MINGAR Fitness Wearables

Exploring the customer base and device performance of MINGAR's fitness wearables.

Report prepared for MINGAR by UofTears

2022-04-07

Contents

Executive summary

Background & Aim

MINGAR is a company that produces fitness tracking wearable devices. As the wearables market expands, the company has released a new line of products at a lower price - 'Active' and 'Advance' - to compete with other businesses in the industry, namely Bitfit. In order to assess MINGAR's current standing in the market, the aim of this study is to determine the user demographic of its new products as well as investigate if any of its products have performance issues related to sleep tracking for users of certain skin colors.

Key Findings

- The new affordable product lines, "Active" and "Advance", have attracted customers that are located in lower income regions and are slightly older in age in comparison to constomers of MINGAR's other lines.
- For every 6.51e-06 Canadian dollars increase in a customer's median region income, the odds of that customer purchasing a device from MINGAR's affordable line over a device from MINGAR's other lines decreases by roughly 96% (see Figure 2).
- For every 0.013 year increase in customer age, the odds of that customer purchasing a device from MINGAR's affordable line over a device from MINGAR's other lines increases by about 46% (see Figure 1).
- The maximum number of flags per sleep session is higher for customers that use darker skin tone emoji modifiers compared to those that use lighter or default skin tone (see Figure 3).
- The average increase in the number of flags per one more minute of sleep across users of default, light, medium-light, and medium skin tone is 24.55% of the average increase across medium-dark and dark skin tone users.

Limitations

- The UofTears team did not have direct access to customer information such as income. Instead, customer income was estimated using the median income for their area via postal code.
- Emoji skin tone is not an absolute indicator of a user's race or ethnicity, which adds uncertainty to conclusions on the relationship between device performance and skin tone.

Visuals

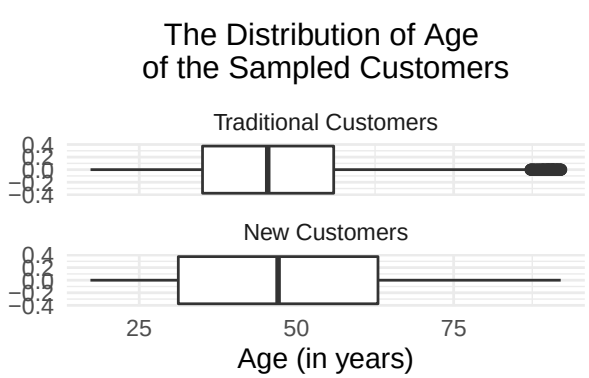

Figure 1: Boxplots displaying ages of customers who purchased MINGAR's new affordable line (i.e. 'New Customers') vs. customers who purchased MINGAR's other lines (i.e. 'Traditional Customers').

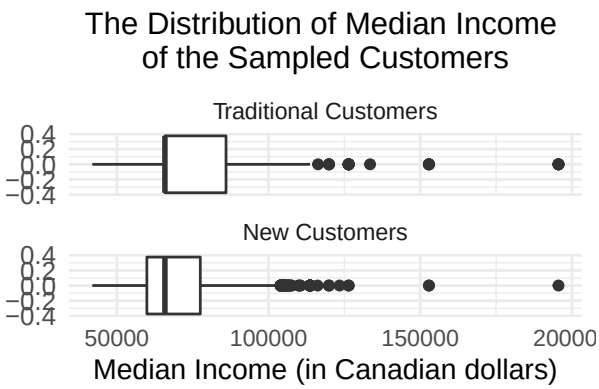

Figure 2: Boxplots displaying incomes of customers purchased MINGAR's new affordable line vs. customers who purchased MINGAR's other lines (i.e. 'Traditional Customers').

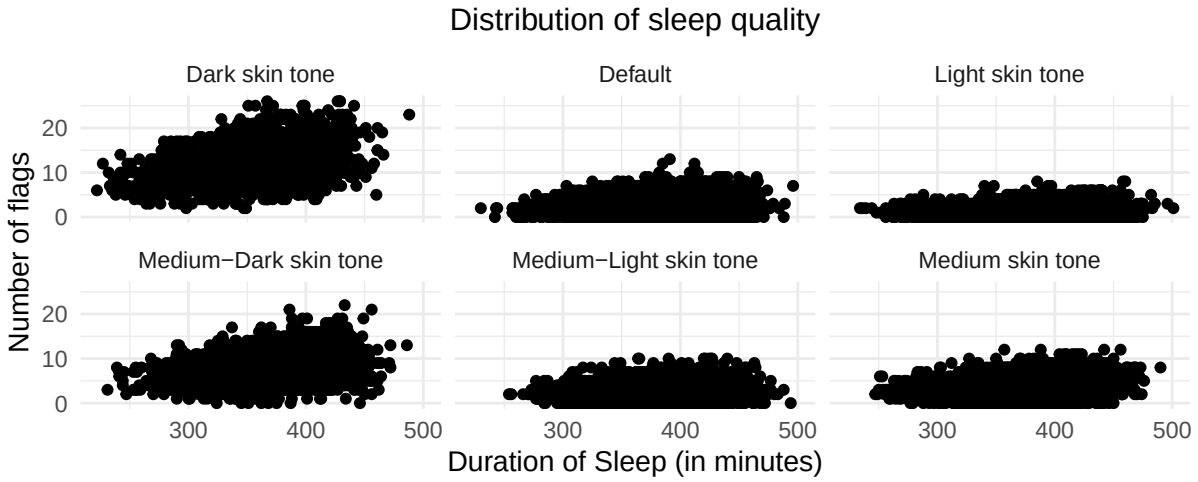

Figure 3: Scatterplots showing the relationship between duration of sleep session and number of flags for each emoji modifier.

