Chinook Database Analysis

Investigation of track, genre, album DataFrames

Rock and Latin genres have the highest frequency
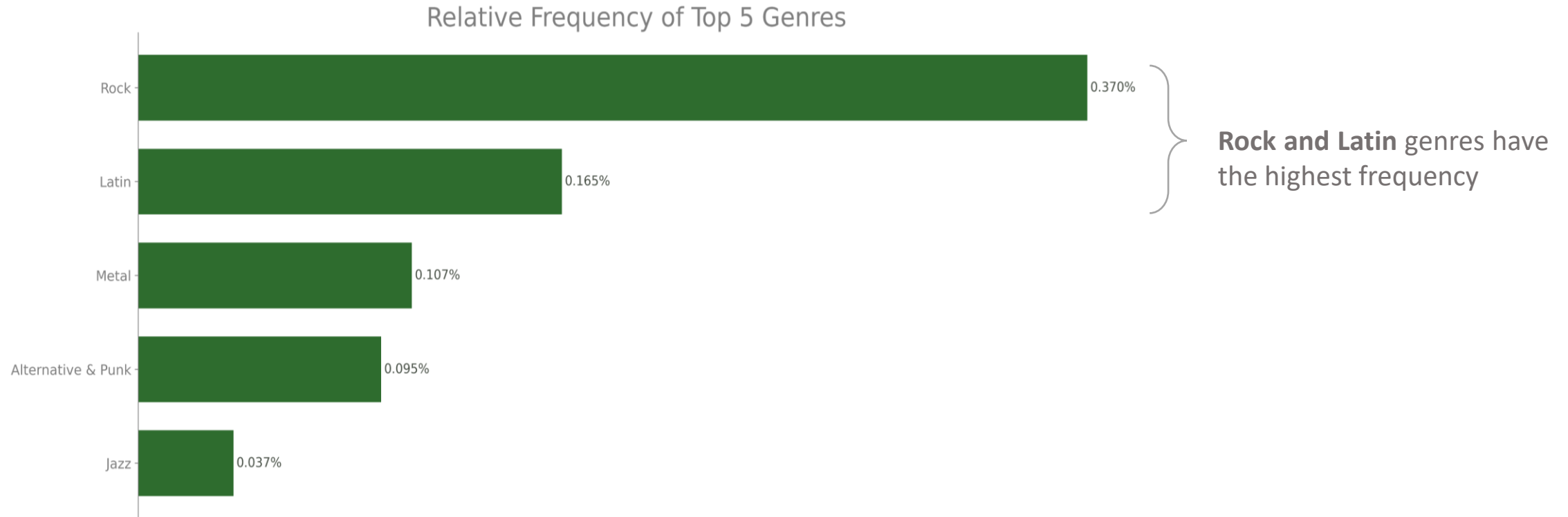