Max Mitchell

Marketing and Revenue Assignment

Introduction:

Since joining the American League in 2013, the Houston Astros have been one of the most competitive in Major League Baseball. Their 2017 and 2022 World Series wins have helped build a loyal fan base, consistently ranking in the top half for attendance figures over the past decade. Still, like many other MLB franchises, the Astros have faced challenges with declining attendance in recent years despite having the second-highest attendance in the American League last year. Factors such as the COVID-19 pandemic and shifts in how fans consume sports have contributed to this trend. In response, the Astros organization is exploring ways to continue boosting ticket sales by examining how promotions, opponents, and scheduling impact attendance. Hence, this report discusses how these variables can be used to predict attendance accurately and whether Astrso should strategically implement a new dynamic pricing model.

Literature Review:

This report is presented in light of recent MLB attendance declines and widespread concerns among league ownership and fans. To help with the analysis, I reviewed three articles: "MLB attendance down 4 percent as 6 ballparks set lows" (ESPN, 2018), "Baseball Saw a Million More Empty Seats" (Allentuck & Draper, 2019), and "As Lockout Drags On, Poll Indicates MLB Could Be Charting Path to Irrelevance" (Foster, 2022), which all discuss the downward trend in league-wide attendance and provide a foundation for this proposal.

All three articles agree that attendance has steadily declined over the past decade. ESPN (2018) notes that average attendance in 2018 reached its lowest level since 2003, with six

stadiums reporting record lows. The article attributes this partially to poor weather and inconsistent team performance but also notes that these explanations only scratch the surface. Allentuck and Draper (2019) continue this discussion by highlighting that even teams with strong playoff contention, like the Astro, aren't immune to the downturn. Moreover, despite having one of the best records in baseball, Houston drew approximately 1,500 fewer fans per game in 2019 compared to previous seasons.

The articles also find different reasons for the drop in league-wide attendance. The article by Allentuck and Draper (2019) in *The New York Times* finds that teams are experimenting with flexible and dynamic ticket pricing and models to appeal to the younger fan demographic who value affordability and spontaneity. Foster's (2022) piece in *The Sporting News* states that the league's attendance and viewership problems extend beyond economics. Foster (2022) argues that MLB is becoming less culturally relevant, especially among younger audiences, based on polling data. This raises questions about sustaining a healthy fan base in the future as well as about packing stadiums today.

One major takeaway from all three articles is that traditional assumptions about what drives attendance no longer apply in the same way they once did. For example, it is no longer safe to assume that winning alone guarantees a packed stadium. This is particularly relevant for the Astros, who have consistently been one of the top-performing teams since 2017 but still saw attendance drop during that time. This points to the importance of understanding broader leagues variables, such as the day of the week, opponent quality, promotional events, weather, and fan engagement strategies, when evaluating ticket sales.

It's also important to give thought to the move toward dynamic pricing and subscription-based ticketing. According to Allentuck and Draper (2019), these more recent pricing schemes

are more concerned with fostering enduring bonds with fans than they are with optimizing immediate income. Foster (2022) notably states that no pricing strategy will reverse the downward attendance trend unless MLB addresses more fundamental issues regarding accessibility and entertainment. For now, literature suggests that the most important thing for the Astros and the MLB to reverse this downward trend in attendance is to focus on the product of the field, with promotions, flexible ticketing and targeting engagement playing a supporting role in the background.

Methods:

To analyze the factors influencing attendance across Major League Baseball, I compiled data from several sources. From *Baseball-Referenc*e, I collected attendance data for all home games from the 2022-2024 seasons. Notably, I didn't collect data going any farther back due to my assumption that Covid-19 might skew attendance data, and I collected more than one year of data to increase the sample size. The datasets for the Astros cover the day of the week, date, opponent, and attendance. I then collected data regarding the Astros promotional schedules from the 2022, 2023, and 2024 seasons. I gathered the 2022 promotional schedule data from *MLB.com*, the 2023 schedule from *MLB.com*, and the 2024 schedule from *Tickpick*.

To compare the Astros against the rest of the MLB, I collected game-level attendance data for all other 29 MLB organizations. I ran a code using Python on Jupyter Notebook to scrape the day of the week, date, opponent, and attendance for all 29 teams for the 2022, 2023, and 2024 seasons from B*aseball-Reference* into a single output Excel spreadsheet. Notably, I was able to download and modify the tables for just the Astros, but I did not want to do this for the 87 other web pages, and hence, I used a loop to obtain the rest of the MLB attendance data. I also left out promotional data for the league-wide attendance data.

