

# NAAN MUDHALVAN DATASCIENCE FUDAMANETAL PROJECT



## **PROJECT TITLE: MALL CUSTOMER SEGMENTATION**

**PRESENTED BY:**

**VENKATESHWARAN. T**

**SACS M.A.V.M.M. ENGINEERING COLLEGE**

**III YR CIVIL DEPT**

# Project statement



- Comprehensive report detailing the segmentation analysis, including methodology, results, and interpretations.
- Visualization of customer segments for easy understanding and decision-making.
- Recommendations for tailored marketing strategies and implementation plans.
- Presentation to stakeholders outlining key findings and proposed actions.

# Problem solution



- Increasingly competitive retail environments, such as shopping malls, necessitate a deeper understanding of customer behavior to tailor offerings and improve marketing strategies.
- The goal of this project is to segment mall customers based on their demographics, purchasing behavior, and visit frequency to identify distinct customer groups.
- By employing machine learning and data analytics techniques on customer data, we aim to provide actionable insights for personalized marketing campaigns, optimized tenant selection, and enhanced customer experiences within the mall environment.”

# System approach



- Building the proposed solution would involve a combination of data processing, python programming and visualization skills.
- **System requirements:**
  1. **Hardware :**
    - A computer with sufficient processing power, preferably with multiple cores or a GPU for faster training of machine learning models.
    - Adequate RAM to handle the size of the dataset and computational requirements.
  2. **Software:**
    - - An operating system compatible with the required python libraries (e.g., windows,linux,macOS).

# System approach – CONT.



- Library Requirements:
- Data processing and analysis:
  - **Pandas** : For data manipulation and analysis.
  - **Numpy** : For numerical operations on data.
- Data visualization:
  - **Matplotlib and seaborn**: For creating visualizations to understand data patterns.
  - **Plotly or Bokeh**: Interactive visualization libraries for more complex visualizations.

# ALGORITHM & DEPLOYMENT



## Algorithm selection

### Data exploration :

- Explore the movie rating structure, features and variables.
- Identify potential patterns, correlations and outliers.

### Problem formulation:

- Define the problem: Predict optimal booking times, likelihood of special requests based on historical data.

### Algorithm selection:

- Regression tasks(e.g., predicting daily rates)
  - Consider linear regression, decision tree, or ensemble methods
- classification tasks(e.g., predicting special requests);
  - Consider logistic regressive, decision trees or random forests.

# ALGORITHM & DEPLOYMENT



## **Data input:**

### **Data collection:**

- Gather historical data including booking dates ,special requests,and relevant details.

### **Data cleaning:**

- Handle missing values, outliers, and any inconsistencies in the dataset.
- Convert categorical variables into numerical representation through encoding techniques.

### **Feature Engineering:**

- Create new features or modify existing ones based on domain knowledge.
- Extract meaningful information from date variables, such as day-of-week or month.

# ALGORITHM & DEPLOYMENT



## Training process:

### Data splitting:

- Divide the dataset into training and testing sets to evaluate the model's performance.

### Feature scaling:

- Standardize or normalize numerical features to ensure they have consistent scale.

