

# Data Translation Assignment

Amrapali Samanta

## Table of contents

<b>Introduction</b>	<b>1</b>
Target audience: . . . . .	2
Goal: . . . . .	2
<b>Data</b>	<b>2</b>
Data Dictionary: . . . . .	2
<b>Summary of Findings:</b>	<b>3</b>
Story: . . . . .	3
<b>Visualizations</b>	<b>3</b>
Visualization 1: Total number of Headphones Sales in Urban Areas . . . . .	3
Visualization 2: Revenue Comparison of Headphone Products . . . . .	5
Visualization 3: Revenue Trend for Apple AirPods by Month . . . . .	6
Visualization 4: Revenue Distribution by State for all Headphones . . . . .	7
Visualization 5: Revenue Distribution by State and City for Apple AirPods . . . . .	8
Visualization 6: Median Household Income by City . . . . .	9
Visualization 7: Sales Distribution by Zip Code for Apple AirPods . . . . .	10
<b>Conclusion</b>	<b>10</b>

## Introduction

This report presents an analysis of headphone sales data from Amazon's tech products division. The objective is to provide insights and recommendations to increase revenues.

The analysis focuses on three types of headphones: **Apple AirPods, Bose SoundSport Headphones, and Wired Headphones** .

**Target audience:**

The target audience for this report is Amazon Technical Product Managers. These individuals are responsible for overseeing the development, launch, and performance of various technical products offered by Amazon. They are interested in gaining insights into the sales and revenue trends of different headphone products, particularly in urban markets, to inform strategic decision-making and future product development efforts.

**Goal:**

The goal of this analysis is to understand the sales and revenue trends of different headphones in the year 2019. We aim to analyze consumer behavior and preferences in urban markets to inform our strategies for the year 2020.

**Note:**

For this analysis, we are only considering a subset of urban ZIP codes from cities such as Dallas, Boston, Los Angeles, San Francisco, Seattle, Atlanta, New York City, Portland, and Austin.

**Data****Data Dictionary:**

We have used two datasets of Sales and ZIP. The description of variables is as below:

**Sales:**

- **Product**, the product that has been ordered
- **Quantity**, how many of the product was ordered (note this is a string, you'll want to fix that!)
- **PriceEach**, the price of each item (note this is a string, you'll want to fix that!)
- **DateTime** and **Date**, when the order was placed. **DateTime** includes both day and time-of-day when the order was placed, while **Date** is just the date
- **ZIP**, the ZIP code where the order was sent to
- **State** and **City**, the city and state where the order was sent to

**Zip:**

- `ZIP`, which is a ZIP code we can use to join this data set with the `sales` data
- `TotalPopulation`, which is the population in that ZIP code
- `MedianHHIncome`, which is the median annual household income in that ZIP. Household income calculates the total income from everyone in a given household, and then finds the median household (Income statistics use 2020 ACS instead of 2018)
- `PCIncome`, which is the annual per-capita (i.e. per-person) income in that ZIP. Per-capita income sums up all the income earned by everyone in the ZIP, and then divides it by the number of people in that ZIP (which may include a lot of non-earners, or children) (Income statistics use 2020 ACS instead of 2018)
- `MedianAge`, the median age of people in the ZIP code
- `Race_*` variables, the number of people of each broad-category race in that ZIP code. Note that races are not mutually exclusive. Someone who is, for example, both White and Asian will be counted once as White and once as Asian
- `Ethnicity_Hispanic`, which is the number of people who are Hispanic in the ZIP code. Ethnicity can overlap with any race, so someone who is, for example, both Hispanic and Black will be counted once as Hispanic and once as Black
- `Citizens`, which is the number of US citizens living in the ZIP code

