BUAN 6320

DATABASE FOUNDATIONS FOR BUSINESS ANALYTICS

PROJECT REPORT

MADE BY-:

PRAVAL GODRE

Praval.Godre@utdallas.edu

Defining the Problem Statement:

The main goal of this report is get a practical exposure of dealing with relational database models.

STEP 1:

Normalization of the provided datasets.

Data Normalization is a technique used to organize the data in the database. The main of database normalization is to reduce minimize data redundancies along with reducing the likelihood of anomalies.

Normalization of the datasets was done as the first step. This step involved picking up the columns from the datasets and pasting them into new files. This step was used to create 10 tables on excel. The columns to be chosen along with name of the respective tables were taken from the data dictionary.

STEP 2:

Cleaning of each table and filling the missing records.

Once the datasets were normalized to the respective tables, the step of data cleaning was carried out. The tables had various missing records, duplicate values and bad characters that were cleaned to match the data format mentioned in the data dictionary.

a. **Removing bad characters:** Certain tables had records that did not match the format mentioned in the data dictionary and were therefore cleaned. For instance, the LGPRODUCT table had few bad characters present in the PROD_QOH column and were removed.

b. **Removing duplicate values:** Few of the tables had duplicate values of the records present leading to data redundancies. Therefore, these duplicate values were removed in excel itself using the "Remove Duplicates" option.

c. **Filling in the blank records:** Tables like the SALARY HISTORY had which had the column SAL_END with characters like – present for the few records. The records having – in

the SAL_END column was filtered out and the – sign was removed from the selected records.

This was done to distinguish the employees currently working in the organization.

d. **Correcting the date format:** There were few tables having the incorrect format on the data. For instance, the SALARY HISTORY table had incorrect date format present for few of its columns (SAL_FROM and SAL_END). The format of these columns was corrected on Excel from **General to Date** format thereby producing meaningful data.

These were the major cleaning and correcting issues that were administered on the 10 tables on Excel itself.

After the cleaning and formatting, the update number of records for each table were as follows:

TABLE NAME	COLUMNS	NUMBER OF RECORDS
LGBRAND	BRAN_ID, BRAND_NAME, BRAND_TYPE	9
LGCUSTOMER	CUST_CODE, CUST_FNAME, CUST_LNAME, CUST_STREET, CUST_CITY, CUST_STATE, CUST_ZIP, CUST_BALANCE	1362
LGDEPARTMENT	DEPT_NUM, DEPT_NAME, DEPT_MAIL_BOX, DEPT_PHONE_EMP_NUME	8
LGEMPLOYEE	EMP_NUM, EMP_FNAME, EMP_LNAME, EMP_EMAIL, EMP_PHONE, EMP_HIREDATE, EMP_COMM, DEPT_NUM	363
LGINVOICE	INV_NUM, INV_DATE, CUST_CODE, INV_TOTAL, EMPLOYEE_ID	1362
LGLINE	INV_NUM, LINE_NUM, PROD_SKU, LINE_QTY, LINE_PRICE	1347

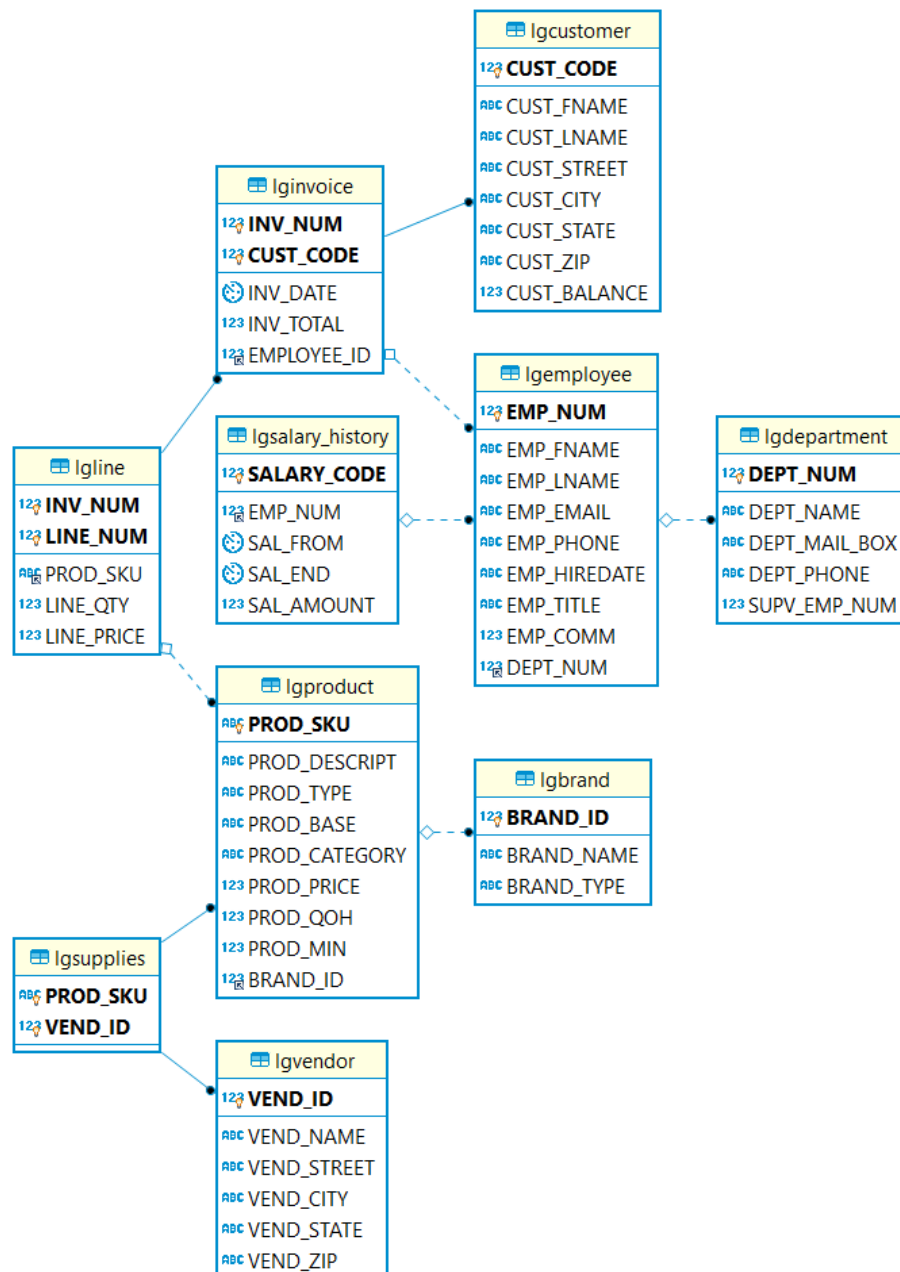|		
LGPRODUCT	PROD_SKU, PROD_DESCRIPT, PROD_TYPE, PROD_BASE, PROD_CATEGORY, PROD_PRICE, PROD_QOH, PROD_MIN, BRAND_ID	252
LGSALARYHISTORY	EMP_NUM, SAL_FROM, SAL_END, SAL_AMOUNT	6710
LGSUPPLIES	PROD_SKU, VEND_ID	252
LGVENDOR	VEND_ID, VEND_NAME, VAND_STREET, VEND_CITY, VEND_STATE, VEND_ZIP	253

