**Fitter-Spot Case Study**

Ong Kian Eng A0052270U  
Shen Chen A0058260J

# Executive Summary

Fitter-Spot is a local chain of fitness classes which offers two types of fitness programs – Yoga and KickBoxing. The management is interested in obtaining information relating to member profiles to identify loyal members and understand member’s preference for fitness programs in order to target new and prospective members.

Using the available datasets with member’s personal details and Dec’18 session schedule, this report serves to:

1. Present a member profiling analysis based on gender, age and preference for Yoga and KickBoxing fitness programmes (including location, time, day of session),
2. identify and leverage on popular sessions (i.e. location, time, day of session),
3. and propose strategies to enable management to expand customer base, retain current loyal customers and increase member’s fitness points contribution.

# Data Selection and Pre-processing

## Data Description & Cleansing Performed

Examining the datasets, **2 major issues** arising from human error in the manual data entry and processing were identified:

1. multiple missing values
2. duplicates

Among these, we have selected 4 variables to clean up and input missing values (if necessary), and these would be subsequently used in our analysis. Variables with issues are summarized as below in Table 1:

Table : Variables which need to be cleaned

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Dataframe** | **Variable** | **Issue** | **Comments** | **To clean** | **To verify** |
| MemberData | MemberName | duplicate values |  | o |  |
| WorkPhone | missing values | not required for analysis |  |  |
| Address | missing values |  |  |
| Gender | missing values | not possible to clean up just by looking at dataset so rows will be ignored in analysis |  |  |
| Height | missing values |  |  |
| Weight | missing values |  |  |
| FitnessPointsUpToNov18 | possible outlier | it is possible to attain 826 points in 11 months given by the member's Dec'18 records and hence it will be considered in analysis |  |  |
| FitnessPointsDec18 | missing values |  | o | o |
| Sessiondata | Location | missing values |  | o |  |
| SessionTime | missing values |  | o |  |
| FitnessChampion | missing values | not required for analysis |  |  |
| MemberMobilePhone | duplicate values (leading from duplicated MemberName) |  |  | o |

* + 1. ***Cleansed variable 1: MemberName***
* **2 sets of duplicates** were found:

1. *Duplicate 1 comes from the same person (Cox, Anastacia Felicia) with 2 mobile numbers, identified by a combination of name and DOB (Date of Birth). As most of the other variables are identical (e.g. never changing variables such as DOB, Gender, MemberSince), we chose to keep the one with the highest points and removed the other entry.*
2. *Duplicate 2 comes from 2 individuals with the same name (Varma, Sneha). We chose to keep both as these two individuals have different DOB and MemberSince.*

* Outcome: the row with lower FitnessPointsUpToNov18 is removed for Duplicate 1.
  + 1. ***Cleansed variable 2: Location***
* **9 missing values** were recovered by matching the SessionDate, SessionTime, and FitnessChampion of a missing Location to that of a known Location. By this, we are able to retrieve the location of each missing Yoga / KickBoxing Session.
* Rationale: Each Fitness Champion can only be physically present at one place at any time.
* Assumption: There are no two or more Fitness Champions going by the same name.

* + 1. ***Cleansed variable 3: SessionTime***
* **20 missing values** were recovered by matching the SessionDate, Location, and FitnessChampion of a missing SessionTime to that of a known SessionTime.
* Rationale: Each Fitness Champion can only be physically present to teach one class at one place at any time.
* Assumption: There are no two or more Fitness Champion going by the same name.

* + 1. ***Cleansed variable 4: FitnessPointsDec18***
* **52 missing values** were recovered by:

1. *re-computing total fitness points for Dec’18 based on the total number of sessions attended for each SessionType in the period. Table* 2 *shows the fitness points distribution for attending Yoga and KickBoxing lessons according to the number of sessions attended. More points were given for participating in more sessions. Those who did not attend any sessions in Dec were excluded from the calculation of fitness points distribution.*
2. *Total points for Dec’18 is then calculated by the sum of fitness points for each SessionType shown in Table* 2

* Assumption: MemberMobilePhone variable in SessionData dataset is correct, which implies that the number of sessions attended for Dec’18 based on MemberMobilePhone would assumed to be correct.

Table 2 Fitness points awarded by total sessions

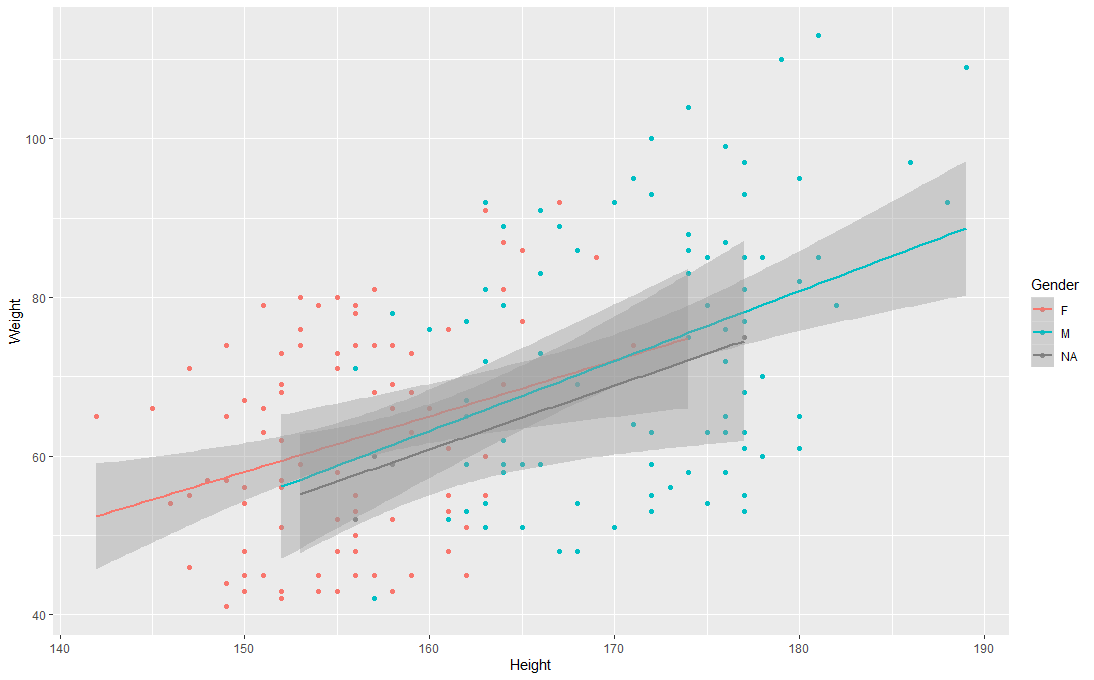
|  |  |  |
| --- | --- | --- |
| **Sessions attended** | **Points** | |
| ***Yoga*** | ***Kickboxing*** |
| 1 | 5 | 3 |
| 2 | 5 | 3 |
| 3 | 5 | 3 |
| 4 | 5 | 3 |
| 5 | 5 | 3 |
| 6 | 5.5 | 3 |
| 7 | 5.86 | 3 |
| 8 | 6.13 | 3 |
| 9 | 6.33 | 3.33 |
| 10 | 6.5 | 3.6 |
| 11 | 6.64 | 3.82 |
| 12 | 6.75 |  |

