Handover Student Assistent

Tim Roessling

Table of contents

Chapter 1

Introduction

Glad to have you on the team!

This handbook is here to help you settle in and get up to speed. Inside, you'll find tips and resources to make your job easier, like:

- **Operational Tasks**: What you'll be doing day-to-day.
- **Dashboards**: How our dashboards work, with examples, screenshots, and code snippets.
- **Specific Explanations**: Explanations for specific features and measures in our dashboards.
- **Google Cloud Platform (GCP)**: How to use the GCP tools we need.
- **Tabular Editor**: How to handle data models with Tabular Editor.
- **Project Ideas**: Things you can work on, including ongoing projects you might want to pick up.
- **Useful Links**: Handy resources and docs.
- **HR Tips**: How to do things like submit timesheets and manage your schedule.

Check back here whenever you need a reminder or want to learn something new.

Starting a new job can feel overwhelming, but you'll pick things up quickly. If you get stuck or have questions, just ask Tomas—he's always happy to help.

All the best, and enjoy your time here!

Tim

Chapter 2

Operational Tasks

This section outlines the key operational tasks I regularly managed. These activities are crucial for keeping our dashboards current and ensuring timely data sharing with stakeholders.

Operational tasks are divided between you and Tomas. You can find and track these tasks on the team's **Asana board**, where they are tagged as *operational*:

- Asana Tasks Board

Additionally, Tomas has prepared a detailed file describing each operational task and the steps required to complete them:

- Tomas Master Document

Below, you will find further explanations about the specific tasks I managed and those you will now be responsible for.

2.1 Weekly Task Overview

Day	Task(s)
Monday	Bloomreach Update, Adobe Data
Tuesday	LiftLab QA
Wednesday	Promo Calendar, Media Spend File (bi-weekly)

2.2 Monday

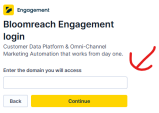
2.2.1 Bloomreach Update

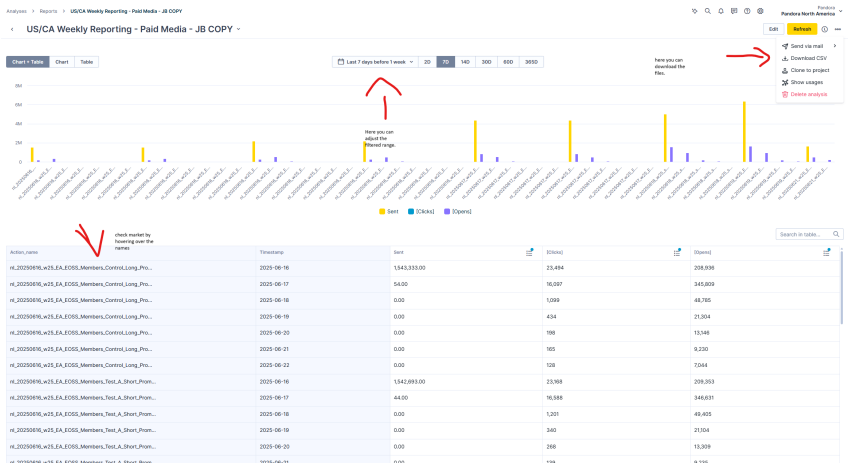
Purpose:
Ensure the latest Bloomreach data is shared with LiftLab. Bloomreach manages our customer engagement and marketing data.

2.2.1.1 Steps

- **Receive Data**
  - Every Monday at 8 AM, you will receive an email with the previous week's Bloomreach data at the market level.
  - Copy this data from the email into a new Excel file.
  - Tomas can assist with initial account setup and email scheduling if needed.

- **Check and Prepare Data**
  - Only the most recent week from the Bloomreach email is needed.
  - For US/CA, the email may only include US data. Log in to Bloomreach to ensure both US and CA are covered for the entire week.
  - Bloomreach Analytics (Login with 'pandora' at Bloomreach Engagement if prompted.)
  - Filter for the last week, ensuring CA is included. Download the CSV. Merge files if needed.
  - Replace the US/CA data in your file with the newly retrieved data (including Canada). Other markets can remain unchanged.





- **Format and Upload Data**
  - Flip the order of columns: **#0 Sent** should come before **#1 timestamp**.
  - Remove any unnecessary headers.
  - Copy the latest week into the Google Sheet, ensuring the format matches requirements.
  - Refer to the Tomas Master Document for detailed steps.

- **Schedule Backfill in GCP**
  - GCP Scheduled Queries

- **Generate and Send CSV**
  - Generate the CSV of the latest week and send it to LiftLab.
  - Email the CSV to:
    `pandora-us_datafeed@liftlab.com`
    CC: `nchalla@liftlab.com`, `jjohn@liftlab.com`

2.2.2 Adobe Data

Purpose:
Update the Adobe data used in our All Tides Rise Dashboard.

