

Caso práctico: ¿Cómo puede hacer una empresa tecnología para el bienestar para tomar decisiones inteligentes?

Tania Gómez

Análisis de Datos para Mejorar el Marketing de Bellabeat

Contexto y Objetivo

Bellabeat es un fabricante de productos de alta tecnología orientados a la salud de la mujer.

Tarea Empresarial

La tarea empresarial consiste en **mejorar la estrategia de marketing de sus productos inteligentes**. Hay que adaptar la estrategia basándonos en datos sobre tendencias actuales y el uso de dispositivos de bienestar, permitiendo así que Bellabeat ajuste su enfoque para atraer a más clientes y maximizar el impacto de en el mercado.

Consideraciones y Áreas de Mejora

- **Fuentes de Datos:** ¿Es suficiente analizar el uso de dispositivos de otras marcas o podríamos recopilar datos adicionales sobre los entornos de los usuarios de Bellabeat?
- **Segmentos Rentables:** Se puede explorar si existen segmentos específicos de usuarios rentables.
- **Frecuencia de Uso:** ¿Cuáles son los patrones de tiempo y frecuencia en el uso de dispositivos similares que podrían enriquecer la experiencia de usuario de Bellabeat?

Preguntas de análisis

- ¿Qué características son más valoradas por los usuarios de dispositivos de bienestar?
- ¿Qué tipo de hábitos y métricas son las más rastreadas (ej., actividad física, sueño, hidratación)?
- ¿Cuáles son las tendencias actuales de uso de estos dispositivos?

Interesados Clave

- **Equipo de liderazgo:** Urška Sršen (cofundadora y directora creativa) y Sandro Mur (matemático y cofundador).
- **Equipo de marketing y análisis computacional de datos:** Equipo responsable del análisis y de validar los hallazgos para la implementación en la estrategia de marketing.

Datos utilizados

Para el análisis se han usado dos fuentes de datos:

- **Datos de seguimiento de actividad física de Fitbit** (cc0: Dominio público, conjunto de datos disponibles a través de Mobius): Este conjunto de datos de Kaggle contiene el seguimiento de la actividad física personal en treinta usuarios de Fitbit. Treinta usuarios elegibles de Fitbit prestaron su consentimiento para el envío de datos personales de seguimiento que incluyen rendimiento de la actividad física en minutos, ritmo cardíaco y monitoreo del sueño. Incluye información sobre la actividad diaria, pasos y ritmo cardíaco que se puede usar para explorar los hábitos de los usuarios.
- **Fitness Consumer Survey Data** (cc0: Dominio público): Este conjunto de datos de Kaggle contiene un seguimiento de actividad física personal de usuarios de dispositivos inteligentes. Consiste en las respuestas a una encuesta de veintiuna preguntas por parte de treinta usuarios. Incluye datos demográficos, uso del dispositivo inteligente y cómo este ha influido en su actividad física.

Los datos se encuentran en archivos CSV y han sido descargados de la plataforma del caso práctico. Se han añadido otros conjuntos de datos relacionados, como el Fitness Consumer Survey Data de Kaggle, para complementar el análisis.

La mayoría de los archivos originales están en formato ancho, lo cual permite ver un solo registro por fila para cada usuario y sus métricas diarias u horarias. Otros archivos también están en formato largo y se recoge en períodos de cinco segundos de diferencia, lo cual hace que sean complicados de manejar en una hoja de cálculo.

Los datos están bien organizados, aunque a fecha actual de 2024, se podrían considerar obsoletos. No son específicos de Bellabeat, de manera que se pueden tomar con cuidado y complementarlos con datos externos para tener una visión más completa. Al no ser datos específicos de Bellabeat, no refleja el uso de sus usuarios, pero son útiles para observar tendencias generales.

Los datos son anónimos y públicos, por lo que no representan un riesgo de privacidad. La seguridad se mantiene al trabajar en un entorno protegido, sin datos de identificación personal.

Para verificar la integridad de los datos, estos se limpiaron de duplicados, se validaron formatos y se verificaron columnas problemáticas como pueden ser caracteres extraños.

Estos datos permiten entender las tendencias de uso de dispositivos de bienestar y compararlas. Ayudan a analizar patrones en el uso de dispositivos e identificar características de uso que podrían aplicarse a la estrategia de marketing de Bellabeat.

Procesamiento de los datos

- 25/10/2024 Descarga, procesamiento y análisis de los datos.
- 28/10/2024 Análisis de datos.
- 29/10/2024 Visualizaciones y hallazgos clave.

Pasos de procesamiento

25/10/2024

- Eliminar duplicados y verificar registros nulos. - Convertir formatos de fecha y hora. - Estandarizar nombres de columnas.

Análisis de datos

25/10/2024

- Cálculo de la media de todos los minutos activos.
- Cálculo del total de los pasos diarios y los minutos activos.
- Cálculo del promedio diario de actividad.
- Cálculo de minutos activos por pasos.

28/10/2024

- Cálculo de usuarios por edad.
- Cálculo de usuarios por género.
- Cálculo de cuántas veces se usa el dispositivo inteligente.
- Cálculo de cuántas veces se registra actividad con el dispositivo inteligente.
- Carga de datos en RStudio con revisión de la estructura.
- Análisis de las tendencias y relaciones.

29/10/2024

- Creación de visualizaciones.
- Resumen del análisis.
- Hallazgos clave.

