Hi Mr….,

Greetings!! I hope you are doing good.

I would like to take this mail as an opportunity to introduce myself.

I am Amal Sharma, and I am working as Data analyst in Fetch Rewards. I am assigned with the task to analyze Reward program for our customers, found out if there is any anomaly and create new robust relational database. Based on the analysis, we will define rules for our new Database which will be used to store records going forward. I have few questions related to data which will help me to understand its context better and I will be grateful if you can help in that regard.

* What are the critical attributes which is important in Receipt, brand and user data?
* What are the KPIs of your department which you monitor and want me to dig more in that area?
* How analysis from data impact operational decision in your data?
* Are we capturing user demographics to do analysis for luring new customers?

Further While analyzing the data, I have found out few issues which is severely impacting the data quality. These issues were found during data exploration. I have identified the actions which will help us in the resolution.

|  |  |  |
| --- | --- | --- |
| Category | Data Quality Issue | Resolution |
| Data Value Missing | Almost 50% of data related to bonus point, purchase date, Purchase count and total price of the receipt is missing which hinders to find the correlation between receipt amount and rewards earned. | Not Null Constraints to be put in the Schema to avoid missing value |
| Date not present | There is no data related to brand in the brand table so we cannot analyze our top brands and their Year of year performance. | Put Foreign key, Not null constraint so that data can be maintenance can be made mandatory |
| Format Standard | Date Format is not as per the standard. So not able to find out the past trends and predict the future growth | Define the Date format standard as per standard Db practice or Develop function which will get the system date & time while entering the record in the table. |

Apart from the initial Analysis, I want to know about compliance, security and privacy of data as it has information about the users which need to be encrypted. Also, what steps you take in present when you face duplicate records issue? These questions will help in optimizing the database we are creating.

Lastly, we need to keep in mind performance and scaling of our RDBMS so that it does not hamper our process and customers. Following actions are required:

|  |  |
| --- | --- |
| Points | Action |
| Data should be maintained properly, and junk data should not be maintained in the system | System rules and constraints to be defined |
| Tables should be maintained in 3 NF | To avoid data quality & duplicity issues |
| Regular monitoring and health hygiene check of systems | Health monitoring tools to be installed which will trigger alerts in case of anomaly |
| Vertical scaling | In future we might need to add more resources to our existing database to keep the performance at par |

I have attached presentation which define my findings in more detail.

Please let me know if you need clarification in any of the points stated above and I will be happy to schedule meeting to detail in discuss.

Thanks,

Amal Sharma