Then, I used Python in a Jupyter notebook to create a merged Excel file with a dummy variable for the Astros versus all other MLB organizations, comparing the extent to which the main variables (day, opponent, time of year) drive attendance. This way, I can perform league-wide analysis related to ticket sales.

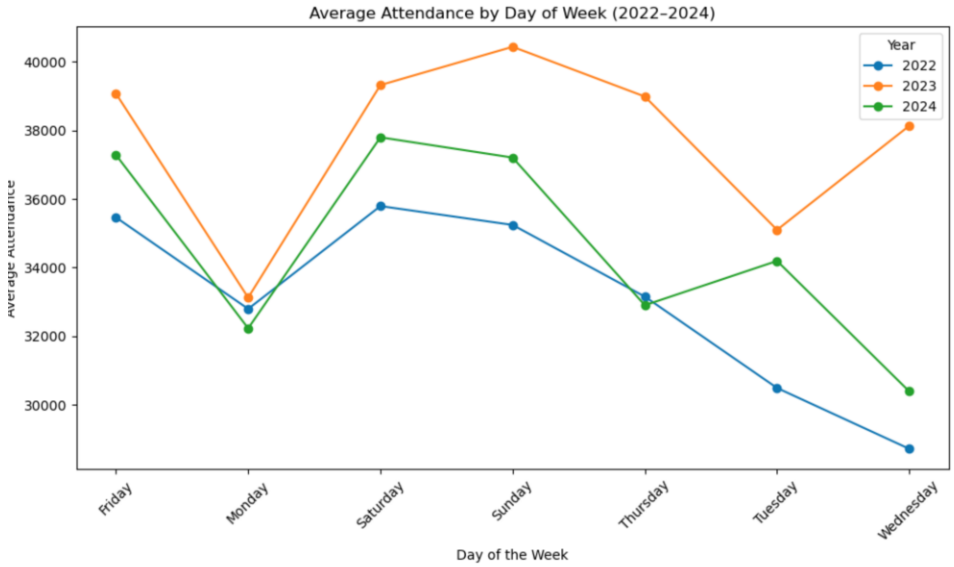
The last data I gathered was the win percentage and season-long franchise season record of the Astros and 29 other MLB franchises from the 2022, 2023, and 2024 seasons, as reported by Baseball Reference. I merged this into the Astros attendance data to see to what extent the other MLB franchises' winning percentages affect Astros home game attendance.

In total, my data includes a 2022, 2023, and 2024 Excel spreadsheet for the Houston Astros attendance information, an Excel spreadsheet for the overall MLB attendance from 2022-2024 for the other 29 teams, a CSV file for the merged league-wide attendance with the Astros dummy variable, and overall separate excel files for the overall MLB records for the 2022, 2023, and 2024 seasons. After collecting all the data, I began statistical analysis to evaluate the extent to which the chosen variables impact ticket sales and attendance.

Results:

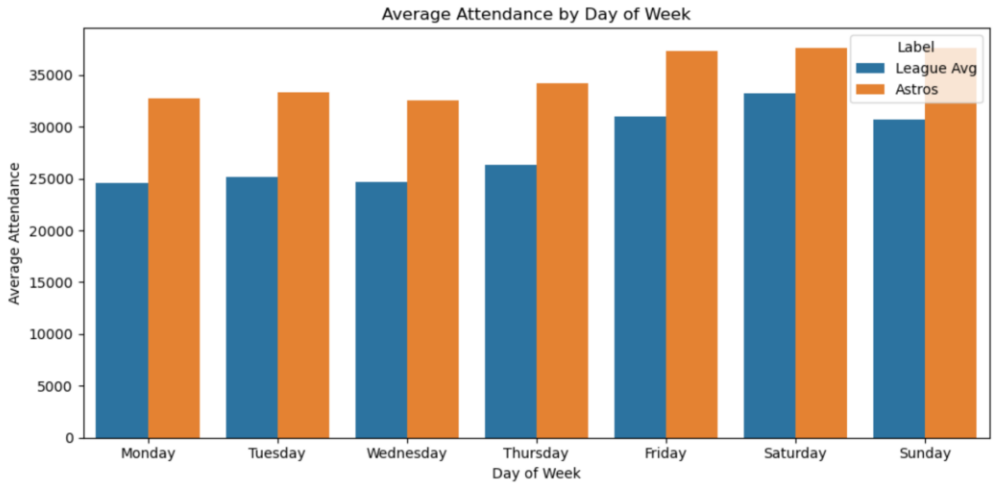
First, I analyzed the predictive value of the chosen variables for attendance among the Astros alone before moving on to a league-wide analysis. Based on Figure 1 below, we can see a clear trend of higher attendance levels at Astros home games over the weekend than during the week from the 2022 to 2024 seasons. Over the three years, Saturdays drew the highest average crowds at approximately 37,635 attendees, closely followed by Sundays at 37,627 and Fridays at 37,273, with weekend games averaging over 37,000 attendees. On the contrary, weekdays experienced lower attendance, with Thursdays averaging 35,014, Tuesdays 33,263, and Mondays 32,719. Wednesdays consistently recorded the lowest average attendance across the three years,

