# DATA ANALYSIS AND VISUALIZATION

Assignment-1

Roll No: 160122737010, 160122737012

Course: Information Technology (H1)

## PROBLEM STATEMENT:

Understanding time dynamics of online food delivery through data analysis and visualization to uncover patterns and trends in delivery times.

TITLE: “ DELIVER ON TIME ”

## ABSTRACT:

## This project aims to analyse the time dynamics of online delivery services through comprehensive data analysis and exploration. Leveraging a dataset containing various attributes related to delivery operations, including delivery personnel information, weather conditions, and order details, the project unfolds in several stages. First, a thorough understanding of the dataset is established, elucidating the meaning and significance of each attribute. Subsequently, data cleaning procedures are employed to address missing values and ensure data integrity. Following this, exploratory data analysis (EDA) techniques are applied to delve into the distribution of the target variable, Time taken, and to uncover insights into the relationships between numerical and categorical attributes and delivery time. Moreover, feature engineering is conducted to create additional features that may impact delivery time, such as time of day and distance between the restaurant and delivery location. Graphical analysis is then employed to assess the relationships between different attributes and delivery time. Through these systematic steps, this project aims to provide valuable insights into the factors influencing delivery time in online delivery services.

## OBJECTIVES AND OUTCOMES:

Objective 1 (Bloom's Taxonomy Level 3 - Application): Apply data analysis techniques to identify and address data quality issues in the online food delivery dataset.

Outcome 1: Demonstrate proficiency in using Pandas to load the dataset, identify missing data, outliers, zero values, and duplicates, and apply appropriate methods to handle these issues effectively.

Objective 2 (Bloom's Taxonomy Level 4 - Analysis): Analyse temporal patterns and trends in delivery times across different variables to gain insights into factors influencing service efficiency and customer satisfaction.

Outcome 2: Utilize statistical methods and visualization techniques (e.g., scatter plots, line plots) to explore relationships between delivery times and variables such as day of the week, time of day, geographical location, and food type, thereby identifying key factors impacting delivery performance.

# TOOLS AND LIBRARIES:

1. Python: As the primary programming language for data analysis tasks.
2. Pandas: For data manipulation and analysis, including loading the dataset, data preprocessing, and feature engineering.
3. NumPy: For numerical operations and array manipulation, often used in conjunction with Pandas for data processing.
4. Matplotlib and Seaborn: For data visualization, including plotting graphs and charts to explore relationships between variables and visualize model results.
5. GitHub: As an interactive development environment for running Python code, exploring data, and documenting the analysis process.

## DATASET :

Dataset from the GitHub platform, which contains relevant information on online food delivery operations, including attributes such as delivery personnel details, weather conditions, and order characteristics. This dataset serves as the foundation for conducting thorough analysis and exploration of the time dynamics in online food delivery services.

## DESIGN FLOW:

