



รายงาน

เรื่อง การวิเคราะห์การยกเลิกบริการของลูกค้าธนาคาร

จัดทำโดย

นาย ธนากรฤทธิ์	อักษรเจริญสุข	66050172
นาย ธรรมรงค์	ภูเจริญ	66050191
นาย จิรายุทธ์	อยู่อุ่ม	66050606

เสนอ

ผู้ช่วยศาสตราจารย์ ดร.พรพิมล ชัยวุฒิคั้กตี

รายงานนี้เป็นส่วนหนึ่งของการศึกษาวิชา 0546164

วิชา การวิเคราะห์ข้อมูลโดยใช้ตัวแบบตารางจัดการ

ภาคเรียนที่ 1 ปีการศึกษา 2568

คณะวิทยาศาสตร์ สาขาวิชาสถิติประยุกต์และการวิเคราะห์ข้อมูล

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

คำนำ

รายงานฉบับนี้ เป็นส่วนหนึ่งของรายวิชา Data Analysis Using Spreadsheet Models มีวัตถุประสงค์เพื่อศึกษาการนำโปรแกรม Microsoft Excel มาประยุกต์ใช้ในการวิเคราะห์ข้อมูลเชิงธุรกิจ โดยใช้ข้อมูลตัวอย่างของลูกค้าธนาคาร (Bank Customer Churn Dataset) เพื่อวิเคราะห์พฤติกรรมและปัจจัยที่มีผลต่อการยกเลิกการใช้บริการของลูกค้าภายในรายงานประกอบด้วยขั้นตอนการเตรียมและทำความสะอาดข้อมูล (Data Preparation) การวิเคราะห์เชิงสำรวจ (Exploratory Data Analysis) การสรุปผลด้วย function ต่างๆ Pivot Table, Chart, และการสร้าง Calculated Field (Measure) เพื่อคำนวณอัตราการยกเลิกบริการ (Churn Rate) รวมถึงการนำเสนอผลลัพธ์ผ่านกราฟและตารางสรุปที่เข้าใจง่าย

ผู้จัดทำหวังว่ารายงานฉบับนี้จะเป็นประโยชน์ต่อผู้ที่สนใจในการวิเคราะห์ข้อมูลด้วยสเปรดชีตและสามารถนำไปใช้กับข้อมูลในบริบทธุรกิจจริงได้ ทั้งนี้ ขอขอบคุณอาจารย์ผู้สอนที่ได้ให้ความรู้ คำแนะนำและแนวทางในการทำรายงานฉบับนี้จนสำเร็จลุล่วง

คณะผู้จัดทำ

6 พฤศจิกายน พ.ศ. 2568

บทคัดย่อ (Abstract)

โครงการนี้มีวัตถุประสงค์เพื่อศึกษาปัจจัยที่มีผลต่อการยกเลิกการใช้บริการของลูกค้าธนาคาร (Bank Customer Churn) โดยใช้ข้อมูลลูกค้าซึ่งประกอบด้วยตัวแปร เช่น อายุ เพศ รายได้ ระยะเวลาการเป็นสมาชิก จำนวนผลิตภัณฑ์ที่ถือครอง และสถานะการยกเลิกบริการ (Churn = Yes/No) มาทำการวิเคราะห์เชิงสถิติและเชิงข้อมูล (Data Analytics) ด้วยโปรแกรม Excel ผ่านกระบวนการ Exploratory Data Analysis (EDA), PivotTable, และ การสร้างสูตรเชิงตรรกะ (IF, COUNTIFS, AVERAGEIFS) เพื่อหาความสัมพันธ์ระหว่างตัวแปรต่าง ๆ กับพฤติกรรมการยกเลิกบริการผลการวิเคราะห์เบื้องต้นพบว่า ลูกค้าที่มีระยะเวลาเป็นสมาชิกสั้น ใช้บริการผลิตภัณฑ์น้อยและมีรายได้ต่ำ มีแนวโน้มยกเลิกบริการสูงกว่ากลุ่มอื่น ๆ ข้อมูลเชิงลึกเหล่านี้สามารถนำไปใช้ในการวางแผนกลยุทธ์ รักษาฐานลูกค้า (Customer Retention) และเพิ่มความภักดีของลูกค้าในระยะยาวได้

สารบัญ

เรื่อง คำนำ	2
บทคัดย่อ (Abstract)	3
สารบัญ	4
บทที่ 1 บทนำ (Introduction)	5
บทที่ 2 เอกสารที่เกี่ยวข้อง	6
2.1 ข้อมูลและวิธีการวิเคราะห์ (Data and Methodology)	6
2.2 เครื่องมือและเทคนิคที่ใช้ในการวิเคราะห์	7
2.3 สรุป	8
บทที่ 3 วิธีการดำเนินการ	9
3.1 การดึงข้อมูล (Data Extraction)	9
3.2 การจัดการข้อมูล (Data Handling)	11
3.3 การสร้างสูตร measure บน pivot table	12
3.4 การวิเคราะห์ข้อมูล EDA	13
3.5 สอนทดสอบด้วย t-test assuming equal variance ตัวแปร exited กับ tenure	14
3.6 ดูความสัมพันธ์ระหว่างอัตราการยกเลิกบริการกับคะแนนความน่าเชื่อถือด้วยฟังก์ชัน correl และสร้าง scatter plot	14
3.7 สร้างpivot table และ conditional formatting เพื่อหา insight จากอัตราการยกเลิกบริการของลูกค้า	15
3.8 สร้างกราฟจาก pivot table และนำไปทำ dashboard	16
3.9 Dashboard	17
3.10 สร้างโมเดล logistic regressionc แบบ imbalance class	17
3.11 สร้างโมเดล logistic regression แบบ undersampling	16
บทที่ 4 ผลการดำเนินงาน	18
บทที่ 5 สรุปผลการวิเคราะห์และข้อเสนอแนะ	22
บรรณานุกรม	25

บทที่ 1

บทนำ (Introduction)

ความสำคัญของการวิเคราะห์ข้อมูลการยกเลิกบริการของธนาคาร

ในยุคของการแข่งขันด้านการเงินและธนาคาร การรักษาฐานลูกค้าเดิมมีความสำคัญไม่น้อยไปกว่าการหาลูกค้าใหม่ การยกเลิกบริการของลูกค้า (Churn) มักส่งผลต่อรายได้โดยตรงขององค์กร การวิเคราะห์ข้อมูลพัฒนาระบบลูกค้าช่วยให้ธนาคารเข้าใจสาเหตุของการเลิกใช้บริการ สามารถคาดการณ์แนวโน้ม และกำหนดกลยุทธ์การตลาดเพื่อป้องกันการสูญเสียลูกค้าได้อย่างมีประสิทธิภาพ ดังนั้น การนำเทคนิคการวิเคราะห์ข้อมูลมาใช้ เช่น การสร้างตารางสรุปข้อมูล (PivotTable), การใช้สูตรทางสถิติ และการหาความสัมพันธ์ของตัวแปร จึงเป็นเครื่องมือสำคัญในการวิเคราะห์ปัญหาเชิงธุรกิจขององค์กร

วัตถุประสงค์ของรายงาน

- เพื่อวิเคราะห์ปัจจัยที่มีผลต่อการยกเลิกบริการของลูกค้าธนาคาร
- เพื่อหาความสัมพันธ์ระหว่างตัวแปรเชิงประชากรและพฤติกรรมการใช้บริการกับสถานะการยกเลิกบริการ
- เพื่อสร้างข้อมูลเชิงลึก (Insight) ที่สามารถใช้สนับสนุนการตัดสินใจด้านกลยุทธ์การตลาดและการบริการลูกค้า

ประโยชน์ที่ได้รับ

- เพื่อให้ธนาคารสามารถเข้าใจสาเหตุของการเลิกใช้บริการจากพฤติกรรมและข้อมูลประชากรของลูกค้า
- เพื่อคาดการณ์แนวโน้มการสูญเสียลูกค้าและระบุลูกค้าที่มีความเสี่ยงสูงต่อการยกเลิกบริการ
- เพื่อใช้เป็นเครื่องมือในการกำหนดกลยุทธ์ด้านการตลาด เช่น การออกแบบบริการใหม่ การทำโปรโมชันเฉพาะกลุ่มและการปรับปรุงคุณภาพการให้บริการ
- เพื่อเพิ่มประสิทธิภาพในการรักษาฐานลูกค้าเดิม ลดต้นทุนในการหาลูกค้าใหม่ และเพิ่มความสามารถในการแข่งขันขององค์กร

บทที่ 2

เอกสารที่เกี่ยวข้อง

2.1 ข้อมูลและวิธีการวิเคราะห์ (Data and Methodology)

ชื่อข้อมูล การวิเคราะห์การยกเลิกบริการของลูกค้าธนาคาร (Churn for Bank Customers) แหล่งข้อมูล (<https://www.kaggle.com/datasets/mathchi/churn-for-bank-customers>)

