**EdTech User Acquisition Analysis Project:**

This is a data analyst project in which I had extracted and analysed data and took out useful inferences from the given dataset. These inferences helped business teams to analyse the team performance and managers can focus on more specific target areas which will lead to increase in overall conversion rates.

**Data Reading and Cleaning:**

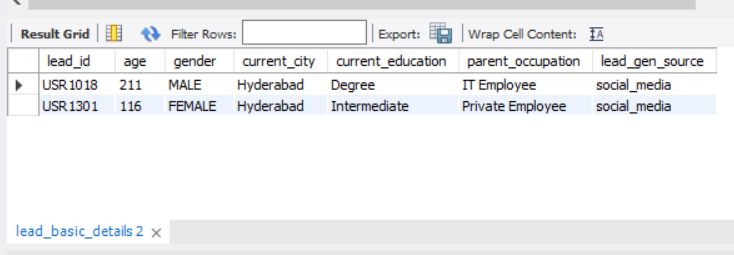
Create a database Edtech in Mysql and import all the given tables in that dataset:

create database edtech;

use edtech;

**Outliers:**

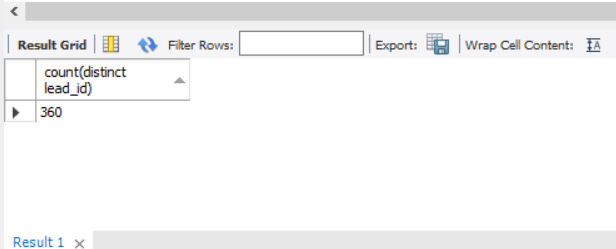
select \* from lead\_basic\_details where age not between 10 and 50;



**INFERENCE – Two outliers are found with wrong age out of total 360 leads.**

**Total Leads:**

select count(distinct lead\_id) from lead\_basic\_details;



**Conversion rates of leads by current education:**

select a.current\_education,count(distinct a.lead\_id) as leads,

count(distinct case when b.lead\_stage='conversion' then b.lead\_id end) as converted\_leads,

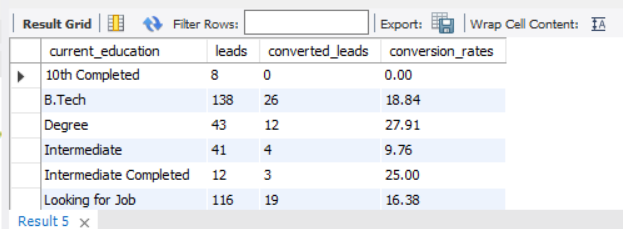
round(count(distinct case when b.lead\_stage='conversion' then b.lead\_id end)\*100.00/count(distinct a.lead\_id),2) as conversion\_rates

from lead\_basic\_details a

join leads\_interaction\_details b

on a.lead\_id=b.lead\_id

group by 1;



**INFERENCE-Conversion rates in Btech,Intermediate,10th completed is low we had to focus in these sections in correlation with other data to improve.**

**Conversion rates city wise:**

select a.current\_city,count(distinct a.lead\_id) as leads,

count(distinct case when b.lead\_stage='conversion' then b.lead\_id end) as converted\_leads,

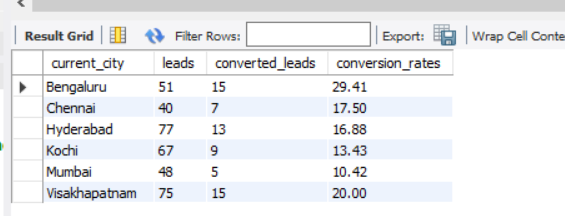
round(count(distinct case when b.lead\_stage='conversion' then b.lead\_id end)\*100.00/count(distinct a.lead\_id),2) as conversion\_rates

from lead\_basic\_details a

join leads\_interaction\_details b

on a.lead\_id=b.lead\_id

group by 1;



**INFERENCE-conversion rates of cities like kochi and mumbai is low we should focus on these cities more to improve conversion rate by corelating with other analysis.**

**Language Wise conversion Rates:**

select a.language,count(distinct a.lead\_id) as leads,

count(distinct case when b.lead\_stage='conversion' then b.lead\_id end) as converted\_leads,