## Other variables that need to be cleansed but not chosen

1. ***Gender, Height, Weight***

* Despite their significance in constructing member profile, we were uncomfortable with assigning values to fill in the missing inputs based on assumptions.
* We cannot determine the gender of the person by looking at their names, which could be gender-neutral names (e.g. Joey or Skyler) and could lead to wrong analysis later on.
* As for the Height and Weight, we considered the possibility of deducing the figures based on the Gender of the individual. However, the spread of the points in the scatterplot (See Figure 1) is too wide. Individuals varies - there are females out there who could be tall, and males who could be short. Therefore, we cannot find a logical way to clean up the missing values.

Figure 1 Correlation between Height and Weight and Gender



* As such, the best approach is to assign the registration staff (i.e. reception) to verify these missing details with the members, rather than imputing an incorrect value which may mess up the analysis subsequently.

1. ***Mobilephone, Address, Fitness Champion***

* These variables are not cleaned as they are not required for our analysis and it is not possible to clean them without verifying each entry personally with the respective member. NAs will thus be ignored in our analysis.

## Accuracy of information

## Verified variable 1: FitnessPointsDec18

* + Verified by matching FitnessPointsDec18 values to fitness points computed based on method mentioned in 1.1.4
  + Outcome: **2 discrepancies were found** and values have been corrected for use in subsequent analysis.

## Verified variable 2: MemberMobilePhone

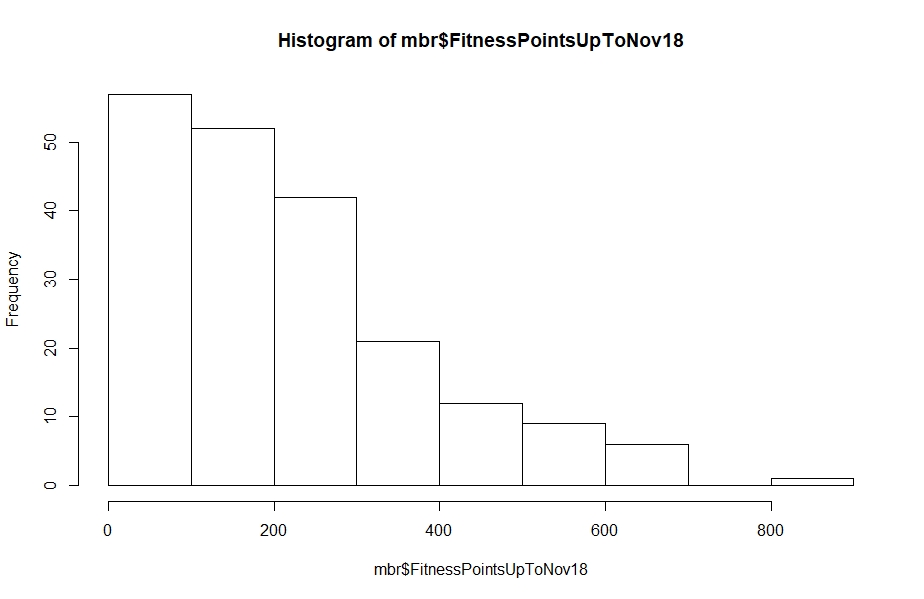
* Verified by matching MemberMobilePhone to MemberName. Each member should only have one unique mobile number used to indicate their attendance for classes.
* Outcome: **7 sets of duplicates** were found and removed– a second mobile phone number belonging to Cox, Anastacia Felicia was used to register for class, resulting in double counting of classes attended by the same member.

## Other variables that need to be verified but not chosen

FitnessPointsUpToNov18

* It is not possible to verify the data without information on member’s attendance data.
* There is a possible outlier of 826 points based on Figure 2. However, it is possible to achieve the figure based on the particular member’s Dec points. Therefore, we assume that the data is clean for use in our analysis.

Figure 2 Fitness points up to Nov'18



## Derived Variables

* The 2 engineered features are as follows:

1. Age Category – for segmentation of member profile
2. Member’s Loyalty – defined by length of membership (mth) and frequency of class participation

## Feature 1: Age Category

1. ***Age of Member***

* Deduct Date (31 Dec 2018) from Date of Birth (DOB)
* Rounded to the nearest year
* Rationale: Find out the Age of the member, which will be useful in identifying the profile of the members and be used for binning to create different age categories.

1. ***Age Categories***

* Age categories of <25, 25 to 29, 30 to 34, 35 to 39, 40- 44, >45 was created ( Figure 3).
* Rationale:

1. to derive greater insights and better visualize the preferences of different age groups (e.g. preference for certain sessions), different ages were grouped together to form age categories.
2. We have considered the age intervals based on gender (Figure 4) to determine the intervals for the age categories. We have also generated numerous iterations of histograms of Age with breaks of 5-12, which showed the same distribution. Thus, a 5-years interval was chosen for a more detailed analysis on member profiling.

