



# Credit Card Approval Analysis and Predictive Modeling

การวิเคราะห์และการสร้างแบบจำลองคาดการณ์การอนุมัติบัตรเครดิต

DS512/513 Data Analytics

DS514/515 Data Science

นำเสนอโดย

อาทิตย์ บูรณสิงห์

ID: 68199160309



# Index of Contents

## **1. Introduction**

- Data Analytics Framework
- Data Science Framework

## **2. Problem Statement/Background**

- What do we know about?
- What problem are you trying to solve?
- What is the business problem?

## **3. Questions/Hypothesis**

- Analytical Questions
- Predictive Hypothesis (What can we predict?)

## **4. Data Sources/Attributes**

- Data sources and Collection
- Data cleaning and Preprocessing

## **5. Analysis/Model Development**

- Analytics Methodology
- Modeling Methodology

## **6. Findings and Insights**

- Business Insights
- Predictive Results

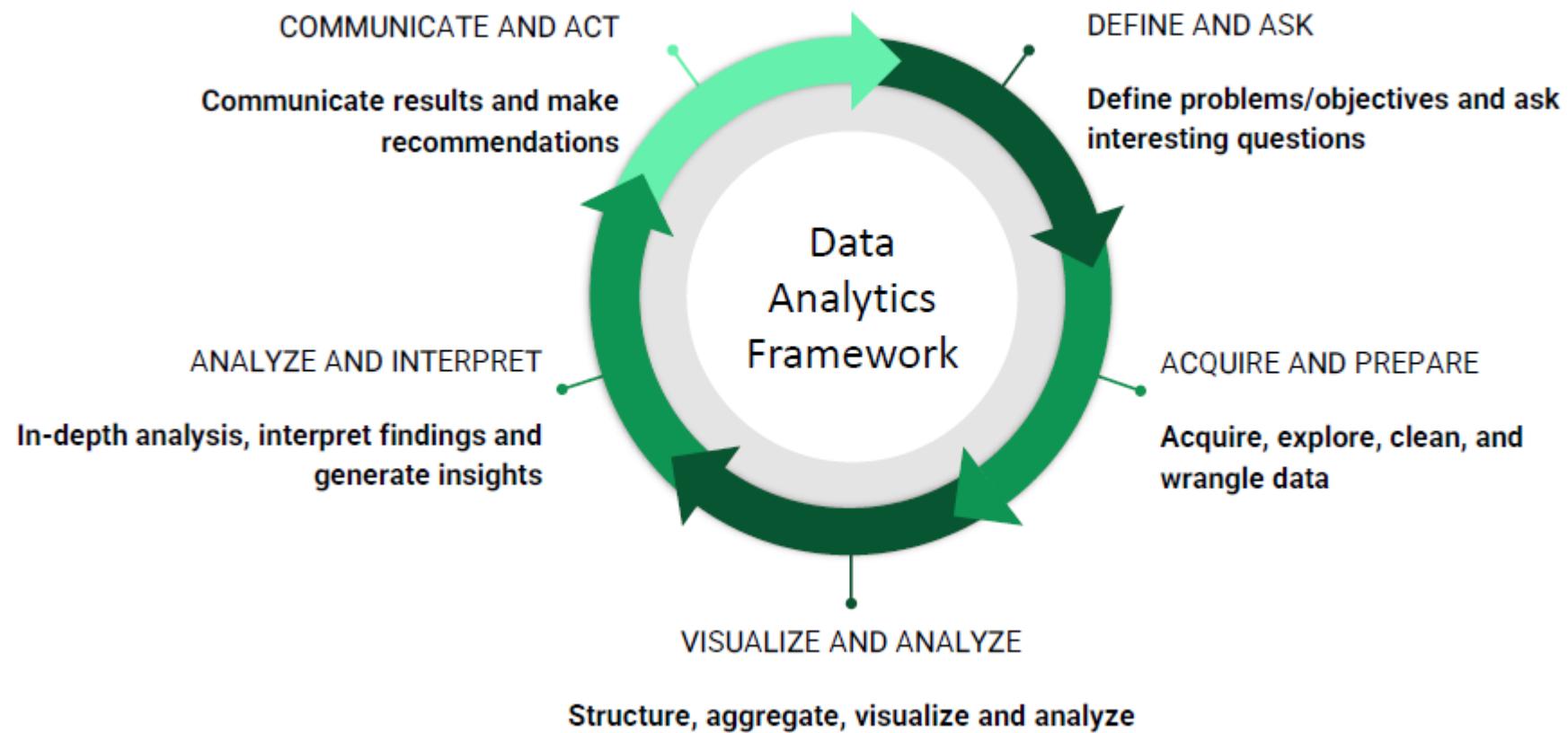
## **7. Conclusion and Recommendation/Action and Impact**

- What should we do with the findings?



## 1. Introduction

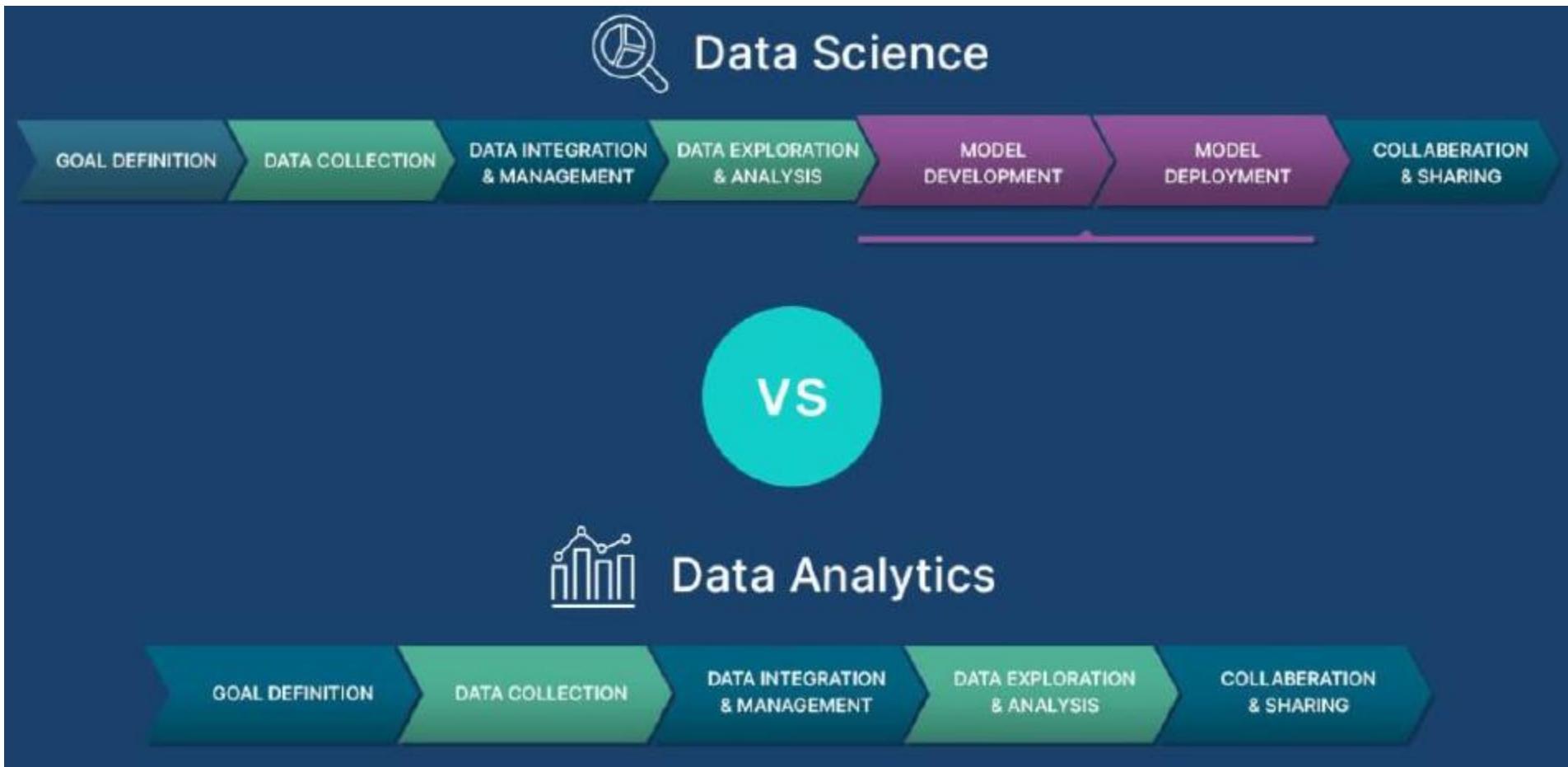
### Data Analytics Framework





## 1. Introduction

### Data Science Framework





# 1. Introduction

## DATA PROJECT CANVAS

### Title: Credit Card Approval Analysis and Predictive Modeling

#### 1. Problem Statement/Background



- What do we know about?

สถาบันการเงินอนุมัติ/ปฏิเสธการสมัครบัตรเครดิต โดยพิจารณาจากข้อมูลผู้สมัคร โดยกระบวนการพิจารณาอาจใช้เวลานานและขึ้นอยู่กับดุลยพินิจของนักวิเคราะห์สินเชื่อ ซึ่งอาจนำไปสู่ความเสี่ยงหรือการสูญเสียโอกาสทางธุรกิจ

- What problem are you trying to solve?

การสร้างแบบจำลองเพื่อทำนายผลการอนุมัติบัตรเครดิตโดยอัตโนมัติ เพื่อเพิ่มความรวดเร็วและความแม่นยำในการตัดสินใจ

- What is the business problem?

การลดความเสี่ยงทางการเงิน (อัตราการผิดนัดชำระหนี้ - Default Rate) และเพิ่มประสิทธิภาพในการดำเนินงาน (Operational Efficiency) ของฝ่ายอนุมัติสินเชื่อ

- Who are the stakeholders?

ธนาคาร/สถาบันการเงิน (ฝ่ายบริหารความเสี่ยง, ฝ่ายการตลาด, ฝ่ายปฏิบัติการ) ทีม Data Scientist/Analyst ผู้สมัครบัตรเครดิต

#### 5. Analysis/Model Development



- Analytics Methodology

- EDA techniques: การวิเคราะห์ความถี่และค่าสถิติเชิงพรรณนา (Univariate/Bivariate Analysis) เพื่อดูความสัมพันธ์ของแต่ละคุณลักษณะกับผลการอนุมัติ
- Visualization strategy: Heatmap สำหรับ Correlation Matrix
- Segmentation approach: การจัดกลุ่มลูกค้าตามโปรไฟล์ความเสี่ยง (Risk Profiles)

- Modeling Methodology

- Algorithm selection: Classification Algorithms - Logistic Regression

- Training, hyperparameter tuning, evaluation metric:

- \*\* Training: แบ่งชุดข้อมูลเป็น Training, และ Test Sets (Cross-Validation)

- \*\* Tuning: Grid Search เพื่อหา Hyperparameters ที่เหมาะสมที่สุด

Evaluation Metric: Accuracy, Precision, Recall, F1-Score

#### 2. Questions/Hypothesis



- Analytical Questions

- คุณลักษณะใดของผู้สมัคร (เช่น อายุ, รายได้, ประเภทงาน) ที่มีความสัมพันธ์กับการอนุมัติมากที่สุด?
- มีกลุ่มลูกค้า (Segments) ที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันขนาดไหนหรือไม่? เกณฑ์การตัดสินใจแบบเดิมมีข้อบกพร่องตรงไหนบ้าง?

- Predictive Hypothesis (What can we predict?)

- เราสามารถทำนายผลการอนุมัติบัตรเครดิต (Approved หรือ Rejected) จากข้อมูลของผู้สมัครที่ให้มาด้วยความแม่นยำสูง เพื่อช่วยในการตัดสินใจอัตโนมัติ

- SMART Objectives

- เพิ่มยอดสินเชื่อ จากการอนุมัติบัตรเครดิต อย่างน้อย 5% ภายใน 1 ปี ด้วยการปรับกลยุทธ์ด้านสินเชื่อ

#### 3. Value Propositions



- What are we trying to do for the end-user(s) of the system?

มอบเครื่องมือในการตัดสินใจที่รวดเร็ว เป็นกลาง และสม่ำเสมอให้กับเจ้าหน้าที่ สินเชื่อ ทำให้ผู้สมัครได้รับผลการตัดสินใจที่รวดเร็วยิ่งขึ้น

- What objectives are we serving?

เพิ่มประสิทธิภาพการดำเนินงาน (ลดต้นทุนและเวลา) ปรับปรุงการประเมินความเสี่ยง (ลดหนี้เสีย) ขยายฐานลูกค้าอย่างมีคุณภาพ

#### 4. Data Sources/Attributes



- Data sources & collection

- Kaggle Dataset: "Credit Card Approvals (Clean Data)" (ชุดข้อมูลจากชุมชน)

- ข้อมูลในทางปฏิบัติ: ฐานข้อมูลภายในของธนาคาร/สถาบันการเงิน (ข้อมูลการสมัคร, ประวัติสินเชื่อ, ข้อมูลเครดิตตู้ ATM)

- Data cleaning & preprocessing

- นำมาจาก "Clean Data" จะเน้นการตรวจสอบข้อผิดพลาดที่หายไป (Missing Values) ที่ยังอาจมีอยู่

- การเข้ารหัสข้อมูลประเภท Categorical (เช่น One-Hot Encoding) การปรับขนาดข้อมูลตัวเลข (Normalization/Standardization)

- Target variables & feature

- Target Variable: สถานะการอนุมัติ (Approved: '+', Rejected: '-')

- Features: คุณลักษณะของผู้สมัคร เช่น อายุ, รายได้, หนี้สิน, สถานภาพ, ประวัติเครดิต, ประเภทงาน (ตามคุณลักษณะใน Dataset)

- Encoding & scaling strategies

- Encoding: ข้อมูลที่ไม่ใช้ตัวเลข (Non-numerical/Categorical Data) ให้กลายเป็นตัวเลข โดยใช้เทคนิคที่เรียกว่า Label Encoding

- Scaling: Standard Scaler สำหรับตัวเลข (Numerical)

#### 6. Findings and Insights



- Business Insights

- คุณลักษณะหลักที่ขับเคลื่อนผลการอนุมัติ (Feature Importance)
- ค่าตอบ ปัจจัยด้าน ประวัติความเสี่ยง (PriorDefault) และ ความมั่นคง (Employed, CreditScore, YearsEmployed) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต

- ระบุกลุ่มลูกค้าที่มีความเสี่ยงสูง/ต่ำ ที่ดีเจนในชุดข้อมูล

- \* ค่าตอบ ชี้สูงสุด: ปัจจัย Credit Score เป็นปัจจัยขับเคลื่อนหลัก (Primary Driver) ในกระบวนการอนุมัติ ทำให้เกิดการแบ่งกลุ่มลูกค้าเป็นสองขั้วอย่างมีประสิทธิภาพ

- แนวโน้มและรูปแบบของผู้สมัครที่ถูกอนุมัติ/ปฏิเสธ

- \* ค่าตอบ ประวัติความเสี่ยง (PriorDefault) และ ความมั่นคง (Employed, CreditScore, YearsEmployed) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต ซึ่งตอบคำถามเชิงวิเคราะห์ว่า "คุณลักษณะใดของผู้สมัครมีความสัมพันธ์กับการอนุมัติมากที่สุด"

- Predictive Results

- ประสิทธิภาพของแบบจำลองที่ดี (F1-Score = 0.88)
- รายชื่อคุณลักษณะที่มีความสำคัญสูงสุด 4 อันดับแรกในการทำนายผล ประวัติความเสี่ยง (PriorDefault) และ ความมั่นคง (Employed, CreditScore, YearsEmployed)
- ผลการทดสอบกับ Test Set ที่ไม่ได้ใช้ในการฝึกฝน (Score = 0.88)

#### 7. Recommendation/Action and Impact



- What should we do with the findings?

- Action: นำแบบจำลองที่ปรับปรุงแล้วไปใช้งาน (Deploy) ในรูปแบบของ API หรือ Service เพื่อให้ระบบของธนาคารสามารถเรียกใช้ในการตัดสินใจอนุมัติเบื้องต้นโดยอัตโนมัติ

- What are the impacts?

