

Lobster Land Annual Summer Season 2023 Project Proposal

Team Delta

Putra, Steve, Hugi, Yusen, Luke, Paul

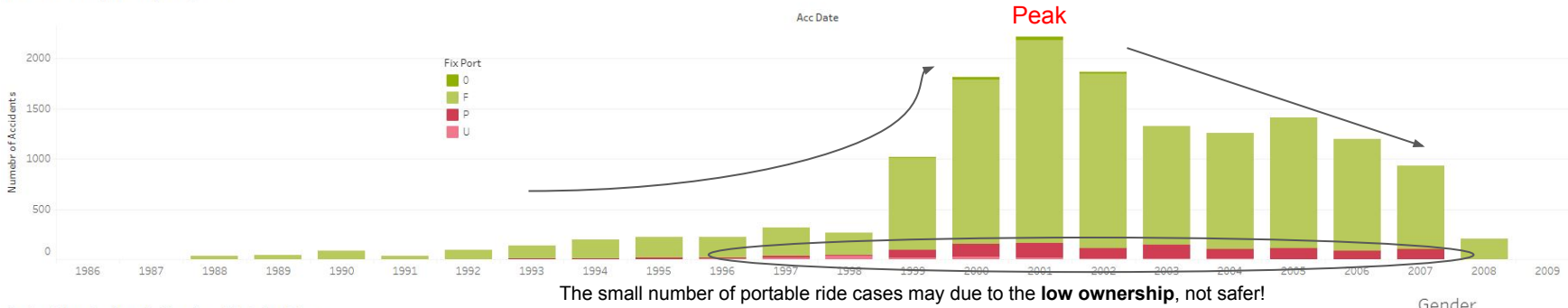
Executive Summary

- Delta Team **performed an analytics consultation** for Lobster Land Management to prepare for an international conference of theme park operators and hoteliers three days prior the Memorial Day.
- The team **delivered seven key outputs**, including visualizations of historical theme park accidents, segmentation and targeting analysis for a ski-themed hotel, strategic memo recommendation, and forecasting of hotel spending.
- Other key deliverables included conjoint analysis for new hotel amenities and rates, a classification model for predicting hotel satisfaction, and statistical testing for email marketing campaigns.

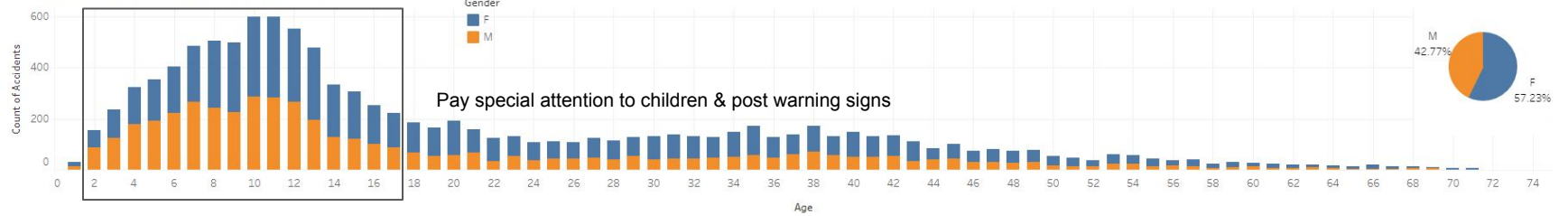
Park Accident: Visualization and Statistics

