Library Management using Analytics and Big Data Technologies

Background:

Showers, B. (2015). *Library analytics and metrics*. London: Facet Publishing.

> "*Libraries, along with archives, museums, and galleries, find themselves ideally placed to exploit the full potential of analytics*."
>
> "*The variety and scope of the data collected and generated by libraries and organizations such as museums and archives is significant: transactional data on catalogue searches, item check-outs, log-ins to online resources and services, swipes through the entrance gates; manually collected statistics on space usage, student satisfaction, external visitors to the library. The applications of the data are equally varied and overlapping, including management functions (collections development and management, usage statistics), impact (demonstrating value, benchmarking, improving learner outcomes) and improving services and meeting user requirements (recommendation services, collections management/development).*"

James Hardiman Library at the National University of Ireland Galway has a functional BI environment, which collects both catalogue and transactional data. However, with limited Analytical tools and reports, the data is under-utilized in producing insights on the routine functioning of the Library and its management.

The following points were identified through our discussions with the stakeholders at the library:

- Books tend to run its course as newer editions are released to the market or newer books come in. There are books that see an increase in demand during certain semesters. Shelf utilisation can be improved and made more efficient based on the seasonal demands of various books
- There are situations where high demand for a book can lead to other books in the same domain to be ignored. This can be countered by means of clustering different *tags* associated with a book to recommend a potential alternative to the user. The clustering of book *tags* may be done based on choices generated from a programme's cohort
- Books are currently loaned for either 1,3 or 81 days. A more dynamic system can be used to gauge future demand for a book and adjust loan days accordingly. This can incur in a book being available for a longer time during lean periods

Objective:

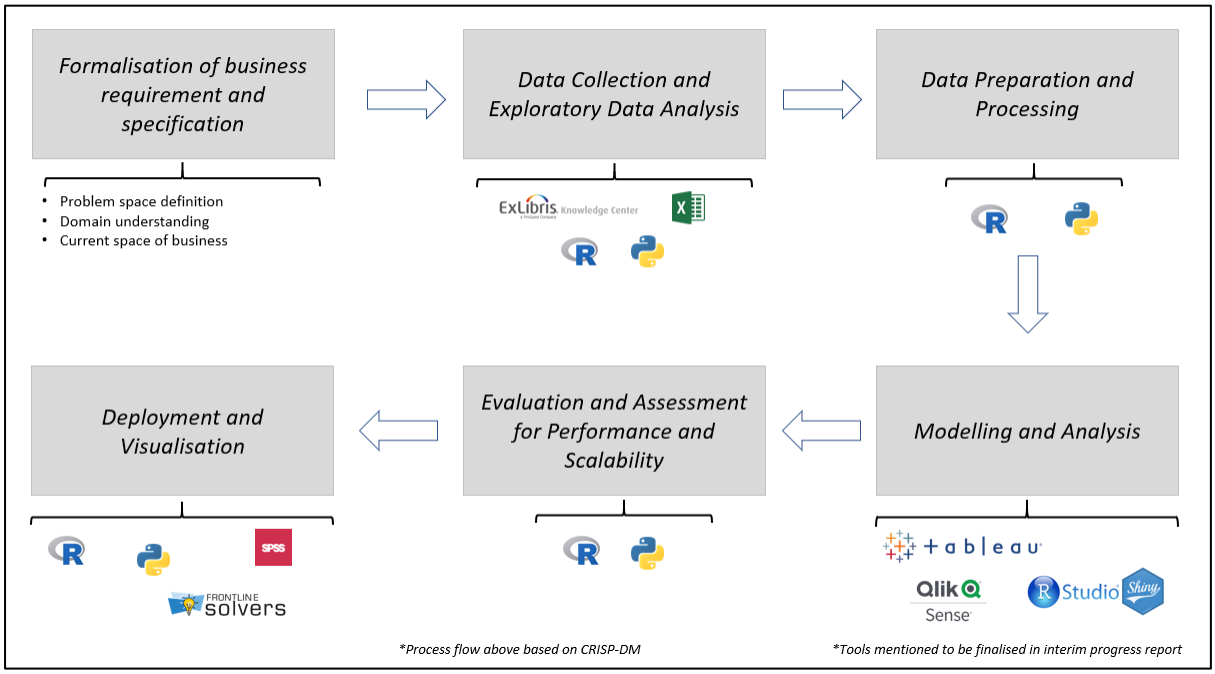
Analyse catalogue and transaction data of a library to gather insights and create recommendations for:

- Categorise books in three different ways-
  - Rarely used
  - Requested a few times last year
  - In-demand books
- Shelving books efficiently based on demand (trends and seasonality)
- Predicting future requirements for books, journals, and articles across different categories
- Prescribing dynamic period for book loans
- Predicting obsolescence of a category/catalogue

Analyse library asset transaction data to gather insights and create recommendations for:

- Online library usage behaviour
- Personalised recommendation services

Tools and Techniques:



Overview of Data:

Expected measures to help achieve project objectives:

ISBN (Identifier)
Actual Loan Period (Number of Days)
Library Loan Period (Number of Days)
Number of Requests (Number)
Number of Renewals (Number)
Early Return? (Categorical)
Recall? (Categorical)
Reading List? (Categorical)
Semester (Categorical)
Number of Books (with same ISBN)

Relevant Variables identified in Dataset:

Tables:

1. **Fulfilment**
2. **Titles**

Variables:

Number of Loans (Inhouse)	Number of Loans (Not in-house)
Returns	Recalls
Renewals	Lost
Claimed returns	Auto-renewals
Reading room at shelf	Reading room at user
WB Change Backward	WB change forward
Bulk change backward	Bulk change forward
Loan days	Return due days
Item ID	Policy Name
Measurement Value	Measurement Unit
Has Active Electronic collections	Has active Physical items
ISBN	Editions
Loan Date	

Stakeholders:

Primary stakeholder: Library staff

Secondary stakeholder(s): University/college students and staff