E-commerce Business Analysis

By: Udeme Umana

Table Of Content

Overview

This project contains datasets that were developed by me from scratch. The main objective of this project is to use data to answer real life business problems. The dataset consists of the following files: Customer_data, Employee_data, job_info, order, product, sales stored procedures, functions and a view.

Link to my SQL codes

Project Case:

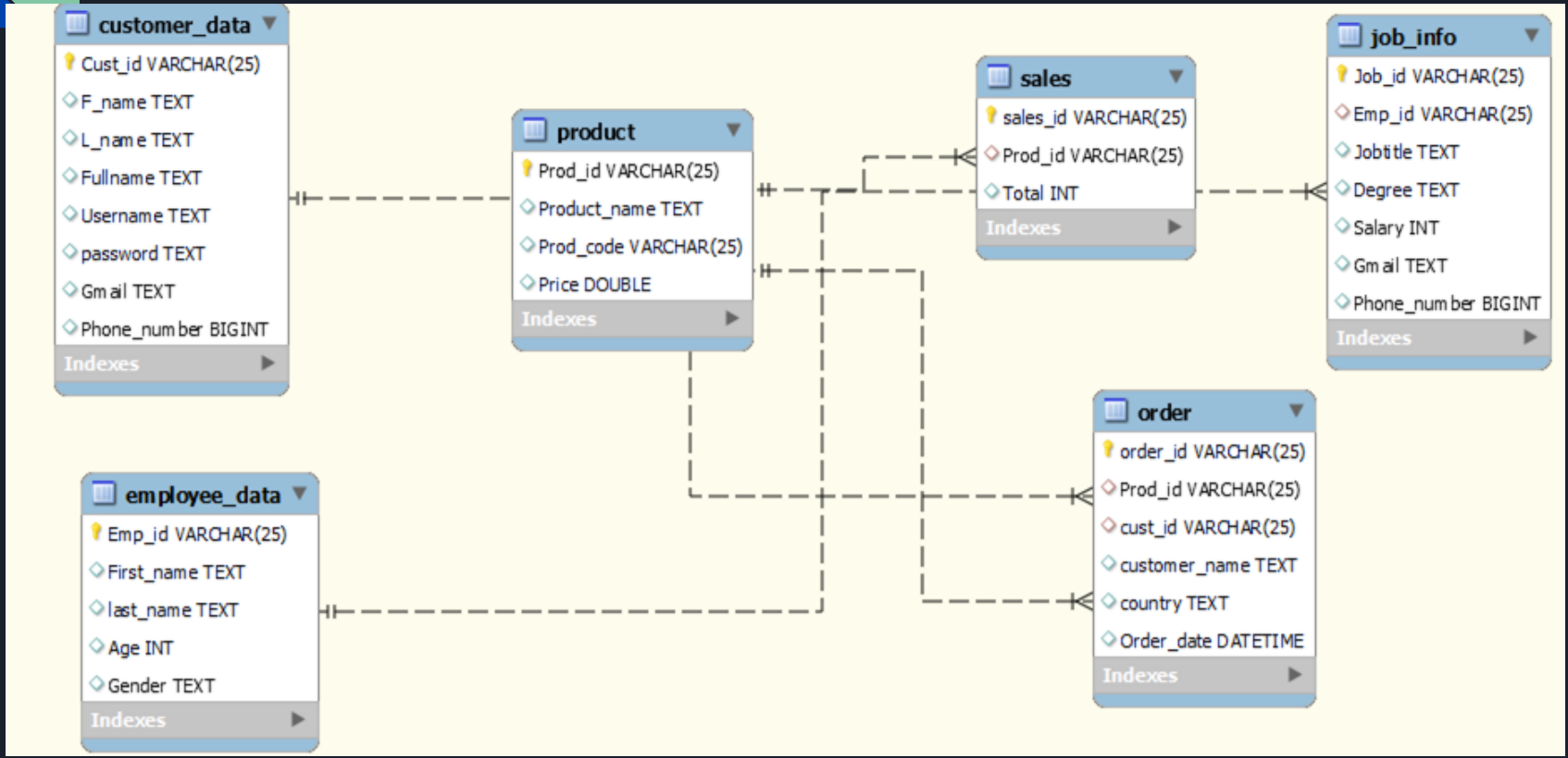
I was appointed a data analyst for Udy_Xpress which is an e-commerce company. They have recently invested in a new technology and are keen to migrate their data into this new capability.

They currently maintain a suite of excel files, which they want to migrate onto a dedicated MYSQL Database. Furthermore, they want to leverage data as a source of information and will also want me to lead on the reporting process of this project.

Procedures and tools used

I focused on cleaning the data, adding new columns where necessary and importing the dataset into MySQL Workbench. I modified each unique keys for each table to a varchar(12) and later converted them into primary and foreign keys in order to get the entity relational diagram. Then I conducted Exploratory data analysis on the datasets before writing queries to answer various business problems.