## Summary of Findings:

### Story:

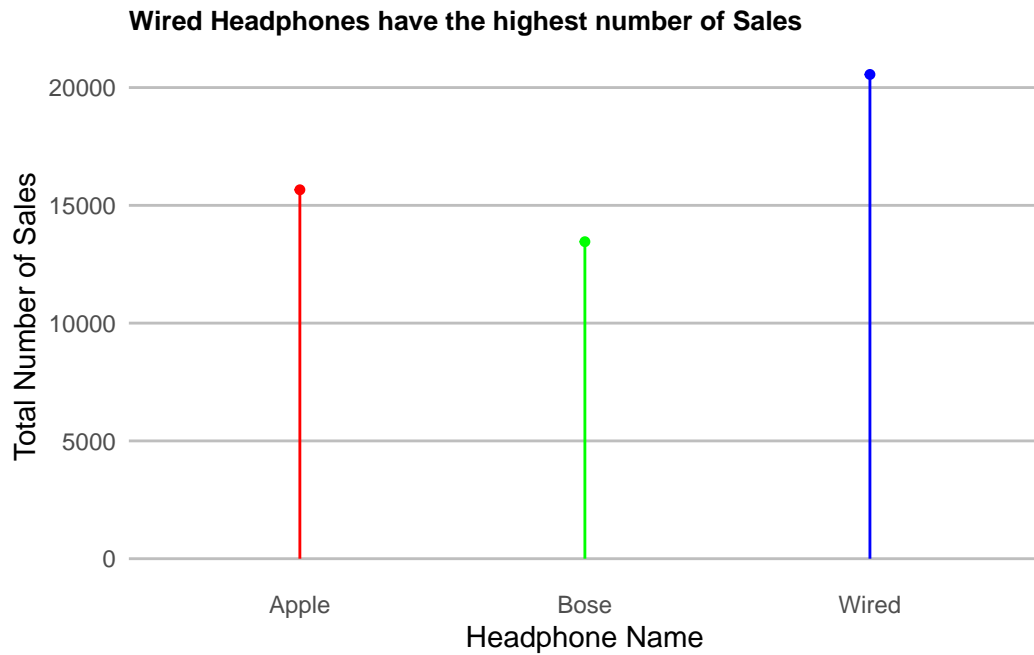
Continue to focus on the sales of ‘Apple AirPods Headphones’ in the city of San Francisco during December month.

For a detailed analysis and visual representation of these findings, please refer to the following sections.

## Visualizations

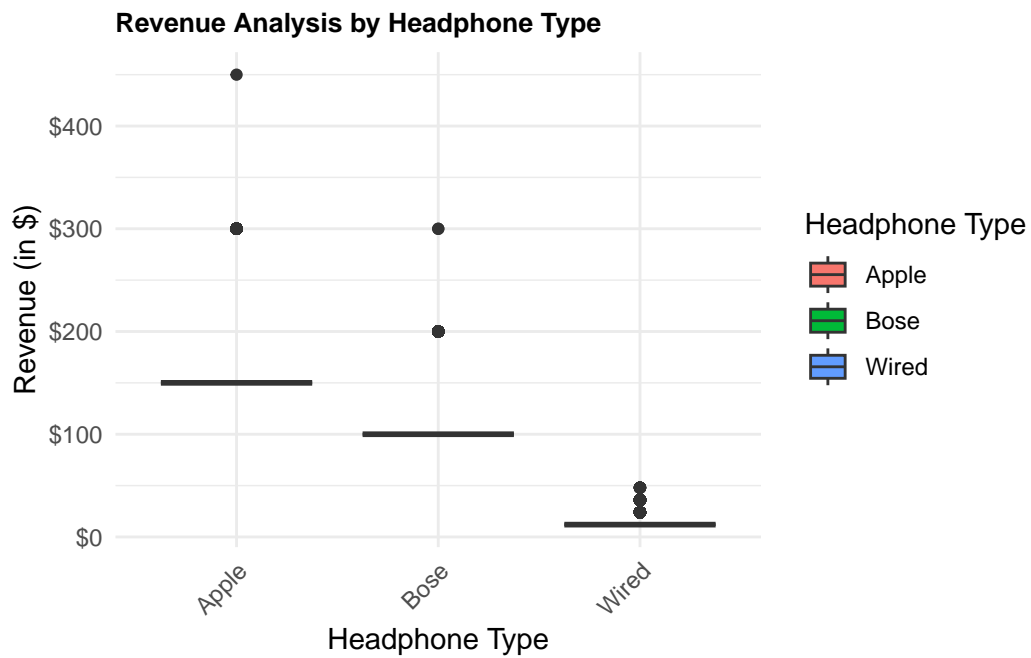
### Visualization 1: Total number of Headphones Sales in Urban Areas

Warning: Using the ``size`` aesthetic with `geom_segment` was deprecated in ggplot2 3.4.0.  
 i Please use the ``linewidth`` aesthetic instead.



