

FORECASTING BROADWAY SALES

Predicting future sales of Broadway tickets

Team Juno

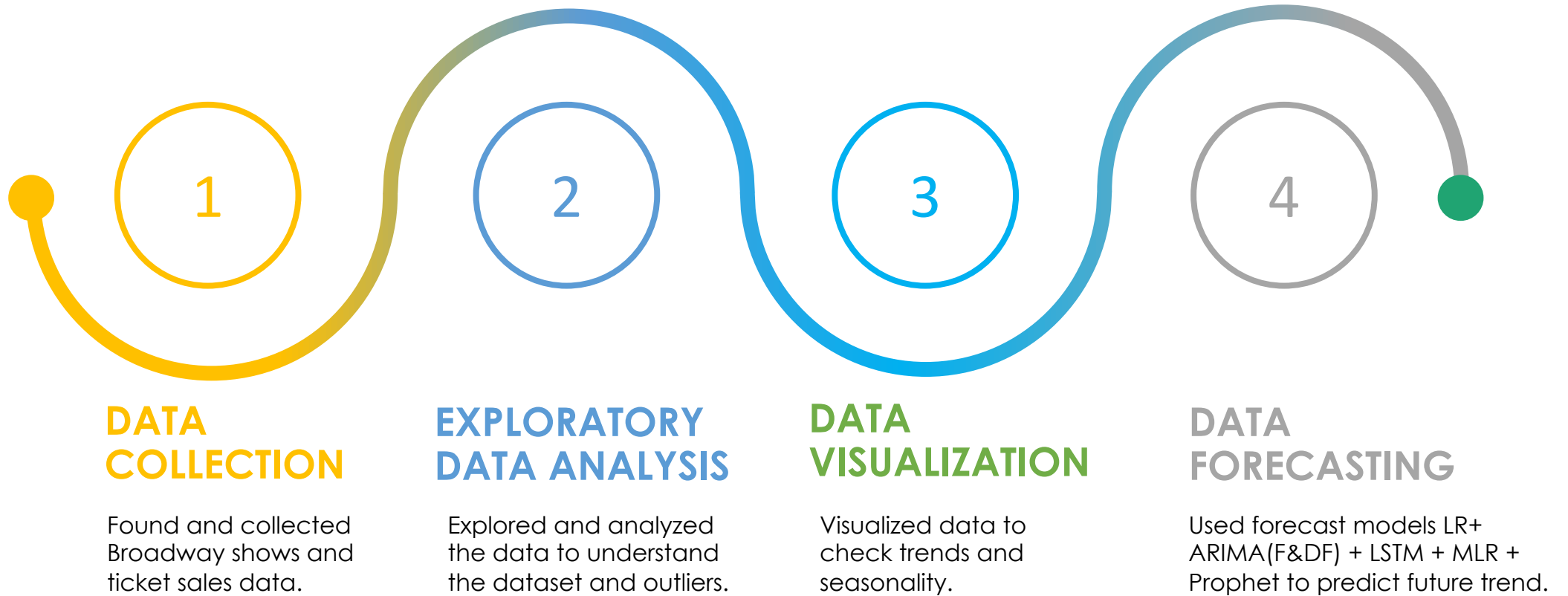
Praveen Shahani

Gautam Prakriya

Michael Amoah Awuah

Moran Shechnick

PROJECT ROADMAP



PUZZLE: FROM IDEA TO EXECUTION

Step 01: Find, web scrape and combine data

Broadway sales

<https://www.playbill.com/grosses>

Broadway reviews:

<https://www.newyorktheatreguide.com/reviews/broadway>

Step 02: Exploratory Data Analysis

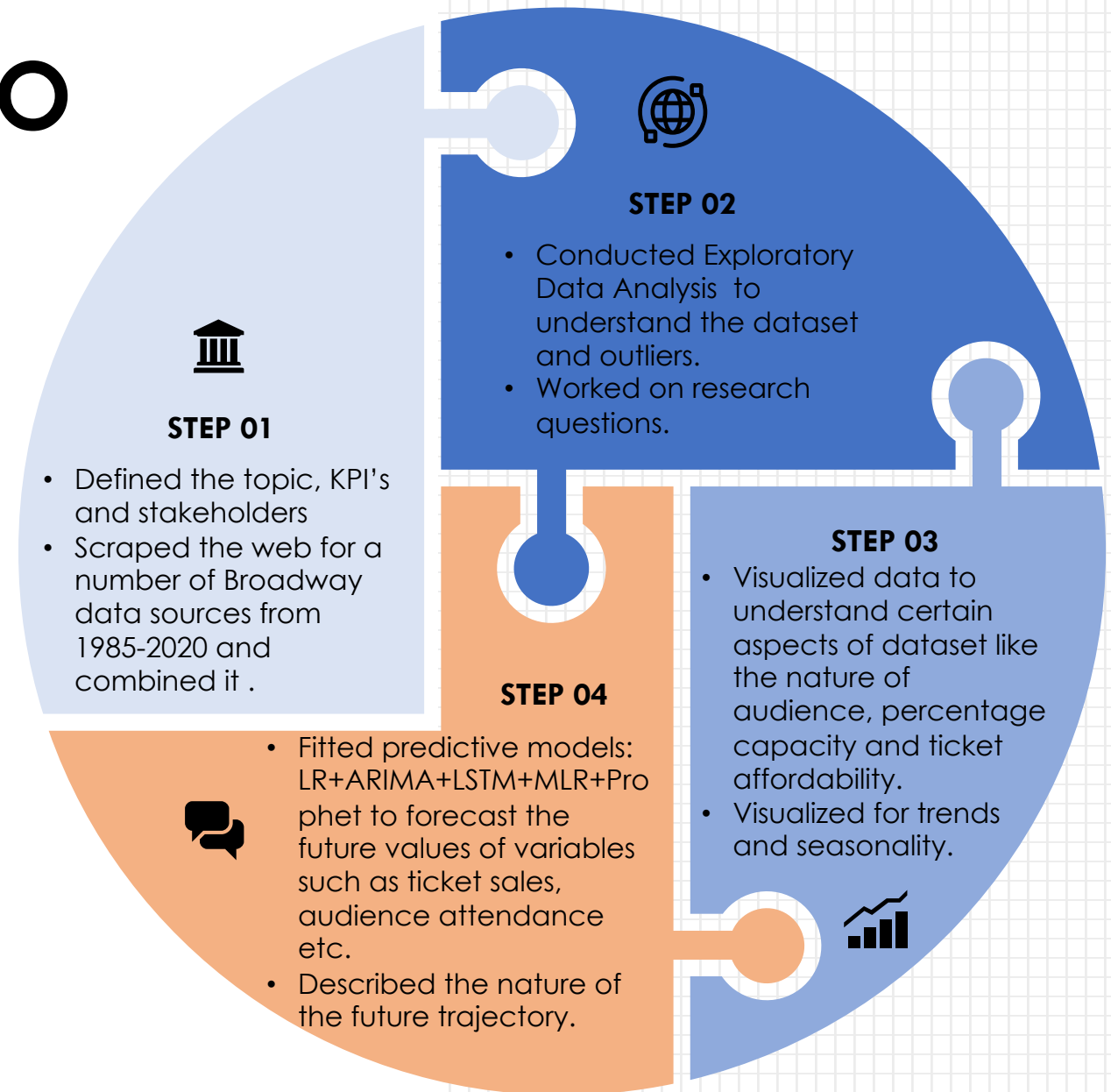
Described the data and conducted EDA to explore the data and its outliers.

Step 03: Visualization of Data

Visualized the data to check for trends and seasonality.

Step 04: Use the ARIMA, LSTM, MLR, and Prophet predictive models on the dataset

Used LR + ARIMA, LSTM, MLR & Prophet for time-series forecasting of revenue and ticket sales.



Introduction

Motivation:

Money for Broadway shows is typically raised by a small group of producers. Besides the initial costs of producing the show, the producers are also responsible for the weekly wages of the cast and crew as well as the rent for the theatre in which the show is played. It typically takes investors 1 year to 18 months to break even. Broadway shows are a high risk/ high reward investments - less than 1 in 5 shows make a profit, however when there is a hit show like Lion King, Phantom of the Opera or Hamilton, the returns are enormous.

