Process\_summary.md

Data Cleaning Process (Detailed)

1. **Initial Inspection**
   1. Found columns with bad formatting: Unnamed: 0, text-based Date
   2. Spotted possible duplicates and negative values
2. **Step-by-Step Cleaning**
   1. Dropped unnecessary column Unnamed: 0
   2. Converted Date column from TEXT to DATE using STR\_TO\_DATE
   3. Added id as a unique identifier
   4. Removed duplicate rows using ROW\_NUMBER ()
   5. Replaced negative prices and volumes with NULL
   6. Standardized type and region using LOWER () and TRIM ()
   7. Added a new column avg\_price to show regional price averages
3. **Final Export**
   1. Exported cleaned dataset as avocado\_cleaned.csv