Park Accident: Tableau Visualization

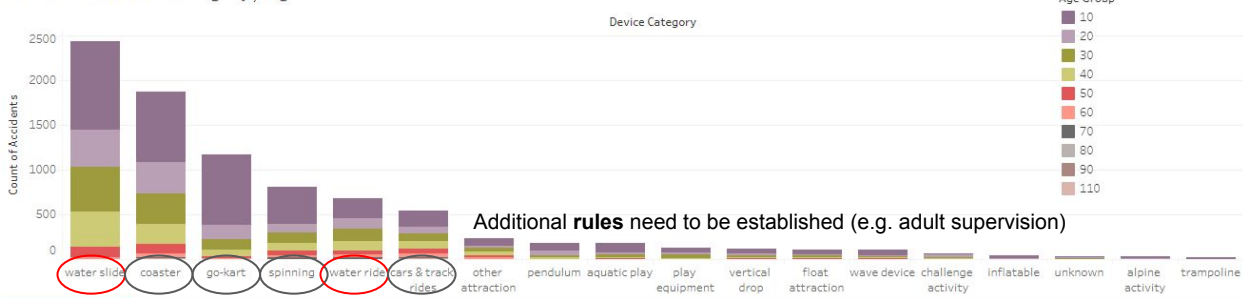
1. Accidents/Year/Fix_Port



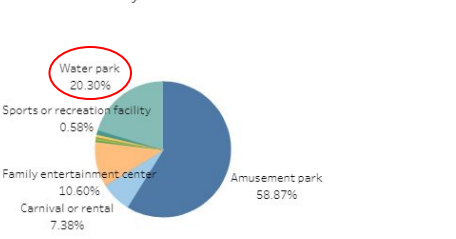
2. Accidents Age & Gender Distribution



3. Accident Device Category / Age

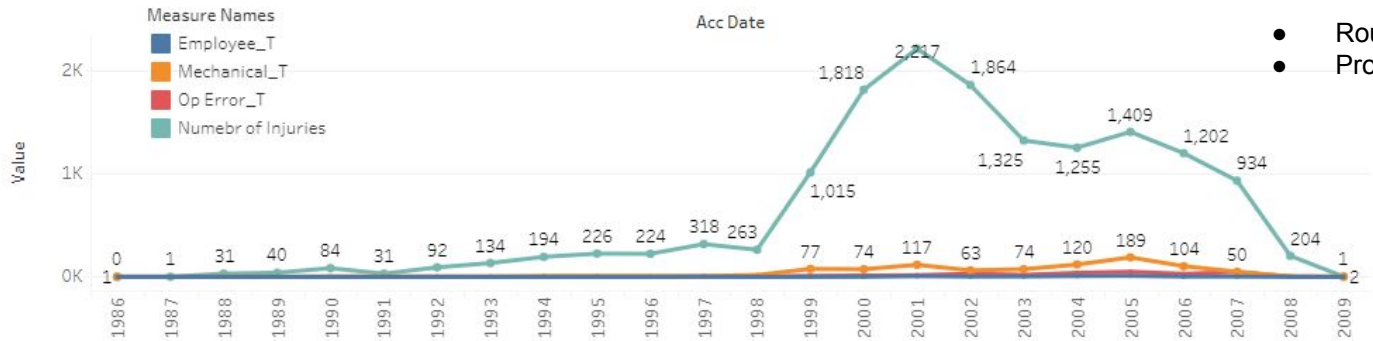


4. Accidents by Park



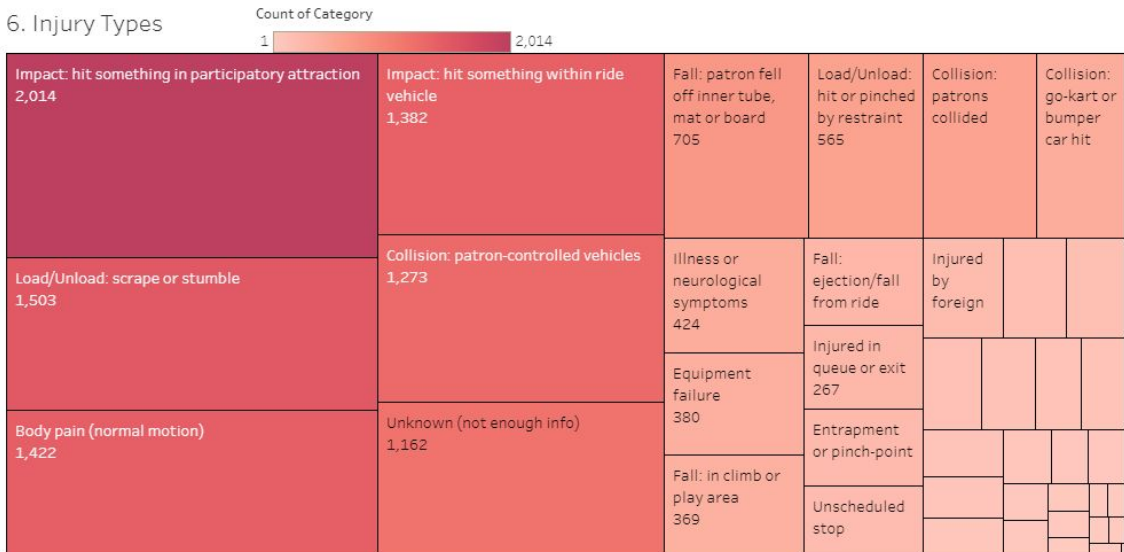
Park Accident: Tableau Visualization

5. Accidents vs. Park Responsibility



- Routine inspections is essential
- Proper ride operation

6. Injury Types



- First-aid stations
- Qualified personnels
- Immediate medical attention
- Emergency response plan

Park Accident: Summary Statistics

Accident by state



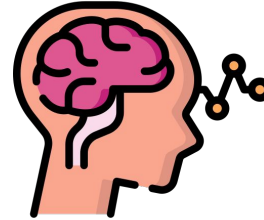
New Jersey has the highest number of accidents (5646) followed by **California** and **Texas**.

Injury Causes



Water slides account for the highest number of accidents (3530), followed by **Coasters** and **Spinning Rides**.

Mean Age by Accident Category



Accidents involving "**Illness or neurological symptoms**" predominantly affect older individuals, with mean age of 38.38 and 25.57 year and categories such as "**Derailment,**" and "**Choking, water inhalation, suffocation**" involve younger individuals with mean ages of 6.98, and 8.82 years, respectively.

Manufacturers of the faulty ride with the highest number of injured people

INTAMIN

Manufacturers with the highest number of injured people include **INTAMIN AG** and **in-house** manufacturers with total people with injuries reach 795 and 1167 respectively.

Ski-Themed Hotel Segmentation and Targeting

Meet Our Ski Hotel Resort Clusters!