Figure 3 Histogram with breaks of 5 years

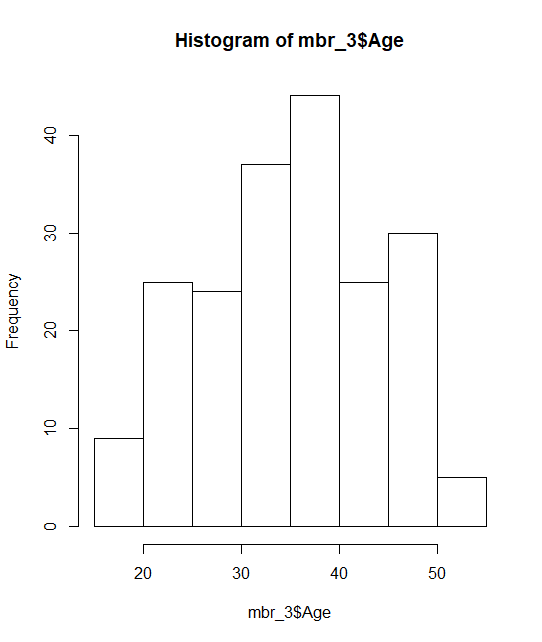
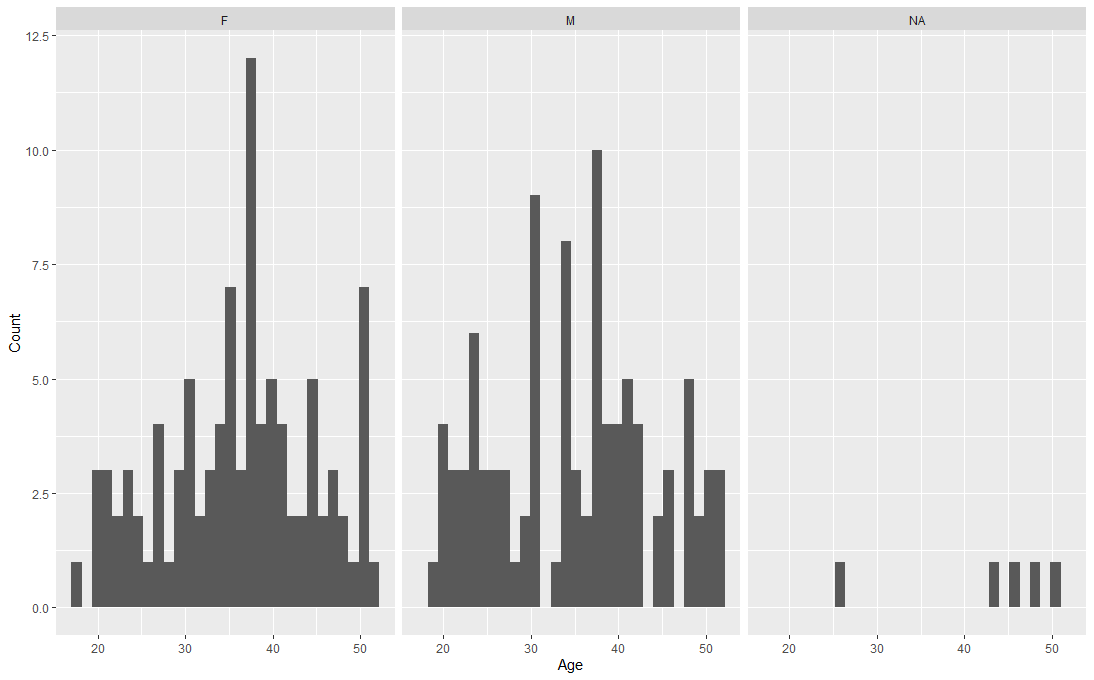


Figure 4 Histogram of Age Category created



## Feature 2: Member’s Loyalty

1. ***Total FitnessPoints for the Year***
   * TotalFitnessPoints obtained by

FitnessPointsUpToNov18 + Fitness Points in Dec (dec\_clean)

* + Rationale: Find out the most active member with highest fitness points for the year

1. ***Length of Membership Up to Dec 2018***
   * LengthMemMth obtained by

Date (31 Dec 2018) – MemberSince

* Rounded to the nearest month
  + Rationale: Find out the length of membership of the member, which will be used for calculation of Average Points earned in each month (AvgPtsPerMth).

1. ***Average Points Per Month***

* Divide TotalFitnessPoints by LengthMemMth (length of membership)
* Rounded to the nearest whole number
* Rationale: Find out the average points earned in each month, which will be used to determine the members who are consistently attending sessions. Based on our analysis, a member with Fitter-Spot for longer periods of time tends attend more sessions and accumulate more points. Hence, the average points earned per month based on length of membership allows us to have a better measure of who are the more active members.

1. ***Average Session Per Month***

* Divide Total number of Sessions (Sum of Number of Yoga and KickBoxing sessions in Dec) by LengthMemMth (length of membership)
* Rounded to the nearest whole number
* Assumption: As the number of sessions attended from Jan to Nov is unknown (only the fitness points are known), we extrapolated Dec session data to the previous month. However, we note the limitation that the average number of sessions attended by each member in a month (Jan - Nov) could not reflect their actual attendance.
* Rationale: Find out the average number of sessions attended in each month, which will be used to determine the members who are consistently attending sessions. The key idea of normalizing it to length of membership is similar to what was discussed under Average Points Per Month.

1. ***Activeness of member***

* Variation (henceforth standard deviation) for Points earned up till Nov was obtained.
* To determine if the member is consistently active, an active member’s Points in Dec should fall within the range of Average Points Per Month (calculated up till Nov) ± Standard Deviation (for points calculated up till Nov).
* Assumption:

1. As session data from Jan to Nov is not available, fitness points was used as a proxy measure of number of sessions.
2. As the points systems is not disclosed, we assume that incremental points for each session type is reset and calculated each month.
3. Members attend similar number of sessions every month.

* Rationale: Determine the activeness and loyalty of member through the number of Points earned. The key idea of normalizing it to length of membership is similar to what was discussed under Average Points Per Month.