STEP 3:

Importing the tables on MSSQL and creating a logical and physical model.

After cleaning the tables on Excel, each of these excel files (tables) was imported into MSSQL using the Import Data option and the respective primary keys and foreign keys were defined for each table.

TABLE NAME	PRIMARY KEY(s)	FOREIGN KEY(s)
LGBRAND	BRAN_ID	-
LGCUSTOMER	CUST_CODE	-
LGDEPARTMENT	DEPT_NUM	EMP_NUM
LGEMPLOYEE	EMP_NUM	DEPT_NUM
LGINVOICE	INV_NUM, CUST_CODE	EMPLOYEE_ID, CUST_CODE
LGLINE	INV_NUM, LINE_NUM	PROD_SKU, INV_NUM
LGPRODUCT	PROD_SKU	BRAND_ID
LGSALARYHISTORY	SAL_CODE	EMP_NUM
LGSUPPLIES	PROD_SKU, VEND_ID	PROD_SKU, VEND_ID
LGVENDOR	VEND_ID	-

We have used Identity Primary Keys (SAL_CODE) in the LGSALARYHISTORY. Also, few tables like LGINVOICE and LGLINE did not have a unique INV_NUM and LINE_NUM respectively, therefore, a composite primary key has created for these two tables.

After the tables were created, a logical and physical model was prepared in MSSQL describing the relationship between in the tables using an Entity Relationship Diagram (ERD).

LOGICAL AND PHYSICAL MODEL USING DBEAVER

STEP 5:

SQL queries and their respective results

This step involved writing SQL queries to retrieve the desired results for the following questions.

1. Write a query to display the current salary for each employee in department 300. Assume that only current employees are kept in the system, and therefore the most current salary for each employee is the entry in the salary history with a NULL end date. Sort the output in descending order by salary amount.

QUERY:

```
SELECT emp.emp_num as EmployeeNo, emp_fname as EmployeeFistName, emp.emp_lname as EmpLastName, s.sal_amount as CurrentSal
FROM lgemployee emp join lgsalary_history s ON emp.emp_num = s.emp_num WHERE sal_end is Null and DEPT_NUM= 300
order by SAL_AMOUNT desc;
```

OUTPUT:

	EmployeeNo	EmployeeFistName	EmpLastName	CurrentSal
1	83746	SEAN	RANKIN	95550
2	84328	FERN	CARPENTER	94090
3	83716	HENRY	RIVERA	85920
4	84432	MERLE	JAMISON	85360
5	83902	ROCKY	VARGAS	79540
6	83695	CARROLL	MENDEZ	79200
7	84500	CHRISTINE	WESTON	78690
8	84594	ODELL	TIDWELL	77400
9	83910	LAUREN	AVERY	76110
10	83359	MERLE	WATTS	72240
11	83790	LAVINA	ACEVEDO	72000
12	83433	RONNA	NORWOOD	68870
13	84521	DELFINA	JUDD	66000
14	83653	LEEANN	HORN	61920
15	83738	PORTER	STACY	58200
16	83788	LANA	DOWDY	56760
17	83867	TRACIE	KELLY	56750
18	84234	LUISA	MINER	54720
19	83637	TANIKA	CRANE	52870
20	83877	STEPHAINE	DUNLAP	52650
21	84035	HAL	FISHER	51600
22	83729	CORRINA	RAMEY	48500
23	83732	SAMMY	DIGGS	44720
24	83644	WILLA	MAXWELL	43200
25	83312	ROSALBA	BAKER	42400

Total record count: 25

2. Write a query to display the starting salary for each employee. The starting salary would be the entry in the salary history with the oldest salary start date for each employee. Sort the output by employee number.

QUERY:

```
SQLQuery5.sql - LA...P-T9G6EED7\a (69))  + X
SELECT emp.EMP_NUM, emp.EMP_FNAME, emp.EMP_LNAME, sal.SAL_AMOUNT from LGEMPLOYEE emp inner join LGSALARY_HISTORY sal
on emp.EMP_NUM= sal.EMP_NUM where SAL_FROM = (Select min(sal_from) from
LGSALARY_HISTORY sal WHERE emp.EMP_NUM=sal.EMP_NUM) order by emp.EMP_NUM;
```

OUTPUT:

#	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
1	83304	TAMARA	MCDONALD	19770
2	83308	CONNIE	LOVE	11230
3	83312	ROSALBA	BAKER	39260
4	83314	CHAROLETTE	DAVID	15150
5	83318	DARCIE	PECK	22330
6	83321	ANGELINA	FARMER	18250
7	83332	WILLARD	LONG	23380
8	83341	CHRISTINE	CORTEZ	14510
9	83347	QUINTIN	WINN	17010
10	83349	JENNIFFER	SINGH	21220
11	83359	MERLE	WATTS	25370
12	83366	PHOEBE	BLEDSOE	23200
13	83371	ROXANE	MATHEWS	18140
14	83372	CLAUDINE	DAHL	25780
15	83374	DARRON	TILLEY	16940
16	83378	FELICIA	DUNHAM	34050
17	83382	STELLA	CONKLIN	12270
18	83385	BRODERICK	COLBERT	17350
19	83398	ZACK	GILES	34950
20	83403	FELICITA	PONCE	11700
21	83404	LIZ	FENTON	32360
22	83411	ROGELIO	REDMOND	8990
23	83413	KYRA	MARINO	19250
24	83415	CHASITY	WHALEN	32930
25	83419	EUGENE	SNEED	15940
26	83423	LINDSAY	GOOD	26890
27	83428	SHERRILL	CRUM	10640

