

Questions for Domain Experts - Initial Data Inspection

General request: a Data Dictionary could be really useful, providing information about each column.

Customers

- What is the difference between `customer_id` and `customer_unique_id`?

Geolocation

- How do we actually get the `latitude` and `longitude`? Is it by the full zip-code? Is it by IP? Something else?

Order Items

- What is the currency? Is it USD? Is it BRL? Is it always the same currency?
- Should there be any price limits? Are prices somehow being monitored?
- Is there any monitoring for illegal or abusive products / goods?
- Who dictates the freight price? Is it a contractor (like freight-forwarder)? Is it the seller?
- Is there no flexibility regarding freight, like lower freight charge for slower delivery or vice versa?
- Is delivery only by land transport or also by shipping?

Orders

- Does it make sense that the cancellation rate is less than 1%?
- Are we missing data for Sep 2018?

Products

- Is product photos data available?
- Are product details (size and weight) self reported or verified?
- Is it not mandatory to upload a photo for each product?
- Are photos provided by the seller, buyers or both?

- Is there a cap for the number of products per seller?
- Is there quality/feedback/review data for products?
- Are there any limitations to product sizes or weights?

Sellers

- Can anyone become a seller?
- Are there any verification processes for sellers?