ข้อมูลที่ใช้ในการวิเคราะห์ได้มาจากชุดข้อมูลลูกค้าธนาคาร (Bank Customer Churn Dataset) ซึ่งประกอบด้วยข้อมูลลูกค้าจำนวนหลายพันรายการ เก็บในรูปแบบไฟล์ .xlsx โดยมีรายละเอียดตัวแปรที่สำคัญได้แก่ รหัสลูกค้า อายุ เพศ รายได้ ระยะเวลาการเป็นสมาชิก จำนวนผลิตภัณฑ์ที่ถือครอง คะแนนเครดิต สถานะการเป็นสมาชิกต่างประเทศ และสถานะการยกเลิกบริการ (Churn = Yes/No) ข้อมูลนี้สะท้อนพฤติกรรมของลูกค้าธนาคารในแต่ละช่วงเวลา และใช้เป็นพื้นฐานในการวิเคราะห์แนวโน้มการคงอยู่ของลูกค้า

ข้อมูลทั่วไปของลูกค้า (Customer Information)

ตัวแปรที่ใช้ในการวิเคราะห์

- **CustomerId** — รหัสประจำตัวของลูกค้า (เป็นค่าแบบสุ่ม ไม่มีผลต่อการตัดสินใจยกเลิกบริการ)
- **Surname** — นามสกุลของลูกค้า (ไม่มีผลต่อการยกเลิกบริการของลูกค้า)
- **Geography** — ภูมิภาคหรือประเทศที่ลูกค้าอาศัยอยู่ เช่น France, Germany, Spain ซึ่งอาจมีผลต่อแนวโน้มการยกเลิกบริการ
- **Gender** — เพศของลูกค้า (Male/Female) เพื่อสำรวจว่ามีความแตกต่างระหว่างเพศหรือไม่

ข้อมูลทางการเงิน (Financial Information)

- **CreditScore** — คะแนนเครดิตของลูกค้า ยิ่งคะแนนสูงยิ่งมีแนวโน้มคงอยู่กับธนาคาร
- **Balance** — ยอดเงินคงเหลือในบัญชีของลูกค้า ลูกค้าที่มีเงินคงเหลือสูงมักมีแนวโน้มยกเลิกน้อยกว่า
- **EstimatedSalary** — รายได้โดยประมาณของลูกค้า รายได้ที่ต่ำอาจสัมพันธ์กับการยกเลิกบริการ

ข้อมูลด้านเวลาและความสัมพันธ์กับธนาคาร (Customer Relationship)

- **Age** — อายุของลูกค้า โดยที่ว่าไปลูกค้าที่อายุน้อยมักยกเลิกบริการมากกว่ากลุ่มอายุสูง
- **Tenure** — ระยะเวลาการเป็นลูกค้าธนาคาร (หน่วยเป็นปี) ลูกค้าที่อยู่นานมักมีความภักดีต่อธนาคารมากกว่า

ข้อมูลพฤติกรรมการใช้บริการ (Service Usage Information)

- **NumOfProducts** — จำนวนผลิตภัณฑ์ทางการเงินที่ลูกค้าใช้ เช่น บัญชี บัตรเครดิต สินเชื่อ ฯลฯ
- **HasCrCard** — การถือบัตรเครดิต (1 = มี, 0 = ไม่มี) ลูกค้าที่มีบัตรเครดิตมักมีความสัมพันธ์กับธนาคารมากกว่า
- **IsActiveMember** — สถานะการใช้งานบัญชี (1 = Active, 0 = Inactive) ลูกค้าที่ใช้งานบ่อยมีแนวโน้มคงอยู่กับธนาคาร

ข้อมูลความพึงพอใจและประสบการณ์ลูกค้า (Customer Experience)

- **Complain** — ระบุว่าลูกค้ามีการร้องเรียนหรือไม่ (1 = มี, 0 = ไม่มี)
- **Satisfaction Score** — คะแนนความพึงพอใจของลูกค้า โดยคะแนนต่ำอาจสัมพันธ์กับอัตราการยกเลิกที่สูงขึ้น
- **Card Type** — ประเภทของบัตรที่ลูกค้าใช้ เช่น Silver, Gold, Platinum, Diamond
- **Points Earned** — คะแนนสะสมจากการใช้บัตรเครดิต ยิ่งคะแนนมากอาจแสดงถึงการใช้บริการอย่างต่อเนื่อง

ตัวแปรผลลัพธ์ (Target Variable)

- **Exited** — ระบุสถานะของลูกค้าว่ายังใช้บริการธนาคารอยู่หรือไม่ (1 = ยกเลิกการใช้บริการธนาคาร 0 = ยังคงใช้บริการอยู่)

2.2 เครื่องมือและเทคนิคที่ใช้ในการวิเคราะห์

การวิเคราะห์ชุดข้อมูลนี้ใช้โปรแกรม Microsoft Excel เป็นหลัก โดยใช้เครื่องมือและเทคนิคสำคัญดังนี้:

- Pivot Table: ใช้สรุปจำนวนลูกค้าที่ Churn ตามตัวแปรต่าง ๆ เช่น เพศ อายุ จำนวนผลิตภัณฑ์
- กราฟ (Bar, Pie, Line Chart): สร้างกราฟเปรียบเทียบกลุ่มลูกค้าตามเงื่อนไขที่สนใจ

- พังก์ชัน COUNTIFS: ใช้ในการนับจำนวนลูกค้าตามเงื่อนไขที่กำหนด
- พังก์ชัน AVERAGEIFS: หาค่าเฉลี่ยรายได้หรือคะแนนเครดิตของลูกค้าที่ Churn
- พังก์ชัน IF, PERCENTAGE, RANK: เพื่อคำนวณอัตราการยกเลิกบริการและอันดับความเสี่ยง
- Conditional Formatting: เน้นสีเซลล์ลูกค้าที่มีแนวโน้มยกเลิกสูง
- Dashboard: นำเสนอบรรบกราฟและตัวเลขสำคัญของการวิเคราะห์ในแผ่นงานเดียว
- Data analysis: ทำ t test ดูค่าเฉลี่ยผู้ที่ยกเลิกบริการว่ามีระยะเวลาการใช้บริการต่างจากผู้ที่ยังใช้บริการอยู่ไหม
- ทดสอบความสัมพันธ์ระหว่างตัวแปรด้วย credit score
- โมเดล logistic regression คำนวณการยกเลิกบริการของลูกค้า

2.3 สรุป

บทนี้ได้นำเสนอแหล่งข้อมูล ตัวแปรที่ใช้ในการวิเคราะห์ เครื่องมือที่นำมาใช้ในการสรุป และตีความแนวโน้มการยกเลิกบริการของลูกค้าธนาคาร เพื่อเป็นข้อมูลพื้นฐานสำคัญในการวิเคราะห์และข้อเสนอแนะแนวทางแก้ไขปัญหาต่อไปในบทถัดไป

บทที่ 3

วิธีการดำเนินการ

3.1 การดึงข้อมูล (Data Extraction)

สำหรับรายงานเรื่อง การวิเคราะห์การยกเลิกการใช้บริการของลูกค้าธนาคาร (Bank Customer Churn Analysis) ได้ใช้ชุดข้อมูลชื่อว่า “Bank Customer Churn Dataset” ซึ่งแหล่งข้อมูลต้นฉบับมาจากเว็บไซต์ Kaggle โดยชุดข้อมูลนี้ประกอบด้วย 10,000 แล้ว และ 17 คอลัมน์ ครอบคลุมข้อมูลลูกค้าธนาคาร ทั้งด้านประชากรศาสตร์ ข้อมูลทางการเงิน พฤติกรรมการใช้บริการ และความพึงพอใจของลูกค้า วัตถุประสงค์ของข้อมูลคือ เพื่อศึกษาปัจจัยที่มีผลต่อการ “ยกเลิกการใช้บริการธนาคาร (Exited = 1)” ของลูกค้าแต่ละราย

ข้อมูลทั่วไปของลูกค้า (Customer Information)

- **CustomerId** — รหัสประจำตัวของลูกค้า (เป็นค่าแบบสุ่ม ไม่มีผลต่อการตัดสินใจยกเลิกบริการ)
- **Surname** — นามสกุลของลูกค้า (ไม่มีผลต่อการยกเลิกบริการของลูกค้า)
- **Geography** — ภูมิภาคหรือประเทศที่ลูกค้าอาศัยอยู่ เช่น France, Germany, Spain ซึ่งอาจมีผลต่อแนวโน้มการยกเลิกบริการ
- **Gender** — เพศของลูกค้า (Male/Female) เพื่อสำรวจว่ามีความแตกต่างระหว่างเพศหรือไม่

ข้อมูลทางการเงิน (Financial Information)

- **CreditScore** — คะแนนเครดิตของลูกค้า ยิ่งคะแนนสูงยิ่งมีแนวโน้มคงอยู่กับธนาคาร
- **Balance** — ยอดเงินคงเหลือในบัญชีของลูกค้า ลูกค้าที่มีเงินคงเหลือสูงมักมีแนวโน้มยกเลิกน้อยกว่า
- **EstimatedSalary** — รายได้โดยประมาณของลูกค้า รายได้ที่ต่ำอาจสัมพันธ์กับการยกเลิกบริการ

ข้อมูลด้านเวลาและความสัมพันธ์กับธนาคาร (Customer Relationship)

- **Age** — อายุของลูกค้า โดยทั่วไปลูกค้าที่อายุน้อยมักยกเลิกบริการมากกว่ากลุ่มอายุสูง

