

Banking Customers Churn Analysis



ที่มาและความสำคัญ

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



วัตถุประสงค์

1. เพื่อวิเคราะห์ปัจจัยที่มีผลต่อการยกเลิกบริการของลูกค้าธนาคาร



2. เพื่อหาความสัมพันธ์ระหว่างตัวแปรเชิงประชากรและพฤติกรรมการใช้บริการกับสถานะการยกเลิกบริการ

3. เพื่อสร้างข้อมูลเชิงลึก (Insight) ที่สามารถใช้สนับสนุนการตัดสินใจด้านกลยุทธ์การตลาดและการบริการลูกค้า

ประโยชน์ที่คาดว่าจะได้รับ

1. เพื่อให้ธนาคารสามารถเข้าใจสาเหตุของการเลิกใช้บริการจากพฤติกรรมและข้อมูลประชากรของลูกค้า
2. เพื่อคาดการณ์แนวโน้มการสูญเสียลูกค้าและระบุลูกค้าที่มีความเสี่ยงสูงต่อการยกเลิกบริการ
3. เพื่อใช้เป็นเครื่องมือในการกำหนดกลยุทธ์ด้านการตลาด เช่น การออกแบบบริการใหม่ การทำโปรโมชั่นเฉพาะกลุ่ม และการปรับปรุงคุณภาพการให้บริการ
4. เพื่อเพิ่มประสิทธิภาพในการรักษาลูกค้าเดิม ลดต้นทุนในการหาลูกค้าใหม่ และเพื่อเพิ่มความสามารถในการแข่งขันขององค์กร



ข้อมูลและวิธีการวิเคราะห์ (Data and Methodology)

ข้อมูลทั่วไปของลูกค้า (Customer Information)

CustomerId — รหัสประจำตัวของลูกค้า (เป็นค่าแบบสุ่ม ไม่มีผลต่อการตัดสินใจยกเลิกบริการ)

Surname — นามสกุลของลูกค้า (ไม่มีผลต่อการยกเลิกบริการของลูกค้า)

Geography — ภูมิภาคหรือประเทศที่ลูกค้าอาศัยอยู่ France, Germany, Spain ซึ่งอาจมีผลต่อแนวโน้มการยกเลิกบริการ

Gender — เพศของลูกค้า (Male/Female)

ข้อมูลทางการเงิน (Financial Information)

CreditScore — คะแนนเครดิตของลูกค้า

Balance — ยอดเงินคงเหลือในบัญชีของลูกค้า

EstimatedSalary — รายได้โดยประมาณของลูกค้า

ข้อมูลความพึงพอใจและประสบการณ์ลูกค้า (Customer Experience)

Complain — ระบุว่าลูกค้ามีการร้องเรียนหรือไม่ (1 = มี, 0 = ไม่มี)

Satisfaction Score — คะแนนความพึงพอใจของลูกค้า โดยคะแนนต่ำอาจสัมพันธ์กับอัตราการยกเลิกที่สูงขึ้น

Card Type — ประเภทของบัตรที่ลูกค้าใช้ เช่น Silver, Gold, Platinum

Points Earned — คะแนนสะสมจากการใช้บัตรเครดิต ยิ่งคะแนนมากอาจแสดงถึงการใช้บริการอย่างต่อเนื่อง

ข้อมูลพฤติกรรมการใช้บริการ (Service Usage Information)

NumOfProducts — จำนวนผลิตภัณฑ์ทางการเงินที่ลูกค้าใช้

HasCrCard — การถือบัตรเครดิต (1 = มี, 0 = ไม่มี)

IsActiveMember — สถานะการใช้งานบัญชี (1 = Active, 0 = Inactive)

ข้อมูลด้านเวลาและความสัมพันธ์กับธนาคาร (Customer Relationship)

Age — อายุของลูกค้า โดยทั่วไปลูกค้าที่อายุน้อยมักยกเลิกบริการมากกว่ากลุ่มอายุสูง

Tenure — ระยะเวลาการเป็นลูกค้าธนาคาร (หน่วยเป็นปี) ลูกค้าที่อยู่มานานมักมีความภักดีต่อธนาคารมากกว่า

ตัวแปรผลลัพธ์ (Target Variable)

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

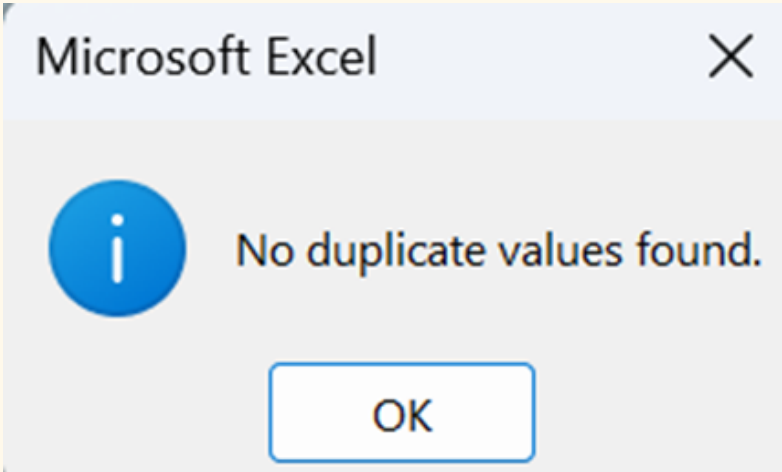
Data Cleaning เช็คค่า missing

123 CustomerId	AB_C Surname	123 CreditScore	AB_C Geography	AB_C Gender	123 Age	123 Tenure
<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>

Balance	123 NumOfProducts	123 HasCrCard	123 IsActiveMember	1.2 EstimatedSalary	123 Exited	123 Complain
<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>

123 Satisfaction Score	AB_C Card Type	123 Point Earned	AB_C Age Group	AB_C Tenure group	AB_C Credit score Group
<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>	<div><div>Valid</div><div>Error</div><div>Empty</div></div> <div>100%</div> <div>0%</div> <div>0%</div>

เช็ค Duplicates



ผลการวิเคราะห์เบื้องต้น

Section 1: ข้อมูลภาพรวมของลูกค้า				
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	เฉลี่ยรายได้	อัตราการChurn (%) เทียบกับ ประเทศอื่นๆ	ยอดเงินคงเหลือเฉลี่ย
France	5014	99899	16.17%	62093
Spain	2477	99441	16.67%	61818
Germany	2509	101113	32.44%	119730
Section 4: ระดับสมาชิก (Card Category)				
ประเภทบัตร	จำนวนลูกค้า	รายได้เฉลี่ย	อัตราChurn (%) เทียบกับ บัตรอื่นๆ	
Silver	2496	101093	20.11%	
Gold	2502	100551	19.26%	
Platinum	2495	100198	20.36%	
Diamond	2507	98526	21.78%	
Section 5: พฤติกรรมลูกค้า				
ตัวแปร	ค่าเฉลี่ย			
คะแนนความพึงพอใจ	3.01			
รายได้โดยประมาณ	100090			
คะแนนสินเชื่อ	650.53			
Section 6: ความสัมพันธ์ระหว่างปัจจัย				
ตัวชี้วัด	ค่าเฉลี่ย			
อายุเฉลี่ยลูกค้า	38.92			
อายุเฉลี่ยของลูกค้าที่ลาออก	44.84			
อายุเฉลี่ยของลูกค้าที่คงอยู่	37.41			
Section 7: การแบ่งช่วงอายุ				
ช่วงอายุ	จำนวนลูกค้า	ร้อยละ		
18-25 ปี	611	6.11%		
26-35 ปี	3542	35.42%		
36-45ปี	3,736	37.36%		
46-60ปี	1647	16.47%		
60ปีขึ้นไป	464	4.64%		
Section 8: Active vs Non-Active Customers				
สถานะ	ร้อยละ	อัตราChurn(%)		
Active (ยังใช้บริการอยู่)	51.50%	14.27%		
Non-Active (ลาออก/ไม่ใช้บริการ)	48.50%	26.87%		

ผลลัพธ์จากการใช้ function ใน Excel

=COUNT

=COUNTIF

=COUNTIFS

=AVERAGE

=AVERAGEIFS



เพื่อคำนวณอัตราการลาออก (Churn Rate) และสัดส่วนลูกค้าแต่ละกลุ่ม โดยไม่ต้องใช้สูตรนอก Pivot

Measure

Table Name:

Table1

Measure Name:

Total Customers

Value Description:

Formula:

fx

Check DAX Formula

=countrows(Table1)

Category:

General

Date

Number

Currency

TRUE\FALSE

Measure

Table Name:

Table1

Measure Name:

Churn customers

Value Description:

Formula:

fx

Check DAX Formula

=CALCULATE(countrows(table1),Table1[Exited] = 1)

Category:

General

Date

Number

Currency

TRUE\FALSE

Measure

Table Name:

Table1

Measure Name:

Churn rate

Value Description:

Formula:

fx

Check DAX Formula

=DIVIDE([Churn customers], [Total Customers],0)

Category:

General

Date

Number

Currency

TRUE\FALSE

เกณฑ์ อัตราการยกเลิกการใช้บริการ และ conditional formatting

Churn rate	20.38%
------------	--------

Conditional Formatting Rules Manager

Show formatting rules for:

This Worksheet

New Rule...

Edit Rule...

Delete Rule

Duplicate Rule

Rule (applied in order shown)	Format	Applies to
Cell Value < 0.1	AaBbCcYyZz	=I\$7:I\$8,\$O\$7:\$O\$11,\$I\$16:\$I\$18,\$O\$17:\$O\$20 <div></div>
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 <div></div>
Cell Value between 0.2 and 0.29999	AaBbCcYyZz	=I\$7:I\$8,\$O\$7:\$O\$11,\$I\$16:\$I\$18,\$O\$17:\$O\$20 <div></div>
Cell Value > 0.3	AaBbCcYyZz	=I\$7:I\$8,\$O\$7:\$O\$11,\$I\$16:\$I\$18,\$O\$17:\$O\$20 <div></div>

OK

Close

Colors of churn rate		
	> 30%	
	20 - 29.99%	
	10 - 19.99%	
	< 10%	

ผลลัพธ์จากการนำข้อมูล Pivot table เพื่อหาความสัมพันธ์การยกเลิกบริการของลูกค้าธนาคาร

[illegible]

