Customer Demographics and Product Performance of MINGAR devices

Potential performance bias for users with different emoji modifiers

Report prepared for MINGAR by MANGO

2022-04-11

Contents

Executive summary

Background and aim

MANGO was hired to generate insights for MINGAR's marketing and product considerations. The purposes of this report are threefold: (1) to identify demographic characteristics of buyers of MINGAR's two newer, more affordable lines of fitness tracking wearables, "Active" and "Advance"; (2) to investigate potential performance issues with sleep scores in darker skin users; and (3) to understand product features that drive sales.

The rest of this summary presents key findings and highlights limitations.

Key findings

- Customers of the new affordable "Active" and "Advance" lines are most populated in Ontario, followed by Alberta and Quebec (Table 1).
- The new "Active" and "Advance" product lines sell most in lower-income neighborhoods, meeting marketing goals.
- Female customers make up the majority of the customer case across all product lines (Table 1).
- Customers with darker skin emoji-modifiers have a larger mean number of flags during a sleep session when compared to those with lighter skin emoji-modifiers (Figure1).
- Despite recent launches, lower price points, and less features, the combined sales of "Active" and "Advance" top all other product lines.

Limitations

- Emoji-modifiers were used to predict customer skin tone, therefore limiting the ability to make generalizations on the basis of race and ethnicity.
- Due to the lack of data, customer income was made equivalent to the median household income based on postal code, therefore may not be reflective of an individual customer's true income level.

Table 1: Comparison of old and new line

	Old Line	New Line
Location	Ontario, followed by Quebec, Alberta and British Columbia	Ontario, followed by Alberta, Quebec and British Columbia
Gender	58.6% female	57.8% female
Mean Income	$73166.18	$68815.47
Avg Retail Price	$378.49	$73.32
Sales	8,579	10,674

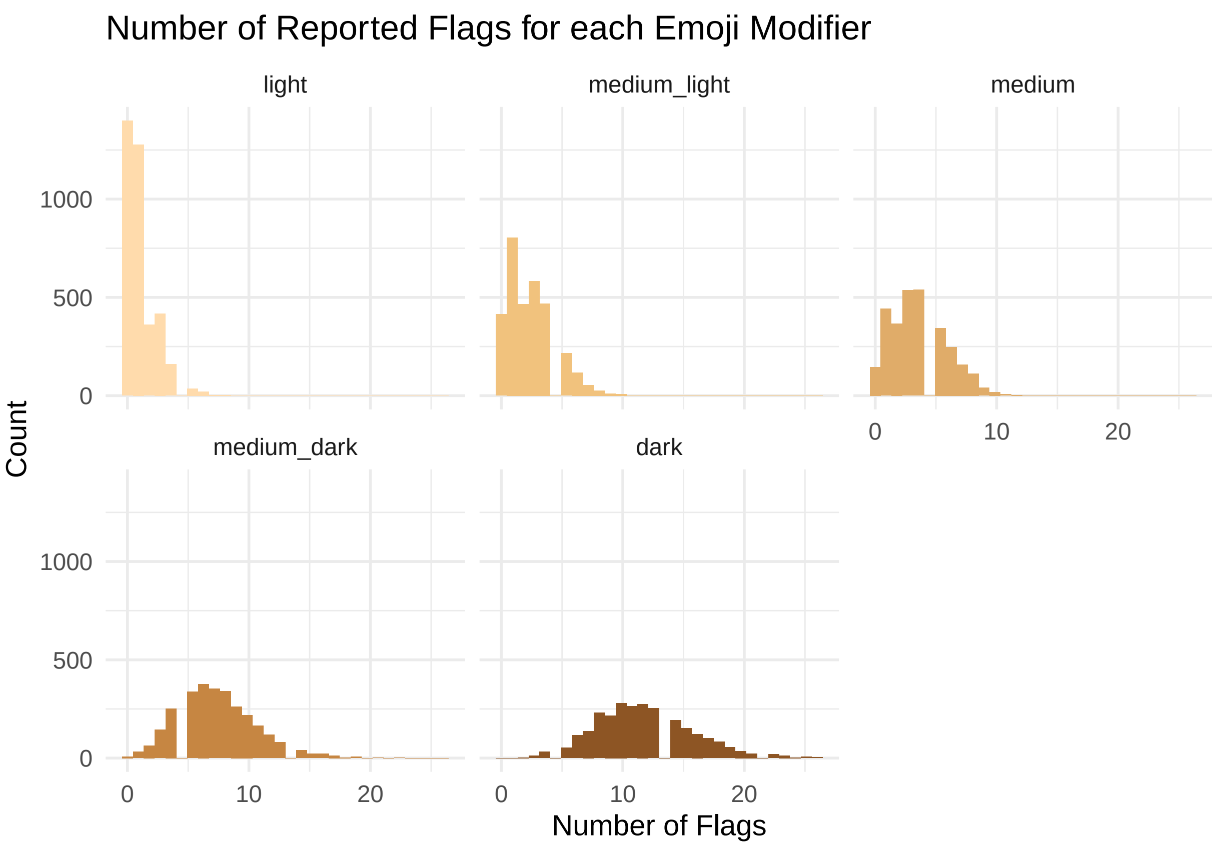

Figure 1: Distribution of Flags for each Emoji Modifier

Technical report

Introduction

Since the early 2000s, MINGAR has been growing its business in the fitness tracker space. Recently, MINGAR has expanded its business to launch two new lines of products, “Active” and “Advance”, both set at a more approachable price point than the traditional high-end product lines. In addition to helping MINGAR understand its new customer base and potential issues around product features better, we conduct data-driven customer and product analyses with goals to inform marketing strategies, help MINGAR stay competitive in the fitness tracker space, and expand its market share in the Canadian market.

