Enhancing Guest Experience through Data-Driven Journey Mapping and Analysis

Project Context:

Attractions and entertainment venues often struggle to provide consistently excellent guest experiences due to a lack of comprehensive understanding of the guest journey, preferences, and pain points. Traditional methods of gathering and analyzing guest data often fall short in providing actionable insights to improve operations, marketing strategies, and overall guest satisfaction.

This project aims to develop a data-driven system that maps and analyzes the entire guest journey, from pre-visit planning to post-visit feedback. By leveraging advanced data analysis, machine learning, and predictive modeling techniques, the project seeks to identify bottlenecks, optimize guest flow, personalize experiences, and ultimately boost guest satisfaction while potentially increasing revenue and operational efficiency.

Scenario:
You are a team of 8 data scientists joining a cross-functional team of business analysts, marketing managers, and operations experts working to enhance the guest experience at a major attraction or entertainment venue.

The team asks for your help with the following business questions:

Collaborative Phase (All Team Members):

Data Collection and Analysis Framework:
1. Identify and collect relevant guest journey datasets from provided sources and available systems (e.g., ticketing, point-of-sale, queue management, guest feedback).
2. Design a comprehensive survey to gather data on guest expectations, experiences, and satisfaction levels throughout their journey.
3. Conduct the survey and collect responses from a representative sample of guests.
4. Create a plan for generating synthetic data to:
   - Augment the collected data if needed.
   - Replace the data if unable to collect sufficient responses.
5. Clean and preprocess the collected datasets.
6. Generate synthetic guest journey data based on patterns in public datasets and reasonable assumptions to either augment existing data or replace it entirely if necessary.
7. Document all data sources, cleaning procedures, and synthetic data generation methods.

After completing the collaborative phase, split into two subgroups to address the following business questions:

Subgroup A: Guest Journey Analysis and Segmentation

1. What are the key factors influencing guest satisfaction throughout their journey?
   - Analyze survey responses and operational data to identify main drivers of satisfaction and dissatisfaction at each touchpoint.
   - Propose metrics for tracking guest satisfaction over time.

2. How can we segment our guests based on their behavior, preferences, and demographics?
   - Develop a guest segmentation model using the collected data.
   - Identify unique needs, preferences, and pain points for each guest segment.

3. What are the most common guest journey patterns, and how do they vary across different segments?
   - Analyze and visualize journey patterns across different attractions, times, and guest segments.
   - Identify opportunities for personalization and service improvements based on these patterns.

Subgroup B: Experience Optimization and Predictive Modeling

1. How can we accurately predict guest flow and demand for various attractions and services?
   - Develop a demand forecasting model using historical data, guest journey patterns, and external factors (e.g., weather, events).
   - Identify key factors influencing demand fluctuations.

2. What changes to attraction layouts, schedules, or operations would optimize the guest experience?
   - Create an algorithm to optimize guest flow based on predicted demand and guest preferences.
   - Simulate the impact of proposed changes on wait times, guest satisfaction, and revenue.

3. How can we better allocate resources (staff, capacity) across different attractions and services to meet varying demand?
   - Develop a model to optimize resource allocation across the venue.
   - Incorporate factors such as peak hours, special events, and seasonal variations.

Optional Bonus Questions (for higher grades) - answer any or all to boost your score:

Subgroup A: Guest Journey Analysis and Segmentation

1. How can we implement a real-time sentiment analysis system to process guest feedback and social media mentions?
   - Develop an NLP pipeline to analyze and categorize guest comments and social media posts.
   - Create a dashboard to visualize emerging themes and sentiment trends in real-time.

2. What is the potential impact of implementing a dynamic pricing model based on guest segments and demand patterns?
   - Design and implement a prototype dynamic pricing system.
   - Conduct simulations to assess the impact on revenue, guest behavior, and satisfaction.

3. How can we predict and mitigate the effects of unexpected events (e.g., attraction breakdowns, weather changes) on guest satisfaction?
   - Develop a machine learning model to predict the impact of various types of disruptions.
   - Create a real-time alert system and recommend mitigation strategies.