- Operational Impact: ลดเวลาในการตัดสินใจจากชั่วโมง/วัน เหลือเพียงวินาที

- Business Impact: ลดการเกิดหนี้เสีย (ลดความเสี่ยง) เมื่อเวลาผ่านไป แม่นยำขึ้น แต่เพิ่มโอกาสในการอนุมัติลูกค้าที่มีคุณภาพเริ่วขึ้น (เพิ่มกำไร)

- Customer Impact: ผู้สมัครได้รับผลการตัดสินใจรวดเร็วขึ้น ทำให้ประสบการณ์การใช้งานดีขึ้น



## 2. Problem Statement/Background

### 2. Problem Statement/Background

- **What do we know about?**

สถาบันการเงินอนุมัติ/ปฏิเสธการสมัครบัตรเครดิต โดยพิจารณาจากข้อมูลผู้สมัคร โดยกระบวนการพิจารณาอาจใช้เวลานานและซึ่งอยู่กับดุลยพินิจของนักวิเคราะห์สินเชื่อซึ่งอาจนำไปสู่ความเสี่ยงหรือการสูญเสียโอกาสทางธุรกิจ

- **What problem are you trying to solve?**

การสร้างแบบจำลองเพื่อทำนายผลการอนุมัติบัตรเครดิตโดยอัตโนมัติ เพื่อเพิ่มความรวดเร็วและความแม่นยำในการตัดสินใจ

- **What is the business problem?**

การลดความเสี่ยงทางการเงิน (อัตราการผิดนัดชำระหนี้ - Default Rate) และเพิ่มประสิทธิภาพในการดำเนินงาน (Operational Efficiency) ของฝ่ายอนุมัติสินเชื่อ

- **Who are the stakeholders?**

ธนาคาร/สถาบันการเงิน (ฝ่ายบริหารความเสี่ยง, ฝ่ายการตลาด, ฝ่ายปฏิบัติการ) ทีม Data Scientist/Analyst ผู้สมัครบัตรเครดิต



### 3. Questions/Hypothesis

#### 3. Questions/Hypothesis

- **Analytical Questions**

- คุณลักษณะใดของผู้สมัคร (เช่น อายุ, รายได้, ประวัติงาน) ที่มีความสัมพันธ์กับการอนุมัติมากที่สุด?
- มีกลุ่มลูกค้า (Segments) ที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันชัดเจนหรือไม่? เกณฑ์การตัดสินใจแบบเดิมมีข้อบกพร่องตรงไหนบ้าง?

- **Predictive Hypothesis (What can we predict?)**

- เราสามารถทำนายผลการอนุมัติบัตรเครดิต (Approved หรือ Rejected) จากข้อมูลของผู้สมัครที่ให้มาด้วยความแม่นยำสูง เพื่อช่วยในการตัดสินใจอัตโนมัติ



## 4. Data Sources/Attributes

### 4. Data Sources/Attributes

- **Data Sources / Collection**

- **Kaggle Dataset:** "Credit Card Approvals (Clean Data)"  
(ชีว์จำลองมาจากการสมัครบัตรเครดิต)

- **ข้อมูลในทางปฏิบัติ:**  
ฐานข้อมูลภายในของธนาคาร/สถาบันการเงิน (ข้อมูลการสมัคร, ประวัติสินเชื่อ,  
ข้อมูลเครดิตบูโร)

- **Data cleaning / Preprocessing**

- เนื่องจากเป็น "Clean Data" จะเน้นการตรวจสอบขั้นสุดท้าย การจัดการกับค่าที่  
หายไป (Missing Values) ที่ยังอาจมีอยู่

- **Target Variables & Feature**

- **Target Variable:**  
สถานะการอนุมัติ (Approved: '1', Rejected: '0')

- **Features:**  
คุณลักษณะของผู้สมัคร เช่น อายุ, รายได้, หนี้สิน, สถานภาพ, ประวัติเครดิต, ประเภทงาน (ตามคุณลักษณะ  
ใน Dataset)

- **Encoding & Scaling strategies**

- **Encoding:** ข้อมูลที่ไม่ใช่ตัวเลข (Non-numerical/Categorical Data) ให้กลายเป็นตัวเลข โดยใช้  
เทคนิคที่เรียกว่า Label Encoding
- **Scaling:** Standard Scaler สำหรับตัวแปรตัวเลข (Numerical)



## 4. Data Sources/Attributes

### Credit Card Approvals - Clean version of UCI dataset (Raw Data)

| Gender | Age   | Debt   | Married | BankCustomer | Industry              | Ethnicity | YearsEmployed | PriorDefault | Employed | CreditScore | DriversLicense | Citizen      | ZipCode | Income | Approved |
|--------|-------|--------|---------|--------------|-----------------------|-----------|---------------|--------------|----------|-------------|----------------|--------------|---------|--------|----------|
| 1      | 30.83 | 0      | 1       | 1            | Industrials           | White     | 1.25          | 1            | 1        | 1           | 0              | ByBirth      | 202     | 0      | 1        |
| 0      | 58.67 | 4.46   | 1       | 1            | Materials             | Black     | 3.04          | 1            | 1        | 6           | 0              | ByBirth      | 43      | 560    | 1        |
| 0      | 24.5  | 0.5    | 1       | 1            | Materials             | Black     | 1.5           | 1            | 0        | 0           | 0              | ByBirth      | 280     | 824    | 1        |
| 1      | 27.83 | 1.54   | 1       | 1            | Industrials           | White     | 3.75          | 1            | 1        | 5           | 1              | ByBirth      | 100     | 3      | 1        |
| 1      | 20.17 | 5.625  | 1       | 1            | Industrials           | White     | 1.71          | 1            | 0        | 0           | 0              | ByOtherMeans | 120     | 0      | 1        |
| 1      | 32.08 | 4      | 1       | 1            | CommunicationServices | White     | 2.5           | 1            | 0        | 0           | 1              | ByBirth      | 360     | 0      | 1        |
| 1      | 33.17 | 1.04   | 1       | 1            | Transport             | Black     | 6.5           | 1            | 0        | 0           | 1              | ByBirth      | 164     | 31285  | 1        |
| 0      | 22.92 | 11.585 | 1       | 1            | InformationTechnology | White     | 0.04          | 1            | 0        | 0           | 0              | ByBirth      | 80      | 1349   | 1        |
| 1      | 54.42 | 0.5    | 0       | 0            | Financials            | Black     | 3.96          | 1            | 0        | 0           | 0              | ByBirth      | 180     | 314    | 1        |
| 1      | 42.5  | 4.915  | 0       | 0            | Industrials           | White     | 3.165         | 1            | 0        | 0           | 1              | ByBirth      | 52      | 1442   | 1        |
| 1      | 22.08 | 0.83   | 1       | 1            | Energy                | Black     | 2.165         | 0            | 0        | 0           | 1              | ByBirth      | 128     | 0      | 1        |
| 1      | 29.92 | 1.835  | 1       | 1            | Energy                | Black     | 4.335         | 1            | 0        | 0           | 0              | ByBirth      | 260     | 200    | 1        |
| 0      | 38.25 | 6      | 1       | 1            | Financials            | White     | 1             | 1            | 0        | 0           | 1              | ByBirth      | 0       | 0      | 1        |
| 1      | 48.08 | 6.04   | 1       | 1            | Financials            | White     | 0.04          | 0            | 0        | 0           | 0              | ByBirth      | 0       | 2690   | 1        |
| 0      | 45.83 | 10.5   | 1       | 1            | Materials             | White     | 5             | 1            | 1        | 7           | 1              | ByBirth      | 0       | 0      | 1        |
| 1      | 36.67 | 4.415  | 0       | 0            | Financials            | White     | 0.25          | 1            | 1        | 10          | 1              | ByBirth      | 320     | 0      | 1        |
| 1      | 28.25 | 0.875  | 1       | 1            | CommunicationServices | White     | 0.96          | 1            | 1        | 3           | 1              | ByBirth      | 396     | 0      | 1        |
| 0      | 23.25 | 5.875  | 1       | 1            | Materials             | White     | 3.17          | 1            | 1        | 10          | 0              | ByBirth      | 120     | 245    | 1        |
| 1      | 21.83 | 0.25   | 1       | 1            | Real Estate           | Black     | 0.665         | 1            | 0        | 0           | 1              | ByBirth      | 0       | 0      | 1        |

### Credit Card Approvals - Clean version of UCI dataset

Data Source: <https://www.kaggle.com/datasets/samuelcortinhas/credit-card-approval-clean-data/>

Missing values have been filled and feature names and categorical names have been inferred, resulting in more context and it being easier to use. Your task is to predict which people in the dataset are successful in applying for a credit card.



## 4. Data Sources/Attributes

### Cleansing Data

#### • Cleaning Data with Power Query

- เนื่องจากเป็นข้อมูลที่ได้มาเป็น "Clean Data" จะเน้นการตรวจสอบขั้นสุดท้าย การจัดการกับค่าที่หายไป (Missing Values) ที่ยังอาจมีอยู่ / Data Type ที่เหมาะสม รวมถึงการเพิ่ม Field เป็นต้น

|    | Gender               | Age                      | Debt                     | Married              | BankCustomer          | Industry              | Ethnicity               | YearsEmployed        | PriorDefault |
|----|----------------------|--------------------------|--------------------------|----------------------|-----------------------|-----------------------|-------------------------|----------------------|--------------|
|    | 2 distinct, 0 unique | 350 distinct, 170 unique | 215 distinct, 106 unique | 2 distinct, 0 unique | 14 distinct, 0 unique | 5 distinct, 0 unique  | 132 distinct, 56 unique | 2 distinct, 0 unique |              |
| 1  | 1                    | 30.83                    | 0                        | 1                    | 1                     | Industrials           | White                   | 1.25                 |              |
| 2  | 0                    | 58.67                    | 4.46                     | 1                    | 1                     | Materials             | Black                   | 3.04                 |              |
| 3  | 0                    | 24.5                     | 0.5                      | 1                    | 1                     | Materials             | Black                   | 1.5                  |              |
| 4  | 1                    | 27.83                    | 1.54                     | 1                    | 1                     | Industrials           | White                   | 3.75                 |              |
| 5  | 1                    | 20.17                    | 5.625                    | 1                    | 1                     | Industrials           | White                   | 1.71                 |              |
| 6  | 1                    | 32.08                    | 4                        | 1                    | 1                     | CommunicationServices | White                   | 2.5                  |              |
| 7  | 1                    | 33.17                    | 1.04                     | 1                    | 1                     | Transport             | Black                   | 6.5                  |              |
| 8  | 0                    | 22.92                    | 11.585                   | 1                    | 1                     | InformationTechnology | White                   | 0.04                 |              |
| 9  | 1                    | 54.42                    | 0.5                      | 0                    | 0                     | Financials            | Black                   | 3.96                 |              |
| 10 | 1                    | 42.5                     | 4.915                    | 0                    | 0                     | Industrials           | White                   | 3.165                |              |
| 11 | 1                    | 22.08                    | 0.83                     | 1                    | 1                     | Energy                | Black                   | 2.165                |              |
| 12 | 1                    | 29.92                    | 1.835                    | 1                    | 1                     | Energy                | Black                   | 4.335                |              |
| 13 | 0                    | 38.25                    | 6                        | 1                    | 1                     | Financials            | White                   | 1                    |              |
| 14 | 1                    | 48.08                    | 6.04                     | 1                    | 1                     | Financials            | White                   | 0.04                 |              |
| 15 | 0                    | 45.83                    | 10.5                     | 1                    | 1                     | Materials             | White                   | 5                    |              |
| 16 | 1                    | 36.67                    | 4.415                    | 0                    | 0                     | Financials            | White                   | 0.25                 |              |
| 17 | 1                    | 28.25                    | 0.875                    | 1                    | 1                     | CommunicationServices | White                   | 0.96                 |              |
| 18 | 0                    | 23.25                    | 5.875                    | 1                    | 1                     | Materials             | White                   | 3.17                 |              |
| 19 | 1                    | 21.83                    | 0.25                     | 1                    | 1                     | Real Estate           | Black                   | 0.665                |              |
| 20 | 0                    | 19.17                    | 8.585                    | 1                    | 1                     | InformationTechnology | Black                   | 0.75                 |              |



## 4. Data Sources/Attributes

### Cleaned Data

#### • Cleaning Data with Power Query

- เนื่องจากเป็นข้อมูลที่ได้มาเป็น "Clean Data" จะเน้นการตรวจสอบขั้นสุดท้าย การจัดการกับค่าที่หายไป (Missing Values) ที่ยังอาจมีอยู่ / Data Type ที่เหมาะสม รวมถึงการเพิ่ม Field เป็นต้น

| Gender | Age   | Debt  | Married | BankCustomer | Industry | Industry_Name         | Ethnicity | YearsEmployed | PriorDefault | Employed | CreditScore | DriversLicense | Citizen      | ZipCode | StateName            | Income | Approved |
|--------|-------|-------|---------|--------------|----------|-----------------------|-----------|---------------|--------------|----------|-------------|----------------|--------------|---------|----------------------|--------|----------|
| 1      | 30.83 | 0.00  | 1       | 1            | 1        | Industrials           | White     | 1.25          | 1            | 1        | 1.00        | 0              | ByBirth      | 00202   | District of Columbia | 0.00   | 1        |
| 1      | 27.83 | 1.54  | 1       | 1            | 1        | Industrials           | White     | 3.75          | 1            | 1        | 5.00        | 1              | ByBirth      | 00100   | Michigan             | 3.00   | 1        |
| 0      | 15.83 | 0.59  | 1       | 1            | 7        | Energy                | Black     | 1.50          | 1            | 1        | 2.00        | 0              | ByBirth      | 00100   | Michigan             | 0.00   | 1        |
| 1      | 23.92 | 0.67  | 1       | 1            | 7        | Energy                | White     | 0.17          | 0            | 0        | 0.00        | 0              | ByBirth      | 00100   | Michigan             | 0.00   | 1        |
| 0      | 49.00 | 1.50  | 1       | 1            | 14       | Research              | Other     | 0.00          | 1            | 0        | 0.00        | 1              | ByBirth      | 00100   | Michigan             | 27.00  | 0        |
| 1      | 22.75 | 11.00 | 1       | 1            | 2        | Materials             | White     | 2.50          | 1            | 1        | 7.00        | 1              | ByBirth      | 00100   | Michigan             | 809.00 | 1        |
| 1      | 20.42 | 1.84  | 1       | 1            | 7        | Energy                | White     | 2.25          | 1            | 1        | 1.00        | 0              | ByBirth      | 00100   | Michigan             | 150.00 | 1        |
| 1      | 43.00 | 0.29  | 0       | 0            | 5        | InformationTechnology | Black     | 1.75          | 1            | 1        | 8.00        | 0              | ByBirth      | 00100   | Michigan             | 375.00 | 1        |
| 0      | 58.67 | 4.46  | 1       | 1            | 2        | Materials             | Black     | 3.04          | 1            | 1        | 6.00        | 0              | ByBirth      | 00043   | Ohio                 | 560.00 | 1        |
| 1      | 20.17 | 5.63  | 1       | 1            | 1        | Industrials           | White     | 1.71          | 1            | 0        | 0.00        | 0              | ByOtherMeans | 00120   | Indiana              | 0.00   | 1        |
| 0      | 23.25 | 5.88  | 1       | 1            | 2        | Materials             | White     | 3.17          | 1            | 1        | 10.00       | 0              | ByBirth      | 00120   | Indiana              | 245.00 | 1        |
| 0      | 27.42 | 14.50 | 1       | 1            | 9        | Utilities             | Black     | 3.09          | 1            | 1        | 1.00        | 0              | ByBirth      | 00120   | Indiana              | 11.00  | 1        |
| 1      | 26.67 | 4.25  | 1       | 1            | 5        | InformationTechnology | White     | 4.29          | 1            | 1        | 1.00        | 1              | ByBirth      | 00120   | Indiana              | 0.00   | 1        |
| 1      | 31.42 | 15.50 | 1       | 1            | 7        | Energy                | White     | 0.50          | 1            | 0        | 0.00        | 0              | ByBirth      | 00120   | Indiana              | 0.00   | 0        |
| 1      | 25.00 | 12.00 | 1       | 1            | 6        | Financials            | White     | 2.25          | 1            | 1        | 2.00        | 1              | ByBirth      | 00120   | Indiana              | 5.00   | 0        |
| 0      | 25.00 | 11.00 | 0       | 0            | 12       | ConsumerStaples       | White     | 4.50          | 1            | 0        | 0.00        | 0              | ByBirth      | 00120   | Indiana              | 0.00   | 0        |
| 0      | 24.75 | 12.50 | 1       | 1            | 12       | ConsumerStaples       | White     | 1.50          | 1            | 1        | 12.00       | 1              | ByBirth      | 00120   | Indiana              | 567.00 | 1        |
| 0      | 23.50 | 9.00  | 1       | 1            | 2        | Materials             | White     | 8.50          | 1            | 1        | 5.00        | 1              | ByBirth      | 00120   | Indiana              | 0.00   | 1        |
| 0      | 24.50 | 0.50  | 1       | 1            | 2        | Materials             | Black     | 1.50          | 1            | 0        | 0.00        | 0              | ByBirth      | 00280   | North Carolina       | 824.00 | 1        |

