Milestone 1
31/10

Each teammate chooses a task to work individually on. We strongly believe that task 3 will be the most challenging and we plan on all focusing on it once we will be done with task 2 and 3

05/11

All tasks concerning **data extraction is finished** : we have cleaned the data to work with.

Task 1.a and 1.b are done. We have enriched the definition of task c. with possible reformulation.

Task 2.a and 2.b are done.

Task 3.a. we have tackled a good load of it, we have defined the good way try to detect fraudulent reviews. We have an initial pipeline that allows to score reviews (on an arbitrary scale) to reveal how likely they are to be fraudulent. We tried to find a clever way to get the origin country of the reviewer (might scrap the user profile but not all of them have a country indication) and we assessed the feasability of it.

Initial Data Exploration
31/10

Research tasks are already well completed.

Task 1.c. is finished and we should discuss whether the findings are relevant to be put in the final presentation. We do not yet know if there is such a price gap that can influence reviews or not but think that users still face a tradeoff when it comes to e.g. two TV's where one is premium with excellent reviews and the second is a cheaper option with average reviews.

Task 2.c. is finished or nearly finished we will have a discussion on the findings and the possible remaining ideas on what parameters are of interest for the helpfulness of a review.

Task 3.a. is done.

19/11

All **tasks are finished and only final merging and proofreading are necessary**.

The full team switched to Task 3 after finishing the details of 1.c and 2.c in order to have one student on each subtask b. c. and d. We discuss the findings and how we want to tell a preliminary story in the notebook.

26/11

Milestone 2
28/10

All the **merging has been successfully conducted** and we can submit the notebook.