From Occasional to Steady: Habit Formation Insights From a Comprehensive Fitness Study

Ege Demirci[1†], Efe Tüzün[1,2†], Ahmet Furkan Ün[1,2†], Taner Giray Sönmez[1,2†], Onur Varol[1,3*]

[1]Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, 34956, Türkiye.
[2]Mars Athletic Club, Istanbul, 34340, Türkiye.
[3]Center of Excellence in Data Analytics, Sabanci University, Istanbul, 34956, Türkiye.

*Corresponding author: onurvarol@sabanciuniv.edu;
†These authors contributed equally to this work.

Abstract

Exercising regularly is widely recognized as a cornerstone of health, yet the challenge of sustaining consistent exercise habits persists. Understanding the factors that influence the formation of these habits is crucial for developing effective interventions. This study utilizes data from Mars Athletic Club, Türkiye's largest sports chain, to investigate the dynamics of gym attendance and habit formation. The general problem addressed by this study is identifying the critical periods and factors that contribute to the successful establishment of consistent exercise routines among gym-goers. Here we show that there are specific periods during which gym attendance is most crucial for habit formation. By developing a survival metric based on gym attendance patterns, we pinpoint these critical periods and segment members into distinct clusters based on their visit patterns. Our analysis reveals significant differences in how various subgroups respond to interventions, such as group classes, personal trainer sessions, and visiting different clubs. Using causal inference analysis, we demonstrate that personalized guidance and social dynamics are key drivers of sustained long-term engagement. By systematically examining these variables and considering the specific characteristics of different clusters, our research demonstrates the importance of a tailored, multi-dimensional approach to promoting exercise habits, which integrates social dynamics, personalized guidance, and strategic interventions to sustain long-term engagement.

1 Introduction

Many people struggle to make exercise a consistent part of their lives, despite the well-documented benefits of physical activity. Regular exercise is associated with numerous health advantages, including improved cardiovascular health, enhanced mental well-being, and a reduced risk of chronic diseases such as obesity, diabetes, and certain cancers [1, 2]. The World Health Organization estimates that nearly 25% of adults globally do not meet the recommended levels of physical activity, which contributes significantly to the rising prevalence of chronic conditions like obesity, cardiovascular disease, and diabetes [3]. Moreover, approximately 9% of premature deaths are attributed towards sedentary lifestyles globally, which shows the critical importance of encouraging more widespread participation in regular exercise [4].

This gap between awareness of the benefits of exercise and the actual adoption of a regular workout routine is influenced by a complex interplay of behavioral, psychological, demographic, and environmental factors [5–7]. Psychological barriers, such as present bias and time inconsistency often lead individuals to favor immediate comforts over long-term health benefits, resulting in procrastination and a lack of consistent exercise [8]. Additionally, the process of transforming sporadic exercise into a consistent routine generally unfolds in two phases: the initiation phase, driven by immediate motivations like health concerns or aesthetic goals, and the execution phase, where exercise becomes ingrained in one's lifestyle [9]. In addition, consistent routines, simple exercise activities, and positive emotional experiences are significant predictors of successful habit formation. [10]. Milkman also talks about the effect of positive emotional experiences on habit formation, who suggests that temptation bundling is a useful strategy that combines the tasks one needs to do with those that one enjoys. For instance, listening to a favorite podcast only while working out can help overcome the inertia of undesirable habits and foster long-term adherence to beneficial routines like regular exercise [11].

Small incentives can play a surprisingly large role in encouraging exercise adherence. For instance, research has shown that financial incentives can increase gym attendance, especially for those who did not previously attend regularly [12]. One study found that rewarding participants for coming back to the gym after they had missed a previous workout paved the way for an increase of 0.4 more weekly gym visits, with just a bonus of 125 points that would be worth almost 10 cents [13]. These findings align with broader research indicating that even very small incentives can have a disproportionately large impact on behavior [14].

Moreover, social dynamics play a critical role in reinforcing exercise habits. Group exercises, for example, provide motivational feedback and foster a sense of community, which can drive individuals to increase their workout intensity and frequency [15–17]. Small-group, individualized personal training sessions can support basic psychological needs, autonomous exercise motivation, and exercise self-efficacy [18]. Peer influence and social norms are also powerful factors that can significantly impact how regularly and intensely people exercise, which demonstrates that people's behaviors are often shaped by the actions and expectations of those around them [9, 19].

Personalizing exercise programs is also a crucial factor in encouraging long-term physical activity. Studies have shown that personalized training sessions with certified personal trainers help maintain exercise behavior and motivation, particularly in groups such as female college students [20]. Furthermore, supervised, periodic exercise programs, which are often customized to the individual's fitness level and goals, have been shown to lead to greater improvements in physical fitness and adherence compared to self-directed training [21]. This is because the presence of a trainer provides both accountability and expert guidance, which are critical for sustaining motivation over time. Additionally, supportive interactions from exercise instructors, who can offer personalized feedback and encouragement, have been found to promote long-term exercise participation in gym settings significantly, further emphasizing the importance of personalization in fitness programs [22]. Personalized exercise programs that combine moderate and high-intensity training are more effective at improving cardiorespiratory fitness and metabolic health than standardized programs [23]. This suggests that tailoring the intensity and type of exercise to an individual's needs can yield better health outcomes than a one-size-fits-all approach. Moreover, resistance training with a personal trainer often results in participants selecting higher exercise intensities compared to when they train alone, demonstrating the motivational impact of personalized supervision [24]. The ability of personal trainers to adjust workout intensity based on real-time feedback helps individuals push their limits safely, thereby achieving better results.

Despite the comprehensive understanding of these individual factors, there remains a notable gap in the literature concerning holistic intervention models that integrate these insights into cohesive, computational frameworks for practical application. Current research often focuses on isolated aspects of behavior modification without a unified model that accounts for the complex interplay of psychological, social, and environmental influences in a real-world setting. Our research utilizes comprehensive data from Mars Athletic Club, Türkiye's largest sports chain. First, we developed a survival metric based on the frequency and consistency of gym visits to define what constitutes a lasting exercise habit. This metric is a foundational tool to assess the transformation from sporadic to regular exercise patterns among gym-goers. Additionally, we identified distinct clusters of gym members based on their visit patterns, which revealed significant differences in how various subgroups respond to interventions. With this survival metric and cluster-based segmentation established, we applied causal inference methodologies to precisely analyze the impact of various interventions. Propensity score matching allows us to control for confounding variables and isolate the effects of factors such as participation in group classes, personal trainer sessions, and social interactions like inviting a friend to the gym. By systematically examining these variables and considering the specific characteristics of different clusters, our research aims to discover which strategies are most effective in enabling consistent exercise behaviors. The ultimate goal is to integrate these findings into actionable strategies that can be implemented by gyms and health clubs to enhance public health outcomes.

2 Results

2.1 Clustering Based on Visits

We hypothesize that users may differ in terms of their exercise habits. These differences are often attributed to age and gender; however, external factors like exercise environment and daily routines also determine when people find opportunities for these activities. We used Non-negative Matrix Factorization (NMF) to identify distinct behavioral groups when studying daily and weekly gym visits. The idea here was that different groups may have varying tendencies and independent habit formations, and we could incorporate the group-specific differences. The detailed explanations regarding the methodology are presented in Section 4.2. The clustering led us to the five apparent clusters in all clubs, shown in Figure 1(a). Firstly, there is the morning cluster, who visit the gyms predominantly around 8-9 AM. Then, there is the noon cluster, who go to the gym around 12-13. After that, there is the afternoon cluster, who tend to visit the most approximately between 15-17. Next is the evening cluster who go to the establishments around 18-19. Finally, there is the night cluster, who like to visit the gym when it is mostly dark outsize, around 20-21. We notice morning and night clusters exhibit different patterns for weekend days, because the gyms operate between the hours 6-23 on the weekdays and 8-20 on the weekends.

We classified and named these behavioral groups according to the hours that they most visited. The density of the visits presented in the heatmap to reflect the most preferred times for each cluster. The first 6 weeks of visitation patterns for each user is used for clustering, but we show the assigned clusters are consistent when compared to 17^{th} week.

After we established these clusters, we wanted to find out how much each cluster deviated from our overall customer in multiple facets. For that, we looked at the gender and age groups and their occurrence rate in particular clusters. Figure 1(b) shows the deviations from the overall behavior for that group. The deviation measures how much more or less likely to observe particular type of users within that cluster. We calculate deviation as the difference in probabilities that we calculate for different groups. As an example, 0.01 deviation for females in the first subfigure in Fig1(b) implies that in the Morning cluster, it is 1% more likely to find a woman than overall. Conversely, we are 53% less likely to pick a person between the ages 14-20 in the Morning cluster, depicted in the same subfigure. Detailed methodology for the calculations are available in Section 4.2.