Entity Relational Diagram



Data Observation With SQL

The tables were imported into Udy_Xpress database in MySQL Workbench, then I wrote queries to interact with the data.

```sql
#Get top 21 rows of most data to observe
SELECT * FROM customer_data as cu
JOIN `order` as od ON cu.Cust_id= od.cust_id
JOIN product as pr ON od.Prod_id = pr.Prod_id
JOIN sales as sa ON pr.Prod_id = sa.Prod_id
LIMIT 21;
```

F_name	L_name	Fullname	Username	password	Gmail	Phone_number	order_id	Prod_id	cust_id	customer_name	count
James	Smith	James Smith	James12	James12@Smith1	JamesSmith@gmail.com	23420202021	OR121	PD121	UD10	James Smith	Nigeria
Emily	Johnson	Emily Johnson	Emily12	Emily12@Johnson1	EmilyJohnson@gmail.com	23420202022	OR122	PD121	UD11	Emily Johnson	Nigeria
William	Brown	William Brown	William12	William12@Brown1	WilliamBrown@gmail.com	23420202023	OR123	PD121	UD12	William Brown	Nigeria
Charlotte	Davis	Charlotte Davis	Charlotte12	Charlotte12@Davis1	CharlotteDavis@gmail.com	23420202024	OR124	PD122	UD13	Charlotte Davis	Nigeria
Thomas	Taylor	Thomas Taylor	Thomas12	Thomas12@Taylor1	ThomasTaylor@gmail.com	23420202025	OR125	PD122	UD14	Thomas Taylor	Nigeria
Olivia	Wilson	Olivia Wilson	Olivia12	Olivia12@Wilson1	OliviaWilson@gmail.com	23420202026	OR126	PD122	UD15	Olivia Wilson	Nigeria
Benjamin	Robinson	Benjamin Robinson	Benjamin12	Benjamin12@Robinson1	BenjaminRobinson@gmail.com	23420202027	OR127	PD122	UD16	Benjamin Robinson	Nigeria
Ava	Wright	Ava Wright	Ava12	Ava12@Wright1	AvaWright@gmail.com	23420202028	OR128	PD123	UD17	Ava Wright	Nigeria
Daniel	Thompson	Daniel Thompson	Daniel12	Daniel12@Thompson1	DanielThompson@gmail.com	23420202029	OR129	PD123	UD18	Daniel Thompson	Nigeria
Mia	White	Mia White	Mia12	Mia12@White1	MiaWhite@gmail.com	23420202031	OR130	PD123	UD19	Mia White	Nigeria

Business Questions

Your company is having a layoff and your boss tells you create a list of employees that does not have a degree.

```sql
#Finding the names,jobtitle,degree of employees that do not have a degree
SELECT em.first_name,em.last_name,degree, jb.Jobtitle,
count(Jobtitle) OVER (partition by Jobtitle) as jb1
FROM employee_data as em
JOIN job_info as jb
ON em.Emp_id = jb.Emp_id
WHERE jb.DEGREE = 'No';
```

first_name	last_name	degree	Jobtitle	jb1
Phillip	Peter	No	Business Anlayst	2
Goodness	David	No	Business Anlayst	2
Joseph	Austin	No	Customer Care	1
Glory	Francis	No	Web Developer	2
Grace	Okon	No	Web Developer	2

Business Questions

My boss wants to know the product that generated the most sales from the data we have.

```sql
#Finding the product with the highest number of sales
SELECT pd.Product_name,Max(total) as mx_total
FROM product pd
JOIN sales sa
ON pd.Prod_id = sa.Prod_id
GROUP BY Product_name
ORDER BY mx_total desc;
```

Product_name	mx_total
Bone Straight wig	625
Iphone X	405
Portable blender	225
Handbags	180
Bucket caps	159
Earphones	150
Sneaker	140
Antiray lenses	138
Headsets	136
WristWatch	132
Comb	120

Business Problem

With my experience working as a data analyst at Udy_xpress, I noticed that in order to get the price of a product you have to manually search for the product code and price from the database. So I created a stored procedure that gives you the price of product when you input the product code(This is similar to scanning of a product barcode to get the price). Below is an illustration of me inputting the product code and getting the price.

```
call udy_express.`Get price`('PD121');
```

PRICE
50

Business Questions

I was asked to create a database where the names and emails of customers who subscribed to be receiving weekly newsletters are stored.

