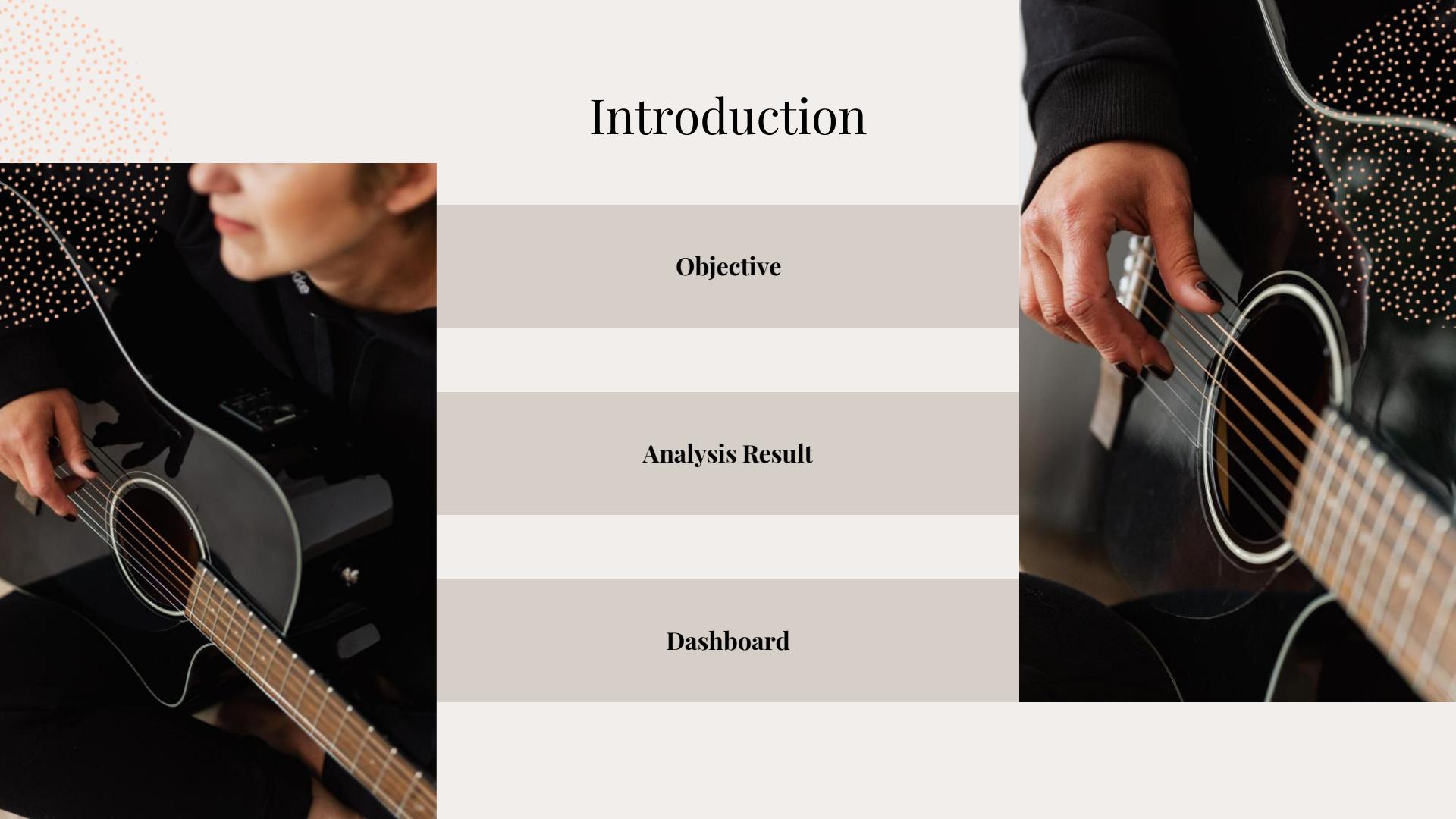


Chinook Sales Analysis

Farhaneh Nazemi



Augest 2024



Introduction

Objective

Analysis Result

Dashboard

Objective

The objective of this project is to analyze the Chinook Database to gain insights into the digital media store's operations. We will:

Understand Data Structure : Explore tables and relationships.

Perform SQL Queries : Extract meaningful information.

Generate Insights : Analyze customer behavior and sales trends.

Visualize Data : Create visual representations to support data-driven decisions.

This analysis will help in making informed business decisions and strategies.

Analysis Result



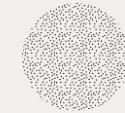
Q1. Get data from My SQL DB

- Install SQL Alchemy package
- Create .json file include DB configuration
- Get all table into separate data frames

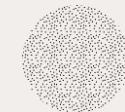
Q2. Initial data examination

```
{df.shape}  
{df.columns.tolist()}  
{df.dtypes}  
{df.isnull().sum()}  
{df.describe(include='all')}
```

