Album Financiero: A Full-Scale Personal Financial Analysis and Reporting Solution using Power BI

Author:
Juan Carlos Larios Coronado

Date:
August 2025

Tools Used:
Power BI, Microsoft Excel, Power Query, DAX, Data Modeling, Google Data Analytics Framework, Open AI.

Project Type:
Capstone Project – Personal Portfolio

Contentsp

Analysis Framework

Introduction

This project presents a personal budget analytics report developed using Excel. The objective is to capture the steps taken to transform an initially unstructured and manually managed dataset into a fully optimized, structured, and automated data model. This dataset contains over 10 years of personal financial records, originally maintained without formal data cleaning, normalization, or standardized categories. This documentation outlines the process of working with inconsistent data entry, missing data and lack of data integrity, through the design of a clean, normalized structure ready for analysis. The project applies the complete data analysis process to demonstrate core skills in data preparation, transformation, modeling, and visualization using Excel and Power BI.

I started my personal budget "Album Financiero" as soon as I was accepted in the job I was looking for in 2015. I was going to start receiving a bigger and constant income and did not want to lose control of my personal finances and starting to set myself goals for the next years, starting a spreadsheet was the best option.

Project Background

The analyst identified an opportunity to transform over a decade of manually maintained personal financial records into a comprehensive data analytics case study for portfolio demonstration. The project began with an unstructured Excel file with challenges typical of real-world datasets. Rather than presenting raw monetary values, the analysis emphasizes anonymized, proportional metrics such as percentages, ratios, and categorized summaries. This approach preserves data privacy while effectively demonstrating analytical methodologies, data modeling techniques, and dashboard development.

Over the last 10 years, the spreadsheet has had many changes to upgrade the quality of the metrics that were obtained from the expenses added into the document. For this project I wanted to finally show a professional dashboard that summarizes the important metrics of my data that are distributed in the spreadsheet, simplifying the capture of the data and using my knowledge obtained this last year from my courses of Data Analyst.

A clear problem statement / research question

The primary goal of this project is to design and build an interactive dashboard that provides clear visibility into spending patterns, saving habits, and category trends over the past decade.

This project addresses two key problems:

1. The original financial tracking system — a manually maintained Excel file — lacked structure, was prone to errors, and did not support meaningful analysis beyond basic summaries.
2. The absence of an efficient, user-friendly system for both entering data and generating actionable financial insights in real time.

The solution aims to optimize both the data management process (through structured, clean, and scalable datasets) and the analytical output (via interactive reports and dashboards). This enables smarter financial decision-making while serving as a comprehensive demonstration of data analysis, data modeling, and visualization skills.

Project Objectives

Develop a clean, scalable, and user-friendly financial tracking system that powers a responsive Power BI dashboard and supporting analytical reports, summarizing key spending, income, and budgeting trends over the past 10 years.

- Normalize and restructure core Excel tables of the “Album Financiero” spreadsheet to follow analytical best practices, ensuring long-term scalability and reliability.
- Enhance Excel usability by minimizing ongoing manual data entry and transitioning most calculated fields and aggregations into the Power BI data model using DAX measures.
- Build a comprehensive Power BI solution that includes:
    - Interactive analytical reports for exploring detailed trends in spending, savings, income, debt, investments, and budgeting performance.
    - A high-level dashboard summarizing key financial indicators.
- Design a mobile-optimized version for quick expense lookup and high-level metrics.
- Document the full transformation process as a portfolio case study that demonstrates end-to-end skills in data cleaning, analysis, and visualization.

My goal with this project is to realize the full data analytics workflow applied to my "Album Financiero" dataset (a real-world dataset), demonstrating technical proficiency, attention to data quality, and the ability to generate actionable insights through both detailed reports and executive dashboards.

Project Scope and Limitations

This project focuses solely on data from a personal financial records over the past 10 years. It will not incorporate external market data or make predictive forecasts. The scope includes data cleaning, normalization, data modeling, and the creation of interactive reports and dashboards for descriptive and diagnostic analysis. To maintain privacy, all reported metrics will be anonymized and presented as percentages or ratios rather than absolute values.

As this is the analyst's first personal capstone project, several limitations should be noted:

- **Early-stage professional experience:** The analyst is still developing familiarity with standard professional workflows commonly used in collaborative data projects, such as version control, formal documentation standards, and peer review practices.
- **Limited applied use of programming languages:** Although the analyst has foundational knowledge of SQL, Python, and R, these tools will not be integrated into the current project. The focus remains on building expertise in Microsoft Excel, Power BI, and DAX measures before expanding to broader automation and scripting capabilities.
- **Manual ETL processes:** While proficient in Power Query, the analyst will conduct much of the data transformation manually in Excel and Power BI. This decision prioritized transparency and learning, but it limited the project's automation potential and efficiency.

These limitations reflect my current stage in the learning journey and serve as motivation for further development in future projects. While it is true that this project is for personal matters in nature, the processes, techniques, and methodologies applied are fully transferable to business analytics, financial reporting, and operational dashboards in professional settings.

Project Timeline Consideration

As this project forms part of professional development in data analysis and visualization for the analyst, it is decided to work without a fixed deadline. This approach allowed to dedicate the necessary time to explore advanced tools, deepen

the understanding of Power BI, and apply best practices in data modeling, transformation, and dashboard design.

This deliberate approach ensured that the final deliverable is not only a demonstration of technical skills but also reflects a thoughtful, process-driven methodology consistent with professional data analysis standards and obtaining the solutions I expect to obtain, not only as a analyst, but also as the final stakeholder of the project.

