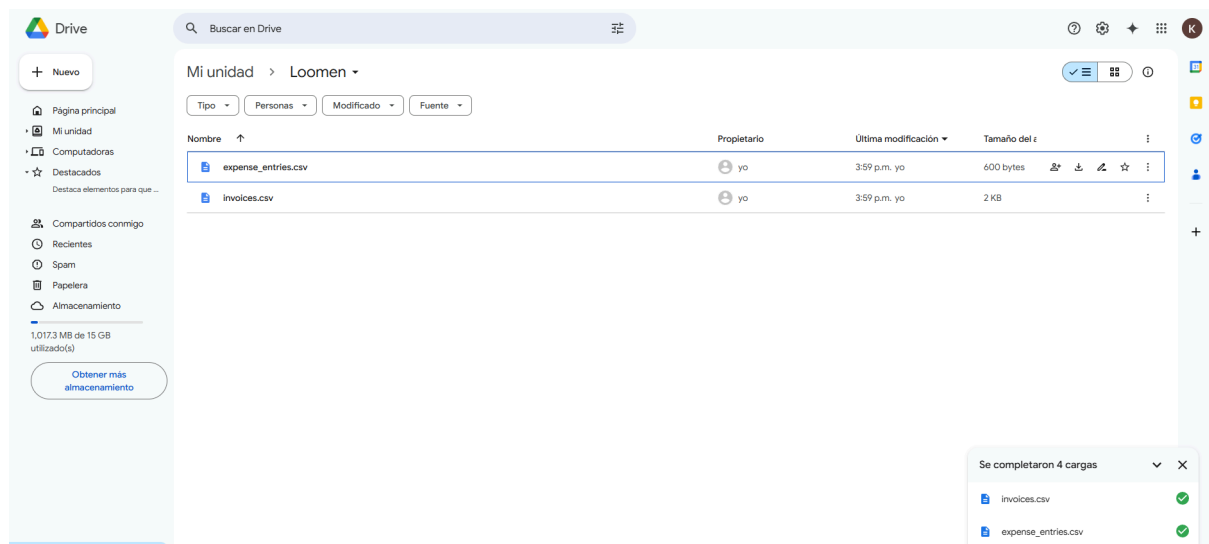


Project Documentation: Automated Data Pipeline and Dashboard Integration

1. CSV File Management

The process begins with the extraction of CSV files containing raw data. These files are uploaded and stored in a structured folder system within Google Drive to maintain centralized access and data version control.



2. Coupler.io Data Automation

Two **Data Flows** were configured using Coupler.io, a tool that automates data transfers between applications:

- The data source is **Google Drive**, where the CSV files are stored.
- The destination is **Google BigQuery**, where a cloud data warehouse is created.
- Each data flow is configured to run automatically **every hour from Monday to Friday, between 9:00 AM and 6:00 PM**.

Coupler.io

Personal

Home

Data flows

Sources

Destinations

My dashboards

Templates

Settings

7 days left on your Business trial before you return to the free plan. Upgrade now!

UPGRADE

Data flows

+ CREATE

Search for data flows, apps

Data flow	Status	Actions
<div><div></div><div>Loomen Expense Entries</div><div>2 data connections, 2 data sources, 1 destination</div></div> <div>Completed</div> <div>25 minutes ago • Next in 18 minutes</div>		<div></div> <div></div>
<div><div></div><div>Loomen_Invoice</div><div>1 data connection, 1 data source, 1 destination</div></div> <div>Completed</div> <div>16 minutes ago • Schedule</div>		<div></div> <div></div>

Need help?

Contact support

Book onboarding

Drive

Nuevo

Página principal

Mi unidad

Computadoras

Destacados

Compartidos conmigo

Recientes

Spam

Papelera

Almacenamiento

Buscar en Drive

Mi unidad > Loomen

Tipo Personas Modificado Fuente

Nombre	Propietario	Última modificación	Tamaño del c
expense_entries.csv	yo	3:59 p.m. yo	600 bytes
invoices.csv	yo	3:59 p.m. yo	2 KB

1.017.3 MB de 15 GB utilizado(s)

Obtener más almacenamiento

Se completaron 4 cargas

invoices.csv

expense_entries.csv

Import was successful

Now you can work with the file.

