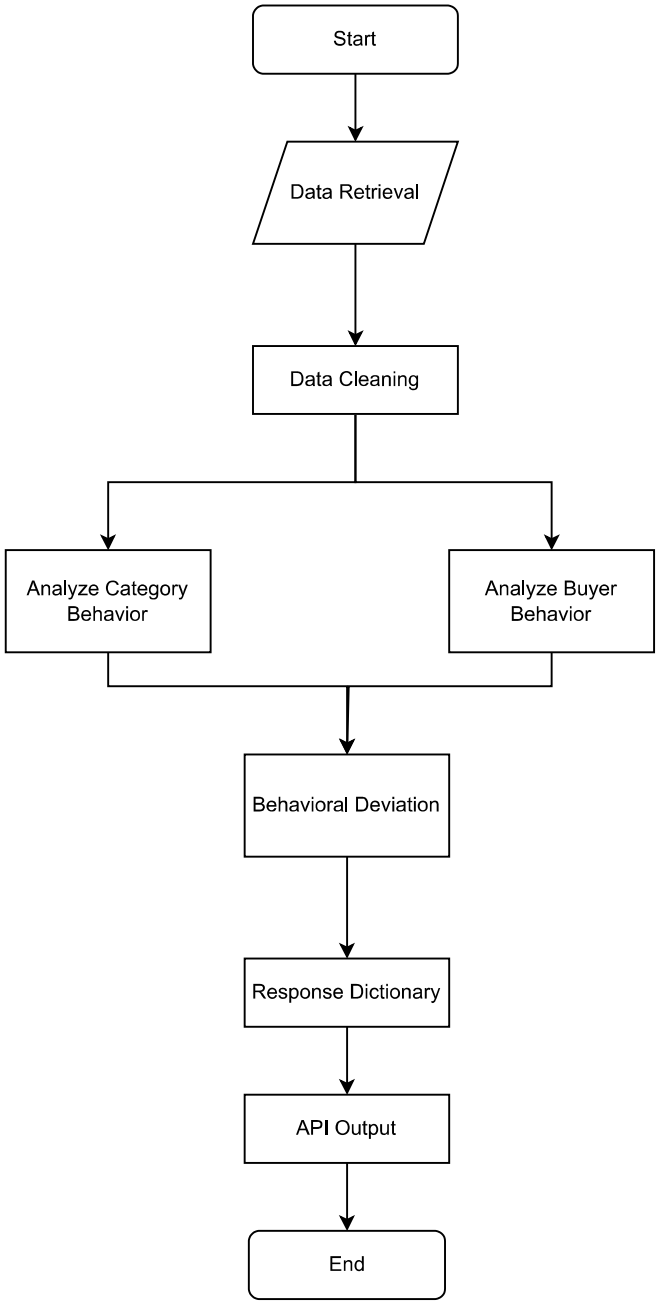
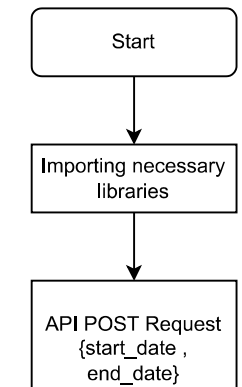