Younger people (ages 14-27) tend to do less visits in the morning and noon, as indicated by their negative deviation in the two upper subfigures in Fig1(b). Older people (ages 49+); however, make more visits in the earlier parts of the day, as their morning deviations are in the positive, and the rest are in the negative. This can be attributed to visiting just before work to allow themselves the proper free time after work. Their deviations also stand out, which we associate with less number of people in those groups. Women are more prominent in the noon and afternoon clusters, while men, as they make up the majority of people, do not deviate too much, but visit more often in the afternoon and at night. Most starkly, the people in the ages of 14-20 make

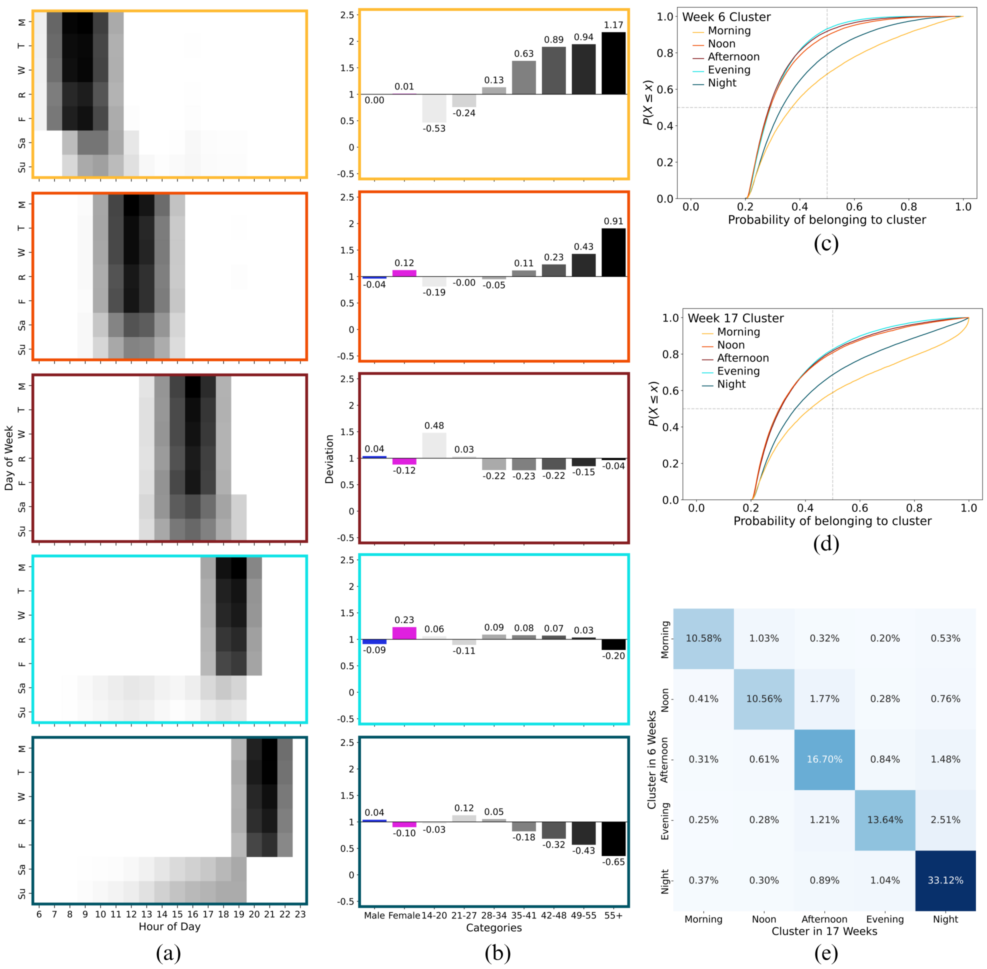


Fig. 1: **Behavioral clusters and preferences by demographics**. Non-negative matrix factorization points to 5 distinct exercise groups. These groups differ mainly by their preference to exercise during the day (a). Some of these clusters are more preferred by some gender and age groups (b). Since the cluster assignments were determined by user activities in the first 6 weeks, we measured the confidence of the NMF model for different models (c,d). Finally, we check the consistency of assignment when data for the first 6 and 17 weeks are considered (e).

most of their visits in the afternoon. This may be due to their schedules allowing them to visit the gym after school/university.

Figures 1(c) and (d) show the cumulative distribution function for the probability of belonging to a cluster for each different cluster, signified by their unique color.

Looking at these figures, it is clear that the NMF decomposition produces more confident results for the morning and night clusters, while the probabilities produced for the people in the rest of the clusters are less confident. This is also evident by the heatmaps produced in Fig1(a), as there are a lot of overlaps for the rest of the clusters, while the morning and night clusters are more separated. We theorize that people that work out in the morning and at night have more fixed schedules, preventing them from visiting at different hours. The rest of the clusters have more flexible schedules, allowing them to change their gym hours more frequently, thus making them harder to assign to a single cluster.

Figure 1(e) presents the cluster transition matrix for the important weeks of 6 and 17. On both axes, we have the 5 clusters, and the annotations represent some percent of the population. Although the previous two figures showed that morning and night clusters were more confidently assigned than the rest, this figure proves that clusters assignments are robust for an extended period of time, as the diagonal values account for more than 84% of the population, which indicates that people do not transition between different behavioral clusters. Additionally, most of the transitions are from clusters with closer times, which is to be expected.

2.2 Proposition on a Measurement of Habit Formation

Understanding the underlying mechanisms of habit formation is key to developing targeted interventions that motivate individuals to attend the gym consistently. We developed the survival metric to measure and analyze the persistence of gym habits. Our approach to this analysis was inspired by the work of Harris and Kessler (2019), who explored habit formation and activity persistence in the context of gym equipment usage. Their study demonstrated that frequent early activity leads to more persistent exercise behavior, suggesting that interventions aiming for behavioral change need to involve higher frequency or longer duration activities to be effective [25]. This metric was defined as the number of consecutive weeks a member went to the gym at least once. We added a tolerance of one week of absence called gap week, recognizing that occasional interruptions (e.g., travel or illness) should not disrupt habit formation. The rationale behind creating the survival metric was to ensure that habit attendance could be measured. This metric could then be used to identify factors that influence the habit formation process.

Our analysis of the survival metric revealed clear patterns in gym attendance behavior and allowed us to identify essential milestones that indicate the progression of habit formation. As shown in Figure 2(a), the cumulative distribution of members' survival series shows a rapid increase in gym attendance in the first weeks. This rapid increase may reflect high initial motivation. However, we observed that approximately 50% of members did not maintain their attendance streak at the 6-week point. This 6-week milestone emerged as a critical point in the habit formation process where members either solidify or begin to abandon their gym habit. When we extended our analysis to users who survived even longer, we observed a steady decline in the number of members maintaining their attendance streak. At the 17-week point, only 20% of the member population maintained their streak, indicating an increasing difficulty in

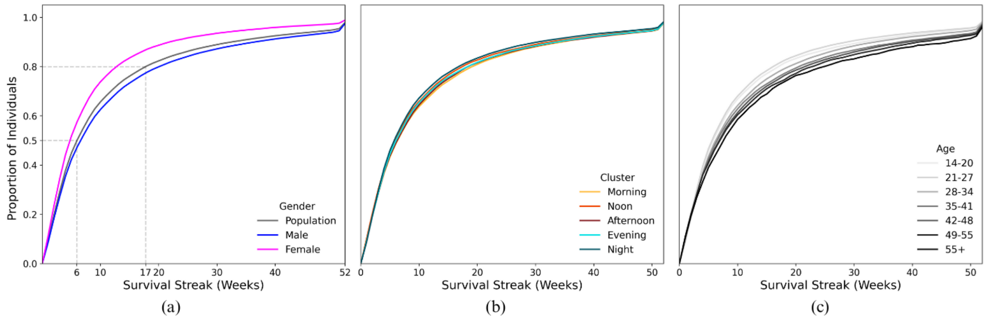


Fig. 2: **Cumulative distributions of the members' survival streaks by clusters and demographics**. We looked at the cumulative distributions of the length of the survival streaks for the whole population and genders (a), different behavioral clusters (b), and stratified groups by age (c).

maintaining long-term attendance. Members with a survival streak between these two weeks are habit-holders.

In Figure 2(b), we examined the survival metric across different user clusters. Our analysis revealed no significant difference in habit persistence across these clusters. As shown in Figure 2(a), when we analyzed the survival metric by gender, we observed a significant difference between male and female members. Male members showed longer survival streaks, indicating marginally better performance in gym attendance over time. Finally, Figure 2(c) examines the survival metric across age groups. The analysis shows that older members tend to maintain their gym habits longer than younger ones. This comparison across different groups suggests that when people exercise during the day, it has less effect on habit formation than their demographics. We will further investigate these dimensions in our causal analysis.

Figure 3(a) shows that while there is a slight decrease, the distribution of gap week usage remains relatively uniform overall. Although the general trend might seem to decrease, as seen in Figure 3(b), the reason for the apparent higher gap usage in the initial weeks is due to the significant proportion of individuals with short survival streaks, as previously discussed, with 50% having a survival value of less than 6 weeks. However, as illustrated in Figure 3(b), for all survival subgroups, the rate of gap usage seems to increase as they approach the end of their survival streak, indicating a higher likelihood of churn. Figure 3(c) reveals that most users do not use more than a few gaps. Moreover, users' reliance on gap weeks decreases as they achieve longer survival streaks. As the gradient shading suggests in Figure 3(c), the likelihood of members using gap weeks decreases as survival streaks lengthen, which could imply that habit strength increases with consistent attendance.