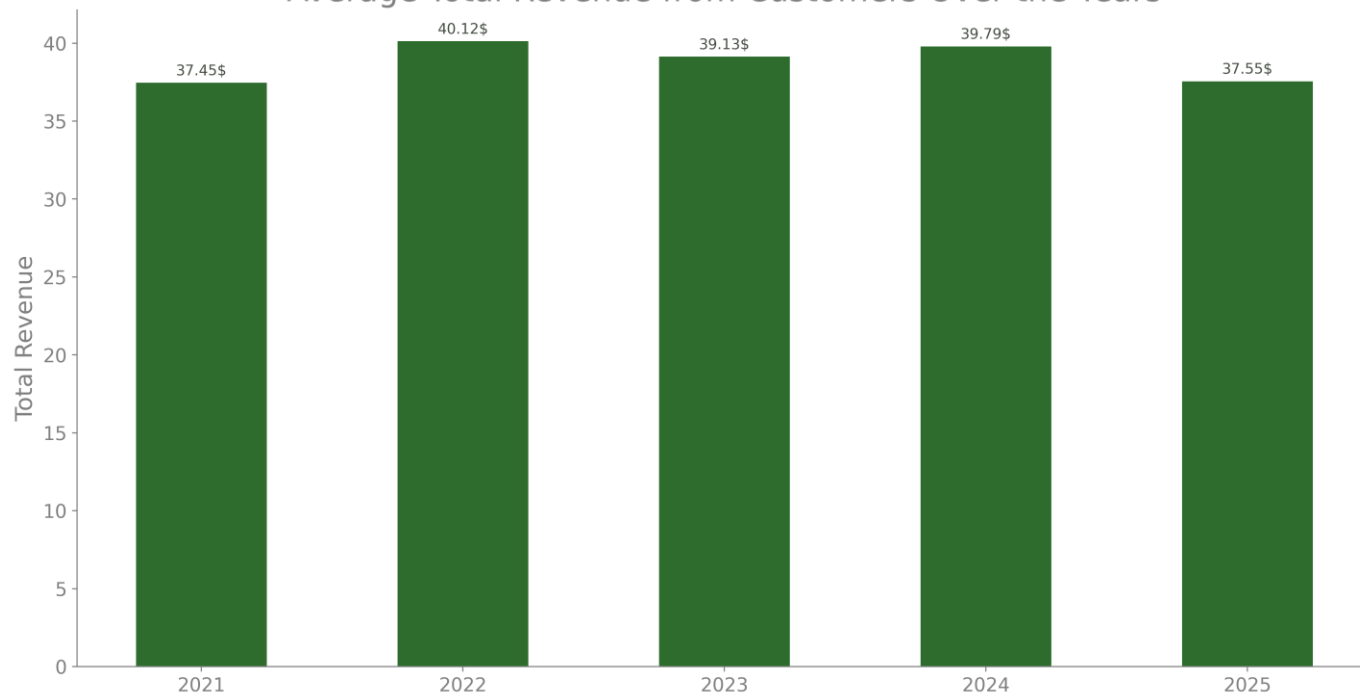
Investigation of track, genre, album DataFrames
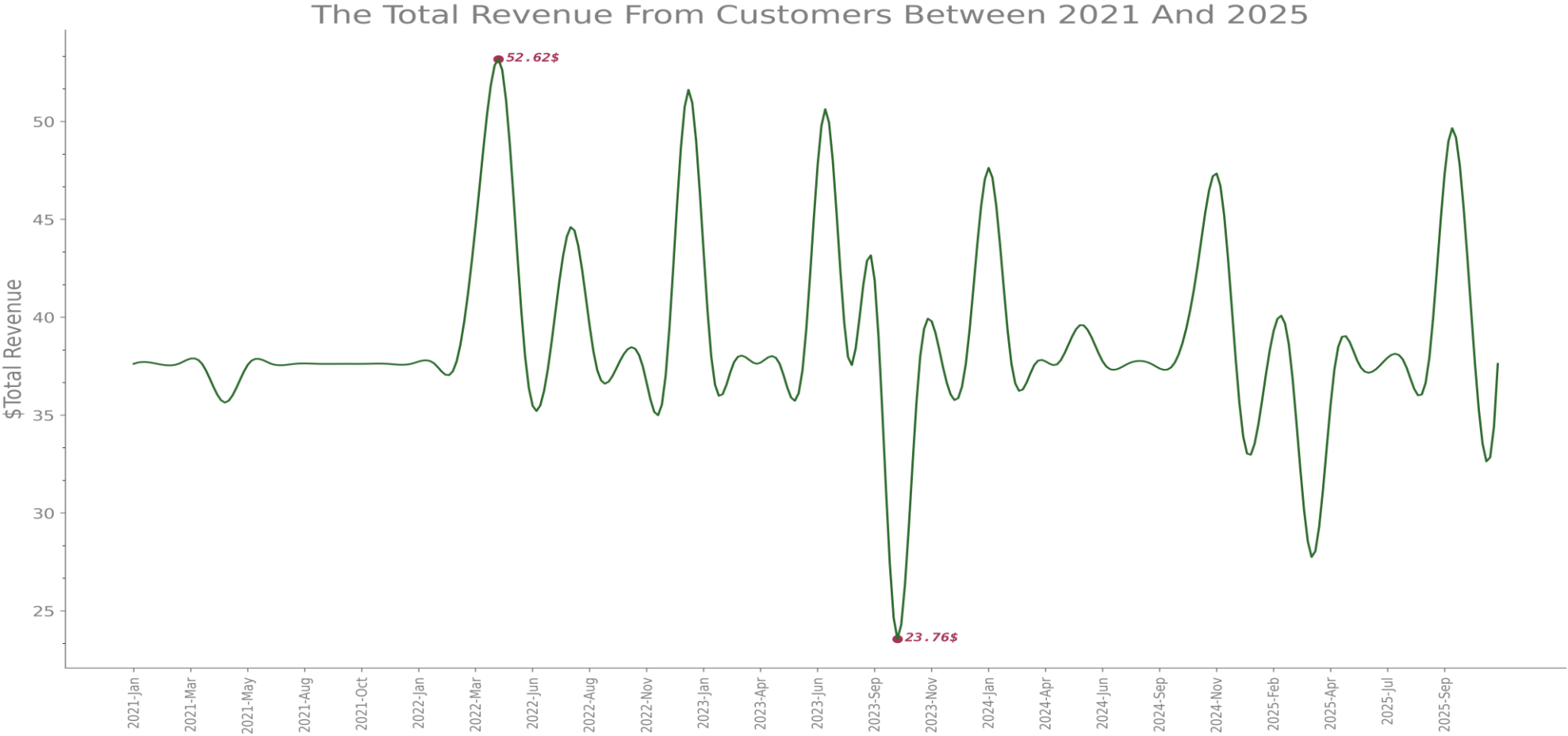