18 Attributes / 690 Samples



## 4. Data Sources/Attributes

### Data Dictionary

| Attribute      | Description                                         | Data Type          | Valid Range/Example                                                                                                                                                                                               |
|----------------|-----------------------------------------------------|--------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Gender         | Gender                                              | Nominal            | 0=Female, 1=Male                                                                                                                                                                                                  |
| Age            | Age (Years)                                         | Interval           | [20, 80]                                                                                                                                                                                                          |
| Debt           | Outstanding debt (Feature has been scaled)          | Ratio (Continuous) | [0, Infinity)                                                                                                                                                                                                     |
| Married        | Married Status                                      | Nominal            | Married, 0=Single/Divorced/etc, 1=Married                                                                                                                                                                         |
| BankCustomer   | Bank Customer Status                                | Nominal            | 0=Does not have a bank account, 1=Has a bank account)                                                                                                                                                             |
| Industry       | Industry (Job sector of current or most recent job) | Nominal            | [0, 14]<br>[CommunicationServices, ConsumerDiscretionary, ConsumerStaples, Education, Energy, Financials, Healthcare, Industrials, InformationTechnology, Materials, Real Estate, Research, Transport, Utilities] |
| Industry_Name  | Industry Name                                       | Nominal            | White, Black, Asian, Latino, Other                                                                                                                                                                                |
| Ethnicity      | Ethnicity                                           | Nominal            | [0, Infinity)                                                                                                                                                                                                     |
| PriorDefault   | Prior default                                       | Asymmetric Binary  | 0=No prior defaults, 1=Prior default                                                                                                                                                                              |
| Employed       | Employed Status                                     | Asymmetric Binary  | 0=Not employed, 1=Employed                                                                                                                                                                                        |
| CreditScore    | Credit score (Feature has been scaled)              | Interval           | [0, 100]                                                                                                                                                                                                          |
| DriversLicense | Drivers license                                     | Symmetric Binary   | 0>No license, 1=Has license                                                                                                                                                                                       |
| Citizen        | Citizenship                                         | Nominal            | ByBirth, ByOtherMeans, Temporary                                                                                                                                                                                  |
| ZipCode        | ZipCode                                             | Nominal            | 5 digit number                                                                                                                                                                                                    |
| StateName      | StateName                                           | Nominal            | [100 = Michigan, etc.]                                                                                                                                                                                            |
| Income         | Income (Feature has been scaled)                    | Ratio (Continuous) | [0, Infinity)                                                                                                                                                                                                     |
| Approved       | Approved Status                                     | Asymmetric Binary  | 0=Not approved, 1=Approved                                                                                                                                                                                        |



## 5. Analysis/Model Development

### Analytics Methodology

- **Exploratory Data Analysis - EDA techniques:** การวิเคราะห์ข้อมูลเบื้องต้น เช่น การวิเคราะห์ความถี่ และค่าสถิติเชิงพรรณนา (Descriptive Statistics) เพื่อดูความสัมพันธ์ของแต่ละคุณลักษณะกับผลการอนุมัติ การหาความสัมพันธ์ระหว่างตัวแปร (Correlation)
- **Visualization strategy:** Heatmap สำหรับ Correlation Matrix
- **Segmentation approach:** การจัดกลุ่มลูกค้าตามโปรไฟล์ความเสี่ยง (Risk Profiles)



## 5. Analysis/Model Development

### Analytics Methodology

- EDA techniques:** การวิเคราะห์ความถี่และค่าสถิติเชิงพรรณนา เพื่อถูความสัมพันธ์ของแต่ละคุณลักษณะกับผลการอนุมัติ

|              | Gender     | Age        | Debt       | Married    | BankCustomer | Industry   | YearsEmployed | PriorDefault | Employed   | CreditScore | DriversLicense | ZipCode     | Income        | Approved   |
|--------------|------------|------------|------------|------------|--------------|------------|---------------|--------------|------------|-------------|----------------|-------------|---------------|------------|
| <b>count</b> | 690.000000 | 690.000000 | 690.000000 | 690.000000 | 690.000000   | 690.000000 | 690.000000    | 690.000000   | 690.000000 | 690.000000  | 690.000000     | 690.000000  | 690.000000    | 690.000000 |
| <b>mean</b>  | 0.695652   | 31.514116  | 4.758725   | 0.760870   | 0.763768     | 6.960870   | 2.223406      | 0.523188     | 0.427536   | 2.40000     | 0.457971       | 180.547826  | 1017.385507   | 0.444928   |
| <b>std</b>   | 0.460464   | 11.860245  | 4.978163   | 0.426862   | 0.425074     | 3.802822   | 3.346513      | 0.499824     | 0.495080   | 4.86294     | 0.498592       | 173.970323  | 5210.102598   | 0.497318   |
| <b>min</b>   | 0.000000   | 13.750000  | 0.000000   | 0.000000   | 0.000000     | 1.000000   | 0.000000      | 0.000000     | 0.000000   | 0.000000    | 0.000000       | 0.000000    | 0.000000      | 0.000000   |
| <b>25%</b>   | 0.000000   | 22.670000  | 1.000000   | 1.000000   | 1.000000     | 3.000000   | 0.165000      | 0.000000     | 0.000000   | 0.000000    | 0.000000       | 60.000000   | 0.000000      | 0.000000   |
| <b>50%</b>   | 1.000000   | 28.460000  | 2.750000   | 1.000000   | 1.000000     | 7.000000   | 1.000000      | 1.000000     | 0.000000   | 0.000000    | 0.000000       | 160.000000  | 5.000000      | 0.000000   |
| <b>75%</b>   | 1.000000   | 37.707500  | 7.207500   | 1.000000   | 1.000000     | 10.000000  | 2.625000      | 1.000000     | 1.000000   | 3.00000     | 1.000000       | 272.000000  | 395.500000    | 1.000000   |
| <b>max</b>   | 1.000000   | 80.250000  | 28.000000  | 1.000000   | 1.000000     | 14.000000  | 28.500000     | 1.000000     | 1.000000   | 67.00000    | 1.000000       | 2000.000000 | 100000.000000 | 1.000000   |

สถิติเชิงพรรณนา (Descriptive Statistics)



## 5. Analysis/Model Development

### สถิติเชิงพรรณนา (Descriptive Statistics)

| Income             |               | Debt               |          | CreditScore        |          | Age                |           |
|--------------------|---------------|--------------------|----------|--------------------|----------|--------------------|-----------|
| Mean               | 1,017.39      | Mean               | 4.76     | Mean               | 2.40     | Mean               | 31.51     |
| Standard Error     | 198.35        | Standard Error     | 0.19     | Standard Error     | 0.19     | Standard Error     | 0.45      |
| Median             | 5.00          | Median             | 2.75     | Median             | -        | Median             | 28.46     |
| Mode               | -             | Mode               | 1.50     | Mode               | -        | Mode               | 28.46     |
| Standard Deviation | 5,210.10      | Standard Deviation | 4.98     | Standard Deviation | 4.86     | Standard Deviation | 11.86     |
| Sample Variance    | 27,145,169.08 | Sample Variance    | 24.78    | Sample Variance    | 23.65    | Sample Variance    | 140.67    |
| Kurtosis           | 214.67        | Kurtosis           | 2.27     | Kurtosis           | 50.83    | Kurtosis           | 1.20      |
| Skewness           | 13.14         | Skewness           | 1.49     | Skewness           | 5.15     | Skewness           | 1.17      |
| Range              | 100,000.00    | Range              | 28.00    | Range              | 67.00    | Range              | 66.50     |
| Minimum            | -             | Minimum            | -        | Minimum            | -        | Minimum            | 13.75     |
| Maximum            | 100,000.00    | Maximum            | 28.00    | Maximum            | 67.00    | Maximum            | 80.25     |
| Sum                | 701,996.00    | Sum                | 3,283.52 | Sum                | 1,656.00 | Sum                | 21,744.74 |
| Count              | 690.00        | Count              | 690.00   | Count              | 690.00   | Count              | 690.00    |

### Statistic Description

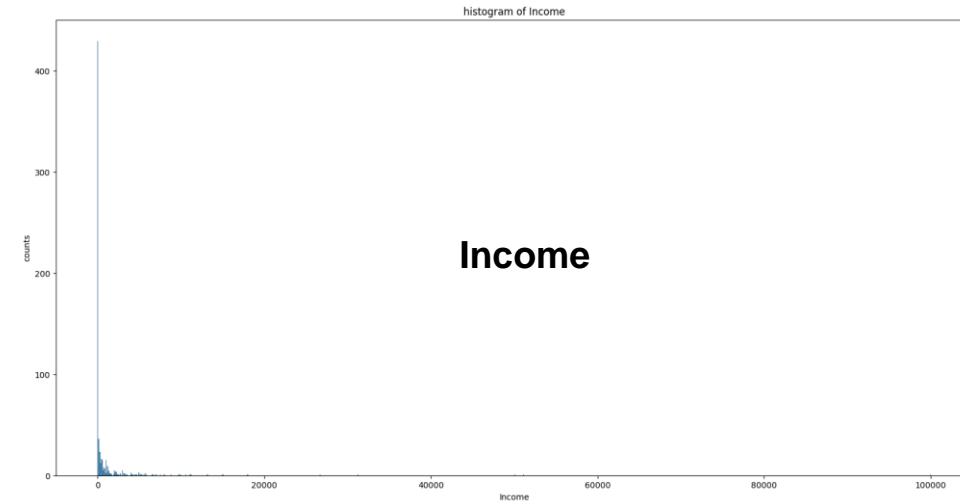
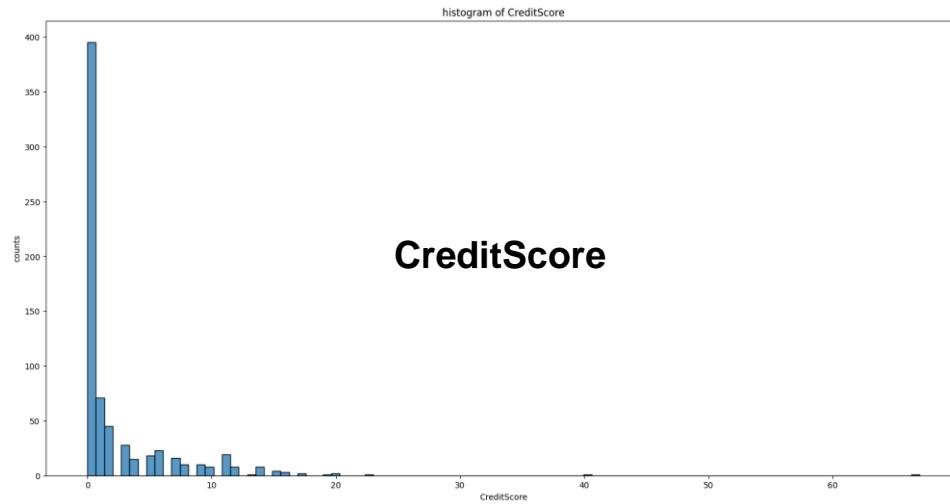
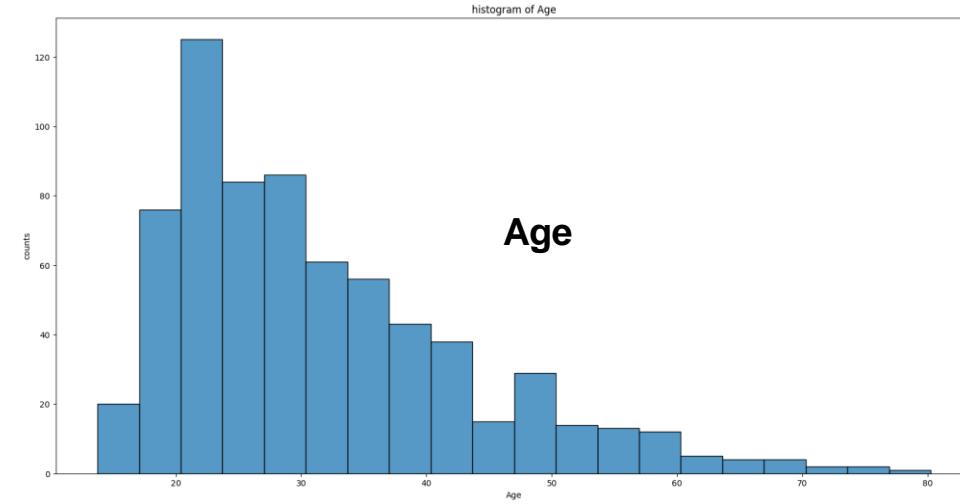
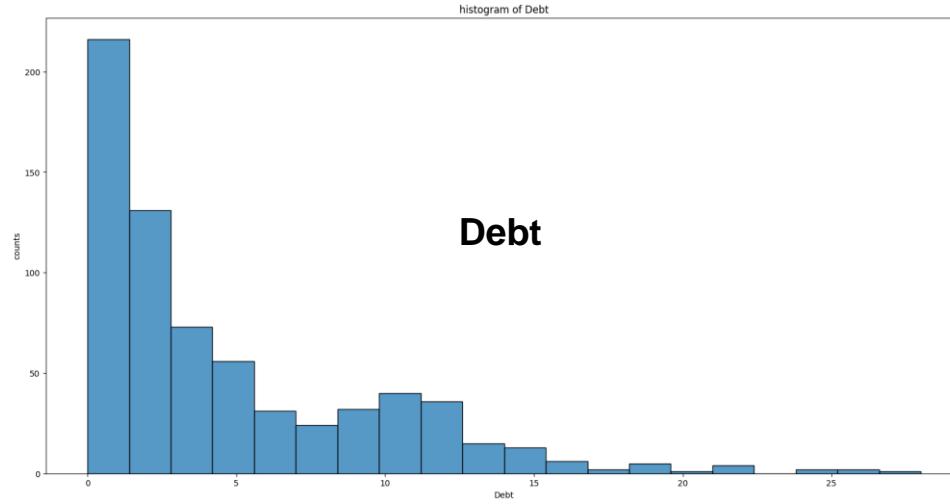
**Income** ชุดข้อมูลประกอบด้วยข้อมูลสังเกตทั้งหมด 690 รายการ ค่ารายได้มีช่วงตั้งแต่ต่ำสุด 0 K ถึงสูงสุด 100,000 K โดยมีช่วงทั้งหมด (Range) 100,000 ผลรวมรายได้ทั้งหมดคือ 701,996 ข้อมูลมีความเบ้สูง รายได้เฉลี่ยอยู่ที่ประมาณ 1,017 K ขณะที่ค่ามัธยฐาน (Median) อยู่ที่เพียง 5 และค่าฐานนิยม (Mode) อยู่ที่ 0 ความแตกต่างที่มากระหว่างค่าเฉลี่ย ค่ามัธยฐาน และค่าฐานนิยมนี้บ่งชี้ว่าข้อมูลมีความเบี้ปีทางข้าวย่างมาก ซึ่งหมายความว่ารายได้สูงมากบางค่าเดียวอย่างมีนัยสำคัญ ปัจจัยนี้ยังได้รับการสนับสนุนจากค่าเบี้ย (Skewness) ที่สูงมากที่ 13.14 และค่าความโด่ง (Kurtosis) ที่สูงที่ 214.67 ซึ่งแสดงถึงค่าผิดปกติจำนวนมากและการกระจายตัวที่จุดสูงสุด ค่าเบี้ยงเบนมาตรฐาน (SD) ที่ 5,210.10 ค่อนข้างสูง ซึ่งยิ่งยืนยันการกระจายตัวที่กว้างของข้อมูลและการมีอิทธิพลของค่าผิดปกติที่มีนัยสำคัญ