สถิติทดสอบ t-test

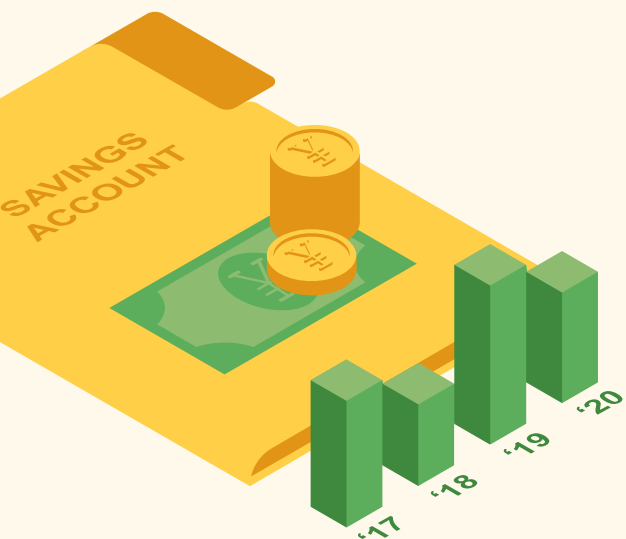
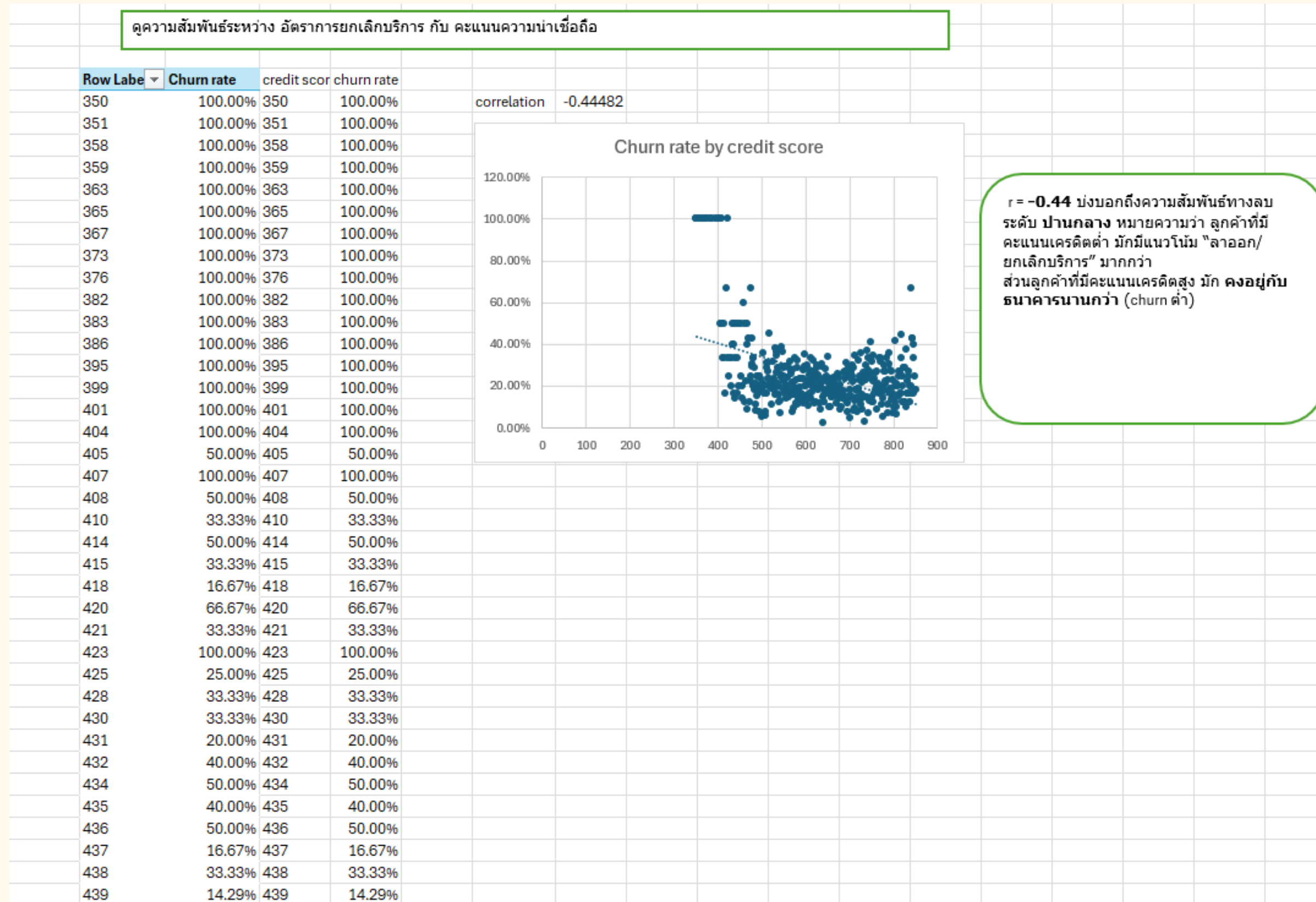
ต้องการทดสอบว่า ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่ หรือไม่?

Tenure (Retained)	Tenure (churned)
1	2
1	8
2	8
7	4
4	1
2	4
6	3
3	0
10	8
5	2
7	9
3	9
9	8
6	1
6	4
8	2
8	2
3	3
5	4
3	5
2	1
9	4
3	9
0	9
7	8
9	9
6	1
9	0

ต้องการทดสอบว่า ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่ หรือไม่?			
ทดสอบความแปรปรวน 2 กลุ่ม ด้วย F test			
H0: ค่าความแปรปรวนระยะเวลาใช้บริการของ ผู้ที่ยกเลิกบริการแล้ว ไม่แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่			
H0: ค่าความแปรปรวนระยะเวลาใช้บริการของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่			
F-Test Two-Sample for Variances			
	<i>Tenure (Exited=1)</i>	<i>Tenure (Exited=0)</i>	
Mean	4.934739941	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	ยอมรับH0 Assume variance
F Critical one-tail	1.058864057		

ทำการทดสอบ t test two sample แบบ assume varinece			
H0: ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว ไม่แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่			
H1: ระยะเวลาใช้บริการเฉลี่ยของ ผู้ที่ยกเลิกบริการแล้ว แตกต่างจาก ผู้ที่ยังคงใช้บริการธนาคารอยู่			
t-Test: Two-Sample Assuming Equal Variances			
	<i>Tenure (Churned)</i>	<i>Tenure (Retained)</i>	
Mean	4.934739941	5.032780708	
Variance	8.624605013	8.297254646	
Observations	2038	7962	
Pooled Variance	8.363949255		
Hypothesized Mean Difference	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 ผู้ที่ยังคงใช้บริการธนาคารอยู่		