#	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
28	83432	SAL	BENAVIDES	20530
29	83433	RONNA	NORWOOD	27850
30	83434	GERALD	WHALEY	40620
31	83437	JUNE	SWANSON	53450
32	83445	QUINTIN	STAHL	46560
33	83446	CHARLINE	CRAFT	20010
34	83451	ROSALIE	ELLIS	10440
35	83453	TAMARA	CARPENTER	31060
36	83465	HELEN	TRUJILLO	39000
37	83476	DULCE	MCCULLO...	19220
38	83477	VELIA	DOTSON	23250
39	83480	DOLLY	KING	22560
40	83495	ABRAHAM	STAHL	7520
41	83503	GERARD	CAMP	17100
42	83504	RICHELLE	WALKER	20770
43	83509	FRANKLYN	STOVER	72360
44	83511	SHERMAN	MICHEL	12900
45	83517	SONDRA	ALBRIGHT	23060
46	83519	TAMMI	KEYES	18840
47	83521	PALMER	DODGE	44210
48	83527	LAN	STROUD	9260
49	83529	LASHANDA	BRITT	18260
50	83532	STUART	KENDRICK	17550
51	83534	BOBBY	NADEAU	24580
52	83537	CLEO	ENGLISH	58700
53	83542	TOSHA	LONGORIA	6940
54	83543	ALICIA	OROZCO	10250

Results | **Messages**

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
55	83545	JOVITA	TAYLOR	16370
56	83547	LON	HARRISON	6520
57	83555	LUANNE	SPEARS	11510
58	83558	CRISTAL	STALLINGS	30500
59	83562	FELICITA	MANUEL	68950
60	83564	EDGARDO	CONWAY	9060
61	83565	LOURDES	ABERNATHY	48080
62	83569	ROSALINE	TOMLINSON	34210
63	83573	VALARIE	BLEDSOE	23400
64	83575	LINDSEY	GILLIAM	26490
65	83583	MARGARETE	ROLLINS	19710
66	83593	ROSANNE	NASH	55250
67	83595	LEEANN	CLINTON	24180
68	83602	BECKY	SPEARS	33120
69	83603	TANIKA	PADGETT	29320
70	83607	ALENE	PEACOCK	24290
71	83608	DARREL	MAYFIELD	15900
72	83609	SHAUN	GIBBS	81950
73	83614	ARIELLE	FARLEY	25070
74	83615	IGNACIO	PECK	17460
75	83618	CHRISTIAN	DUNLAP	47500
76	83621	FONDA	GONZALEZ	50960
77	83624	GERTRUDE	ELLIOT	17870
78	83628	FELICITA	WALTERS	12900
79	83630	BRENDAN	DOVE	9960
80	83636	ROSALIE	SOSA	14690
81	83637	TANIKA	CRANE	11380

Results | **Messages**

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
82	83639	EVETTE	MIRANDA	38220
83	83644	WILLA	MAXWELL	15580
84	83645	BRYON	BRADY	24270
85	83646	NIGEL	BURNS	14080
86	83647	FRANCIS	HALE	20310
87	83649	DELMA	JACOB	51680
88	83650	GERALYN	SHELDON	24090
89	83652	GEORGINA	EASON	20690
90	83653	LEEANN	HORN	42660
91	83656	NATHALIE	BOYLE	21130
92	83658	SANTIAGO	MARINO	15330
93	83661	DAN	FINN	15100
94	83664	STEPHENIE	FINN	24220
95	83669	MITCHELL	RUSHING	6400
96	83677	HERB	MANNING	51010
97	83683	MAYA	STONE	54630
98	83691	ROY	HEWITT	15190
99	83695	CARROLL	MENDEZ	29430
100	83696	MIMI	GONZALEZ	64350
101	83703	TONEY	HELTON	8830
102	83705	JOSE	BARR	70610
103	83707	ARMAND	ALDRICH	6640
104	83713	KRISTEN	LAIRD	40830
105	83716	HENRY	RIVERA	35080
106	83719	JETTIE	HICKS	15020
107	83721	KERRI	EASON	19820
108	83723	MARYLIN	MCCLELLAN	17360

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
109	83724	SHERRYL	PUCKETT	28630
110	83728	SAMANTHA	KERNS	12540
111	83729	CORRINA	RAMEY	16450
112	83731	SHERON	VARGAS	43740
113	83732	SAMMY	DIGGS	23040
114	83733	YASMIN	HORTON	26290
115	83734	INEZ	ROCHA	50910
116	83735	MAVIS	TAYLOR	17380
117	83737	FAUSTO	MICHEL	31000
118	83738	PORTER	STACY	35360
119	83739	YESENIA	BRIGGS	10390
120	83741	VEDA	LOPEZ	29650
121	83745	DWAIN	SPICER	56020
122	83746	SEAN	RANKIN	22200
123	83748	CHRISTOPER	WEIR	9010
124	83757	KENDALL	MORIN	37730
125	83758	OLIVIA	DELEON	23950
126	83759	THOMAS	CHARLES	18020
127	83763	JAIME	FELTON	52570
128	83767	ELLA	WILLARD	32570
129	83770	HOLLIE	CRAIG	35240
130	83773	ARETHA	BAKER	7160
131	83777	KARRIE	COOLEY	15170
132	83783	AARON	CARROLL	32380
133	83785	ELISHA	MCMULLEN	44440
134	83788	LANA	DOWDY	33260
135	83790	LAVINA	ACEVEDO	24000