Data Preparation and Cleaning

Focuses on ensuring data quality, privacy, and readiness for analysis. Key tasks included:

- Converting sensitive monetary values into anonymized relative metrics.
- Standardizing and addition of category names, labels, and descriptions.
- Identifying and correcting inconsistent, incomplete, duplicated, or unnecessary data.
- Review and decision-making regarding missing or excessive information in the document, to achieve proper and efficient normalization of the final solution.

In the last decade, I learned about the importance of correctly using information that really informs the final user in the easiest way possible. Since the creation of my budget document, I have deleted information that did not apport any helpful insight into my objectives and only added extra work calculating the metric and/or adding the data.

Analysis and Tools to be Use

- Power BI (including Power Query, DAX, and dashboard design)
- Power BI Service (for publishing and mobile layout adaptation)
- Excel (Data Analysis Tool pack, PivotTables, formulas/functions, charts)
- OpenAI (ChatGPT for support with DAX, translations, and documentation)

Making this re-analysis of the "Album Financiero" dataset, to build a complete solution that shows more the insights that can be obtained in a more comprehensive way, is a good challenge for me, allowing me to put into practice all the knowledge I have acquired over the years.

Data Analysis Process

This project demonstrates core business intelligence skills, including data cleaning, dimensional modeling, reports and dashboard development, by transforming a decade-long personal finance tracking system into a structured, scalable reporting solution using Excel and Power BI.

Initially focused on optimizing a single transactional tab (**"C_diaria"**), the project has since expanded to improve the entire structure of the "Album Financiero" spreadsheet, this to obtain the maximum improvement require for the final solution. From five tabs, now a total of eight tabs serve as the main tables of the data model, supporting a more organized, efficient, and analysis-ready foundation for reporting and decision-making.

This project was developed following the Google Data Analytics framework, as this methodology was considered to clearly and effectively reflect how responsibilities are distributed in the work of a data analyst. This choice helped guide each phase of the process with a consistent approach aligned with widely recognized standards in the field:

- Ask: Define the problem and identify key questions to answer.
- Prepare: Collect and organize the data for analysis.
- Process: Clean and transform the data to ensure quality and consistency.
- Analyze: Explore the data, identify patterns, and develop insights.
- Share: Communicate findings clearly through visualizations and storytelling.
- Act: Apply insights to make informed decisions or recommendations.

Ask

Stakeholder: Juan Carlos Larios Coronado (also the analyst)

Objective: Design a robust BI reporting system by:

1. Structuring and normalizing the "Album Financiero" spreadsheet.
2. Integrating the cleaning data into a relational Power BI data model.
3. Developing detailed reports for each core table, along with an interactive dashboard that enables users to:

- Identify dominant expense categories over time.
- Monitor savings ratios and spending behavior trends.
- Reference for expected budgets vs actual performance.

Key Financial questions to be resolved:

1. How much money do I have in my payment methods and in my credits?
2. How much do I have budgeted for the month? How much remains to be spent?
3. What day is my final credit card payment due and how much should I pay? Have I arranged the money to pay the debt on time?
4. Where and during which trip did I spend the money? How much did I spend per clave?
5. How much money do I currently have in my savings accounts? How have my investments changed over time? How many months can I sustain my current lifestyle based on my average monthly budget and my savings and available investments?

Prepare

The data source consists of more than ten years of personal finance records manually recorded in Excel and is distributed between five core sheets, a definition sheet, and two-dimensional sheets:

1. Each row in **"c_diaria"** represents a transaction, serving as the project's fact table.
2. The **"presupuesto"** tab contains monthly and annual budgets, which were initially static but are now being restructured for normalization and integration into Power BI.
3. The **"credito"** tab documents both settled and outstanding credit debts.
4. The **"viajes"** tab provides a historical log of trips taken over the years, including travel dates and brief descriptions or reasons for each trip.
5. The **"inversiones"** tab details the investments that have existed in the life of the "Album Financiero" spreadsheet.
6. The **"definiciones"** tab consolidates definitions and recommendations for the columns "razon_uso," "clave," and "*metodo_pago*." It will be completed with updated information from the "Guía de Llenado" spreadsheet.
7. The **"parametros"** tab was created to inform the total available amount for each managed credit account.

8. The **“geografia”** tab was created to store complete geographical information for each place visited and recorder in the **“viajes”** tab.

Process

The process phase was divided by spreadsheet, as each required a detailed and focused transformation to create professionally normalized datasets. This involved renaming columns, creating supporting fields, restructuring entire spreadsheets, applying new calculations, building pivot tables, and leveraging Power Query to remove duplicates and blank spaces, standardize inconsistent descriptions, and perform other necessary adjustments.

Process – **“C_diaria”**

1. Setup & Initial Cleaning

- Working copies of “Album Financiero” and “Guia de Llenado” spreadsheets were created.
- These copies were iteratively updated to reflect all transformation logic and definitions.
- No duplicate rows or structural inconsistencies were found (aside from payment method columns, which were under review).
- In the *“razon_uso”* column, two outdated categories (PRESOP and PRESEJE) were merged into PRESOM.

2. Category Consolidation

- The *“razon_uso”* categories were reduced from twenty to fourteen by merging similar categories for consistency:
    - MENSUALIDAD merged into SUSCRIPCION.
    - EJERCICIO merged into NECESIDAD.
    - LECTURA merged into EXTRAS, then renamed LIBRE due to psychological framing.
    - TRABAJO was renamed to INGRESO.
    - PRESCAR, PRESDOC, PRESEDU merged into P.FORMAL.
    - PRESREG merged with PRESREC that is renamed to P.RECREACION.
    - PRESMANUAL renamed to P.MANUAL.