Correlation ดูความสัมพันธ์ระหว่างอัตราการยกเลิกบริการกับคะแนนความน่าเชื่อถือ



Logistic Regression

สร้างโมเดล logistic regression (imbalance class)
โมเดลเพื่อใช้ทำนายการยกเลิกบริการของลูกค้า(y)จาก ตัวแปรต่างๆ (x)

Logistic/Probit Regression

Input Range: a1:M10001

☒ Column headings included with data
☐ Show item by item details in output
☒ Hide ROC table

Regression Type
☒ Logistic ☐ Probit

Input Format
☒ Raw data ☐ Summary data

Analysis Type
☒ Newton's method ☐ Solver

Alpha: 0.05

Classification Cutoff: 0.2038

of Iterations (Newton's method only): 20

Output Range: AG22

	<i>coeff</i>	<i>s.e.</i>	<i>Wald</i>	<i>p-value</i>	<i>exp(b)</i>	<i>lower</i>	<i>upper</i>		LL	-4297.4
intercept	-3.5949	0.25248	202.731	0	0.0274629				LL0	-5056.3
CreditScore	-0.0007	0.00028	5.8052	0.015978821	0.9993258	0.99878	0.99987		Chi-sq	1517.73
Geography	0.35442	0.03361	111.174	0	1.4253532	1.33448	1.52242		df	12
Gender	-0.5295	0.05434	94.936	0	0.5888932	0.52939	0.65508		p-value	0
Age	0.07272	0.00257	800.895	0	1.0754248	1.07002	1.08085		R-sq (L)	0.15008
Tenure	-0.0152	0.00934	2.6408	0.104150771	0.9849378	0.96707	1.00313		R-sq (CS)	0.14082
Balance	3.5E-06	4.8E-07	51.8985	5.84421E-13	1.0000035	1	1		R-sq (N)	0.22133
HasCrCard	-0.0392	0.05918	0.43965	0.507293101	0.9615182	0.85621	1.07978		AIC	8620.78
IsActiveMember	-1.0805	0.05757	352.266	0	0.3394143	0.3032	0.37996		BIC	8714.51
EstimatedSalary	4.9E-07	4.7E-07	1.08428	0.297741694	1.0000005	1	1			
Satisfaction	-0.0093	0.01924	0.23156	0.630366758	0.9907861	0.95413	1.02885			
Card Type	0.01978	0.02423	0.66676	0.414182758	1.0199793	0.97268	1.06958			
Point Earned	-0.0001	0.00012	1.17792	0.277778926	0.9998695	0.99963	1.00011			

ตัวแปรที่ pvalue < 0.05 ซึ่งมีผลต่อการยกเลิกบริการ ได้แก่ credit score geography gender age tenure balance is active member

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

คำนวณค่า prob จากสมการโมเดล และทำนาย ให้exited=0 กรณี prob<0.05 (retained) , exited=1 กรณี>0.05(Churned)

Logistic regression evaluation

การประเมินโมเดล

Logistic regression
imbalance class

Classification Table				
	churn		retainted	
	Obs	Suc	Obs	Fail
Pred Suc	437		273	
Pred Fail	1601		7689	
Total	2038		7962	
Accuracy	0.21443		0.96571	
Cutoff	0.5			
AUC	0.76511			

ข้อมูลนี้มีปัญหา Class Imbalance ค่อนข้างมาก คือมีลูกค้าที่ ยกเลิกบริการจริง เพียง 20% (2,038 จาก 10,000) model bias ไปทาง ผู้ที่ยังไม่ยกเลิกบริการ ค่อนข้างมาก เราจึงแก้ไขด้วยการ undersampling ในหน้าถัดไป