## Other derived variables

***Day of Week for Session in Dec***

* Use lubridate function to derive day of week of the session in Dec (NB: 0 is Sunday, 1 is Monday).
* Rationale: Find out the days that are most popular for members instead of using exact dates.

# Member Profiling

## Overall Summary

* Majority of the members consist of adults (likely to be working adults) with **median age of 37**.
* There are equal numbers of male and female members.
* Members between **age 30 and 40** forms the largest age group among females (44%) and males (40%) (Table 3).

Table 3 Distribution of gender in various age categories

|  |  |  |  |
| --- | --- | --- | --- |
| **Age Category** | **All members** | **Females**  (% of all females) | **Males**  (% of all males) |
| **<25** | 34 (18%) | 14 (14%) | 20 (21%) |
| **25-29** | 23 (12%) | 12 (12%) | 11 (11%) |
| **30-34** | 37 (19%) | 18 (19%) | 19 (20%) |
| **35-40** | 44 (23%) | 24 (25%) | 20 (21%) |
| **40-44** | 24 (12%) | 13 (13%) | 11 (11%) |
| **>45** | 32 (16%) | 16 (16%) | 16 (16%) |
| **Total** | 194 | 97 | 97 |

* + 1. ***Limitations***

As several members’ data on Gender was not available (i.e. NA), they were categorized under NA and these data were ignored in our analysis. Their other data (e.g. fitness points, classes attended etc), however, are still used in analysis where Gender information is not required, since these data still provide insights about the members’ preference (e.g. preference for type of class, time, location) regardless of Gender.

## Distribution of Gender and Age

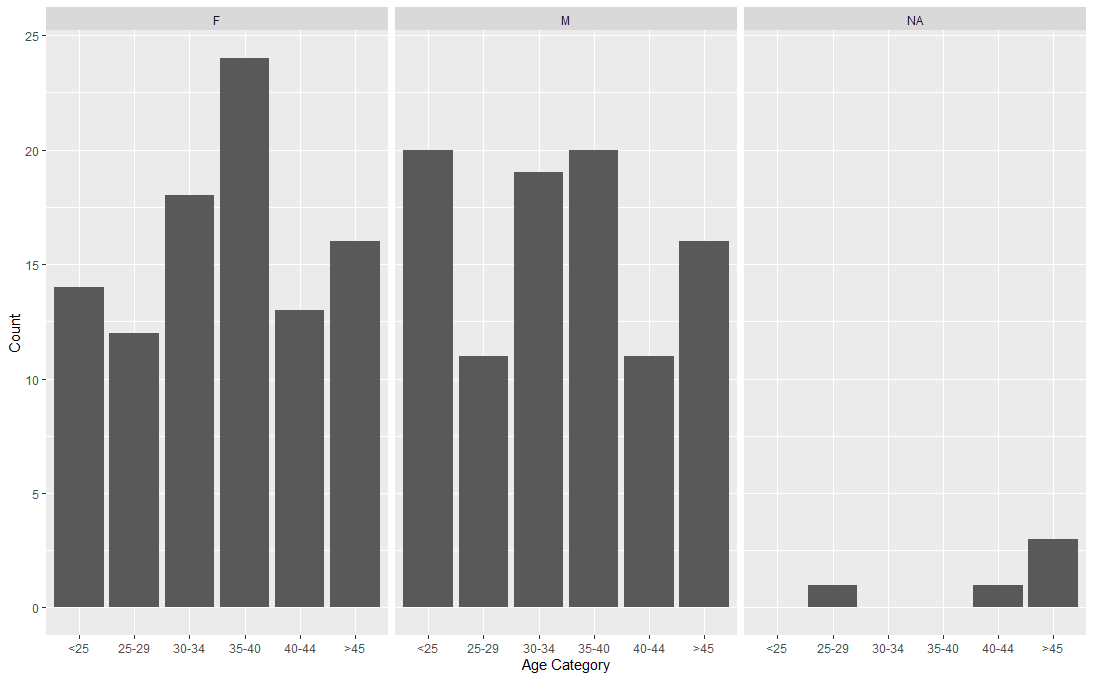
* + 1. ***Summary***
* Based on age bracket, **35 – 40 years form the majority** of both female and for males members (Table 4 & Figure 5).

Table 4 Member age profiling

|  |  |  |  |
| --- | --- | --- | --- |
|  | **All members** | **Females** | **Males** |
| **Youngest age** | 18 (1) | 18 (1) | 19 (1) |
| **Median age** | 37 | 37 | 35 |
| **Mode age** | 38 (13) | 35 (7) | 34 (8) |
| **Oldest age** | 52 (4) | 52 (1) | 52 (3) |

*NB: Number in brackets is the total number of members*

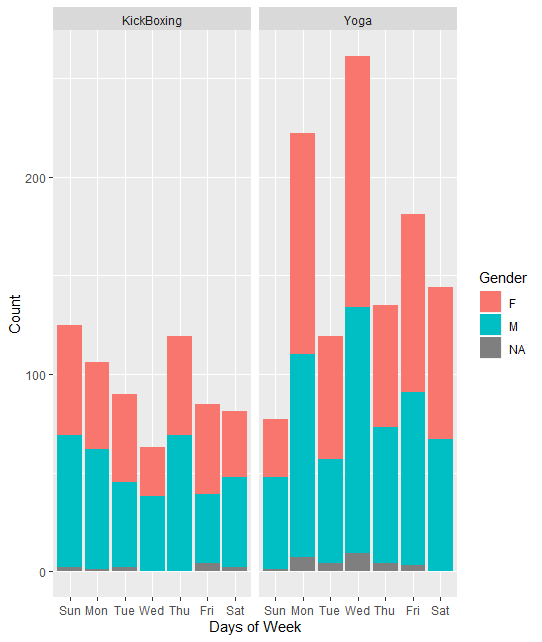
Figure 5 Distribution of gender in different age categories