Total Unit Price of Top 11 Album Names
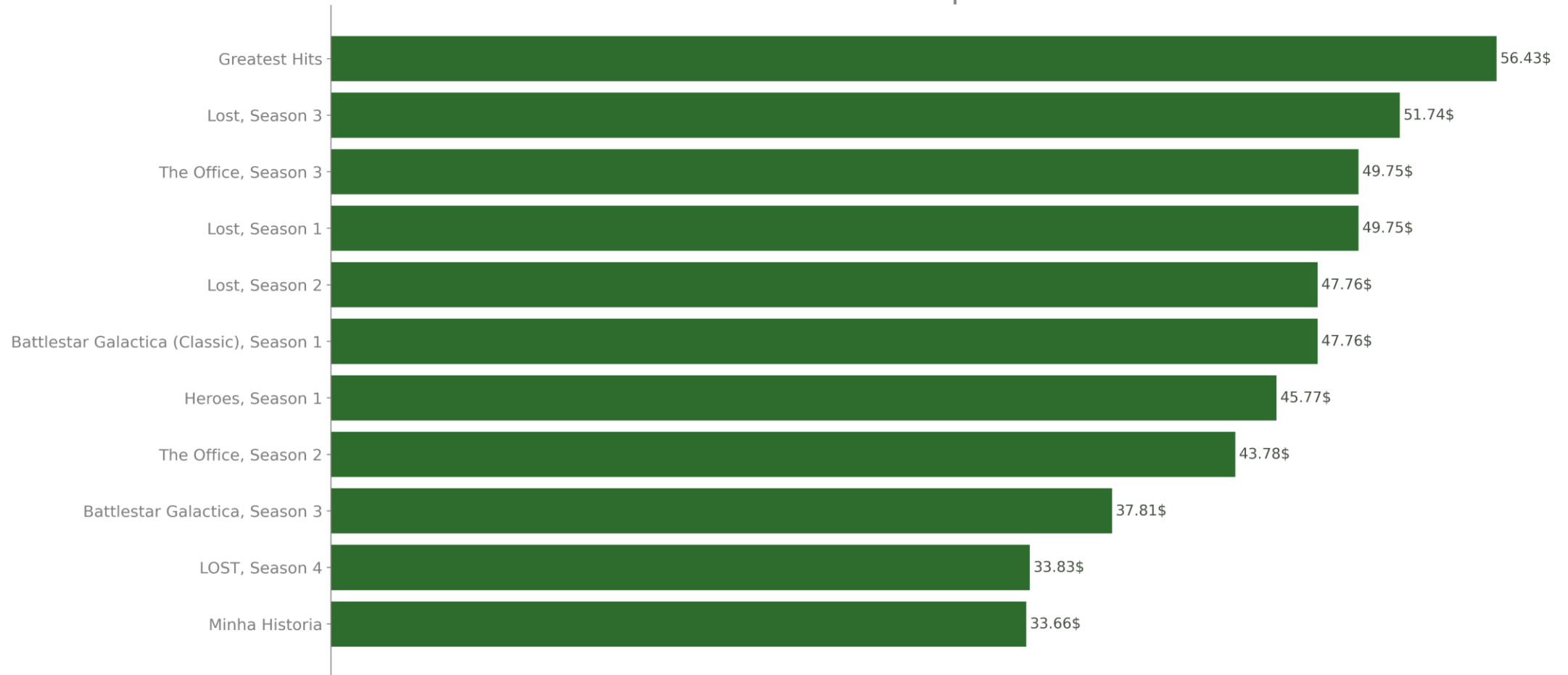
Investigation of invoice DataFrame