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2     3.5.1      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.1
## v purrr       1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(hrbrthemes)
library(tinytex)
library(viridis)
```

```
## Cargando paquete requerido: viridisLite
```

```
Caso_Bellabeat <- read_csv("Caso Bellabeat.csv")
```

```
## New names:
## Rows: 940 Columns: 20
## -- Column specification
## ----- Delimiter: "," chr
## (11): ActivityDate, TotalDistance, TrackerDistance, LoggedActivitiesDist... dbl
## (8): Id, TotalSteps, VeryActiveMinutes, FairlyActiveMinutes, LightlyAct... lgl
## (1): ...16
## i Use 'spec()' to retrieve the full column specification for this data. i
## Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## * ' -> '...16'

Fitness_Data <- read_csv("████████████████████/Caso Bellabeat - Fitness Consumer Data.csv")
```

```
## Rows: 30 Columns: 22
## -- Column specification -----
## Delimiter: ","
## chr (22): Timestamp, Age, Gender, Degree, Occupation, Exerciseweek, Timeusin...
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

```
View(Caso_Bellabeat)
View(Fitness_Data)
```

```
## Revisión de la estructura general y filas de muestra
str(Caso_Bellabeat)
```

```
## spc_tbl_ [940 x 20] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ Id : num [1:940] 1.5e+09 1.5e+09 1.5e+09 1.5e+09 1.5e+09 ...
## $ ActivityDate : chr [1:940] "4/12/2016" "4/13/2016" "4/14/2016" "4/15/2016" ...
## $ TotalSteps : num [1:940] 13.16 10.73 10.46 9.76 12.67 ...
## $ TotalDistance : chr [1:940] "45.420" "696.999.979.019.165" "673.999.977.111.816" "628.0
## $ TrackerDistance : chr [1:940] "45.420" "696.999.979.019.165" "673.999.977.111.816" "628.0
## $ LoggedActivitiesDistance: chr [1:940] "0,00" "0,00" "0,00" "0,00" ...
## $ VeryActiveDistance : chr [1:940] "187.999.999.523.163" "157.000.005.245.209" "244.000.005.72
## $ ModeratelyActiveDistance: chr [1:940] "550.000.011.920.929" "689.999.997.615.814" "400.000.005.96
## $ LightActiveDistance : chr [1:940] "605.999.994.277.954" "471.000.003.814.697" "391.000.008.58
## $ SedentaryActiveDistance : chr [1:940] "0,00" "0,00" "0,00" "0,00" ...
## $ VeryActiveMinutes : num [1:940] 25 21 30 29 36 38 42 50 28 19 ...
## $ FairlyActiveMinutes : num [1:940] 13 19 11 34 10 20 16 31 12 8 ...
## $ LightlyActiveMinutes : num [1:940] 328 217 181 209 221 164 233 264 205 211 ...
## $ SedentaryMinutes : num [1:940] 728 776 1.22 726 773 ...
## $ Calories : num [1:940] 1.99 1.8 1.78 1.75 1.86 ...
## $ ...16 : logi [1:940] NA NA NA NA NA NA ...
## $ TotalActiveMinutes : num [1:940] 122 86 74 91 89 74 97 115 82 79 ...
## $ TotalDailyActivity : chr [1:940] "13.284,00" "10.820,67" "10.534,00" "9.852,67" ...
## $ AverageDailyActivity : chr [1:940] "84.000.000.367.828" "180.799.997.649.235" "158.866.664.967
## $ RatioActiveMinutes : chr [1:940] "0,01" "0,01" "0,01" "0,01" ...
## - attr(*, "spec")=
## .. cols(
## .. Id = col_double(),
## .. ActivityDate = col_character(),
```

```
## .. TotalSteps = col_double(),
## .. TotalDistance = col_character(),
## .. TrackerDistance = col_character(),
## .. LoggedActivitiesDistance = col_character(),
## .. VeryActiveDistance = col_character(),
## .. ModeratelyActiveDistance = col_character(),
## .. LightActiveDistance = col_character(),
## .. SedentaryActiveDistance = col_character(),
## .. VeryActiveMinutes = col_double(),
## .. FairlyActiveMinutes = col_double(),
## .. LightlyActiveMinutes = col_double(),
## .. SedentaryMinutes = col_double(),
## .. Calories = col_double(),
## .. ...16 = col_logical(),
## .. TotalActiveMinutes = col_double(),
## .. TotalDailyActivity = col_character(),
## .. AverageDailyActivity = col_character(),
## .. RatioActiveMinutes = col_character()
## .. )
## - attr(*, "problems")=<externalptr>
```

```
head(Caso_Bellabeat)
```

```
## # A tibble: 6 x 20
##       Id ActivityDate TotalSteps TotalDistance TrackerDistance
##       <dbl> <chr>         <dbl> <chr>         <chr>
## 1 1503960366 4/12/2016          13.2 45.420         45.420
## 2 1503960366 4/13/2016          10.7 696.999.979.019.165 696.999.979.019.165
## 3 1503960366 4/14/2016          10.5 673.999.977.111.816 673.999.977.111.816
## 4 1503960366 4/15/2016           9.76 628.000.020.980.835 628.000.020.980.835
## 5 1503960366 4/16/2016          12.7 815.999.984.741.211 815.999.984.741.211
## 6 1503960366 4/17/2016           9.70 648.000.001.907.349 648.000.001.907.349
## # i 15 more variables: LoggedActivitiesDistance <chr>,
## #   VeryActiveDistance <chr>, ModeratelyActiveDistance <chr>,
## #   LightActiveDistance <chr>, SedentaryActiveDistance <chr>,
## #   VeryActiveMinutes <dbl>, FairlyActiveMinutes <dbl>,
## #   LightlyActiveMinutes <dbl>, SedentaryMinutes <dbl>, Calories <dbl>,
## #   ...16 <lgl>, TotalActiveMinutes <dbl>, TotalDailyActivity <chr>,
## #   AverageDailyActivity <chr>, RatioActiveMinutes <chr>
```

```
summary(Caso_Bellabeat)
```

```
##       Id           ActivityDate      TotalSteps      TotalDistance
## Min.   :1.504e+09 Length:940      Min.    : 0.000 Length:940
## 1st Qu.:2.320e+09 Class :character 1st Qu.: 4.489 Class :character
## Median :4.445e+09 Mode  :character Median : 7.912 Mode  :character
## Mean   :4.855e+09          Mean   : 19.622
## 3rd Qu.:6.962e+09          3rd Qu.: 11.202
## Max.   :8.878e+09          Max.    :980.000
## TrackerDistance LoggedActivitiesDistance VeryActiveDistance
## Length:940      Length:940          Length:940
## Class :character Class :character    Class :character
## Mode  :character Mode  :character    Mode  :character
```

```
##
##
##
## ModeratelyActiveDistance LightActiveDistance SedentaryActiveDistance
## Length:940          Length:940          Length:940
## Class :character    Class :character    Class :character
## Mode  :character    Mode  :character    Mode  :character
##
##
##
## VeryActiveMinutes FairlyActiveMinutes LightlyActiveMinutes SedentaryMinutes
## Min.   : 0.00      Min.   : 0.00      Min.   : 0.0      Min.   : 0.000
## 1st Qu.: 0.00      1st Qu.: 0.00      1st Qu.:127.0     1st Qu.: 1.192
## Median : 4.00      Median : 6.00      Median :199.0     Median : 1.440
## Mean   : 21.16     Mean   : 13.56     Mean   :192.8     Mean   :316.180
## 3rd Qu.: 32.00     3rd Qu.: 19.00     3rd Qu.:264.0     3rd Qu.:708.250
## Max.   :210.00     Max.   :143.00     Max.   :518.0     Max.   :993.000
##      Calories      ...16      TotalActiveMinutes TotalDailyActivity
## Min.   : 0.000     Mode:logical    Min.   : 0.00     Length:940
## 1st Qu.: 1.841     NA's:940        1st Qu.: 49.00     Class :character
## Median : 2.159                                Median : 82.00     Mode  :character
## Mean   : 5.729                                Mean   : 75.85
## 3rd Qu.: 2.804                                3rd Qu.:106.00
## Max.   :928.000                                Max.   :184.00
## AverageDailyActivity RatioActiveMinutes
## Length:940          Length:940
## Class :character    Class :character
## Mode  :character    Mode  :character
##
##
##
```

