

0404: Supervisor Meeting

Date/Time : 4 April 2018, 2:00PM – 3:00PM

Attendees : Shubhangi, Arushi, Tanushree

Sr. No.	Notes	Actors	Follow up Action
1	<ul style="list-style-type: none">- Cluster Analysis may not be best suited for the problem we are trying to solve for our sponsor. Decision tree might be more useful. Even notifications or mailing will have to be dependent on who has a higher propensity to respond therefore you first need to set a target group of people which decision tree allows to set. Then the recommendations can be developed as a scheme- Executing decision tree on JMP- Preparing the data for decision tree: Set target by recoding the International Cuisine as 1, others 0. Then value order in column info.- Predictive Modelling - Partition - Y response target -> Internal and Others and X factor ->	All	Decide on whether we want to stick to cluster analysis or switch to decision tree. We decided to stick to cluster analysis

	<p>Booking Day, Booking Month, Booking Time, Booking_on_same_day Made on time, Days in Adv, Booking Discount Grouped, Diners Ungrouped. Method - Decision Tree</p> <ul style="list-style-type: none"> - Validation - 0.4 (60% for training and 40% for validation) - Click Go 		
	<p>Research Paper Comments:</p> <ol style="list-style-type: none"> 1. EDA summarize, everything else is fine. 1. What is clustering 2. Methods of clustering (Hierarchical vs. K-Means) and their adv./disadv. 3. Literature Review: <p>Shubhangi pulled out the paper on (Customer segmentation with purchase channels and media touchpoints). Refer to this to structure our own paper.</p> <ul style="list-style-type: none"> - Do a previous literature on bookings - Describe clustering methods - K-Means Clustering (Mathematics + Data Transformation requirements → Should not skew, etc.) 	<p>All</p>	<p>Keep in Mind while writing the paper. Divide the work accordingly for the report:</p> <p>Arushi: Literature review and Introduction Shubhangi: Clustering bits Tanu: profiling bits Everybody: Write Conclusion</p>

	<ul style="list-style-type: none"> - Describe the data - Summary statistics of all the data parameters - Results: How to select the best model. - Optimal CCC Table: Choose the table you use for the range you choose. (If you always touch maximum, then keep increasing until the maximum stops. If maximum is at 12, increase to 20) - Talk about the method you used to get the z-scores, table shows the tables - Then show the full cluster output (Interpreting the cluster. First sort within cluster, and use that to see the highest and lowest and then look across. So look for high-high, low-low.) - Cluster Descriptions: Section, Description, Cluster Label. Very important to label clusters. <p>Table Name and Graph call Figure.</p> <ul style="list-style-type: none"> - Usage of Model: Advised to calibrate the model yearly to ensure that the whole thing is up to date to 		
--	--	--	--

	<p>keep up to date with the customer behaviour.</p> <ul style="list-style-type: none">- Conclusion: Summary. We received this, did this. And then mention, that since it's unsupervised use this.- Advised to calibrate the model yearly to ensure that the whole thing is up to date to keep up to date with the customer behaviour. <p>SMU Library - Customer Segmentation Using K- Means etc - Follow style for Lit Review</p>		
--	--	--	--