Data Analysis for Sales Group Inc. March, 21, 2024

Analyst: Anthony Grieco

Introduction

Sales Group Inc. is an organization that specializes in selling electronic devices and accessories all across the United States. Sales Group Inc. maintains a database containing a considerable amount of information about every customer transaction that they have ever processed. Accordingly, the organization now intends to use that same information to better gauge who the bulk of their customers are and how to better market their products to them.

Sales Group Inc. has requested that the CS345 Analytics Team, or A-Team, conduct an analysis of their data in order to answer the following questions:

1. How much revenue was generated each month in sales across all products? Which month has the highest revenue and which month has the least revenue?
2. Our marketing team breaks age into the following groups: 12-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, and 75-84. Based on these groups, what does the age demographic look like in terms of total items sold per age group?
3. Our “USB-C Charging Cable” is a good seller according to our sales people. What does the total amount (count) of this product sales look like across all the states in which we sell? Where do we sell the most of this product?
4. In the top selling state for the “USB-C Charging Cable”, what age groups buy the greatest number of this product? Should we market more to specific age groups to maximize our marketing expenditures?
5. Across all of our products, which age groups buy the most items from us?
6. We need to focus are marketing for “Apple Airpods Headphones” and “Wired Headphones”. Currently we market them to all age groups. Is there a relationship between age and these two products? Is gender a factor? How could we focus our marketing to maximize our marketing expenditures?
7. What other interesting things can you tell us about our sales data?

Description of the Data

The data provided to the A-Team includes transactional data from a subset of states that Sales Group Inc. has done business in over the course of the past year. The data set, provided in the form of access to a Sales Group Inc. SQL database, includes the following data:

* customer\_id: The code that uniquely identifies each customer who made a purchase from Sales Group Inc.
* street: The street of the address that the customer wants their order shipped to
* city: The city of the address that the customer wants their order shipped to
* st: The state of the address that the customer wants their order shipped to
* zip: The zip code of the address that the customer wants their order shipped to
* gender: The gender of the customer who made the purchase
* age: The age of the customer who made the purchase
* order\_id: The code that uniquely identifies every order processed by Sales Group Inc.
* order\_date: The date in which the customer’s order was processed
* order\_total: The total amount due for the given customer’s total purchase
* product\_id: The code that uniquely identifies every product customers could order
* quantity: The number of a particular product being ordered by the customer
* price\_ea: The unit price of an individual product
* product\_desc: The name of the product

No data transformations were required. Similarly, no data integration was required because all of the necessary attributes to complete the analysis were present in the original data set that Sales Group Inc. provided.

Detailed Analysis

The data was analyzed using both R version 4.3.2[[1]](#footnote-1) and MySQL version 10.3.28-MariaDB[[2]](#footnote-2). The Tidyverse[[3]](#footnote-3) library was used along with the MariaDB[[4]](#footnote-4) library. The analysis environment used was R-Studio version 2023.12.0+369 “Ocean Storm” Release[[5]](#footnote-5).

In answer to “How much revenue was generated each month in sales across all products? Which month has the highest revenue and which month has the least revenue?”, an analysis was conducted to filter the dataset so that it would reveal the total sales revenue for each individual month over the course of the last year. Based on this data, the goal was then to determine which months had the most and least respective revenue. Results were rounded to the nearest penny. The month of December had the highest sales for Sales Group Inc. with $4,608,220 and January had the fewest sales with $1,821,413 (see Appendix A).

In answer to “Our marketing team breaks age into the following groups: 12-17, 18-24, 25-34, 35-44, 45-54, 55-64, 65-74, and 75-84. Based on these groups, what does the age demographic look like in terms of total items sold per age group?”, an analysis was conducted to select the total number of items sold to each of the above age groups as defined by Sales Group Inc. (see Appendix B).

In answer to “Our ‘USB-C Charging Cable’ is a good seller according to our sales people. What does the total amount (count) of this product sales look like across all the states in which we sell? Where do we sell the most of this product?”, an analysis was conducted to determine how many total units were sold and which individual states sold the most USB-C Charging Cables over the course of the last year. There was a total of 21,855 USB-C Charging Cables sold across all eight states represented in the provided dataset. The state of California had 8803 units sold to it, representing the greatest number of USB-C Charging Cables sold to any state in the given dataset (See Appendix C).

