

TEAM V
ANLY482 SUPERVISOR MEETING
MINUTES 3

Date	18 January 2017
Time	17:20 - 18:00
Venue	SIS Meeting Room 4.3
Attendees	Team V: Andrew, Sarah Supervisor: Prof Kam
Agenda	<ol style="list-style-type: none"> 1. Update supervisor of project with Vanitee 2. Update supervisor of data exportation progress 3. Update and discuss with supervisor on EDA and our findings till date 4. Discuss with supervisor on ways to get Vanitee to share more about their business model

S/N	Item Discussed	Remarks
1	Current progress	<ul style="list-style-type: none"> - Andrew updated Prof Kam about the progress of the project with regards to submission of proposal and submission of all minutes (sponsor, supervisor and internal) on e-learn. The team also have updated the wiki accordingly. - Andrew updated Prof Kam that the team has a Project Log where our action items are captured. Andrew clarified with Prof Kam if the log should only be updated after the tasks are done since at times the team may not anticipated/planned for an action item to happen. Prof Kam advised the team to include planned tasks and if certain tasks are not planned but done, the team should include it in the log as well along the way.
2	Data	<ul style="list-style-type: none"> - Andrew told Prof Kam that the team has access to Vanitee's data via MongoDB and have extracted the data to .csv files. - Prof Kam told the team that JMP can be connected to MongoDB and advised the team to learn how to

		<p>do so. JMP has a query builder that can be used to interact with MongoDB. The team will explore this method as it would be useful to our EDA.</p> <ul style="list-style-type: none"> - Andrew updated Prof Kam that the team will use data from Aug 2016 to Dec 2016, with the end of Dec 2016 as the cut off. Prof Kam is alright with this. - Andrew updated Prof Kam that the column is_test indicates whether a row is test data or not. Andrew also mentioned that apart from removing data that is indicated to be test data, there are other test data which are done in the early testing days that is not indicated as test will be removed too. Prof Kam is alright with this. - As datetimes are recorded in UTC format, Andrew updated Prof Kam that the team has converted it to SG datetime format, GMT+8. - Andrew updated Prof Kam that some bookings in the dataset will have a child booking that will indicate the status of the booking. Knowing this, the team used the data that is a master booking and managed to get approximate the same number of bookings that is on the Vanitee dashboard. - Andrew informed Prof Kam that there are two ways to derive the frequency of of the number of bookings per customer. One is by using the column booking_count in the user table to get the frequency, the other is to use the bookings table to retrieve the number of bookings. - The team is however unsure if the booking_count column is updated each time a customer makes a new booking. The team told Prof Kam that we will clarify this with Vanitee.
3	EDA	<ul style="list-style-type: none"> - Andrew showed Prof Kam some EDA that the team has done. From our findings on the number of manual booking and online booking were almost equal. This raised the question whether Vanitee is generating revenues. Prof Kam mentioned that maybe the team could clarify with Vanitee if they are charging a fee to the beauty professionals for using their application and if they also earn revenue from the manual bookings. - Additionally, when doing this EDA, booking made directly with the beauty professionals are not taken into account. - Andrew suggested to Prof Kam that the team might not look at manual bookings in our analysis since (if) it does not generate income for Vanitee.

4	JMP	<ul style="list-style-type: none"> - Prof Kam advised the team to use JMP to generate charts for our powerpoint slides in future as we have used the charts in Microsoft Excel. - Prof Kam advised the team to use bar charts (horizontal or vertical) in future as rather than pie charts as it is easier view comparisons and labels can be shown on the axis than as a legend. - Andrew told Prof Kam that for every step in cleaning up the data, a new version of the .csv data file is created. Prof Kam said that instead of creating a new version, the team can create new columns instead. - For the chart on the number of manual bookings and online bookings, the team can create a new column stating if it is manual or online and plot the chart from there. - Prof Kam taught the team to plot a graph easily using JMP. The chart plotted is a bar chart of booking status. <ul style="list-style-type: none"> - We can change the value and label fields. Each number (0 - 4) represents a booking status. We can do this by: right clicking > column properties > insert value labels > edit the value and label field accordingly. - During the time of demonstration, the booking status was in numeric form, we can change it to categorical form to do the analysis. By right clicking > column info > change from numeric to character form. The bar chart will be able to be plotted now. - Right click on bar axis > sort by ascending or descending order. - (for future use) Right click > edit > copy graph > paste graph in slides - If we had a column to state if a booking is manual or online, when plotting the chart, it can be dragged to “wrap” to split the chart by the different forms of a booking
5	EDA	<ul style="list-style-type: none"> - Andrew showed Prof Kam the frequency by bookings chart. There is significantly more users with only one booking. - Prof Kam mentioned that even though it does not look encouraging, the team should explore more. - Prof Kam gave ideas such as finding out the recency of users with one booking. If it is a recent booking, it means that Vanitee can that follow up with these customers in order to retain them.

		<ul style="list-style-type: none"> - Prof Kam also said that we can look at how users with one booking spread out in a year. This is EDA on bookings by year/month/week. This will also tell us if they are declining in new customers or not. It means that they should look more into recruiting more customers. - Andrew suggested to Prof Kam that our team would like to look into the duration a user makes their first booking from the time they sign up. - Prof Kam agreed on this and asked us to go ahead. - Prof Kam mentioned that from our analysis and insights, we can pick out if there is a loss of revenue and to detect where their leakages are.
6	Meeting with Sponsor	<ul style="list-style-type: none"> - The team discussed with Prof Kam on how we can approach our sponsor in our next meeting on how they are earning their revenue. - Prof Kam suggested we can show them the chart on the number of manual and online bookings, then mention that based on the number of manual bookings, if the sponsor can retain just 1% margin from it, \$xxx can be the amount they are able to gain in revenue. - Through this, our sponsor might be able to tell the team more about their business model and if they are earning some form of revenue from the beauty professionals through their listing.
7	EDA	<ul style="list-style-type: none"> - Prof Kam suggested to the team that a column to show when new customers joined would be meaningful to our analysis. If frequencies are high recently, it is a good sign to the business. - Andrew clarified with Prof Kam if we should include users affiliated to Vanitee (their employees or family members) as including these users may skew the distribution and they would be more inclined to use Vanitee. Prof Kam confirmed that we should exclude these records.

S/N	Action Item	Action By	Deadline
1	Arrange meeting with Vanitee to update them on project scope, clarify doubts the team have with the data and to find out more about their business model	Andrew	By 19 Jan 2017

2	Prepare slides for sponsor meeting	Andrew, Sarah	By 22 Jan 2017
3	List out the EDA that we want to do	Andrew, Sarah	By 24 Jan 2017
4	Continue with EDA to get more insights	Andrew, Sarah	By 25 Jan 2017