```
str(Fitness_Data)
```

```
## spc_tbl_ [30 x 22] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ Timestamp      : chr [1:30] "2023/03/30 9:43:19 PM GMT+5:30" "2023/03/31 5:07:46 PM GMT+5:30" ...
## $ Age            : chr [1:30] "18-24" "Under 18" "18-24" "25-34" ...
## $ Gender         : chr [1:30] "Male" "Male" "Female" "Female" ...
## $ Degree         : chr [1:30] "Some college or associate degree" "Bachelor's degree" "Bachelor" ...
## $ Occupation     : chr [1:30] "Student" "Student" "Student" "Employed part-time" ...
## $ Exerciseweek   : chr [1:30] "5 or more times a week" "5 or more times a week" "Less than once" ...
## $ Timeusing      : chr [1:30] "Less than 6 months" "Less than 6 months" "Less than 6 months" ...
## $ Frequentuse    : chr [1:30] "Daily" "3-4 times a week" "Rarely" "3-4 times a week" ...
## $ Trackfitness   : chr [1:30] "Every day" "Once a week" "Rarely" "Every day" ...
## $ Impactfitnessroutine: chr [1:30] "Positively impacted my fitness routine" "Positively impacted my" ...
## $ ImpactMotivation : chr [1:30] "Strongly agree" "Neutral" "Strongly agree" "Strongly agree" ...
## $ ImpactEnjoying  : chr [1:30] "Strongly agree" "Agree" "Agree" "Agree" ...
## $ EngagingWearable : chr [1:30] "Very engaged" "Somewhat engaged" "Not very engaged" "Somewhat engaged" ...
## $ ConectionCommunity : chr [1:30] "Agree" "Agree" "Strongly agree" "Strongly agree" ...
## $ FitnessGoal     : chr [1:30] "No impact on achieving my goals" "Helped me achieve my goals some" ...
## $ ImpactHealth    : chr [1:30] "No impact on my overall health" "Improved my overall health some" ...
## $ SleepPatterns   : chr [1:30] "Agree" "Agree" "Strongly agree" "Agree" ...
## $ ImprovedWellbeing : chr [1:30] "Neutral" "Strongly agree" "Strongly agree" "Strongly agree" ...
## $ IncluceExercise : chr [1:30] "Strongly agree" "Agree" "Agree" "Agree" ...
```

```
## $ IncluncePurchase : chr [1:30] "Neutral" "Neutral" "Agree" "Disagree" ...
## $ Influencegym : chr [1:30] "Agree" "Neutral" "Strongly agree" "Neutral" ...
## $ InfluenceDiet : chr [1:30] "Agree" "Neutral" "Agree" "Agree" ...
## - attr(*, "spec")=
## .. cols(
## .. Timestamp = col_character(),
## .. Age = col_character(),
## .. Gender = col_character(),
## .. Degree = col_character(),
## .. Occupation = col_character(),
## .. Exerciseweek = col_character(),
## .. Timeusing = col_character(),
## .. Frequentuse = col_character(),
## .. Trackfitness = col_character(),
## .. Impactfitnessroutine = col_character(),
## .. ImpactMotivation = col_character(),
## .. ImpactEnjoying = col_character(),
## .. EngagingWearable = col_character(),
## .. ConectionCommunity = col_character(),
## .. FitnessGoal = col_character(),
## .. ImpactHealth = col_character(),
## .. SleepPatterns = col_character(),
## .. ImprovedWellbeing = col_character(),
## .. InclunceExercise = col_character(),
## .. IncluncePurchase = col_character(),
## .. Influencegym = col_character(),
## .. InfluenceDiet = col_character()
## .. )
## - attr(*, "problems")=<externalptr>
```

```
head(Fitness_Data)
```

```
## # A tibble: 6 x 22
## Timestamp Age Gender Degree Occupation Exerciseweek Timeusing Frequentuse
## <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
## 1 2023/03/30 ~ 18-24 Male Some ~ Student 5 or more t~ Less tha~ Daily
## 2 2023/03/31 ~ Unde~ Male Bache~ Student 5 or more t~ Less tha~ 3-4 times ~
## 3 2023/03/31 ~ 18-24 Female Bache~ Student Less than o~ Less tha~ Rarely
## 4 2023/03/31 ~ 25-34 Female Some ~ Employed ~ 3-4 times a~ 6-12 mon~ 3-4 times ~
## 5 2023/03/31 ~ 18-24 Male Bache~ Student 1-2 times a~ Less tha~ Daily
## 6 2023/03/31 ~ 18-24 Female Maste~ Employed ~ 5 or more t~ 1-2 years Daily
## # i 14 more variables: Trackfitness <chr>, Impactfitnessroutine <chr>,
## # ImpactMotivation <chr>, ImpactEnjoying <chr>, EngagingWearable <chr>,
## # ConectionCommunity <chr>, FitnessGoal <chr>, ImpactHealth <chr>,
## # SleepPatterns <chr>, ImprovedWellbeing <chr>, InclunceExercise <chr>,
## # IncluncePurchase <chr>, Influencegym <chr>, InfluenceDiet <chr>
```

```
summary(Fitness_Data)
```

```
## Timestamp Age Gender Degree
## Length:30 Length:30 Length:30 Length:30
## Class :character Class :character Class :character Class :character
## Mode :character Mode :character Mode :character Mode :character
```

```
## Occupation      Exerciseweek      Timeusing      Frequentuse
## Length:30      Length:30      Length:30      Length:30
## Class :character Class :character Class :character Class :character
## Mode :character Mode :character Mode :character Mode :character
## Trackfitness    Impactfitnessroutine ImpactMotivation ImpactEnjoying
## Length:30      Length:30      Length:30      Length:30
## Class :character Class :character Class :character Class :character
## Mode :character Mode :character Mode :character Mode :character
## EngagingWearable ConectionCommunity FitnessGoal      ImpactHealth
## Length:30      Length:30      Length:30      Length:30
## Class :character Class :character Class :character Class :character
## Mode :character Mode :character Mode :character Mode :character
## SleepPatterns   ImprovedWellbeing InclunceExercise IncluncePurchase
## Length:30      Length:30      Length:30      Length:30
## Class :character Class :character Class :character Class :character
## Mode :character Mode :character Mode :character Mode :character
## Influencegym     InfluenceDiet
## Length:30      Length:30
## Class :character Class :character
## Mode :character Mode :character
```

```
## Acortar nombres de algunos archivos
```

```
cb <- Caso_Bellabeat
fd <- Fitness_Data
```

