MSA 8040 Data Management for Analytics: MySQL

The goal of this project is to enhance students' abilities of using MySQL. Details of the project can be found in following table:

Project	Topics	Release Date	Due Date
Part I (70)	Query	09/14/2021	10/10/2021
Part II (30)	Table Structure, Procedure, and Trigger	09/14/2021	10/10/2021

Instructions:

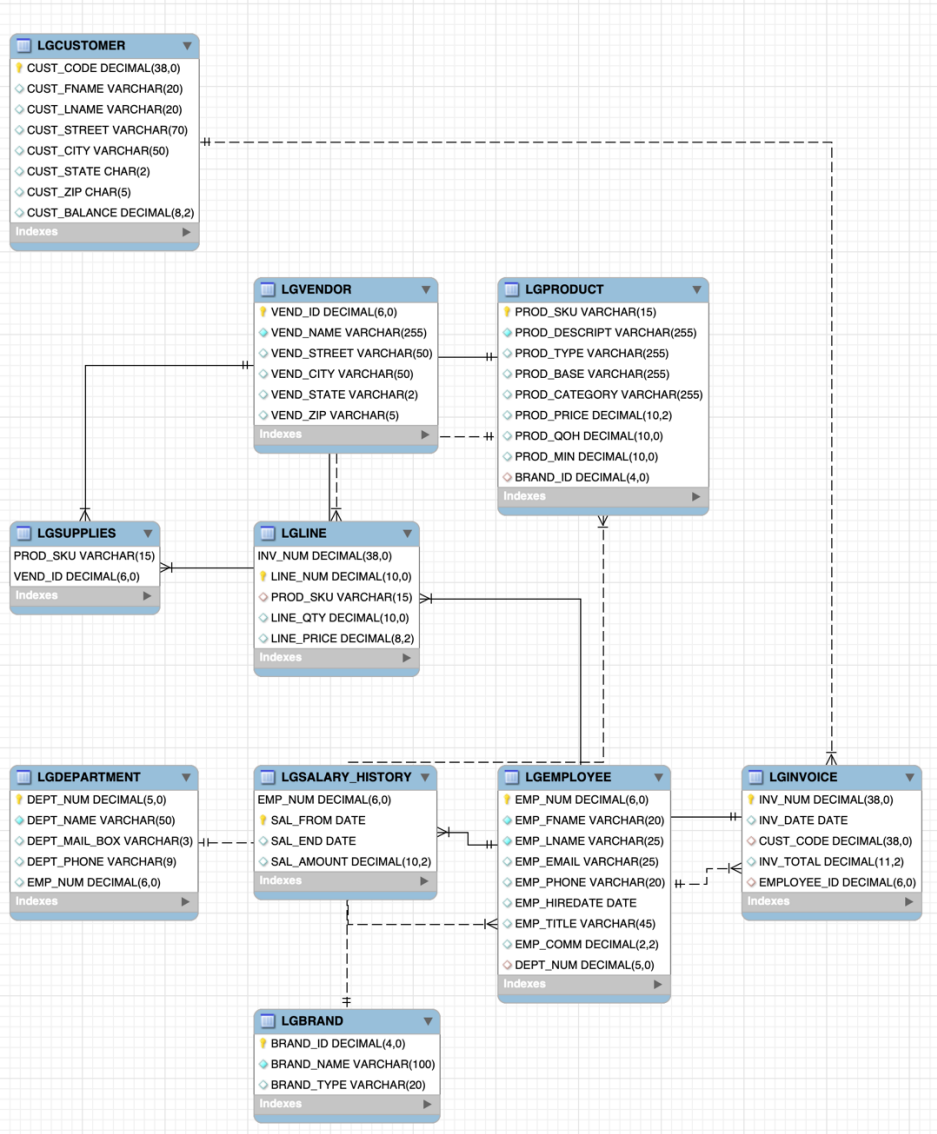
1. This is **individual project**. Please refer to the Syllabus for Integrity policy.
2. Please refer to the Syllabus for late submission policy
3. Each question has the objectives and scores followed by the questions. Each question can be correctly implemented in different way. As long as you can obtain the correct answers, you will get full credits. If you fail to answer the question correctly, you will get NONE.
4. Individual feedbacks will be available after 10 /17/2021 per requests.

Things need to be submitted via iCollege:

1. Your code contains all queries for Q1~Q14.
   Name of your code should be in the following format: **firstname_lastname_studentid.sql.** If not correctly name your file, you will lose 5 points.

Part I

The LargeCo database (see the following figure) stores data for a company that sells paint products. The company tracks the sale of products to customers. The database keeps data on customers (LGCUSTOMER), sales (LGINVOICE), products (LGPRODUCT), which products are on which invoices (LGLINE), employees (LGEMPLOYEE), the salary history of each employee (LGSALARY_ HISTORY), departments (LGDEPARTMENT), product brands (LGBRAND), vendors (LGVENDOR), and which vendors supply each product (LGSUPPLIES). Some of the tables contain only a few rows of data, while other tables are quite large; for example, there are only eight departments, but more than 3,300 invoices containing over 11,000 invoice lines.