Our project predicts how the Broadway industry can maximize profit on ticket sales and theatre attendance.

Stakeholders:

- **Primary Stakeholders:** Investors and Producers
- **Secondary Stakeholders:** Theatre owners & managers, Donors and Crew

Main Question:

What are the trends in Broadway show attendance, ticket sales, playing weeks/months and expected profits given historical data of Broadway shows?

Our Approaches:

- **Trial:**

Linear Regression

- **5 Forecast Models:**

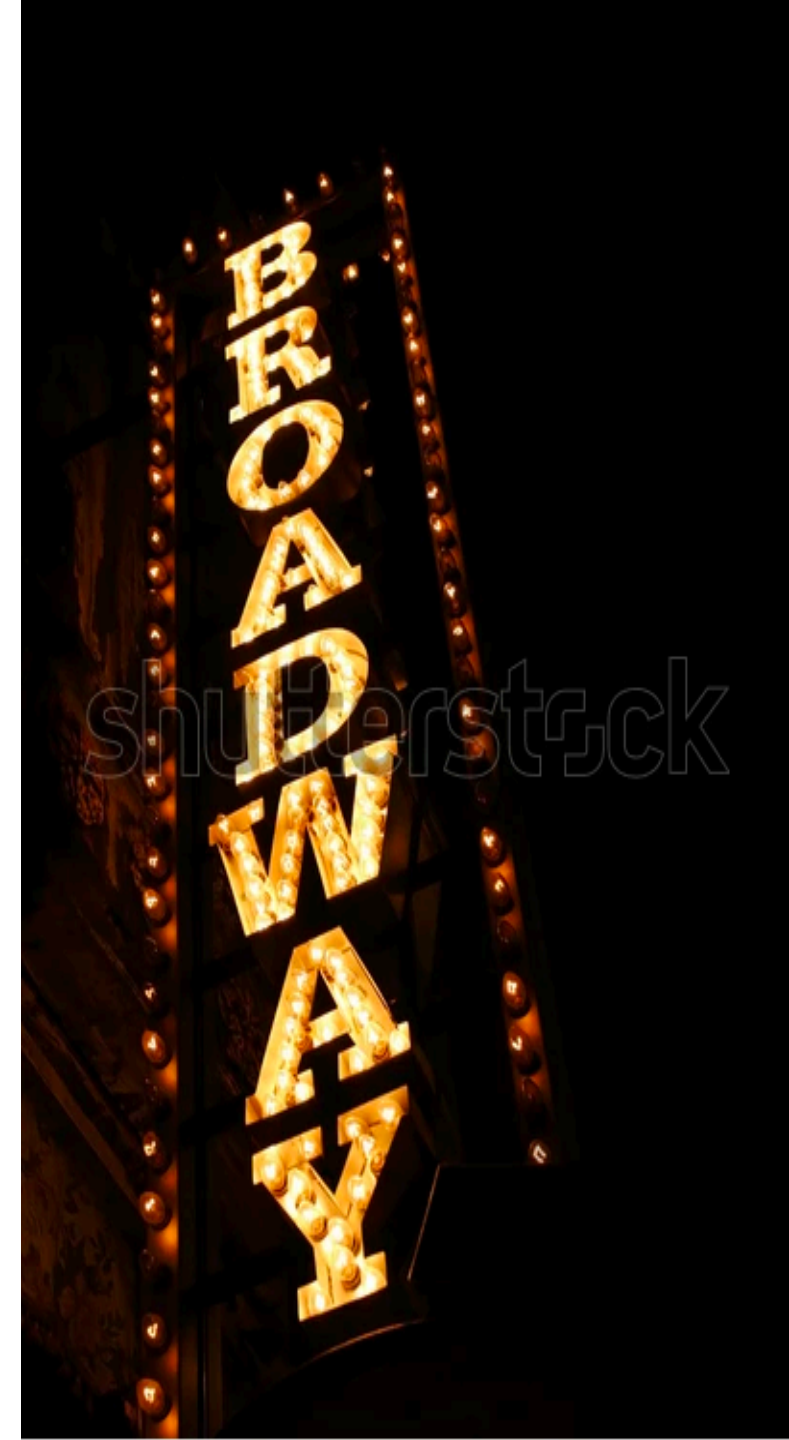
ARIMA – Forecasting and Dynamic Forecasting

LSTM

Multiple Linear Regression

FB Prophet

(NB: These 5 models were used in order to recommend the best model for time-series forecasting of the data to check for future trends and seasonality.)



