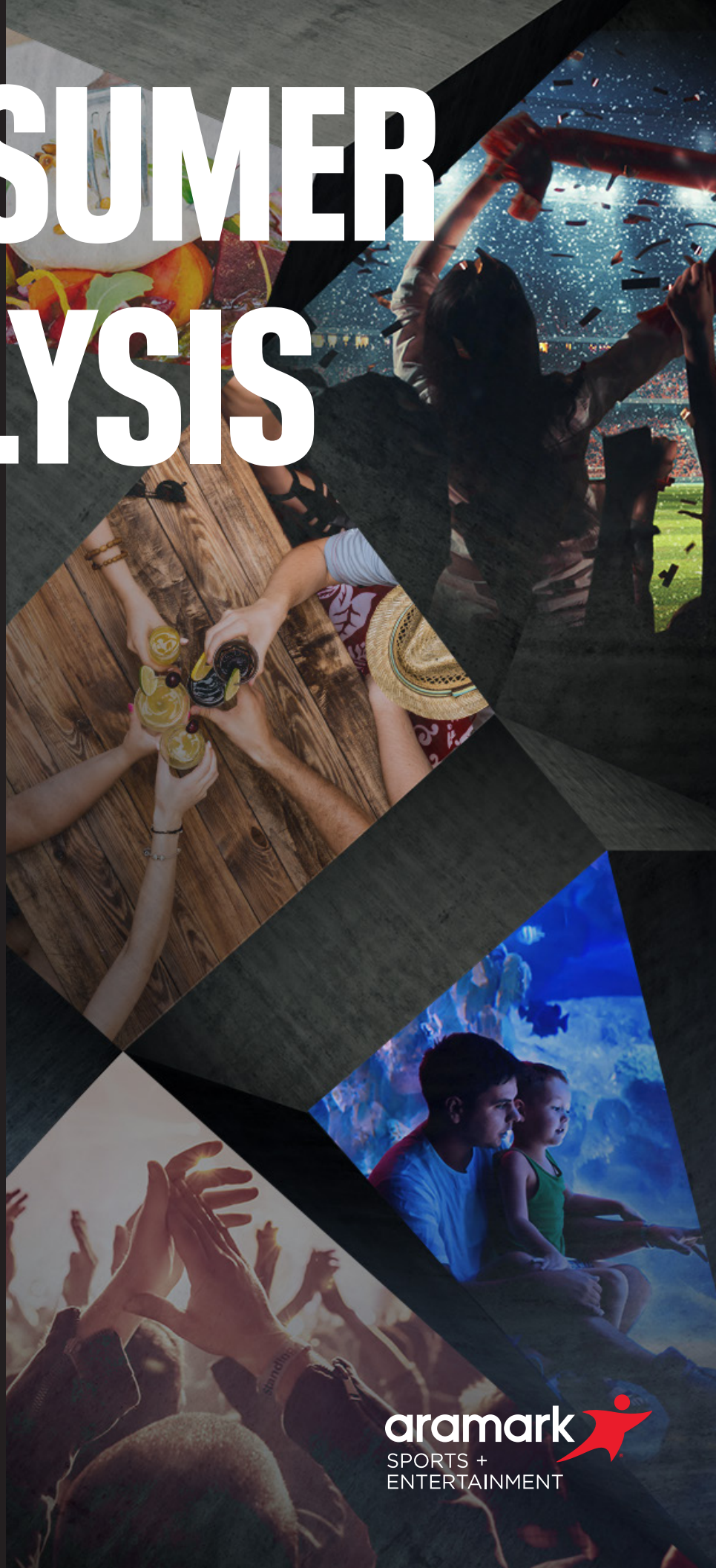
CONSUMER ANALYSIS POC



DS

DATA › SCIENCE

APRIL 8, 2024

Authored by: Jacob M. Pearl Ph.D.

aramark SPORTS + ENTERTAINMENT

INTRODUCTION

At Aramark Sports & Entertainment we are committed to thoroughly understanding how business decisions and external factors impact fan experiences and client outcomes. To provide the most rigorous insights, our data science team goes beyond transaction level data and interrogates behavior change at the individual consumer level. Combining several sources of data, we identify and track customers at our venues throughout their game day experiences, between events, and even across multiple seasons. This approach surfaces insights previously obfuscated by anonymous transaction level data. For example, we are now able to know the preferences and characteristics of individuals by tying together their purchase history. Furthermore, we can now interrogate whether these preferences are stable across events, and whether these preferences vary in response to changes introduced at our venues.

