

# Interpretación de Resultados del Comparador

**Propósito.** Este documento explica, cómo interpretar las salidas del comparador entre una **tabla de referencia (REF)** y una **tabla nueva (NEW)**. Incluye definiciones de métricas, tipos de duplicados, opciones de configuración.

**Incluyo en esta explicación** dos casos reales de pruebas de ejecuciones en PRO desde el 2025-09-18 hasta 2025-09-24. La comparación es de REF contra REF. Se utilizó la siguiente **clave compuesta**:

geo, uetr, type, time, in\_out, segregation, sender\_reference

---

## 1) Tablas de salida y para qué sirven

### 1. Tabla Ejecutiva: bd.results\_summary

block	metric	universe	numerator	denominator	pct
KPIS	Unique IDs	REF	55058	-	-
KPIS	Unique IDs	NEW	55058	-	-
KPIS	Total rows REF	ROWS	55841	-	-
KPIS	Total rows NEW	ROWS	55841	-	-
KPIS	Total diff(new-ref)	ROWS	0	55841	0.0000%
KPIS	Global quality	REF	55050	55058	99.9855%
EXACT MATCH	1:1 (all columns)	BOTH	55058	55058	100.0000%
PARTIAL MATCH	1:1 (match & no_match co...	BOTH	0	55058	0.0000%
GAP	1:0 (only in ref)	REF	0	55058	0.0000%
GAP	0:1 (only in new)	NEW	0	55058	0.0000%
DUPS	duplicates (both)	BOTH	8	55058	0.0145%
DUPS	duplicates (only in ref)	REF	0	55058	0.0000%
DUPS	duplicates (only in new)	NEW	0	55058	0.0000%

Resumen ejecutivo de la ejecución. No muestra detalle por fila. Incluye:

## BLOQUE 1:

- **Unique IDs**: número de identificadores únicos según la clave definida.
- **Total rows**: conteo de filas totales que llegan (incluye duplicados).
- **Total diff**: diferencia de filas entre NEW y REF ( $NEW - REF$ ).
- **Global Quality**: métrica global de calidad. **Si existen duplicados no alcanza el 100%**, porque los dupes forman parte de su cálculo.
- **Samples por métrica (hasta 6 IDs)**: para cada métrica relevante (p. ej., diferencias y duplicados) se guardan **hasta 6 identificadores de ejemplo**. Sirven para localizar rápidamente esos casos tanto en `results_differences` como en `results_duplicates`.

## BLOQUE 2:

### Tabla de Diferencias: `bd.results_differences`: `bd.results_differences**`

Detalle de las divergencias a nivel de identificador y columnas. Puede incluir (según configuración): - **EXACT\_MATCH (1:1)**: todas las columnas coinciden (se muestra si está activada la opción de guardar exactos).

- **PARTIAL\_MATCH (1:1)**: el identificador coincide, pero hay diferencias en una o más columnas.
- En esta situación, pueden guardarse **MATCH** y **NO\_MATCH** si la opción está habilitada. - **GAP**: elementos presentes en una tabla y ausentes en la otra. Se muestran **hasta 6 ejemplos** de identificadores por tipo de hueco.

## BLOQUE 3:

### Tabla de Duplicados: `bd.results_duplicates`

Detalle de duplicidades de identificadores:

- **both**: el mismo identificador se duplica **en REF y en NEW**. *Nota: cuando ocurre en ambos orígenes, se descuenta de los contadores independientes de cada lado.*
- **duplica REF**: duplicado solo en la referencia.
- **duplica NEW**: duplicado solo en la nueva.

Además, se distinguen dos subtipos en la tabla de duplicados:

- **Duplicados exactos**: filas idénticas.
- **Duplicados con variación**: el identificador coincide, pero alguna(s) columna(s) difiere(n). Es habitual cuando algún campo usado como identificador viene **NULL**.

La columna **variaciones** muestra el detalle de la diferencia por columna, p. ej.: `bic_1: [valor_1, valor_2]` (y así sucesivamente para el resto de los campos afectados).

**Nota de producto:** En el futuro, es posible unificar la categoría visual de duplicados en una sola etiqueta (**both**), coherente con el resumen.