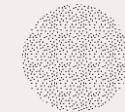
Q3,4. Visualizing Key Variables



Revenue

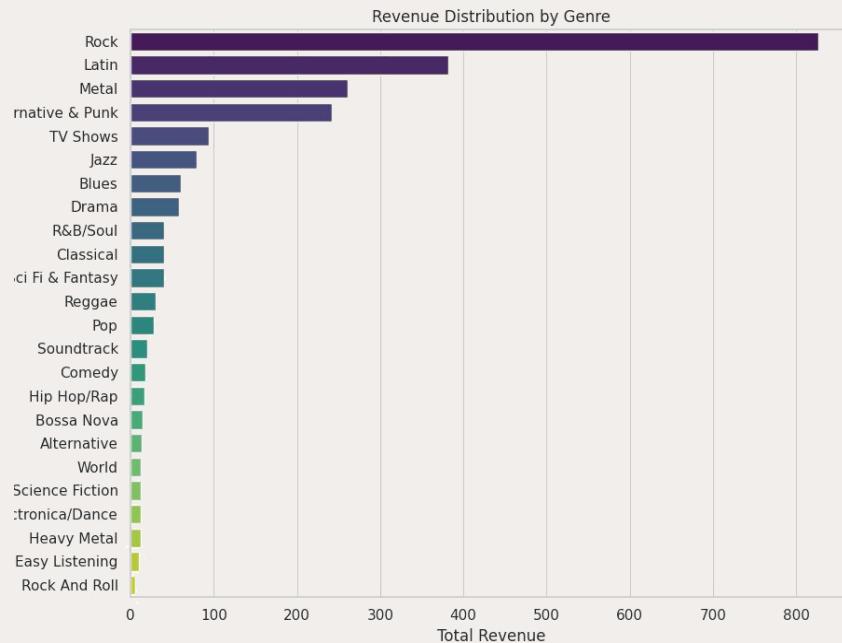


Artist

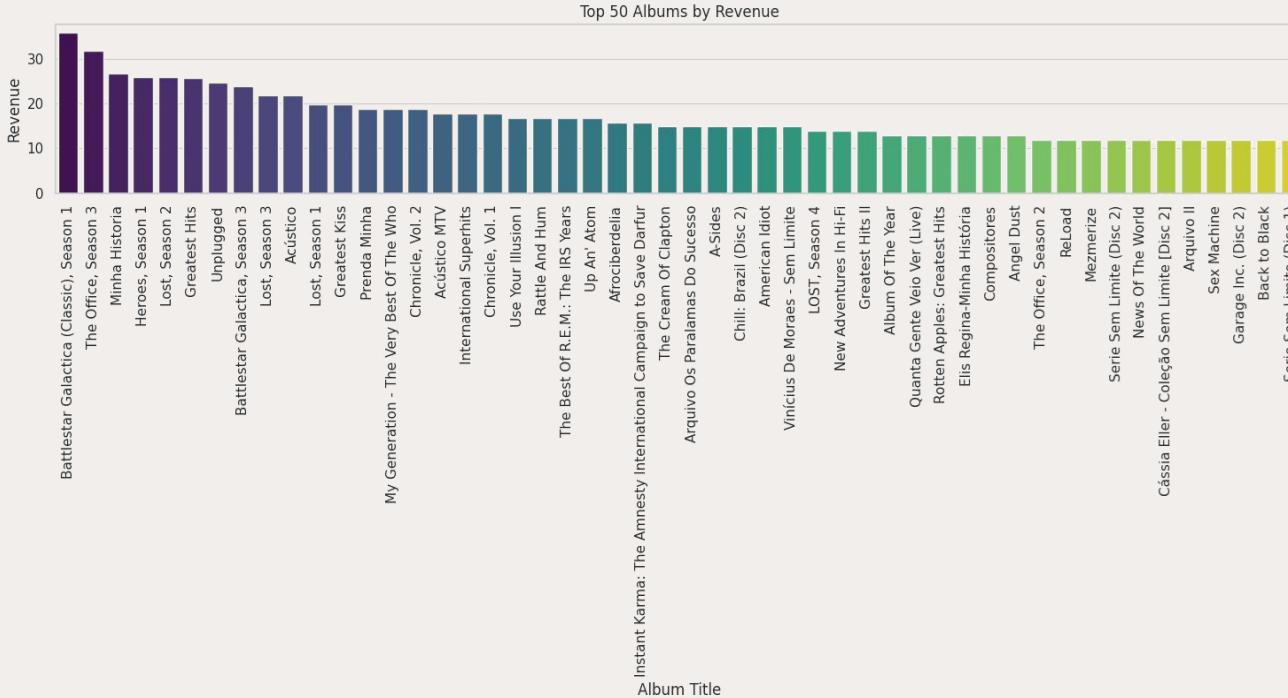


Customer

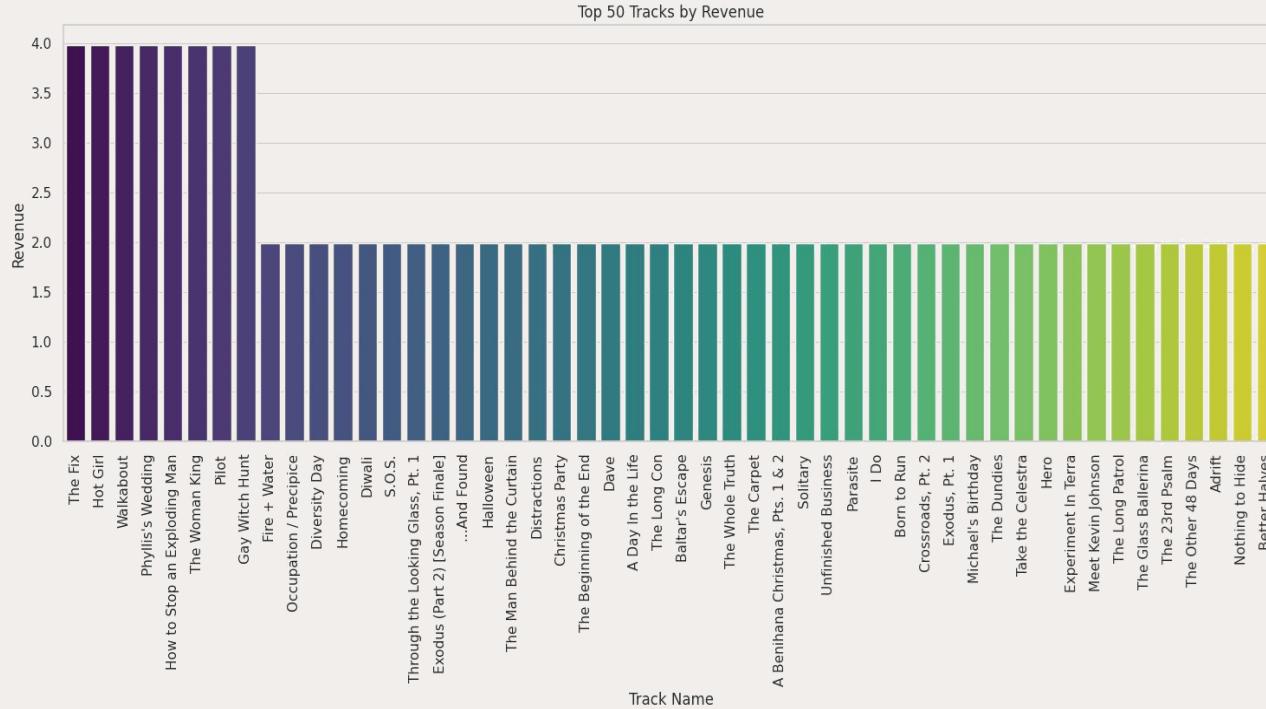
Q3,4. Revenue Analysis



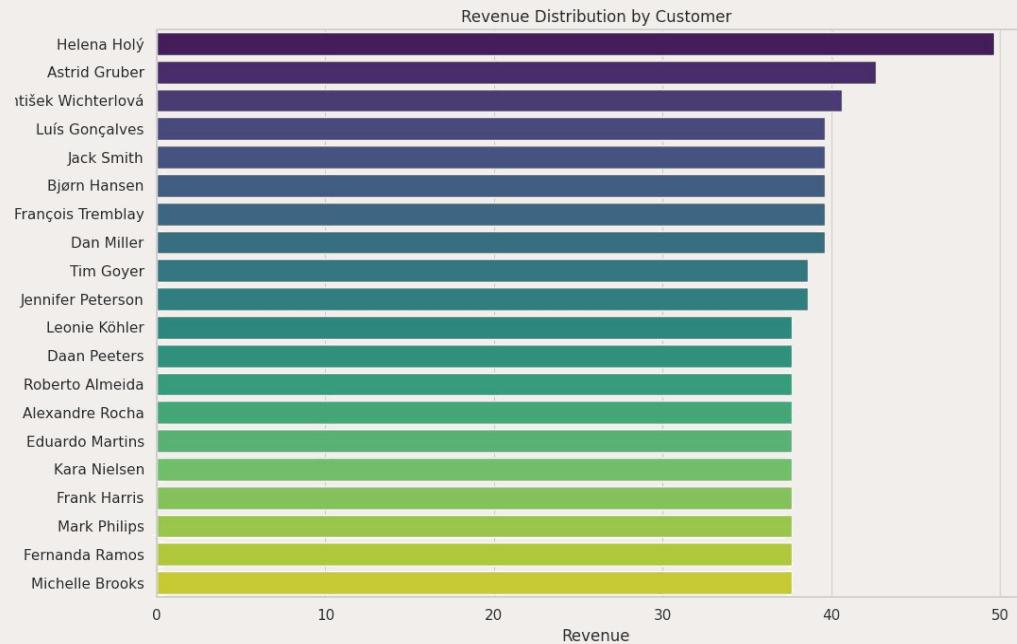
Q3,4. Revenue Analysis



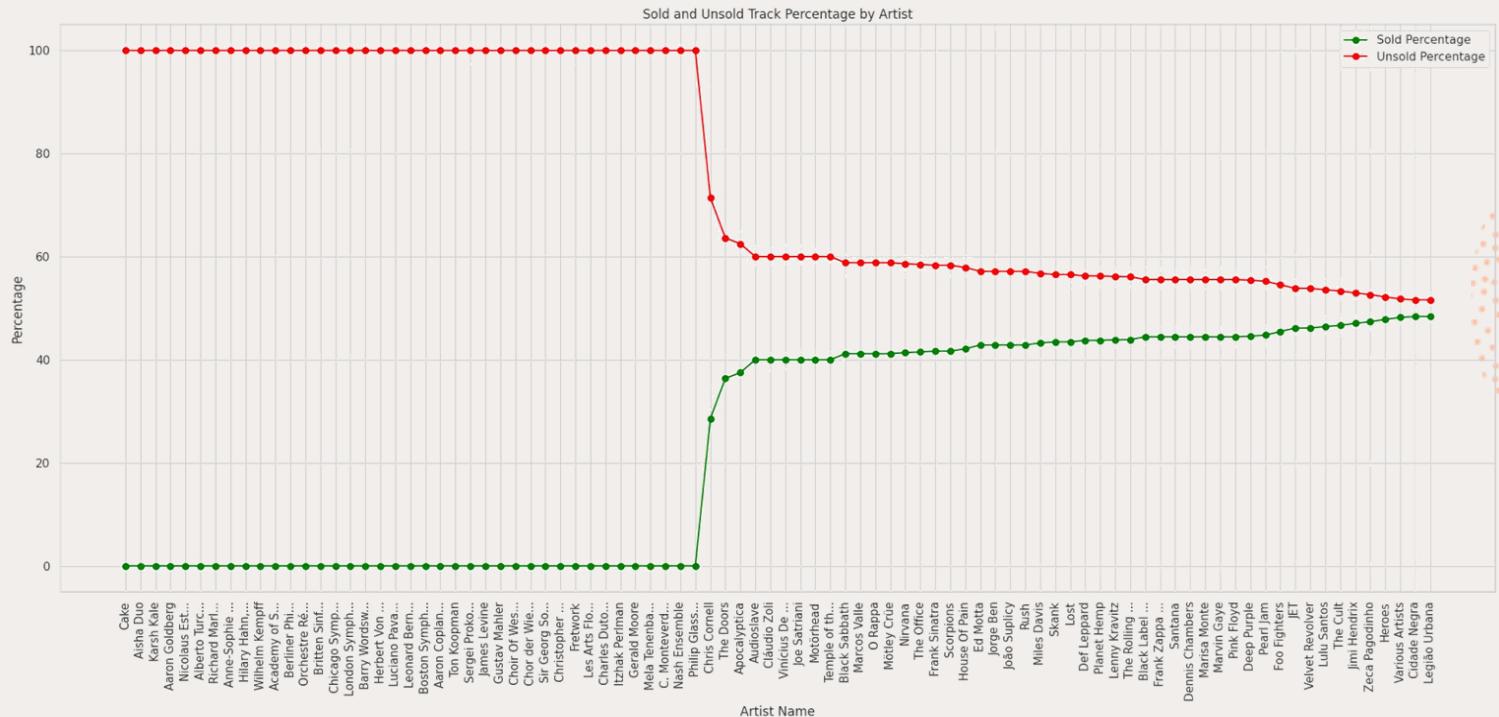
