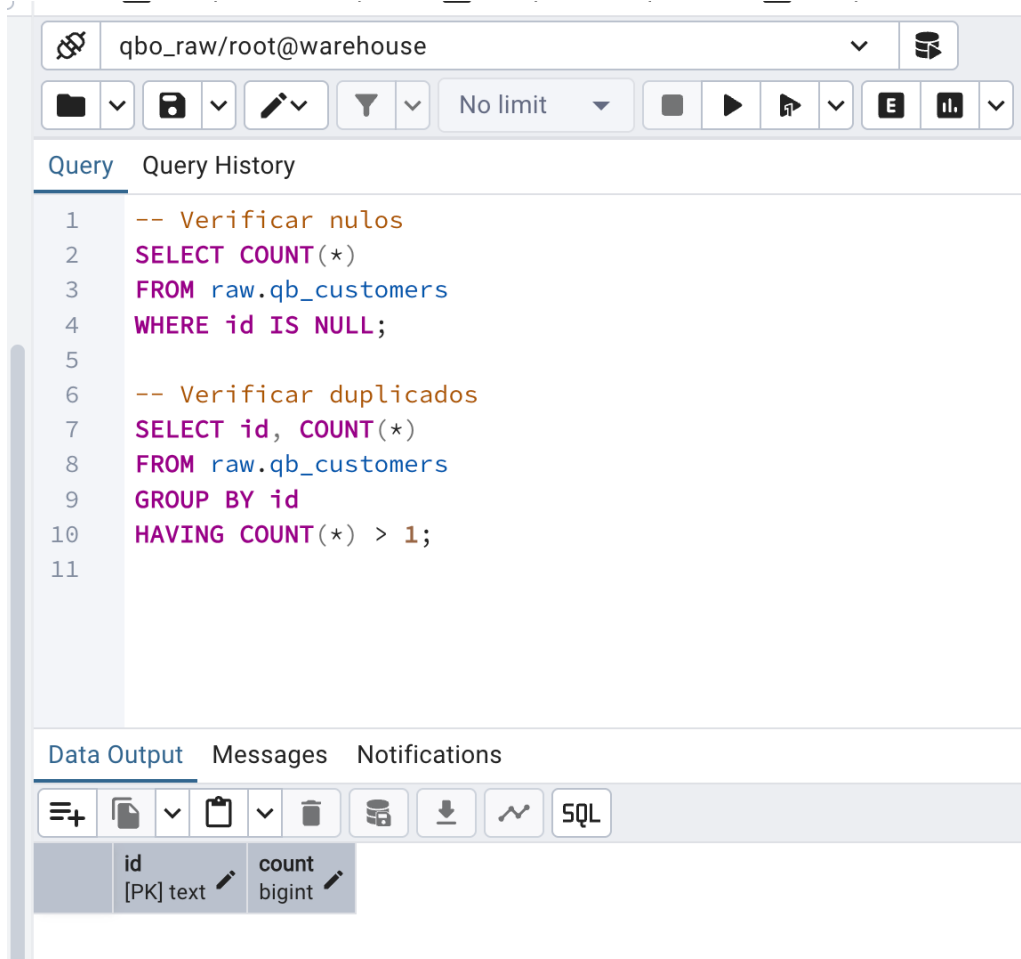


Pruebas/validaciones de calidad y guía para ejecutarlas

Con el objetivo de garantizar la calidad de los datos cargados en el esquema raw, se definieron y ejecutaron las siguientes validaciones:

1. Integridad

Se verificó que las claves primarias (id) no contengan valores nulos ni duplicados.



The screenshot shows a SQL query editor interface. At the top, the connection is 'qbo_raw/root@warehouse'. Below the connection bar is a toolbar with icons for file operations, filters, and execution. The main area is titled 'Query' and contains two SQL queries. The first query checks for null values in the 'id' column of the 'raw.qb_customers' table. The second query checks for duplicate values in the 'id' column. The bottom of the interface shows tabs for 'Data Output', 'Messages', and 'Notifications'. The 'Data Output' tab is active, displaying a table with two columns: 'id' (type [PK] text) and 'count' (type bigint).