Our review and assessments involve internal data on customer demographics provided by MINGAR, as well as external data, including Census Canada postal code conversion file (PCCF) of 2021, household median income data acquired from the 2016 Canadian census, and industry standard device and feature data acquired from the Fitness Tracker Info Hub.

In the following sections, the aforementioned datasets will be used to answer three research questions, with each detailing data manipulation, statistical methods, and conclusions. A summary of findings across the three questions will be presented in the Discussion section, along with the strengths and limitations of the analysis.

Research questions

- How do the demographic for the traditional lines compare with that of the “Active and”Advance" lines?
- How do skin tone emoji modifiers affect performance accuracy of wearable devices?
- What are the main differences between different lines of the device?

Q1: What are the demographic characteristics of MINGAR customers across different product lines?

Location of customers

To create the plot that demonstrate the location of users in different production line, we combined the user dataset with web script Census Canada Postal Code Conversion dataset by postcode. We can therefore generate the province that the user belongs to by extracting the first two digit of CSDuid (Census subdivision unique identifier) from the Census dataset.

In Figure 1, the size of the bubble is proportional to percentage of customers in each province in that line and the position of each bubble represents the latitude and longtitude of that province. From the plot, we can notice that Ontario has the most users in all four lines. In Active line and Advance line, The users in Quebec is slightly less than ones in Ontario, but in Run line, the number of users in Alberta exceeds the number of users in Quebec. From the result, we are able to see where the target customers is located for product in different lines.

