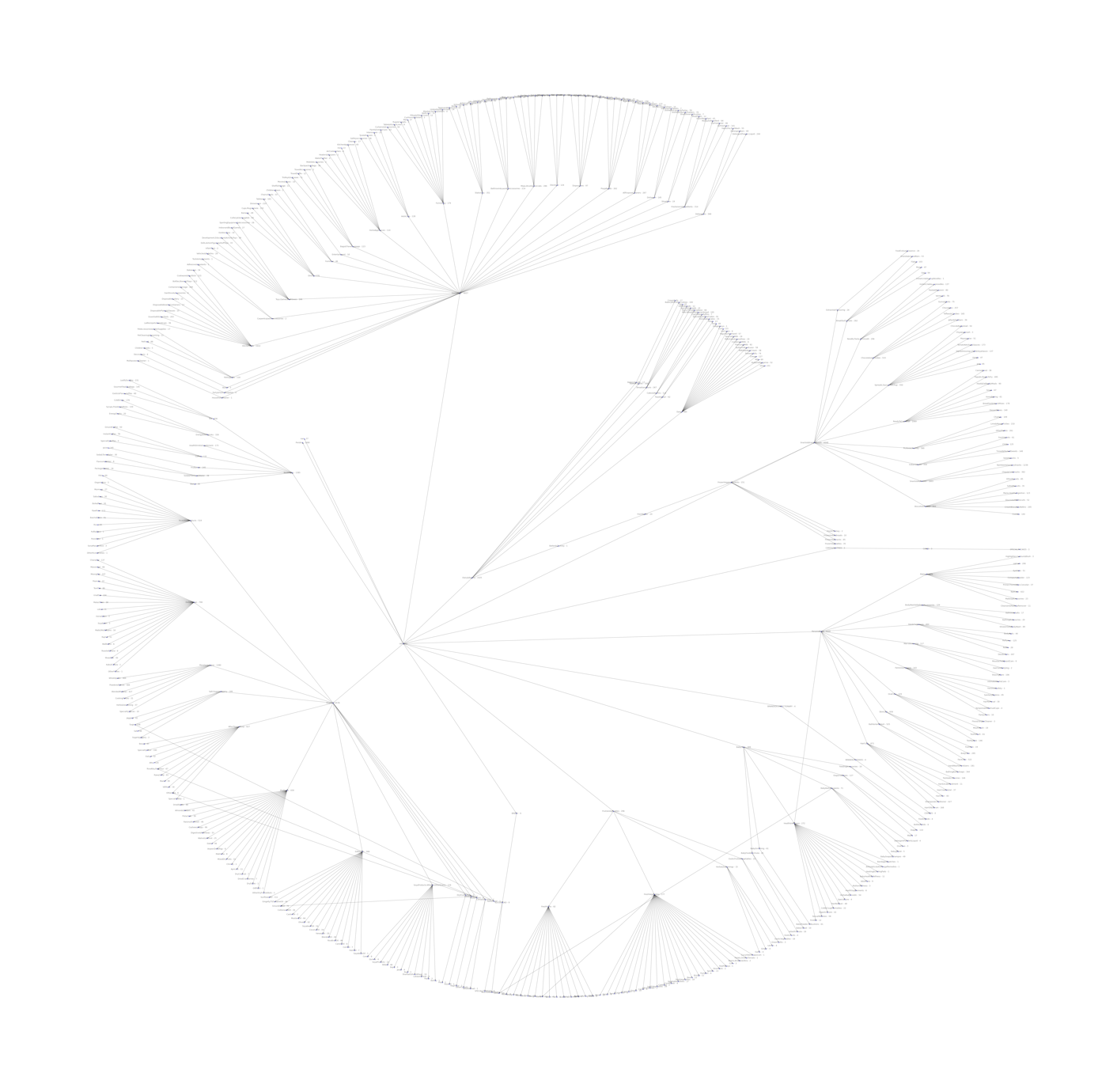
**Hierarchical Tree Representation of Categories(L1, L2 and L3) for AJIO and Jiomart**

Below is the radial BFS tree diagram of L1, L2 and L3 categories for Jiomart catalog data.

Main purpose of this tree is visualize the product distribution across different categories and sub-categories.(Below image is not clear due to larger data density. Can be shared as separate image file).