Instructions:

Follow the steps outlined in the Tomas Master Document.

Tip

Troubleshooting:
- If the Adobe Report Builder refresh takes unusually long, try refreshing the page and running the report again.
- If problems persist, ask Tomas to attempt the refresh from his end.

2.3 Tuesday

2.3.1 LiftLab QA

Purpose:
Perform a quality assurance (QA) check on the data displayed in the LiftLab ROI Dashboard to ensure spend data matches our internal sources.

Instructions:

- Ensure you have access to the LiftLab ROI Dashboard.
- The dashboard updates every Tuesday; complete your checks on this day.

What to Check:
For each market, verify the spend data for these channels:

- **Paid Social**
- **Paid Search**
- **TV**

Note

If the spend difference is greater than 2.5%, contact `nchalla@liftlab.com` and ask for clarification. You may also loop in Kasper if necessary. Use this template for comparing the spend: LiftLab QA

2.3.1.1 QA Steps

- **Paid Social**
  - Retrieve spend data from Smartly.io.
  - Smartly reports TikTok spend in USD and other channels in local currency. All spend must be converted to local currency for consistency.
  - Use the provided Python script to: Python Script
    - Upload Smartly data.
    - Convert currencies as needed.

    - Aggregate spend by channel and market.
    - Transfer the results to the designated Google Sheet for review.
  - For questions about Smartly.io or account setup, contact the Paid Social team.
- **Paid Search**
  - Retrieve spend data directly from the Google Ads interface.

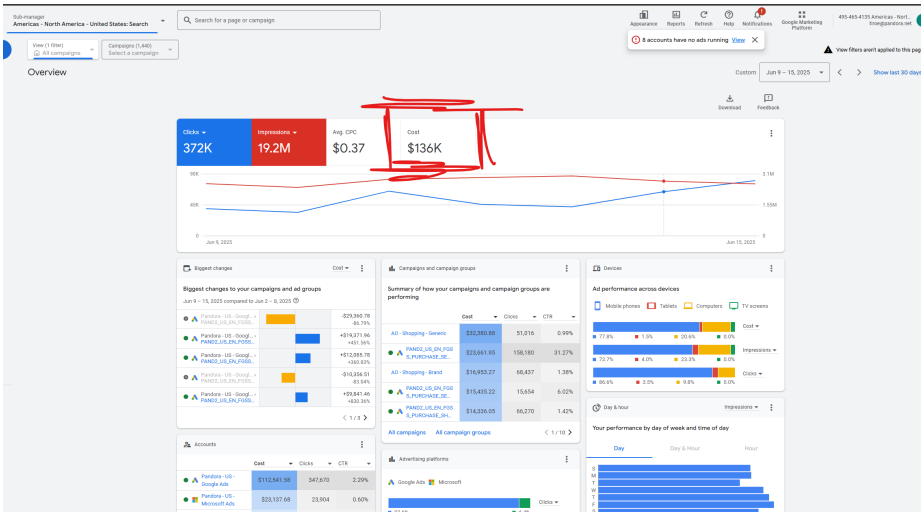


Figure 2.1: Google Ads Interface

  –

- **TV**
  - Use the Media Spend File to check TV spend.
  - Filter for TV and compare against the planned budget.

> **Tip**
>
> - Always ensure currency conversions are correct, especially for TikTok spend.
> - Use the Google Sheet to compare and document any discrepancies.
> - Automation is possible but currently limited by login requirements and the need to fetch aggregated reports across markets. LiftLab, in particular, requires market-by-market checking.

2.4 Wednesday

2.4.1 Promo Calendar

Purpose:
Share the promo calendar received from merchandising with LiftLab on a weekly basis.

> **Instructions:** Follow the steps from the master document to complete this task.

2.4.2 Media Spend File (Bi-Weekly)

Purpose:
Update the media spend file bi-weekly (usually Wednesday).

Instructions:

- Tomas first refreshes the file.
- You must ensure the dashboards are updated by feeding the latest data into the Google Docs, updating the exchange rates, and later sharing the file with LiftLab.
- Use a BI file to refresh spend in the correct format and export the CSV from there.
- Send the CSV to the usual email address:
  `pandora-us_datafeed@liftlab.com`
  CC: `nchalla@liftlab.com`, `jjohn@liftlab.com`

For step-by-step instructions, always refer to the Tomas Master Document.

Chapter 3

Dashboards

Welcome! This section will familiarize you with our dashboards. As a student assistant, you will support Tomas with a variety of tasks, including data updates, new visualizations, bug fixes, and feature development. Rather than a step-by-step Power BI guide, this section gives a brief overview on the dashboards, their important features, and the data sources used in our models.