## 2) Existen 2 Casos de métricas que se dan en PRO (casos reales)

### Caso 1 — Igualdad perfecta (2025-09-21)

- **Summary** igualdad total. **KPI Global quality** y **EXACT MATCH** estarán al **100%**
- El resto de las métricas deben estar a 0.

block	metric	universe	numerator	denominator	pct	samp
KPIS	Unique IDs	REF	1572	-	-	-
KPIS	Unique IDs	NEW	1572	-	-	-
KPIS	Total rows REF	ROWS	1572	-	-	-
KPIS	Total rows NEW	ROWS	1572	-	-	-
KPIS	Total diff(new-ref)	ROWS	0	1572	0.0000%	-
KPIS	Global quality	REF	1572	1572	100.0000%	-
EXACT MATCH	1:1 (all columns)	BOTH	1572	1572	100.0000%	> BR_008
PARTIAL MATCH	1:1 (match & no_match co...	BOTH	0	1572	0.0000%	-
GAP	1:0 (only in ref)	REF	0	1572	0.0000%	-
GAP	0:1 (only in new)	NEW	0	1572	0.0000%	-
DUPS	duplicates (both)	BOTH	0	1572	0.0000%	-
DUPS	duplicates (only in ref)	REF	0	1572	0.0000%	-
DUPS	duplicates (only in new)	NEW	0	1572	0.0000%	-

- **Differences:** solo veríamos **EXACT\_MATCH** si la opción de guardarlos estuviera activada; si no, la tabla queda vacía.

	column	value_ref	value_n...	results
1	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
2	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
3	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
4	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
5	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
6	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
7	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
8	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
9	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
0	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
1	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
2	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
3	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
4	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH
5	ALL_COLUMNS	ALL_MATCH	ALL_MATCH	EXACT_MATCH

**Duplicates:** debe estar vacía.

## Caso 2 — Escenario típico (funcionalmente correcto) (2025-09-24)

- **Global Quality ≈ 99.9%:** no llega al 100% porque hay duplicados.
- Se observan **8 IDs únicos** con ejemplos duplicados **en ambos orígenes** (ver detalle y tipología en results\_duplicates).

block	metric	universe	numerator	denominator	pct	samp
KPIS	Unique IDs	REF	55058	-	-	-
KPIS	Unique IDs	NEW	55058	-	-	-
KPIS	Total rows REF	ROWS	55841	-	-	-
KPIS	Total rows NEW	ROWS	55841	-	-	-
KPIS	Total diff(new-ref)	ROWS	0	55841	0.0000%	-
KPIS	Global quality	REF	55050	55058	99.9855%	-
EXACT MATCH	1:1 (all columns)	BOTH	55058	55058	100.0000%	> BJJ_23
PARTIAL MATCH	1:1 (match & no_match co...	BOTH	0	55058	0.0000%	-
GAP	1:0 (only in ref)	REF	0	55058	0.0000%	-
GAP	0:1 (only in new)	NEW	0	55058	0.0000%	-
DUPS	duplicates (both)	BOTH	8	55058	0.0145%	> FRK_N
DUPS	duplicates (only in ref)	REF	0	55058	0.0000%	-
DUPS	duplicates (only in new)	NEW	0	55058	0.0000%	-

- **Differences:** muestra PARTIAL\_MATCH (si procede) y, si la opción está activa, también EXACT\_MATCH.
- **Duplicados:** muestra los duplicados por origen (ref | new) y el identificador que hayamos decidido.