## Member’s Preference for Type of Class

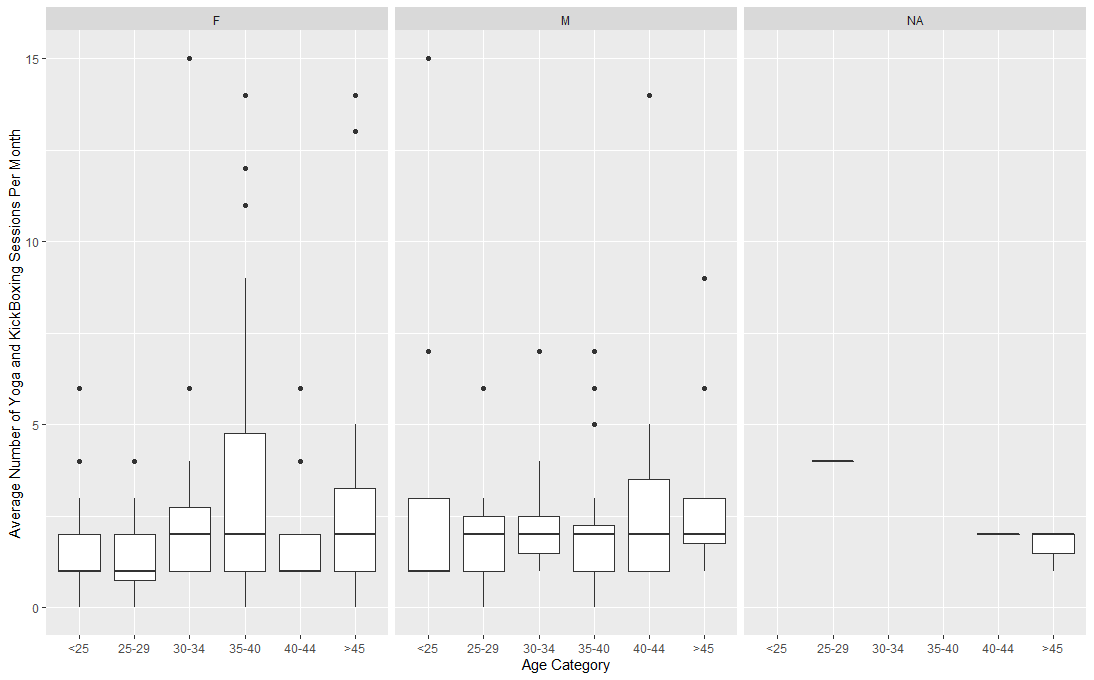
* + 1. ***Summary***
* In line with Fitter-Spot management’s current strategy, **Yoga is more preferred** than KickBoxing among all members, especially among females between 35–40 ( Figure 6).
* *Yoga is particularly popular on Mondays and Wednesdays*
* *Kickboxing is particularly popular on Sundays and Thursdays*

Figure 6 Number of members for different classes in a week



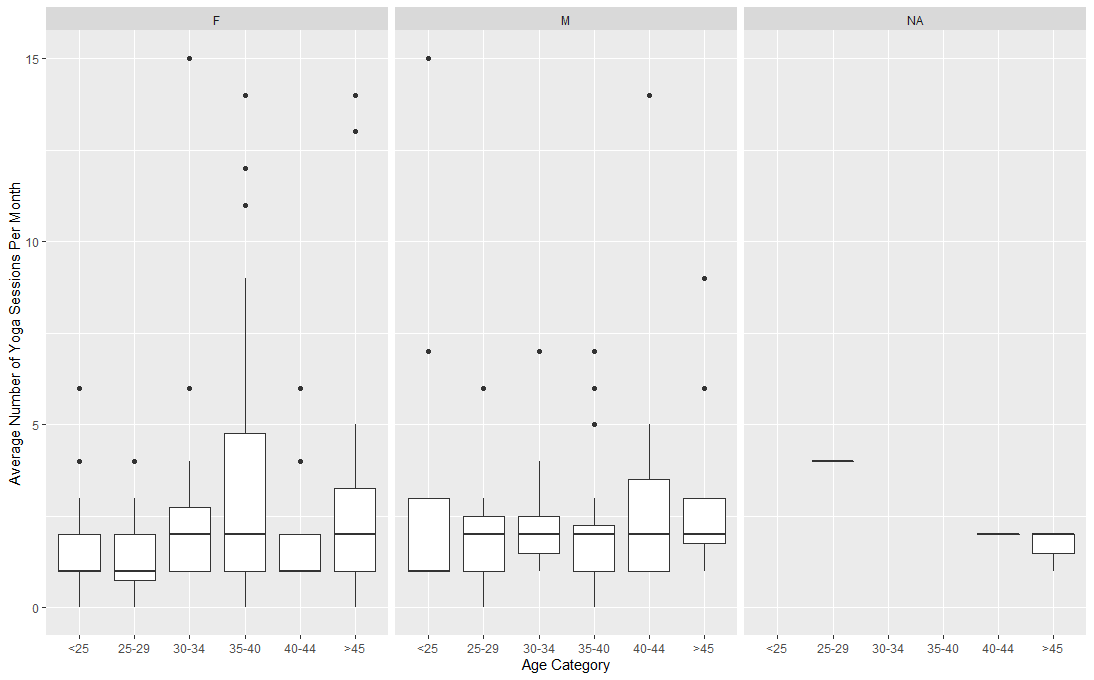
* + 1. ***Distribution of Average Number of ALL Sessions Per Month based on Age Categories for different Genders***
* Based on median values (Figure 7),
  1. **females of ages 30-34, 35-40 and >50** attended more sessions (average of 2 sessions per month)
  2. **males of ages of above 25** attended more sessions (average of 2 sessions per month)

Figure 7 Distribution of Average Number of Yoga and KickBoxing Sessions in various age categories



