

DSO 552: SQL Databases for Business Analysts

Final Exam (Exam 1)

This exam is timed. You will have 3 hours and 15 minutes as soon as you opened this exam on Blackboard to complete this final exam. Please upload a `.sql` file to Blackboard with your answers before the time limit expires. You will need to connect to the `yelp` and `entertainment_agency` databases.

Since this is an asynchronous test, if you have any questions, make your best assumptions and write them as comments in the SQL file. We will make sure to consider your assumptions in our grading if the answer is incorrect.

Also note that depending on the database you connect to (Virginia versus Hong Kong), your answers may vary slightly from the sample output. That does not mean the answer is wrong, so do not spend too much time trying to make your answer align 100% with the sample output.

You get 1 point for free. There are 10 questions, and you can choose to submit any 9 questions you'd like. We will not provide extra credit. If you do all 10 questions, we will pick the first 9 to grade.

Since the `yelp` database is slightly different for Hong Kong than for Virginia, we have included sample outputs for both databases. The `entertainment_agency` is the same across both databases.

Note: Make sure to submit your answers by the deadline on Blackboard. Late submissions will not be accepted.

Also make sure

- your column names match the sample output
- you do not `LIMIT 5` unless the question directly asks you to do so
- unless the question explicitly asks for an ordering, you are free to return the results in any order you want.

Yelp Database

1. *(1 point)* Which users tend to have the highest average rating per reviews? List the top 5 users in terms of star ratings on reviews. List only users with more than 2 total reviews.

Expected (Virginia Database):

user_id	average_star_rating
-LG_49tfkstPBzXB-C6R2g	5
I-4KVZ9lqHhk8469X9FvhA	5
lMTmBNddR47i35jgOnIdIA	5
WLde9fTArPw_L6KkuZ60-w	5
5hfrKyck8DBN96EN9VczUw	5

Expected (Hong Kong Database):

user_id	average_star_rating
CKRzIkY1N0eFCDkShZi67A	5.000000
-LG_49tfkstPBzXB-C6R2g	5.000000
I-4KVZ9lqHhk8469X9FvhA	5.000000
WLde9fTArPw_L6KkuZ60-w	5.000000
ZRnsUs9q6DbeJbOLjCY4ZA	4.333333

2. (1 point) Which category of business tend to have the highest ratings? Calculate the weighted average number of stars for a category - this is defined as the sum of the review count x stars divided by the sum of review_count column. Show only categories that have more than 3 businesses. Sort by the weighted average number of stars.

Expected (Virginia Database):

category	weighted_average_stars	num_businesses
Health & Medical	4.278571	5
Local Services	4.090000	6
Shopping	3.998000	10
Professional Services	3.938272	4
Restaurants	3.625767	19
Sandwiches	3.572581	4
Event Planning & Services	3.504287	4
Food	3.311475	5

Expected (Hong Kong Database):

category	weighted_average_stars	num_businesses
Nightlife	4.465659	4
Health & Medical	4.358209	4
Local Services	4.294118	4
Shopping	3.982036	10
Professional Services	3.938272	4
Restaurants	3.922591	17
Home Services	3.877551	4
Event Planning & Services	3.504287	4
Food	3.342466	6
Fashion	3.065217	4

3. (1 point) List all users that left both a tip and a review for a particular business. List the business' name, and the number of tips and reviews left for that particular business. Sort by business_name.

Expected (Virginia Database):

username	businessname	tips_left	reviews_left
Alicia	99 Cents Only	3	1
Fay	99 Cents Only	2	1
Matt	99 Cents Only	1	1
Mike	Bark 'N' Scratch	1	1
Mary	Burlington Coat Factory	1	1
Brent	Cholla Prime Steakhouse & Lounge	1	1
Comfy	Cholla Prime Steakhouse & Lounge	16	2
Jacqueline	Cholla Prime Steakhouse & Lounge	1	1
Jae And Nui	Cholla Prime Steakhouse & Lounge	1	1
Jolly	Cholla Prime Steakhouse & Lounge	1	1

Expected (Hong Kong Database):

username	businessname	tips_left	reviews_left
Alicia	99 Cents Only	3	1
Fay	99 Cents Only	2	1
Matt	99 Cents Only	1	1
Kevin	Antojitos Y Neveria Chayo's	1	1
Mike	Bark 'N' Scratch	1	1
Mary	Burlington Coat Factory	1	1
Comfy	Cholla Prime Steakhouse & Lounge	16	2
Jacqueline	Cholla Prime Steakhouse & Lounge	1	1
Jae And Nui	Cholla Prime Steakhouse & Lounge	1	1
Jolly	Cholla Prime Steakhouse & Lounge	1	1

4. (1 point) List all users who have at least 5 fans that have left a review for Burlington Coat Factory (a business).