The Slope Saver



Difficulty



Price



The Alpine Haven



Difficulty



Price



The Glacier Elite



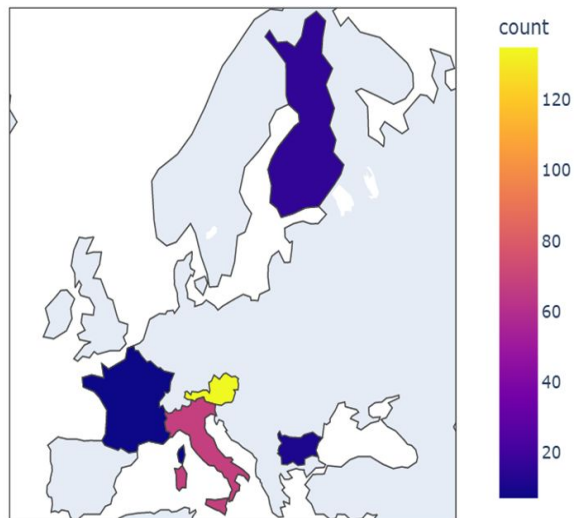
Difficulty



Price

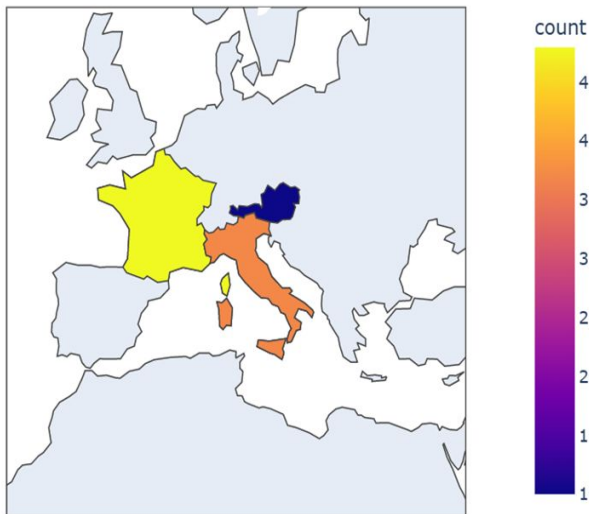
Where are the cluster located?

The Slope Saver: Country Count



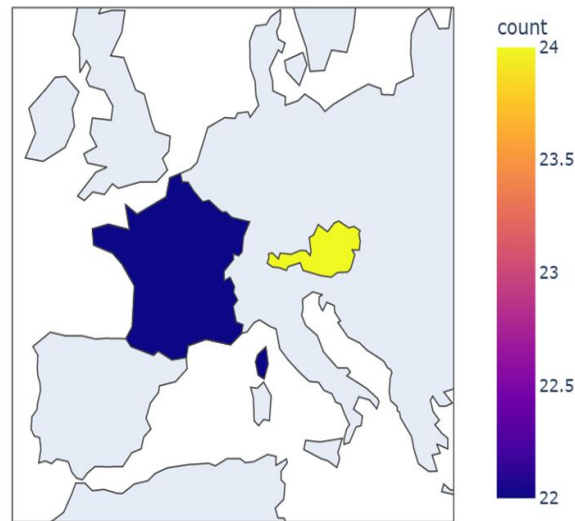
**Austria, Italy, Bulgaria,
Finland, France**

The Alpine Haven: Country Count



France, Italy, and Austria

The Glacier Elite: Country Count



France and Austria

Hotel Amenities Conjoint Analysis

Conjoint Analysis Result for Lobsterland

Amenity	Level	Cost	Coefficient
Wifi Network	Best in Class	19.15	1.73
Breakfast	Full Buffet	22.45	0.73
Flex Check	Yes	12	0.48
Pool Temp.	84F	45	0.26
VIP Shop	Yes	12	0.22
Air Purification	Yes	12.85	0.08
Gym	Basic	10	0.02
Parking	Open Lot	15	0
Shuttle Bus	No	0	0
Jacuzzi	No	0	0

Maximum Budget: \$150.00

Total Expected Cost: \$148.45

Best Coefficient: 3.51

Recommendations

- Focus on the best value per marginal cost
- Continuous operation throughout the year
- Flexible operation highly recommended

Strategic Memo

Strategic Memo

Golden Arch SWOT Analysis

S

Strengths ✓

- Strong brand recognition of McDonald's worldwide.
- Competitive pricing strategy.
- Technologically advanced amenities for the time.

W

Weaknesses ✓

- Inconsistent customer experience across various aspects of the hotel stay.
- Incongruent brand image between McDonald's and a four-star hotel.
- Limited food options and poor ambiance in the bar and restaurant area.

O

Opportunities ✓

- Capitalize on the established McDonald's brand to attract customers.
- Collaboration with airlines, tour operators, and business travelers to increase occupancy rate.
- Leverage the hotel's proximity to the airport and exhibition area.

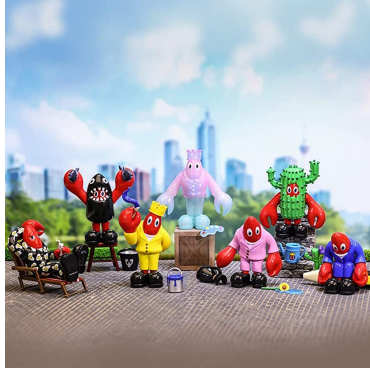
SWOT

T

Threats ✓

- Strong competition from other established hotel chains.
- Negative customer experiences impacting the hotel's reputation.
- Potential language and cultural barriers.