Subgroup B: Experience Optimization and Predictive Modeling

1. How can we incorporate external factors (e.g., local events, social media trends, economic indicators) into our demand forecasting model to improve accuracy?
   - Implement an automated data ingestion system for relevant external data sources.
   - Enhance the existing demand forecasting model to incorporate these new variables.

2. What is the potential impact of implementing an AI-driven real-time queue management system?
   - Develop a prototype AI system for real-time queue and capacity adjustments.
   - Set up a simulation environment to test and quantify the benefits of this system.

3. How can we use reinforcement learning to optimize the sequencing of attractions and experiences for different guest segments?
   - Implement a reinforcement learning model for personalized itinerary recommendations.
   - Create a visualization tool to demonstrate the model's decision-making process and potential impact on guest satisfaction.

Collaborative Deliverable:

As a team, synthesize your findings and recommendations into a comprehensive strategy for enhancing the guest experience.

Your final deliverables should include:

1. A 10-minute video presentation for CxO senior stakeholders.
2. A slide deck (8-12 slides) supporting the video presentation.
3. [Optional] An interactive dashboard showcasing your results and key metrics.

A Git repository containing:
1. Production-ready Python code:
   - Modular Python scripts for data cleaning, analysis, modeling, and visualization
   - A main.py file that orchestrates the entire data pipeline and analysis process

- A config.py file for all configuration parameters
- A utils.py file for utility functions used across multiple scripts
- SQL scripts for data extraction and transformation (stored in a 'sql' directory)
- A simple API (using Flask or FastAPI) to serve model predictions and key insights
- A requirements.txt file listing all dependencies
- Comprehensive docstrings for all functions, classes, and modules
- [Optional] Unit tests for all critical functions
- [Optional] A logging system for tracking the execution of the code

2. Docker-related files:
- A Dockerfile to containerize the application
- [Optional] A docker-compose.yml file if multiple services are required

3. Documentation:
- A README.md file with:
  - Project overview
  - Instructions for setting up the environment and running the code
  - Description of the repository structure
  - Data sources and any necessary data preparation steps
  - Instructions for building and running the Docker container(s)
  - API documentation (endpoints, request/response formats)
- [Optional] API documentation using Swagger/OpenAPI specification
- A data dictionary explaining all variables used in the analysis

Optional Deliverables (for higher grades) - You don't need to do them all:

1. Advanced Analytics:
- Implement a machine learning pipeline for continuous model training and deployment.
- Develop a computer vision model for analyzing guest behavior in common areas.
- Create a simulation model to test various optimization strategies under different scenarios.

2. Enhanced Visualization:
- Develop an interactive, real-time dashboard using tools like PowerBI, Tableau, or Streamlit.
- Create a 3D visualization of guest flow through the venue using tools like Three.js or Unity.
- Implement a heat map visualization of guest density and movement patterns.

3. Advanced Software Engineering:
- Implement a microservices architecture for different components of the system.
- Set up a CI/CD pipeline using tools like Jenkins or GitLab CI.
- Implement automated data quality checks and model performance monitoring.

4. Big Data Processing:
- Set up a data pipeline using Apache Airflow for orchestration.
- Implement batch and stream processing capabilities using Apache Spark.

- Design a data lake architecture for storing and processing large volumes of guest data.

5. Advanced API and Mobile Application:
   - Develop a full-fledged mobile app for guests with personalized recommendations and real-time updates.
   - Implement real-time updates using WebSockets.
   - Create a chatbot interface for guest inquiries and feedback.

6. Extended Documentation and Testing:
   - Develop comprehensive API documentation using tools like Swagger UI.
   - Implement integration tests and end-to-end tests in addition to unit tests.
   - Create a detailed technical design document outlining the system architecture and data flow.

7. Advanced Containerization and Deployment:
   - Implement a Kubernetes deployment for scalability and easier management.
   - Set up monitoring and logging using tools like Prometheus and ELK stack.
   - Implement blue-green deployment strategy for zero-downtime updates.

8. Ethics and Privacy Enhancements:
   - Develop a comprehensive data anonymization and privacy protection strategy.
   - Implement differential privacy techniques for sensitive guest data analysis.
   - Create an ethics review process for new features and analyses.

