Unit 05 Problem Set Submission Form

Overview

Your Name	Bhavya Shah
Your SU Email	bhshah@yr.edu

Instructions

Put your name and SU email at the top. Answer these questions all from the lab. When asked to include screenshots, please follow the screen shot guidelines from the first lab.

Remember as you complete the problem sets it is not only about getting it right / correct. We will discuss the answers in class so it's important to articulate anything you would like to contribute to the discussion in your answer:

- If you feel the question is vague, include any assumptions you've made.
- If you feel the answer requires interpretation or justification provide it.
- If you do not know the answer to the question, articulate what you tried and how you are stuck.

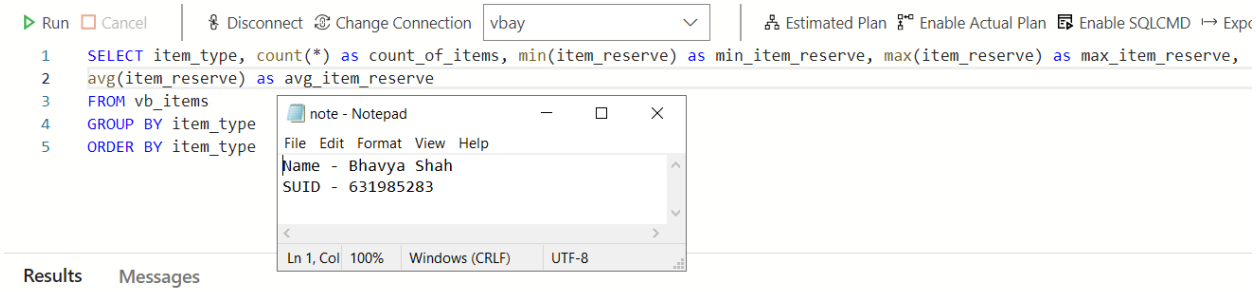
This how you receive credit for answering questions which might not be correct.

Questions

Answer these questions using the problem set submission template. You will need to consult the logical model in the overview section for details. For any screenshots provided, please follow the guidelines for submitting a screenshot.

Write the following as SQL queries. If the query is ambiguous, fill in the gaps yourself and justify your reasoning. For each, include the SQL as a screenshot with the output of the query.

1. How many item types are there? Perform an analysis of each item type. For each item type, provide the count of items in that type, the minimum, average, and maximum item reserve prices for that type. Sort the output by item type.

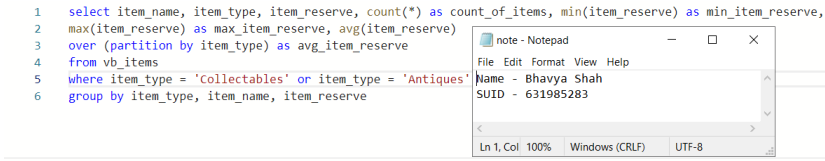


Results Messages

	item_type	count_of_items	min_item_reserve	max_item_reserve	avg_item_reserve
1	All Other	4	0.99	10000000.00	2500004.86
2	Antiques	6	9.00	250.00	81.5833
3	Books	3	4.50	10.99	8.48
4	Collectables	14	5.00	500.00	105.3828
5	Electronics	1	15.00	15.00	15.00
6	Jewelry	2	6.95	599.99	303.47
7	Sporting Goods	2	12.50	12.75	12.625
8	Tickets	2	5.00	750.00	377.50

There are total of 8 item types.

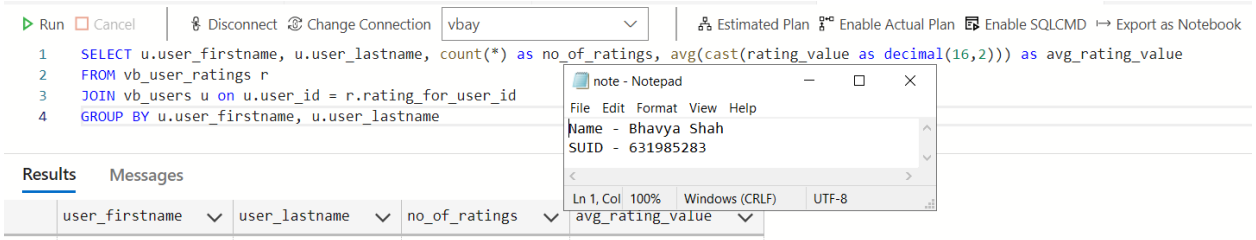
2. Perform an analysis of each item in the “Antiques” and “Collectables” item types. For each item display the name, item type and item reserve. Include the min, max and average item reserve over each item type so that the current item reserve can be compared to these values.