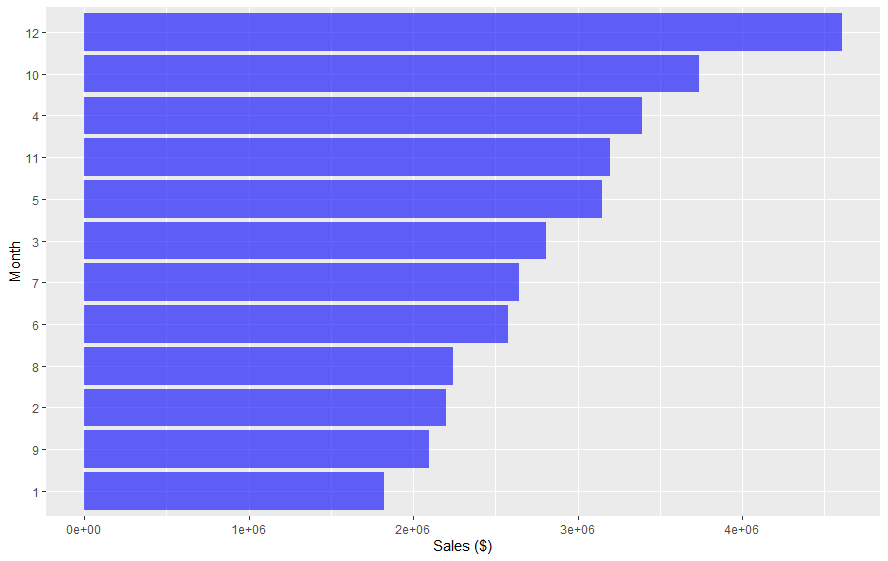
In answer to “In the top selling state for the “USB-C Charging Cable”, what age groups buy the greatest number of this product? Should we market more to specific age groups to maximize our marketing expenditures?”, an analysis was conducted to determine which age groups (as defined above) within the state of California (see Appendix C) purchased the greatest number of USB-C Charging Cables. The 25-34 age group (1,632 units) and 18-24 age group (1,448 units) purchased the greatest number of USB-C Charging Cables in California (see Appendix D). As such, marketing USB-C Charging Cables more to these two age groups specifically would likely have the greatest chance of maximizing profit for Sales Group Inc..

In answer to “Across all of our products, which age groups buy the most items from us?”, an analysis was conducted to select the age group specified above that has purchased the most products from Sales Group Inc.. The 24-35 age group purchased the greatest number of products from Sales Group Inc. with 30,571 units sold (see Appendix B).

