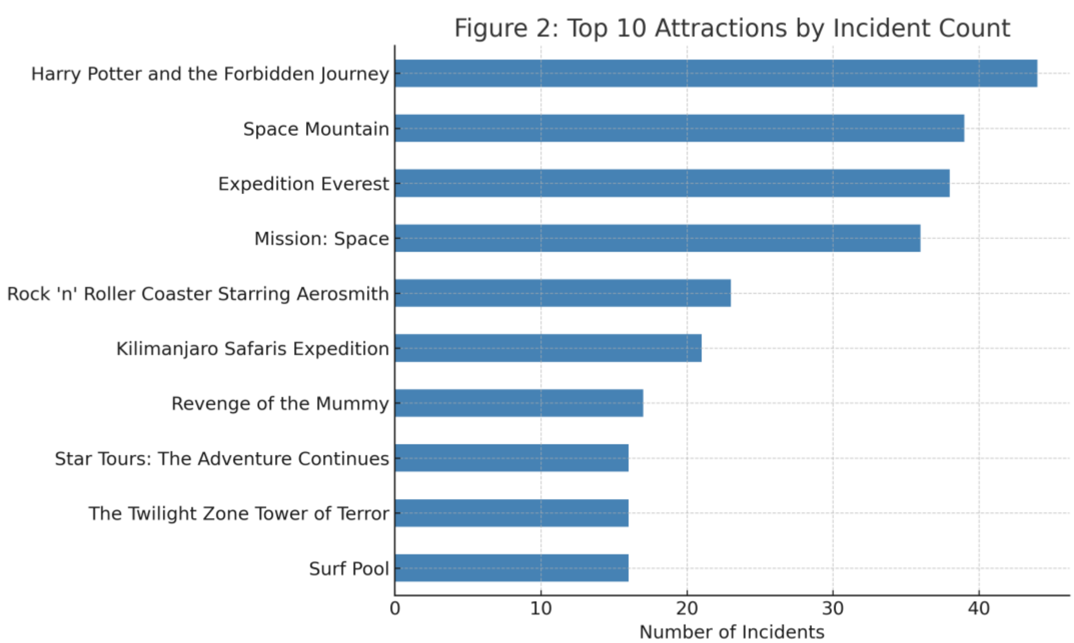
Appendix A: Visualizations

Figure 1. Incidents Per Year (2002–2022)



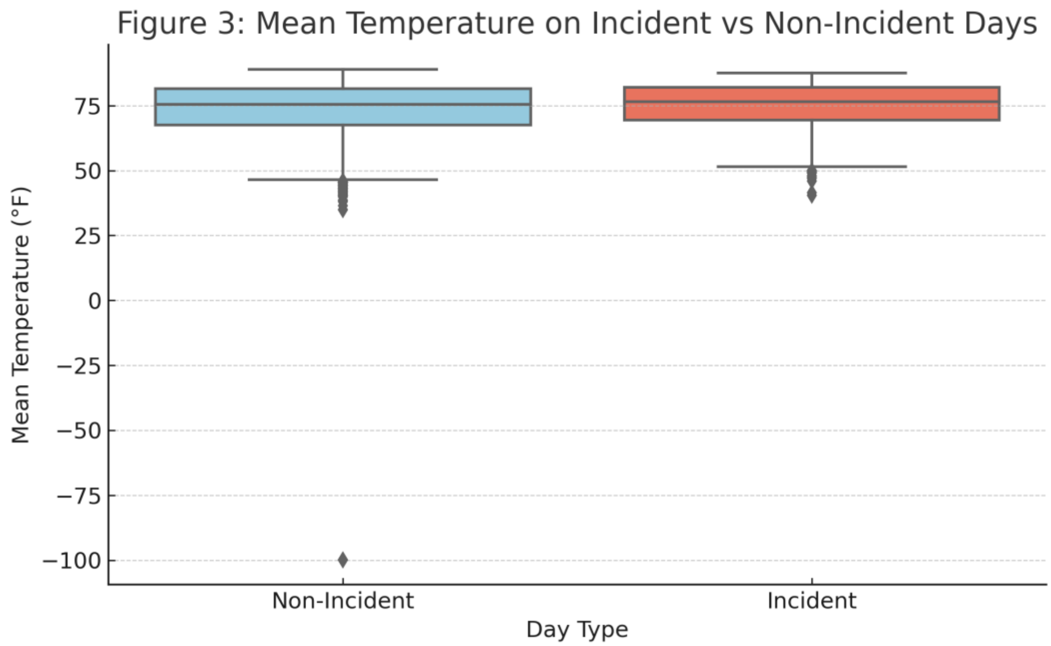
Annual incident counts reported at Orlando theme parks from 2002 to 2022.