Although image is not clear, but we will go through some key observations from this image.

Most of the radial graph is organized except few random edges inside the graph.

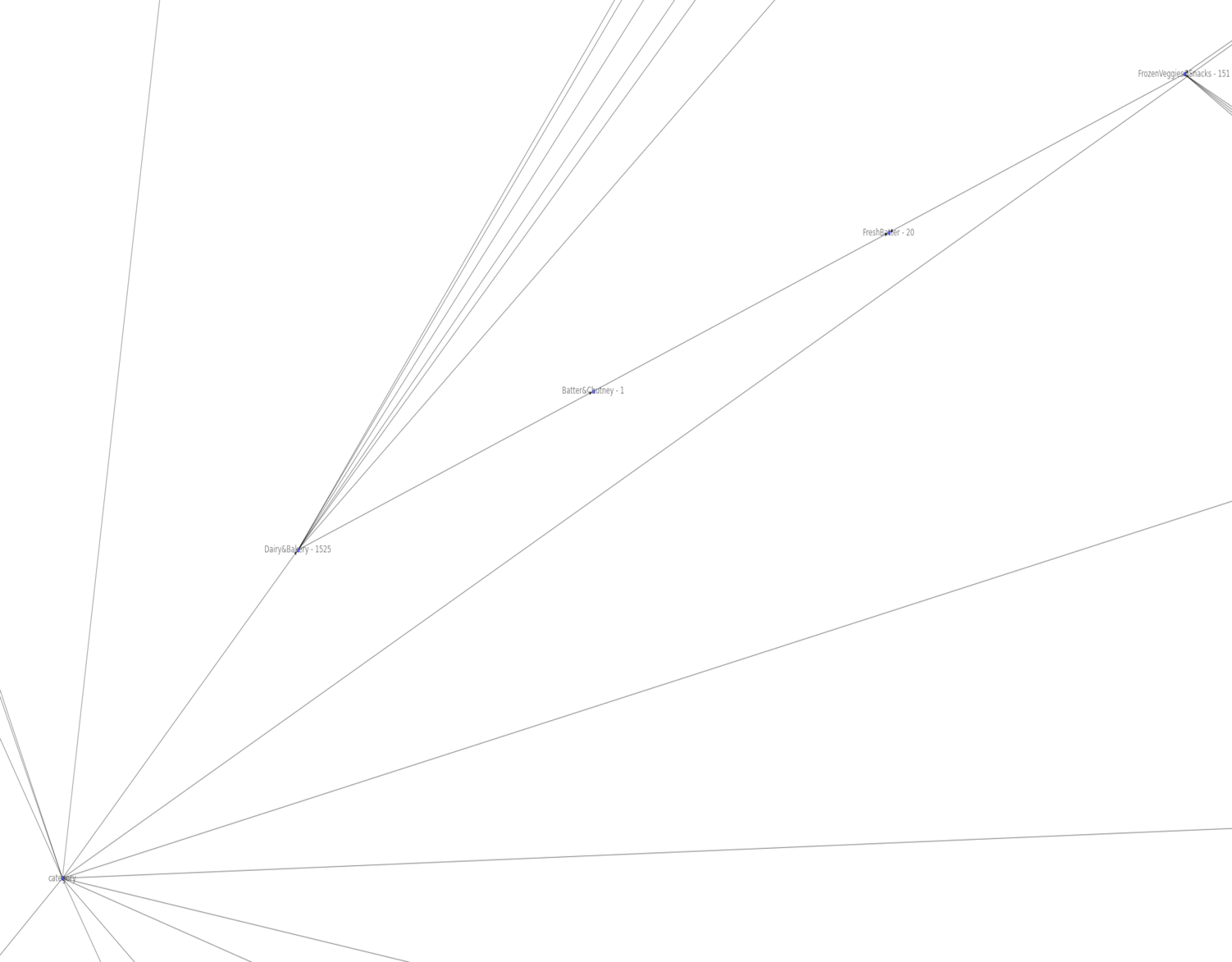
We will cover each of them one by one.

“Category” is the central node and all L1 categories are connected to that. All L2 categories are connected to their corresponding L1 and all L3 categories are connected to their corresponding L2.

