Edge Cases

Below are the edge cases considered in the solution implemented.

    A) Data Quality
- **Null-to-default handling**
- **Filtering invalid records**
- **Deduplication at the natural grain**
- **Type Parsing**
- **Derived columns existence check**
- **String normalization (by trimming & lowercasing)**
- **Happy path + boundary separation for window joins (table joins) & enrichment**

    B) Operational (Databricks & pytest execution constraints)
- **SparkSession injection into pytest**
- **Writable test execution path to avoid collisions/permission denials across concurrent runs**
- **No pycache writes into Workspace, preventing pollution & permission errors**