**CreditScore** ชุดข้อมูลประกอบด้วยข้อมูลการสังเกต 690 รายการ คะแนนเครดิตมีช่วงตั้งแต่ต่ำสุด 0 ถึงสูงสุด 67 โดยมีช่วงคะแนนรวม 67 คะแนน ผลรวมของคะแนนเครดิตทั้งหมดคือ 1,656 ข้อมูลมีความเบ้สูง คะแนนเครดิตเฉลี่ยอยู่ที่ 2.4 ขณะที่ค่ามัธยฐานและค่าฐานนิยมอยู่ที่ 0 ทั้งคู่ ความแตกต่างที่สำคัญนี้ชี้ให้เห็นว่าข้อมูลการสังเกตจำนวนมากอยู่ที่หรือใกล้ศูนย์ และมีค่าสูงบางค่าที่ทำให้ค่าเฉลี่ยสูงขึ้น ค่าความเบ้ที่สูงที่ 5.15 ยืนยันว่ามีการเบี้ปีทางข้าวยัดเจน ค่าเบี้ยงเบนมาตรฐานอยู่ที่ 4.86 ซึ่งบ่งชี้ถึงการกระจายตัวของข้อมูลอย่างกว้างขวาง ค่าความโด่งที่สูงมากที่ 50.83 บ่งชี้ถึงการกระจายตัวที่มีจุดสูงสุดและหางที่มาก ซึ่งอาจเป็นผลมาจากการค่าจำนวนมาก ณ จุดหนึ่ง (ศูนย์) และค่าผิดปกติที่มีนัยสำคัญมาก



## 5. Analysis/Model Development

### Graph – Visualization – Pivot Table

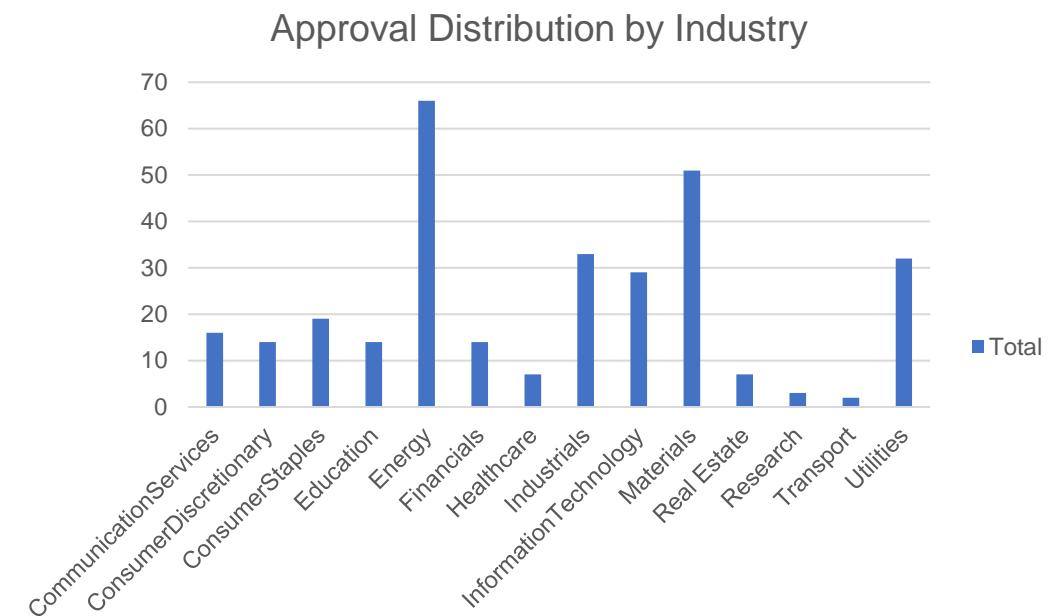




## 5. Analysis/Model Development

### Graph – Visualization – Pivot Table

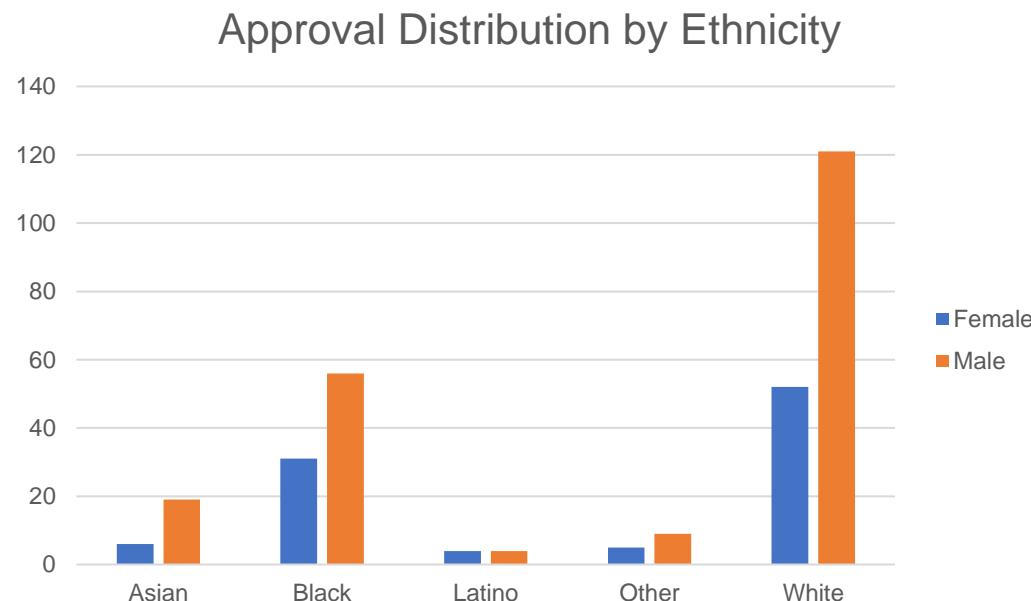
| Industry Name         | Number of Approved |
|-----------------------|--------------------|
| CommunicationServices | 16                 |
| ConsumerDiscretionary | 14                 |
| ConsumerStaples       | 19                 |
| Education             | 14                 |
| Energy                | 66                 |
| Financials            | 14                 |
| Healthcare            | 7                  |
| Industrials           | 33                 |
| InformationTechnology | 29                 |
| Materials             | 51                 |
| Real Estate           | 7                  |
| Research              | 3                  |
| Transport             | 2                  |
| Utilities             | 32                 |
| Grand Total           | 307                |





## 5. Analysis/Model Development

Graph – Visualization – Pivot Table





## 5. Analysis/Model Development

### Pearson Correlation (สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน)

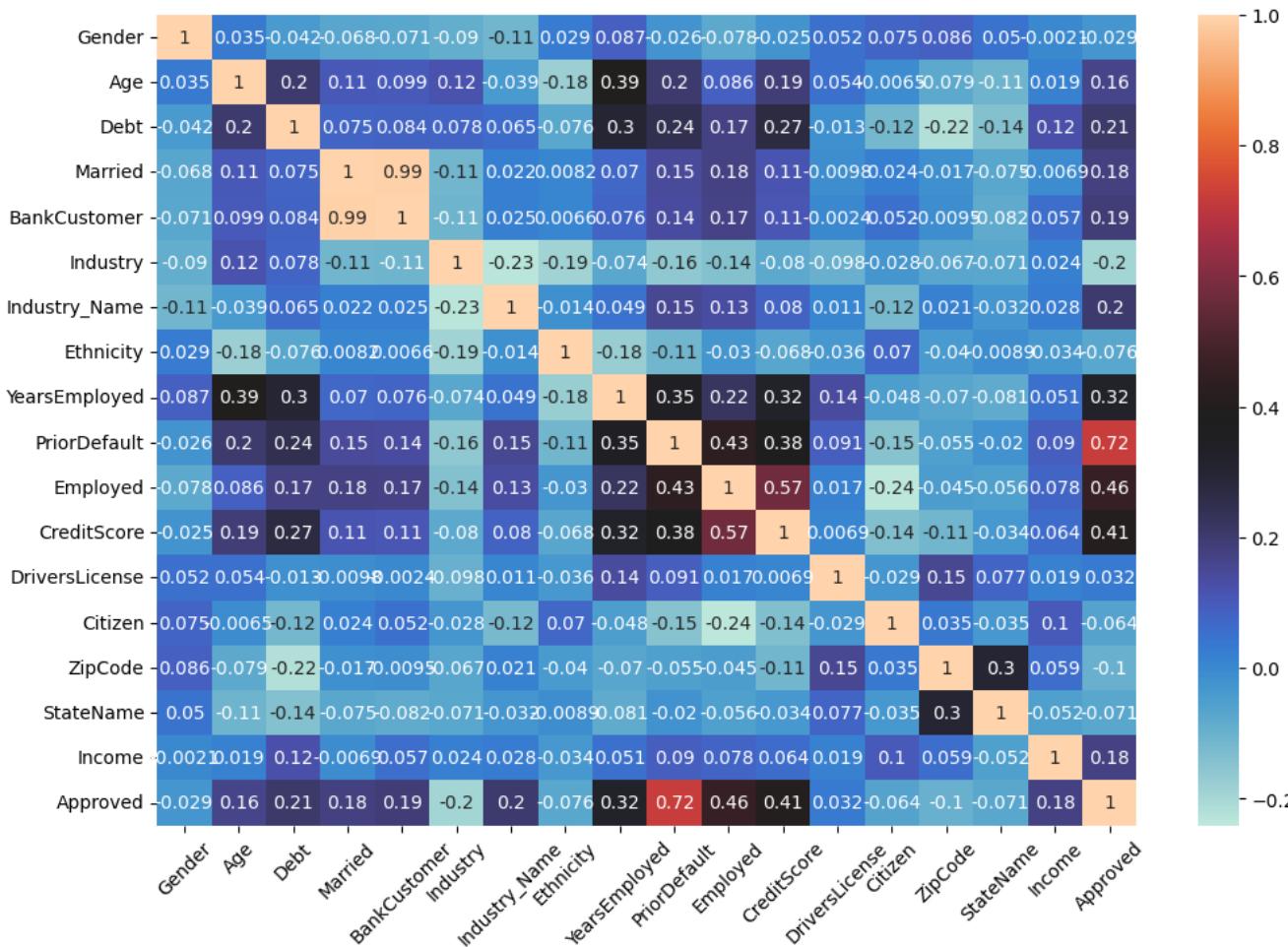
- Pearson correlation (สัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน) คือ ค่าที่ใช้วัดความสัมพันธ์เชิงเส้น (Linear Relationship) ระหว่างตัวแปรเชิงปริมาณ (Quantitative Variables) สองตัว มีค่าดังตาราง

|                       | Gender    | Age       | Debt      | Married   | BankCustomer | Industry  | Industry_Name | Ethnicity | YearsEmployed | PriorDefault | Employed  | CreditScore | DriversLicense | Citizen   | ZipCode   | StateName | Income    | Approved  |
|-----------------------|-----------|-----------|-----------|-----------|--------------|-----------|---------------|-----------|---------------|--------------|-----------|-------------|----------------|-----------|-----------|-----------|-----------|-----------|
| <b>Gender</b>         | 1.000000  | 0.035044  | -0.041746 | -0.068062 | -0.071250    | -0.089697 | -0.111889     | 0.029492  | 0.086544      | -0.026047    | -0.077784 | -0.024630   | 0.051674       | 0.075413  | 0.086007  | 0.049724  | -0.002063 | -0.028934 |
| <b>Age</b>            | 0.035044  | 1.000000  | 0.202177  | 0.106929  | 0.099477     | 0.120104  | -0.038746     | -0.179534 | 0.391464      | 0.204434     | 0.086037  | 0.187327    | 0.053599       | -0.006481 | -0.078690 | -0.107888 | 0.018719  | 0.164086  |
| <b>Debt</b>           | -0.041746 | 0.202177  | 1.000000  | 0.074649  | 0.083781     | 0.078030  | 0.064553      | -0.075789 | 0.298902      | 0.244317     | 0.174846  | 0.271207    | -0.013023      | -0.116975 | -0.217903 | -0.141502 | 0.123121  | 0.206294  |
| <b>Married</b>        | -0.068062 | 0.106929  | 0.074649  | 1.000000  | 0.992033     | -0.111277 | 0.021652      | 0.008226  | 0.069945      | 0.145073     | 0.175428  | 0.113968    | -0.009784      | 0.024319  | -0.017074 | -0.075021 | -0.006899 | 0.180583  |
| <b>BankCustomer</b>   | -0.071250 | 0.099477  | 0.083781  | 0.992033  | 1.000000     | -0.108083 | 0.024677      | 0.006648  | 0.075905      | 0.138535     | 0.170268  | 0.111077    | -0.002402      | 0.052141  | -0.009513 | -0.082177 | 0.057273  | 0.188964  |
| <b>Industry</b>       | -0.089697 | 0.120104  | 0.078030  | -0.111277 | -0.108083    | 1.000000  | -0.232826     | -0.190879 | -0.074235     | -0.161784    | -0.143740 | -0.079754   | -0.097701      | -0.027940 | -0.067478 | -0.071252 | 0.023707  | -0.196453 |
| <b>Industry_Name</b>  | -0.111889 | -0.038746 | 0.064553  | 0.021652  | 0.024677     | -0.232826 | 1.000000      | -0.013881 | 0.048689      | 0.154645     | 0.134769  | 0.080107    | 0.011010       | -0.122211 | 0.020551  | -0.032218 | 0.027820  | 0.202158  |
| <b>Ethnicity</b>      | 0.029492  | -0.179534 | -0.075789 | 0.008226  | 0.006648     | -0.190879 | -0.013881     | 1.000000  | -0.177111     | -0.114148    | -0.030250 | -0.068072   | -0.035688      | 0.070235  | -0.040003 | -0.008858 | -0.034251 | -0.075558 |
| <b>YearsEmployed</b>  | 0.086544  | 0.391464  | 0.298902  | 0.069945  | 0.075905     | -0.074235 | 0.048689      | -0.177111 | 1.000000      | 0.345689     | 0.222982  | 0.322330    | 0.138139       | -0.047522 | -0.070495 | -0.080999 | 0.051345  | 0.322475  |
| <b>PriorDefault</b>   | -0.026047 | 0.204434  | 0.244317  | 0.145073  | 0.138535     | -0.161784 | 0.154645      | -0.114148 | 0.345689      | 1.000000     | 0.432032  | 0.379532    | 0.091276       | -0.145357 | -0.055010 | -0.020499 | 0.090012  | 0.720407  |
| <b>Employed</b>       | -0.077784 | 0.086037  | 0.174846  | 0.175428  | 0.170268     | -0.143740 | 0.134769      | -0.030250 | 0.222982      | 0.432032     | 1.000000  | 0.571498    | 0.017043       | -0.240789 | -0.044834 | -0.055918 | 0.077652  | 0.458301  |
| <b>CreditScore</b>    | -0.024630 | 0.187327  | 0.271207  | 0.113968  | 0.111077     | -0.079754 | 0.080107      | -0.068072 | 0.322330      | 0.379532     | 0.571498  | 1.000000    | 0.006944       | -0.138341 | -0.112816 | -0.033683 | 0.063692  | 0.406410  |
| <b>DriversLicense</b> | 0.051674  | 0.053599  | -0.013023 | -0.009784 | -0.002402    | -0.097701 | 0.011010      | -0.035688 | 0.138139      | 0.091276     | 0.017043  | 0.006944    | 1.000000       | -0.029087 | 0.154924  | 0.077415  | 0.019201  | 0.031625  |
| <b>Citizen</b>        | 0.075413  | -0.006481 | -0.116975 | 0.024319  | 0.052141     | -0.027940 | -0.122211     | 0.070235  | -0.047522     | -0.145357    | -0.240789 | -0.138341   | -0.029087      | 1.000000  | 0.035488  | -0.035296 | 0.102000  | -0.063556 |
| <b>ZipCode</b>        | 0.086007  | -0.078690 | -0.217903 | -0.017074 | -0.009513    | -0.067478 | 0.020551      | -0.040003 | -0.070495     | -0.055010    | -0.044834 | -0.112816   | 0.154924       | 0.035488  | 1.000000  | 0.301966  | 0.059234  | -0.099598 |
| <b>StateName</b>      | 0.049724  | -0.107888 | -0.141502 | -0.075021 | -0.082177    | -0.071252 | -0.032218     | -0.008858 | -0.080999     | -0.020499    | -0.055918 | -0.033683   | 0.077415       | -0.035296 | 0.301966  | 1.000000  | -0.052210 | -0.071361 |
| <b>Income</b>         | -0.002063 | 0.018719  | 0.123121  | -0.006899 | 0.057273     | 0.023707  | 0.027820      | -0.034251 | 0.051345      | 0.090012     | 0.077652  | 0.063692    | 0.019201       | 0.102000  | 0.059234  | -0.052210 | 1.000000  | 0.175657  |
| <b>Approved</b>       | -0.028934 | 0.164086  | 0.206294  | 0.180583  | 0.188964     | -0.196453 | 0.202158      | -0.075558 | 0.322475      | 0.720407     | 0.458301  | 0.406410    | 0.031625       | -0.063556 | -0.099598 | -0.071361 | 0.175657  | 1.000000  |