Question 1. Write a query to display the SKU (stock keeping unit), description, type, base, category, and price for all products that have a PROD_BASE of Water and a PROD_CATEGORY of Sealer. And, use aliases to each attribute. For instance, PROD_DESCRIPT → 'product description'. (Logical/Comparison Operator) (5%)

yourname1	yourname2	yourname3	yourname4	yourname5	yourname6
1403-TUY	Sealer, Water Based, for Concrete Floors	Interior	Water	Sealer	42.99

Question 2. Write a query to display the employee number, last name, first name, salary "from" date, salary end date, and salary amount for employees 83731, 83745, and 84039. Sort the output by employee number and salary "from" date. (Logical/Comparison Operator) (5%)

EMP_NUM	EMP_LNAME	EMP_FNAME	SAL_FROM	SAL_END	SAL_AMOUNT
83731	VARGAS	SHERON	2014-07-15	2015-07-14	43740.00
83731	VARGAS	SHERON	2015-07-14	2016-07-13	48110.00
83731	VARGAS	SHERON	2016-07-14	2017-07-14	49550.00
83731	VARGAS	SHERON	2017-07-15	NULL	51040.00
83745	SPICER	DWAIN	2011-08-02	2012-08-01	56020.00
83745	SPICER	DWAIN	2012-08-02	2013-08-02	57700.00
83745	SPICER	DWAIN	2013-08-03	2014-08-01	63470.00
83745	SPICER	DWAIN	2014-08-02	2015-08-01	68550.00
83745	SPICER	DWAIN	2015-08-01	2016-07-31	71980.00
83745	SPICER	DWAIN	2016-08-01	2017-08-01	74140.00
83745	SPICER	DWAIN	2017-08-02	NULL	76360.00
84039	COLEMAN	HANNAH	2014-06-28	2015-06-27	47380.00
84039	COLEMAN	HANNAH	2015-06-27	2016-06-26	51170.00
84039	COLEMAN	HANNAH	2016-06-27	2017-06-27	52700.00
84039	COLEMAN	HANNAH	2017-06-28	NULL	54280.00

Question 3. Write a query to display the first name, last name, street, city, state, and zip code of any customer who purchased a Foresters Best brand top coat between July 15, 2017, and July 31, 2017. If a customer purchased more than one such product, display the customer's information only once in the output. Sort the output by sate, last name, and then first name. (JOIN, Logical/Comparison Operator) (5%)

CUST_FNAME	CUST_LNAME	CUST_STREET	CUST_CITY	CUST_STATE	CUST_ZIP
LUPE	SANTANA	1292 WEST 70TH PLACE	Phenix City	AL	36867
HOLLIS	STILES	1493 DOLLY MADISON CIRCLE	Snow Hill	AL	36778
LISETTE	WHITTAKER	339 NORTHPARK DRIVE	Montgomery	AL	36197
DEANDRE	JAMISON	1571 HANES STREET	Miami	FL	33169
CATHLEEN	WHITMAN	1712 NORTHFIELD DRIVE	Marshallville	GA	31057
SHERIE	STOVER	640 MOUNTAIN VIEW DRIVE	Parksville	KY	40464
BRYCE	HOGAN	1860 IMLACH DRIVE	Newbury	MA	01951
SHELBY	SALAS	486 SUSITNA VIEW COURT	North Tisbury	MA	02568
JERMAINE	HANCOCK	1627 SAUNDERS ROAD	Ellicott City	MD	21041
WHITNEY	WHITFIELD	1259 RHONE STREET	Phippsburg	ME	04567
MONROE	ALLISON	272 SCHODDE STREET	Kalamazoo	MI	49002
DARLEEN	PARRA	561 COLLIE HILL WAY	Madison	MS	39130
CLINTON	AGUIRRE	1651 VANGUARD DRIVE	Franklinville	NC	27248
TOMMIE	PALMER	933 ELCADORE CIRCLE	Arapahoe	NC	28510
JEFFEREY	MCBRIDE	1043 ROCKRIDGE DRIVE	Glenwood	NJ	07418
SIDNEY	GARZA	772 SHEPPARD DRIVE	Fair Harbor	NY	11706
TAMELA	GUIDRY	1873 BAXTER ROAD	Brooklyn	NY	11252
KAREN	LEVINE	1534 PALMER COURT	Cincinnati	OH	45218
STEPHENIE	MCKENZIE	1039 DELAWARE PLACE	Wilkes Barre	PA	18763
LAN	NICHOLS	367 LAKEVIEW DRIVE	Pittsburgh	PA	15262
KASEY	SOSA	975 WEST 96TH AVENUE	Kinzers	PA	17535
SHELBY	THAYER	1634 RUANE ROAD	Bordeaux	SC	29835
WILSON	BELL	1127 CUNNINGHAM STREET	Louisville	TN	37777
RENATE	LADD	652 LEWIS STREET	Crystal City	VA	22202
MELONIE	JIMENEZ	848 DOWNEY FINCH LANE	East Monkton	VT	05443

