Xiaohui Tan

Data Challenge Summer Fall 2022

Deliverable 1

Oct 3rd, 2022

**Question 1:**

From the table above, we can conclude that the evening game do have higher average attendance than afternoon games.

A picture containing chart

Description automatically generatedTable

Description automatically generatedBelow is the average per game by month: it is surprised that January has the lowest attendance.

**Question 2:**

Note that number of fans in arena:

* 2hr prior to start count if ticket scan time <= 2 hrs prior event time.
* 1hr prior to start count if ticket scan time <= 1 hr prior event time.
* at start count if ticket scan time <= event time
* A picture containing text

  Description automatically generatedpeople who being late are not counted in this question.

From the table above, it shows the result that the average number of fans in arena 2-hrs prior to start time is only **2** fans; the average number of fans in arena 1-hr prior to start time is **706**; the average number of fans in arena at start time is **3575**.

**Question 3:**

The diagram below shows all the account with ticket bought >=5 and scans >=3.

* Table

  Description automatically generatedThe total number of accounts is **3882**
* **A picture containing table

  Description automatically generated**Of those, **2411** are member accounts

Table

Description automatically generated**Question 4:**

The top 10 sections that were most filled on average are listed above:

* 106, 102, 105, 114, 107, 113, 115, 109, 101, 116

**Question 5:**

* Chart, scatter chart

  Description automatically generatedThe correlation between the opponent team’s Vegas over/under win total and attendance is **0.37**, the correlation graph is shown below:
* Even though the correlation is below 0.5, which indicates not very strong, there is still a positive correlation between the higher Vegas over/under win total and higher attendance.

**Question 6:**

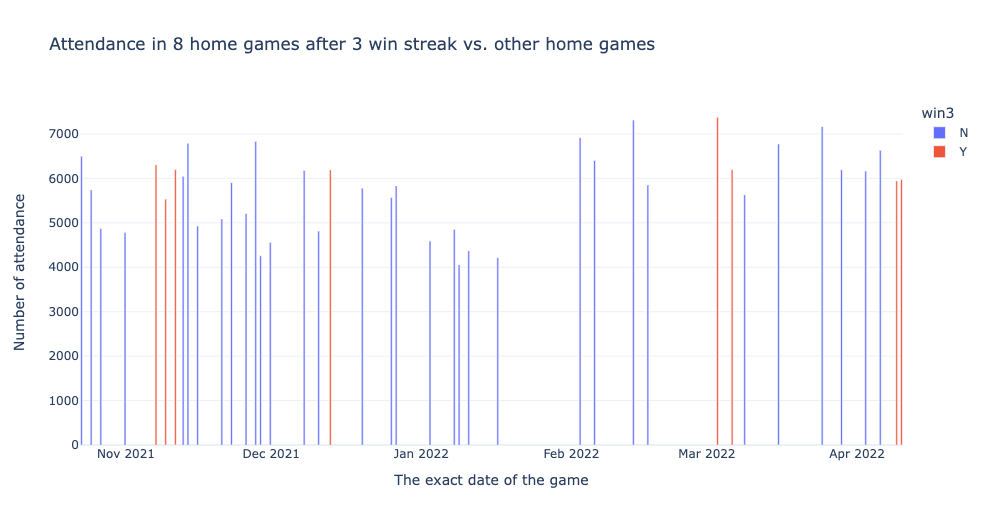
**Timeline

Description automatically generated**Below is the code for scraping Basketball Reference and pulling in data:

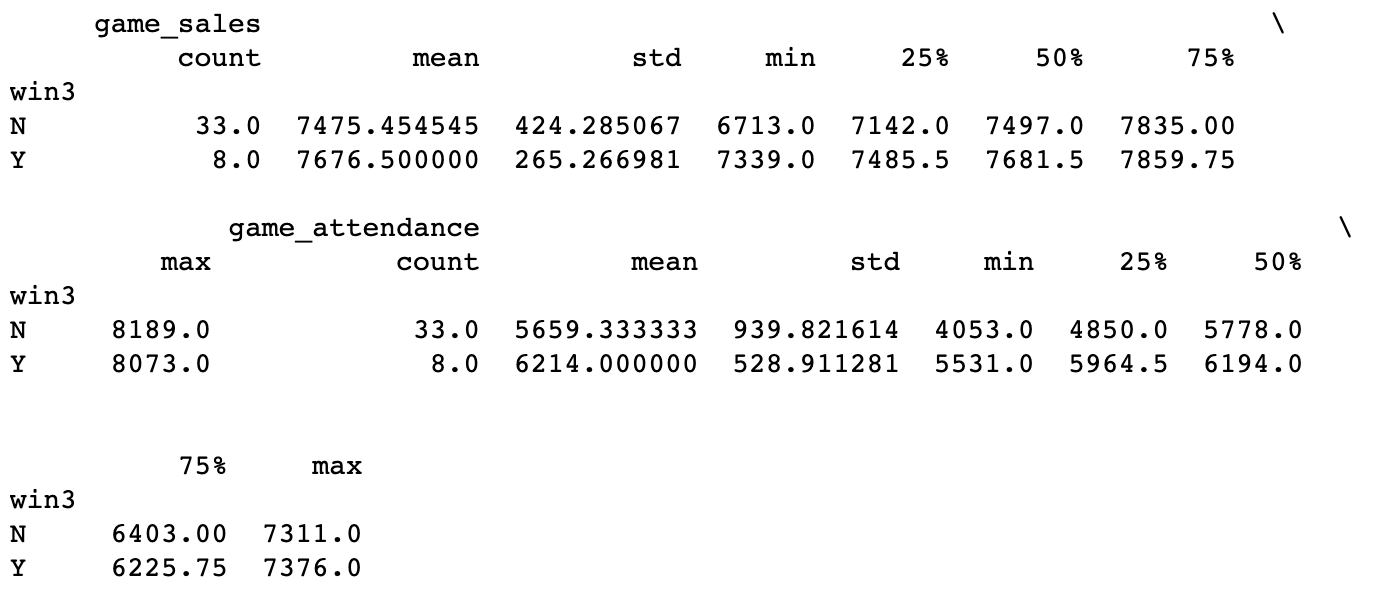
|  |  |  |
| --- | --- | --- |
|  | Average ticket sales | Average attendance |
| 8 games after Win 3 | 7676 | 6214 |
| 33 other games | 7475 | 5659 |

The table above shows the average attendance and average ticket sales of the 8 games that occurred after one of the win 3 streaks and the remaining 33 games. We can see that both the attendance and ticket sales of those 8 games are higher than the other 33 games.

The different between ticket sales is small (201), but it between attendance (555) is obvious.



Above is the graph to show each of the attendance of the 8 games after win3 streaks and the remaining 33 games. The red lines represent the 8 games, and the blue lines are the remain 33 games. It shows more detail for the attendance

Here is the summary statistic of Q6: