CASE STUDY PRESENTATION

PREPARED BY: DANIYAL ARIF

Data Investigation

Table 1: Group Members
Contains group join & leave events since group creation by teacher

Table 2: Session Log
Contains details about live sessions of attendances

Table 1: Group Members (Data Variables)

GROUP MEMBERS
student_id
group_id
joining date

- Observations: **14,798**
- Unique Students by Student ID: **14,365**
- Observations take place when a student **joins** or **leaves** a group
- Thus, there can be Student ID recurring. For instance, if a students joins 3 groups then that student id should be captured as three separate observations
- I merged student_id if with group_id to see if certain student has joined and left a certain group. There was no such observation
- Thus, frequency of student_id recurring shows when student joined various group
- Student ID (Recurrence) - Frequency
  1 – 13,893; 2 – 344; 3 – 29; 4 – 7; 6 - 2
- This means that there were **344** students that joined two groups

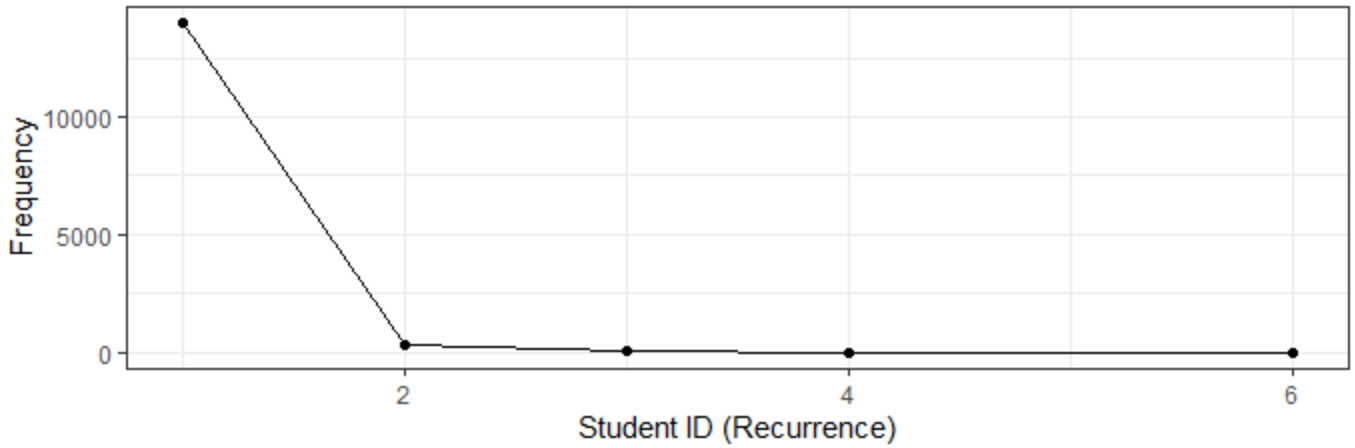
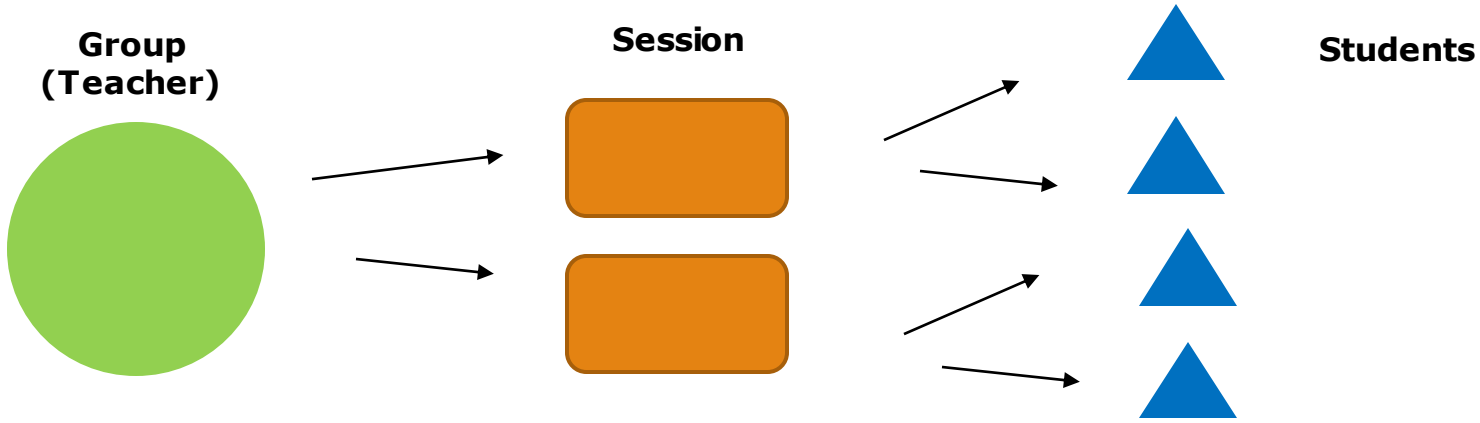