at roughly 32,415. Hence, Astros game attendance exhibits a clear pattern strongly influenced by the day of the week.



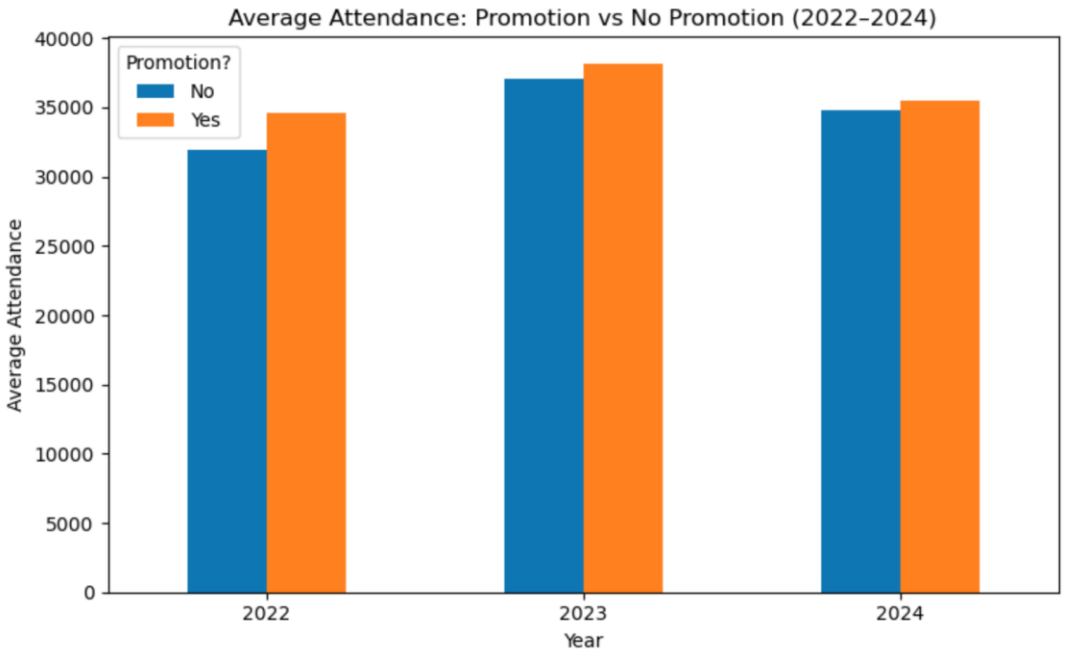
(Figure 1)

When I analyzed the impact of attendance per day of the week for all 29 other MLB teams, excluding the Astros, I found a similar trend of significantly higher average attendance for weekend games compared to weekdays. Both the Astros and the entire MLB experience higher attendance on Fridays through Sundays, with Saturdays being the peak.