To further assess and track habit formation, we established specific attendance milestones corresponding to critical survival durations. These milestones mark a member's progress in their gym attendance journey. Each milestone is defined by a critical visit count that members must achieve within the specified period. For example, to reach the 6^{th} week milestone, a member must attend the gym at least 9 times within

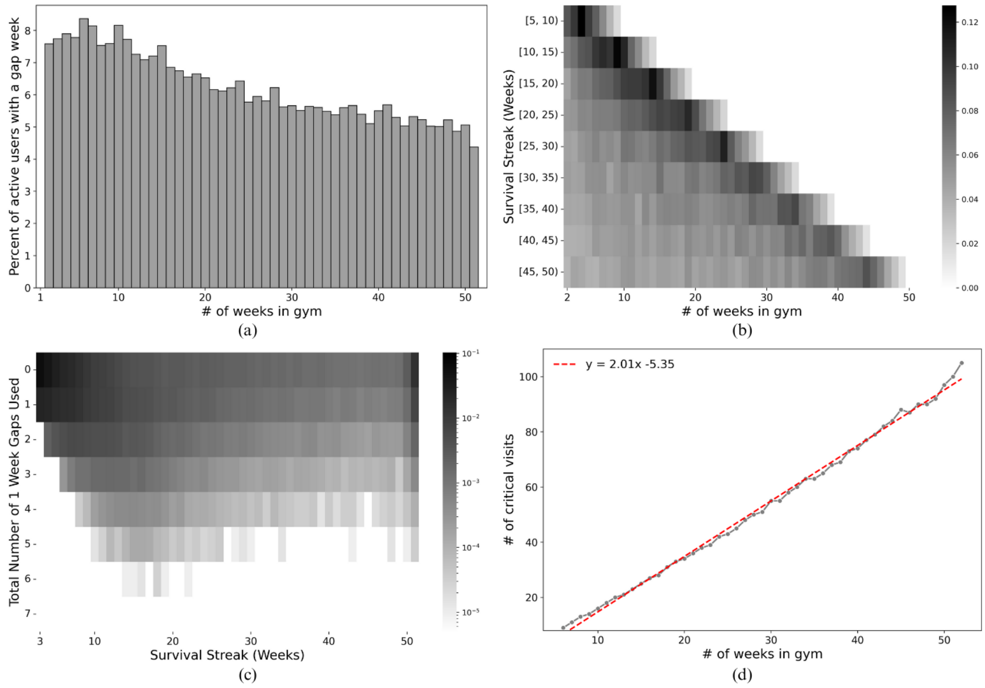


Fig. 3: **1 Week Allowance Analysis and Critical Visit to Reach Milestones**. Since members are allowed to take 1-week gaps during their 52-week journey, we count gaps used for each week (a). To see in which weeks the members take a 1-week gap, we divided them into subpopulations of different survival bins and analyzed the distribution of gaps (b). Members are allowed to use gaps more than once, and we checked the distribution of the total number of used gaps by the length of the survival streaks (c). Survival analysis results in critical visits for each week in the member's journey (d)

the first six weeks of their membership. Figure 3(d) shows an almost linear trend of the critical visits to reach milestones. We fitted a linear model $y = 2.01x - 5.35$, where $m \approx 2.01$, indicating that members need to visit the gym roughly twice per week to meet their milestones. This suggests that attending the gym at least twice a week is important for maintaining consistent habits. The decision to create these milestones was due to the need for a structured approach to evaluate habit formation over time. A user can still attend gyms for an extended period of time, but we use milestones to separate users with the same length of membership into two categories. By focusing on these milestones, we aimed to simplify the tracking of member progress and to identify points at which interventions might be most effective. The details of how these milestones were achieved are discussed in Section 4.4.

2.3 Impact of Interventions on Habit Formation

Deriving meaningful insights in fields such as public health, economics, and social sciences often relies on identifying the underlying causal relationships between various interventions and outcomes. To achieve this, we implemented a causal inference model designed to estimate the impact of specific gym-related interventions on the consistency of exercise habits among gym members. We use propensity score matching (PSM) framework, which allows us to isolate the effect of particular interventions, such as participation in group lessons, personal trainer sessions, visiting different clubs, and invitation credit usage, while controlling for confounding variables that might otherwise introduce bias into the results. In our model, we first identified the treatment and control groups for each intervention. The treatment group consisted of gym members who participated in the specific intervention, while the control group comprised members who did not engage in that intervention but were otherwise similar based on matched propensity scores. Propensity scores were calculated using a variety of covariates, including demographic factors such as age and gender, self-reported experience level, the type of cluster they were in, body mass index, and membership type. These covariates were carefully selected to account for factors that could simultaneously affect both the likelihood of receiving the treatment and the outcome, which in this case is the formation of consistent exercise habits. By matching gym members in the treatment and control groups based on their propensity scores, we aimed to reduce confounding and better approximate the conditions of a randomized controlled trial.

To assess the impact of interventions on habit formation, we focused on the first six weeks of gym attendance, a period identified as critical for habit formation in Section 2.2. There are several variables within these six weeks that we use to analyze their influence on sustained exercise habits. This approach allowed us to observe how early interventions and behaviors contribute to the long-term success of habit formation. Critical visit counts for each week were accepted as the reference point for determining habit formation. By integrating critical visit measure into our model, we were able to provide a more precise estimation of how early interventions, as measured by the critical visits, affect the likelihood of members developing consistent exercise habits.

Once the groups were matched, we estimated the treatment effects by analyzing the differences in outcomes (habit formation) between the treated and control groups. The coefficients derived from this analysis represent the estimated effect size of the interventions. Positive coefficients suggest that the intervention has a beneficial impact on the development of consistent exercise habits, while negative coefficients indicate a detrimental effect. Coefficients close to zero suggest a negligible impact. This method provided us with a robust framework to evaluate the significance of early interventions in shaping long-term exercise habits.

In our study, we employed DoWhy [26], an end-to-end library for causal inference, to implement this causal model. DoWhy facilitated the identification, estimation, and validation of causal effects, allowing us to rigorously assess the impact of different gym-related interventions on habit formation. The overall effects of various gym-related interventions on habit formation, shown in Fig4(a).

The effect of **attending group lessons** starts relatively high but decreases over time. Despite this decline, the impact remains substantial. Notably, members who

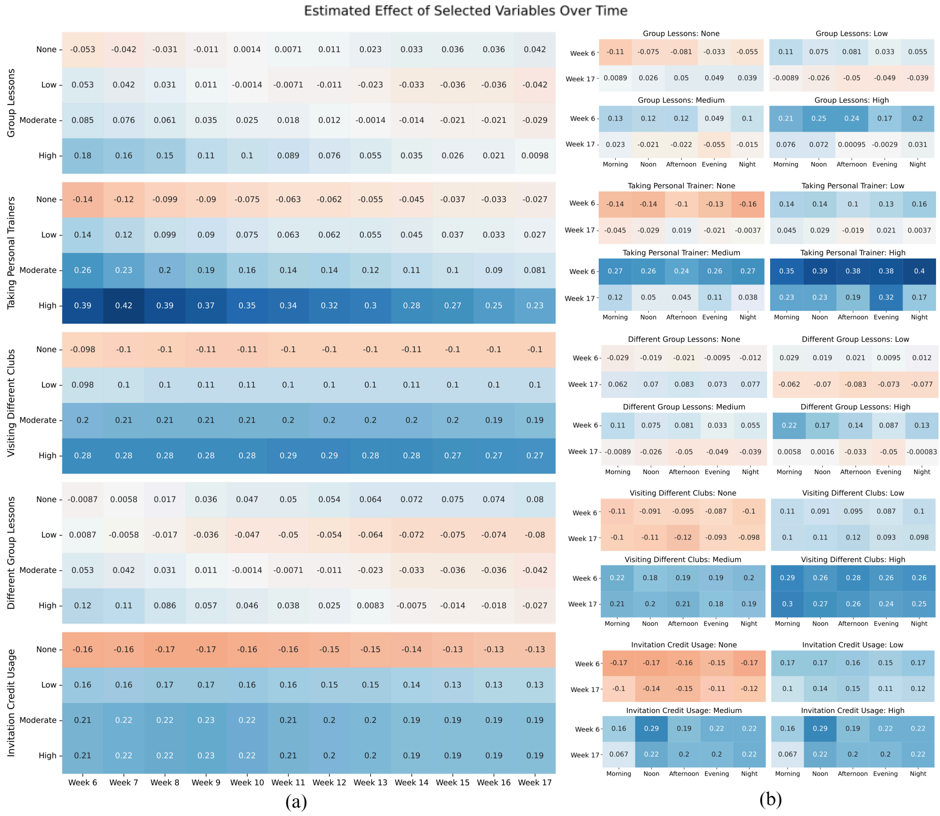


Fig. 4: **Estimated effects of interventions on habit formation**. The overall effects of various gym-related interventions on habit formation, shown from the 6$^{\text{th}}$ to the 17$^{\text{th}}$ week (a). The impact of these interventions across different clusters, specifically highlighting the effects observed at the 6$^{\text{th}}$ and 17$^{\text{th}}$ weeks (b).

attend group lessons at the highest level exhibit significantly greater positive effects on habit formation compared to those at lower levels. The communal aspect of group lessons likely provides both social support and accountability, which are critical in the early stages of habit formation. Additionally, **the number of different group classes attended** shows a similar pattern, with a positive effect on habit formation, though slightly less pronounced compared to the overall attendance at group lessons. This suggests that while variety in group class participation contributes to habit formation, the sheer frequency of participation plays a more dominant role.