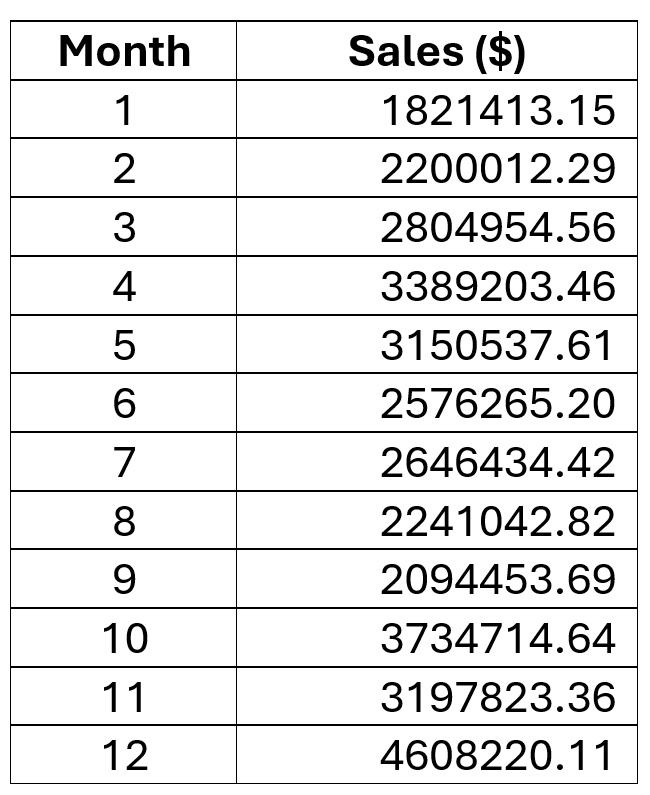
In answer to “We need to focus our marketing for ‘Apple Airpods Headphones’ and ‘Wired Headphones’. Currently we market them to all age groups. Is there a relationship between age and these two products? Is gender a factor? How could we focus our marketing to maximize our marketing expenditures?”, an analysis was conducted to determine whether either age or gender played a role in both “Apple Airpods Headphones” and “Wired Headphones” sales. Age was a factor as individuals in both the 18-24 age group (8398 units) and the 12-17 age group (7,159 units) were responsible for purchasing a combined total of 15,557 units whereas the other five age groups combined purchased 18,815 units of the same two products (see Appendix E). With the 12-17 and 18-24 age groups accounting for roughly 45.26% of the total sales, age certainly played a factor in getting both ‘Apple Airpods Headphones’ and ‘Wired Headphones’ off the shelves. The same can be said for gender, where males purchased 23,025 units, the equivalent of 66.99% of all sales, compared to the 11,347 units purchased by females (see Appendix F). In order to maximize marketing expenditures, Sales Group Inc. should focus their marketing efforts more on males between the ages of 12-24 in order to increase the likelihood of making a profit in the future.

In answer to “What other interesting things can you tell us about our sales data?”, an analysis was conducted to see whether there was a relationship between gender and the total number of products that were sold by Sales Group Inc.. Over the course of the past year, males (101,774 units) bought more products than females (83,865 units). Additionally a total of 185,639 products were sold by Sales Group Inc. over the course of the last year (see Appendix G).Due to that fact that males accounted for 54.82% of all sales made, Sales Group Inc. should focus their marketing efforts more towards the male demographic in order to maximize sales in the future.

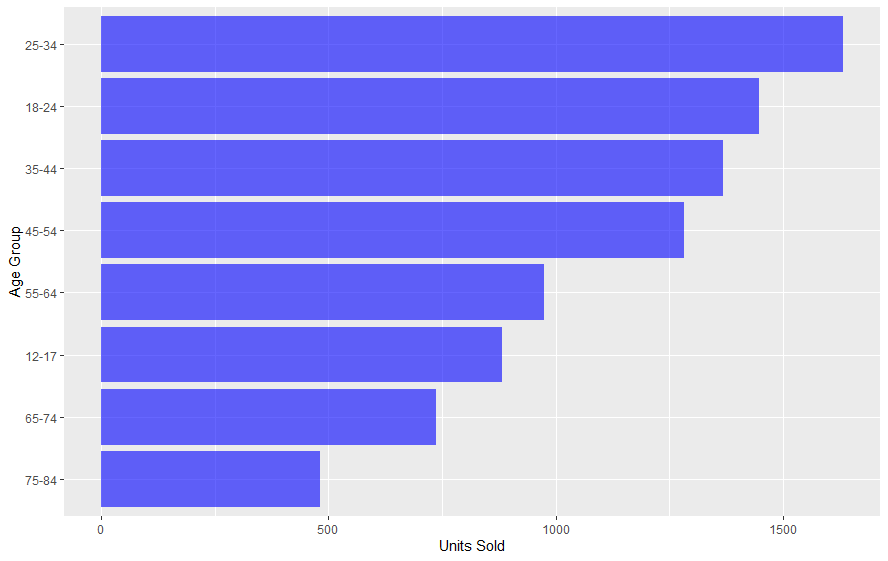
Appendix A: Total Sales by Sales Group Inc. Per Month



