Living Desert Park Attendance Pre & Post Covid-19

Jorge Daniel Atuesta, Arjun Mahanty, Kevin Howard, Omar Girbau & Juan P. Villegas

THE LIVING DESERT
ZOO | GARDENS

DFSG CONSULTING
Est. 2022

Agenda

1. Project Overview

2. Results

3. Data Insights

4. Q&A

"The Living Desert's mission is desert conservation through preservation, education, and appreciation."

Project Overview

DFSG CONSULTING GROUP SPOTLIGHT

Exploratory analysis of the data from The Living Desert to help discover key insights on attendance in the 6 past years.

The company seeks to understand if they are back on track with park attendance after the Global Pandemic Covid-19.

The total ticket quantity was the main indicator for park attendance by day since 2017 to 2022.
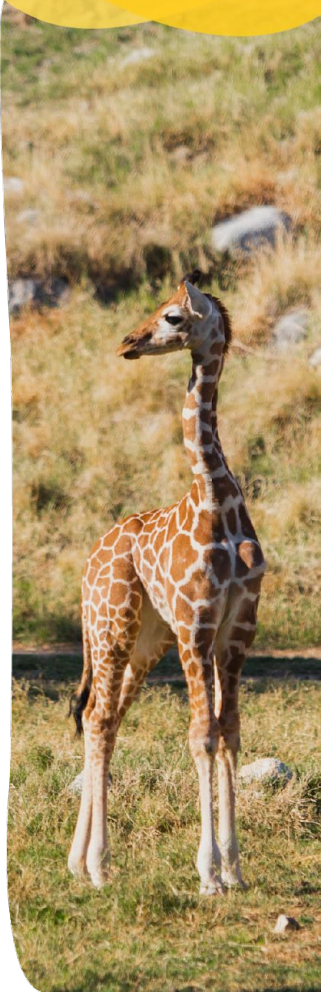
Results - Attendance After Covid-19 announcement by the CDC

★ Our model explains that currently the zoo is expected an attendance of ≈1730 compared to historical trend ≈723, before CDC's Covid announcement in March 20th 2020 *(when all other variables are offset).*

★ Attendance is in an upward trend but faces many external challenges.

★ The park attendance is seasonal with more down peaks.
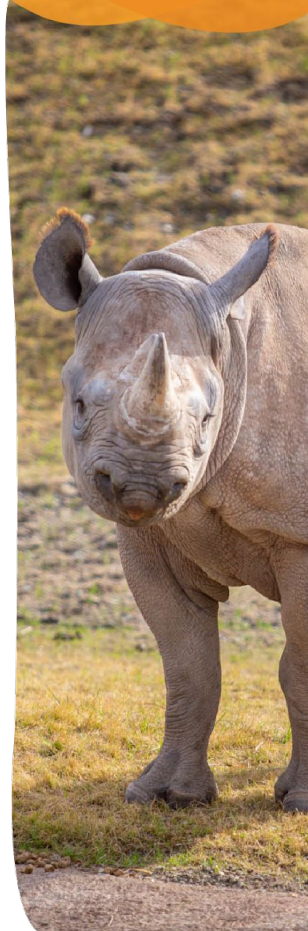
Results - Linear Regression Model

Based on <u>descriptive analytics</u> we aimed to solve the business question, "**Have we recovered from Covid?**" Our model consists of the following variables:

Variable	Effect on Total Attendance
Month	
Pre Covid vs Post Covid	*1006 more* in total attendance during post-covid relative to pre-covid
Weekend vs Weekday	*535 more* in total attendance on Weekends relative to Weekdays
Holiday and Attend Type Regular	*1 more* in total attendance when the ticket if used during a holiday and it is a regular attend type
Temperature	*4 less* people attend the zoo for every 1 degree increase in high temperature
Covid and Attend	*2 more* people in total attendance when it is post - covid and not an attraction ticket
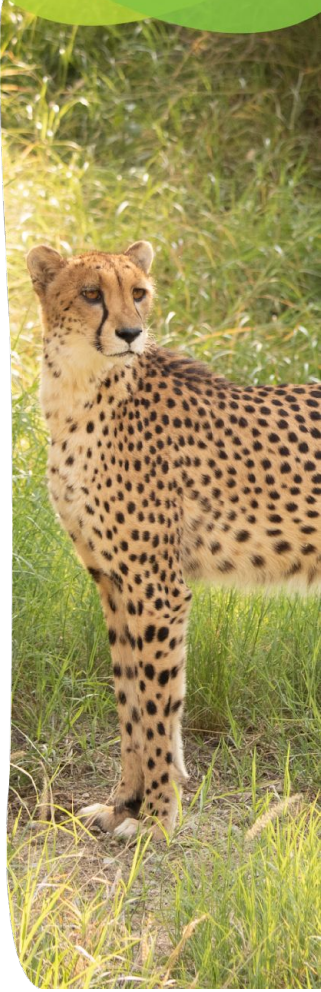
Results

