DIPLOMA: INFORMATION TECHNOLOGY

Subject   : Business Analysis 3 Module 2/ Business Analysis 3.2

Subject code   : AIBUY3A

Assessment   : Module ONE (Team Project)

Lecturer   : Mr. Matsela

Examiner   : Mr. Matsela

External Moderator   : Ms. S Matyila

Due date   :06 September 2023

Table of Contents

PROJECT BACKGROUND

Theme: Integrated Library System

From the theme that we were given we chose to help the community of Vanderbijlpark by creating an Integrated Library System for their library. Our ILS will use four modules which are circulation, cataloging, security & privacy and OPAC. The system will enable patrons to use the self-service machine to check-in and check-out books. ILS will be handling the circulation of library items, returns, holds, renewals, and fines. Members of the library will be able to use facial recognition to get themselves inside the library this will be part of the security and safety for the library. When this system is implemented, it will address critical issues of time and resource management in the library by enhancing efficiency, accurate library data management and user experience of the community library system. The security personnel at the entrance and exit may also profit from it. Our integrated library system (ILS) is a breakthrough tool designed to empower libraries of all sizes in an era where libraries act as crucial centers of knowledge and information.

Main Objective

We want to develop an ILS for the Vanderbijlpark public library to enhance the efficiency of library operations:

- Increase operational efficiency:
  The system will automate various tasks, reducing manual work and enabling library staff to focus on providing better service and assisting patrons.

- Data management:
  Our integrated library system will centralize data related to library collection, patrons and circulation allowing libraries to maintain accurate records and generate useful reports.

- Resources organization
  It will help organize and catalogue library materials in a structured manner, making it easier for patrons to discover and access items.

- Enhanced User Experience:
  Provide patrons with better search and discovery tools, personalized recommendations, and easy-to-use interfaces.

- Streamlined Circulation:
  It simplifies circulation processes, such as checking items in and out, renewals, and managing holds, resulting in smoother and faster transactions.

PROBLEMS

South African public libraries have grappled with numerous challenges. Addressing these issues is crucial to fostering equitable access to knowledge and enabling libraries to evolve in a rapidly changing landscape.

➢ Staffing Challenges

Due to limited budgets community libraries have, the library often struggles with hiring enough qualified and motivated staff.

➢ Lack of Relevant Content

The library's inability to engage with the patrons results in having material/Books irrelevant or having too many books that are no longer needed on shelves.

➢ Language and Cultural Relevance

Libraries do not cater to the linguistic and cultural diversity of the community.

➢ Reliability and security

Library users can easily steal or leave without verifying / confirming that they are borrowing the book, trying to avoid long queues.

➢ Decreased Efficiency

Library staff are unable to keep proper records of all visitors, Books received, Books needed, Books that are supposed to leave the shelves and the ones borrowed because of manually details entering, of which will result in loss of data.

➢ Time Consuming

Patrons spent time queuing when entering, borrowing, and returning a book. Most of the library staff spend time recording data daily therefore limiting their active hours.
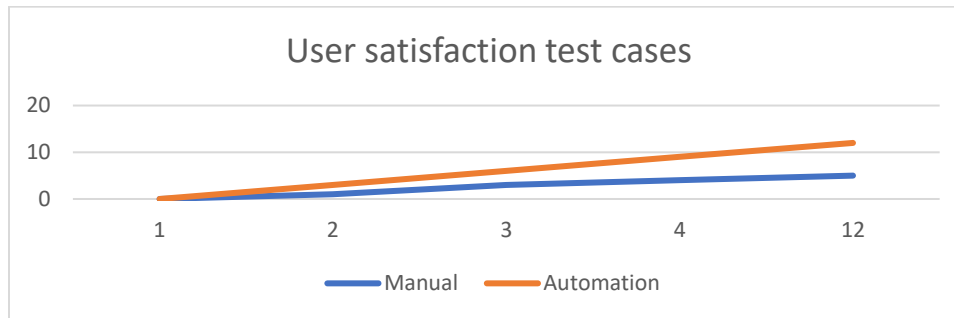
BENEFITS

Applying AI into a virtual library assistant can bring significant benefits to patrons, staff, and the library. Here is how:

- The system improves reliability and security in libraries through predictive maintenance, anomaly detection, access control, and data security measures. It also aids in optimizing inventory, user authentication, fraud detection, and resource allocation while considering ethical implications and ongoing maintenance.
- It is able manage a significant volume of interactions, reducing the need for additional staff resources, especially during peak times.
- Insights from System can inform strategic decisions, resource allocation, and service improvements, leading to a more data-informed library operation.
- Everyday duties like responding to simple inquiries and supplying information to the staff are assigned to the virtual assistant (Self-Service Kiosk), The use of our system enables staff to concentrate on more intricate responsibilities and interactions.
- Faster interaction thus reducing the time spent queuing.

Business Success Criteria

- **User satisfaction**

  According to the number of cases tested, it shows that the system has an intuitive   and user-friendly interface that patrons find easier to use or interact.

- **Scalability and efficiency**

  The system can handle growing amounts of data and users, which will create accurate records, easy search, and retrieval of resources.

- **Cost effectiveness**

  With the system operating faster and managing too much work compared to manual operation, the number of staff needed is gradually thus saving the library from costs.

- **Easy integration**

   It is easily integrated with the old system for seamless operation.

Requirements, Constraints and Risks

Requirements

- Efficient circulation system for check-in and check-out, holds, and renewals.
- Analytics and reporting capabilities for data-driven decision-making.
- Search and discovery interface accessible all the time.
- Patron management system for user registration, preferences, and history.

Constraints

- Budget

  The system must be completed within a budget that is suitable for the community and approved by the municipality/ library.

- Integration

  The new system must integrate seamlessly with the existing system.

- Time

  Limited time frame to implement the system to minimize disrupting the library.
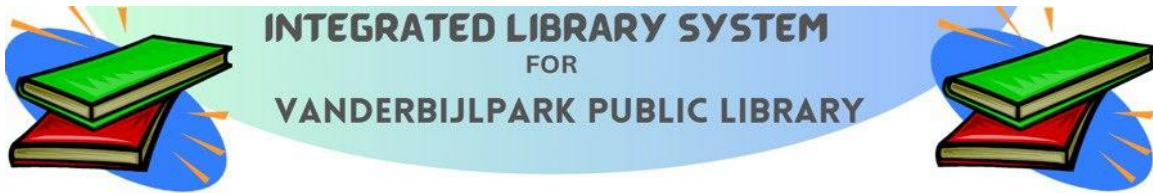
Risks

- Transferring data from an old system to the new integrated library system without errors is critical.
- Patrons and library staff being unable to adapt to the new system.
- Technical challenges during integration with legacy systems.
- Data migration risks leading to loss or corruption of information.

<u>Initial Assessment of Tools and Techniques</u>

- Data migration Tools
  - ➤ Use data migration tools to ensure a smooth transition of existing data from the old system to the new integrated system.

- Training and Change Management
  - ➤ Provide comprehensive training for staff and create change management strategies to ease the transition process.

- Authentication and Security
  - ➤ Our biometric authentication protects patron data and ensures authorized access to the system.

- Data Analytics Tools
  - ➤ Implement tools for collecting and analyzing usage data, patron behavior, popular items, and resource usage patterns to make informed decisions.

- Discovery and Search Tools
  - ➤ Implement a search and discovery layer that allows patrons to easily find physical and digital resources within the library's collection.

Poster

THEORETICAL ASPECT:

Machine Learning Approach

- ➢ This is a branch of artificial intelligence that empowers computers to learn patterns in the data and make decisions without physical programming.

Types of Machine Learning

Within machine learning, there are three categories to base our system on,

They are:

1. Supervised machine learning
2. Unsupervised machine learning
3. Reinforced machine learning

- ➢ With that explained, within these kinds of machine learning approaches, we chose to go with the supervised machine learning approach. This is because we intend to have our system use classification and regression to carry out its duties. For example: We will use classification to group books in genres based on their content, then regression to predict books that the user is likely to read.
- ➢ The system will use classification to group books together in a sort of relevance. It divides the population or data points into various groups to tackle grouping issues.

Data:

Integrated library system (ILS) is an enterprise resource planning system for a library, used to track items owned, orders made, bills paid and patrons who have borrowed books. It is used to increase operational efficiency, provide access to a library's collection, and provide access to external resources. This is how data will be processed in the system:

Step1: Start with Exploratory Data Analysis

- ➢ To provide valuable insights into various aspects of library operations and patron behavior and acquaint the ILS with the dataset.
- ➢ Introduce the system to the dataset.

Step2: Data cleansing

- ➢ We know that library data comes from different sources, the system will utilize automated software tools that are designed for data cleansing, these tools will identify patterns of noise, such as incorrect or inconsistent data and provide suggestions for correction, the system will use pandas which will also help in rectifying issues like missing data, duplicates, and errors.

Step3: Data transformation

- ➢ Data transformation in terms of aggregation, it will combine and summarize data to create more meaningful and insightful presentations, for instance it will generate reports on popular genres, most borrowed books, or library usage patterns. Aggregation data will also help the library staff and patrons have a clearer understanding of the library's collection, user behavior and the overall impact of the library.
- ➢ In terms of discretization, a large amount of data will be converted into smaller values to make them easier to analyze and manage, for an example converting

continuous publication years into decades intervals e.g. (2000-2009,2010-
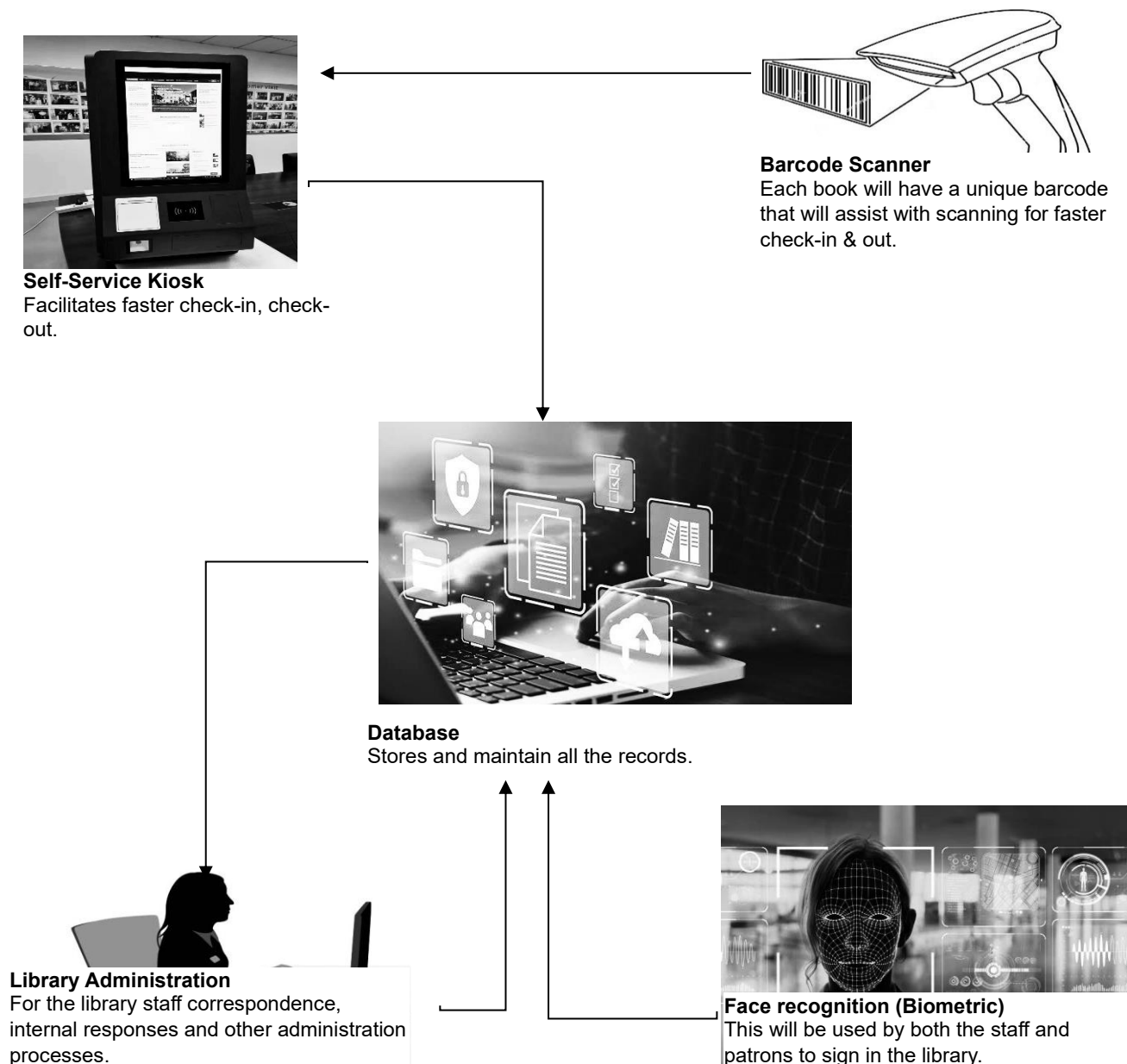2019….) to analyze trends in the library's collection time.

