Course Title:	Fundamentals of Data Engineering
Course Number:	COE 848
Semester/Year (e.g.F2016)	W2021

Instructor:	Dr. Faezeh Ensan

Assignment/Lab Number:	4
Assignment/Lab Title:	Manipulating Data

Submission Date:	March 16, 2021
Due Date:	March 16, 2021

Student LAST Name	Student FIRST Name	Student Number	Section	Signature*
Shreekant	Vatsal	500771363	01	VS

Database Values:

The following values and data types were built in conjunction with labs 1, 2 and 3.

Customer:

Customer_ID	Customer_First_Name	Customer_Last_Name	Num_Of_Orders
1	Elizabeth A.	Andersen	2
2	Bryn Paul	Dunton	4
3	Stuart	Glasson	6
4	Satomi	Hayakawa	8
5	Amaya	Hernandez-Echevarria	10
6	Madeleine	Kelley	12
7	Mikael	Sandberg	14
8	Naoki	Sato	16
9	Luis	Sousa	18
10	Cornelia	Weiler	20

Figure 1: Table of Customers

Order_Des:

Order_ID	Customer_ID	Order_Cost	Order_Status	Order_Date
1	1	100	Shipped	2020-12-07
2	2	200	Delivered	2020-11-02
3	3	300	In-Progress	2020-07-28
4	4	400	Delivered	2021-03-08
5	5	500	Shipped	2021-02-14
6	6	600	Shipped	2021-01-19
7	7	700	Delivered	2020-10-13
8	8	800	In-Progress	2020-09-11
9	9	900	Shipped	2021-03-11
10	10	1000	Shipped	2021-03-01

Figure 2: Table of Order Description

Product:

Product_ID	Order_ID	Department	Subscription_O	Product_Name	Product_Descriptio	Product_Avg_Ra	Product_Company
1	1	Books	Yes	Greenlights	From the Academy	4	Matthew McConaugh
2	2	Video Games	No	Grand Theft Auto	Grand Theft Auto V	4	Rockstar Games
3	3	Music	Yes	After Hours	The Weeknd's lates	5	The Weeknd
4	4	Electronics	No	SAMSUNG 27-inch	Monitor with Mobil	5	Samsung
5	5	Health & Beaut	Yes	CALVIN KLEIN CKI	Eau De Toilette For	4	Calvin Klein
6	6	Grocery	Yes	McCafe K-Cup po	McCafe Premium R	5	McDonalds
7	7	Men's Clothing	No	Flannel Shirt	Amazon Essentials	3	Amazon Essentials
8	8	Sports	No	Golf Balls	PrideSports Practic	3	PrideSports
9	9	Automotive	No	Michelin Tires	Michelin Anakee III	5	Michelin
10	10	Electronics	No	Samsung Ear Buds	Samsung Galaxy Bu	5	Samsung

Product_Availabi	Warehouse_Offering	Listing_Date
No	Yes	2020-12-31
Yes	No	2020-12-24
No	No	2020-06-30
No	Yes	2020-08-04
Yes	No	2020-08-05
No	No	2020-08-08
Yes	Yes	2020-08-09
Yes	Yes	2020-08-10
No	No	2020-08-01
Yes	No	2020-08-06

Figure 3: Table of Products

Deals:

Deal_ID	Product_ID	Product_Sell_Pe	Price	Dicount_Rates(%)
1	1	Best-Seller	90	10
2	2	Best-Seller	180	10
3	3	Avg-Seller	240	20
4	4	Best-Seller	360	10
5	5	Worse-Seller	300	40
6	6	Best-Seller	560	10
7	7	Avg-Seller	560	20
8	8	Best-Seller	720	10
9	9	Avg-Seller	720	20
10	10	Best-Seller	700	30

Figure 4: Table of Deals

Review:

Review_ID	Customer_ID	Product_ID	Rating_Stars	Comment_Sentiment
1	1	1	4	Good
2	2	2	4	Moderate
3	3	3	5	Good
4	4	4	5	Good
5	5	5	4	Moderate
6	6	6	5	Good
7	7	7	3	Moderate
8	8	8	3	Bad
9	9	9	5	Good
10	10	10	3	Moderate

Figure 5: Table of Reviews

Billing_Detail:

Billing_ID	Customer_ID	Street_Num	Street_Name	City	Province	Postal_Code	Country
1	1	123	1st Street	Seattle	WA	99999	USA
2	2	123	2nd Street	Boston	MA	99999	USA
3	3	123	3rd Street	Los Angelas	CA	99999	USA
4	4	123	4th Street	New York	NY	99999	USA
5	5	123	5th Street	Minneapolis	MN	99999	USA
6	6	123	6th Street	Milwaukee	WI	99999	USA
7	7	123	7th Street	Boise	ID	99999	USA
8	8	123	8th Street	Portland	OR	99999	USA
9	9	123	9th Street	Salt Lake City	UT	99999	USA
10	10	123	10th Street	Chicago	IL	99999	USA

Figure 6: Table of Billing Detail

Places:

Customer_ID	Order_ID	Num_Of_Order	Order_Date
1	1	2	2020-12-07
2	2	4	2020-11-02
3	3	6	2020-07-28
4	4	8	2021-03-08
5	5	10	2021-02-14
6	6	12	2021-01-19
7	7	14	2020-10-13
8	8	16	2020-09-11
9	9	18	2021-03-11
10	10	20	2021-03-01

Figure 7: Table of 'Places' Relationship

Contains:

Order_ID	Product_ID
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Figure 8: Table of 'Contains' Relationship

Display:

Deal_ID	Product_ID
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Figure 9: Table of 'Display' Relationship

Has:

Product_ID	Review_ID
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Figure 10: Table of 'Has' Relationship

Write:

Review_ID	Customer_ID
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10

Figure 11: Table of 'Write' Relationship

The following are the 10 tasks/queries listed in lab 1:

1. Products offered on the company website, their description, their category and their manufacturer.
   The product table is used here to display the 5 attributes listed below.

Product_ID	Product_Name	Product_Description	Department	Product_Company
1	Greenlights	From the Academy Award®–winning acto	Books	Matthew McConaughey
2	Grand Theft Auto V	Grand Theft Auto V Premium Online Editi	Video Games	Rockstar Games
3	After Hours	The Weeknd's latest chapter, After Hours	Music	The Weeknd
4	SAMSUNG 27-inch M5 Smart Monitor	Monitor with Mobile Connectivity, FHD,	Electronics	Samsung
5	CALVIN KLEIN CKIN2U	Eau De Toilette For Him 50MI	Health & Beauty	Calvin Klein
6	McCafe K-Cup pod	McCafe Premium Roast Medium Dark Sin	Grocery	McDonalds
7	Flannel Shirt	Amazon Essentials Men's Slim-Fit Long-Sl	Men's Clothing	Amazon Essentials
8	Golf Balls	PrideSports Practice Golf Balls, Foam, 12-	Sports	PrideSports
9	Michelin Tires	Michelin Anakee III Radial Tire - 120/70R	Automotive	Michelin
10	Samsung Ear Buds	Samsung Galaxy Buds Live - Black, SM-R1	Electronics	Samsung

Figure 12: Table of Task #1

```sql
SELECT Product.Product_ID, Product.Product_Name,
Product.Product_Description, Product.Department,
Product.Product_Company
FROM Product;
```

Figure 13: SQL Code of Task #1

2. Products listed by their ratings.
   The product table is used here to display the 3 attributes listed below. The product average rating are displayed in a descending order.

Product_ID	Product_Name	Product_Avg_Ratings
10	Samsung Ear Buds	5
9	Michelin Tires	5
6	McCafe K-Cup pod	5
4	SAMSUNG 27-inch M5 Smart Monitor	5
3	After Hours	5
5	CALVIN KLEIN CKIN2U	4
2	Grand Theft Auto V	4
1	Greenlights	4
8	Golf Balls	3
7	Flannel Shirt	3

Figure 14: Table of Task #2

```
1  SELECT Product.Product_ID, Product.Product_Name,
2  Product.Product_Avg_Ratings
3  FROM Product
4  ORDER BY Product.Product_Avg_Ratings DESC;
```

Figure 15: SQL Code of Task #2

3. Are highly rated products built by the same manufacturer?
   The highly rated products are displayed here with their respective attributes and the ratings are yet again sorted in a descending order to clarity.

Product_ID	Product_Name	Product_Company	Product_Avg_Ratings
10	Samsung Ear Buds	Samsung	5
9	Michelin Tires	Michelin	5
6	McCafe K-Cup pod	McDonalds	5
4	SAMSUNG 27-inch M5 Smart Monitor	Samsung	5
3	After Hours	The Weeknd	5
5	CALVIN KLEIN CKIN2U	Calvin Klein	4
2	Grand Theft Auto V	Rockstar Games	4
1	Greenlights	Matthew McConaughey	4
8	Golf Balls	PrideSports	3
7	Flannel Shirt	Amazon Essentials	3

Figure 16: Table of Task #3

```
1  SELECT Product.Product_ID, Product.Product_Name,
2  Product.Product_Company, Product.Product_Avg_Ratings
3  FROM Product
4  ORDER BY Product.Product_Avg_Ratings DESC;
```

Figure 17: SQL Code of Task #3

4. The products that are bought by the customers on a daily basis.
The product table is used here where the subscription option is equated to a 'Yes'.

Product_ID	Product_Name	Subscription_Option
1	Greenlights	Yes
3	After Hours	Yes
5	CALVIN KLEIN CKIN2U	Yes
6	McCafe K-Cup pod	Yes

Figure 18: Table of Task #4

```
1   SELECT Product.Product_ID, Product.Product_Name,
2   Product.Subscription_Option
3   FROM Product
4   WHERE (((Product.Subscription_Option)='Yes'));
```

Figure 19: SQL Code of Task #4

5. In the event of a sale/discount day, which products are the least bought?
In order to display the least bought products, the items are sorted by having them equated to 'Worse-Seller' entry.

