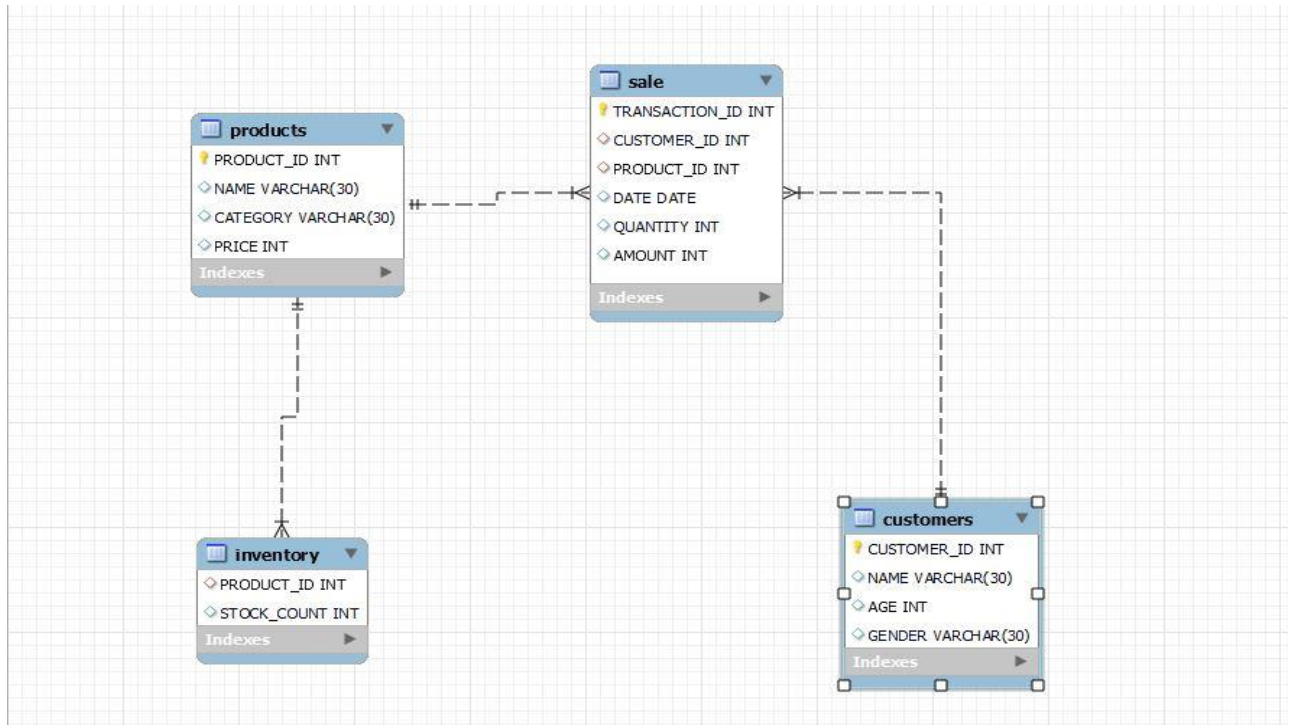


# Apnibus Assignment

- For this assignment, you will be working with a dataset containing information about a retail company's sales and inventory transactions.