## 5. Analysis/Model Development

**Correlation Heatmap (แผนที่ความร้อนสหสัมพันธ์)**



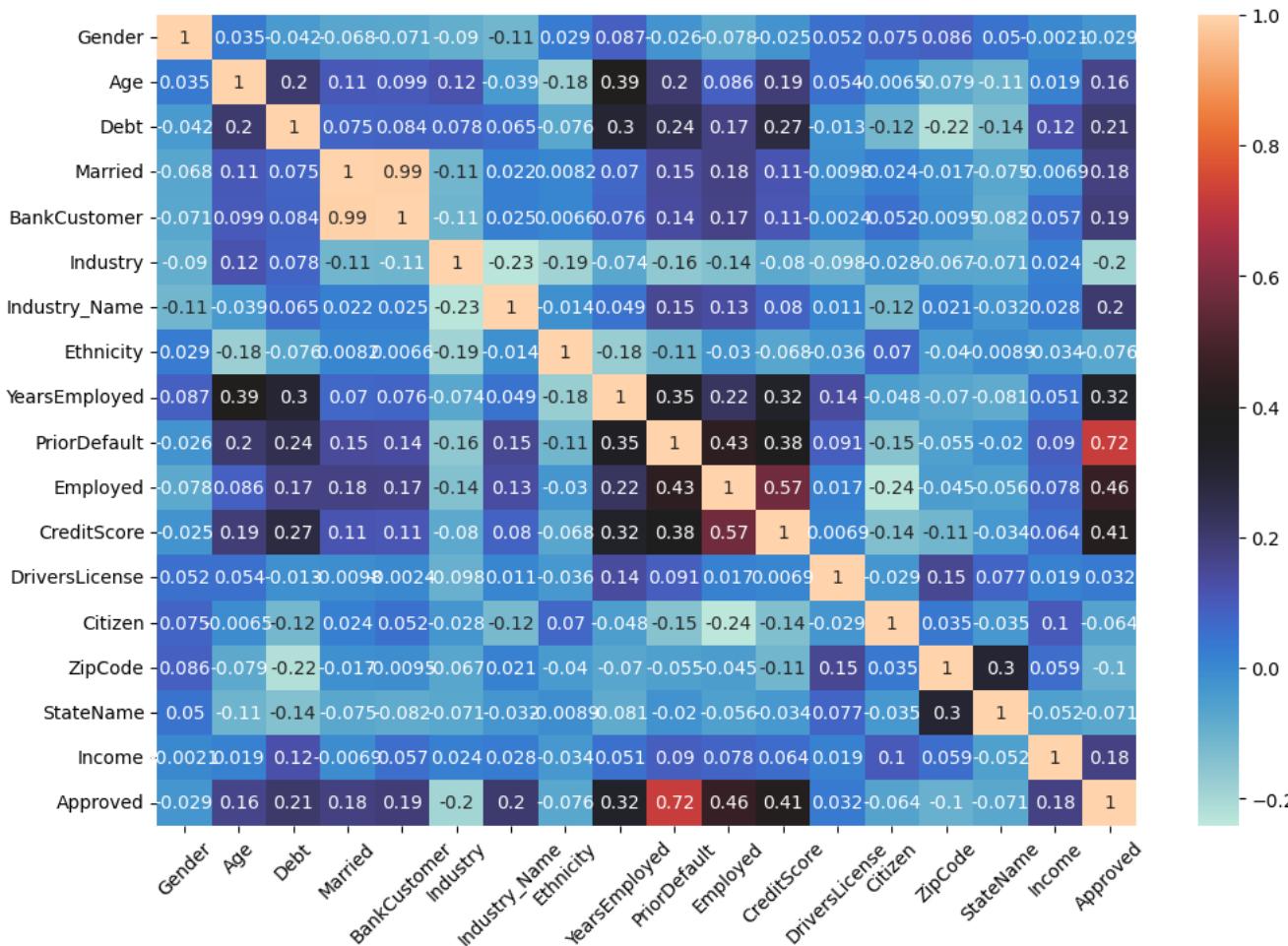
ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) ที่มีค่ามากกว่า 0.3 (ทั้งค่าบวกและลบ) หมายความว่ามีความสัมพันธ์เชิงเส้นระหว่างตัวแปรทั้งสองในระดับ ปานกลางถึงค่อนข้างมาก โดยมีค่าดังนี้

Married and BankCustomer = 0.992  
 PriorDefault and Approved = 0.720  
 Employed and CreditScore = 0.571  
 Employed and Approved = 0.458  
 PriorDefault and Employed = 0.432  
 CreditScore and Approved = 0.406  
 Age and YearsEmployed = 0.391  
 PriorDefault and CreditScore = 0.380  
 YearsEmployed and PriorDefault = 0.346  
 YearsEmployed and Approved = 0.322  
 YearsEmployed and CreditScore = 0.322  
 ZipCode and StateName = 0.302



## 5. Analysis/Model Development

### Correlation Heatmap (แผนที่ความร้อนสหสัมพันธ์)



ค่าสัมประสิทธิ์สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) ที่มีค่ามากกว่า 0.3 (ทั้งค่าบวกและลบ) หมายความว่ามีความสัมพันธ์เชิงเส้นระหว่างตัวแปรทั้งสองในระดับ ปานกลางถึงค่อนข้างมาก โดยมีค่าดังนี้

Married and BankCustomer = 0.992  
PriorDefault and Approved = 0.720  
Employed and CreditScore = 0.571  
Employed and Approved = 0.458  
PriorDefault and Employed = 0.432  
CreditScore and Approved = 0.406  
Age and YearsEmployed = 0.391  
PriorDefault and CreditScore = 0.380  
YearsEmployed and PriorDefault = 0.346  
YearsEmployed and Approved = 0.322  
YearsEmployed and CreditScore = 0.322  
ZipCode and StateName = 0.302



## 5. Analysis/Model Development

### การวิเคราะห์ข้อมูลเบื้องต้น (Analytical Insights)

ผลจากการวิเคราะห์ความสัมพันธ์ของฟีเจอร์ (Correlation Analysis) เป็นไปตามสมมติฐานที่ว่าคุณลักษณะบางอย่างของผู้สมัครมีความสัมพันธ์อย่างมากกับผลการอนุมัติ

| ความสัมพันธ์ที่โดดเด่น     | ค่า Correlation | ความหมายเชิงธุรกิจ                                                                                                            |
|----------------------------|-----------------|-------------------------------------------------------------------------------------------------------------------------------|
| PriorDefault and Approved  | 0.720           | สำคัญที่สุด: ผู้สมัครที่มีประวัติไม่เคยผิดนัดชำระหนี้ ( <b>PriorDefault=0</b> ) มีโอกาสได้รับการอนุมัติสูงมาก (ความเสี่ยงต่ำ) |
| Employed and Approved      | 0.458           | ผู้สมัครที่มีงานทำ ( <b>Employed=1</b> ) มีแนวโน้มได้รับการอนุมัติสูงกว่า                                                     |
| CreditScore and Approved   | 0.406           | คะแนนเครดิตสูง เป็นปัจจัยหลักในการอนุมัติ                                                                                     |
| YearsEmployed and Approved | 0.322           | ระยะเวลาการทำงานที่ยาวนาน มีความสัมพันธ์กับการอนุมัติ ซึ่งบ่งชี้ถึงความมั่นคงทางการเงิน                                       |

บทสรุปเชิงวิเคราะห์: ปัจจัยด้าน ประวัติความเสี่ยง (PriorDefault) และ ความมั่นคง (Employed, CreditScore, YearsEmployed) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต ซึ่งตอบคำถามเชิงวิเคราะห์ที่ว่า "คุณลักษณะใดของผู้สมัครมีความสัมพันธ์กับการอนุมัติมากที่สุด"



## 5. Analysis/Model Development

### การวิเคราะห์ข้อมูลเบื้องต้น (Analytical Insights)

การเปรียบเทียบ Credit Score (CreditScore) ระหว่างผู้ที่ได้รับการอนุมัติ (Approved=1) และผู้ที่ไม่ได้รับการอนุมัติ (Approved=0) สามารถทำ t-test ได้เนื่องจาก CreditScore เป็นตัวแปรเชิงปริมาณ (Numerical/Continuous) และเราต้องการเปรียบเทียบค่าเฉลี่ยของตัวแปรนี้ระหว่างสองกลุ่มอิสระ (Approved=1 และ Approved=0) การใช้ Independent Samples T-test (การทดสอบที่สำหรับกลุ่มตัวอย่างอิสระ) จึงเป็นวิธีที่เหมาะสมที่สุด

#### Independent Samples T-test

สมมติฐาน

$H_0$  (สมมติฐานว่าง): ค่าเฉลี่ย CreditScore ของผู้ที่ได้รับการอนุมัติ ไม่แตกต่าง จากผู้ที่ไม่ได้รับการอนุมัติ

$H_a$  (สมมติฐานทางเลือก): ค่าเฉลี่ย CreditScore ของผู้ที่ได้รับการอนุมัติ แตกต่าง จากผู้ที่ไม่ได้รับการอนุมัติ

ค่าเฉลี่ย CreditScore กลุ่ม Approved (1): 4.6059 (307 Samples)

ค่าเฉลี่ย CreditScore กลุ่ม Not Approved (0): 0.6319 (383 Samples)

ค่า T-statistic: 10.6384

ค่า T ที่สูงมากนี้บ่งชี้ว่ามีความแตกต่างระหว่างค่าเฉลี่ยของทั้งสองกลุ่มค่อนข้างมากเมื่อเทียบกับความผันแปรภายในกลุ่ม

ค่า P-value: 0.0000

**P-value < 0.05:** ปฏิเสธ  $H_0$  สรุปได้ว่า Credit Score เฉลี่ยแตกต่างกันอย่างมีนัยสำคัญ

ผลจากการวิเคราะห์ความสัมพันธ์ของฟีเจอร์ (Correlation Analysis) เป็นไปตามสมมติฐานที่ว่ามีกลุ่มลูกค้า (Segments) ที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันชัดเจน หรือไม่? -- สรุป ค่าเฉลี่ย Credit Score ของผู้ที่ได้รับการอนุมัติ แตกต่าง จากผู้ที่ไม่ได้รับการอนุมัติ



## 5. Analysis/Model Development

### บทสรุปเชิงกลยุทธ์: ความเสี่ยงและประสิทธิภาพของ Credit Score

จากการวิเคราะห์ T-test พบว่า Credit Score เป็นตัวแปรที่มีผลลัพธ์ในการจำแนกสูงมากระหว่างกลุ่มลูกค้าที่ได้รับอนุมัติและไม่ได้รับอนุมัติ ข้อมูลนี้ชี้ให้เห็นถึงกลไกการตัดสินใจที่เข้มแข็ง (Strong Decision Mechanism) แต่ก็แฝงไว้ด้วยข้อจำกัดที่ควรนำไปสู่การปรับปรุงเชิงกลยุทธ์

#### 1. สมมติฐานการแบ่งกลุ่มลูกค้าตามความเสี่ยง (Customer Segmentation)

คำออบ: มีกลุ่มลูกค้าที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันอย่างชัดเจน

กลุ่มความเสี่ยงต่ำ (Low-Risk/Approved Group): ลูกค้าที่ได้รับอนุมัติมี Credit Score เฉลี่ยสูงถึง 4.61 สะท้อนให้เห็นว่ากลุ่มนี้มีประวัติทางการเงินที่ดีและตรงตามเกณฑ์หลักของธนาคารอย่างชัดเจน

กลุ่มความเสี่ยงสูง (High-Risk/Not Approved Group): ลูกค้าที่ไม่ได้รับอนุมัติมี Credit Score เฉลี่ยต่ำมากที่ 0.63 แสดงให้เห็นว่ากลุ่มนี้มีprofileความเสี่ยงสูงมากและถูกปฏิเสธอย่างชัดเจน

ข้อสรุปเชิงธุรกิจ: ปัจจุบัน Credit Score เป็นปัจจัยขับเคลื่อนหลัก (Primary Driver) ในการอนุมัติ ทำให้เกิดการแบ่งกลุ่มลูกค้าเป็นสองขั้วอย่างมีประสิทธิภาพ



## 5. Analysis/Model Development

### บทสรุปเชิงกลยุทธ์: ความเสี่ยงและประสิทธิภาพของ Credit Score

#### 2. ข้อบกพร่องในเกณฑ์การตัดสินใจแบบเดิม

คำตอบ: หากเกณฑ์การตัดสินใจพึ่งพา CreditScore เพียงอย่างเดียวหรือใช้ค่า Cutoff ที่ไม่ยืดหยุ่น อาจเกิดข้อบกพร่องดังต่อไปนี้

##### ข้อบกพร่องที่อาจเกิดขึ้น

การพึ่งพาตัวแปรเดียวมากเกินไป (Over-reliance on CreditScore)

การใช้ค่า Cutoff แบบแข็งตัว

การไม่พิจารณา Interaction Effects

##### นัยยะทางธุรกิจ

อาจละเลยศักยภาพของลูกค้าที่มี CreditScore ปานกลาง (Gray Zone) แต่มีปัจจัยชุดเซย์อิน ๆ ที่แข็งแกร่ง (เช่น Income สูงมาก หรือ YearsEmployed ยาวนาน) ซึ่งอาจถูกมองเป็นลูกค้าที่ดีในอนาคต

การใช้ค่า Cutoff แบบตายตัวอาจทำให้เสียโอกาส (Lost Opportunities)  
ในการอนุมัติลูกค้าที่เพิ่งผ่านเกณฑ์มาเพียงเล็กน้อย หรือปฏิเสธลูกค้าที่อยู่ต่ำกว่าเกณฑ์นิดเดียวแต่มีความเสี่ยงต่ำเมื่อรวมปัจจัยอื่น

เกณฑ์แบบเดิมอาจไม่ได้ให้ความสำคัญกับการทำงานร่วมกันของปัจจัย (เช่น ลูกค้าที่มี CreditScore ต่ำ แต่มีประวัติการจ้างงานมั่นคงยาวนาน) ทำให้การตัดสินใจขาดความแม่นยำในกลุ่มลูกค้าที่ซับซ้อน



## 5. Analysis/Model Development

### บทสรุปเชิงกลยุทธ์: ความเสี่ยงและประสิทธิภาพของ Credit Score

#### บทสรุปและการดำเนินการที่แนะนำ (Executive Summary & Recommendation)

Credit Score เป็นมาตรวัดที่มีประสิทธิภาพสูงและเป็นตัวชี้วัดความเสี่ยงที่เชื่อถือได้ในกระบวนการอนุมัติสินเชื่อปัจจุบัน ซึ่งสะท้อนให้เห็นถึงการแบ่งกลุ่มลูกค้าที่มีความเสี่ยงต่ำและสูงอย่างชัดเจน อย่างไรก็ตาม เพื่อเพิ่มประสิทธิภาพในการอนุมัติและลดความเสี่ยงจากการพลาดโอกาสทางธุรกิจ (Opportunity Loss)

