

# التقرير النهائي لتحليل البيانات " كوفيد-19"

## 1. جمع البيانات:

تم الاعتماد على مجموعة البيانات العالمية COVID-19 – Our World in Data المتاحة على منصة Kaggle، والتي تتضمن مؤشرات الحالات، الوفيات، التطعيمات، والعوامل الاقتصادية والديموغرافية عبر أكثر من 230 دولة خلال الفترة من 1 جانفي 2020 إلى 5 مارس 2022. تم تحميلها مباشرة عبر واجهة API Kaggle ملف تنظيف البيانات والمعالجة: owid-covid-data.csv.

## 2. نظيف البيانات : (Data Cleaning)

### التحقق الأولى

- فحص أنواع الأعمدة وعدد القيم المفقودة.
- التأكد من أن عمود date منسق كتواريخ.
- تحديد الفترة الزمنية وعدد المواقع (238 دولة).

### تنقية البيانات

- حذف الصفوف الخاصة بالتجمیعات الإقليمية. (World, Africa, Europe, ...).
- الاحتفاظ فقط بسجلات الدول.
- حذف الأعمدة التي تتجاوز نسبة القيم المفقودة فيها 70%.
- تعويض القيم المفقودة في الأعمدة الرقمية بال وسيط، وفي النسبة بكلمة "Unknown".
- حذف التكرارات وتوحيد أسماء الأعمدة لتكون بصيغة minuscule ومن دون مسافات.

### النتيجة بعد التنظيف

- عدد الصفوف النهائية: ≈ 158,000.
- عدد الأعمدة المحافظ عليها: ≈ 50.
- خلو الجدول من التكرارات والقيم الناقصة المهمة.

## 3. التحليل الاستكشافي (Exploratory Data Analysis)

تم إجراء تحاليل مبدئية للتوضیح:

- تطور عدد الحالات والوفيات عبر الزمن.
- المقارنة بين القارات والدول.
- العلاقة بين التطعيمات وانتشار العدوى.
- الفوارق المرتبطة بالعوامل الاقتصادية.

## Power BI Visualizations .4

بعد تجهيز الملف clean\_covid\_data.csv ، تم تحليلها والتوصل الى أبرز النقاط :

- متوسط العمر للمصابين كانوا من فئة الشباب.
- وصلت الإصابات ذروتها في حدود 2022 حيث كانت في تزايد المستمر قبل ذلك مع الازد
- بعين الاعتيار أن جمع البيانات توقف بتاريخ 5 مارس من نفس السنة، رغم أن عدد اللذين
- نلقوا اللقاح وصل ذروته قبل ذلك أي في حدود نهاية 2021 .
- الدول التي تتمتع بالقدرة الاقتصادية استطاعت تحسين البنية الصحية التحتية فتحكموا بذلك
- على عدد الوفيات على عكس الدول التي تعاني من هشاشة في احديهما أو كليهما بحيث ساهم
- الإهمال أو عدم القدرة على توفير الرعاية بزيادة الإصابات وحتى الوفيات .
- الدول التي تضم عدد كبير من المسنين غالباً ما كانت تسجل عدد وفيات أكبر وذلك بسبب
- الإصابة المسبقة للمسنين بالأمراض كالسكري وأمراض القلب، بحيث يكون تأثير الفايروس
- أقوى عليهم.