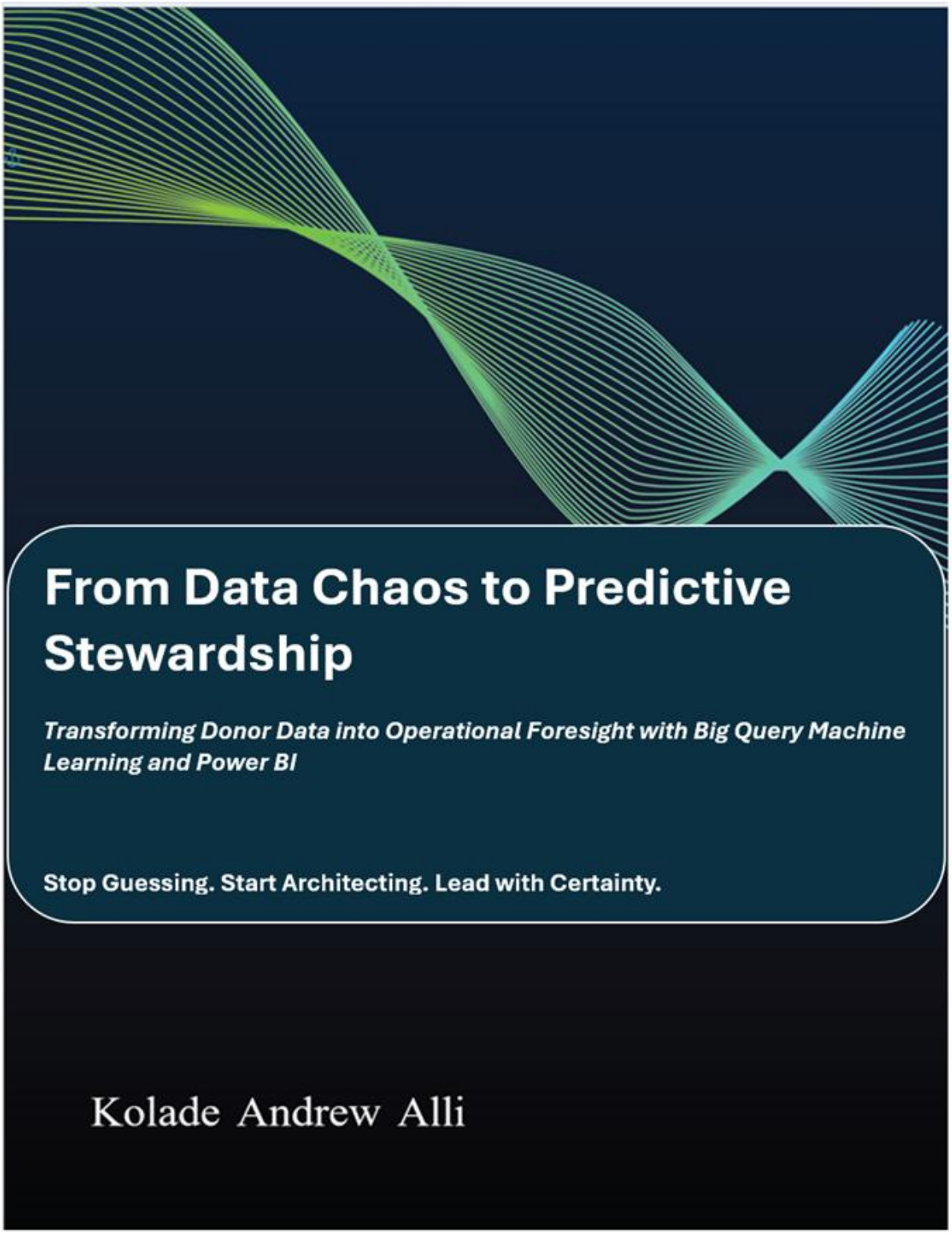

From Data Chaos to Predictive Stewardship
Transforming Donor Data into Operational Foresight with Big Query Machine Learning and Power BI
Stop Guessing. Start Architecting. Lead with Certainty.
Kolade Andrew Alli

Table of Contents

PART II: THE TECHNICAL MANUAL (The Execution)

The Transition: From Strategy to Execution

- **A Note to the COO/Executive Reader:** Leading the Change
- **A Note to the Technical Lead:** Architecting the Propensity Engine
- **[AUTHOR'S NOTE]** How ARK BI Supports Your Transformation

The Technical Roadmap

- **The Case Study:** The Transformation of Second Chance Futures
- **Learning Resources:** Workbooks, Videos, and Code References
- **A Note on Scalability:** From Small Trials to Enterprise CRM Integration

Phase 1: Environment Configuration (The Foundation)

- **Cloud Identity:** Setting up Google Cloud Projects and Project IDs
- **Workspace Activation:** Accessing the BigQuery "Brain" and Power BI "Interface"
- **Global Settings:** Disabling Autodetect for Manual Model Control

Phase 2: BigQuery SQL Abstraction (The Intelligence Layer)

- **The Single Source of Truth:** Creating the Salesforce_Data Mirror View
- **De-duplication Logic:** Using Window Functions for Clean Donor Master Lists
- **The Universal Timeline:** Manufacturing the calendar_lkup for Audit-Proof Trends

Phase 3: AI Donor Propensity (The Linear Regression Brain)

- **Feature Engineering:** Building RFM (Recency, Frequency, Monetary) Signals
- **Training the Brain:** Teaching the Model to Predict 2024 Behavior
- **The Science Audit:** Decoding R-Squared, MAE, and Explained Variance

Phase 4: Predictive Horizons (Generating 2025 Forecasts)

- **ML.PREDICT:** Generating Propensity Scores for Every Individual
- **The Confidence Gate:** Applying the 0.6 Threshold for Actionable Lists
- **Interpreting Extrapolation:** Understanding Super Supporters (Scores >1)

Phase 5: Relational Modelling & DAX (The Impact Injection)

- **The Skeleton:** Connecting BigQuery Views into a Star Schema
- **Time Intelligence:** Building YoY Growth and Momentum Measures
- **The Stewardship Metric:** Implementing Advanced Donor Retention DAX
- **Forecasting Revenue:** Calculating Likely Donation (2025) Totals

Phase 6: Visualizing Success (The Executive Dashboard)

- **The 8-Card Header:** Historical Context vs. Predictive Forecasts
- **The Performance Mix:** Area Charts for Growth and Scatter Charts for Propensity
- **The Tactical Matrix:** Creating Drill-Downs for Channel and Campaign ROI
- **Maintenance:** Real-time Refreshes and Operational Rhythm

Back Matter

- **Technical Reference:** The SQL Script & DAX Formula Library
- **Troubleshooting Appendix:** Resolving Common Cloud and Connector Errors
- **Looking Ahead:** Specialized Business Intelligence Consulting with ARK BI

Executive Summary: The Stewardship Transformation

In the modern non-profit landscape, **transparency is your currency**, and data is your most underutilised asset. Most organisations operate in a state of **Data Chaos**—paralysed by **Data Silos** where fundraising, finance, and operational impact are locked in separate, incompatible systems. This fragmentation forces teams into manual "spreadsheet wrestling" that is prone to error, lacks an audit trail, and provides only a disconnected, "rearview mirror" look at the health of the mission.

The Predictive Stewardship Blueprint provides a technical and strategic escape path. By leveraging **Google BigQuery** as a central "Silo-Killer" and **Power BI** as your executive interface, your organisation transitions from retrospective reporting to **AI-driven foresight**, achieving:

- **Funding Authority:** The ability to break down silos and visualise donor propensity alongside future revenue forecasts. You move from "hoping" for a successful year to leading with a high-integrity **Predictive Baseline** for your budget.
- **Operational Sustainability:** The power to quantify supporter loyalty into actionable metrics. By identifying the exact friction points in your **Donor Retention**, you move beyond anecdotal concerns to prove exactly where re-engagement strategies are required to protect your frontline services.
- **The Stewardship Dividend:** A robust reporting output that demonstrates a specific **Business Case for Investment**. By using AI to identify high-probability supporters, you maximise your team's ROI—converting raw data signals into **Reclaimed Staff Hours** and **Optimised Campaign Revenue**.

The Goal: To eliminate the 40+ hours a month currently wasted on manual data consolidation and redirect that energy toward your true calling: **Amplify your impact, one data-driven decision at a time.**

The Predictive Stewardship Blueprint

Transforming Donor Data into Operational Foresight with BigQuery ML and Power BI

Dedication

To the tireless staff, passionate volunteers, and visionary leaders of mission-driven organisations worldwide. You dedicate your lives to making a profound difference in the world; may this blueprint empower you to protect your team, sustain your funding, and amplify your impact—one data-driven decision at a time.

Who Am I and Why I Wrote This?

I'm Andrew, and I believe data is a catalyst for change.

My journey into the heart of non-profit analytics started with a clear purpose: to empower those who serve others. I have seen countless mission-driven organisations struggling under the weight of **Data Chaos**—fragmented silos, rising administrative burdens, and the constant, nagging feeling that resources are being stretched to the breaking point. Too often, talented teams spend their precious time wrestling with disconnected numbers instead of focusing on their vital mission of care and advocacy.

That's when I knew I'd found my calling: to bridge the gap between raw data and real-world impact. My gift lies in taking complex, messy operational data and transforming it into clear, actionable knowledge. Using the "Brain" of **Google BigQuery** and the "Interface" of **Power BI**, I create visually compelling environments that tell your organisation's true story—revealing hidden patterns that allow you to reach more people with your existing assets.

Every day spent wrestling with manual spreadsheets is a day lost in pursuing your mission. I am here to help you reclaim that time, silence the noise of data silos, and lead your organisation with **Predictive Stewardship**.

My Promise to You

By following this blueprint, you will move from **Data Chaos to Confident Decision-Making**. I promise to provide you with a technically robust, audit-proof framework that eliminates "spreadsheet wrestling" and ensures your reporting is always aligned with your mission. We will bridge the gap between fragmented silos and strategic foresight, allowing you to lead with operational certainty.

Who Should Read This Book?

- **Non-Profit CEOs & COOs:** Seeking to stabilize funding, optimize operations, and justify resource increases to the board.
- **Finance Managers & Accountants:** Who need to quantify the ROI of organizational efficiency and ensure rigorous grant and donor compliance.
- **Data Analysts & IT Leads:** Tasked with building the "Brain" (AI Models) and "Interface" (Dashboards) that serve as the organization's single source of truth.
- **Trustees & Board Members:** Who want to ensure the highest standards of financial stewardship, transparency, and impact reporting.

Author's Note: The Second Chance Futures Framework

Throughout this eBook, you will see frequent references to **"Second Chance Futures."** This is a fictionalized case study used to protect the privacy of real-world organizations while preserving the absolute integrity of the operational data and strategic outcomes discussed.

The funding crises, retention challenges, and financial transformations described herein are not theoretical; they are based on real-world implementation results using **Google BigQuery** and **Microsoft Power BI**. While the name is a pseudonym, the AI logic and the "Stewardship Dividend" are entirely authentic.

Chapter 1: The Stewardship Mandate

In the world of non-profit organisations, navigating through the complexities of donor data can often feel like traversing a maze. Many leaders find themselves grappling with **Data Chaos**—a state defined by fragmented silos of information where marketing engagement lives in one tool, financial history in another, and impact metrics are scattered across disparate, manual spreadsheets.

This fragmentation does more than just slow down operations; it hinders the ability to make informed strategic decisions. Too often, leadership is forced to drive by looking only in the rearview mirror, relying on "Post-Mortem" reporting—analysing what was raised last quarter or which donors lapsed last year. While this historical data is accurate, it is fundamentally reactive. By the time a spreadsheet reveals a funding gap or a drop in donor retention, the damage is already done.

The **Stewardship Mandate** addresses the pressing need to move from this state of reactive survival to **Predictive Stewardship**. By leveraging advanced tools like **Google BigQuery ML** (the "Brain") and **Power BI** (the "Interface"), non-profit CEOs and COOs can transform their organisations into data-driven entities capable of anticipating donor behaviour rather than simply documenting it after the fact.

1. Predictive Fundraising: The Foundation of Financial Sustainability

Predictive fundraising is the bedrock of financial sustainability. In the traditional model, fundraising metrics are siloed from operational outcomes, leading to a "guessing game" during budget season. When data is integrated, the conversation shifts from "What did we get?" to "What can we expect?"

Central to this transformation is the **'Stewardship Dividend.'** This represents the potential to monetise efficiency gains—such as reclaiming dozens of staff hours lost to manual reporting—and utilising AI-driven insights to enhance donor engagement. By understanding the **propensity of donors to give**, organisations can stabilise their funding. This allows us to move beyond simple totals and begin monitoring the "health" of the donor base through AI-driven propensity scoring. When you can predict who is likely to give and who is at risk of churning, you transform fundraising from a high-stress gamble into a stable, engineered process.

2. The Shift: Moving from "Rearview Mirror" Reporting to 2026 Foresight

The core of this shift is active foresight. This chapter explores how finance managers and accountants can harness data to create robust financial strategies that align with the evolving landscape of donor expectations. In the old world, a COO might spend 40 hours a month consolidating reports from three different departments; in the world of Predictive Stewardship, that time is reclaimed for the mission.

AI and machine learning play pivotal roles in **donor segmentation**, allowing organisations to tailor their strategies to increase loyalty. By employing these technologies, non-profits can identify patterns in donor behaviour—specifically **RFM (Recency, Frequency, Monetary)** signals. This enables leadership to craft customised communication that resonates with individual supporters. This targeted approach significantly increases the likelihood of repeat donations, contributing to long-term success. Using the **BigQuery "Cloud Brain"**, we can now forecast revenue horizons with statistical integrity.

3. Securing Major Gifts: Using Radical Transparency to Invite Investment

The modern donor is no longer moved by generic stories alone; they are looking for **Radical Transparency**. They want to know exactly how their investment is being managed and what the projected outcome is. The integration of BigQuery ML and Power BI facilitates a comprehensive analysis, enabling organisations to visualise trends and identify areas for improvement instantly.

This data-driven methodology empowers data analysts and IT leads to collaborate effectively. As non-profits embrace these technologies, they move away from assumptions and toward decisions founded on **empirical evidence**. By leveraging the **Stewardship Blueprint**, you can provide a **"Confidence Scorecard."** Presenting a board or a major donor with evidence-based projections turns a "general donation" into a "targeted investment." It replaces the "risk" of a donation with the "certainty" of an impact, inviting stakeholders into a calibrated, professional, and transparent system.

Conclusion: The Stewardship Dividend

The journey from data chaos to predictive stewardship is not merely a technological upgrade; it represents a fundamental shift in how non-profit organisations operate. It is the move from being "data-burdened" to "data-empowered." By embracing the **Stewardship Dividend** and leveraging AI-driven tools, organisations create a sustainable model of funding that is resilient to economic changes.

This chapter sets the stage for a transformative approach that anticipates the challenges of tomorrow, ensuring a brighter future for non-profits and their critical missions. By embracing this mandate, you ensure your organisation is not just surviving the present but is mathematically prepared for the future.

Chapter 2: Navigating Data Chaos

In the realm of non-profit organisations, the challenge of **Data Chaos** is omnipresent. Many leaders grapple with fragmented silos of information—where fundraising data lives in one universe, financial records in another, and operational impact metrics are buried in a third. This disorganisation is more than a technical nuisance; it is a strategic barrier that leads to significant inefficiencies, misallocation of resources, and ultimately a dilution of the mission. For a CEO or COO, identifying the true cost of this chaos is paramount, as it directly impacts operational foresight and the ability to steer the organisation with confidence.

1. Identifying the High Cost of Data Silos

A data silo occurs when information is trapped within a specific department or software tool, inaccessible to the rest of the organisation. In a typical non-profit, the "Fundraising Silo" (CRM) and the "Finance Silo" (Accounting Software) rarely speak the same language.

The financial implications of this silence extend beyond mere frustration. Non-profits often invest in multiple, overlapping systems without realising the cumulative cost of these redundancies. For finance managers and accountants, this represents not just wasted subscription fees, but also **"Insight Leakage"**—missed opportunities for maximising donor contributions because the data needed to trigger a timely ask was locked in a different department's spreadsheet. By quantifying these costs—in lost staff hours and missed revenue—organisations can better appreciate the urgent need for a cohesive data strategy.

2. BigQuery: The "Silo-Killer"

To break these walls, we must move beyond traditional manual consolidation. This is where **Google BigQuery** acts as the **"Silo-Killer."** Unlike a standard database, BigQuery is a cloud-based data warehouse designed to ingest vast amounts of information from disparate sources—Salesforce, Xero, Excel, and even Social Media metrics—and "handshake" them into a single, unified framework.

By centralising these records, BigQuery eliminates the "version control" nightmares that plague board meetings. Data analysts and IT leads no longer spend 70% of their time cleaning data; instead, they use BigQuery to create an automated pipeline. This allows the organisation to see, for the first time, a **360-degree view of the donor**. You can finally correlate a specific marketing spend with an actual donation and, crucially, with the operational impact that donation achieved.

3. From Redundancy to the Stewardship Dividend

Transitioning from data chaos to predictive stewardship requires this technical foundation to support a strategic shift. When silos are removed, AI and machine learning techniques can finally be applied to the entire dataset. IT leads can then segment donors more effectively, tailoring strategies that resonate with their specific interests and behaviours across all touchpoints.

This is where the concept of the **'Stewardship Dividend'** matures. It encapsulates the idea of monetising the efficiency gains achieved through better data management. For trustees and board members, understanding this dividend is essential; it proves that the investment in a modern data architecture isn't just an IT expense—it is a revenue-generating asset. By reclaiming the hours previously lost to manual data entry and "spreadsheet wrestling," your team can focus on fostering deeper connections with donors, thereby increasing loyalty and long-term financial sustainability.

4. Building the "Silo-Free" Narrative

The path to predictive stewardship is paved with the insights generated when data flows freely. Organisations that harness this potential find themselves at the forefront of the non-profit sector's evolution. They are the ones who can answer a donor's question about impact in seconds, rather than days. They are the ones who can see a funding gap on the horizon and pivot before it becomes a crisis.

Chapter 3: The ROI of Predictive Propensity

In the modern non-profit landscape, understanding donor behaviour has never been more critical. Chapter 3 delves into the transformative power of predictive analytics through **BigQuery ML**, illustrating how artificial intelligence can help quantify donor loyalty and move your fundraising from "begging" for support to "inviting investment" in a proven, calibrated mission.

1. Activating the AI Brain: BigQuery ML

BigQuery ML stands out as a powerful tool that provides non-profits with the means to perform sophisticated data analyses without requiring a PhD in data science. Traditionally, predictive modelling was locked behind expensive consultancy fees or complex coding. BigQuery ML democratises this, allowing your data analysts to use standard SQL to train machine learning models directly within the data warehouse.

By leveraging AI, organisations can finally move away from **Data Chaos**. Instead of looking at a donor's total lifetime value in a vacuum, BigQuery ML looks at the "Signals" between the data—analysing the mathematical relationship between how recently they gave, how often they give, and the channels they prefer. This moves the organisation toward a cohesive framework of Predictive Stewardship.

2. The Stewardship Dividend: Monetising Propensity

A fundamental concept introduced in this chapter is the **"Stewardship Dividend."** This term encapsulates the idea of monetising efficiency through AI-driven insights. By employing machine learning algorithms like **Linear Regression**, organisations can identify which donors are most likely to continue their support with a high degree of statistical accuracy.

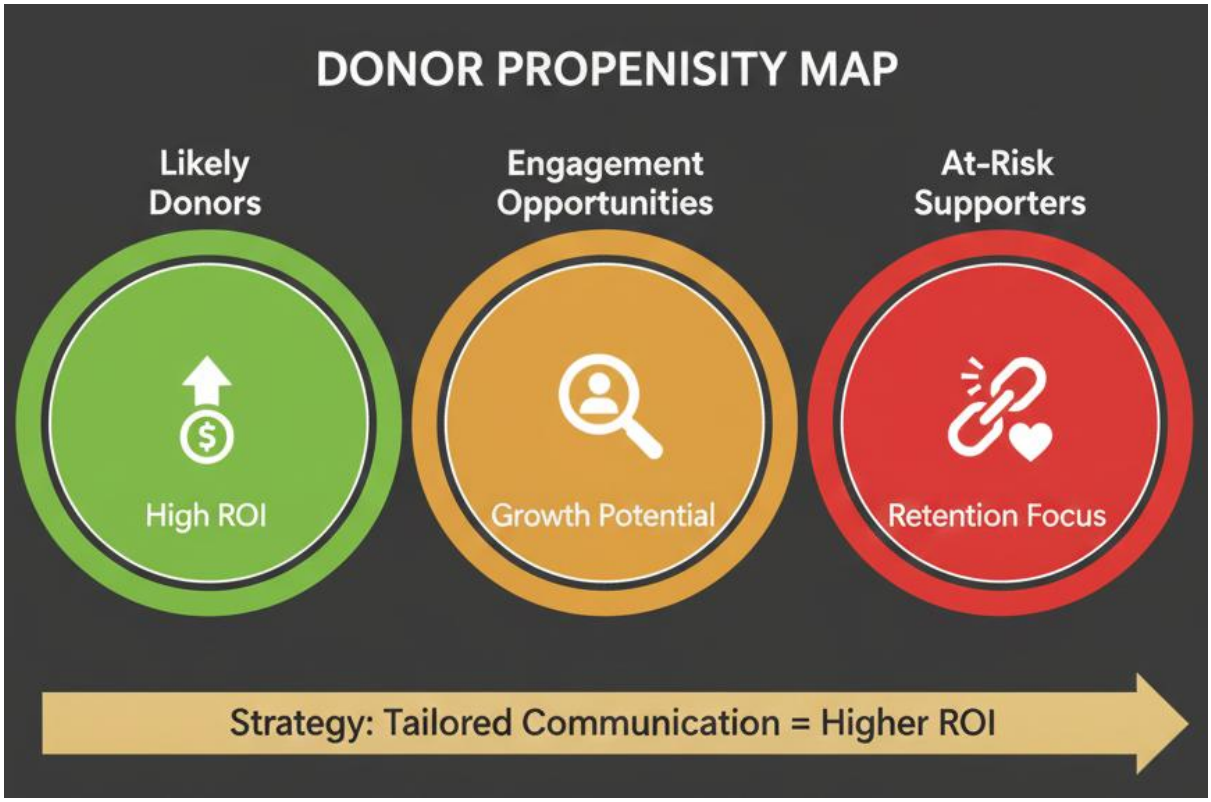
This capability allows leaders to stop wasting marketing budget on low-propensity segments and instead focus their resources on nurturing relationships with high-potential donors. For the COO, the "Dividend" is the reclaimed budget and staff time that can now be reinvested into frontline services. This shift not only stabilises funding but also allows organisations to tailor their engagement strategies effectively, ensuring every touchpoint increases loyalty.

3. Segmenting with Precision

The application of AI in donor segmentation allows non-profits to craft personalised strategies that resonate with different cohorts. We move beyond simple "Major Donor" vs. "Regular Giver" labels. With BigQuery ML, we can create segments based on **Predicted Likelihood**:

- **The Bedrock:** High-propensity donors who require consistent stewardship to protect long-term stability.
- **The Fence-Sitters:** Moderate-propensity donors who need a specific "Impact Story" or personalized nudge to move toward a commitment.
- **The Churn Risk:** Formerly active donors whose behavioral signals (long gaps between gifts) suggest they are drifting away.

By understanding the unique motivations of each group, organisations can implement targeted campaigns that significantly enhance engagement. This tailored approach encourages donors to contribute more consistently, thereby ensuring the financial stability of the organisation.



4. Visualising the 12-Month Horizon

By integrating BigQuery ML with **Power BI**, organisations can visualise their predictive data in meaningful ways. Instead of presenting the Board with a static list of names, you present a **12-Month Predictive Horizon**. This makes it easier for board members and trustees to comprehend donor trends and make informed decisions about future spending.

When you can show a Board that your model explains **78% of donor behaviour** (using the R-Squared metric), you provide the **"Audit-Proof Certainty"** required for large-scale strategic expansion. You are no longer guessing your budget; you are leading with empirical evidence.

Conclusion: A Strategic Imperative

The journey from data chaos to predictive stewardship is not merely a technological transformation; it is a strategic imperative. By harnessing the capabilities of BigQuery ML, organisations create a more robust funding model that is resilient to economic shifts. This cultivates a culture of informed decision-making, paving the way for sustainable growth and a deeper, more profound impact in the community.

Chapter 4: The Executive Interface

In the contemporary landscape of non-profit management, visualisation tools like **Power BI** serve as pivotal lenses through which Boards and executive teams gain a 360-degree view of their organisation's operational health. While BigQuery serves as the **"Brain,"** Power BI is the **"Executive Interface"**—the layer that translates raw data chaos into a cohesive narrative of mission impact and financial sustainability.

1. Power BI: The Lens for the Board

The ability to transform complex, multi-silo data into intuitive visuals allows board members to quickly grasp critical metrics without getting lost in technical minutiae. As non-profit leaders navigate the intricacies of fundraising and donor engagement, Power BI becomes an essential ally, translating fragmented data from your CRM, finance software, and impact trackers into actionable intelligence.

This transition is not merely a technological upgrade; it is a strategic imperative. It fosters a culture of transparency where the Board no longer asks "What happened?" but instead discusses "What is next?" By presenting data in real-time, leadership can shift from reactive survival to proactive stewardship.

2. Visualizing the "Stewardship Dividend"

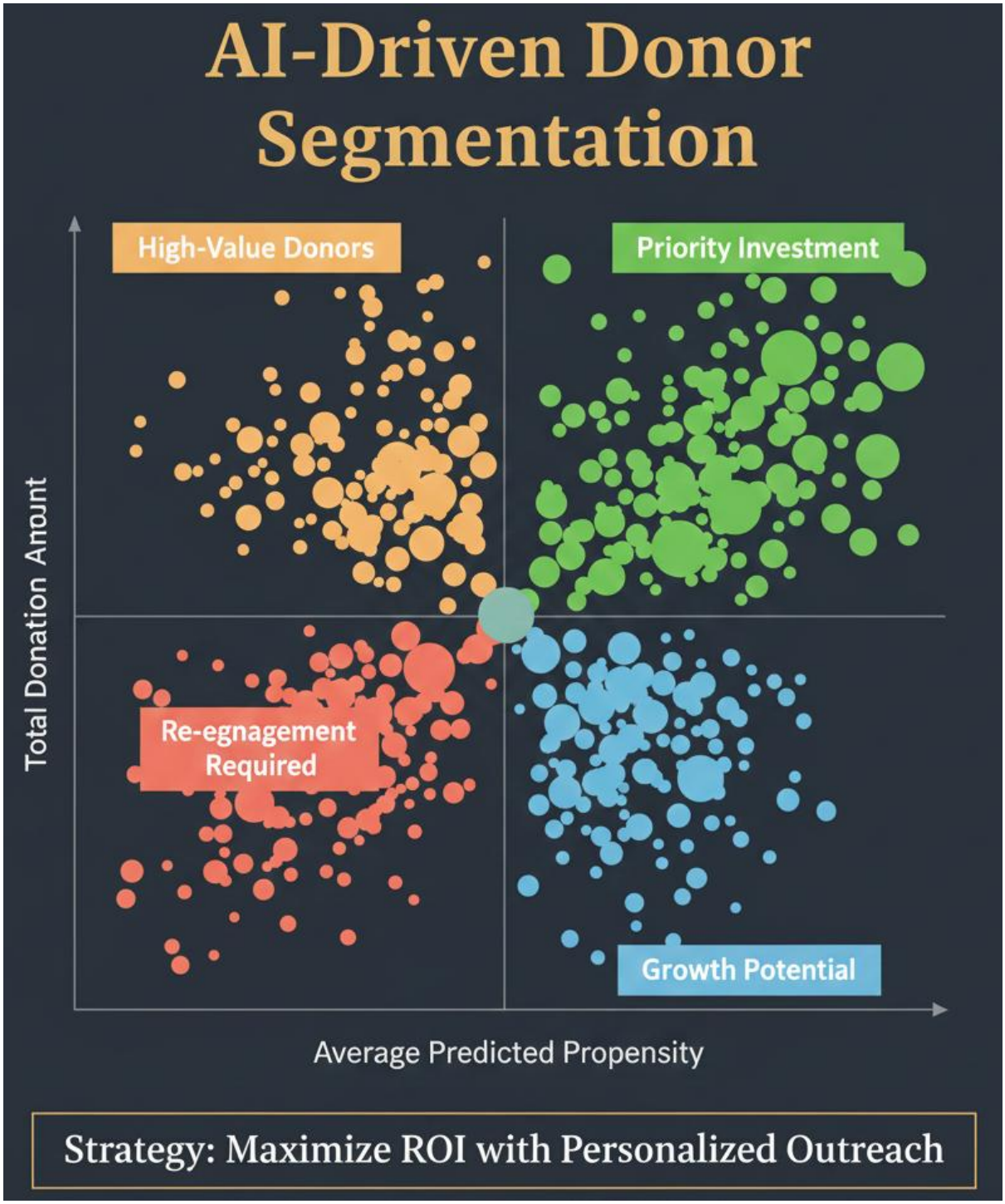
The concept of the **"Stewardship Dividend"** emerges as a cornerstone of effective governance. By leveraging data visualisation, Boards can identify trends that monetise efficiency—such as identifying which fundraising channels provide the highest ROI or pinpointing exactly where staff time is being lost to manual, repetitive reporting.

By implementing BigQuery ML alongside Power BI, organisations create a robust framework that enables precise **Donor Segmentation**. Instead of a flat, static list of names, the Board sees a dynamic **"Propensity Map."** This targeted strategy allows the organisation to allocate limited resources to high-potential relationships, fostering the long-term loyalty essential for sustained organisational success.

3. Real-Time AI and Machine Learning Insights

The power of AI in donor segmentation cannot be overstated. Power BI's visualisation capabilities allow Boards to monitor these AI-driven strategies in real-time. For instance, a scatter chart can instantly reveal which campaigns are meeting the nuanced preferences of diverse donor groups and which are failing to resonate.

