Central Information System – St. Clair

(Data Analytics for Business)

Business Problem

Data Analytics for Business program at St Clair College is facing a challenge in effectively analyzing student trends and sentiments due to a **lack of a centralized system for collecting and managing student data**. This lack of organization leads to a lack of visibility into student issues, which may result in a lack of satisfaction among students.

The project aims to improve the student experience by creating a centralized system for collecting and utilizing student data, as well as a mechanism to monitor and understand student needs and their problems, including challenges related to accommodation, transportation, educational support, the college schedule, health services, and other matters.

Goals

The main objective is to **collect and conduct a thorough analysis of data pertaining to all current students enrolled in the program**. This analysis will be used to identify key issues and potential solutions, which will be further presented through an interactive dashboard for management review.

Additionally, the data will be **analyzed using machine learning techniques to gain valuable insights and predictions**. A final goal is to create a template that can be used by other departments for similar data collection and analysis.

Stakeholders and Impact

- **Management** will benefit from a streamlined and comprehensive solution for understanding student's data and concerns.
- **Students** will have access to a platform for raising and addressing their concerns.
- **Project team** will have the opportunity to work on a practical problem and provide an effective solution.

Intended Insights

The system intends to provide solutions/answers to the below mentioned.

- What is the current **student enrollment** for each semester?
- What is the **demographic** information for students?
- What is the breakdown of education and work **background** for students?
- What is the distribution of student housing **accommodations**?
- Do students **work**, and if so, where?

Group 12
Amit Sharma	Suhail Ahmed	Rajvi Mehta	Harshil Patel	Jayraj Radadiya
0794488	0789949	0788372	0791261	0789984

- How are **college services** being utilized by students?
- What is the impact of **academics** on students?
- What is the overall **sentiment** of students towards the college?

Success Criteria

Measuring the success of a descriptive and exploratory dashboard project can be done through.

- **Data Accuracy**: Ensure that the data displayed on the dashboard is accurate and up to date.
- **Feedback**: Collect feedback from users on the dashboard's design, usability, and usefulness.
- **Adoption Rate**: Measure the number of users who adopt the dashboard as their primary tool for data analysis and decision-making.
- **User Engagement**: Track the number of students participating and provide insightful data.
- **Trends and Analysis**: Presence of some trends and connections between concerns and services

It is important to note that measuring success will be specific to the goals and objectives of the project, and will vary depending on the context, audience and data being displayed.

Methods

Collection and Analysis of data will be conducted in a step-by-step approach.

Phase 1: Primary Data collection

- Collection of Primary Data set through Survey
- Setting up the Pipeline and initial data landing zone
- Data engineering and EDA through Transformation engine or Python
- Initial Phase 1 Dashboard through Tableau.

Phase 2: Secondary Data work

- Collection of Secondary Data sets through Survey and linked data
- Creation of a mapping mechanism for primary and secondary sets
- Phase 2 Dashboard through Tableau.

Phase 3: Machine Learning and Analysis

- Use of Machine Learning (NLP) on entire data set for Sentimental analysis

Skills

The following skills are required for this project:

- Knowledge of the student's specific data domain
- Proficiency in Python & Machine Learning, specifically natural language processing (NLP)

Group 12
| Amit Sharma	Suhail Ahmed	Rajvi Mehta	Harshil Patel	Jayraj Radadiya
0794488	0789949	0788372	0791261	0789984

- Experience with Tableau for data visualization
- Familiarity with cloud-based technologies for creating Data Pipelines
- Understanding of data ethics and best practices
- Experience in conducting surveys and collecting data.

Data Sources

- Primary set: Data collected as a part of Phase 1 survey.
- Secondary set:
  - Data collected as a part of Phase 2 survey.
  - International Department Student Data
  - Data Related to Services

Pipeline

A data pipeline, typically involve the ingestion or extraction of data from one or more sources **(Surveys and International department data)**, the transformation of that data to fit the needs of downstream systems **(Data Engineering/EDA using Python or Cloud Tool)**, and the loading of the data into a destination for storage and/or analysis **(Tableau and Machine Learning)**.
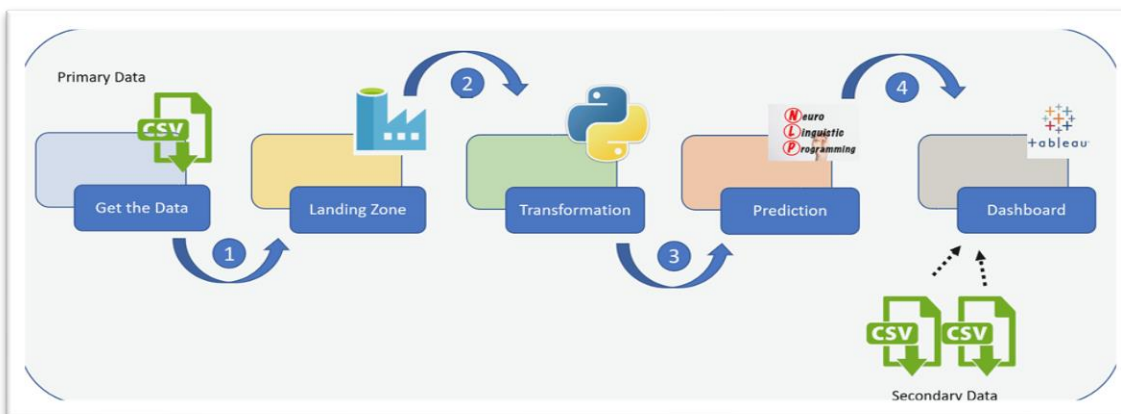
References

- **Tableau**: Tableau Public, https://www.tableau.com/learn, Datacamp
- **Python**: Data_science_life_hacks (linkedin.com), https://keras.io, https://numpy.org, NLP - Datacamp
- **Project Management:** Project management: A priceless skill (linkedin.com)
- **Data Ethics:** Content from semester 3, Ethics for analytics course – DAB 302
- **Student Management:** https://www.stclaircollege.ca/sites/default/files/policies/scc-policy-7-1.pdf

Group 12

Amit Sharma	Suhail Ahmed	Rajvi Mehta	Harshil Patel	Jayraj Radadiya
0794488	0789949	0788372	0791261	0789984