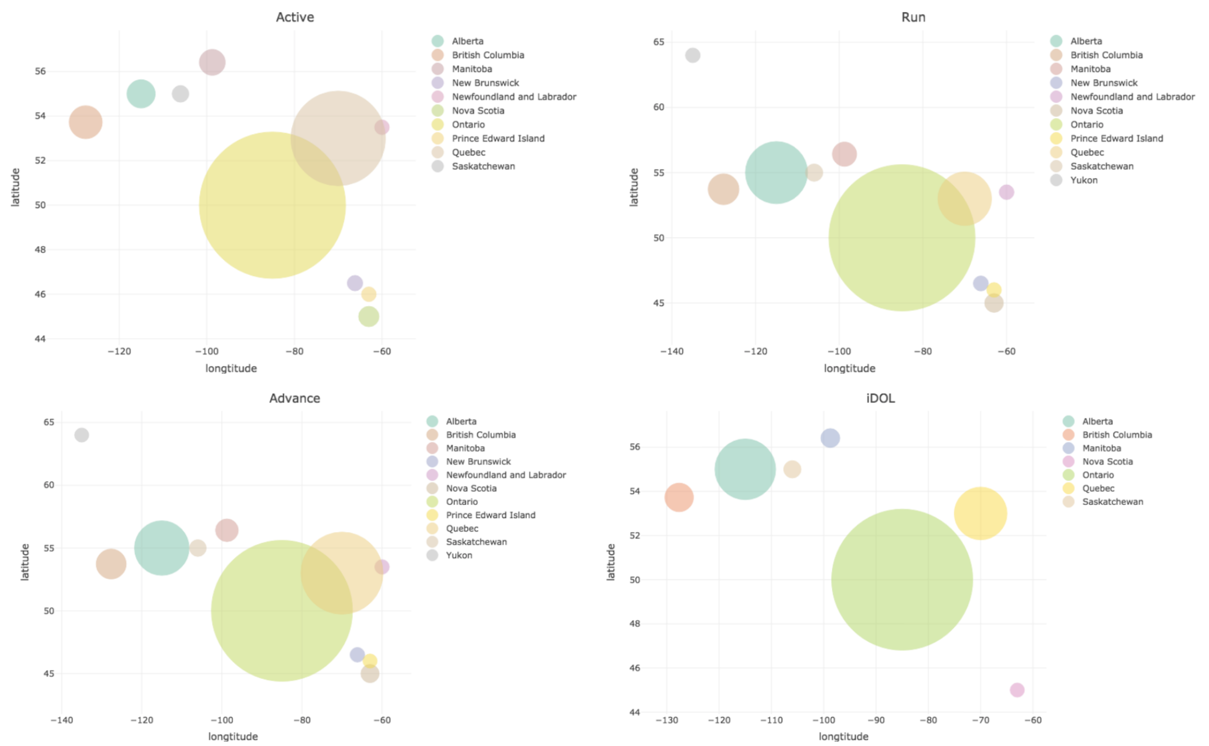

Figure 2: Customer location by device line

Age of customers

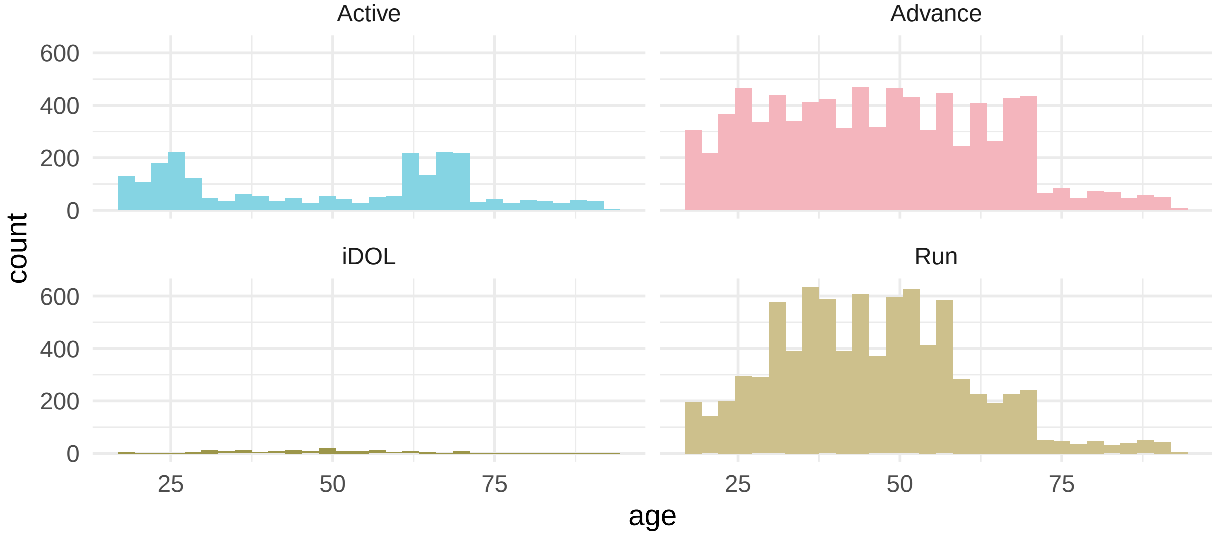

Figure 3: Age of customer by device line

To get a sense of how the age of customers is distributed in different lines, we created a histogram plot of the Age of customer by device line. In the plot above, we can see that Advance, Run and iDOL line have a similar distribution that is slightly right-skewed and there is a bimodal distribution in Active line. We deside to use Kruskal Wallis test since the response do not have a normal distribution and there are more than two categories for the independent variable, `line`.

Assumption Checking

- It satisfies the assumptions that there are two or more levels in the independent variable `line`, the dependent variable,`age`, is on a ratio scale and the observations are independent.
- The fourth assumption is that all groups should have same shape distributions. In our case, three of the groups have same shape distribution, while Active line have a bimodal distribution, which may be caused by the lack of data. It is a limitation in the model assumption checking process.

Kruskal Wallis Test After running the Kruskal Wallis Test, we get a p-value of 5.206^{-13}. It is smaller than the significant level of 0.05, thus we can conclude that the customers in the four lines do not have the same median age.

Gender of customers

To test whether the customers gender in new and traditional lines are different. We created a new variable `line_binary` which have a value `new` for active and advance line, and `old` for other lines.From the bar plot below, we can notice that there are more female customers in both new line and old lines. Since there are more females in either case, we created a new variable `sex_binary` which have a value 1 for female and 0 for other genders, in order to do the two sample t-test.

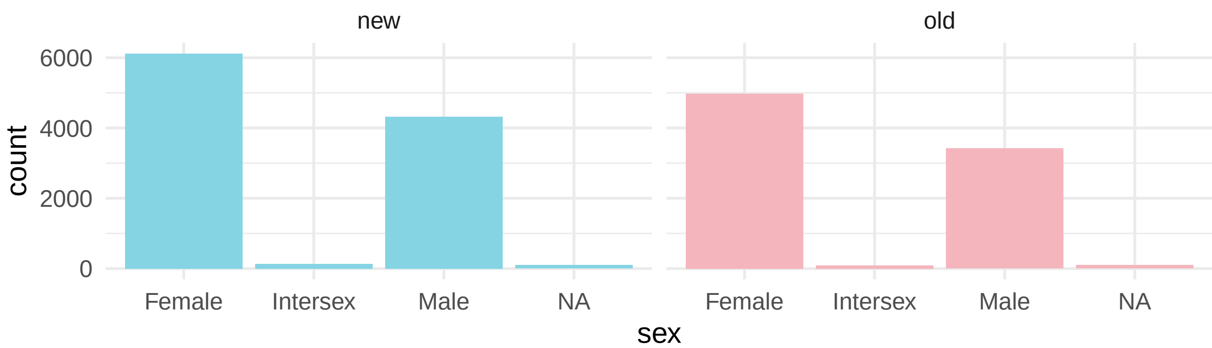

Figure 4: Sex of customer by device line

Two Sample t-test We conducted a two sample t-test on it. The mean in group `new` and group `old` is 0.578 and 0.586 respectively. Which indicates that there are approximatly 58% customers that are female in all lines. Since the p-value is 0.2598, we do not reject the hypothesis that the difference in mean gender equal to 0 between two groups of lines. The result indicates that the gender of customers do not very for old and new lines.