Personal trainer sessions exhibit a strong positive impact on maintaining regular exercise habits. The coefficients for personal training are consistently higher than those for other interventions, indicates the significant role of personalized guidance and accountability in habit formation. Personal trainers offer customized workout plans and motivational support, which appear to be significant in helping members establish and maintain their exercise routines.

The effect of **visiting different clubs** is also substantial initially and remains one of the stronger effects among the interventions. This suggests that variety in gym environments significantly supports the development of exercise habits. The exposure to different facilities and equipment likely adds an element of novelty and excitement, preventing the monotony that can sometimes lead to dropout.

Using invitation credits shows a consistent positive effect on gym attendance over time. Unlike other interventions like taking personal trainer and group lessons whose impact may diminish as individuals establish their routines, the effect of inviting friends to the gym remains relatively stable. This suggests that social interactions and shared experiences continue to play an important role in motivating regular attendance, even as individuals settle into their exercise habits. The sustained impact of invitation credit usage highlights the importance of utilizing social connections to maintain engagement and support long-term habit formation.

The results of the cluster analysis also indicates that specific interventions have diverse impacts on distinct subgroups of gym members. The impact of these interventions across different clusters, specifically highlighting the effects observed at the 6^{th} and 17^{th} weeks shown in Fig4(b).

Morning cluster members show more positive responses to taking different group lessons and personal training sessions compared to other clusters. In this cluster, older adults are over-represented relative to the general gym-going population as presented in an earlier section. The variety offered by different group lessons plays a crucial role in maintaining their engagement, as it may introduce novelty and keeps the routine from becoming monotonous. These members benefit greatly from structured and personalized interventions that provide both variety and individualized support. The consistent early morning routine likely helps in establishing a disciplined approach to exercise, which is particularly important for older adults who may prioritize health and wellness.

Noon cluster members exhibit a significantly higher positive response to invitation credit usage and group lessons compared to other clusters. This group tends to be younger adults, with a higher-than-average proportion of females. Since the noon cluster hours coincide with working hours in Türkiye, it is possible that noon cluster members have more flexible schedules. The flexible schedules of the noon cluster may allow them to view the gym as a social hub, where they can combine exercise with socializing, which is particularly appealing to younger members. Additionally, these members show a sustained positive effect from taking personal trainer sessions, particularly by week 17. This indicates that personalized guidance and accountability provided by personal trainers continue to support consistent exercise habits over time, making it an important intervention for this group.

Afternoon cluster members are predominantly within the 14-20 age group, making younger gym-goers over-represented in this cluster compared to others. This demographic is often more experimental and seeks variety in their activities, which explains their strong positive response to visiting different clubs. Interestingly, the effectiveness of working with a personal trainer in this cluster is quite high initially but shows a strong decrease over time. This may indicate that younger members quickly benefit from the initial guidance and motivation provided by a personal trainer but may lose interest or feel more confident in continuing on their own after the early weeks. Additionally, group lessons have a consistently positive impact on this cluster, providing a structured environment that may help sustain their engagement by offering social interaction and variety.

Evening cluster members have a higher-than-average proportion of females, which influences their response to interventions. Unlike other clusters, this group shows a lower positive response to group lessons. Although the female presence is significantly higher, similar to the noon cluster, the cluster behavior differs significantly. The overall effect of group lessons on the evening cluster is negligible, with the impact being lower than the entire population (0.18), suggesting that group workouts may not be as appealing to them. However, this cluster shows the highest and most sustained positive effect from personal trainer sessions by the 17^{th} week, making it the cluster that best maintains its benefits from personalized training. This indicates that for evening exercisers, particularly in a female-represented cluster, personalized attention and tailored workouts play a crucial role in fostering long-term exercise commitment, especially during the evening hours after a day's work.

Night clusters benefits from a mix of interventions, including group lessons and personal training sessions. The late-night hours are typically suitable for those with unconventional schedules, such as students or shift workers. The fact that this group is predominantly younger, with older individuals being strongly underrepresented also supports this. Working with a personal trainer has the most significant effect on habit formation within this cluster during the first 6 weeks. However, this effect diminishes considerably by the 17^{th} week. This decline may suggest that younger members, who form the majority of this cluster, initially benefit from the structure and guidance provided by a personal trainer but may choose to transition to more independent workouts as they gain confidence and familiarity with their routines. The initial high impact of personal training likely helps them establish a solid exercise habit, but as their need for external motivation decreases, they may feel less inclined to continue using these services.

Figure 5 presents the estimated effects of self-reported variables like form level, experience level, and estimated visit frequency over time. These self-reported variables provide insights into how individual characteristics influence the formation of consistent exercise habits, offering a different perspective compared to the direct intervention variables analyzed earlier. These variables are gathered from users through a survey on the mobile app at the start of their membership.

The effect of the **initial form level** on habit formation shows a gradual and consistent increase over time. Members who reported higher form levels at the start were more likely to develop and maintain consistent exercise habits. This contrasts with the impact of direct interventions like group lessons or personal training sessions, where the effect tends to peak early and may diminish over time. The steady rise in the effect of form level suggests that physical fitness provides a robust foundation for sustaining long-term exercise routines, potentially due to greater confidence and physical capability, which enable members to stick to their routines more effectively.

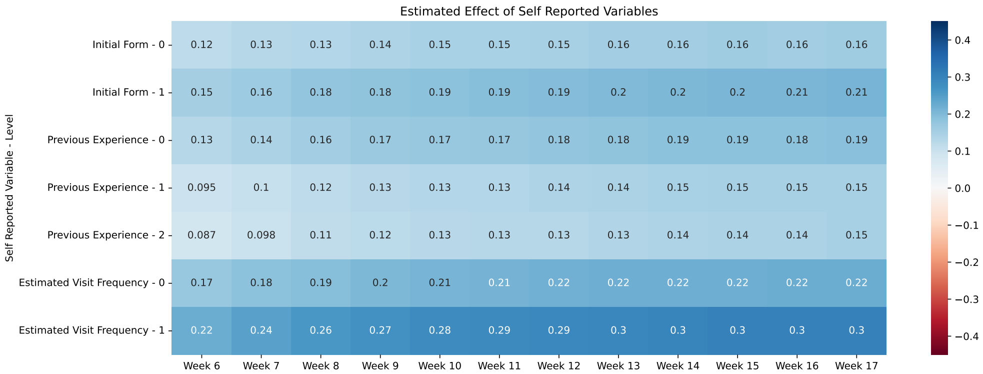


Fig. 5: **Estimated effects of self-reported variables on habit formation.** User-reported characteristics separated by their intensity for previous form, experience and visit frequencies.

Prior experience with fitness exhibits a steadily growing effect on gym attendance consistency over time as well. Interestingly, we do not observe that same clear "more is better" trend among those claiming to have exercised previously: in fact, people with no experience end up slightly surpassing the other experience levels. This phenomenon may hints at a beginner's boost, wherein newcomers may benefit from heightened early support or enthusiasm that temporarily exceeds the advantages one might expect from prior fitness background. Another possible explanation is that individuals who have never engaged in fitness before may find it easier to establish new habits compared to those who have started and stopped exercising in the past, as the latter group might face psychological barriers or carry over negative associations. Despite this, we still see consistent effects among members with previous fitness experience.

The initial **estimated visit frequency** is another self-reported variable that demonstrates a continuously increasing effect on habit formation. Members who were estimated to visit the gym more frequently at the outset were more likely to sustain their exercise routines. The positive trajectory of visit frequency's effect suggests that regularity in the initial stages of gym attendance is critical for embedding exercise as a long-term habit. Unlike some interventions that provide an initial push, a high estimated frequency of visits appears to build a momentum that carries through the critical period of habit formation.

When comparing these self-reported variables to the direct intervention variables analyzed earlier, it is evident that while the effect of most interventions peaks early and diminishes over time, self-reported variables such as form level, experience level, and estimated visit frequency show a steady increase in their influence. This indicates that intrinsic factors related to the members' prior fitness levels and behaviors play a crucial role in sustaining long-term engagement with exercise routines. These findings suggest that while interventions can be effective in jump-starting habit formation, the long-term success of these habits is significantly bolstered by the members' initial fitness levels and commitment to regular exercise.

3 Discussion

This study offers a detailed examination of the factors that influence the formation of consistent exercise habits among gym members. By applying advanced methodologies such as causal inference analysis and cluster-based segmentation, our findings provide nuanced insights that can inform the design of more effective interventions to promote regular physical activity.

One of the key distinctions of our study lies in the development and application of a survival metric to define and measure habit formation. This metric allowed us to capture the persistence of gym attendance over time more precisely than traditional methods. Previous studies that primarily relied on gym equipment data [25], here we offer a survival metric that identifies critical milestones in the early weeks of gym membership specifically, the 6-week and 17-week marks where the likelihood of maintaining regular attendance significantly declines using comprehensive visitation data. This approach provided a structured framework for understanding when and how gym habits are most vulnerable, thus offering clear targets for intervention.