origin	id	exact_duplicates	dups_vs_variations	occurrences	variations
new	MLN_NOT PROVIDED_103_0812_IJT_NULL	0	170	171	> bic_1: [BPPITR8COPBITMMJSUNITMM]   sender_bic: [BPPITR8COPBITMMJSUNITMM]   session_sequence: [02bf085-7527-4a85-a3da-a75fe117c6b034ed58-6f76-4395-a0e9... > bic_1: [BACRIT22BAPIT22BAPIT22BAPIT22BPPITR8COPBITMMJSUNITMM]   receiver_bic: [BACRIT22BAPIT22BAPIT22BPPITR8COPBITMMJSUNITMM]   transaction_id: [02bf085-7527-4a85-a3da-a75fe117c6b034ed58-6f76-4395-a0e9... > bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   transaction_instructed_amount_eur: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]
new	MLN_NOT PROVIDED_103_0812_IJT_NULL	0	578	579	> bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]
new	MLN_RCT_NULL_0614_IJT_SCL	0	1	2	> bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]
new	FRK_NOT PROVIDED_103_2012_I_DE_NULL	0	1	2	> bic_1: [AARBEDESW0001MBKDEM0000]   sender_bic: [AARBEDESW0001MBKDEM0000]   session_sequence: [8cc6b3f1-b3be-4bdf-a28c-079034eff198.92d467f-ce7a-457e-ba7... > bic_1: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   sender_bic: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   session_sequence: [36e7a2-16ab-4582... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
new	MLN_NOT PROVIDED_103_1408_IJT_NULL	0	10	11	> bic_1: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   sender_bic: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   session_sequence: [36e7a2-16ab-4582... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
new	> PT_f65f7062-235d-41df-8367-87ad17d6...	0	1	2	> bic_1: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   sender_bic: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   session_sequence: [00de965d-9dda-4412-a2b9-42b97a34bfb0cf.. > bic_1: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   sender_bic: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   session_sequence: [00de965d-9dda-4412-a2b9-42b97a34bfb0cf..
new	MLN_NOT PROVIDED_103_1508_IJT_NULL	0	16	17	> bic_1: [DEUTDEFF_PASCTIMURZSBf128]   sender_bic: [DEUTDEFF_PASCTIMURZSBf128]   session_sequence: [2a3a91fc-2074-477d-8798-efbcb3a9c20f.4071f1ea-dcbb-4389-8440... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
new	MLN_NOT PROVIDED_103_1011_IJT_NULL	0	6	7	> bic_1: [DEUTDEFF_PASCTIMURZSBf128]   sender_bic: [DEUTDEFF_PASCTIMURZSBf128]   session_sequence: [2a3a91fc-2074-477d-8798-efbcb3a9c20f.4071f1ea-dcbb-4389-8440... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
ref	> PT_f65f7062-235d-41df-8367-87ad17d6...	0	1	2	> bic_1: [DEUTDEFF_PASCTIMURZSBf128]   sender_bic: [DEUTDEFF_PASCTIMURZSBf128]   session_sequence: [2a3a91fc-2074-477d-8798-efbcb3a9c20f.4071f1ea-dcbb-4389-8440... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
ref	MLN_NOT PROVIDED_103_1011_IJT_NULL	0	6	7	> bic_1: [DEUTDEFF_PASCTIMURZSBf128]   sender_bic: [DEUTDEFF_PASCTIMURZSBf128]   session_sequence: [2a3a91fc-2074-477d-8798-efbcb3a9c20f.4071f1ea-dcbb-4389-8440... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
ref	MLN_RCT_NULL_0614_IJT_SCL	0	1	2	> bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]
ref	MLN_NOT PROVIDED_103_1408_IJT_NULL	0	10	11	> bic_1: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   sender_bic: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   session_sequence: [36e7a2-16ab-4582... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
ref	MLN_NOT PROVIDED_103_1508_IJT_NULL	0	16	17	> bic_1: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   sender_bic: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   session_sequence: [00de965d-9dda-4412-a2b9-42b97a34bfb0cf.. > bic_1: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   sender_bic: [BVRUEBCJCRATIRJSUNITMMJVRAULT2V]   session_sequence: [00de965d-9dda-4412-a2b9-42b97a34bfb0cf..
ref	MLN_NOT PROVIDED_103_0812_IJT_NULL	0	578	579	> bic_1: [BPPITR8COPBITMMJSUNITMM]   sender_bic: [BPPITR8COPBITMMJSUNITMM]   session_sequence: [02bf085-7527-4a85-a3da-a75fe117c6b034ed58-6f76-4395-a0e9... > bic_1: [BACRIT22BAPIT22BAPIT22BPPITR8COPBITMMJSUNITMM]   receiver_bic: [BACRIT22BAPIT22BAPIT22BPPITR8COPBITMMJSUNITMM]   transaction_id: [02bf085-7527-4a85-a3da-a75fe117c6b034ed58-6f76-4395-a0e9... > bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]
ref	MLN_NOT PROVIDED_103_2012_I_DE_NULL	0	170	171	> bic_1: [AARBEDESW0001MBKDEM0000]   sender_bic: [AARBEDESW0001MBKDEM0000]   session_sequence: [8cc6b3f1-b3be-4bdf-a28c-079034eff198.92d467f-ce7a-457e-ba7... > bic_1: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   sender_bic: [BVAESMM_CRAAT2RDEUTDEFF_PASCTIMURVUALT2V]   session_sequence: [36e7a2-16ab-4582... > session_sequence: [0944784720.9944784722]   bic_2: [CGDITPXXXXCHASUS33XXXX]   receiver_bic: [CGDITPXXXXCHASUS33XXXX]   mur: [TOTAS059241604240_TOTAS0592416...
ref	FRK_NOT PROVIDED_103_2012_I_DE_NULL	0	1	2	> bic_1: [BPPITR8COPBIT2P]   sender_bic: [BPPITR8COPBIT2P]   session_sequence: [000000499/025924061405772]   priority: [BPPITR8COPBIT2P]   transaction_instructed_amount: [5.75000.5.96000]   ordering_customer_name: [5.75.96.0]   ordering_institution_bic: [CALAMANDEI VALTER TUMARELLI MARTIN NACELINA]   ordering_institution_name: [BPPITR8COPBIT2P]   source: [SEPA_BB1C0D90A83F41628A8DFA096250DSEPA_91883A34623263EBC53CB63C4B]