So I decided to create a view, The reason for using a view is each time a customer signs up and registers to be receiving weekly newsletters on the company's website, their information is being saved in the customer table and it'll automatically be added to the subscribers_table.

```
#Displaying the data from the subscribers_table
    SELECT *
    FROM subscribers_email;
```

Fullname	Gmail
James Smith	JamesSmith@gmail.com
Emily Johnson	EmilyJohnson@gmail.com
William Brown	WilliamBrown@gmail.com
Charlotte Davis	CharlotteDavis@gmail.com
Thomas Taylor	ThomasTaylor@gmail.com
Olivia Wilson	OliviaWilson@gmail.com
Benjamin Robinson	BenjaminRobinson@gmail.com
Ava Wright	AvaWright@gmail.com
Daniel Thompson	DanielThompson@gmail.com
Mia White	MiaWhite@gmail.com
Alexander Hall	AlexanderHall@gmail.com

Business Questions

In the month of march the company had a problem with the price software and it overcharged people that placed orders. I was asked to create a list of people that purchased a product/products in the month of march. The main aim of doing this is to enable the company know the people that made a purchase during that period so that a refund can be given to them.

```
#Writing a subquery to find the names of people that purchased a particular product in the month of march
SELECT Cust_id,F_name,L_name
FROM customer_data as cu
WHERE cust_id IN
(SELECT cust_id
FROM `order`
WHERE order_date BETWEEN '2022-03-01 00:00:00' AND '2022-03-31 00:00:00');
```

Cust_id	F_name	L_name
UD48	Mohammed	Lawand
UD49	Joel	Beacon
UD50	Joel	Ian
UD51	Penguin	Random
UD52	Wiley	White
UD53	Harper	Collins
UD54	Rowman	Littlefield
UD55	DK	Johnson
UD56	Ebury	polish
UD57	Louis	Charles
UD58	Charles	Brown

Business Questions

Business analysts and customer_care has done well in their jobs and I was asked to give them 0.6 raise for business analyst and 0.8 raise for customer care and give the rest 0.3 raise.

```
SELECT em.First_name,em.last_name,jb.salary,jb.jobtitle,
CASE
    WHEN jb.jobtitle = 'Customer care' THEN Salary + (salary*.8)
    WHEN jb.jobtitle = 'Business Analyst' THEN Salary +( salary*.6)
    ELSE Salary +( salary*.3)
END AS salary_after_raise
FROM employee_data em
join Job_info jb
ON em.emp_id =jb.emp_id;
```

First_name	last_name	salary	jobtitle	salary_after_raise
Peter	Wing	50000	Customer Care	90000.0
Paul	Jonah	52000	Customer Care	93600.0
John	Paul	51000	Customer Care	91800.0
Isreal	ken	50000	Customer Care	90000.0
Stone	Von	50000	Customer Care	90000.0
Joseph	Austin	48000	Customer Care	86400.0
Isaac	Peter	62000	Digital Operation Manager	80600.0
Umana	Udeme	64000	Digital Operation Manager	83200.0
Inyang	Gabriel	61000	Digital Operation Manager	79300.0
Johnson	Luke	112000	Web Developer	145600.0
Forever	Mathew	102000	Web Developer	132600.0

Business Problem

So before udy_express website launched I created a stored procedure that'll allow customers to access their information from the database using their username and password. Below shows the user inputting his login details as well as his information being displayed:

```
call udy_express.User_login('James12', 'James12@Smith1');
```

Cust_id	F_name	L_name	Gmail	Phone_number
UD10	James	Smith	JamesSmith@gmail.com	23420202021

Business questions

The manager realised that some departments are short staffed and wants you to find the job title with less than 3 people working so that he can hire more people to work in that department:

```
SELECT jobtitle,COUNT(jobtitle) as jbt
FROM job_info
GROUP BY jobtitle
HAVING jbt <3;
```

jobtitle	jbt
Digital Marketing Manger	2
Finance and Accounting	2

Business Problem

I realized that I usually need to find the cust_id, country and order date from Udy_Xpress database on a daily basis, so I created a CTE to store the query in order for me to be able to run it with few lines of codes.

I was asked to find the country with the highest customers in the month of February as seen below:

```
#Using the with statement to find the country with the highest number of customers in the month of february
SELECT country,count(cust_id) as number_of_customer
FROM CustomerOrder_CTE
WHERE order_date BETWEEN '2022-01-12 00:00:00' AND '2022-02-20 00:00:00'
GROUP BY country;
```

country	number_of_customer
Nigeria	22
United Kingdom	10

Business Problem

I decided to write a function in SQL to format a number that has a Nigerian country code. It removes the country code and adds 0 in place of it. (This was just a bonus query)

```
select udy_express.format_phone_number('2348020222023');
```

udy_express.format_phone_number('2348020222
08020222023

Conclusion

- I was able to understand the concept of databases as well as foreign keys and primary keys.
- I understood how to join multiple tables and how it worked.
- I was able to know more about the functionalities of a stored procedure.
- For records that you need to update on a regular basis, Views is more suitable to use.
- Lastly, I was able to understand how to use functions and its syntax.
- SQL codes

Thank you!