1. Project Overview:

# Title:

**Data Analysis using Python**

Description:

**Objective:**

This project aims to analyze employee data from a corporate environment to gain insights into the relationship between housing, transportation, and employee performance and satisfaction. By leveraging Python's data analysis libraries, we intend to identify how factors such as owning a house, owning a car, salary levels, and location influence employee engagement, job satisfaction, and retention.

**Technologies Used:**

I Used Software Technologies Such as Python Modules like NumPy, Pandas, Matplotlib, Seaborn

and python frameworks like Django, etc.

**Data Sources**: The data sources I used utilized for the project are :

* SQL Server,
* MySQL,
* Excel,
* text files

**2. Data Acquisition and Preparation (Django Faked Data):**

**Django Setup**: Created a Django project for generating fake data. .

**Data Generation**: For generating realistic and representative fake data using Django's built-in faker module or other data generation libraries.

**Data Storage:** The generated fake data is stored (in a database, as a text file, or and sqlserver).

**3. Data Cleaning and Exploration (NumPy):**

**Data Loading:**

The data from different sources (SQL Server, MySQL, Excel, text files) into Python's memory using appropriate modules like pandas.read\_sql, pandas.read\_excel, text etc.

**Data Cleaning:**

**Handling missing values**: Used some commands methods like deletion, imputation (filling in missing values), or creating new categories..

**5. Data Visualization (Matplotlib):**

**Plot Types**:

I created line charts, bar charts, histograms, heatmaps, etc. to convey insights gained from the data

**Dashboard Design**: I built data dashboards

**7. Conclusion**:

I analyzied employee data from a corporate environment to gain insights into the relationship between housing, transportation, and employee performance and satisfaction.