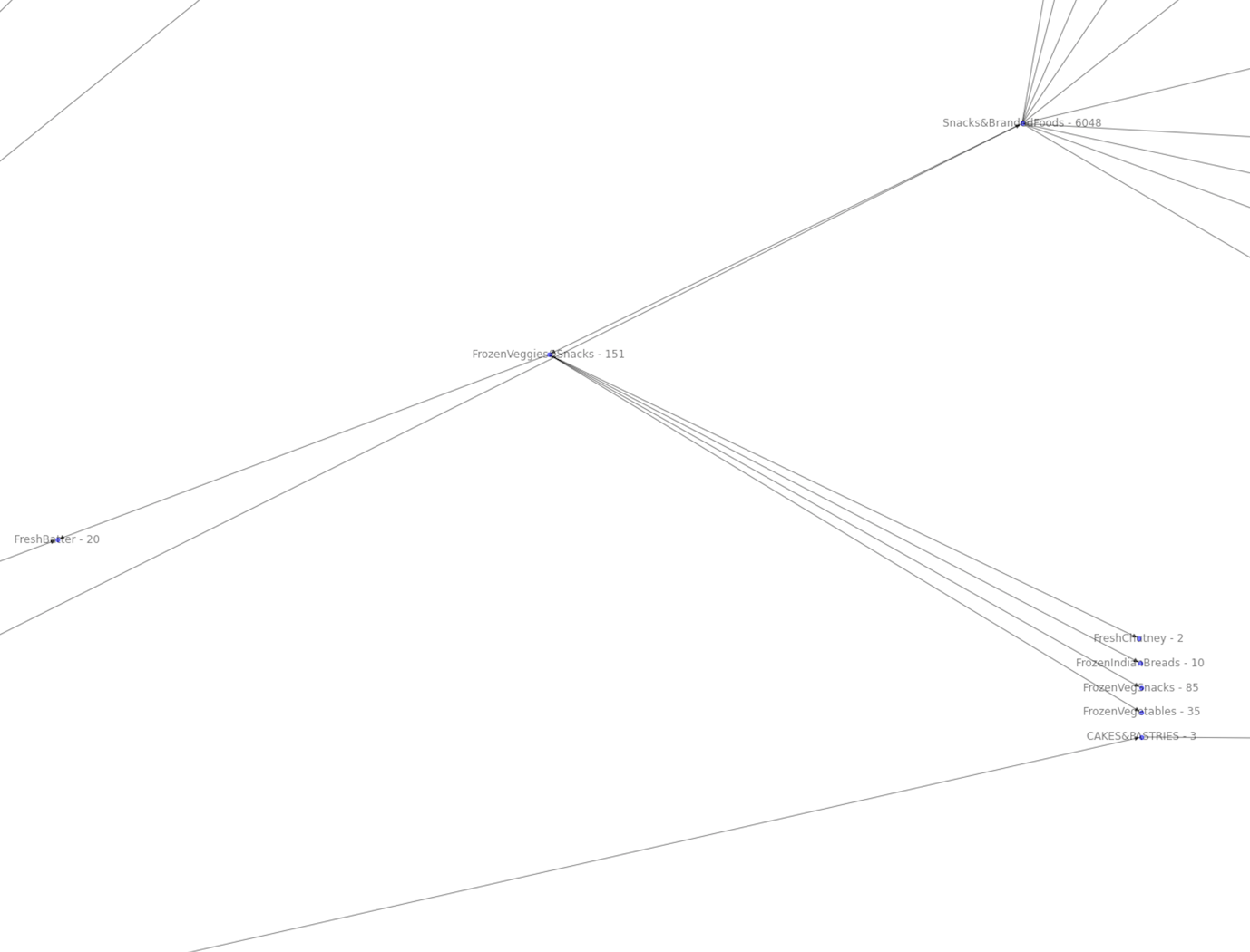
“Category” 🡪 L1 🡪 L2 🡪 L3

Let us go through some small observations.