Results | **Messages**

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
136	83792	WALLY	ANDERSEN	53030
137	83797	GALE	KIRK	18750
138	83808	JACINTO	TAYLOR	46380
139	83813	JERRI	BENAVIDES	13970
140	83818	KRIS	BRIGHT	16660
141	83820	TOM	HOBSON	44320
142	83824	PHYLISS	BOUCHARD	23030
143	83827	SHELBY	PRESLEY	40290
144	83838	LEN	FLORES	16930
145	83839	RONNY	TANNER	40510
146	83841	EDDY	GOULD	27020
147	83844	THOMAS	HILTON	34920
148	83845	MICHEAL	MAYBERRY	9340
149	83850	RUSTY	MILES	47590
150	83858	DAMARIS	MCQUEEN	46350
151	83867	TRACIE	KELLY	41640
152	83868	ROBIN	CAMPBELL	14240
153	83870	MALISSA	CLARK	17880
154	83873	DORTHA	MAYBERRY	24580
155	83877	STEPHAINE	DUNLAP	34590
156	83878	PAT	MARIN	53170
157	83879	CORNELIA	FOOTE	12580
158	83886	CHRISTINE	LEBLANC	35300
159	83893	SHERRILL	CHATMAN	30840
160	83894	GALE	DEWITT	55820
161	83895	RASHIDA	MCNEAL	46500
162	83896	JAMAAL	KENNEDY	33000

Results | Messages

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
163	83901	ZORA	WOODALL	8620
164	83902	ROCKY	VARGAS	58890
165	83903	LEE	CONNOR	52220
166	83904	ISREAL	WHITTAKER	29200
167	83906	DELLA	SIMONS	50300
168	83907	IRVING	SYLVESTER	25550
169	83910	LAUREN	AVERY	13500
170	83914	REYNA	DUKE	21700
171	83917	PALMER	NADEAU	11880
172	83923	GERALD	DOTY	32950
173	83926	JAE	JIMENEZ	34400
174	83929	RETA	SALTER	11160
175	83936	BRADFORD	BRAY	57030
176	83941	TRACIE	OBRIEN	15800
177	83943	TILLIE	DODD	32240
178	83948	SHANNA	BOOKER	26190
179	83952	MONROE	GOODWIN	19890
180	83957	WANDA	HOLLINGS...	29670
181	83961	SHERRILL	FOOTE	36310
182	83962	VALENCIA	WATTS	16190
183	83963	MARVIN	HAYNES	15200
184	83964	HAILEY	SWEENEY	47750
185	83973	EMILY	GREENWO...	10670
186	83977	JUSTIN	SHEARER	16290
187	83978	KASEY	CASH	33730
188	83990	LACEY	HINKLE	67950
189	83991	PANSY	BOLTON	9290

Results | Messages

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
190	83993	SANG	CORTES	45440
191	83995	KASEY	CRAFT	10040
192	83996	DOTTIE	WESTON	9340
193	83998	EVELIA	MOCK	38220
194	84001	MICHAELA	FARMER	23610
195	84005	ALIDA	BLACKWELL	20850
196	84007	LELA	PETERS	6400
197	84009	KERI	DUBOIS	18770
198	84011	LOUELLA	ROBERSON	8920
199	84017	IGNACIO	WALDRON	28570
200	84021	JAROD	DICKINSON	49200
201	84023	JAROD	KEYES	40970
202	84024	SANJUANITA	FIELDS	99510
203	84031	GAYLE	PECK	8520
204	84035	HAL	FISHER	23700
205	84039	HANNAH	COLEMAN	47380
206	84041	DEE	STILES	14850
207	84042	LORRI	PETTIT	38140
208	84046	FRANCESCO	ELLIOT	40570
209	84047	EMMITT	HOPPER	15840
210	84049	LANE	BRANDON	64880
211	84052	RUBY	FORD	23730
212	84055	PHILLIS	SHAW	70990
213	84058	CHRISTOPER	HAWKINS	16480
214	84067	SAMANTHA	ALBRIGHT	74600
215	84078	DIEGO	ERWIN	42160
216	84085	JOSEFA	MCGHEE	24700

Results | Messages

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
217	84093	JUSTIN	HULL	34390
218	84094	YONG	MCDONALD	32550
219	84098	STELLA	PHELPS	56500
220	84100	GILLIAN	PADGETT	20040
221	84105	BERTRAM	MOCK	50840
222	84106	FELICE	SAMUEL	47480
223	84110	BRAIN	ROY	49000
224	84114	CLEMENT	VOGT	10980
225	84116	CARLOTTA	ASH	13600
226	84120	ALANA	FOREMAN	13450
227	84123	MARGOT	PLATT	10960
228	84130	ADA	JOYNER	12910
229	84132	BRENDAN	GUERRA	12740
230	84133	QUINN	ROSEN	39790
231	84134	ROSALIE	GARLAND	43540
232	84156	ISSAC	CORTEZ	15780
233	84162	GIL	OSBORN	41630
234	84163	GWEN	EASLEY	38830
235	84178	ALYSON	WILLARD	29230
236	84180	JANNETTE	CURRIE	36310
237	84184	DEANDRE	MOSLEY	32380
238	84185	MADGE	KANE	13480
239	84186	SOFIA	BARTLETT	8750
240	84187	VILMA	BUNCH	27500
241	84191	ROXANA	HOLBROOK	41760
242	84193	ANTOINETTE	RUIZ	40300
243	84196	LELIA	ALVARADO	9650

Results | Messages

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
244	84199	NAN	CORNETT	9670
245	84202	CLEVELAND	SAMUELS	8690
246	84204	EZEKIEL	JAMISON	52310
247	84205	CESAR	COLE	16810
248	84206	NANCY	HEALY	9960
249	84208	GIL	BRUNER	10090
250	84213	LES	FIELD	13920
251	84214	CECELIA	LACEY	40700
252	84219	THURMAN	WILKINSON	40350
253	84223	ANDY	FARRELL	19500
254	84224	ELTON	LOVE	83070
255	84233	ALIDA	WELCH	9990
256	84234	LUISA	MINER	21250
257	84235	JANNETTE	HARRISON	26120
258	84236	ALISHA	CRUM	61830
259	84238	EDNA	WILEY	10520
260	84240	REID	COLEMAN	23080
261	84248	DANICA	CASTLE	44520
262	84249	CORRINA	BURKE	15020
263	84251	MONROE	KNUTSON	8020
264	84253	SADIE	COVINGTON	36230
265	84254	CLINTON	SEYMOUR	8810
266	84256	EDGAR	DODGE	40710
267	84259	CORINA	CUNNINGH...	38100
268	84263	BRYON	DANIEL	17180
269	84264	MARCOS	LEVINE	26730
270	84265	PAM	CASH	6530
271	84266	ALANNA	WISE	33520
272	84268	TRISHA	ALVAREZ	61270
273	84276	ROSALIND	VILLARREAL	38770