➤ In terms of data smoothing, algorithms to remove distorted and meaningless data
from the dataset will be used to make it easier for staff and the system to spot
clear patterns in the dataset and reduce noise or variability in data.

Step4: Data reduction

➤ It will summarize user borrowing history to understand the types of resources
users prefer over time.

➤ Aggregate data by combining multiple records into summary statistics, for
instance aggregate circulation data by month or year to analyze trends over time.

➤ Feature selection, choose a subset of relevant features or attributes from your
data while discarding less important ones, for example select the most important
features related to book characteristics for analysis, while omitting less impactful
features.

➤ Clustering and grouping, use clustering algorithms to group similar resources or
users, allowing you to work with summary data for each cluster instead of
individuals.

➤ Apply data cube aggression to resource usage data to identify popular resources,
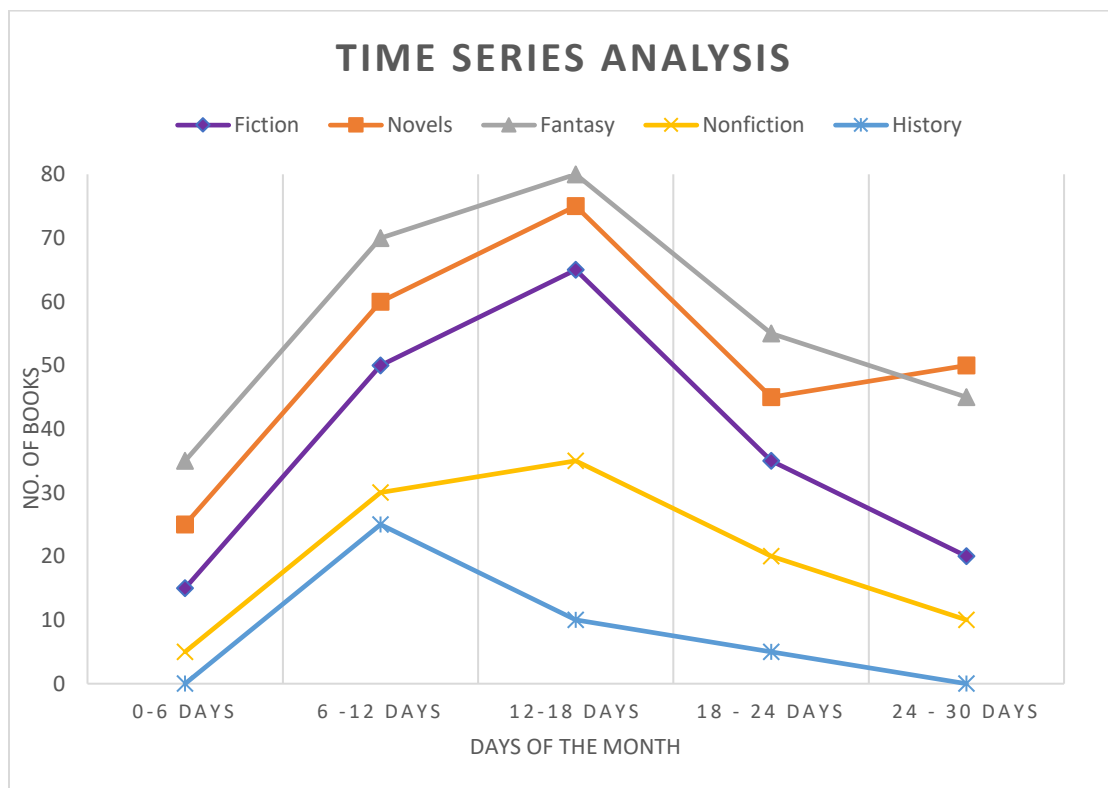genres, and formats among patrons.

MODEL

Related to the diagram below, we are using both Classification and regression model to categorizes library items or activities into predefined groups, like sorting books into genres or classifying patrons by membership status, utilizing algorithms such as decision trees or support vector machines, system predicts numerical outcomes, such as estimating library visitor counts or budget forecasts, using techniques like linear regression.

Barcode Scanner
Each book will have a unique barcode that will assist with scanning for faster check-in & out.

Self-Service Kiosk
Facilitates faster check-in, check-out.

Database
Stores and maintain all the records.

Library Administration
For the library staff correspondence, internal responses and other administration processes.

Face recognition (Biometric)
This will be used by both the staff and patrons to sign in the library.

Time Series Analysis on Data

In the time series below, we are analyzing the number of people who go to the library and the hours they spend. We can observe that when the library opens, there is less facial detection in the library and the peak hours are between 10 and 12 am.

TIME SERIES ANALYSIS

Fiction — Novels — Fantasy — Nonfiction — History

NO. OF BOOKS

DAYS OF THE MONTH

0-6 DAYS | 6 -12 DAYS | 12-18 DAYS | 18 - 24 DAYS | 24 - 30 DAYS

Solution Techniques

As the background has explained, the system is going to help our public library enormously. The challenges faced by the library are hiring the right staff and time spent performing a single task because of the current system. There are several solution techniques and components involved in designing and implementing an effective Integrated Library System.

The two that our system will be focusing on are:

➢ Circulation and Patron Management
➢ Online Public Access Catalog (OPAC)

With the circulation solution patrons are going to able to sign themselves into the library without the help of the staff using the facial recognition that will be implemented. Since the library is public and free the least that we can do as a community is to keep it safe.

A publicly accessible data collection from the library in the form of books, authors, genres, is what we call an OPAC, and it is found online. With these patrons can login to the system and search for what they want, for example books, journals and more. This will provide users with item availability and library holdings, also the location of specific items.

Natural Language Processing, Speech Recognition or Speech Synthesis

1. NLP: Natural Language Processing

Relevance to Theme: NLP deals with how computers and human language interact. It can be applied to a variety of tasks, including sentiment analysis, text classification, linguistic translation and chatbots. NLP may be very pertinent to your issue if it entails processing and comprehending textual data. Achievability: NLP technology has improved greatly in recent years, and developers may now construct NLP solutions using a variety of libraries, frameworks, and APIs. Many NLP applications are fairly feasible, even though more difficult jobs like sophisticated machine translation may necessitate significant resources.

2. Speech Recognition:

Relevance to Theme: Speech recognition software transforms spoken words into text. Speech recognition may be particularly pertinent to your theme if it includes recording spoken information and putting it into a format that may be used. It is utilized by voice assistants, transcribing services, and other systems. Achievability: Software development kits (SDKs) and APIs are extensively used to access speech recognition technology, which has advanced significantly. Although rigorous training and adaptation to certain accents and languages may be necessary to achieve high accuracy, basic speech-to-text functionality is doable for a range of applications.

Deep Learning Models

1. Neural Network Structure

➢ To accomplish circulation, cataloguing, security and privacy and OPAC, our ILS data was collected through image data (barcodes).

➢ We used Convolutional Neural Networks, CNN to be able to leverage its feature extraction capabilities for image data.

➢ CNN can learn features directly from raw data.

➢ We constructed the CNN with four layers; an Input Layer for pre-processed data, Convolutional Layer for feature extraction, a ReLU Layer for introducing non-linearity and a fully connected layer for accurate predictions.