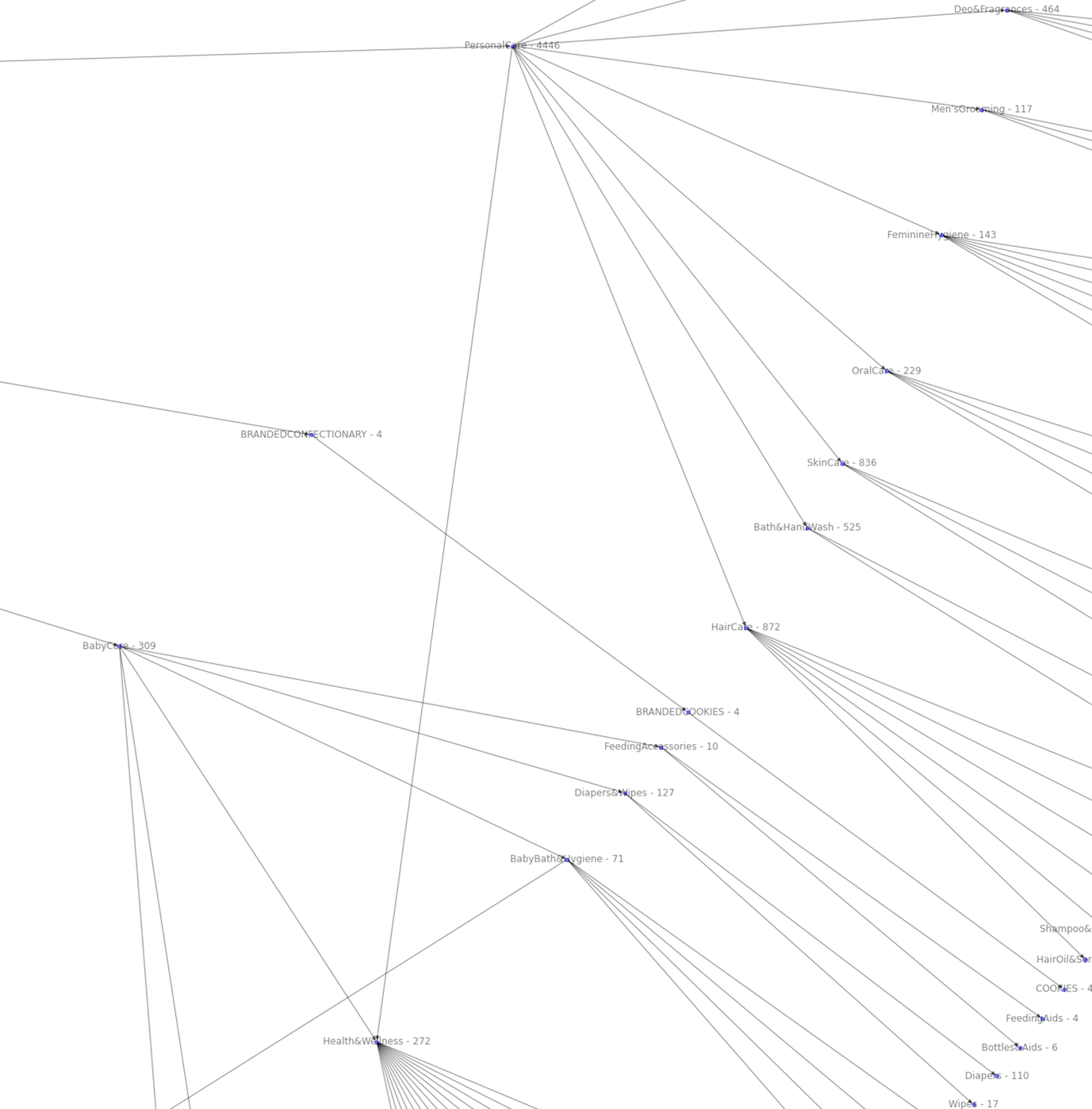
1. Batter&Chutney(L2) is connected Dairy&Bakery(L1).
2. Fresh Batter is connected to FrozenVeggies&Snacks.



