Optimizing Guest Experience: Data-Driven Forecasting & Capacity Strategies for Euro-Park

Gong Hei Fat Choy Consulting (Team #4)

Prashant ARYA, Piangpim CHANCHARUNEE, Ruxi HE,
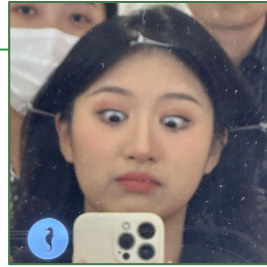I-Hsun LU, Hanqi YANG, Bowei ZHAO

February 14, 2025

The team

Piangpim is a master's student in data science and business analytics with experience in management consulting.

Hanqi is a master's student in data science and business analytics with experience in data visualization for business management.

Bowei is a master's student in data science and business analytics with experience in pricing and sales forecasting.

Ruxi is a master's student in data science and business analytics with experience in machine learning modeling and forecasting.

Prashant is a master's student in data science and business analytics with experience in engineering.

I-Hsun is a master's student in data science and business analytics with experience in data analytics and optimization.
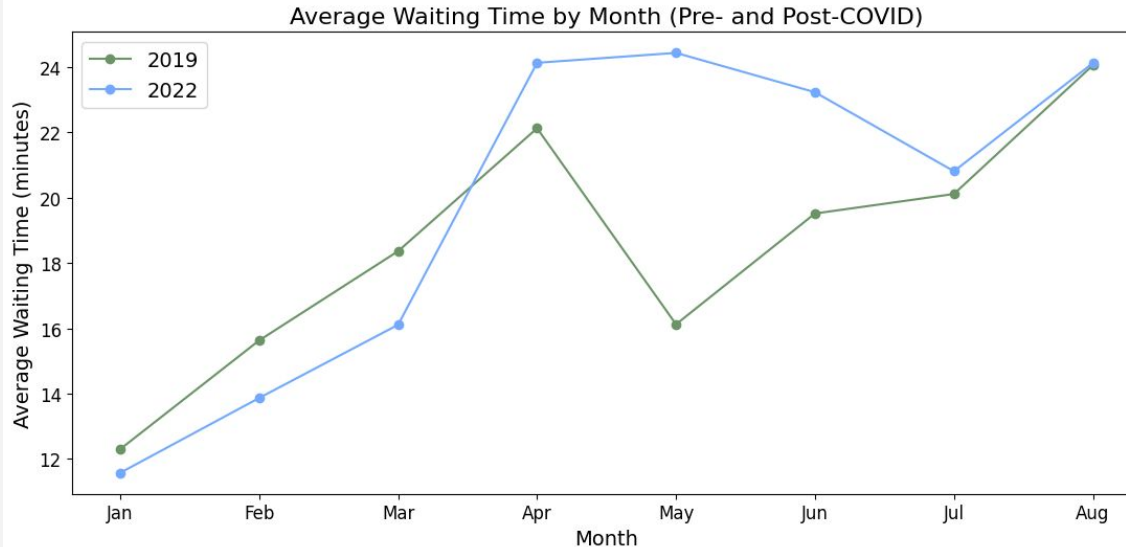
1. Problem Description

Waiting times have significantly increased post-COVID, resulting in negative impacts on customer satisfaction as well as efficiency and effectiveness of operations

Attendance decreased by almost 10% in 2022 compared to 2019, but waiting times have increased significantly…

…leading to negative implications in four aspects

Average Waiting Time by Month (Pre- and Post-COVID)

	Internal	External
Financial	*Resource Inefficiency*	*Guest Spending*
Non-financial	*Operational Strain*	*Guest Satisfaction*

Four main negative implications of increased waiting times can be classified into four categories