- PRESVIAJE was renamed to P.VIAJE.
- PRESOM renamed to P.GUSTOS.
- PRESSAL renamed to P.SALUD.

3. Column Normalization

- An *"id"* column was added to assign a unique identifier to each transaction.
- The *"id_viaje"* column was created and linked via an array to the *"id_viaje"* column in the **"viajes"** tab. As a result, *"ubicacion_viaje"* and *"numero_viaje"* columns became redundant and were removed to reduce table size.
- The *"clave"* column underwent label standardization:
  - Apoyo Casa and Hogar Propio merged into hogar.
  - Carro Propio merged into transporte.
  - Entrada Dinero and Salida Dinero merged into deudas.
  - New labels created: regalo, evento and transaccion.
  - The Viajes label was removed, and its entries were reassigned to more appropriate categories.

4. Payment Method Normalization

- Transactions involving two payment methods (typically transfers between personal accounts) were duplicated, one for each method.
- Summary of how many rows were found per "metodo_pago" usage:
  - 13,508 transactions had one payment method.
  - 1,129 transactions had two methods.
  - 6 transactions had no payment method recorded and were removed.
  - Final total: 15,766 transactions.

5. Master Tab Creation

Previously, eight separate columns captured payment methods. These were consolidated with the Master Tab, into two normalized columns:

- *"metodo_pago"*
- *"monto"*

The steps followed were the next ones:

- A helper column using COUNTA() was created to identify and manage multi-method entries.
- The new spreadsheet **"**Master Table" was created. Separate tabs were created for each payment method, containing *"id"*, *"monto"*, and "*metodo_pago*" columns from the original **"C_diaria"** tab. Each tab included only the corresponding method's transactions.
- Rows with zero amounts were removed from each tab.
- These cleaned tables were consolidated into a single Master tab, filtered by column *"id"*.
- The remaining columns from the **"C_diaria"** tab were added to the "Master" tab using VLOOKUP().
- Total values between **"C_diaria"** and the Master tab were compared and confirmed to match.
- The *"id"* column in the Master tab was sorted to start from 1, numbering sequentially.
- The Master Tab was copy/pasted into "Album Financiero" spreadsheet. The Master table was renamed to **"c_diaria"** and the original **"C_diaria"** was deleted.

6. Dropdown option

- Dropdown menus were created for the *razon_uso, clave*, and *metodo_pago* columns to prevent typographical errors and ensure only valid categories are used in the dataset.

Process – **"presupuesto"**

1. Initial Setup & Table Structure Design

- A new working tab titled **"presupuesto"** was created, and the original version was renamed to **"old_Presupuesto."**
- A normalized structure was implemented to store budget data in a tabular format.
- Two date-related, helping columns, *"año"* and *"mes,"* were added to capture the year and month name for each budget entry.

- Two additional columns, *"inicio_mes"* and *"final_mes,"* were created to inform the first and last day of the corresponding period assigned.

2. Data Transcription and Standardization

- All budget data from **"old_Presupuesto"** was transcribed into the new structure.
- Budget categories and concepts were recorded in the *"razon_uso"* and "*descripcion"* columns.
- Budgeted amounts were stored in the column *"monto_presupuesto."*

3. Automation of Actual Spending Calculation

- A helping column, *"monto_gastado,"* was created to automatically calculate actual spending for each defined period (from *"inicio_mes"* to *"final_mes"*).
- This calculation was performed using the SUMIFS() function, replacing the previous manual input method and ensuring consistency with the transaction data in the **"C_diaria"** tab.
- The **"old_Presupuesto"** was deleted.

Process – **"deudas," "viajes,"** & **"geografia"**

1. Initial Cleaning

- It was decided to change the name of the **"deudas"** tab into **"creditos"** as this name focused more on the data managed between the tab.

2. Column preparation – "creditos" tab

- The helping column *"fecha_inicio"* and the columns *"fecha_corte"* and *"fecha_limite_pago"* retained their original structure and logic.
- The helping column *"retiro_programado"* was automated using the following logic:
    - If *"listo_para_pago"* = "SI" and *"fecha_corte"* > TODAY() = Programado a la fecha de pago
    - If *"listo_para_pago"* ≠ "SI" and *"fecha_corte"* > TODAY() = Pendiente de programar
    - If *"listo_para_pago"* ≠ "SI" and *"fecha_corte"* ≤ TODAY() = No se puede programar

- This logic was based on the columns *"fecha_corte"* and the newly created *"listo_para_pago"* and using IF() and TODAY() functions.
- For the helping columns *"monto_pagado"* and *"monto_pendiente_pagar"*, were updated to reference the normalized **"c_diaria"** tab for accurate calculations.

3. Column Preparation – "viajes" & "geografia" Tabs

- The **"viajes"** tab tracks spending associated with each trip, using the *"ubicacion_viajes"* and the helping *"numero_viaje"* columns, which correspond directly to fields in the **"c_diaria"** tab.
- For the cancelled trips, the *"numero_viaje"* column was updated from "x" to "-1."
- For the trips involving a change of residence, the *"numero_viaje"* column was updated from "MUDANZA" to "0.0."
- As Power BI aggregate expenses dynamically by trip and payment method, the eight columns representing individual payment methods were removed to avoid redundancy and improve maintainability.
- A new column *"id_viaje"* was created to assign a unique identifier to each trip, starting from V001, V002, V003, and so on.
- A *"grupo_viaje"* column was added to identify grouped trip, with IDs starting from G001, G002, G003, and so on.
- The columns *"ubicacion_viajes"*, *"continente", "pais", "estado_region", "latitud" and "longitud"* created in the tab **"geografia"**, which stores detailed location data for all trips.