Q3,4. Revenue Analysis



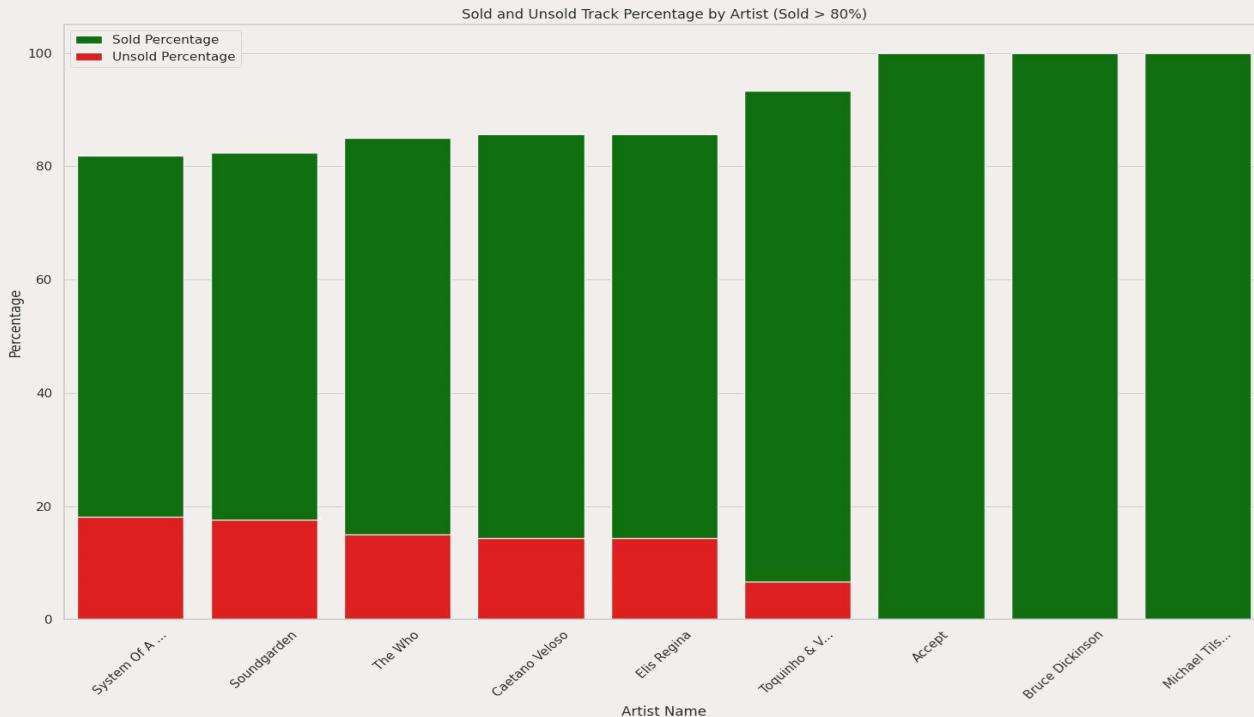
Q3,4. Revenue Analysis



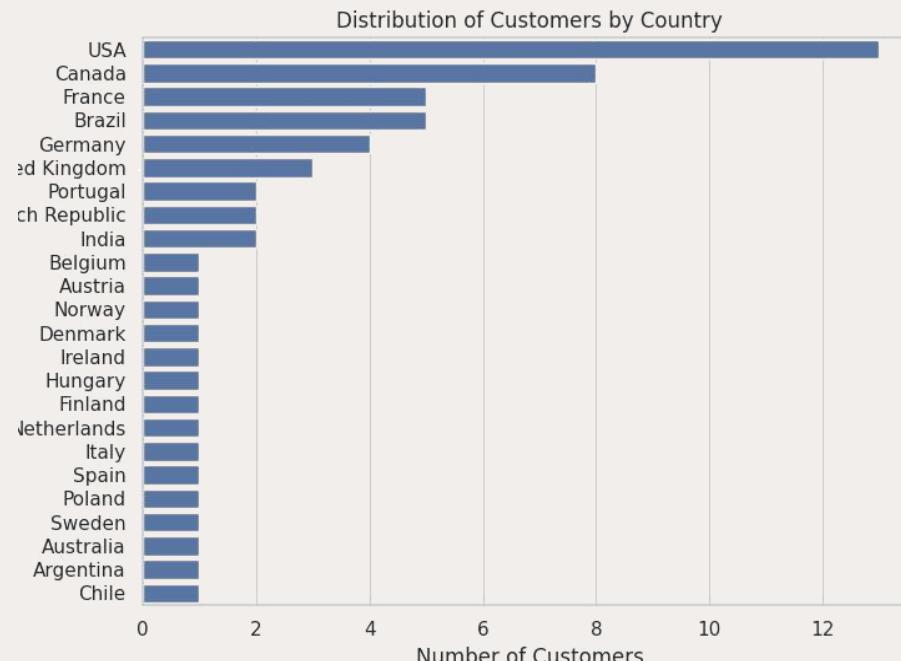
Q3,4. Artist Analysis



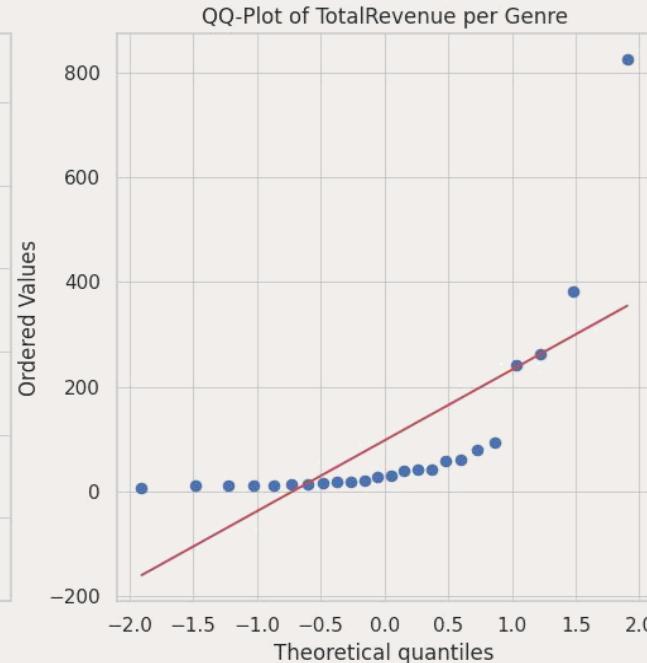
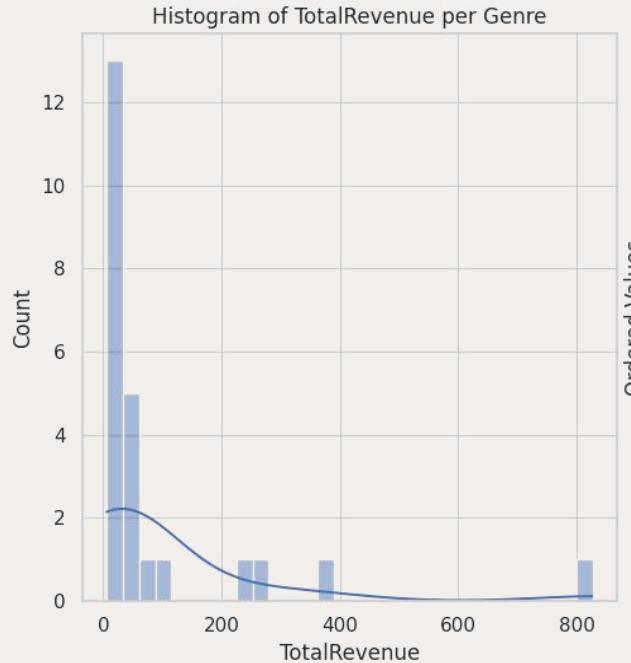
Q3,4. Artist Analysis



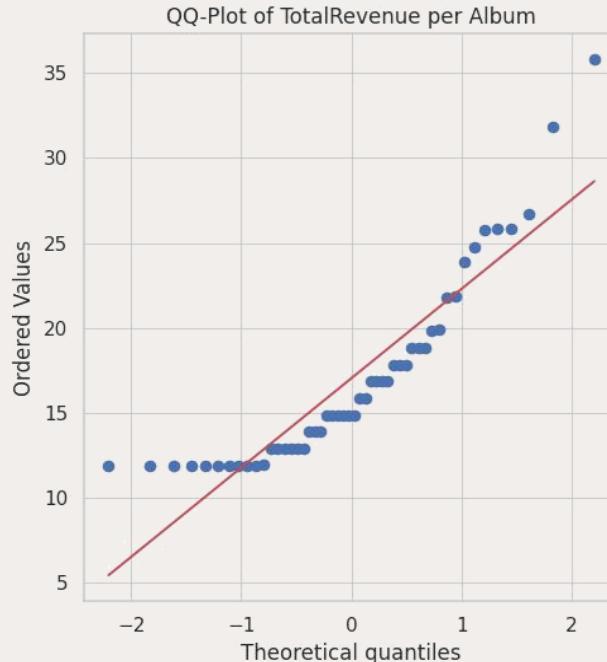
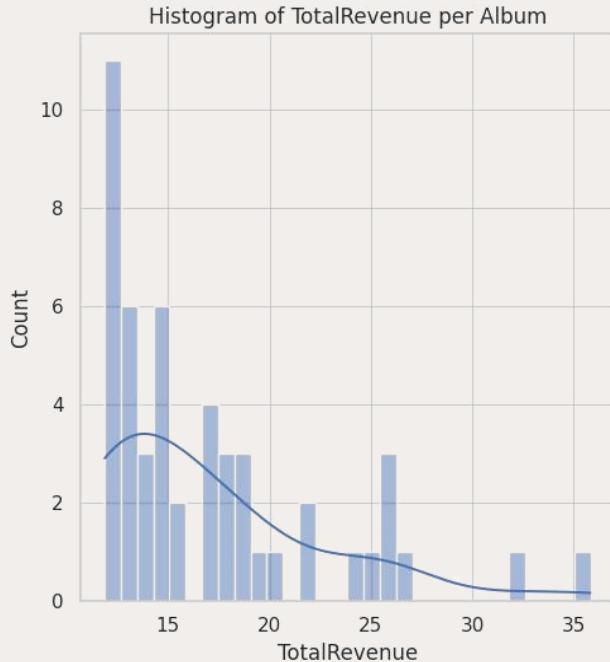
Q3,4. Customer Analysis