Income of customers

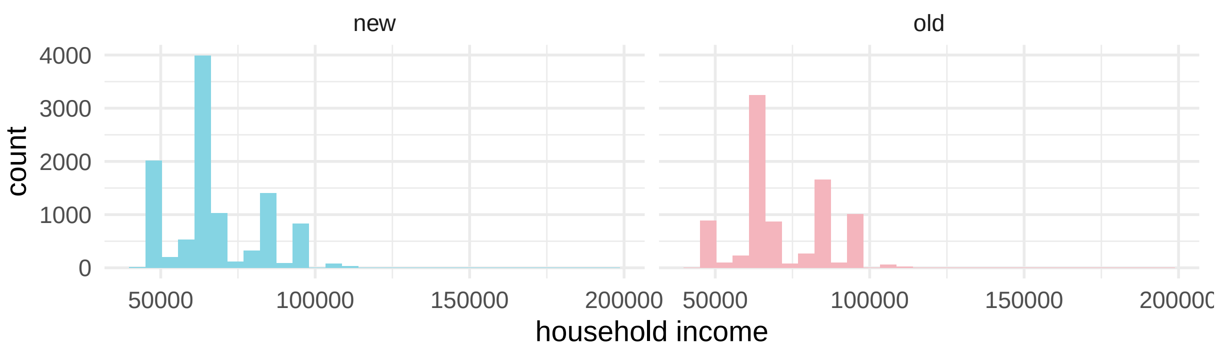

Figure 5: Household income of customer by device line

Two Sample t-test From the two sample t-test, the mean income in group `new` and group `old` is 68815.47 and 73166.18 respectively with a p-value of 2.2^{-16}. Thus, we reject the hypothesis that the two groups have the same mean in income and we can see that the customers who buy products from the new line have around 4000 less mean income than those who buy from the old line.

Q2 How does emoji modifier color affect the accuracy of wearable devices?

Data Manipulation

The dataset for this question is obtained through combining the customer profiles, the sleep tracking, and the device data. The names are cleaned and only variables relating to quality and customer attributes are selected. The newly added variable "skin_tone" divides the users to different skin tone categories depending on their selected emoji modifier.

Exploratory Plots

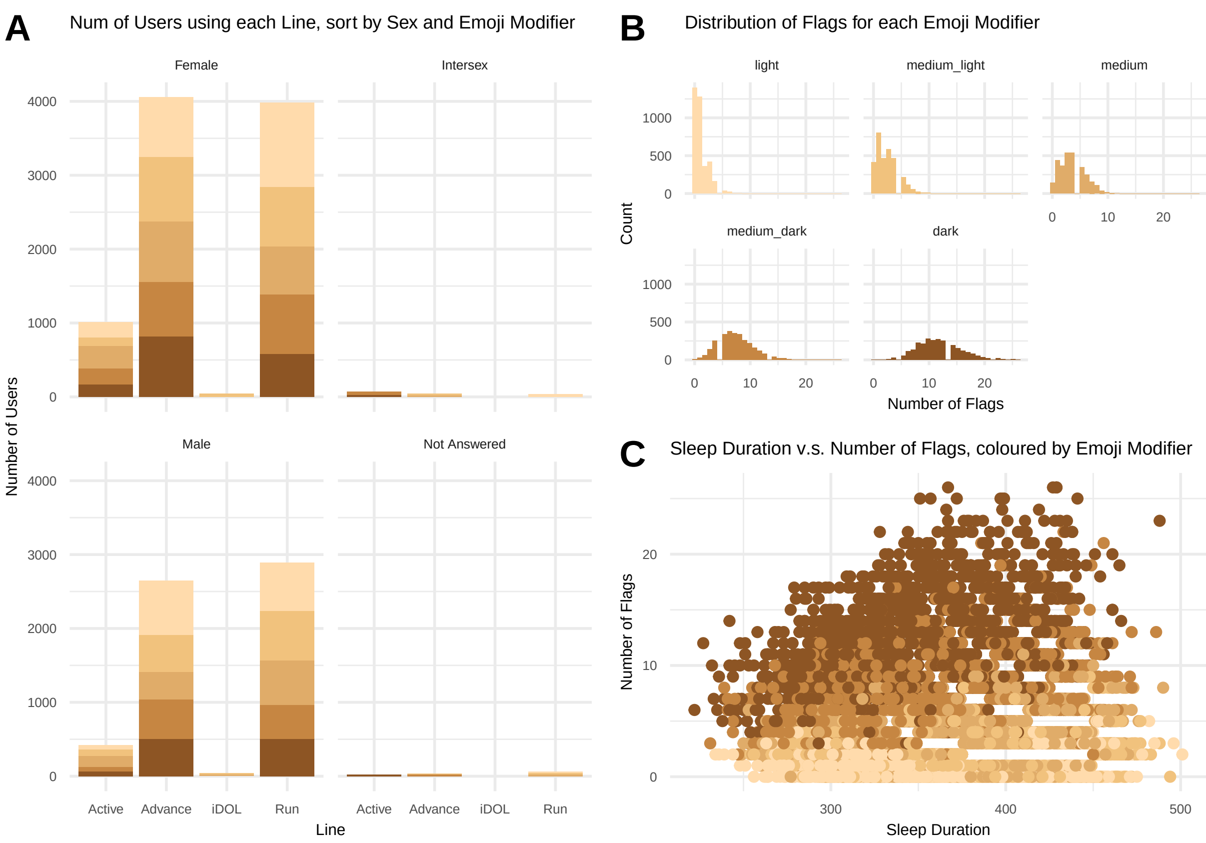

Figure 6: Exploratory plots for users with different emoji modifier

Figure 6A shows that there are more light emoji modifier users than dark emoji modifier users and there are more females than males. Most users use the Advance or Run lines.

