**MIDTERM REPORT**

**TICKET ANALYSIS – GIRL GENERATION DATA BY GROUP 4**

**Group member:**

|  |  |
| --- | --- |
| **Name - ID** | **Contribution** |
| **蔡沛騰 - 408850310** | **25%** |
| **郭鎮維 - 408850419** | **25%** |
| **貝萊亞** - **611785519** | **25%** |
| **阮俊宇 - 407855088** | **25%** |

**Tasks:**

(1) Based on the dataset, find any insight about number of sales of different zone between males and females, member and non-member in this concert.

(2) Please define the “sale speed” of each zone. Provide the detail about how your group will define the “sale speed” carefully, and be creative!!: -Use your defined “sale speed’ to compare the ticket sale of different zones. find any interesting insight from the result.

(3) Write a report < 10 pages to address the tasks and submit these files:

-Report file (PDF/Word file)

-Python notebook (\*.ipynb) file26

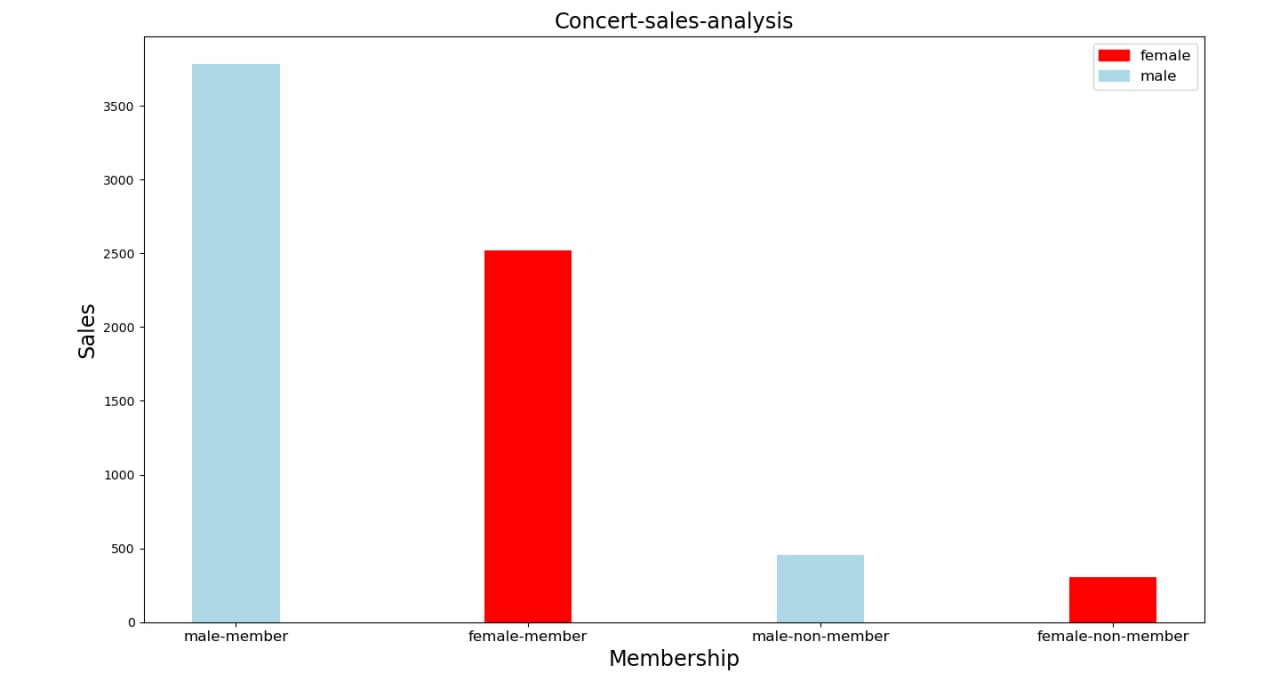
**(1) Sale insight base on analyzed dataset**

For the first question of finding insights about the number of sales for different zones, as well as for male/female & member/non-member. We decided to divide the first question into two parts:

1. Overall sale analysis
2. Zone analysis

For each part we separated member into smaller categories based on their gender and membership status – therefore have an insight into the influence of mentioned factor to the ticket sales.