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
274	84278	IRENA	TOBIN	26880
275	84280	LIZ	AYERS	32520
276	84286	PATRICA	PATE	13210
277	84287	FRANK	PETTY	16150
278	84291	ELEANORE	MACIAS	7130
279	84294	KIETH	MCKENZIE	10070
280	84298	JENNA	LYLES	7670
281	84300	ALEXANDRA	SEAY	10880
282	84302	KIP	OGDEN	16210
283	84306	ROWENA	MEDINA	42290
284	84312	KURTIS	REDMOND	36180
285	84316	BRIDGET	NIXON	22670
286	84320	LUCIO	CAUDILL	26240
287	84327	RUPERT	STONE	9700
288	84328	FERN	CARPENTER	37040
289	84329	MARGOT	HATFIELD	13290
290	84330	DENNA	CRAFT	15270
291	84333	PHILLIS	CONKLIN	48590
292	84334	LINNIE	GOLDMAN	37530
293	84340	LATOYA	KEENAN	51340
294	84343	ELDA	THORPE	20200
295	84356	BERNITA	CONWAY	38930
296	84363	AWILDA	MORRIS	25900
297	84364	LESLIE	VARNER	43540
298	84369	FERNANDO	GRANGER	34760
299	84372	JOSEPHINA	FAULKNER	71940
300	84374	DESSIE	PAULSON	45760
301	84380	ALICIA	SHEARER	12840
302	84383	LUCRETIA	WASHING...	17500
303	84386	DELFINA	RIVERA	13900

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMO
304	84389	STEFANIE	DUKE	8750
305	84392	ALEJANDRA	WHALEY	38960
306	84394	PARIS	STONE	8100
307	84396	ROBERTA	NEFF	30360
308	84397	SUE	ASH	28820
309	84405	LAWANDA	RUCKER	27010
310	84406	KAREN	HINKLE	6190
311	84417	ADELA	CHU	15680
312	84419	MITCHELL	ROLAND	27330
313	84420	DOUG	CAUDILL	40090
314	84432	MERLE	JAMISON	53340
315	84441	BURT	HEALY	12930
316	84442	ANGEL	GREGORY	17890
317	84448	WALTON	BATEMAN	11230
318	84453	SANTIAGO	BRIGGS	51000
319	84458	JERRY	WEIR	13140
320	84459	EMILIO	GILLIAM	25750
321	84460	LAVERNE	TANNER	7550
322	84462	ELYSE	HOPKINS	7300
323	84463	LUANNE	LOVE	18320
324	84465	GRAIG	BLOCK	32440
325	84476	FERNE	BLANKENS...	10720
326	84477	GALEN	DUGAN	31590
327	84484	MARVIN	GLOVER	50850
328	84488	SHANA	CORNELL	41720
329	84494	NICOLETTE	FIGUEROA	10290
330	84500	CHRISTINE	WESTON	36940
331	84501	DENNA	MCGRAW	47830
332	84502	SAL	FITZPATRI...	14000
333	84506	AISHA	ERICKSON	77390

	EMP_NUM	EMP_FNAME	EMP_LNAME	SAL_AMOUNT
334	84510	KRISTINA	TRUJILLO	16080
335	84511	ILEANA	CHURCH	106640
336	84515	TONEY	TANNER	21510
337	84520	ALVA	NICHOLS	49440
338	84521	DELFINA	JUDD	62210
339	84526	FERNE	LASSITER	35210
340	84530	HAL	HINKLE	70490
341	84543	KYLA	GOODMAN	22630
342	84544	GERTRUDE	BAUER	11940
343	84545	IRENA	BURKETT	45950
344	84554	MICKEY	CRAFT	30340
345	84557	PRUDENCE	TANNER	44740
346	84563	STEPHAINE	YODER	10940
347	84564	WILLY	OAKES	31520
348	84565	CHELSEA	GAY	44330
349	84576	DARELL	MERCER	16170
350	84580	LORENZO	PECK	25200
351	84583	LINDSEY	YAZZIE	44660
352	84585	MARIAM	HANCOCK	7730
353	84586	REYNA	BERNAL	7650
354	84587	LORRI	SHANNON	26860
355	84589	TRICIA	CONNORS	27010
356	84591	LEOPOLDO	SMALL	25430
357	84592	CRISTAL	CALLAHAN	24260
358	84593	TONJA	PERKINS	34160
359	84594	ODELL	TIDWELL	22790
360	84596	PRECIOUS	FARMER	36350
361	84597	WILFORD	BURGOS	23460
362	84598	EDNA	SARGENT	126000
363	84599	ELEANORE	FENTON	34850

Total Record Count: 363

3. Write a query to display the invoice number, line numbers, product SKUs, product

    descriptions, and brand ID for sales of sealer and top coat products of the same brand

    on the same invoice.

QUERY:

```
SELECT l.inv_num, l.line_num, p.prod_sku, p.prod_descript, p.brand_id FROM (lgline l inner join lgproduct p
ON l.prod_sku = p.prod_sku) inner join (lgline l2 inner join lgproduct p2 ON l2.prod_sku = p2.prod_sku) ON l.inv_num = l2.inv_num
WHERE p.brand_id = p2.brand_id AND p.prod_category = 'Sealer' AND p2.prod_category = 'Top Coat';
```

OUTPUT:

Results	Messages			
inv_num	line_num	prod_sku	prod_descript	brand_id

Total Record Count: 0

The query returned no results. There a visual inspection was carried out to confirm that

there are no records present of sealer and top coats product category of the same brand on

the same invoice number.

