SQL Question

Instructions: Submit a Notepad file (.sql file) or GitHub repo link containing the SQL queries via Email.

Note: We expect your answers to be based on the dataset snapshots provided – there is no need to use external data sources or download anything. You can make assumptions if you feel something is missing, but do specify the assumptions you take in the submission.

Tips for Success:

●Make sure to include all information and/or insights that you find relevant

●Make sure code is well commented, has a good readability and is understood easily

Data: We have two tables available: transactions and items

Table: transactions

buyer_id	purchase_time	refund_item	store_id	item_id	gross_transaction_value
3	2019-09-19 21:19:06.544 UTC		a	a1	$58
12	2019-12-10 20:10:14.324 UTC	2019-12-15 23:19:06.544 UTC	b	b2	$475
3	2020-09-01 23:59:46.561 UTC	2020-09-02 21:22:06.331 UTC	f	f9	$33
2	2020-04-30 21:19:06.544 UTC		d	d3	$250
1	2020-10-22 22:20:06.531 UTC		f	f2	$91
8	2020-04-16 21:10:22.214 UTC		e	e7	$24
5	2019-09-23 12:09:35.542 UTC	2019-09-27 02:55:02.114 UTC	g	g6	$61

Table: items

store_id	item_id	item_category	item_name
a	a1	pants	denim pants
a	a2	tops	blouse
f	f1	table	coffee table
f	f5	chair	lounge chair
f	f6	chair	armchair
d	d2	jewelry	bracelet
b	b4	earphone	airpods

Questions:

1. What is the count of purchases per month (excluding refunded purchases)?
2. How many stores receive at least 5 orders/transactions in October 2020?
3. For each store, what is the shortest interval (in min) from purchase to refund time?
4. What is the gross_transaction_value of every store's first order?
5. What is the most popular item name that buyers order on their first purchase?
6. Create a flag in the transaction items table indicating whether the refund can be processed or not. The condition for a refund to be processed is that it has to happen within 72 of Purchase time.

Expected Output: Only 1 of the three refunds would be processed in this case

7. Create a rank by buyer_id column in the transaction items table and filter for only the second purchase per buyer. (Ignore refunds here)

Expected Output: Only the second purchase of buyer_id 3 should the output

8. How will you find the second transaction time per buyer (don’t use min/max; assume there were more transactions per buyer in the table)
Expected Output: Only the second purchase of buyer_id along with a timestamp