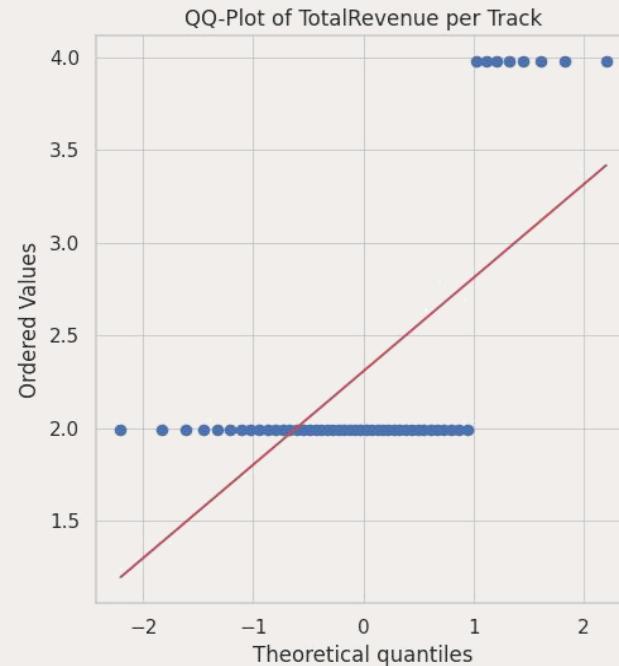
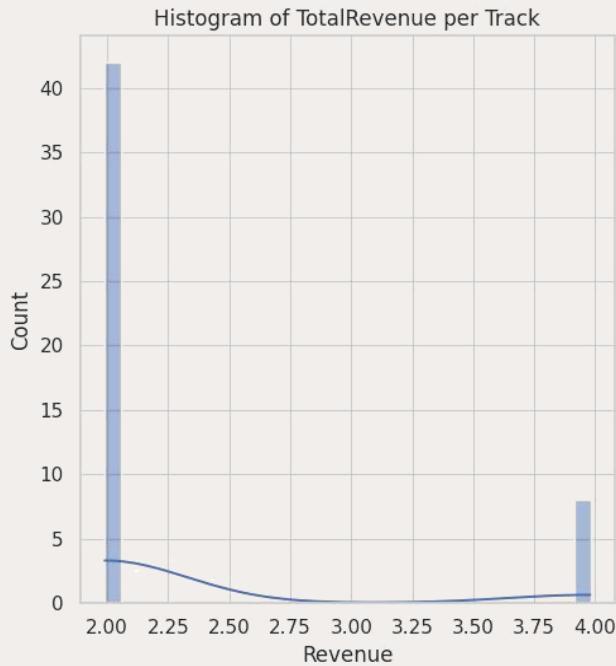
Q5. Check Normality of Data Using Plots



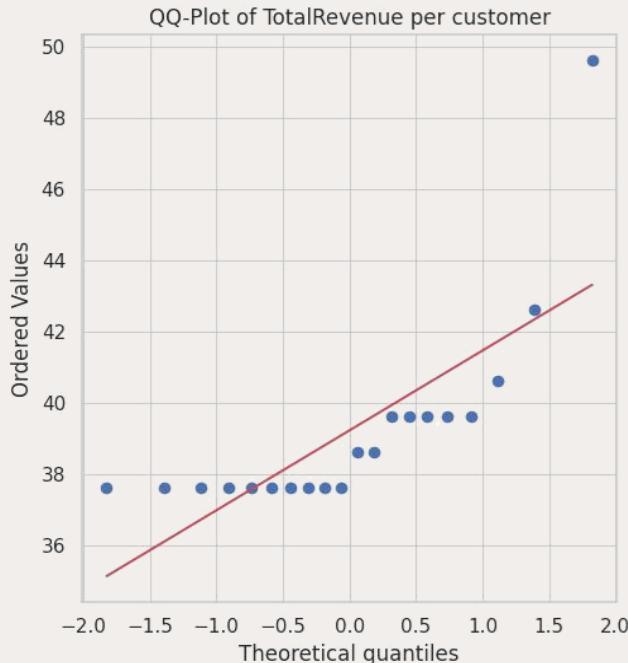
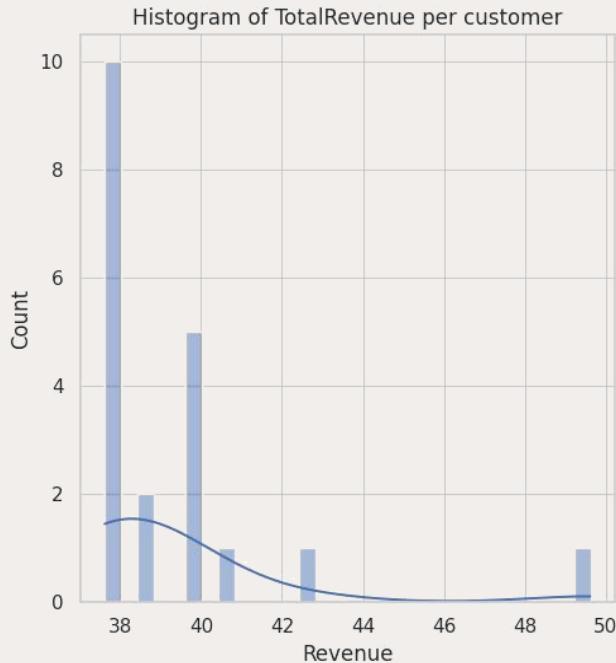
Q5. Check Normality of Data Using Plots



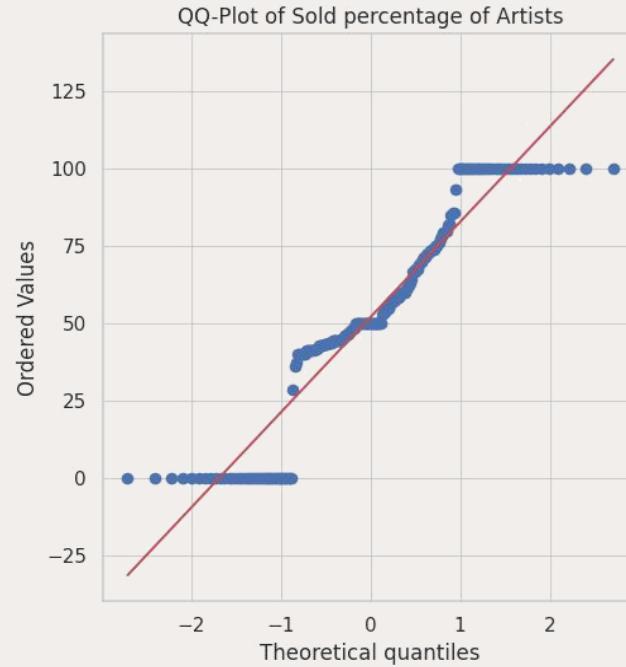
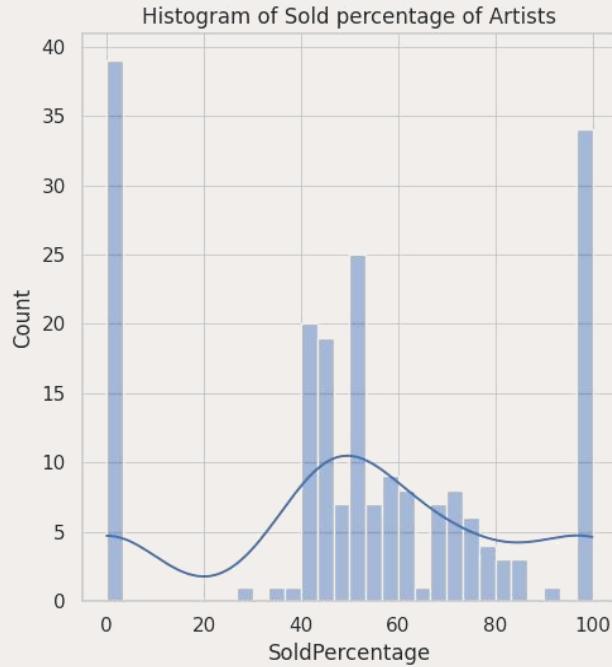
Q5. Check Normality of Data Using Plots



Q5. Check Normality of Data Using Plots



Q5. Check Normality of Data Using Plots



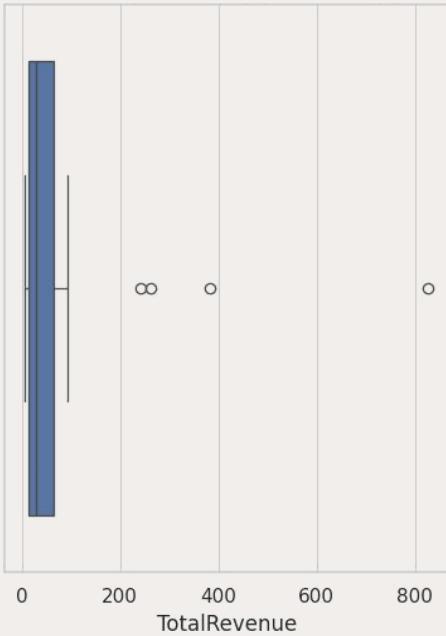
Q5. Check Normality of Data Using Statistical Tests

- Shapiro-Wilk Test
- Kolmogorov-Smirnov Test
- D'Agostino's K² Test

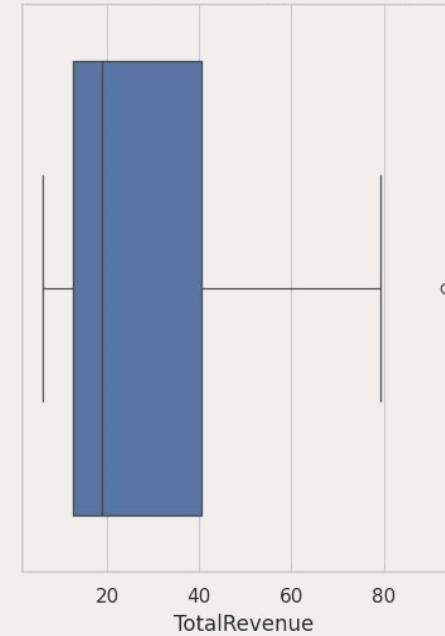
Reject the null hypothesis: The data does not appear to be normally distributed.