DATASETS

	Year	Show_name	Potential_Gross	Difference	Average_ticket	Seats_Sold	Seats_in_theater	Previews	%cap	diff_cap
0	09-06-1985	42nd Street	0.0	0.00	30.42	9281	1655	8	70.10	0.00
1	09-06-1985	A Chorus Line	0.0	0.00	27.25	8167	1472	8	69.35	0.00
2	09-06-1985	Aren't We All?	0.0	0.00	33.75	7386	1088	8	84.86	0.00
3	09-06-1985	Arms and the Man	0.0	0.00	20.87	4586	682	8	84.05	0.00
4	09-06-1985	As Is	0.0	0.00	20.78	2938	684	8	53.69	0.00
...
47715	08-05-2022	The Phantom of the Opera	742308.2	117127.78	94.60	7847	1605	8	61.11	9.18
47716	08-05-2022	The Skin of Our Teeth	174481.0	33370.00	52.81	3304	1058	8	39.04	-8.27
47717	08-05-2022	Tina: The Tina Turner Musical	1065578.2	193232.80	122.56	8694	1478	8	73.53	14.13
47718	08-05-2022	Wicked	1341127.0	144364.00	119.06	11264	1807	8	77.92	4.53
47719	08-05-2022	for colored girls who have considered suicide/...	159092.9	-91081.60	65.71	2421	773	8	39.15	-11.77

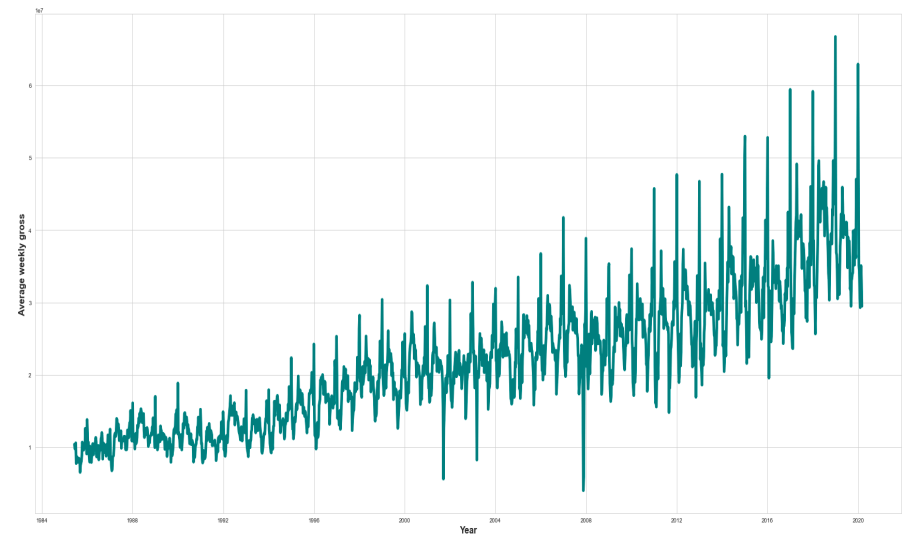
Source: <https://www.playbill.com/grosses>

