**Product review analysis**

We kindly ask you to prepare a case at home. As your role is all about working with data, can you answer at least two of the questions below using Amazon Review data?

Case Questions

1)          Is there a relation between reviews from products which are bought together?

2)          What is the relation between the reviews and the helpfulness?

3)          What is the review behavior among different categories?

4)          Is there a relation between price and reviews?

5)          Optional - Any other hypothesis you think is interesting, as long it would have business value for Amazon

Data

The data is available at:

•            Reviews: <https://s3-eu-west-1.amazonaws.com/bigdata-team/job-interview/item_dedup.json.gz>

•            Metadata: <https://s3-eu-west-1.amazonaws.com/bigdata-team/job-interview/metadata.json.gz>

To extract and load the data into a data frame you can use the sample code and instructions on this page: <http://jmcauley.ucsd.edu/data/amazon/links.html>

Results

Please present your**analysis and suggestions with actionable insights** as if you were to present to management level audience. You can use whatever tools/languages in order to solve the questions. Please make the product in such a way so that you are sure it could run in a productionised setting. You get 20-30 minutes to present your results to the team where after they will ask detailed questions. Furthermore, please send in your code and presentation via email at the **latest by Tuesday 12th of September at 15.00 CET**.