Figure 6B shows that darker emoji modifier users seem to have a larger spread and larger mean of number of flags than compared to lighter emoji modifier users, suggesting some evidence of a performance bias.

Figure 6C shows that there is not much relationship between duration of sleep and number of flags, indicating that duration may not be an important predictor for the number of flags.

Fitting a GLMM with Poisson Link Function

Model Assumption Checks

1. Independence of Subjects: the independence of subjects assumption is satisfied as we assume that every customer is independent from each other and are randomly selected from the population
2. Random Effects (each individual) come from a normal distribution
3. Random Effect errors have constant variance:

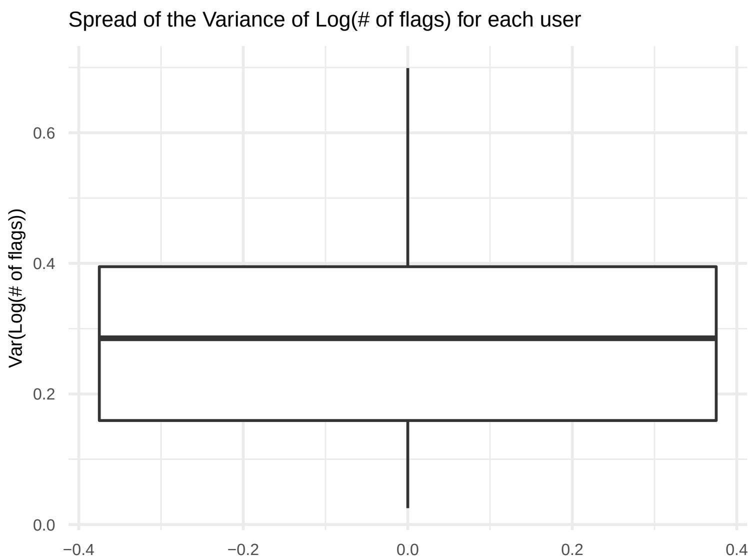

Figure 7: Boxplot of variance of log(# of flags) for all users

We see that the variance of the number of flags transformed by the link function (log) is mostly the same with little spread across individuals. This suggests that the homoscedasiticy of random effect errors assumption is satisfied.

4. Poisson link function is appropriate:

From the Figure B above, we see that the number of flags (response variable) ranges from 0 to over 20, is right skewed, and can be modeled by Poisson distribution. We also check if the mean and variance of number of flags is equal for each skin tone:

Table 2: Compare the mean and variance of number of flags for each emoji modifier group

Emoji Modifier	Mean(# of flags)	Var(# of flags)	Number of Observations
light	1.156522	1.682480	3680
medium_light	2.502364	3.596858	3173
medium	3.653950	4.982877	2962
medium_dark	7.432215	10.163701	2862
dark	11.792333	16.089398	2687

We do observe some evidence of a violation of the mean=variance assumption; however, any violations are modest. Hence, the Poisson model is appropriate in this case.

Model Fitting

$$Y_{irs} \sim Poisson(\lambda_{irs})$$
$$\log(\lambda_{irs}) = \beta_0 + X_{irs}\beta + U_i$$
$$U_i \sim N(0, \sigma^2)$$

- Y_{irs} is the number of flags for individual i with emoji modifier r and sex s
- X_{irs} has indicator variables for emoji modifier and sex
- U_i is the individual level random effect

Table 3: Exponential of the confidence intervals for GLMM model

	2.5%	97.5%
(Intercept)	1.1397195	1.2263361
skin_tonemedium_light	2.0651302	2.2639955
skin_tonemedium	3.0099785	3.2920486
skin_tonemedium_dark	6.1408829	6.6919149
skin_tonedark	9.7524545	10.6208188

	2.5%	97.5%
sexIntersex	0.7986575	1.0322661
sexMale	0.9097382	0.9571648
sexNot Answered	0.8353090	1.1004299

From the model, we see that the general trend is the darker one's emoji modifier is, the higher the log of the mean number of flags their devices report. For example, compared to the base line of the light emoji modifier, we are 95% confident that, compared to light emoji modifier users, the mean number of flags for a user with the dark emoji modifier changes by a factor of between (9.77, 10.65), which is about ten times more. The p-values for each level of emoji modifier predictors are extremely small, suggesting strong evidence against the fact that the mean number of flags are equal for different types of emoji modifiers. We also observe that there is a (1 - 0.91) x 100% = 9% decrease in the mean number of flags for males compared to females; however, this result may just be due to the relatively fewer data for male users than for female users.

Q3 What are the main differences between different lines of the device?

In this research question, we are interested in the differences of the **performance/price** among the different lines of services, combined with the different demographic of the users between lines, we can give a report summary/advice for what has changed and how the change affects the target audience.

Data Manipulation

We used three database in this part. The first one is dataset for device, which is retrieved by web scraping from the fitness tracker info hub. This database includes all the information on the devices produced by the company Mingar and Bitfit. The second database includes the data for individuals who bought the devices from Mingar. This database is obtained by merging the customer and devices ids database and customer info database. The third database used in this question includes the percentage of the devices having some functionality in each line. This dataset is built by group the first dataset by line and functionality.