Furthermore, the causal inference analysis in our study marked a significant advancement in understanding the causal impact of specific gym-related interventions on habit formation. Traditional observational studies often encountered challenges with confounding factors, making it difficult to accurately determine the true effects of various interventions [17, 20]. In contrast, our use of propensity score matching allowed us to more effectively isolate the effects of interventions such as group lessons, personal training sessions, visiting different clubs, and invitation credit usage. The analysis revealed that while interventions like personal training consistently demonstrated strong and sustained positive effects on habit formation, other interventions such as group lessons showed diminishing returns over time. However, it is important to highlight that interventions such as visiting different clubs and the use of invitation credits exhibited a more stable impact, with little decline over time. This suggests that while some social and motivational interventions may have temporally limited effectiveness, personalized and varied experiences like those offered through personal training and the opportunity to explore different gym environments maintain their influence over a longer duration. This finding shows the importance of offering a diverse range of interventions to cater to different aspects of habit formation, a detail that has been underexplored in previous studies, particularly those focusing on smaller social exercise settings [18].

Our cluster analysis also introduced a novel way to segment gym members based on their visit patterns, allowing for a more targeted examination of how different subgroups respond to various interventions. Unlike previous studies that often treated gym-goers as a homogeneous group, our segmentation into morning, noon, afternoon, evening, and night clusters revealed significant differences in engagement and responsiveness to interventions. For example, we found that the morning cluster, the peak-time hours for older adults, benefited most from structured and personalized interventions, such as personal training and diverse group lessons. In contrast, the afternoon cluster, dominated by younger adults, responded more positively to interventions that offered variety and novelty, such as visiting different clubs. This highlights the importance of designing interventions that are not only personalized but also adaptable to the specific preferences and schedules of different gym-goer demographics.

Additionally, the analysis of self-reported variables, such as form level, prior experience with fitness, and estimated visit frequency, provided further insights into the intrinsic factors that contribute to habit formation. Unlike the direct interventions, which often showed diminishing returns over time, these self-reported variables exhibited a steadily increasing impact on the likelihood of maintaining consistent exercise habits. This underscores the importance of intrinsic motivation and baseline fitness levels in sustaining long-term engagement.

In synthesizing these findings, it becomes evident that successful habit formation in a gym setting requires a multi-dimensional approach that combines both extrinsic and intrinsic factors. The survival metric highlights the critical periods where interventions can be most effective, while the causal inference analysis provides robust evidence of the causal impact of specific interventions. The cluster analysis underscores the need for tailored interventions that cater to the distinct needs of different member groups, and the role of self-reported variables highlights the value of supporting members' intrinsic motivation.

Building on the insights gained from this study, several avenues for future research and practical implementation can be explored. Firstly, this research aligns with several United Nations Sustainable Development Goals (SDGs) [27], particularly SDG 3 (Good Health and Well-being) and SDG 11 (Sustainable Cities and Communities). Promoting consistent exercise habits contributes to improved public health and well-being, supporting the development of healthier and more resilient communities. Encouraging sustained exercise routines, fitness centers, and public health initiatives can help reduce the incidence of these diseases, contributing to SDG 3's target of reducing premature mortality from NCDs by one-third by 2030 [27].

Secondly, integrating wearable technology and fitness-tracking apps can provide real-time feedback and support a sense of accountability, which are critical components in sustaining exercise habits. Studies have shown that real-time feedback from wearable devices can significantly boost motivation and adherence to physical activity programs by reinforcing positive behaviors and helping users set realistic goals [28]. By providing continuous monitoring and personalized insights, these digital tools can help individuals stay on track with their fitness routines, adjusting their goals and strategies as needed to maintain engagement over time. This approach could be particularly effective when integrated with existing gym-based interventions, offering a hybrid model that utilizes both the precision of digital health tools and the motivational support of personal trainers.

Moreover, strengthening the social aspects of fitness through community-building activities and group challenges can enhance motivation. Investigating the role of social support networks and peer influence in habit formation can provide further insights into effective strategies, especially given our findings on the temporal limitations of social interventions like group lessons. Additionally, conducting long-term studies to monitor the persistence of exercise habits and the effectiveness of interventions over extended periods will provide deeper insights into the dynamics of habit formation and maintenance. Longitudinal studies can help identify the critical factors that contribute to sustained engagement and inform the design of more effective interventions.

Finally, developing comprehensive intervention models that integrate psychological, social, and environmental factors can offer a more holistic approach to promoting regular exercise. Collaborating with behavioral scientists and public health experts can enhance the effectiveness of these models, ensuring that they address the multi-dimensional nature of habit formation. Additionally, machine learning can identify context variables associated with habit formation, which can inform targeted interventions to increase gym attendance, thereby optimizing the personalization and efficacy of these interventions [29].

In conclusion, this study provides valuable insights into the factors that influence the formation of consistent exercise habits. By integrating causal inference techniques and focusing on both structured interventions and individual characteristics, we offer practical recommendations for gyms and health organizations. Future research should continue to explore innovative strategies and technologies to support sustained physical activity, contributing to improved public health and well-being in alignment with the Sustainable Development Goals.

4 Methods

4.1 Data Set

Throughout this project, we analyzed anonymized data provided by Mars Athletic Club. Mars Athletic Club (MAC) is the biggest gym chain in Türkiye, and has over 100 clubs in 13 provinces in the country, with most being in Istanbul, the most populated province. Founded in 2007, the number of locations has been growing steadily over the years.

Due to the COVID-19 pandemic, a lot countries had to set drastic measures. Türkiye was one of the countries that implemented lockdowns in 2020 and 2021, meaning that people experienced lockdowns and unable to exercise indoors. These measures affected a lot businesses, one of which was gyms. As people were not allowed outside for extended durations, their gym memberships became obsolete for the period of pandemic.

Due to the extreme situation surrounding the data, we decided to analyse the data starting from 2022 until the end of 2023. Following the insights of project partners at MAC, we have decided that it would be better if we analyses only the first paid contract for all customers. This was done to mitigate the fact that some users may be on their second contracts during this time frame, and this would mean that they would already have had a chance to achieve some habit formation before, which would mean giving them a second chance. Additionally, we decided to limit the contract type to annual, instead of having different contract types such as 6-month or monthly contracts. This would give every user an equal chance to form their habits in a reasonable time frame or fail to do so. Furthermore, we also filtered out non-paid contracts, on the basis that these were given to employees. Crucially, all data used in this study were fully anonymized to safeguard personal information in compliance with Türkiye's Personal Data Protection Law (KVKK) [30]. This anonymization ensures the privacy of

MAC's members while allowing for a comprehensive analysis of gym usage trends and behaviors without compromising individual confidentiality. No identifiable information was available to the research team at any stage of the analysis.

The variables that we have used to analyse people's habitual differences are listed below.

- **Main club**: Since some contract types allowed for visiting multiple clubs, we chose the club that they signed their contract in as their main club.
- **Membership Category**: Different membership plans offer varying opportunities to customers. These change based on the club type and the package they offer.
- **Number of group classes attended**: Total number of group class participations by a customer. These classes may vary from club to club, but they are unified through various categories such as GFX, core, and cardio.
- **Number of sessions with a personal trainer**: The number of times that a customer worked with a personal trainer. These workouts are sold as sets of 10, and the customer may choose any trainer they wish.
- **Number of different clubs visited**: As some contract types allowed for visiting multiple clubs, we counted the number of clubs visited during their contract lifetime.
- **Number of different Group classes attended**: The number of different group lesson participations made by a customer.
- **Number of invitation credits used**: Total number of credits that the customer has used to invite their friends for a single use. Different numbers of credits are issued for different membership categories every month.
- **Previous experience level**: Self-reported prior experience level of the customer with fitness from MAC+ mobile application. Levels of 0 through 3. Available for around 40.2% of the customers.
- **Form Level**: Self-reported current form level of the customer. For males, it is measured as the number of push-ups they can do in a minute, for females, it is the number of bodyweight squats they can do in a minute from MAC+ mobile application. Levels of 0 through 2. Available for approximately 33.6% of the customers.
- **Estimated Frequency of Exercise**: Self-reported frequency of exercises for the last 3 months. Levels of 0 through 2 from MAC+ mobile application. Available for more than 40% of the customers.
- **Body Mass Index (BMI)**: Calculated from the self-reported weight and height from MAC+ mobile application. Available for all customers in our study.

4.2 Clustering Customers

To understand how the visit patterns of our customers were, we cluster the customers based on their visits. These clusters would serve as one of the fundamental factors in our later analyses. As the first step, we conducted customer clustering on a single club that was representative of the whole customer base, and we later test the generalization. For that, we decided to utilize one of the most popular MacFit branch in Istanbul, which has around 4,500 members. This was a suggestion from the company since they knew beforehand that it resembled the general customer population.

We used Non-Negative Matrix Factorization to identify different behavioural groups based on their visitation patterns. We vectorized all of the visits of customers during their first six weeks by grouping them by the day of a week and hours of a day.

Having seen that the clustering algorithm provided good separation when the number of clusters was 5, we decided to conduct a robustness analysis by experimenting with other clubs. Those experiments also points similar outcomes and we convinced the robustness of this approach, so we replicate the results of the clustering using the whole set of customers. Their visit patters are presented in Fig1(a).

At first, we decided that it was best to categorize every customer strictly, where they could only be part of one cluster. However, later on, we decided that since we would be using propensity score matching, having their probabilities of belonging to a cluster would be better, as we could use the probabilities for the matching algorithm. Thus, we utilized the output of the NMF algorithm to get the probabilities. NMF algorithm tries to approximate the Equation 1.

$$A \approx WH \tag{1}$$

where A is the matrix with features and observations, W is the feature (basis) matrix, and H is the coefficient matrix [31]. Using the feature matrix, it is possible to assign people to clusters, or to calculate probabilities. We used a softmax function to calculate the probability of belonging to the clusters mentioned above. The cumulative distribution functions for the probabilities of weeks 6 and 17 are presented in Figures 1(c,d).

