Python Project

Name of the Project:

Housing Society

Done By: Tanmay Saha

Introduction

The housing society has provided us with a comprehensive dataset named "Final\_Data.csv," which contains essential information about its residents, flats, confirmed member and donation details.

Objective

Our primary objective is to analyse the data that empowers the housing society to make data-driven decisions, optimize event planning, and ensure the efficient utilisation of resources. By leveraging the decorator and caterer data, we aim to enhance the event experience, minimise costs, and maximise resident satisfaction. The selection of the best caterer and decorator based on customer satisfaction and cost-effectiveness will play a crucial role in achieving these goals and ensuring the success of the grand event.

Building the code

* Data Loading and Overview
* Getting a Summary of the Dataset
* Descriptive Statistics
* Count Number of Unique Values
* Checking for Unique Values
* Check for Duplicate Values
* Replace "\_" with NaN in the entire DataFrame
* Handling Missing Values
* Data Type Conversion
* Impute Missing Values in 'Maintenance Amt' , ‘Flat Vacancy’, 'Donation', ‘No of Resident' , 'Confirmed Members', 'Availability of owner' and 'Origin of Owner'
  + - To address missing values in the 'Maintenance Amt' column. For each missing value, we impute it based on the corresponding 'Flat Area (sq.mt)' value. We reference a previously calculated dictionary, 'avg\_maintenance\_amt,' to fill in these values
    - We handle missing values in the 'Flat Vacancy' column by using a conditional approach. If 'Availability of owner' is 'Yes,' we assume the flat is owned ('Flat Vacancy' is 'Owned'), and if 'Availability of owner' is not 'Yes,' we consider the flat as vacant ('Flat Vacancy' is 'Vacant')
    - For missing values in the 'Donation', ‘No of Resident' , 'Confirmed Members' column, we impute them with the median value of the respective column.
    - Finally, we handle missing values in the 'Availability of owner' and 'Origin of Owner' columns by imputing them with the mode (most frequent value) of their respective columns
* Calculate "Outsiders"
* Standardize Columns

Analysis

The 'Outsiders' column helps identify residents who are not part of the housing society but have confirmed their participation in the event, potentially impacting event logistics.

We conducted a demographic analysis of the housing society dataset, including insights into residents' origins, flat characteristics, and ownership structure

We provided event planning recommendations, which are crucial for efficient event management, cost optimization, and ensuring a successful grand event. These analyses help the housing society make informed decisions, allocate resources effectively, and create an enjoyable and memorable event for the residents.

By following these recommendations, the housing society can ensure that they have both a highly-rated decorator and a cost-effective caterer for the grand event. This balanced approach should help optimize the event experience while minimizing costs, ultimately contributing to the success of the event.

In conclusion, this analysis provides a solid foundation for event planning, taking into account resident demographics, participation rates, budgeting, and vendor selection. By following these recommendations, the housing society can ensure a well-organized and enjoyable event that caters to the preferences and needs of its diverse resident population while making efficient use of available resources.