Expected (Virginia Database):

user_id	fans	username	text
hWDybu_KvYLSdEFzGrniTw	659	Bruce	This location has been turned over so many times and vacant for so
e9x9H_wGR430eCqvTgDgPw	41	Eren	Burlington Coat factory is easily one of my favorite places to shop h
l9i6eupjzPDyfvJoUrHyOw	7	Stefanie	This new location is just wonderful. It's a lot cleaner than the other
9eFi-CKKmdlwpohVja9Y6g	21	Cameron	Burlington Coat factory is hit or miss. I've been here a handful of t
nyl_1VcRIAyI55bb_scpdw	61	Mary	I love Burlington Coat Factory. Years ago, when I lived in Rocheste

Expected (Hong Kong Database):

user_id	fans	username	text
hWDybu_KvYLSdEFzGrniTw	659	Bruce	This location has been turned over so many times and vacant for so
e9x9H_wGR430eCqvTgDgPw	41	Eren	Burlington Coat factory is easily one of my favorite places to shop h
l9i6eupjzPDyfvJoUrHyOw	7	Stefanie	This new location is just wonderful. It's a lot cleaner than the other
9eFi-CKKmdlwpohVja9Y6g	21	Cameron	Burlington Coat factory is hit or miss. I've been here a handful of t
nyl_1VcRIAyI55bb_scpdw	61	Mary	I love Burlington Coat Factory. Years ago, when I lived in Rocheste

5. (1 point) Do users who have left at least one tip tend to have higher average star ratings than users who have not left any tips?

Expected (Virginia Database):

type_of_user	avg_stars
Users without tips	3.566143
Users with tips	3.955307

Expected (Hong Kong Database):

type_of_user	avg_stars
Users without tips	3.560077
Users with tips	3.958084

6. (1 point) Perform an analysis that shows the average star ratings of businesses that 1) accept credit cards or 2) offers take-out (RestaurantsTakeOut in attributes). If the attribute is not available for the business, default it to False. Sort by the average number of stars.

Expected (Virginia Database):

accepts_credit_cards	offers_takeout	average_stars
True	False	3.625000
False	True	3.600000
False	False	3.437500
True	True	3.192308

Expected (Hong Kong Database):

accepts_credit_cards	offers_takeout	average_stars
False	True	3.900000
True	False	3.860000
True	True	3.363636
False	False	3.166667

Entertainment Agency Database

7. (1 point) The agency has a theory that the first five engagements that entertainers book tend to be lower contract value than the later engagements. Produce an analysis that shows the average contract price of the first five engagements for an entertainer versus the average contract price of the 6th engagement and beyond.

Expected (Virginia Database):

engagement_category	avg
6th and Beyond Engagements	1419.118
First Five Engagements	1136.250

8. (1 point) For all engagements that have a contract price greater than the global average contract price across all engagements, list the average contract price (for these above average engagements), the most recent start date, and the number of engagements.

Expected (Virginia Database):

avg_contract_price	most_recent_start_date	num_engagements
2284.302	2018-03-04	43

9. (1 point) Show the counts for each combination of musical style and entertainer member's gender. Only show combinations where there are more than 4 distinct entertainer members. Order by the number of members.

Expected (Virginia Database):

gender	stylename	num_members
F	Standards	8
M	60's Music	8
F	Contemporary	5
F	Top 40 Hits	5
M	Classic Rock & Roll	5
M	Country	5

10. (1 point) Show how much each customer has paid total to our agency - which is reflected in the contractprice. However, if a customer books 3 or more engagements in the same month, we charge only 90% of the contract price for the 3rd engagement that month and beyond (until the start of a new month). Do not include the agent Karen Smith's engagements at all here in this computation.

Expected (Virginia Database):

customerid	total_paid
10005	24635
10006	12970
10014	12669
10004	10618
10010	10470
10002	10393
10001	8855
10015	7080
10009	5690
10012	5357
10013	4580
10007	4485
10003	3130