(Figure 2)

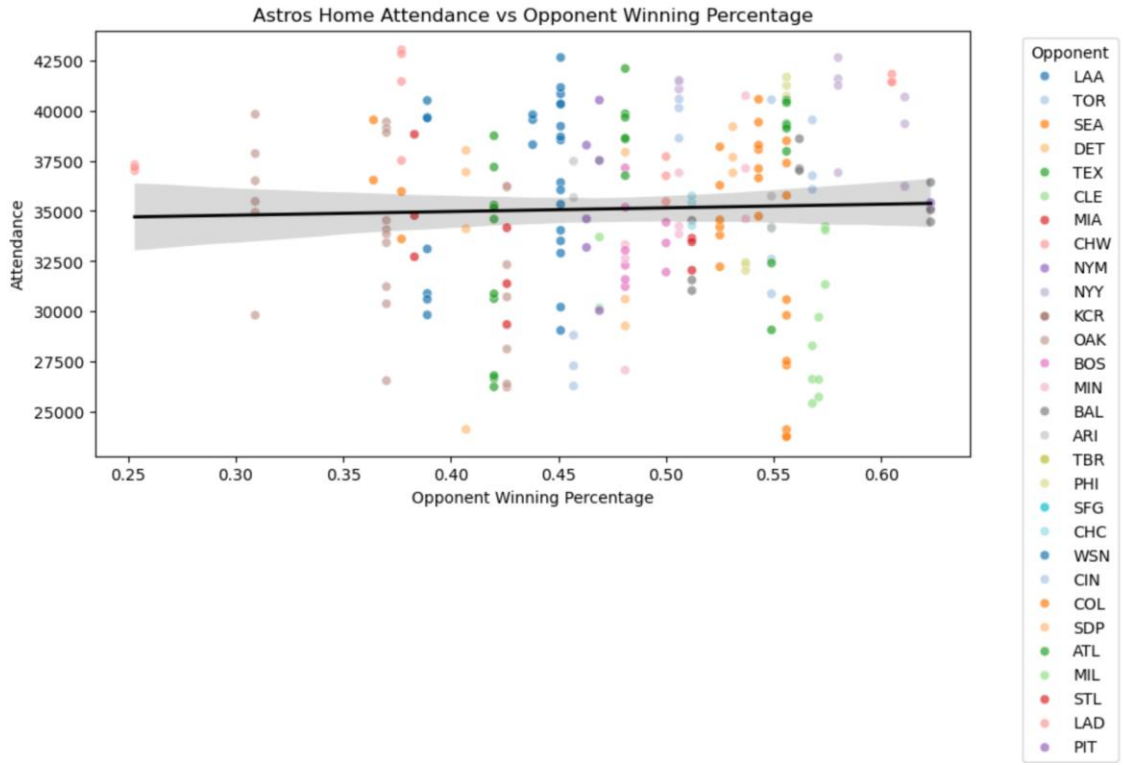
Based on Figure 3, Astros games featuring promotions consistently drew higher average attendance than those without promotions across 2022, 2023, and 2024. In 2022, promotional events (averaging 34,585 attendees) resulted in a notable increase compared to non-promotional games (averaging 31,971 attendees). While promotions continued to boost attendance in subsequent years, the magnitude of this effect diminished. In 2023, promotional games averaged 38,187 attendees, compared to 37,102 attendees for non-promotional games, representing a smaller increase of approximately 1,085. Similarly, in 2024, promotional games averaged 35,479 attendees versus 34,790 attendees for non-promotional games. Despite these year-over-year fluctuations in overall attendance, the positive impact of promotional schedules on drawing larger crowds remained a consistent pattern.



(Figure 3)

The last variable that I analyzed to evaluate the extent in which factors can predict attendance for the Houston Astros was their opponent and win percentage. For this variable, I created a scatter plot instead of a barplot. The scatter plot reveals a very weak positive

relationship between Astros home attendance and opponent winning percentage. The slight upward slope of the regression line suggests that games against stronger opponents may attract slightly larger crowds; however, the effect is minimal. Attendance figures are widely scattered across all levels of opponent winning percentage, indicating that other factors—such as day of the week, promotions, rivalries, and opponent popularity—likely play a much more significant role in driving attendance. The broad confidence interval around the regression line further confirms the low predictive power of the opponents winning percentage alone.