Results Messages

	item_name	item_type	item_reserve	count_of_items	min_item_reserve	max_item_reserve	avg_item_reserve
1	a Toaster	Antiques	20.00	1	20.00	20.00	81.5833
2	Antique Desk	Antiques	250.00	1	250.00	250.00	81.5833
3	Brass French Press	Antiques	45.50	1	45.50	45.50	81.5833
4	case of vintage tube socks	Antiques	9.00	1	9.00	9.00	81.5833
5	Dusty Vase	Antiques	100.00	1	100.00	100.00	81.5833
6	Original Coke Bottle from 19…	Antiques	65.00	1	65.00	65.00	81.5833
7	Alf Alarm Clock	Collectables	5.00	1	5.00	5.00	105.3828
8	Autographed Mik Jagger Poster	Collectables	75.00	1	75.00	75.00	105.3828
9	Carlos Villalba BobbleHead	Collectables	49.95	1	49.95	49.95	105.3828
10	Dukes Of Hazard ashtray	Collectables	149.99	1	149.99	149.99	105.3828
11	Farrah Fawcet poster	Collectables	50.00	1	50.00	50.00	105.3828
12	Joe Montanna Figurine	Collectables	200.00	1	200.00	200.00	105.3828
13	Kleenex used by Dr. Dre	Collectables	500.00	1	500.00	500.00	105.3828
14	Mike Fudge BobbleHead	Collectables	49.95	1	49.95	49.95	105.3828
15	PacMan Fever lunchbox	Collectables	29.99	1	29.99	29.99	105.3828

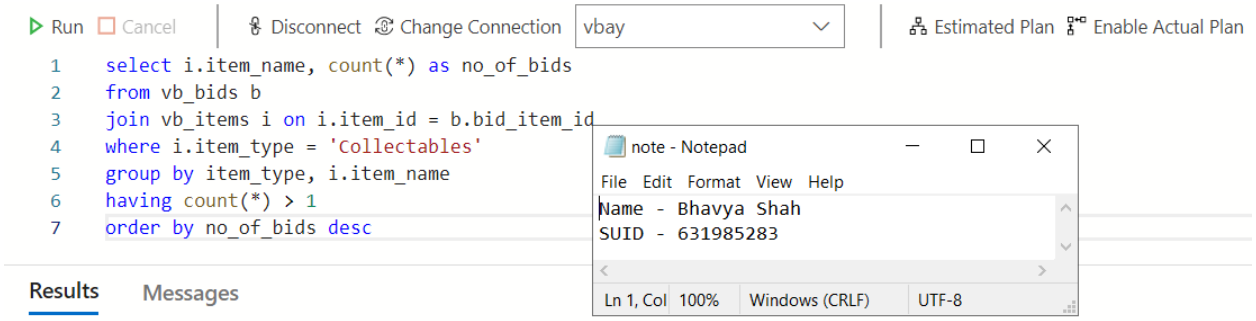
3. Write a query to include the names, counts (number of ratings) and average seller ratings (as a decimal) of users. For reference, User Carrie Dababbbi has 4 seller ratings and an average rating of 4.75.



Results | Messages

	user_firstname	user_lastname	no_of_ratings	avg_rating_value
1	Rose	Abov-Duresst	3	1.000000
2	Ty	Anott	2	2.500000
3	Barb	Barion	4	4.250000
4	Carrie	Dababbi	4	4.750000
5	Barry	DeHatchett	2	5.000000
6	Martin	Eyezing	2	2.500000
7	Isabelle	Gunnering	2	4.000000
8	Les	Ismoore	8	1.875000
9	Anita	Job	1	3.000000
10	Abby	Kuss	5	4.200000
11	Mary	Melator	4	4.500000
12	Victor	Rhee	1	4.000000
13	Gus	Toffwind	2	3.000000

4. Create a list of “Collectable” item types with more than 1 bid. Include the name of the item and the number of bids making sure the item with the most bids appear first.