4. The Binder Prime Company wants to recognize the employee who sold the most of

    their products during a specified period. Write a query to display the employee

    number, employee first name, employee last name, e-mail address, and total units

    sold for the employee who sold the most Binder Prime brand products between

    November 1, 2015, and December 5, 2015. If there is a tie for most units sold, sort the

    output by employee last name.

QUERY:

```
SELECT INV.EMPLOYEE_ID, COUNT(*) AS "TOTAL UNITS SOLD" FROM LGINVOICE INV
INNER JOIN LGLINE L ON L.INV_NUM = INV.INV_NUM INNER JOIN LGPRODUCT P ON P.PROD_SKU = L.PROD_SKU
INNER JOIN LGBRAND B ON B.BRAND_ID=P.BRAND_ID
WHERE (INV.INV_DATE >= CONVERT(VARCHAR(20), '11/01/2015', 120) AND
INV.INV_DATE <= CONVERT(VARCHAR(20), '12/05/2015', 120))
AND UPPER (B.BRAND_NAME) ='BINDER PRIME'
GROUP BY INV.EMPLOYEE_ID;
```

OUTPUT:

Results	Messages
EMPLOYEE_ID	TOTAL UNITS SOLD

Total Record Count: 0

The query returned no results. There a visual inspection was carried out to confirm that

there are no records present of sealer and top coats product category of the same brand on

the same invoice number.

5. Write a query to display the customer code, first name, and last name of all customers

    who have had at least one invoice completed by employee 83649 and at least one

    invoice completed by employee 83677. Sort the output by customer last name and

    then first name.

QUERY:

```
SELECT C.CUST_CODE, C.CUST_FNAME, C.CUST_LNAME FROM LGCUSTOMER C
INNER JOIN LGINVOICE INV
ON C.CUST_CODE= INV.CUST_CODE
WHERE INV.EMPLOYEE_ID IN (83649, 83677)
ORDER BY CUST_LNAME, CUST_FNAME;
```

OUTPUT:

	CUST_CODE	CUST_FNAME	CUST_LNAME
1	651	ERASMO	ARELLANO
2	511	CHRISTEN	ASH
3	156	JONAS	BARRERA
4	124	LINDSAY	BARRERA
5	684	WENDI	BEAN
6	499	JERRI	BLAIR
7	640	PHIL	BOWMAN
8	1078	MAURO	BRITTON
9	1154	SONDRA	BROUSSARD
10	1330	BARBARA	BUCKLEY
11	135	CHRISTINE	CAMACHO
12	197	XAVIER	CARDENAS
13	209	SIDNEY	CARR
14	1023	LEEANN	CARSON
15	211	GERALD	CAUDILL
16	206	DENNY	CHANDLER
17	512	JETTIE	CHURCH
18	248	ROXANNE	CONKLIN
19	702	LUCIE	CONWAY
20	1131	CARMA	CORNETT
21	818	HANK	CRAFT
22	445	WILMA	CRAIG
23	347	MARYELLEN	DALTON
24	850	THERESA	DAVENPORT
25	127	CORRINA	DENNIS
26	292	VALARIE	DILLARD
27	1161	ELISA	DOWDY
28	310	EDIE	EASON
29	1436	TRICIA	EDDY
30	1373	MARSHA	ELLIOTT

	CUST_CODE	CUST_FNAME	CUST_LNAME
31	293	CLAIR	ERICKSON
32	1300	FABIOLA	FERREIRA
33	1119	TOD	FINN
34	967	FRANK	FLYNN
35	563	ADELAIDA	FOOTE
36	576	MODESTA	FRAZIER
37	763	PORTER	GILES
38	938	LUISA	GILLIAM
39	815	ROLANDO	HAGEN
40	1062	SUZY	HARDING
41	1223	KYRA	HARTMAN
42	227	KATIE	HEARN
43	489	MARINA	HILTON
44	1359	ISIAH	HORN
45	887	SIDNEY	JENNINGS
46	793	BRODERICK	JOHNSON
47	321	DONA	KEENE
48	399	ROSALIA	KELLY
49	1388	MARCOS	KENDALL
50	1434	JUNE	KEYS
51	1027	KERI	KNOX
52	909	KENDALL	LAM
53	1429	RUSTY	LAWRENCE
54	371	KRISTINE	LONG
55	253	NUMBERS	LOUIS
56	241	MOSES	LUNSFORD
57	522	ALBERTA	LYON
58	1098	JUDITH	MACIAS
59	307	PHILLIS	MATOS
60	437	HILTON	MATTHEWS

	CUST_CODE	CUST_FNAME	CUST_LNAME
61	269	JOSEPHINA	MAURER
62	1146	ANGELO	MAYER
63	652	ELEANORE	MCMULLEN
64	133	RASHIDA	MEDINA
65	194	JORGE	MILLER
66	319	CLEO	MORA
67	1152	FRANK	NORWOOD
68	1175	GEORGIANA	OGDEN
69	605	JACKIE	PEARCE
70	231	ALYCIA	PECK
71	1138	NORMA	PHELPS
72	1140	MA	POOLE
73	851	JAMISON	PRESTON
74	425	ROYCE	PUTNAM
75	439	WANDA	RANSOM
76	896	MURRAY	RHOADES
77	806	AGNES	ROSE
78	1189	ERNA	SAENZ
79	542	SAMMY	SALTER
80	1308	CLEMENT	SHANNON
81	1255	FAUSTO	SLAUGHTER
82	619	CLARENCE	SPICER
83	756	EARL	STATON
84	899	VINCE	STONE
85	433	SHERIE	STOVER
86	1307	APRIL	SUAREZ
87	826	STEFANIE	SWANSON
88	660	FELICITA	SYLVESTER
89	1286	BARNEY	TEAGUE
90	1298	MAURICIO	TRIPP

91	485	DARLINE	VENTURA
92	1423	TANIKA	VILLALOBOS
93	317	RICHARD	VINCENT
94	642	JENNA	WEISS
95	991	TRACEY	WELCH
96	1040	JACKIE	WELLS
97	1360	MICHEL	WHALEN
98	152	LISETTE	WHITTAKER
99	301	BRIDGET	WILLIAMSON

Total Record Count: 99

6. 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 Alabama. Be certain to include any customers in Alabama who have never made a purchase (their invoice dates should be NULL and the invoice totals should display as 0).

