

تقرير تحليل بيانات الرعاية الصحية باستخدام Python و Power BI

1. مصدر البيانات وجمعها

تم استخدام بيانات الرعاية الصحية (Healthcare Dataset) والتي تتضمن معلومات عن المرضى، المستشفيات، حالات الدخول، الفواتير، ومدة الإقامة.

2. تنظيف وتحويل البيانات

شملت عملية التنظيف المراحل التالية:

- إزالة القيم المفقودة والمكررة.
 - توحيد أسماء الأعمدة وتنسيقات التاريخ.
 - معالجة القيم غير المنطقية في الأعمدة العددية (مثل مدة الإقامة أو المبالغ).
 - تحويل أنواع البيانات لتناسب التحليل في Power BI.
- بعد التنظيف، تم حفظ الملف كـ `healthcare_dataset_cleaned.csv` ورفعته إلى GitHub لاستخدامه في Power BI عبر الرابط المباشر (Raw link).

3. التحليل والرؤى المستخلصة

من خلال استكشاف البيانات وتحليلها عبر Power BI تم الوصول إلى النتائج التالية:

- **مواسم الذروة في الفواتير الإجمالية (Total Billing)** تكون في أشهر الصيف، وبشكل خاص في شهر أوت (أغسطس)، ما قد يُعزى إلى زيادة الأمراض في هذه الفترة أو السفر والعطلات التي تؤدي إلى ارتفاع الطلب على الخدمات الطبية.
- كما نرى أن الفواتير الإجمالية كانت مرتفعة جداً من سنة 2021 إلى 2023 وهذا يفسر بسبب مرض الكورونا.
- **أكثر الحالات شيوعاً بين المرضى هي التهاب المفاصل (Arthritis).**
- أغلب الفواتير مصنفة ضمن الفئات High و Medium.
- **قيمة الفواتير ليست مرتبطة بمدة الإقامة في المستشفى، بل تعتمد على المستشفى نفسه ونوع الحالة الطبية، حيث وُجد أن مستشفى Johnson PLC من بين الأكثر تكلفة، خاصة في حالات Arthritis.**