Table 2: Session Log (Data Variables)

SESSION LOG
student_id
session_id
group_id
student_arrival_time
student_exit_time
session_arrival_time
session_exit_time



- Observations: **20,120**
- Observations take place at student's **live session attendance**
- **11,036** unique student_id in this file meaning **14,365 – 11,036 = 3,329** students joined a group but did not attend a live session
- **9** unique group ID meaning there are nine different groups
- Assuming that each teacher has one group and each group has one teacher, there has to be 9 teachers
- There were **341** unique live session
- **Three new variables** were created. *(1) Student Time Spent = Student Exit Time – Student Arrival Time (2) Session Time = Session Exit Time – Session Arrival Time (3) Student Time Per Session = Student Time Spent / Session Time * 100 (Conveys the percentage of time student attended session)*

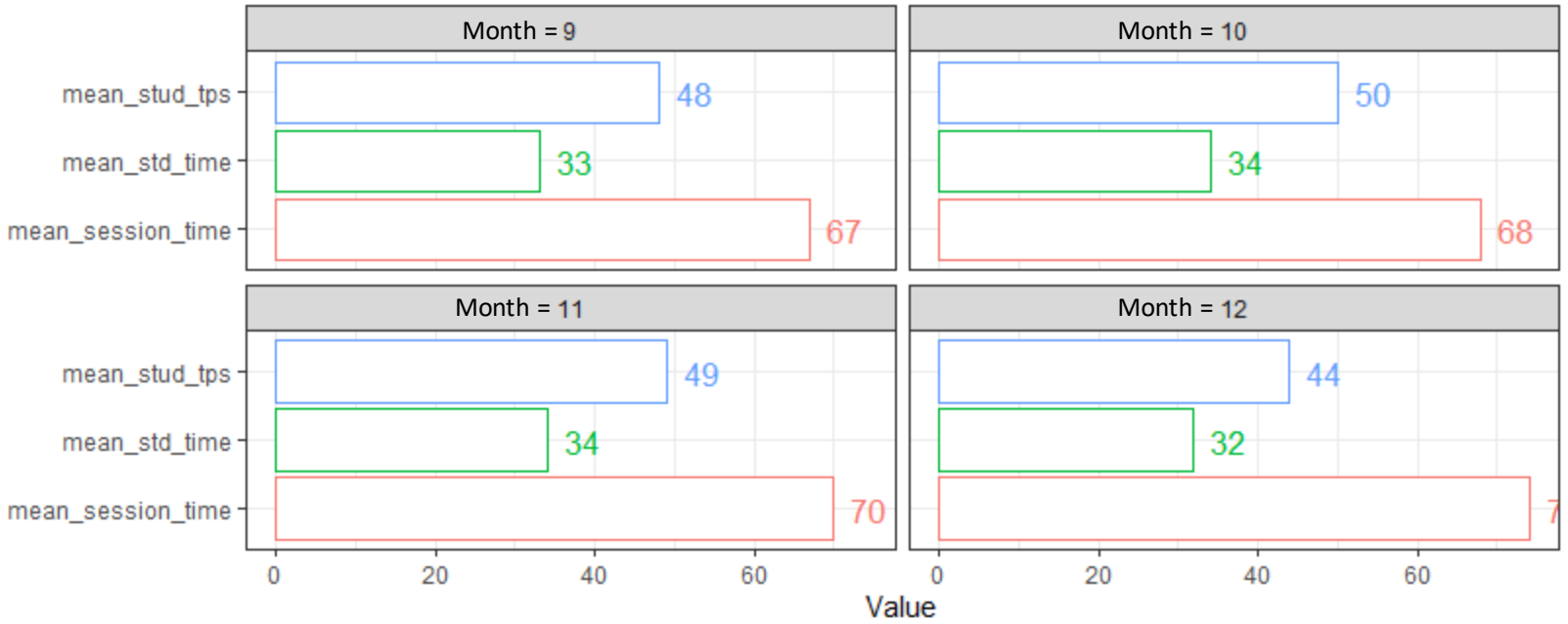
Mean Session Attendance Trend (Overall & By Group)

