

# Bike Sales Analysis

by Nicholas Legaspi

[https://github.com/nicholaslegaspi/personal\\_projects/tree/main/bike\\_sales](https://github.com/nicholaslegaspi/personal_projects/tree/main/bike_sales)

## Introduction

The following project investigates synthetic sales data of a bike store. This analysis will cover when the company makes the most money in the scope of the year-to-year, seasonal trends, month-to-month, and time of day.

## Methodology

In the project there were three datasets used. (each found in the *data* folder)

- *bike\_share\_yr\_0.csv*

|        |        |    |      |    |         |         |            |            |           |      |       |     |           |            |        |
|--------|--------|----|------|----|---------|---------|------------|------------|-----------|------|-------|-----|-----------|------------|--------|
| dteday | season | yr | mnth | hr | holiday | weekday | workingday | registered | weather_s | temp | atemp | hum | windspeed | rider_type | riders |
|--------|--------|----|------|----|---------|---------|------------|------------|-----------|------|-------|-----|-----------|------------|--------|

- *bike\_share\_yr\_1.csv*

|        |        |    |      |    |         |         |            |            |           |      |       |     |           |            |        |
|--------|--------|----|------|----|---------|---------|------------|------------|-----------|------|-------|-----|-----------|------------|--------|
| dteday | season | yr | mnth | hr | holiday | weekday | workingday | registered | weather_s | temp | atemp | hum | windspeed | rider_type | riders |
|--------|--------|----|------|----|---------|---------|------------|------------|-----------|------|-------|-----|-----------|------------|--------|

- *cost\_table.csv*

|    |       |      |
|----|-------|------|
| yr | price | COGS |
|----|-------|------|

The files, *bike\_share\_yr\_0.csv* and *bike\_share\_yr\_1.csv* are the main datasets which reported the amount of sales at each hour given the day and certain other conditions. Within these datasets the main columns which were most crucial to the project were the following: *dteday*, *season*, *yr*, *weekday*, *hr*, *rider\_type*, and *riders*.

The remaining file, *cost\_file.csv* is a dataset which has the price point and COGS for the given year, to be referenced.

A SQL query (found in GitHub as *bikesales.sql*) was written to be used in PowerBI (found in GitHub as *Bike Sales Dashboard.pbix*) to combine the tables from both years and calculate the revenues and profits in reference to the cost table.

## Recommendations

**Conservative Price increase:** In consideration of the price increase, we can see that comparing the years 2021 and 2022 the amount of riders has still doubled, resulting in an 106.25% increase in revenue year-over-year. Considering that the jump from the \$3.99 price in 2021 to \$4.99 in 2022 was a significant price increase, a more conservative price increase in the 10%-15% range for future years could test the market response without significant loss to the customer base.