3.1 Dashboards Overview

You will primarily work with two dashboards:

3.1.1 1. Weekly Media Plans (WMP)

- WMP Dashboard Link
- Based on the global media template, ensuring consistency and alignment across markets.

3.1.2 2. All Tides Rise (ATR)

- ATR Dashboard Link
- Provides global and local views of performance indicators and business estimates based on marketing impact.
- Focuses on actionable insights at the market and weekly level.
- Developed in collaboration with multiple internal teams to support trade conversations with GMs and above.

  Tip: All other dashboards can be found in the workspace of the Paid Media team. Golabl Paid Media - Dashboards

---

3.2 Getting Started

To begin, familiarize yourself with the dashboards by understanding:

- Measures
- Data sources and dataflows
- Data model and relationships

Just click through the dashboards and try to understand how

3.3 Data Modeling in Power BI

- **Linking Tables:** Relationships are created using primary and foreign keys. Selecting the correct cardinality (one-to-one, one-to-many, many-to-many) is crucial.
- The most used tables are the **Master Date Table** and the **Market Table**. These connect to most other tables, providing a common date and market context.
- **Backend Adjustments:** Most backend data model adjustments can be done in Tabular Editor, which is faster than Power BI Desktop as it does not load frontend visualizations. However, we currently work mostly in the **Power Query Editor** to manage tables and queries.
- Find more information on Tabular Editor in the respective section of this book.

3.4 Data Sources

Especially for ATR, we manage a variety of data sources in different formats:

- **Retail Cube:** Company-wide retail data and metrics, connected directly.
- **Adobe Data:** Stored and connected from BigQuery.
- **Customer Insights Team:** Data from Customer Acquisition, often in Excel files.
- **Global Media Template:** Excel file, updated weekly.
- **Asana:** Task process data fetched into the dashboard (for WMP and ATR).
- **Brand Interest & Category Interest:** Data from Google, stored in BigQuery.
- **Exchange Rates, Dates, Market Info:** Also stored in BigQuery.

  There are more sources you will encounter as you work with the dashboards. Data comes in various formats, offering opportunities for optimization and standardization to reduce maintenance.

WMP uses fewer sources, mainly the global media template, exchange rates, dates, and market information.

3.5 Important Features

3.5.1 Spend Conversion to DKK

- A calculated column in the planned media spend table converts spend from the GMT to DKK.
- This is done by multiplying the spend with exchange rates and fees for the respective week and market.
- Fees are stored in an Excel file.

3.5.2 Actualization Logic

- The GMT reports both planned and actual spend.
- To determine actualization, we compare reported spend at the most granular weekly level (Market, Channel, Subchannel, etc.).
- If `actual spend * 0.5 >= planned spend`, it is considered actualized; otherwise, not actualized.
- Once a week is actualized, all previous weeks are assumed actualized.
- Managed with a calculated column on the planned media spend table.

3.6 Publishing Changes or Data Refreshes

1. Click the **Publish** button and follow the prompts.
2. Select the workspace **Global Paid Media**.

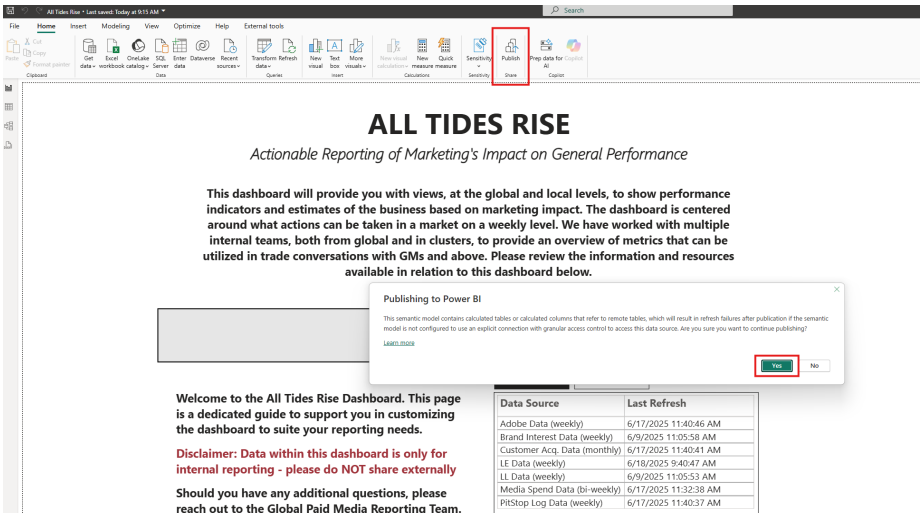


Figure 3.1: Publish changes

Chapter 4

Measures

This section outlines some some features i have created and it is useful to have some documentation for them.

4.1 BrandInterest & Category Interest in BigQuery and PowerBI

To enhance performance and reliability, both BrandInterest and Category Interest calculations are now performed in BigQuery.

4.1.1 BrandInterest in BigQuery

There are two main views for BrandInterest in BigQuery:

- **BrandInterestAGG**: The primary view containing all key metrics for all markets across the full date range.
- **BrandInterestAGG_2**: A fallback view that always contains the latest complete week for each market (one row per market). This ensures that if a user selects a week with incomplete data, the dashboard will automatically display the latest available complete week for each market. If the selected week has complete data, the dashboard will show that week as expected. This approach avoids complex workarounds or heavy computations in PowerBI and ensures users always see the most relevant and complete data.

The same approach is applied to CategoryInterest data.

- BrandInterestAGG View
- BrandInterestAGG_2 View
- CategoryInterestAGG View
- CategoryInterestAGG_2 View

These views are connected to PowerBI, allowing you to use simple measures directly in your visuals. The tables are linked to the master dates table using `Week_ForForFive` and to the market dimension using the market key.

Below is the SQL query for the **BrandInterestAGG** view, with explanations:

```
-- Step 1: Aggregate weekly data per country and map country names to market codes
WITH Aggregated AS (
    SELECT
        b.country,
        -- Map country names to 2-letter market codes
        CASE b.country
            WHEN 'Australia' THEN 'AU' WHEN 'Canada' THEN 'CA' WHEN 'Germany' THEN 'DE
            WHEN 'Spain' THEN 'ES' WHEN 'France' THEN 'FR' WHEN 'Italy' THEN 'IT'
            WHEN 'Poland' THEN 'PL' WHEN 'United Kingdom' THEN 'UK' WHEN 'United State
            WHEN 'China' THEN 'CN' WHEN 'Mexico' THEN 'MX' ELSE 'N/A'
        END AS Market,
        MIN(b.date) AS FirstDayOfTheWeek, -- Get the earliest date for the week
        CAST(m.Week_ForForFive AS INT64) AS Week_ForForFive, -- Week identifier
        AVG(b.queries) AS Sum_BI, -- Average branded queries for the week
        MAX(m.WeekDay) AS Max_WeekDay, -- Last day of the week
        AVG(b.queries / b.total_searches) AS avg_share_of_searches -- Average share of
    FROM `pandora-dashboard-data.google_offline_data.branded_searches_backfill` b
    JOIN `pandora-dashboard-data.Master_Date_Table.master_dates` m
        ON b.date = PARSE_DATE('%Y-%m-%d', m.DW_TS_From) -- Join on date
    WHERE b.Brand = "Pandora" -- Filter for Pandora brand
    GROUP BY b.country, m.Week_ForForFive
),

-- Step 2: Calculate rolling averages for the current year (CY)
CY AS (
    SELECT
        Market,
        Week_ForForFive,
        Week_ForForFive - 100 AS Week_ForForFive_LY, -- Calculate last year's week
        Max_WeekDay,
        FirstDayOfTheWeek,
        Sum_BI,
        avg_share_of_searches,
        -- 4-week rolling average of Sum_BI (current and next 3 weeks)
        AVG(Sum_BI) OVER (PARTITION BY Market ORDER BY Week_ForForFive DESC ROWS BETWE
        -- 4-week rolling average of avg_share_of_searches
        AVG(avg_share_of_searches) OVER (PARTITION BY Market ORDER BY Week_ForForFive
        -- 12-week rolling average of Sum_BI
        AVG(Sum_BI) OVER (PARTITION BY Market ORDER BY Week_ForForFive DESC ROWS BETWE
    FROM Aggregated
```

```
),

-- Step 3: Calculate rolling averages for the last year (LY)
LY AS (
    SELECT
        Market,
        Week_ForForFive,
        Sum_BI,
        avg_share_of_searches,
        -- 4-week rolling average for LY
        AVG(Sum_BI) OVER (PARTITION BY Market ORDER BY Week_ForForFive DESC ROWS BETWEEN CURRENT
        -- 12-week rolling average for LY
        AVG(Sum_BI) OVER (PARTITION BY Market ORDER BY Week_ForForFive DESC ROWS BETWEEN CURRENT
    FROM Aggregated
)

-- Step 4: Join CY and LY, calculate YOY metrics and lagged values
SELECT
    cy.Market,
    cy.Week_ForForFive,
    cy.Max_WeekDay,
    FORMAT_DATE('%Y%m%d', cy.FirstDayOfTheWeek) AS DateKey, -- Format date for reporting
    cy.Sum_BI,
    cy.avg_share_of_searches,
    cy.Four_Week_Avg_CY,
    cy.Four_Week_Avg_Share_CY,
    cy.Twelve_Week_Avg_CY,
    ly.Four_Week_Avg_LY,
    ly.Twelve_Week_Avg_LY,
    ly.avg_share_of_searches as avg_share_of_searches_ly,

    -- Year-over-year (YOY) comparisons (as percentages)
    (cy.Four_Week_Avg_CY / NULLIF(ly.Four_Week_Avg_LY, 0)) * 100 AS Four_Week_YOY,
    (cy.Twelve_Week_Avg_CY / NULLIF(ly.Twelve_Week_Avg_LY, 0)) * 100 AS Twelve_Week_YOY,
    (cy.Sum_BI / NULLIF(ly.Sum_BI, 0)) * 100 AS YOY,

    -- Lagged YOY values (next week's YOY, for trend analysis)
    LEAD((cy.Four_Week_Avg_CY / NULLIF(ly.Four_Week_Avg_LY, 0)) * 100) OVER (PARTITION BY cy.Mark
    LEAD((cy.Twelve_Week_Avg_CY / NULLIF(ly.Twelve_Week_Avg_LY, 0)) * 100) OVER (PARTITION BY cy.
    LEAD((cy.Sum_BI / NULLIF(ly.Sum_BI, 0)) * 100) OVER (PARTITION BY cy.Market ORDER BY cy.Week_

    -- Lagged values for share of searches and rolling averages
    LEAD(cy.avg_share_of_searches) OVER (PARTITION BY cy.Market ORDER BY cy.Week_ForForFive DESC)
    LEAD(ly.avg_share_of_searches) OVER (PARTITION BY cy.Market ORDER BY cy.Week_ForForFive DESC)
    LEAD(cy.Four_Week_Avg_Share_CY) OVER (PARTITION BY cy.Market ORDER BY cy.Week_ForForFive DESC
```

```
FROM CY
LEFT JOIN LY
    ON cy.Market = ly.Market
    AND cy.Week_ForForFive_LY = ly.Week_ForForFive -- Join CY with LY by market and wee
ORDER BY cy.Market ASC, cy.Week_ForForFive DESC;
```