Process – **"inversiones"**

1. Simplification, Tab Setup & Table Structure Design (Normalized Monthly Snapshot)

- A new tab titled **"inversiones"** was created.
- The objective was to shift all logic and calculations to Power BI, leaving the Excel sheet focused solely on clean, structured data entry.
- A normalized table format was implemented to represent monthly snapshots of each investment. Each row corresponds to a single investment for a specific month, allowing for the monitoring of changes over time.
- This structure enables Power BI to calculate monthly growth, trends over time, and display current totals using dynamic measures.

2. Historical Data Adaptation and Integration

- Over the past four years, periodic records maintained to track changes in total investment amounts. As the objective of the normalized tab is to know how each investment changed over time, the next procedure was made to obtain an estimated value for each one per month:
    - Previous versions of the "Album Financiero" spreadsheet were analyzed to obtain which investments have been used over the past four years and capture in helping columns with all the available values
    - As each version of the spreadsheet was saved once per year, many periods of time were not possible to obtain the actual value corresponding to each investment, so the decision was made to calculate an estimated growth, without considering the actual changes happened financially in market prices and international stock changes over time.
- They are grouped by two groups: investments readily available for use (Disponible), and those restricted due to maturity terms or withdrawal restrictions (No Disponible).
- This historical data has been reviewed, adapted, and integrated into the normalized **"inversiones"** tab, enabling long-term trend analysis in Power BI while preserving the distinction between accessible and non-accessible funds.
- A key adaptation involved manually estimating missing months by assigning rounded values based on amounts from surrounding months, ensuring continuity in the historical timeline
- The original **"inversiones"** table was deleted.

Analysis

As outlined in the framework, the objective is to develop a financial tracking system powered by a responsive Power BI dashboard and reports. For that, the analysis procedure was focused on answering the five key statements defined in the Ask phase.

- A calendar table named **"Calendario"** was created and related to the *"fecha"* column in the **"c_diaria"** table to enable time-based analysis.
- A dimension table **"dim_razon_uso"** was created to link the *"razon_uso"* column between the **"c_diaria"** and **"presupuesto"** tables.
- A dimension table **"dim_tipo_inversion"** was created so was possible to relate the **"c_diaria"** with **"inversiones"** tables.

- Filtro por Fecha: Slicer created in all the pages, to filter the visuals by year, quarter, month and/or day, in exception of "Reporte Creditos".
- Filtro por Periodo: Slicer created in the page "Reporte Creditos", considering the column *"fecha_corte",* to filter the visuals by year, month and/or day.

How much money do I have in my payment methods and in my credits?

- A Dashboard page called **"Resumen Financiero"** was created, where the visual "Balance Metodo Pago" was developed. This is explained in more detail in the Share Process.

How much do I have budgeted in the month? How much remains to be spent?

- A report page called **"Reporte Presupuesto"** was developed.

The following visuals were developed, and the DAX measures used in each of them (if applicable) will be explained:

- Montos Totales: Display the budgeted amount and the spending amount by the selected period.
    - Monto Presupuestado: Calculate the sum of the total budgeted amount filtered by the selected period.
    - Monto Gastado Dinamico: Calculates the total spending, excluding the "DINERO" and "INGRESO", filtered by the selected period. This replace the helping *"monto_gastado"* column from the **"presupuesto"** tab.
- Restante para Gastar: Display how much money is available to be spent by the selected period.
    - Restante para Gastar: Computes the difference between the budgeted and actual spending.
- Ingresos: Display the total amount received for any concept categorized as "INGRESO" in *"razon_uso".*
- Comparativa por Presupuesto: Displays the comparative between the budgeted amount vs. the total spending made by the selected period. "INGRESO" and "DINERO" were excluded.
    - Monto Presupuestado(Visual): Informs the absolute values of the Monto Presupuesto measure.
    - Monto Gastado (Visual): Informs the absolute values of the Monto Gastado Dinamico measure.
- Descripcion por Presupuesto Especifico: Displays the description assigned to each specific budget category that was assigned in the selected period.

    - Mostrar Descripcion: Inner filter that helps to retrieve the appropriate description of each *"razon_uso"* informed in the **"presupuesto"** tab, considering active filters.
- Detalle de Transacciones: Displays detailed transactions based on selections made in the summary visual.

What day is my final credit card payment due and how much should I pay? Have I arranged the money to pay the credit debt on time?

- A report page **"Reporte Creditos"** was created.

The following visuals were developed, and the DAX measures used in each of them (if applicable) will be explained:

- Gasto en Credito por Periodo de Facturacion: Displays the total amount of the billing period, between a specific period. A dotted line represents auto imposed limit amount informed in *"limite_autoimpuesto"*
    - Monto Asociado: Calculates the total amount of the billing period, between a specific period. This measure consolidates data previously informed separately in helping columns *"monto_pagado"* and *"monto_pendiente_pagar"* in **"deudas"** tab.
- Fecha Limite de Pago: Displays the date of payment of the selected billing period. Surpassing this date, the credit generates interest.
- Estatus Pago del Credito: Display a symbol if the specific billing period has been already programmed to be paid or not.
    - *simbolo_listo_para_pago*: A calculated column in the **"creditos"** table was created to show a symbol if the payment is programmed to be paid or not.
- Monto Disponible vs. Limite Personal: Displays how much money we have available before reaching the self-imposed limit amount informed in *"limite_autoimpuesto"*
    - Disponible antes del Limite: Informs how much money we have available before reaching the self-imposed limit amount informed in *"limite_autoimpuesto"*
- Fecha de Inicio: Displays the initial date of the selected billing period.
    - *fecha_inicio*: A calculated column in the **"creditos"** table was created to replace the helping *"fecha_inicio"* column from the **"creditos"** tab.
- Fecha de Corte: Displays the ending date of the selected billing period.
- Detalle de Transacciones: Displays detailed credit-related transactions filtered by selections made in the summary visual.
    - Seleccion Fecha Inicio: Helps the inner filter to return the earliest (MIN) date within the selected range.

