Data Analysis & Infrastructure Project Brief

About the company

You are working with Insurance data. The data consists of the following.

1. Portfolio Data
2. Policy Data
3. Policy Transactions
4. Sales Organization Data

1. Portfolio data.
   This is a data set that contains information about the current and historical portfolio for the insurance company. The following information about an insurance policy exists in this data.

Column Name	Data Type	Description
Key_Policy	Numeric	Surrogate Key to the policy data
Sales_Date	Date	The date when the sale was made
Key_SS_Org	Numeric	Surrogate Key to the Sales organization data
PaymentDate	Date	The date the payment was made for the annual premium
Payment_Status	Text	The status of payment. Paid or not paid
Annual_Premium	Numeric	The annual premium amount for the policy
Cancel_Before_Vld_Fm	Numeric	If the policy was cancelled before it was in effect, there would be a 1 otherwise 0
Cancel_After_Vld_Fm	Numeric	If the policy was cancelled after it was in effect, there would be a 1 otherwise 0
Product_Name	Text	The name of the product offered by the insurance company
No_Of_Sold_Policies	Numeric	This will always be 1
No_Of_Cancelled_Policies	Numeric	0 if the policy is active, 1 if it was cancelled
No_Of_Paid_Policies	Numeric	0 if the premium was not paid, 1 if paid
Vld_Fm_Tms	DateTime	The date and time from when this information is effective
Vld_To_Tms	DateTime	The date and time when this information will cease to be effective.

2. Policy Data
   This data set tells information about the policy. The data set can be joined with the portfolio data via key_policy. There should be only one row per key_policy in this table. The table consists of the following column.

Column Name	Data Type	Description
Key_Policy	Numeric	Surrogate key
Ext_Refr	Text	Policy Number
Term	Numeric	Term number of the policy
Vrsn	Numeric	Version of the policy
Sub_Vrsn	Numeric	Subversion of the policy
Incp_Dt	Date	The date when the policy was first created
Bnd_Dt	Date	Date from which the policy starts being effective
Pln_End_Dt	Date	Planned end date for the policy
Renew_Dt	Date	Planned renewal date for the policy
Cncl_Dt	Date	In case the policy was cancelled, the date.
Sts_Dt	Date	The date when the status for the policy was changed to the current status
Agrm_Sts_Cd	Text	Status of the policy. Active or Cancelled
Prtn	Text	If the policy is connected to a partner, then information about the partner is here.
Vld_Fm_Tms	DateTime	The date and time from when this information is effective
Vld_To_Tms	DateTime	The date and time when this information will cease to be effective.

3. Policy Transactions
   While working in a policy system, there is one transaction that is created for every action that is done on a policy. The transactions can be of any nature. In practice, one policy will have many transactions in its lifetime however the policy information may not change. The transaction data included here is very basic and include the following information. The policy transactions can be joined with policy data through policy number.

Column Name	Data Type	Description
Pev_CreatedAt	DateTime	The date time at the time of the transaction.
Pev_PortfolioResponsibleCode	Text	The portfolio responsible code for the transaction
Po_No	Text	Policy number
Pev_Id	Numeric	Surrogate Key

4. Sales Organization Data
   This table consists of data about the sales agent and their sales channel. The sales data can be joined with portfolio data through key_ss_org. The following is included in the sales organization data

Column Name	Data Type	Description
Key_SS_Org	Numeric	Surrogate Key
Fld_Rep_Cd	Text	Internal code for sales agent
Org_Lvl_Cd	Text	The code for the sales channel the agent belongs to.
Org_Lvl_Nm	Text	The name for the sales channel the agent belongs to.

Agreement with Wild West AB

There is an active agreement with Wild West AB where the aforementioned are authorized to earn commissions based on referrals they send towards the insurance company. For every sale, the insurance company pays a commission to Wild West AB. But there are certain rules.

Your task is to create a report for each month of the year, where the commissions are reported.

All policies that fulfill the following criteria are eligible for commissions.

1. The policy has been marked with the portfolio code as ‘WILDWEST-2’ or ‘WILDWEST-3’
2. The policy has been marked with the portfolio code for the entire duration of its existence.
3. The premium for the policy is paid.
4. The following products are within the scope of the agreement. No other product sold even with the portfolio code, will be valid for a commission:
   a. ‘Product 1’ through ‘Product 8’
   b. ‘Product 13’ through ‘Product 31’
5. Moreover, the policy sale should fulfil the following criteria:
   a. Outbound sales: portfolio code = ‘WILDWEST-3’
   b. Internet Sales: portfolio code = ‘WILDWEST-3’
   c. Inbound Sales: portfolio Code = ‘WILDWEST-2’

   No other sales channel would be eligible for commissions.

The commission is paid as per the following rules:

1. For the first 1500 policies, the commission is 12% of the annual premium.
2. From the 1501-th policy, the commission is 14% of the annual premium.

Clawbacks:
In case of cancellation of a policy mid-term, there needs to be a clawback. This means only the percentage corresponding to the active validity of the policy, of the commission will be paid and the rest will be returned to the insurance company.

The Deliverable

The deliverable for this project is a presentation where your group will explain the logic that you have implemented by showing the SQL code and how you have interpreted the requirements. (Including any assumptions that you've made).

The submission will also include data aggregated data for monthly reports from Jan 2024 – Aug 2024. (8 months).

Additional points, if you can come up with additional analysis on the monthly reported data.