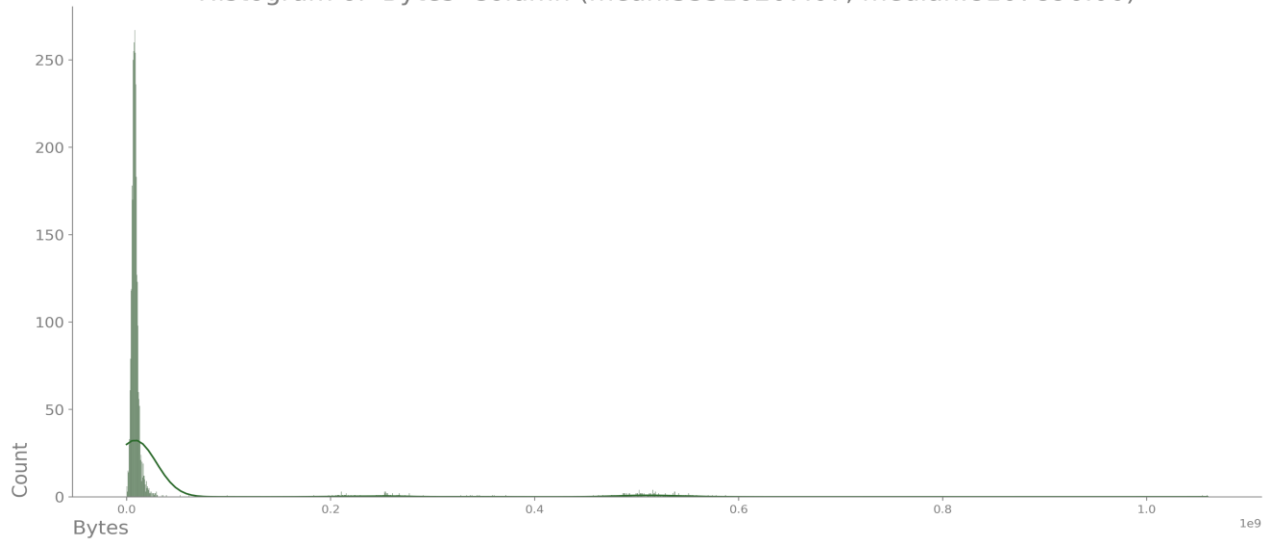Average Total Revenue from Customers Over the Years

- Total revenue over the years is almost the same, but in **2022**, it differs slightly from the others.

Investigation of invoice DataFrame

The Total Revenue From Customers Between 2021 And 2025

Normality of Numerical Features

Histogram of 'Bytes' Column (mean:33510207.07, median:8107896.00)

	UnitPrice	Bytes
Shapiro-Wilk	p= 0.000	p= 0.000
Anderson-Darling	value= 0.78600 (p= 0.05)	value= 0.78600 (p= 0.05)
Kolmogorov-Smirnov	p= 0.000	p= 0.000

Box Plot of 'Bytes' Column (mean:33510207.07, median:8107896.00)
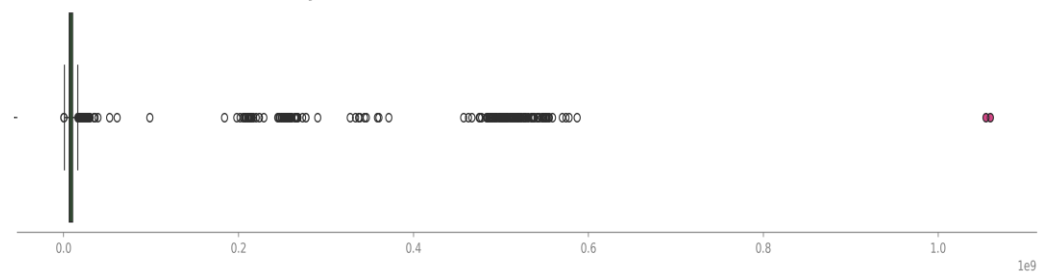
Box Plot of 'Bytes' Column (mean:32925531.35, median:8106249.00)

Normality of Numerical Features

Histogram of 'Milliseconds' Column (mean:393599.21, median:255634.00)