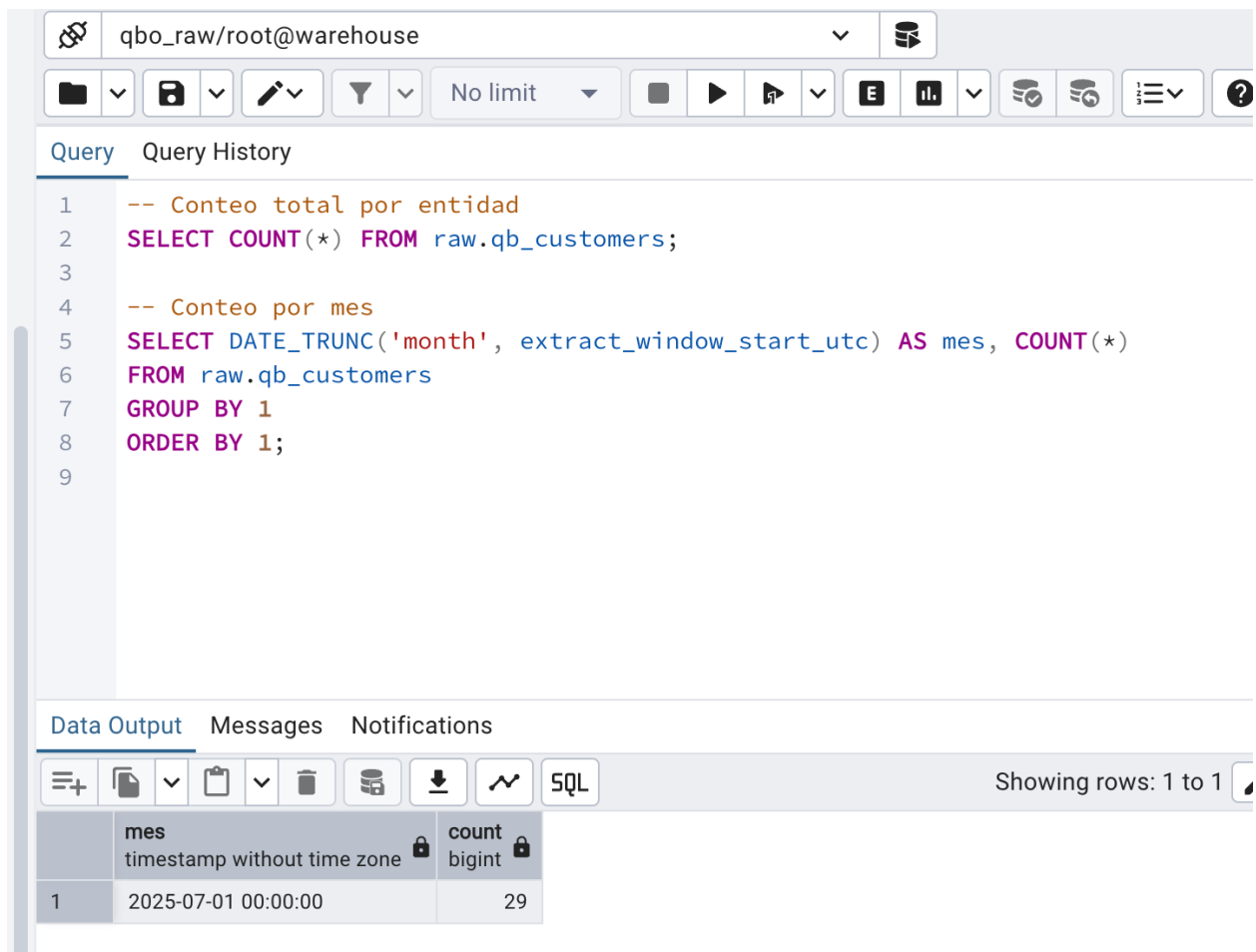
```
1  -- Verificar nulos
2  SELECT COUNT(*)
3  FROM raw.qb_customers
4  WHERE id IS NULL;
5
6  -- Verificar duplicados
7  SELECT id, COUNT(*)
8  FROM raw.qb_customers
9  GROUP BY id
10  HAVING COUNT(*) > 1;
11
```