Technical report

Introduction

This report provides customer and product analytics on MINGAR's fitness tracking wearable devices with the goal of delivering insights for better product strategizing. To better grasp the customer market of MINGAR's new affordable lines ('Active' and 'Advance'), the UofTears team details demographic characteristics of new customers that these products have attracted thus far- with a particular focus on the targeted lower-income customer segment. Additionally, this report discusses the discrepancy in product performance for customers of diverse races and highlights underlying issues that this presents. The analysis was conducted with respect to ethical concerns and a summary of UofTear's ethical code of conduct is included in the report.

Research questions

- What are the demographic characteristics associated with customers of the new affordable product lines ("Active" and Advance") versus customers of previous fitness wearable lines? This question aims to investigate whether the affordable product lines managed to successfully attract new customers of a lower income target market.
- Are there unusual discrepancies in the quality of the data measured by MINGAR fitness wearables for customers of different races (specifically, those with darker skin tones) and differing amounts of sleep?

Customers of MINGAR's New Afforable Lines

For research question 1, we used the `cancensus` API to retrieve 2016 median incomes for each census subdivision defined by its unique CSDuid - the Canadian census subdivision ID. Next, we downloaded 2016 Census Canada Postal Code Conversion Files from the University of Toronto portal- which indicates the CSDuid for various postcodes. Joining these datasets, we were able to determine median income for each postal code. After that, we joined this newly joined dataset with our customer data on postcode to estimate our customers' incomes. Note that there are customers with more than 1 postcode (and hence, are matched with more than 1 income). For such customers, we assigned them the mean of the multiple incomes. Moreover, we added a new variable "age" to our customer data to inform us of the customer's current age. We then merged this customer information data with customer-device matchup data on customer ID. This allowed us to get information on customers and the specific ID of their device. This dataset was then merged with our device dataset on device ID to retrieve information on each customer's

device. After that, we selected the relevant variables such that our customer's sensitive data are not made public in our report.

As a final data cleaning step, we created a new binary variable in the dataset that takes value 1 if a customer is considered "new" (i.e. is a buyer of our line of "Active" or "Advance" products) and 0 otherwise. This allowed us to build a model that predicts if, given some information about a customer, this customer belongs to the camp of "new" customers or if they belong to the camp of "traditional" customers, which helped us answer the question of who our new customers are and how they differ from those we consider to be our more "traditional" customers.

To start with some exploratory data analysis, we examined the distribution of regional median incomes among the customers from whom we have collected data:

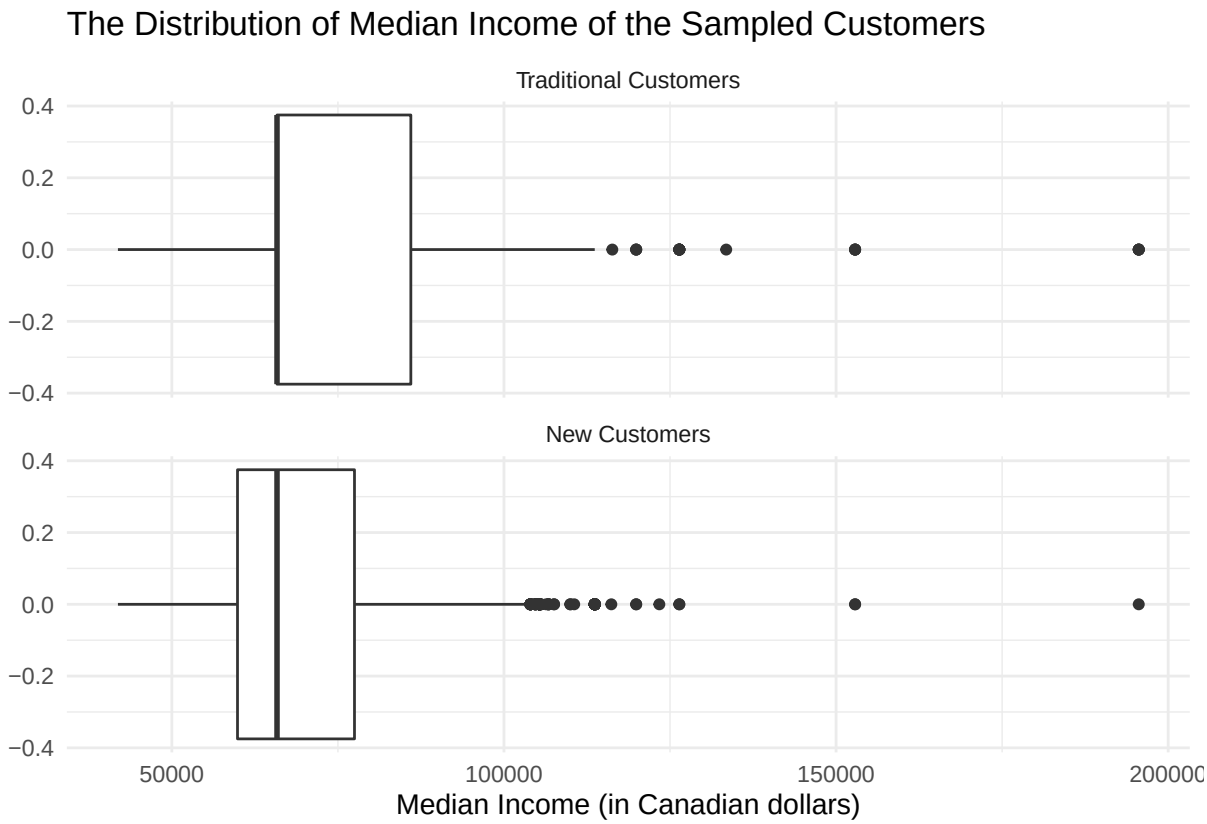

Figure 1: Box plots displaying incomes of customers who purchased MINGAR's new affordable line (i.e. 'New Customers') vs. customers who purchased MINGAR's other lines (i.e. 'Traditional Customers').

From the above boxplots, we can see first that for both traditional and new customers, the distribution of median incomes appears to be skewed to the right with a lot of outliers on

the right tail, which is expected as income distributions tend to skew to the right in most settings. Another similarity between the two types of customers is that the median of the median income for both groups is around the same as seen by the placement of the middle bars in the box. However, one major contrast between the two types of customers is that the interquartile range of the median income for new customers appears to be narrower than that of the traditional customers. This suggests that although the two groups of customers have similar median incomes, the new customers may have slightly less variation in median income as seen by the narrower interquartile range. Therefore, the variable of median income could possibly play a role in separating customers belonging to the new and traditional camps.

Next, we looked at the distribution of the ages of these customers:

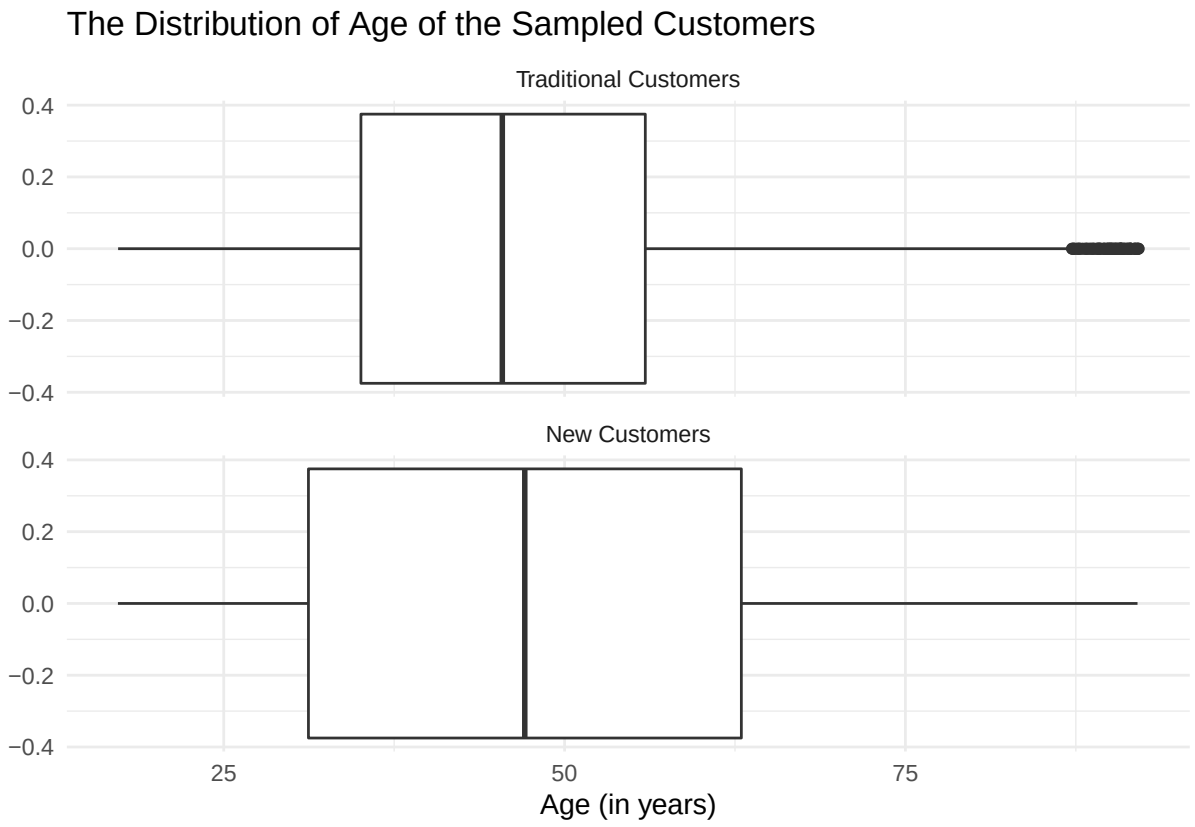

Figure 2: Box plots displaying ages of customers who purchased MINGAR's new affordable line (i.e. 'New Customers') vs. customers who purchased MINGAR's other lines (i.e. 'Traditional Customers').