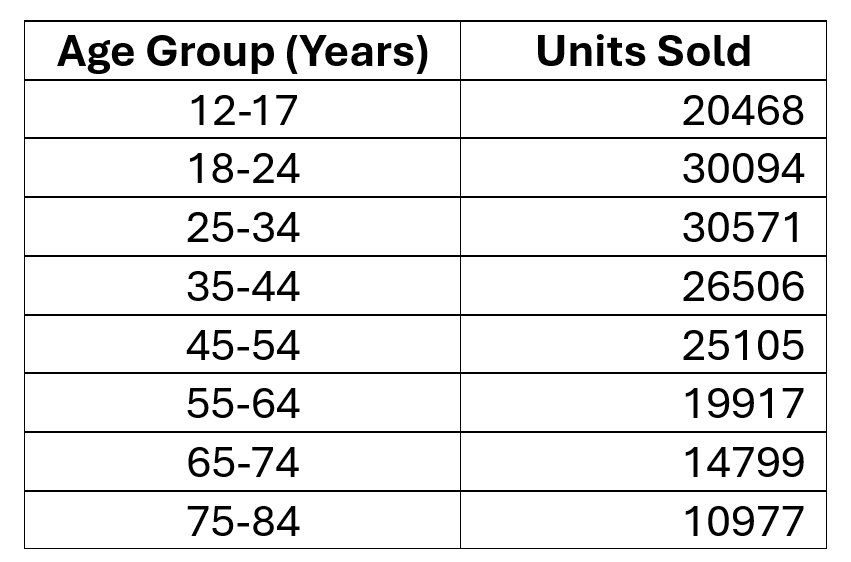
Tabular Data:



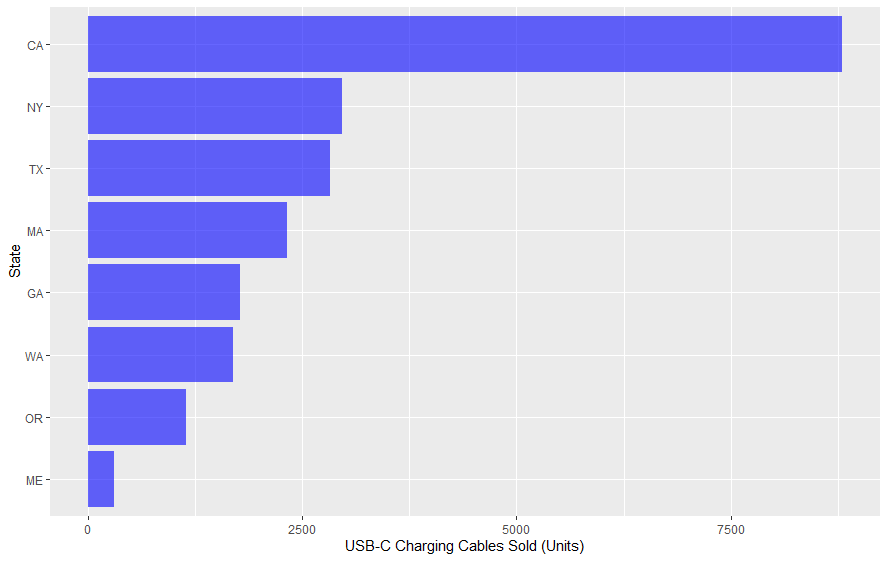
Appendix B: Total Number of Products Sold Per Age Group