Recommendation for Lobsterland to enter new market!



Clear **brand identity** and **differentiation**



Aligning **brand expectations**



Consistency in **customer experience**



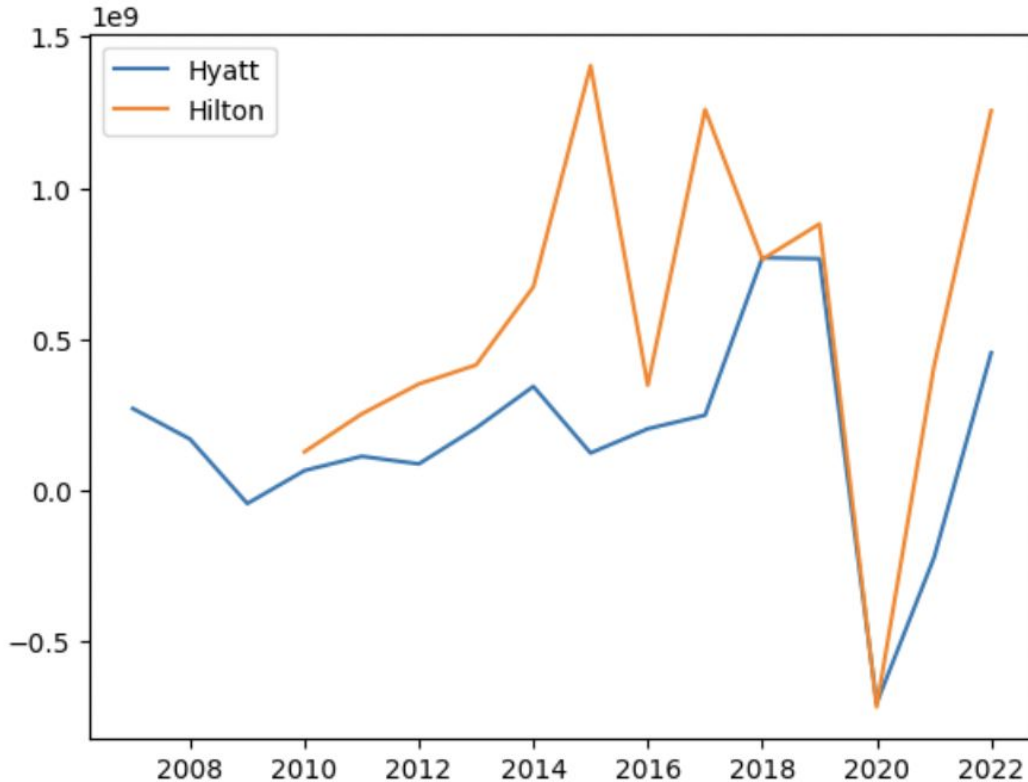
Leveraging **technology**



Community involvement and **sustainability**

Hotel Spending Forecast

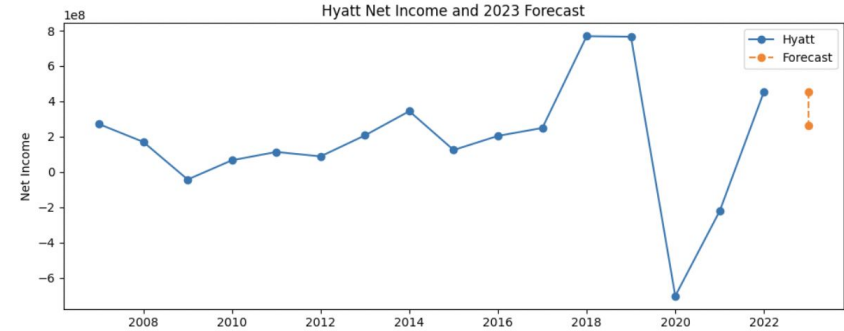
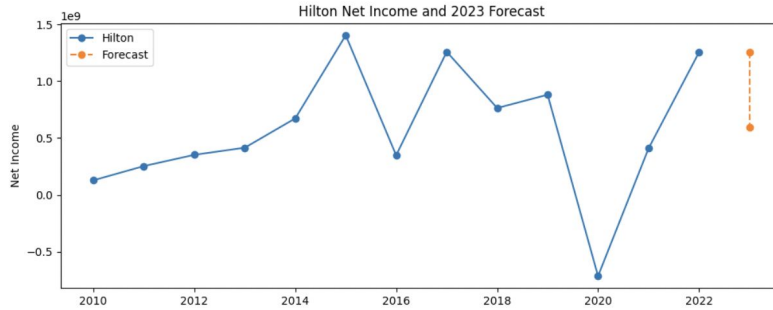
Net Income Forecast for Hyatt and Hilton



Net Income past data for Hyatt gathered from 2007 to 2022 and Hilton gathered from 2010 to 2022.

What will be the net income for 2023?

