**📞 Customer Call Data Cleaning Project**

This mini-project focuses on **cleaning and preparing customer call data** for analysis using **Python**. Data cleaning is a crucial step in data science to ensure the dataset is free from errors, inconsistencies, and missing values, making it ready for further analysis and visualization. 🚀

**🛠️ Project Workflow:**

1. **Loading Data** 📂
   * Imported the dataset using pandas.
   * Performed an initial check for structure and anomalies with .info(), .describe(), and .head().
2. **Handling Missing Values** 🤔
   * Identified missing data using .isnull().
   * Replaced or removed null values where necessary with techniques like mean/mode imputation and dropping rows/columns.
3. **Removing Duplicates** ♻️
   * Checked for duplicate rows and removed redundant entries.
4. **Data Standardization** 🧹
   * Standardized column names (e.g., lowercase and underscores for consistency).
   * Formatted inconsistent date and time formats.
5. **Outlier Detection and Treatment** 📊
   * Used statistical methods (IQR, z-scores) to detect outliers in columns like call duration and customer satisfaction scores.
6. **Categorical Encoding** 🔢
   * Encoded categorical variables using one-hot encoding and label encoding for compatibility with machine learning models.
7. **Final Clean Dataset** ✅
   * Exported the cleaned dataset for further use in analysis and modeling.

**📌 Key Python Libraries Used:**

* pandas for data manipulation
* numpy for numerical operations
* matplotlib and seaborn for visualizing patterns during cleaning

**✨ Outcome:**

The dataset is now clean, structured, and ready for advanced analytics or machine learning! This project strengthened my skills in handling messy real-world data. 💪

**🌟 Future Scope:**

* Perform exploratory data analysis (EDA) on this cleaned dataset 📊.
* Build predictive models to gain deeper insights! 🤖

**🛠️ *Tech Stack:***

* **Language:** Python 🐍
* **Libraries:** pandas, numpy, matplotlib, seaborn