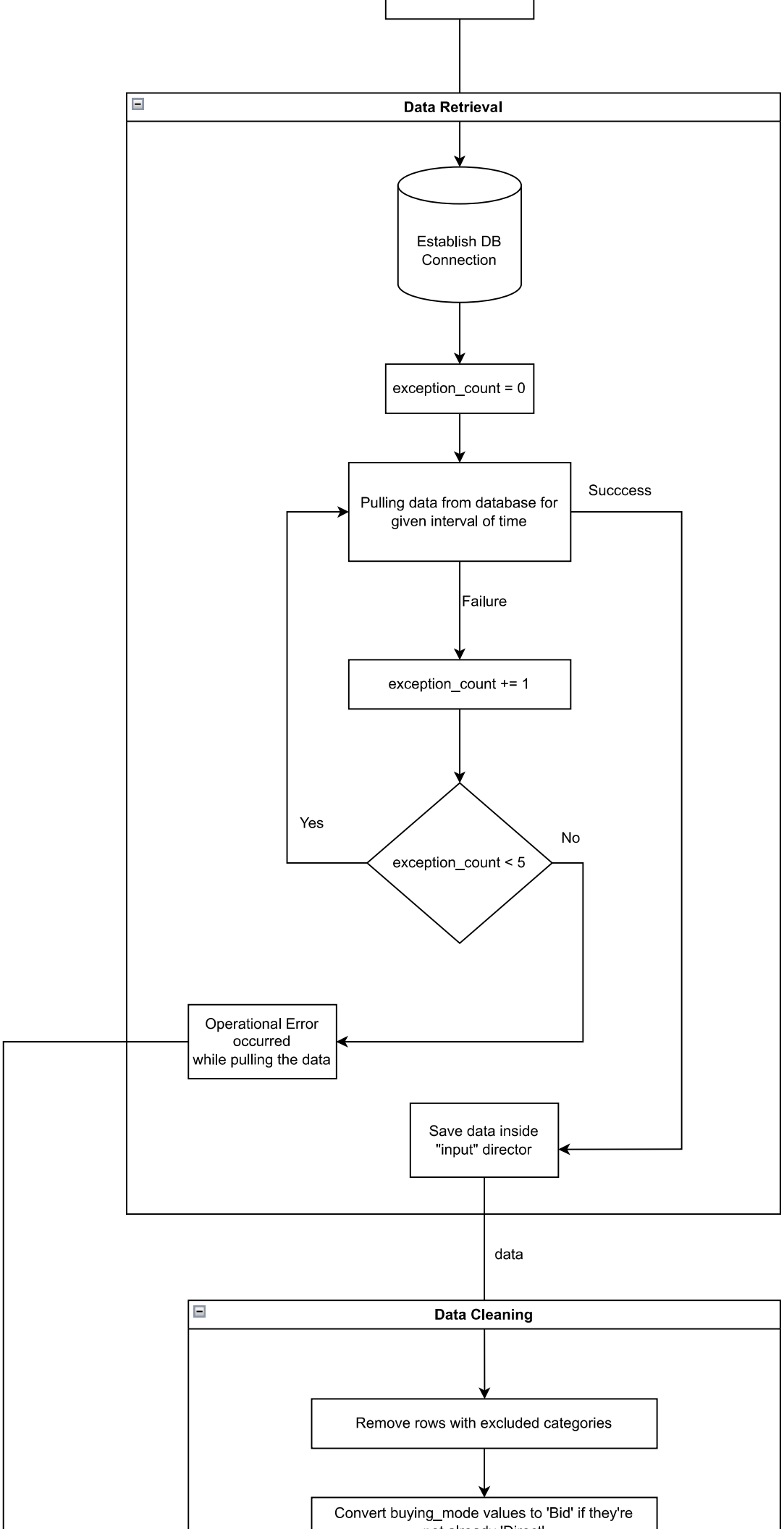
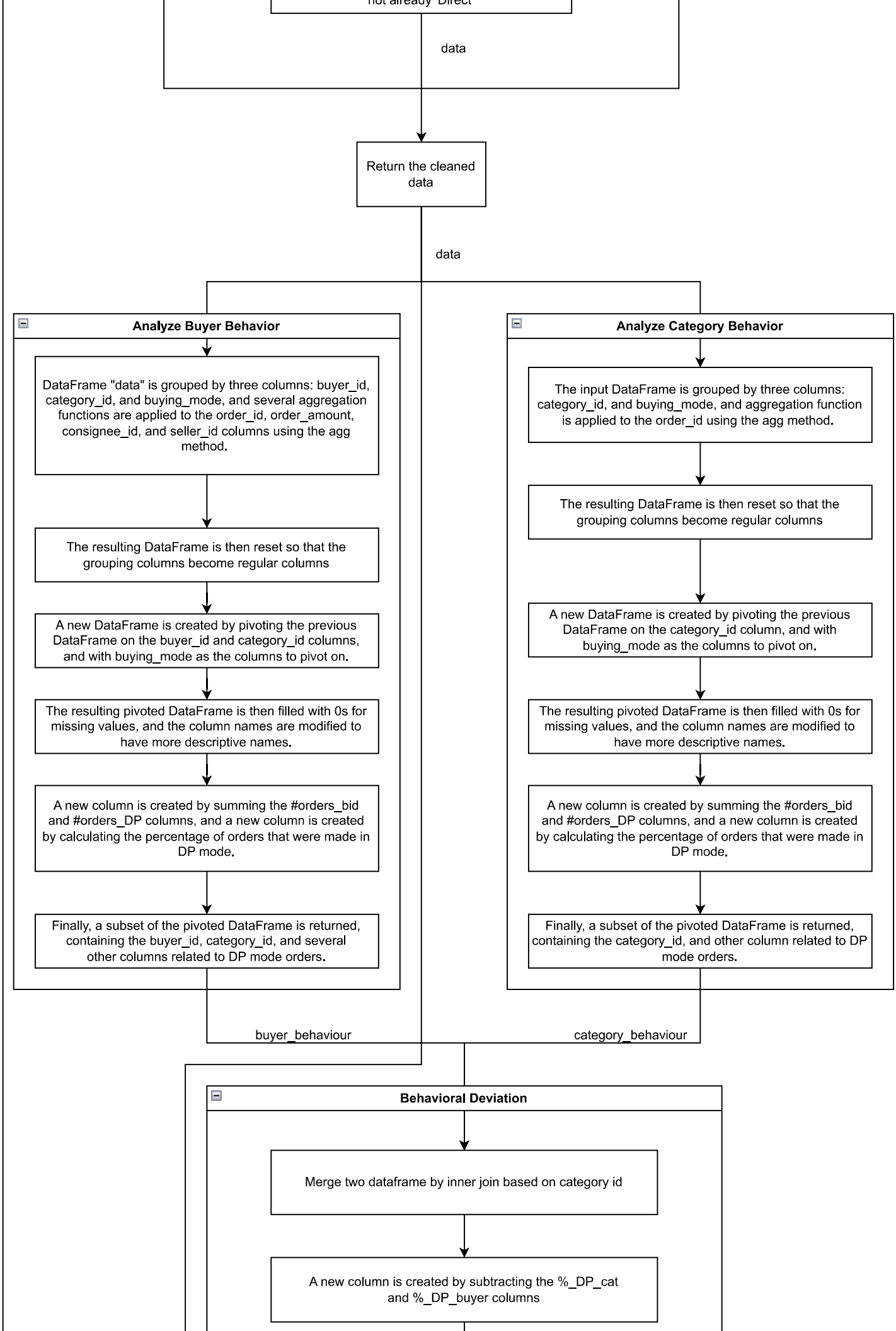
AoA_FD

