I wanted to list down all the anomalies in the data and the transformations I have made.  
  
Before writing anomalies, I wanted to describe my process in few steps.

1. Joining all the tables is the difficult thing as we didn’t have the similar joining key in the tables.
2. Firstly, I have joined **users table** and the **recommendations table** based on the user\_id primary key.
3. Now, the challenge is joining the **user\_recommendation table** and the moderator table. To tackle this challenge, I have followed the proportional approach based on the idea where one moderator can assist different users.
4. So, we can generate moderators **by (chat\_moderators\_managed/total moderated\_sessions)\*100.**
5. Now we have list of moderators and we are joining them with the **users\_recommendation** table.

Now I’m writing the anomalies.

1. After joining the data, I found a lot of missing joins for recommendations table. Its 606 users, if there is less missing join rate, we can be able to provide better platform for user satisfaction.

A blue and red pie chart

Description automatically generated

1. In Moderator data, some moderators responded quickly, but still received the less ratings. There is no proper correlation between the.

A graph of blue bars

Description automatically generated

1. In the user engagement data, there is also no proper correlation between messages\_Sent and the session\_length. It means users who have spent more time on the platform and sent more messages and still received less ratings. This means they are spending time passively on the platform which could even decrease the engagement score. Here is the plot  
   A chart of different colored dots

   Description automatically generated
2. The Final one is no proper joining column in moderators with the other two tables.