➢ **March** has the highest amount of attendance at <u>750</u> more than December attendance

➢ **August** has the lowest amount of attendance at <u>903</u> less than December attendance
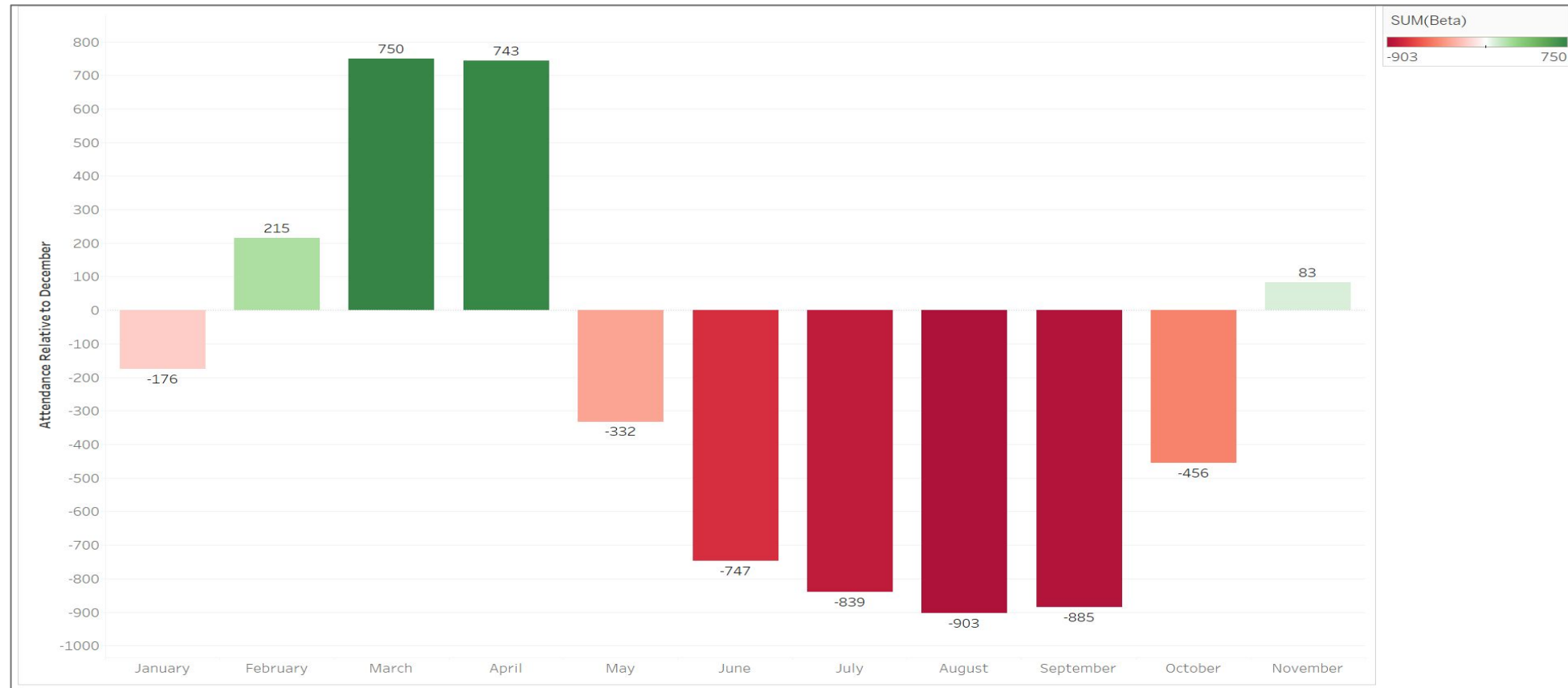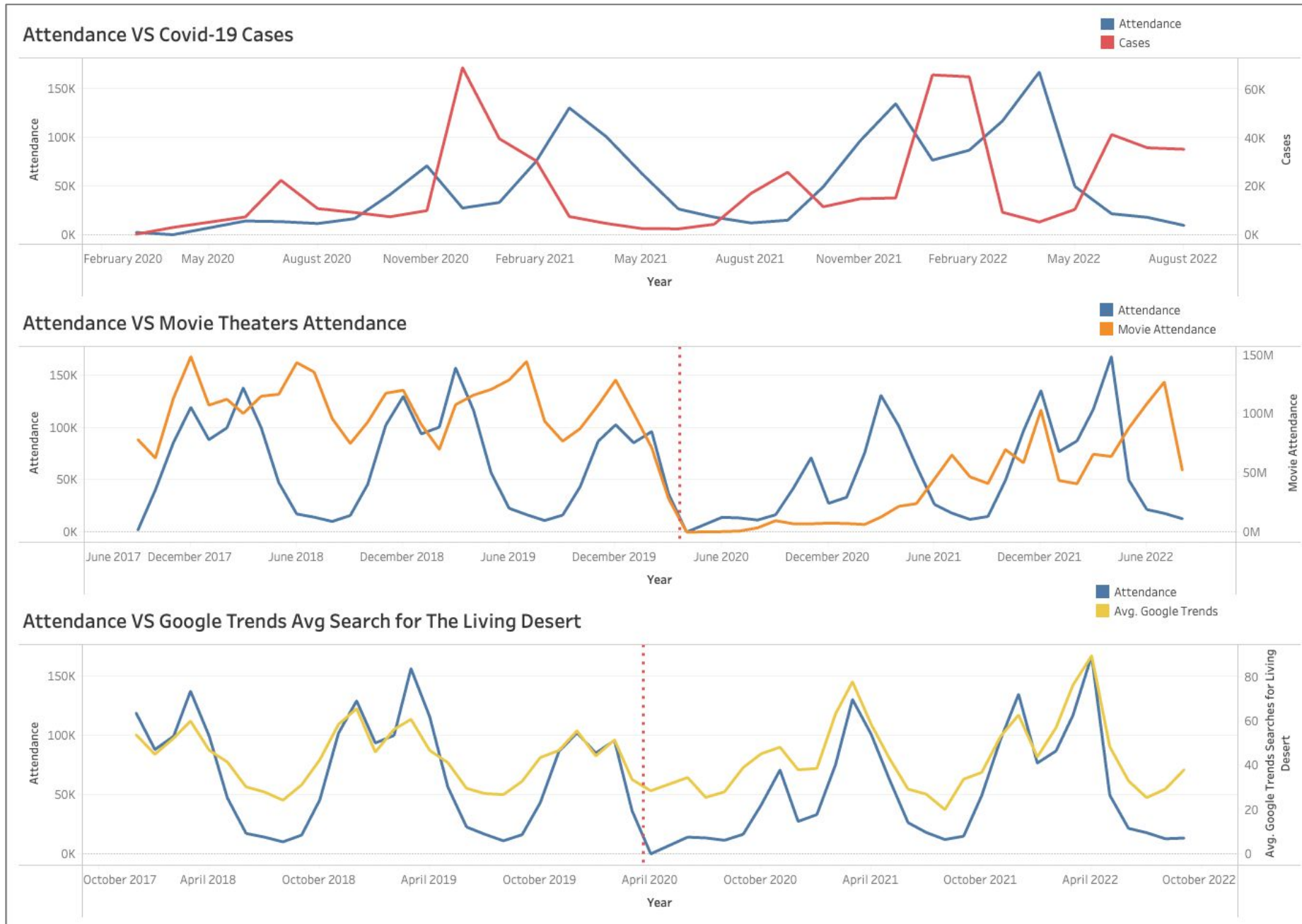
Contextual Data Insights