id	count
[PK] text	bigint

2. Volumetría

Se revisaron los conteos de registros por entidad y por tramo de fechas para confirmar que coincidan con los logs del data loader.

Como podemos observar, hay 29 filas únicas (las mismas que fueron cargadas desde la api)



```
1  -- Conteo total por entidad
2  SELECT COUNT(*) FROM raw.qb_customers;
3
4  -- Conteo por mes
5  SELECT DATE_TRUNC('month', extract_window_start_utc) AS mes, COUNT(*)
6  FROM raw.qb_customers
7  GROUP BY 1
8  ORDER BY 1;
9
```

	mes timestamp without time zone	count bigint
1	2025-07-01 00:00:00	29

3. Coherencia temporal

Se validó que los timestamps de ingesta (ingested_at_utc) y las ventanas de extracción (extract_window_start_utc y extract_window_end_utc) sean consistentes con los filtros utilizados en el data loader.

q... × raw.qb_items/qbo... × raw.qb_customers... × raw.backfill_report... × r

qbo_raw/root@warehouse

Query Query History

```
1 SELECT MIN(extract_window_start_utc), MAX(extract_window_end_utc)
2 FROM raw.qb_items;
3
```

Data Output Messages Notifications

Showing r

	min timestamp without time zone	max timestamp without time zone
1	2025-07-22 00:00:00	2025-08-04 23:59:59

Coinciden con los filtros determinados en el pipeline:

Global Variables			+ New
Press Enter or Return to save changes.			
	end_date	2025-08-04	
	start_date	2025-07-22	

Reprocesos seguros (idempotencia)

Se volvió a ejecutar un mismo tramo de fechas y se verificó que el número de registros en la tabla raw no cambió, confirmando que la carga es idempotente. Además, en la tabla de reportes backfill_report aparecen múltiples ejecuciones del mismo rango con estado success.

Reporte de volumetría por tramo y evidencia de idempotencia

Al ejecutar el bloque **data loader**, es posible observar en los logs cómo avanza el proceso: la posición inicial, el número de intentos y la cantidad de *chunks* necesarios para procesar toda la información según el rango de fechas definido.

```
PY DATA LOADER ingest_qb_customers ←X Edit parents

Nuevo access_token obtenido vía refresh_token
Chunk 1: 2025-07-01T00:00:00Z → 2025-07-07T23:59:59Z
Request API: 2025-07-01T00:00:00Z → 2025-07-07T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 1 devolvió 0 customers
Chunk 2: 2025-07-08T00:00:00Z → 2025-07-14T23:59:59Z
Request API: 2025-07-08T00:00:00Z → 2025-07-14T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 2 devolvió 0 customers
Chunk 3: 2025-07-15T00:00:00Z → 2025-07-21T23:59:59Z
Request API: 2025-07-15T00:00:00Z → 2025-07-21T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 3 devolvió 0 customers
Chunk 4: 2025-07-22T00:00:00Z → 2025-07-28T23:59:59Z
Request API: 2025-07-22T00:00:00Z → 2025-07-28T23:59:59Z, posición=1, intento 1
Posición 1 → 7 customers
Chunk 4 devolvió 7 customers
Chunk 5: 2025-07-29T00:00:00Z → 2025-08-04T23:59:59Z
Request API: 2025-07-29T00:00:00Z → 2025-08-04T23:59:59Z, posición=1, intento 1
Posición 1 → 22 customers
Chunk 5 devolvió 22 customers
Chunk 6: 2025-08-05T00:00:00Z → 2025-08-11T23:59:59Z
Request API: 2025-08-05T00:00:00Z → 2025-08-11T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 6 devolvió 0 customers
Chunk 7: 2025-08-12T00:00:00Z → 2025-08-18T23:59:59Z
Request API: 2025-08-12T00:00:00Z → 2025-08-18T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 7 devolvió 0 customers
Chunk 8: 2025-08-19T00:00:00Z → 2025-08-25T23:59:59Z
Request API: 2025-08-19T00:00:00Z → 2025-08-25T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
```

