

---

# تقرير تحليل بيانات قروض المشاريع الصغيرة

إعداد: عفيفي ياسر عفيفي  
المسمى الوظيفي: عالم بيانات (Data Scientist)  
الملفان المستخدمان:

1. **Notebook 1 – Data Cleaning**  
يتضمن عمليات التنظيف، التعويض، وإعادة التصنيف للبيانات.

2. **Notebook 2 – Data Analysis**  
يتضمن تحليل العلاقات بين الأعمدة المختلفة ونتائجها.

---

## ملخص عمليات تنظيف البيانات:

- تم حذف الصفوف التي تحتوي على قيم مفقودة في الأعمدة التي يقل عدد القيم المفقودة فيها عن 5000، بهدف تنظيف البيانات من الفجوات الصغيرة دون التأثير على حجمها.
- تم تعويض قيم  $NAICS = 0$  بالاعتماد على اسم الشركة، ثم استخدام **KNN Imputer** لإكمال القيم المتبقية.
- تم حذف عمود **ChgOffDate** لأنه مكرر في المعنى مع **MIS\_Status** و **ChgOffPrinGr**.

---

## تحليل العلاقات والنتائج:

### 1. التحليل حسب المنطقة (Region):

- الشمال الشرقي أقل المناطق تعثرًا بنسبة **20.36%**، والجنوب الأعلى بنسبة **25.77%**.
  - التوصية: دعم البنوك المحلية في الجنوب لتقليل التعثر.
-

## 2. حسب الولاية (BankState):

- ولايات مثل CA, FL, NC, SC, VA لديها نسب تعثر أعلى من المتوسط.
  - التوصية: تحسين نظم السداد وتسهيل إجراءات الدفع في هذه الولايات.
- 

## 3. حسب نوع النشاط (NAICS):

- حوالي 200,000 سجل غير مصنف (NAICS = 0).
  - أعلى نسب تعثر كانت في الأنشطة الزراعية وتربية الأحياء المائية.
  - التوصية: تطبيق سياسات تقييم مخاطر أكثر دقة لهذه القطاعات.
- 

## 4. حسب السنة المالية (ApprovalFY):

- ارتفاع كبير في التعثر بين 1972-1982، انخفاض في 1997-1998، ثم ارتفاع تدريجي حتى 2006-2008.
  - التوصية: دراسة تأثير التغيرات الاقتصادية في تلك الفترات لتحسين سياسات الإقراض.
- 

## 5. حسب مدة القرض (Term\_Years):

- أعلى تعثر في القروض قصيرة الأجل (حتى 7 سنوات)، ثم انخفاض تدريجي بعد ذلك.
  - التوصية: تحسين شروط القروض القصيرة لتقليل المخاطر.
- 

## 6. حسب حجم الموظفين (Emp\_Category):

- المشاريع الصغيرة (1-10 موظفين) الأعلى تعثرًا بنسبة 19.9%.
  - التوصية: توفير برامج دعم وتمويل موجهة للمشروعات الصغيرة.
-

## 7. حسب نوع المشروع (Franchise):

- مشاريع Franchise أقل تعثرًا (12.6%) من المشاريع المستقلة (34.3%).
  - التوصية: تشجيع الشراكات مع العلامات التجارية الكبرى لتقليل المخاطر.
- 

## 8. حسب نوع المنطقة (Area\_Type):

- المناطق الريفية أقل تعثرًا (18.8%) من الحضر (24.5%).
  - التوصية: تعزيز التمويل للمشاريع الريفية.
- 

## 9. حسب وجود خط انتمان (RevLineCr):

- المشاريع التي تمتلك خط انتمان (Y) تعاني من تعثر أعلى (26.8%).
  - التوصية: مراقبة استخدام خطوط الائتمان وتقنينها.
- 

## 10. حسب نوع القرض (LowDoc):

- قروض (Y LowDoc) تحقق سدادًا أعلى (91%) رغم قلة المستندات.
  - التوصية: توسيع هذا النوع مع معايير تقييم صارمة.
- 

## 11. حسب حجم مبلغ الصرف (DisbursementGross):

- القروض الصغيرة جدًا أكثر عرضة للتعثر (24%) مقابل 8-9% في القروض الكبيرة.
  - التوصية: دراسة أسباب ضعف السداد في القروض الصغيرة.
-

## 12. حسب مبلغ التعثر (ChgOffPrinGr):

- القروض التي بها أي تعثر ( $0 <$ ) لديها نسب فشل تفوق 89%.
- التوصية: تطوير سياسات الإنذار المبكر لمتابعة العملاء قبل التعثر.

## 13. حسب المبلغ الموافق عليه من البنك (GrAppv):

- القروض الصغيرة جدًا (Very Small) تتعثر بنسبة 29% مقابل 10% للقروض الكبيرة جدًا.
- التوصية: مراجعة آليات الموافقة على القروض الصغيرة.

## 14. حسب المبلغ الموافق عليه من (SBA (SBA\_Appv:

- القروض الصغيرة جدًا تتعثر بنسبة 29% مقابل 10% للقروض الكبيرة.
- التوصية: تحسين تقييم الجدارة الائتمانية للقروض الصغيرة من SBA.

## النتائج العامة والتوصيات النهائية:

1. القروض الصغيرة والمشروعات الصغيرة أكثر عرضة للتعثر.
2. القروض في المناطق الريفية تؤدي أداءً أفضل.
3. الانتماء إلى علامة تجارية (Franchise) يعزز الاستقرار المالي.
4. يجب إعادة تصميم نماذج تقييم المخاطر لتأخذ في الاعتبار حجم المشروع، نوع النشاط، والمنطقة الجغرافية.
5. يوصى بزيادة برامج التوعية المالية والإدارة للمشروعات الصغيرة والمتوسطة.

إعداد:

عفيفي ياسر عفيفي

المسمى الوظيفي: عالم بيانات (Data Scientist)