

## Phase 2: Data Cleaning & Preprocessing - README

Project Title: Electric Vehicle Population Analysis - Washington State

Phase: 2 - Data Cleaning & Preprocessing

Tools Used: Microsoft Excel (2021)

Dataset: Electric Vehicle Population Dataset (Washington State)

Original Format: CSV

Cleaned Format: Excel (.xlsx)

-----  
Key Cleaning Steps Performed:

1. Renamed Columns:

- Replaced spaces with underscores and removed special characters for consistency.

2. Removed Duplicates:

- Ensured unique records based on 'VIN (1-10)'.

3. Handled Missing Values:

- Postal Code: Filled with 'Unknown'.
- City & County: Filled with 'Unknown'.
- Electric Range and Model Year: Defaulted to 0 if missing.

4. Standardized Text:

- Converted all relevant columns (Make, Model, City, County, etc.) to uppercase and removed extra spaces.

#### 5. Converted Data Types:

- Ensured numeric fields like Model Year, Electric Range, and MSRP are set to integer/decimal types.

- Preserved Postal Code as Text to maintain formatting.

#### 6. Extracted Latitude & Longitude:

- Split from 'Vehicle Location' and stored as separate numeric fields.

#### 7. Dropped Redundant Columns:

- Removed 'Vehicle Location' after extracting coordinates.

-----

#### Final Output:

1. Cleaned Excel File: PowerBI\_Cleaned\_EV\_Data.xlsx
2. Date Table Excel File: PowerBI\_Date\_Table.xlsx
3. Cleaning Report PDF: Data\_Cleaning\_Report\_EV\_Population.pdf
4. This README File

Prepared By: [Your Name]

Date: [Insert Date of Submission]