Question 4. Write a query to display the employee number, last name, email address, title, and department name of each employee whose job title ends in the word "ASSOCIATE". Sort the output by department name and employee title. (JOIN, Logical/Comparison Operator) (5%)

Totally, it should be 168 rows. Top several rows are shown as that in the following figure.

EMP_NUM	EMP_LNAME	EMP_EMAIL	EMP_TITLE	DEPT_NAME
84386	RIVERA	D.RIVERA76@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
83661	FINN	D.FINN87@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
83583	ROLLINS	M.ROLLIN99@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
84526	LASSITER	F.LASSIT8@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
83517	ALBRIGHT	SO.ALBRI96@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
83378	DUNHAM	F.DUNHAM5@LGCOMPANY.COM	ASSOCIATE	ACCOUNTING
83341	CORTEZ	C.CORTEZ85@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84240	COLEMAN	R.COLEMA91@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84023	KEYES	J.KEYES10@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84383	WASHINGTON	L.WASHIN98@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84265	CASH	P.CASH79@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84007	PETERS	L.PETERS80@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84476	BLANKENSHIP	F.BLANKE85@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
83415	WHALEN	C.WHALEN5@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84599	FENTON	E.FENTON3@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
84394	STONE	P.STONE75@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...
83721	EASON	K.EASON99@LGCOMPANY.COM	ASSOCIATE	CUSTOMER SERVI...

Question 5. Write a query to display the number of products within each base and type combination, sorted by base and then by type. (Aggregation functions, group by) (5%)

PROD_BASE	PROD_TYPE	NUMPRODUCTS
Solvent	Exterior	67
Solvent	Interior	83
Water	Exterior	39
Water	Interior	63

Question 6. Write a query to display the brand ID, brand name, and average price of products of each brand. Sort the output by brand name. Results are shown with the average price rounded to two decimal places. (join, aggregation) (5%)

BRAND_ID	BRAND_NAME	AVGPRICE
33	BINDER PRIME	16.12
29	BUSTERS	22.59
23	FORESTERS BEST	20.94
27	HOME COMFORT	21.80
35	LE MODE	19.22
30	LONG HAUL	20.12
28	OLDE TYME QUALITY	18.33
25	STUTTENFURST	16.47
31	VALU-MATTE	16.84

Question 7. Write a query to display the department number, department name, department phone number, employee number, and last name of each department manager. Sort the output by department name. (Join) (5%)

DEPT_NUM	DEPT_NAME	DEPT_PHONE	EMP_NUM	EMP_LNAME
600	ACCOUNTING	555-2333	84583	YAZZIE
250	CUSTOMER SERVI...	555-5555	84001	FARMER
500	DISTRIBUTION	555-3624	84052	FORD
280	MARKETING	555-8500	84042	PETTIT
300	PURCHASING	555-4873	83746	RANKIN
200	SALES	555-2824	83509	STOVER
550	TRUCKING	555-0057	83683	STONE
400	WAREHOUSE	555-1003	83759	CHARLES

Question 8. Write a query to display the brand ID, brand name, brand type, and average price (round to 2 decimal places) of products for the brand that has the largest average product price. Results are shown with the average price rounded to two decimal places. (having) (7%)

BRAND_ID	BRAND_NAME	BRAND_TYPE	AVGPRICE
29	BUSTERS	VALUE	22.59

Question 9. Write a query to display the manager name, department name, department phone number, employee, customer name, invoice date, and invoice total for the department manager of the employee who made a sale to a customer whose last name is Hagan on May 18, 2017. (Join) (7%)

EMP_FNAME	EMP_LNAME	DEPT_NAME	DEPT_PHONE	EMP_FNAME	EMP_LNAME	CUST_FNAME	CUST_LNAME	INV_DATE	INV_TOTAL
FRANKLYN	STOVER	SALES	555-2824	THURMAN	WILKINSON	DARELL	HAGAN	2017-05-18	315.04

Question 10. The Binder Prime Company wants to recognize the employee who sold the most of its products during a specified period. Write a query to display the employee number, employee first name, employee last name, email address, and total units sold for the employee who sold the most Binder Prime brand products between November 1, 2017, and December 5, 2017. If there is a tie for most units, sort the output by employee last name. (nested query) (7%)

emp_num	emp_fname	emp_lname	emp_email	total
84134	ROSALIE	GARLAND	G.ROSALI98@LGCOMPANY.COM	23
83850	RUSTY	MILES	M.RUSTY95@LGCOMPANY.COM	23

