Data Analyst Assignment - 5

Overview

Please understand the below-mentioned real-life scenario and try to solve the assignment.

Business Scenario

You are a data analyst, and your client has a large e-commerce company in India (let's call it ShopX). ShopX gets a thousand orders via their website on a daily basis, and they have to deliver them as fasts they can. For delivering the goods ordered by the customers, ShopX has tied up with multiple courier companies in India as delivery partners who charge them some amount per delivery.

The charges are dependent upon two factors:

- Weight of the product
- Distance between the warehouse/StoreHouse (pickup location) and customer's delivery address(destination location)

On an average, the delivery charges are Rs. 100 per shipment. So if ShopX ships 2,00,000 orders per month, they have to pay approximately Rs. 2 crore to the courier companies on a monthly basis as charges.

So ShopX wants to verify the accuracy of the charges imposed by their delivery partners for each order, as the current amount they have to pay to courier companies is excessively high.

Input Data

- **ShopX Data (ShopX's internal data spread across three reports)**

1) **Company ShopX - Order Report**

   a. **Order ID:** This column represents the unique identifier assigned to each order.
   b. **Product Code:** This column contains the unique identification code assigned to each product
   c. **Units Ordered:** This column represents the quantity of each product that was ordered.

   Note: Order ID is common identifier between **ShopX's** order report and courier company invoice

2) **Company ShopX - Product Weight**

   a) **Product Code:** This column represents the unique identification code assigned to each product.
   b) **Product Weight (g):** This column represents the weight of each product in grams.

Data Analyst Assignment - 5

Note: These columns should be used to calculate total weight of each order and during analysis compare against one reported by courier company in their CSV invoice per Order ID. The courier company calculates weight in slabs of 0.5 KG multiples, so first you have to figure out the total weight of the shipment and then figure out applicable weight slabs.

For example:
- If the total weight is 400 gram then weight slab should be 0.5
- If the total weight is 950 gram then weight slab should be 1
- If the total weight is 1 KG then weight slab should be 1
- If the total weight is 2.2 KG then weight slab should be 2.5

3) Company ShopX - Warehouse&Customer Pin Code and Zone details

a) **Store House Pincode:** This column represents the pincode of the storehouse or warehouse where the products are stored.
b) **Customer Area Code:** This column represents the area code of the customer's location or delivery address.
c) **Delivery Zone:** This column represents the assigned delivery zone for each customer's location

Note: Store House Pincode to All India pincode mapping -(this should be used to figure out delivery zone (a/b/c/d/e) and during analysis compare against one reported by the courier company in their CSV invoice per Order ID

➢ courier company invoice in CSV file

1) Courier Company – Invoice

a) AWB Code (Airway Bill Number): courier company's own internal ID
b) Order ID: company ShopX's order ID
c) **Chargeable Weight:** This column represents the weight of the shipment that is used to calculate the shipping charges.
d) **Store House Pincode:** Store house pickup Pincode
e) **Customer Area Code:** This column represents the area code of the customer's location.
f) Delivery Zone: This column indicates the specific zone or region where the shipment is scheduled to be delivered
g) **Freight Type:** This column specifies the type of freight or shipment, such as Forward charges, Forward and RTO charges etc.
h) **Total Amount (Rs.):** This column represents the total amount charged for the shipment in Indian Rupees (Rs.). It includes the shipping charges and any additional fees or costs associated with the order.

2) Courier Company – Rates

a) Courier charges rate card at weight slab and pincode level. If the invoice mentions "Forward charges" then only forward charges ("fwd") should be applicable as per zone and fixed & additional weights based on weight slabs. If the invoice mentions "Forward and rto charges" then forward charges ("fwd") and RTO charges ("rto") should be applicable as per zone and fixed & additional weights based on weight slabs.
b) For the first 0.5 KG, "fixed" rate as per the slab is applicable. For each additional 0.5 KG, "additional" weight in the same proportion is applicable. Total charges will be "fixed" + "total additional" if any

Data Analyst Assignment - 5

Output Data 1

Create a resultant CSV/Excel file with the following columns:
- Order ID
- AWB Number
- Total weight as per ShopX (KG)
- Weight slab as per ShopX (KG)
- Total weight as per Courier Company (KG)
- Weight slab charged by Courier Company (KG)
- Delivery Zone as per ShopX
- Delivery Zone charged by Courier Company
- Expected Charge as per ShopX (Rs.)
- Charges Billed by Courier Company (Rs.)
- Difference Between Expected Charges and Billed Charges (Rs.)

Output Data 2

Create a summary table

	Count	Amount (Rs.)
Total orders where ShopX has been correctly charged	<count>	<total invoice amount>
Total Orders where ShopX has been overcharged	<count>	<total overcharging amount>
Total Orders where ShopX has been undercharged	<count>	<total underchargingamount>

Submission

Please submit the result in Excel with two workbooks (summary table in one and order level calculation in another) and your code in Python programming language etc.

Please zip the files and share it on email.

Data Analyst Assignment - 5