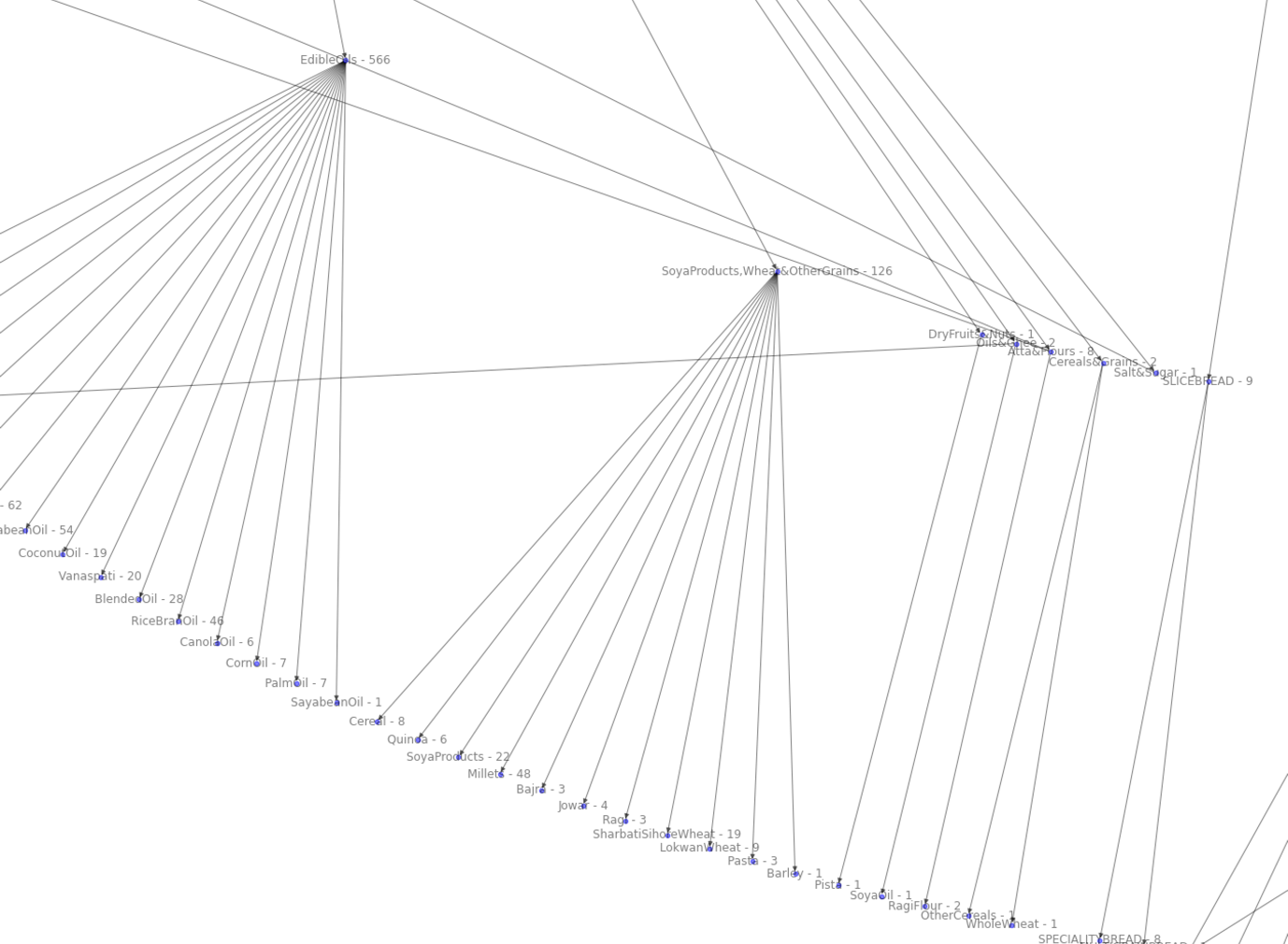
1. FrozenVeggies&Snacks is present in both L1 and L2 categories.