LINEAR REGRESSION

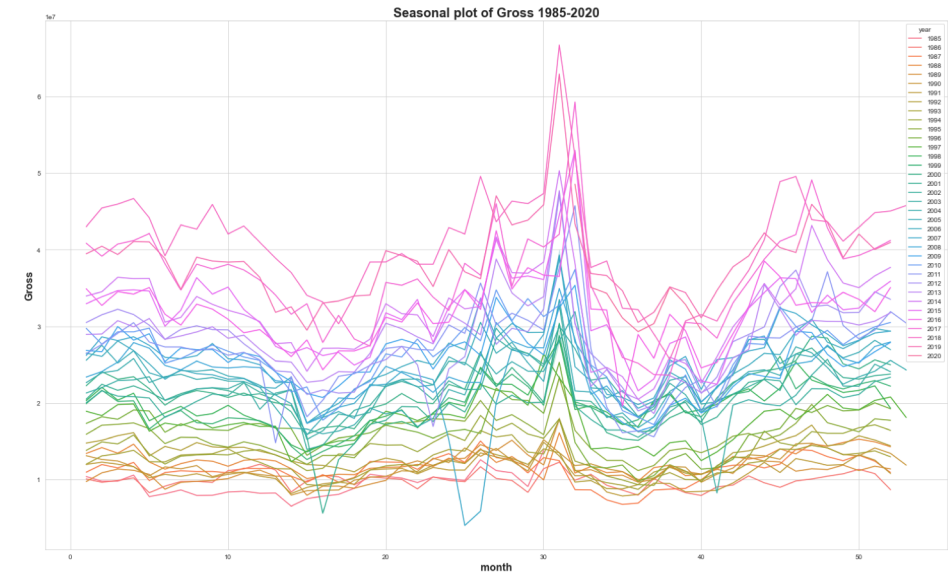
Linear Regression			
	Show Length	Total Gross: First 6 Months	Average % Capacity: First 6 Months
Average Value	35 weeks	\$11,319,833	76%
Mean Square Error	47 weeks	\$2,566,693	4%
Hamilton (Predicted)	110 weeks	\$36,502,666	93.90%
Hamilton (Actual)	>240 weeks	\$45,726,544	101%