### Modeling training:

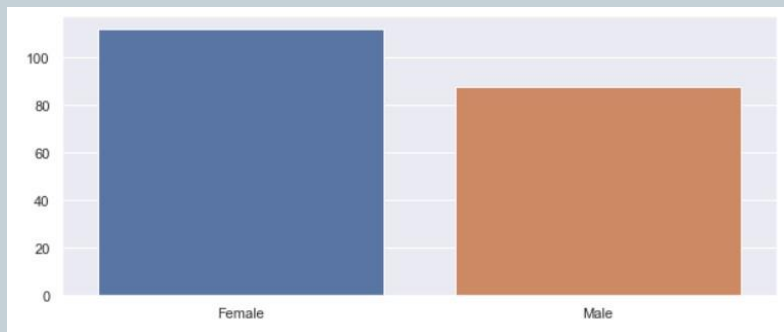
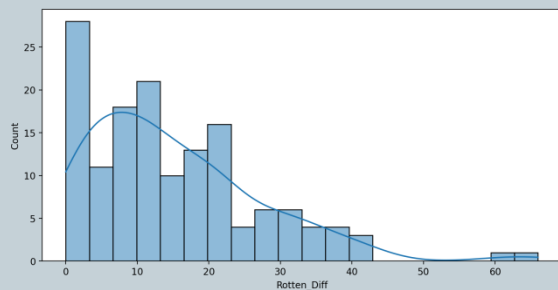
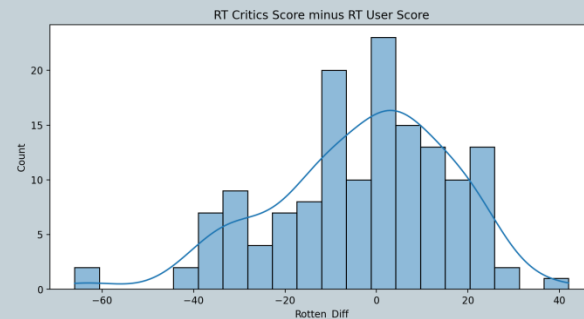
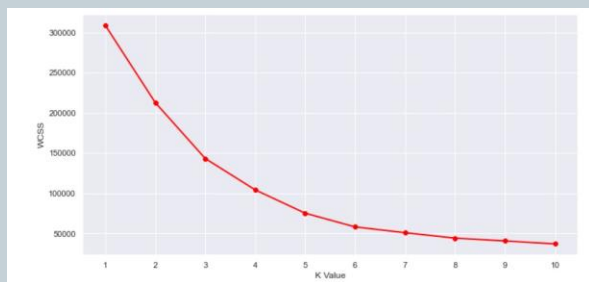
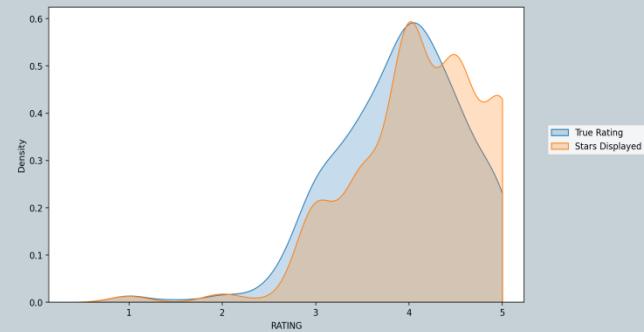
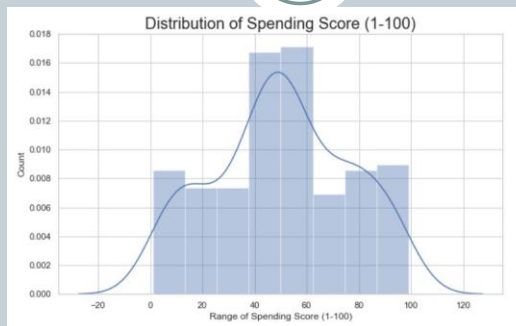
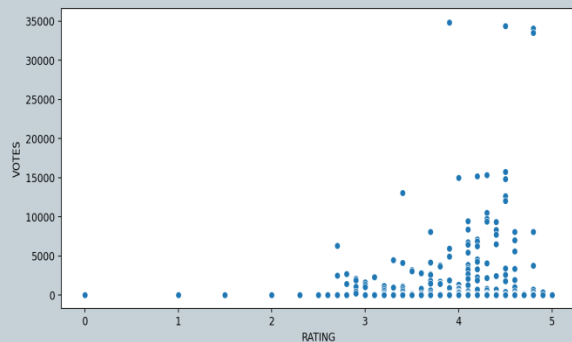
- Use the selected algorithm to train the model on the all sites scores dataset.
- Adjust hyperparameters to optimize model performance.

### Model evaluation:

- Evaluate the model on the dataset using appropriate metrics(e.g., Mean Squared Error for regression, accuracy, precision, recall for classification).
- Fine-tune the model if necessary.



# RESULT



# CONCLUSION



- Offers a powerful approach to understanding the diverse needs and preferences of visitors. By analyzing demographic data, purchasing behavior, and visit frequency, malls can identify distinct customer segments and tailor their offerings accordingly.
- This facilitates targeted marketing efforts, optimized tenant selection, and personalized experiences, ultimately enhancing customer satisfaction and driving business success in today's competitive retail landscape.

# REFERENCES



- <http://www.kaggle.com/datasets>
- <http://pandas.pydata.org/pandas-docs/stable/userguide/index.html>
- <http://seaborn.pydata.org/>
- <http://matplotlib.org/stable/contents.html>

Thank you