To: Multnomah County Library

From: Chukwuma Ajomiwe-Ukaegbu

30 June 2021

Email chumaxi@yahoo.co.uk

**MULTNOMAH COUNTY LIBRARY: DATA ANALYSIS TO DETERMINE THE RATE OF LATE RETURNS.**

**1.0 Introduction**

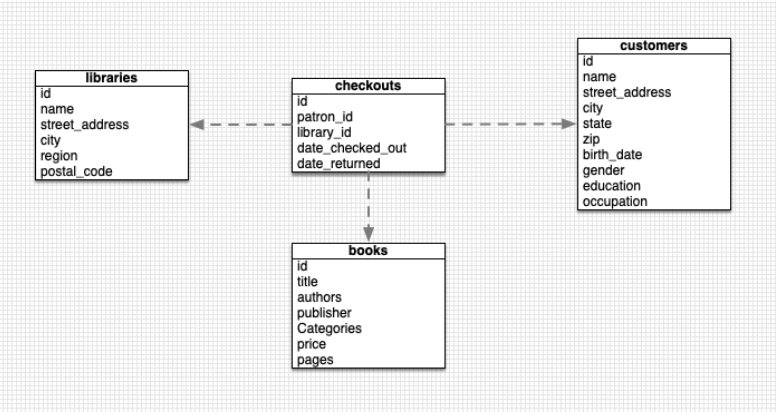
# The Data Analyst has reviewed the various sets of data provided by Multnomah County Library (the Company) in order to give a comprehensive report on the Rate of Late Returns of Library books and provided solutions to mitigate such risks.

**2.0 Engagement Objective**

# The objective of the data analysis is to understand what the rate of late returns is and factors that relate to late returns.

# 3.0 Data Source

The review carried out by the Data Analyst was to use the various tables to arrive at a decision. Data extraction was performed by the client.



* Fig 1.1

# 4.0 Work Done

In order to transform the data obtained to the format required for our analysis, the following procedures were performed to ensure that the quality, format and structure of the data is appropriate.

The structure of the data obtained from the client was in .csv format.

* Data Cleaning
* Data Processing
* Data Visualization

Data Cleaning means the process of identifying the incorrect, incomplete, inaccurate, irrelevant or missing part of the data and then modifying, replacing or deleting them according to the necessity. Data cleaning is considered a foundational element of the basic data science.

Data Processing are the various techniques in the conversion of raw data into a user friendly and readable format.

Data Visualization can be described as the representation of information and data using graphs and charts.

***#Kindly note that the cleaning, processing and visualization processes have been included in the scripts for easy reading and understanding of the processes.***

# 5.0 Result

After completion of our analysis, we were able to determine that the Rate of books being returned later than 28 days after they have been checked out is 33.97%. This means that 1 in every 3 books that has been checked out is going to be returned later than the allocated time frame.

This can be caused as a result of improper data entry of the date format. This could lead to the misrepresentation and interpretation of data and the assertion – **Accuracy.**

# Solution

To combat the issue of improper data entry, it is advised to change from a manual system to an automated system such as the use of a scanner upon Library books being checked out and when they are returned. The use of a scanner and a library ID card would suffice.

Upon checking out of a library book, the book is scanned to keep accurate dates of checkout and the Library ID card too is scanned to keep records of the customer. All data would be kept on a relational database.

An automation would then be created by our team in order to track the length of time the book has been with the user and an automatic reminder would be sent to the users right about the expiry of the allocated time. This would help ensure that the customers are aware and kept updated on the expiry of allocated time frame.

**6.0 Conclusion**

We noted a 33% late return rate from the data provided. Fields such Occupation, Education should be kept distinct in order to provide a more accurate analysis report using those parameters. We can also conclude that customers who have a graduate degree and work in Sales are more likely to hold on to books. See attached below



We were also able to pinpoint that books in the following categories are likely to be returned in the allocated time period;

* Airplanes
* Barrier Island
* History
* Juvenile Nonfiction
* Mechanical engineering
* Physicians

Kindly find attached below the link to understanding the process and the scripts written to arrive at our analysis.

[Data-Analysis/Data Analyst Exercise.ipynb at main · Chukwuma-Ajomiwe/Data-Analysis (github.com)](https://github.com/Chukwuma-Ajomiwe/Data-Analysis/blob/main/Data%20Analyst%20Exercise.ipynb)

Happy to hear from you soon.