

EXCEL DATA CLEANING BASED PROJECT

Jimoh, Sekinat Eniola

March 6, 2025

Tables of Content

1. Introduction
2. Data Cleaning Process
 - Initial Assessment
 - Handling Duplicates and Missing Values
 - Column-Specific Cleaning
 - Date Column
 - Vehicle Registration Number
 - Make/Model Column
 - Item Description
 - Assigned to Column
 - Service KM
 - Workshop Name
 - Invoice Number
 - Amount (NGN)
 - Remarks
 - Final Verification Using Power Query
3. Conclusion
4. Pictures
5. Link to the Spreadsheet

INTRODUCTION

The Project involved cleaning a structured dataset using Microsoft Excel to improve its accuracy and usability. The dataset contained 3,833 records with various columns, including Date, Vehicle Registration Number, Make/Model, Item Description, Assigned to, Service KM, Workshop Name, Invoice Number, Amount (NGN), and Remarks. The primary objective was to identify and resolve inconsistencies, missing values, and redundant information to ensure a well-structured dataset ready for further analysis and decision-making.

Data Cleaning Process

Initial Assessment

- The dataset contained 3,833 records with various structured fields.
- An image was embedded in the file, which was removed.
- Unnecessary and empty rows were deleted, including section titles and project details, ensuring the dataset started with actual column headers. A total of 7 empty rows were removed.

Handling Duplicates and Missing Values

- The dataset was reduced to 3,830 records after removing 6 duplicate entries using the **Remove Duplicates** feature in the **Data** tab.
- Missing values in numerical columns were replaced with appropriate values (e.g., 0 in Service KM where data was absent).

Column-Specific Cleaning

a. Date Column

- The original date format displayed as numbers (e.g., "42064").
- Converted to proper date format using the **Format Cells** option (e.g., "01-03-15").
- Identified and manually corrected two inconsistencies in date formatting.

b. Vehicle Registration Number

- Inconsistencies such as values enclosed in brackets were identified.
- Used the **Find and Replace** function to remove unnecessary brackets, affecting 67 values.

c. Make/Model Column

- The column was split using the **Text to Columns** function, resulting in four columns.
- The Make column remained unchanged, while Columns 2, 3, and 4 contained fragmented model names.
- Used the **Concatenate** function to merge them back into a single column.
- Applied **Find and Replace** to remove brackets and correct typographical errors.

d. Item Description

- Standardized the item descriptions into five categories using **IF, OR, ISNUMBER, and SEARCH** functions:
 - Maintenance
 - Body Repair
 - Electrical Repairs
 - Tyres and Wheels
 - Other Repairs
- Example in using the function “= (IF (OR (ISNUMBER (SEARCH (“OIL”, E2)), ISNUMBER (SEARCH (“FILTER”, E2)), ISNUMBER (SEARCH (“BELT”, E2))), “MAINTENANCE”, ...)” **Simply meaning if either oil, filter or plugs are found in column E2, group it as maintenance.**
- Replaced the original description column with a newly created **Item Category** column.

e. Assigned to Column

- Standardized name formats by sorting alphabetically and applying **Find and Replace**.
- Example: "A. A. Warsi," "A. Warsi," and "Ahfad Warsi" were unified to maintain consistency.

f. Service(km)

- Changed data type to numerical format using **Format Cells**.
- Missing values replaced with "0" for consistency.
- Applied number formatting to include commas for readability.

g. Workshop Name

- Some entries contained state names; these were separated using **Text to Columns**, and the redundant state column was deleted.
- Sorted alphabetically and corrected typographical errors using **Find and Replace**.
- Applied **TRIM** function to remove unnecessary spaces and **PROPER** function for uniform casing.

h. Invoice Number

- Contained numeric, alphanumeric, and alphabetic values with some missing data.
- Changed data type to text since invoice numbers do not require calculations.
- Standardized values by replacing "Petty Cash" entries with "PC001".
- Missing values replaced with "NA" (Not Available) for clarity.

i. Amount (NGN)

- Changed data type to currency.
- Customized formatting to display the **Naira (₦) symbol**.
- No missing values were detected in this column.

j. Remarks

- The column was removed due to being empty.


Final Verification Using Power Query

- Loaded the cleaned dataset into **Power Query** for additional validation.
- Identified and corrected minor typographical errors.
- Reloaded the refined dataset back into Microsoft Excel for final use.