Histogram of 'Total' Column (mean:5.65, median:3.96)
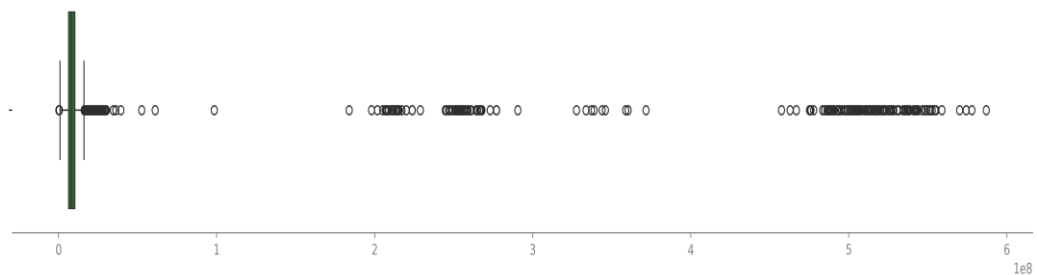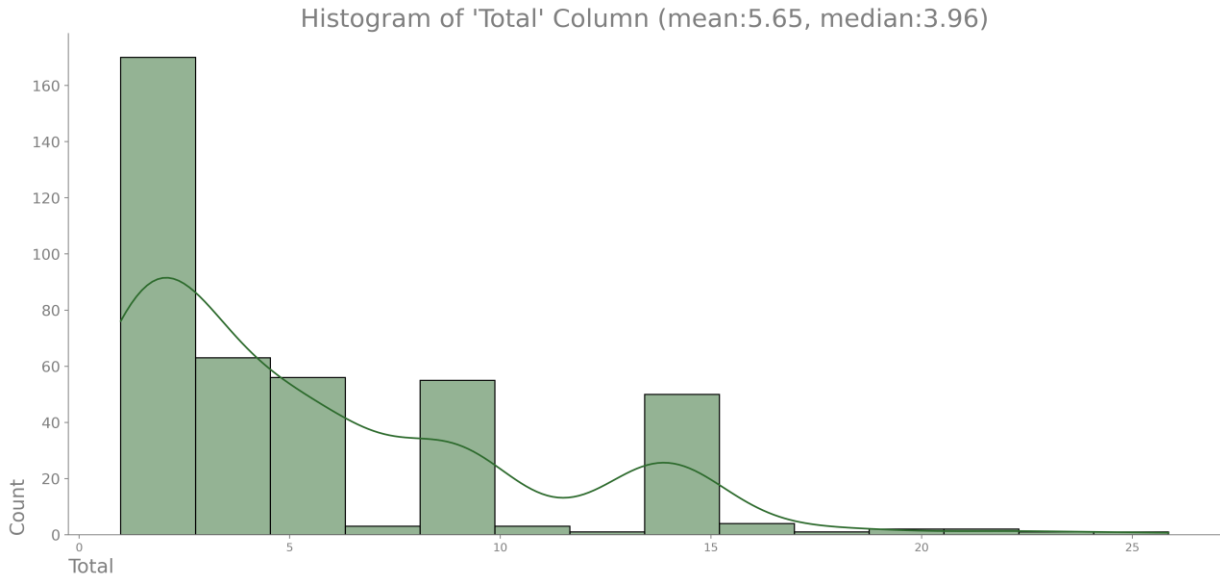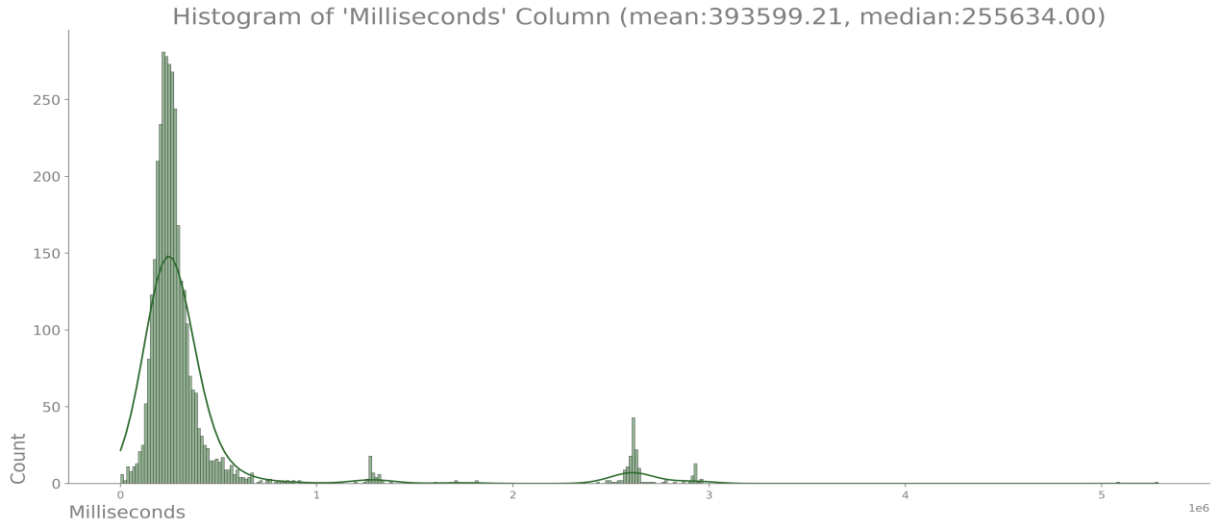
	Milliseconds	**Total**
Shapiro-Wilk	p= 0.000	p= 0.000
Anderson-Darling	value= 0.78600 (p= 0.05)	value= 0.78000 (p= 0.05)
Kolmogorov-Smirnov	p= 0.000	p= 0.000