- **Tenure** — ระยะเวลาการเป็นลูกค้าธนาคาร (หน่วยเป็นปี) ลูกค้าที่อยู่นานนานมากมีความภักดีต่อธนาคารมากกว่า

ข้อมูลพฤติกรรมการใช้บริการ (Service Usage Information)

- **NumOfProducts** — จำนวนผลิตภัณฑ์ทางการเงินที่ลูกค้าใช้ เช่น บัญชี บัตรเครดิต สินเชื่อ ฯลฯ
- **HasCrCard** — การถือบัตรเครดิต (1 = มี, 0 = ไม่มี) ลูกค้าที่มีบัตรเครดิตมักมีความสัมพันธ์กับธนาคารมากกว่า
- **IsActiveMember** — สถานะการใช้งานบัญชี (1 = Active, 0 = Inactive) ลูกค้าที่ใช้งานบ่อยมีแนวโน้มคงอยู่กับธนาคาร

ข้อมูลความพึงพอใจและประสบการณ์ลูกค้า (Customer Experience)

- **Complain** — ระบุว่าลูกค้ามีการร้องเรียนหรือไม่ (1 = มี, 0 = ไม่มี)
- **Satisfaction Score** — คะแนนความพึงพอใจของลูกค้า โดยคะแนนต่ำอาจสัมพันธ์กับอัตราการยกเลิกที่สูงขึ้น
- **Card Type** — ประเภทของบัตรที่ลูกค้าใช้ เช่น Silver, Gold, Platinum, Diamond
- **Points Earned** — คะแนนสะสมจากการใช้บัตรเครดิต ยิ่งคะแนนมากอาจแสดงถึงการใช้บริการอย่างต่อเนื่อง

ตัวแปรผลลัพธ์ (Target Variable)

- **Exited** — ระบุสถานะของลูกค้าว่ายังใช้บริการธนาคารอยู่หรือไม่ (1 = ยกเลิกการใช้บริการธนาคาร, 0 = ยังคงใช้บริการอยู่)

3.2 การจัดการข้อมูล (Data Handling)

การตรวจสอบและประเมินข้อมูลถูกดำเนินการเพื่อให้มั่นใจว่าข้อมูลที่ใช้ในการวิเคราะห์มีความถูกต้อง ครบถ้วน และเหมาะสมกับวัตถุประสงค์ของการวิเคราะห์ก่อนเข้าสู่ขั้นตอนการสรุปผลและสร้างกราฟวิเคราะห์ ได้มีการดำเนินการจัดการข้อมูลตามขั้นตอน ดังนี้

1. ตรวจสอบค่าที่เป็นค่าว่างหรือค่าที่หายไป (Null หรือ Missing Values)

มีการตรวจสอบข้อมูลทุก colum ที่เป็นค่าว่าง (Missing) ผ่าน power query

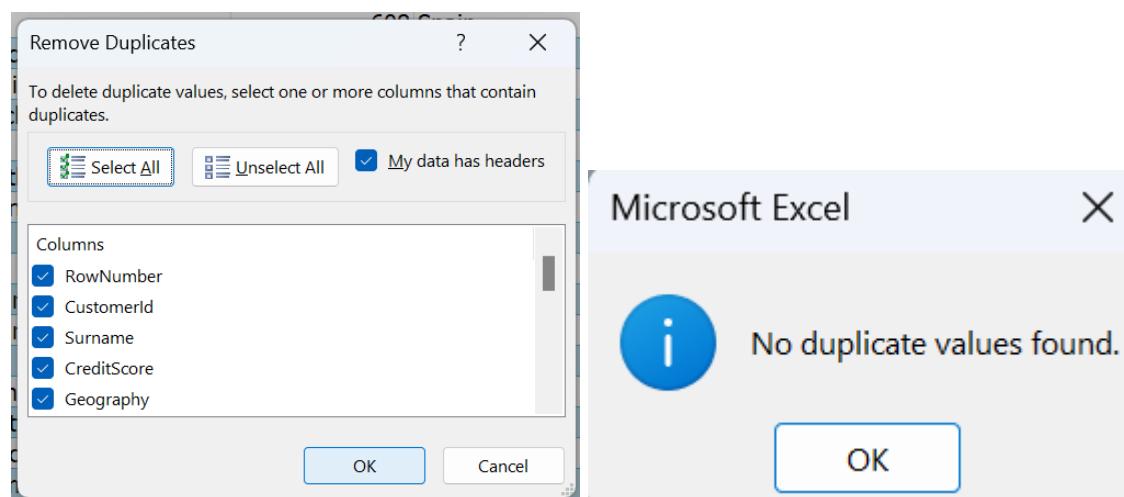
Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)												
CustomerID	100%	0%	0%	Surname	100%	0%	0%	CreditScore	100%	0%	0%	Geography	100%	0%	0%	Gender	100%	0%	0%	Age	100%	0%	0%	Tenure	100%	0%	0%
	Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty
	●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●

Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)												
Balance	100%	0%	0%	NumOfProducts	100%	0%	0%	HasCrCard	100%	0%	0%	IsActiveMember	100%	0%	0%	EstimatedSalary	100%	0%	0%	Exited	100%	0%	0%	Complain	100%	0%	0%
	Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty
	●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●

Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)	Column	Valid (%)	Error (%)	Empty (%)												
Satisfaction Score	100%	0%	0%	Card Type	100%	0%	0%	Point Earned	100%	0%	0%	Age Group	100%	0%	0%	Tenure group	100%	0%	0%	Credit score Group	100%	0%	0%				
	Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty		Valid	Error	Empty
	●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●		●	●	●

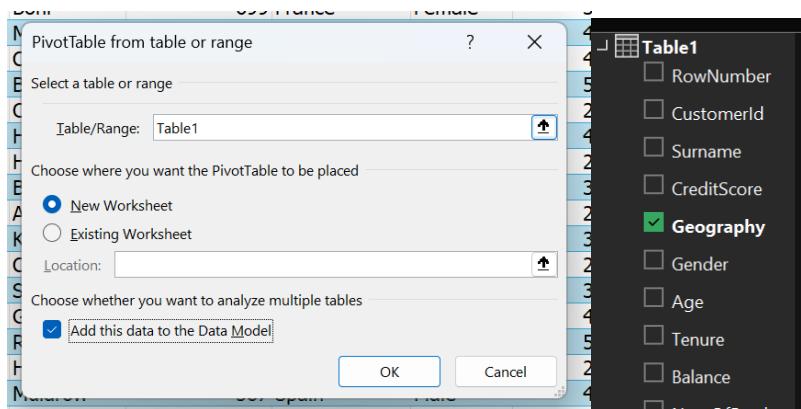
2. ตรวจสอบข้อมูลซ้ำ (Duplicate Records)

ตรวจสอบว่ามีແຕวข้อมูลไม่พบค่าที่ซ้ำกัน



3.3 การสร้างสูตร measure บน pivot table

นำข้อมูลไปสร้าง PivotTable โดยใช้ Insert --> PivotTable และสร้างเป็น table1 กด add this data to data model เพื่อให้สามารถคำนวณขั้นสูง (Measure, Calculated Field) และเชื่อมหลายตารางได้



สร้างสูตรคำนวณ measure บน table1 ให้กับ totalcustomer, churn customer, churn rate, retained customer , retainted customer เพื่อให้วิเคราะห์และสร้างกราฟได้สะดวกขึ้น

Measure Name	Formula
Retained rate	=Divide(Table1[Retained customers],Table1[Total Customers])
Total Customers	=countrows(Table1)
Churn customers	=CALCULATE(countrows(table1),Table1[Exited] = 1)
Churn rate	=DIVIDE([Churn customers], [Total Customers],0)
Retained customers	=CALCULATE(countrows(table1),Table1[Exited] = 0)

3.4 การวิเคราะห์ข้อมูล EDA

ฟังก์ชันที่ใช้ในการวิเคราะห์ได้แก่ countif, countifs, averageif, averageifs และดึงข้อมูลผ่าน table 1

Section 1: ข้อมูลการรวมของลูกค้า				
รายการ	ค่า			
จำนวนลูกค้าทั้งหมด	10000			
จำนวนลูกค้าที่ลาออก	2038			
อัตราการลาออก (%)	20.38%			
Section 2: เพศของลูกค้า				
เพศ	จำนวน	ร้อยละ		
ชาย	5457	54.57%		
หญิง	4543	45.43%		
Section 3: ประเทศของลูกค้า				
ประเทศ	จำนวน	เมืองที่อยู่	อัตรา Churn (%) เทียบกับ ประเทศเดิม	ยอดเงินลงทุนเดิม
France	5014	99899	16.17%	62093
Spain	2477	99441	16.67%	61818
Germany	2509	101113	32.44%	119730
Section 4: ประเภทบัตร (Card Category)				
ประเภทบัตร	จำนวนบัตร	รายได้เฉลี่ย	อัตรา Churn (%) ต่อ... ร. ๙... ๔...	
Silver	2496	101093	20.11 Churn rate	
Gold	2502	100551	19.26 Value: 25.07%	
Platinum	2495	100198	20.36 Category: Female	
Diamond	2507	98526	21.78 Series: Churn rate	
Section 5: คะแนนความพึงพอใจ				
ล้วงไป	ต่อเดือน	การเดินทาง		
คะแนนความพึงพอใจ	3.01	ต่ำกว่าค่าเฉลี่ย — อาจไม่พอใจในบริการ		
รายได้โดยประมาณ	100090	รายได้ต้องซื้อสิ่งของทุกครั้ง		