Product_ID	Product_Name	Product_Sell_Performance
5	CALVIN KLEIN CKIN2U	Worse-Seller

Figure 20: Table of Task #5

```
1   SELECT Product.Product_ID, Product.Product_Name,
2   Deals.Product_Sell_Performance
3   FROM Product
4   INNER JOIN Deals ON Product.Product_ID = Deals.Product_ID
5   WHERE (((Deals.Product_Sell_Performance)='Worse-Seller'));
```

Figure 21: SQL Code of Task #5

6. What products are offered through the company warehouse and not from the manufacturer?
The products available at the warehouse are listed below by sorting the 'product' table list in terms of a Boolean value of Warehousing_Offering as 'Yes'.

Product_ID	Product_Name	Warehouse_Offering
1	Greenlights	Yes
4	SAMSUNG 27-inch M5 Smart Monitor	Yes
7	Flannel Shirt	Yes
8	Golf Balls	Yes

Figure 22: Table of Task #6

```
1  SELECT Product.Product_ID, Product.Product_Name,
2  Product.Warehouse_Offering
3  FROM Product
4  WHERE (((Product.Warehouse_Offering)='Yes'));
```

Figure 23: SQL Code of Task #6

7. What products are out of stock and in customer demand?
   In order to display the products out of stock, the Product_Availability attribute as part of
   the products table is equated to 'No'.

Product_ID	Product_Name	Product_Availability	Rating_Stars	Comment_Sentiment
1	Greenlights	No	4	Good
3	After Hours	No	5	Good
4	SAMSUNG 27-inch M5 Smart Monitor	No	5	Good
6	McCafe K-Cup pod	No	5	Good
9	Michelin Tires	No	5	Good

Figure 24: Table of Task #7

```
1  SELECT Product.Product_ID, Product.Product_Name,
2  Product.Product_Availability, Review.Rating_Stars,
3  Review.Comment_Sentiment
4  FROM Review
5  INNER JOIN Product ON Review.Product_ID = Product.Product_ID
6  WHERE (((Product.Product_Availability)='No'));
```

Figure 25: SQL Code of Task #7

8. Products that offer a subscription service in lieu of a one-time purchase.
   The products with a subscription service are displayed with a 'Yes' and conversely with a
   'No' where there is no subscription.

Product_ID	Product_Name	Subscription_Option
1	Greenlights	Yes
2	Grand Theft Auto V	No
3	After Hours	Yes
4	SAMSUNG 27-inch M5 Smart Monitor	No
5	CALVIN KLEIN CKIN2U	Yes
6	McCafe K-Cup pod	Yes
7	Flannel Shirt	No
8	Golf Balls	No
9	Michelin Tires	No
10	Samsung Ear Buds	No

Figure 26: Table of Task #8

```
1    SELECT Product.Product_ID, Product.Product_Name,
2    Product.Subscription_Option
3    FROM Product;
```

Figure 27: SQL Code of Task #8

9. Products that are released 1 day before the stat holidays.
   The products released before the stat holidays are sorted by having the Listing_Date attribute compared with the major stat holiday date ranges. For instance, July 4th, December 25th and January 1st.

Product_ID	Product_Name	Listing_Date
1	Greenlights	2020-12-31
2	Grand Theft Auto V	2020-12-24
3	After Hours	2020-06-30

Figure 28: Table of Task #9

```
1     SELECT Product.Product_ID, Product.Product_Name,
2     Product.Listing_Date
3     FROM Product
4     WHERE
5     Product.Listing_Date >= Convert(dateTime, '12/01/2020') AND
6     (((Product.Listing_Date)> 12/23/2020 AND
7     (Product.Listing_Date)< 12/25/2020)) OR
8     (((Product.Listing_Date)> 12/30/2020 AND
9     (Product.Listing_Date)< 1/1/2021)) OR
10    (((Product.Listing_Date)>6/29/2020 AND
11    (Product.Listing_Date)<7/1/2020));
```

Figure 29: SQL Code of Task #9

10. The products associated with concerning reviews and high rating.
    The products with concerning and high ratings are displayed by inner joining the product and review tables. The Product.ID in Review table acts as a foreign key.

Product_ID	Product_Name	Rating_Stars	Comment_Sentiment
1	Greenlights	4	Good
2	Grand Theft Auto V	4	Moderate
3	After Hours	5	Good
4	SAMSUNG 27-inch M5 Smart Monitor	5	Good
5	CALVIN KLEIN CKIN2U	4	Moderate
6	McCafe K-Cup pod	5	Good
7	Flannel Shirt	3	Moderate
8	Golf Balls	3	Bad
9	Michelin Tires	5	Good
10	Samsung Ear Buds	3	Moderate

Figure 30: Table of Task #10

```
1    SELECT Product.Product_ID, Product.Product_Name,
2    Review.Rating_Stars, Review.Comment_Sentiment
3    FROM Product
4    INNER JOIN Review ON Product.Product_ID = Review.Product_ID;
```

Figure 31: SQL Code of Task #10