Sales of Wired headphones is the highest. So we should also check if Wired headphones also generates highest revenue.

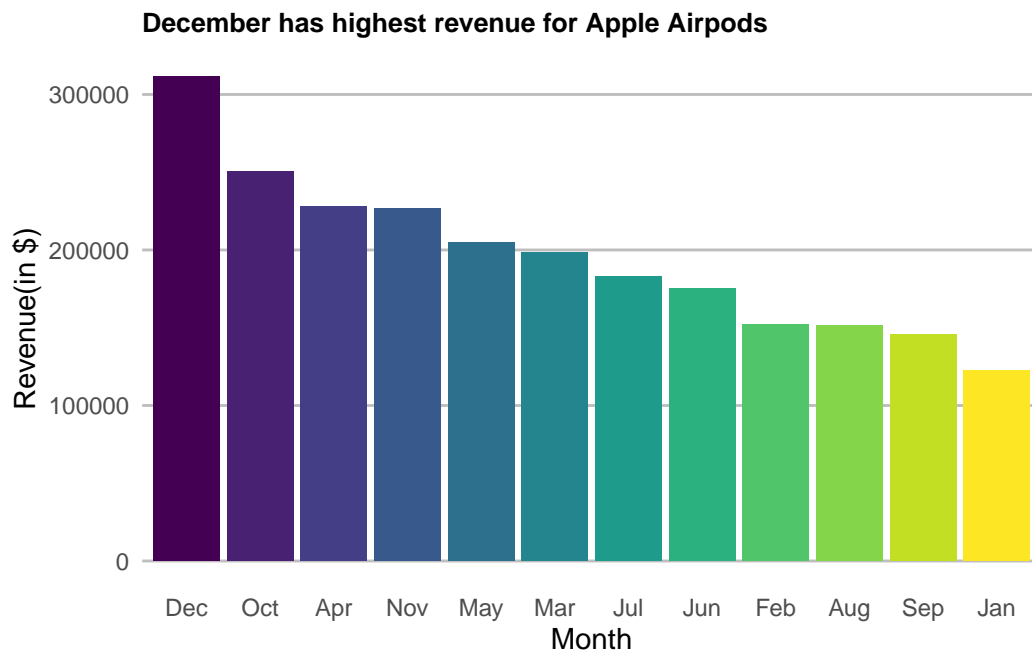
## Visualization 2: Revenue Comparison of Headphone Products



Though Wired Headphones are sold more, Apple AirPods are generating more revenue overall.

Because Apple AirPods gives us most revenue now we will look for insights focussing on AirPods so that we can keep on increasing revenue.

### Visualization 3: Revenue Trend for Apple AirPods by Month



The highest revenue is generated in the month of December.

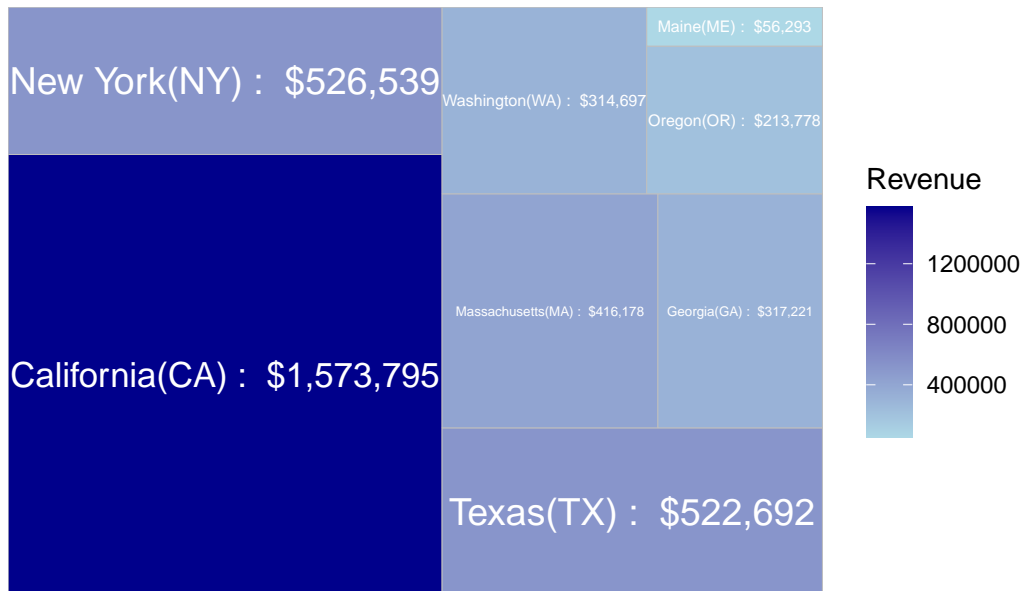
So now we know that we should focus on the sales of AirPods in December for greater revenues.