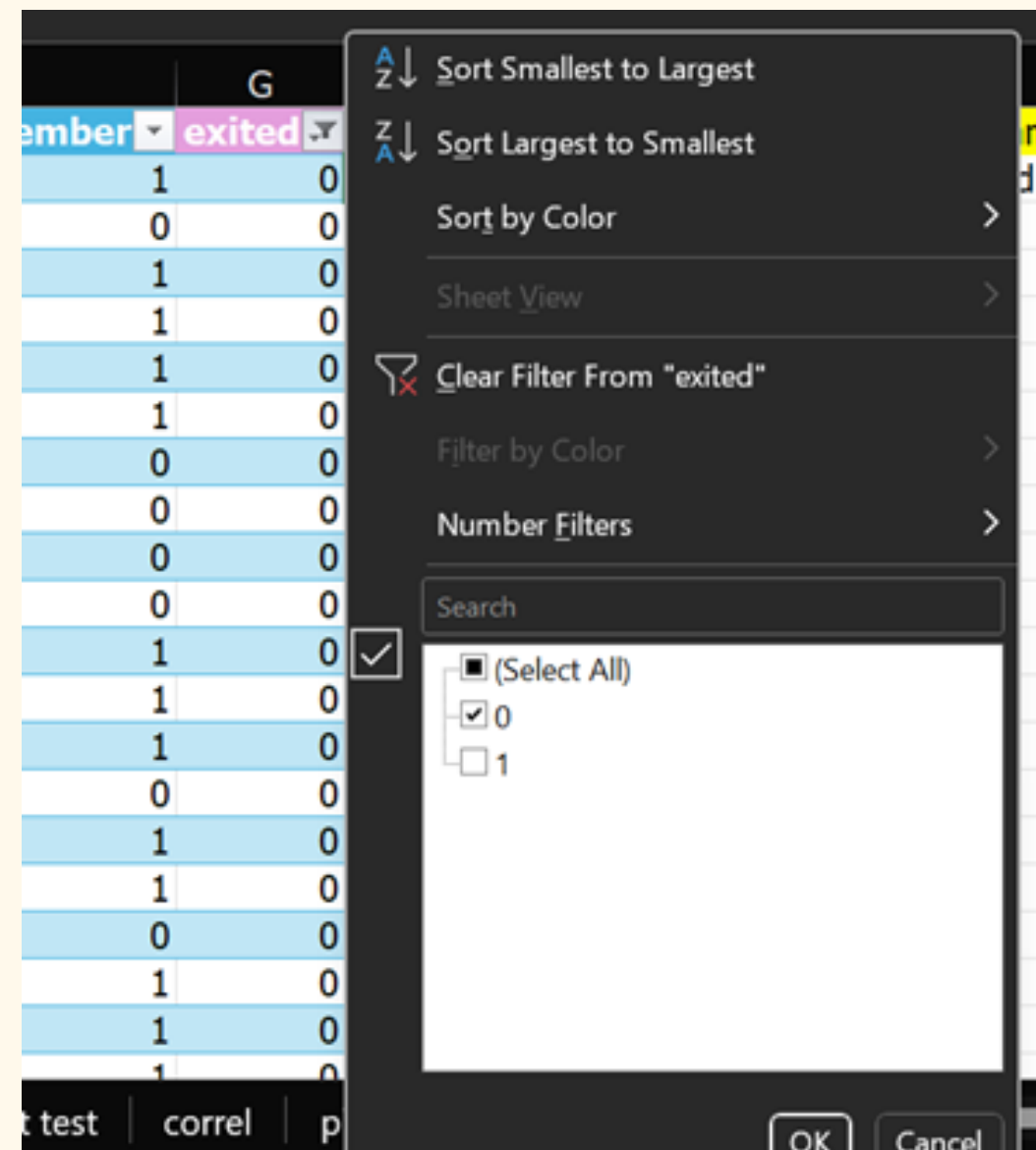
balance class ด้วย
เทคนิค Random
Under sampling