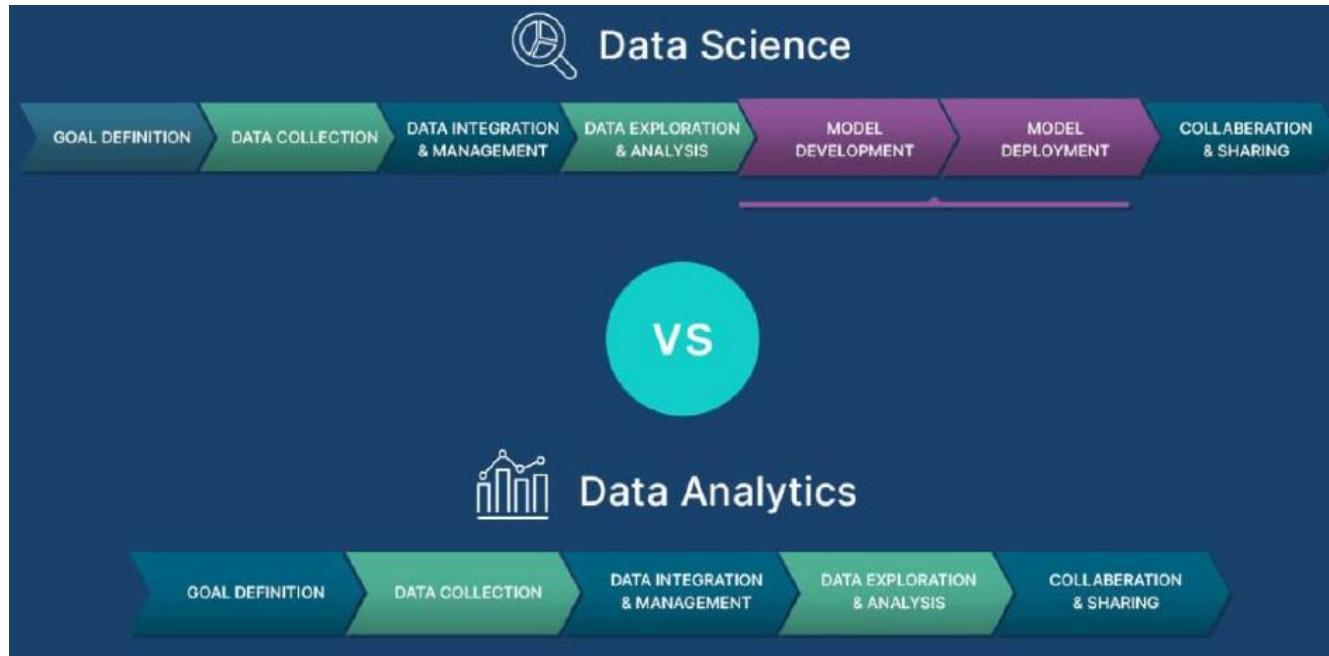
#### การดำเนินการที่แนะนำ:

- วิเคราะห์ค่า Cutoff ใหม่: ทำการศึกษาเชิงลึกเพื่อหา CreditScore ค่า Cutoff ที่เหมาะสมที่สุด (Optimal Threshold) เพื่อให้แน่ใจว่าไม่ได้ปฏิเสธลูกค้าที่มีคุณสมบัติเพียงพอโดยไม่จำเป็น
- สร้างแบบจำลองคะแนนรวม (Composite Scoring Model): พัฒนาระบบการให้คะแนนที่ผนวก CreditScore เข้ากับตัวแปรสำคัญอื่น ๆ เช่น Income YearsEmployed และ Debt โดยมีการถ่วงน้ำหนัก (Weighted Score) เพื่อให้การตัดสินใจมีความแม่นยำและครอบคลุมกลุ่มลูกค้าในเขตสีเทา (Gray Zone) มากขึ้น
- พัฒนากลยุทธ์การอนุมัติ: ใช้ความแตกต่างของ CreditScore นี้เป็นหลักฐานในการปรับปรุงนโยบายสินเชื่อ โดยให้ความยืดหยุ่นมากขึ้นสำหรับลูกค้าที่อาจมี CreditScore ต่ำกว่าเกณฑ์เล็กน้อย แต่มีปัจจัยด้านอื่น ๆ ที่ชดเชยความเสี่ยงได้



## 5. Analysis/Model Development

### Modeling Methodology



### Model Development

- **Algorithm selection:** Classification Algorithms: Logistic Regression
- **Training, hyperparameter tuning, evaluation metric:**  
Training: แบ่งข้อมูลเป็น Training, Validation และ Test Sets  
(Cross-Validation) Tuning: Grid Search เพื่อหา Hyperparameters ที่เหมาะสมที่สุด
- **Evaluation Metric:** Accuracy, Precision, Recall, F1-Score



## 5. Analysis/Model Development

### Workflow– Model Development

```
▶ x = df1.drop('Approved', axis = 1)  
y = df1['Approved']
```

```
print(x.shape)  
x.head()
```

```
(690, 17)
```

|   | Gender | Age   | Debt  | Married | BankCustomer | Industry | Industry_Name | Ethnicity | YearsEmployed | PriorDefault | Employed | CreditScore | DriversLicense | Citizen | ZipCode | StateName | Income |
|---|--------|-------|-------|---------|--------------|----------|---------------|-----------|---------------|--------------|----------|-------------|----------------|---------|---------|-----------|--------|
| 0 | 1      | 30.83 | 0.000 | 1       | 1            | 1        | 7             | 4         | 1.250         | 1            | 1        | 1           | 0              | 0       | 202     | 8         | 0      |
| 1 | 1      | 27.83 | 1.540 | 1       | 1            | 1        | 7             | 4         | 3.750         | 1            | 1        | 5           | 1              | 0       | 100     | 22        | 3      |
| 2 | 0      | 15.83 | 0.585 | 1       | 1            | 7        | 4             | 1         | 1.500         | 1            | 1        | 2           | 0              | 0       | 100     | 22        | 0      |
| 3 | 1      | 23.92 | 0.665 | 1       | 1            | 7        | 4             | 4         | 0.165         | 0            | 0        | 0           | 0              | 0       | 100     | 22        | 0      |
| 4 | 0      | 49.00 | 1.500 | 1       | 1            | 14       | 11            | 3         | 0.000         | 1            | 0        | 0           | 1              | 0       | 100     | 22        | 27     |

```
print(y.shape)  
y.head()
```

```
(690,)
```

```
Approved
```

|   |   |
|---|---|
| 0 | 1 |
| 1 | 1 |
| 2 | 1 |
| 3 | 1 |
| 4 | 0 |

```
dtype: int64
```



## 5. Analysis/Model Development

### Workflow – Model Development

```
# train_test_split จากไลบรารี scikit-learn เพื่อแบ่งข้อมูลทั้งหมด (x คือฟีเจอร์, y คือตัวแปรเป้าหมาย) ออกเป็น 4 ส่วน สำหรับใช้ในการวนการฝึกฝนและประเมินผลโมเดลแม่ข่ายนิ่ง
```

```
from sklearn.model_selection import train_test_split

x_train, x_test, y_train, y_test = train_test_split(x, y, test_size = 0.25, random_state = 0)
```

```
# ทำให้ชุดข้อมูลการฝึก (x_train) และชุดข้อมูลการทดสอบ (x_test) มีมาตรฐานเดียวกัน (Scale)
```

```
from sklearn.preprocessing import StandardScaler
```

```
sc = StandardScaler()
x_train = sc.fit_transform(x_train)
x_test = sc.transform(x_test)
```

```
# Build Machine Learning Model - Model name: Logistic Regression
from sklearn.linear_model import LogisticRegression

log_reg = LogisticRegression(random_state = 0)
log_reg.fit(x_train, y_train)
```

```
LogisticRegression(random_state=0)
```

```
y_pred = log_reg.predict(x_test)
print("Train Score: {:.5f}".format(log_reg.score(x_train, y_train)))
print("Test Score: {:.5f}".format(log_reg.score(x_test, y_test)))
```

```
Train Score: 0.87041
Test Score: 0.86705
```



## 5. Analysis/Model Development

### Workflow– Model Development (Parameter / Hyperparameter Tuning)

Parameter / Hyperparameter Tuning การค้นหาชุดของค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับโมเดล

```
# โน๊ต เคล Logistic Regression กារណົດຄ່າພາຣາມີເຕອຣ C = 0.1 ເປັນຄວາມພກຜົນຂອງຄ່າຄວາມແຮງຂອງການປັບໄທເປັນຮະບັບນິຍົມ (Regularization)
# ກារណົດປະເທດຂອງ Regularization ເປັນ L1 Regularization L1 Regularization ມັກເຮືອກວ່າ Lasso
#ກາຮັດວຽກລົກອະນຸມີນີ້ໃຫ້ໃນການຫາຄ່າທີ່ດີທີ່ສຸດ (Optimization Algorithm)
```

```
log_reg1 = LogisticRegression(C=0.1, penalty='l1', solver='liblinear', random_state = 0)
log_reg1.fit(x_train, y_train)
```

```
LogisticRegression
LogisticRegression(C=0.1, penalty='l1', random_state=0, solver='liblinear')
```

```
y_pred1 = log_reg1.predict(x_test)
print("Train Score: {:.5f}".format(log_reg1.score(x_train, y_train)))
print("Test Score: {:.5f}".format(log_reg1.score(x_test, y_test)))
```

```
Train Score: 0.85880
Test Score: 0.84971
```



## 5. Analysis/Model Development

### Workflow– Model Development (Parameter / Hyperparameter Tuning with GridsearchCV)

Parameter / Hyperparameter Tuning การค้นหาชุดของค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับโมเดล

```
# กำหนดพารามิเตอร์
paragrid = { 'C': [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000], 'penalty':['l1', 'l2'], 'solver':['liblinear']}

from sklearn.model_selection import GridSearchCV
grid_search = GridSearchCV(LogisticRegression(), param_grid=paragrid, cv=5, scoring='accuracy')

grid_search.fit(x_train, y_train)

# คุณภาพที่ดีที่สุด
print("Best Parameters:", grid_search.best_params_)
# คะแนนความแม่นยำ (Accuracy Score) เกลี้ยงสูงสุด ที่ได้จากการประเมินโมเดลในกระบวนการ Cross-Validation (CV)
print("Best Score:{:.5f}".format(grid_search.best_score_))

# คะแนนความแม่นยำ (Accuracy Score) ที่ได้จากการประเมินโมเดลสุดท้าย (โมเดลที่ดีที่สุดจากการท่า Grid Search) บนชุดข้อมูล x_train ทั้งหมด
print("Grid Search - Train Score: {:.5f}".format(grid_search.score(x_train, y_train)))
print("Grid Search - Test Score: {:.5f}".format(grid_search.score(x_test, y_test)))
```

Best Parameters: {'C': 10, 'penalty': 'l1', 'solver': 'liblinear'}

Best Score: 0.87018

Grid Search - Train Score: 0.87621

Grid Search - Test Score: 0.87861



## 5. Analysis/Model Development

### Workflow – Model Development (Parameter / Hyperparameter Tuning with Pipeline and GridsearchCV)

Parameter / Hyperparameter Tuning การค้นหาชุดของค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับโมเดล

```
import pandas as pd
from sklearn.model_selection import GridSearchCV
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression

# 1. กำหนดขั้นตอนใน Pipeline
# -----
# ใช้ StandardScaler และ LogisticRegression
steps = [
    ('scaler', StandardScaler()), # ขั้นตอนที่ 1: Data Preprocessing (Standardization)
    ('model', LogisticRegression(random_state=0)) # ขั้นตอนที่ 2: Model (Logistic Regression)
]

pipeline = Pipeline(steps)
# -----

# 2. กำหนด Hyperparameters สำหรับ Grid Search ในรูปแบบ Pipeline
# -----
# กำหนดค่าพารามิเตอร์ใน Pipeline
paragrid = {
    'model__C': [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000],
    'model__penalty':['l1', 'l2'],
    'model__solver':['liblinear']
}
# -----
```

```
# 3. สร้าง Grid Search และทำการฝึกฟัน (Fit)
grid_search = GridSearchCV(
    pipeline, # ใช้ Pipeline
    param_grid=paragrid,
    cv=5,
    scoring='accuracy',
    n_jobs=-1
)

grid_search.fit(x_train, y_train)

# 4. แสดงผลลัพธ์
print("Best Parameters:", grid_search.best_params_)
print(f"Best Score: {grid_search.best_score_:.5f}")

# ตรวจสอบประสิทธิภาพบนชุดข้อมูลจริง
print(f"Grid Search - Train Score: {grid_search.score(x_train, y_train):.5f}")
print(f"Grid Search - Test Score: {grid_search.score(x_test, y_test):.5f}")

...
Best Parameters: {'model__C': 10, 'model__penalty': 'l1', 'model__solver': 'liblinear'}
Best Score: 0.87018
Grid Search - Train Score: 0.87427
Grid Search - Test Score: 0.87861
```



## 5. Analysis/Model Development

### Workflow– Model Development (Parameter / Hyperparameter Tuning with Pipeline and GridsearchCV – CV Experiment - 1)

Parameter / Hyperparameter Tuning การค้นหาชุดของค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับโมเดล

```
import pandas as pd
from sklearn.model_selection import GridSearchCV
from sklearn.pipeline import Pipeline
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LogisticRegression

# 1. กำหนดขั้นตอนใน Pipeline (เหมือนเดิม)
steps = [
    ('scaler', StandardScaler()),
    ('model', LogisticRegression(random_state=0))
]
pipeline = Pipeline(steps)

# 2. กำหนด Hyperparameters สำหรับ Grid Search (เหมือนเดิม)
paragrid = {
    'model__C': [0.0001, 0.001, 0.01, 0.1, 1, 10, 100, 1000],
    'model__penalty':['l1', 'l2'],
    'model__solver':['liblinear']
}

# 3. กำหนดรายการค่า CV ที่ต้องการทดสอบ
cv_values = [2, 3, 4, 5, 6, 7, 8, 9, 10]
```

```
# 4. วนลูปเพื่อรัน GridSearchCV สำหรับแต่ละค่า CV
# -----
best_results = []

for cv_val in cv_values:
    print(f"\n--- Running GridSearchCV with cv = {cv_val} ---")

    # สร้าง Grid Search สำหรับค่า cv_val นั้นๆ
    grid_search = GridSearchCV(
        pipeline,
        param_grid=paragrid,
        cv=cv_val, # เปลี่ยนค่า cv ตามลูป
        scoring='accuracy',
        n_jobs=-1
    )

    # ฝึกฝนและค้นหาพารามิเตอร์ที่ดีที่สุดสำหรับ cv_val นี้
    grid_search.fit(x_train, y_train)

    # เก็บผลลัพธ์ที่ดีที่สุดของรอบนี้
    best_results.append({
        'cv': cv_val,
        'best_params': grid_search.best_params_,
        'best_score': grid_search.best_score_,
        'train_score': grid_search.score(x_train, y_train),
        'test_score': grid_search.score(x_test, y_test)
    })

# แสดงผลลัพธ์ย่อย
print(f"Best Parameters (cv={cv_val}):", grid_search.best_params_)
print(f"Best Score (cv={cv_val}): {grid_search.best_score_:.5f}")
```

```
# 5. แสดงผลลัพธ์สรุปทั้งหมด
# -----
print("\n=====")
print("สรุปผลลัพธ์การค้นหา CV ทั้งหมด")
print("=====")

summary_df = pd.DataFrame(best_results)
print(summary_df[['cv', 'best_score', 'best_params', 'test_score']])

# ตัวอย่างการหาค่า cv ที่ให้ Test Score สูงสุด
best_overall = summary_df.loc[summary_df['test_score'].idxmax()]
print("\n*** Overall Best Result (Based on Test Score) ***")
print(f"Optimal CV Value: {best_overall['cv']}")
print(f"Best Hyperparameters: {best_overall['best_params']}")
print(f"Highest Test Score: {best_overall['test_score']:.5f}")
```