Resource Inefficiency

Under-or over-utilized resources, e.g. staff and capacity

> Wasted or increased payroll costs

> Increased per-unit operating costs

Guest Spending

Decrease in guest spending

> Decreased customer retention rate and/or lost opportunities to competitors

> Reduced time for in-park spending, e.g. food and merchandise

Operational Strain

Less efficient crowd control

> Risk of increased staff turnover due to staff strain

> Overcrowding and safety issues

Guest Satisfaction

Decrease in guest satisfaction

> Increased complaints to staff, increasing workload

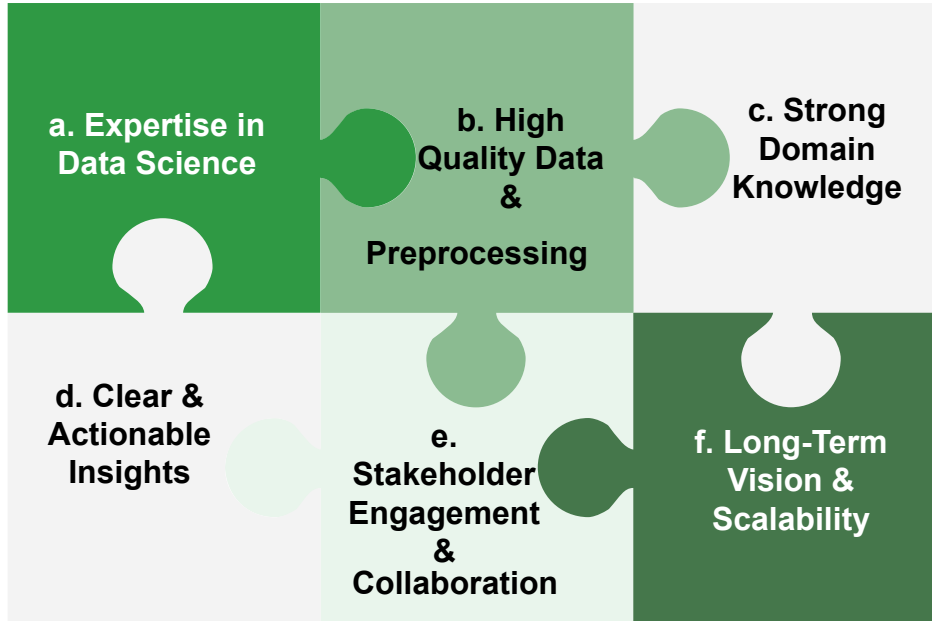> Negative park reviews and impacts on brand reputation

Proposed solution is to build a dashboard displaying more accurate predictions to derive key insights and facilitate decision making

Phase	Key Activities	Mon	Tue	Wed	Thur	Fri
Part 1: Data Inspection and Modeling	Study data and perform data cleaning	■				
	Discussion and agreement on methodology, strategies, evaluation metrics, and performance indicators	■				
	Implement and test chosen models, including performing parameter optimization		■	■		
	Improve models as needed			■		
Part 2: Dashboard Creation and Insight Gathering	Design layout of dashboard, such as determining relevant features and KPIs	■				
	Develop full dashboard using model developed in Part 1		■	■	■	
	Retrieve insights from dashboard to develop actionable solutions				■	
	Develop and present final report on methodology and insights				■	★

Deliverables submission

6

There are six key factors to ensuring successful implementation

a. Expertise in Data Science
b. High Quality Data & Preprocessing
c. Strong Domain Knowledge
d. Clear & Actionable Insights
e. Stakeholder Engagement & Collaboration
f. Long-Term Vision & Scalability

a. Expertise in Data Science: Our experienced team is fully committed to developing the right model adapted to key variables with continuous improvement.

b. High Quality Data & Preprocessing: We have ensured that the data is reliable, clean, and complete.

c. Strong Domain Knowledge: Our team is comprised of members with industry experience and extensive domain knowledge.

d. Clear & Actionable Insights: Not only do we provide predictions, but also actionable recommendations and visualization tools to add business value as well.

e. Stakeholder Engagement & Collaboration: We emphasize the importance of fully understanding stakeholder goals, frequent feedback, and constant communication.

f. Long-Term Vision & Scalability: We ensure that our predictions can be used in multiple areas and business functions such as marketing, process optimization, and pricing as well as other parks.