Net Income Forecast for Hyatt and Hilton



Past Net Income data downloaded
from **Yahoo Finance** annual
Report



There are no Trend or Seasonality.
Therefore **Single Exponential
Smoothing (SES)** is used.

HYATT®

Forecasted Net Income
for Hyatt in 2023 is
\$263,902,512.75



Forecasted Net Income
for Hilton in 2023 is
\$592,369,792.45

Hotel Customer Satisfaction Classification Model

Classification Analysis of Hotel Satisfaction Data

Objective of this Analysis:

*“Get an **insight into** what drives **consumer satisfaction** in the **hospitality** aspect of the hotel business”*

Data:

id	Gender	Age	purpose_of_travel	Type of Travel	Type Of Booking	Hotel wifi service	Departure/Arrival convenience	Ease of Online booking	Hotel location
				Food and drink	Stay comfort	Common Room entertainment	Checkin/Checkout service	Other service	Cleanliness satisfaction

Classification Analysis of Hotel Satisfaction Data

Steps to Building our Logistic Regression Model:

1. Preprocess:

- Identify Categorical and Numerical Variables and Outcome Variable ('satisfaction')
- Inspect Correlations
- Remove variables
- Dummify categorical variables

2. Build Model:

- Iteration #1 → remove variables based on significance,
- Iteration #2
- Iteration #3

3. Evaluate Coefficients of Regression Model to gain Insight



Logit Regression Results									
Dep. Variable:	unemployment	Number of Observations:	4052						
Model:		Sum of Squared Residuals:	421.0						
Date:	01/01/2013	Time:	08:21						
File:	130101unl	Unrestricted Model:	-124.0						
Unrestricted:		Restricted Model:	-124.0						
Unrestricted:		LR Statistic:	0.00						
Unrestricted:		Prob (Chi-Square):	1.000						
Unrestricted:		Asymptotic Significance:	1.000						
Constant	-4.9592	0.5756	-8.2711	0.000	-3.868	0.5756	-8.2711	0.000	
Age	0.0317	0.0161	19.7030	0.000	0.0317	0.0161	19.7030	0.000	
Male	-0.0712	0.0226	-3.1459	0.002	-0.0712	0.0226	-3.1459	0.002	
Married with spouse	-0.2054	0.0999	-2.0556	0.041	-0.2054	0.0999	-2.0556	0.041	
Dependent/unemployed	-0.0004	0.0009	-0.4000	0.692	-0.0004	0.0009	-0.4000	0.692	
Unemployed	0.0000	0.0000	0.0000	1.000	0.0000	0.0000	0.0000	1.000	
Ever married	0.0001	0.0161	0.6760	0.500	0.0001	0.0161	0.6760	0.500	
Married	0.0000	0.0000	0.0000	1.000	0.0000	0.0000	0.0000	1.000	
Charless/Charless spouse	-0.0889	0.0500	-1.7800	0.076	-0.0889	0.0500	-1.7800	0.076	
Charless	0.0000	0.0000	0.0000	1.000	0.0000	0.0000	0.0000	1.000	
Spouse	0.0000	0.0000	0.0000	1.000	0.0000	0.0000	0.0000	1.000	
Classmate	-0.1319	0.1014	-1.2945	0.195	-0.1319	0.1014	-1.2945	0.195	
Classmate (Personal Network)	-0.0000	0.0000	0.0000	1.000	-0.0000	0.0000	0.0000	1.000	
Time	-0.0000	0.0000	0.0000	1.000	-0.0000	0.0000	0.0000	1.000	
Time (Personal Network)	-0.0000	0.0000	0.0000	1.000	-0.0000	0.0000	0.0000	1.000	
Time (Personal Network)	-0.1389	0.0448	-3.0749	0.002	-0.1389	0.0448	-3.0749	0.002	



Regression Output

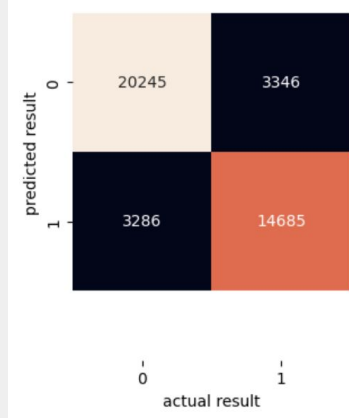
Coefficients of Regression:

	coef
-----	-----
const	-4.6992
Age	0.0137
Hotel wifi service	0.5717
Departure/Arrival convenience	-0.0556
Hotel location	-0.1748
Food and drink	-0.1355
Stay comfort	0.2501
Common Room entertainment	0.3693
Checkin/Checkout service	0.3398
Other service	0.1937
Cleanliness	0.1725
Type of Travel_Personal Travel	-1.8802
Type Of Booking_Individual/Couple	-1.4892
Type Of Booking_Not defined	-1.3169
=====	=====

Test Set Stats:

Accuracy: 0.8404

Balanced Accuracy: 0.8377



- Generalizes well to unseen data
- 103904 records, 17 columns
- **60% training, 40% test**

Insights

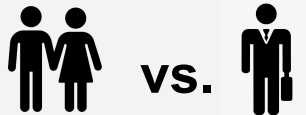
Positive Impact:

- Most important: Wifi!
- Convenience related variables follow:
 - Room Entertainment
 - Check-in/Check-out
 - General Convenience



Negative Impact:

- Type of Booking impacts satisfaction



Recommendations

- Expand and Advertise Wifi on booking website
- Ensure heightened convenience by communicating offerings upon check-in (HBO, directions to hotel, additional convenience items in room)
- Further inspect specific type of booking needs



Promotional Picture Statistical Tests

Statistical Test - Promo_pics Dataset

Objective / Success Criteria:

- Recommend one of these three pictures; Sunset, Main St, or Waterslide, for the **next round of invites for a conference** (email format)
- Conduct A/B test to determine which picture **performs better** and whether it is **statistically significant**

Steps:

1. Summary Statistics - checking for normal distribution
2. Graph inputs to observe distributions
3. T-test on pic_seen and user spend to determine statistical significance
4. T-test on pic_seen and duration to determine statistical significance
5. Binomial Z-Test on pic_seen and registration to determine statistical significance
6. Calculate average people who register, amount spent, and duration



Statistical Test Results

T Test - Picture and Spend

- Main St to Waterslide: statistic = -48.20, p-value = 0.0
- Sunset to Main St: statistic = 40.53, pvp-value = 2.433e-270
- Sunset to Waterslide: statistic = -23.6, p-value = 3.5328e-110

T Test - Picture and Site Duration

- Main St to Waterslide: statistic = 112.1, p-value = 0.0
- Sunset to Main St: statistic = -9.911, p-value = 1.0724e-22
- Sunset to Waterslide: statistic = 180.719, p-value = 0.0

Binomial Z Test - Picture and Registration

- Main St to Waterslide: statistic = 5.393, p-value = 6.89e-08
- Sunset to Main St: statistic = -2.36818, p-value = 0.01787
- Sunset to Waterslide: statistic = -2.983, p-value = 0.0028

Mean Spend (\$)

- Main St: 14.016289
- Sunset: 16.781892
- Waterslide: 18.606778

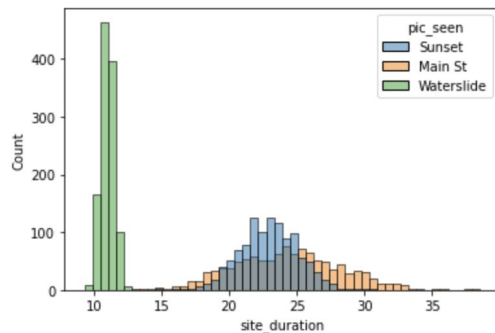
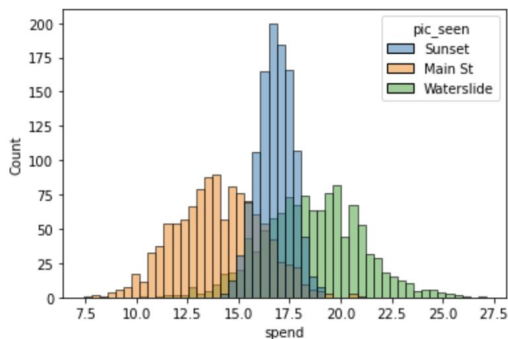
Mean Duration

- Main St: 24.29
- Sunset: 22.949
- Waterslide: 10.975

Mean Registration Rate

- Main St: 0.341463
- Sunset: 0.440541
- Waterslide: 0.348511

Reject the null hypothesis that there is no difference in customer's registering based on the photo and accept that there is a statistical difference



	recipient	site_duration	spend	register
count	3400.000000	3400.000000	3400.000000	3400.000000
mean	1700.500000	19.381494	16.461041	0.376176
std	981.639785	6.559998	2.707941	0.484496
min	1.000000	9.300000	7.400000	0.000000
25%	850.750000	11.300000	14.900000	0.000000
50%	1700.500000	21.500000	16.600000	0.000000
75%	2550.250000	24.400000	17.820000	1.000000
max	3400.000000	38.510000	27.120000	1.000000

Conclusion and Recommendations

Conclusions

- Insights from analysis help create an appealing, successful park
- Cater to target audience while maintaining brand identity
- Use data for future improvements & strategizing