PART 1: Overall ticket sales



Based on the given dataset, we are able to identify the significant differences between the sales between members and non-members. As shown in the graph, we can see that the majority of the tickets sold are purchased by members of the tickets system – this could be contributed by numerous factors such as,

* Ticket availability based on membership status, i.e., that initial ticket sales were only offered to members first, before being released to non-members.
* The outreach of the concert marketing toward members. Most likely, members are enrolled in an e-mail marketing list which makes them highly informed on concerts and when tickets are released.
* Members are most likely people who have purchased tickets before and thus, became members. These people might therefore have an overall greater interest in concerts than non-members.

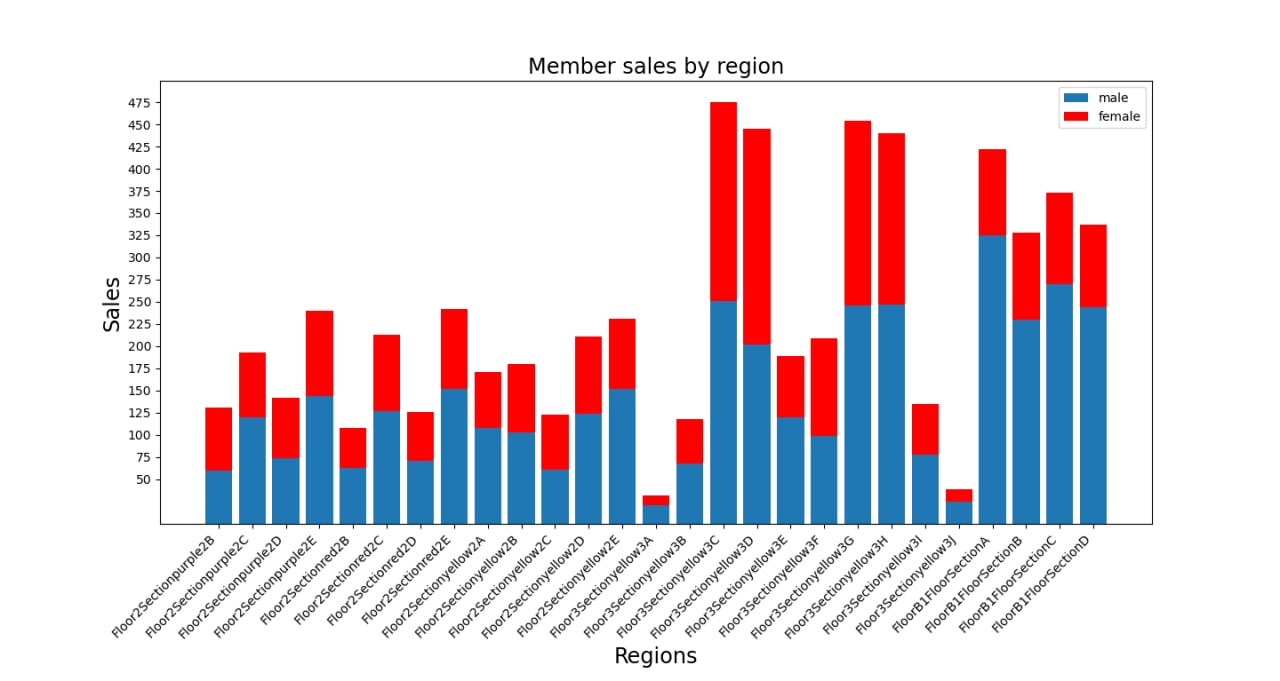
The graph also shows the gap between the number of males participants and female counterparts – with 30% and 20% higher revenue generated by male members and male non-members, respectively. This trend persists despite the difference in membership status suggests that the concert attracted more male participants than female.