➢ We implemented our CNN using Microsoft Azure Machine Learning to allow users to run multiple training processes with different data at once.

1. Input Layer:

➢ For image data, this layer receives the raw pixel values, image size and categorizes the objects of the image based on the three color channels, Red, Green and Blue.

➢ The input data is then passed to the convolutional network.

2. Convolutional Layer:

➢ The convolutional layer is the core building block of CNN responsible for feature extraction of the input image and the size of the image is reduced through pooling techniques to avoid overfitting.

➢ This layer has a set of learnable filters that are responsible for detecting various features in the input data like edges, corners, and patterns.

➢ The type of features extracted is moderated by adjusting the number of filters, size and strides in this layer.

3. Rectified Linear Unit Layer:

➢ The ReLU activation function follows the convolutional layer to introduce nonlinearity in the hidden layers of the neural network.

➢ It transfers positive values of the precious layer to the next unchanged and sends the negative and zero values in the feature maps to zero.

➢ It helps the neural network to learn complex patterns, improving feature extraction.

3. Fully Connected Layer:

> ➤ This is the last layer of CNN which shows the desired network's output of the image classification tasks.
> ➤ Desired feature representation is flattened and fed to a fully connected neural network.

References

Zaher Talab (2021) 'Machine Learning vs Deep Learning', at https://www.developer.com/guides/machine-learning-vs-deep-learning/ (Accessed:03 August 2023)

Rajat Sharma 'ML | Natural Language Processing using Deep Learning' at https://www.geeksforgeeks.org/ml-natural-language-processing-using-deep-learning/?ref=ml_lbp (Accessed:27 August 2023)

ABSR 'Classification vs Regression ' at https://www.geeksforgeeks.org/ml-natural-language-processing-using-deep-learning/?ref=ml_lbp (Accessed:27 August 2023)

MK Gurucharan 'Basic CNN Architecture' at https://www.upgrad.com/blog/basic-cnn-architecture/

'Integrated Library System-Wikipedia' at https://en.wikipedia.org/wiki/integrated_library_system (accessed:23 August 2023)

GRAMMALLY REPORT

Report: CodeCrafters Business Analysis Project-1

CodeCrafters Business Analysis Project-1

by Thabiso Sydney

General metrics

17,997	2,648	231	10 min 35 sec	20 min 22 sec
characters	words	sentences	reading time	speaking time

Score

84

This text scores better than 84% of all texts checked by Grammarly

Writing Issues

114	✓	**114**
Issues left	Critical	Advanced

Unique Words

Measures vocabulary diversity by calculating the percentage of words used only once in your document

31%

unique words

grammarly

Report: CodeCrafters Business Analysis Project-1

Rare Words

Measures depth of vocabulary by identifying words that are not among the 5,000 most common English words.

43%

rare words

Word Length

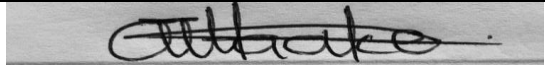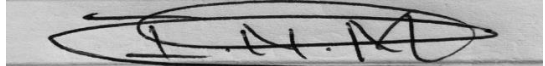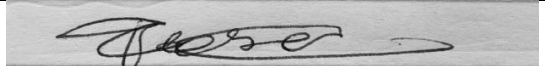Measures average word length
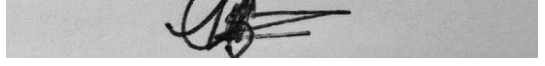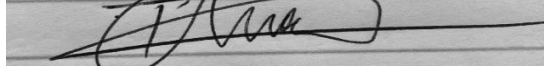
5.5

characters per word

Sentence Length

Measures average sentence length

11.5

words per sentence

DECLARATION

We declare that this assignment is an original work submitted by the following group members who have all actively contributed. Any other work of a similar nature has been appropriately referenced in this assignment.

Surname and Initials	Student Number	Signature
Maake TS	218590822	
Monyeseala I	221768963	
Lebeko T	221686274	
Mofokeng T	221025138	
Sereme TH	221442375	
Tshavhani NG	221347801	
Mafanywa V	221138846	
Marunga TJ	221766960	