In this report, we provide an initial case study that represents the capabilities of this novel approach. Specifically, in this report we provide a rigorous and statistically robust investigation of consumer behavior at four MLB venues before and after the introduction of the pitching clock.

In 2023 the MLB implemented several rule changes, including a new 20 second pitching clock that had a direct impact on rate-of-play and duration of games; on average 9-inning MLB games in 2023 were 24 minutes shorter than previous seasons (Grant, 2023). While the effect of this rule change on game play is established, it is not yet known what effect shorter and faster-paced games have on consumer behavior.

Aramark Sports & Entertainment (S&E) is directly invested in understanding consumer behavior in order to better serve both our clients and MLB consumers. Thus, **it is of utmost importance that our organization provide rigorous and meaningful analysis of the effects of the 2023 rule changes on consumer behavior.**

Included in this initial POC analysis are hypotheses relating to repeat purchasing behavior. The questions and hypotheses interrogated here are non-inclusive of all analysis planned for this work. Throughout this report, we make note of specific limitations that may be overcome by additional partnership with our MLB teams and more thorough quasi-experimental design. We acknowledge that the observational nature of this data and lack of a control group makes causal inference difficult; we discuss potential confounds where appropriate and maintain a relational and circumscribed interpretation of our results.

DATA

Total Sample. Through electronic point-of-sale (POS) systems, Aramark S&E collects and stores large amounts of feature-rich data related to consumer transactions. Additionally, payment system information (e.g. Freedompay credit card reader) is collected at many of our venues, which allows partial identification of consumers. For this analysis we combined all available POS data from the 2022 and 2023 MLB regular season at four venues with payment system information to tie transactions to unique individuals.

Not all POS transactions were tied to individual consumers. Furthermore, many transactions were associated with default bank information (e.g. Visa CreditCard Holder, Chase Cardholder), or were associated with partial credit card numbers (first 4 and last 6 digits) that were linked with many different cardholder names. We removed transactions from our final dataset that were tied with default bank information or were tied to partial credit card numbers associated with 4 or more names. Table 1 provides a summary of our total transaction sample size before pre-processing and after.

Venue	Season	Total Transactions	Filtered Transactions	% Filtered
Venue 1	2022	982,089	282,915	29%
Venue 1	2023	1,284,241	460,370	36%
Venue 2	2022	590,043	135,737	23%
Venue 2	2023	514,612	143,345	28%
Venue 3	2022	2,057,931	576,509	28%
Venue 3	2023	1,931,292	648,343	34%
Venue 4	2022	1,985,810	537,507	27%
Venue 4	2023	1,975,067	597,009	30%
TOTAL		**11,321,085**	**3,381,735**	**30%**

Table 1. POC sample.

It is important to acknowledge here a significant limitation in our analysis. Specifically, the data used in this analysis is censored, in that we are only able to analyze the behavior of consumers who *did* make a purchase at an event, and not those who abstained. As such, we are not able to answer the critical question of whether the 2023 rule changes effected the likelihood of a consumer making *any* purchases at an event. To elaborate, we are not able to directly infer that the 2023 rule changes were responsible for any individual purchasing concessions in 2022 and not 2023; this individual simply may have not attended a game in 2023. We believe that with further partnership with our MLB clients, we may obtain additional information about attendees and answer this critical question.

TREATMENT VALIDITY: GAME DURATION

Prior to testing any hypotheses, we must first verify that the assumed treatment effect of the 2023 rule change is valid. More precisely, while across the MLB game duration has on average shortened, it is an assumption that game times have indeed shortened as a result of the rule changes imposed by the MLB at the four specific venues used in this POC analysis. Here we assess the validity of this assumption (Analysis here is limited to games with no extra innings).