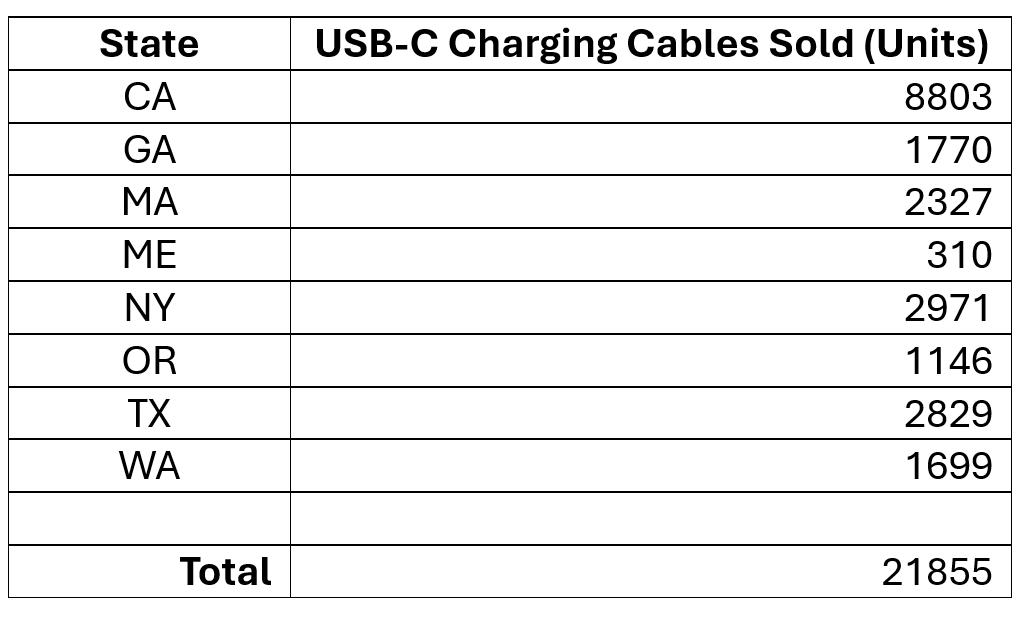
Tabular Data:



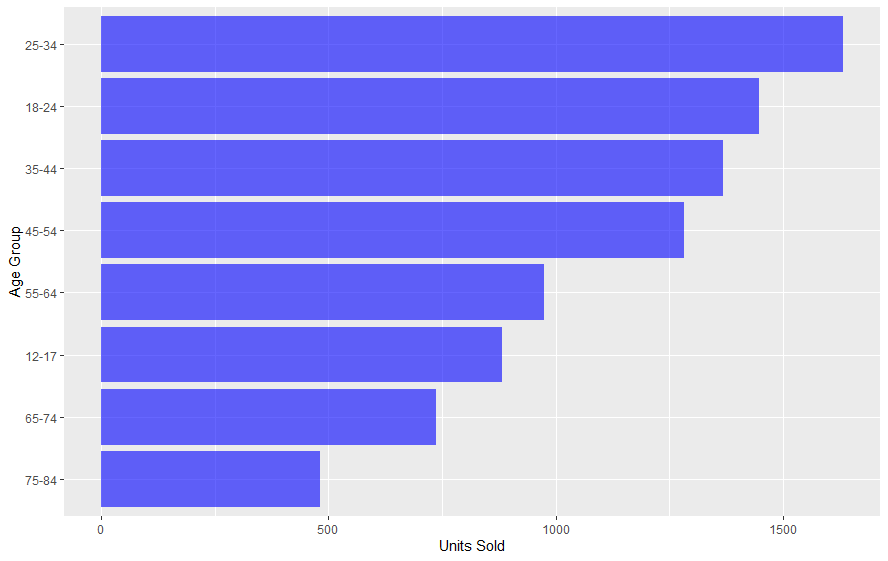
Appendix C: Total USB-C Charging Cables Sold Per State



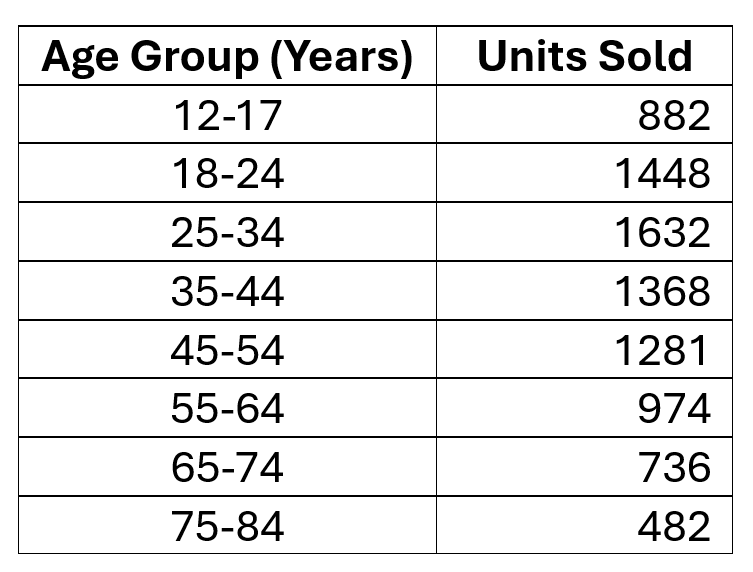
Tabular Data:



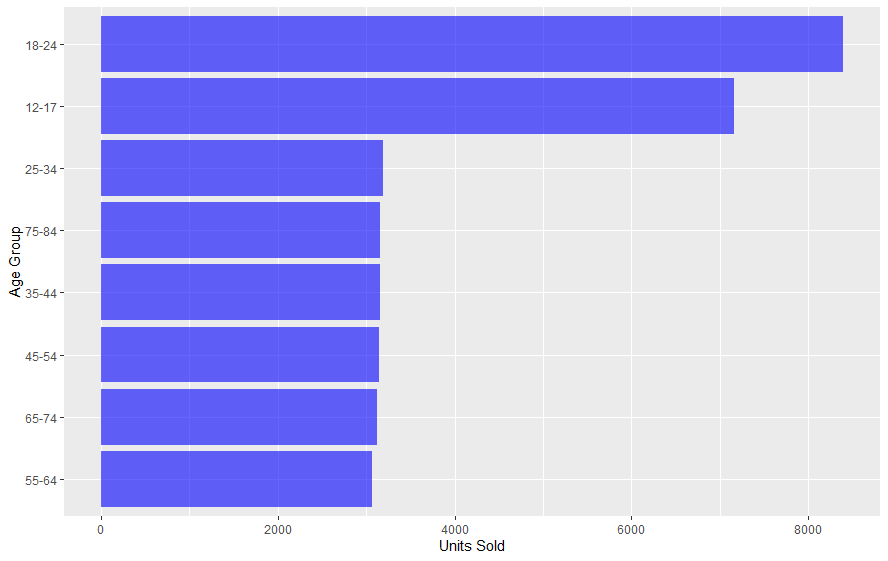
Appendix D: Total USB-C Charging Cables Sold Per Age Group in California



Tabular Data:



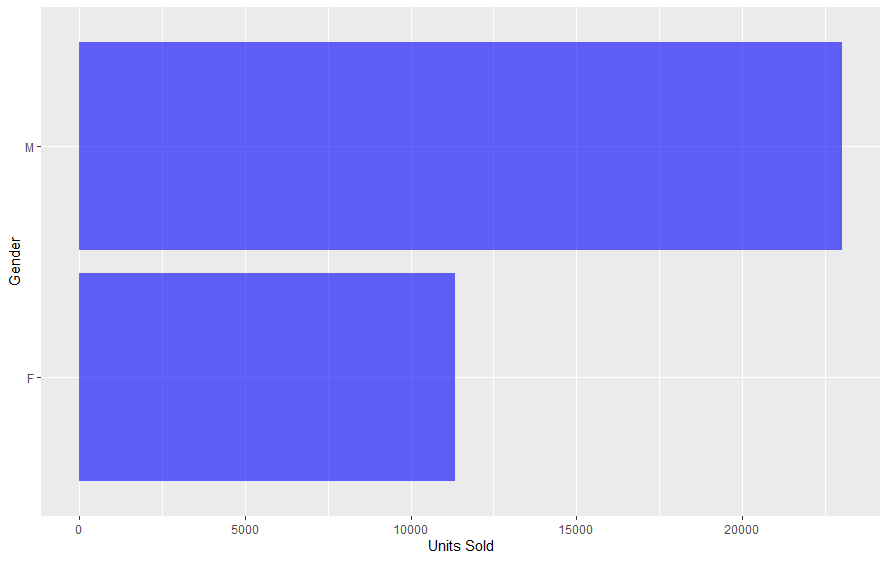
Appendix E: Total “Apple Airpods Headphones” and “Wired Headphones” Sold By Age Group



