

السيناريو والمهام

الموقف التخيلي

تعمل في شركة تجارة إلكترونية (وهمية) باسم **SwiftShop**.

تبيع الشركة ثلاثة أصناف رئيسية من المنتجات:

- إلكترونيات (Electronics)
- ملابس (Clothing)
- ومستلزمات المنزل (Home Goods)

تم تزويدك بملف CSV باسم `swiftshop_sales_data.csv`، يحتوي على بيانات المعاملات (الطلبات) والمعلومات الأساسية عن العملاء خلال آخر 18 شهرًا.

يحتاج فريقا التسويق والمبيعات إلى لوحة معلومات (Dashboard) داخلية للإجابة عن أسئلة مهمة حول

- الأداء البيعي، والمنتجات الأكثر مبيعًا
- والاتجاهات على مستوى الزمن وسلوك العملاء.
- كما يرغبون في إمكانية تصفية النتائج حسب المنطقة الجغرافية والفترة الزمنية، وفئة المنتج.

مهمتك هي بناء هذه اللوحة باستخدام **python**، وبالتحديد **pandas** للمعالجة وتحليل البيانات، و**Plotly** لإنشاء الرسوم البيانية التفاعلية، و**Dash** لبناء واجهة لوحة المعلومات.

تفاصيل البيانات

الملف (`swiftshop_sales_data.csv`) يحتوي على الأعمدة التالية:

1. `order_id` – معرف فريد لكل طلبية.
2. `order_date` – تاريخ الطلبية (على هيئة YYYY-MM-DD).
3. `customer_id` – معرف فريد لكل عميل.
4. `customer_region` – المنطقة الجغرافية للعميل (مثل "North"، "East"، "South"، "West").
5. `product_id` – معرف فريد للمنتج الذي تم شراؤه.
6. `product_name` – الاسم المقروء للمنتج (مثل: "Wireless Headphones Model X").
7. `category` – فئة المنتج (Electronics, Clothing, Home Goods).
8. `unit_price` – سعر الوحدة الواحدة من المنتج.
9. `quantity` – عدد الوحدات المشتراة في الطلبية.
10. `total_amount` – حاصل ضرب (`unit_price * quantity`).

11. – payment_method طريقة الدفع المستخدمة) مثل "Credit Card" ، "PayPal" ، "Cash on Delivery".

12. – customer_rating تقييم العميل (من 1 إلى 5) بعد استلام الطلبية (قد تكون بعض الطلبيات بلا تقييم).

يفترض أن حجم الملف يتراوح بين 5,000 و 10,000 صف؛ أي أنه كافٍ لاختبار الكفاءة في التعامل مع البيانات دون أن يكون ضخماً جداً.

متطلبات المهمة

1. استيراد البيانات ومعالجتها

- اقرأ الملف (CSV) إلى Pandas DataFrame.
- تحقق من القيم المفقودة أو غير الصالحة. على سبيل المثال، قد لا يوجد تقييم لبعض الطلبات. قرر كيف ستتعامل مع هذه القيم.
- أنشئ أي أعمدة إضافية ضرورية مثل (استخراج الشهر أو السنة من حقل order_date، أو حساب هامش الربح إن لزم، إلخ)

2. التحليل الأساسي والمؤشرات الرئيسية (KPIs)

- احسب إجمالي المبيعات اليومية أو الشهرية (total_amount. مجموع)
- احسب متوسط قيمة الطلبية (Average Order Value) بحسب الزمن أو بحسب المنطقة.
- اعرض المنتجات الأكثر مبيعاً والفئات الأعلى عائداً (بحسب إجمالي المبيعات أو عدد الوحدات المباعة).
- احسب متوسط تقييم العملاء (إن كان التقييم متوفراً) حسب المنتج أو الفئة.

3. الرسوم البيانية باستخدام Plotly

أنشئ على الأقل ثلاثة رسوم تفاعلية:

1. المبيعات عبر الزمن: (Sales Over Time) رسم بياني (خطي أو أعمدة) يُظهر الإيرادات على مستوى اليوم أو الشهر.
2. أداء فئة المنتج: (Product Category Performance) رسم بياني دائري (Pie) أو عمودي (Bar) يوضح حصة كل فئة (الالكترونيات، ملابس، مستلزمات منزل) من إجمالي المبيعات.
3. توزيع تقييمات العملاء: (Customer Rating Distribution) رسم بياني مثل Histogram أو Bar لإظهار توزيع التقييمات من 1 إلى 5.

يُفضّل إتاحة خصائص تفاعلية في Plotly مثل معلومات التلميح (Tooltips) عند تمرير المؤشر فوق العناصر.

4. المرشحات والتفاعلات (Filters & Interactions)

باستخدام أدوات Dash ، وقِّر على الأقل نوعين من أدوات التصفية/التحكم:

1. **Date Range Picker** لاختيار الفترة الزمنية من-إلى، لتصفية النتائج
 2. **قائمة منسدلة لتحديد المنطقة والفئة (Dropdown)** بحيث يتمكن المستخدم من عرض البيانات المتعلقة بمنطقة جغرافية أو فئة محددة أو الاثنين معاً.
- يجب أن تتحدّث الرسوم والمؤشرات (KPIs) تلقائياً عند تغيير المرشحات.

