

Meeting minutes

- Team SkyTrek

Date (dd.mm.yyyy)	06.01.2016
Attendee(s)	Jedaiah, Aseem, Huy
Absentee(s)	-

Contents

Item	Discussion point	Remark
1	<ul style="list-style-type: none">- Project Key Goals - Topic Modelling Timeline - Google Trends Timeline	<ul style="list-style-type: none">- One key goal is to validate the popular trends over the year. 2 methods will be used to do this:<ul style="list-style-type: none">o Topic modelling timelineo Google trends timeline - to generate topics based on snapshot taken for a given duration- URLs will be binned with the topic which the algorithm determines has the highest probability- Create timeline to see how the topic frequency varies throughout the year - Check if google trends has API to allow extraction of trends based on user input. If not possible, we can also look at extracting information based on web scrape easily since URL is constantly being updated based on user inputs
2	Content Theme Classification	<p>Skyscanner is classifying their contents into the following themes:</p> <ol style="list-style-type: none">1. Inspirational (top festivals, top foods, traditions, etc.)2. City Guides3. Domestic/Local (Singapore related)4. Trending (Game of thrones, CNY, etc.)5. Deals (Discounts, Prices, etc.)6. Practical/Tips7. Product (Skyscanner related feature) <p>We are to study which kinds of article perform the best in each content theme.</p> <p>URLs will be manually sorted into categories based on tokens extracted from URL topic title</p>

Actions

Task	Assigned to	Status
Put topic modelling and google trends research and implementation into Gantt chart timeline	Jedaiah	Created
Input meeting with client in gantt chart for: <ol style="list-style-type: none"> 1. 25th Jan 2. 2nd Feb 3. 11th Feb 	Jedaiah	Created
Create flow chart documenting process how skyscanner team plans for the next quarter's content postings. To send it to client for verification	Aseem	Created
Test Google API public holiday pulling	Huy	Created
Create a mockup for the following visualization: <ul style="list-style-type: none"> o To plot Unique page views views (count) vs Week of article publishing o Clicking this particular bar of interest (week has high count) would drill into the next bar chart (Number of unique page views, vs identifying URL) o Further clicking on the URL would reveal characteristics of the article (e.g. number of images, videos, etc, bounce/exit rate, etc) 	Jedaiah	Created
Create a mockup for the following visualization: <ul style="list-style-type: none"> o Identify key features of popular posts (MLR against unique page views) 	Huy	Created
Create a mockup for the following visualization: <ul style="list-style-type: none"> o Where do they come from (On the whole country news page level, where do the traffic come from) o How long do they spend at the page (average time on page) o Do they visit other pages? (bounce rate vs exit rate) o How effective are the recommended pages? (If GA does not track the unique page views, then this analysis is not going to be implemented) o [Just to prove a logical point] Correlation test between number of shares vs unique page views 	Jedaiah	Created
Retrieve URL share count	Jedaiah	Created
Deal with proposal segments: <ol style="list-style-type: none"> 1. Sponsor and Background Information 2. Motivation 3. Objectives 4. Data 5. Scope of Work 	Aseem	Created
Characteristics to scrape from article: <ol style="list-style-type: none"> 1. Number of words (remove stopwords) 2. Number of links 3. Images 4. Videos 5. Number of shares of article (by Jed) 	Huy	Created

Questions for Prof Kam: 1. Ask Kam about requirement to post sample date. Concern: data confidentiality 2. What is required for scope of work? (different from methodology)	All	Created
-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----	---------

The minutes were prepared by: Jedaiah Tan and Nguyen Viet Huy