Q6. Cleaning Outlier Data

TotalRevenue per Genre (IQR Cleaned) (Before)

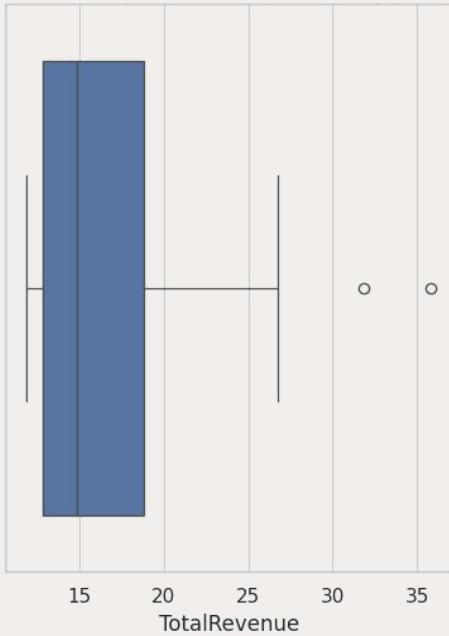


TotalRevenue per Genre (IQR Cleaned) (After)

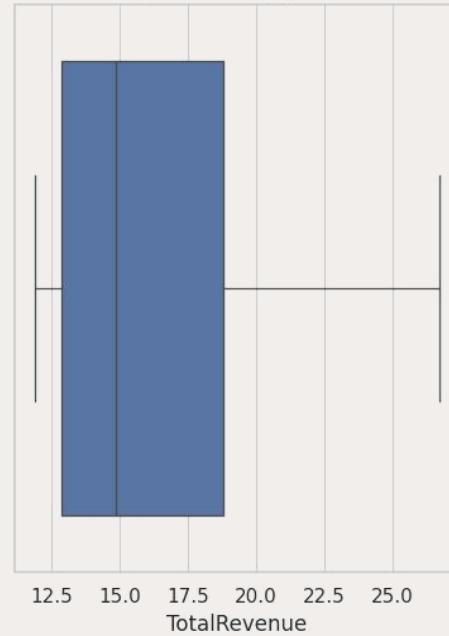


Q6. Cleaning Outlier Data

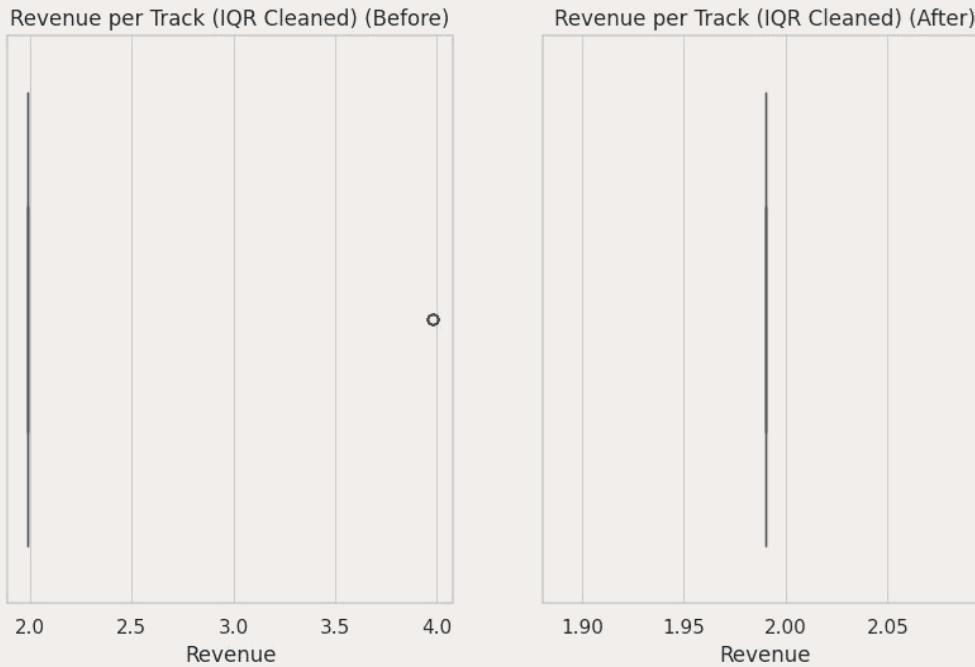
TotalRevenue per Album (IQR Cleaned) (Before)



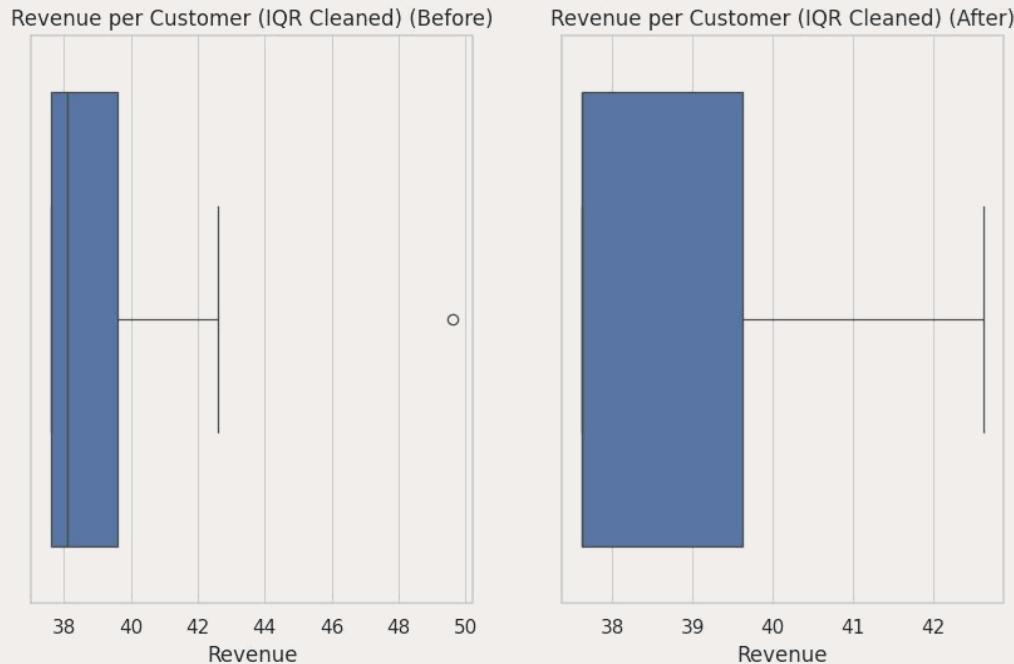
TotalRevenue per Album (IQR Cleaned) (After)



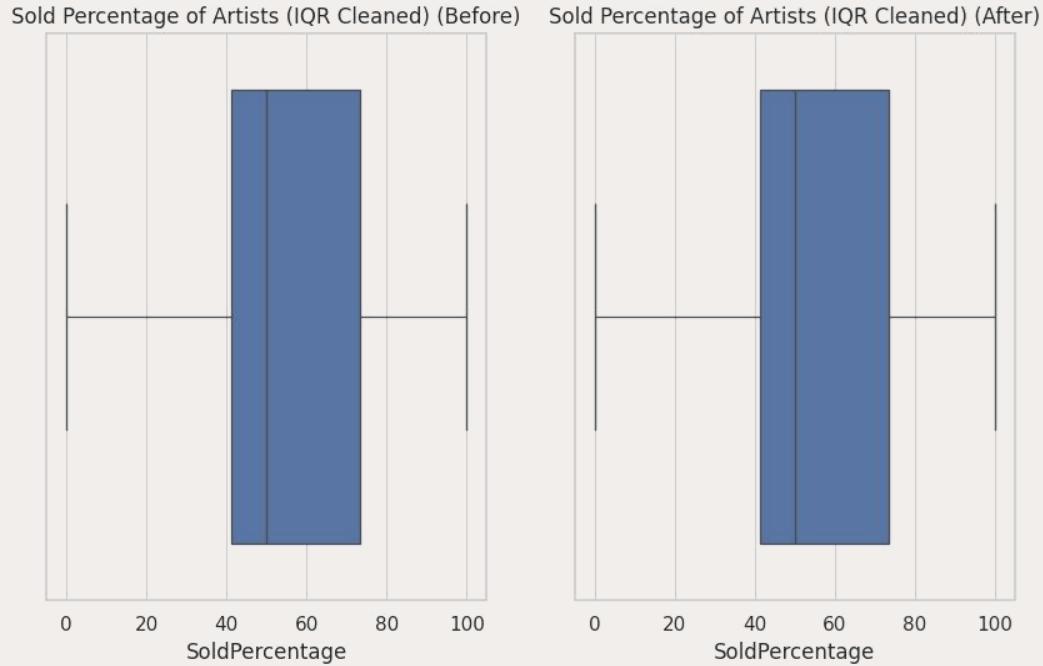
Q6. Cleaning Outlier Data



Q6. Cleaning Outlier Data



Q6. Cleaning Outlier Data



Q7. Correlation Between Mean Unit Price among the top Genres

Name	Total Revenue	Unit Price
Rock	826.65	0.99
Latin	382.14	0.99
Metal	261.36	0.99

It means that there's no variability in the prices, and hence, statistical tests like ANOVA or Kruskal-Wallis won't be useful since they require variability to compare means.

Q7. Correlation Between Track Length and Unit Price

Pearson

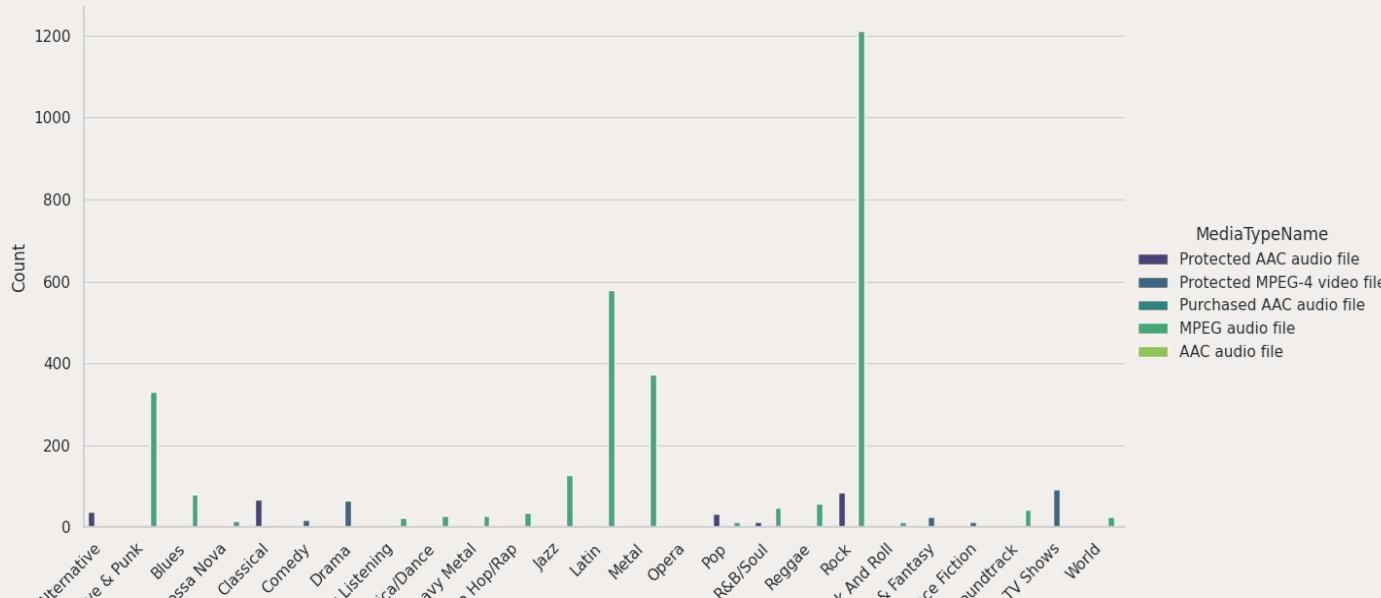
Reject the null hypothesis: There is a significant correlation between track length and unit price.



Q7. Association Between Genre and Media Type

Chi²

Reject the null hypothesis: There is a significant association between genre and media type.



Q7. Additional Findings

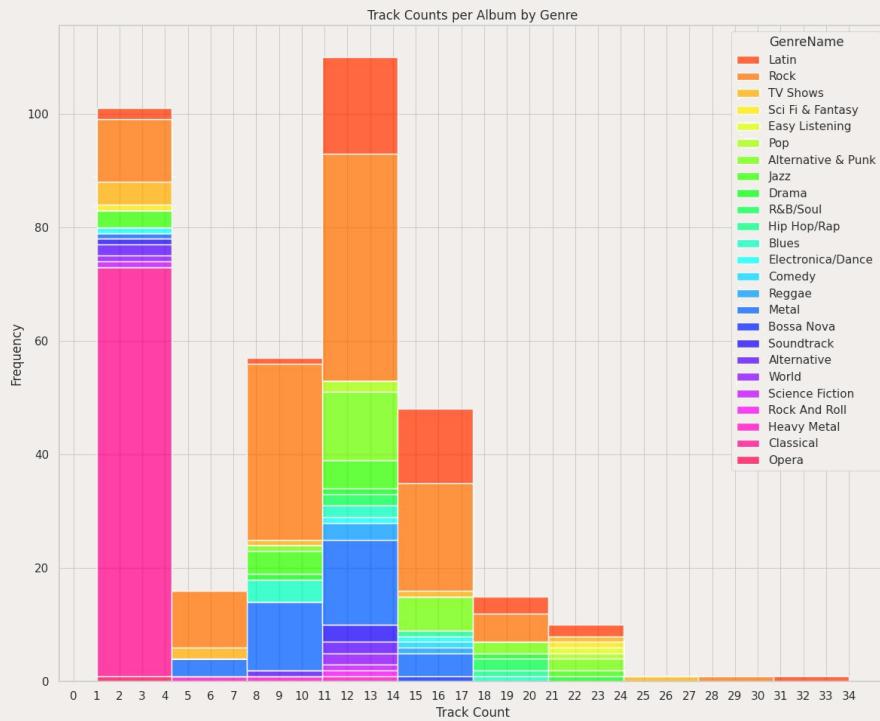
1. Relationship between Track Count and Genre

2. Correlation Between Track Count and Revenue of album

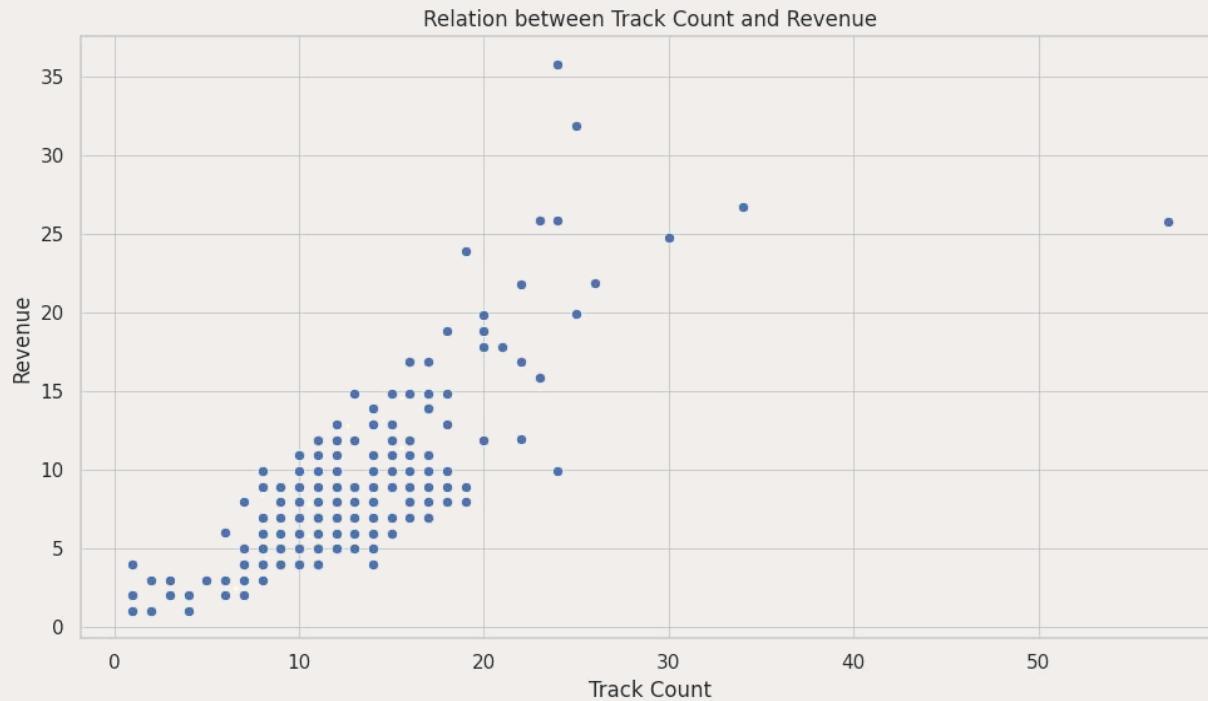
3. Correlation between Track Duration and Revenue of top tracks



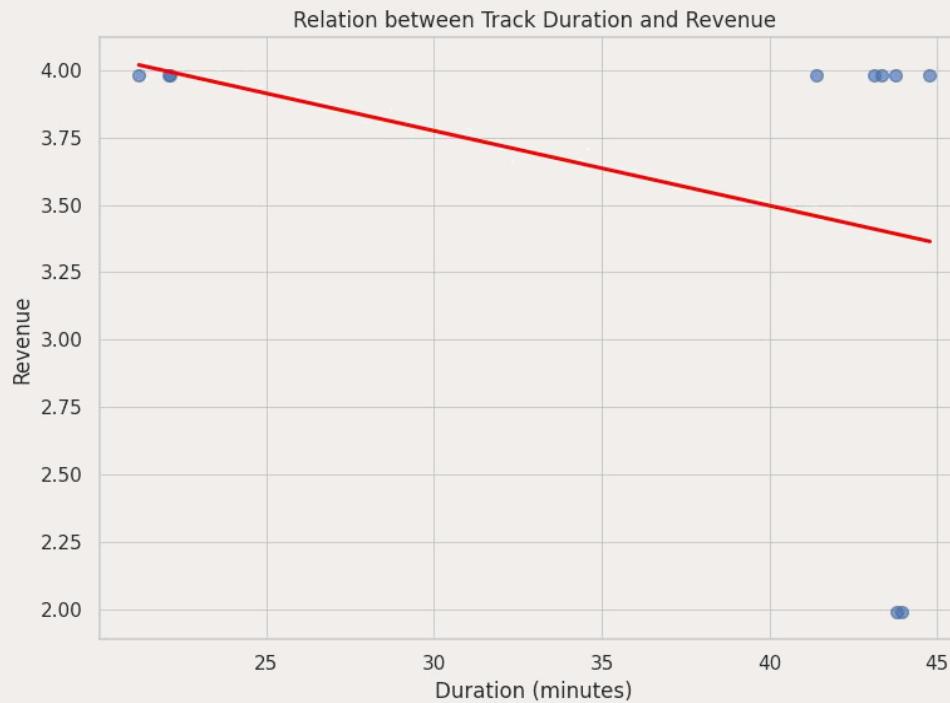
Q7. There is a significant relationship between Track Count and Genre. (Chi-Square)