2. Methodology

KPIs have been identified to potentially address these implications

	Decrease in guest spending	Decrease in guest satisfaction	Under-or over-utilized resources	Less efficient crowd control
Average Wait Time (mean of WAIT_TIME_MAX)	🏰	🏰		
Peak Wait Time (max of WAIT_TIME_MAX)	🏰	🏰		🏰
Attraction Capacity Utilization (ratio of guests carried and capacity)		🏰	🏰	🏰
Peak Hour(s) (maximum of average wait times)	🏰	🏰		🏰

1.1 Model Development (1/2)

Data Cleaning & Assumptions

Feature	Assumption
Parade/Night show	• No show on dates not in parade_night_show • Every show is 15 minutes long
Opening times	• Ride is open if not on entity_schedule.csv
Up/Down Times	• **"Normal operation"** means no down times
Dates	• Exclude entries during COVID (2020-03-14 to 2021-06-10) • Cut off date at 2022-07-26 to align with attendance.csv

Attendance Prediction

- **XGBoost** with lag features, rolling windows, exponential moving average
- **Features:** year, month, day, weekday, dayofyear, weekofyear
- **Findings:** most important factors are **weekly trends** and whether the day is a **weekday/weekend**

XGBoost Model - Attendance Prediction for Next 60 Days

— Actual Attendance
--- Predicted Attendance (Next 60 Days)
▨ Excluded Period
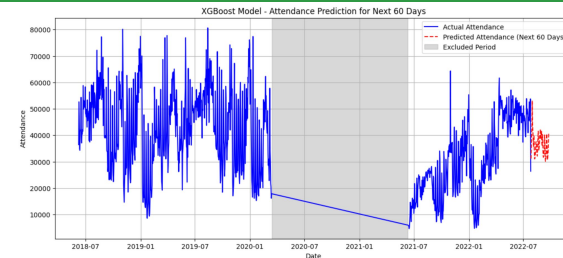
1.1 Model Development (2/2)

Legend:
- 🟥 Predicted
- 🟦 Test

Model #1: LSTM

⚠️ Advantages

Deep learning model which can make accurate predictions by inputting multiple features
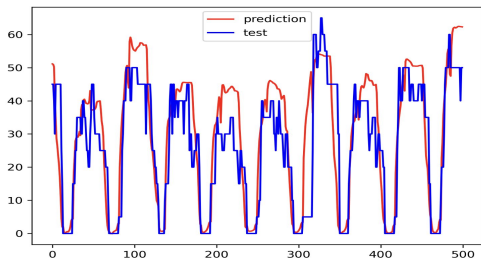
❓ Disadvantages

Requires knowledge/predictions of features in the targeted time frame

Results & Conclusion

Error: 10 minutes

✔ Can produce highly accurate results; however, requires input of external features

Model #2: VAR

⚠️ Advantages

Time series model which can make predictions without external features; able to capture dependencies of one entity on another
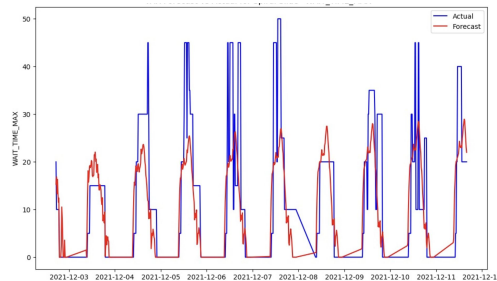
❓ Disadvantages