EXPLORATORY DATA ANALYSIS: TRENDS + SEASONALITY

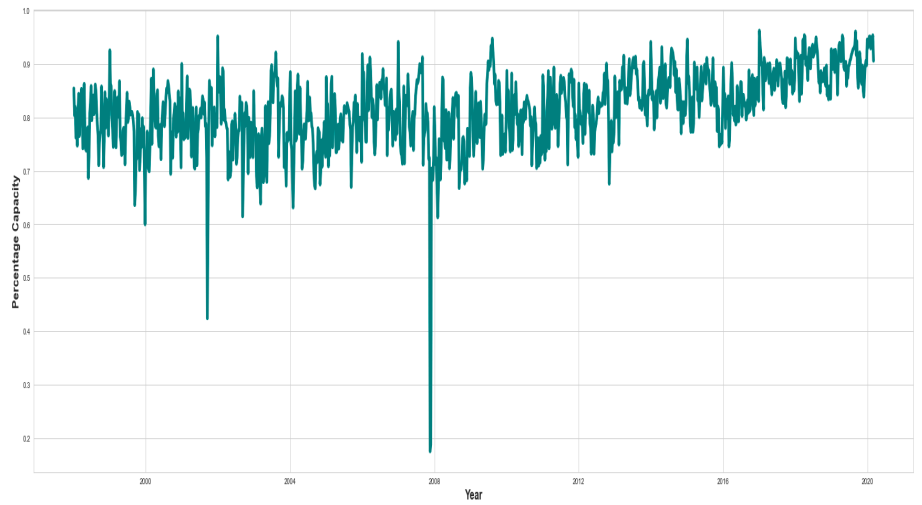
A



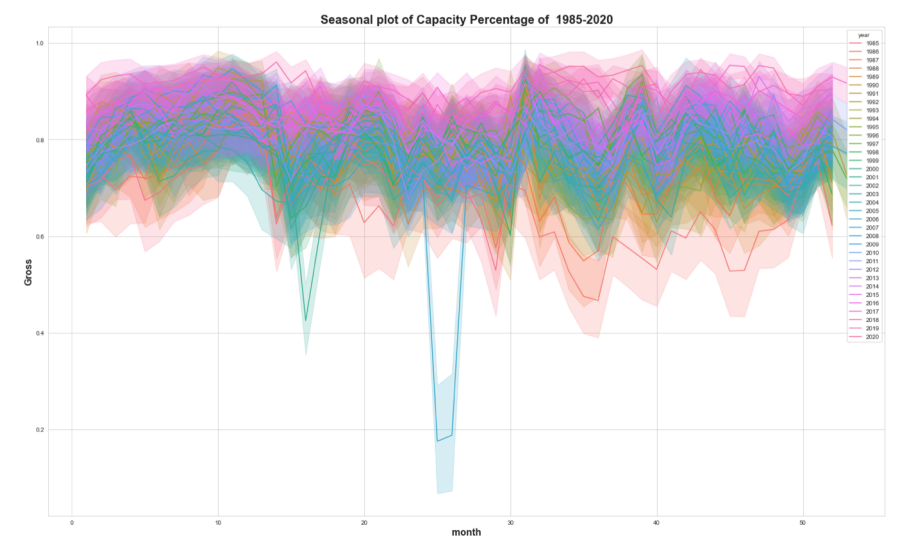
C



B

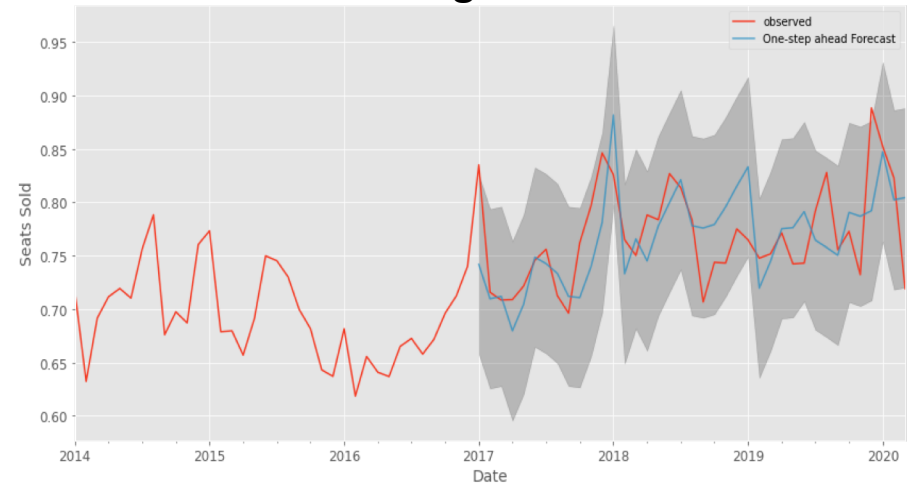


D

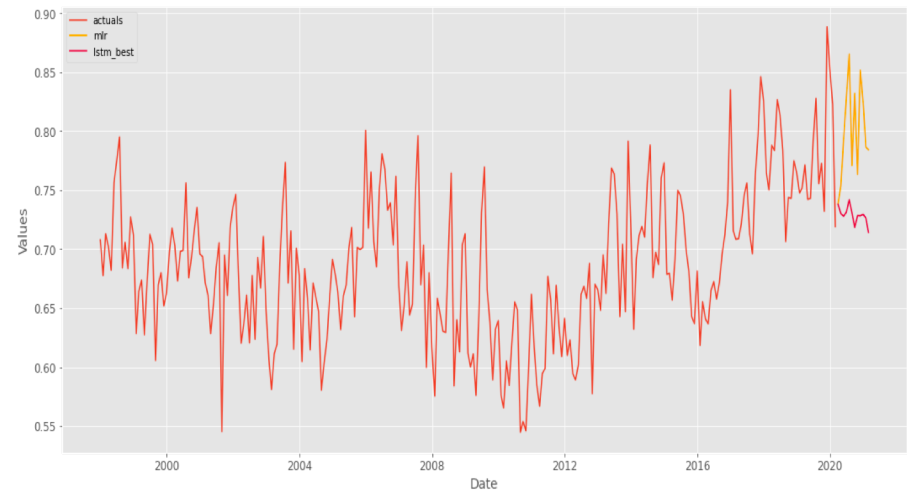
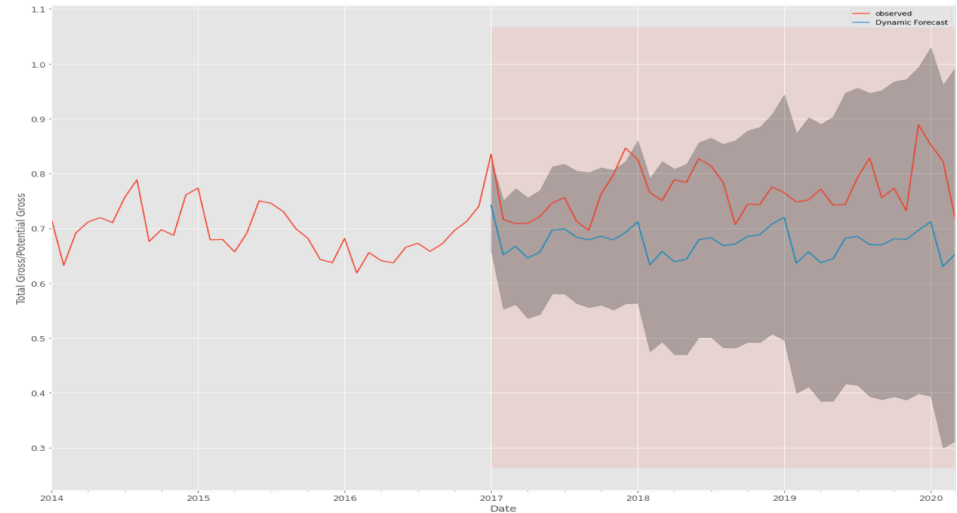