- Session attendance has remained about 50% approximately but in Dec there was a declining trend of student session attendance.
- There were continuous live session throughout the time for group 1341 while for group 5475 there were sessions just for 2 days in November, 6 days for group 7369 in December

I. What kind of use cases students have on the platform, based on this data?

- Noon's tutoring

Overall Mean Time Spent By Student (Min)	Overall Mean Session Time (Min)	% student attended live session (%)
33.4	71.2	47.4 %



- We see student session attendance clearly declined in December

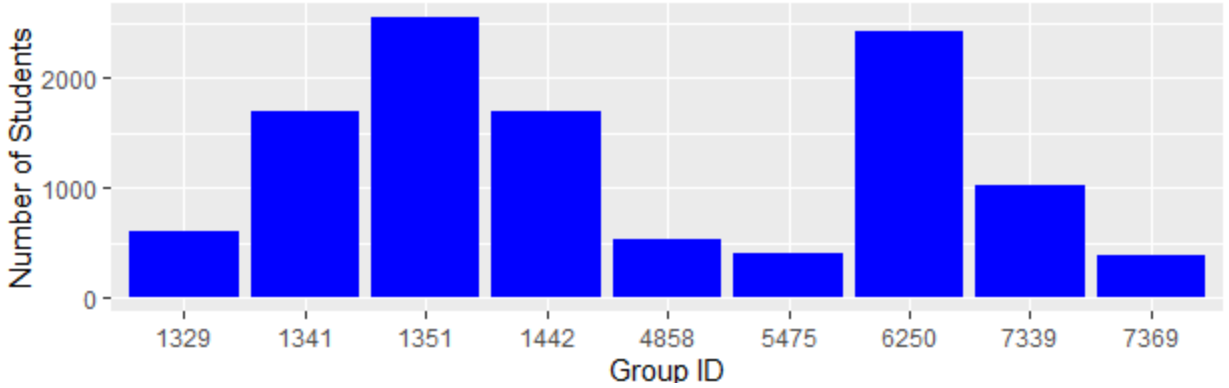
- Mean_Session_Time = Average Time of Session (min)
- Mean_Std_Time = Average Time of Session Attendance by Student (min)
- Mean_Stud_TPS = % of session attended by the student (%)
- Mean_Stud_TPS = Mean_Std_Time / Mean_Session_Time

a mean_session_time
a mean_std_time
a mean_stud_tps

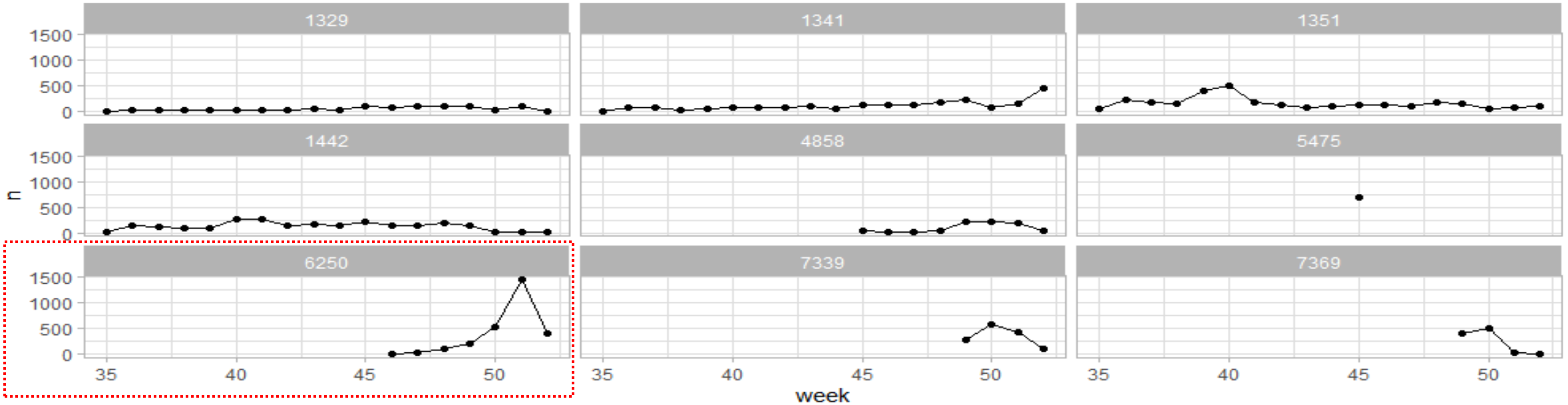
II. What teachers should we promote at the platform? (assume that each teacher has one group and each group has one teacher)