Assumes linear relationships between variables; may not capture complex and nonlinear dynamics
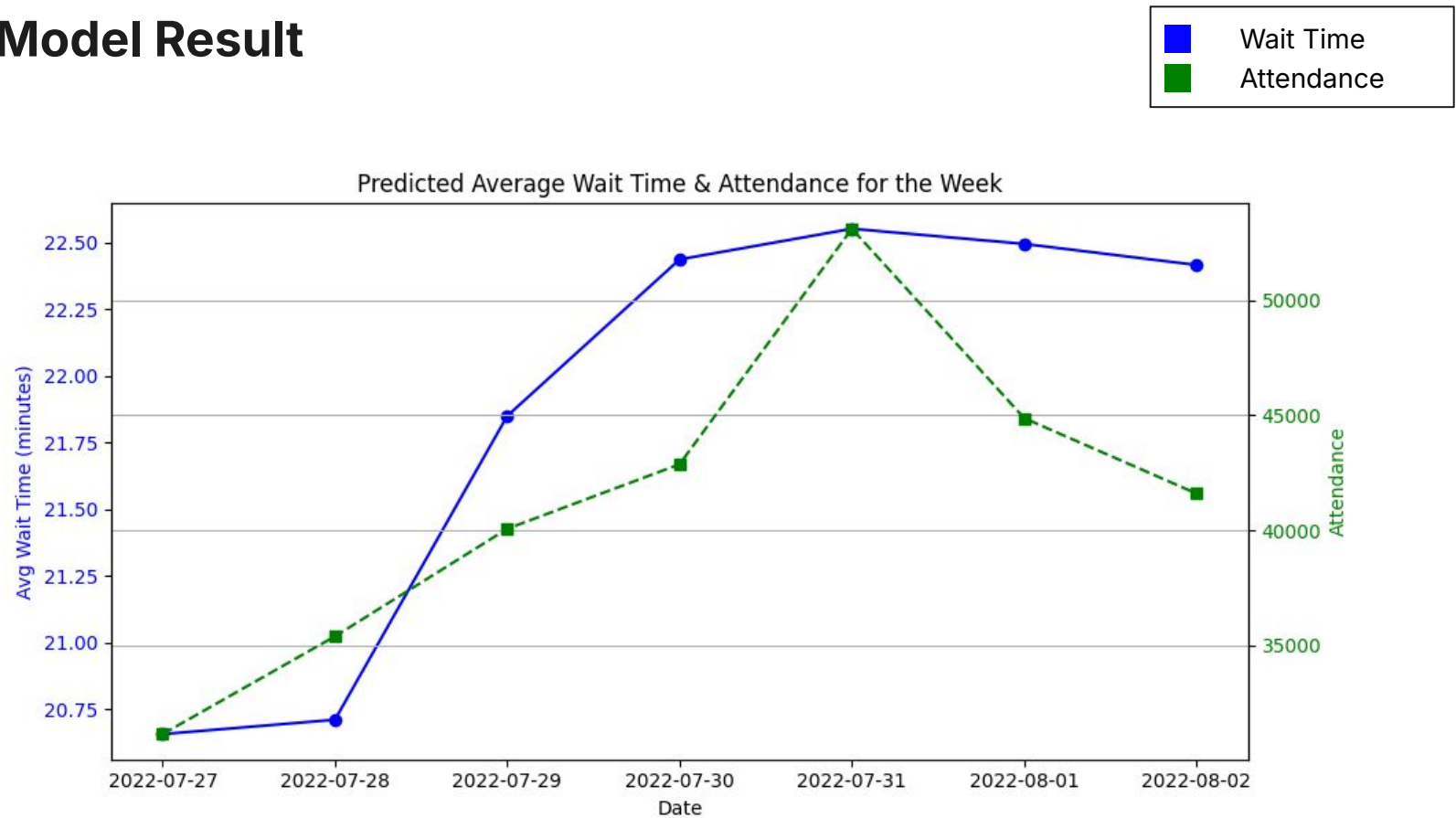
Results & Conclusion

Error: 11 minutes

✔ Can produce accurate results solely based on features at hand; **chosen for its suitability to the situation**

Sample graph based on "Spiral Slides" ride, global mean average error.

1.2 Model Result

Predicted Average Wait Time & Attendance for the Week

2. Dashboard

Scan the QR code or use the link for the demo!

https://crew-malaysia-mambo-routes.trycloudflare.com

3. Initiatives & Solutions

There are three main initiatives to address the identified KPIs

	Optimize Capacity	Personalize Notifications	Maximize Queue Revenue
Initiative	**Optimize Capacity**	**Personalize Notifications**	**Maximize Queue Revenue**
Description	Ensure that the capacity of each ride is always at 85%*.	Allow guests to be notified when certain rides have low wait times.	Deploy food or merchandise carts to attractions with long waiting times.
Impact	✔ Improved operating cost management ✔ Enhanced guest flow management ✔ Ideal capacity utilization	✔ Elevated guest experience ✔ Increased time for guest spending	✔ Transformed idle time into revenue streams ✔ Higher revenue per visitor
Related KPIs	3	1	2 4
Related Issues	1 2 4	3 4	2 3

Issues Identified

1	Less efficient crowd control
2	Under-or over-utilized resources
3	Decrease in guest spending
4	Decrease in guest satisfaction

KPIs Utilized

1	Average Wait Time
2	Peak Wait Time
3	Attraction Capacity Utilization
4	Peak Hour(s)

*Assumption for purpose of calculation only. Subject to change.