Section 1: ข้อมูลการรวมของลูกค้า		
รายการ	ค่า	
จำนวนลูกค้าทั้งหมด	=COUNT(Table1[CustomerId])	
จำนวนลูกค้าที่ลาออก	=COUNTIF(Bank_churnN2:N10001,1)	
อัตราการลาออก (%)	=((B4/B3))	
Section 2: เพศ		
เพศ	จำนวน	ร้อยละ
ชาย	=COUNTIFS(Table1[Gender],"Male")	=B9/B3
หญิง	=COUNTIFS(Table1[Gender],"Female")	=B10/B3
Section 3: ประเทศ		
ประเทศ	จำนวน	เมืองที่อยู่
France	=COUNTIF(Table1[Geography],"France")	=AVERAGEIFS(Table1[EstimatedSalary],Table1[Geography],"France")
Spain	=COUNTIF(Table1[Geography],"Spain")	=AVERAGEIFS(Table1[EstimatedSalary],Table1[Geography],"Spain")
Germany	=COUNTIF(Table1[Geography],"Germany")	=AVERAGEIFS(Table1[EstimatedSalary],Table1[Geography],"Germany")
Section 4: ประเภทบัตร		
ประเภทบัตร	จำนวนบัตร	รายได้เฉลี่ย
Silver	=COUNTIFS(Table1[Card Type],"SILVER")	=AVERAGEIF(Table1[Card Type],"SILVER",Table1[EstimatedSalary])
Gold	=COUNTIFS(Table1[Card Type],"GOLD")	=AVERAGEIF(Table1[Card Type],"GOLD",Table1[EstimatedSalary])
Platinum	=COUNTIFS(Table1[Card Type],"PLATINUM")	=AVERAGEIF(Table1[Card Type],"PLATINUM",Table1[EstimatedSalary])
Diamond	=COUNTIFS(Table1[Card Type],"DIAMOND")	=AVERAGEIF(Table1[Card Type],"DIAMOND",Table1[EstimatedSalary])
Section 5: คะแนนความพึงพอใจ		
ล้วงไป	ต่อเดือน	การเดินทาง
คะแนนความพึงพอใจ	=AVERAGE(Table1[Satisfaction Score])	ต่ำกว่าค่าเฉลี่ย — อาจไม่พอใจในบริการ
รายได้โดยประมาณ	=AVERAGE(Table1[EstimatedSalary])	รายได้ต้องซื้อสิ่งของทุกครั้ง

3.5 สถิติทดสอบด้วย t-test assuming equal variance ตัวแปร exited กับ tenure

Add ins datanalysis

ต้องการทดสอบว่า ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการนานาครอช หรือไม่?	
ทดสอบความแปรปรวน 2 กลุ่ม ด้วย F test	
H0: ค่าความเบี่ยงเบนระหว่างเวลาใช้บริการของ ผู้ที่ยกเลิกบริการแล้ว ในแต่ละต่อจาก ผู้ที่ยังคงใช้บริการนานาครอช หรือไม่?	
H1: ค่าความเบี่ยงเบนระหว่างเวลาใช้บริการของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการนานาครอช	
F-Test Two-Sample for Variances	
Mean	Tenure (Exited=1) 4.934739941 Tenure (Exited=0) 5.032780708
Variance	8.624605013 8.297254646
Observations	2038 7962
df	2037 7961
F	1.039452853
P(F<=f) one-tail	0.132852492
P(F<=f) two-tail	0.265704985 >0.05 ยอมรับ H0 Assume variance
F Critical one-tail	1.058864057

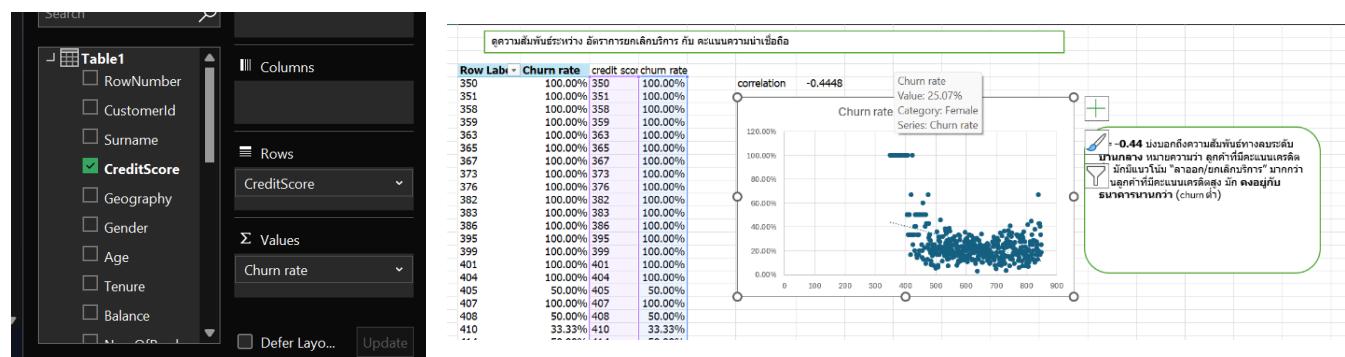
ทดสอบ variance 2 กลุ่ม

ทำการทดสอบ t test two sample แบบ assume varinace	
H0:	ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว ไม่แตกต่างจาก ผู้ที่ยังคงใช้บริการนานาครอช
H1:	ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการนานาครอช
t-Test: Two-Sample Assuming Equal Variances	
Mean	Tenure (Churned) 4.934739941 Tenure (Retained) 5.032780708
Variance	8.624605013 8.297254646
Observations	2038 7962
Pooled Variance	8.363949255
Hypothesized Mean差	0
df	10000
t Stat	-1.365570679
P(T<=t) one-tail	0.086052238
t Critical one-tail	1.645006049
P(T<=t) two-tail	0.172104475 >0.05 ยอมรับ h0
t Critical two-tail	1.960201287

ยอมรับ h0

สรุปผลการทดสอบ ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว ไม่ต่างจาก ผู้ที่ยังคงใช้บริการนานาครอช ทดสอบค่าเฉลี่ยทั้ง 2 กลุ่ม และผลการทดสอบ

3.6 ดูความสัมพันธ์ระหว่างอัตราการยกเลิกบริการกับคะแนนความน่าเชื่อถือด้วยฟังก์ชัน correl และสร้าง scatter plot



3.7 สร้าง pivot table และ conditional formatting เพื่อหา insight จากอัตราการยกเลิกบริการของลูกค้า

ลูกค้า

Rule (applied in order shown)	Format	Applies to	Stop If True
Cell Value < 0.1	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value between 0.1 and 0.1999	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value between 0.2 and 0.2999	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value > 0.3	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>

The screenshot illustrates the process of creating a pivot table and applying conditional formatting to analyze customer churn rates.

Conditional Formatting Rule:

Rule (applied in order shown)	Format	Applies to	Stop If True
Cell Value < 0.1	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value between 0.1 and 0.1999	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value between 0.2 and 0.2999	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>
Cell Value > 0.3	AaBbCcYyZz	= \$I\$7:\$I\$8:\$O\$7:\$O\$11:\$I\$16:\$I\$18:\$O\$17:\$O\$20	<input type="checkbox"/>

PivotTable Fields (Top Left):

- Choose fields to add to report: RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary.
- Drag fields between areas below:
 - Columns: Exited
 - Rows: Card Type
 - Sigma Values: Count of Complain

PivotTable Fields (Top Right):

- Choose fields to add to report: RowNumber, CustomerId, Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary.
- Drag fields between areas below:
 - Columns: Exited
 - Rows: IsActiveMember
 - Sigma Values: Count of IsActiveMember

PivotTable Fields (Bottom Left):

- Choose fields to add to report: RowNumber, CustomerId, Surname, CreditScore, Geography, **Gender**, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary.
- Drag fields between areas below:
 - Columns: Exited
 - Rows: Gender
 - Sigma Values: Count of Gender

PivotTable Fields (Bottom Middle):

- Choose fields to add to report: Age, **Age Group**, Balance, Balance Group, Card Type, Complain, Credit score Group, CreditScore, CustomerId, EstimatedSalary.
- Drag fields between areas below:
 - Columns: Exited
 - Rows: Age Group
 - Sigma Values: Count of Age Group

PivotTable Fields (Bottom Right):

- Choose fields to add to report: RowNumber, CustomerId, Surname, CreditScore, **Geography**, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary.
- Drag fields between areas below:
 - Columns: Exited
 - Rows: Geography
 - Sigma Values: Count of Location

Pivot Table Data:

	Exited	Retained (0)	Churned (1)	Grand Total	Churn rate
3404	1139	4543	25.07%		
4558	899	5457	16.47%		
7962	2038	10000			

G7/H7

ผลลัพธ์:

Geography:

Geography	Exited	Churned	Grand Total	Churn rate
France	3404	1139	4543	25.07%
Germany	4558	899	5457	16.47%
Spain	2064	413	2477	16.67%
Grand Total	7962	2038	10000	

Gender:

Gender	Exited	Churned	Grand Total	Churn rate
Female	3404	1139	4543	25.07%
Male	4558	899	5457	16.47%
Grand Total	7962	2038	10000	

Location:

Location	Exited	Churned	Grand Total	Churn rate
France	4203	811	5014	16.17%
Germany	1695	814	2509	32.44%
Spain	2064	413	2477	16.67%
Grand Total	7962	2038	10000	

card type:

card type	Exited	Churned	Grand Total	Churn rate
DIAMOND	1950	546	2507	21.28%
GOLD	2020	482	2502	19.26%
PLATINUM	1987	508	2495	20.36%
SILVER	1994	502	2496	20.11%
Grand Total	7962	2038	10000	

Age:

Age	Exited	Churned	Grand Total	Churn rate
18-25	565	46	565	25.00%
26-35	3241	301	3542	8.50%
36-45	3002	734	3736	19.65%
46-60	805	842	1647	51.13%
60+	349	115	464	24.78%
Grand Total	7962	2038	10000	

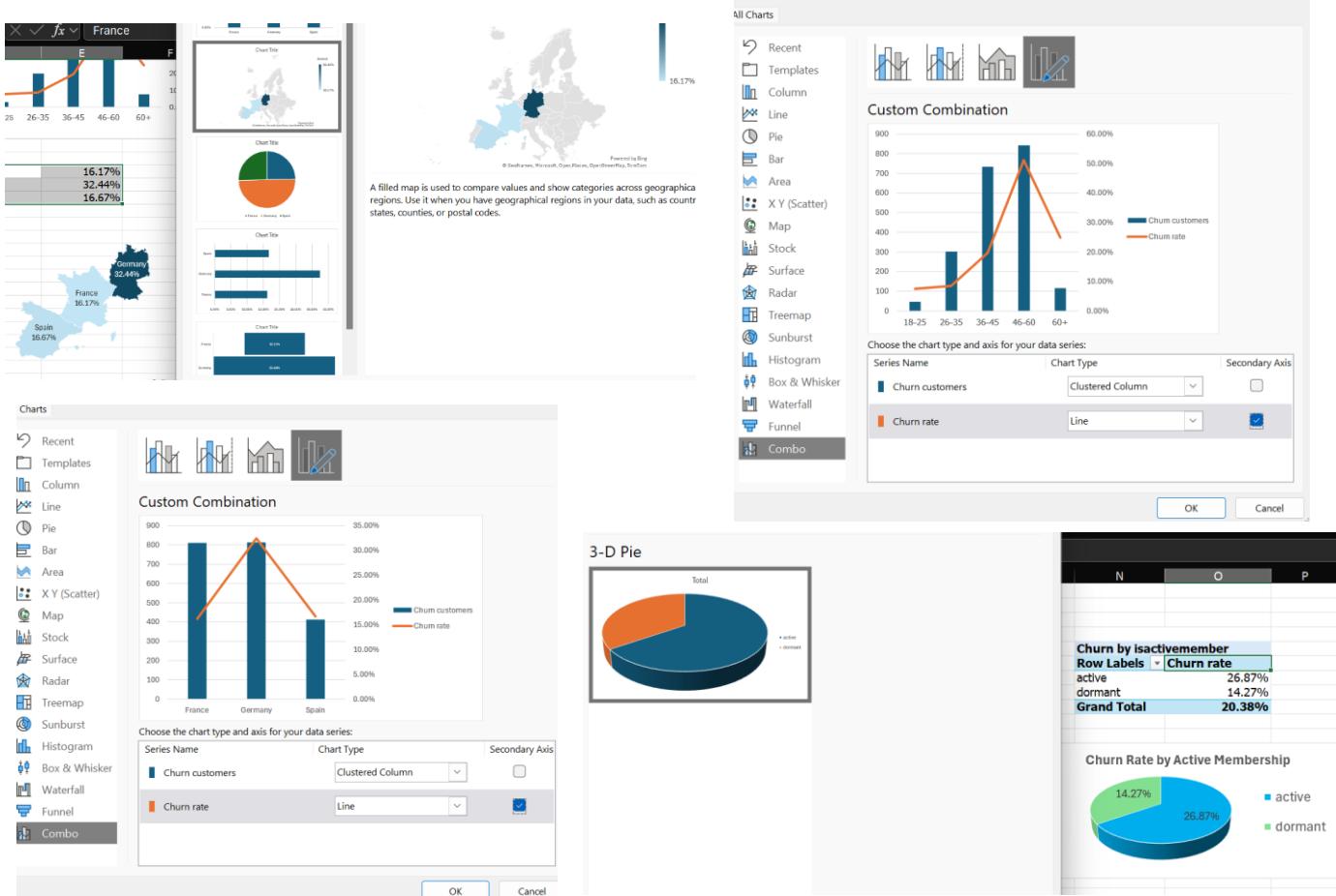
Tenure:

Tenure	Exited	Churned	Grand Total	Churn rate
<1 月	318	95	413	23.00%
>5 月	4412	1094	5506	19.87%
1-3 月	1650	433	2083	20.79%
3-5 月	1582	416	1998	20.82%
Grand Total	7962	2038	10000	

IsActiveMember:

IsActiveMember	Exited	Churned	Grand Total	Churn rate
0	3546	1303	4849	28.07%
1	4416	735	5151	14.27%
Grand Total	7962	2038	10000	

3.8 สร้างกราฟจาก pivot table และนำมายัง dashboard

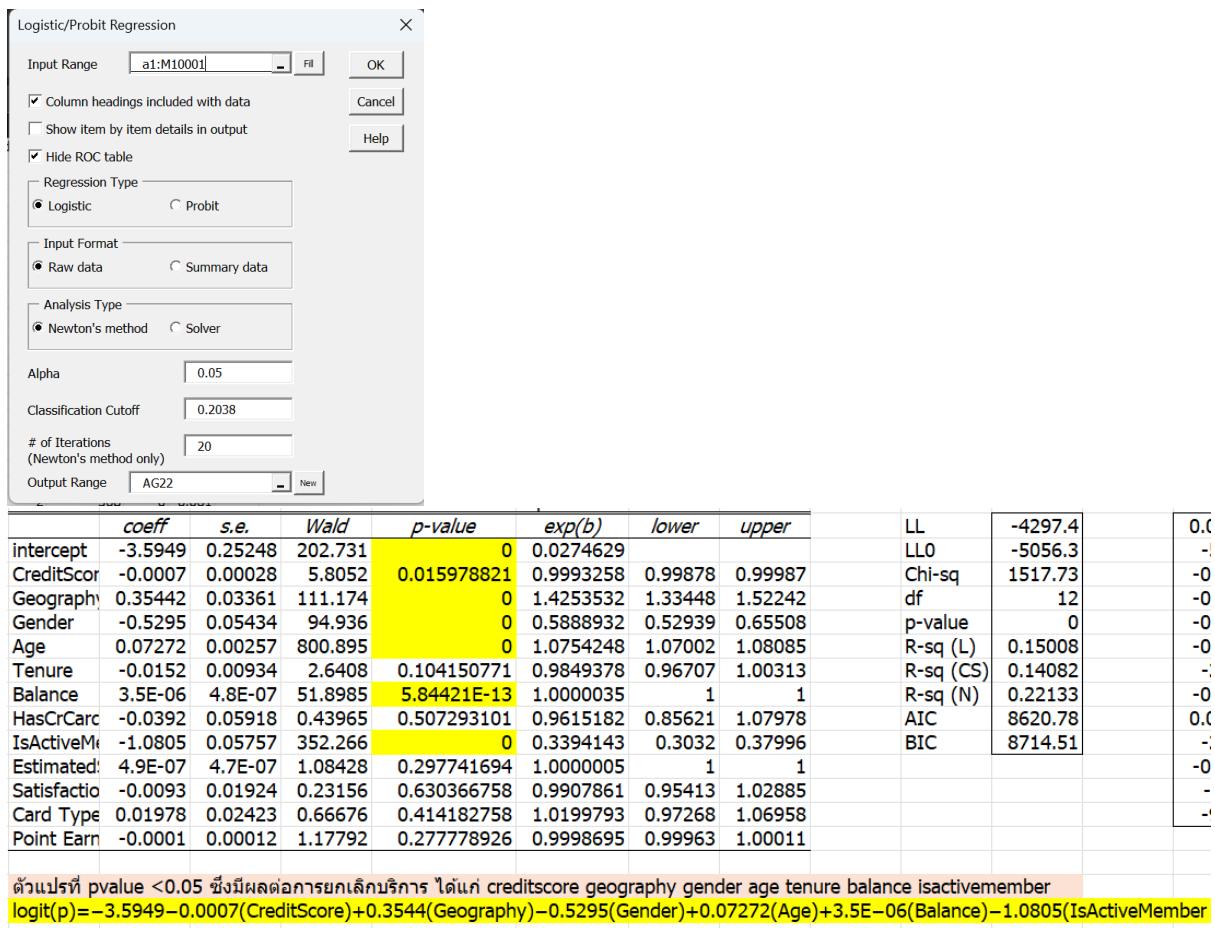


3.9 Dashboard



3.10 สร้างโมเดล logistic regression (imbalance class)