This level of personalised engagement increases the likelihood of donor retention. If a campaign's engagement score dips, the executive team sees it on the dashboard immediately—not three months later in a "post-mortem" report. This ensures that adjustments can be made swiftly in response to shifting donor dynamics.



4. Fostering a Data-Centric Culture

As non-profit CEOs and COOs adopt this data-driven approach, they must also lead a cultural shift. Embracing a data-centric mindset requires a commitment to **Data Literacy** across the entire team. Board members play a crucial role in championing this transformation, setting the tone for how data is perceived and utilised.

By prioritising data literacy, organisations empower their teams to leverage insights effectively. When a Fundraiser can look at the dashboard and see exactly which donors are **"At-Risk,"** they are no longer guessing who to call; they are acting on empirical evidence. This drives a more informed and strategic approach to resource allocation and donor engagement.

Conclusion: Redefining Strategic Decisions

In conclusion, the integration of Power BI into the executive interface represents a significant advancement for non-profit boards. Visualising data through this lens provides a clear and compelling evidence-base for understanding organisational performance. The journey from data chaos to predictive stewardship is not just about technology; it is about redefining how non-profits engage with their stakeholders and make strategic decisions. By embracing these innovations, non-profit leaders position their organisations for sustainability and growth in an increasingly competitive landscape.

Chapter 5: Radical Transparency

In the increasingly competitive world of non-profit funding, "good intentions" are no longer the primary currency for securing multi-year grants or board approval for major initiatives. Stakeholders now demand **Stewardship Authority**. Chapter 5 explores how to leverage high-integrity data to transform donor and board relationships from simple transactions into deep, strategic partnerships founded on **Radical Transparency**.

1. The Evidence-Based Narrative

Radical transparency is pivotal for establishing trust. By leveraging evidence-based impact narratives, organisations can articulate their mission and the results of their programmes with absolute clarity. This shift moves the conversation away from anecdotal success stories toward a cohesive narrative of measurable change.

The integration of data-driven narratives allows for a profound connection between the organisation’s goals and stakeholder expectations. When a CEO can show exactly how a specific campaign directly funded a service expansion, the organisation’s credibility in the eyes of donors and the community is exponentially enhanced.

2. The Confidence Scorecard: Model Accuracy (R-Squared)

The most powerful tool in your fundraising arsenal is the ability to prove that your projections aren't just guesses. This is where we move from the "what" to the "how well." In the world of analytics, the **R-Squared value** provides a quantitative measure of how well your data supports the narratives you present.

High R-Squared values (such as the 0.778 achieved in our case study) signify that your model explains a significant portion of donor behaviour. Presenting this to a board signifies that your strategies are grounded in solid data, not assumptions. This **"Confidence Scorecard"** ensures that all strategic decisions are audit-proof and statistically significant, fostering a culture of accountability and continuous improvement.

3. Monetising Efficiency through the Stewardship Dividend

The concept of the **"Stewardship Dividend"** plays a critical role in this transformation. It encapsulates the idea of monetising efficiency through enhanced data management and predictive analytics. By accurately forecasting donor propensity, non-profits can stop "spraying and praying" with their marketing budget and instead allocate resources where they will have the highest ROI.

This proactive approach reinforces the organisation's mission to deliver measurable results to its beneficiaries. When a finance leader can show the board that an investment in BigQuery ML led to a "Dividend" of reclaimed staff hours or optimized gift processing, they prove that data engineering is a direct contributor to the mission's bottom line.

Conclusion: Securing Sustainable Funding

In conclusion, radical transparency through evidence-based narratives and model accuracy is essential for building unbreakable trust. Non-profit CEOs and finance leaders must embrace these innovative data solutions to transition from fragmented data management to a paradigm of **Predictive Stewardship**.

This transition does more than just optimise operations; it empowers organisations to articulate their impact compellingly. By showing donors and boards the "math" behind the mission, you enhance donor loyalty and secure the sustainable funding necessary to navigate the challenges of tomorrow with confidence.

True donor engagement happens when the entire organisation—from the frontline staff to the boardroom—embraces a culture of **Data-Driven Decision-Making**. This shift fundamentally changes the nature of the conversation you have with your supporters and stakeholders.

- **From "Stories" to "Evidence-Based Narratives":** While anecdotes move hearts, data-driven evidence moves pens to sign cheques. We transition from saying "We are doing our best" to saying: *"Our AI models identified 479 likely supporters, and by retaining 80% of them, we can fund 1,200 additional hours of service delivery this year."*
- **Context-Aware Engagement:** Using **Power BI Slicers**, you can show a donor exactly how their specific grant impacted a particular Campaign or Channel. This granular visibility turns a "general donation" into a "targeted investment."

- **Proactive Stewardship:** Instead of reporting on a deficit after it happens, you use **Strategic Foresight** to tell a donor: *"Our propensity model shows a potential dip in support for 2026; your early funding today will act as a stabilisation fund to ensure our mission remains uninterrupted."*

The ROI of Radical Transparency

When you provide donors with this level of clarity, the **"Stewardship Dividend"** is immediate. You move away from "begging" for funds and toward "inviting investment" in a calibrated, professional system.

Donor Concern	Your Data-Driven Response
"Is my money being wasted?"	Show the **Avg Donation Frequency** and **ROI per Channel**.
"Are you reaching enough people?"	Show the **Donor Growth trend (+9.3%)** vs. **Income growth (+34.9%)**.
"Do you have a plan for the future?"	Show the **Likely Donation (2026) AI forecast (£114.8k)**.

The Case Study: The Transformation of Second Chance Futures

The Fundraising Crisis: Before adopting the **Stewardship Blueprint**, Second Chance Futures was struggling with an invisible ceiling. The team was managing donor relationships through "Intuition and Spreadsheets," unaware that their **Donor Retention** was slipping toward a dangerous 61.8%. Because they relied on manual, retrospective tracking, they couldn't identify that while they were raising more money, they were becoming overly dependent on a shrinking pool of "Super Supporters."

The Operational Crisis: They were "Data Rich but Insight Poor." While they held thousands of rows of donation history in Salesforce, they lacked the ability to translate those rows into **Predictive Propensity**. Without an AI model, they couldn't justify the need for a dedicated Stewardship Manager to the Board. They were operating blindly, unable to see that a minor improvement in retention could "manufacture" the equivalent of a major corporate grant without a single new acquisition.

The Stewardship Transformation: By transitioning to **Predictive Donor Modeling**, Second Chance Futures moved from reactive "fire-fighting" to proactive growth. They stopped simply counting gifts and began using **Linear Regression** models in BigQuery to forecast future income. By building an **Executive Dashboard** in Power BI, they achieved a 34.9% increase in total income, allowing the COO to foresee and mitigate funding "Danger Zones" twelve months in advance.

The Result: Today, Second Chance Futures leads with operational authority. Every campaign is optimized, every churn risk is foreseen, and every strategic decision—from hiring clinical staff to launching a winter appeal—is backed by high-integrity, automated AI forecasting.

A Note to the Reader

If you are a **Chief Operating Officer (COO)**, **CEO**, or **Board Member**, you have reached the summit of the **"Strategic Model."** The preceding chapters have outlined the *Why* and the *Where*—the destination of Predictive Stewardship and the high-level outcomes of financial stability.

Your Role in this Journey: Your primary focus is the **Outcomes**: the dashboards, the 2025 revenue projections, and the strategic recovery of donor retention. You do not need to write the SQL code to benefit from the power of the model.

What follows is the **"How."** The next section of this book is a rigorous, technical manual designed for the **"Technician"**—the Data Analyst, the IT Lead, or the Finance Manager. It is the step-by-step blueprint they need to build the **"Propensity & Retention Engine"** described in the training.

How to Use the Technical Section

- **If you are the Executive:** You may choose to pause here. Your next step is to hand this book to your data lead and say: *"I want the outcomes shown in the report; here is the technical manual to build the system that delivers them."*
- **If you are the Technician:** This is where your work begins. You are the architect of the **Stewardship Blueprint**. The following phases are designed to be followed sequentially. I have provided the exact BigQuery SQL, AI Model logic, and Power BI DAX required to transform your donation data into a professional forecasting system.

Important: Maximizing Your Learning Before you begin the technical implementation, it is highly recommended that you watch the accompanying training video for this module. While this manual provides the exact code and steps, the video offers critical context and visual demonstrations of the interface.

The Technical Roadmap: From Data Chaos to AI Foresight

This master blueprint ensures the technician understands the flow of data across the **mlprojecttraining-arkbi2025** environment before beginning the granular build.

Phase 1: The Foundation (Fact & Lookup Views)

- **Structural Integrity:** Staging the data in BigQuery. We establish the "Single Source of Truth" by creating abstraction views for your Salesforce data and building the relational "Lookup" tables (Donors, Campaigns, Channels, and Calendar) that prevent double-counting.

Phase 2: The Brain (BigQuery ML)

- **Activating the AI:** This is the Machine Learning layer. We perform **RFM Feature Engineering** to create donor signals and deploy **Linear Regression** models to calculate propensity scores, moving the organization from retrospective reporting to 2025 forecasting.

Phase 3: The Skeleton (The Relational Skeleton)

- **Establishing Relational Integrity:** We transition to the Power BI interface to architect the **Star Schema**. By connecting our BigQuery Lookups to our Fact tables, we ensure the entire dashboard is "Context-Aware" and audit-proof.

Phase 4: The Intelligence (DAX Measures)

- **Moving Beyond Simple Counts:** We deploy Power BI DAX to calculate the "Muscles" of the report. This includes **Time Intelligence** (YoY Growth), **Donor Retention** percentages, and the **Predictive Revenue** metrics that consume our AI scores.

Phase 5: The Interface (Executive Summary)

- **The Visual Payoff:** The final assembly. We build the high-impact report using interactive slicers and KPI cards, allowing the COO and the Board to drill down into specific campaigns and channels without losing the strategic narrative.

Technical Setup: Final Confirmation

Phase	Core Tool	Outcome
1 & 2	Google BigQuery	Relational Views & AI Model Training
3 & 4	Power BI Desktop	Data Modeling & Intelligence (DAX)
5	Power BI Service	Executive Storytelling & Strategic Decision-Making

Course Onboarding: The Legacy Foundation

Welcome to the course! To build your **Legacy Architecture**, you must first prepare your digital workspace. This guide is divided into two phases: **Setting up your Cloud Environment** and **Securing your Project Assets**.

Phase 1: Environment Configuration

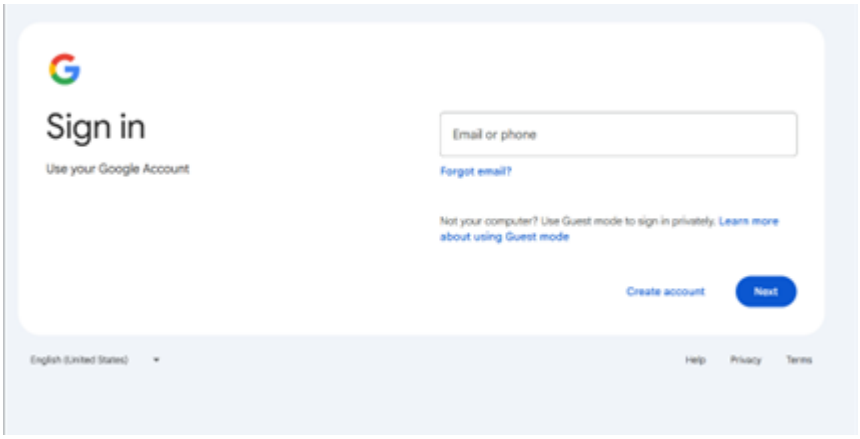
Before downloading your data templates, you must establish the "Brain" and "Interface" where your models will live.

Step 1: Set Up Your Google Accounts

First, we will set up your new, dedicated Gmail and Google Cloud accounts. It is highly recommended that you create a new Gmail account specifically for this course to keep your personal data separate and to manage your cloud projects more cleanly.

1.1 Creating a Dedicated Gmail Account

1. **Navigate to the Creation Page:** Open your web browser and go to the Gmail creation page at https://accounts.google.com/signin .

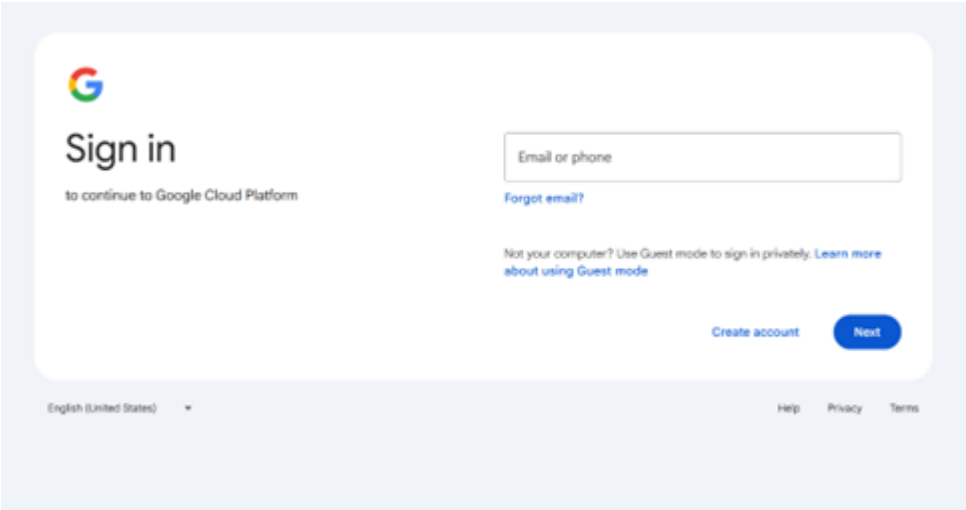


2. **Start Account Creation:** On the sign-in page, click the **"Create account"** button at the bottom of the form. A menu will pop up; select **"For my personal use"** to begin the process.

3. **Fill in Your Details:**
    - You will be asked to enter your first and last name.
    - On the next screen, enter your birth date and gender.
    - Choose a unique username for your new email address (e.g., yourname.datacourse.2025@gmail.com).
    - Create a strong password and confirm it. It's a good practice to use a combination of uppercase and lowercase letters, numbers, and symbols.
4. **Add Recovery Information:** For security, you will be prompted to add a recovery email address or phone number. This will help you regain access to your account if you forget your password.
5. **Review and Agree:** Read through Google's Privacy and Terms, and if you agree, click **"I agree"** to finalize the account creation. You will now be taken to your new, empty Gmail inbox.