QUERY:

```sql
SELECT c.cust_code, cust_fname, cust_lname, cust_street, cust_city, cust_state, cust_zip,inv_date,
inv_total AS "Total of the Largest Purchase" FROM lgcustomer c JOIN lginvoice i ON c.cust_code = i.cust_code WHERE cust_state = 'AL'
AND inv_total = (SELECT Max(inv_total) FROM lginvoice i2 WHERE i2.cust_code = c.cust_code)
UNION SELECT cust_code, cust_fname, cust_lname, cust_street, cust_city, cust_state, cust_zip, NULL, 0 FROM lgcustomer
WHERE cust_state = 'AL' AND cust_code NOT IN (SELECT cust_code FROM lginvoice);
```

OUTPUT:

	cust_code	cust_fname	cust_lname	cust_street	cust_city	cust_state	cust_zip	inv_date	Total of the Largest Purchase
1	89	MONICA	CANTRELL	697 ADAK CIRCLE	Loachapoka	AL	36865	2014-01-12 ...	314
2	152	LISETTE	WHITTAKER	339 NORTHPARK DRIVE	Montgomery	AL	36197	2013-11-19 ...	139
3	169	ROSS	LANG	1991 EASTWIND COURT	Higdon	AL	35979	2013-11-14 ...	177
4	188	LUANNE	GOODWIN	293 KIANA AVENUE	Pinegrove	AL	36507	2013-08-11 ...	202
5	218	LUPE	SANTANA	1292 WEST 70TH PLACE	Phenix City	AL	36867	2013-10-31 ...	270
6	219	CATHI	WHITEHEAD	760 WOODCLIFF DRIVE	Huntsville	AL	35893	2013-11-19 ...	273
7	286	JEANNE	STEINER	1974 SCHUSS DRIVE	Carrollton	AL	35447	2013-03-20 ...	226
8	295	DORTHY	AUSTIN	829 BIG BEND LOOP	Diamond Shamrock	AL	36614	2013-11-01 ...	86
9	304	GERTRUDE	CONNORS	1042 PLEASANT DRIVE	Georgiana	AL	36033	2013-12-29 ...	376
10	364	DELLA	MAYO	543 STELIOS CIRCLE	Birmingham	AL	35214	2013-07-08 ...	94
11	367	HOLLIS	STILES	1493 DOLLY MADISON CIRCLE	Snow Hill	AL	36778	2013-07-17 ...	312
12	380	ALBINA	ENGLE	670 UPPER BOWERY LANE	Clanton	AL	35045	2013-04-12 ...	124
13	393	FOSTER	BERNAL	1299 EAST 3RD AVENUE	Birmingham	AL	35280	2013-09-15 ...	158
14	416	TATIANA	HOWE	1650 ALL STAR CIRCLE	Sunny South	AL	36769	2013-09-05 ...	75
15	458	ELOISA	VALLE	182 BRANDON STREET	Abel	AL	36258	2013-12-19 ...	224
16	487	ROSENDO	REYNA	1009 HAMPSHIRE BOULEVA...	Gunter Afs-Eci	AL	36118	2013-12-01 ...	41
17	538	CHIQUITA	CALDWELL	1501 BRIGGS COURT	Normal	AL	35762	2013-05-26 ...	143
18	585	LORRAINE	HANNAH	173 LAUREL STREET	Orange Beach	AL	36561	2013-12-31 ...	113
19	643	NINA	ALLEN	680 RED TALON DRIVE	Robertsdale	AL	36574	2013-06-21 ...	11
20	696	ALISHA	TOMLINSON	1985 EAST 52ND AVENUE	Catherine	AL	36728	2013-07-07 ...	37
21	738	ALIDA	HANSEN	792 FERGY CIRCLE	Furman	AL	36741	2013-08-02 ...	110
22	753	CECIL	ESPARZA	1928 VALLEY VISTA CIRCLE	Mulga	AL	35118	2013-07-19 ...	179
23	780	LARISSA	POOL	574 ADAK CIRCLE	Decatur	AL	35609	2014-01-06 ...	243
24	798	LETICIA	HEBERT	1244 DEER PARK DRIVE	Shorterville	AL	36373	2013-07-29 ...	41