- **Duplicados exactos:** filas idénticas.
- **Duplicados con variación:** el identificador coincide, pero alguna(s) columna(s) difiere(n). Es habitual cuando algún campo usado como identificador viene **NULL**.

La columna **variaciones** muestra el detalle de la diferencia por columna, p. ej.:  
 bic\_1: [valor\_1, valor\_2] (y así sucesivamente para el resto de campos afectados).

	<b>exact_duplicates</b>	<b>dups_w_variations</b>	<b>occurrences</b>	<b>variations</b>
1 PROVIDED_103_08:12_I_IT_NULL	0	170	171	> bic_1: [BPPIITRR,CIPBITMM,ISUN
2 ?ROVIDED_103_08:12_I_IT_NULL	0	578	579	> bic_1: [BACRIT22,BAPPIT22,BITA
3 _NULL_06:14_I_IT_SCL	0	1	2	✓ bic_1: [BPPIITRR,CRPPIT2P]   sen [BPPIITRRXXX,CRPPIT2P]   trans ordering_institution_bic: [CALAN [SEPA_881C03D960A83F41B28A

### 3) Definición de la clave compuesta (ejecuciones mostradas)

Se utilizó la siguiente **clave compuesta**:

geo, uetr, type, time, in\_out, segregation, sender\_reference

**Ejemplo de identificador** (formado con la clave anterior):

BR\_0007589b-71e8-4d5c-9f69-134e9037631a\_103\_09:14\_0\_3565729148

---

### 4) Opciones y comportamiento

- **checkDuplicates:** si está **desactivado**, la evaluación de duplicados se omite y results\_duplicates queda **vacía incluso si existen duplicados**.
- **includeEqualsInDiff:** si está **activado**, results\_differences incluirá también los **EXACT\_MATCH** y los **MATCH** de los **PARTIAL MATCH**. Si está **desactivado**, esa parte no se escribe y la tabla puede quedar vacía (en igualdad total).

## 5) Configuración actual en Variable de Airflow (PRO)

- **Tablas a comparar:** Se definen refTable (tabla de referencia) y newTable (tabla nueva).
- **Rangos de ventanas independientes por tabla**, refWindowDays y newWindowDays s. ejemplos:
  - "refWindowDays": "0..0", Date actual
  - "newWindowDays": "-1..0", Desde d actual hasta d-1
  - "newWindowDays": "-1..+1", Desde d actual hasta d-1 y d+1
  - "newWindowDays": "0..+2", Desde d actual hasta d+2

*Todas las posibilidades asc y desc para comparar ventanas independientes por tabla.*

