LDA stuff

If new set of data comes in -> how do you make sure that the topics are the same?

The data are roughly the same,

Will have a backend thing a not?

Is it expected?

After passing it to them, are we required to give any support?

Make it black and white

After showing sponsors, what was their feedback?

Hard to do prescriptive cause very hard for us

Just stay with eCommerce on competitor

Try to eyeball it

Even if you answer their marketing and promotion, it will be quite hard for them in future

Average ratings compared across brands

Group product by the price range

Loreal: between different category, the price range might be too high

Innisfree what are the top rated vs Loreal what are the top rated

How did we come up with these subcategories

Sub Category have lip and lips – cause of the way the website categorize their product

Loreal: cannot see the most popular products

The issue is how do you organize this whole thing

The most we can do is subcategory – product

Make it two panel and list the product

Number of review

Dynamic

Ask them to select themselves

She say she like the promo code across the month

Sorted by the

Number of promotions and number of orders have correlation

No promotion applied on check out by right should be straight (then shouldn’t run promotion)

Types of category that uses the product

Which product uses the voucher codes

Please change the header

Think of adding smth to add

Whatever you have done for sentiment, just do the same for the competitor and get the labels out.

You will need to submit a report (around 10 pages but there are no rules) – 40%

1. Project deliverables (project management), completed dashboard and all the other programmes
2. Evidence of sufficient planning
3. Who have done which part
4. Testing within the team
5. At least put at appendix if no space
6. How to read your project schedule
7. Need to talk about lifecycle frame?
8. How many sprint we do if like agile
9. Need gantt chart
10. Technical (data processing model)
11. Solution design blah blah is dashboard
12. Why we design our dashboard this way
13. But here you will already state the objective for each of the dashboard
14. User guide for … must be attached and then add in the findings.
15. UAT and schedule

Presentation (7th April – Afternoon) – 30%

Track Position – 30%

Faster wrap by week 13

For stopword all remove, see why those that are predicted wrongly , then see if can explain

Or just use a super simple stopword list: is a …..

For the count vectorizer, don’t really need to explain but you can try

Which are the words that are the support vectors, then you can see which is positive or negative

Use sklearn

Report more technical, Presentation more business-ish

Story telling part is there but

1. Technical side (coherence score and all)
2. There’s a lot of gap
3. How do you show that you know what you doing?
4. Just now what she say for stopword
5. Coherence score was not used properly
6. Vizualisation
7. Don’t know what is the main message that you can’t view
8. Total number of reviews (the log one)