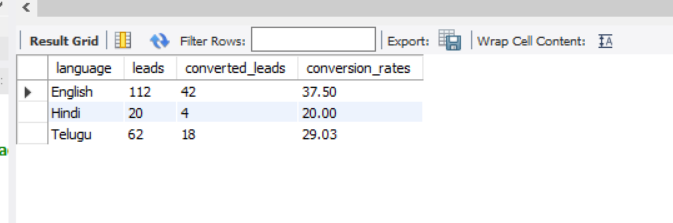
round(count(distinct case when b.lead\_stage='conversion' then b.lead\_id end)\*100.00/count(distinct a.lead\_id),2) as conversion\_rates

from leads\_demo\_watched\_details a

join leads\_interaction\_details b

on a.lead\_id=b.lead\_id

group by 1;



**INFERENCE-Conversion rate in Hindi is less so we hed to make more quality courses of Hindi.**

**Cycle wise convertion rates:**

select a.cycle,count(distinct a.lead\_id) as leads,

count(distinct case when b.lead\_stage='conversion' then b.lead\_id end) as converted\_leads,

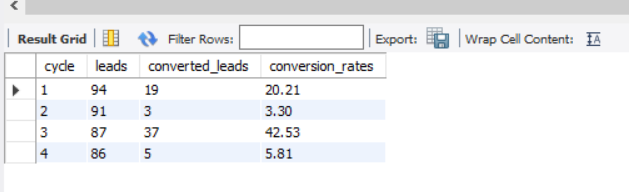
round(count(distinct case when b.lead\_stage='conversion' then b.lead\_id end)\*100.00/count(distinct a.lead\_id),2) as conversion\_rates

from sales\_manager\_assigned\_leads\_details a

join leads\_interaction\_details b

on a.lead\_id=b.lead\_id

group by 1;



**INFERENCE- conversion rates of cycle 3 and 4 are extremely low so managers should focus more on team performance at that cycle time.**

**Conversion rates from every lead generation source:**

with base as (SELECT a.\*,b.lead\_gen\_source

FROM edtech.leads\_interaction\_details a

left join lead\_basic\_details b on a.lead\_id=b.lead\_id)

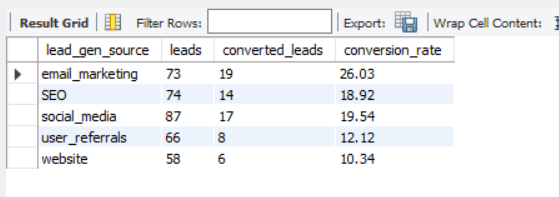
Select lead\_gen\_source,count(distinct lead\_id) as leads,

count(distinct case when lead\_stage='conversion' then lead\_id end) as converted\_leads,

round(count(distinct case when lead\_stage='conversion' then lead\_id end)\*100/count(distinct lead\_id),2) as conversion\_rate

from base

group by 1;



**INFERENCE-Conversion rates of user\_referrals and website are low so business team should focus on these lead gen source to bring more quality leads and side by side also correlate with other analysis.**

**Lead stage Journey-Dropout rates at every lead stage:**

select jnr\_sm\_id,count(distinct lead\_id) as total\_leads, count(distinct case when lead\_stage='lead' then lead\_id end) as leads,

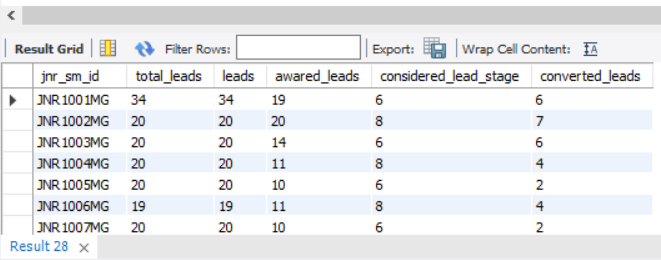
count(distinct case when lead\_stage='awareness' then lead\_id end) as awared\_leads,

count(distinct case when lead\_stage='consideration' then lead\_id end) as considered\_lead\_stage,

count(distinct case when lead\_stage='conversion' then lead\_id end) as converted\_leads

from leads\_interaction\_details

group by 1;



with base as(select jnr\_sm\_id,count(distinct lead\_id) as total\_leads, count(distinct case when lead\_stage='lead' then lead\_id end) as leads,

count(distinct case when lead\_stage='awareness' then lead\_id end) as awared\_leads,

count(distinct case when lead\_stage='consideration' then lead\_id end) as considered\_lead\_stage,

count(distinct case when lead\_stage='conversion' then lead\_id end) as converted\_leads

from leads\_interaction\_details

group by 1)

