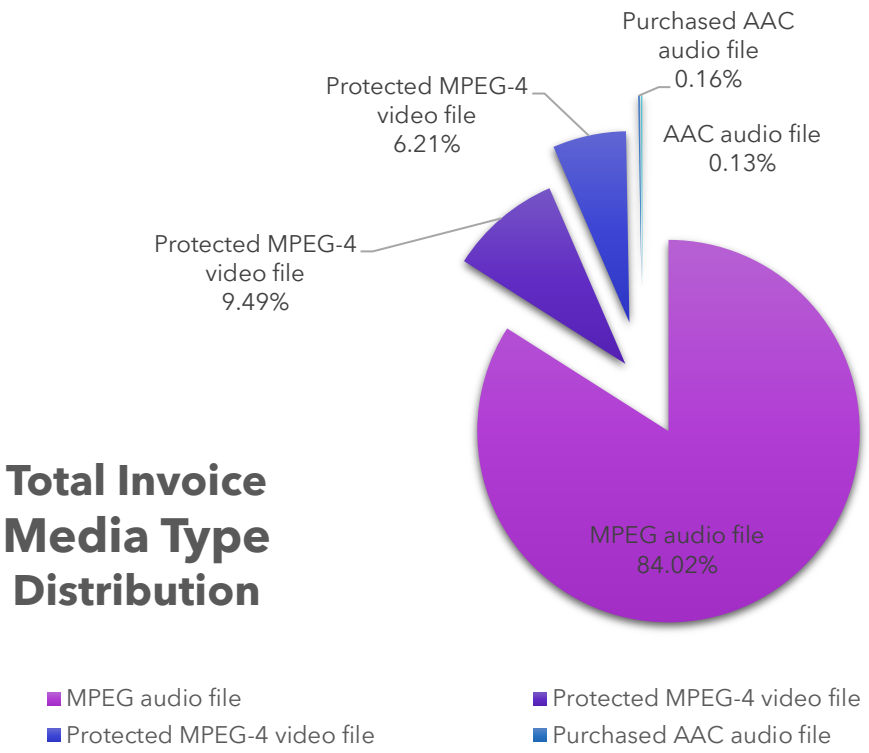
Insights Into a Digital Music Store Database

Project 3 of UDACITY's Business Analytics NanoDegree

By

Hisham D Macaraya

What are the total invoice distribution of each media type?

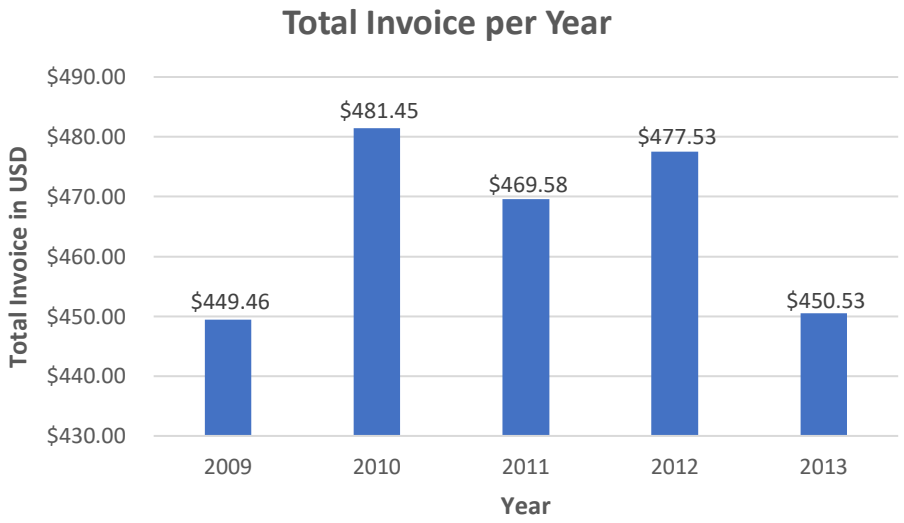


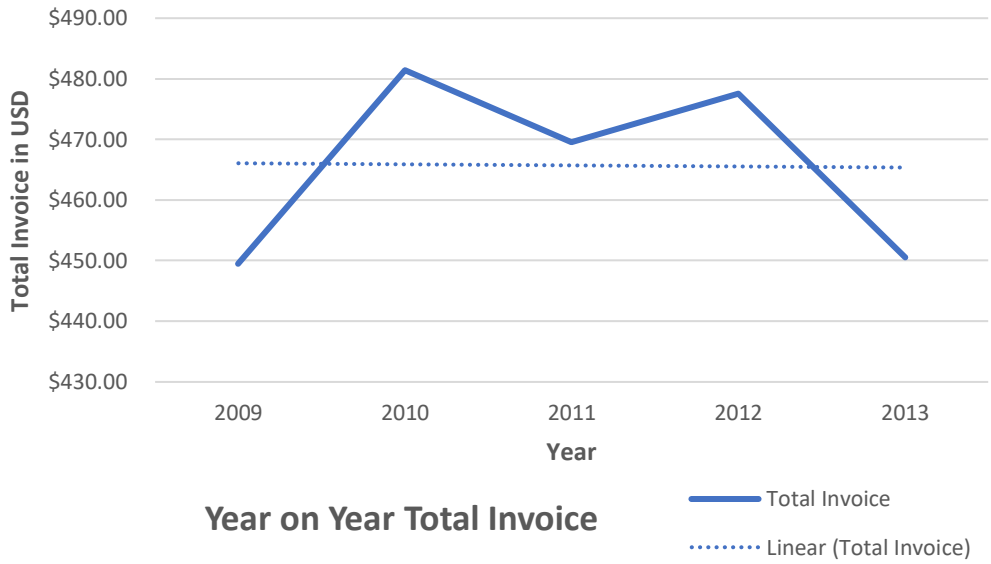
The overall total invoice from 2009 to 2013 is $2,328.30. The distribution of each media type on total invoice are listed on the table below.

The majority (84.02%) of the media type sold are in MPEG audio file format, though it might not be the case in few more years as the digital store started to offer a newer file formats like MPEG-4 and AAC file formats.

Media Type	Total Invoice	Percentage
MPEG audio file	$1,956.24	84.02%
Protected MPEG-4 video file	$220.89	9.49%
Protected MPEG-4 video file	$144.54	6.21%
Purchased AAC audio file	$3.69	0.16%
AAC audio file	$2.94	0.13%
	$2,328.30	100.00%

What is the most profitable and least profitable years of the Digital Music Store? Is there a trend?

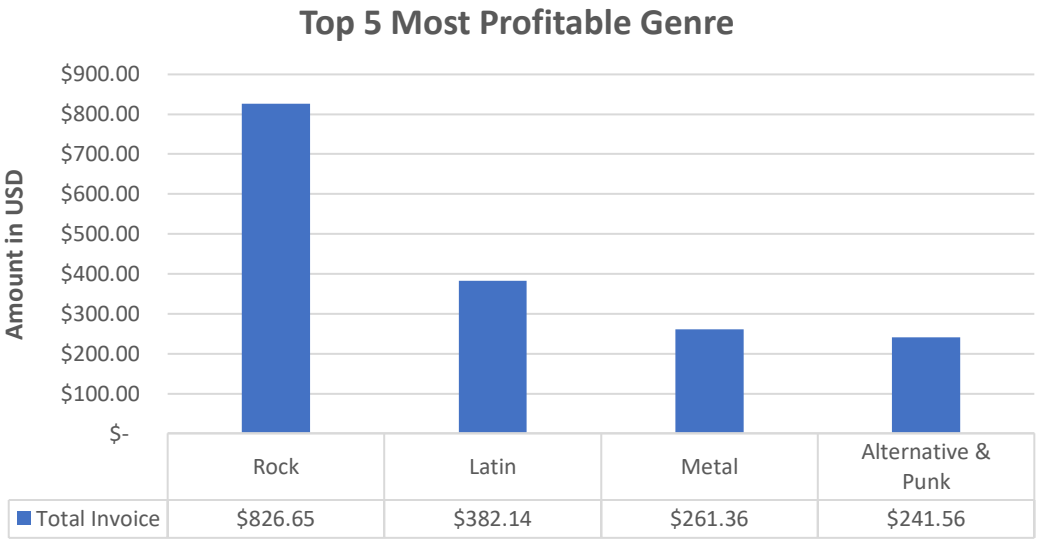


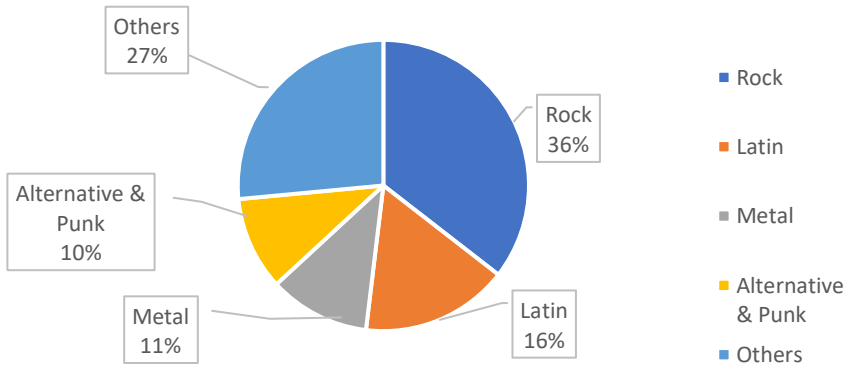


The most profitable year of the Digital Music Store was on its 2nd year, 2010, where the total invoice was $481.45. While the least profitable was on the store's first year, 2009, where the total invoice was $449.46 – which is usually anticipated on any businesses on their early years.

There is a noticeable slight linear downtrend on the year-on-year total invoice, as seen on the right table. This is a bad sign for the store as it means that the store's year on year total invoice are stagnant. The store needs to ramp up marketing and advertising efforts to improve the total sales. We want to see a consistent growth in total invoice of the store for it to thrive.

What are the top 4 most profitable genre sold by the Music Store? What are the total invoice distribution of these top 4 most profitable genre?





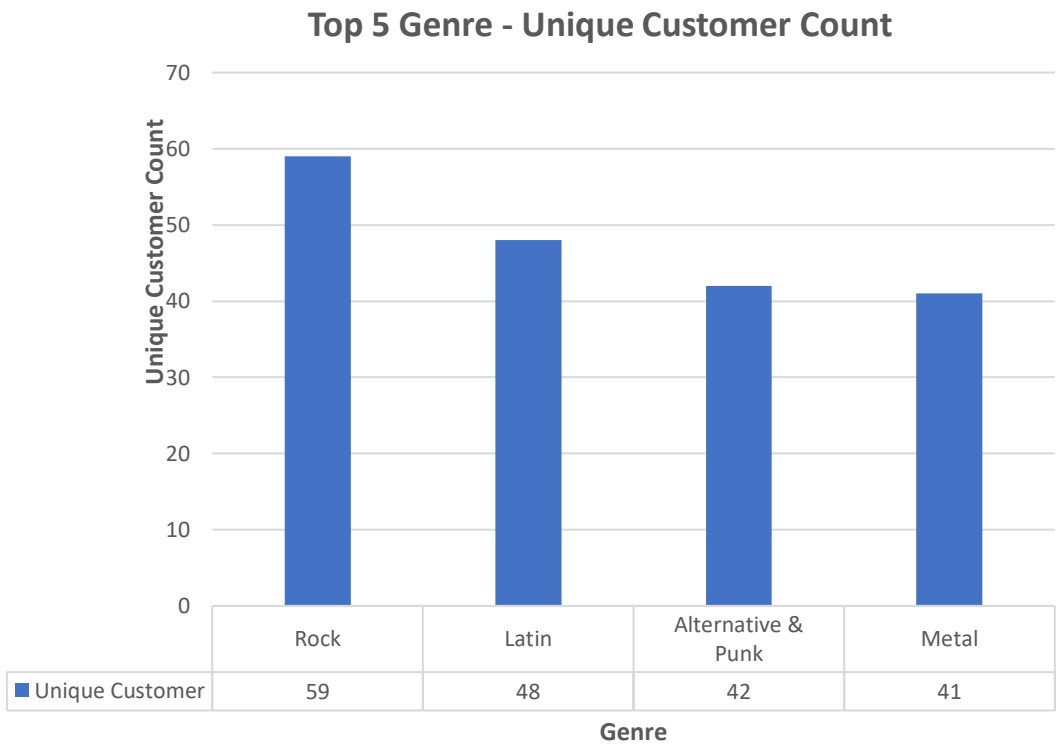
Total Invoice Distribution of Genre

The Music Store's top 4 most profitable genre, out of 25 genre, from 2009 to 2013 are the following:

1. Rock , $826.65 – 36% of Total Invoice
2. Latin, $382.14 – 16% of Total Invoice
3. Metal, $261.36 – 11% of Total Invoice
4. Alternative & Punk, $241.56 – 10% of Total Invoice

Summing it up, the top 4 genre makes up 73% of the Music Store's total invoice. We can therefore advise the Music Store to offer more of these genres.

How many unique customers ordered from each of the top 5 genre?



All customers from the Music Store's database ordered a song in Rock genre. This shows how popular and significant the genre is. Assuming this, the Music Store can further expand the selection of songs in rock genre.

With this data and the data from the pervious slide (top 4 most profitable genre), we can assume that the Music Store can take advantage of this insight by expanding the selections of the top 4 genre: rock, Latin, alternative & punk, metal, and jazz. The Music Store's marketing and advertising effort should also revolve around this 4 genre.

Queries used:

/*Query #1*/

```
SELECT m.name AS media_type, SUM(i.unitprice * i.quantity) AS total_invoice
FROM mediatype AS m
JOIN track AS t
ON m.mediatypeid = t.mediatypeid
JOIN invoiceline AS i
ON t.trackid = i.trackid
JOIN invoice AS inv
ON i.invoiceid = inv.invoiceid
GROUP BY 1
ORDER BY 2 DESC;
```

/*year check for Query 1*/

```
SELECT MIN(DATE_PART('YEAR', invoicedate)) AS start_year,
MAX(DATE_PART('YEAR', invoicedate)) AS latest_year
FROM invoice;
```

/*Query #2*/

```
SELECT (DATE_PART('YEAR', inv.invoicedate)) AS year, SUM(i.unitprice * i.quantity)
AS total_invoice
FROM invoice AS inv
JOIN invoiceline AS i
ON i.invoiceid = inv.invoiceid
JOIN track AS t
ON t.trackid = i.trackid
WHERE inv.invoicedate BETWEEN '2009-01-01' AND '2014-01-01'
GROUP BY 1
ORDER BY 1 DESC;
```

/*Query #3*/

```
SELECT CASE
                WHEN g.name = 'Rock' THEN '1_Rock'
                WHEN g.name = 'Latin' THEN '2_Latin'
                WHEN g.name = 'Metal' THEN '3_Metal'
                WHEN g.name = 'Alternative & Punk' THEN '4_Alternative &
Punk'
                WHEN g.name = 'Jazz' THEN '5_Jazz'
                                ELSE '6_Others' END AS top_genre,
                SUM(inv.total) AS total_invoice
FROM genre AS g
JOIN track AS t
ON g.genreid = t.genreid
JOIN invoiceline AS i
ON t.trackid = i.trackid
JOIN invoice AS inv
ON i.invoicelineid = inv.invoiceid
GROUP BY 1
ORDER BY 1;
```

Queries used:

```
/*Alternative to Query #3, notice that SUM invoice of others genre is included on the first query*/


SELECT g.name, SUM(inv.total) AS total_invoice
FROM genre AS g
JOIN track AS t
ON g.genreid = t.genreid
JOIN invoiceline AS i
ON t.trackid = i.trackid
JOIN invoice AS inv
ON i.invoicelineid = inv.invoiceid
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;


/*Distict Genre Count*/
SELECT COUNT(DISTINCT name)
FROM genre;
```

```
/*Query #4*/


SELECT g.name, COUNT(DISTINCT c.customerid)
FROM genre AS g
JOIN track AS t
ON g.genreid = t.genreid
JOIN invoiceline AS i
ON t.trackid = i.trackid
JOIN invoice AS inv
ON i.invoicelineid = inv.invoiceid
JOIN customer AS c
ON c.customerid = inv.customerid
GROUP BY 1
ORDER BY 2 DESC
LIMIT 5;
```

```
/*Alternative Query for #4*/
SELECT g.name, COUNT(DISTINCT c.customerid)
FROM genre AS g
JOIN track AS t
ON g.genreid = t.genreid
JOIN invoiceline AS i
ON t.trackid = i.trackid
JOIN invoice AS inv
ON i.invoicelineid = inv.invoiceid
JOIN customer AS c
ON c.customerid = inv.customerid
WHERE g.name IN('Rock', 'Latin', 'Metal', 'Alternative & Punk', 'Jazz')
GROUP BY 1
ORDER BY 2 DESC;



                    /*Distict Genre Count*/
SELECT DISTINCT COUNT(customerid)
FROM customer;
```

References:

https://www.udacity.com/