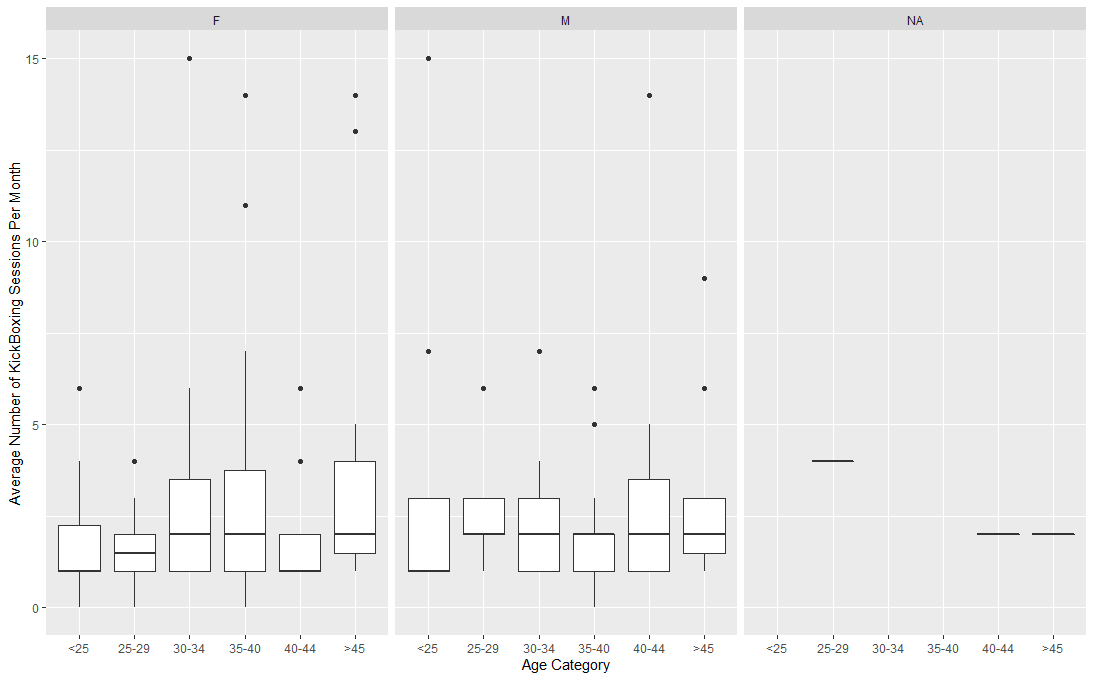
* + 1. ***Distribution of Average Number of Yoga Sessions Per Month based on Age Categories for different Genders***
* The profiling of members attending Yoga class (Figure 8) **supports the overall** observation above (Figure 7)

Figure 8 Distribution of Average Number of Yoga Sessions in various age categories



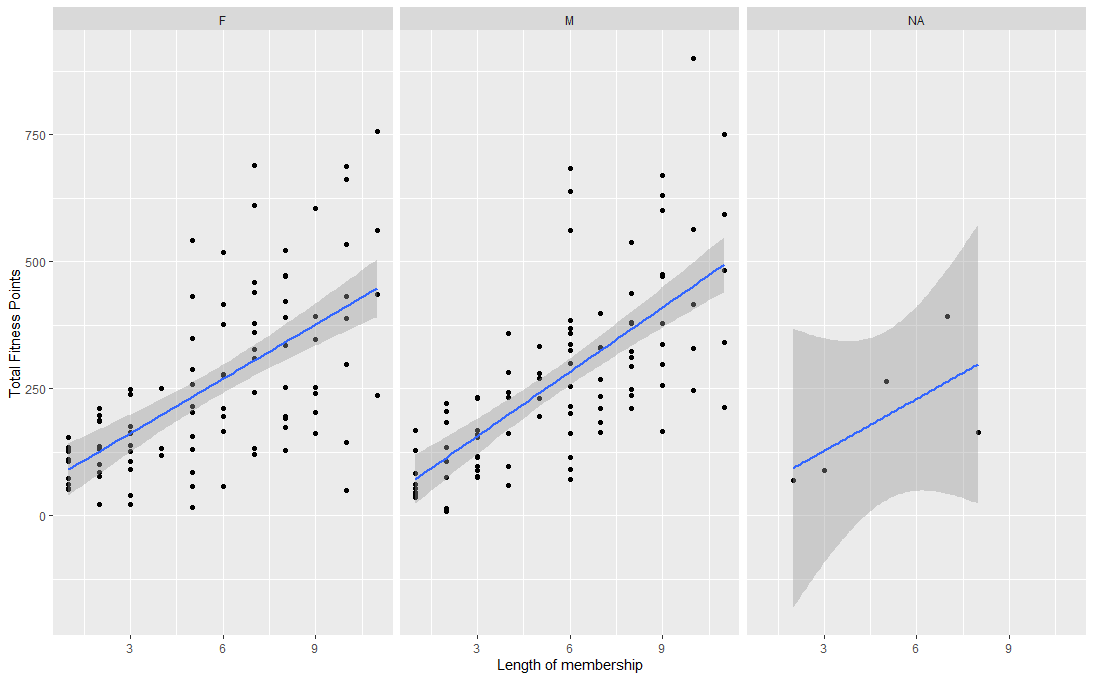
* + 1. ***Distribution of Average Number of KickBoxing Sessions Per Month based on Age Categories for different Genders***
* The profiling of members attending Kickboxing class (Figure 9Figure 8) **supports the overall** observation above (Figure 7).

Figure 9 Distribution of Average Number of KickBoxing Sessions in various age categories



* + 1. ***Relationship between length of membership and total fitness points***
* There is a **positive correlation** between length of membership and points earned (Figure 10). However, there is great variation as shown in the scatterplot.

Figure 10 Relationship between length of membership and Total Fitness Points



## Information about popular session, period and location in Dec

* Yoga is the most popular class for all age categories and gender (Figure 12)
* Popular day for Yoga is Wednesday and KickBoxing is Sunday (Table 5 & Figure 11).
* Popular day and time for Yoga is Saturday 14:00-15:00 and KickBoxing is Sunday 15:00-16:00.
* Popular location for Yoga is Hougang Mall and KickBoxing is Punggol 21.