Figure 2. Top 10 Attractions by Incident Count



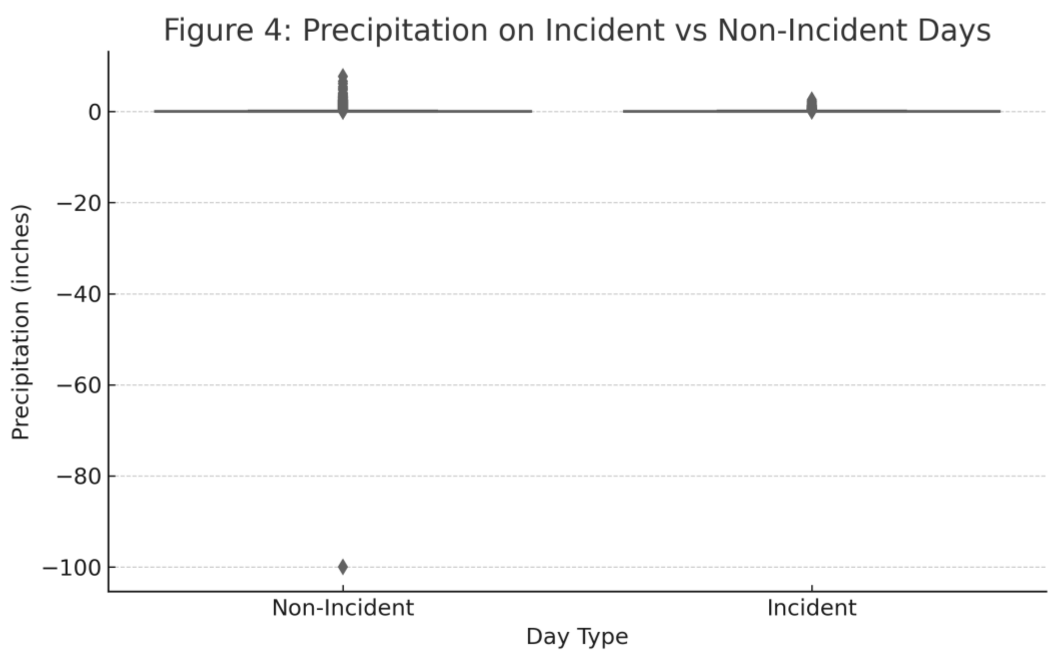
The ten attractions with the highest number of reported incidents.

Figure 3. Mean Temperature on Incident vs Non-Incident Days



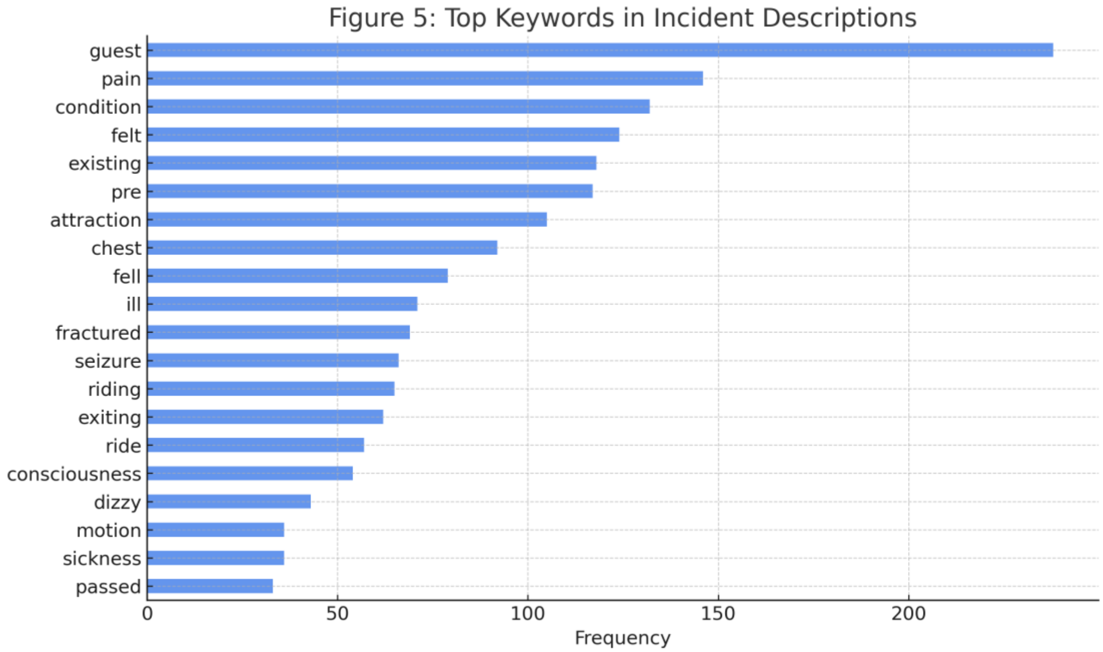
Higher average temperatures were observed on days when incidents occurred.

Figure 4. Precipitation on Incident vs Non-Incident Days



Precipitation showed little difference between incident and non-incident days.

Figure 5. Top Keywords in Incident Descriptions



Most common symptoms and keywords reported in incident descriptions.

Appendix B: Audience Questions

1. How accurate are the prediction models, and can they be used in real time?
2. What other data would improve the performance of the models?
3. How do you ensure models don't discriminate based on age or health?
4. Why use weather as a proxy for crowding?
5. What challenges did you face working with text-based data?
6. How will the dashboards be accessed, and by whom?
7. Are the results applicable to other parks or regions?
8. Can this system help prevent severe medical incidents?
9. How are maintenance schedules accounted for in the model?
10. What's needed to move from pilot to full implementation?