From the above boxplots, we can see that for both traditional and new customers, the distribution of ages appears to be roughly skewed to the right. Additionally, the median and spread of the ages for both groups are around the same as seen by the placement of the middle bars

in the boxes. The only major difference between the two distributions is that ages considered to be outliers for traditional customers are not considered outliers for new customers - which suggests that the frequency of ages for traditional customers gradually levels off as age increases while the frequency of older ages for new customers suddenly drops off at a certain point. This indicates that most features of the two distributions, such as median and skew, are the same, and therefore, we must be wary of deciding whether to use age for predicting if a customer belongs to the traditional or the new group.

Finally, we examined the frequencies of the different pronouns used by these customers:

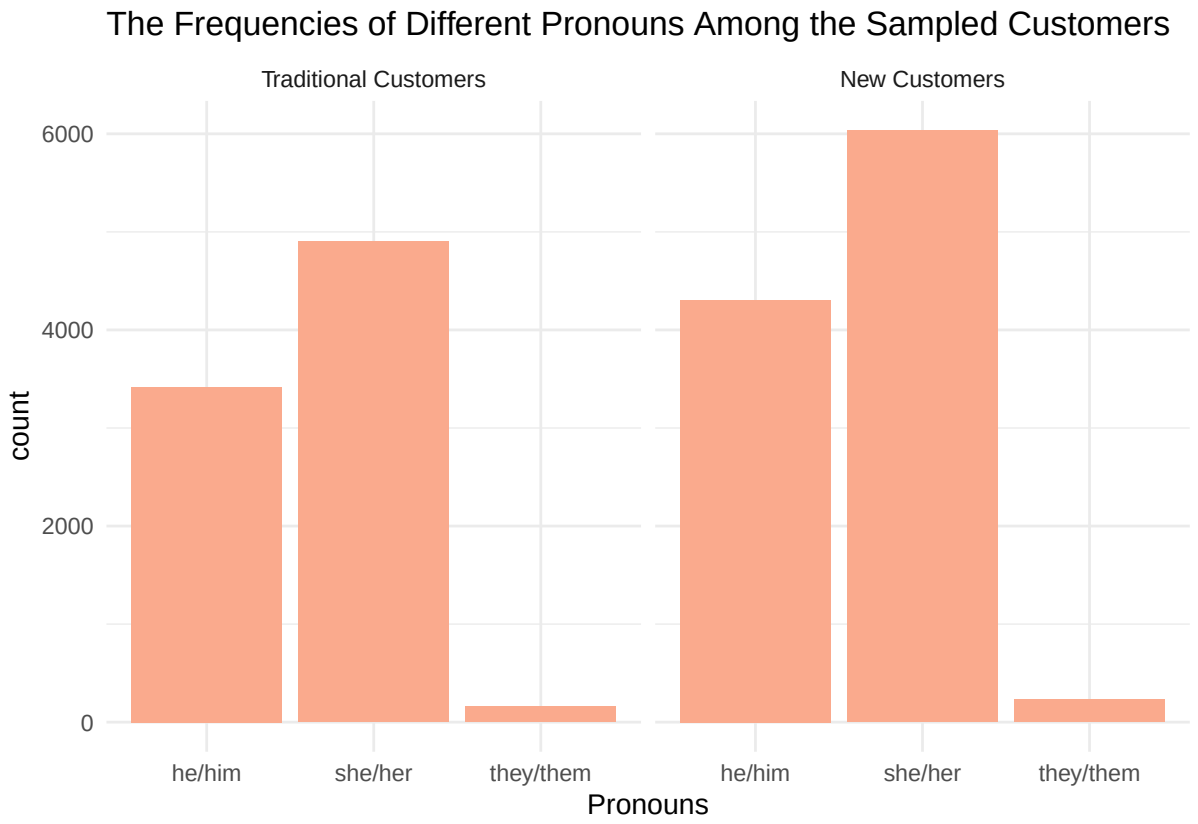

Figure 3: Bar plots displaying pronouns of customers who purchased MINGAR's new affordable line (i.e. 'New Customers') vs. customers who purchased MINGAR's other lines (i.e. 'Traditional Customers').

From the above bar plots, we can see that for both traditional and new customers, the pronouns of "she/her" are most commonly used, followed by "he/him" and "they/them". We can also observe that the bars in the plot corresponding to the new customers are taller than those in the other plot, but that is expected as we know that there are 8476 customers in the traditional camp and 10569 in the new camp. Due to this, we must also be wary of deciding whether to

Table 1: Estimates of Odds Ratios and Corresponding 95 Percent Confidence Intervals

	Estimate	95% CI
Baseline Odds	1.93	(1.78, 2.10)
Rescaled Age	1.46	(1.29, 1.66)
Rescaled Median Income	0.04	(0.03, 0.06)

use pronouns as a predictor for determining whether a customer belongs to the traditional or the new group, given how similar the frequencies of pronouns are across both groups.

Before building preliminary models, recall that as we are trying to gain insight as to what who tends to buy our "Active" and "Advance" line products and how they differ from our traditional customers who buy other lines, we know that the colour of the emoji that customers use has no direct or quantifiable (but still some) impact on whether they belong to the new or traditional camp, which is why we will include this as a random effect.