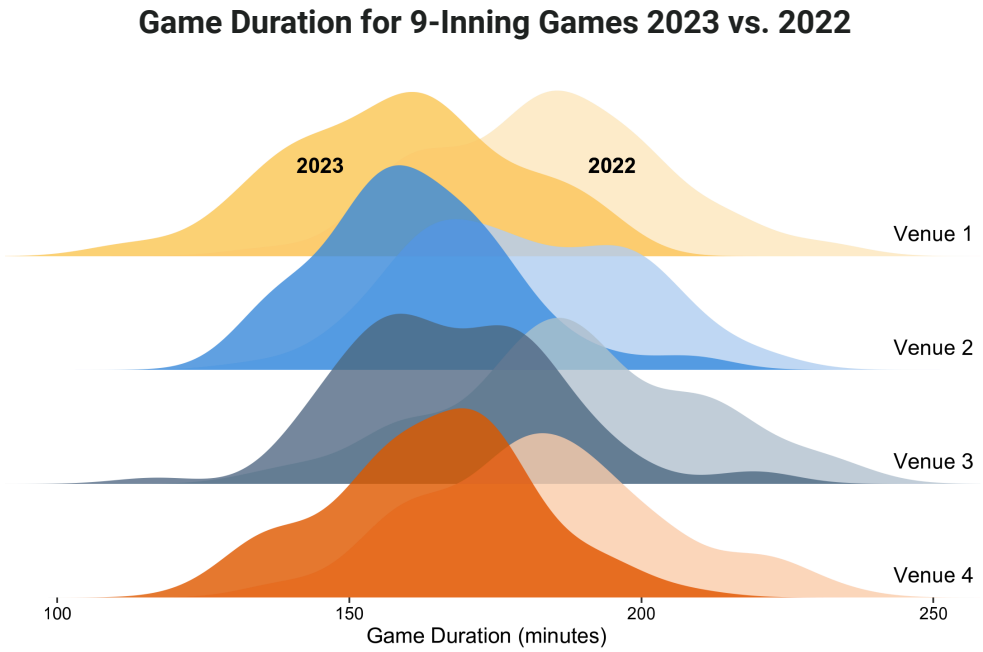


Figure 1. Distributions of 9-inning game duration.

Indeed, there is a significant decrease in game duration for all four of the venues investigated in this study (see appendix for statistics). This analysis validates the underlying assumption of this study: The MLB rule changes shortened regular season games in 2023.

HYPOTHESIS TESTS

H_1 - CONSUMERS WERE LESS LIKELY TO MAKE MULTIPLE PURCHASES IN 2023 COMPARED TO 2022

An intuitive question to begin with is whether shorter game times influence the number of transactions an individual consumer will make at a single event. We hypothesize that shorter games reduce the likelihood that an individual will make multiple purchases at an event.

To perform this analysis, we categorize consumers into 'repeat purchaser' or 'single purchaser'. Specifically, we categorize a consumer as a repeat purchaser if they made more than one transaction at *any* event within a season, and then perform a simple chi-squared test to understand if there was a significant decrease in the proportion of repeat purchasers from 2022 to 2023. We perform this test for all data together and for each venue separately, understanding that there may be individual venue-level variation in these proportions.

OVERALL

Season	Repeat Purchaser False	Repeat Purchaser True
2022	529,939 (61%)	339,487 (39%)
2023	644,645 (62%)	387,715 (38%)

VENUE 1

Season	Repeat Purchaser False	Repeat Purchaser True
2022	94,554 (60%)	62,912 (40%)
2023	161,009 (63%)	94,164 (37%)

VENUE 2

Season	Repeat Purchaser False	Repeat Purchaser True
2022	51,528 (65%)	28,166 (35%)
2023	57,088 (66%)	28,899 (34%)

VENUE 3

Season	Repeat Purchaser False	Repeat Purchaser True
2022	188,350 (58%)	133,535 (42%)
2023	232,770 (63%)	135,241 (37%)

VENUE 4

Season	Repeat Purchaser False	Repeat Purchaser True
2022	195,507 (63%)	114,874 (37%)
2023	193,778 (60%)	129,411 (40%)

Table 2. Probability of repeat purchase, overall and broken out by venue.

We find that overall, there *was not* a meaningful change in the probability of a consumer purchasing more than once in a game from 2022 and 2023. There were, however, important differences within some venues. Specifically, Venue 1 saw a 3% decrease in the probability of repeat purchasing behavior from 2022 to 2023, and Venue 3 saw a 4% decrease. (For overall and venue level chi-sq statistics and alpha values see appendix table 2A.)

H_2 - CONSUMERS ARE MAKING THEIR FIRST PURCHASE AT A DIFFERENT TIME IN 2023 COMPARED TO 2022

Here we test whether consumers are making their first purchase at different times relative to the duration of a game. More specifically, to account for differences in game duration we operationalize the time of the first transaction as the proportion of a games total time that has passed (i.e. time from start of game / total game duration). Simply calculating 'time since start of game' would naturally result in later times for 2022 because these games are known to be longer (e.g. a transaction in the 9th inning in 2022 will naturally be further from the start of the game than in 2023). Thus, "earlier" or "later" is here a relative term that innately accounts for the duration of a game.