1.2 Activating Your Google Cloud Account

1. **Go to the Google Cloud Console:** In a new browser tab, go to the Google Cloud Console at https://console.cloud.google.com/



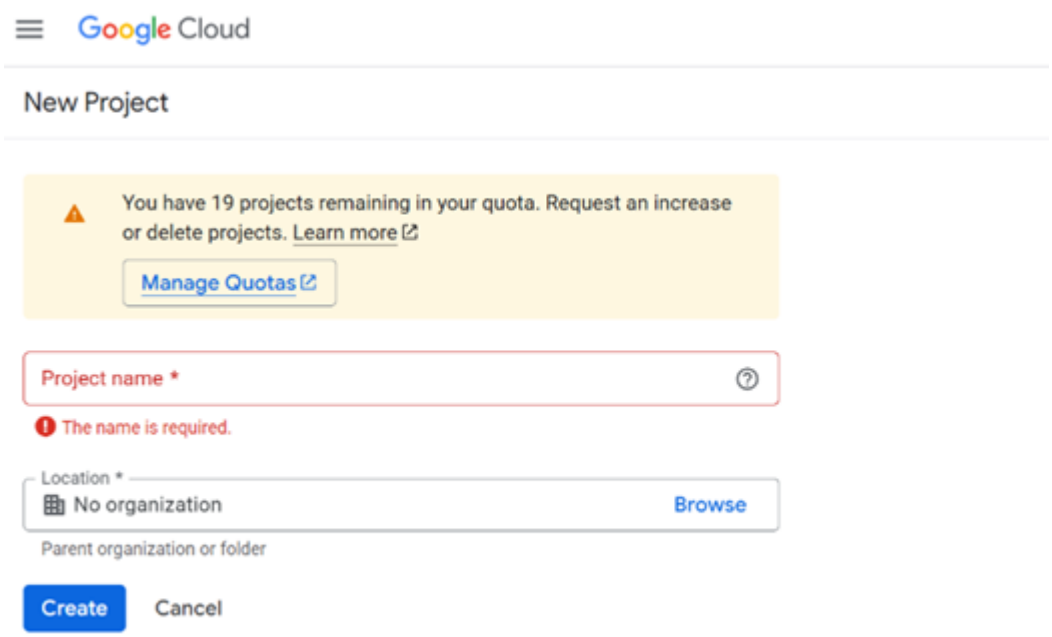
2. **Sign In:** Sign in using the new Gmail account you just created.
3. **Activate Free Trial:** If this is your first time using Google Cloud, you will be automatically prompted to activate a free trial. Click **"Activate"** or **"Start Free Trial"**.
4. **Complete Billing Setup:** You will be taken to a billing setup page. You must fill out the form, including your contact information and payment method. **You will not be charged unless you manually upgrade to a paid account after the trial period ends.** Complete this process to gain access to all the services we need for the course.

Step 2: Create a Project and Enable APIs

In Google Cloud, a project is like a dedicated, isolated folder for all your cloud services and data. Creating a new project keeps our course work organized and separate from any other potential cloud work.

2.1 Creating a New Google Cloud Project

1. **Navigate to the Project Menu:** Once you are in the Google Cloud Console dashboard, look for the project dropdown menu at the top of the screen.
2. **Select "New Project":** Click on the project name, then click **"New Project"** from the menu that appears.

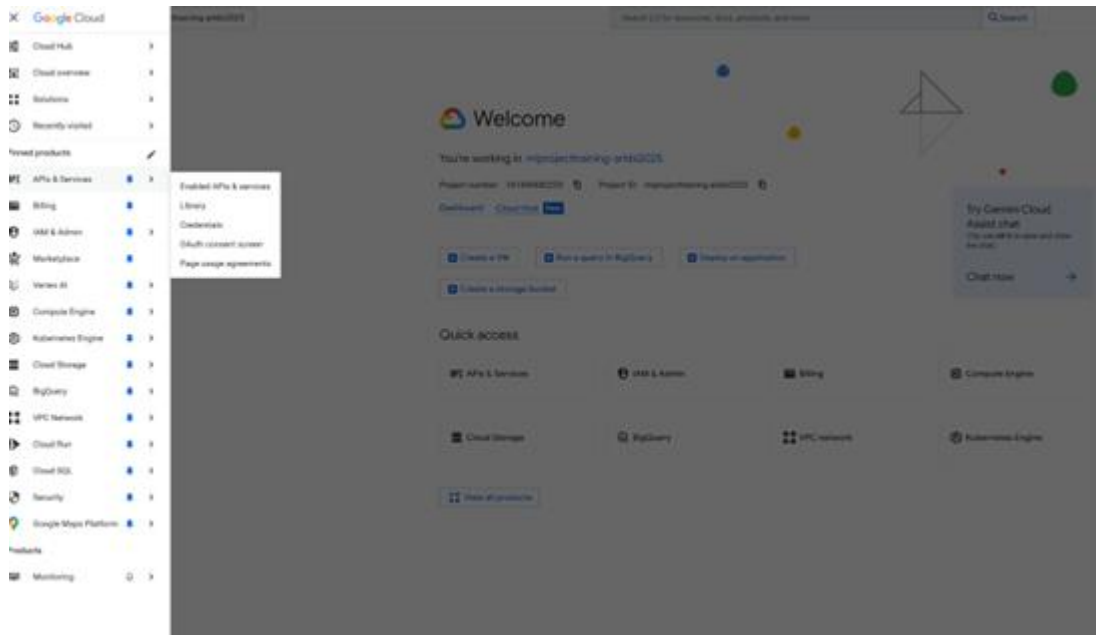


3. **Name Your Project:** In the "New Project" form, enter a unique name and Project ID for your project (e.g., yourname-ml-project-2025). **Remember:** Project IDs must be unique globally, so if your first choice is taken, you will need to try a different ID.
4. **Create the Project:** Click **"Create"**. It may take a minute or two for your project to be created and for you to be redirected to its dashboard.

2.2 Enabling Necessary APIs

APIs (Application Programming Interfaces) are what allow different services to communicate. For this course, we need to enable two key APIs.

1. **Go to the API Library:** In the left-hand navigation menu of the Google Cloud Console, search for **"APIs & Services"** and click on **"Library"**.



2. **Enable BigQuery API:** In the API search bar, type **"BigQuery API"**. Click on the search result, and then click the **"Enable"** button.
3. **Enable Gemini Cloud Assist:** Go back to the API Library search bar. This time, search for **"Gemini Cloud Assist"**. Click on the search result, and then click the **"Enable"** button.

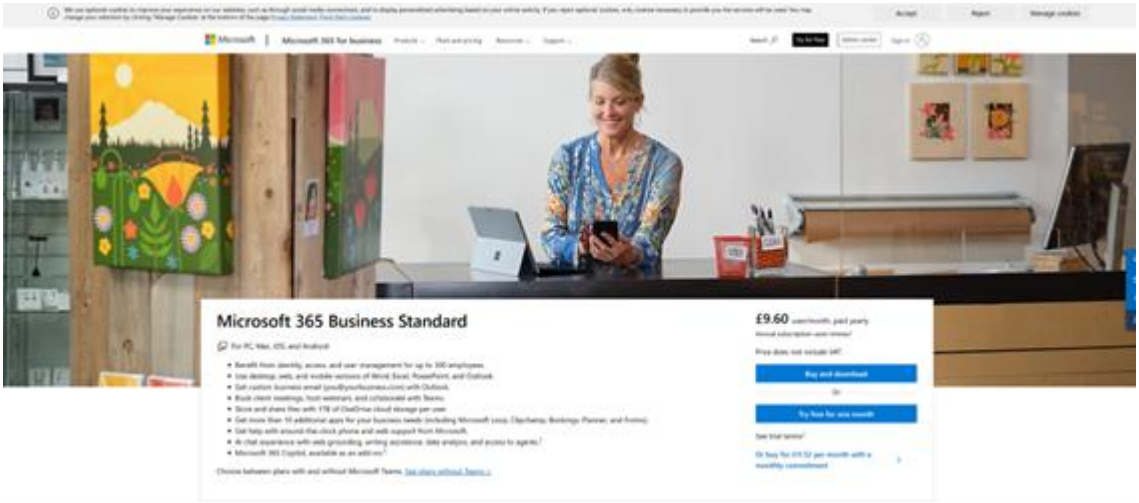
Step 3: Set Up Power BI

For this course, you will need both the Power BI Service (the online portal for sharing and collaboration) and the Power BI Desktop app (the application you install on your computer for building reports).

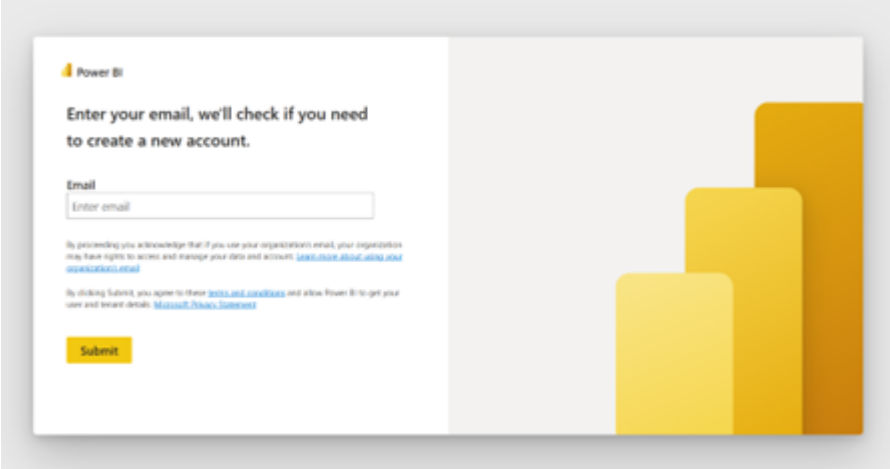
3.1 Creating Your Power BI Service Account

The Power BI Service requires a work or school email address. If you don't have one, the easiest workaround is to sign up for a **Microsoft 365 free trial**. This will create an @onmicrosoft.com email address for you, which is perfect for accessing Power BI.

1. **Start the Microsoft 365 Trial:** Go to the Microsoft 365 Enterprise plans page to get started. https://www.microsoft.com/en-gb/microsoft-365/business/microsoft-365-business-standard?msockid=141f76a2ac896a0537ed60a3adcc6b67&activetab=pivot:overviewtab Find a suitable plan (e.g., Microsoft 365 Business Standard) and click the button to **"Try free for one month"**.



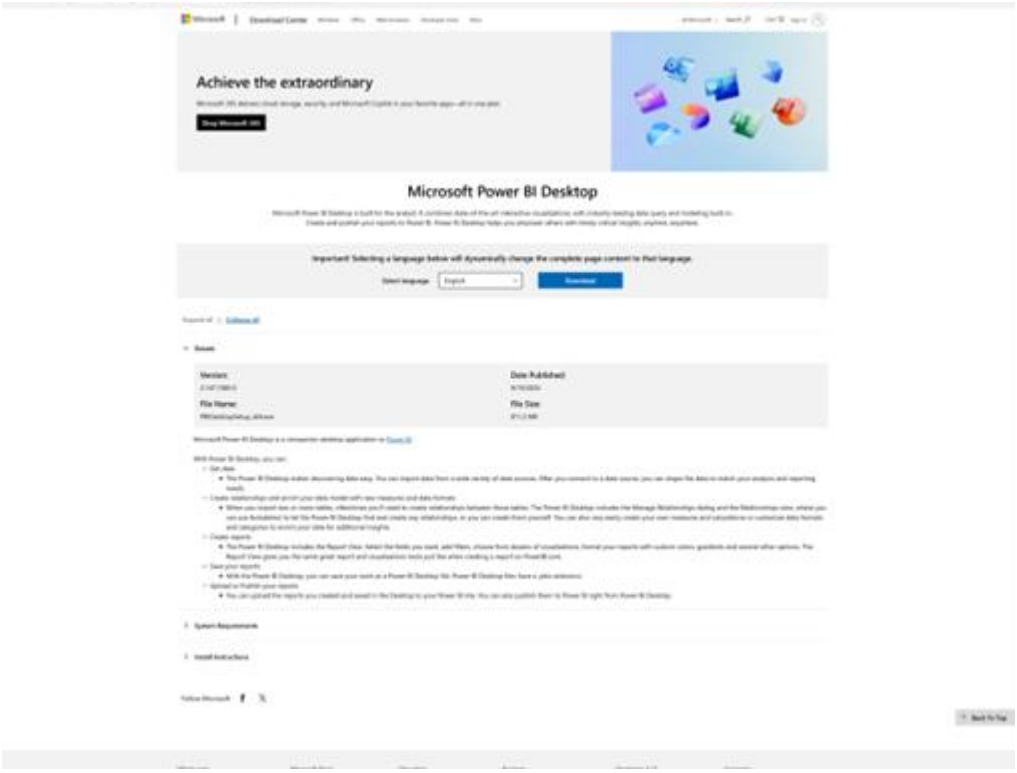
2. **Complete the Sign-Up:** Follow the on-screen prompts to set up your account. You will be asked for your personal details and will be required to provide a credit card, but **you will not be charged** unless you continue the subscription after the trial. Microsoft will automatically create a new business account for you with an email address ending in @onmicrosoft.com.
3. **Access the Power BI Service:** Once your new @onmicrosoft.com email is created, go to the Power BI Service login page. https://app.powerbi.com/ Enter your new email address and click **"Submit"**.



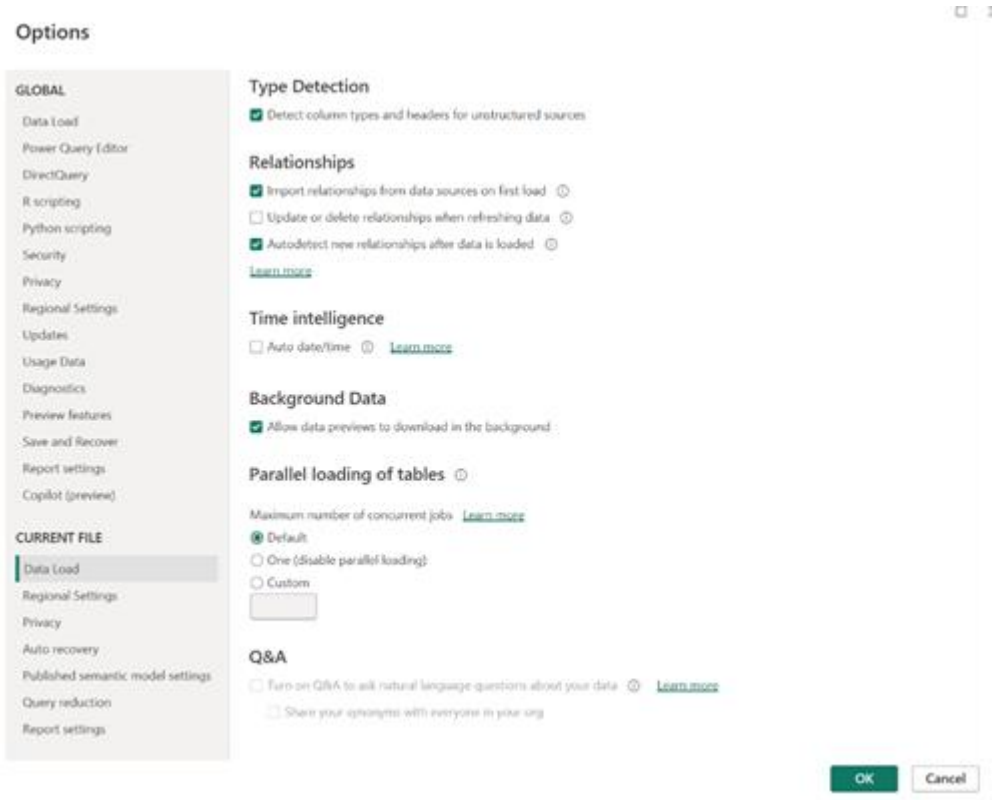
4. **Create a Workspace:** Once logged in, you will be in your personal Power BI workspace. To keep your work organized, we will create a dedicated workspace for this course.
    - Look at the left-hand navigation pane and click on **"Workspaces"**.
    - In the top-left corner, click **"+ New workspace"**.
    - Give your workspace a name like mlproject-arkbi2025 and then click **"Save"**. This is where we will be storing all our data and reports.