Dismiss

VIEW RESULTS

Settings

Overview History Webhooks

Loomen_Invoice

1 data connection, data source, destination

Created by: keineruniversidad@gmail.com

Last run

Next run

Last modified

in a few seconds

in 18 minutes

in a few seconds

Automatic data refresh

Interval

Days of week

Time preferences

Timezone

Need help?

Contact support

Book onboarding

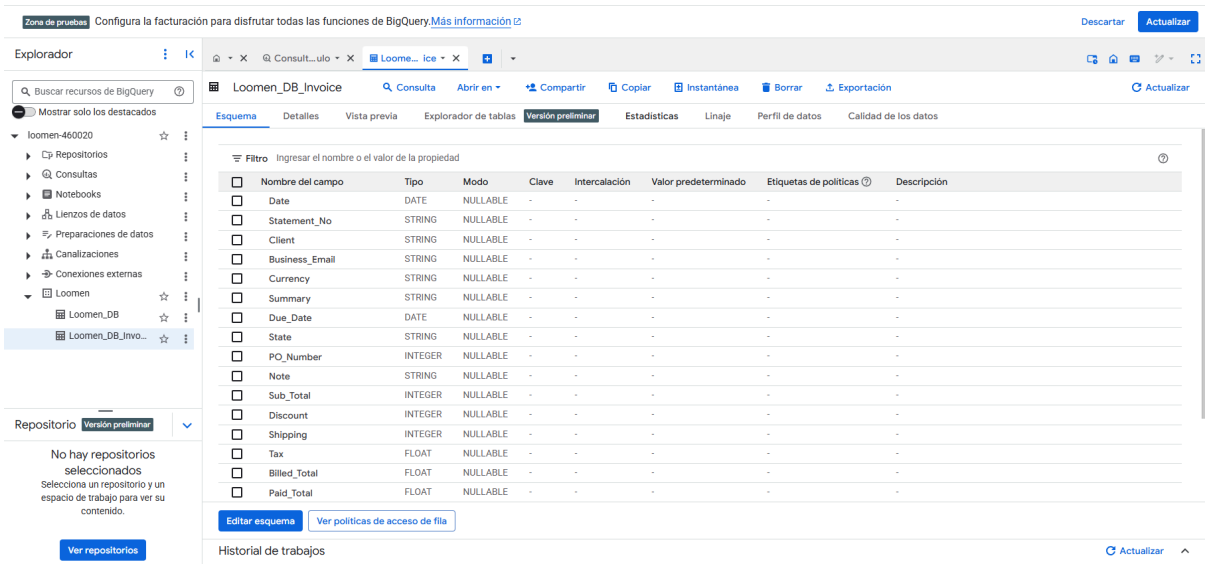
BACK TO DESTINATIONS SAVE AND CLOSE SAVE AND RUN

Using Coupler.io, we created a flow that exports data from our Google connections to BigQuery, thus creating a data warehouse. In BigQuery, we can use SQL to perform the ETL process, but since we have a small amount of data, we will use Power Query directly within Power BI.

This setup ensures timely ingestion of the latest data without the need for manual uploads or intervention.

3. Data Warehouse in BigQuery

A centralized **data warehouse** was created in **Google BigQuery** to store and manage the structured data imported from the CSV files via Coupler.io. This warehouse supports advanced querying, scalable storage, and integration with business intelligence tools.



4. Automated Updates

Any changes or updates to the original CSV files in Google Drive (e.g., added records, corrections, or new uploads) are automatically reflected in the BigQuery warehouse, as the Coupler.io data flows are set to refresh on an **hourly schedule** during business days.

