NAME : AATMIC TIWARI

E-mail: aatmict2@illinois.edu

First off all let us understand what basically is going on in all the four files vaguely , so lets say a user(users table) bought a product (brands table) and got a receipt (receipt table) and based on the products purchased they will get further rewards on any other items(receipt_items table) like coupons or something like this.

Now lets start with the questions

1. 1) Review CSV data and diagram a new structured relational data model

Here is the below diagram of ER diagram,I could have used Lucid chart but went with the handwritten way for better understanding of the assignment by myself.

Here primary keys are as follows for the following tables:

Brands : id

Users: id

Receipts : id

Receipt_items : reviews_receipt_items_id

In the above diagram * meant many relations so receipts can have have many rewards on different receipt_items and similarly one user can get many or different receipts so it is also a one to many relationships.

User_id can act as a foreign key which will help further to connect with the users and modify_date will have one to one relationship or common attribute which can help to join the tables,likewise with brandcode as well.

Please have a look at the diagram below.


E.R diagram
Brands
id (P.K.)
CPG_id
Barcode
Brand Code
Category
Category_Code
Name
Romance_Related
Brand_IDS
gets
users
id (P.K)
Created Date
Birth date
Gender
Last_Rewards_login
State
Sign_up Platform
Sign_up_Source
receipts
id (P.K)
user_id (F.K)
store_name
Purchase_date
Purchase_time
Date Scanned
Total_Spent
Rewards_receipt_status
User_viewed
Purchased_item_count
Created_date
Modify date
Processed date
Finished date
Rejected date
Need Fetch Review
Digital_Receipt
Deleted
has
receipt_items
Review_Receipt_item_id (P.K)
Review_Receipt_id
Item_idex
Barcode
Brandcode (F.K)
Quantity_Purchased
Total_Final_Price
Points Earned
Rewards_group
Original_Receipt_item_text
Modify_date

Did this answer the questions asked in part 2, absolutely :

1). We can take the review_items table for finding brand names using brand code and total quantity for finding the most dollars spent in June.

2). We can use user_id and SUM(TOTAL) from the receipts table to find which user spent the most money in august

by extracting the month.

3). For this question, you can join table receipt and receipt_items on id and review_item_id, and then one can easily find the item brand as well described.

4). For the most expensive purchase, we can use the receipt_items table directly.

5). This can also be directly answered by using the receipt table.