3.2 Installing and Configuring Power BI Desktop

1. **Download the Installer:** Navigate to the Power BI Desktop download page https://www.microsoft.com/en-us/download/details.aspx?id=58494. Select your language from the dropdown menu and click **"Download"**.



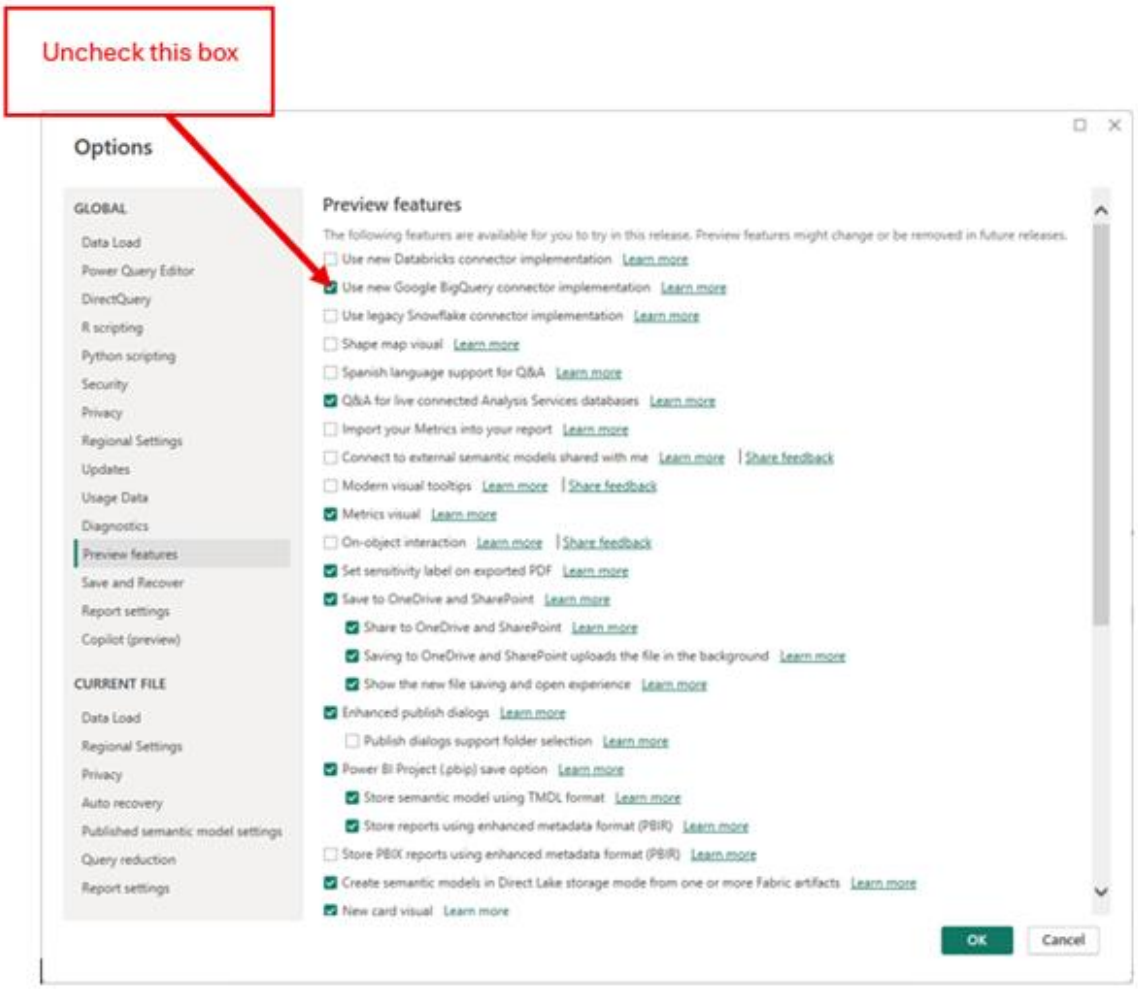
2. **Install the Application:** Locate the downloaded installer file and double-click it. Follow the on-screen installation wizard to install Power BI Desktop on your computer.
3. **Configure a Crucial Setting:** Once Power BI Desktop is installed, open it and follow these steps to prevent a common issue that can arise in complex data models:
    - Go to **"File"** in the top-left corner.
    - Select **"Options and settings"**, then click **"Options"**.
    - In the options window, select **"Data Load"** from the left-hand menu.
    - Find the option **"Autodetect new relationships after data is loaded"** and **uncheck the box** next to it.
    - Click **"OK"** to save your changes. This setting gives you manual control over your data model, which is a best practice for advanced data analysis.



3.3 Configuring the Big Query Connector

Some users encounter an ADBC error when connecting to BigQuery due to a newer connector implementation. To prevent this, you'll need to disable a preview feature in Power BI.

- **Open Power BI Desktop.**
- Go to **File** > **Options and settings** > **Options**.
- In the options window, select **Preview features** from the left-hand menu.
- Find and **uncheck** the box next to "**Use new Google BigQuery connector implementation**".
- Click **OK** to save your changes and then **restart** Power BI Desktop to apply the new setting.


Uncheck this box
Options
GLOBAL
Data Load
Power Query Editor
DirectQuery
R scripting
Python scripting
Security
Privacy
Regional Settings
Updates
Usage Data
Diagnostics
Preview features
Save and Recover
Report settings
Copilot (preview)
CURRENT FILE
Data Load
Regional Settings
Privacy
Auto recovery
Published semantic model settings
Query reduction
Report settings
Preview features
The following features are available for you to try in this release. Preview features might change or be removed in future releases.
Use new Databricks connector implementation Learn more
Use new Google BigQuery connector implementation Learn more
Use legacy Snowflake connector implementation Learn more
Shape map visual Learn more
Spanish language support for Q&A Learn more
Q&A for live connected Analysis Services databases Learn more
Import your Metrics into your report Learn more
Connect to external semantic models shared with me Learn more | Share feedback
Modern visual tooltips Learn more | Share feedback
Metrics visual Learn more
On-object interaction Learn more | Share feedback
Set sensitivity label on exported PDF Learn more
Save to OneDrive and SharePoint Learn more
Share to OneDrive and SharePoint Learn more
Saving to OneDrive and SharePoint uploads the file in the background Learn more
Show the new file saving and open experience Learn more
Enhanced publish dialogs Learn more
Publish dialogs support folder selection Learn more
Power BI Project (.pbip) save option Learn more
Store semantic model using TMDL format Learn more
Store reports using enhanced metadata format (PBIR) Learn more
Store PBIX reports using enhanced metadata format (PBIR) Learn more
Create semantic models in Direct Lake storage mode from one or more Fabric artifacts Learn more
New card visual Learn more
OK
Cancel

Phase 2: Project Assets & Resources

With your environment staged, you can now secure the raw materials required for the build.

Important: Maximising Your Learning

Before you begin the technical implementation, **it is highly recommended that you watch the accompanying training video for this module.** While this manual provides the exact code and steps, the video offers critical context, visual demonstrations of the interface, and the "why" behind each architectural decision. Combining both resources will ensure you build a robust, audit-proof solution.

Your Learning Resources

These provide the visual context and the raw data required to execute the phases below.

Download resources Link

Preparation: Essential Resources & Assets

To successfully execute this blueprint, ensure you have secured the following files. These assets provide the raw material and the "Gold Standard" reference points required for your build.

- **Salesforce_Donor_Data (CSV Snippet):** This is your raw input file. It contains anonymized historical donation records, donor IDs, and campaign categories. This data serves as the "fuel" for your BigQuery ML models and the foundation for your relational model.
- **BlankTemplate—Building a Predictive Donor (.pbix):** Your primary workspace. This Power BI file is pre-configured with the necessary theme settings but contains no data or logic. Use this to follow the step-by-step assembly of the Star Schema and DAX measures.
- **CompletedTemplate—Building a Predictive Donor (.pbix):** The "Gold Standard" reference. This file contains the finalized architecture, including the end-to-end Star Schema, all advanced DAX measures, and the high-impact executive dashboards. Use this to audit your work and verify your results as you progress through the phases.

Step 2: Access the Training Modules

This guide is a companion to the video training series. Each technical chapter corresponds to a specific video phase. For the best results, watch the modules sequentially as you interact with the manual.

Step 3: Integrated Technical Workbook

The technical logic—including exact **SQL scripts**, **Power Query M-Code**, and **DAX formulas**—is embedded directly within each relevant technical chapter of this manual. This "Inline Workbook" approach ensures you have the code at your fingertips precisely when you need to copy and paste it into your BigQuery or Power BI environments.

A Note on Technical Scalability & Professional Support

The case study presented in this manual is a streamlined, foundational version of an enterprise-grade solution. In a production environment, the **Second Chance Futures** project utilised highly sophisticated **ETL (Extract, Transform, Load)** pipelines to manage data orchestration.

The full implementation involved integrating diverse data streams, including live foreign exchange rates for international grant compliance and multi-layered third-party CRM systems. To ensure you master the **Core Architecture of Power Query and the Power BI Data Model** without unnecessary complexity, I have distilled these elements into this blueprint. By completing this course, you will possess the structural knowledge required to lead more complex data engineering initiatives and architect professional-grade business intelligence solutions in the future.

Transitioning to Execution: Launching the BigQuery Workspace

Now that your Google Cloud Project is established and your raw data is staged, it is time to engage the "Brain." We will perform the remainder of our data engineering within the **BigQuery SQL Workspace**. This is where we transform raw CSV data into intelligent, relational views and train our AI models.

How to access your workspace:

1. **Open the Console:** Go to console.cloud.google.com.
2. **Select Your Project:** Ensure your Project ID (e.g., mlprojecttraining-arkbi2025) is selected in the top blue header.
3. **Search for BigQuery:** In the search bar at the top, type **"BigQuery"** and select the first result.
4. **Pin the Workspace:** Once the BigQuery page loads, you will see the **Explorer** pane on the left and the **Query Editor** in the center. This is your primary development environment.

Expert Tip: Keep this browser tab open throughout the following chapters. You will be toggling between this manual and the BigQuery editor to copy, paste, and run your SQL scripts.

Phase 1: The Foundation (BigQuery SQL Abstraction)

In this first technical step, we transition from **Data Chaos** to **Structural Integrity**. We do not work directly on raw tables; instead, we build an **Abstraction Layer**. This ensures that if your underlying Salesforce export changes in the future, you only have to update it in one place without breaking your entire Power BI suite.

1. Creating the Salesforce_Data View

The goal here is to create a "Mirror View." This acts as your official **Single Source of Truth** for the entire project. By using a View, we can change the logic in the cloud later (like renaming columns or filtering dates) without the Executive ever seeing a "broken" dashboard in Power BI.

Detailed Instructions

1. **Open BigQuery:** Go to your Google Cloud Console and navigate to the **BigQuery SQL Workspace**.
2. **Locate your Dataset:** Ensure you have already created a dataset named `DonorAnalysis`.
3. **Compose New Query:** Click the **+** (Compose New Query) button to open a clean editor tab.
4. **Update the Project ID:** In the code block below, you **must** replace `mlprojecttraining-arkbi2025` with your unique Google Cloud Project ID.
5. **Run the Script:** Paste the updated code into the editor and click the **Run** button.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID before clicking 'Run'.
*/

CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.Salesforce_Data` AS
SELECT
    * FROM
    `mlprojecttraining-arkbi2025.DonorAnalysis.SalesforceData`;
```

Code Description

- **CREATE OR REPLACE VIEW**: This is a "Stewardship Standard." It ensures the view is created—or updated if it already exists—allowing you to refine your data without encountering "Already Exists" errors.
- **SELECT ***: This command pulls every column and row from your raw SalesforceData table into the new view.
- **The Abstraction Strategy**: We are taking the raw table and giving it a professional alias (Salesforce_Data). From this point forward, every chart, AI model, and DAX measure will refer to this **View**, not the raw table. This creates a "safety buffer" for your architecture.

Gemini AI Strategy Tip

Once you have successfully run the code, you can use the Gemini sidecar in BigQuery to audit your data instantly.

Try this Prompt:

"Look at my Salesforce_Data view. Show me the top 10 rows to confirm the data loaded correctly, and tell me which campaign has the highest total donation amount so far."

2. Creating the donor_lkup View

In a real-world CRM like Salesforce, a single donor may appear dozens of times—once for every donation they have ever made. If we used this "flat" list for our analysis, we would accidentally over-count our supporters.

This step creates a **Dimension Table** (a unique master list). We use a professional "Window Function" to de-duplicate the data, ensuring we only keep the **most recent** contact details for every individual donor.

Detailed Instructions

1. **Open BigQuery:** Ensure you are in the SQL Workspace.
2. **Open a New Tab:** Click the **+** button to start a new query.
3. **Update the Project ID:** Just as before, replace mlprojecttraining-arkbi2025 with your unique Project ID.
4. **Run the Script:** Copy the block below, paste it into the editor, and click **Run**.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
This script removes duplicates by picking the
LATEST record for each donor.
*/

CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_lkup` AS
SELECT
  donor_id,
  full_name,
  email,
  postal_code,
  country
FROM (
  SELECT
    sfd.donor_id,
    sfd.full_name,
    sfd.email,
    sfd.postal_code,
    sfd.country,
    ROW_NUMBER() OVER (PARTITION BY sfd.donor_id ORDER BY sfd.donation_date
DESC) as rn
  FROM
    `mlprojecttraining-arkbi2025.DonorAnalysis.Salesforce_Data` AS sfd
)
WHERE rn = 1;
```

Code Description

- **ROW_NUMBER() OVER**: This is a powerful analytical function. It assigns a rank (1, 2, 3...) to every row in your data.
- **PARTITION BY donor_id**: This tells BigQuery to "group" the rows by each specific person.
- **ORDER BY donation_date DESC**: Within those groups, we sort by date, putting the newest donation at the very top (Rank #1).
- **WHERE rn = 1**: Finally, we throw away all the older records (Rank 2 and below), leaving us with a clean, de-duplicated master list of donors.

Gemini AI Strategy Tip

Once the code has finished running, you can use Gemini to expand the metadata of your donors.

Try this Prompt:

"How can I modify my donor_lkup view to also include a new column called 'total_donations_count' that shows the total number of times each donor has given in the past?"

3. Creating the campaign_lkup and channel_lkup Views

To build a professional **Star Schema** in Power BI, we need "Lookup" tables that contain every unique category in our data. In this step, we isolate every unique **Campaign** (e.g., Annual Appeal, Winter Gala) and every unique **Channel** (e.g., Email, Social Media, Direct Mail).

These views allow the Executive to filter the entire dashboard by a specific marketing effort or platform without any data being "double-counted."