5. Data Structure and Organization

In BigQuery, datasets are organized into clear table structures to support analytics. Each table represents a specific data source or metric. Fields are standardized and pre-validated to reduce errors during analysis.

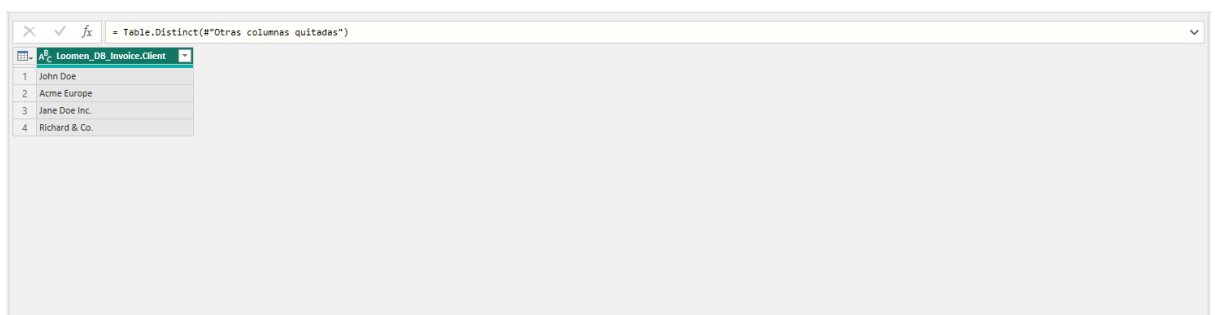
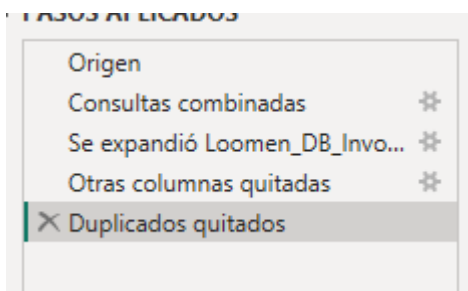
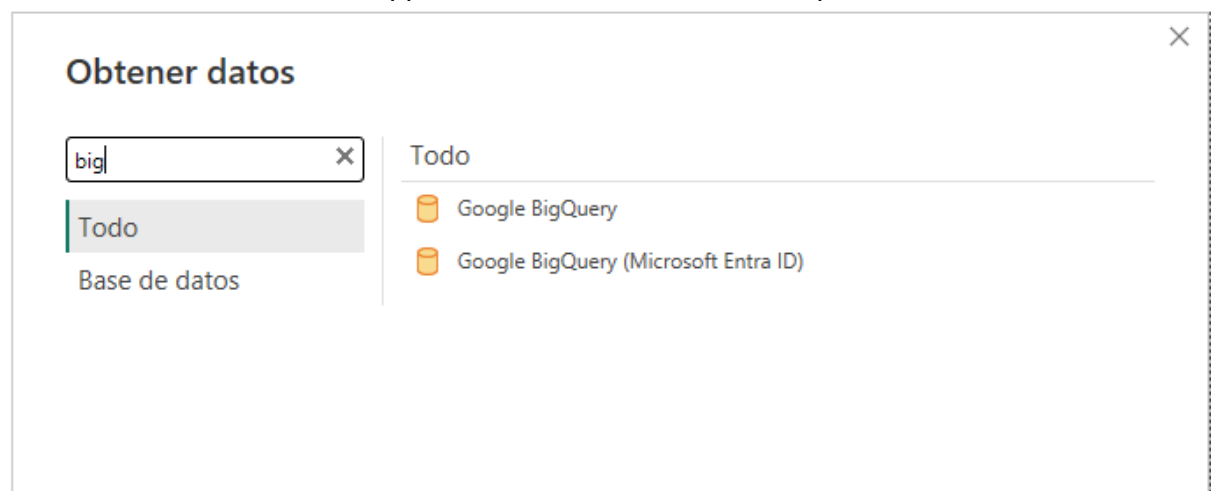
6. Scalability and Maintenance

The current system allows for additional data sources to be added easily. Each new CSV source can be integrated by setting up an additional data flow in Coupler.io, following the same logic, and mapped to new or existing tables in BigQuery. This makes the infrastructure scalable and adaptable to the client's evolving data needs.