Now, we have the tools to build our preliminary models. Since the response is binary and we have a random effect, we must use a generalized linear mixed model (GLMM) of the binomial family for this data. We start by fitting a GLMM with all of the potential predictors that we have so far, including pronouns, median income, and age. Then, we create an auxiliary model that doesn't include pronouns as a predictor and another that doesn't include age as a predictor, since in the previous exploratory data analysis we could not determine whether these would be good predictors based on observing the box plots.

Note that the assumptions for the GLMM models are met:

- Our subjects are independent as they are individual customers.
- Random effects are normally distributed.
- All features have constant variance.
- The chosen linked function (logit) is appropriate given the binary response.

By running likelihood ratio tests between the model containing all the predictors so far and the two auxiliary models, we conclude that while a model without age as a predictor will NOT explain the data as well as the model containing all the predictors (due to the LRT p-value of 6.426e-09), a model without pronouns as a predictor WILL explain the data as well as the model containing all the predictors (due to the LRT p-value of 0.1258). Thus, we conclude that pronouns can be discarded as a predictor in our final model.

To interpret the final model whose estimates and corresponding confidence intervals are in Table 1, first note that we were forced to rescale age and median income in order to allow R to properly build the model. These variables were rescaled with the following formulae:

$$\text{age}_{\text{rescaled}} = \frac{\text{age}_{\text{raw}} - \min(\text{age})}{\max(\text{age}) - \min(\text{age})}$$

$$\text{income}_{\text{rescaled}} = \frac{\text{income}_{\text{raw}} - \min(\text{income})}{\max(\text{income}) - \min(\text{income})}$$

where age_{raw} and $\text{income}_{\text{raw}}$ correspond to the original, "raw" age (in years) or median income (in Canadian dollars) value that is to be rescaled, $\min(\text{age}), \min(\text{income}), \max(\text{age})$, and $\max(\text{income})$ correspond to the minimum and maximum age (in years) and median income (in Canadian dollars) values among all customers in the dataset, and $\text{age}_{\text{rescaled}}$ and $\text{income}_{\text{rescaled}}$ correspond to the resulting rescaled age and median income values.

In Table 1, we can first see that the estimate of the baseline odds ratio is 1.93, which means that when both rescaled age and rescaled median income are 0, the odds of a customer belonging to the new camp is 1.93 to 1. Also, by the 95% CI in this row, we are 95% confident that the true odds that a customer belongs to the new camp lies between 1.78:1 and 2.10:1, given that they have a rescaled age of 0 and also a rescaled median income of 0. Since a rescaled age of 0 and a rescaled median income of 0 correspond to the age of 17 years and the median income of 41880 Canadian dollars respectively, we can say that the odds ratio of a customer belonging to the new camp is 1.93 to 1 when that customer is 17 years old and lives in a region where the median income is 41880 Canadian dollars.

In the next row of Table 1, we have that the estimate of the odds ratio corresponding to the rescaled age is 1.46. This means that for every unit increase in rescaled age, the odds of a customer belonging to the new camp increase by about 46%. Also, by the 95% CI in this row, we are 95% confident that for every unit increase in rescaled age, the true change in the odds that a customer belongs to the new camp lies between a 29% increase and a 66% increase. By definition, we have that a unit increase in rescaled age corresponds to a $\frac{1}{\max(\text{age}) - \min(\text{age})}$ increase in raw, unscaled age (in years), so we conclude that for every approximately 0.013 year increase in a customer's age (this is approximately 4.9 days), the odds of that customer belonging to the new camp increases by about 46%.

In the last row of Table 1, we have that the estimate of the odds ratio corresponding to the rescaled median income is 0.04, which means that for every unit increase in rescaled median income, the odds of a customer belonging to the new camp decreases by about 96%. Also, by the 95% CI in this row, we are 95% confident that for every unit increase in rescaled median income, the true change in the odds that a customer belongs to the new camp lies between a 97% decrease and a 94% decrease. Using the same logic as previously, we have that a unit increase in rescaled median income corresponds to a $\frac{1}{\max(\text{income}) - \min(\text{income})}$ increase in raw, unscaled median income (in Canadian dollars), so we conclude that for every approximately 6.51e-06 Canadian

dollar increase in the median income of the region where a customer resides, the odds of that customer belonging to the new camp decreases by about 96%.

These estimates of the odds ratios corresponding to each of the variables in the model tell us that the youngest and lowest income customers have almost 2 to 1 odds of buying our newer and more affordable Active and Advance line products, and this trend is made even more clear by the fact that by even the smallest of increases in median income, the odds of a customer purchasing a product from one of these lines goes down by a staggering 96%. This allows us to conclude that among our lower income customers, our Active and Advance line products have indeed been able to attract considerable attention, and that our new customers mainly reside in regions of lower median income.

