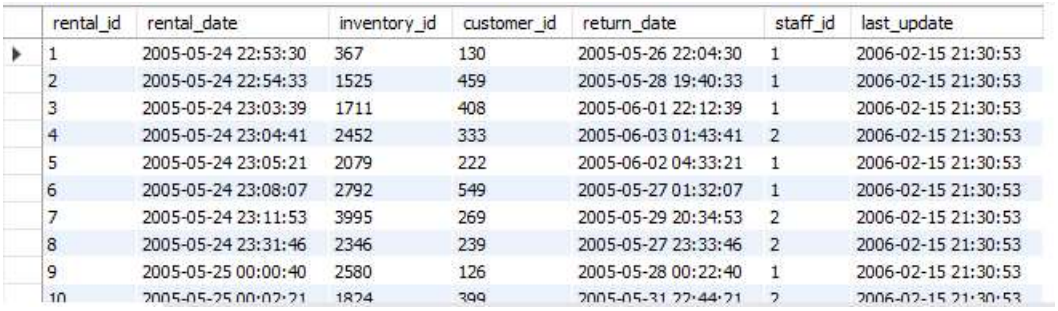
Lab | SQL Queries 9

Instructions

#In this lab we will find the customers who were active in consecutive months of May and June. Follow the steps to complete the analysis.

-- Create a table `rentals_may` to store the data from rental table with information for the month of May.

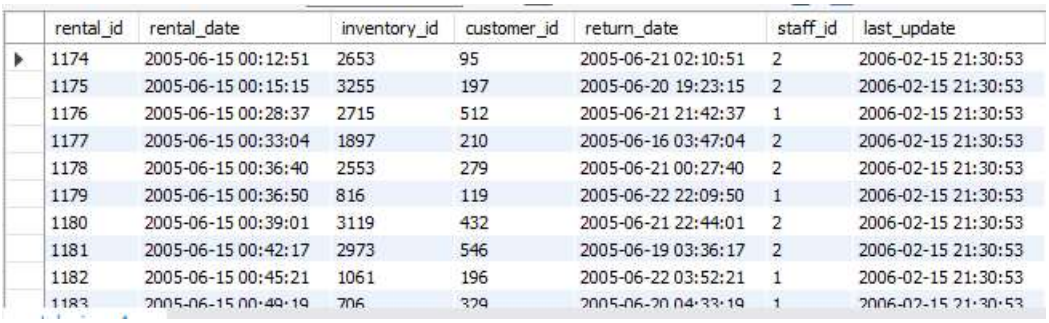
-- Insert values in the table `rentals_may` using the table rental, filtering values only for the month of May.



rental_id	rental_date	inventory_id	customer_id	return_date	staff_id	last_update
1	2005-05-24 22:53:30	367	130	2005-05-26 22:04:30	1	2006-02-15 21:30:53
2	2005-05-24 22:54:33	1525	459	2005-05-28 19:40:33	1	2006-02-15 21:30:53
3	2005-05-24 23:03:39	1711	408	2005-06-01 22:12:39	1	2006-02-15 21:30:53
4	2005-05-24 23:04:41	2452	333	2005-06-03 01:43:41	2	2006-02-15 21:30:53
5	2005-05-24 23:05:21	2079	222	2005-06-02 04:33:21	1	2006-02-15 21:30:53
6	2005-05-24 23:08:07	2792	549	2005-05-27 01:32:07	1	2006-02-15 21:30:53
7	2005-05-24 23:11:53	3995	269	2005-05-29 20:34:53	2	2006-02-15 21:30:53
8	2005-05-24 23:31:46	2346	239	2005-05-27 23:33:46	2	2006-02-15 21:30:53
9	2005-05-25 00:00:40	2580	126	2005-05-28 00:22:40	1	2006-02-15 21:30:53
10	2005-05-25 00:02:21	1824	399	2005-05-31 22:44:21	2	2006-02-15 21:30:53

-- Create a table `rentals_june` to store the data from rental table with information for the month of June.
create table if not exists rentals_june like sakila.rental;

-- Insert values in the table `rentals_june` using the table rental, filtering values only for the month of June.



rental_id	rental_date	inventory_id	customer_id	return_date	staff_id	last_update
1174	2005-06-15 00:12:51	2653	95	2005-06-21 02:10:51	2	2006-02-15 21:30:53
1175	2005-06-15 00:15:15	3255	197	2005-06-20 19:23:15	2	2006-02-15 21:30:53
1176	2005-06-15 00:28:37	2715	512	2005-06-21 21:42:37	1	2006-02-15 21:30:53
1177	2005-06-15 00:33:04	1897	210	2005-06-16 03:47:04	2	2006-02-15 21:30:53
1178	2005-06-15 00:36:40	2553	279	2005-06-21 00:27:40	2	2006-02-15 21:30:53
1179	2005-06-15 00:36:50	816	119	2005-06-22 22:09:50	1	2006-02-15 21:30:53
1180	2005-06-15 00:39:01	3119	432	2005-06-21 22:44:01	2	2006-02-15 21:30:53
1181	2005-06-15 00:42:17	2973	546	2005-06-19 03:36:17	2	2006-02-15 21:30:53
1182	2005-06-15 00:45:21	1061	196	2005-06-22 03:52:21	1	2006-02-15 21:30:53
1183	2005-06-15 00:49:19	706	329	2005-06-20 04:33:19	1	2006-02-15 21:30:53

-- Check the number of rentals for each customer for May.

customer_id	num_rentals_may
1	2
2	1
3	2
5	1
6	2
7	3
9	1
14	5
16	2
17	2

-- Check the number of rentals for each customer for June.

customer_id	num_rentals_june
1	7
2	1
3	4
4	6
5	5
6	4
7	5
8	3
9	2
10	5

-- Create a Python connection with SQL database and retrieve the results of the last two queries (also mentioned below) as dataframes:

-- Check the number of rentals for each customer for May

	customer_id	num_rentals_may
0	1	2
1	2	1
2	3	2
3	5	1
4	6	2

-- Check the number of rentals for each customer for June

	customer_id	num_rentals_june
0	1	7
1	2	1
2	3	4
3	4	6
4	5	5

-- Write a function that checks if customer borrowed more or less films in the month of June as compared to May.

	customer_id	num_rentals_may	num_rentals_june	diff
0	1	2	7	5
1	2	1	1	0
2	3	2	4	2
3	5	1	5	4
4	6	2	4	2