Detailed Instructions

1. **Open BigQuery:** Ensure you have the SQL Workspace active.
2. **Open a New Tab:** Click the **+** button for a fresh editor.
3. **Update the Project ID:** Replace mlprojecttraining-arkbi2025 with your actual Project ID in both scripts below.
4. **Run the Scripts:** You can copy both blocks and run them together in one window.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
*/

-- 1. Create the campaign_lkup view
CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.campaign_lkup` AS
SELECT
    DISTINCT campaign
FROM
    `mlprojecttraining-arkbi2025.DonorAnalysis.Salesforce_Data`;

-- 2. Create the channel_lkup view
CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.channel_lkup` AS
SELECT
    DISTINCT channel
FROM
    `mlprojecttraining-arkbi2025.DonorAnalysis.Salesforce_Data`;
```

Code Description

- **SELECT DISTINCT**: This is the key command. It scans the thousands of rows in your Salesforce data and returns only one instance of each name it finds.
- **Dimensional Integrity**: By creating these as separate views, we ensure that when you select "Email" in your final dashboard, Power BI knows exactly which records to display across every chart on the page.

Gemini AI Strategy Tip

Now that you have these views, you can use Gemini to perform a "Quick Audit" of your marketing performance directly in BigQuery.

Try this Prompt:

"Using my campaign_lkup view joined back to Salesforce_Data, show me the total donation amount and the total number of donors for each campaign. Sort the results so the most profitable campaign is at the top."

4. Creating the calendar_lkup View

In high-performance analytics, a **Universal Calendar** is non-negotiable. Without it, your charts will have "gaps" on days where no donations were received. This script "manufactures" a continuous timeline, ensuring your donor retention and growth trends are mathematically perfect and audit-proof.

Detailed Instructions

1. **Open BigQuery:** Ensure you are in your SQL Workspace.
2. **Open a New Tab:** Click the **+** button.
3. **Update the Project ID:** Replace mlprojecttraining-arkbi2025 with your unique Project ID.
4. **Run the Script:** Copy the block below, paste it into the editor, and click **Run**.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
This script creates an unbroken 6-year timeline
for accurate time-series forecasting.
*/

CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.calendar_lkup` AS
SELECT
  date,
  FORMAT_DATE('%b %Y', date) AS Month_Year,
  CAST(FORMAT_DATE('%Y%m', date) AS INT64) AS Month_Year_No,
  EXTRACT(YEAR FROM date) AS Calendar_Year,
  CONCAT('Q', CAST(CEIL(EXTRACT(MONTH FROM date) / 3) AS STRING), ' ',
EXTRACT(YEAR FROM date)) AS Quarter_Year,
  (EXTRACT(YEAR FROM date) * 100) + (CAST(CEIL(EXTRACT(MONTH FROM date) /
3) AS INT64)) AS Quarter_Year_No
FROM
  UNNEST(GENERATE_DATE_ARRAY('2020-01-01', '2025-12-31', INTERVAL '1' DAY))
AS date;
```

Code Description

- **GENERATE_DATE_ARRAY**: This is the engine of the script. It creates a list containing every single day between January 1st, 2020, and December 31st, 2025.
- **UNNEST**: This transforms that list of dates into a column that BigQuery can process as a table.
- **FORMAT_DATE & EXTRACT**: These functions create the "attributes" you need for your dashboard, such as **Month_Year** (e.g., Jan 2024) and **Calendar_Year**.
- **Sorting Logic**: The columns ending in _**No** (like Month_Year_No) are created so that your charts sort chronologically (January, February, March) rather than alphabetically.

Gemini AI Strategy Tip

You can use Gemini to customise this calendar for specific organisational needs, such as adding a "Day of the Week" column to see if donors give more on weekends.

Try this Prompt:

"How can I modify the calendar_lkup view to include a new column for the day of the week, formatted as a full name (e.g., 'Monday'), and a boolean column called 'is_weekend' that is true for Saturdays and Sundays?"

Integrating Gemini AI with BigQuery: Your AI Thought Partner

BigQuery Studio now integrates Gemini AI directly into your workflow. This isn't just a "search bar"—it is a co-pilot that helps you write, explain, and debug the very SQL and Machine Learning queries you are building in this course. By using natural language prompts, you can reclaim hours of development time.

Core Functionality

- **SQL Generation:** Describe what you want in plain English (e.g., *"Join my donors to my donations and group by month"*), and Gemini will write the syntax for you.
- **Code Explanation:** If you are looking at a complex window function like ROW_NUMBER() OVER (...) and need a refresher, highlight the code and ask Gemini to explain it.
- **Code Refinement:** Request instant modifications, such as: *"Add a filter to only show donors from the 'SO1' postal code."*
- **BigQuery ML Support:** Ask Gemini to generate the correct CREATE MODEL syntax for **Linear Regression** or **Boosted Trees**, including the required options.

Usage Example: Engineering New Insights

You can use Gemini to generate new views or features beyond what is in this manual. For instance, to dive deeper into campaign performance, you could use this prompt in the BigQuery editor:

Prompt: *"Write a BigQuery SQL query that joins the 'donor_features_view' and the 'campaign_lkup' view to show the average donation amount for all donors who participated in the 'Annual Appeal 2023' campaign."*

⚠️ **Important: The Stewardship Standard for AI**

While Gemini is highly capable, it is an **assistive tool**, not a replacement for human oversight. To maintain "Audit-Proof" accuracy, follow these guidelines:

1. **Review the Syntax:** AI can occasionally introduce small errors in highly complex or niche edge cases.
2. **Verify Context:** Ensure Gemini hasn't misinterpreted a column name (e.g., using donation_val instead of your actual column donation_amount).
3. **Check for Efficiency:** For very large datasets, ensure the AI isn't using "sub-optimal" SQL that could increase your processing costs.

Stewardship Recommendation: Always perform a quick "Sanity Check" on AI-generated code. Use it to build the **draft**, but use your **expert eyes** to verify the project IDs and table names before clicking "Run."

Part 2: BigQuery ML SQL

1. Creating the donor_features_view (RFM Analysis)

To move from simple reporting to **AI-driven forecasting**, we must first perform **Feature Engineering**. This script creates the "signals" that the Machine Learning model needs to learn. We focus on **RFM** (Recency, Frequency, Monetary) metrics, which are the gold standard for predicting donor behavior.

This view collapses thousands of individual donations into a single, high-intelligence row for every donor.

Detailed Instructions

1. **Open BigQuery:** Ensure you are in your SQL Workspace.
2. **Open a New Tab:** Click the **+** button.
3. **Update the Project ID:** Replace mlprojecttraining-arkbi2025 with your unique Project ID in **three** places within the script.
4. **Run the Script:** Copy the block below, paste it into the editor, and click **Run**.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
This script performs Feature Engineering to create
the input signals for our AI Model.
*/

CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_features_view` AS
SELECT
    donor_id,
    MAX(donation_date) AS Last_Donation_Date,
    DATE_DIFF(DATE '2024-12-31', MAX(donation_date), DAY) AS
Days_Since_Last_Donation,
    COUNT(*) AS Donation_Frequency,
    AVG(donation_amount) AS Average_Donation_Amount,
    SUM(donation_amount) AS Total_Donation_Amount,
    ARRAY_AGG(donation_amount ORDER BY donation_date DESC LIMIT
1)[OFFSET(0)] AS Last_Donation_Amount,
    MAX(CASE WHEN EXTRACT(YEAR FROM donation_date) = 2024 THEN 1 ELSE 0
END) AS Donated_in_2024,
    AVG(annual_donation_count) AS Annual_Donation_Frequency
FROM
    `mlprojecttraining-arkbi2025.DonorAnalysis.Salesforce_Data` AS sd
LEFT JOIN
    (SELECT
        donor_id AS did,
        EXTRACT(YEAR FROM donation_date) AS donation_year,
        COUNT(*) AS annual_donation_count
     FROM
        `mlprojecttraining-arkbi2025.DonorAnalysis.Salesforce_Data`
     WHERE
        donation_date BETWEEN '2021-01-01' AND '2024-12-31'
     GROUP BY
        did, donation_year) AS annual_counts
ON
    sd.donor_id = annual_counts.did
    AND EXTRACT(YEAR FROM sd.donation_date) = annual_counts.donation_year
WHERE
    sd.donation_date BETWEEN '2021-01-01' AND '2024-12-31'
GROUP BY
    sd.donor_id;
```

Code Description

- **Recency (Days_Since_Last_Donation)**: Calculates how many days have passed since the donor's last gift. Generally, the more recent the gift, the more likely they are to give again.
- **Frequency (Donation_Frequency)**: Counts the total number of times they have supported the organization.
- **Monetary (Total_Donation_Amount)**: Aggregates the lifetime value of the donor.
- **The Target Variable (Donated_in_2024)**: This is a binary "Flag" (1 or 0). This tells the AI what happened in 2024 so it can learn the patterns of who stayed and who lapsed.
- **Subquery & Join**: We calculate the Annual_Donation_Frequency to understand the consistency of the donor's support over several years, smoothing out one-off anomalies.

Gemini AI Strategy Tip

You can use Gemini to deepen your feature engineering by adding growth-based metrics that track donor "momentum."

Try this Prompt:

"In the donor_features_view, how can I calculate a new feature called 'average_annual_growth' that measures the percentage change in Total_Donation_Amount from the first two years (2021-2022) to the last two years (2023-2024)?"

2. Training the Linear Regression Model

With our features engineered, we now move into the most transformative phase: **Activating the AI Brain**. In this step, we use BigQuery ML to train a **Linear Regression** model. Instead of just looking at historical totals, the model will analyze the complex relationships between Recency, Frequency, and Monetary value to calculate a **Propensity Score**.

This score represents a continuous estimate of how likely a donor is to give again, allowing you to move from "Rearview Mirror" reporting to **Strategic Foresight**.

Detailed Instructions

1. **Open BigQuery:** Ensure you are in your SQL Workspace.
2. **Compose New Query:** Click the **+** button.
3. **Update the Project ID:** Replace mlprojecttraining-arkbi2025 with your unique Project ID in **two** places within the script.
4. **Run the Script:** Copy the block below, paste it into the editor, and click **Run**. This may take a minute as BigQuery processes the machine learning iterations.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
This script trains the AI to find patterns between
past behavior and future donation likelihood.
*/

CREATE OR REPLACE MODEL `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_propensity_model`
OPTIONS(
  model_type = 'LINEAR_REG',
  input_label_cols = ['Donated_in_2024']
) AS
SELECT
  Days_Since_Last_Donation,
  Donation_Frequency,
  Average_Donation_Amount,
  Total_Donation_Amount,
  Last_Donation_Amount,
  Donated_in_2024,
  Annual_Donation_Frequency
FROM
  `mlprojecttraining-arkbi2025.DonorAnalysis.donor_features_view`;
```

Code Description

- **CREATE OR REPLACE MODEL**: The standard command to build your AI model within your DonorAnalysis dataset.
- **model_type = 'LINEAR_REG'**: This specifies **Linear Regression**. In this context, it is used to predict a "Propensity Score"—a number typically between 0 and 1 that indicates donation probability.
- **input_label_cols**: This identifies **Donated_in_2024** as the "Answer Key." The model looks at this column to learn which combinations of features (like high frequency or recent gifts) resulted in a successful donation.
- **SELECT ... FROM donor_features_view**: This feeds the "signals" we created in the previous step into the model's training engine.

Gemini AI Strategy Tip

Linear Regression is excellent for continuous scores, but for "Yes/No" binary classifications, you might want to explore **Logistic Regression**.

Try this Prompt:

"How can I change the CREATE MODEL script to train a logistic regression model instead of a linear regression model? Also, what is the key difference between these two model types in the context of donor churn?"

3. Evaluating the Model

Before we trust our AI to predict the future, we must audit its performance. In this step, we use the ML.EVALUATE function to see how well the model "learned" the patterns in your donor data. This provides the **Audit-Proof Certainty** required to present these findings to your Board or Executive Team.

Detailed Instructions

1. **Open BigQuery:** Go to your SQL Workspace.
2. **Compose New Query:** Click the **+** button.
3. **Update the Project ID:** Replace mlprojecttraining-arkbi2025 with your unique Project ID in **two** places.
4. **Run the Script:** Copy the block below and click **Run**. You will receive a single row of data containing several statistical scores.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
This script audits the AI to ensure its predictions
are statistically significant.
*/

SELECT
  *
FROM
  ML.EVALUATE(MODEL `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_propensity_model`,
  (
    SELECT
      Days_Since_Last_Donation,
      Donation_Frequency,
      Average_Donation_Amount,
      Total_Donation_Amount,
      Last_Donation_Amount,
      Donated_in_2024,
      Annual_Donation_Frequency
    FROM
      `mlprojecttraining-arkbi2025.DonorAnalysis.donor_features_view`
  ));
```

Code Description

- **ML.EVALUATE**: This function compares the values the model *predicted* against the *actual* historical outcomes (the "Ground Truth").
- **r2_score (R-Squared)**: This is your "Hero Metric." It ranges from 0 to 1. A score of 0.778, for example, means your model explains roughly 78% of donor behavior variability. The higher this number, the more confident you can be in your 2025 forecast.
- **mean_absolute_error (MAE)**: This tells you the average "miss" of the model. In a propensity model (0 to 1), a low MAE indicates high precision in predicting the likelihood of a gift.

Gemini AI Strategy Tip

Raw statistical scores can be difficult to explain in a boardroom. You can use Gemini to translate these numbers into a compelling "Stewardship Narrative."

Try this Prompt:

"Explain what an R-squared score of 0.778 means for my donor_propensity_model in simple, non-technical terms that I can use in an executive summary for my Board of Trustees."

Step 3.1: Decoding the Science—Model Evaluation Results

Once you run the ML.EVALUATE script, BigQuery will return a table of statistical markers. These aren't just "math"—they are the **Confidence Scorecard** of your AI brain. In the world of **Predictive Stewardship**, these numbers prove that your 2025 strategy is built on a solid foundation of evidence, not guesswork.

The Model Performance Scorecard

Model Performance Metric	Value	Stewardship Interpretation
r2_score (R-Squared)	**0.778**	**The Hero Metric.** 77.8% of donor behaviour is explained by this model.
mean_absolute_error (MAE)	**0.189**	On average, the AI's "guess" is within 0.18 units of the actual outcome.
explained_variance	**0.778**	Confirms the R-Squared; the model successfully captures the "why" behind the data.
median_absolute_error	**0.158**	50% of predictions have an error even lower than the average (0.15).

What These Numbers Mean for Your Strategy

1. High R-Squared (0.778): The "Fit"

An R-squared value of approximately **0.778** indicates that about **77.8%** of the variability in whether donors donated in 2024 can be accounted for by the patterns captured in our model. In the non-profit sector, any score above **0.7** is considered a **"Strong Fit."** This gives you high confidence that the "signals" we engineered (Recency, Frequency, and Monetary value) are the correct drivers of your mission's growth.

2. Error Metrics (MAE & MedAE): The "Precision"

- **MAE (0.189):** This measures the average magnitude of the errors. Since our target is binary (0 for no gift, 1 for a gift), an error of only 0.189 means the model is exceptionally close to the **"Ground Truth."**
- **Median Absolute Error (0.159):** Because the Median error is lower than the Mean (MAE), we know our model isn't being "fooled" or skewed by a few unusual "outlier" donors (e.g., a one-time massive legacy gift). It is performing consistently and reliably across your entire database.

Strategic Insight for Leadership

When presenting these results to your Board or Executive Team, focus on the **Explained Variance**. You can confidently state:

"Our AI model has successfully identified the underlying patterns of 78% of our donor base. This allows us to move away from reactive fundraising and toward a proactive, predictive engagement strategy for 2025."

Step 4: Generating the 2025 Predictions

Now that we have **Statistically Validated** our model, we can apply it to the future. This script creates the final view that identifies exactly who is likely to support **Second Chance Futures** in 2025. This is the moment where "Data Chaos" officially becomes **Actionable Foresight**.

Detailed Instructions

1. **Open BigQuery:** Ensure you are in your SQL Workspace.
2. **Update Project ID:** Replace the placeholder mlprojecttraining-arkbi2025 with your unique Project ID in **three** places within the script.
3. **Thresholding:** Note that we have set a threshold of **0.6**. This means we only label a donor as "will donate" if the AI is 60% sure or higher. This "Confidence Gate" ensures your marketing resources are spent on the highest-probability leads.

SQL