Device Performance for Customers of Different Races

To address research question 2, we first retrieved device data scraped from the Fitness Tracker info hub (which can be found here), which gave us information on the device's name, line, retail price, and features. All data was scraped after confirming that we have permission from the website to do so. We also added a few potentially helpful variables to this dataset, including device retail price category and age based on its release date. Using the client provided data on device, we joined the two datasets to determine the device information corresponding to each device ID. From this, we selected only the devices that can track sleep, as sleep score is our main variable of interest. We then join this dataset with sleep data for each customer to match each sleep session to the associated device. From this, we selected only variables related to the investigation to protect customer privacy. We also dropped sessions with missing values for sex, pronouns, and/or postcode because there were not a lot to represent a significant trend, and they are not the predictors we wish to focus on for this question. The emoji-modifier unicodes were also converted to more interpretable labels.

With the data that we had, the only way to measure the quality of the data for a given sleep session was to measure the number of flags that were raised in the session. Furthermore, as we do not directly collect data on race or ethnicity, we were forced to use a proxy variable– using emoji-modifiers to predict each customer's skin tone. The following figure depicts the distribution in the duration of sleep sessions with differing number of flags raised for people with different emoji-modifiers.

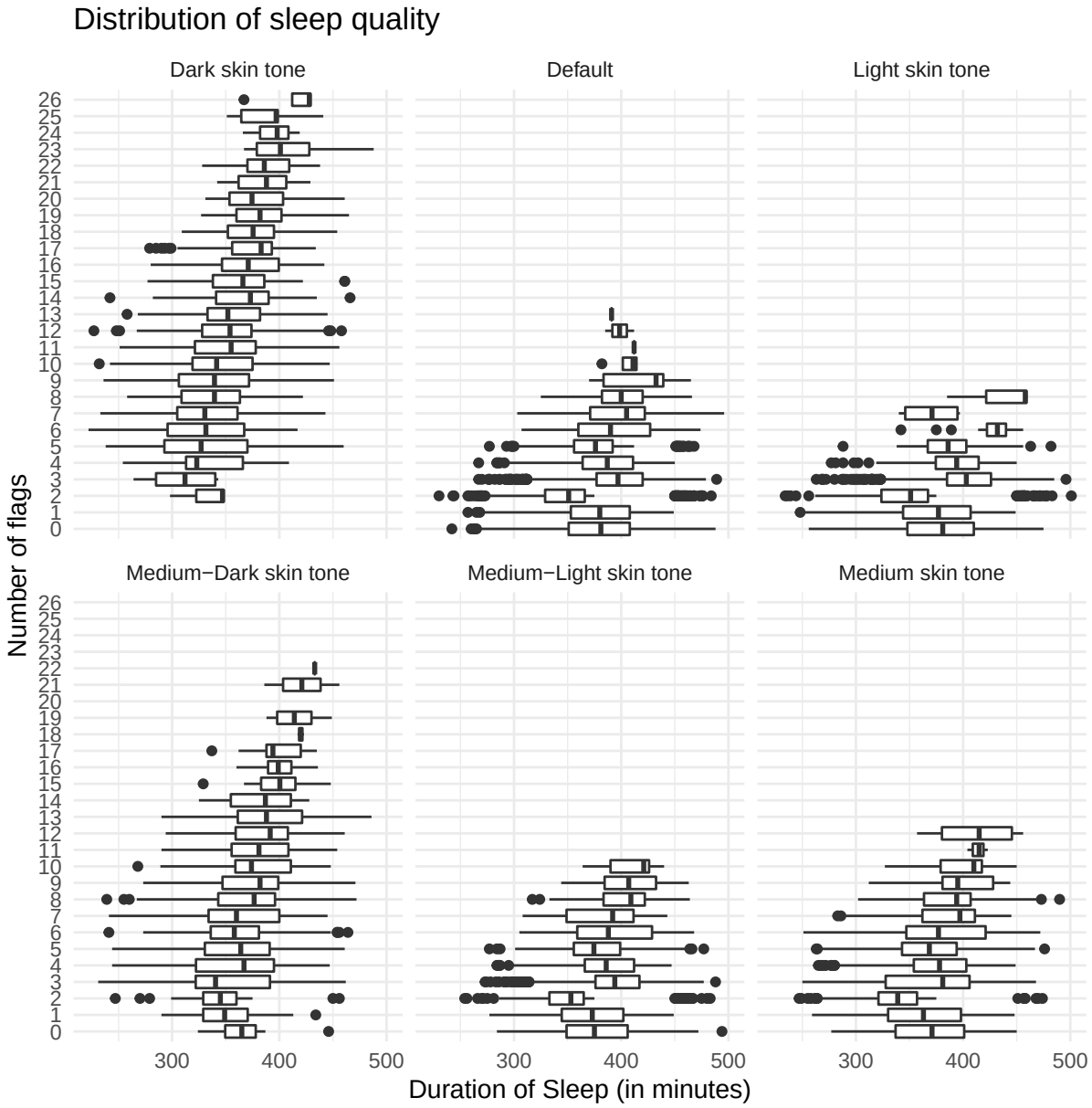

Figure 4: Box-plots that show the distribution of the sleep quality for different number of flags and different emojis used.

As we are trying to understand if there are discrepancies between the performance of the device (i.e., quality of the data measured) for different sleep sessions, our predicted variable was the number of flags raised for a given sleep session. We were primarily interested in how the skin tone and the sleep quality affected these measurements, so we set the emoji-modifiers and the duration of a sleep session as the fixed effects. An ANOVA test led us to include the interaction

between these two variables in our model. This was our baseline model. After performing many likelihood ratio tests on models with and without various random effects (i.e., the customer's age and gender, as well as the line and battery life of the device), we decided to include another model with age as a random effect.