Conclusion

The data cleaning process significantly enhanced the dataset's accuracy, structure, and usability. By systematically addressing inconsistencies, removing redundant information, and applying appropriate formatting techniques, the dataset is now well-structured for further analysis. The refinements improve data integrity, ensuring reliable insights for decision-making. The final dataset is free from duplicates, contains correctly formatted values, and adheres to standardized naming conventions, making it more efficient for reporting and analysis.

PICTURES

A	B	C	D	E	F	G	H	I	J	
1		NIGERIAN RAILWAY MODERNISATION PROJECT: LAGOS - IBADAN Section								
2		VEHICLE MAINTENANCE HISTORY								
3		Papalanto, 31								
4										
5										
6										
7										
8	Date	VEHICLE REG. No.	MAKE / MODEL	ITEM DESCRIPTION	ASSIGNED TO	SERVICE KM	WORKSHOP NAME	INVOICE No.	AMOUNT (NGN)	REMARKS
9	42064	FG 275 A19	MITSUBISHI (PAJERO-p)	Service LHS Rear bulb, Rear Brake pads, Engine oil, (Mr. C.Gullini	Head Office	146137	Mitsubishi Motors		43470	
10	42064	FG 280 A19	MITSUBISHI (ASX)	Replacement of new compressor, a/c condenser, ex	Pool (IDU)		MTA Company	411	263000	
11	42810	FG 286 A19	TOYOTA (HILUX)	Repair of front bomber	Dr L. Portanova		Andrew M. M.W.	372	6000	
12	42830	FG 228 A19	TOYOTA (PRADO)	Service-Front Brake pad, Back Brake pads	Stefano Orioli	151491	MTA Company	422	113000	
13	42856	FG 280 A19	MITSUBISHI (ASX)	Fan belt	Dr L. Portanova	156877	Odika Sons Motors, Lagos	152	5000	
14	42870	FG 228 A19	TOYOTA (PRADO)	Case Gold Battery 12V/75A	Dr L. Portanova	202058	Madu F. Obinna, Lagos	293	27000	
15	42877	FG 228 A19	TOYOTA (PRADO)	Back Wiper Blade	Dr L. Portanova	202689	Narde Bros Ent., Lagos	2454	1500	
16	42877	FG 228 A19	TOYOTA (PRADO)	Set of Front Wiper Blades	Dr L. Portanova	202689	Batson Investment Co., Lagos	445	1500	
17	42877	FG 280 A19	MITSUBISHI (ASX)	Set of Front Wiper Blades	Stefano Orioli	202815	Batson Investment Co., Lagos	445	1500	
18	42877	FG 275 A19	MITSUBISHI (PAJERO-p)	Set of Front Wiper Blades	Claudio Gullini	194929	Batson Investment Co., Lagos	445	1500	
19	42881	FG 228 A19	TOYOTA (PRADO)	Repair of A/C, panel beater work	Dr L. Portanova	202892	Mon-Obi & Company Ent., Lagos	233	20000	
20	42885	FG 280 A19	MITSUBISHI (ASX)	Steering cover, sewing front seat driver	Stefano Orioli	157778	Narde Bros Ent., Lagos	2461	1700	
21	42922	FG 275 A19	MITSUBISHI (PAJERO-p)	Fixing No. 2 tyre BRAVO A/T MAXXIS 700 (265-70-16	Claudio Gullini	194350	Conoil Service Station, Lagos		5000	
22	42893	FG 275 A19	MITSUBISHI (PAJERO-p)	Purchasing of No. 2 tyre BRAVO A/T MAXXIS 700 (26	Claudio Gullini	194350	Innomakson Int. Ltd., Lagos	5790	80000	
23	42893	FG 228 A19	TOYOTA (PRADO)	Repair and fill A/C compressor with gas	Dr L. Portanova	203470	(Intel's Eko Atlantic), Lagos	petty cash	8500	
24	42896	FG 275 A19	MITSUBISHI (PAJERO-p)	Control front break and steering thyroid	Claudio Gullini	194572	(Intel's Eko Atlantic), Lagos	petty cash	2500	
25	42901	FG 281 A19	TOYOTA (HILUX)	Repair and sewing driver seat cover	Claudio Teagno	211219	Jamui Hassan, Lagos	352	4000	