## Task 1: Data Exploration and Cleaning(In MYSQL)

```
CREATE DATABASE APNIBUS;  
USE APNIBUS;
```

```
CREATE TABLE CUSTOMERS(  
CUSTOMER_ID INT PRIMARY KEY,  
NAME VARCHAR(30),  
AGE INT,
```

```
GENDER VARCHAR(30));  
INSERT INTO CUSTOMERS VALUES  
( 1,"John Doe",35,"Male"),  
(2,"Jane Smith",28,"Female"),  
(3,"Alex Kim",42,"Male"),(4,"Emily  
Wong",31,"Female"),(5,"Mark Ford",45,"Male"),  
(6,"Lisa Chen",39,"Female"),(7,"Paul  
Brown",51,"Male"),(8,"Emma Davis",24,"Female"),  
(9,"Adam Lee",33,"Male"),(10,"Amy  
Kim",27,"Female"),(11,"Mike Smith",42,"Male"),  
(12,"Lily Wang",29,"Female"),(13,"Tom  
Li",36,"Male"),(14,"Anna Chen",31,"Female"),(15,"Jack  
Wong",48,"Male"),  
(16,"Mia Hall",NULL,"Male");
```

```
CREATE TABLE PRODUCTS(  
PRODUCT_ID INT PRIMARY KEY,  
NAME VARCHAR(30),  
CATEGORY VARCHAR(30),  
PRICE INT);
```

```
INSERT INTO PRODUCTS  
VALUES(1,"Tshirt","Apparel",20),(2,"Jeans","Apparel",50),(3,"Sn  
eakers","Footwear",80),  
(4,"Watch","Accessory",150),(5,"Backpack","Accessory",30),(6,"  
Laptop","Electronics",800),  
(7,"Chair","Furniture",120),(8,"Desk","Furniture",200),(9,"Book  
","Books",10),
```

```
(10,"Bottle","Kitchen",5),(11,"Plants","Garden",15),(12,"Shoes",  
"Footwear",60),  
(13,"Hat","Accessory",25),(14,"Coat","Apparel",100),(15,"Socks  
","Apparel",8);
```

```
CREATE TABLE SALES(  
TRANSACTION_ID INT PRIMARY KEY,  
CUSTOMER_ID INT,  
PRODUCT_ID INT,  
`DATE` DATE,  
QUANTITY INT,  
AMOUNT INT),  
FOREIGN KEY(CUSTOMER_ID) REFERENCES  
CUSTOMERS(CUSTOMER_ID),  
FOREIGN KEY (PRODUCT_ID) REFERENCES  
PRODUCTS(PRODUCT_ID));
```

```
INSERT INTO SALES VALUES (1,1,4,'2023-06-01',2,99.98),  
(2,2,10, '2023-06-02' ,1, 399.99),  
(3,3,1, '2023-06-03' ,3, 62.97),  
(4,4,6, '2023-06-04' ,1, 99.99),  
(5,5,8, '2023-06-05' ,2, 59.98),  
(6,6,12, '2023-06-06' ,2, 139.98),  
(7,7,14, '2023-06-07' ,1, 19.99),  
(8,8,18, '2023-06-08' ,1, 89.98),  
(9,9,3, '2023-06-09' ,5, 79.95),  
(10,10,16, '2023-06-10' ,1, 79.99),  
(11,11,2, '2023-06-11' ,1, 599.99),  
(12,12,5, '2023-06-12' ,2, 79.98),
```

```
(13,13,9, '2023-06-13' ,1, 79.99),  
(14,14,15, '2023-06-14' ,1, 34.99),  
(15,15,20,' 2023-06-15' ,1,24.99);
```

```
CREATE TABLE INVENTORY(  
PRODUCT_ID INT,  
STOCK_COUNT INT),  
FOREIGN KEY(PRODUCT_ID) REFERENCES  
PRODUCTS(PRODUCT_ID));
```

```
INSERT INTO INVENTORY VALUES  
(1,50),(2,100),(3,20),(4,40),(5,30),(6,10),(7,15),(8,25),(9,5),  
(10,70),(11,20),(12,60),(13,10),(14,30),(15,35);
```

### **Cleaning :**

- **In Customer table Customer\_id number 16 have null in their Age so we have to removing that row.**

```
DELETE FROM CUSTOMERS  
WHERE CUSTOMER_ID=16;
```

- **In sales table transaction\_id number 4 have sold 1 quantity lets change the quantity of transaction\_id 4**

```
UPDATE SALES  
SET QUANTITY=4 WHERE TRANSACTION_ID=4;
```