(Figure 4)

I then ran a multivariable regression to see the extent to which promotions, day of the week, and opponent winning percentage impact home attendance levels at Astros games. The regressions result is as follows:

Dep. Variable:	Attendance	R-squared:	0.232
Model:	OLS	Adj. R-squared:	0.202

Method:	Least Squares	F-statistic:	7.817
Date:	Sun, 08 Jun 2025	Prob (F-statistic):	3.66e-09
Time:	09:20:02	Log-Likelihood:	-2093.8
No. Observations:	216	AIC:	4206.
Df Residuals:	207	BIC:	4236.
Df Model:	8		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	3.527e+04	1831.722	19.258	0.000	3.17e+04	3.89e+04
Promotion?	567.2369	579.621	0.979	0.329	-575.481	1709.955
Opponent Win %	3301.1848	3479.905	0.949	0.344	-3559.415	1.02e+04
Day of Week_Monday	-4389.9994	1058.393	-4.148	0.000	-6476.611	-2303.388
Day of Week_Saturday	207.6919	972.435	0.214	0.831	-1709.453	2124.837
Day of Week_Sunday	151.2732	973.181	0.155	0.877	-1767.345	2069.891
Day of Week_Thursday	-2776.4711	1165.350	-2.383	0.018	-5073.947	-478.995
Day of Week_Tuesday	-3559.8571	975.389	-3.650	0.000	-5482.827	-1636.888
Day of Week_Wednesday	-4497.7864	980.739	-4.586	0.000	-6431.303	-2564.270

Omnibus:	1.946	Durbin-Watson:	0.959
Prob(Omnibus):	0.378	Jarque-Bera (JB):	1.787
Skew:	-0.124	Prob(JB):	0.409
Kurtosis:	2.630	Cond. No.	18.0

The regression analysis reveals that approximately 23.2% of the variability in Astros home attendance can be explained by the included factors, with the overall model being statistically significant ($p<0.001$). Of the variables examined, Day of the Week is the most significant factor influencing Astros home attendance. Specifically, games played on weekdays (Monday, Tuesday, Wednesday, Thursday) are associated with a statistically significant decrease in attendance compared to Fridays (the baseline day). The model also suggests that attendance on Saturdays and Sundays is not statistically different from Fridays, indicating that the biggest attendance dips occur during the core weekdays. Hence, there is a statistically significant difference between weekdays and weekends when predicting Astros home game attendance.

Notably, promotions and opponent winning percentage didn't have a statistically significant impact on attendance in this model. Despite promotions and opponent winning percentages showing positive coefficients, these relationships are not significant in the regression model. Therefore, the Day of the Week plays the most prominent role in predicting Astros home attendance among the three factors considered in this regression, with specific weekdays showing a consistent and statistically significant negative effect on attendance.

I then wanted to see if this trend of Day of the Week having the most significant impact on attendance was also found throughout the other 29 MLB organizations. Hence, I performed an ANOVA test. The ANOVA results suggest a statistically significant difference in average attendance across the different days of the week for the 29 other MLB teams. The F-statistic of 117.138 is very high, suggesting that there's a substantial amount of variability in attendance that can be attributed to the day of the week. Moreover, The p-value (PR(>F)) of approximately 1.965×10^{-141} is extremely small. Since this p-value is far below common significance levels (e.g., 0.05 or 0.01), we can confidently reject the null hypothesis that there is no difference in average attendance among the different days of the week. In practical terms, this means that the day on which an MLB game is played has a clear and measurable impact on the number of attendees for these teams. As I did not collect promotional data for all other 29 times over the 2022, 2023, and 2024 seasons, and opponent winning percentage didn't play a significant role in affecting Astros attendance, I can confidently conclude that day of the week plays the biggest role in driving attendance not only for the Houston Astros but also across the rest of the MLB.

ANOVA Results for Day of Week Impact on MLB Attendance (Excluding Astros):

	sum_sq	df	F	PR(>F)
C(Q("Day of Week"))	8.004396e+10	6.0	117.138473	1.965369e-141
Residual	7.952795e+11	6983.0	NaN	NaN

(Figure 5)

Dynamic Ticket Pricing

Now that the Houston Astros are winning and demand is always high—especially for weekend games and big series—dynamic pricing for home games is a no brainer. This approach allows the Astros to adjust prices in real time based on demand, opponent, day of the week and other factors to optimize attendance and revenue. For example, games against marquee opponents or those late in the season with playoff implications could be priced at a premium, while midweek matchups against lower-profile teams might be priced more affordably to boost turnout.

Dynamic pricing also encourages earlier ticket purchases, as fans are motivated to buy before prices potentially rise. This purchasing behavior provides the organization with valuable forecasting insights for staffing, concessions, and other operational needs. Moreover, pricing decisions can be informed by historical attendance trends—like the clear weekday vs. weekend disparity revealed in this study—as well as other variables such as weather forecasts and team performance momentum.