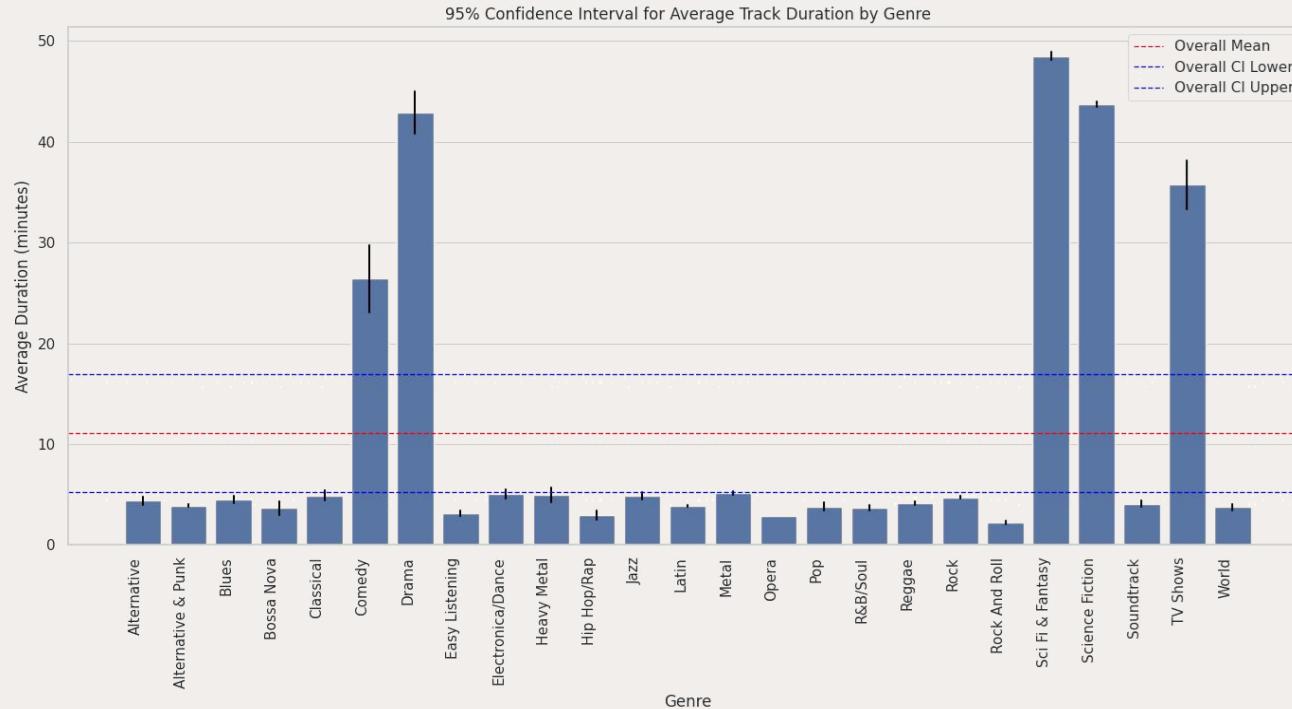
Q7. Correlation between Track Count and Revenue of album : 0.82



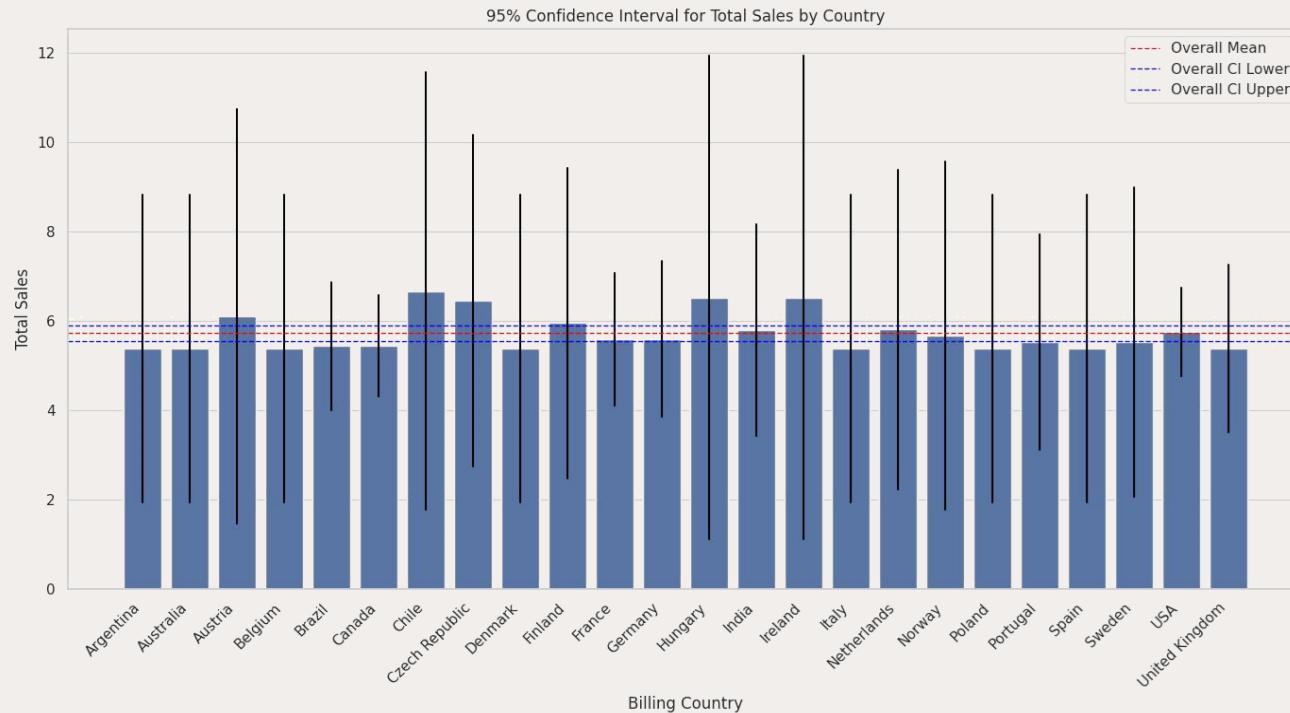
Q7. Correlation between Track Duration and Revenue of top tracks: -0.35



Q8. 95% Confidence Interval for Average Track Duration by Genre

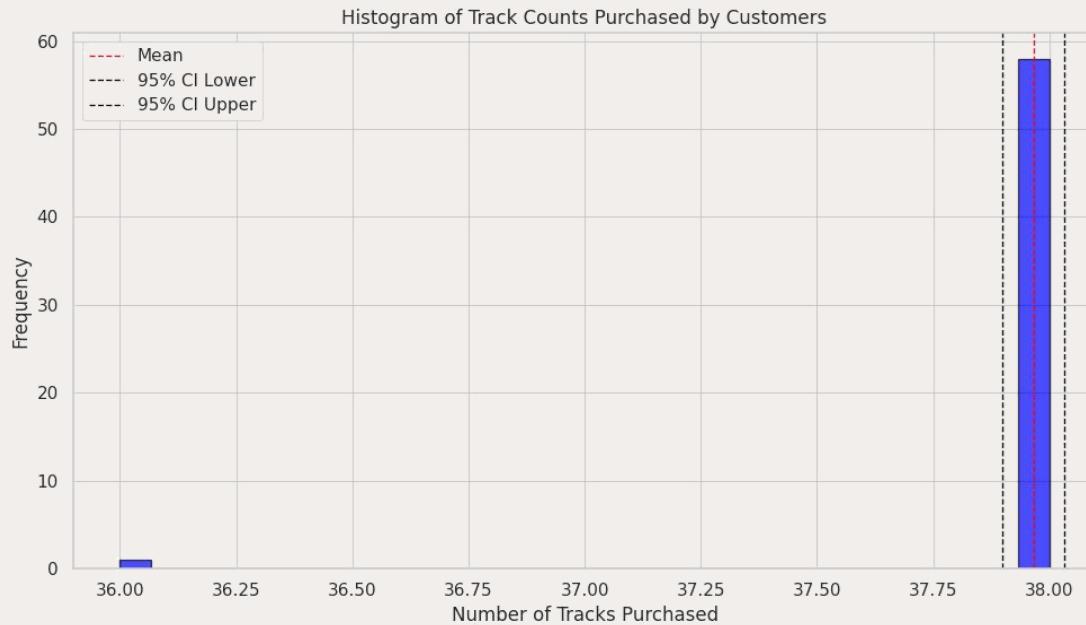


Q8. 95% Confidence Interval for Total Sales by Country



Q8. 95% Confidence Interval for Track Counts Purchased by Customers

Mean Track Count: 37.97 , 95% Confidence Interval: (37.90, 38.03)