Next steps

Prototype Development (0 - 3 months)

- Implement and test prototype with additional data, e.g. yearly
- Gather feedback

Iteration & Optimization (3 - 6 months)

- Refine model based on feedback
- Implement improved prototype
- Financial modeling

Scalability & Stakeholder Alignment (6 - 12+ months)

- Add features to model, e.g. unexpected events
- Add features to dashboard, e.g. staff allocation optimization
- Constantly align with stakeholders on needs and use cases through ongoing engagement

Thank you

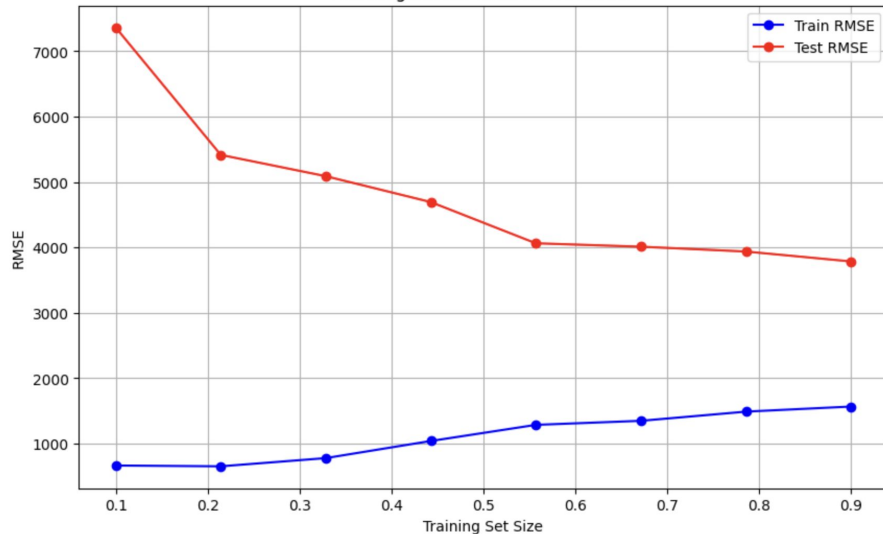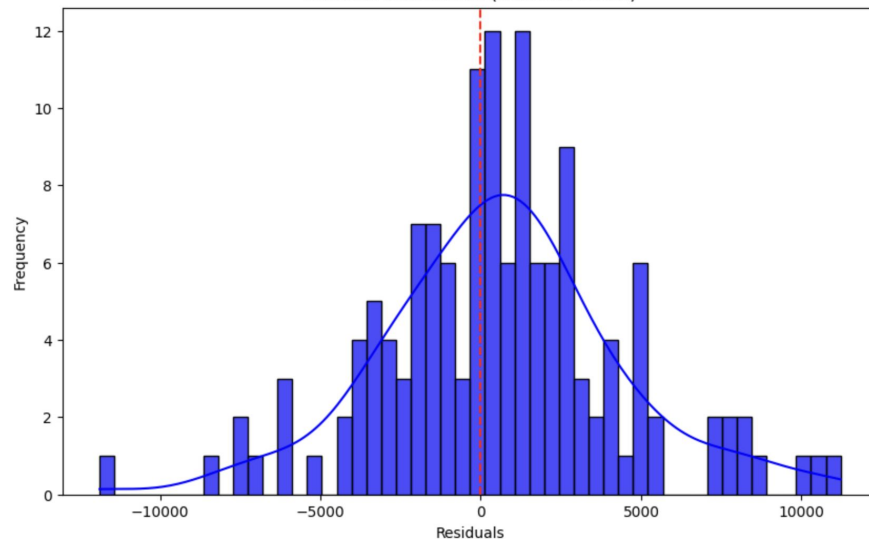- https://www.linkedin.com/in/prashantarya01250502/
- https://www.linkedin.com/in/piangpim-chancharunee/
- https://www.linkedin.com/in/ruxi-he/
- https://www.linkedin.com/in/i-hsun-lu/
- https://www.linkedin.com/in/hanqi-yang-0064431b2/
- https://www.linkedin.com/in/bowei-zhao-ph-d-59658a80/