The Astros should also consider the potential drawbacks of implementing a new dynamic pricing system, which could cause fans to feel overcharged if it is not implemented transparently. To address this, I propose that the Astros adopt a phased strategy, starting with the games with the highest demand or simply using tiered pricing based on the day of the week the home game is. This would allow the Astros to assess the effectiveness of the dynamic ticket pricing and obtain feedback before expanding their model.

Ultimately, dynamic pricing fits with the patterns found in this analysis, especially the statistically significant impact of the day of the week on attendance. Through the adoption of a fully transparent and data-driven ticket sales pricing model, the Astros have the opportunity to

enhance ticket sales revenue and continue to be one of the highest attending crowds in Major League Baseball.

Recommendations, Limitations, and Future Directions

Based on the findings from this analysis, a dynamic ticket pricing strategy is recommended for the Houston Astros. This approach would allow the team to adjust ticket prices based on game-specific factors, with a particular emphasis on the day of the week—identified as the most significant driver of attendance. Weekday games, particularly those on Mondays through Wednesdays, should be priced more competitively to increase demand, while weekend games, which usually attract larger audiences, can be priced higher. A staged rollout might be used to progressively use this method, beginning with games that have historically had significant demand or attendance swings. Throughout this phased rollout, the Astros can increase revenue efficiency and maintain more seats filled at Minute Maid Park throughout the season.

This analysis also presents limitations due to a lack of acquired data. The main limitation in my eyes is the lack of promotional data for the other 29 Major League teams. I could have gone through the promotional schedules for all 29 teams over the past three seasons, but concluded that this task was far too time consuming and not necessary for my current analysis as the promotional schedule didn't have a statistically significant impact on attendance for the Astros. I also didn't include factors such as weather, game start times, and star player injuries which could have expanded my model. Lastly, as the results are generally focused on the Astros, it doesn't fully capture fan behavior which might vary for each MLB organization and city.

To address these limitations, future research should broaden the dataset to include more factors that might influence attendance. As mentioned above, these factors should include a complete league-wide promotional schedule for all teams, game times, player availability, and

more to provide a more accurate and complete model. Moreover, the model can be expanded over a longer period of time, such as five to ten years, to provide more statistical significance in regression models. Furthermore, I could have used more advanced statistical techniques to fully account for the nonlinear effects in the model. Future research should also include fan surveys and sentiment analysis to analyze fans perceptions of dynamic ticket pricing models and their purchase behavior and decisions to provide more context than what is solely captured in the numerical data.

To begin this future research, the next steps include expanding the current dataset by gathering the missing promotional and environmental variables across the league. A centralized database should be developed to integrate all relevant information from 2022 through 2024, or even an expanded timeline of 2020 through 2024 or 2015 through 2024. Once complete, new models can be tested to compare performance with the existing regression and to identify additional significant predictors of attendance. Collaboration with MLB ticketing offices or third-party vendors like Ticketmaster could also help access more granular ticket sales data. Finally, a pilot program could be launched using dynamic pricing for select Astros games to assess the real-world impact and refine the strategy prior to full-scale implementation.

Sources:

Aboreden, A.-A. (2024, March 12). Houston Astros promotions and Giveaway Schedule . Tickpick. https://www.tickpick.com/blog/astros-promotional-and-giveaway-schedule/

Allentuck, D., & Draper, K. (2019, September 29). Baseball saw a million more empty seats. does it matter? - The New York Times. The New York Times. https://www.nytimes.com/2019/09/29/sports/baseball/mlb-attendance.html.

ESPN Internet Ventures. (n.d.). MLB attendance down 4 percent as 6 ballparks set lows. ESPN. https://www.espn.com/espn/wire?section=mlb&id=24874152.

Foster, J. (2022, January 2). As lockout drags on, poll indicates MLB could be charting path to irrelevance. Sporting News. https://www.sportingnews.com/us/mlb/news/as-lockout-drags-on-poll-indicates-mlb-could-be-charting-path-to-irrelevance/1qbmfwrcjfdsr16izao285afnw.

Houston Astros schedule. MLB.com. (n.d.). https://www.mlb.com/astros/schedule.