**تحليل كود إدراج البيانات (Data insert.sql)**

يتضمن أوامر **INSERT INTO ... SELECT DISTINCT** لسحب البيانات من مصدر بيانات موجود **([Superstore-Sales])** وإدراجها في الجداول. إليك ما يفعله كل جزء:

1. **إدراج العملاء**

sql

نسختحرير

INSERT INTO Customers (Customer\_ID, Customer\_Name, Segment)

SELECT DISTINCT Customer\_ID, Customer\_Name, Segment

FROM [Superstore-Sales];

* + يضمن إدراج كل عميل مرة واحدة باستخدام **DISTINCT** لمنع التكرارات.

1. **إدراج بيانات الفروع**

sql

نسختحرير

INSERT INTO Store\_Branches (City, State, Postal\_Code, Region)

SELECT DISTINCT City, State, Postal\_Code, Region

FROM [Superstore-Sales];

* + يتم إدخال كل فرع مرة واحدة بناءً على الموقع الجغرافي الخاص به.

1. **إدراج بيانات طرق الشحن**

sql

نسختحرير

INSERT INTO Ship\_Modes (Ship\_Mode\_Name)

SELECT DISTINCT Ship\_Mode

FROM [Superstore-Sales];

* + يضمن أن كل طريقة شحن يتم تسجيلها مرة واحدة فقط.

1. **اكتشاف التكرارات في المنتجات**

sql

نسختحرير

SELECT Product\_ID, Product\_Name, COUNT(\*) AS Count

FROM [Superstore-Sales]

GROUP BY Product\_ID, Product\_Name

HAVING COUNT(\*) > 1;

* + يبحث عن المنتجات التي قد تكون مسجلة أكثر من مرة بنفس **Product\_ID** ولكن بأسماء مختلفة.

1. **التعامل مع المنتجات المكررة باستخدام CTE و ROW\_NUMBER**

sql

نسختحرير

WITH Duplicates AS (

SELECT \*, ROW\_NUMBER() OVER (PARTITION BY Product\_ID ORDER BY (SELECT NULL)) AS RowNum

FROM [Superstore-Sales]

)

SELECT \* FROM Duplicates

WHERE Product\_ID IN (

SELECT Product\_ID

FROM [Superstore-Sales]

GROUP BY Product\_ID

HAVING COUNT(DISTINCT Product\_Name) > 1

);

* + يستخدم نافذة **ROW\_NUMBER()** لترقيم الصفوف داخل كل **Product\_ID**.
  + يحدد المنتجات التي لها معرف واحد ولكن بأسماء مختلفة.

**النتيجة والتأثير على البيانات**

* يضمن إنشاء الجداول وتنظيم البيانات داخلها بطريقة صحيحة دون تكرارات.
* يضمن أن البيانات المدخلة نظيفة ومتسقة باستخدام **DISTINCT** والفلاتر الأخرى.
* يساعد في تحسين جودة البيانات من خلال الكشف عن الأخطاء المحتملة مثل التكرارات في أسماء المنتجات.