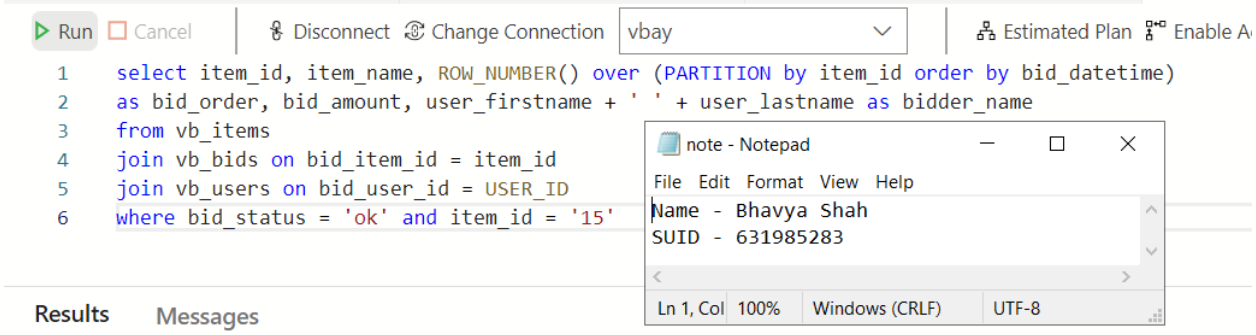


Results | Messages

	item_name	no_of_bids
1	Dukes Of Hazard ashtray	9
2	Autographed Mik Jagger Poster	6
3	Shatner's old Toupee	5
4	Rare Mint Snow Globe	3
5	Farrah Fawcet poster	3
6	Pez dispensers	2

5. Generate a valid bidding history for any given item of your choice. Display the item id, item name a number representing the order the bid was placed, the bid amount and the bidder's name. Here's an example showing the first 3 bids on item 11.

item_id	item_name	bid_order	bid_amount	bidder
11	Dukes Of Hazard ashtray	1	150.0000	Dan Delyons
11	Dukes Of Hazard ashtray	2	175.0000	Al Fresco
11	Dukes Of Hazard ashtray	3	200.0000	Carrie Dababbi

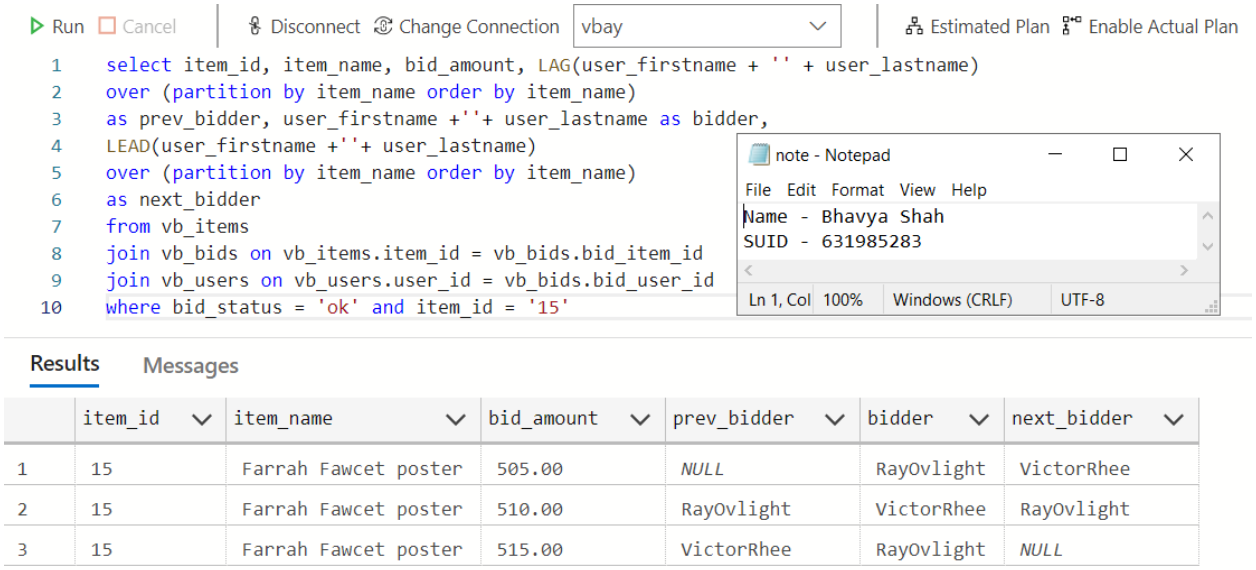


Results Messages

	item_id	item_name	bid_order	bid_amount	bidder_name
1	15	Farrah Fawcet poster	1	505.00	Ray Ovlight
2	15	Farrah Fawcet poster	2	510.00	Victor Rhee
3	15	Farrah Fawcet poster	3	515.00	Ray Ovlight

