# INTRODUCTION MARGINAL-WORKERS-CLASSIFIED-AGE-INDUSTRIAL

# **INTRODUCTION**

This table gives the break-up of the population by their economic activity status as 'main workers', 'marginal workers', 'non-workers' and 'marginal and non-worker' seeking/available for work cross classified with educational level and sex. This table gives the data for India/States/ UTs./Districts and City. This table is separate for SCs upto District level. It allows organizations to make informed decisions related to inventory levels, procurement, pricing, and marketing strategies.

# PREREQUISITES FOR BUILDING A MARGINAL-ORKERS-CLASSIFIED-AGE-INDUSTRIAL-MODEL.

* The data is obtained from https://tn.data.gov.in/catalog/marginal-workers-classified-age-industrial-category-and-sex-census-2011-india-and-states

Have the following libraries installed —

1. Importing OS (data.gov)
2. Numpy and Pandas libraries
3. Matplotlib
4. Seaborn

* Columns Required from dataset
* TABLE CODE
* STATE CODE
* DISTRICT CODE
* AREA NAME
* AGE GROUP

# UNDERSTAND THE SEGMENTATION DATA

Before starting any data science project, it is vital to explore the dataset and understand each variable.

* Libraries Imported :

1. Numpy
2. Pandas
3. Matplotlib
4. Seaborn

* Loading the Data

## df=pd.read\_csv(‘/kaggle/input/mall-customers/Mall\_Customers.csv’)

* let’s look at the head of the dataframe:

**df.head()**

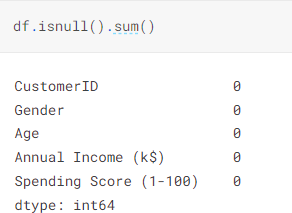


# PREPROCESSING DATA FOR SEGMENTATION

The raw data we downloaded is complex and in a format that cannot be easily ingested by customer segmentation models. We need to do some preliminary data preparation to make this data interpretable.

* Description

## IMG-20231018-WA0001.jpgdf.describe()

* Null Values

# TRAIN AND TEST

Testing the data by importing sklearn.linear\_modal from Linear Regression with ensuring the plot range and axis labels producing the values, scattering the data by mean\_absolute\_error and producing 3D plot. Training the dataset by describe(), isnull().sum(), drop(), show(), and by using Linear Regression algorithm we train the data

Testing the data by importing sklearn.linear\_modal from Linear Regression with ensuring the plot range and axis labels producing the values, scattering the data by mean\_absolute\_error and producing 3D plot.

# REST OF THE EXPLANATIONS

## Data Collection

The process involves gathering products data, which includes information about their purchase history, demographics, and interaction patterns.

## Data Preprocessing

The task involves preparing and cleaning data, handling missing values, and converting categorical features into numerical representations.

## Feature Engineering

Data preparation and cleaning, handling missing values, and the transformation of categorical features into numerical representations are all part of the task.

Modal Evaluation

Evaluate the model's performance on the test set using appropriate evaluation metrics. Common metrics for demand prediction include Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

ALGORITHMS USED

Apply clustering algorithms like K-Means, DBSCAN, or hierarchical clustering to segment customers.

Visualization: Visualize the customer segments using techniques like scatter plots, bar charts, and heatmaps. Interpretation: Analyze and interpret the characteristics of each customer segment to derive actionable insights for marketing strategies.

**Source code**

 <result>

    <index\_name>cdfc7695-7683-430d-ac29-46186f3615b5</index\_name>

    <title>Marginal Workers Classified by Age, Industrial Category and Sex, 2011 - Tamil Nadu</title>

    <desc>Marginal Workers Classified by Age, Industrial Category and Sex, 2011 - Tamil Nadu</desc>

    <org\_type>Central</org\_type>

    <org>

      <item>Ministry of Home Affairs</item>

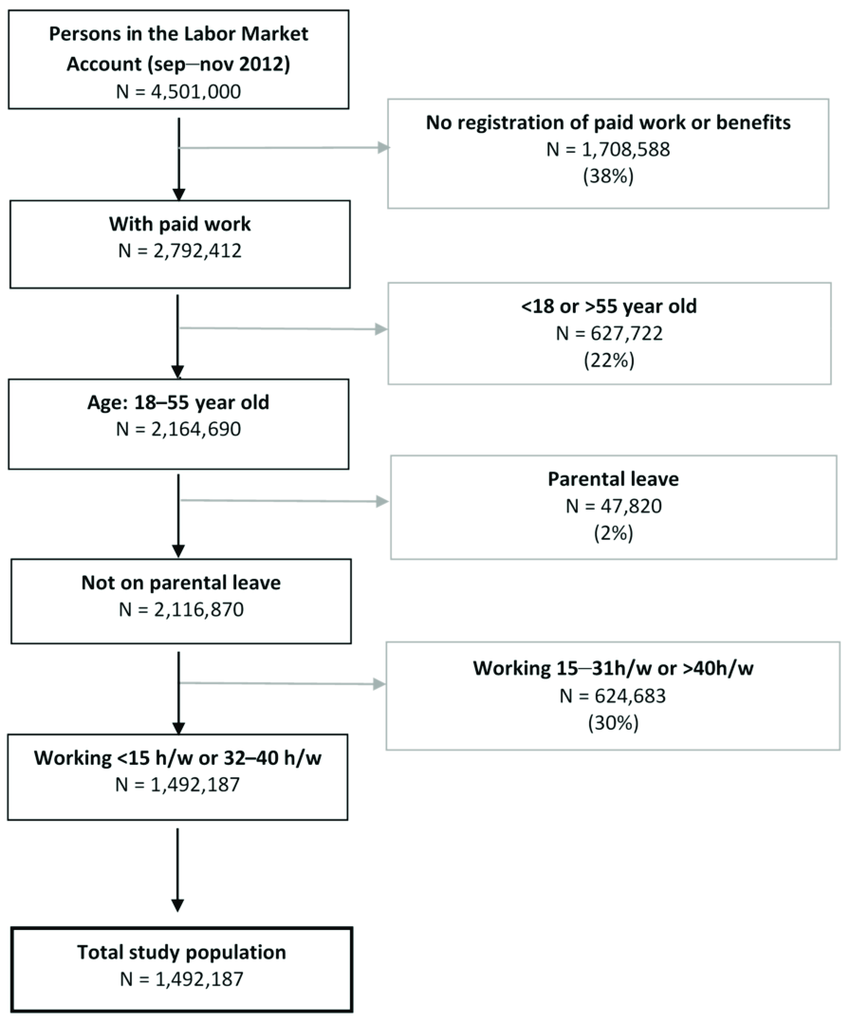
      <item>Department of Home</item>

      <item>Registrar General and Census Commissioner, India</item>

    </org>

# DATA FLOW OF CUSTOMER MODEL

1. Physical Flow



Figure

1. Logical Flow