CHINOOK MUSIC SERVICE

Normality of Numerical Features

Box Plot of 'Milliseconds' Column (mean:393599.21, median:255634.00)
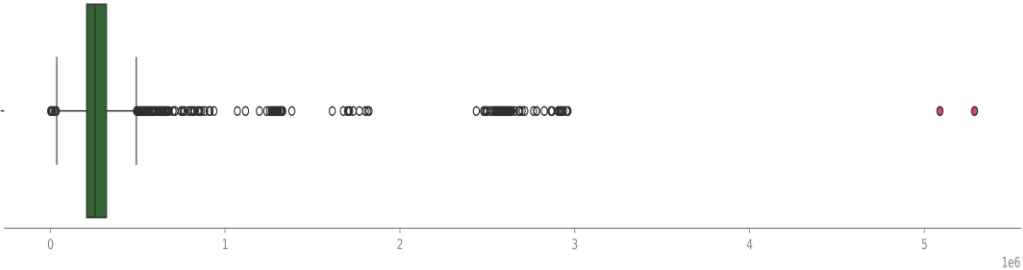
Box Plot of 'Milliseconds' Column (mean:390860.40, median:255608.00)
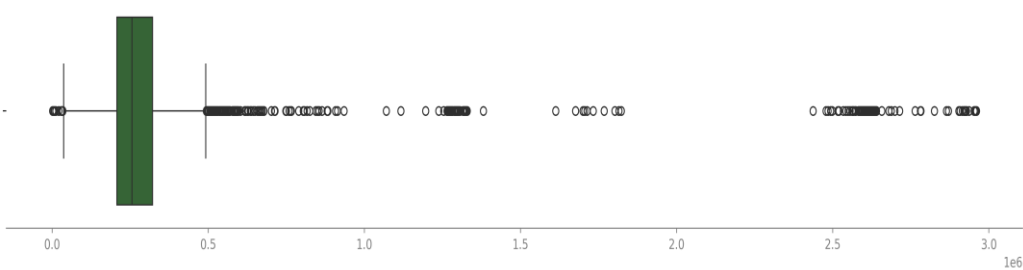
Box Plot of 'Total' Column (mean:5.65, median:3.96)
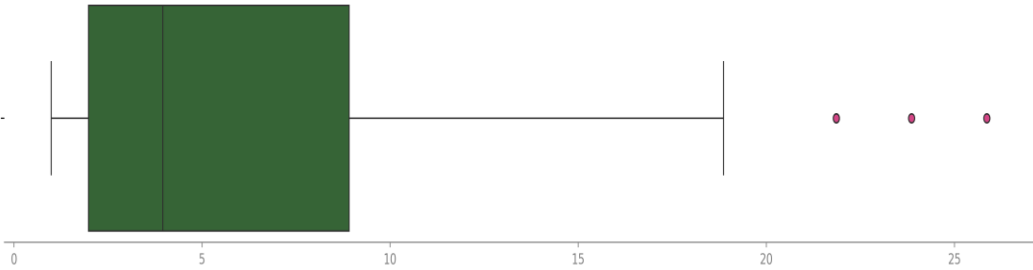
Box Plot of 'Total' Column (mean:5.48, median:3.96)
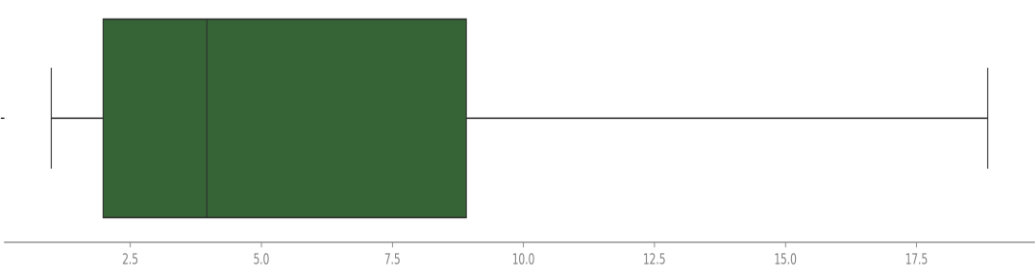
Popular genres

Genre	SumTotalPrice ($)	MeanTotalPrice ($)
Rock	7720.02	9.246
Latin	3472.55	8.996
Metal	2093.13	7.929

Test	Genre	
	Rock – Latin - Metal	Result
Kruskal-Wallis	p=0.00047	Significantly difference

Number of songs purchased by male and female

CleanedGender	TotalBuy
Female	684
Male	1444

- We use people's names to guess their gender.

Dependency between Columns

Scatter Plot of 'Milliseconds' vs 'UnitPrice'
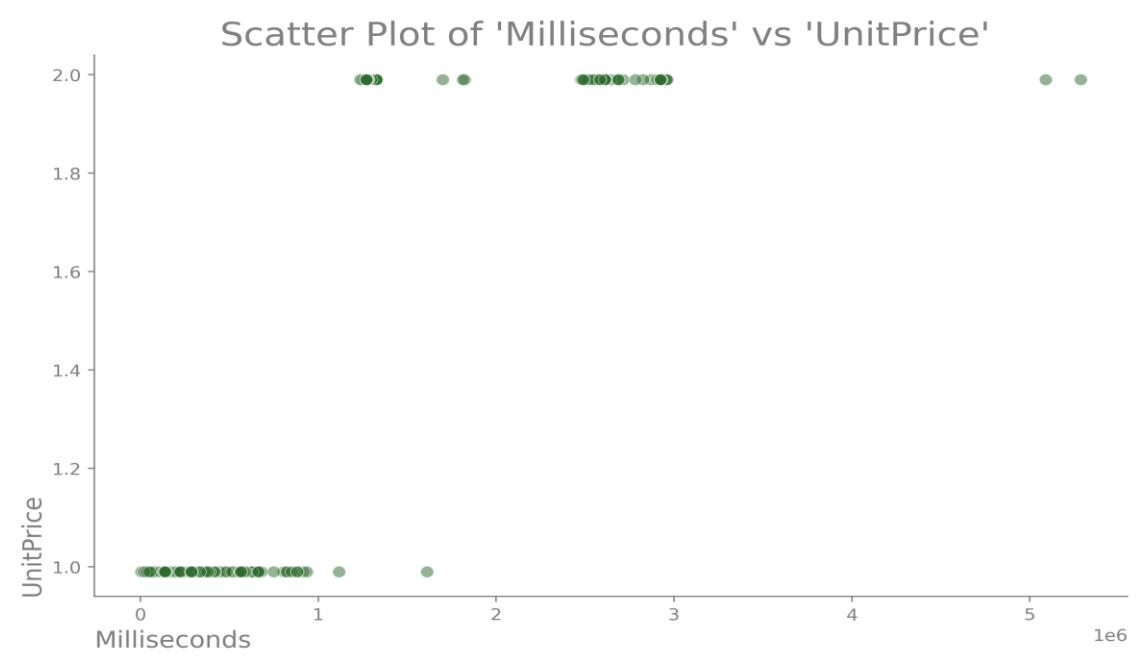
- There is weak dependency between two distributions.

