
1. ما هي الأعمدة في الجدول السابق؟

- Customer_Name
- Age
- Email
- Join_Date
- Total_Purchase

2. لماذا يحتاج الجدول إلى تنظيف قبل التحليل؟

- Age, Email, Join_Date, Total_Purchase في أعمدة مثل (NaN) لأنه يحتوي على قيم مفقودة.
- وجود صفوف مكررة مثل أحمد علي (1001 و 1006).
- إلى تاريخ بدل Join_Date بعض الأعمدة تحتاج تحويل نوع البيانات مثل نص.
- هذه المشاكل تؤثر على دقة التحليل والنتائج.

3. ما هو أثر الصفوف المكررة على التحليل؟

- تضخم القيم النهائية (مثل مجموع المبيعات أو متوسط العمر).
- يعطي صورة غير صحيحة عن العملاء (عميل واحد يظهر أكثر من مرة).

4. ما هي أفضل طريقة للتعامل مع القيم المفقودة؟

- نستخدم المتوسط لملء القيم → (Age) إذا كان التحليل يعتمد على العمر الناقصة.
- 0 يمكن وضع → (Total_Purchase) إذا كان التحليل يعتمد على المبيعات "إذا كان يعني" لا يوجد شراء.
- إذا كان البريد الإلكتروني مفقودًا → قد نتركه أو نحذف الصف حسب أهمية التحليل.

5. إلى تاريخ؟ Join_Date لماذا من المهم تحويل العمود

- حتى نستطيع استخدامه في التحليل الزمني مثل معرفة العملاء حسب الأشهر أو حساب مدة بقاء العميل.
- بدون تحويله إلى تاريخ سيظل كنص ولا يمكن إجراء عمليات حسابية أو مقارنة زمنية عليه.

6. ما هي الملاحظات التي يمكن استخراجها عن العملاء والشراء من الجدول؟

- أعمار العملاء تتراوح بين 28 و 42 سنة.
- يوجد تكرار لنفس العميل (أحمد علي) مما يشير إما إلى عملية شراء متكررة أو خطأ في البيانات.

- المبيعات الكلية تتفاوت (من 150 حتى 400).
 - (NaN) يوجد عملاء لم يسجلوا عمرهم أو لم يشتروا شيئاً.
 - العملاء بدأوا ينضمون للنظام ابتداءً من يناير 2025.
-