

# **Analysis of – Zekelman School of IT**

## **(All Programs of IT)**

**Group 1:** Archit Navadiya(0815897), Dhruvil Patel(0813662), Bhumi Navadiya(0812548), Felix Saji(0803791), Arbaz Shaikh(0823127)

### **Data Sources:**

This dataset will be collected by means of an online survey, which focuses on gathering information related to their current situation, concerns, and requirements of students with respect to the course that they are enrolled in. This is self-reported data with scope for response bias.

**Survey Categories:** The survey will be categorized into three main divisions:

- Demographics – to better understand the population characteristics.
- Convenience of living – to understand the housing and rental situation.
- Employment and Scholastics – to understand the job and educational opportunities that the students are either aware of or looking for with respect to their current course.

**Link for the online survey:** <https://forms.office.com/r/Kvz0yW41Pu>

**Data Collection Process:** Data collection will commence after a comprehensive review of the survey questions by the Project Coordinator and all program coordinators of the Zekelman School of IT.

To ensure that the data collected is accurate and useful, proper cleaning and filtering of data will be the initial responsibility after data collection.

- **Pre-processing:** First, we will gather raw data in Microsoft Excel. After that, we are going to make a data pipeline that will automatically fetch data from Excel to the Google Cloud platform with the help of a big query. Then we will sort our data, according to programs in Big Query, and convert it into a functional format.
- **Data Cleaning:** In the data cleaning we will convert answers given by students based on category. We will remove the outlier and fill the null value with the help of feature engineering.
- **Data Transformation:** Accurate analysis requires transforming data into a suitable format. Techniques like normalization, standardization, and scaling are employed. Normalization ensures data falls within a specific range, while standardization scales data to have a mean of zero and a standard deviation of one.
- **Data Encoding:** To analyze the data effectively, categorical variables are transformed into a numerical format using encoding techniques. This step is crucial for meaningful analysis.

## Initial Data Overview and Dataset Features:

First, the data will be input into a Raw Data table, which is filled directly from the survey data sheets. Subsequently, this data would be transferred to a Data table, where appropriate column names and data types would be assigned.

Initially, our data had various features which are around 20 based on the answers provided by the student. There are some questions that only pop up when students select specific options in the survey.

- The **Academic** table includes columns for intake and current semester to monitor the student's progress in the enrolled course and any prior experience.
- **Demographics** for collecting demographic data, aiding in understanding the student population and diversity.
- **The accommodation** part contains what kind of rooms they prefer to live in, rent, and a few questions based on the Likert scale.
- **Employment** sections ask the student about their average income and kind of employment status.
- **Academic Feedback** contains five columns enabling students to rate their scholastic experience.
- **Services** focus on college services, gathering data on 3 main points *awareness*, *usage*, and *satisfaction* to pinpoint areas needing enhancement.
- **Feedback** constitutes a specific dataset wherein students can provide feedback regarding their experiences. This information is valuable for analysis, identifying areas for enhancement, and making informed decisions to enhance the student experience based on the data collected.

## Ethical Assessment:

When collecting and analyzing student experience survey data, ethical considerations play a significant role. These concerns are spread in several key areas:

**Informed Consent:** Securing informed consent from the student participants is vital. This involves providing clear details about the survey's purpose, data usage, and potential risks or benefits, ensuring participants understand before their data is collected.

**Confidentiality:** Respecting participants' privacy rights is essential to ethical integrity. Personal information must be kept confidential, and responses anonymized to safeguard participant identities.

**Bias:** Avoiding biases in survey design and administration is crucial. The survey must be neutral, unbiased, clear, and unambiguous questions to ensure accurate and reliable data is at the disposal for analysis.

**Data Usage:** Collected data should only be used for its intended purpose. Any additional usage must be communicated clearly to participants and align with their original consent. In the event that there is a change in usage of data, the student participants must be informed about the same.

**Data Ownership:** Participants should be informed about data ownership and usage rights. Transparency regarding who owns the collected data and its intended use is essential.

**Data Security:** Given the potential sensitivity of survey data, especially demographic information or personal experiences, robust data security measures are vital. These measures ensure secure storage and prevent unauthorized access, disclosure, or data leaks.

## **Future data needs**

**The second phase of the Survey:** More elaborated data will be collected related to college services and student issues. This data will be needed for further analysis regarding issues and services.

## **Next Steps:**

The form will be circulated amongst the targeted student audience via means of college email.

Coordination with the program coordinators for Zekelman School of Business to spread the word regarding the survey and its importance.

Visiting the students (target audience) during class hours to ensure in-person announcements are made to capture more engagement since emails and announcements can be sometimes overlooked.

## **References:**

- Documentation of the initial project proposal and implementation for guidance.
- Google for reference regarding ethical data collection.