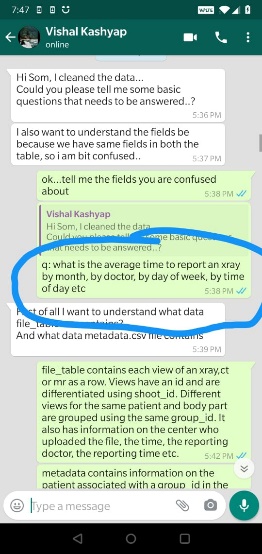
**Data analyst internship learnings**

**About application**

1. Application was the common platform that clients(eg : clinics, hospitals) and Doctors(eg : xrays specialist, bone specialist etc) were using which was generating the data.
   1. Client uploads an xray of patient and Doctors use to report their observation on an xrays.

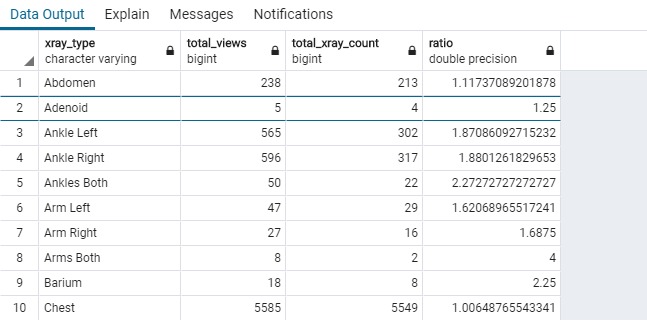
**RAW DATA**

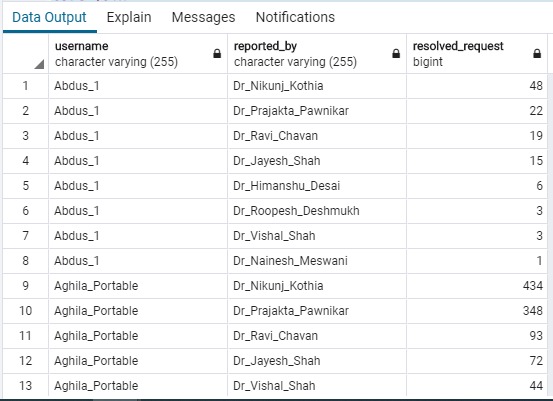
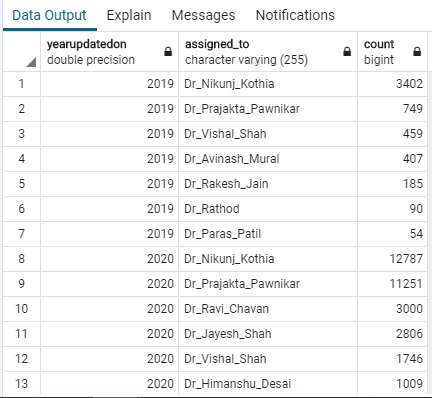
For data please refer to **table\_xraymetadata.csv** and **table\_files.csv .** These two table was joined with 2 columns **table\_xraymetadata.id = table\_files.group\_id.**

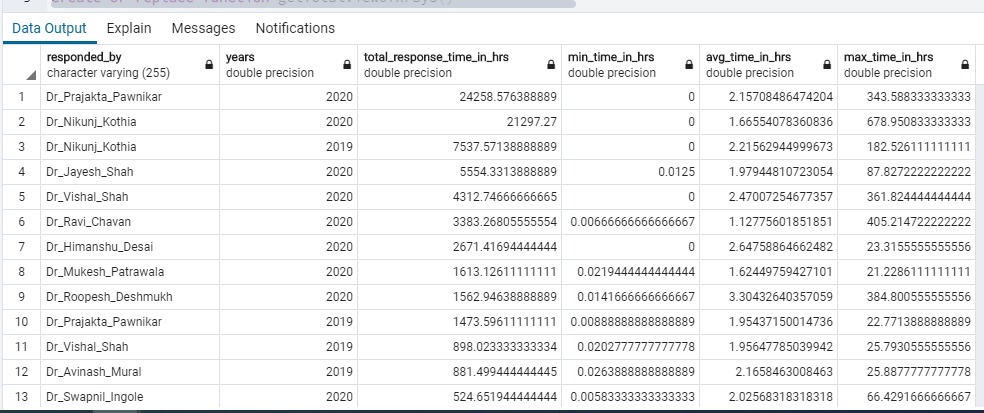
1. **table\_xraymetadata.csv** contains actual data like patient\_complain, xray\_type etc.
2. **table\_files.csv** contains views of xrays,ct or mr.
   1. For eg : All side xrays of skull will be 4 xrays, 1 for front,1 for back etc. These 4 xrays belongs to 1 patient so 4 entries will be made in this table which will have same group\_id as they belongs to 1 patient only. When client uploads these xrays on portal, table\_xraymetadata is the table which contains “how many xrays are been uploaded for 1 patient”.
   2. 

**We need to answer below question for business.**

1. What is the xray\_count and xray\_views by xray\_type? OR what is the ratio of xray count and xray views by body type(xray\_type)? Is it leg takes more xrays views?



1. How many clients viewed the reports and who reported it?
2. Who responds/resolves more requests by year?
3. What is the total,min,avg and max time taken for resolving a request?



1. What is the average time to report an xray, by month, by doctors, by day of week, by time of day etc?
   1. Please refer Cleaned\_formated\_data.xlsx file, in this file all the important **EDA** is done which helped in finding good insights.