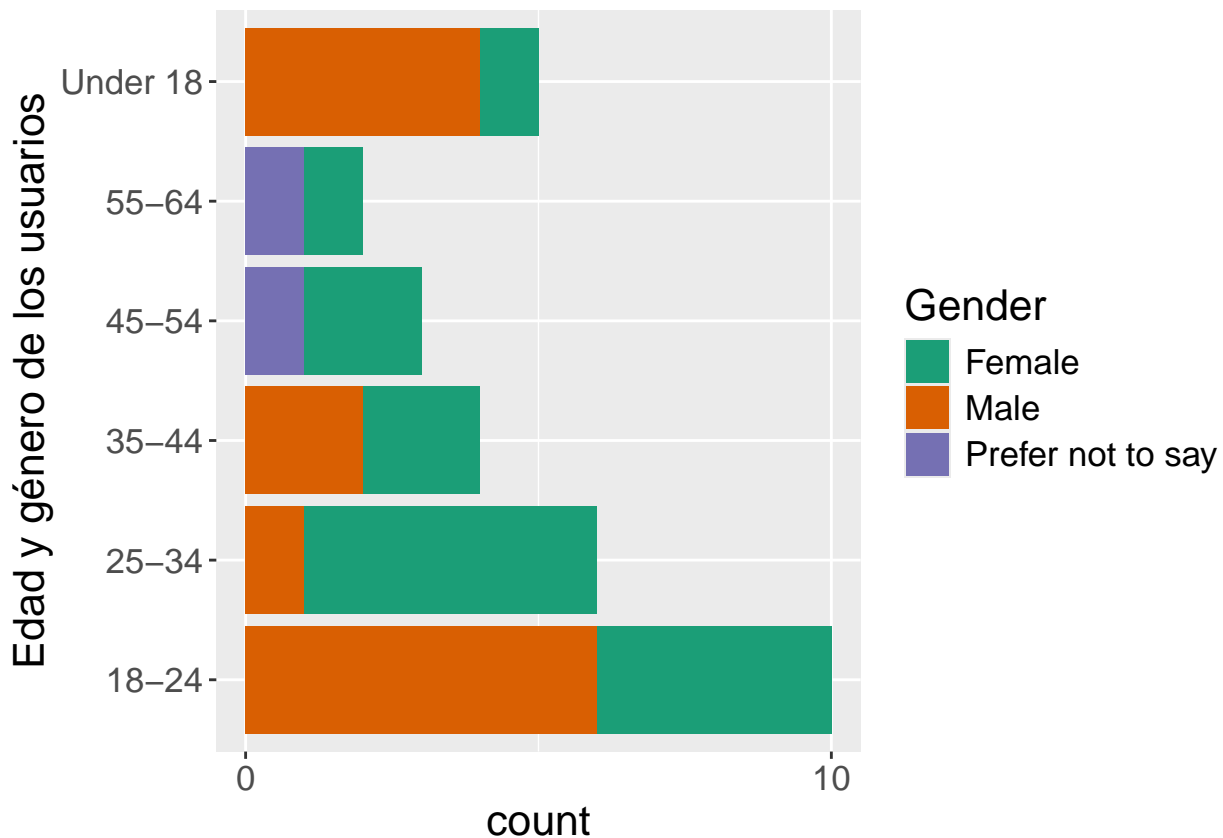
Análisis de variables específicas:

```
sample(fd)
```

```
## # A tibble: 30 x 22
##   ImprovedWellbeing InfluenceDiet InclunceExercise ImpactMotivation
##   <chr>             <chr>             <chr>             <chr>
## 1 Neutral          Agree             Strongly agree    Strongly agree
## 2 Strongly agree   Neutral          Agree             Neutral
## 3 Strongly agree   Agree            Agree             Strongly agree
## 4 Strongly agree   Agree            Agree             Strongly agree
## 5 Strongly agree   Strongly agree   Agree             Neutral
## 6 Strongly agree   Strongly agree   Strongly agree    Strongly agree
## 7 Strongly agree   Neutral          Neutral           Strongly agree
## 8 Agree            Agree            Agree             Agree
## 9 Agree            Agree            Neutral           Strongly agree
## 10 Neutral         Strongly agree   Strongly agree    Strongly agree
## # i 20 more rows
## # i 18 more variables: Frequentuse <chr>, Timeusing <chr>, Timestamp <chr>,
## #   Influencegym <chr>, Age <chr>, Exerciseweek <chr>, IncluncePurchase <chr>,
## #   ImpactHealth <chr>, Occupation <chr>, SleepPatterns <chr>, Degree <chr>,
## #   ConectionCommunity <chr>, Gender <chr>, ImpactEnjoying <chr>,
## #   Impactfitnessroutine <chr>, EngagingWearable <chr>, FitnessGoal <chr>,
## #   Trackfitness <chr>
```



```
ggplot(fd, aes(x = Age, fill = Gender)) +
  geom_bar() +
  scale_x_discrete() +
  xlab("Edad y género de los usuarios") +
  coord_flip() +
  theme(text = element_text(size=16)) +
  scale_y_continuous(breaks=seq(-40,40,10),labels=abs(seq(-40,40,10))) +
  scale_fill_brewer(palette="Dark2")
```