โมเดลเพื่อใช้คำนวณการยกเลิกบริการของลูกค้า(y)จาก ตัวแปรต่างๆ (x)



สมการ logistic regression ที่ได้

$\text{logit}(p) = -3.5949 - 0.0007(\text{CreditScore}) + 0.3544(\text{Geography}) - 0.5295(\text{Gender}) + 0.07272(\text{Age}) + 3.5E - 06(\text{Balance}) - 1.0805(\text{IsActiveMember})$

L	M	N	O	P	Q	มีด
Point	Earned	exited	prob	pred		imbalance class
464	1	0	0.111	0	imbalance class	Lo
456	0	1	0.191	0	retained	7962
377	1	1	0.391	0	churned	2038
350	0	0	0.218	0		int
425	0	0	0.19	0		Cr
484	1	1	0.322	0		Ge
206	0	0	0.089	0		Ge

คำนวณค่า prob จากสมการโมเดล และทำนาย ให้ exited=0 กรณี prob<0.05 (retained) ,

exited=1 กรณี >0.05(Churned)

การประเมินโมเดล พร้อมคำอธิบาย

Classification Table			
	churn	retained	
	Obs Suc	Obs Fail	Total
Pred Suc	437	273	710
Pred Fail	1601	7689	9290
Total	2038	7962	10000
Accuracy	0.21443	0.96571	0.8126
Cutoff	0.5		
AUC	0.76511		

ข้อมูลนี้มีปัญหา Class imbalance ค่อนข้างมาก คือมีลูกค้าที่ยกเลิกบริการจริงเพียง 20% (2,038 จาก 10,000) model bias ไปทางผู้ที่ยังไม่ยกเลิกบริการ ค่อนข้างมาก เราจึงแก้ไขด้วยการ undersampling ในหน้านี้ ดังไป

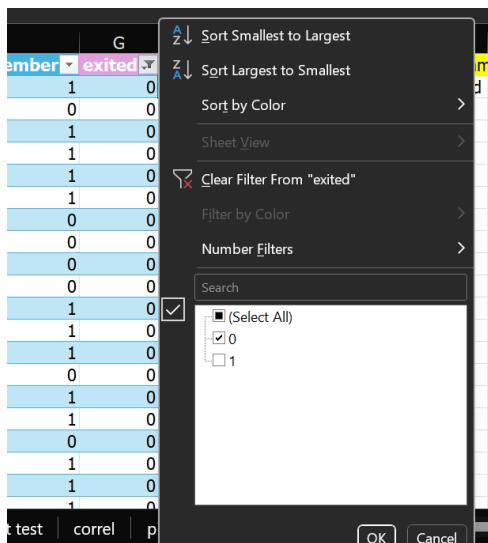
3.11 Logistic Regression (undersampling)

จาก 3.10 พบว่ามี ตัวแปร Creditscore Geography Gender Age Balance isactivemember ที่มีค่า pvalue <0.05 แสดงว่าส่งผลต่อการยกเลิกบริการของลูกค้าธนาคาร จึงใช้ข้อมูลในส่วนนี้ และ ตัวแปรตาม Exited มาทำการ สร้าง Logistic regression แบบ undersampling เพื่อให้ class balance กัน

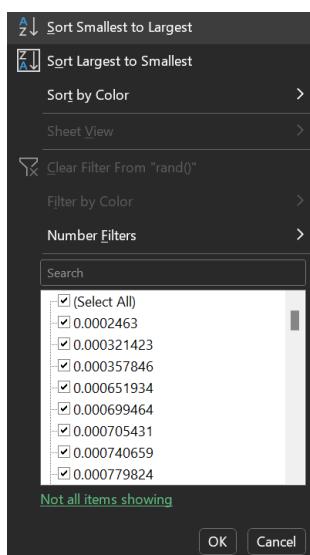
1.สร้างข้อมูล random ด้วยฟังก์ rand() เป็น columnใหม่เขียนมา

A	B	C	D	E	F	G	H
CreditScore	Geography	Gender	Age	Balance	IsActiveMember	Exited	=RAND()
619	0	0	42	0	1	1	0.761223
608	1	0	41	83807.86	1	0	0.043615
502	0	0	42	159660.8	0	1	0.240294
699	0	0	39	0	0	0	0.200757
850	1	0	43	125510.82	1	0	0.807253
645	1	1	44	113755.78	0	1	0.363323
822	0	1	50	0	1	0	0.612023
376	2	0	29	115046.74	0	1	0.914077
501	0	1	44	142051.07	1	0	0.924571
684	0	1	27	134603.88	1	0	0.73544
528	0	1	31	102016.72	0	0	0.057308
497	1	1	24	0	0	0	0.713851

2.filter เลือก classที่ใหญ่กว่า exited=0



3.sortค่า rand()ให้เรียงจากน้อยไปมาก และทำการตัดข้อมูลส่วนนั้น ให้ class ใกล้เคียง exited=1



4. หลังจากตัดแล้วข้อมูลแต่ละ class ใกล้เคียงกันมากขึ้น churned=2038 retained=2061

F	G	H	I	J	K	L	M	N	O	P	Q	R	S
IsActiveMember	Exited	prob	pred										
1	0	0.385747	0										
1	1	0.521427	1										
1	0	0.273506	0										
0	0	0.593788	1										
1	0	0.117993	0										

Under sampling ใน imbalance เมื่อจาก excel ท่า smote หรือ toomeks link ไม่ได้ แบบ python R
Retained 2061 exited =0
churned 2038 exited =1
รวม 4099

5. คำนวณค่า prob และ สร้างเงื่อนไข if prob>0.5 = 1 (churned) prob<0.05=0 (retained)

D	E	F	G	H	I	J	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
=IF([@prob]>0.5,1,0)							=1/(1+EXP(-(\$L\$10+\$L\$11*A3+\$L\$12*B3+														
ge	Balance	IsActiveMember	Exited	prob	pred		ge	Balance	IsActiveMember	Exited	prob	pred									
40	95624.36	1	0	0.385747	0		40	95624.36	1	0	0.385747	0									
50	0	1	1	0.521427	1		50	0	1	1	0.521427	1									
32	131081.66	1	0	0.273506	0		32	131081.66	1	0	0.273506	0									
40	0	0	0	0.593788	1		40	0	0	0	0.593788	1									
25	127728.24	1	0	0.117993	0		25	127728.24	1	0	0.117993	0									
30	117356.19	0	1	0.569247	1		30	117356.19	0	1	0.569247	1									
31	118899.45	0	0	0.362941	0		31	118899.45	0	0	0.362941	0									

สมการโมเดลที่ได้

Logistic Regression		# Iter	20	Alpha	0.05	LL statistics		Covari	
coeff	s.e.	Wald	p-value	exp(b)	lower	upper	LL	-2283.51	0.089
intercept	-3.19558	0.299448	113.8821	0	0.040943		LL0	-2841.15	-8.8
CreditScore	-0.00097	0.000371	6.885439	0.00869	0.999028	0.998303	Chi-sq	1115.273	-0.00
Geography	0.388449	0.044496	76.21198	0	1.474692	1.351532	df	6	-0.00
Gender	-0.63513	0.07232	77.31414	0	0.529868	0.459922	p-value	0	-0.00
Age	0.097637	0.004151	553.2411	0	1.102562	1.093628	R-sq (L)	0.196272	-2.6
Balance	3.11E-06	6.29E-07	24.37995	7.91E-07	1.000003	1.000002	R-sq (CS)	0.23821	-0.00
IsActiveMe	-0.87597	0.073121	143.5119	0	0.416458	0.360853	R-sq (N)	0.317616	
							AIC	4581.019	

สมการโลจิสติก
 $\text{Logit}(P) = -3.2004 - 0.0009(CreditScore) + 0.38937(Geography) - 0.6347(Gender) + 0.09734(Age) + 0.0000031(Balance) - 0.8731(IsActiveMember)$

ผลประเมินเดล imbalance class vs balance class(undersampling)

Classification Table		Churn		retained		Churned		retained		Total	Churned		retained		Churned		retained		Churn		retained	
		Obs Suc	Obs Fail	Total		Obs Suc	Obs Fail	Total		Total	Obs Suc	Obs Fail	Total		Obs Suc	Obs Fail	Total		Obs Suc	Obs Fail	Total	
Pred Suc	437	273	710			1432	555	1987			1432	555	1987		1432	555	1987		1432	555	1987	
Pred Fail	1601	7689	9290			606	1506	2112			606	1506	2112		606	1506	2112		606	1506	2112	
Total	2038	7962	10000			2038	2061	4099			2038	2061	4099		2038	2061	4099		2038	2061	4099	
Accuracy	0.21443	0.96571	0.8126			0.70265	0.730713	0.71676			0.70265	0.730713	0.71676		0.70265	0.730713	0.71676		0.70265	0.730713	0.71676	
Cutoff	0.5					0.5					0.5				0.5				0.5			
AUC	0.76511					0.787856					0.787856				0.787856				0.787856			