Table 5 Popularity of sessions

|  |  |  |
| --- | --- | --- |
|  | **Yoga** | **KickBoxing** |
| **Peak day**  (Sum of members for each weekday) | Wednesday (261) | Sunday (125) |
| **Peak day and time**  (Sum of members for each weekday with consideration of time) | Saturday 14:00-15:00 (95) | Sunday 15:00-16:00 (125) |
| **Popular location** | Hougang Mall (281) | Punggol 21 (212) |
| **Mode Age** | Female: 35 (42)  Males: 31 (43) | Females: 35 (29)  Males: 34 (37) |
| **Mode Age Category** | Females: 35-40 (153)  Males: <25 (109) | Females: 35-40 (74)  Males: <25 (82) |

*NB: Numbers in brackets are number of members*

Figure 11 Distribution of members for different types of sessions in a week

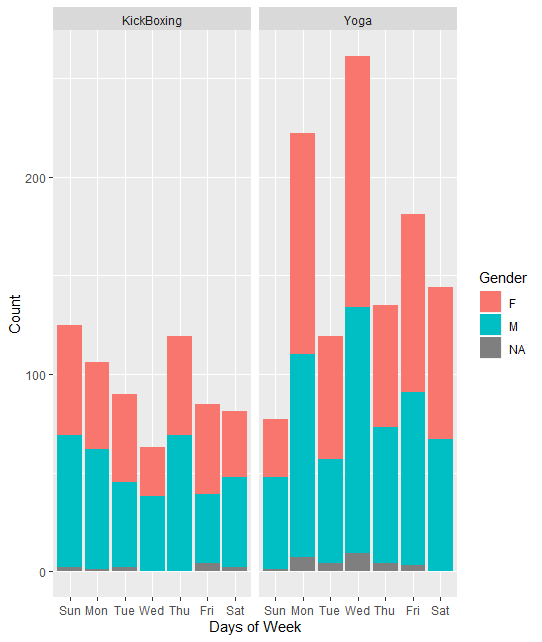
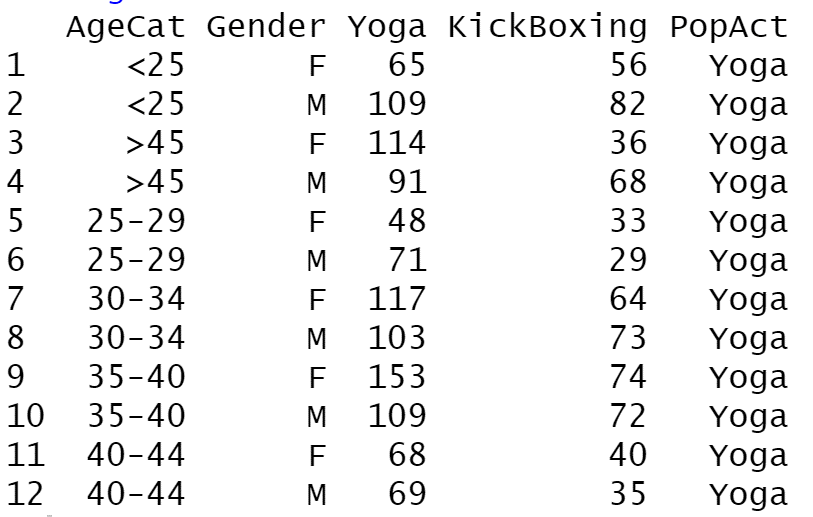


Figure 12 Most popular activity (PopAct) across different age categories and genders

## Valuable and active members

* Fitter-Spot’s most valued members can be defined in terms of **length of membership (loyal)** and **consistent attendance of sessions (highly active).** To be considered loyal, members have to fulfil the criteria below

1. Length of membership: Membership > 6 mth
2. Activeness of member ship: Top 10% ranked on fitness points *(*Table 6*)*

* The **most loyal member profile coincides with Fitter-Spot’s customer base (30 – 40 years old adults)** as concluded in Sections 2.1 & 2.2.Overall Summary
* Rational: Members with longer membership, especially those who have acquired membership for 6 months and above, are more active (Figure 13).