6. Re-Write your query in the previous question to include the names of the next and previous bidders, like this example again showing the first 3 bids for item 11.

item_name	bid_order	bid_amount	prev_bidder	bidder	next_bidder
Dukes Of Hazard ashtray	1	150.0000	*NULL*	Dan Delyons	Al Fresco
Dukes Of Hazard ashtray	2	175.0000	Dan Delyons	Al Fresco	Carrie Dababbi
Dukes Of Hazard ashtray	3	200.0000	Al Fresco	Carrie Dababbi	Gus Toffwind



Results Messages

	item_id	item_name	bid_amount	prev_bidder	bidder	next_bidder
1	15	Farrah Fawcet poster	505.00	*NULL*	RayOvlight	VictorRhee
2	15	Farrah Fawcet poster	510.00	RayOvlight	VictorRhee	RayOvlight
3	15	Farrah Fawcet poster	515.00	VictorRhee	RayOvlight	*NULL*

7. Find the names and emails of the users who give out the worst ratings (lower than the overall average rating) to either buyers or sellers (no need to differentiate whether the user rated a buyer

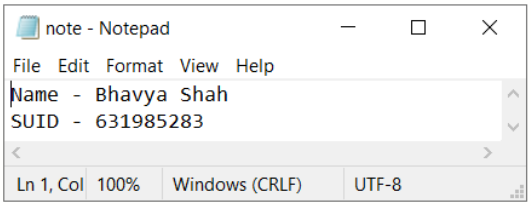
or seller), and only include those users who have submitted more than 1 rating.

Run | Cancel | Disconnect | Change Connection | vbay | Estimated Plan

```
with ratings_tables as (select rating_id, u.user_firstname as user_name, u.user_email,
avg(rating_value) over() as overall_average_rating, count(u.user_firstname)
over (partition by u.user_firstname) as count_of_reviews
from vb_users u
join vb_user_ratings r on u.user_id = r.rating_for_user_id
group by r.rating_id, u.user_firstname, u.user_email, rating_value)
select USER_NAME, user_email
from ratings_tables
join vb_user_ratings r on r.rating_id = ratings_tables.rating_id
where r.rating_value < overall_average_rating and count_of_reviews > 1
```

Results | Messages

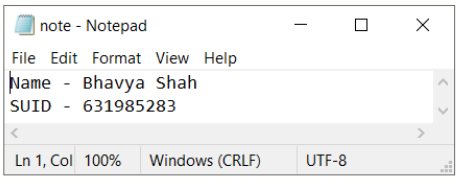
	USER_NAME	user_email
1	Barb	bbarion@mail.org
2	Les	lismoore@mail.org
3	Les	lismoore@mail.org
4	Les	lismoore@mail.org
5	Les	lismoore@mail.org
6	Les	lismoore@mail.org
7	Martin	meveyzing@mail.org
8	Rose	rabovdu@mail.org
9	Rose	rabovdu@mail.org
10	Rose	rabovdu@mail.org
11	Ty	tanott@mail.org



8. Produce a report of the KPI (key performance indicator) user bids per item. Show the user's name and email total number of valid bids, total count of items bid upon and then the ratio of bids to items. As a check, Anne Dewey's bids per item ratio is 1.666666

Run | Cancel | Disconnect | Change Connection | vbay | Estimated Plan | Enable Actual Plan | Enable S(