- From graph, it can be seen that teacher from group 1329, 1341, 1442 and 1351 have conducted sessions throughout the time period so they should be promoted for more students to get to know of the sessions and join
- It makes more sense to promote them as the content they produce is more regular

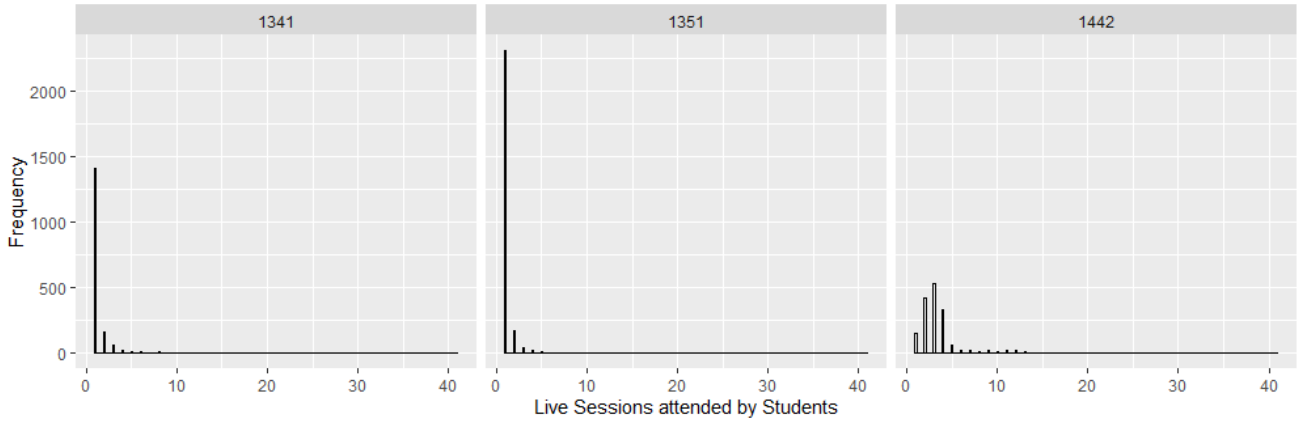
III. What groups would you monetize and why?



