

Question : Script to calculate the moving average and standard deviation of the last 10 transactions for each card_id for the data present in Hadoop and NoSQL database. If the total number of transactions for a particular card_id is less than 10, then calculate the parameters based on the total number of records available for that card_id. The script should be able to extract and feed the other relevant data ('postcode', 'transaction_dt', 'score', etc.) for the look-up table along with card_id and UCL.

Step-1

1. Load card_transactions to dataframe
 1. `df_card_transactions = pd.read_csv('card_transactions.csv')`
2. Create a rank column with group by cardid and date sorted in descending order
 1. `df_card_transactions["Rank"] = df_card_transactions.groupby("card_id")["transaction_dt2"].rank(ascending=False, method='dense').astype(int)`
 2. `df_card_transactions.sort_values("Rank")`
3. Get only the top 10 rank for each user
 1. `df_card_transactions_top10 = df_card_transactions[(df_card_transactions.Rank <=10)]`

Step-2

1. Calculate mean and standard deviation by amount for each member
 1. `df_ucl_sd = df_card_transactions_top10.groupby(['member_id'])['amount'].agg(['mean', 'std'])`
 2. `df_ucl_sd.rename(columns = {'mean':'Avg1', 'std':'SD'}, inplace = True)`
2. Calculate UCL per the rules
 1. `df_ucl_sd['ucl'] = df_ucl_sd.Avg1 + 3*df_ucl_sd.SD`
3. Keep Only 1st Rank
 1. `lstTxnPostCodeTxDate = df_card_transactions[df_card_transactions.Rank ==1]`
4. Merge the tables to get other columns – postode, transaction date and score
 1. `tmp1 = df_ucl_sd.merge(lstTxnPostCodeTxDate, on='member_id', how='left')`
 2. `tmp2 = df_member_score.merge(tmp1, on='member_id', how='left')`
 3. `columns = ['amount', 'pos_id', 'status', 'transaction_dt', 'status', 'Rank', 'Avg1', 'SD']`
 4. `tmp2.drop(columns, inplace=True, axis=1)`
 5. `lookup_df = tmp2`
 6. `lookup_df = lookup_df.drop_duplicates(subset=['member_id'])`
 7. `lookup_df = lookup_df.reset_index()`
 8. `final_lookup_df = lookup_df[['member_id','score','ucl','card_id','postcode','transaction_dt2']]`