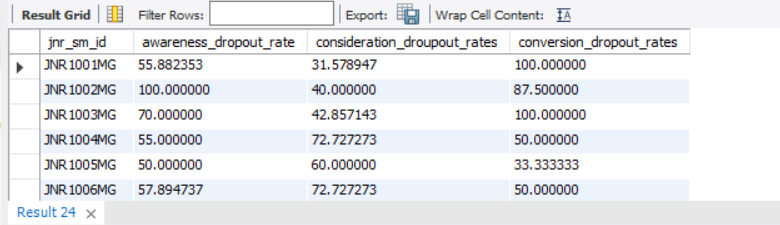
select jnr\_sm\_id,

awared\_leads\*100.00/leads as awareness\_dropout\_rate,

considered\_lead\_stage\*100.00/awared\_leads as consideration\_droupout\_rates,

converted\_leads\*100.00/considered\_lead\_stage as conversion\_dropout\_rates

from base group by jnr\_sm\_id;



**INFERENCE- Business team can see the conversion or dropout rates every stage that at which stage most leads are backing off also business teams can see the lead stage condion every junior manager.**

**Team Performance-Conversion rate of leads under every junior manager and senior manager:**

select s.snr\_sm\_id,l.jnr\_sm\_id,

count(distinct l.lead\_id) as leads,

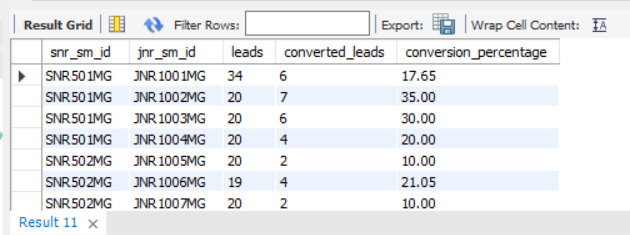
count(distinct case when l.lead\_stage='conversion' then l.lead\_id end) as converted\_leads,

round(count(distinct case when lead\_stage='conversion' then l.lead\_id end)\*100.00/count(distinct l.lead\_id),2) as conversion\_percentage

from leads\_interaction\_details l

left join sales\_manager\_assigned\_leads\_details s on l.lead\_id=s.lead\_id

group by l.jnr\_sm\_id,s.snr\_sm\_id;



**INFERENCE- With this analysis business heads can know which team is most converting leads into final conversion. This is an indepth analyis in which we are getting individual performance report of junior and senior sales manger.**

**conversion rates by lead generation source and city:**

with base as(select a.\*,b.lead\_gen\_source,b.current\_city

from leads\_interaction\_details a

left join lead\_basic\_details b

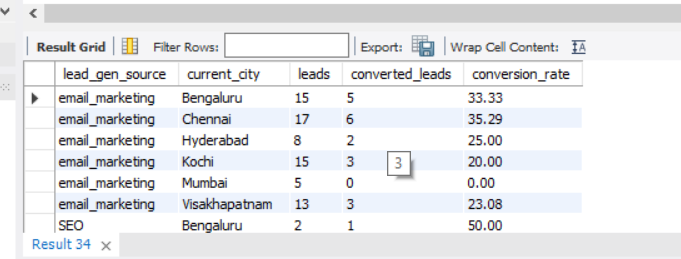
on a.lead\_id=b.lead\_id)

select lead\_gen\_source,current\_city,count(distinct lead\_id) as leads,

count(distinct case when lead\_stage='conversion' then lead\_id end) as converted\_leads,

round(count(distinct case when lead\_stage='conversion' then lead\_id end)\*100/count(distinct lead\_id),2) as coversion\_rate

from base group by 1,2



**INFERENCE-With this analysis we are getting conversion rates per lead gen source and city. This is a very deep analysis and can be used to know in which city which lead gen source is not giving good results.**