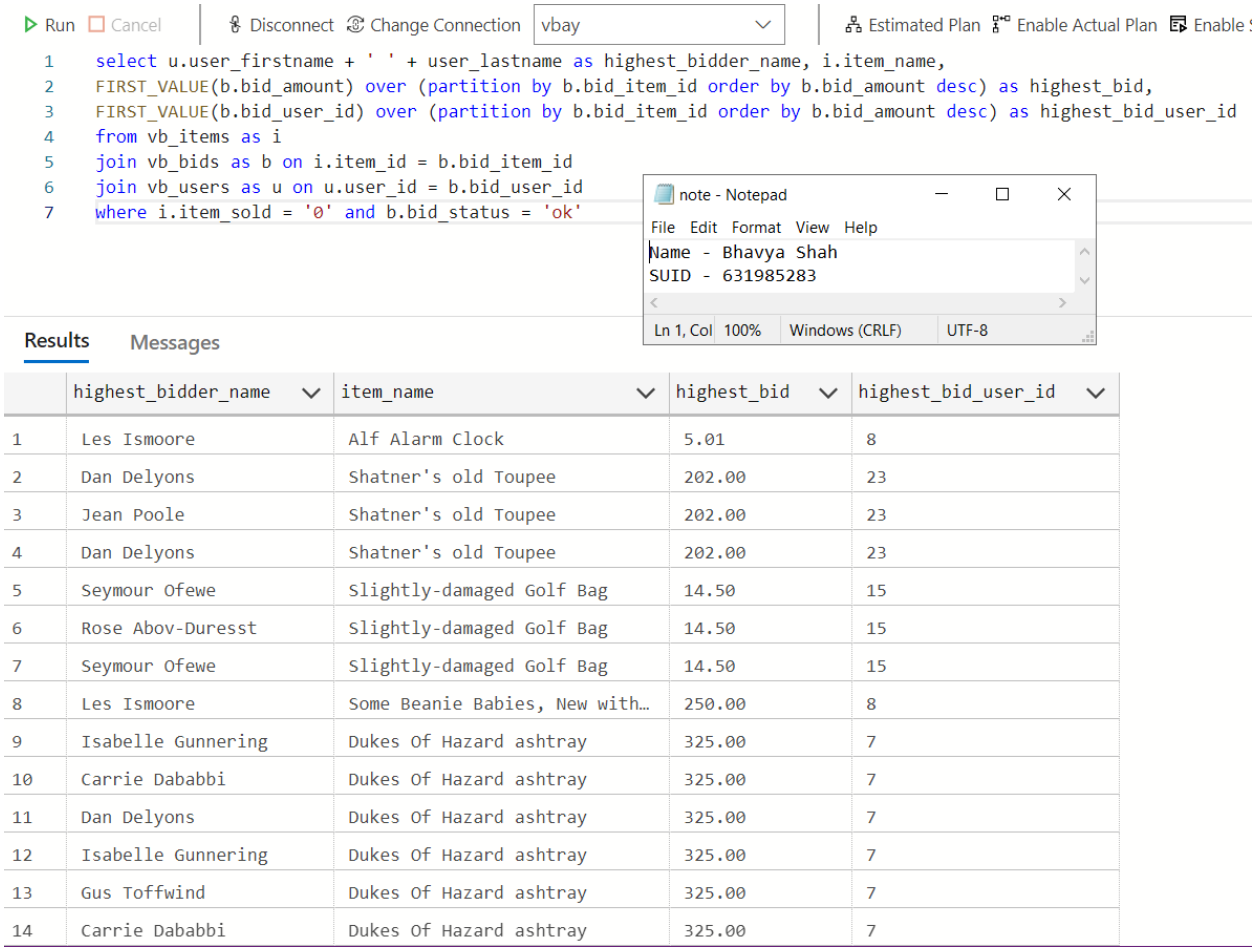
```
1  with user_bid_items as (select u.user_firstname + ' ' + u.user_lastname as user_name, u.user_email,
2  count (b.bid_user_id) as total_count_of_valid_bids,
3  count (distinct b.bid_item_id) as total_items_bid
4  from vb_users u
5  join vb_bids as b on u.user_id = b.bid_user_id
6  where b.bid_status = 'ok'
7  group by u.user_firstname + ' '+ u.user_lastname, u.user_email
8  )
9  select user_name, total_count_of_valid_bids, total_items_bid,
10 cast(total_count_of_valid_bids as decimal) / cast(total_items_bid as decimal) as KPI from user_bid_items
```



Results | Messages

	user_name	total_count_of_valid_bids	total_items_bid	KPI
1	Abby Kuss	3	1	3.000000000000000000
2	Anne Dewey	5	3	1.666666666666666666
3	Barb Barion	3	2	1.500000000000000000
4	Barry DeHatchett	5	1	5.000000000000000000
5	Bo Enarreau	2	2	1.000000000000000000
6	Gus Toffwind	2	2	1.000000000000000000
7	Isabelle Gunnering	7	2	3.500000000000000000
8	Les Ismoore	3	3	1.000000000000000000
9	Martin Eyezing	1	1	1.000000000000000000
10	Rose Abov-Duresst	2	2	1.000000000000000000
11	Ray Ovlight	6	3	2.000000000000000000
12	Victor Rhee	2	2	1.000000000000000000
13	Seymour Ofewe	2	1	2.000000000000000000
14	Pete Moss	2	2	1.000000000000000000