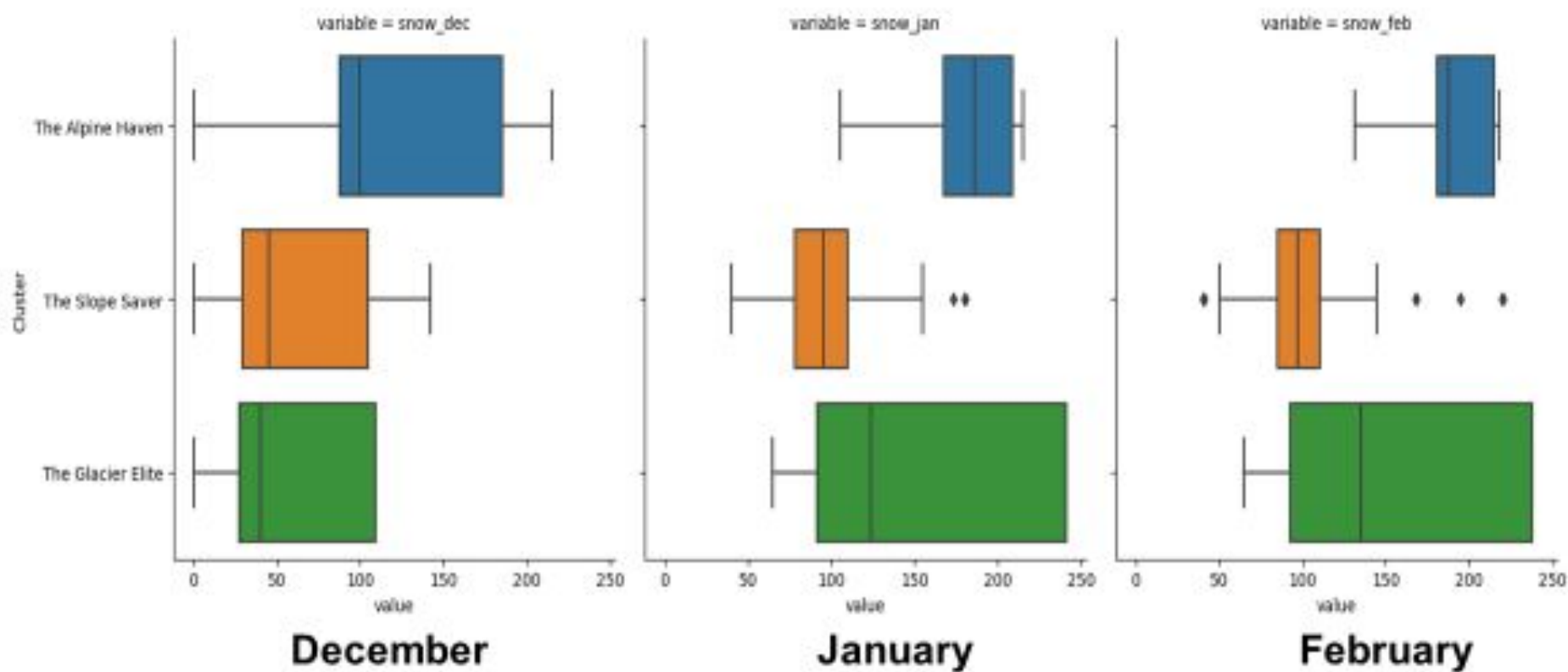
Recommendations:

- Emphasize USP: seaside atmosphere, affordability, family-friendly
- Provide valuable amenities, scalable services, and quality experience
- Engage local community, create seasonal events, promote sustainability

Thank You!

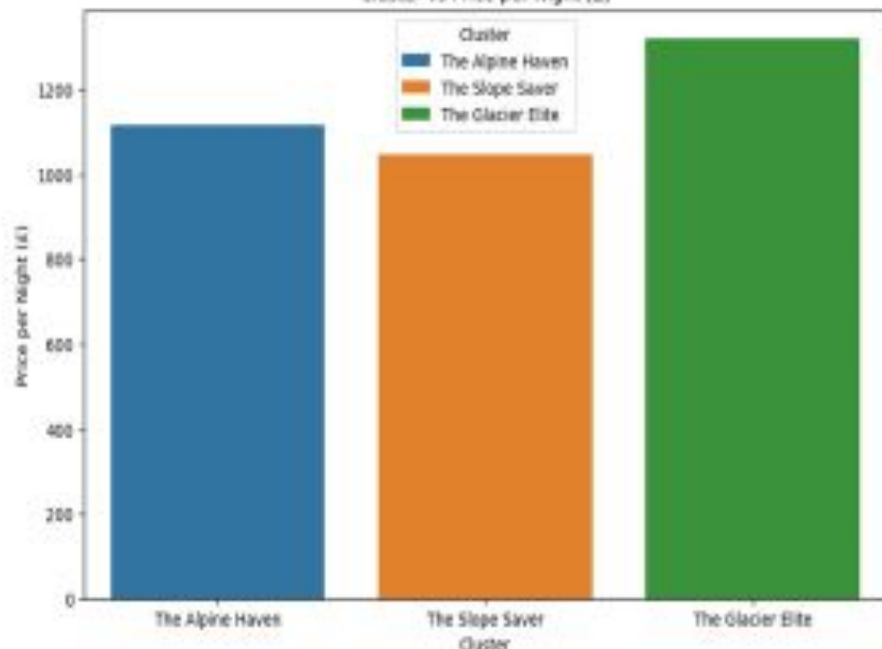
AD 654 Final Project.ipynb -
Colaboratory (google.com)

What about the snow?



