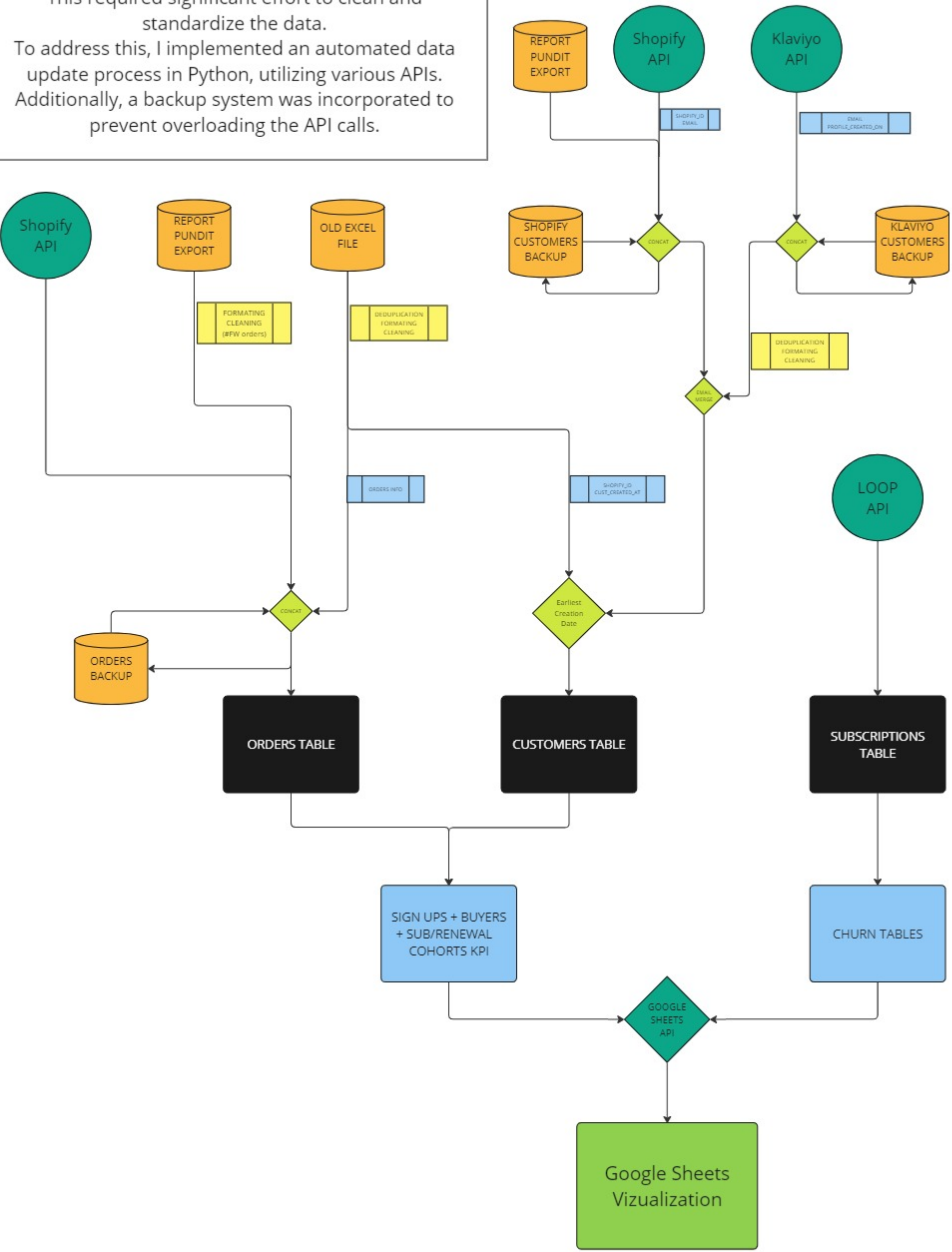
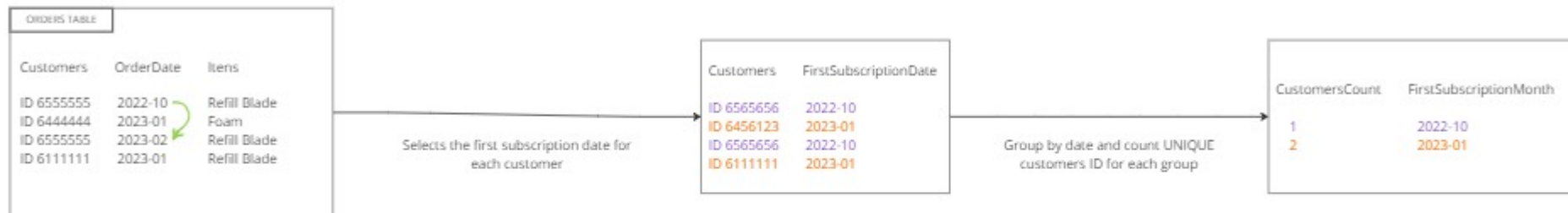


Depending on the maturity of a company's data analysis infrastructure, an analyst may need to prepare the data before it can be effectively used. In this project, my client's data was scattered across several platforms, with inconsistent formatting, duplicate entries, and some records that were invalid. This required significant effort to clean and standardize the data.

To address this, I implemented an automated data update process in Python, utilizing various APIs. Additionally, a backup system was incorporated to prevent overloading the API calls.



Subscriptions Cohort Sizes



Issue 1:

Some customers have more than one subscription, and in the orders table there is nothing to distinguish between these subscriptions. So, when the COUNT method is made, two or more orders from the same customer are counted as only one.

When these orders come from the same subscription, the count is correct. But when they come from different subscriptions, they are counted as one, so this may be a problem.

LOOP

Shopify

Subscription ID	Status	Frequency	Customer name	Created on	Next order date
#8210350182	ACTIVE	2 months	Antea Franzoni antea.franzoni@gmail.com	May 31, 2024	July 31, 2024
#7807271014	ACTIVE	6 months	Antea Franzoni antea.franzoni@gmail.com	April 16, 2024	October 16, 2024

Order	Date	Customer	Total	Items	Payment status	Fulfillment status	Tags
#FLER-326782	May 31 at 10:43 am	Antea Franzoni	€11.90	1 item	Paid	Fulfilled	Basic Subscription
#FLER-323248	Apr 16 at 1:28 pm	Antea Franzoni	€21.90	1 item	Paid	Fulfilled	Standard Subscription

Solution for Issue 1:

Simply using the cohorts size from the Subscription Table. Then I can use it to calculate Monthly Revenue.

I can distinguish between subscriptions using the Subscription ID.

Issue 2:

Some subscriptions have different creation dates compared to the customers' first subscription orders listed in the Excel file.

Old export orders

Nome cliente	Email cliente	Status c	Tipo di ordine	Data ordine
Chiara Bianchini	chiara.bianchini95@gmail.com	shipped	subscription	13/10/2021

Loop subscriptions

Subscription ID	Status	Frequency	Customer name	Created on
#7579861094	ACTIVE	3 months	Chiara Bianchini chiara.bianchini95@gmail.com	December 15, 2021

Solution for Issue 2:

Decide where to pick the date.