Before fitting the final model, we checked that the assumptions were met. From the figure, note that we can see that when we are given the emoji-modifier of a customer, the relationship between the duration of sleep and the number of flags raised is roughly linear and homoscedastic. Since MINGAR devices all use the same MINGAR sleep-measuring technology and each observation in our dataset is a recorded sleep session, it is fair to assume that devices is not a random effect and thus the units are independent. Moreover, both the random effects and within-unit residuals follow normal distributions. Hence, we decided to fit a linear mixed model. The reason why we avoided generalized additive models is because we sought to maximize interpretability, as we needed to explain our model to our employers.

Table 2: Average Increase in number of Flags for 1 more minute of sleep

Emoji Modifiers	LM Model Increase	LMM Model Increase
Default	0.006494	0.006487
Light Skin Tone	0.003046	0.003052
Medium-Light Skin Tone	0.007467	0.007457
Medium Skin Tone	0.011354	0.011361
Medium-Dark Skin Tone	0.022366	0.022352
Dark Skin Tone	0.035384	0.035391

There were some very interesting findings from our final models, as well as from our boxplot. First, from the boxplot, we can clearly see that only customers that use the dark skin tone emoji-modifier have experienced a sleep session with more than 22 flags. Similarly, only customers with dark skin tone or medium-dark skin tone emoji-modifiers have experienced a sleep session with more than 13 flags.

On to our models. The most meaningful thing to look at are the slopes of the models, which are shown in the table below. The LM model is the linear model with duration and the emoji-modifier (and its interaction) while the LMM model is the linear mixed-model with age as another random effect.

We can see that for both models, as we move from a lighter emoji skin tone modifier to a darker emoji skin tone modifier, the average increase in number of flags for 1 more minute of sleep, which strongly suggests that the device performs more poorly when used by someone with darker skin. This implies that the complaint that the MINGAR social media team is dealing with is indeed valid.

Discussion

Taking a look at the customer base of the new affordable lines (“Active” and Advance" products) through our first research question, we discover that for every 6.51e-06 Canadian dollars increase in a customer’s median region income, the odds of that customer purchasing a device from MINGAR’s affordable line over a device from MINGAR’s other lines decreases by roughly 96%. Moreover, for every 0.013 year increase in customer age, the odds of that customer purchasing a device from MINGAR’s affordable line (‘Active’ or ‘Advance’) over a device from MINGAR’s other lines increases by about 46%. Additionally, customers who are 17 years old and live in a region where the median income is 41880 have 1.93 odds of purchasing a device from MINGAR’s affordable lines.

From this, we can conclude that the new affordable lines have attracted customers that are located in lower income regions and are slightly older in age in comparison to the consumer base

of MINGAR's other lines. Hence, the affordable lines have successfully attracted new customers of a lower income target market.

Next, we examined the relationship between emoji skin tone modifier & sleep session duration with the number of flags present during that session, which could occur due to device error or data quality issues. We found that customers that use dark skin tone modifiers can experience more than 22 flags per sleep session whereas those using default or lighter skin tones experience at most 12 flags. On average, the number of flags for users with dark skin tone emoji modifiers increase by 0.035 with an additional minute in sleep duration. In comparison, those with light skin tone only see an average increase of 0.0030 in number of flags for one more minute of sleep.

These results suggest that there is a indeed an association between the choice of skin tone for emoji modifiers and the quality of sleep data collected for MINGAR's devices, specifically that there seems to be more issues collecting data for longer sleep sessions of customers with darker skin tone emoji modifiers.

Strengths and limitations

The UofTears team prides ourselves on interpretability of our models. Whenever possible, the team makes conscious effort to work with models that are human-understandable. In fact, all finalized models in this report are easily interpretable. Interpretability is important as understanding the decisions an algorithm takes can help to identify underlying prejudice in the algorithm.

With that said, the UofTears team faced a few limitations during this analysis. In the first research question, the team did not have direct access to customer incomes data and was required to make assumptions on a customer's income based on their postal code. While this generalization allows for reasonable analysis, for a more accurate report on market demographics, the team would need direct customer information. With that said, the UofTears team acknowledges that income is sensitive data and collection of such information may cause user privacy concerns. Moreover, the median income data used in this analysis was attained in 2016. Hence, this number may not be entirely reflective of the customer's current income.

With regards to the second research question, one limiting factor is the presence of missing values for user's emoji modifier, which signifies that they used the default skin tone. The UofTears team is thus unable to determine the race/ethnicity of those users, which adds uncertainty and limits the conclusions we were able to make about the relationship between device flags and user race. Similarly, a certain choice of emoji skin tone does not necessarily correspond to the user's actual racial identity.

A suggestion to remedy these limitations is for MINGAR to send out an optional customer survey where customers can volunteer to provide data such as race about themselves. Note that MINGAR must disclose the usage of this data as well as any other possible ethical concerns. Customers must also have the ability to opt out of the survey if desired.