ข้อมูลนี้มีปัญหา Class imbalance ค่อนข้างมาก คือมี กลุ่มที่ ยกเลิกบริการเริ่งเพียง 20% (2,038 จาก 10,000) model bias ไปทาง ผู้ที่ยังไม่ยกเลิกบริการ ค่อนข้างมาก เราจึงแก้ไขด้วยการ undersampling ในหน้าตัดไป

ใช้ Random Undersampling ด้วยฟังก์ชัน RAND() เพื่อรับตัวเลขอุ่นหนา (50:50) ทว่าใน accuracy ของกลุ่มยกเลิกบริการ เท่ากับ accuracy ของกลุ่มที่ยังไม่ยกเลิก แม้ accuracy รวมจะลดลงให้ไม่เต็มเป็นเท่าที่ควร

Imbalanceclass

Balanace class (random undersampling)

หลังการทำ Random Undersampling ค่า Recall ของกลุ่มลูกค้าที่ยกเลิกเพิ่มขึ้นจาก 437/2038 = 0.21 (21%) เป็น 1432/2038 = 0.70 (70%) แสดงว่าโมเดลสามารถตรวจจับลูกค้าเสี่ยงได้มากขึ้น แม้ Accuracy โดยรวมจะลดลงเล็กน้อยแต่ผลลัพธ์มีความสมดุลและสะท้อนประสาทที่แท้จริงมากขึ้น.

บทที่ 4

ผลการดำเนินงาน

การวิเคราะห์ข้อมูลและการตีความเพื่อหา insight



-จำนวนข้อมูลลูกค้า 10000 คน

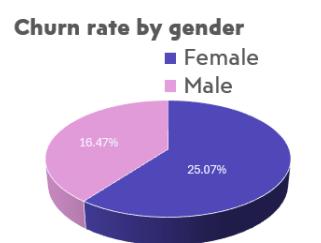
-ยังคงอยู่ 7962 คน

-ยกเลิกบริการแล้ว 2038 คน

-อัตราการยกเลิกบริการ 20.38%

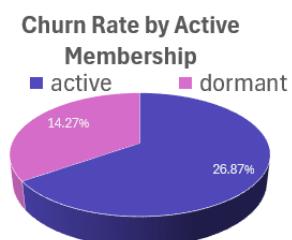
-ระดับความพึงพอใจเฉลี่ย 3.01 เต็ม 5

1. การเปรียบเทียบอัตราการยกเลิกบริการของแต่ละเพศ



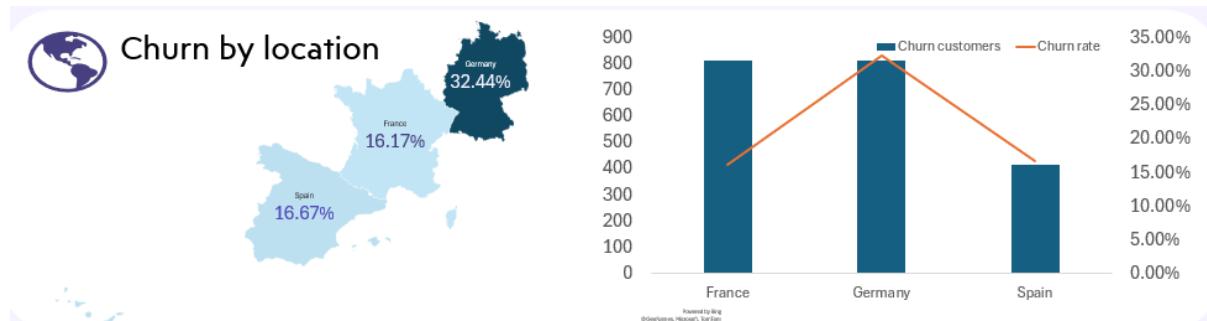
ผู้หญิงมีอัตราการยกเลิกบริการสูงกว่าผู้ชายอย่างชัดเจน และนำไปสู่เพิ่มการดูแลด้านประสบการณ์ลูกค้าและมุ่งมั่นพัฒนาสิทธิประโยชน์ที่ตอบโจทย์ เพื่อช่วยลด Churn ในกลุ่มลูกค้าผู้หญิง

2. การเปรียบเทียบอัตราการยกเลิกบริการของความต่อเนื่องในการใช้งาน



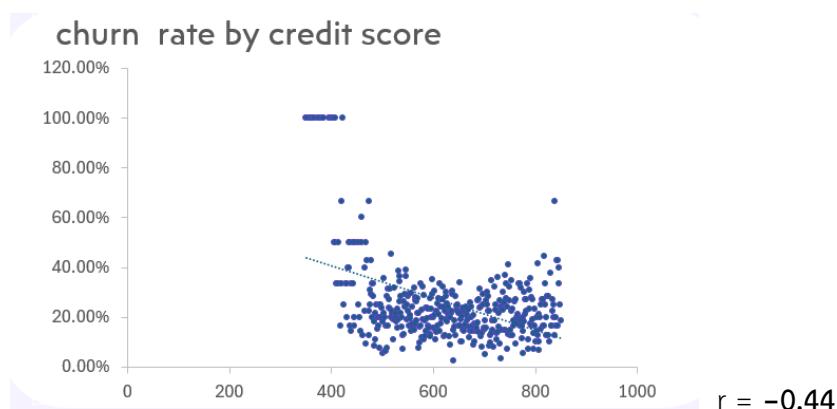
ลูกค้าที่ไม่ค่อยใช้งานมีโอกาสยกเลิกบริการมากกว่าลูกค้าที่ใช้งานสม่ำเสมอเกือบสองเท่า ดังนั้นกลยุทธ์ต้องเน้นกระตุ้นการใช้งาน (Re-engagement) สำหรับกลุ่ม Dormant และสร้างรางวัล/คะแนนสะสมเพื่อรักษากลุ่ม Active

3. การเปรียบเทียบอัตราการยกเลิกบริการของแต่ละประเทศ



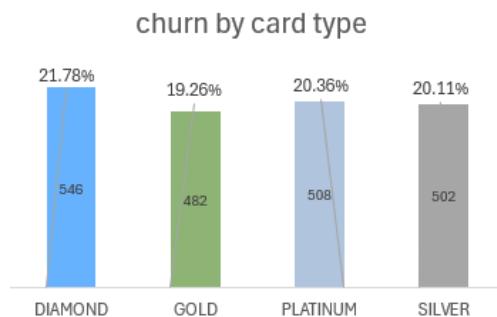
เยอรมนีเป็นตลาดที่มีปัญหา Churn สูงที่สุด จึงต้องวิเคราะห์ประสบการณ์การให้บริการเชิงพื้นที่อย่างเร่งด่วน ขณะที่ฝรั่งเศสเป็นตลาดสำคัญที่ต้องรักษาไว้ เพราะมีฐานลูกค้าขนาดใหญ่ แม้อัตราการลาออกจากต่ำก็ตาม

4. การเปรียบเทียบความสัมพันธ์ของอัตราการลาออกและคะแนนความน่าเชื่อถือ



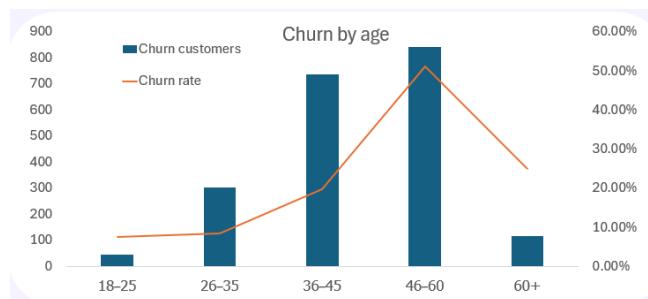
ลูกค้าที่มีคะแนนเครดิตต่ำมีอัตราการยกเลิกสูงกว่ากลุ่มอื่นอย่างเห็นได้ชัด ซึ่งสะท้อนว่าสถานะทางการเงินและต้นทุนที่ลูกค้ารับรู้เป็นปัจจัยสำคัญในการตัดสินใจเลิกใช้บริการ ดังนั้นกลยุทธ์ควรเน้นช่วยลดภาระและการสนับสนุนมากกว่าการขายเพิ่ม