## 5. Analysis/Model Development

### Workflow– Model Development (Parameter / Hyperparameter Tuning with Pipeline and GridsearchCV – CV Experiment - 2)

Parameter / Hyperparameter Tuning การค้นหาชุดของค่าพารามิเตอร์ที่เหมาะสมที่สุดสำหรับโมเดล

```
*** --- Running GridSearchCV with cv = 2 ---  
Best Parameters (cv=2): {'model_C': 0.1, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=2): 0.88388  
  
--- Running GridSearchCV with cv = 3 ---  
Best Parameters (cv=3): {'model_C': 0.1, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=3): 0.87420  
  
--- Running GridSearchCV with cv = 4 ---  
Best Parameters (cv=4): {'model_C': 1, 'model_penalty': 'l1', 'model_solver': 'liblinear'}  
Best Score (cv=4): 0.87036  
  
--- Running GridSearchCV with cv = 5 ---  
Best Parameters (cv=5): {'model_C': 10, 'model_penalty': 'l1', 'model_solver': 'liblinear'}  
Best Score (cv=5): 0.87018  
  
--- Running GridSearchCV with cv = 6 ---  
Best Parameters (cv=6): {'model_C': 0.1, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=6): 0.86646  
  
--- Running GridSearchCV with cv = 7 ---  
Best Parameters (cv=7): {'model_C': 0.01, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=7): 0.87026  
  
--- Running GridSearchCV with cv = 8 ---  
Best Parameters (cv=8): {'model_C': 0.01, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=8): 0.86442  
  
--- Running GridSearchCV with cv = 9 ---  
Best Parameters (cv=9): {'model_C': 0.1, 'model_penalty': 'l2', 'model_solver': 'liblinear'}  
Best Score (cv=9): 0.86234  
  
--- Running GridSearchCV with cv = 10 ---  
Best Parameters (cv=10): {'model_C': 1, 'model_penalty': 'l1', 'model_solver': 'liblinear'}  
  
=====  
สรุปผลลัพธ์การค้นหา CV ทั้งหมด  
=====  
best_params \\\n 0 2 0.883878 {'model_C': 0.1, 'model_penalty': 'l2', 'mod...\\n 1 3 0.874199 {'model_C': 0.1, 'model_penalty': 'l2', 'mod...\\n 2 4 0.870364 {'model_C': 1, 'model_penalty': 'l1', 'model...\\n 3 5 0.870183 {'model_C': 10, 'model_penalty': 'l1', 'mode...\\n 4 6 0.866457 {'model_C': 0.1, 'model_penalty': 'l2', 'mod...\\n 5 7 0.870260 {'model_C': 0.01, 'model_penalty': 'l2', 'mo...\\n 6 8 0.864423 {'model_C': 0.01, 'model_penalty': 'l2', 'mo...\\n 7 9 0.862338 {'model_C': 0.1, 'model_penalty': 'l2', 'mod...\\n 8 10 0.864329 {'model_C': 1, 'model_penalty': 'l1', 'model...\\n  
test_score\\n 0 0.855491\\n 1 0.855491\\n 2 0.872832\\n 3 0.878613\\n 4 0.855491\\n 5 0.826590\\n 6 0.826590\\n 7 0.855491\\n 8 0.872832  
*** Overall Best Result (Based on Test Score) ***  
Optimal CV Value: 5  
Best Hyperparameters: {'model_C': 10, 'model_penalty': 'l1', 'model_solver': 'liblinear'}  
Highest Test Score: 0.87861
```



## 5. Analysis/Model Development

### Workflow– Model Evaluation

```
from sklearn.metrics import confusion_matrix, classification_report

# 1. ໂມເດລທີ່ດີ່ສຸດຈາກ Grid Search
best_model = grid_search.best_estimator_

# ໃຊ້ get_params() ເພື່ອດູ Hyperparameters ທັງໝາດຂອງໂມເດລ
print("Hyperparameters ທັງໝາດຂອງໂມເດລທີ່ດີ່ສຸດ:")
print(best_model.get_params())

# 2. ທ່ານຍົກລວມຂັ້ນຂໍ້ມູນທັດສອນ
y_pred = best_model.predict(x_test)

# 3. ສ້າງ Confusion Matrix
cm = confusion_matrix(y_test, y_pred)
print("\nConfusion Matrix:\n", cm)

# 4. ແສດ Classification Report
report = classification_report(y_test, y_pred)
print("\nClassification Report:\n", report)
```

Confusion Matrix:  
[[83 15]  
 [ 6 69]]

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.93      | 0.85   | 0.89     | 98      |
| 1            | 0.82      | 0.92   | 0.87     | 75      |
| accuracy     |           |        | 0.88     | 173     |
| macro avg    | 0.88      | 0.88   | 0.88     | 173     |
| weighted avg | 0.88      | 0.88   | 0.88     | 173     |

Hyperparameters ທັງໝາດຂອງໂມເດລທີ່ດີ່ສຸດ:

```
{'memory': None, 'steps': [('scaler', StandardScaler()), ('model', LogisticRegression(C=10, penalty='l1', random_state=0, solver='liblinear'))], 'transform_input': None, 'verbose': False, 'scaler': StandardScaler(), 'model': LogisticRegression(C=10, penalty='l1', random_state=0, solver='liblinear'), 'scaler__copy': True, 'scaler__with_mean': True, 'scaler__with_std': True, 'model__C': 10, 'model__class_weight': None, 'model__dual': False, 'model__fit_intercept': True, 'model__intercept_scaling': 1, 'model__l1_ratio': None, 'model__max_iter': 100, 'model__multi_class': 'deprecated', 'model__n_jobs': None, 'model__penalty': 'l1', 'model__random_state': 0, 'model__solver': 'liblinear', 'model__tol': 0.0001, 'model__verbose': 0, 'model__warm_start': False}
```



## 5. Analysis/Model Development

### Workflow – Model Evaluation

#### Confusion Matrix และ Classification Report

Confusion Matrix:

```
[[83 15]
 [ 6 69]]
```

Classification Report:

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.93      | 0.85   | 0.89     | 98      |
| 1            | 0.82      | 0.92   | 0.87     | 75      |
| accuracy     |           |        | 0.88     | 173     |
| macro avg    | 0.88      | 0.88   | 0.88     | 173     |
| weighted avg | 0.88      | 0.88   | 0.88     | 173     |

#### ประสิทธิภาพและการทำงานของแบบจำลอง (Model Performance and Discussion)

##### ประสิทธิภาพโดยรวม (Overall Performance)

ผลลัพธ์จากโมเดลการพยากรณ์นี้แสดงให้เห็นว่าโมเดลมีความสามารถในการจำแนกผู้ขอสินเชื่อบัตรเครดิตได้อย่างแม่นยำและสมดุล โดยมี Accuracy รวมอยู่ที่ 88% จาก 173 ตัวอย่างในชุดทดสอบ ซึ่งสูงกว่าเกณฑ์ที่ยอมรับได้สำหรับงานจำแนกประเภท

##### การประเมินความเสี่ยงจาก Confusion Matrix

|                     |     |                                                                                                        |
|---------------------|-----|--------------------------------------------------------------------------------------------------------|
| ผลลัพธ์             | ค่า | ความหมายเชิงธุรกิจ (Business Problem: ลดความเสี่ยง)                                                    |
| True Negative (TN)  | 83  | ไม่เดลగานายถูกว่า 'ปฏิเสธ' และค่าจริงคือ 'ปฏิเสธ' (ตีมาก: ป้องกันการอนุมัติที่ไม่ควรเกิด)              |
| True Positive (TP)  | 69  | ไม่เดลganayถูกว่า 'อนุมัติ' และค่าจริงคือ 'อนุมัติ' (ตีมาก: เพิ่มโอกาสทางธุรกิจ)                       |
| False Positive (FP) | 15  | ความเสี่ยงทางการเงิน: ไม่เดลganayผิดว่า 'อนุมัติ' ก็ต้องค่าจริงคือ 'ปฏิเสธ' (นำไปสู่ พนักเสียงด้วยตรง) |
| False Negative (FN) | 6   | การสูญเสียโอกาส: ไม่เดลganayผิดว่า 'ปฏิเสธ' ก็ต้องค่าจริงคือ 'อนุมัติ' (ทำให้พลาดลูกค้าดีๆ)            |



## 5. Analysis/Model Development

### Workflow – Model Evaluation

#### Confusion Matrix และ Classification Report

Confusion Matrix:

```
[[83 15]
 [ 6 69]]
```

Classification Report:

|              | precision | recall | f1-score | support |
|--------------|-----------|--------|----------|---------|
| 0            | 0.93      | 0.85   | 0.89     | 98      |
| 1            | 0.82      | 0.92   | 0.87     | 75      |
| accuracy     |           |        | 0.88     | 173     |
| macro avg    | 0.88      | 0.88   | 0.88     | 173     |
| weighted avg | 0.88      | 0.88   | 0.88     | 173     |

#### ประสิทธิภาพแยกตามคลาส (Classification Report Details)

| Metric    | Class 0<br>(ปัจจີເສດ) | Class 1<br>(ອຸນຸມັດີ) | ອົກປາຍເໜຶງອຸງກິດ                                                                                                        |
|-----------|-----------------------|-----------------------|-------------------------------------------------------------------------------------------------------------------------|
| Precision | 0.93                  | 0.82                  | ເມື່ອໂມເດລທໍານາຍ 'ປັບປຸງ' ວຽກແນ່ນຢ່າງສູງມາກ (93%) ແຕ່ເມື່ອທໍານາຍ 'ອຸນຸມັດີ' ວຽກແນ່ນຢ່າດຄວງເໜີຂອງ 82% (ສະກັນຄໍາ FP = 15) |
| Recall    | 0.85                  | 0.92                  | ໂມເດລສາມາດຮະບຸລູກຄ້າທີ່ຄວາມ 'ອຸນຸມັດີ' ໄດ້ດັ່ງ 92% (FN ຕໍ່: 6) ແຕ່ກາຣະບຸລູກຄ້າທີ່ຄວາມ 'ປັບປຸງ' ອູ້ກໍ່ 85% (FP/FN: 15/6) |
| F1-Score  | 0.89                  | 0.87                  | ຄ່າສົມຄລກທີ່ດີ ແຕ່ປະສົກທີ່ກຳໄຊໃນການທໍານາຍຄລາສ 'ປັບປຸງ' (Class 0) ຍັງຄົງສູງກວ່າເລີກນັ້ນຍັງ                               |

### บทสรุปของໂມເດລ

ໂມເດລນີ້ມີປະສົກທີ່ແລະສາມາດຖັນໄປ Deploy ເພື່ອໃຊ້ໃນການຕັດສິນໃຈອຸນຸມັດີເບື້ອງດັນໂດຍອັດໂນມັດີໄດ້ຕາມ Value Proposition ຂອງໂຄຮງການ ອ່າງໄຮກຕາມ ຄວາມເນັ້ນການປັບປຸງ Precision ຂອງ Class 1 ໃຫ້ສູງຂຶ້ນເພື່ອລົດຄວາມເສີ່ງທີ່ເສີ່ຍ (FP) ໃຫ້ສອດຄລອງກັບວັດຖຸປະສົງຄໍທ່າງຮູກກິຈທຳກັນຂອງສານັກເງິນ



## 5. Analysis/Model Development

### Workflow– Model Development (Selected Parameter with Pipeline and GridsearchCV )

| <pre>print(x.shape) x.head()</pre>                                                                                                                                                                                                                                                                                                                                                                                                                                                             | <pre>print(y.shape) y.head()</pre> | Confusion Matrix:<br>[[78 20]<br>[ 5 70]] |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------|-------------------------------------------|--------------|-------------|-------------|---|-------|---|---|---|---|-------|---|---|---|---|-------|---|---|---|---|-------|---|---|---|---|-------|---|---|---|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|----------|---|---|---|---|---|---|---|---|---|---|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--|-----------|--------|----------|---------|---|------|------|------|----|---|------|------|------|----|----------|--|--|------|-----|-----------|------|------|------|-----|--------------|------|------|------|-----|
| <pre>(690, 4)</pre> <table><thead><tr><th></th><th>YearsEmployed</th><th>PriorDefault</th><th>Employed</th><th>CreditScore</th></tr></thead><tbody><tr><td>0</td><td>1.250</td><td>1</td><td>1</td><td>1</td></tr><tr><td>1</td><td>3.750</td><td>1</td><td>1</td><td>5</td></tr><tr><td>2</td><td>1.500</td><td>1</td><td>1</td><td>2</td></tr><tr><td>3</td><td>0.165</td><td>0</td><td>0</td><td>0</td></tr><tr><td>4</td><td>0.000</td><td>1</td><td>0</td><td>0</td></tr></tbody></table> |                                    | YearsEmployed                             | PriorDefault | Employed    | CreditScore | 0 | 1.250 | 1 | 1 | 1 | 1 | 3.750 | 1 | 1 | 5 | 2 | 1.500 | 1 | 1 | 2 | 3 | 0.165 | 0 | 0 | 0 | 4 | 0.000 | 1 | 0 | 0 | <pre>(690,)</pre> <table><thead><tr><th></th><th>Approved</th></tr></thead><tbody><tr><td>0</td><td>1</td></tr><tr><td>1</td><td>1</td></tr><tr><td>2</td><td>1</td></tr><tr><td>3</td><td>1</td></tr><tr><td>4</td><td>0</td></tr></tbody></table> |  | Approved | 0 | 1 | 1 | 1 | 2 | 1 | 3 | 1 | 4 | 0 | Classification Report:<br><table><thead><tr><th></th><th>precision</th><th>recall</th><th>f1-score</th><th>support</th></tr></thead><tbody><tr><td>0</td><td>0.94</td><td>0.80</td><td>0.86</td><td>98</td></tr><tr><td>1</td><td>0.78</td><td>0.93</td><td>0.85</td><td>75</td></tr><tr><td>accuracy</td><td></td><td></td><td>0.86</td><td>173</td></tr><tr><td>macro avg</td><td>0.86</td><td>0.86</td><td>0.86</td><td>173</td></tr><tr><td>weighted avg</td><td>0.87</td><td>0.86</td><td>0.86</td><td>173</td></tr></tbody></table> |  | precision | recall | f1-score | support | 0 | 0.94 | 0.80 | 0.86 | 98 | 1 | 0.78 | 0.93 | 0.85 | 75 | accuracy |  |  | 0.86 | 173 | macro avg | 0.86 | 0.86 | 0.86 | 173 | weighted avg | 0.87 | 0.86 | 0.86 | 173 |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | YearsEmployed                      | PriorDefault                              | Employed     | CreditScore |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1.250                              | 1                                         | 1            | 1           |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 3.750                              | 1                                         | 1            | 5           |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1.500                              | 1                                         | 1            | 2           |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 0.165                              | 0                                         | 0            | 0           |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 0.000                              | 1                                         | 0            | 0           |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | Approved                           |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1                                  |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1                                  |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 2                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1                                  |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 3                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 1                                  |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 4                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 0                                  |                                           |              |             |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | precision                          | recall                                    | f1-score     | support     |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 0                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 0.94                               | 0.80                                      | 0.86         | 98          |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| 1                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 0.78                               | 0.93                                      | 0.85         | 75          |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| accuracy                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                    |                                           | 0.86         | 173         |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| macro avg                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 0.86                               | 0.86                                      | 0.86         | 173         |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |
| weighted avg                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 0.87                               | 0.86                                      | 0.86         | 173         |             |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |   |       |   |   |   |                                                                                                                                                                                                                                                     |  |          |   |   |   |   |   |   |   |   |   |   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |  |           |        |          |         |   |      |      |      |    |   |      |      |      |    |          |  |  |      |     |           |      |      |      |     |              |      |      |      |     |

Hyperparameters ทั้งหมดของโมเดลที่ดีที่สุด:

```
{'memory': None, 'steps': [('scaler', StandardScaler()), ('model', LogisticRegression(C=0.01, penalty='l1', random_state=0, solver='liblinear'))], 'transform_input': None, 'verbose': False, 'scaler': StandardScaler(), 'model': LogisticRegression(C=0.01, penalty='l1', random_state=0, solver='liblinear'), 'scaler__copy': True, 'scaler__with_mean': True, 'scaler__with_std': True, 'model__C': 0.01, 'model__class_weight': None, 'model__dual': False, 'model__fit_intercept': True, 'model__intercept_scaling': 1, 'model__l1_ratio': None, 'model__max_iter': 100, 'model__multi_class': 'deprecated', 'model__n_jobs': None, 'model__penalty': 'l1', 'model__random_state': 0, 'model__solver': 'liblinear', 'model__tol': 0.0001, 'model__verbose': 0, 'model__warm_start': False}
```



## 5. Analysis/Model Development

### Workflow– Model Development (Selected Parameter with Pipeline and GridsearchCV )

| print(x.shape)<br>x.head() |   | print(y.shape)<br>y.head() |   | Confusion Matrix:<br>[[94  4]<br>[37 38]] |        |          |         |
|----------------------------|---|----------------------------|---|-------------------------------------------|--------|----------|---------|
| (690, 1)                   |   | (690,)                     |   | Classification Report:                    |        |          |         |
| CreditScore                |   | Approved                   |   | precision                                 | recall | f1-score | support |
| 0                          | 1 | 0                          | 1 | 0                                         | 0.72   | 0.96     | 0.82    |
| 1                          | 5 | 1                          | 1 | 1                                         | 0.90   | 0.51     | 0.65    |
| 2                          | 2 | 2                          | 1 | accuracy                                  |        | 0.76     | 173     |
| 3                          | 0 | 3                          | 1 | macro avg                                 |        | 0.81     | 0.73    |
| 4                          | 0 | 4                          | 0 | weighted avg                              |        | 0.80     | 0.76    |