FORECASTS

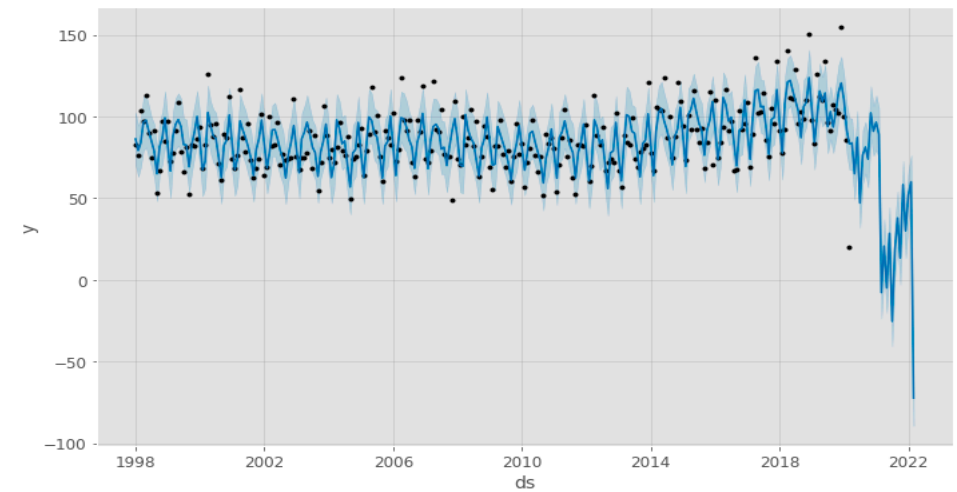
ARIMA: Forecasting



ARIMA: Dynamic Forecasting



LSTM + Multiple Linear Regression



Prophet

MODEL COMPARISON

RMSE VALUES							
	ARIMA-FORECASTING	ARIMA-DYNAMIC FORECASTING	LSTM	MLR	PROPHET		
Total Gross/Potential Gross	0.04	0.1	0.105492	0.037487	12.503776		
Average Ticket Price	8.09	17.24	11.02593	8.556401	1824.676015		
Seats Sold	358.93	519.28	673.439724	436.492349	133495.0822		

KEY TAKEAWAYS

TOP 2 PERCENTILE OF THEATERS BRING IN MORE SALES

The theaters that make up more than 2% of potential gross accounts for 68.318% of weekly actual gross of ticket sales in the Broadway Industry.

“THE PHANTOM OF THE OPERA”