- **Clave compuesta (compositeKeyCols):** Define los campos que forman el identificador único.
- **Columnas a ignorar (ignoreCols):** Campos que no se tienen en cuenta en la comparación.
- **Opciones funcionales:**
  - checkDuplicates: Indica si se debe evaluar la existencia de duplicados (establecido como true en la configuración actual).
  - includeEqualsInDiff: Indica si se deben incluir los EXACT\_MATCH (coincidencias perfectas) en la tabla de diferencias.

```
AMLTOOLS_tables_comparator_swift {
    "refTable": "scib_cm_cmplnc_trans_messages_swift_s3.swift_transactions",
    "newTable": "scib_bu_cmplnc_trans_messages_swift_s3.swift_transactions_new",
    "initiativeName": "swift",
    "tablePrefix": "scib_bu_cmplnc_trans_messages_swift_s3.results_",
    "outputBucket": "s3a://scib-pro-bu-aml/internal-tools",
    "partitionSpec": "geo=*/data_date_part=YYYY-MM-dd/",
    "refWindowDays": "0..0",
    "newWindowDays": "-1..0",
    "compositeKeyCols": ["geo", "uetr"],
    "ignoreCols": ["original_message", "data_timestamp_part"],
    "checkDuplicates": true,
    "includeEqualsInDiff": true
}
```

**Observación:** En el ejemplo de ejecuciones se amplió la clave efectiva (*geo*, *uetr*, *type*, *time*, *in\_out*, *segregation*, *sender\_reference*). Ajusta compositeKeyCols cuando necesites esa granularidad.

---

## 6) Configuración utilizada en Postman (ejemplo)

POST: [https://hybridbalancer.scib.gs.corp/v2/execution/aml\\_internal\\_tools](https://hybridbalancer.scib.gs.corp/v2/execution/aml_internal_tools)

```
{  
  "preferences": [ "azure_databricks_v2", "cloudera"],  
  "size": "S",  
  "parameters": [  
    "refTable=scib_cm_cmplnc_trans_messages_swift_s3.swift_transactions",  
    "newTable=scib_cm_cmplnc_trans_messages_swift_s3.swift_transactions",  
    "initiativeName=swift",  
    "tablePrefix=scib_bu_cmplnc_trans_messages_swift_s3.results_",  
    "outputBucket=s3a://scib-pro-bu-aml/internal-tools",  
    "executionDate=2025-09-30",  
    "partitionSpec=data_date_part=YYYY-MM-dd/",  
  
    "refWindowDays=0..0",  
  
    "newWindowDays=-1..0",  
    "compositeKeyCols=geo,uetr,type,time,in_out,segregation,sender_reference"  
,  
    "ignoreCols=original_message,data_timestamp_part",  
    "checkDuplicates=true",  
    "includeEqualsInDiff=true"  
  ]  
}
```

---

## 7) Lectura rápida / Checklist

1. Mira **results\_summary** para el estado general: Unique IDs, Total rows, Total diff, Global Quality, y usa **Samples por métrica (hasta 6 IDs)** para pinchar ejemplos al vuelo en results\_differences o results\_duplicates.
2. Abre **results\_differences** para el **detalle de diferencias** (y exactos si los guardas).

3. Revisa **results\_duplicates** para ver **dónde** y **cómo** se duplcan los identificadores (both / duplica REF / duplica NEW, exactos o con variación).
  4. Recuerda que **duplicados afectan a Global Quality** y que **las opciones de configuración** pueden hacer que algunas tablas aparezcan vacías intencionalmente.

## 8) Glosario rápido

- **REF / NEW:** tabla de referencia / tabla nueva a comparar.
  - **EXACT\_MATCH:** coincidencia 1:1 perfecta (todas las columnas).
  - **PARTIAL\_MATCH:** coincidencia 1:1 con diferencias en alguna columna.
  - **GAP:** elementos presentes en una tabla y ausentes en la otra (se listan ejemplos).
  - **Duplicados exactos:** filas 100% idénticas.
  - **Duplicados con variación:** mismo identificador con diferencias en otras columnas.
  - **Global Quality:** métrica global que no llega a 100% si hay duplicados.

## 9) Job Control-M

Fin