5. การเปรียบเทียบการยกเลิกบริการธนาคารแต่ละประเภทบัตรเครดิต



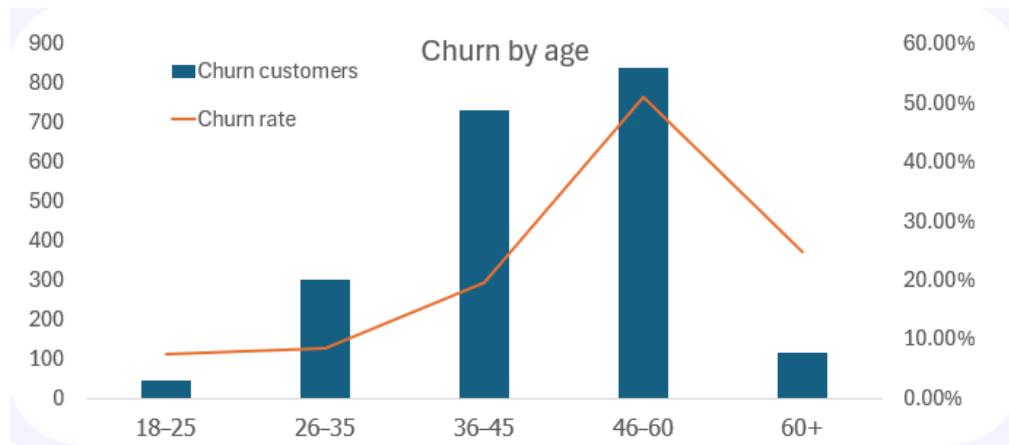
กลุ่มลูกค้าบัตรระดับสูงอย่าง Diamond มีอัตราการยกเลิกสูงที่สุด แม้เป็นกลุ่มที่มีมูลค่าสูง ดังนั้นควรเพิ่มสิทธิประโยชน์และดูแลแบบเฉพาะบุคคล เพื่อป้องกันไม่ให้หายไปใช้บริการของคู่แข่ง

6. การเปรียบเทียบการยกเลิกบริการแต่ละช่วงอายุ



แม้ลูกค้าที่อยู่กับธนาคารมากกว่า 5 ปี จะมี Churn Rate ต่ำที่สุด แต่เป็นกลุ่มที่มีจำนวนลูกค้าที่ลาออกมากที่สุด เพราะเป็นกลุ่มลูกค้าฐานใหญ่ ส่งผลกระทบต่อรายได้รวมอย่างมีนัยสำคัญ จึงเป็นกลุ่มที่ควรให้ความสำคัญที่สุดในการรักษาและลูกค้าที่เป็นสมาชิกน้อยกว่า 1 ปี มี Churn Rate สูงที่สุด แสดงว่าช่วง Onboarding หรือเริ่มต้นใช้งานมีปัญหา หรือลูกค้ายังไม่รู้สึกถึง Value ของบริการ

7. การเปรียบเทียบการยกเลิกบริการแต่ละช่วงอายุ



อัตราการลาออกสูงที่สุดอยู่ในกลุ่มอายุ 46–60 ปี ซึ่งเป็นกลุ่มลูกค้าที่มีความคาดหวังสูงและมีผลต่อรายได้รวมมาก ดังนั้นควรให้ความสำคัญกับการดูแลลูกค้ากลุ่มนี้แบบเฉพาะบุคคล เช่น การให้คำแนะนำด้านการเงินและสิทธิพิเศษเพิ่มเติม

บทที่ 5

สรุปผลการวิเคราะห์และข้อเสนอแนะ

จากการวิเคราะห์ข้อมูลลูกค้าธนาคาร พบปัจจัยที่มีผลต่อการยกเลิกบริการ (Customer Churn) หลายด้าน เช่น เพศ ความถี่ในการใช้งาน ประเทศ คะแนนเครดิต ประเภทบัตร ระยะเวลาที่เป็นสมาชิก และ ช่วงอายุ ซึ่งแต่ละปัจจัยมีแนวโน้มแตกต่างกันและสามารถใช้วางแผนกลยุทธ์เพื่อลดการลาออกจากลูกค้าได้ดังนี้

1. เพศ (Gender)

ผลการวิเคราะห์: ผู้หญิงมีอัตราการยกเลิกบริการสูงกว่าผู้ชายอย่างเห็นได้ชัด
แนวทางแก้ไข:

- เพิ่มช่องทางบริการที่รวดเร็วและเป็นมิตร เช่น แอปพลิเคชัน / สายด่วนเฉพาะกลุ่มลูกค้าผู้หญิง
- จัดแคมเปญสิทธิพิเศษที่ตรงกับไลฟ์สไตล์ เช่น ส่วนลดร้านอาหาร สุขภาพ และความงาม
- ทำคอนเทนต์สั้น ๆ เกี่ยวกับ “วางแผนการเงินง่าย ๆ สำหรับผู้หญิง” เพื่อสร้างความผูกพัน

2. ความถี่ในการใช้งาน (Activity Level)

ผลการวิเคราะห์: ลูกค้าที่ไม่ค่อยใช้งานมีโอกาสลาออกจากเกือบสองเท่าของกลุ่มที่ใช้งานบ่อย
แนวทางแก้ไข:

- ส่งโปรโมชั่นปลุกการใช้งาน เช่น “ใช้บริการครบ 3 ครั้ง/เดือน รับแต้มเพิ่ม”
- จัดระบบสะสมแต้มสำหรับลูกค้าที่ใช้งานต่อเนื่อง
- ส่งแจ้งเตือน (Notification) ให้กลับมาใช้งานหากไม่มีการเคลื่อนไหวเกิน 30 วัน

3. ประเทศ (Country)

ผลการวิเคราะห์: ประเทศเยอรมนีมีอัตราการลาออกจากสูงที่สุด ส่วนปริมาณ Churn ต่อฐานลูกค้าใหญ่:
แนวทางแก้ไข:

- วิเคราะห์ประสบการณ์การใช้งานของลูกค้าในเยอรมนี (เช่น ค่าธรรมเนียม ภาษา หรือการบริการหลังการขาย)

- ปรับปรุงให้เหมาะสมกับพฤติกรรมของแต่ละประเทศ
- สำหรับผู้เชส ควรเน้นรักษาฐานลูกค้าด้วยโปรแกรมสะสมแต้มระยะยาว

4. คะแนนเครดิต (Credit Score)

ผลการวิเคราะห์: พบรความสัมพันธ์เชิงลบ ($r = -0.44$) ระหว่างคะแนนเครดิตกับอัตราการลาออกจาก ลูกค้าที่มีเครดิตต่ำมานี้มีผลลัพธ์มากกว่า

แนวทางแก้ไข:

- จัดโปรแกรม “ช่วยเหลือลูกค้าการเงิน” เช่น ผ่อนจ่าย หรือพักหนี้ระยะสั้น
- ให้คำปรึกษาด้านการวางแผนการเงินฟรีสำหรับกลุ่มเครดิตต่ำ
- ลดการเสนอขายสินเชื่อซ้ำในกลุ่มนี้และเน้นการช่วยให้กลับมาใช้บริการได้บ่อยๆ

5. ประเภทบัตรเครดิต (Card Type)

ผลการวิเคราะห์: บัตร Diamond มีอัตราการลาออกจากสูงที่สุด แม้เป็นกลุ่มลูกค้าระดับพรีเมียม
แนวทางแก้ไข:

- ตั้งทีมดูแลลูกค้า Diamond เฉพาะ (Personal Banker)
- จัดสิทธิพิเศษเฉพาะ เช่น ห้องรับรอง กิจกรรมสุดพิเศษ หรือ Cashback เพิ่ม
- วิเคราะห์สิทธิ์คู่แข่งและปรับสิทธิประโยชน์ให้เหนือกว่า

6. ระยะเวลาการเป็นสมาชิก (Tenure)

ผลการวิเคราะห์:

- ลูกค้าที่อยู่กับธนาคารเกิน 5 ปี มีอัตราออกต่ำแต่ “จำนวนยกเลิก” สูง เพราะฐานใหญ่
- ลูกค้าใหม่ (<1 ปี) มีอัตรายกเลิกสูงสุด เพราะยังไม่เห็นคุณค่าของบริการ

แนวทางแก้ไข:

- ลูกค้าเก่า: ให้ลูกค้าเก่าที่มีการใช้งานต่อเนื่องเลื่อนระดับสถานะ เช่น จาก Silver → Gold → Diamond พร้อมสิทธิพิเศษเพิ่ม เช่น ดอกเบี้ยพิเศษ พรีค่าธรรมเนียมบางรายการ หรือสิทธิ์กดเงิน หรือ รูดบัตรในต่างประเทศโดยไม่เสียค่าใช้จ่าย
- ลูกค้าใหม่: เมื่อเป็นสมาชิกครบทุก 1 ปี จะได้รับคะแนนพิเศษเพิ่ม หรือของขวัญขอบคุณประจำปี (Gift Card ส่วนลดร้านค้า สิทธิพิเศษต่าง ๆ)
- ส่งแบบสอบถามหลัง 1 เดือน เพื่อเก็บ Feedback ปรับปรุงทันที

7. ช่วงอายุ (Age Group)

ผลการวิเคราะห์: กลุ่มอายุ 46–60 ปี มีอัตราการยกเลิกสูงที่สุด และเป็นกลุ่มรายได้หลักของธนาคาร แนวทางแก้ไข:

- สร้างโปรแกรม “วางแผนการเงินวัยทำงานปลาย” เช่น การลงทุน หรือประกันชีวิต
- เพิ่มช่องทางบริการแบบมีที่ปรึกษา (Financial Advisor)
- สื่อสารผ่านช่องทางที่กลุ่มนี้เข้าถึงง่าย เช่น อีเมล หรือ SMS พร้อมข้อความสั้น ชัด เข้าใจง่าย

บรรณานุกรม

Bank Customer Churn Dataset. (2024). *Kaggle – Customer Churn Modeling Dataset*.

Retrieved from <https://www.kaggle.com/datasets/radheshyamkollipara/bank-customer-churn>