Let's now see which location generates highest revenue for AirPods.

Note: We only have data for 8 states.

#### Visualization 4: Revenue Distribution by State for all Headphones

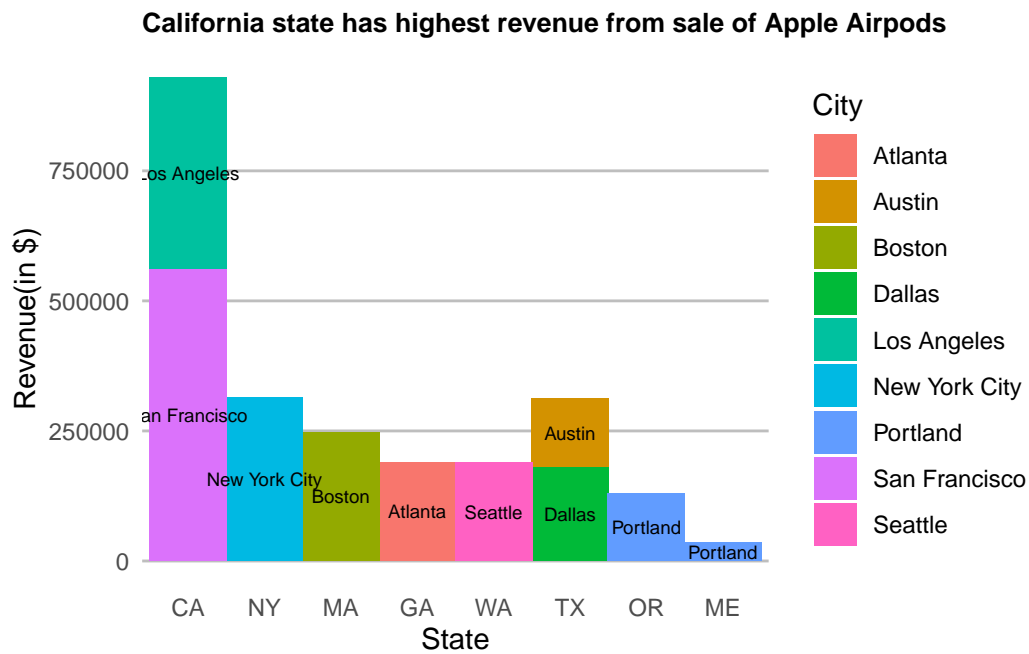
**California state has highest revenue from Headphones**



We understand that state of California(CA) has highest revenue overall for all types of headphones.

Let's see now if it is true for Apple AirPods as well?