5. بناء لوحة المعلومات (Dash)

- **التصميم العام (Layout):** على الأقل ينبغي أن تحتوي اللوحة على:
 - عنوان في الأعلى (Header) مثل: "SwiftShop Sales Dashboard"
 - قسم لعناصر التحكم (Filters) مثل اختيار التاريخ والقوائم المنسدلة.
 - قسم رئيسي للرسوم البيانية والعناصر التفاعلية.
- استخدم مكونات Dash المناسبة مثل dcc.Graph, dcc.Dropdown, dcc.DatePickerRange, إلخ.
- يمكنك أيضاً إضافة جدول (Dash DataTable) إذا أردت عرض بعض البيانات بشكل مُجدول.

6. ميزات إضافية (اختيارية لتحدي إضافي)

- دعم اختيار عدة فئات في نفس الوقت (مثلاً، مشاهدة نتائج فئتي الإلكترونيات والملابس معاً).
- عرض أفضل 10 منتجات بالأرباح في رسم بياني عمودي محدّث حسب المرشحات.
- إضافة زر تصدير (Export) للبيانات المفلترة على هيئة CSV.
- إضافة رسائل توضيح أو تلميحات حول كل رسم أو مقياس كي يفهمه المستخدم غير التقني.
- عرض مقارنة سنة بسنة (Year-over-Year) أو شهر بشهر (Month-over-Month) إذا امتدت البيانات عبر فترة طويلة.

ما يجب تسليمه

1. مستودع مشروع GitHub أو ما شابه يتضمّن:
 - ملف requirements.txt أو pyproject.toml يحدّد حزم python المطلوبة مثل pandas, Dash, Plotly
 - ملف python رئيسي مثل app.py يقرأ ملف CSV، ويقوم بالمعالجة اللازمة، ويُشغّل تطبيق Dash
 - أي ملفات أو وحدات (Modules) مساعدة أخرى.

2. خطوات واضحة لتشغيل لوحة المعلومات (بما في ذلك تهيئة بيئة بايثون وكيفية إطلاق التطبيق).

الهدف من التقييم

يركّز هذا التمرين على تقييم مدى قدرة المرشح/ة على:

- فهم استيراد البيانات وتنظيفها باستخدام pandas مثل التعامل مع البيانات المفقودة، عمليات التجميع، وحساب المؤشرات.
- بناء مشروع صغير (ولكن متكامل) في بايثون بخطوات متعددة.
- إنشاء رسوم بيانية تفاعلية باستخدام Plotly
- تطوير واجهة مستخدم بسيطة وفعالة بواسطة Dash مع توفر المرشحات وعناصر التحكم.
- تنظيم الكود والتعليق عليه واتباع أفضل ممارسات بنية المشاريع.

شجّع المرشحين على الإبداع وتجاوز الحد الأدنى من المتطلبات إذا كان لديهم الوقت، لإظهار مهاراتهم في التصميم والبرمجة والاهتمام بالتفاصيل.

عينة من البيانات (Dataset Sample)

فيما يلي مثال صغير (10 صفوف فقط) على صيغة ملف swiftshop_sales_data.csv يمكنك استخدامه للعرض أو الاختبار، أو يمكنك إنشاء ملف أكبر وأشمل بالهيكل نفسه:

```
order_id,order_date,customer_id,customer_region,product_id,product_name,category,unit_price,quantity,total_amount,payment_method,customer_rating
1001,2024-01-05,501,North,2001,"Smartphone A",Electronics,299.99,1,299.99,Credit Card,5
1002,2024-01-07,502,East,3001,"Men's T-Shirt",Clothing,19.99,2,39.98,PayPal,4
1003,2024-01-10,501,North,3002,"Women's Jeans",Clothing,49.99,1,49.99,Cash on Delivery,3
1004,2024-02-02,503,South,2002,"Wireless Headphones",Electronics,89.99,1,89.99,Credit Card,4
1005,2024-02-10,504,West,4001,"Ceramic Vase",Home Goods,25.00,2,50.00,PayPal,
1006,2024-02-15,505,East,4002,"Table Lamp",Home Goods,35.50,1,35.50,Credit Card,5
1007,2024-03-01,506,North,2003,"Laptop B",Electronics,799.99,1,799.99,Credit Card,4
```

ملاحظة:

- بعض الأسطر (مثل الطلبية رقم 1005) لا تحتوي على تقييم (customer_rating) ؛ هذا مثال على القيم المفقودة.
 - يمكنك تعديل هذه البيانات أو إضافة مزيد من الصفوف/الأعمدة وفق حاجة الاختبار.
-

كيفية استخدام العينة

1. حمل الملف المرفق واحفظه في ملف swiftshop_sales_data.csv.
2. اقرأ الملف في بايثون عبر:

```
import pandas as pd

df = pd.read_csv('swiftshop_sales_data.csv')
```

3. ابدأ بتجربة وتنظيف البيانات وإنشاء لوحة معلومات بسيطة وفق التعليمات المطلوبة.
-

بالتوفيق في التقييم!