9. Among items not sold, show highest bidder name and the highest bid for each item. Make sure to include only valid bids.



	highest_bidder_name	item_name	highest_bid	highest_bid_user_id
1	Les Ismoore	Alf Alarm Clock	5.01	8
2	Dan Delyons	Shatner's old Toupee	202.00	23
3	Jean Poole	Shatner's old Toupee	202.00	23
4	Dan Delyons	Shatner's old Toupee	202.00	23
5	Seymour Ofewe	Slightly-damaged Golf Bag	14.50	15
6	Rose Abov-Duresst	Slightly-damaged Golf Bag	14.50	15
7	Seymour Ofewe	Slightly-damaged Golf Bag	14.50	15
8	Les Ismoore	Some Beanie Babies, New with…	250.00	8
9	Isabelle Gunnering	Dukes Of Hazard ashtray	325.00	7
10	Carrie Dababbi	Dukes Of Hazard ashtray	325.00	7
11	Dan Delyons	Dukes Of Hazard ashtray	325.00	7
12	Isabelle Gunnering	Dukes Of Hazard ashtray	325.00	7
13	Gus Toffwind	Dukes Of Hazard ashtray	325.00	7
14	Carrie Dababbi	Dukes Of Hazard ashtray	325.00	7

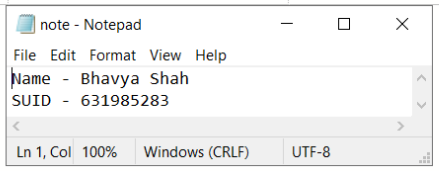
10. Write a query with output similar to question 3, but also includes the overall average seller rating, and the difference between each user's average rating and the overall average. For reference, the overall average seller rating should be 3.2.

Run | Cancel | Disconnect | Change Connection | vbay | Estimated Plan | Enable Actual Plan | Enable SQL

```
with users_info as (select u.user_firstname +' ' + u.user_lastname as "user_name",
count(rating_id) over (partition by rating_for_user_id) as no_of_ratings_for_user,
AVG(CAST(rating_value as float)) over (partition by rating_for_user_id) as avg_of_ratings,
AVG(CAST(rating_value as float)) over() as overall_avg from vb_user_ratings as r
join vb_users as u on u.user_id = r.rating_for_user_id
where r.rating_astype = 'Seller'
)
select user_name, no_of_ratings_for_user, avg_of_ratings, avg_of_ratings - overall_avg as delta from users_info
group by user_name, no_of_ratings_for_user, overall_avg, avg_of_ratings
```

Results | Messages

	user_name	no_of_ratings_for_user	avg_of_ratings	delta
1	Abby Kuss	3	4.333333333333333	1.1333333333333329
2	Anita Job	1	3	-0.20000000000000018
3	Barb Barion	2	3.5	0.2999999999999998
4	Carrie Dababbi	4	4.75	1.5499999999999998
5	Les Ismoore	2	2.5	-0.7000000000000002
6	Martin Eyezing	2	2.5	-0.7000000000000002
7	Rose Abov-Duresst	3	1	-2.2
8	Ty Anott	2	2.5	-0.7000000000000002
9	Victor Rhee	1	4	0.7999999999999998



Reflection

Use this section to reflect on your learning. To achieve the highest grade on the assignment you must be as descriptive and personal as possible with your reflection.

1. What are the key things you learned through the process of completing this assignment?
   Ans - window functions and select queries

2. What were the challenges or roadblocks (if any) you encountered on the way to completing it?
   Ans – Window functions is a completely was a new concept for me, so it took some time for me to understand it.

3. Were you prepared for this assignment? What can you do to be better prepared?
   Ans - I watched Prof. Fudge's videos on window functions. But few more exercise questions would have helped me do better in the assignment.

4. Now that you have completed the assignment rate your comfort level with this week's material. This should be an honest assessment: (choose one)

   4 ==> I understand this material and can explain it to others.
   3 ==> I understand this material.
   2 ==> I somewhat understand the material but sometimes need guidance from others.
   1 ==> I understand very little of this material and need extra help.