- Seleccion Fecha Corte: Helps the inner filter to return the latest (MAX) date within the selected range.
- Mostrar Transaccion: Inner filter that informs all transactions related to credit payments that occurred between the selected MIN and MAX dates.

Where and during which trip did I spend the money? How much did I spend per clave?

- A report page called **"Reporte Viajes"** was created.
- A separate report page "Tooltip Viajes" was created to display Tooltip information.

The following visuals were developed, and the DAX measures used in each of them (if applicable) will be explained:

- Filtro por Ubicacion: A hierarchy slicer was created to filter by continent, country, location (city or town) and the "*id_viaje"*.
  - *id_viaje_descripcion*: A calculated column created in the **"viajes"** table, that combines *"id_viaje"* and *"descripcion"* in a single descriptive field for better understanding.
  - Filtro Ubicacion de Viaje: Inner filter that filter out any *"ubicacion_viaje"* that wasn't used in the corresponding period.
- Viajes Agrupados: A hierarchical matrix visual to display and filter grouped trips (with more than one related trip under the same *"grupo_viaje"* value) informing all the travels that were made because of the principal location visited.
  - *grupo_viaje_descripcion:* A calculated column that combines *"grupo_viaje"* and "*viaje_principal*" into a single description.
  - *id_viaje_descripcion*.
  - *grupo_conteo*: A calculated column in the **"viajes"** table that counts the number of trips between each *"grupo_viaje*". It is an inner filter, to only show grouped trips with more than one trip are displayed (filter applied).
- Mapa Gastos de Viaje: A map showing all locations where transactions were made during the selected trip.
- Gasto Total por Clave: Displays total spendings per label informed in the *"clave"* column, for the selected period and trip.
  - Total Gastado Clave: A measure that sums all the expenses for the selected trip and label.
- Detalles Viajes: Display detailed transactions filtered by selection in the summary visual.

- Total Gastado Clave.

The report page **"Tooltip Viajes"** provides the tooltip visuals for the Mapa Gastos de Viaje visual.

- Duracion Viaje: Displays the trip duration for the selected *"id_viaje".*
  - Duracion del Viaje: Calculates the duration between *"desde*" and *"hasta"* dates from the **"viajes"** table.
- Lugar Visitado: Displays the name of the selected place.
  - ubicacion_completa: Informs the place, state or region and the country of the selected place.
- Numero de Viajes: Displays the total number of trips recorded for the selected place.

How much money do I currently have in my savings accounts? How have my investments changed over time? How many months can I sustain my current lifestyle based on my average monthly budget and my available savings?

- The report page **"Reporte Inversiones"** was created.
- Two separate report pages **"Tooltip Inversiones"** and **"Tooltip Inversiones - Balance"** were created to display Tooltip information.

The following visuals were developed, and the DAX measures used in each of them (if applicable) will be explained:

- Tendencia Acumulada de Inversiones: Shows the accumulated value of investments over the selected period.
- Variacion Mensual de Inversiones: Displays month-over-month changes in investment values for the selected period.
  - Monto Mensual: Informs the total value for each *"tipo_acceso"* category.
  - Monto Mes Anterior: Informs the total value for each of *"tipo_acceso"* from the previous month.
  - Diferencia Mensual: Calculates the difference between Monto Mensual and Monto Mes Anterior.
- Diferencia Periodo Seleccionado: Display the changes in amount between the selected period by *"tipo_acceso".*
  - Diferencia Periodo Seleccionado: Informs the changes in amount between the selected period by *"tipo_acceso".*
- Tipo de Inversion: Used to filter the Generado Inversiones Visual.
- Balance Financiero Mensual e Inversiones: Displays the financial result and the current investment balance, corresponding to the selected period.

- Saldo Actual Inversiones: Informs the current available balance for the selected investment.
- Resultado Financiero a la Fecha: Informs of the total profits or losses for the selected investment, considering the accumulative money invested through the years and the available amount in the investment.
- *primer_dia_mes*: A calculated column was created in **"inversiones"** tab to inform the *"fecha_monto"* date with the first day of the reference month.
- *tipo_inversion_detectado*: A calculated column was created in **"c_diaria"** tab to inform the transfers that correspond to an investment.

The report page **"Tooltip inversiones"** provides the tooltip information for both Tendencia Acumulada de Inversiones and Variacion Mensual de inversiones.

- Monto por Inversion: Displays the amount for the selected month and the difference from the most recent previous total available.
  - Monto Reciente por Acceso: Informs the corresponding available amount per *"tipo_acceso."*
  - Diferencia Mensual.
- Disponible: Displays the total available investments for the selected month.
- No Disponible: Displays the total available investments for the selected month.

The report page **"Tooltip inversiones - Balance"** provides the tooltip information for the visual Balance Financiero Mensual e Inversiones.