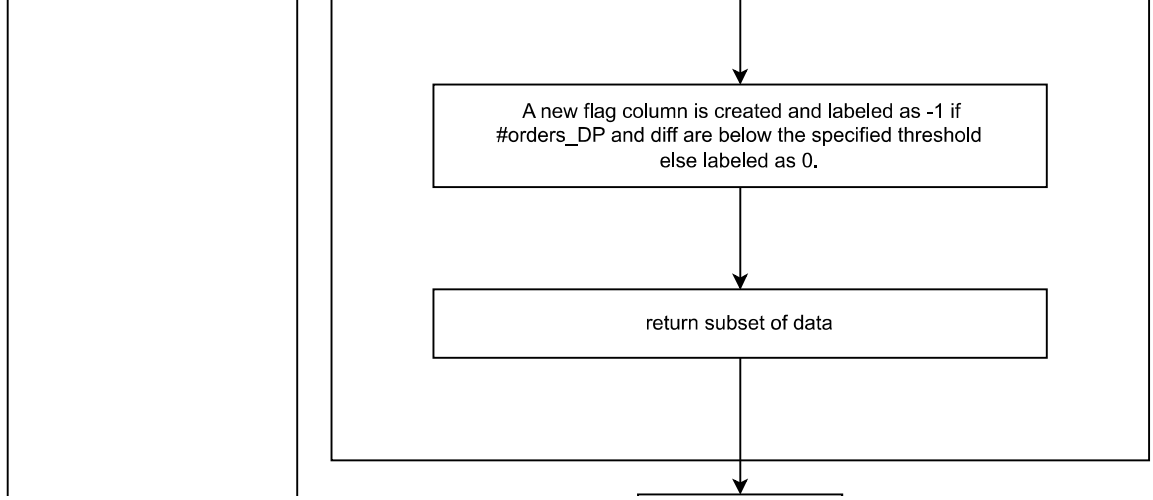


AoA_Detailed_FD




Data Retrieval
Establish DB Connection
exception_count = 0
Pulling data from database for given interval of time
Succcess
Failure
exception_count += 1
Yes
No
exception_count < 5
Operational Error occurred while pulling the data
Save data inside "input" director
data
Data Cleaning
Remove rows with excluded categories
Convert buying_mode values to 'Bid' if they're


not already 'Direct'
data
Return the cleaned data
data
Analyze Buyer Behavior
DataFrame "data" is grouped by three columns: buyer_id, category_id, and buying_mode, and several aggregation functions are applied to the order_id, order_amount, consignee_id, and seller_id columns using the agg method.
The resulting DataFrame is then reset so that the grouping columns become regular columns
A new DataFrame is created by pivoting the previous DataFrame on the buyer_id and category_id columns, and with buying_mode as the columns to pivot on.
The resulting pivoted DataFrame is then filled with 0s for missing values, and the column names are modified to have more descriptive names.
A new column is created by summing the #orders_bid and #orders_DP columns, and a new column is created by calculating the percentage of orders that were made in DP mode.
Finally, a subset of the pivoted DataFrame is returned, containing the buyer_id, category_id, and several other columns related to DP mode orders.
Analyze Category Behavior
The input DataFrame is grouped by three columns: category_id, and buying_mode, and aggregation function is applied to the order_id using the agg method.
The resulting DataFrame is then reset so that the grouping columns become regular columns
A new DataFrame is created by pivoting the previous DataFrame on the category_id column, and with buying_mode as the columns to pivot on.
The resulting pivoted DataFrame is then filled with 0s for missing values, and the column names are modified to have more descriptive names.
A new column is created by summing the #orders_bid and #orders_DP columns, and a new column is created by calculating the percentage of orders that were made in DP mode.
Finally, a subset of the pivoted DataFrame is returned, containing the category_id, and other column related to DP mode orders.
buyer_behaviour
category_behaviour
Behavioral Deviation
Merge two dataframe by inner join based on category id
A new column is created by subtracting the %_DP_cat and %_DP_buyer columns


A new flag column is created and labeled as -1 if
#orders_DP and diff are below the specified threshold
else labeled as 0.
return subset of data