4.3 Cluster-level Deviations for Demographic Groups

In each behavioral cluster, user demographics may differ and certain age and gender groups are more likely to be observed in these clusters. To distinguish the prior probably of observing certain user demographics (D_i) from their preferential occurrences, we calculated the likelihood ratios between conditional probably of being in certain cluster (C_i) following the Eq.2

$$\text{Likelihood ratio for Demographic i's presence in Cluster j} = \frac{P(D_i \mid C_j)}{P(D_i)} \tag{2}$$

This value will be 1 if cluster assignment has no effect on observing certain demographic groups. Deviation from this will indicate whether certain group is more (> 1) or less (< 1) likely to be observed. We obtain the deviations by subtracting 1 from this value.

4.4 Survival Metric and Critical Visit Counts

As briefly explained in the results section, we developed the survival metric to measure gym attendance and habituation. This metric measures the number of consecutive weeks a member attends the gym at least once. The survival metric accommodates real-world interruptions by providing a one-week tolerance for absences, preventing short-term lapses in attendance from prematurely ending a streak. The decision to include this tolerance was based on an analysis of intermediate periods between consecutive gym attendance periods, which can be found in the Appendix D that showed that a one-week gap was most frequently associated with sustained attendance.

However, due to the inclusion of this tolerance period, the number of weeks it took for individuals to reach each survival milestone (i.e., 6, 7, ... 17 weeks) varied across users. As a result, when creating attendance milestones, it was imperative to assess users' habit formation based on the number of visits made during a fixed period rather than the survival longevity. This approach also provided a fixed period for later analysis of users' behaviors and allowed us to investigate the impact of early actions on long-term persistence.

To determine the critical minimum number of visits required to reach a 6-week survival milestone, we compared the distribution of the number of visits made in 6 weeks by the members who survived 6 weeks or less and who survived more than 6 weeks in Figure 6(a). We analyzed the cumulative distribution of visits for both groups, visualized in Figure 6(b), examined the difference between the cumulative distribution functions (CDFs), and determined the number of visits that maximized the difference between the two groups, as shown in Figure 6(c). The critical number of visits was nine within the first six weeks, as this threshold maximized the difference between those who maintained their gym habit and those who did not. This approach allowed us to identify the number of visits that best demonstrated sustained engagement and habit formation during the critical early weeks of membership. Critical visit counts for each week, weeks 6 through 52, are included in Figure 3(d).

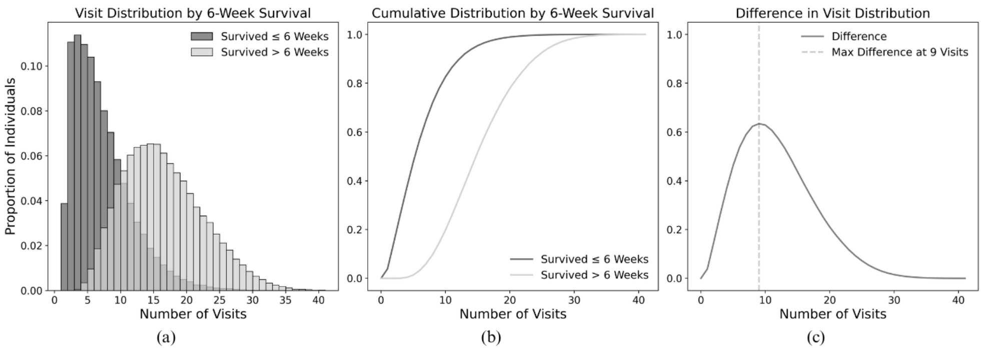


Fig. 6: **Estimating critical number of visits**. We compared the visit distributions of the members who survived less than or equal to 6 weeks and survived more than 6 weeks (a) by looking at their cumulative distributions (b) to determine the critical visit that maximizes the difference between them (c).

4.5 Causal Inference Framework

The objective of our study is to discern the causal effects of various gym-related interventions on the establishment of consistent exercise habits among gym members.

Traditional correlation-based methodologies, though informative, are limited in their ability to distinguish between mere associations and actual causal relationships. To address this limitation, we employed causal inference methods, which are specifically designed to estimate the impact of an intervention (or treatment) on an outcome while rigorously accounting for confounding variables that could otherwise introduce bias. This approach allows us to better understand the true effects of specific interventions and to make more informed recommendations based on our findings.

A key focus of our causal analysis was on the first six weeks of gym attendance, a period identified as critical in our survival analysis. This early phase is crucial for determining long-term habit formation, as we have determined in our survival analysis. To rigorously assess the impact of these early interventions, we specifically analyzed the number of interventions administered within these first six weeks and evaluated their effects over a more extended period, up to 17 weeks. We utilized critical visit counts for each week during this period as reference points, allowing us to observe how early interventions influenced habit formation in the long term. By focusing on this critical window, we aimed to capture the pivotal moments where interventions could make the most significant difference in establishing consistent exercise habits.

To implement our causal inference framework, we utilized DoWhy, an end-to-end library for causal inference [26]. DoWhy is grounded in the formal language of causal graphs and structural causal models, which allowed us to explicitly define our causal assumptions through graphical models. These graphs specified the relationships between treatments, outcomes, and confounders, providing a clear framework for identifying and estimating causal effects. Once the causal assumptions were established, we applied Propensity Score Matching (PSM) to estimate the causal effects of the interventions. Finally, robustness checks were performed to validate our causal assumptions and estimates, ensuring that our findings were not artifacts of model misspecification or unobserved biases.

Propensity Score Matching (PSM) is a robust statistical technique that simulates a randomized controlled trial by accounting for confounding variables, thereby reducing bias in the estimation of treatment effects. In PSM, the propensity score models the probability that a gym member receives a particular treatment given their background characteristics and gym usage patterns. By matching individuals based on their propensity scores, we ensured that the treatment and control groups were comparable in terms of observed characteristics, allowing us to isolate the effect of the treatment on the outcome more effectively. This methodology provided a clearer understanding of which gym-related interventions are most effective in promoting consistent exercise habits.

In our study, the variables used for matching were carefully selected to control for potential confounders, thereby providing a robust and unbiased estimation of treatment effects. The selected variables included demographic factors, membership details, and behavioral indicators, which are detailed in Table 1.

For the first five interventions (group lessons attended, sessions with a personal trainer, invitation credits used, different group lessons attended, and visiting different clubs), all matching variables were utilized. In contrast, for the self-reported variables (estimated visit frequency, form level, and experience level), we excluded experience

	Demographic Variables			Membership Details			Experience and Behavior	
Interventions	Age	Gender	BMI	Contract start	Main club	Membership category	Experience Level	Cluster
Group Lessons	✓	✓	✓	✓	✓	✓	✓	✓
Personal Trainer Sessions	✓	✓	✓	✓	✓	✓	✓	✓
Invitation Credit Usage	✓	✓	✓	✓	✓	✓	✓	✓
Different Club Visits	✓	✓	✓	✓	✓	✓	✓	✓
Different Group Lessons	✓	✓	✓	✓	✓	✓	✓	✓
Self-Reported Variables								
Form Level	✓	✓	✓	✓	✓	✓		✓
Experience Level	✓	✓	✓	✓	✓	✓		✓
Estimated Visit Frequency	✓	✓	✓	✓	✓	✓		✓

Table 1: Variables used in the causal inference framework.

from the matching variables, as these variables are user-specified and self-reported. Furthermore, the analysis on self-reported variables was conducted on the 33.6% of users who provided responses to all relevant questions. For the cluster analysis, the cluster type was excluded from the matching variables to prevent bias.

The matching process employed nearest-neighbor matching without replacement, ensuring that each treated member was paired with a control member who did not receive the treatment but had the closest propensity score. This strict matching process was critical in balancing the treatment and control groups with respect to observed characteristics, thereby improving the accuracy of the estimated treatment effects.

Our analysis centered on the critical period from the 6th to the 17th week of gym membership, as identified in our survival analysis. To facilitate the analysis, we binarized the variables into four categories based on their distribution: `None` for values equal to 0, `Low` for values in the bottom 33rd percentile, `Moderate` for values between the 33rd and 66th percentiles, and `High` for values in the top 66th to 100th percentiles. This binarization approach simplified the analysis by categorizing the variables into distinct groups, making it easier to compare the effects of different levels of the treatment variables. Self-reported variables were binarized differently due to their unique level systems. For these variables, we applied binarization using n − 1 discrete thresholds, where n is the number of levels for the self-reported variable. For instance, **Previous Form - 0** refers to the binarization where individuals with a form level of 0 were grouped separately from those with higher form levels. Similarly, **Previous Form - 1** refers to binarization where individuals with form levels of 0 and 1 were grouped together, with form level 2 constituting a separate group.

To estimate the impact of gym-related interventions on habit formation, we adhered to the following systematic steps:

1. For each treatment-outcome pair, we specified a causal model. The treatments included various interventions such as group lessons, personal training sessions, different club visits, and the use of invitation credits, measured over 6 to 17 weeks.
2. We identified the causal effect using the backdoor criterion to ensure that all common causes (confounders) were accounted for. This step involved defining the relationship between the treatment, outcome, and common causes.
3. We estimated the effect using the Propensity Score Matching (PSM) method provided by DoWhy. This involved calculating the propensity scores and matching treated individuals with similar control individuals who did not receive the treatment.