When COVID-19 cases rise, park attendance falls

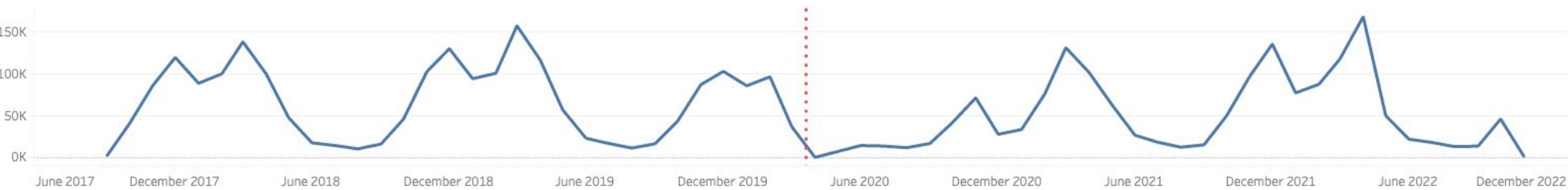LD appears to be recovering faster than movie industry and exceeding 2019 YoY attendance*

Dips in attendance relative to organic search trends present more efficient opportunity for marketing spend

Covid Recovery for The Living Desert

Upward trend for park attendance
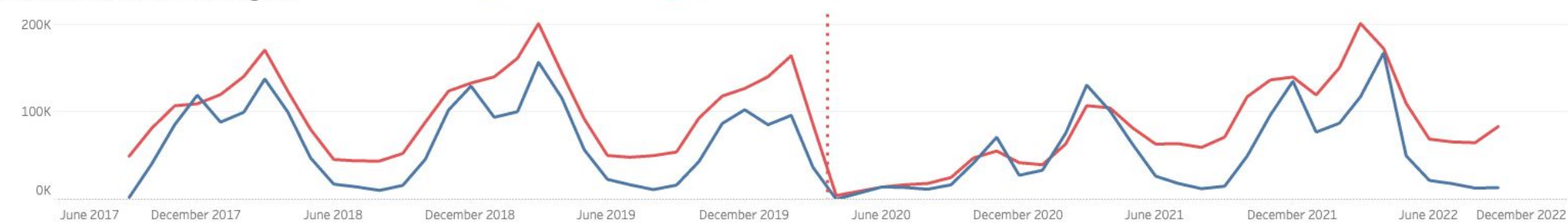
Web Sale spiked after COVID-19, but decreasing. "Other Sale" surging

Inbound flights & attendance are highly correlated. Gaps represent opportunity

	Peak Season Sept - Mar	**Non - Peak Season** Apr - Aug
Baseline	372,165 Sept 2018 - Mar 2019	139,809 Apr 2019 - Aug 2019
Comparison Baseline + 1	-16% Sept 2019 - Mar 2020	-78% Apr 2020 - Aug 2020
Comparison Baseline + 2	-17% Sept 2020 - Mar 2021	17% Apr 2021 - Aug 2021
Comparison Baseline + 3	16% Sept 2021 - Mar 2022	45% Apr 2022 - Aug 2022

Summary Dashboard

Year-over-Year attendance at the Living Desert has increased every year since COVID-19 began

Results - Key Takeaways

1. The ticket quantity in comparison to when Covid-19 was announced by CDC in March 2020 is back to normality (2019)

2. Statistical evidence proving seasonality and the effects on ticket quantity

3. The information gathered can be used in combination with marketing data to allocate marketing budget for peak and non-peak months

4. Traveling data (deplane) has a direct relationship with ticket quantity
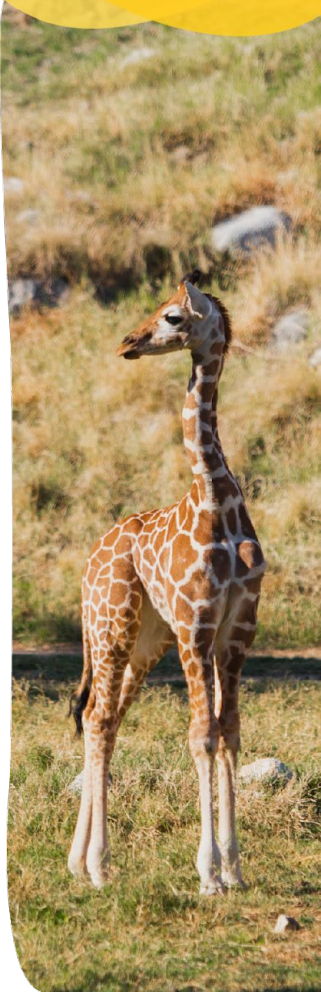
Future Considerations & Limitations