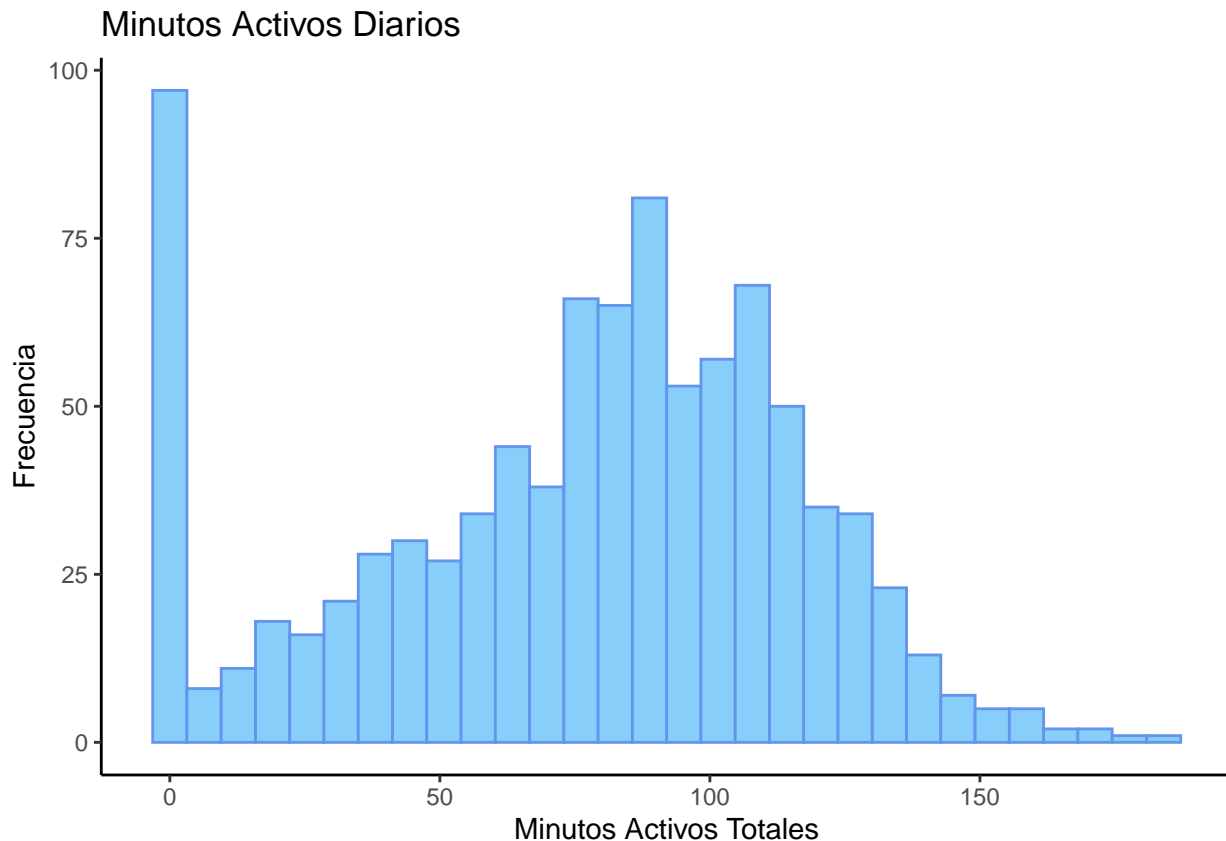


```
sample(fd)
```

```
## # A tibble: 30 x 22
##   ConectionCommunity Impactfitnessroutine FitnessGoal Occupation Exerciseweek
##   <chr>               <chr>                <chr>      <chr>      <chr>
## 1 Agree               Positively impacted m~ No impact ~ Student    5 or more t~
## 2 Agree               Positively impacted m~ Helped me ~ Student    5 or more t~
## 3 Strongly agree      Positively impacted m~ Helped me ~ Student    Less than o~
## 4 Strongly agree      I don't know          Helped me ~ Employed ~ 3-4 times a~
## 5 Agree               Positively impacted m~ Helped me ~ Student    1-2 times a~
## 6 Strongly agree      Positively impacted m~ Helped me ~ Employed ~ 5 or more t~
## 7 Strongly agree      Positively impacted m~ Helped me ~ Student    Less than o~
## 8 Strongly agree      No impact on my fitne~ Helped me ~ Student    Less than o~
## 9 Neutral             Positively impacted m~ Helped me ~ Employed ~ 1-2 times a~
## 10 Agree              Positively impacted m~ Helped me ~ Employed ~ Less than o~
## # i 20 more rows
## # i 17 more variables: Frequentuse <chr>, ImpactMotivation <chr>,
```