Question 11. LargeCo is planning a new promotion in Alabama (AL) and wants to know about the largest purchases made by customers in that state. Write a query to display the customer code, customer first name, last name, full address, invoice date, and invoice total of the largest purchase made by each customer in AL. Be certain to include any customers in AL who have never made a purchase; their invoice date should be NULL and the invoice totals should display as 0. Sort the results by customer last name, and then first name. (union) (7%)

Totally, there are 50 rows. The top several rows are shown as that in the following figure.

cust_code	cust_fname	cust_lname	cust_street	cust_city	cust_state	cust_zip	inv_date	Largest Invoice
903	ROBIN	ADDISON	323 LORETTA PLACE	Mobile	AL	36693	2017-08-26	230.63
643	NINA	ALLEN	680 RED TALON DRIVE	Robertsdale	AL	36574	2017-06-21	11.99
295	DORTHY	AUSTIN	829 BIG BEND LOOP	Diamond Shamrock	AL	36614	2017-04-24	589.75
393	FOSTER	BERNAL	1299 EAST 3RD AVENUE	Birmingham	AL	35280	NULL	0.00
853	GAYLORD	BOLTON	1069 LUGENE LANE	Montgomery	AL	36131	2017-11-25	372.68
925	ALANA	BOOKER	1874 I STREET	Mccullough	AL	36502	2017-12-12	208.85
1248	LISA	BRADY	491 LOWLAND AVENUE	Daphne	AL	36577	2017-12-05	414.47
538	CHIQUITA	CALDWELL	1501 BRIGGS COURT	Normal	AL	35762	2017-05-26	143.90
89	MONICA	CANTRELL	697 ADAK CIRCLE	Loachapoka	AL	36865	2017-03-31	516.58
1233	NATHALIE	CHURCH	1802 SNOWY OWL CIRCLE	Napier Field	AL	36303	2017-11-24	160.96
304	GERTRUDE	CONNORS	1042 PLEASANT DRIVE	Georgiana	AL	36033	2017-12-29	376.32
1131	CARMA	CORNETT	767 CHISANA WAY	Killen	AL	35645	2017-10-25	265.12
1407	FELICIA	CRUZ	643 TURNAGAIN PARKW...	Coalburg	AL	35068	2018-01-06	387.93
1068	ELIZA	CURRIE	778 LOUDERMILK CIRCLE	Panola	AL	35477	2017-10-10	365.84
820	MARCELA	DUGAN	1785 DORIS PLACE	Sylacauga	AL	35150	2017-08-04	195.34

Question 12. The purchasing manager is still concerned about the impact of price on sales. Write a query to display the brand name, brand type, product SKU, product

description, and price of any products that are not a premium brand, but that cost more than the most expensive premium brand products. (subqueries) (7%)

brand_name	brand_type	prod_sku	prod_descript	prod_price
LONG HAUL	CONTRACTOR	1964-OUT	Fire Resistant Top Coat, for Interior Wood	78.49

Part II

Use the database, MOVIECO, to finish the following two questions.

Question 13. Trigger (15%) : Create a trigger named **trg_mem_balance** that will maintain the correct value in the membership balance in the MEMBERSHIP table when videos are returned late. The trigger should execute as an AFTER trigger when the due date or return date attributes are updated in the DETAILRENTAL table. The Trigger should satisfy the following conditions:

a. Calculate the value of the late fee prior to the update that triggered this execution of the trigger. The value of the late fee is the days late multiplied by the daily late fee. If the previous value of the late fee was null, then treat it as zero (0).
b. Calculate the value of the late fee after the update that triggered this execution of the trigger. If the value of the late fee is now null, then treat it as zero (0).
c. Subtract the prior value of the late fee from the current value of the late fee to determine the change in late fee for this video rental.
d. If the amount calculated in Part c is not zero (0), then update the membership balance by the amount calculated for the membership associated with this rental.

Question 14. Procedure (15%): Create a stored procedure named **prc_new_rental** to insert new rows in the RENTAL table. The procedure should satisfy the following conditions:

a. The membership number will be provided as parameter.
b. Use a count() function to verify that the membership number exits in the MEMBERSHIP table. If it does not exist, then a message should be displayed that the membership does not exist and no data should be written to the database.
c. If the membership does exist, then retrieve the membership balance and display a message that the balance is the previous balance. (e.g., if the membership has a balance of 5.00, then display “previous balance: $ 5.00.)
d. Insert a new row in the rental table using the rent_num_seq sequence created to generate the value for RENT_NUM, the current system date for the

RENT_DATE value, and the membership number provided as the value for MEM_NUM.