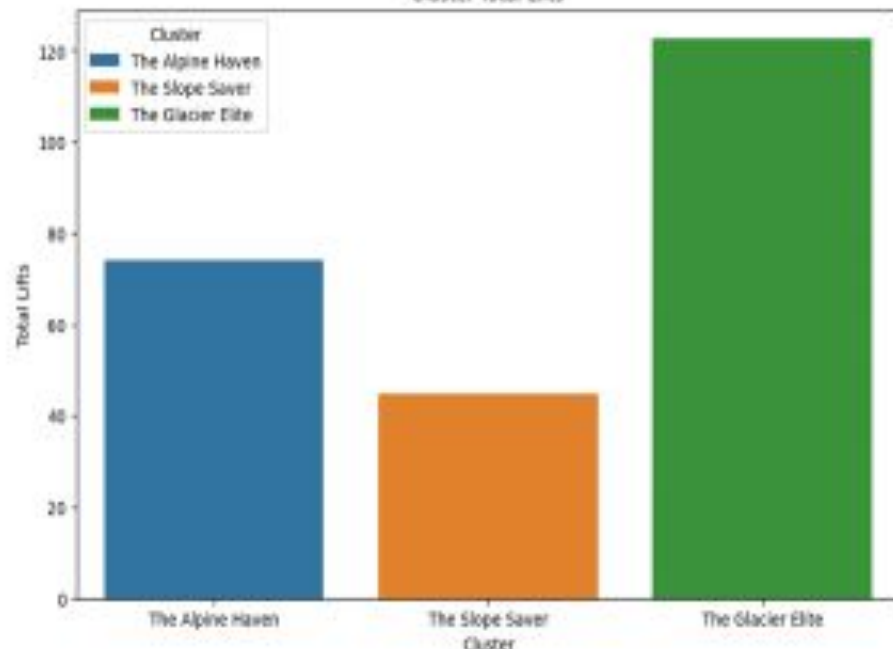
How much do i need to pay?

Cluster vs Price per Night (£)



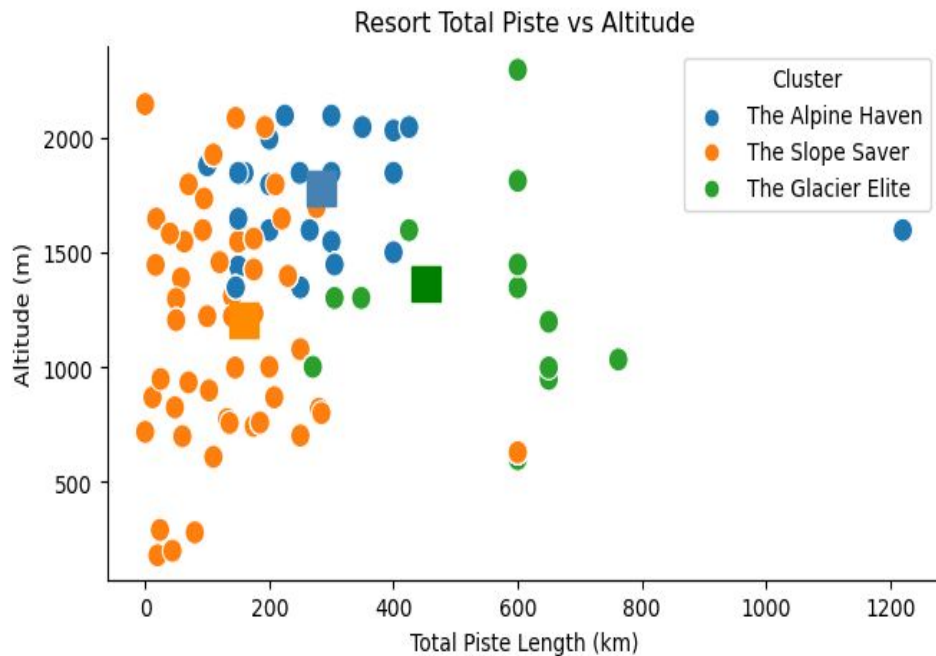
Cluster Price

Cluster Total Lifts



Cluster Total Lifts

How challenging is the cluster?



Total Piste vs Altitude



Total Piste vs Total Runs

Backup

Logit Regression Results

```
=====
Dep. Variable:    satisfaction_satisfied    No. Observations:    62342
Model:                Logit    Df Residuals:    62331
Method:                MLE    Df Model:    10
Date:                Mon, 08 May 2023    Pseudo R-squ.:    inf
Time:                00:35:24    Log-Likelihood:    -inf
converged:                True    LL-Null:    0.0000
Covariance Type:    nonrobust    LLR p-value:    1.000
=====
```

```
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const              -5.9751      0.066     -90.988      0.000      -6.104      -5.846
Age                 0.0173      0.001      26.030      0.000       0.016       0.019
Hotel wifi service   0.6017      0.009      67.096      0.000       0.584       0.619
Departure/Arrival convenience -0.3234      0.008     -41.899      0.000      -0.339      -0.308
Hotel location       -0.0515      0.009      -5.824      0.000      -0.069      -0.034
Food and drink       -0.1813      0.011     -16.767      0.000      -0.202      -0.160
Stay comfort         0.3565      0.011      33.132      0.000       0.335       0.378
Common Room entertainment 0.4031      0.013      30.622      0.000       0.377       0.429
Checkin/Checkout service 0.3500      0.008      41.396      0.000       0.333       0.367
Other service        0.2476      0.010      23.895      0.000       0.227       0.268
Cleanliness          0.0921      0.012       7.542      0.000       0.068       0.116
=====
```