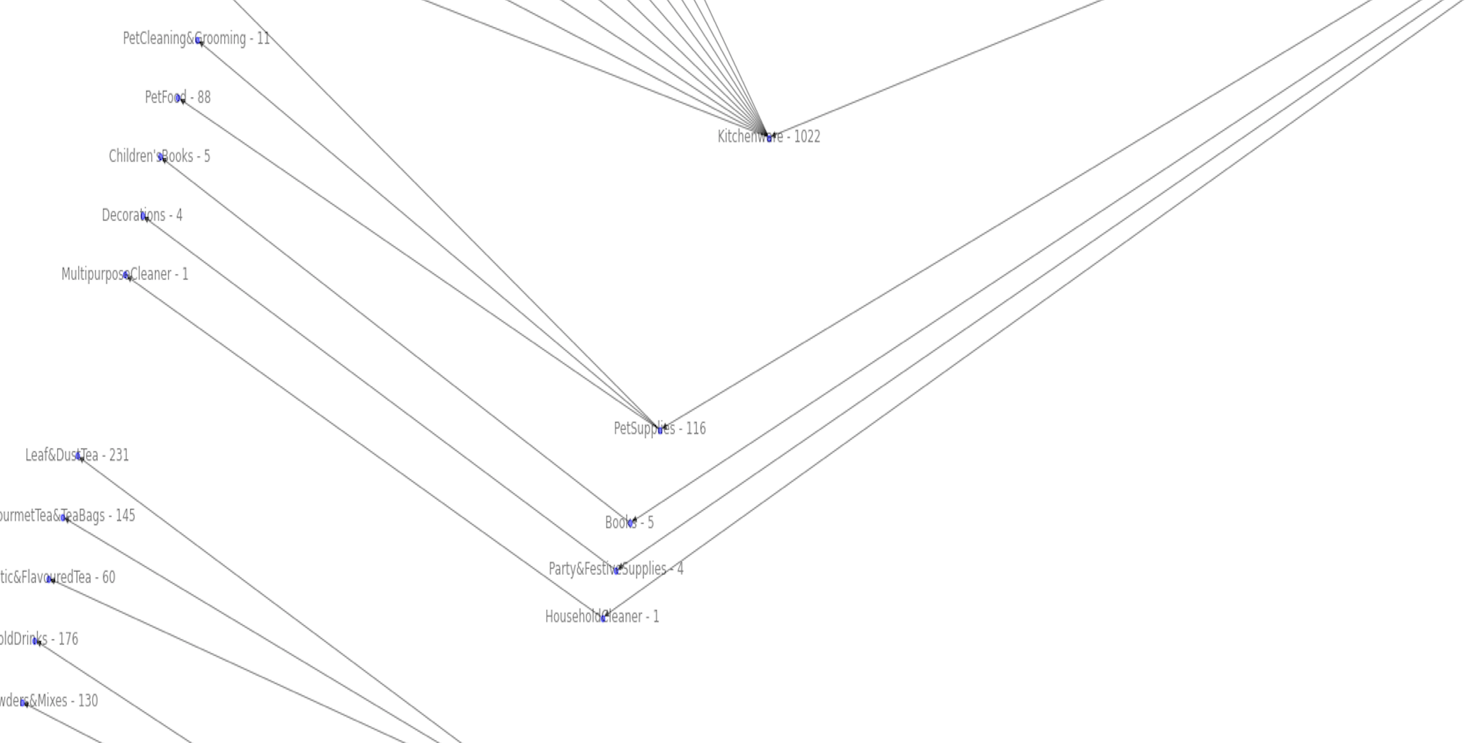


1. Branded\_Confectionary is a different L1 category although this category already exists with different name (can be renamed).