4.2 Customer Acquisition

With the integration of Customer Acquisition (CA) into the dashboards, several changes have been made. We now use monthly data for CA reporting.

```
CA_Month_YoY_Simplified =
VAR CurrentYear = [Current Year Number WA]
VAR SelectedWeek = SELECTEDVALUE('Date'[Week Number int 2])

-- Step 1: Get the MonthIndex for the selected week
VAR SelectedMonthIndex =
    CALCULATE(
        MAX('Date'[YearMonth445]),
        ALL('Date'),
        'Date'[Year] = CurrentYear,
        'Date'[Week Number int 2] = SelectedWeek
    )

-- Step 2: Find latest populated MonthIndex from the fact table
VAR LatestMonthNr =
CALCULATE(
    MAX('Customer_Acquisition_Monthly'[MonthYR]),
    TOPN(
        2,
        ALL('Customer_Acquisition_Monthly'),
        'Customer_Acquisition_Monthly'[Year], DESC,
        'Customer_Acquisition_Monthly'[Month Nr], DESC
    )
)
-- Step 3: Determine which month to use based on the simplified logic
VAR UseMonth =
    IF(
        SelectedMonthIndex <= LatestMonthNr + 1,
        SelectedMonthIndex - 1,
        LatestMonthNr
    )
```

```
-- Step 5: Get values for CY and LY
VAR CY_Value =
    CALCULATE(
        SUM(Customer_Acquisition_Monthly[NbrCustomers]),
        Customer_Acquisition_Monthly[MonthYR] = UseMonth
    )

VAR LY_Value =
    CALCULATE(
        SUM(Customer_Acquisition_Monthly[NbrCustomers]),
        Customer_Acquisition_Monthly[MonthYR] = UseMonth - 100
    )

-- Step 6: Return YoY %
RETURN
    DIVIDE(CY_Value * 100, LY_Value, 0)
```

4.2.1 How the "CA_Month_YOY_Simplified" Measure Works

The measure "CA_Month_YOY_Simplified" calculates the Year-over-Year (YoY) change in customer acquisition for a selected week. It determines the relevant month based on the selected week and the latest available data, then compares the number of customers acquired in that month for the current year and the previous year.

4.2.1.1 Step-by-Step Logic

1. **Determine the Month for the Selected Week**
   The measure identifies which month (using the `YearMonth445` field) the selected week falls into.
2. **Find the Latest Populated Month in the Fact Table**
   It checks the fact table to find the most recent month (`MonthYR`) for which customer acquisition data is available (typically the last two months).
3. **Decide Which Month to Use for Reporting**
   The measure applies the following logic:
   - If the selected month index is less than or equal to the latest available month index plus one, use the selected month (or the previous month if the selected week is ahead by one).
   - If the selected month index is greater than the latest available month index plus one, use the latest available month with data.