Appendix

Evaluation of Attendance Model
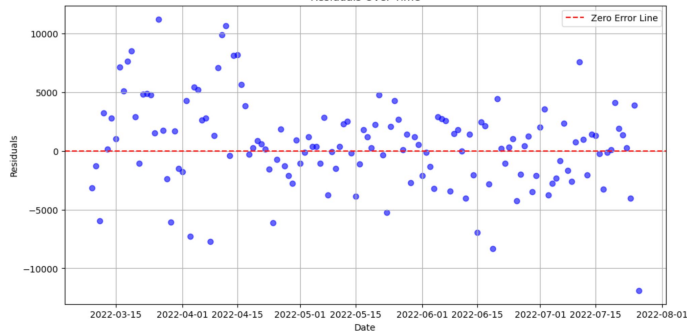
Evaluation of LSTM Model
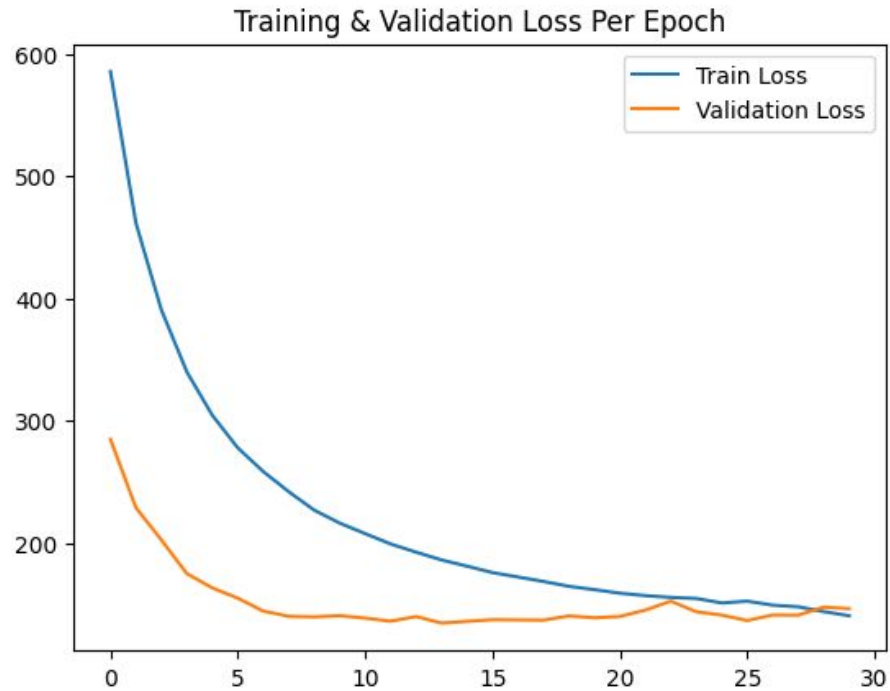
Comparison of Attendance in 2019 and 2022