Exploratory plots

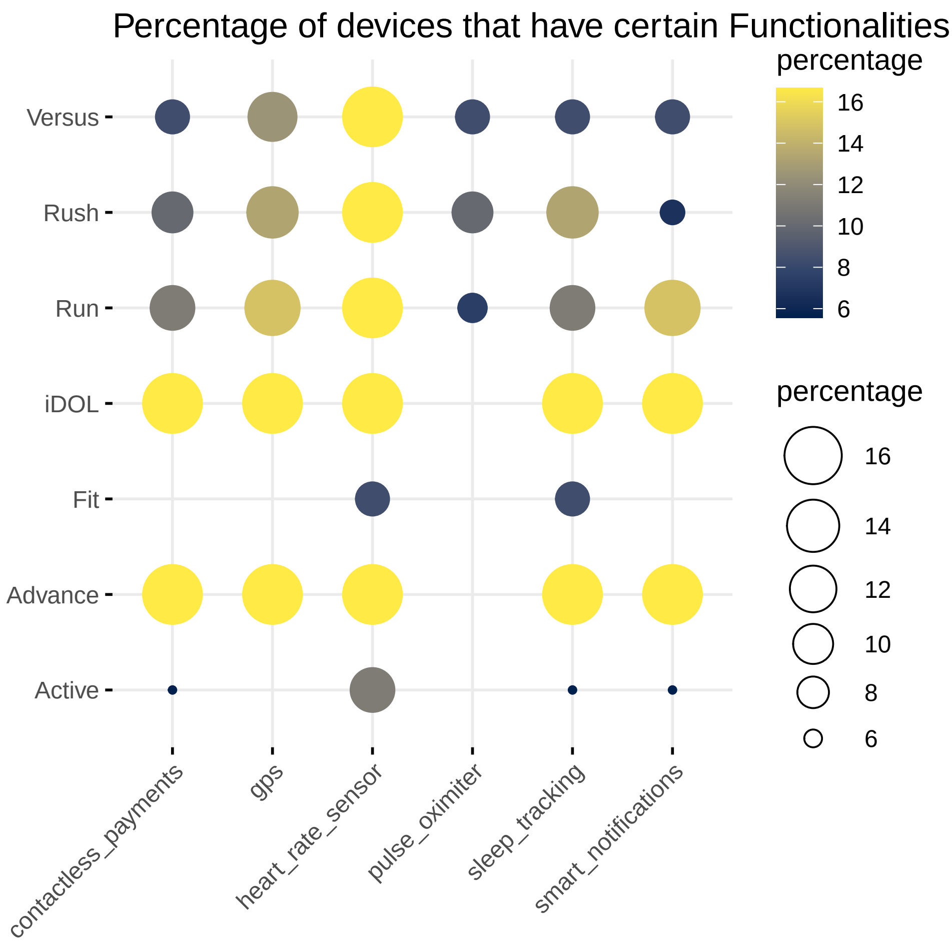

Figure 8: Percentage of devices that have certain Functionalities for each Line

The above plot shows the performance of different lines. The size of the circle is the percentage of devices in the line on x-axis that has the functionality on y-axis. Based on the plot, the line *Run* has all of the functionality, while line *Active* and *Advanced* has all except the pulse oximiter. By the size of each circle, the line *Run* has a more diverse functionality than the other two lines. Each device in line *Advanced* almost has the same functionality.