This ensures the measure always uses the most relevant and available month for calculations, even if the selected week is ahead of the latest data.

4.2.1.2 Example Scenarios

Scenario	Selected Month Index	Latest Data Month Index	Logic Applied	Month Used for CA
1. User selects 3rd week of July, latest data is June	7	6	7 6 + 1 → Use previous month: 7 - 1 = 6	June
2. User selects a week in August, latest data is June	8	6	8 6 + 1 → False, use latest available month	June
3. User selects 2nd week of February, latest data is June	2	6	2 6 + 1 → True, use selected month	February

This logic ensures that:

- If the user selects a week in a month for which data is not yet available, the measure defaults to the latest available month.
- If the user selects a week in a month for which data is available, the measure uses that month.
- If the user selects a week in a past month (for which data is available), the measure uses the selected month.

This approach guarantees that customer acquisition reporting is always accurate and up-to-date, reflecting the most relevant data available.

4.3 Total Spend Calculated Column

The Total Spend calculated column is used in both dashboards (ATR and WMP). It displays either the actual or planned spend for each row, or 0, depending on the context—specifically, whether actual or planned spend should be shown. This is necessary because the GMT table contains both planned and actual spend values in the same table.

Using a calculated column instead of a measure provides greater consistency and flexibility: the result is always the same for each row, regardless of any filters applied in visuals. In contrast, a measure could change its value depending on the filters, which may not be desirable for reporting purposes.

4.4 The Index Logic to allow for dynamic year changes

To support dynamic year changes and accurate time-based calculations, we have implemented an index logic in the master date table. Each unique combination of week and year is assigned a sequential index (e.g., `WeekIndex`). This index is referenced in measures and calculations, enabling dashboards to seamlessly handle week-over-week or year-over-year comparisons—even when weeks span across different years.

4.4.1 Why Use an Index?

When calculating rolling averages (such as 4-week or 12-week averages), simply subtracting a fixed number from the selected week number does not work if the calculation crosses a year boundary. The index solves this by providing a continuous sequence across years.

4.4.2 How It Works

- The master date table contains a column (e.g., `WeekIndex`) that uniquely identifies each week, incrementing across years.
- Measures use the selected week's index to determine the correct range for calculations.
- For example, to calculate a 4-week average, the measure filters data where `WeekIndex` is between the selected index and (selected index - 3), regardless of year.

This approach ensures that time-based calculations remain accurate and robust, even as users select weeks from different years.

Example:

Suppose a user selects the first week of a new year. The index logic allows the measure to include the last weeks of the previous year in rolling calculations, ensuring continuity and correctness.

This method is essential for scenarios where rolling averages or comparisons span multiple years, providing consistent and reliable results in your dashboards.

4.5 Integrating Asana into Power BI

You can connect Asana to Power BI to automatically import and visualize project tasks.

4.5.1 How to Connect Asana as a Data Source

1. In Power BI Desktop, click **Home > Get Data > More...**.

2. Search for “Asana” and select it as a data source.
3. Paste the link to the specific Asana project you want to connect to.
4. Follow the prompts to authenticate and load your data.

4.5.2 Managing Tasks for Power BI Dashboards

- To include a new task in your dashboard, simply create it in Asana and apply the appropriate tag.
- The following tags are used to categorize tasks:
  - `Patch Notes ATR`
  - `Patch Notes WMP`
  - `Refresh`
- Use the “Patch Notes” tags for patch note items and the “Refresh” tag for refresh-related tasks.

The Power BI setup is configured to automatically fetch all tasks with these tags, ensuring your dashboards are always up to date with the latest relevant items from Asana.

Chapter 5

Tabular Editor

Tabular Editor is a powerful tool for developing and maintaining Power BI data models.
I have explored its capabilities using the free trial version, and I recommend purchasing **two individual licenses**—one for you and one for Tomas—to unlock the full feature set.

You can launch Tabular Editor directly from Power BI via the **External Tools** section.
This integration allows you to load your current data model and make changes seamlessly, without the need to export or re-import files.

5.1 Introduction to Tabular Editor

Tabular Editor offers an efficient interface for editing tabular data models.
It is especially valuable for backend dashboard management, enabling quick updates, advanced scripting, and automation.

5.2 Key Features Explored

5.2.1 Best Practice Analyzer (BPA)

- I have run the Best Practice Analyzer on the All Tides RIS model using a Best Practices JSON file from GitHub.
- This tool automatically scans the model for common issues and optimization opportunities, helping us adhere to recommended standards.

- **Best Practices JSON file:**
  Best Practices JSON

5.2.2 Tabular Object Model (TOM) Access

- Tabular Editor provides direct access to the Tabular Object Model (TOM), allowing us to inspect and edit the structure of our data models.
- This enables advanced modifications and efficient automation of repetitive tasks.

