Exploratory Data Analysis Summary

Day of Week Analysis

The first area I explored was how attendance and sales varied by day of the week. I observed a slight uptick in average spend per customer toward the end of the week and on weekends, but overall, there were no strong trends.

Sample Data:

	weekday	weekday_name	avg_total_sales_per_event	avg_spend_per_customer	avg_tickets_sold_per_event	avg_tickets_sold_per_customer	total_tickets_all_events
0	1	monday	359933.75	259.44	4879.75	3.52	19519
1	2	tuesday	297690.00	264.61	3900.00	3.47	3900
2	3	wednesday	372923.33	264.61	4934.00	3.51	14802
3	4	thursday	566550.00	267.87	7439.00	3.52	7439
4	5	friday	364758.75	271.09	4733.25	3.53	18933
5	6	saturday	257885.00	269.26	3361.00	3.51	6722

Temperature Impact

Next, I examined how average daily temperature influenced attendance and sales. Events held on days where the average temperature was above or equal to freezing (0°C) had significantly higher attendance, averaging 5,629 tickets sold per event compared to 4,317 on colder days below freezing. This indicates that temperature may play a meaningful role in attendance #'s.

Sample Data:

	temp_category	avg_total_sales_per_event	avg_spend_per_customer	avg_tickets_sold_per_event	avg_tickets_sold_per_customer	total_tickets_all_events
0	Above or equal 0°C	430546.0	269.64	5629.4	3.53	28147
1	Below 0°C	324482.0	263.88	4316.8	3.51	43168

Venue Capacity by Section

I then analyzed the percentage of venue capacity filled by section for each city and event. There were no consistent trends or capacity issues, though one anomaly stood out: the upper bowl section in New York on 2025-03-12 showed 103.66 percent capacity, suggesting possible overbooking or a data irregularity.

Sample Data:

	event_date	home_city	section	section_capacity	total_tickets_sold	percent_full
0	2025-01-04	boston	club	1147	552	48.13
1	2025-01-31	boston	club	1134	420	37.04
2	2025-01-04	boston	lower bowl	2677	1185	44.27
3	2025-01-31	boston	lower bowl	2646	1082	40.89
4	2025-01-04	boston	standing room	382	135	35.34
...	...	...	...	...	...	...
70	2025-02-24	vancouver	club	1219	696	57.10
71	2025-02-24	vancouver	lower bowl	2844	1768	62.17
72	2025-02-24	vancouver	standing room	406	329	81.03
73	2025-02-24	vancouver	suite	812	541	66.63
74	2025-02-24	vancouver	upper bowl	2844	1788	62.87

Market Comparison

Next, I took a broader look at overall sales performance across markets. Seattle emerged as the strongest market, leading in both tickets sold per event and average spend per customer. Boston, on the other hand, showed weaker performance, averaging 3,374 attendees per event compared to 6,180 in Seattle. Despite lower attendance, Boston had the second-highest average ticket price, which may be contributing to sales issues. It may be worth exploring promotional offers or pricing adjustments to increase engagement in the Boston market.

Sample Data:

	home_city	avg_total_sales_per_event	avg_spend_per_customer	avg_tickets_sold_per_event	avg_tickets_sold_per_customer	avg_ticket_price	total_tickets_all_events
0	seattle	479977.5	273.48	6180.0	3.52	77.68	12360
1	new york	409767.5	266.67	5393.0	3.51	75.87	10786
2	vancouver	380130.0	261.44	5122.0	3.52	74.22	5122
3	toronto	375127.5	265.87	4997.0	3.54	75.05	9994
4	minnesota	350175.0	259.23	4682.5	3.48	74.51	9365
5	ottawa	332195.0	260.48	4457.5	3.48	74.77	8915
6	montreal	301555.0	264.07	4012.5	3.51	75.19	8025
7	boston	259912.5	272.97	3374.0	3.55	77.00	6748

Sales Channel Distribution

I also examined sales distribution across purchase channels by city. Sales were evenly distributed across all channels, including online, mobile, and box office, with each accounting for roughly 20 percent of total sales.

Sample Data:

	home_city	purchase_channel	total_tickets_sold	percent_of_city_tickets
0	boston	online	1381	20.47
1	boston	group sales	1351	20.02
2	boston	box office	1343	19.90
3	boston	mobile app	1340	19.86
4	boston	season ticket	1333	19.75
5	minnesota	mobile app	2011	21.47
6	minnesota	group sales	1873	20.00
7	minnesota	online	1856	19.82
8	minnesota	season ticket	1840	19.65
9	minnesota	box office	1785	19.06

Repeat Customer Analysis

Lastly, I looked at repeat customer counts by city. There were no repeat account sales in this dataset. In a real-world scenario, I would recommend a focus on customer retention.

Sample Data:

	home_city	repeat_customers	total_customers	repeat_customer_pct
0	minnesota	0	2693	0.0
1	boston	0	1904	0.0
2	toronto	0	2825	0.0
3	ottawa	0	2562	0.0
4	new york	0	3068	0.0
5	montreal	0	2284	0.0
6	seattle	0	3514	0.0
7	vancouver	0	1454	0.0

How the Data Model Supports Future Reporting and Analytics

The data model uses a star schema with a central fact table for ticket sales and supporting dimension tables for date, venue, weather, customer, and channel. This structure makes it easy to analyze performance across different dimensions such as city, day, or weather conditions. It also supports efficient querying and clear reporting for future analytics or dashboard needs.