Attendance Pre- and Post-COVID
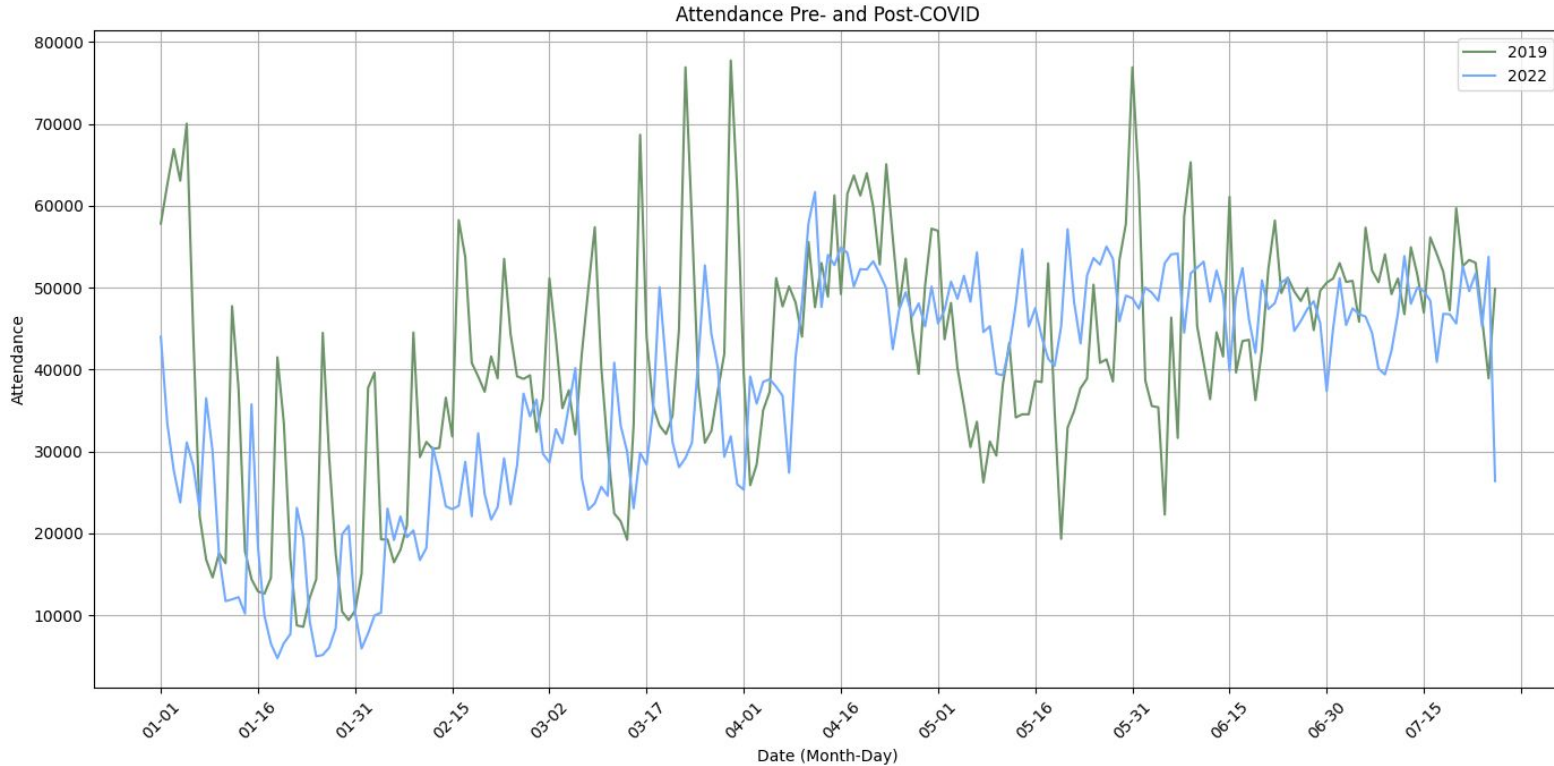
Net change in attendance: -9.87%

Data Cleaning Steps

Issues	Solutions
Negative values in columns	• Calculate moving averages • Replace values with zeros • Replace values with previous valid values
Dataset requires certain assumptions to be made	• If a date is not in parade_night_show, assume no show on that date • If no entry for a ride in entity_schedule, assume open • Assume every show is 15 minutes long
Certain values must be excluded to assume **"normal operation"**	Exclude entries during COVID period (2020-03-14 to 2021-06-10) and disregard UP/DOWN TIMES
Dates on attendance.csv and waiting_times.csv do not end at same time	Cut the data to attendance.csv end date (2022-07-26)

1

Context & Objectives: what do you understand from the project? Why is the context very complex?

✓ Present your understanding of the situation
✓ Include the process, the needs, the potential ROI, etc.

➢ c. 2/3 slides

2

Key Success Factors: what would be the reason of your success?

KSF	WHY?	HOW?
Business oriented		
Hands-on approach		
etc.		

✓ Explain the reasons of success: strong interaction with client? Deep expertise in AI?

➢ c. 1 slide

3

Proposed methodology: what could you do for the potential client?

Roadmap and touchpoints

Step 1 → Step 2 → Step 3 → Step 4 → Step 5
Workshop   Management meeting

Proof of concept

Code   Dashboard   Analyses

✓ What you could do, how, what you have already done, any relevant insight to prove your expertise and how much time do you need for the mission!