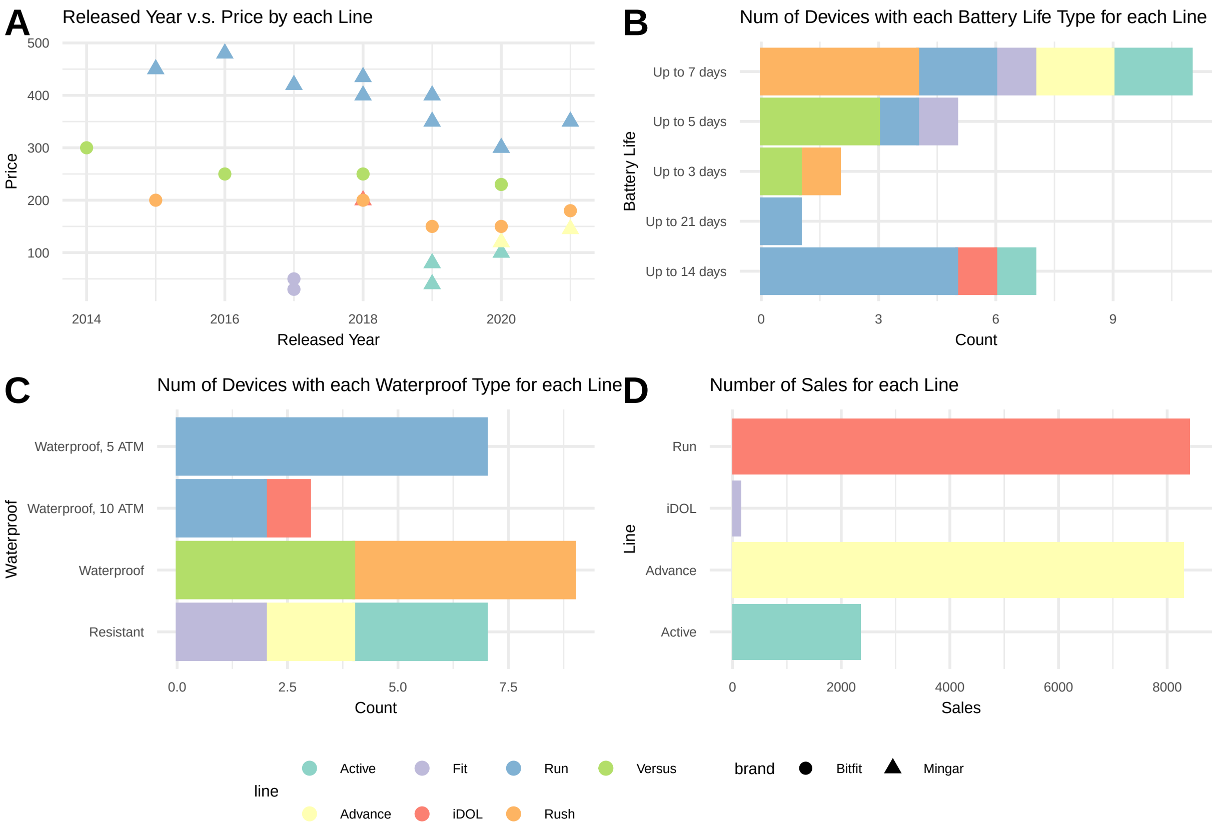

Figure 9: Comparing different functionalities for different lines

Plot 9A above is a scatter plot showing the recommended retail price, released year and the lines. According to the plot, the line *Run* has the highest price in the range of $300 - $500, while line *Advanced* and *Active* has a much lower price from $50 to $150. The average price of all lines in *Mingar* is $284.66. And the average price of all lines in *Bitfit* is $180.90. By the plot A, the released year for the line *Run* is spread through the x-axis, while the line *Advanced* and *Active* only came out in the most recent 3 years.

Plot 9B is a bar chart showing the battery life for each line. Based on the plot, we noticed that line *Run* usually has the longest battery life, since most of the devices of *Run* has a battery life up to 14 days and more. However, the battery life of line *Advanced* and *Active* is shorter, which is about 5-14 days.

Plot 9C is a also a bar chart showing the performance of waterproof. Line *Run* has a better waterproof performance, all devices from *Run* has a waterproof of 5ATM-10ATM, while devices from line *Advanced* and *Active* are only water resistant, which is the lowest level of waterproof.

Plot 9D shows the popularity of each line in *Mingar*. From the bar chart, we can conclude that the line *Run* and *Advanced* are the two most popular line in *Mingar*. And the sales of *Run* is slightly greater than *Advanced.*

Conclusions

Above all, we noticed that the line *Run* and *Advanced* are the most popular lines at *Mingar.* Line *Run* has an overall better performance and a higher price among all other lines. Run is one of the earliest lines the company started producing, thus, line *Run* has 9 devices, which is the line having most products. Line *Active* and *Advanced* are two new lines that came out in the most recent 3 years. There are only 5 devices for the line *Advanced* and *Active.* However, the total sales of line *Advanced* and *Active* is more than that of line *Run.* The result implies that when the devices with sufficient functionality, price is the major factor for number of sales.

Discussion

Findings from research question 1 suggest that buyers of MINGAR's newer affordable lines "Active" and "Advance" tend to come from a lower-income region, therefore meeting the marketing goal of attracting customers outside of the traditionally higher-income base. Further, customers across all four product lines do not have the same median age. Specifically, the "Active" line appeals most to users in two distinctive age groups: mid-20s and 60s. Moreover, Ontario tops the customer base across all lines, followed by Alberta and Quebec in the two new lines. Therefore, we suggest MINGAR to continue expanding its marketing efforts in these provinces.

Results from research question 2 indicate that MINGAR's devices are significantly less accurate at tracking sleep scores of dark skin emoji-modifiers than those of light skin emoji-modifiers, implying that the devices may have great potential to perform more poorly on dark skin individuals. In anticipation of future complaints, we suggest MINGAR prepare contingency plans, such as creating a Frequently Asked Questions page to assure users that this is a known technical challenge MINGAR is actively working to resolve.

Lastly, from research question 3, we observe "Run" and "Advanced" to be the most popular lines among users of MINGAR. However, despite the recent launch of "Active" and "Advance", their combined sales have exceeded the sales of "Run", the traditional high-end line with most features. These results imply that customers tend to favour approachable price points over functionality when choosing a device.

Strengths and limitations

At MANGO, one of our greatest strengths is that we take ethical data practices extremely seriously. In particular, we have a track record of handling identity-based data with sensitivity, integrity, and tact, which is demonstrated in all our past and current collaboration projects and also consistent with MINGAR's values.

In addition, we were meticulous about checking all model assumptions before making a prediction. We also used one of the most considerate models – generalized linear mixed models – to not only take account of the speciality in our response variable and the individual random effect differences, being considerate and careful in the conclusions that we make.

However, there are several limitations in our analysis. The first limitation is that the Kruskal Wallis Test, used in the first research question to identify age distributions among different product lines, does not meet the model assumption of same shape distributions, as the age distribution of the "Active" line resembles a binomial distribution, while the remaining three product lines appear relatively normal. We can potentially attribute this violation to the lack of available data collected on the "Active" line.

Another important caveat for interpreting our study is that emoji-modifiers were used as a proxy for ethnicity due to lack of data. MANGO suggests MINGAR to introduce a personalized cartoon avatar profile feature, similar to Apple's Memoji feature, where users can choose to create realistic avatar profiles that look like themselves. This avatar feature will likely be a more powerful proxy for race and ethnicity than the current skin tone emoji-modifier.