7. Power BI Integration for Data Visualization

After the data is securely stored in BigQuery, it is connected to **Power BI** for visualization and reporting:

- We load data from both tables already created in BigQuery into Power BI.
- We perform the ETL process in Power Query. Since the data volume is low, we only confirm that the data is in the correct format.
- We create a table where we append client data and remove duplicates.



a. Direct Connection to BigQuery

- Power BI uses the built-in Google BigQuery connector.
- Data can be imported or queried live via DirectQuery.

b. Selecting Tables and Fields

- Specific tables or views are selected for reporting.
- Data can be transformed or filtered either in BigQuery or within Power BI for performance optimization.

c. Dashboard Creation

- Custom dashboards are created using KPIs, charts, tables, and maps.
- Reports support filtering by dimensions such as time, category, or region, based on business requirements.

d. Data Refresh

- Power BI refreshes data in sync with BigQuery updates.
- Automatic refresh is scheduled **hourly on weekdays between 9:00 AM and 6:00 PM**.

e. Secure Sharing

- Dashboards are published on Power BI Service.
- Stakeholders receive access based on defined roles and permissions to maintain data security and compliance.

8. Date Table Creation Using DAX in Power BI

We created a dedicated **Date Table** in Power BI using a DAX formula to generate a comprehensive calendar. This calendar table includes all relevant dates and is connected

(related) to the date fields in our databases to enable accurate time-based analysis and filtering.

The Date Table supports:

- Time intelligence calculations (YTD, QTD, MTD, etc.)
- Filtering and slicing reports by days, weeks, months, quarters, and years
- Ensuring consistency across all date-related data points

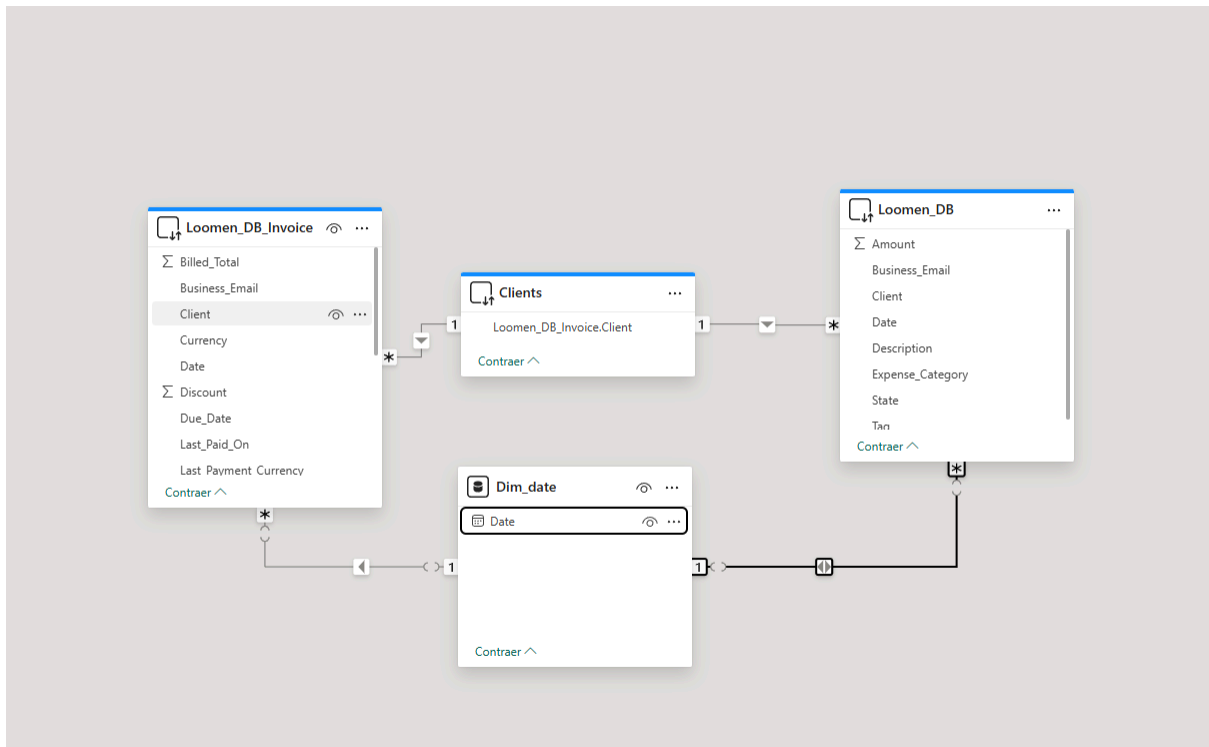
By establishing relationships between this calendar table and the date columns in our datasets, we enable more powerful and flexible reporting capabilities in Power BI.