4. This process was iteratively repeated for each treatment and outcome pair, covering different weeks and intervention types. Specifically, we focused on treatments administered from the 6th to the 17th week, estimating the effects of all interventions sequentially.

As part of our efforts to ensure the accuracy of our estimates, we conducted robustness checks. These involved introducing random common causes into the model to examine the persistence of estimated effects under potential unobserved conditions. Additionally, we employed p-values to gauge the statistical significance of the estimates. A p-value nearing 1.0 signifies non-random observed effects. The obtained p-values for tests with introduced random effects yielded a value of 1.0, instilling confidence in the stability of the results and ruling out random variations as a factor.

Following the identification of clusters, a key aspect of our analysis was assessing the impact of gym-related interventions within each cluster. We applied the `DoWhy` framework independently for each cluster, allowing us to evaluate the causal effects of interventions tailored to the characteristics and behaviors of each subgroup. As we mentioned previously, during this cluster-specific analysis, the cluster variable was excluded from the set of matching variables in the PSM process. This exclusion allowed us to focus on isolating the effect of the interventions within each subgroup, without introducing bias from cluster assignment.

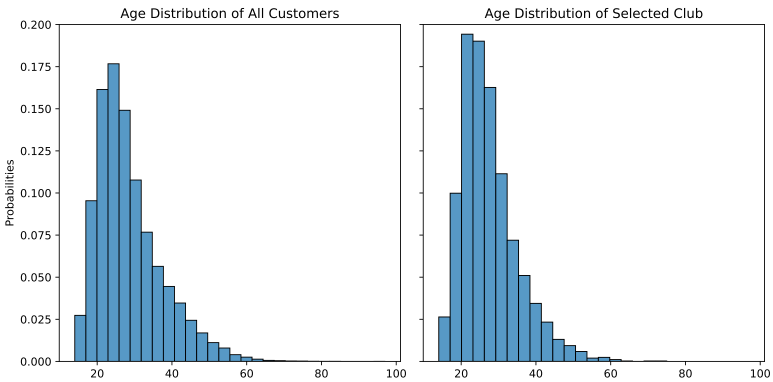
By integrating these methodologies, our analysis provided a comprehensive understanding of how different interventions influence the formation of consistent exercise habits, both overall and within specific clusters of gym members.

5 Acknowledgements

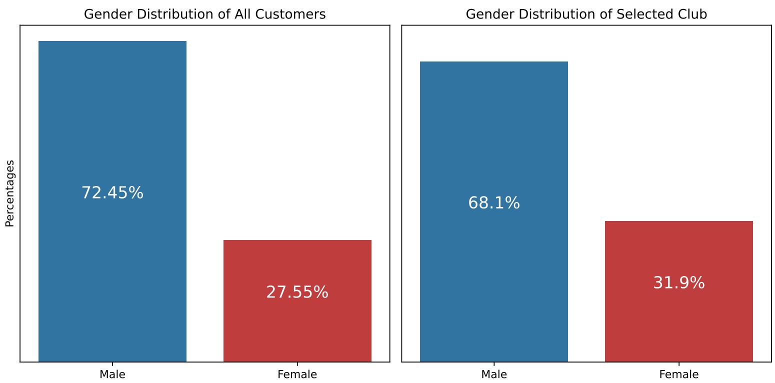
We would like to express our heartfelt gratitude to Mümtaz Demirci, co-CEO of Mars Athletic Club, for his contributions to the study design and interpreting the data. We also thank Mehmet Cenk Bursalı and Emre Çakmak for their insights. Furthermore, we extend our thanks to the members of the VRL lab for discussion.

Appendix A

We hypothesized that users may differ in terms of their exercise habits, which we attributed to various variables. To test this idea, we decided to start small and cluster one of the clubs. The idea to start at a particular club was provided to us by our correspondents at Mars Athletic Club. They mentioned that one of the clubs resembled the overall population closely in terms of age and gender distribution and had a high number of people, which would give us enough data to analyze and feed into the NMF algorithm to produce meaningful results. Below in FigA1, the age and gender distributions of this club are showed. Since these statistics showed little deviation from overall, we went ahead and applied the same procedures to all customers. This first experiment was essential in the process of clustering, as it provided us with what to expect from other clubs as well and shaped the later iterations to come up with the final results.



(a) Age distribution of the customer population vs. selected club



(b) Gender distribution of the customer population vs. selected club

Fig. A1: Age and gender distributions.

Appendix B

To cluster all customers, we needed to vectorize their visits so that we could feed them into the NMF algorithm. Here, we show how the vectorization of all customers' visits were achieved. All customers have their own unique dictionary which represents their weekly visit habits. The dictionaries are then fed into the NMF to produce the distances to cluster centers. Since the gyms worked less hours in the weekends, we set the thresholds differently.

Algorithm 1 Pseudocode for vectorization script.

```
Initialize hours as a dictionary with keys as hourly intervals from 6:00 to 24:00
and values set to 0.
Initialize hours_template as a dictionary with keys as days of the week (0-6)
and values as copies of hour_temp.
procedure PreparePatterns(user_data, entry, exit, day)
    if day is a weekend day then
        for hour in range from entry to min(exit+1, 20) do
            Increment user_data[day][f"hour:00-hour+1:00"] by 1
        end for
    else
        for hour in range from entry to min(exit+1, 23) do
            Increment user_data[day][f"hour:00-hour+1:00"] by 1
        end for
    end if
end procedure
```

Appendix C

When considering the conditions under which a survival streak is preserved or lost, it is essential to recognize the concept of tolerance for occasional absences. Specifically, the metric permits members to miss gym visits for up to one week without breaking their survival streak. This allowance recognizes that individuals may encounter situations that temporarily restrict their ability to attend the gym, such as travel obligations, illness, or exceptionally demanding work or personal commitments. As a result of this tolerance period, if a member fails to visit the gym for two consecutive weeks, their survival streak is broken. This threshold signifies a departure from consistent gym attendance habits, indicating a potential disruption in forming long-term exercise routines.

To determine that a two-week absence is the threshold for breaking the survival streak, we analyzed all intermediate periods between consecutive gym attendance periods, as shown in Figure C1. The Figure displays the cumulative distribution of each intermediate gap between streaks, with 50% of all intermediate gaps being one week long. Therefore, we decided to tolerate these one-week gaps between streaks.

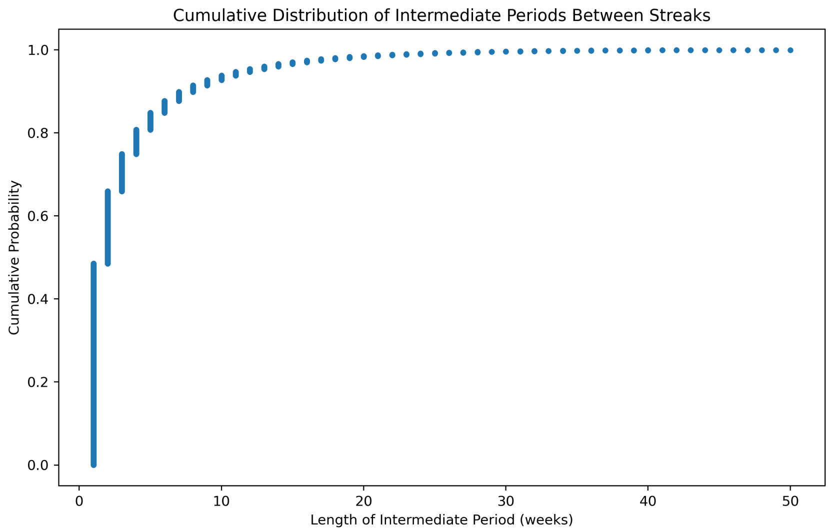


Fig. C1: Cumulative distribution of the intermediate periods between streaks.

Appendix D

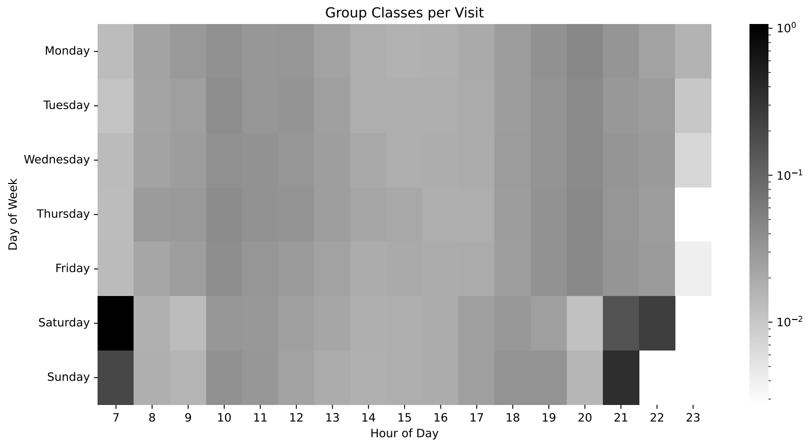
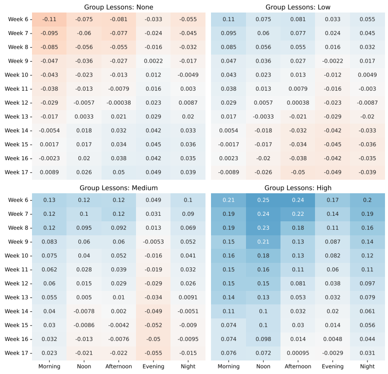
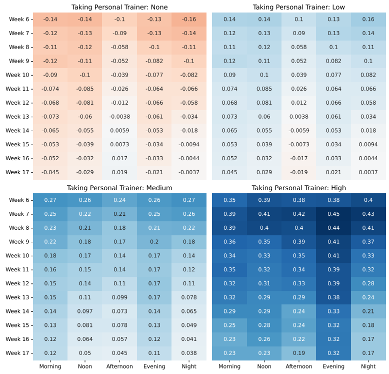


Fig. D1: Group classes opened per visit.

