# Individual Company customer Data

## Abstract

Customer data analysis is a detailed analysis of a company's ideal customers. It helps businesses better understand their customers and makes it easier for them to modify products according to the specific needs, behaviors, and interests of different types of customers.

Customer data analysis helps the company to identify the types of target groups in terms of age, gender, residence, country,... etc. and to adjust its products based on its target customers from different types of customer segments. For example, instead of spending money to market a new product to each customer in the company's database, the company can analyze which customer segment is likely to buy the product and then market the product to only that specific segment

## Design

This project is one of the T5 Data Science BootCamp requirements. Data provided by Kaggle has been used in this project. Create powerful marketing campaigns that are more personalized and timely. Helping the sales team. From the results of these analyzes, the team can make better products by knowing which features customers prefer and what they do not like and target groups prefer

**Q1\ what most common of fam_income customer**

**Q2\ what most common gender on the customer¶**

**Q3\ What is the most common educatione on the customer?**

**Q4\ What is the most common marrige on the customer?**

**Q5\ What is the most common you have child or not on the customer?**

**Q6\ What are the age groups for customer avarge sales?**

**Q7 \ What region sold the most?**

**Q8\ What the least occupation ?**

**Q9\ customer has bought the target product or not?**

**Q10\ What percentage of customers return the purchase experience?**

# Data

- They are the customer of an individual company and information about the customer in terms of family income has been collected to find out more ways to attract the customer and what ages, marital status and educational qualifications are affected by sales and through which statistics are made and predicted the most buying and most researched customers ,

- The dataset is provided in **.csv ,**and dataset includes about 40,000 rows and 15 columns Each row corresponds to a customer infomation, and includes the variables:

    1. flag: Whether the customer has bought the target product or not
    2. gender: Gender of the customer
    3. education: Education background of customer
    4. house_val: Value of the residence the customer lives in
    5. age: Age of the customer by group
    6. online: Whether the customer had online shopping experience or not
    7. customer_psy: Variable describing consumer psychology based on the area of residence
    8. marriage: Marriage status of the customer
    9. children: Whether the customer has children or not
    10. occupation: Career information of the customer
    11. mortgage: Housing Loan Information of customers
    12. house_own: Whether the customer owns a house or not
    13. region: Information on the area in which the customer are located
    14. car_prob: The probability that the customer will buy a new car(1 means the maximum possible）
    15. fam_income: Family income Information of the customer(A means the lowest, and L means the highest)

# Algorithms

- When working on analyzing data, you'll likely come across data that is miss ing (also called null values or NaNs) .**Data cleaning by used drop null val ue of column you have null [** `education, marriage , house_owner` **]** is an important part of the data analysis pipeline and making sure that it's all tidy up will make your analysis much stronger.

- rename column names is one of the most often applied data manipulations, **rename name [val_house ]to [value_house]**

- DataFrame will contain columns that are not useful to your analysis. Such columns should be dropped from the DataFrame **to** make it easier for you to focus on the remaining columns, **Drop columns [car_prob , customer_psy , mortgage ]**

# Tools

- Numpy and Pandas for data manipulation
- Matplotlib and Seaborn for plotting
- clean data
- Streamlit for interactive visualizations