A final limitation is that the postal codes extracted from PCCF are dated (2016). Plus, they represent mailing addresses, not necessarily residential addresses of customers, and therefore may not be accurate when being used as a proxy to infer the income levels of customers.

Consultant information

Consultant profiles

Xinyi Chen. Xinyi is a senior data scientist with MANGO. She specializes in data modeling. Xinyi earned her Bachelor of Science, Specialist in Computer Science from the University of Toronto in 2020.

Rui Miao. Rui is a senior data engineer with MANGO. She specializes in data wrangling. Rui earned her Bachelor of Science, Majoring in Computer Science from the University of Toronto in 2021.

Xinpang Shan. Xinpang is a junior consultant with MANGO. She specializes in reproducible analysis. Xinpang earned her Bachelor of Science, Specialist in Data Science from the University of Toronto in 2021.

Iris Shao. Iris is an associate consultant intern with MANGO. She specializes in statistical communication. Iris earned her Bachelor of Science, Majoring in Statistics and Minoring in Digital Humanities from the University of Toronto in 2022.

Code of ethical conduct

- MANGO strictly adheres to privacy laws or regulations surrounding the collection, storage, and reporting of data and results.
- MANGO maintains the highest standards of ethical statistical practice in satisfying the requirements of the client and disclosing any potential conflict of interest in a timely manner.
- MANGO strives to refrain from procedural bias and any misrepresentation of findings.

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/

Alboukadel Kassambara (2020). ggpubr: "ggplot2" Based Publication Ready Plots. R package version 0.4.0. https://rpkgs.datanovia.com/ggpubr/

C. Sievert. Interactive Web-Based Data Visualization with R, plotly, and shiny. Chapman and Hall/CRC Florida, 2020.

Colvonen, P. J., DeYoung, P. N., Bosompra, N.-O. A., & Owens, R. L. (2020). Limiting racial disparities and bias for wearable devices in Health Science Research. Sleep, 43(10). https://doi.org/10.1093/sleep/zsaa159

D. Kahle and H. Wickham. ggmap: Spatial Visualization with ggplot2. The R Journal, 5(1), 144-161. URL http://journal.r-project.org/archive/2013-1/kahle -wickham.pdf

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.

Device and customer data. MINGAR(2022).

Dmytro Perepolkin (2019). polite: Be Nice on the Web. R package version 0.1.1. https://github.com/dmi3kno/polite

Erich Neuwirth (2022). RColorBrewer: ColorBrewer Palettes. R package version 1.1-3.

Gregory R. Warnes, Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber, Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz and Bill Venables (2020). gplots: Various R Programming Tools for Plotting Data. R package version 3.1.1. https://github.com/talgalili/gplots

H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

Hadley Wickham (2021). rvest: Easily Harvest (Scrape) Web Pages. https://rvest.tidyverse.org/, https://github.com/tidyverse/rvest.

Joe Cheng, Bhaskar Karambelkar and Yihui Xie (2022). leaflet: Create Interactive Web Maps with the JavaScript "Leaflet" Library. R package version 2.1.0. https://rstudio.github.io/leaflet/

Karthik Ram and Hadley Wickham (2018). wesanderson: A Wes Anderson Palette Generator. R package version 0.3.6. https://github.com/karthik/wesanderson

Kirill Müller and Hadley Wickham (2021). tibble: Simple Data Frames. https://tibble.tidyverse.org/, https://github.com/tidyverse/tibble.

Original S code by Richard A. Becker and Allan R. Wilks. R version by Ray Brownrigg. (2018). mapdata: Extra Map Databases. R package version 2.3.0.

Original S code by Richard A. Becker, Allan R. Wilks. R version by Ray Brownrigg. Enhancements by Thomas P Minka and Alex Deckmyn. (2021). maps: Draw Geographical Maps. R package version 3.4.0.

Population density 2016. Census Mapper. (n.d.). Retrieved April 7, 2022, from https://censusmapper.ca/

Postal code conversion file: 2016 census geography. Statistics Canada. (2016). Retrieved April 7, 2022, from: https://mdl.library.utoronto.ca/collections/ numeric-data/census-canada/postal-code-conversion-file/2016

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

STA303/1002 winter 2022 final project. (2022). Fitness tracker info hub. Retrieved April 7, 2022, from https://fitnesstrackerinfohub.netlify.app/

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

We have respected the web scraping policies by providing our contact information to the web scraping agent. We have also respected crawl limit of 12 seconds. We have not used the scraped data in a way that reveals user privacy information. The scraped fitness devices data was only modified by some name changes and was solely used to compare between the different lines for MINGAR.

Accessing Census data on median household income

API access

We used APIs to get the data from the Canadian census website. We signed up for the cancensus API through the website and get the API key for our account, and used the API key to fetch the income data. The income dataset was then joined with the exisiting customer dataset by customer ID to ensure that we only work with information about our customers.

Ethical Professional consideration

Based on the policy on the website, data on CensusMapper can be freely linked or downloaded through the provided interface. And we are allowed to explore census variables ans data for specific geographic. In our research, we respects the policy posted on the website, we used the data to reproduce new dataset, and analyze our research questions based on the data.

Accessing postcode conversion files

We chose to download the census data from August,2021. We want the most recent postcode data from the Census since we want the most up to date information about where people are living. We joined only filtered out the relevant postcode conversion data that matches with customers in our data.

Final advice: KNIT EARLY AND OFTEN!