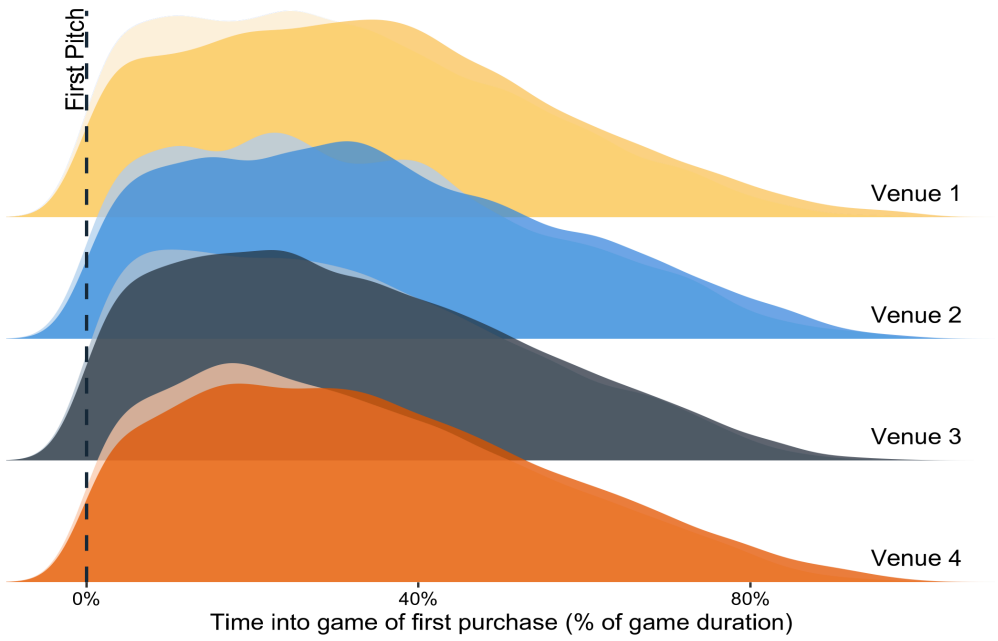


Figure 2. Relative time of first purchase after first pitch 2022, 2023.

Counter to our expectations we found that first purchases were made at a slightly *later* time in 2023. Specifically, consumers made their first purchase (or only purchase) 2% later in 2023 compared to 2022 ($t = 13.04$, $p < 0.001$). Across all four venues, we see that this general trend holds. Thus, we find support for our hypothesis, that time of first purchase changed from 2022 to 2023. We believe that while statistically significant, this overall 2% shift is not particularly meaningful in the context of consumer behavior.

$H_{2.2}$ - CONSUMERS ARE MAKING MORE PURCHASES BEFORE FIRST-PITCH IN 2023 COMPARED TO 2022

Across all four of the venues in this analysis, we see that consumers are making their first purchase slightly later after first pitch. This result is counterintuitive – we expected that consumers would be making purchases earlier to accommodate the shorter game time. One reason for this result may be the omission of orders made prior to first-pitch. Thus, we also interrogate whether consumers were making their first purchase earlier by operationalizing the question in a different way: whether or not first purchase was made before first-pitch. We hypothesize that consumers in 2023 were more likely to make their first (or only) purchase prior to first-pitch relative to 2022.

OVERALL

Season	Before first pitch: False	Before first pitch: True
2022	523,148 (63%)	314,901 (37%)
2023	642,500 (59%)	450,152 (41%)

VENUE 1

Season	Before first pitch: False	Before first pitch: True
2022	82,013 (53%)	71,758 (47%)
2023	151,626 (54%)	132,525 (46%)

VENUE 2

Season	Before first pitch: False	Before first pitch: True
2022	53,585 (65%)	27,662 (35%)
2023	58,070 (64%)	31,945 (36%)

VENUE 3

Season	Before first pitch: False	Before first pitch: True
2022	177,587 (60%)	120,583 (40%)
2023	210,383 (55%)	16,611 (45%)

VENUE 4

Season	Before first pitch: False	Before first pitch: True
2022	210,963 (69%)	94,898 (31%)
2023	222,421 (65%)	119,671 (35%)

Table 3. Probability of repeat purchase, overall and broken out by venue.