Test		'Milliseconds'	'UnitPrice'
Spearman	'Milliseconds'	1.000000	0.375827
Mutual Information	'Milliseconds'	1.000000	0.2293

CHINOOK MUSIC SERVICE

Dependency between Columns

Test	Categorical Columns		
	Genre - MediaType	Genre - Track	Track - MediaType
Chi2	p= 0.00	p= 0.00	p= 0.00

- They are independent

The average length of songs by genre

Mean Length (Milliseconds) with 95% Confidence Intervals Per Genre
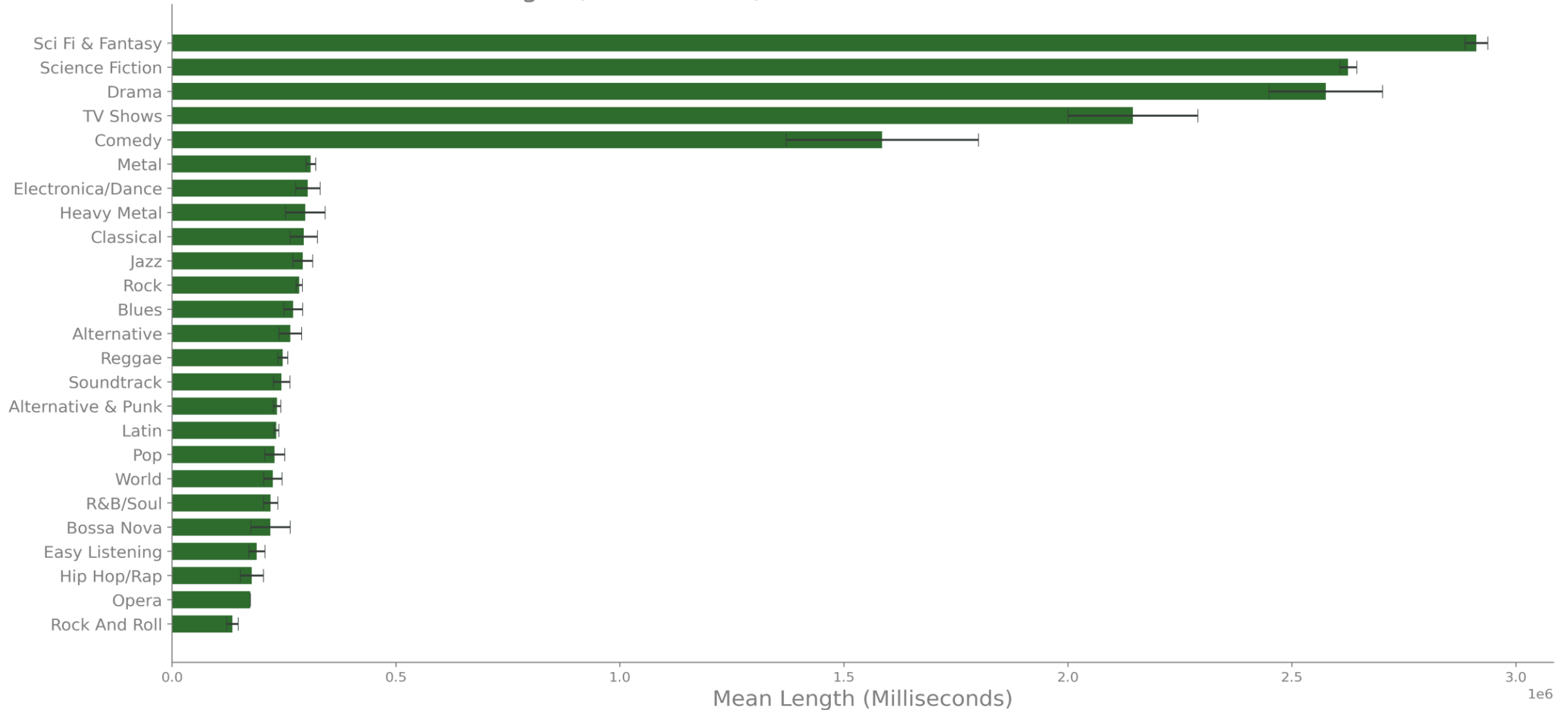
Average sales in different countries

Average Sells with 95% Confidence Intervals Per Country

Country	Value
Chile	6.66$
Hungary	6.52$
Ireland	6.52$
Czech Republic	6.45$
Austria	6.09$
Finland	5.95$
Netherlands	5.80$
India	5.79$
USA	5.75$
Norway	5.66$
Germany	5.59$
France	5.57$
Sweden	5.52$
Portugal	5.52$
Brazil	5.43$
Canada	5.43$
Australia	5.37$
United Kingdom	5.37$
Denmark	5.37$
Italy	5.37$
Argentina	5.37$
Belgium	5.37$
Poland	5.37$
Spain	5.37$

CHINOOK MUSIC SERVICE

The End