## **Task 2: Data Analysis**

## Q1. Calculate the total revenue generated by the company for each product category.

-To calculate the total revenue generated by the company for each product category, we can join the "Sales" table with the "Products" table on the "product\_id" column.

```
SELECT P.CATEGORY, SUM(S.AMOUNT) AS TOTAL_REVENUE FROM
SALES S
INNER JOIN PRODUCTS P ON S.PRODUCT_ID=P.PRODUCT_ID
GROUP BY P.CATEGORY
ORDER BY TOTAL_REVENUE DESC;
```

	CATEGORY	TOTAL_REVENUE
►	Apparel	718
	Kitchen	400
	Footwear	220
	Accessory	180
	Electronics	100
	Books	80
	Furniture	60

## Q2 Determine the top 5 customers who have made the highest total purchases, considering the customer's age and gender.

- we have to join the "Sales" table with the "Customers" table on the "customer\_id" column. Then, we can group the data by the "customer\_id", "age", and "gender" columns and calculate the sum of the "amount" column for each customer.

```

SELECT C.CUSTOMER_ID,C.NAME,C.AGE,C.GENDER,SUM(S.AMOUNT)
AS TOTAL_PURCHASE FROM
CUSTOMERS C INNER JOIN SALES S ON
C.CUSTOMER_ID=S.CUSTOMER_ID
GROUP BY C.CUSTOMER_ID,C.NAME,C.AGE,C.GENDER
ORDER BY TOTAL_PURCHASE DESC
LIMIT 5;

```

	customer_id	NAME	AGE	GENDER	TOTAL_PURCHASE
▶	11	Mike Smith	42	Male	600
	2	Jane Smith	28	Female	400
	6	Lisa Chen	39	Female	140
	1	John Doe	35	Male	100
	4	Emily Wong	31	Female	100

### Q3 Identify the most profitable product category by calculating the average revenue per unit sold.

- we have to join the "Sales" table with the "Products" table on the "product\_id" column. Then, we can calculate the average revenue per unit sold by dividing the total revenue (sum of the "amount" column) by the total quantity sold (sum of the "quantity" column) for each product category.

```

SELECT P.CATEGORY,ROUND(SUM(S.AMOUNT)/SUM(S.QUANTITY),2)
AS AVG_REVENUE_PER_UNIT
FROM SALES S INNER JOIN PRODUCTS P ON
S.PRODUCT_ID=P.PRODUCT_ID
GROUP BY P.CATEGORY
ORDER BY AVG_REVENUE_PER_UNIT DESC
LIMIT 1;

```

	CATEGORY	AVG_REVENUE_PER_UNIT
▶	Kitchen	400.00

#### **Q4 Analyze the inventory data and identify products that need restocking (stock count less than a specified threshold).**

- I have query the "Inventory" table and filter the products where the stock count is less than the specified threshold. I assuming that threshold for stocking is less than 20.

```
SELECT * FROM INVENTORY
WHERE STOCK_COUNT<20;
```

	PRODUCT_ID	STOCK_COUNT
▶	6	10
	7	15
	9	5
	13	10

### **TASK 3: Advanced Analysis and Reporting**

#### **Q1. Write a SQL query to calculate the average age of customers for each product category.**

we need to join the "Customers" table with the "Sales" and "Products" tables using the appropriate foreign key relationships

```
SELECT P.CATEGORY ,AVG(C.AGE) AS AVG_AGE_CUSTOMER FROM  
CUSTOMERS  
C INNER JOIN SALES S ON C.CUSTOMER_ID=S.CUSTOMER_ID  
INNER JOIN PRODUCTS P ON S.PRODUCT_ID=P.PRODUCT_ID  
GROUP BY 1;
```

	CATEGORY	AVG_AGE_CUSTOMER
▶	Accessory	32.0000
	Kitchen	28.0000
	Apparel	41.5000
	Electronics	31.0000
	Furniture	45.0000
	Footwear	36.0000
	Books	36.0000

This query assumes that the "Customers" table has an "age" column and the "Sales" table has a "customer\_id" column linking to the "customer\_id" column in the "Customers" table, as well as a "product\_id" column linking to the "product\_id" column in the "Products" table.