9. Table Relationships

We established relationships between our tables to ensure data integrity and enable cross-table analysis. In this project, we decided **not to use the star schema model**. Instead, we designed the relationships based on the specific needs and structure of our datasets.

This approach allows us to maintain flexibility and simplicity in the data model while supporting the necessary queries and reports in Power BI.



10. Expense Control Dashboard

We created an **Expense Control Dashboard** in Power BI to provide clear insights regarding sales performance. The dashboard includes:

- **Total Paid** and **Outstanding Balance** metrics
- **Client filters** to view individual customer performance
- **Date filters** for dynamic time-based analysis
- **Clear KPIs** to quickly assess payment status and financial progress

This dashboard helps stakeholders monitor financial data efficiently and make informed decisions based on up-to-date, visual insights.



11. Invoice Overview – Pivot Table

In another tab, we included a **Pivot Table** that provides a detailed general overview of all invoices. This table allows users to:

- Drill down into invoice data with multiple levels of detail
- Analyze billing information by client, date, or status
- Gain a structured summary of all financial activity related to invoicing

This dynamic view enhances visibility and supports better financial tracking and reporting.

Loomen_DB_Invoice.Client	Billed_Total	Paid_Total	Total_Due	Discount	Shipping
<div> <div></div> <div>John Doe</div> </div> <div> <div>partial</div> <div>draft</div> <div>sent</div> </div>	\$65.930,88 \$31.792,94 \$2.345 \$31.792,94	\$20.000,94 \$20.000,94 \$0 \$0	\$45.929,94 \$11.792,00 \$2.345,00 \$31.792,94	\$334 \$67 \$200 \$67	99 12 75 12
<div> <div></div> <div>Jane Doe Inc.</div> </div> <div> <div>paid</div> <div>sent</div> <div>viewed</div> </div>	\$9.505,75 \$1.500 \$2.455,75 \$5.550	\$1.500 \$1.500 \$0 \$0	\$6.697,55 \$0,00 \$1.147,55 \$5.550,00	\$235 \$0 \$35 \$200	0 0 0 0
<div> <div></div> <div>Acme Europe</div> </div> <div> <div>draft</div> </div>	\$12.505 \$12.505	\$0 \$0	\$12.505,00 \$12.505,00	\$0 \$0	0 0
<div> <div></div> <div>Richard & Co.</div> </div> <div> <div>draft</div> <div>sent</div> </div>	\$29.914,3 \$2.494,3 \$27.420	\$0 \$0 \$0	\$29.914,30 \$2.494,30 \$27.420,00	\$700 \$200 \$500	996 676 320
Total	\$117.855,93	\$21.500,94	\$95.046,79	\$1.269	1095