- Total Invertido: Displays the total amount invested in the selected investment.
  - Total Inversiones: Informs the total amount invested in the selected investment.
- Saldo Acumulado: Displays the current available balance for the selected investment.
  - Saldo Actual Inversiones.
- Resultado Financiero: Displays the total profits or losses for the selected investment, considering the accumulative money invested through the years and the available amount in the investment.
  - Resultado Financiero a la Fecha.
  - Generado % Inversiones: Informs the percentage of profits or losses for the selected investment

Additionally, The Meses y Días Sostenibilidad visual was created in the Dashboard page **"Resumen Financiero"**. This visual is described in more detail in the Share Process.

*Originally, this project was intended to meet my main requirements, which did not include such a complex investments report. However, after analyzing and drafting the initial version of the "**Reporte Inversiones**", I decided to invest time and effort into this section. Here, I put into practice my knowledge of normalization, modeling, and design to deliver professional-quality work focused on financial topics. I consider the result a major achievement, as I was able to find the logic to fill in incomplete information and create insightful metrics that I didn't think would be possible to build.*

Share

After answering the key questions created to give form to the Power BI reports and tooltips, the dashboard page was ready to be constructed as a summary of what was wanted to see in the Power BI Solution.

Share – "Resumen Financiero" Dashboard

- The dashboard page **"Resumen Financiero"** was created.
- Separate report pages **"Tooltip Resumen - Presupuesto"** and **"Tooltip Resumen - Viajes"** were created to display Tooltip information.

The next visuals were created with their respective DAX measures:

- Balance Metodo de Pago: Displays the current amount available in each payment method.
    - Saldo Actual: Shows the current amount for each payment method that has a balance different from zero.
- Estatus Linea de Credito: Displays the total credit limit and the available amount after substracting pending payments charged to the credit.
    - Saldo Credito: Informs the amount spent on each credit payment method.
    - Credito Disponible: Shows the remaining available credit.
    - Limite Credito: Informs the total credit limit per the **"parametros"** tab.
- Horizonte de Sustenabilidad y Fecha Limite: Displays how long current investments and available debit can cover living expenses for the future.
    - Total Disponible Inversiones: Total amount of investments currently available for spending.
    - Saldo Credito.
    - Total Presupuesto Especifico: Total amount of one-time planned budget expenses.
    - Total Presupuesto Regular: Total amount of recurring (monthly) planned budget expenses.

    - Meses de Sostenibilidad: Informs the estimated number of months (including fractions) that current funds can cover all expenses.
    - Meses y dias de Sostenibilidad (con Fecha): Informs the same value, split into months and days.
- Resumen Estado de Pago del Credito: Summarizes the payment status, total amount and due dates for the current payment method and the accumulated payment status corresponding to the next due date (if available).
    - Resumen Estado Pago General: Displays payment status, amount, and due date for the current payment period.
    - Resumen Estado Pago Proximo Mes: Shows the same information for the next payment period.
    - Resumen Pago Combinado: Combines both measures for proper display in the visual.
- Razon de Uso: Provides a breakdown by *"razon_uso"* allowing sorting and filtering alongside the date slicer.
    - Filtro Razon de Uso: Inner filter that filter out any *"razon_uso"* that wasn't used in the corresponding period.
- Porcentaje Utilizado del Presupuesto: Displays the percentage of the budget that has been used in the selected period.
    - % Presupuesto: Calculates the percentage of the total budget utilized.
- Gasto Asociados a Viajes: Displays total spending during travels in the selected period, grouped by location.
    - *ubicacion_viaje_ajustado*: A calculated column in **c_diaria** that labels transactions not associated with any trip as "Sin Viaje."
    - Gasto Real Viajes: Informs of the total amount spent on travel.
    - Gasto Viajes (Treemap): Informs the expenses made per travel, and when there is no travel in the month, show the legend "Sin Viaje".
- Tendencia Acumulada por Inversion: Displays the current value of each investment in the selected period, grouped by *"tipo_inversion"*.
    - Monto Inversiones Mensual: Informs the amount per month by each available *"tipo_inversion"* in the selected month. If a longer period is selected, the visual shows zero values.
- Tendencia Acumulada por Acceso: Displays the current value of each investment in the selected period, grouped by availability status.
    - % Monto por inversiones: Informs the percentage represented by each *"tipo_acceso"* category.

The report page **"Tooltip Resumen - Presupuesto"** provides the tooltip information for the visual Porcentaje Utilizado del Presupuesto.

- Total Presupuestado: Displays the current value of the total budgeted amount and the amount spent in the selected period.
    - Monto Gastado Dinamico.
    - Monto Presupuestado.

The report page **"Tooltip Resumen - Viajes"** provides the tooltip information for the visual Gasto Asociados a Viajes.

- Monto Utilizado: Displays the total amount spent per selected travel and period and informs the place where was spent.
    - Gasto Real Viajes.

Share – "Version Mobil"

- The dashboard page **"Resumen Financiero"** was duplicated. This new page was renamed **"Version Mobil".**
- The Slicer "Filtro por Fecha" was divided in two separate slicers. One informing the year following the style between, and the second one informing the months following the style Tile.
- It is distributed in two sections divided by a blue line. The first section cannot be filtered, so the slicers were inserted in the second section with all the visuals that are affected by this.

