EAS 509: Statistical Learning II
Group Project Proposal

GROUP 12:

Kishan Patel (kpatel39)
Geethika Bedadhala (gbedadha)
Ish Kumar (ishkumar)

TITLE: *Employee Turnover Prediction*

DESCRIPTION:

We will use the below dataset to predict an employee's risk of quitting with a Survival Analysis model. Usually, Logistic Regression is used for these types of situations, but we get accurate results only for short-term situations where the employee has worked for more or less than three months. It is better to use Survival analysis to predict individual quitting risks.

The dataset contains 16 columns like profession, gender, age, experience, etc.

IMPORTANCE:

The dataset we are working on is a real dataset shared from Edward Babushkin's blog used to predict an employee's risk of quitting. Today it is important for companies to understand the individual's risk of quitting. This model if further developed can be used by the companies individually to predict their employees' risk of quitting. It can also be used on a large scale to identify any underlying patterns and coming up with certain counter-measures.

TOPICS COVERED IN THE PROJECT:

1. Exploratory Data Analysis.
2. Data Visualization.
3. Survival Analysis.
4. Other supervised and unsupervised methods as per the requirements.

DATASET LINK: https://www.kaggle.com/datasets/davinwijaya/employee-turnover