	cust_code	cust_fname	cust_lname	cust_street	cust_city	cust_state	cust_zip	inv_date	Total of the Largest Purchase
25	820	MARCELA	DUGAN	1785 DORIS PLACE	Sylacauga	AL	35150	2013-08-04 ...	195
26	853	GAYLORD	BOLTON	1069 LUGENE LANE	Montgomery	AL	36131	2013-11-25 ...	372
27	855	AUBREY	GLOVER	907 GOLD CLAIM DRIVE	Honoraville	AL	36042	2013-08-15 ...	300
28	881	LIZ	MCQUEEN	1203 VALLEY STREET	Brewton	AL	36426	2013-08-26 ...	318
29	886	ROSARIO	STOKES	959 SUNRISE DRIVE	Hightower	AL	36263	2013-08-23 ...	120
30	903	ROBIN	ADDISON	323 LORETTA PLACE	Mobile	AL	36693	2013-10-24 ...	128
31	915	CARLOTTA	KNIGHT	335 HESTERBERG ROAD	Winslow	AL	36003	2013-11-14 ...	212
32	925	ALANA	BOOKER	1874 I STREET	Mccullough	AL	36502	2013-12-12 ...	208
33	931	PANSY	MAYBERRY	728 LYNKERRY CIRCLE	Enterprise	AL	36330	2013-09-10 ...	79
34	979	IMOGENE	MAYES	1017 HARCA STREET	Sylacauga	AL	35150	2013-09-15 ...	245
35	1029	JOHNETTA	ROY	1163 GIROUX CIRCLE	Dauphin Island	AL	36528	2013-09-28 ...	224
36	1068	ELIZA	CURRIE	778 LOUDERMILK CIRCLE	Panola	AL	35477	2013-10-20 ...	249
37	1100	ELEANORE	SAUNDERS	820 QUARTZ AVENUE	Silver Cross	AL	36538	2013-10-18 ...	326
38	1131	CARMA	CORNETT	767 CHISANA WAY	Killen	AL	35645	2014-01-14 ...	237
39	1172	ADELE	PERKINS	1192 RICHARDSON VISTA R...	Sylacauga	AL	35150	2013-12-19 ...	193
40	1233	NATHALIE	CHURCH	1802 SNOWY OWL CIRCLE	Napier Field	AL	36303	2013-11-24 ...	160
41	1248	LISA	BRADY	491 LOWLAND AVENUE	Daphne	AL	36577	2014-01-08 ...	371
42	1253	TRENTON	PRESLEY	1905 ANCHOR PARK CIRCLE	Montgomery	AL	36131	2013-11-28 ...	348
43	1264	MICHAELA	RICHARD	44 RASMUSSON STREET	Georgetown	AL	36521	2013-12-02 ...	143
44	1275	ELEANORE	NEFF	556 CACHE DRIVE	Saint Stephens	AL	36569	2013-12-06 ...	428
45	1350	LATONYA	KAY	61 LOUSSAC DRIVE	Seaboard	AL	36529	2013-12-24 ...	275
46	1407	FELICIA	CRUZ	643 TURNAGAIN PARKWAY	Coalburg	AL	35068	2014-01-06 ...	387
47	1412	EVALYN	HEWITT	293 TIMOTHY CIRCLE	Tallassee	AL	36078	2014-01-07 ...	200
48	1439	SHANNA	PRITCHARD	1448 ARLENE DRIVE	Lavaca	AL	36911	2014-01-14 ...	71
49	1440	LONNIE	MANNING	644 PARKER PLACE	Big Cove	AL	35763	2014-01-15 ...	377
50	1443	ALYSON	SELF	772 LUPIN DRIVE	Motley	AL	36276	2014-01-16 ...	114

Total Record count: 50

7. One of the purchasing managers is interested in the impact of product prices on the sale of products of each brand. Write a query to display the brand name, brand type, average price of products of each brand, and total units sold of products of each brand. Even if a product has been sold more than once, its price should only be included once in the calculation of the average price. However, you must be careful because multiple products of the same brand can have the same price, and each of those products must be included in the calculation of the brand's average price.

QUERY:

```
SELECT brand_name, brand_type, Round(avgprice,2) AS "Average Price", "Units Sold" FROM lgbrand b JOIN (SELECT brand_id, Avg(prod_price)
FROM lgproduct GROUP BY brand_id) sub1 ON b.brand_id = sub1.brand_id JOIN (SELECT brand_id, Sum(line_qty) AS "Units Sold"
FROM lgproduct p JOIN lgline l ON p.prod_sku = l.prod_sku GROUP BY brand_id) sub2 ON b.brand_id = sub2.brand_id
ORDER BY brand_name;
```

OUTPUT:

Total Record Count: 9

8. 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.

QUERY:

```
SELECT brand_name as BrandName, brand_type as BrandType, prod_sku as SKU_Number, prod_descript as ProdDescription, prod_price as Price
FROM lgproduct p JOIN lgbrand b ON p.brand_id = b.brand_id WHERE brand_type <> 'PREMIUM'
AND prod_price > (SELECT Max(prod_price) FROM lgproduct p JOIN lgbrand b ON p.brand_id = b.brand_id WHERE brand_type = 'PREMIUM');
```

OUTPUT:

Total number of record: 1

9. Using SQL descriptive statistics functions calculate the value of the following items:

    a. What are the products that have a price greater than $50?

QUERY:

```
Select * from LGPRODUCT where PROD_PRICE>50;
```

OUTPUT:

	PROD_SKU	PROD_DESCRIPT	PROD_TYPE	PROD_BASE	PROD_CATEGORY	PROD_PRICE	PROD_QOH	PROD_MIN	BRAND_ID
1	1021-MTI	Elastomeric, Exterior, Industrial Grade, Water Ba...	Exterior	Water	Top Coat	62	22	25	35
2	1964-OUT	Fire Resistant Top Coat, for Interior Wood	Interior	Solvent	Top Coat	78	120	10	30
3	3694-XFJ	Epoxy-Modified Latex, Interior, Semi-Gloss (MPI ...	Interior	Water	Top Coat	54	39	25	27

Total Record Count: 3

    b. What is total value of our entire inventory on hand?

QUERY:

```
Select SUM(PROD_PRICE*PROD_QOH) as TOTAL_INVENTORY_VALUE from LGPRODUCT;
```

OUTPUT:

	TOTAL_INVENTORY_VALUE
1	345200

Total Record Count: 1

    c. How many customers do we presently have and what is the total of all customer balances?

QUERY:

```
Select count(CUST_CODE) AS Total_No_Of_Customers, SUM(CUST_BALANCE) as Total_Customer_Balance FROM LGCUSTOMER ;
```

OUTPUT:

	Total_No_Of_Customers	Total_Customer_Balance
1	1362	786704

Total Record Count: 1

      d. What are the top three states that buy the most product in dollars from the

           company?

QUERY:

```
SELECT TOP 3 CUST_STATE AS "TOP STATES", SUM(INV_TOTAL) AS TOTAL_INVOICE_AMOUNT
FROM LGCUSTOMER CUST INNER JOIN LGINVOICE INV
ON CUST.CUST_CODE = INV.CUST_CODE
GROUP BY CUST_STATE
ORDER BY TOTAL_INVOICE_AMOUNT DESC;
```

OUTPUT:

	TOP STATES	TOTAL_INVOICE_AMOUNT
1	PA	38512
2	NY	32156
3	NC	19563

Total Record Count: 3

    10. Using predictive statistics calculate what the predicted forecast of sales for the next

        year based on the INV_DATE (independent) and INV_TOTAL (dependent).

        Remember that you will need to convert the INV_DATE from the MS SQL Server

        stored date value to the expect Julian date, since numbers in MS SQL are stored as

the number of days since 1/1/1900 with the fraction as the portion of a day (if you

are using a different DBMS use the appropriate code for conversion.)