Share – album_financiero

- The design of the Excel spreadsheets has changed from tables to range, as there is no need to have them limited with that format.
- A new spreadsheet *"definiciones_en"* was created with the concepts and definitions translated in English.
- The helping columns of the next spreadsheets were deleted, as they are no necessary for the correct function of the Power BI solution:
    - Spreadsheet **"presupuesto"**: "*año"*, "*mes*" and "*monto_gastado*".
    - Spreadsheet **"creditos"**: "*fecha_inicio*", "*retiro_programado*", "*monto_pagado*" and "*monto_pendiente_pagar*".
    - Spreadsheet **"viajes"**: "*numero_viaje"*.
- Buttons with bookmarks were added in the top right hand of each page in Power BI to function as:

    - For **“Resumen Financiero”, “Reporte Presupuesto”** and **“Reporte Viaje”** a refresh button is available to filter back the visuals to the current month. For the moment, this needs to be manually updated.
    - For **“Reporte Creditos”** and **“Reporte Inversiones”** a refresh button is available to filter back the visuals to the current year. For the moment, this needs to be manually updated.
    - A Return button is available on all the pages to return to the Dashboard Page.
- Both the Excel and Power BI files were moved to their final file location. Both were renamed into “album_financiero”.
- Both documents follow the same format (Accessible City Park) to maintain uniformity. Also, the titles, values, totals charts and slicers follow the same logic in format and style.
- The Power BI solution was already published in Power BI Service and is available to be seen via online and in a smartphone.
- The option to automate the refresh and Publish option of Power BI cannot be done without a Business License. As we are working with a personal license and it was decided not to proceed with the change of license, this procedure will be done manually.

Share – album_financiero_publico

- Copies of the original Power BI (pbix) and Excel (xls) files were created and renamed to album_financiero_publico for publication purposes.
- All monetary values were modified to protect personal information. Generally speaking, the following logic was applied:
    - The sign of each amount (positive or negative) was preserved, to maintain consistency with the original transaction type.
    - All values were systematically scaled and rounded, ensuring the analytical logic and proportions remain intact while obfuscating actual figures.
    - Minor manual adjustments were made to remove small non-representative amounts introduced during transformation that affected the Power BI metrics. With those minor changes, the overall logic or results weren’t affected by the changes.
- Text columns were also adjusted in each table :
    - In pages **“viajes”** and **“geografia”**: Travel locations and descriptions were randomly changed using formulas and Google Maps.
    - In pages **“definiciones”**, **“C_diaria”** and **“parametros”**: The “*metodo_pago*” categories were renamed as “entidad” showing whether the method was debit, credit or cash.

- In page **“C_diaria”** : The “*lugar_de_uso*” column was renamed using a pivot table and the function VLOOKUP() function, assigning generic names as lugar_1, lugar_2, and so on.
- In pages **“inversiones”** and **“c_diaria”** : The *“tipo_inversion*” column was renamed to generic labels like inversion_1, inversion_2, etc. These changes also affected the associated descriptions.
- In page **“c_diaria”** : The “*descripcion*” column, a helper column was created to flag transactions that contained key word affecting Power BI measures. Then, using a pivot table and VLOOKUP() function, the transactions with no flags were renamed to standardized codes like alimento_1, deudas_1, bebida_1 , as they have in the *“clave”* column.
- In page **“presupuesto”**: The “*descripcion*” column was renamed based on its *razon_uso*, using a pivot table and VLOOKUP() function, to assign names like pviaje_1, necesidad_1, etc.

This entire anonymization and transformation process was done using a mix of formulas, pivot tables, and manual review. While it could have been automated with tools like Power Query (Excel/Power BI), SQL, Python, or R, I intentionally carried out the process manually as part of my learning path. The goal is to eventually automate ETL (Extract-Transform-Load) procedures using the tools I'm progressively learning.

Act - Recommendations and Next Steps

During the development of this project, several potential improvements and enhancements were identified that could further optimize the Power BI solution. However, after careful analysis and reflection, it was determined that these ideas fall outside the original scope and objectives defined for this capstone. As such, they have been reserved for consideration in a possible second phase of this project.

To continue building on the insights generated and ensure sustained value from this work, the following actions are recommended:

Technical Improvements

- Apply SQL, Python, and R to replicate or extend the metrics using alternative tools or pipelines.
- Create *“id_presupuesto”* column, to improve data relationships and tracking of pending items. The goal is to improve the visual “Descripcion por Presupuesto Especifico” and be able to:
  - Watch the exact amount spent per budgeted item.

    - Improve the budget-to-actual alignment by ensuring transactions are matched to the originally budgeted period rather than the month they were recorded. This would avoid inflating or misrepresenting budget categories due to late reimbursements or delayed payments.
      Example, if a reimbursement of +100,000 related to a May travel expense is received in June, it should be reflected against the May budget to preserve analytical consistency.
- Create an "*id_pendientes"* in the "**c_diaria"** tab, so each pending payments can be identified more exactly and can be applied in a key visual in Dashboard **Resumen Financiero** to inform which purchases are pending to be completed.
- Automate "*desde"* and *"hasta"* columns in the **"viaje"** table, to reduce manual errors in Excel and standardize data input in Power BI.

Reporting Enhancements

- Add a **"Reporte Ingresos"** report page in Power BI to analyze incomes separately from expenses.
- Enhance tooltips in the **"Reporte Viajes"** page by including a photo took from each travel (when available), enriching the personal experience.

Data Quality & Clarity

- Standardize and clarify descriptions in the *"descripcion"* columns in the album_financiero Excel document, to improve readability and insights and make them easier to understand.
- Analyzed if is possible to add all the past periods of time to the **"Reporte Deudas"** tab, this to have available the most complete information about the credit debts.

