Data Warehousing Problem Statement:

**Abstract:**

Every organization implements CRM as a strategy that integrates concepts of data mining, and data warehousing in order to support an organization’s decision-making process to retain long- term and profitable relationships with its customers.

Retail Bazaar Inc. is a hypothetical leading retail company in India with more than 5000 retail stores and hypermarkets across the country.

Company has set a dedicated team to handle after sales customer care services which handles customer complains and resolves them to increase customer retention, loyalty and decrease customer attrition.

**Below is an abstract of end to end process:**

* Company has multiple call centers across India to handle the customer requests.
* Customers can reach to the care team over different channels. Calls or Chat.
* CCR(customer care representative) will register the complaint with the all the provided details.
* A case can have status like open, closed and reopened.
* A case can have states like Escalated, Withdrawn, Abandoned etc.
* Each case can be given a priority score and depending on the score, CCR has an SLA to close the case in specific amount of time/days.
* Once case is closed, customer is sent a survey link to rate overall process.
* Customer can take a survey or leave it unattended. He can fill scores from 1-10 on various questions around how his experience was during customer care process.
* Survey response is captured for that particular case.
* Some of other dimensions of the CRM process is case category, product, language, channel, market, supplier etc.
* A case can belong to any of the categories like ‘subscription, ‘Purchase, ‘shipment’ etc.

This complete CRM process generate data and store in database OLTP tables.

Company wants to do some analysis around this data and capture below KPIs to further enhance and optimize the CRM process.

**KPIs:**

**Case KPIs:**

What is the total number of open/closed/reopened cases in a day/week/month and should be further sliced and diced by employee, product, priority, region, category and call centre dimensions?

**Chat KPIs:**

If case was handled through a chat with representative, then below KPIs for each employee

* Total number of chats handled by customer care representative in a day/week/month and should be further sliced and diced by employee, product, priority, region, category and call-centre dimensions.
* Total chat time of an CCR for same dimensions as above.
* Total time when representative was online but not chatting
* Total wait time for the representative. Wait time can be calculated as time when customer sent first message minus time when representative sent first response.
* Total time representative was offline.

**Survey KPIs:**

For surveys, if any question had 1-4 score from customer then it is negative response, 5-8 score is neutral and 9-10 is positive response.

* Total number of positive/negative/neutral responses in a day/week/month and further sliced and diced by employee, product, priority, region, category and call centre dimensions.

After understanding these KPIs, we need to think of possible facts and dimensions as well as the columns/information these tables might have to build a warehouse which can help to derive these KPIs.

A possible record in case fact table could be:

|  |  |
| --- | --- |
| Column name | Desc |
| transaction\_no | Unique value for a case |
| created\_employee\_key | Representaive ID who create the case |
| modified\_employee\_key | Representaive ID who modified the case |
| create\_date | Case create date |
| modified\_date | case modified date |
| call\_center\_code | call centre where case has been registered |
| region\_code | customer region |
| language\_code | customer language |
| category\_code | case category |
| product\_code | product for which case has been registered |
| priority\_code | case priority code |
| status\_code | current case status |
| state\_code | current case state |
| close\_date | case close date |