ENDTERM PROJECT:

TOPIC PROPOSAL

By TEAM 7

HAEYEON JEONG

SAI RACHANA KANDIKATTU

ABIRHAM GETIE

YONATHAM SHIMELIS

Introduction to Data Mining: DS 6103 Under the guidance of,

Prof. Divya Pandove

(TA: Bhumika Mallikarjunhorapet)

# RESEARCH TOPIC:

**EXPLORING AND MODELLING DATA ON MENTAL HEALTH DATA**

Nowadays, mental health issues are increasingly apparent in various settings. Our research will conduct exploratory data analysis (EDA) and modeling on mental health data using different classification models. We aim to investigate how factors such as gender, family history, treatment, top occupations with rising mental stress, and days spent indoors influence mental health outcomes. By employing various models, we seek to uncover insights that can inform interventions and improve our understanding of these critical issues.

# SMART QUESTION

“How do gender, family history, treatment status, top occupations, days spent indoors collectively impact associated with growing mental stress, as assessed by various classification models such as Logistic Regression, SVM, KNN, Random Forest, and Ridge Regression, during the years 2014 to 2016?

# SOURCE OF THE DATASET (NO. OF OBS)

The dataset is taken from Kaggle ([link](https://www.kaggle.com/datasets/bhavikjikadara/mental-health-dataset)). There are 292364 observations of 17 features in the taken dataset.

**GIT REPOSITORY LINK:** [**github**](https://github.com/sairachanak/6103-Team7.git)