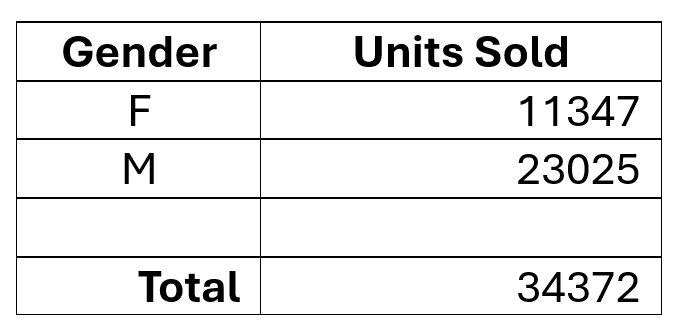
Tabular Data:



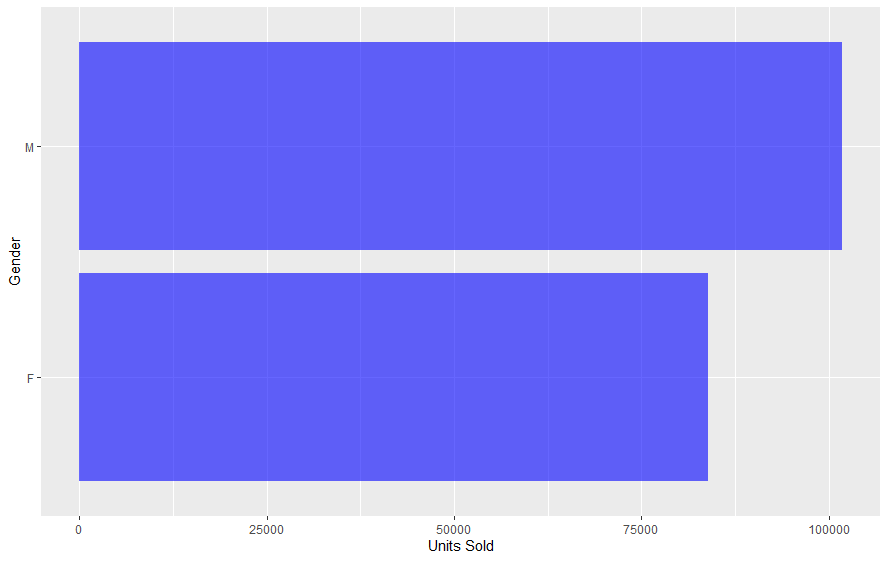
Appendix F: Total “Apple Airpods Headphones” and “Wired Headphones” Sold By Gender



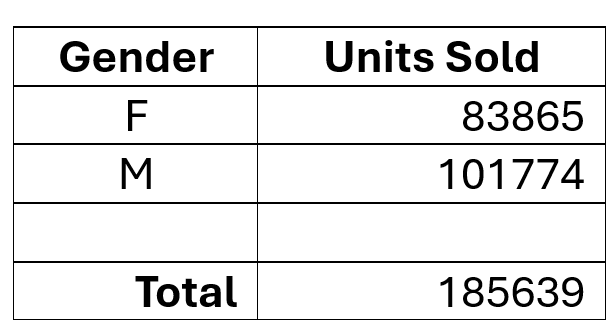
Tabular Data:



Appendix G: Total Products Purchased Over the Last Year By Gender



Tabular Data:



1. R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical

   Computing, Vienna, Austria. https://www.R-project.org/. [↑](#footnote-ref-1)
2. *MySQL*, Oracle, 2024, www.mysql.com/. [↑](#footnote-ref-2)
3. Wickham et al., (2019). Welcome to the tidyverse. Journal of Open Source Software, 4(43), 1686,

   https://doi.org/10.21105/joss.01686. [↑](#footnote-ref-3)
4. Müller, Kirill, et al. *RMariaDB: Database Interface and MariaDB Driver*, 26 Oct. 2023, cran.r-project.org/web/packages/RMariaDB/RMariaDB.pdf. [↑](#footnote-ref-4)
5. RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA, http://www.rstudio.com/. [↑](#footnote-ref-5)