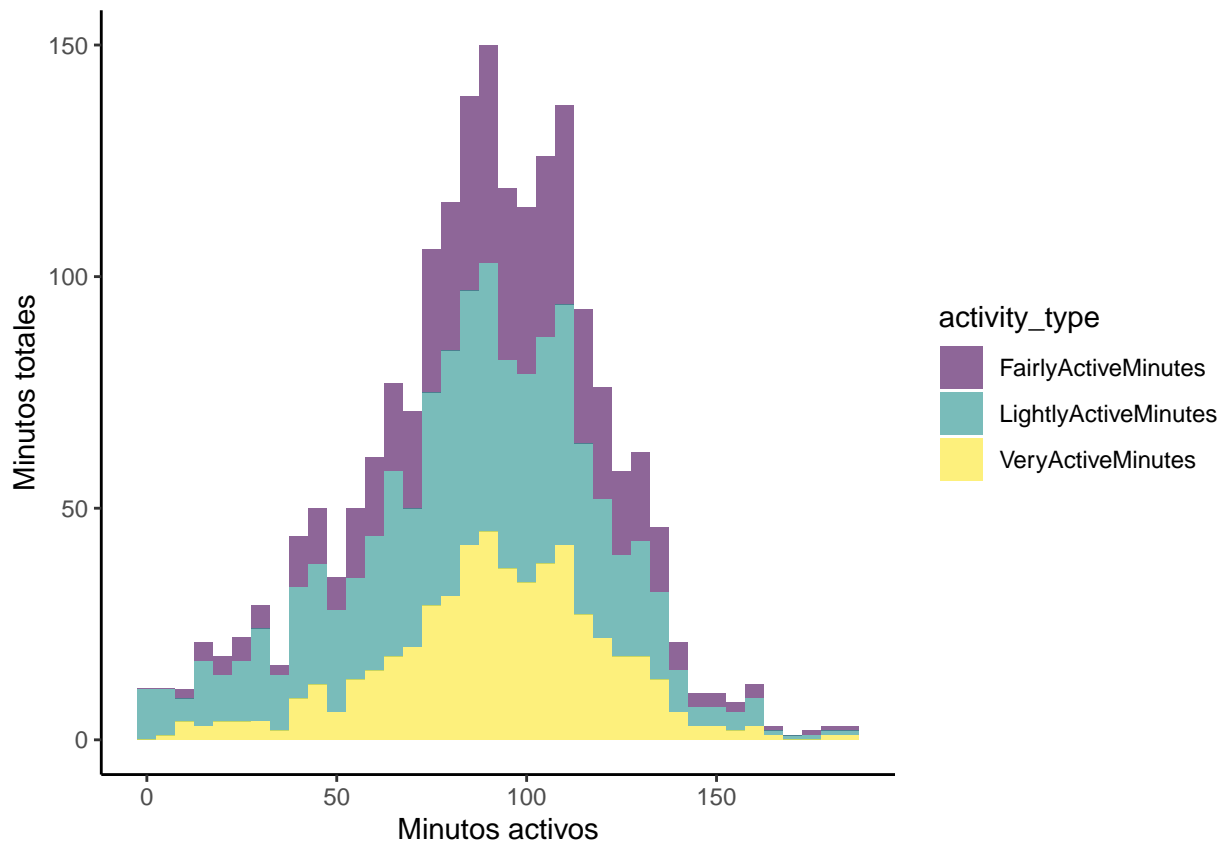
```
## # SleepPatterns <chr>, Influencegym <chr>, Timeusing <chr>,
## # ImpactHealth <chr>, ImpactEnjoying <chr>, Timestamp <chr>, Age <chr>,
## # InlucenceExercise <chr>, Trackfitness <chr>, Gender <chr>, Degree <chr>,
## # ImprovedWellbeing <chr>, InfluenceDiet <chr>, EngagingWearable <chr>,
## # InlucencePurchase <chr>
```

```
ggplot(cb, aes(x = TotalActiveMinutes)) +
  geom_histogram(bins = 30, fill = "LightSkyBlue", color = "CornflowerBlue") +
  theme_classic() +
  labs(title = "Minutos Activos Diarios", x = "Minutos Activos Totales", y = "Frecuencia")
```



```
cb_long <- cb %>%
  pivot_longer(
    cols = c(VeryActiveMinutes, FairlyActiveMinutes, LightlyActiveMinutes),
    names_to = "activity_type",
    values_to = "minutes"
  )
cb_long_filtered <- cb_long %>%
  filter(minutes > 0)

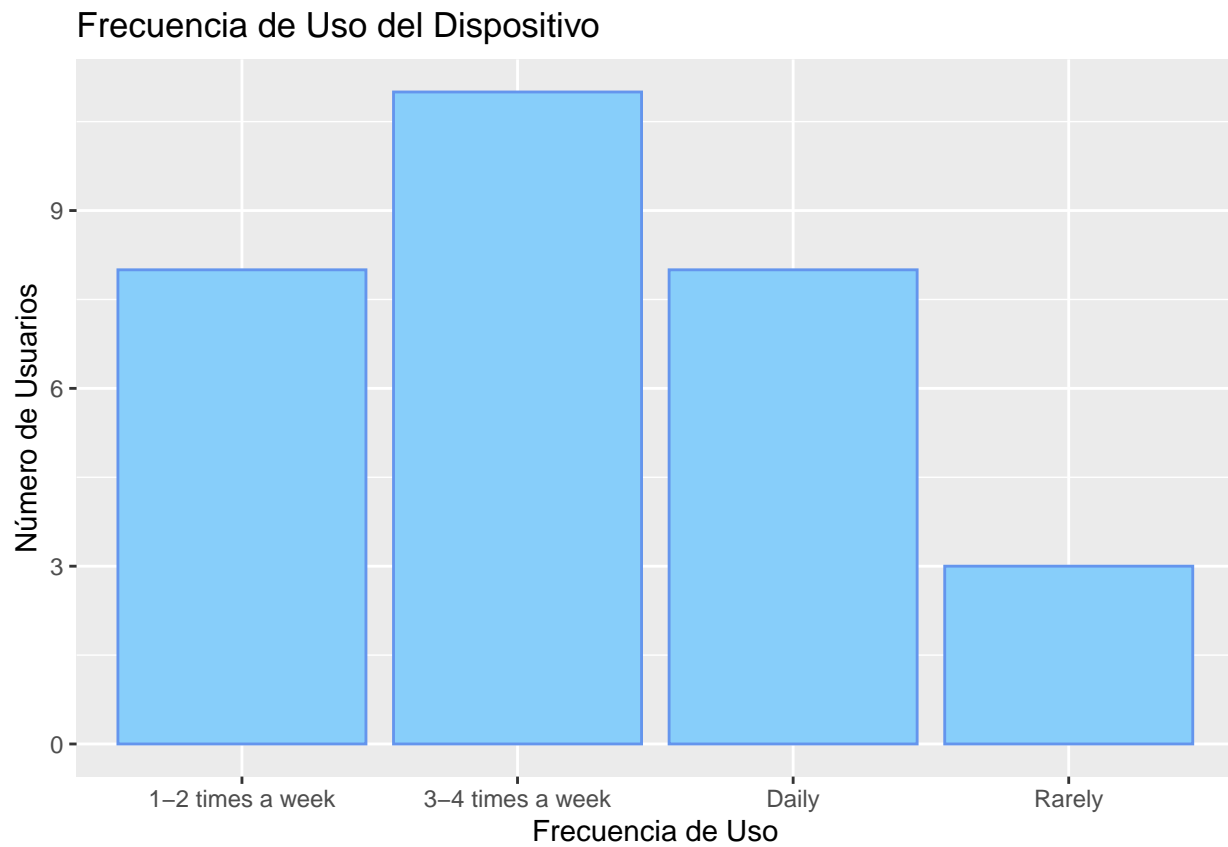
ggplot(cb_long_filtered, aes(x = TotalActiveMinutes, fill = activity_type)) +
  geom_histogram(alpha = 0.6, binwidth = 5) +
  scale_fill_viridis(discrete = TRUE) +
  theme_classic() +
  xlab("Minutos activos") +
  ylab("Minutos totales")
```



