

Excel DA Projects Documentation

Coffee Beans and Sales Raw Dataset:

kaggle datasets download -d saadharoon27/coffee-bean-sales-raw-dataset

Key Functions and Formulas used in this Project:

1. XLOOKUP, VLOOKUP
2. IF, IFS
3. INDEX, MATCH, XMATCH
4. TRIM, TEXT
5. Date Formatting and Number Formatting
6. Pivot Tables
7. Line Chart, Bar Chart, Slicers

Data Cleaning performed in this project:

1. In the “Customer” sheet we perform data cleaning operations like:
 - a. Creating a new “Email-clean” column and using the formula “=IF(\$D2="", "Email not found", \$D2)” to fill the empty cells in the column.
2. In the “Products” sheet we perform data cleaning operations like:
 - a. Creating a new “Coffee Type - clean” column and using the formula “=IFS(\$C2="Ara", "Arabica", \$C2="Exc", "Excelsa", \$C2="Lib", "Liberica", \$C2="Rob", "Robusta)” to replace the short names with full coffee names.
 - b. Creating a new “Roast Type - clean” column and using the formula “=IFS(\$E2="L", "Light", \$E2="M", "Medium", \$E2="D", "Dark)” to replace the initials with full roast types.
3. In the “Orders” sheet we perform data cleaning operations like:
 - a. Removing duplicates from “Order ID” column.
 - b. Created “Order Month” column using “=TEXT(B2, "mmmm)” and “Order Year” column using “=TEXT(B2, "yyyy)”
 - c. Inserted entries in “Customer Name” using “=XLOOKUP(\$E2, clean_customers!\$A\$2:\$A\$1001, TRIM(clean_customers!\$B\$2:\$B\$1001), "Customer Not Found", 0)”
 - d. Inserted entries in “Email” using “=VLOOKUP(\$E2, clean_customers!\$A\$2:\$C\$1001, 3, FALSE)”
 - e. Inserted entries in “Country” using “=VLOOKUP(\$E2, clean_customers!\$A\$2:\$H\$1001, 8, FALSE)”
 - f. Inserted entries in “Coffee Type” using

"=XLOOKUP(\$F2,clean_products!\$A\$2:\$A\$49,clean_products!\$B\$2:\$B\$49,"
Coffee Type Not Found",0)"

- g. Inserted entries in "Roast Type" using

"=VLOOKUP(\$F2,clean_products!\$A\$2:\$E\$49,4,FALSE)"

- h. Inserted entries in "Size" using

"=CONCATENATE(INDEX(clean_products!\$A\$2:\$H\$49,MATCH(clean_orders!\$F2,clean_products!\$A\$2:\$A\$49,0),XMATCH(M\$1,clean_products!\$A\$1:\$H\$1,0,1))," kg")"

- i. Inserted entries in "Unit Price" using

"=INDEX(clean_products!\$A\$2:\$H\$49,MATCH(clean_orders!\$F2,clean_products!\$A\$2:\$A\$49,0),XMATCH(N\$1,clean_products!\$A\$1:\$H\$1,0,1))"

- j. Inserted entries in "Sales" using "=\$G2*\$N2"

- k. Inserted entries in "Loyalty" using

"=XLOOKUP(\$E2,clean_customers!\$A\$2:\$A\$1001,clean_customers!\$J\$2:\$J\$1001,"No",0)"