5.2.3 Dependency Visualization

- The dependency visualization feature helps us understand relationships between tables, measures, and columns.
- It is useful for identifying potential issues before making changes.

5.2.4 Diagram Mode

- Diagram Mode offers a clear graphical overview of our data model, making it easier to communicate design choices and understand the overall structure.

5.2.5 C# Scripting

- Tabular Editor supports C# scripting, which enables automation of repetitive tasks and customization of model management processes.
- With C# scripts, we can batch-create measures, automate formatting, enforce naming conventions, and apply best practices across multiple models efficiently.
- Leveraging C# scripting will be especially valuable for our future dashboard management, as it allows us to standardize processes, reduce manual errors, and accelerate development.
- As our models grow in complexity, scripting will help us maintain consistency and implement changes at scale with minimal effort.

5.3 Getting Started

Watch this useful introduction video:
Tabular Editor for Power BI

Chapter 6

Google Cloud Platform (GCP)

This section provides an overview of the Google Cloud Platform (GCP) tools we use for data management and analysis. GCP is essential for our data operations, enabling us to store, process, and analyze large datasets efficiently.

6.1 Introduction to Google Cloud Platform (GCP)

Google Cloud Platform (GCP) is a suite of cloud computing services offered by Google. It provides a comprehensive set of tools for computing, storage, networking, big data, machine learning, and application development. GCP allows organizations to build, deploy, and scale applications and data solutions efficiently and securely.

6.1.1 Key Functions of GCP

- **Compute**: Virtual machines, containers, and serverless computing for running applications and services.
- **Storage**: Scalable object storage, file storage, and databases for managing structured and unstructured data.
- **Networking**: Secure, high-performance networking solutions, including virtual private clouds and load balancing.
- **Big Data & Analytics**: Tools for data processing, analysis, and visualization, such as BigQuery and Dataflow.
- **Machine Learning & AI**: Pre-built and custom machine learning models, APIs, and infrastructure for AI workloads.

- **Security & Identity**: Comprehensive security, identity, and access management features to protect data and resources.

GCP's integrated services streamline workflows, improve collaboration, and support data-driven decision-making across the organization.

BigQuery Tutorial (YouTube)

6.2 How We Use GCP

We leverage GCP for:

- Storing and managing our data
- Data manipulation and transformation (using views)
- Automating data updates with scheduled queries
- Machine learning tasks and coding with Vertex AI
- Serverless computing tasks with Cloud Functions
- Data warehousing and analytics with BigQuery

6.3 Your Main Activities in GCP

As part of your role, your primary activities in GCP include:

- Backfilling data as part of operational tasks
- Creating and managing scheduled queries to automate data updates
- Managing tables and views in BigQuery
- Using Vertex AI for machine learning tasks and coding
- Maintaining Cloud Functions for serverless computing tasks

Chapter 7

Project Ideas

Below are some project ideas that could be useful for the team or for your own learning. These are suggestions—feel free to pick up any that interest you.

7.1 Dashboard Optimization

I have started working on optimizing dashboard loading times, especially for **All Tides Rise (ATR)**. After analyzing the loading times, I identified the main contributors to slow performance.

7.1.1 What has been done so far

- **Performance Analysis:**
  I analyzed dashboard loading times and identified bottlenecks. You can run your own performance analysis using the Power BI Performance Analyzer to see which visuals or queries take the most time to load.
- **Backend Optimization:**
  Calculations (e.g., 4-week averages, 12-week averages, long table filtering) were previously handled within Power BI, which slowed down performance. I have started moving these calculations and aggregations to BigQuery, using SQL scripts and views.
  For example, brand and category interest data are now calculated in BigQuery and pulled directly into Power BI, resulting in a more efficient data model and faster loading times.
- **Data Model Cleanup:**
  During development, redundant measures often accumulate in the Power BI data model. I have started deleting unused measures (about 30% so

far) in the ATR model.
I used Tabular Editor to identify all measures and then manually checked if they were used in visuals, as Tabular Editor is mainly backend-focused and cannot fully analyze frontend usage.
I have documented my steps and will link a file so you can continue this work.

Tabular Editor Measure Cleanup

7.1.1.1 Opportunities for further work

- Continue cleaning up unused measures in ATR and other dashboards.
- Explore optimizing other dashboards, though ATR is currently the most complex.
- Use Tabular Editor and C# scripts to work with the data model (see the chapter on Tabular Editor for more details).
- Consider moving additional heavy calculations and aggregations outside of Power BI.