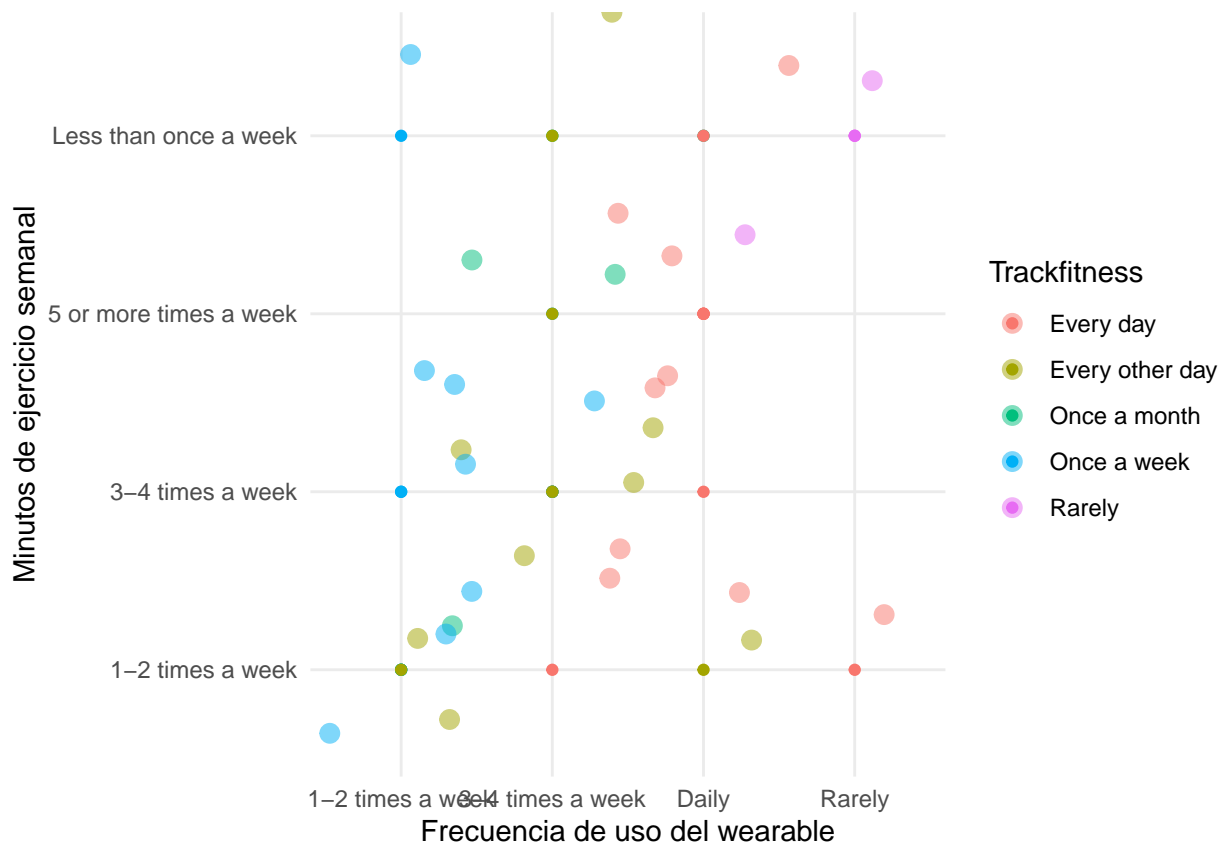
```
sample(fd)
```

```
## # A tibble: 30 x 22
##   ConnectionCommunity ImpactEnjoying Age Occupation Trackfitness Timestamp
##   <chr> <chr> <chr> <chr> <chr> <chr>
## 1 Agree Strongly agree 18-24 Student Every day 2023/03/~
## 2 Agree Agree Under 18 Student Once a week 2023/03/~
## 3 Strongly agree Agree 18-24 Student Rarely 2023/03/~
## 4 Strongly agree Agree 25-34 Employed p~ Every day 2023/03/~
## 5 Agree Agree 18-24 Student Every other~ 2023/03/~
## 6 Strongly agree Strongly agree 18-24 Employed f~ Every day 2023/03/~
## 7 Strongly agree Strongly agree 18-24 Student Once a week 2023/03/~
## 8 Strongly agree Agree 18-24 Student Once a month 2023/04/~
## 9 Neutral Agree 18-24 Employed p~ Every other~ 2023/04/~
## 10 Agree Strongly agree 35-44 Employed f~ Every day 2023/04/~
## # i 20 more rows
## # i 16 more variables: InfluencePurchase <chr>, Influencegym <chr>,
## # ImprovedWellbeing <chr>, ImpactMotivation <chr>, Degree <chr>,
## # Impactfitnessroutine <chr>, ImpactHealth <chr>, Exerciseweek <chr>,
## # FitnessGoal <chr>, Gender <chr>, SleepPatterns <chr>, Timeusing <chr>,
## # InfluenceExercise <chr>, EngagingWearable <chr>, Frequentuse <chr>,
## # InfluenceDiet <chr>
```

```
ggplot(fd, aes(x = Frequentuse)) +
  geom_bar(fill = "LightSkyBlue", color = "CornflowerBlue") +
  labs(title = "Frecuencia de Uso del Dispositivo", x = "Frecuencia de Uso", y = "Número de Usuarios")
```



```
ggplot(fd, aes(x = Frequentuse, y = Exerciseweek, color = Trackfitness)) +  
  geom_point() +  
  geom_jitter(width = 0.8, height = 0.8, alpha = 0.5, size = 3) +  
  labs(x = "Frecuencia de uso del wearable",  
       y = "Minutos de ejercicio semanal",  
       size = "Número de registros") +  
  theme_minimal()
```