**Q2 Write a SQL query to retrieve the top 3 product categories that have the highest average transaction amount.**

we can join the "Sales" table with the "Products" table based on the "product\_id" column.

Then, we can calculate the average transaction amount for each product category and sort the results in descending order.



```

SELECT P.CATEGORY,AVG(S.AMOUNT) AS AVG_TRNX_AMOUNT FROM
SALES S
INNER JOIN PRODUCTS P ON S.PRODUCT_ID=P.PRODUCT_ID
GROUP BY 1
ORDER BY 2 DESC
LIMIT 3;

```

	CATEGORY	AVG_TRNX_AMOUNT
▶	Kitchen	400.0000
	Apparel	179.5000
	Footwear	110.0000

**Q3. Create a comprehensive report summarizing your findings from Task 2 and Task 3. Include relevant tables, charts, and explanations to present your analysis clearly.**

Task 2: Calculate the total revenue generated by the company for each product category

	CATEGORY	TOTAL_REVENUE
▶	Apparel	718
	Kitchen	400
	Footwear	220
	Accessory	180
	Electronics	100
	Books	80
	Furniture	60

From the table, we can see that the "Apparel" category has generated the highest total revenue of \$718, followed by "Kitchen" with \$400.

Task 3: Identify the top 3 product categories with the highest average transaction amount.

	CATEGORY	AVG_TRNX_AMOUNT
▶	Kitchen	400.0000
	Apparel	179.5000
	Footwear	110.0000

From the table, we can observe that the "Kitchen" category has the highest average transaction amount of \$400, followed by "Apparel" with \$179.5 and "Footwear" with \$110.

### **Conclusion:**

Based on the findings from Task 2 and Task 3, we can draw the following conclusions:

- 1.The " Apparel " category has generated the highest total revenue of \$718, indicating that it is the most profitable category for the company.
2. In terms of the average transaction amount, the "Kitchen" category also leads with an average transaction amount of \$400. This suggests that customers spend more on Kitchen products compared to other categories.
3. The " Apparel " and " Footwear " categories also perform well in terms of average transaction amounts, with \$179.5 and \$110, respectively.

Based on these findings, the company can focus on maximizing sales and revenue by strategically promoting and expanding its product offerings in the "Apparel" category, which seems to be the most profitable. Additionally, attention can be given to the "Apparel" and "Footwear" categories, as they have shown potential in generating higher average transaction amounts.

To further analyze and make informed decisions, it would be beneficial to investigate the specific products within these categories that contribute significantly to the revenue and average transaction amounts. This analysis can help identify high-performing products and enable targeted marketing and inventory management strategies.

**Bonus Task (Optional): Perform a customer segmentation analysis to identify different customer groups based on their purchasing behavior, age, and gender. Provide insights on each customer segment and suggest personalized marketing strategies for each segment.**

Here's a high-level approach to customer segmentation based on purchasing behavior, age, and gender:

1. Fashion and lifestyle Shoppers:

Characteristics: Customers who frequently purchase products in the "Clothing" and "Accessories" categories.

Marketing Strategy:

Offer personalized style guides, fashion tips, and exclusive discounts on trendy clothing, accessories, and beauty products. Collaborate with fashion influencers and utilize social media platforms like Instagram for targeted fashion campaigns.

## 2.Diverse General Shoppers:

### Characteristics:

Customers with varied purchasing behavior across multiple categories.

### Marketing Strategy:

Provide personalized recommendations based on previous purchases and promote cross-category discounts or bundles. Utilize targeted email marketing campaigns, loyalty programs, and personalized product suggestions on the company's website to increase engagement and repeat purchases.