

### Team Name:

Shopaholics Anonymous (SA)

### Team Members

Saleha Siddique

### Project Proposal

#### Text Analytics & Visualization of Amazon Customer Reviews

Customer reviews play an important role in influencing purchase decisions in e-commerce. Many products have 100s of reviews though, and going through all of them takes too much time. I would like to analyse the reviews of a product to count the most frequently occurring words, and display them in a chart (similar to how we did it for the course scraping exercise).

Another chart could display the most common words in positive reviews, and the most common words in negative reviews.

Nice to have functionality could include:

- Analysing the sentiment of particular phrases and visualizing that
- Not all reviews are equal – so it would be good to weigh those that more people have found useful more heavily
- Per product different customers will find different things important – allowing the customer to choose what they find important and filtering reviews based on that would be good

#### Modules/Technologies that I plan to use:

I have broken these down by the problems that need to be solved.

1. *How to get the customer reviews* – this should be dynamic – i.e you enter a product URL and the program pulls out the reviews for you.

While Amazon has an API to return details about a product, the API only returns a URL to the reviews – so I will have to use curl to scrape the reviews.

The API could be useful for displaying the picture of the product (a visual validation that it fetched the right product's reviews). There's a JS wrapper for the API available at:

<https://github.com/livelycode/aws-lib>

2. *How to analyse the reviews:*

Node.js + I'd need to shortlist from NLP libraries available on NPM:

<https://www.npmjs.com/browse/keyword/nlp>

3. *How to visualize the analysis:*

- d3

## Biggest Risks

- Timelines on this project are too tight
- I'm new to javascript so even simple things take me longer
- d3 in particular is quite complicated to use
- Also not sure how to "dockerize" at the end as docker doesn't install on my Windows PC