Resumen del análisis

El análisis de los datos sobre los dispositivos inteligentes de bienestar de Bellabeat revela patrones clave que pueden guiar la estrategia de marketing. Se destacan varias tendencias en cuanto a características valoradas, frecuencia de uso y preferencias demográficas, de forma que Bellabeat enfoque sus campañas para atraer segmentos específicos del mercado y promover el uso sostenido de sus productos.

Hallazgos clave

Perfil demográfico

La tecnología de Bellabeat está orientada a la salud de la mujer y estas representan la mayoría de los usuarios, aún si hay una presencia significativa de personas que prefieren no especificar su género, especialmente en grupos de edad media. La edad dominante de los usuarios están en el rango de 18-34 años. Hay una oportunidad para campañas específicas que promuevan los beneficios para la salud en personas mayores de 45 años, dado que el uso de estos dispositivos es menor en ese segmento.

Patrones de uso y frecuencia

La mayoría de los usuarios utiliza el dispositivo de 3 a 4 veces por semana, lo cuál indica que este es el punto óptimo de compromiso con el dispositivo. También existe un grupo fiel de usuarios diarios, mientras que los usuarios ocasionales son minoría. La mayoría de estos registra entre 50 y 120 minutos de actividad diaria, con un pico de 80-100 minutos, lo que nos sugiere que el público objetivo valora el monitoreo de actividad física moderada, sin ser extremo.

Correlación entre actividad y frecuencia de uso

Los usuarios que utilizan el dispositivo a diario muestran patrones de actividad consistentes y elevados, lo cual indica una relación positiva entre el uso regular del dispositivo y el compromiso con la actividad física. En cambio, aquellos usuarios ocasionales se inclinan a patrones de actividad más irregulares. Podrían beneficiarse de características que formenten la regularidad en el uso.

Estrategia de marketing

Los usuarios de dispositivos de bienestar nos han demostrado que el monitoreo de la actividad es una de las funciones más usadas y valoradas, de manera que Bellabeat debería enfatizarlo a la hora de comunicar su producto. La flexibilidad en la frecuencia de uso y la posibilidad de mantener rutinas consistentes son características atractivas para el público objetivo.

El enfoque principal debería ser el mercado joven (18-34 años), mientras que se pueden desarrollar estrategias específicas para atraer a usuarios mayores de 45 años. Otra oportunidad de marketing es la de enfatizar la facilidad de uso del dispositivo para mantener rutinas de actividad física, en específico en el uso moderado y constante, y así aumentar su adopción.

Se podría considerar la incorporación de funciones que motiven el uso regular y adapten el dispositivo a diferentes niveles de actividad, con características específicas para cada grupo demográfico. Teniendo en cuenta que Bellabeat se enfoca en la salud de la mujer y son la mayoría de usuarios, se podría tener en cuenta factores físicos que influyen a la hora de hacer actividad física.

Recomendaciones

- Campañas dirigidas al segmento de 18-34 años en las que se destaquen los beneficios para la salud y el bienestar.
- Crear contenido educativo para los usuarios menos activos y fomentar así el compromiso con el dispositivo.
- Enfatizar la flexibilidad de uso en los mensajes publicitarios para adaptarse a la vida cotidiana de los usuarios.
- Añadir características que incentive el uso regular, como recordatorios y recompensas.
- Crear programas de actividad física adaptados a diferentes niveles y objetivos.
- Implementar funciones específicas orientadas a mujeres, alineándose con el enfoque de Bellabeat en el bienestar de la mujer.
- Desarrollar una comunidad de usuarios para compartir experiencias y progresos.
- Implementar programas de iniciación para nuevos usuarios.

Conclusión

Bellabeat tiene una clara oportunidad en el mercado de dispositivos de bienestar, especialmente entre mujeres jóvenes que buscan mantener una rutina de actividad física consistente. La clave del éxito estará en desarrollar productos que faciliten y motiven el uso regular, con características adaptadas a diferentes niveles de actividad y compromiso, y en comunicarlos de forma eficaz a la audiencia principal.