There is no certain way to make a conclusion on why male members and male non-members make for the largest portion of ticket purchasers. A possible explanation for this might be that the groups fanbase is mostly male. Another explanation might be that males are the ones purchasing tickets for themselves, as well as for others. It should be noted that these assumptions are highly speculative since the data does not show any evidence of this nor gives an explanation.

Our guess for this phenomenon, based on the given dataset and graph, is that this was a concert performed by an all-female group which can attract a more male-oriented fanbase. This, however, could also mean that male participants are willing to spend more money on the ticket compare to the female fans. We will discuss this further in the later part.

PART 2: Detailed zone sales analysis

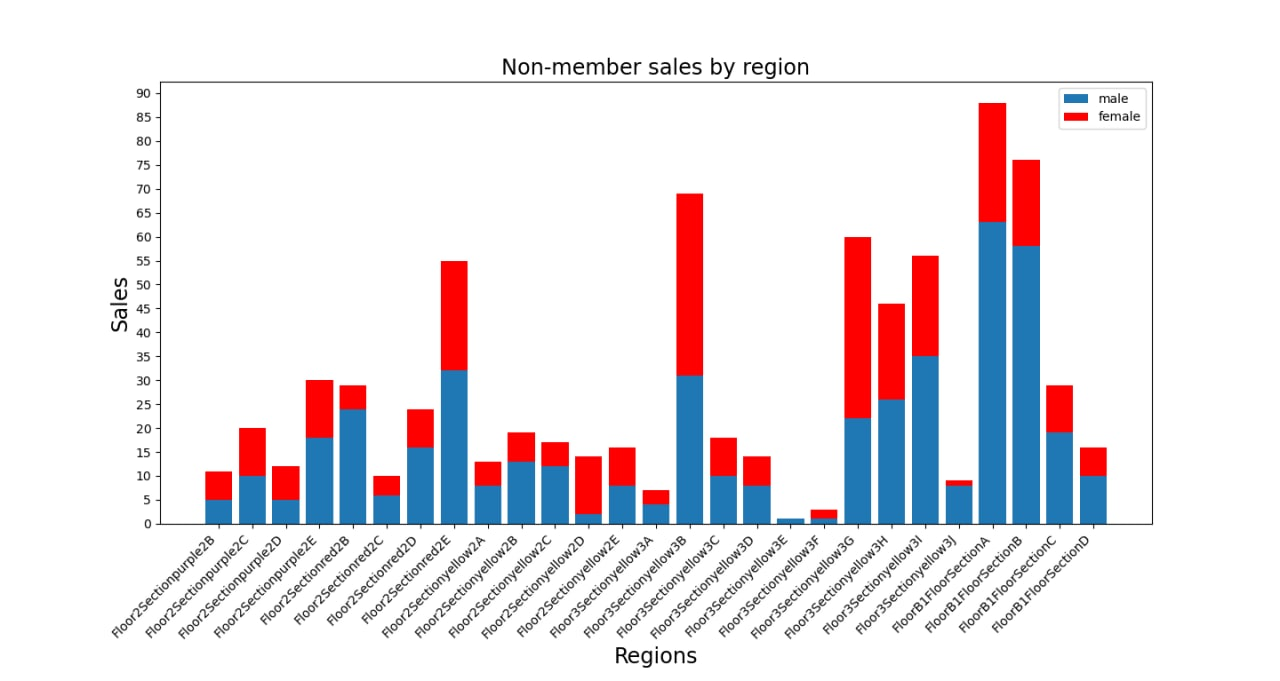
In the second part, we decided to divide it into two graph bases on two classifications: member and non-member and to analyze the zones. We used stacked bar charts to represent two genders, with blue for male and red for female.