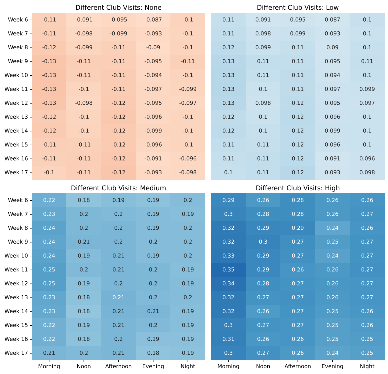
Appendix E



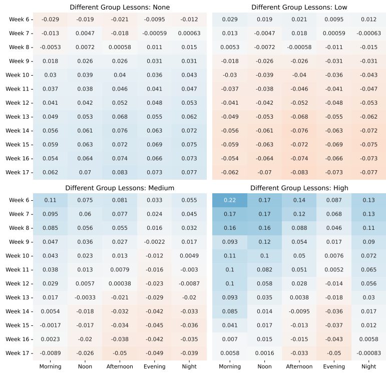
(a) Group Lessons



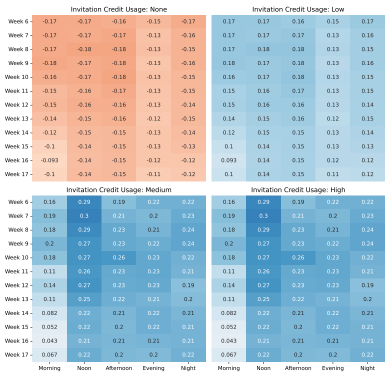
(b) Taking Personal Trainer



(c) Different Club Visits



(d) Different Group Lessons



(e) Invitation Credit Usage

Fig. E1: Estimated effects of interventions on habit formation (individual clusters).

References

[1] Warburton, D. E., Nicol, C. W. & Bredin, S. S. Health benefits of physical activity: the evidence. *CMAJ* **174**, 801–809 (2006). URL https://www.cmaj.ca/content/174/6/801.

[2] Reiner, M., Niermann, C., Jekauc, D. & Woll, A. Long-term health benefits of physical activity – a systematic review of longitudinal studies. *BMC Public Health* **13**, 813 (2013). URL https://doi.org/10.1186/1471-2458-13-813.

[3] World Health Organization. Global action plan on physical activity 2018-2030: more active people for a healthier world (2018). URL https://www.who.int/publications/i/item/9789241514187. Accessed: 2024-08-14.

[4] Lee, I.-M. *et al.* Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet* **380**, 219–229 (2012). URL https://doi.org/10.1016/S0140-6736(12)61031-9.

[5] Quindry, J., Yount, D., O'Bryant, H. & Rudisill, M. Exercise engagement is differentially motivated by age- dependent factors. *American journal of health behavior* **35**, 334–45 (2011).

[6] Sørensen, M. & Gill, D. Perceived barriers to physical activity across norwegian adult age groups, gender and stages of change. *Scandinavian journal of medicine & science in sports* **18**, 651–63 (2008).

[7] Louw, A., Biljon, A. & Mugandani, S. Exercise motivation and barriers among men and women of different age groups: Sport psychology. african journal for physical health education. *Recreation and Dance* **18**, 759–768 (2012).

[8] Royer, H., Stehr, M. & Sydnor, J. Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a fortune-500 company. *American Economic Journal: Applied Economics* **7**, 51–84 (2015). URL http://www.jstor.org/stable/24739047.

[9] Rhodes, R. E. & Rebar, A. L. *Physical Activity Habit: Complexities and Controversies*, 91–109 (Springer International Publishing, Cham, 2018). URL https://doi.org/10.1007/978-3-319-97529-0_6.

[10] Kaushal, N. & Rhodes, R. E. Exercise habit formation in new gym members: A longitudinal study. *Journal of Behavioral Medicine* **38**, 652–663 (2015).

[11] Milkman, K. L. & Duckworth, A. *How to change: The science of getting from where you are to where you want to be* (Portfolio/Penguin, an imprint of Penguin Random House LLC, 2021).

[12] Charness, G. & Gneezy, U. Incentives to exercise. *Department of Economics, UC Santa Barbara, University of California at Santa Barbara, Economics Working*

Paper Series **77** (2008).

[13] Milkman, K. L. *et al.* Megastudies improve the impact of applied behavioural science. *Nature* **600** (2021).

[14] Homonoff, T. A. Can small incentives have large effects? the impact of taxes versus bonuses on disposable bag use. *American Economic Journal: Economic Policy* **10**, 177–210 (2018). URL https://www.aeaweb.org/articles?id=10.1257/pol.20150261.

[15] Zeng, L., Almquist, Z. W. & Spiro, E. S. Let's workout! exploring social exercise in an online fitness community. *iConference 2017 Proceedings Vol. 2* (2017).

[16] Gabay, M. & Oravitan, M. The factors affecting adherence to physical activity in fitness facility settings: a narrative review. *Timisoara Physical Education and Rehabilitation Journal* **15**, 46–61 (2022). URL https://doi.org/10.2478/tperj-2022-0013.

[17] Burke, S., Carron, A., Eys, M., Ntoumanis, N. & Estabrooks, P. Group versus individual approach? a meta-analysis of the effectiveness of interventions to promote physical activity. *International Review of Sport and Exercise Psychology* **2**, 13 (2005).

[18] Wayment, H. A. & McDonald, R. L. Sharing a personal trainer: Personal and social benefits of individualized, small-group training. *The Journal of Strength & Conditioning Research* **31** (2017). URL https://journals.lww.com/nsca-jscr/fulltext/2017/11000/sharing_a_personal_trainer__personal_and_social.24.aspx.

[19] Condliffe, S., Işgın, E. & Fitzgerald, B. Get thee to the gym! a field experiment on improving exercise habits. *Journal of Behavioral and Experimental Economics* **70**, 23–32 (2017). URL https://www.sciencedirect.com/science/article/pii/S2214804317300861.

[20] Fischer, D. & Bryant, J. Effect of certified personal trainer services on stage of exercise behavior and exercise mediators in female college students. *Journal of American college health : J of ACH* **56**, 369–76 (2008).

[21] Storer, T., Dolezal, B., Berenc, M., Timmins, J. & Cooper, C. Effect of supervised, periodized exercise training vs. self-directed training on lean body mass and other fitness variables in health club members. *Journal of strength and conditioning research / National Strength & Conditioning Association* **28** (2013).

[22] Rodrigues, F., Teixeira, D. S., Cid, L. & Monteiro, D. Promoting physical exercise participation: The role of interpersonal behaviors for practical implications. *Journal of Functional Morphology and Kinesiology* **4**, 40–0 (2019). URL https://www.mdpi.com/2411-5142/4/2/40.

[23] Byrd, B. R. *et al.* Personalized moderate-intensity exercise training combined with high-intensity interval training enhances training responsiveness. *International Journal of Environmental Research and Public Health* **16** (2019). URL https://www.mdpi.com/1660-4601/16/12/2088.

[24] Ratamess, N., Faigenbaum, A., Hoffman, J. & Kang, J. Self-selected resistance training intensity in healthy women: The influence of a personal trainer. *Journal of strength and conditioning research / National Strength & Conditioning Association* **22**, 103–11 (2008).

[25] Harris, M. C. & Kessler, L. M. Habit formation and activity persistence: Evidence from gym equipment. *Journal of Economic Behavior & Organization* **166**, 688–708 (2019). URL https://www.sciencedirect.com/science/article/pii/S0167268119302586.

[26] Sharma, A. & Kiciman, E. Dowhy: An end-to-end library for causal inference. https://arxiv.org/abs/2011.04216 (2020).

[27] Nations, U. Sustainable development goals. https://sdgs.un.org/goals (2015).

[28] Patel, M., Asch, D. & Volpp, K. Wearable devices as facilitators, not drivers, of health behavior change. *JAMA* **313** (2015).

[29] Buyalskaya, A. *et al.* What can machine learning teach us about habit formation? evidence from exercise and hygiene. *Proc. Natl. Acad. Sci. U. S. A.* **120**, e2216115120 (2023).

[30] Personal Data Protection Authority. Personal data protection law (2016). URL https://www.kvkk.gov.tr/Icerik/6649/Personal-Data-Protection-Law. Accessed: 2024-08-23.

[31] Kim, J. & Park, H. Sparse nonnegative matrix factorization for clustering. Tech. Rep., Georgia Institute of Technology (2008).