1. Limited data for Pre-covid (Feed the model more data to get more accurate results).

2. Create new interaction variables.

3. Create two separate models based on summer months and winter months.

1. Consider partnerships with local event holders during seasonal lows to boost attendance.

2. Consider analyzing seasonal organic search trends to optimize marketing campaigns for LD and partner campaigns

3. Build a predictive model to forecast future park attendance.

4. Splitting the "Other" Sales node into specific categories.

Q&A

Thank you

Appendix

Appendix A

Linear Regression

Model Summary - TotalQty

Model	R	R²	Adjusted R²	RMSE
H₀	0.000	0.000	0.000	1430.420
H₁	0.877	0.770	0.767	689.915

ANOVA

Model		Sum of Squares	df	Mean Square	F	p
H₁	Regression	2.770e+9	18	1.539e+8	323.355	< .001
	Residual	8.287e+8	1741	475982.685		
	Total	3.599e+9	1759			

Note. The intercept model is omitted, as no meaningful information can be shown.

Coefficients

Model		Unstandardized	Standard Error	Standardized[a]	t	p	Tolerance[a]	VIF[a]
H₀	(Intercept)	1426.977	34.096		41.851	< .001		
H₁	(Intercept)	1729.708	133.339		12.972	< .001		
	pre-covid/post-covid_pre-covid (1)	−1006.548	42.307		−23.792	< .001		
	CovidDummyXAttend	1.476	0.035	0.705	42.774	< .001	0.795	1.259
	Temp High	−3.868	1.664	−0.046	−2.324	0.020	0.817	1.223
	Month_Name_Apr (1)	742.982	91.786		8.095	< .001		
	Month_Name_Aug (1)	−903.005	101.417		−8.904	< .001		
	Month_Name_Feb (1)	215.382	82.602		2.607	0.009		
	Month_Name_Jan (1)	−175.574	81.047		−2.166	0.030		
	Month_Name_Jul (1)	−839.205	100.635		−8.339	< .001		
	Month_Name_Jun (1)	−746.736	99.680		−7.491	< .001		
	Month_Name_Mar (1)	749.990	83.836		8.946	< .001		
	Month_Name_May (1)	−332.129	92.392		−3.595	< .001		
	Month_Name_Nov (1)	83.053	83.416		0.996	0.320		
	Month_Name_Oct (1)	−456.031	85.638		−5.325	< .001		
	Month_Name_Sep (1)	−885.239	98.127		−9.021	< .001		
	Weekday/Weekend_Weekend	535.622	37.387	0.169	14.327	< .001	0.966	1.035
	HolidayXAttendType_Regular	0.266	0.070	0.077	3.766	< .001	0.968	1.033
	Holiday_yes/no_yes (1)	−26.215	182.093		−0.144	0.886		

[a] Standardized coefficients and collinearity statistics can only be computed for continuous predictors.

Appendix B

Ticket Quantity = 1729 – 4 (Temp) - 1007(Dummy Covid (Pre-covid)) + 1.48 (Covid Dummy X Attend) + 743 (Dummy April) - 903 (Dummy August) – 215 (Dummy February) – 176 (Dummy January) - 839 (Dummy July) – 747 (Dummy June) + 750 (Dummy March) – 332 (Dummy March) + 83 (Dummy November) – 456 (Dummy October) – 885 (Dummy September) + 536 (Dummy Weekend)– 26 (Dummy Holiday (YES)) +0.266 (Holiday X Attend type regular)

Appendix C

Extract, Transform & Load (ETL)

1. The data was provided by the client at the beginning of the project
2. Identified COVID announcement day CDC March 20th 2020.
3. Created Dummy variables pre and post covid announcement
4. Created Dummy variables for all months
5. Created Dummy variables for Holiday dates in the USA
6. Extracted high temperature data and appended to the data set
7. Created Linear Regression Model
8. Gather contextual data Air traffic and Marketing data
9. Loaded data into JASP, Python/Jupyter Lab and Tableau to perform analysis