9. Business Intelligence:
   - Develop a set of KPIs and create an executive dashboard for tracking guest experience performance.
   - Conduct a cost-benefit analysis of proposed optimizations and create a prioritized roadmap.
   - Perform a competitive analysis comparing the proposed system to other leading attractions worldwide.

Instructions:

1. Each subgroup should create a 5-7 slide mini-presentation addressing their specific business questions.
2. The full group should then collaborate to create the main 8-12 slide presentation that synthesizes the key findings and recommendations from both subgroups.
3. Record a 10-minute video presentation suitable for CxO senior stakeholders, using the main slide deck as visual support.
4. Create a Wiki for your project.
5. Develop production-ready Python code, SQL scripts, and an API. Set up a Git repository with all code, documentation, and non-sensitive data files. Ensure it's well-organized, follows PEP 8 style guidelines, and includes clear instructions for use.
6. Create a Dockerfile and docker-compose.yml (if necessary) to containerize your application.

Grading:

Your assignment will be judged according to:

- The analytical approach and clarity of your graphs, tables, visualizations,
- The data decisions you made and reproducibility of the analysis,
- Strength of recommendations, prioritizations, and rationale behind those,
- The narrative of your presentation and ability to effectively communicate to non-technical stakeholders,
- Quality and organization of the production-ready Python code,
- Effective use of SQL for data extraction and transformation,
- Design and implementation of the API,
- Proper containerization of the application using Docker,
- Comprehensive documentation and testing of the code,
- The effectiveness of your group collaboration and integration of different aspects of the analysis.
- Implementation and quality of optional deliverables (for higher grades)

About Data:

For this project, our focus will be on a major attraction or entertainment venue. You may choose to model your analysis on a theme park, museum, or other similar venue.

Should you require synthetic data to replicate or augment real-world data, here are some resources you can utilize:

1. Theme Park Attendance Data:
   - Description: Contains information on annual attendance for major theme parks worldwide.
   - Link: https://aecom.com/theme-index/

2. TripAdvisor Reviews for Popular Attractions:
   - Description: Includes customer reviews and ratings for various attractions.
   - Link: https://www.tripadvisor.com/ (You may need to scrape data following their terms of service)

3. Weather Data:
   - Description: Historical weather data that can impact attraction attendance and guest experience.
   - Link: https://www.ncdc.noaa.gov/cdo-web/

4. Social Media Data:
   - Description: Public posts and sentiment related to attractions (e.g., Twitter API).
   - Link: https://developer.twitter.com/en/docs

For survey data, which is not typically publicly available due to privacy concerns:

Design a mock survey using Google Forms or SurveyMonkey, and then either collect responses from colleagues or generate synthetic data based on logical assumptions about guest behavior. In this case, real data collected using the survey will score higher than just synthetic data. If you go for the real-data route, there are things you need to consider (and explain): Minimum survey size, keep it well documented as the data (anonymized) would have to be submitted.

Instructions for Data Use:

1. Choose one or more of the provided datasets as a base for your analysis.
2. You may need to clean, transform, or combine datasets to suit the needs of this project.
3. For any missing aspects (e.g., detailed guest preferences), create synthetic data that aligns with the patterns observed in the real data.
4. Clearly document any assumptions made or synthetic data created in your final report.

Additional Considerations:

- Given that you're working with data from different sources, you may need to adjust your analysis and recommendations to a hypothetical venue that combines aspects of these datasets.
- Pay special attention to data privacy and anonymization, even when working with public datasets.
- When creating synthetic data or making assumptions, base them on patterns observed in the real datasets and document your reasoning clearly.

Bonus / Extra Credit Opportunity:

After completing the main project, teams can earn extra points by creating a Minimum Viable Product (MVP) for an interactive guest experience enhancement system:

1. Design and implement a basic mobile app interface for guests.
2. Incorporate at least one AI-powered feature (e.g., personalized recommendations, real-time crowd prediction, virtual assistant).
3. Create a small-scale prototype that demonstrates how the interactive system could enhance guest experience.
4. Provide a brief report (2-3 pages) outlining:
   - The key features of your MVP
   - How it addresses common guest pain points
   - Potential benefits for guest satisfaction and business performance
   - Implementation challenges and how you addressed them
   - Ethical considerations and data privacy aspects of your interactive system
5. Include a demo video of your MVP in action.