```
/* INSTRUCTIONS:
Replace 'mlprojecttraining-arkbi2025' with your
unique Google Cloud Project ID.
*/

CREATE OR REPLACE VIEW `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_predictions_2025` AS
SELECT
    dfv.donor_id,
    p.predicted_Donated_in_2024 AS predicted_donation_propensity,
    CASE
        WHEN p.predicted_Donated_in_2024 > 0.6 THEN 'will donate in 2025'
        ELSE 'will not donate in 2025'
    END AS donation_prediction_2025,
    dfv.Last_Donation_Amount,
    dfv.Days_Since_Last_Donation
FROM
    ML.PREDICT(MODEL `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_propensity_model`,
        (SELECT
            donor_id,
            Days_Since_Last_Donation,
            Donation_Frequency,
            Average_Donation_Amount,
            Total_Donation_Amount,
            Last_Donation_Amount,
            Donated_in_2024,
            Annual_Donation_Frequency
        FROM `mlprojecttraining-
arkbi2025.DonorAnalysis.donor_features_view`)) AS p
JOIN
    `mlprojecttraining-arkbi2025.DonorAnalysis.donor_features_view` AS dfv
    ON p.donor_id = dfv.donor_id;
```

Important: Interpreting the "Propensity Score"

Unlike other models, Linear Regression is looking for a "best fit" line. This leads to two unique scenarios you may see in your results that actually provide deeper insight into donor loyalty:

- **Scores > 1.0 (e.g., 1.09):** The AI sees such a strong historical pattern (very recent and very frequent) that it "extrapolates" beyond the binary. Treat these as your **"Super Supporters"**—they are the bedrock of your mission.
- **Scores < 0 (e.g., -0.04):** The AI identifies patterns that are the mathematical opposite of a donor (e.g., they haven't given in several years). These are your **"High-Churn Risks"** or lapsed donors.

Next Steps for Stewardship

Once this view is created, your data is no longer static; it is a tactical roadmap. You can now:

1. **Rank Donors:** Instantly identify the top 100 people most likely to support your mission this year.
2. **Segment Marketing:** Create a "Middle-of-the-Road" segment for donors with scores between **0.4 and 0.6**. These individuals are on the fence; a personalized re-engagement campaign can push them "over the line" into a 2025 contribution.
3. **Financial Planning:** Use the Last_Donation_Amount for the "will donate" group to calculate a realistic baseline budget for the coming year.

Part 4: The Interface Handshake (Power BI Desktop)

Now that the "Brain" (BigQuery) has processed the data and generated the AI forecasts, it is time to move to the "Interface." In this phase, you will import your BigQuery views into Power BI, architect the relational skeleton, and use **DAX (Data Analysis Expressions)** to create dynamic calculations. Unlike standard spreadsheet formulas, these measures are "Context-Aware," meaning they will instantly recalculate whenever you filter your dashboard by medic, date, or procedure.

Step 1: Data Ingestion & Transition

Before building the intelligence, we must bring our "Single Source of Truth" views from the cloud into our local Power BI workspace.

1. **Launch Power BI Desktop:** Open the application on your computer. If you are using the course template, ensure it is open.
2. **Connect to Data:** On the Home ribbon, click **Get Data > Google BigQuery**.
3. **Sign In:** Click "Sign in" and use your dedicated Gmail credentials to authenticate your access to the Google Cloud Project.
4. **Select Your Views:** In the Navigator window, locate your **mlprojecttraining-arkbi2025** project and the **DonorAnalysis** dataset. Check the boxes for the following views:
    - `Salesforce_Data` (Your Fact Table)
    - `donor_lkup`
    - `campaign_lkup`
    - `channel_lkup`
    - `calendar_lkup`
    - `donor_predictions_2025` (Your AI Output)
5. **Load Data:** Click **Load**. Power BI will now establish a secure link to your BigQuery data.
6. **Verification:** Click on the **Table View** icon (the grid icon) on the far-left sidebar. Click through each table to confirm the data has successfully landed and that your AI propensity scores are visible.

Step 2: Building the Relational Skeleton (Power BI Relationships)

The Strategy: Establishing the "Star Schema"

In this phase, we connect our **Lookup Tables (Dimensions)** to our **Fact Tables**. This architectural structure is the secret to a professional dashboard. It ensures that a single selection in a "Lookup" (e.g., selecting a specific marketing **Channel**) filters every single chart and AI prediction simultaneously.

Step-by-Step Relationship Mapping

1. **Enter Model View:** On the far-left sidebar of Power BI Desktop, click the **Model View** icon (the icon with three small boxes).
2. **Arrange Your Canvas:** For a professional, scannable layout, place your **Lookup Tables** (calendar_lkup, donor_lkup, campaign_lkup, channel_lkup) along the top. Place your **Fact Tables** (Salesforce_Data, donor_predictions_2025, donor_features_view) along the bottom.
3. **Create the Link:** To create a relationship, click and hold the column name in the Lookup Table and drag it on top of the matching column in the Fact Table.

The Relationship Master Table

Configure the following connections. Verify that each relationship shows a **"1"** on the Lookup side and a **"*" (Many)** on the Fact side.

From: Lookup Table (The "1" Side)	To: Fact Table (The "Many" Side)	Column Join Link
calendar_lkup	Salesforce_Data	date → donation_date
donor_lkup	Salesforce_Data	donor_id → donor_id
campaign_lkup	Salesforce_Data	campaign → campaign
channel_lkup	Salesforce_Data	channel → channel
donor_lkup	donor_predictions_2025	donor_id → donor_id
donor_lkup	donor_features_view	donor_id → donor_id

Verifying the "Stewardship Standards"

Double-click any relationship line to ensure the following properties are set correctly. This ensures your "Context-Aware" calculations are accurate:

- **Cardinality:** One-to-many (1:*).
- **Cross-filter direction:** Single.
- **Make this relationship active:** Checked.

⚠️ Architect's Warning

If you see a **"Many-to-Many"** warning, it means your BigQuery Lookup views contain duplicate values. If this happens, return to BigQuery and ensure your DISTINCT keyword or ROW_NUMBER() logic was used correctly. A **"Predictive Stewardship Architecture"** requires a clean *1: link** to ensure your donor counts and financial projections never double-count.

Part 4: Power BI DAX Measures

The Strategy: Building the Intelligence Layer

Now that your "Skeleton" (the Data Model) is built, we need to add the "Muscles"—the **DAX Measures**. Unlike standard columns, DAX measures are dynamic; they only calculate when you interact with the dashboard. This keeps your report fast and responsive, even with large datasets.

Step-by-Step: How to Add Measures in Power BI

To keep your workspace organized and follow the **Stewardship Standard**, we will store all our calculations in a dedicated "Measure Table."

1. **Create a Measure Table (Optional but Recommended):** On the **Home** tab, click **Enter Data**. Name the table _DAX Measures and click **Load**.
2. **Create a New Measure:** Right-click your _DAX Measures table (or any table in the Data pane) and select **New measure**.
3. **Enter the Formula:** The **Formula Bar** will appear at the top. Copy and paste the code snippet from the table below.
4. **Commit the Code:** Press **Enter** or click the checkmark icon next to the formula bar.
5. **Format Your Measure:** While the measure is selected, go to the **Measure tools** tab at the top. Set the format (e.g., "Currency" for Donations or "Whole Number" with a comma for Donors).

1. Foundational Metrics

These metrics provide the core volume and behavior baseline for your fundraising analysis.

Code Block Name	Code Type	Code Snippet
Donation	DAX	Donation = SUM('Salesforce_Data'[donation_amount])
Donors	DAX	Donors = DISTINCTCOUNT('Salesforce_Data'[donor_id])
Donation Frequency per Donor	DAX	Donation Frequency per Donor = DIVIDE(COUNTROWS('Salesforce_Data'), DISTINCTCOUNT('Salesforce_Data'[donor_id]), 0)
Avg Donation per Donor	DAX	Avg Donation per Donor = AVERAGEX(DISTINCT(Salesforce_Data[donor_id]), CALCULATE(AVERAGE(Salesforce_Data[donation_amount])))

Narrative Logic & Description

- **Donation**: Uses the SUM function to calculate the total financial footprint. **Format this as Currency (£).**
- **Donors**: Uses DISTINCTCOUNT. This is critical for **Predictive Stewardship** because it ensures that even if a supporter gives ten times, they are only counted as one unique donor.
- **Donation Frequency per Donor**: Uses the DIVIDE function. We prefer DIVIDE(a, b, 0) over a standard / because it includes an "alternate result" (0). This prevents your dashboard from showing errors (DIV/0) on days with no data.
- **Avg Donation per Donor**: This uses an **Iterator** (AVERAGEX). It first calculates the average for each individual donor and then averages those results. This ensures that high-value "Super Supporters" are treated with the same weight as smaller donors when looking at behavioral averages.

2. Time Intelligence Metrics

The Strategy: Measuring Growth and Momentum

To prove that your fundraising strategies are working, you must move beyond "Static Totals" and show **Trend Analysis**. Time Intelligence DAX allows you to compare current performance against the same period last year (**Year-over-Year**). This removes seasonal bias (e.g., comparing December's high donations to January's low donations) and provides a "Like-for-Like" audit of your progress.

Step-by-Step: Implementing Time Intelligence

1. **Right-click** your _DAX Measures table and select **New measure**.
2. **Copy and paste** the formulas from the table below.
3. **Format the Growth %:** For the "% YoY Change" measures, ensure you go to the **Measure tools** tab and set the format to **Percentage (%)** with one or two decimal places.

Code Block Name	Code Type	Code Snippet
Prior Year Donation	DAX	Prior Year Donation = CALCULATE([Donation], DATEADD('calendar_lkup'[date], -1, YEAR))
% YoY Change in Donation	DAX	% YoY Change in Donation = VAR PY_Donation = [Prior Year Donation] RETURN DIVIDE([Donation] - PY_Donation, PY_Donation, 0)
Prior Year Donors	DAX	Prior Year Donors = CALCULATE([Donors], DATEADD('calendar_lkup'[date], -1, YEAR))
% YoY Change in Donors	DAX	% YoY Change in Donors = VAR PY_Donors = [Prior Year Donors] RETURN DIVIDE([Donors] - PY_Donors, PY_Donors, 0)

Narrative Logic & Description

- **CALCULATE and DATEADD**: This is the "Time Machine" of DAX. CALCULATE tells Power BI to ignore the current date on your chart and instead use the date provided by DATEADD (which is exactly 365 days in the past).
- **The VAR/RETURN Pattern**: We use **Variables (VAR)** to store the "Prior Year" result. This makes the code faster because Power BI only has to calculate the past value once, rather than multiple times within the same formula.
- **DIVIDE for Growth**: To find the percentage change, we take **(Current - Past) / Past**. By using the DIVIDE function, we ensure that if there were zero donations last year, the formula returns a 0% rather than an "Infinity" error.

3. Retention Metric

The Strategy: The "Stewardship Indicator"

Donor Retention is the ultimate health check for a non-profit. While acquiring new donors is vital, keeping existing ones is far more cost-effective. This measure identifies how many supporters who gave last year have returned to give again this year. A high retention rate proves that your **Donor Stewardship** and engagement strategies are resonating with your community.

Step-by-Step: Implementing the Retention Rate

1. **Right-click** your _DAX Measures table and select **New measure**.
2. **Paste the formula** from the block below into the formula bar.
3. **Set Formatting:** This is a percentage metric. Select the measure, go to the **Measure tools** tab, and set the format to **Percentage (%)** with 1 decimal place.

Code Block Name	Code Type	Code Snippet
Donor Retention	DAX	Donor Retention = VAR CurrentPeriodDonors = CALCULATE(DISTINCTCOUNT(Salesforce_Data[donor_id])) VAR PreviousPeriodDonors = CALCULATE(DISTINCTCOUNT(Salesforce_Data[donor_id]), DATEADD('calendar_lkup'[date], -1, YEAR)) VAR RetainedDonors = CALCULATE(DISTINCTCOUNT(Salesforce_Data[donor_id]), FILTER(VALUES(Salesforce_Data[donor_id]), CALCULATE(NOT ISBLANK([Donation]), DATEADD('calendar_lkup'[date], -1, YEAR)))) RETURN DIVIDE(RetainedDonors, PreviousPeriodDonors, 0)

Narrative Logic & Description

This is an "Advanced" measure that uses multiple variables to perform a complex set of operations:

- **CurrentPeriodDonors**: Captures everyone who gave in the current timeframe (e.g., 2024).
- **PreviousPeriodDonors**: Uses the DATEADD "Time Machine" to look back at everyone who gave in the prior year (e.g., 2023).
- **RetainedDonors**: This is the "Brain" of the formula. It uses FILTER and VALUES to look at the list of donors today and cross-references it with the list from last year. It only counts the donor if they appear in **both** lists.
- **DIVIDE**: Finally, it takes the number of "Returned Supporters" and divides them by the "Total Supporters from Last Year" to give you your percentage.

4. Predictive Metrics

The Strategy: Turning AI into Actionable Foresight

The power of your BigQuery ML model is finally realized here. These measures translate the raw "Propensity Scores" into **Boardroom Strategy**. Instead of guessing how much money you *might* raise, you are now quantifying the specific revenue potential and donor volume for 2025 based on high-probability signals.

Step-by-Step: Implementing Predictive DAX

1. **Right-click** your _DAX Measures table and select **New measure**.
2. **Paste the formula** from the table below into the formula bar.
3. **Format the Output:**
    - **Likely Donation (2025):** Format as **Currency (£)** with 0 decimal places.
    - **Likely Donors (2025):** Format as **Whole Number** with a comma separator.
    - **Avg Predicted Propensity:** Format as a **Decimal** (2 places) or **Percentage**.

Code Block Name	Code Type	Code Snippet
Likely Donors (2025)	DAX	Likely Donors (2025) = CALCULATE(DISTINCTCOUNT(donor_predictions_2025[donor_id]), FILTER(donor_predictions_2025, donor_predictions_2025[donation_prediction_2025] = "will donate in 2025"))
Likely Donation (2025)	DAX	Likely Donation (2025) = CALCULATE(SUM(donor_predictions_2025[Last_Donation_Amount]) , FILTER(donor_predictions_2025, donor_predictions_2025[donation_prediction_2025] = "will donate in 2025"))
Avg Predicted Propensity (2025)	DAX	Average Predicted Propensity (2025) = AVERAGE(donor_predictions_2025[predicted_donation_propensity])
Avg Days Since Last Donation	DAX	Avg Days Since Last Donation = AVERAGE(donor_predictions_2025[Days_Since_Last_Donation])

Narrative Logic & Description

- **Likely Donors (2025)**: This filters your entire database down to only those individuals the AI flagged with a score above 0.6. It provides a concrete target list for your fundraising team.