7.2 General Recommendations

- Moving heavy calculations and aggregations outside of Power BI will benefit future projects.
- Tomas is investigating a long-term migration from BigQuery (currently used only by our team) to Databricks, the company-wide data warehouse.
  BigQuery gives us full ownership and admin rights, but requires manual data maintenance.
  Databricks will be a more sustainable, company-wide solution, though the migration timeline is still unclear.
  Keep this in mind for future projects.

7.3 Automation of Operational Tasks

- Automating repetitive operational tasks is definitely worthwhile.
- Manual logins can make automation challenging.
- I have experimented with different solutions using API keys (e.g., Google Search, Adobe Analytics, Asana).
  I was successful with Asana, but other APIs may require more setup or are not compatible with our environment.
- I encourage you to explore automation opportunities—these tasks recur weekly and can become tedious over time.
  Making them smarter and more efficient will save time in the long run.

7.4 Using Predictive Analytics for Measurement Efforts

I have had a class on predictive analytics. Predictive analytics can enhance our measurement efforts by providing deeper insights and more accurate forecasts. By leveraging historical data and advanced algorithms, we can identify trends, anticipate outcomes, and make data-driven decisions.

> **Tip:**
> If you are interested in exploring predictive analytics—such as forecasting, anomaly detection, or advanced modeling—reach out to Kasper.
> He is open to discussing ideas and would welcome collaboration on integrating these techniques into our measurement work.

7.5 Leveraging AI and Machine Learning

Recently, we completed a project leveraging AI for content classification—including images, articles, and videos—using Vertex AI and BigQuery. This initiative provided valuable insights into how large language models (LLMs) can accelerate workflows and unlock new efficiencies.

> The experience highlighted both the power and potential of AI-driven solutions for our team.
> If you're interested, I can share the presentation we delivered to the broader group, which covers our approach and key learnings.

AI Project Presentation

Code File

Exploring LLMs was both impactful and enjoyable, and I believe there are many more opportunities where these technologies could benefit our work.
If you're interested in applying AI or machine learning to future projects, I encourage you to get involved—there is significant potential for innovation and impact within the team.

7.6 Centralized Documentation for Data Processes

Proposal:
Set up a structured documentation space in Confluence to serve as the single source of truth for our data processes.
This would include:

- Data pipeline overviews and architecture diagrams
- Documentation of recurring data tasks and their owners
- Guides for onboarding new team members to our data stack
- Standard operating procedures for data ingestion, transformation, and reporting
- Troubleshooting guides and FAQs

Benefits:
- Reduces onboarding time for new team members
- Minimizes knowledge loss when people transition roles
- Improves consistency and quality of our data work
- Makes it easier to identify and address process gaps

If you are interested in knowledge management, technical writing, or improving team efficiency, this is a high-impact project that will benefit everyone.

Chapter 8

Useful Links

A collection of helpful resources and references.

Internal Portals

- PandoraNet SharePoint
  Internal documentation and resources.
- Pandora Digital Wiki
  Team knowledge base.
- Global Assets Home
  Asset management portal.

Project Management

- Asana Board
  Task and project tracking.

Coding & Data Warehouse

- Google Colab
  Cloud-based Python notebooks.
- Vertex AI Notebooks
  Managed Jupyter notebooks.
- BigQuery Console
  Data warehouse management.

Dashboards & Analytics

- Power BI
  Business intelligence dashboards.
- LiftLab Dashboard
  Analytics platform.
- Adobe Analytics
  Web analytics.
- Google Search Ads
  Search ads management.

HR & Scheduling

- Learning
  Training portal.
- SameSystem
  Scheduling tool.
- SuccessFactors Performance Manager
  Performance management.

Spreadsheets

- Google Sheet: Media Spend
  Media spend tracking.
- Google Sheet: LL QA
  QA documentation.

Chapter 9

HR

This chapter provides essential information about HR processes at Pandora.

9.1 Timesheet Submission

- **Monthly Timesheet:**
  You are required to submit a monthly timesheet, logging your working hours for each day.
- **Pay Periods:**
  Pay periods typically run from the 11th of one month to the 10th of the next month.
  Ensure all hours are logged by the 10th of each month.
- **Approval:**
  Kasper will review and approve your submitted hours.

9.2 Payroll

- **Payment Date:**
  Salary is usually paid on the 25th of each month.
  If the 25th falls on a weekend or holiday, payment may be processed on a different day.
- **Example:**
  On September 25th, you will receive payment for hours worked from August 11th to September 10th.

9.3 Learning Activities

- At the beginning of your employment, you will be assigned learning activities.

- New chapters and materials will be sent to you periodically.

9.4 Goals and Feedback

It makes sense for you to define your own goals after some time in the role, once you have a better understanding of your tasks and responsibilities. It is not mandatory for student assistants but it is definitely encouraged.

9.5 Support

- If you have questions, you can:
  - Email: DKHR@pandora.net
  - Ask in the student Teams chat
  - Ask Tomas