Knowledge Transfer & Documentation

- Develop a video walkthrough of the project, showcasing the evolution of the dataset and dashboard: before, during, and after the capstone process. This will help communicate project goals, methodology, and outcomes to external reviewers or future collaborators.

Version Control & Future Replication

- Now, three Excel files and one Power BI version prior to the final personal documents were maintained and saved. These will remain saved for internal purposes.
- For the public version, only the final versions will be uploaded, as these documents were made into their public version, until all the changes were made in both excel and power BI documents.
- Will be established a future versioning strategy (e.g., date-based naming, GitHub for future scripts in Python/R) to facilitate transparency, reproducibility, and scalability for similar financial analytics projects.

Conclusions & Reflections

Over the past 3 months that it took to complete this project (since the analysis Framework until the final edition of this document), I can conclude that the work of a data analyst is both broad and rewarding. There are many opportunities to explore new skills and improve the quality of the insights delivered.

I often revisited my personal notes, rewatched video lessons, and studied various perspectives on storytelling and data visualization. This reinforced my understanding and improved the decisions made throughout the project.

On the Use of Generative AI

Throughout this project, generative AI played a key role in supporting my learning process and problem-solving efforts — particularly in developing complex DAX measures and refining data modeling logic. Acting as a collaborative assistant, AI tools allowed me to overcome technical challenges faster, deepen my understanding of Power BI capabilities, and focus more effectively on building a coherent and impactful analysis. While all decisions and final implementations were made by me, this assistance significantly enhanced the quality and scope of the project.

Reflections on tool selections

Although tools such as SQL, Python, and R are widely used in the data analyst field and were initially considered for this project, once I began to understand the complexity involved in normalizing a 10-year dataset and adapting it into a scalable model for dynamic analysis in Power BI, I decided not to integrate those tools into this phase of the project. While I was eager to apply the knowledge I have from educational projects and exercises with these tools, I have not yet used them extensively in real-world scenarios and I didn't want to divert my focus of the objective of this project.

Instead, we determined that the best approach was to split the project into two parts:

- The first part focused on achieving the core goal of normalizing the Excel data and delivering a robust Power BI solution using the skills I currently have.
- The second phase will extend the project by incorporating the use of SQL, Python, R, and implementing the recommendations identified in the Act phase.

This two-part structure allows for both fulfilling the original objectives effectively and providing room for future growth and enhancement using more advanced tools.

This approach reflects my current stage in the data analysis journey: committed to deepening my expertise and progressively integrating new tools as foundational skills are solidified.

Development Steps & learnings

As a final personal touch, I divided this project into four steps as in each one I had to rethink again what I was working on and what I really wanted to see in Excel and the Power BI Solution.

1. Goal Definition and Planning: identifying key questions and requirements. My main objective became clear: to professionalize my Excel-based tracking system. This meant transforming it into a more intuitive, visually organized structure, while transferring all calculations, formulas, macros, and pivot tables from Excel to Power BI.
2. The extract, transform, and load (ETL): I cleaned and restructured the dataset to make it ready for analysis. This step also included formulating five main inquiries that would guide the visual design and structure of the Power BI solution.
3. Power BI Development (Analysis): This is where the ideas took form. Through DAX formulas, proper table relationships, and intentional visual design, the dynamic behavior I envisioned came to life—answering the five key questions and unlocking new insights from my historical financial data.
4. Final Edits and Public Version: In the final phase, I focused on reviewing the Power BI solution from a user perspective rather than a developer one. This allowed me to make final improvements—adjusting visuals and measures, titles, and layout—to ensure that the story being told was not only accurate but also engaging and easy to follow. This phase also included the creation of a public-facing version of the project called “album_financiero_publico”.

Having a clear and well-defined goal was essential not only for maintaining direction, but also for adapting effectively when changes to the original work plan were necessary. Throughout the project, new ideas and challenges required flexibility and rethinking of many steps. However, the established objective served as a solid reference point, helping me stay on track despite those changes. This focus proved invaluable in keeping the work structured, intentional, and ultimately successful.

Appendix and comments of the author

1. The following logic was applied to the style and formatting of the names related to the project, to create a more visual distinction of their origin:
    - **Spreadsheet:** Refers to a sheet in the Excel document.
    - *Column*: Refers to a column in the Excel document.
    - CATEGORY: Refers to a category within the razon_uso column.
    - Label: Refers to a key within the clave column.
    - **Page or Table**: Refers to a page or table in the Power BI document.
    - *Column*: Refers to a column in a table within the Power BI document.
2. To maintain consistency and enable future automation with different analysis tools, the columns, categories and labels were standardized using snake_case conventions, except for *"razon_uso"* categories.
3. All sheets, tables, columns, categories, labels, visualizations, and DAX measures names were written without accent marks. This decision was made to simplify their use in DAX measures and to ensure compatibility with future analytics tools.
4. As this is a personal project, sheets, tables, columns, categories, labels, visualizations, and DAX measures names were not translated into English. This document represents the original version, which was developed in English from the start, in collaboration with ChatGPT, as part of an exercise to practice the professional terminology used by data analysts.
5. If a deep analysis of the Dax measures is required, these can be found in the folder DAX Measures, where the next documents are available:
    - dax_mediciones_y_visuales: An Excel document informing the visuals appearing in the Power Bi Solution, their corresponding DAX measures and their formula written.
    - dax_mediciones: A CSV document informing the DAX measures used between the project and their corresponding formulas.
    - visuales: A CSV document informing the visuals appearing in Power Bi and their corresponding DAX measures.