

# Data Cleaning Project Documentation using Pandas in Jupyter Notebooks

## Project Objective:

The primary aim of this project is to standardize and clean the information provided in an Excel document for the sales team. The goal is to extract specific data so that the sales team can contact only customers who have granted permission and possess the necessary information for calls.

## Technologies Used:

- **Programming Language:** Python
- **Key Libraries:** Pandas
- **Development Environment:** Jupyter Notebooks

## Data Cleaning Process:

- 1. Data Loading:** The project begins by loading the provided Excel document.
- 2. Initial Analysis:** Initial review and exploration of the data to understand its structure and characteristics.
- 3. Cleaning Steps:**
  - **Data Standardization:** Processes are carried out to standardize data formats (dates, numbers, text, etc.).
  - **Handling Null Values:** Treatment of missing or null values based on context and relevance to the sales team.
  - **Filtering Eligible Customers:** Identification and filtering of customers who have granted permission and have the necessary information for calls.
  - **Result Generation:** Presentation of the final set of clean data suitable for the sales team.

## Results and Output:

The final outcome of the project is a set of clean and standardized data showcasing only customers who have granted permission to be contacted and have the necessary information for calls. This optimized set facilitates the sales team's task by easily visualizing eligible customers for interactions.

## Conclusions and Final Considerations:

The project achieves its objective by providing clean and focused data, simplifying the process for the sales team. Ethical considerations and data accuracy were maintained at a high level to ensure the validity and relevance of the presented information.

## Additional Notes:

- **Script Highlights:** The steps taken are highlighted within the script itself for clarity and reference.
- **Result Visualization:** To facilitate result comparison with the original raw data, an image showcasing the final table is included in this document.

	CustomerID	First_Name	Last_Name	Phone_Number	Address	Paying Customer	Do_Not_Contact	City	Postal Code	Street_Address	Postal_Code
0	1001	Frodo	Baggins	1235455421	123 Shire Lane, Shire	Yes	No	Shire		123 Shire Lane	
1	1005	Jon	Snow	8766783469	123 Dragons Road	Yes	No			123 Dragons Road	
2	1008	Sherlock	Holmes	8766783469	98 Clue Drive	No	No			98 Clue Drive	
3	1010	Peter	Parker	1235455421	25th Main Street, New York	Yes	No	New York		25th Main Street	
4	1013	Don	Draper	1235432345	2039 Main Street	Yes	No			2039 Main Street	
5	1014	Leslie	Knope	8766783469	343 City Parkway	Yes	No			343 City Parkway	
6	1015	Toby	Flenderson	3047622467	214 HR Avenue	No	No			214 HR Avenue	
7	1016	Ron	Weasley	1235455421	2395 Hogwarts Avenue	No	No			2395 Hogwarts Avenue	
8	1017	Michael	Scott	1236439775	121 Paper Avenue, Pennsylvania	Yes	No	Pennsylvania		121 Paper Avenue	
9	1020	Anakin	Skywalker	8766783469	910 Tatooine Road, Tatooine	Yes	No	Tatooine		910 Tatooine Road	