Cuando la ejecución finaliza, se imprime la cantidad total de registros procesados (ya sea *customers*, *invoices* o *items*).

```

Chunk 7: 2025-08-12T00:00:00Z → 2025-08-18T23:59:59Z
Request API: 2025-08-12T00:00:00Z → 2025-08-18T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 7 devolvió 0 customers
Chunk 8: 2025-08-19T00:00:00Z → 2025-08-25T23:59:59Z
Request API: 2025-08-19T00:00:00Z → 2025-08-25T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 8 devolvió 0 customers
Chunk 9: 2025-08-26T00:00:00Z → 2025-08-31T23:59:59Z
Request API: 2025-08-26T00:00:00Z → 2025-08-31T23:59:59Z, posición=1, intento 1
Posición 1 → 0 customers
Chunk 9 devolvió 0 customers
Total customers: 29

```

Posteriormente, en el bloque **transformer**, al convertir los datos en un DataFrame, en la esquina inferior izquierda de la tabla se puede confirmar el número total de filas, lo que permite validar la congruencia entre lo extraído y lo transformado.

The screenshot shows a data table with 29 rows and 8 columns. The first few rows are visible, showing customer IDs, active status, and various timestamps. The bottom left corner of the table indicates '29 rows x 8 columns'.

Index	Customer ID	Active	Timestamp 1	Timestamp 2	Timestamp 3	Value 1	Value 2	Metadata
0	4	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
1	6	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
2	11	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
3	14	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
4	19	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
5	27	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
6	28	"Active": true,	2025-09-12T22:50:49.010141	2025-07-22T00:00:00Z	2025-07-28T23:59:59Z	1	100	{ "headers": {
7	1	"Active": true,	2025-09-12T22:50:49.010141	2025-07-29T00:00:00Z	2025-08-04T23:59:59Z	1	100	{ "headers": {
8	2	"Active": true,	2025-09-12T22:50:49.010141	2025-07-29T00:00:00Z	2025-08-04T23:59:59Z	1	100	{ "headers": {

De igual manera, en el bloque **exporter**, se muestra cuántas filas fueron exportadas y a qué tabla específica en la base de datos se insertaron.

```

Exportando 29 customers a raw.qb_customers en warehouse:5432/qbo_raw
Exportación completada: 29 customers procesados.

```

Finalmente, en **pgAdmin** se confirma que todas las filas fueron efectivamente exportadas.

Data Output Messages Notifications		
Showing rows: 1 to 29 Page No: 1 of 1		
	id [PK] text	payload jsonb
1	1	{ "id": "1", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 239.0, "Taxable": true, "WebAddr": null, "BillAddr": { "id": "2"
2	10	{ "id": "10", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 629.1, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "
3	11	{ "id": "11", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 0.0, "Taxable": false, "WebAddr": { "URI": "http://gevelberp
4	12	{ "id": "12", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 81.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "1
5	13	{ "id": "13", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 450.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "
6	14	{ "id": "14", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 0.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "14
7	15	{ "id": "15", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": { "FreeFormNumber": "(973) 555-2234", "domain": "QBO", "sparse": false, "Balance": 0.0, "Taxable": false, "
8	16	{ "id": "16", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 75.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "1
9	17	{ "id": "17", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 314.28, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "
10	18	{ "id": "18", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 954.75, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "
11	19	{ "id": "19", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 0.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "20
12	2	{ "id": "2", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 85.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "3", "
13	20	{ "id": "20", "Fax": null, "Job": false, "Level": null, "Active": true, "Mobile": null, "domain": "QBO", "sparse": false, "Balance": 226.0, "Taxable": false, "WebAddr": null, "BillAddr": { "id": "
Total rows: 29 Query complete 00:00:00.321		

