Team Meeting 6

Date:	17 Feb 2017
Time:	3.30pm-6.30pm
Venue:	LKCSB GSR
Attendees:	ALL
Absentees:	
Agenda:	Removal of Duplicates

No.	Details	Action By	Due Date
1	• Whenever someone search for the same thing on separate databases, this must be considered with regards to user search journey. • Found an error whereby users are searching for the same search query on different database, within seconds apart from one another. Could be that the logging was executed with error. **Removal of Duplicates** • Discovered (finally!) how to remove duplicates that have been bugging us since the start of this project: **Steps to remove duplicates** 1. Sort Date by Descending order 2. Group by 'SESSION ID', 'USER ID' (play safe) and 'DATABASE' 3. Keep the top values only (Filter out unique values by taking out the ones with the largest timestamps) 4. Sort Date by Ascending order **Interim Report Discussion** • Report's thought process must include what issues, what did we try, what didn't work. Need include pre-processing. • Report's gap analysis must include: eg. Couldn't find proper inputs and thus couldn't achieve certain tasks. • Can include tentative gap analysis + what we have already done. Just rename it under something else. • Must include what's up for Final report • Full gap analysis to be included in Final report • What I want to deliver, feasibility, data pipeline, how these will benefit my client and facilitate the process. Back it up with foreseen or unforeseen requirements.	All	

| | Tried and didn't pass but still tried. Vision for final presentation. | | |
| | • Need to analyze 40-60% of the database for final deliverables. | | |

The meeting was adjourned at 6.30pm. These minutes will be circulated and adopted if there are no amendments reported in the next three days.

Agenda for next meeting

1. Share deliverables and work again

Prepared by,
Tan Jun Rong

Vetted and edited by,
Bendexter Lim