We find strong support for our hypothesis that consumers were more likely to make their first purchase before first pitch in 2023 relative to 2022. We saw a 5% shift in the probability of a consumer making their order before first pitch. Between venues there was significant variation however, with Venue 1 and Venue 2 having little change, and Venue 3 and Venue 4 having a 5% and 4% shift to more pre-game first purchases respectively. (For overall and venue level chi-sq statistics and alpha values see appendix table A3.)

H_3 - REPEAT TRANSACTION CONSUMERS ARE LESS LIKELY TO VISIT MULTIPLE STANDS IN 2023 RELATIVE TO 2022

Consumers have many concessions options when attending an MLB game. Whether the 2023 rule changes and resulting shorter games affected how flexible consumers feel to explore these options is an open question. We test here whether consumers who purchased more than once at a given event were less likely to complete a transaction at more than one distinct concession stand. Like the initial repeat consumer analysis, we categorize individuals into "multiple stand" consumers if they made a second purchase at more than one stand at *any* event they attended.

OVERALL

Season	Before first pitch: False	Before first pitch: True
2022	46,310 (14%)	293,177 (86%)
2023	54,000 (14%)	333,715 (86%)

VENUE 1

Season	Before first pitch: False	Before first pitch: True
2022	9,098 (14%)	53,807 (86%)
2023	16,956 (18%)	77,188 (82%)

VENUE 2

Season	Before first pitch: False	Before first pitch: True
2022	3,439 (12%)	24,716 (88%)
2023	3,205 (11%)	25,668 (89%)

VENUE 3

Season	Before first pitch: False	Before first pitch: True
2022	19,743 (15%)	113,834 (85%)
2023	20,667 (15%)	114,610 (85%)

VENUE 4

Season	Before first pitch: False	Before first pitch: True
2022	14,039 (12%)	100,820 (88%)
2023	13,172 (10%)	116,249 (90%)

Table 4. Probability of repeat purchaser visiting multiple stands, overall and broken out by venue.

We found that there was no change from 2022 to 2023 in the probability that a repeat purchaser would visit multiple stands. Furthermore, we see little variation across the four venues in this analysis, with the exception of Venue 4 which saw a 2% *increase* in the probability that a repeat purchaser would visit multiple stands. (For overall and venue level chi-sq statistics and alpha values see appendix table A4.)

H_4 - REPEAT PURCHASERS ARE WAITING LESS TIME BETWEEN TRANSACTIONS IN 2023 COMPARED TO 2022

For consumers who wish to make more than one transaction in a game, they may feel that there is less time to make their second purchase. We hypothesize that consumers in 2023 are waiting less time between their first and second purchase as compared to 2022.

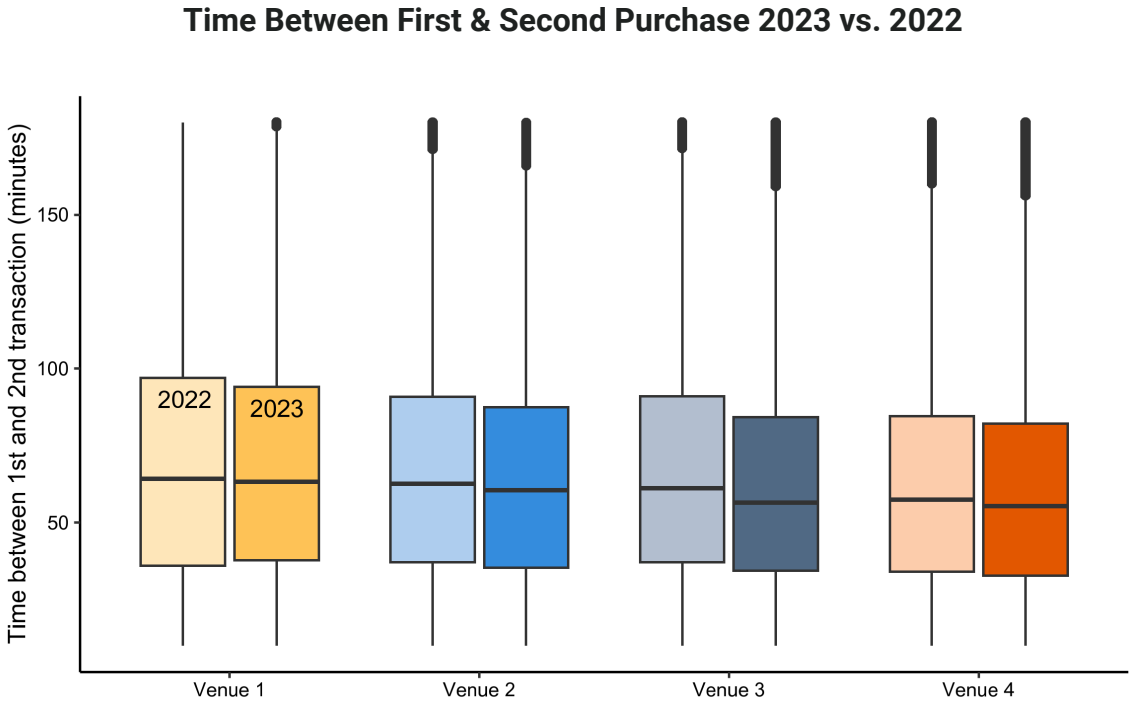


Figure 3. Duration between purchases 2023 vs. 2022.

CONCLUSION

In this initial POC analysis we rigorously interrogated several hypotheses regarding the effect of the 2023 MLB rule changes on consumer behavior. We find that the shorter game times resulting from these rule changes had only a marginal effect on consumer behavior. Indeed, the most substantive difference we found in our analysis was that more consumers were purchasing before first pitch in 2023 compared to 2022. All other effects, while statistically significant, were not meaningful in their interpretation – all differences between 2022 and 2023 for our other analyses were marginal at best. We also saw little variation in our results across the four venues studied here, indicating that the MLB rule changes have generally had little effect on consumer behavior.

Taken together, it appears that the 2023 rule changes did not have a substantive effect on consumer behavior, and that the only substantive change was the buying period. We acknowledge again that there are limitations to our analysis approach, which we hope will be resolved through further information sharing with our MLB partners. We also note that this is only an initial POC for this analysis, and future work will include not only data from our entire MLB portfolio, but also additional hypotheses and exploratory analysis. This report showcases the great potential for innovative analysis that the Aramark Sports and Entertainment Data Science team can provide. We see this POC as a starting point for future analysis targeted at myriad questions relating to consumer behavior, customer retention, loyalty programs, and costumer segmentation.

APPENDIX

	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	183.683	1.126	163.18	p < 0.001
YEAR 2023	-20.82	1.576	-13.22	p < 0.001

Table A1. Regression table for game season on game duration.

VENUE	X-SQUARED	P-VALUE
OVERALL	529.25	p < 0.001
VENUE 1	925.61	p < 0.001
VENUE 2	831.23	p < 0.001
VENUE 3	449.62	p < 0.001
VENUE 4	462.21	p < 0.001

Table A2. Probability of repeat purchase, overall and by venue X-squared test statistics.

VENUE	X-SQUARED	P-VALUE
OVERALL	564.49	p < 0.001
VENUE 1	810.15	p < 0.001
VENUE 2	314.60	p < 0.001
VENUE 3	332.43	p < 0.001
VENUE 4	1111.2	p < 0.001

Table A3. Probability of first purchase before first-pitch, overall and by venue X-squared test statistics.

VENUE	X-SQUARED	P-VALUE
OVERALL	8700.4	p < 0.001
VENUE 1	6700.5	p < 0.001
VENUE 2	7602.3	p < 0.001
VENUE 3	7133.1	p < 0.001
VENUE 4	10923	p < 0.001

Table A4. Probability of repeat purchaser visiting more than one stand, overall and by venue X-squared test statistics.

OVERALL	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	51.069	0.056	909.58	p < 0.001
YEAR 2023	-2.308	0.076	-30.06	p < 0.001

VENUE 1	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	55.163	0.140	393.863	p < 0.001
YEAR 2023	-0.543	0.180	-3.00	p = 0.002

VENUE 2	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	45.819	0.200	228.711	p < 0.001
YEAR 2023	-2.387	0.285	-8.449	p < 0.001

VENUE 3	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	51.897	0.088	584.77	p < 0.001
YEAR 2023	-3.586	0.125	-28.53	p < 0.001

VENUE 4	ESTIMATE	STD. ERROR	T VALUE	P-VALUE
(INTERCEPT)	49.168	0.091	536.1	p < 0.001
YEAR 2023	-2.871	0.124	-23.0	p < 0.001

Table A5. Regression table for length of time between purchases, overall and by venue.