- **Likely Donation (2025)**: This is a "Baseline Revenue" projection. It takes the donors who are likely to give and sums their **Last_Donation_Amount**. This assumes that a retained donor will give at least what they gave last time—a conservative and audit-proof way to forecast income.
- **Average Predicted Propensity**: This calculates the "Mean Likelihood." It's an excellent barometer for the overall health of a specific campaign or channel.
- **Avg Days Since Last Donation**: This is a "Recency Signal." Monitoring this allows you to see if your 2025 "likely" pool is becoming more or less disengaged over time.

Part 5: Building the Interface (Power BI Report Visuals)

This final stage transforms your BigQuery AI models and DAX intelligence into a high-impact visual narrative. We will build a two-page report: an **Executive Summary** for strategic oversight and a **Detailed Page** for tactical donor engagement.

Page 1: Executive Summary

This page provides a high-level overview of donor health, financial trends, and AI-driven 2025 predictions.

1. Setting up Key Performance Indicator (KPI) Cards

- **Objective:** To display primary donor-related metrics prominently.
- **Steps:**
    1. From the **Visualizations** pane, select the **Card** visual icon.
    2. Drag and drop the **Donation** measure from the Data pane into the **Field** well.
    3. **Format the Card:** With the card selected, go to the **Format your visual** pane. Under **Callout value**, set the **Text size** to **20**.
    4. **Duplicate Cards:** Select the card, press **Ctrl+C**, then **Ctrl+V** (Repeat 6 more times until you have 8 cards).
    5. **Assign Measures:** For each duplicated card, replace the field with:
        - **Donors**
        - **Prior Year Donation**
        - **Prior Year Donors**
        - **% YoY Change in Donation**
        - **% YoY Change in Donors**
        - **Likely Donation (2025)**
        - **Likely Donors (2025)**
    6. **Arrange Systematically:** Organize them in two rows. Place the four **Donation** related cards in the top row and the four **Donor** related cards directly below them.

2. Adding the Donor Retention Line Chart

- **Objective:** To visualize the trend of donor retention over time.
- **Steps:**
    1. Select the **Line chart** visual icon. Position it to the right of the KPI cards.
    2. Drag **Donor Retention** to the **Y-axis**.
    3. Drag **Calendar_Year** from calendar_lkup to the **X-axis**.
    4. **Format:** In **Format your visual**, turn **Data labels** to **On** to show the exact percentage on the line.

3. Adding the Calendar Year Slicer

- **Objective:** To allow users to filter the entire report by year.
- **Steps:**
    1. Select the **Slicer** icon. Position it in the top-right corner.
    2. Drag **Calendar_Year** into the **Field** well.
    3. **Format:** Under **Slicer settings > Options**, change the style to **Vertical list**.

4. Background Styling for Top Section

- **Objective:** To create a visual "Header" for the executive metrics.
- **Steps:**
    1. Go to **Insert > Shapes > Rectangle**.
    2. Resize it to cover the KPI cards, Line chart, and Slicer.
    3. In the **Format** pane: Set **Fill color** to a brown shade and **Transparency** to **75%**.
    4. Right-click the shape > **Order > Send to back**.

5. Adding the Area Chart: Donation and Donors Over Time

- **Objective:** To compare financial growth against donor volume on a single timeline.
- **Steps:**
    1. Select the **Area chart** icon. Position it below the KPI section.
    2. Drag **Donation** to the **Y-axis** (Primary).
    3. Drag **Donors** to the **Secondary Y-axis**.
    4. Drag **Calendar_Year** to the **X-axis**.
    5. Drag **Quarter_Year** into the **X-axis** (below Year) to enable the drill-down hierarchy.

6. Adding the Scatter Chart: Propensity vs. Donor Count

- **Objective:** To identify which campaigns are likely to yield the highest 2025 return.
- **Steps:**
    1. Select the **Scatter chart** icon. Position it to the right of the Area chart.
    2. **X-axis:** Drag **Average Predicted Propensity (2025)**.
    3. **Y-axis:** Drag **Donors**.
    4. **Values/Details:** Drag **campaign** from campaign_lkup.
    5. **Legend:** Drag **donation_prediction_2025** from donor_predictions_2025.
    6. **Size:** Drag **Donation**.
- **Interpretation:** Campaigns in the top-right are your "Gold Standard" (High volume, High propensity). Large bubbles represent high historical value.

7. Background Styling for Bottom Section

- **Objective:** Consistency in design.
- **Steps:** Repeat the Rectangle shape process to frame the Area and Scatter charts using the same **75% transparent brown** style.

Page 2: Detailed Page

This page allows fundraisers to find specific donors and analyze performance at the campaign level.

1. Adding the Matrix Visual

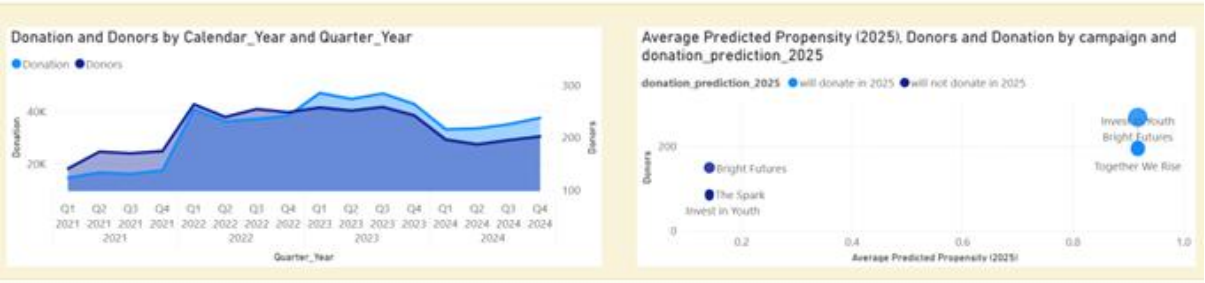
- **Objective:** To provide a drill-down list from Channel to Individual Donor.
- **Steps:**
    1. Select the **Matrix** icon. Expand it to fill the majority of the page.
    2. **Rows (Hierarchy):** Drag **channel**, then **campaign**, then **full_name** into the Rows well.
    3. **Values:** Drag **Donors**, **Donation**, **Avg Donation per Donor**, and **Donation Frequency per Donor**.

2. Adding Slicers for Filtering

- **Objective:** Interactive filtering.
- **Steps:**
    1. **Calendar Year Slicer:** Drag **Calendar_Year** into a new Slicer.
    2. **Campaign Slicer:** Drag **campaign** from campaign_lkup into a new Slicer.
    3. **Channel Slicer:** Drag **channel** from channel_lkup into a new Slicer.
    4. Format all as **Vertical lists** and align them to the right of the matrix.

Executive Narrative: Operational Stewardship & 2025 Growth Strategy

This section translates your technical build into a strategic briefing. As the **Chief Operating Officer (COO)**, this is the "Story" you deliver to the Board and the Executive Team. It moves beyond spreadsheets and into the realm of **Operational Foresight**, ensuring the charity—including the clinic—is financially resilient.



1. Navigating the Future: Donor Landscape & 2025 Projections

Second Chance Future has demonstrated commendable growth in its fundraising efforts over the past four years. Between 2021 and 2024, our total donation income reached a significant **£537,000**, representing a robust **34.9% increase** from the £398,000 raised in the prior period.

Crucially, this financial uplift wasn't solely driven by a massive influx of new supporters. Our donor base expanded by a modest **9.3%** (from 824 to 901 donors). This highlights an encouraging trend: **Donor Depth.** Our existing donors are choosing to contribute more generously, signalling a strengthening relationship with our mission and clinical outcomes.

2. The Path Ahead: 2025 Foresight and the Retention Alert

As we look towards 2025, our AI-driven predictive models offer a high-integrity forecast. Assuming our current engagement strategy holds, we project an estimated **£114,840** in likely donation income from **479 identified "Likely Donors."**

However, from an operational standpoint, the data reveals a critical area for immediate attention: **Donor Retention.** Our retention rate stood at **61.8%** at the end of 2024. This suggests that while our donors are valuable, maintaining their consistent support requires a shift from reactive communication to a proactive, **Predictive Stewardship** model.

3. Actionable Insights for Fundraising & Leadership

For the Fundraising Team: Precision Engagement

- **Deep Dive into "Likely Donors":** Utilize the donor_predictions_2025 view to segment the 479 likely donors by their propensity scores. High-propensity donors should be targeted for major gift asks or clinic tour invitations, while moderately likely donors receive impact-focused "check-in" communications.
- **Target "At-Risk" Segments:** Use the Days_Since_Last_Donation signal to identify supporters who are beginning to lapse. Re-engaging these donors now is significantly more cost-effective than acquiring new ones.
- **Optimize Channel Performance:** Use the Matrix visual on the "Detailed Page" to identify which channels (Email, Social, Direct Mail) deliver the highest **Average Donation per Donor** and re-allocate the marketing budget accordingly.

For Executive Leadership (COO Perspective): Strategic Resource Allocation

- **Baseline Budgeting:** Use the projected **£114.84k** as a "Confidence Floor" for 2025 financial planning. This allows us to commit to clinical staff and program costs with lower risk.
- **Incentivize Loyalty:** Shift our primary success metric from "New Name Acquisition" to "Retention Recovery," aiming to move from 61.8% toward an 80% target through structured donor journeys.
- **Technology & Infrastructure ROI:** This predictive architecture confirms the value of our investment in data. We must plan for regular retraining of these AI models to ensure our forecasting remains sharp as new donor data flows in from the clinic and Salesforce.

Conclusion

By systematically addressing these insights, Second Chance Future can transform "Data Chaos" into a powerful roadmap for sustained fundraising success. We are no longer guessing who might give; we are leading with **Operational Certainty** and **Predictive Stewardship**.

The Final Action Plan: Sustaining Your Stewardship

Building the AI-driven donor propensity model is only the first step. To ensure long-term fundraising sustainability, you must integrate this tool into your organization's operational rhythm.

1. Monthly Maintenance Routine

- **Update the Source:** Import the latest donation records from Salesforce (or your CRM) into your **arkbi** BigQuery dataset.
- **Refresh the Brain:** Run your scheduled SQL scripts to re-train the **Linear Regression** model on the most recent data to capture shifting donor behaviors.
- **Interface Sync:** Open Power BI and click **Refresh** to pull the updated AI propensity scores into your executive dashboards.

2. The Board Presentation Strategy

- **Highlight the "Predictive Baseline":** Use the **Likely Donation (2025)** card to show the board a "Confidence Floor" for next year's budget, moving away from optimistic guesswork.
- **Prove Stewardship ROI:** Use the **Donor Retention %** to justify investment in engagement staff, showing exactly how keeping donors is stabilizing the organization's mission.
- **Drill Down with Authority:** Use the **Campaign and Channel Slicers** during the meeting to answer questions in real-time, proving your granular control over which fundraising strategies are actually delivering results.

3. Scaling for Complexity

The skills you have mastered here are the foundation for high-complexity non-profit engineering. This blueprint is designed to scale, allowing for future integrations of live CRM systems via API or consolidating data across multiple clinical sites and international regional offices.

Final Summary & Author's Note: The Stewardship Mandate

As we conclude this blueprint, reflect on the transformation you have navigated. We began with the challenge of **"Data Chaos"**—fragmented spreadsheets and "brittle" manual tracking that failed to provide a clear view of our supporters' needs. Through the implementation of **Google BigQuery ML** and **Power BI**, you have transitioned from a reactive posture to one of **Predictive Stewardship**.

The Core Pillars of Your New Model

By following this guide, you have established a professional-grade framework built on four immutable pillars:

- **Automation:** You have replaced hours of manual data wrestling with a cloud-based engine that prepares and analyzes donor data with clinical precision.
- **AI Foresight:** Through the **Linear Regression** algorithm, you now possess the ability to identify "High-Propensity" donors 12 months in advance, moving from fire-fighting to proactive growth.
- **Relational Integrity:** By adopting the **Star Schema**, you have connected Donors, Campaigns, and Channels into a unified "Skeleton," ensuring your reporting is always context-aware.
- **Stewardship Empowerment:** With metrics like **Donor Retention** and **Propensity Scores**, you have shifted the narrative from "asking for money" to "fostering a community," directly improving the long-term health of the charity.

Author's Note

The journey of a COO is defined by the balance between **Mission Delivery** and **Operational Capacity**. This manual was designed to be your technical ally in that pursuit. As you move forward, remember that this model is a living asset. Foster a culture of data literacy within your teams and reclaim your time for what matters most: your mission and your patients.

Looking Ahead: Specialized Consulting

If you find yourself facing high-complexity challenges that move beyond this foundational guide—such as:

- **Live CRM/EHR Integrations:** Connecting BigQuery directly to systems like **Salesforce Health Cloud**, **Epic**, or **SystmOne**.
- **Advanced AI Customization:** Building bespoke ML models for donor lifetime value (LTV) or specialized patient pathways.
- **Enterprise-Level Scaling:** Moving to high-volume data orchestration for national healthcare missions.

I am here to support you. I specialize in building bespoke, high-performance clinical and fundraising intelligence solutions tailored to the unique needs of healthcare non-profits.

Start Your Transformation

To discuss a bespoke strategic roadmap for your organization, please complete my project inquiry form. This ensures I can review your specific data environment before our initial consultation.

Let's Connect and Amplify Your Impact

- **Project Inquiry:** [Insert HubSpot Link]
- **LinkedIn:** https://www.linkedin.com/in/andrew-alli/
- **Website:** www.arkbi.co.uk
- **Email:** andrew@arkbi.co.uk

To Your Continued Success,

Andrew Alli Non-Profit BI Consultant

ARK Business Intelligence Ltd


THE ERA OF "REARVIEW MIRROR" REPORTING IS OVER.
In the modern non-profit landscape, transparency is your currency, and data is your most underutilised asset. Yet, most organisations operate in a state of Data Chaos—paralysed by disconnected silos and "spreadsheet wrestling" that provide only a post-mortem look at the health of your mission.
The Predictive Stewardship Blueprint is the masterclass for non-profit leaders who refuse to lead blindly. By leveraging the "Brain" of Google Big Query and the "Interface" of Power BI, this blueprint provides a technical and strategic escape path from reactive survival to AI-driven foresight.
Inside this blueprint, you will discover how to:
• Shatter Data Silos: Deploy Big Query as a "Silo-Killer" to create a 360-degree view of your donor landscape.
• Engineer AI Foresight: Use Linear Regression to calculate donor propensity scores and generate high-integrity 12-month revenue forecasts.
• Audit-Proof Your Strategy: Move beyond anecdotal success to evidence-based narratives backed by R-Squared model accuracy.
• Reclaim 40+ Hours a Month: Automate manual data consolidation and redirect your team's energy toward your true calling—impact.
Don't just report on the past. Architect the future of your mission with Predictive Stewardship.
Kolade Andrew Alli