Dashbord





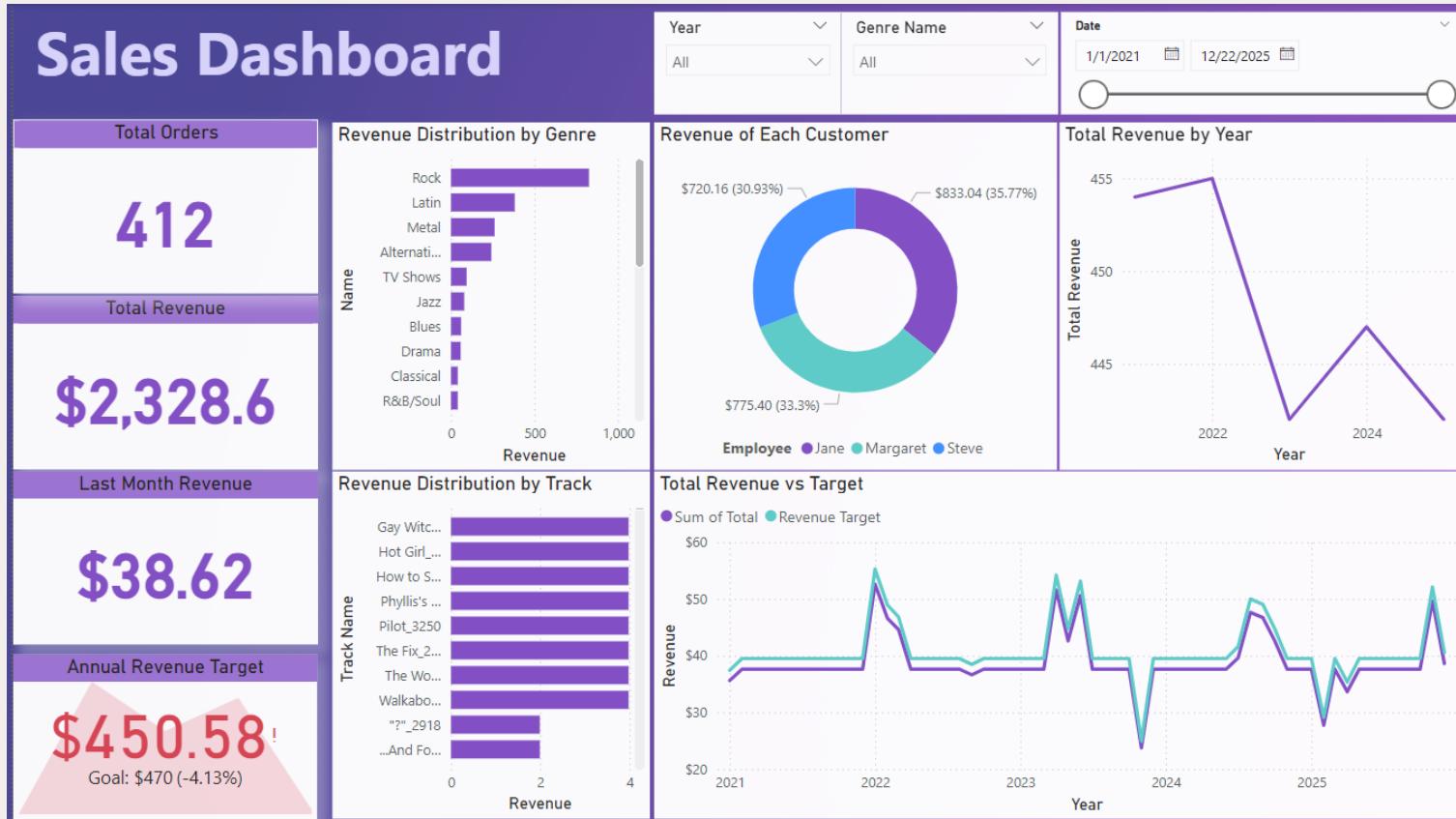
Home Page

Welcome to the Sales Dashboard!

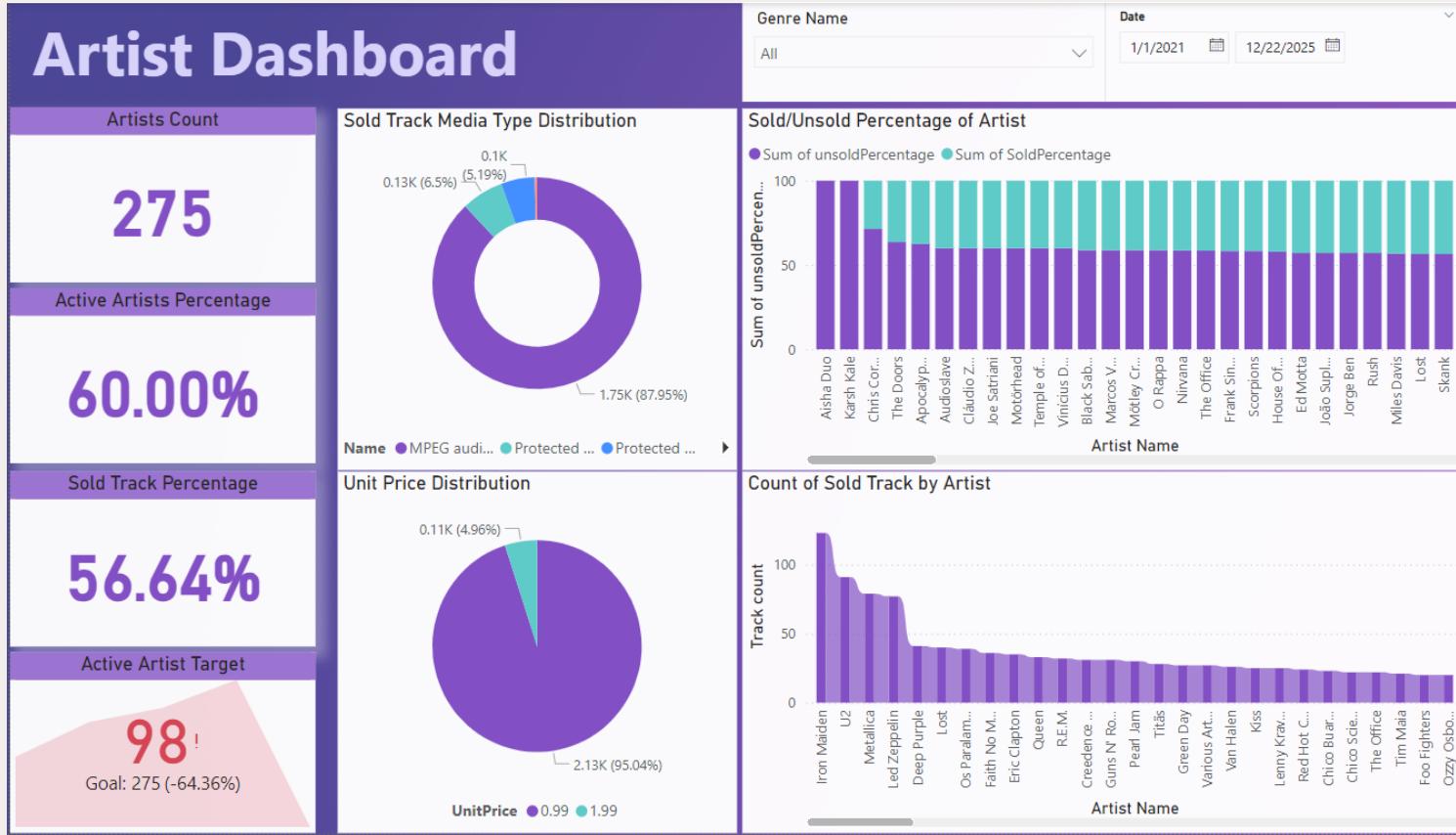
- 💡 This interactive dashboard provides a detailed overview of sales performance, tracks sales trends, and identifies top-performing genres, tracks, artists, and customers, helping you gain valuable insights into various aspects of the business.
 - 💡 The dashboard includes key performance indicators (KPIs) relevant to management's needs. If a new KPI is required, please send your requirements to the email address provided.
 - 💡 This dashboard is accessible via the internal company network at the address provided.
-
- ⓘ The currency unit is "dollars" .
 - ⓘ The unit of measurement for tracks is "quantity".
 - ⓘ Last update : 2025-12-22



Overview



Artist





Customer

Customer Dashboard

Customer

59

Active Customer

21

Customer Attraction

7!

Goal: 40 (-82.5%)

Customers of Each Country

Country	Count of CustomerId
Canada	5
USA	4
Brazil	2
France	2
Germany	2
India	2
United Kingdom	2
Finland	1
Hungary	1
Ireland	1

Count of CustomerId by Year

Year	Count of CustomerId
2021	14
2022	7

Supporter Employee

FirstName
Andrew
Laura
Michael
Robert
Jane
Margaret
Nancy
Steve

InvoiceDate

Year: 2021, 2022

Invoice Date Range: 1/1/2021 - 12/22/2025

Total Revenue by Country and City

Legend: Increase (Green), Decrease (Red), Total (Purple), Other (Yellow)

Country/City	Revenue Type	Revenue Value
Canada	Total	~190
Chicago	Increase	~20
Cupertino	Increase	~20
New York	Increase	~20
Vancouver	Decrease	~20
Montreal	Decrease	~20
Other	Other	~20
USA	Total	~110
Frankfurt	Increase	~20
Berlin	Increase	~20
Other	Decrease	~20
New York	Decrease	~20
Cupertino	Decrease	~20
Chicago	Decrease	~20
Germany	Total	~10
Dijon	Increase	~10
Bordeaux	Increase	~10
Bangalore	Decrease	~10
Other	Decrease	~10
Berlin	Decrease	~10
Frankfurt	Decrease	~10
France	Total	~10
São José do C...	Increase	~10
Rio de Janeiro	Increase	~10
Bangalore	Decrease	~10
Other	Decrease	~10
Bordeaux	Decrease	~10
Dijon	Decrease	~10



Thanks for your attention.