 Graph 1: Member sales by region

Firstly, we will look into member sales in each zone. Overall, based on the bar chart in Graph 1 we can once again see how gender affect sales number – with the majority of sales are generated by male members. If we take a closer look, the sales differences between two genders are not evenly distributed, with certain zones are more male-dominated while some others are equally distributed.

For members, floor 3 section yellow is the most interesting section, specifically zone 3C and 3D where the sales are equally distributed between male and female members as well as being in the top regarding sales volume. This, however, cannot be ruled to a pattern and should be categorized as a coincidence since there is no positional advantage for these two zones to be preferred by female members; thus, we do not have any conclusion on this finding. Also, it’s difficult to look for patterns when it is only one concert being analyzed. In order to determine a pattern, multiple concerts would have to be analyzed.

On the contrary, Floor B1 is the most desired space for male members, as we can see all four sections A, B, C and D have 2-3 times more male members comparing to female. These sections are the closest to the stage, which might explain the sale distribution of males being the most frequents ticket purchasers.



Secondly, we will look into the sales in each zone for non-members. Similar to member sales, most of the revenue is generated by male participants, which a few exceptions in certain zones. Floor 3 section yellow 3B has equal sales for both genders, while the non-member female sale is higher compares to male counterpart in yellow 3G. The most interesting section is on the third floor – section yellow 3E with no non-member female.

What is also important to note is that the sections that are popular with members are also popular with the non-members regarding tickets sales compared to other sections. Even though the volume is higher with members compared to non-members, if you compare zones, within each ticket group, certain zones are more popular than others.

To sum up, we can see that for this concert, males are more willing to spend money on the ticket – whether they have the membership or not.

**(2) Sale speed definition and comparison between zones**

1. Sale speed definition

When defining sales speed, there are a number of things to look at. For example, one could look at how long it took to sell all the tickets, how long it took to sell more than half or any other specific metric used by ticket vendors, or how many tickets are sold at a given time, day, or week.

For our analysis we wanted to examine several important metrics:

* At what time did the ticket sales start and was there a difference between zones?
* At what pace where they selling, for example, did the ticket sales start of strong and then decline drastically?
* We also wanted to investigate at what point did the actual ticket sales decrease.

When answering the questions above, a broad range of information will be made clear. Ticket vendors could use the information to learn more about consumers purchasing habits and thus, adjust their marketing, technical capacity, as well as pricing of the products. We needed to depict a graph where we could first se at approximately what pace where the tickets selling, as well as how many tickets were sold on the specific time and in which floor/zone.

Therefore, we define our sale speed as the time when tickets were sold earliest in each zone – by comparing the time each zone started to sell tickets, at the same time cross-check the sales with other zones in early hours so that we can have a conclusion on which zone are prefered and has the “early bird” privilege.

There are 27 graphs of each zone depicted below, each graph shows the amount of ticket sales for that zone, as well as at what time. If we for example look at Floor B1 and all its sections A, B, C, D. We can see that the sales were relatively high in the first couple of hours and days. Where the sales are almost constantly between 1-4 and with almost no decline except for the final decline a couple of days later when there is a clear majority of the tickets were sold. This can for example be compared to various sections in Floor 2, such as, Section purple 2B/2C, Section red 2B/2C/2D and Section yellow 2A/2C. There we can see that the ticket sales are not as intense as in Floor B1. The trend is furthermore shown in Floor 3 and its yellow sections. Here, the intensity is even weaker and ticket sale are not as high nor rapid as in B1. Some sections of both Floor 2 and Floor 3 are selling tickets relatively quick but not as direct as B1. Considering that the yellow section seats are furthest from the performers, an assumption to make could be that fans are paying for tickets and attending the concert to see the performers live as close as possible. As well as that the music itself is not the most important factor but also the other aspects of the performance. Also, the tickets furthest away are also the cheapest, which shows that price is not an important factor when selling tickets and that fans are not price sensitive.