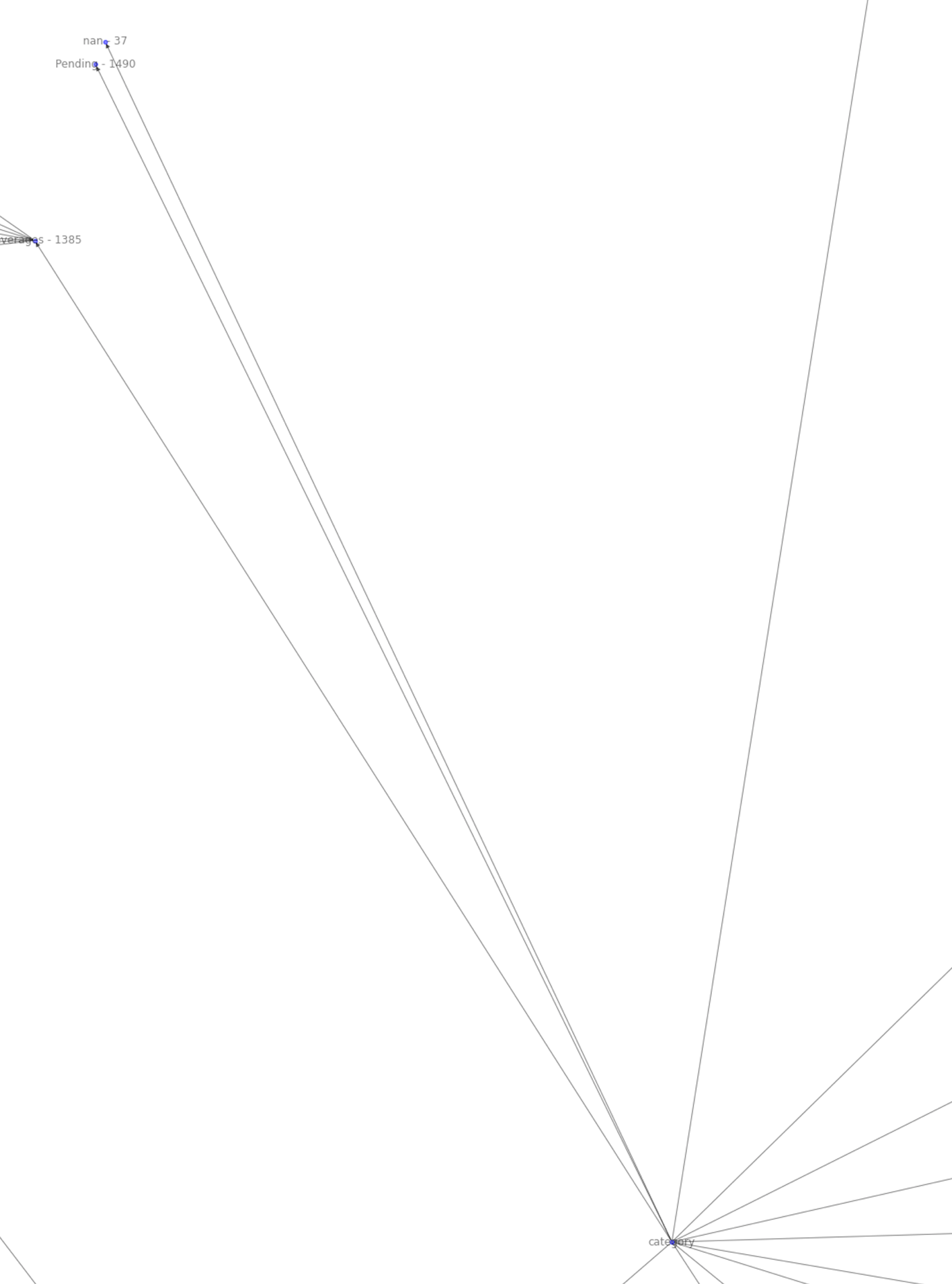


Same redundancy problem exists here also. Different categories like “Dry Druits&Nuts”, “Oils&Ghee”, “Atta&Flour”, “Salt&Sugar” etc. have been created separately although they already exist.



