Python related:

Time series

Michel:

Don’t focus on ETL, warehouse will be star schema

To-do: Manufacturer table with id alone

Ideas: Average ship duration

Ideas: Around 6 countries

Ideas: Customer reviews per country

To-do: Split columns for number available in stock

To-do: Average rating

To-do: Category and sub category split columns

Ideas: Look out for products bought together

Ideas: It’s a plus to use NLP library for customer reviews

If price doesn’t exist, we should not consider it into aggregations