BEFORE CLEANING

A	B	C	D	E	F	G	H	I	J
Date	VEHICLE REG. No.	MAKE	MODEL	ITEM CATEGORY	ASSIGNED TO	SERVICE (km)	WORKSHOP NAME	INVOICE No.	AMOUNT (NGN)
10-06-17	FG 275 A19	mitsubishi	PAJERO-p	BODY REPAIRS	Claudio Gullini	194,572	Intel's Eko Atlantic	PC0001	₦ 2,500
07-06-17	FG 228 A19	TOYOTA	PRADO	MAINTENANCE ONLY	Dr. L. Portanova	203,470	Intel's Eko Atlantic	PC0001	₦ 8,500
31-08-18	FG 286 A19	TOYOTA	HILUX 2.7 WT-I	OTHER REPAIRS	Papalanto Engineering Supervision	204,076	Adio Raji Auto Engineering Service	000369	₦ 6,000
31-05-18	19A 14 FG	NISSAN	URVAN	TYRES AND WHEELS	Staff bus	112,072	Adio Raji Auto Engineering Service	000202	₦ 3,500
04-09-18	19A 90 FG	TOYOTA	HILUX	TYRES AND WHEELS	Ahfad Warsi	34,294	Adio Raji Auto Engineering Service	NA	₦ 205,500
04-09-18	19A 87 FG	TOYOTA	HILUX	TYRES AND WHEELS	Benone Davidescu	46,164	Adio Raji Auto Engineering Service	NA	₦ 205,500
04-09-18	FG 286 A19	TOYOTA	HILUX	TYRES AND WHEELS	Pool	207,992	Adio Raji Auto Engineering Service	NA	₦ 134,000
20-02-18	FG 228 A19	TOYOTA	PRADO	TYRES AND WHEELS	Papalanto Office	222,180	Adio Raji Auto Engineering Service	000032	₦ 5,000
02-02-18	BWR 456 CZ	TOYOTA	HILUX 2.7 WT-I	TYRES AND WHEELS	Pool Lagos Supervision	50,086	Adio Raji Auto Engineering Service	0036	₦ 5,000
31-05-18	19A 15F G	TOYOTA	HILUX 2.7 WT-I	TYRES AND WHEELS	Pool Papalanto	251,105	Adio Raji Auto Engineering Service	00201	₦ 3,500
31-07-18	FG 286 A19	TOYOTA	HILUX 2.7 WT-I	TYRES AND WHEELS	Papalanto Engineering Supervision	201,773	Adio Raji Auto Engineering Service	000207	₦ 3,500
30-06-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	MAINTENANCE ONLY	Engr Muhammad Ali	176,459	Akinkunmi Adisa Enterprises	0032	₦ 50,000
31-05-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	TYRES AND WHEELS	Pool Ibadan Supervision	173,125	Akinkunmi Adisa Enterprises	0017	₦ 3,000
30-06-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	MAINTENANCE ONLY	Pool Ibadan Supervision	175,575	Akinkunmi Adisa Enterprises	0027	₦ 47,000
30-09-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	MAINTENANCE ONLY	Pool Ibadan Supervision	181,298	Akinkunmi Adisa Enterprises	0058	₦ 28,600
30-10-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	BODY REPAIRS	Pool Ibadan Supervision	185,445	Akinkunmi Adisa Enterprises	0070	₦ 33,000
30-10-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	TYRES AND WHEELS	Pool Ibadan Supervision	186,414	Akinkunmi Adisa Enterprises	0083	₦ 16,000
30-10-18	19A 95 FG	TOYOTA	HILUX 2.7 WT-I	MAINTENANCE ONLY	Pool Ibadan Supervision	186,810	Akinkunmi Adisa Enterprises	0079	₦ 83,500
30-10-18	FG 274 A19	TOYOTA	PRADO	OTHER REPAIRS	Alfredo Citta	106,196	Akinkunmi Adisa Enterprises	0082	₦ 51,000
30-07-18	FG 228 A19	TOYOTA	PRADO	MAINTENANCE ONLY	Alvaro Figueredo	233,363	Akinkunmi Adisa Enterprises	0039	₦ 27,000

AFTER CLEANING

LINK

<https://docs.google.com/spreadsheets/d/1TGoD1g7KLcqjrMIFXpBs7-ibfA3c19W/edit?usp=sharing&ouid=106812219724605440864&rtpof=true&sd=true>