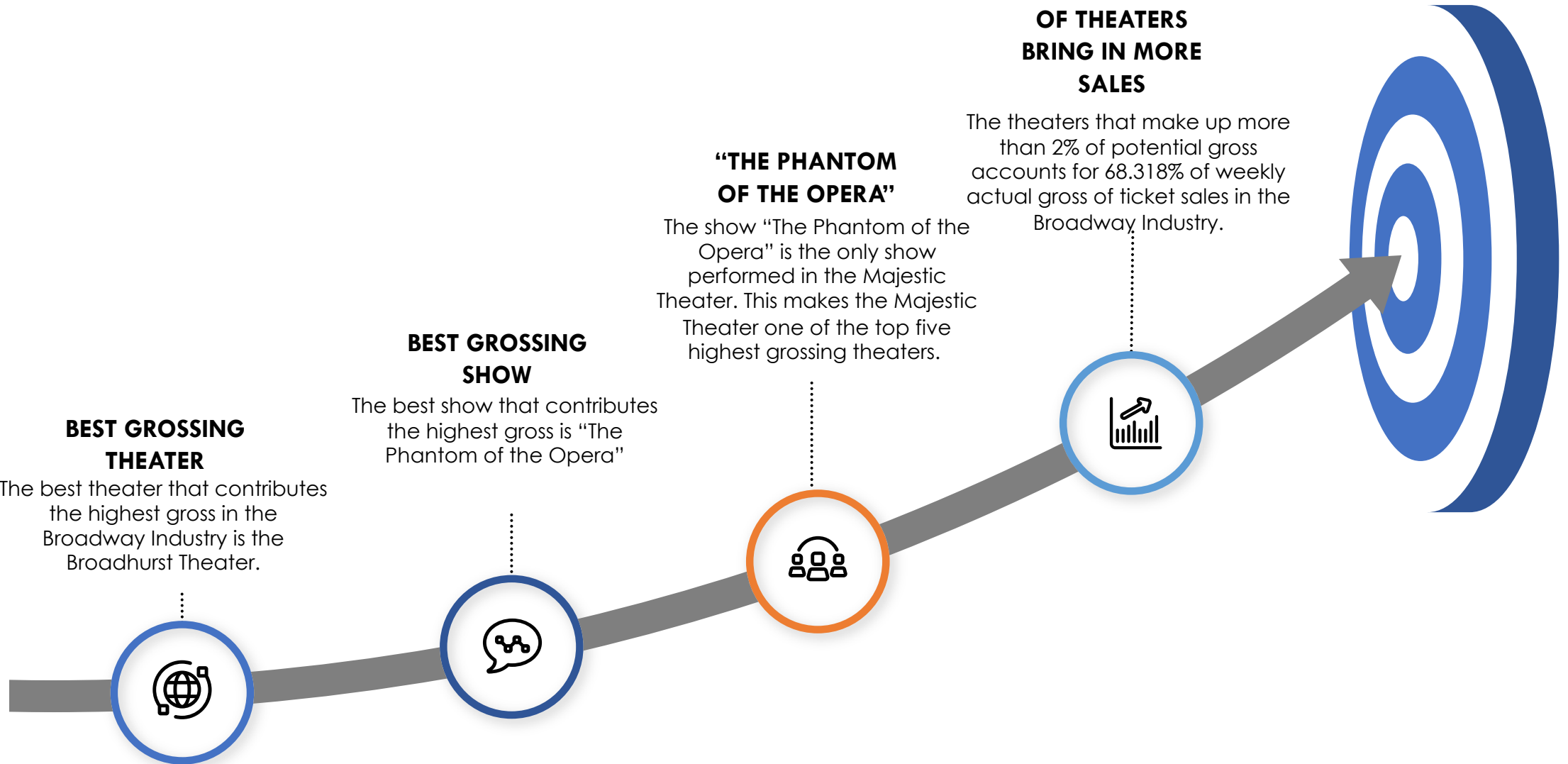
The show “The Phantom of the Opera” is the only show performed in the Majestic Theater. This makes the Majestic Theater one of the top five highest grossing theaters.

BEST GROSSING SHOW

The best show that contributes the highest gross is “The Phantom of the Opera”

BEST GROSSING THEATER

The best theater that contributes the highest gross in the Broadway Industry is the Broadhurst Theater.



FUTURE DIRECTIONS

- Improve Broadway dataset for large scale analysis
- Be able to predict how the winning of a Tony award improves the business of investors

Acknowledgement:

- We are grateful to our project mentor :

Amber Simmons, PhD

- All the team members at the Erdős Institute, especially:

Prof. Roman Holowinsky

Matt Osborne, PhD

THANK YOU FOR YOUR ATTENTION



Team Juno

Praveen Shahani
Gautam Prakriya
Michael Amoah Awuah
Moran Shechnick