Hyperparameters ทั้งหมดของโมเดลที่ดีที่สุด:

```
{'memory': None, 'steps': [('scaler', StandardScaler()), ('model', LogisticRegression(C=0.1, penalty='l1', random_state=0, solver='liblinear'))], 'transform_input': None, 'verbose': False, 'scaler': StandardScaler(), 'model': LogisticRegression(C=0.1, penalty='l1', random_state=0, solver='liblinear'), 'scaler_copy': True, 'scaler_with_mean': True, 'scaler_with_std': True, 'model_C': 0.1, 'model_class_weight': None, 'model_dual': False, 'model_fit_intercept': True, 'model_intercept_scaling': 1, 'model_l1_ratio': None, 'model_max_iter': 100, 'model_multi_class': 'deprecated', 'model_n_jobs': None, 'model_penalty': 'l1', 'model_random_state': 0, 'model_solver': 'liblinear', 'model_tol': 0.0001, 'model_verbose': 0, 'model_warm_start': False}
```



## 6. Findings and Insights

### ข้อสรุปเชิงธุรกิจ (Data Analytics)

- บทสรุปเชิงวิเคราะห์:** ปัจจัยด้าน ประวัติความเสี่ยง (**PriorDefault**) และ ความมั่นคง (**Employed, CreditScore, YearsEmployed**) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต ซึ่งตอบคำถามเชิงวิเคราะห์ที่ว่า "คุณลักษณะใดของผู้สมัครมีความสัมพันธ์กับการอนุมัติมากที่สุด"
- การแบ่งกลุ่มลูกค้าตามความเสี่ยง (Customer Segmentation)** โดยการทดสอบ T-test สามารถตอบคำถามที่ว่า "มีกลุ่มลูกค้าที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันอย่างชัดเจน หรือไม่" ผลลัพธ์คือ
  - กลุ่มความเสี่ยงต่ำ (Low-Risk/Approved Group):** ลูกค้าที่ได้รับอนุมัติมี Credit Score เฉลี่ยสูงถึง 4.61 สะท้อนให้เห็นว่ากลุ่มนี้มีประวัติทางการเงินที่ดีและตรงตามเกณฑ์หลักของธนาคารอย่างชัดเจน
  - กลุ่มความเสี่ยงสูง (High-Risk/Not Approved Group):** ลูกค้าที่ไม่ได้รับการอนุมัติมี Credit Score เฉลี่ยต่ำมากที่ 0.63 แสดงให้เห็นว่ากลุ่มนี้มีประวัติทางการเงินที่ไม่ดีและไม่สามารถชำระหนี้ได้ตามกำหนดเวลา

บทสรุปเชิงวิเคราะห์: **ปัจจัย Credit Score** เป็นปัจจัยขับเคลื่อนหลัก (Primary Driver) ในกระบวนการอนุมัติ ทำให้เกิดการแบ่งกลุ่มลูกค้าเป็นสองขั้วอย่างมีประสิทธิภาพ



## 6. Findings and Insights

### ข้อสรุปเชิงธุรกิจ (Data Science)

#### การเชื่อมโยงกับปัญหาและจุดประสงค์โครงการ

##### 1. การแก้ปัญหาและวัตถุประสงค์ (Problem Solving and Objectives)

- Problem: กระบวนการพิจารณาใช้เวลานานและมีความเสี่ยงจากคลิกพินิจ
- Solution/Value Proposition: โมเดลนี้สามารถให้ผลการตัดสินใจเบื้องต้นที่ รวดเร็ว (Operational Impact) และ เป็นกลาง ด้วย Accuracy 88%, โมเดลสามารถตัดสินใจโดยอัตโนมัติได้อย่างน่าเชื่อถือในกรณีส่วนใหญ่ ลดเวลาในการตัดสินใจจากชั่วโมง/วัน เหลือเพียงวินาที

##### 3. การเพิ่มโอกาสทางธุรกิจ (Business Opportunity)

- โมเดลมี False Negative (FN) = 6 และ Recall สำหรับ Class 1 (อนุมัติ) คือ 0.92
- โมเดลพัฒนาโอกาสที่จะอนุมัติลูกค้าที่มีคุณภาพจริง (FN) เพียงเล็กน้อย (6 ครั้ง) ซึ่ง แสดงว่าโมเดลมีการรักษาลูกค้าที่ดี (Good Customer Acquisition) ได้อย่างมีประสิทธิภาพ

##### 2. การลดความเสี่ยง (Business Problem: Default Rate Reduction)

- วัตถุประสงค์คือ ลดความเสี่ยงทางการเงิน (หนี้เสีย) ซึ่งต้องการให้ False Positive (FP) ต่ำที่สุด
- โมเดลมี FP = 15 และ Precision สำหรับ Class 1 (อนุมัติ) คือ 0.82 หมายความว่าในทุกๆ 100 ครั้งที่โมเดลบอกว่า "อนุมัติ" จะมีประมาณ 18 ครั้ง ที่เป็นการทำนายผิดพลาดและอาจถูกเรียกว่า "เสีย" ซึ่งเป็นจุดที่ควรปรับปรุง
- Action/Recommendation: ควรพิจารณา ปรับ Threshold ของโมเดลให้มีความระมัดระวังมากขึ้น (Bias ไปทาง Class 0/ปฏิเสธ) เพื่อลดค่า FP ลงอีก แม้ว่าจะต้องแลกมาด้วยการเพิ่ม FN เล็กน้อยก็ตาม (Trade-off ระหว่าง Precision และ Recall)

บทสรุปเชิงวิเคราะห์: โมเดลนี้มีประสิทธิภาพดีมากและสามารถนำไป Deploy เพื่อใช้ในการตัดสินใจอนุมัติเบื้องต้นโดยอัตโนมัติได้ตาม Value Proposition ของโครงการ อย่างไรก็ตาม ควรเน้นการปรับปรุง Precision ของ Class 1 ให้สูงขึ้นเพื่อลดความเสี่ยงหนี้เสีย (FP) ให้ลดคล่องกับวัตถุประสงค์ทางธุรกิจหลักของสถาบันการเงิน



## 7. Conclusion and Recommendation/Action and Impact

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <h3>1. Problem Statement/Background</h3> <ul style="list-style-type: none"> <li><b>What do we know about?</b><br/>สถาบันการเงินอนุมัติ/ปฏิเสธการสมัครบัตรเครดิต โดยพิจารณาจากข้อมูลผู้สมัคร โดยกระบวนการพิจารณาอาจใช้ลักษณะและข้อเนื้อ跟 ดูอย่างพิจารณาของนักวิเคราะห์ที่สืบทอด เช่น อาจนำไปสู่ความเสี่ยงหรือการสูญเสียโอกาสทางธุรกิจ</li> <li><b>What problem are you trying to solve?</b><br/>การสร้างแบบจำลองเพื่อทำนายผลการอนุมัติบัตรเครดิตโดยอัตโนมัติ เพื่อเพิ่มความรวดเร็วและความแม่นยำในการตัดสินใจ</li> <li><b>What is the business problem?</b><br/>การลดความเสี่ยงทางการเงิน (อัตราการผิดนัดชำระหนี้ - Default Rate) และเพิ่มประสิทธิภาพในการดำเนินงาน (Operational Efficiency) ของฝ่ายอนุมัติสินเชื่อ</li> <li><b>Who are the stakeholders?</b><br/>ธนาคาร/สถาบันการเงิน (ฝ่ายบริหารความเสี่ยง, ฝ่ายการตลาด, ฝ่ายปฏิบัติการ) ทีม Data Scientist/Analyst ผู้สมัครบัตรเครดิต</li> </ul>                                                                                                                                | <h3>2. Questions/Hypothesis</h3> <ul style="list-style-type: none"> <li><b>Analytical Questions</b> <ul style="list-style-type: none"> <li>- คุณลักษณะใดของผู้สมัคร (เช่น อายุ, รายได้, ประเภทงาน) ที่มีความสัมพันธ์กับการอนุมัติมากที่สุด?</li> <li>- มีกลุ่มลูกค้า (Segments) ที่มีความเสี่ยงสูง/ต่ำ ที่แตกต่างกันขนาดไหนหรือไม่? เกณฑ์การตัดสินใจแบบเดิมมีข้อบกพร่องตรงไหนบ้าง?</li> </ul> </li> <li><b>Predictive Hypothesis (What can we predict?)</b> <ul style="list-style-type: none"> <li>- เราสามารถทำนายผลการอนุมัติบัตรเครดิต (Approved หรือ Rejected) จากข้อมูลของผู้สมัครที่ให้มาด้วยความแม่นยำสูง เพื่อช่วยในการตัดสินใจอัตโนมัติ</li> </ul> </li> <li><b>SMART Objectives</b> <ul style="list-style-type: none"> <li>- เพิ่มยอดสินเชื่อ จากการอนุมัติบัตรเครดิต อย่างน้อย 5% ภายใน 1 ปี ด้วยการปรับกลยุทธ์ด้านสินเชื่อ</li> </ul> </li> </ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | <h3>3. Value Propositions</h3> <ul style="list-style-type: none"> <li><b>What are we trying to do for the end-user(s) of the system?</b><br/>มอบเครื่องมือในการตัดสินใจที่รวดเร็ว เป็นกลาง และสนับสนุนให้เก็บเงินหน้าที่สินเชื่อ ทำให้ผู้สมัครได้รับผลการตัดสินใจที่รวดเร็วขึ้น</li> <li><b>What objectives are we serving?</b><br/>เพิ่มประสิทธิภาพการดำเนินงาน (ลดต้นทุนและเวลา) ปรับปรุงการประเมินความเสี่ยง (ลดหนี้เสีย) ขยายฐานลูกค้าอย่างมีคุณภาพ</li> </ul>                                                                                                                                                                                                                                                                                                                                                     | <h3>4. Data Sources/Attributes</h3> <ul style="list-style-type: none"> <li><b>Data sources &amp; collection</b> <ul style="list-style-type: none"> <li>- Kaggle Dataset: "Credit Card Approvals (Clean Data)" (ชุดข้อมูลมาจากชั้นนำ)</li> <li>- ข้อมูลในเทาบัญชี: ฐานข้อมูลภายในของธนาคาร/สถาบันการเงิน (ข้อมูลการสมัคร, ประวัติสินเชื่อ, ข้อมูลเครดิตติ๊ก)</li> </ul> </li> <li><b>Data cleaning &amp; preprocessing</b> <ul style="list-style-type: none"> <li>- เนื่องจากเป็น "Clean Data" จะไม่ทำการตรวจสอบขั้นสุดท้าย การจัดการกับค่าที่หายไป (Missing Values) ที่ยังคงมีอยู่</li> <li>- การเข้ารหัสข้อมูลประเภท Categorical (เช่น One-Hot Encoding) การปรับขนาดข้อมูลตัวเลข (Normalization/ Standardization)</li> </ul> </li> <li><b>Target variables &amp; feature</b> <ul style="list-style-type: none"> <li>- Target Variable: สถานะการอนุมัติ (Approved: '+', Rejected: '-')</li> <li>- Features: คุณลักษณะของผู้สมัคร เช่น อายุ, รายได้, หนี้สิน, สถานภาพ, ประวัติเครดิต, ประเภทงาน (ตามคุณลักษณะใน Dataset)</li> </ul> </li> <li><b>Encoding &amp; scaling strategies</b> <ul style="list-style-type: none"> <li>- Encoding: ข้อมูลที่ไม่ใช้ตัวเลข (Non-numerical/Categorical Data) ให้กลายเป็นตัวเลข โดยใช้เทคนิคที่เรียกว่า Label Encoding</li> <li>- Scaling: Standard Scaler สำหรับตัวแปรตัวเลข (Numerical)</li> </ul> </li> </ul> |
| <h3>5. Analysis/Model Development</h3> <ul style="list-style-type: none"> <li><b>Analytics Methodology</b> <ul style="list-style-type: none"> <li>- <b>EDA techniques:</b> การวิเคราะห์ความถี่และค่าสถิติเชิงพรรณนา (Univariate/Bivariate Analysis) เพื่อดูความสัมพันธ์ของแต่ละคุณลักษณะกับผลการอนุมัติ</li> <li>- <b>Visualization strategy:</b> Heatmap สำหรับ Correlation Matrix</li> <li>- <b>Segmentation approach:</b> การจัดกลุ่มลูกค้าตามเป้าหมายความเสี่ยง (Risk Profiles)</li> </ul> </li> <li><b>Modeling Methodology</b> <ul style="list-style-type: none"> <li>- <b>Algorithm selection:</b> Classification Algorithms - Logistic Regression</li> <li>- <b>Training, hyperparameter tuning, evaluation metric:</b> <ul style="list-style-type: none"> <li>** Training: แบ่งข้อมูลเป็น Training, และ Test Sets (Cross-Validation)</li> <li>** Tuning: Grid Search เพื่อหา Hyperparameters ที่เหมาะสมที่สุด</li> </ul> </li> </ul> </li> </ul> <p><b>Evaluation Metric:</b> Accuracy, Precision, Recall, F1-Score</p> | <h3>6. Findings and Insights</h3> <ul style="list-style-type: none"> <li><b>Business Insights</b> <ul style="list-style-type: none"> <li>- คุณลักษณะหลักที่ขับเคลื่อนผลการอนุมัติ (Feature Importance)</li> <li>- <b>ค่าตอบแทนลูกค้า</b> ประจำเดือน PriorDefault และ ความมั่นคง (Employed, CreditScore, YearsEmployed) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต</li> <li>- ระดับลูกค้าที่มีความเสี่ยงสูง/ต่ำ ที่ชัดเจนในชุดข้อมูล <ul style="list-style-type: none"> <li>* <b>ค่าตอบแทนลูกค้า</b>: Credit Score เป็นปัจจัยขับเคลื่อนหลัก (Primary Driver) ในกระบวนการอนุมัติ ทำให้เกิดการแบ่งกลุ่มลูกค้าเป็นสองช่วงขึ้นตามมีประสิทธิภาพ</li> </ul> </li> <li>- แนวโน้มและรูปแบบของผู้สมัครที่ถูกอนุมัติ/ปฏิเสธ <ul style="list-style-type: none"> <li>* <b>ค่าตอบแทน</b> ประจำเดือน PriorDefault และ ความมั่นคง (Employed, CreditScore, YearsEmployed) เป็นคุณลักษณะหลักที่ขับเคลื่อนผลการตัดสินใจอนุมัติบัตรเครดิต ซึ่งตอบค่าความเสี่ยงว่า "คุณลักษณะใดของผู้สมัครมีความสัมพันธ์กับการอนุมัติมากที่สุด"</li> </ul> </li> </ul> </li> <li><b>Predictive Results</b> <ul style="list-style-type: none"> <li>- ประสิทธิภาพของแบบจำลองที่ดี (F1-Score = 0.88)</li> <li>รายชื่อคุณลักษณะที่มีความสำคัญสูงสุด 4 อันดับแรกในการทำนายผล ประจำเดือน PriorDefault และ ความมั่นคง (Employed, CreditScore, YearsEmployed)</li> <li>- ผลการทดสอบกับ Test Set ที่ไม่ได้ใช้ในการฝึกฝน (Score = 0.88)</li> </ul> </li> </ul> | <h3>7. Recommendation/Action and Impact</h3> <ul style="list-style-type: none"> <li><b>What should we do with the findings?</b> <ul style="list-style-type: none"> <li>- <b>Action:</b> นำแบบจำลองที่ปรับปุงแล้วไปใช้งาน (Deploy) ในรูปแบบของ API หรือ Service เพื่อให้ระบบของธนาคารสามารถเรียกใช้ในการตัดสินใจอนุมัติเบื้องต้นโดยอัตโนมัติ</li> </ul> </li> <li><b>What are the impacts?</b> <ul style="list-style-type: none"> <li>- <b>Operational Impact:</b> ลดเวลาในการตัดสินใจจากชั่วโมง/วัน เหลือเพียงวินาที</li> <li>- <b>Business Impact:</b> ลดการเกิดหนี้เสีย (ลดความเสี่ยง) เนื่องจากมีความแม่นยำเพิ่มขึ้น และเพิ่มโอกาสในการอนุมัติลูกค้าที่มีคุณภาพเรื้อรัง (เพิ่มกำไร)</li> <li>- <b>Customer Impact:</b> ผู้สมัครได้รับผลการตัดสินใจรวดเร็วขึ้น ทำให้ประสบการณ์การใช้งานดีขึ้น</li> </ul> </li> </ul> |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |



# THANK YOU

[14 มีนาคม 2568]