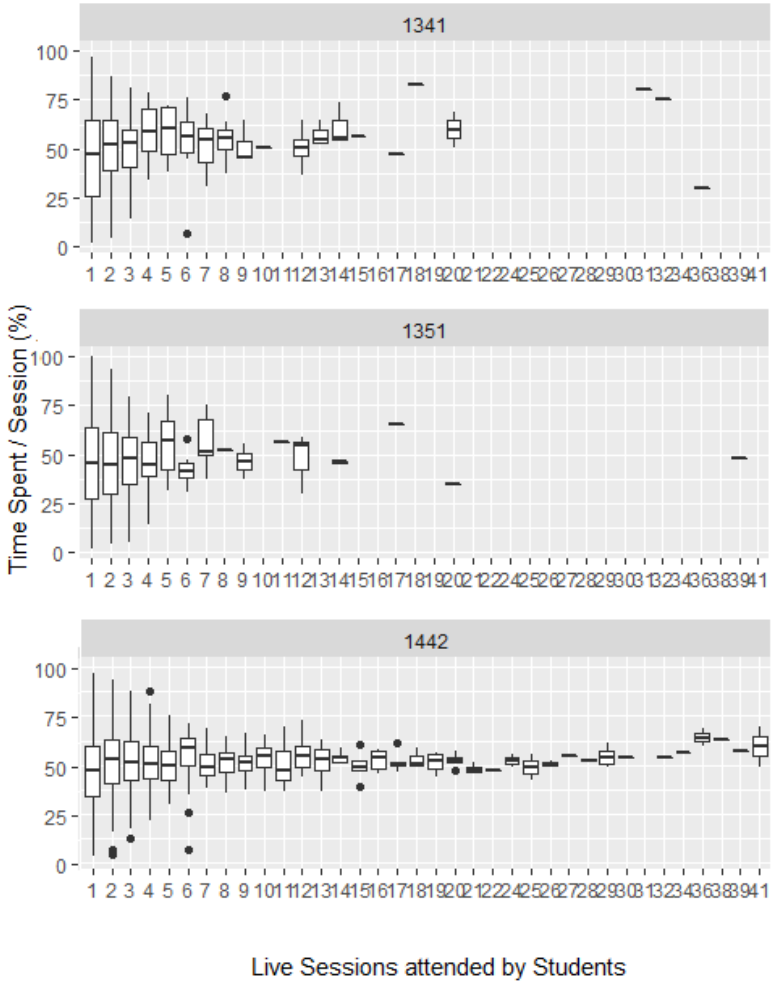
- Group 1351, 1341, 1442 and 6250 are the most popular groups by number of students joined
- It would be giving real value to target audience and thus I would monetize them primarily
- There is a peak at 6250 group as well. But I would not monetize it because most of the students attended session for exam in December and later on the popularity of the group declined



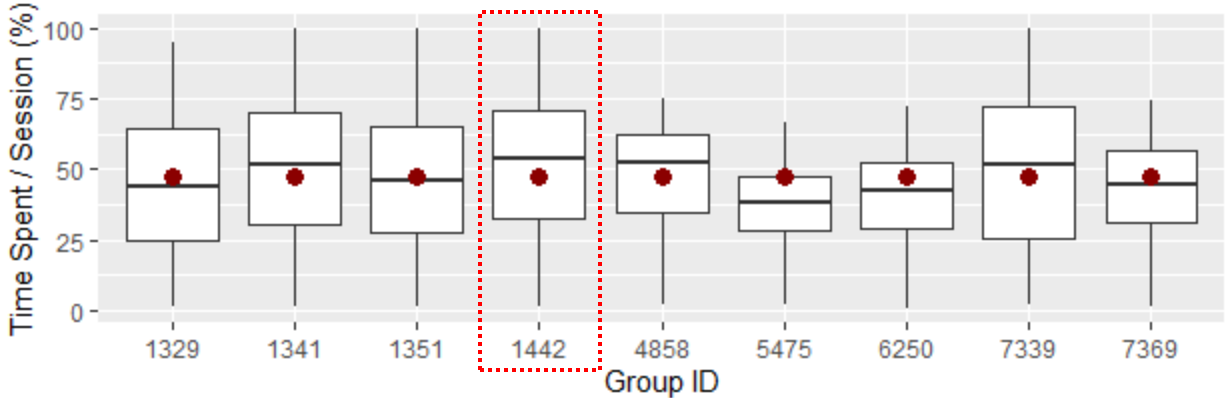
IV. What students would you target for paid groups?



- Students attended 1 – 41 live sessions
- Histogram shows that for group 1341, 1351 majority of the students took one live sessions while for group 1442 it seems majority of students took 3 live sessions
- The median % of session attended by the student was around 50 %
- The trend showed that students who took more than one session from a group attended the session longer (median) which makes sense because only the motivated students came to see another session
- I would target students in these groups based on interval of time that students viewed the session instead of the number of sessions because very few go for more sessions from a group



V. If we would analyze retention, what groups would you learn from the most? What teachers would you interview and study carefully?



- G_join = Number of students who joined a group; T_Attendance = Total session attended in the group; D_Attendance = Distinct student sessions attended in the group;
- It is visibly clear that from group **1442**, total number of session attendance was way more than distinct student attendance. Meaning there were repeat visits. I would say interview and study carefully teacher from this group and try to understand what makes sessions from this teacher worthy for students to make repeats visits

- Red dot is the mean on the box plot and group 1442 also has the highest median value of time spent per session by students