Consultant information

Consultant profiles

Chloe Nguyen. Chloe is a third-year Data Science and Computer Science student at the University of Toronto. She specializes in data science and machine learning. Chloe has experience working in project management and software development at Microsoft's Cloud and AI team- where she collaborated with internal and external stakeholders to create project plans, build prototypes and develop solutions. Chloe is an active member of UofT's Computer Science community, having lead various student unions and clubs-such as President of Women in Computer Science, Treasurer of the Computer Science Student Union and Sponsorship Executive of Neurotech.

Shih-Ting (Cindy) Huang. Cindy is currently a third-year undergraduate studying Data Science at the University of Toronto. With a passion for improving user experience through machine learning, she actively seeks ways to implement different theoretical concepts in real life for business-oriented problems. Her skills in multidisciplinary fields, including data science, marketing, and linguistics, have enabled her to obtain a holistic perspective on multiple projects and devise creative solutions with a user-guided mindset.

Warren Zhu. Warren is a professional Data and Computer Scientist with over 10 years of experience. He was a former researcher at the University of Toronto, specializing in Multilingual Natural Language Processing. Having great understanding and experience with many Mathematical and Statistical concepts, he now serves as a consultant in various fields, including (but not limited to) marketing, health care and linguistics.

Xiaotang (Jeffrey) Zhou. Jeffrey is a third-year undergraduate student at the University of Toronto currently pursuing a joint Data Science Specialist and Mathematics Major degree. With a passion for the manifold applications of data analysis, he has worked on several projects that have involved many different fields, including but not limited to sports and athletics, genetics, economics, and many others. His future goals involve entering the sports industry, where he hopes to use his expertise in data analytics to devise new and creative ways to help his favourite teams succeed.

Code of ethical conduct

The UofTears team is committed to abiding by ethical guidelines for statistical practice when preparing and analyzing data. We promise to avoid disclosing confidential information without prior permission from our clients, such as data on their customers and devices. Moreover, we seek to make careful conclusions about the data and the investigative questions at hand. In this

report, we have listed out all the assumptions we made when constructing the models as well as potential limitations to our research, fully informing the audience of the context surrounding our conclusions. Lastly, we hold responsibility to the society. The goal of this investigation was not only to determine if products could have harmful biases towards certain populations, but also assist the general public in understanding easily how we have obtained those insights.

References

Achim Zeileis, Torsten Hothorn (2002). Diagnostic Checking in Regression Relationships. R News 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Baptiste Auguie (2017). gridExtra: Miscellaneous Functions for "Grid" Graphics. R package version 2.3. https://CRAN.R-project.org/package=gridExtra

David Cooley (2022). geojsonsf: GeoJSON to Simple Feature Converter. R package version 2.0.2. https://CRAN.R-project.org/package=geojsonsf

Dmytro Perepolkin (2019). polite: Be Nice on the Web. R package version 0.1.1. https://CRAN.R-project.org/package=polite

Douglas Bates, Martin Maechler, Ben Bolker, Steve Walker (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1-48. doi:10.18637/jss.v067.i01.

Fitness tracker info hub. (n.d.). Retrieved April 3, 2022, from https://fitnesstrackerinfohub.netlify.app/.

Hadley Wickham (2021). rvest: Easily Harvest (Scrape) Web Pages. R package version 1.0.2. https://CRAN.R-project.org/package=rvest

Hadley Wickham and Evan Miller (2021). haven: Import and Export 'SPSS', 'Stata' and 'SAS' Files. R package version 2.4.3. https://CRAN.R-project.org/package=haven

Hao Zhu (2021). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.3.4. https://CRAN.R-project.org/package=kableExtra

Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10 (1), 439-446, https://doi.org/10.32614/RJ-2018-009

Postal code conversion file. (2016). Toronto. Retrieved April 3, 2022, from https://mdl.library.utoronto.ca/collections/numeric-data/census-canada/postal-code-conversion-file.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Unicode, Inc. (n.d.). Full Emoji Modifier Sequences, v14.0. Unicode, Inc. Retrieved April 4, 2022, from https://unicode.org/emoji/charts/full-emoji-modifiers.html.

von Bergmann, J., Dmitry Shkolnik, and Aaron Jacobs (2021). cancensus: R package to access, retrieve, and work with Canadian Census data and geography. v0.4.2.

Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686, https://doi.org/10.21105/joss.01686

Appendix

Web scraping industry data on fitness tracker devices

Data on fitness tracker devices was scraped from the Fitness tracker info hub. The UofTears team first consulted its scraping permissions of the website, noting that we are allowed to scrape with at least 12 seconds in between each request. We also provided our user agent information to clarify our intentions with the obtained data. Only absolutely necessary data was saved from the webpage. Our team sought not to duplicate the dataset for this study and instead only variables relevant for this study.

Accessing Census data on median household income

Median household income data was attained via the Canadian census API. The UofTears team created an account through CensusMapper, followed the proper guidelines of the API for retrievement and credited CensusMapper in our report. This is open data, available to the public. The use of this data complies with guidelines of the Statistics Canada Open Data Licence.

Accessing postcode conversion files

Postal code conversion data was retrieved from Census Canada via the University of Toronto portal. The UofTears team comprises of University of Toronto students and hence, had the proper credentials to access the data. In downloading the dataset, the UofTears team has agreed to the license agreement. The source for this information has been acknowledged in this report.