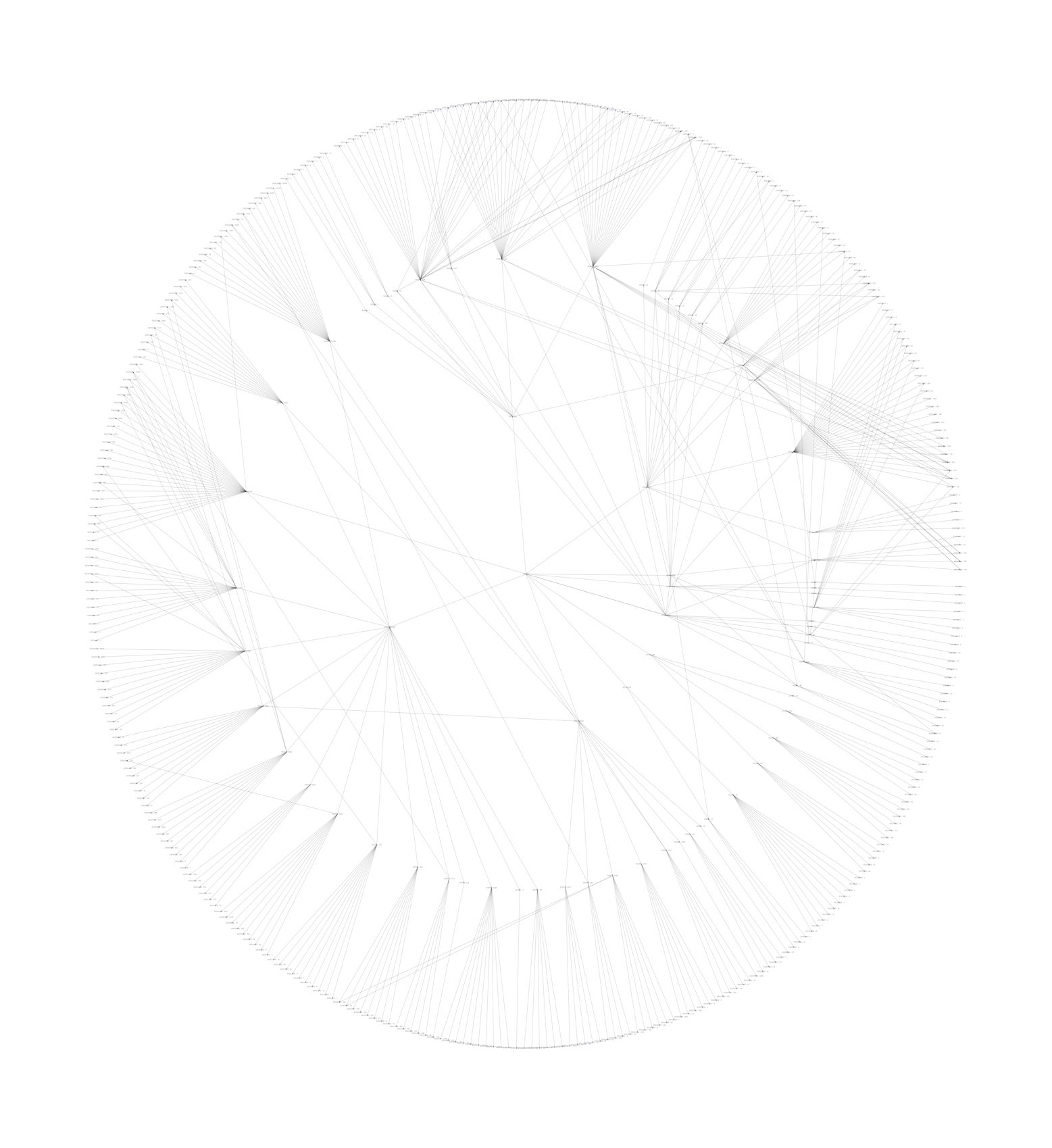


1. Pending Catgeories



Below is the radial BFS tree for AJIO categories.

Some noise can be seen here also but will be clear with the integration of mapping of category\_id with their corresponding category names.



# Product Catalogue Health Checker

Main steps in checking health of the current uploaded product catalogue will be :

* Identification of missing or “Pending” marked L1, L2 and L3 categories for Jiomart and Ajio(can be used for other verticals also) and also auto-suggest their corresponding categories based on the product description.

L1, L2 and L3 categories are must fields as they are the first door to product search and their relevancy.

* Identification of other missing fields.
  + - Description – Missing description can reduce the user trust over product and hence its presence is very important.

Description auto-generation can be done using copying the description of most similar product (if similarity> threshold) .

* + - Images, Price and Brands – Missing images can again reduce user trust and should be present there. Health checker will identify and suggest using similar algorithm like description.
    - Quantity – Missing quantity can confuse user during purchase and hence, that’s an important field. Quantity can be estimated based on the same product value for other quantities considering their brand, quality details and product type.
    - Merchants Name and other fields
* We can also check the validity of each of these fields if these are actually correct or not. It’s kind of verification of each field for a product and identification of any mismatch or error in the catalogue.
  + - One type of verification is kind of produced in the previous section of document related to L1, L2 and L3 categories.
    - As we can see the discrepancies and redundancies in the product categories. Many categories are tagged to multiple super-categories while many of them are present in both of them.

Red indicates that these are must fields and they are affecting business seriously and hence, they should be fixed ASAP.

Orange indicates comparatively lower priority fields.