➢ c. 6/8 slides

4

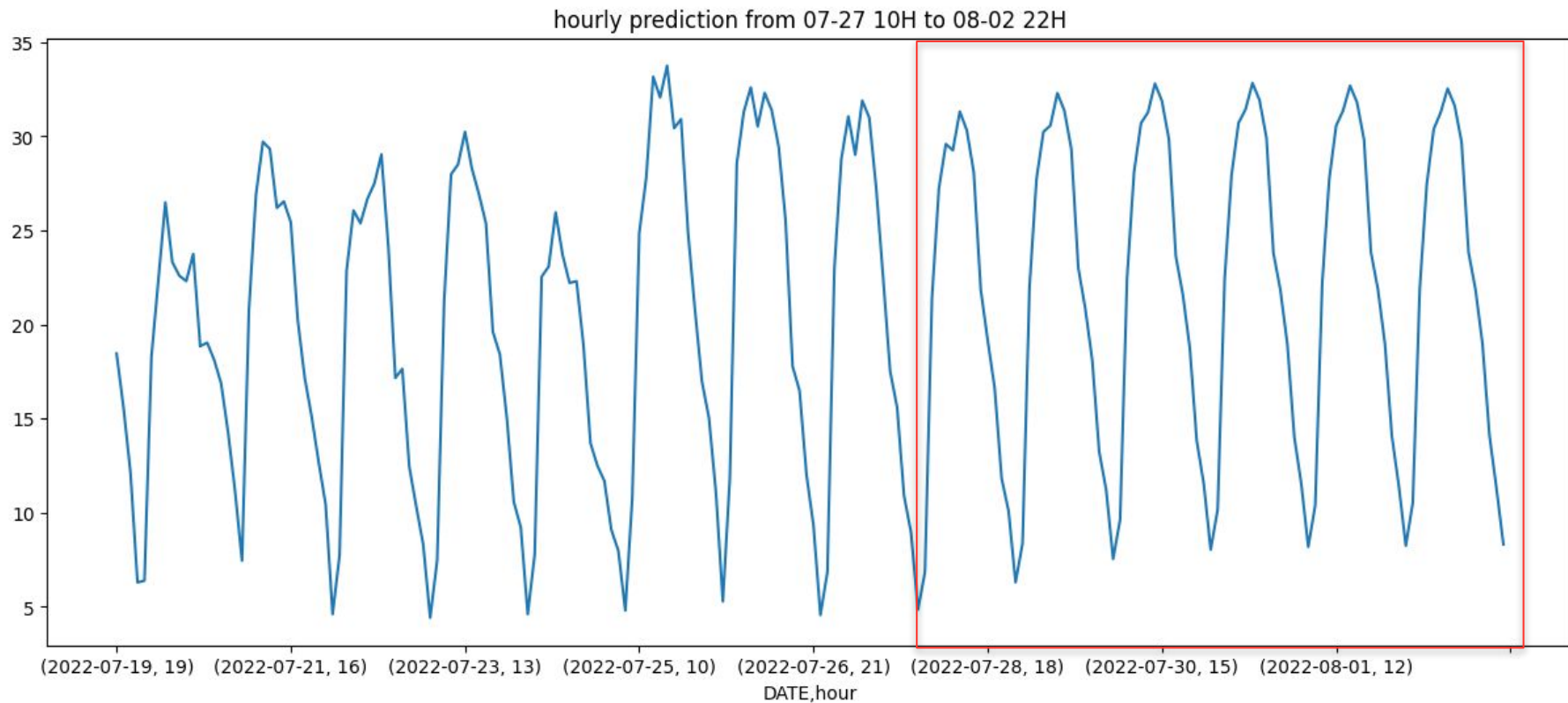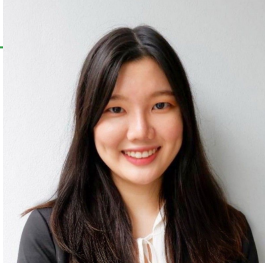**Team**: who are you and what have you done? Why are you relevant to solve the case?

✓ Introduce yourself, your background and explain why you are relevant for this mission

➢ c. 1 slide

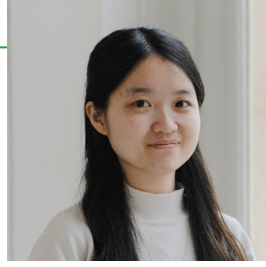hourly prediction from 07-27 10H to 08-02 22H

The team

Piangpim is a master's student in data science and business analytics with experience in management consulting.

I-Hsun is a master's student in data science and business analytics with experience in data analytics and optimization.

Bowei is a master's student in data science and business analytics with experience in pricing and sales forecasting.

Ruxi is a master's student in data science and business analytics with experience in machine learning modeling and forecasting.

Prashant is a master's student in data science and business analytics with experience in engineering.

Hanqi is a master's student in data science and business analytics with experience in data visualization for business management.