## Visualization 5: Revenue Distribution by State and City for Apple AirPods



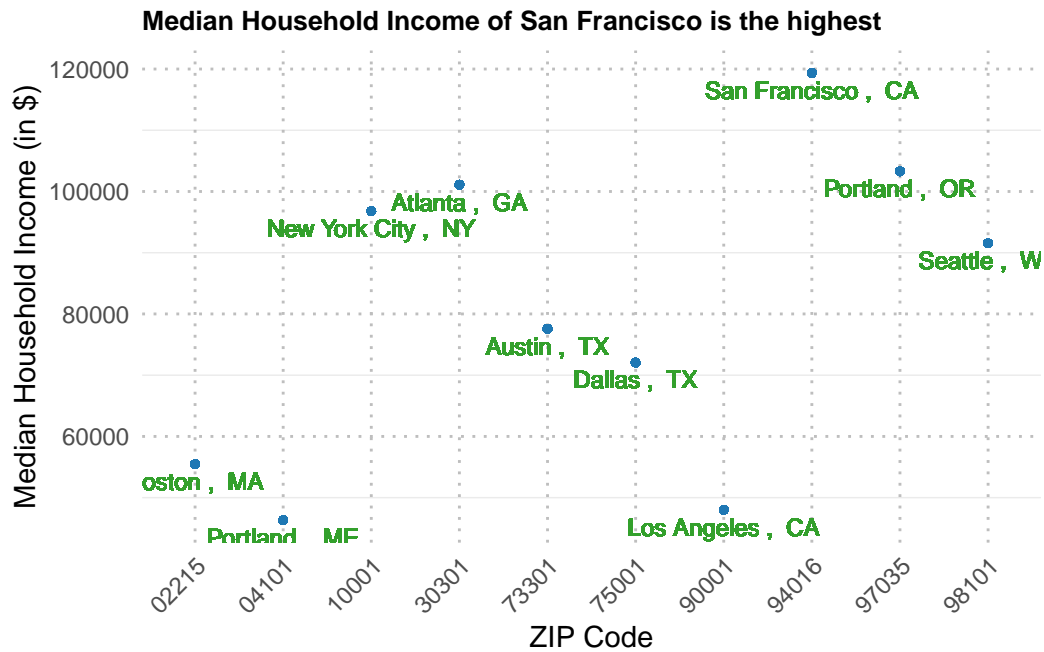
Yes, it seems that California also generates highest revenue for Apple AirPods compared to the 8 states.

Does high income people live in California?

Unfortunately, we have only one urban zipcode per city so let's see which City has highest Median household income.



## Visualization 6: Median Household Income by City

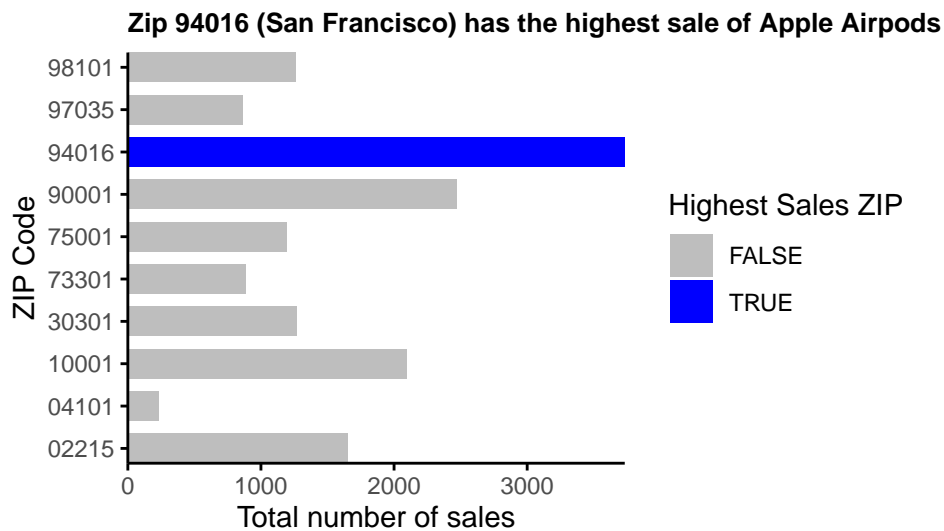


We find that the ZIP code 94016 which belongs to the city ‘San Francisco’ of California state has the highest median household income.

From this we can understand that there is a possibility that people belonging to this zipcode in San Francisco can afford more headphones.

Let’s verify that if sale of Apple AirPods is also highest in San Francisco.

## Visualization 7: Sales Distribution by Zip Code for Apple AirPods



So this looks like zip codes 94016 which belongs to San Francisco does have the highest sales.

We already saw highest revenue is from California and the zipcode '94016' of San Francisco having highest sales reconfirms that.

## Conclusion

Based on the analysis, we are taking a decision to call out Apple AirPods are the top-performing product in the headphones category in terms of revenue, with December being the peak sales month. The state of California, particularly the city of San Francisco contributes significantly to headphone sales. Additionally, zip code 94016 stand out as key areas for headphone sales, particularly for Apple AirPods.