Figure 13 Distribution of Length of Membership for different genders

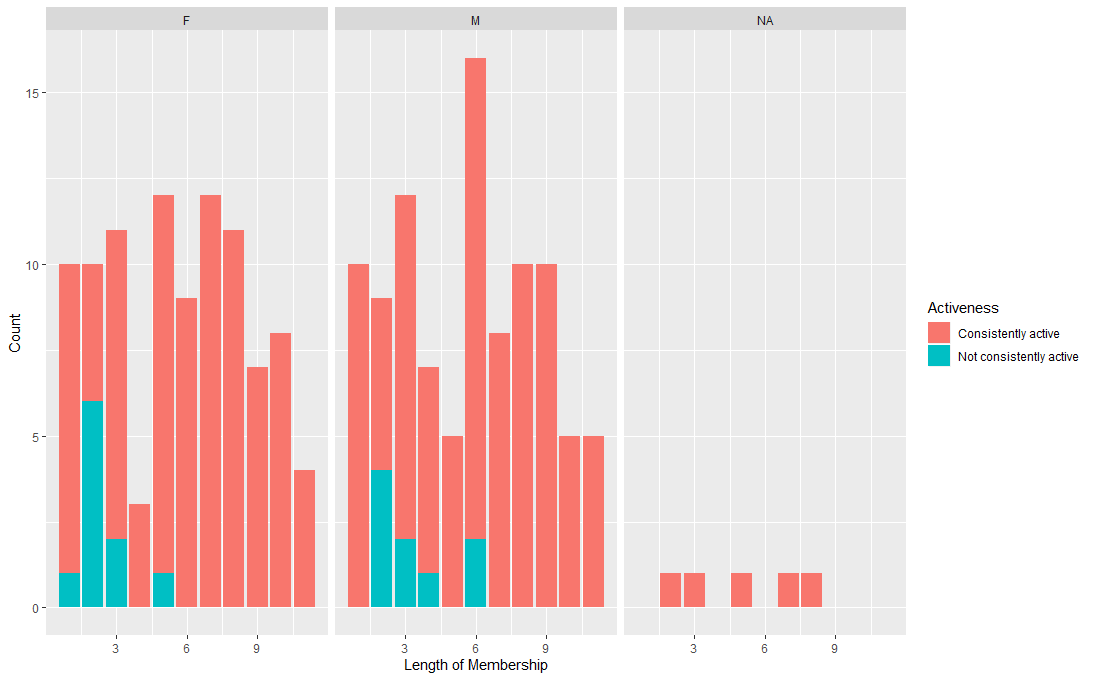


Table 6 Profile of members based on points or activeness of membership

|  |  |  |
| --- | --- | --- |
|  | **Females** | **Males** |
| **Mode Age based on Points** | 35 (2137 points) | 31 (2110 points) |
| **Mode Age based on Average Points Per Month** | 21 (376 points) | 38 (30 points) |
| **Mode Length of Membership based on Points** | 7 months (4394 points) | 6 months (5062 points) |
| **Mode Age of Membership based on Length of Membership** | Yoga: Age 35 and member of 8 months (21 members)  KickBoxing: Age 32 and member of 1 month (12 members) | Yoga: Age 31 and member of 4 months (18 members)  KickBoxing: Age 31 and member of 8 months (14 members) |
| **Points** | | |
| **Top member based on Points** | Lee, Emma (757 points) | Zhao, Alfred (900 points) |
| **Top 10% member based on Points** | * Median age of 34 * 9 months membership * Attended median of 9 Yoga and 3 KickBoxing sessions * Earned median total points of 631 * Most popular location for both Yoga (59 members) and KickBoxing (32 members) is Hougang Mall * Most popular time for Yoga is Mon 12:00-13:00 (19 members) * Most popular time for KickBoxing is Sun 15:00-16:00 (17 members) | |
| **Activeness of Membership** | | |
| **Top member based on Activeness of Membership** | * Lee, Emma (757) | * Zhao, Alfred (900) |
| **Top 10% member based on Length of Membership** | * Median age of 30.5 * 9.5 months membership * Attended median of 8.5 Yoga and 3 KickBoxing sessions * Earned median total points of 608 * Most popular location for both Yoga (59 members) and KickBoxing (32 members) is Hougang Mall * Most popular time for Yoga is Mon 12:00-13:00 (15 members)   Most popular time for KickBoxing is Sun 15:00-16:00 (15 members) | |

# Recommendations

* Based on the findings in Section 2, we propose the following recommendations to meet Fitter-Spot’s business goals:

Table 7 Recommendations

|  |  |  |
| --- | --- | --- |
| **Findings** | **Recommendations** | |
| **Direction** | **Strategy** |
| Majority member base (30 - 40 years old adults) | Increase customer base | 1. Target working adults, especially those aged 30-40 in future marketing campaigns 2. Carry out campaigns at popular locations (Hougang and Punggol) 3. Consider opening new branches at CBD area where there is large pool of working adults of 30-40 years of age |
| Loyal customers (at least 6 mth of membership and top 10% points contribution) | Retain customer loyalty | 1. Introduce membership package with terms more than 6 months (e.g. 12 months) at discounted rates to new members and encourage recently joined members to upgrade 2. Segment the members by introducing concept of loyalty club with exclusive privileges (e.g. free trial classes for friends brought over on weekends) for members who reach a certain length of membership (e.g. 6 months) |
| Increase fitness points contribution | 1. Increase number of Yoga classes, especially on popular days at popular locations and session times, so that members will not be turned away should the popular sessions be full 2. Encourage members to maintain consistent participation per month with incentives (e.g. free gifts) |

## Other considerations

* There are various ways to evaluate our most valued members. Besides length of membership + points contribution as mentioned in Section 2.5, We could also consider the following:

1. **Average Points per Month (from Jan to Dec)**

* A member who is with us longer will earn more points. In order to obtain a more accurate measure, the average points per month is a normalized measure of points based on the length of membership. Points is a proxy for number of sessions attended.

1. **Length of Membership**

* This will allow us to see the profile of our top 10% most loyal member (in terms of length of membership).

1. **Total number of sessions attended in Dec (Note: Based on data obtained in Dec only)**

* This will allow us to see the profile of our top 10% most active member (in terms of number of sessions attended).
* Below are some of the key statistics based on other ways to evaluate our most valued members:

|  |  |  |
| --- | --- | --- |
| **Average Points per Month** | | |
| **Top member based on Average Points per Month** | Jha, Debrita (154 points) | Ray, Ram (168 points) |
| **Top 10% member based on Average Points per Month** | * Median age of 38 * 2 months membership * Attended median of 9 Yoga and 3.5 KickBoxing sessions * Earned median total points of 186 * Earned average points of 107 per month * Most popular location for Yoga is West Coast (48 members) * Most popular location for KickBoxing is Punggol 21 Community Club (19 members) * Most popular time for Yoga is Sat 14:00-15:00 (20 members) * Most popular time for KickBoxing is Mon 13:00-14:00 (9 members) | |
| **Length of Membership** | | |
| **Top member based on Length of Membership** | * Lee, Emma (757) | * Zhao, Alfred (900) |
| **Top 10% member based on Length of Membership** | * Median age of 34 * 9 months membership * Attended median of 9 Yoga and 3 KickBoxing sessions * Earned median total points of 631 * Most popular location for both Yoga (70 members) and KickBoxing (32 members) is Hougang Mall * Most popular time for Yoga is Mon 12:00-13:00 (19 members) * Most popular time for KickBoxing is Sun 15:00-16:00 (17 members) | |
| **Total number of sessions attended in Dec** | | |
| **Top member based on Total number of sessions in Dec** | * Deol, Tanu (15 sessions) * Reddy, Manisha (15 sessions) * Liu, Elizabeth (15 sessions) * Ahuja, Tanvi (15 sessions) | * Lee, Danial (15 sessions) * Ray, Ram (15 sessions) |
| **Top 10% member based on Total number of sessions in Dec** | * Median age of 38 * 6 months membership * Attended median of 9 Yoga and 5 KickBoxing sessions * Earned median total points of 405.5 * Earned average points of 83 per month * Most popular location for both Yoga (70 members) and KickBoxing (40 members) is Hougang Mall * Most popular time for Yoga is Sat 14:00-15:00 (31 members) * Most popular time for KickBoxing is Sun 15:00-16:00 (25 members) | |