Para verificar la **idempotencia**, se ejecuta nuevamente el pipeline para el mismo rango de fechas. En la tabla de reportes (backfill_report) se observa que existen dos ejecuciones con el mismo *window_start* y *window_end*, ambas exitosas. Al revisar nuevamente la tabla de la entidad en cuestión (por ejemplo, qb_customers), se confirma que el número total de filas se mantiene igual, lo que garantiza que no se generaron duplicados y que el proceso es efectivamente idempotente.

Data Output Messages Notifications								
Showing rows: 1 to 3 Page No: 1 of 1								
	id [PK] integer	pipeline_name text	entity text	row_count integer	window_start timestamp without time zone	window_end timestamp without time zone	started_at timestamp without time zone	ended_at timestamp without time zone
1	1	qb_customers_back...	custome...	29	2025-07-22 00:00:00	2025-08-04 23:59:59	2025-09-12 23:19:14.670834	2025-09-12 23:19:14.694807
2	2	qb_customers_back...	custome...	29	2025-07-22 00:00:00	2025-08-04 23:59:59	2025-09-12 23:43:32.020912	2025-09-12 23:43:32.084204
3	3	qb_customers_back...	custome...	29	2025-07-22 00:00:00	2025-08-04 23:59:59	2025-09-12 23:46:04.469444	2025-09-12 23:46:04.533462

			Showing rows: 1 to 29		Page No: 1	of 1
	id [PK] text	payload jsonb				
1	1	{\"Id\": \"1\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 239.0, \"Taxable\": true, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"2\", \"				
2	10	{\"Id\": \"10\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 629.1, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"1				
3	11	{\"Id\": \"11\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": {\"URI\": \"http://gevelberph				
4	12	{\"Id\": \"12\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 81.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"12				
5	13	{\"Id\": \"13\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 450.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"1				
6	14	{\"Id\": \"14\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"14				
7	15	{\"Id\": \"15\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": {\"FreeFormNumber\": \"(973) 555-2234\", \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"V				
8	16	{\"Id\": \"16\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 75.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"16				
9	17	{\"Id\": \"17\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 314.28, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"				
10	18	{\"Id\": \"18\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 954.75, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"				
11	19	{\"Id\": \"19\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"20				
12	2	{\"Id\": \"2\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 85.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"3\", \"				
13	20	{\"Id\": \"20\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 226.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"2				
14	21	{\"Id\": \"21\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 78.6, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"22				
15	22	{\"Id\": \"22\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"23				
16	23	{\"Id\": \"23\", \"Fax\": null, \"Job\": true, \"Level\": 1.0, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 274.5, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"24				
17	24	{\"Id\": \"24\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 362.07, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"				
18	25	{\"Id\": \"25\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 160.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"2				
19	26	{\"Id\": \"26\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 414.72, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"				
20	27	{\"Id\": \"27\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": null, \"BillAddr\": {\"Id\": \"28				
21	28	{\"Id\": \"28\", \"Fax\": null, \"Job\": false, \"Level\": null, \"Active\": true, \"Mobile\": null, \"domain\": \"QBO\", \"sparse\": false, \"Balance\": 0.0, \"Taxable\": false, \"WebAddr\": {\"URI\": \"http://www.dreamr				
Total rows: 29			Query complete 00:00:00.052			L