1.สร้างข้อมูล random ด้วยฟังก์ rand() เป็น column ใหม่ขึ้นมา

=RAND()							
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

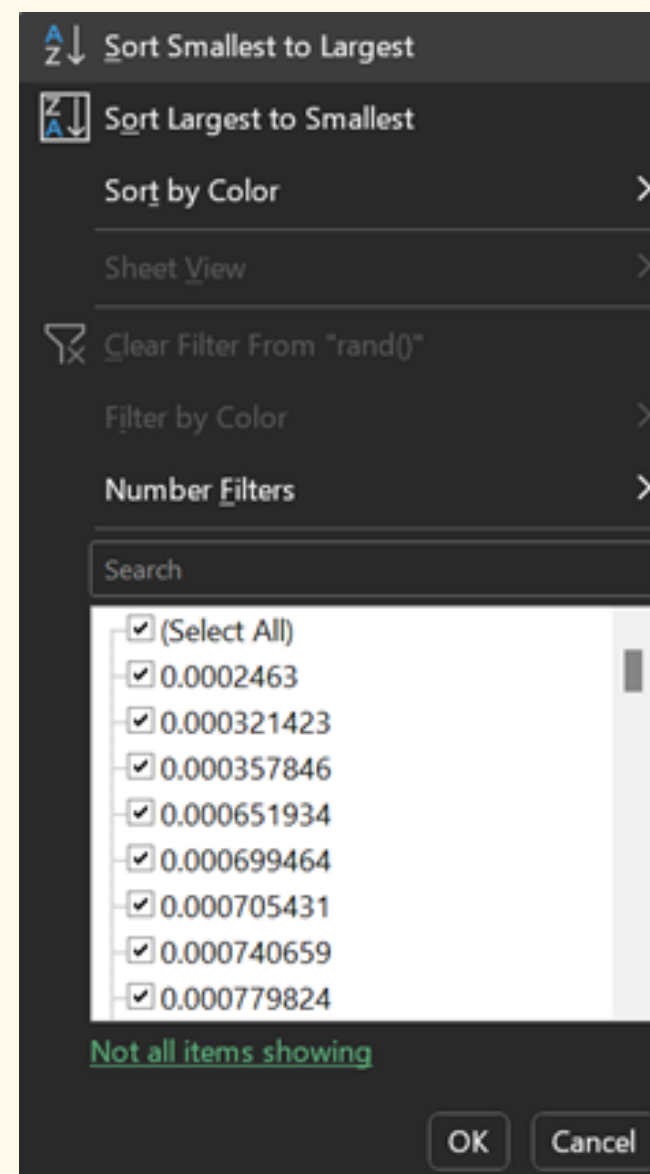
balance class ด้วย
เทคนิค Random
Under sampling

2.filter เลือก classที่ใหญ่กว่า exited=0



balance class ด้วย เทคนิค Random Under sampling

3.sortค่า rand()ให้เรียงจากน้อยไปมา แล้วทำการตัดข้อมูลส่วนนั้น 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		Under sampling ให้ imbalance เนื่องจาก excel ทำ smote หรือ tomes link ไม่ได้ แบบ python R								
1	0	0.385747	0										
1	1	0.521427	1		Retained	2061	exited =0						
1	0	0.273506	0		churned	2038	exited =1						
0	0	0.593788	1		รวม	4099							
1	0	0.117993	0										

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

=1/(1+EXP(-(\$L\$10 + \$L\$11*A3 + \$L\$12*B3 +						
D	E	F	G	H	I	
Age	Balance	IsActiveMember	exited	prob	pred	
40	95624.36	1	0	0.385747	0	
50	0	1	1	0.521427	1	
32	131081.66	1	0	0.273506	0	
40	0	0	0	0.593788	1	
25	127728.24	1	0	0.117993	0	
30	117356.19	0	1	0.569247	1	
31	118899.45	0	0	0.362941	0	

balance class ด้วย เทคนิค Random Under sampling

Logistic Regression								LL statistics		Covari
		# Iter	20		Alpha	0.05				
	<i>coeff</i>	<i>s.e.</i>	<i>Wald</i>	<i>p-value</i>	<i>exp(b)</i>	<i>lower</i>	<i>upper</i>	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	0.999754	Chi-sq	1115.273	-0.00
Geography	0.388449	0.044496	76.21198	0	1.474692	1.351532	1.609076	df	6	-0.00
Gender	-0.63513	0.072232	77.31414	0	0.529868	0.459922	0.610453	p-value	0	-0.00
Age	0.097637	0.004151	553.2411	0	1.102562	1.093628	1.111569	R-sq (L)	0.196272	-2.6
Balance	3.11E-06	6.29E-07	24.37995	7.91E-07	1.000003	1.000002	1.000004	R-sq (CS)	0.23821	-0.00
IsActiveMe	-0.87597	0.073121	143.5119	0	0.416458	0.360853	0.480632	R-sq (N)	0.317616	
								AIC	4581.019	

สมการโลจิสติก

Logit(P)

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

การประเมินโมเดล Logistic regression imbalance class vs balance class

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 คือมีลูกค้าที่ ยกเลิกบริการจริงเพียง 20% (2,038 จาก 10,000) model bias ไปทาง ผู้ที่ยังไม่ยกเลิกบริการ ค่อนข้างมาก เราจึงแก้ไขด้วยการ undersampling ในหน้าถัดไป

Imbalanceclass

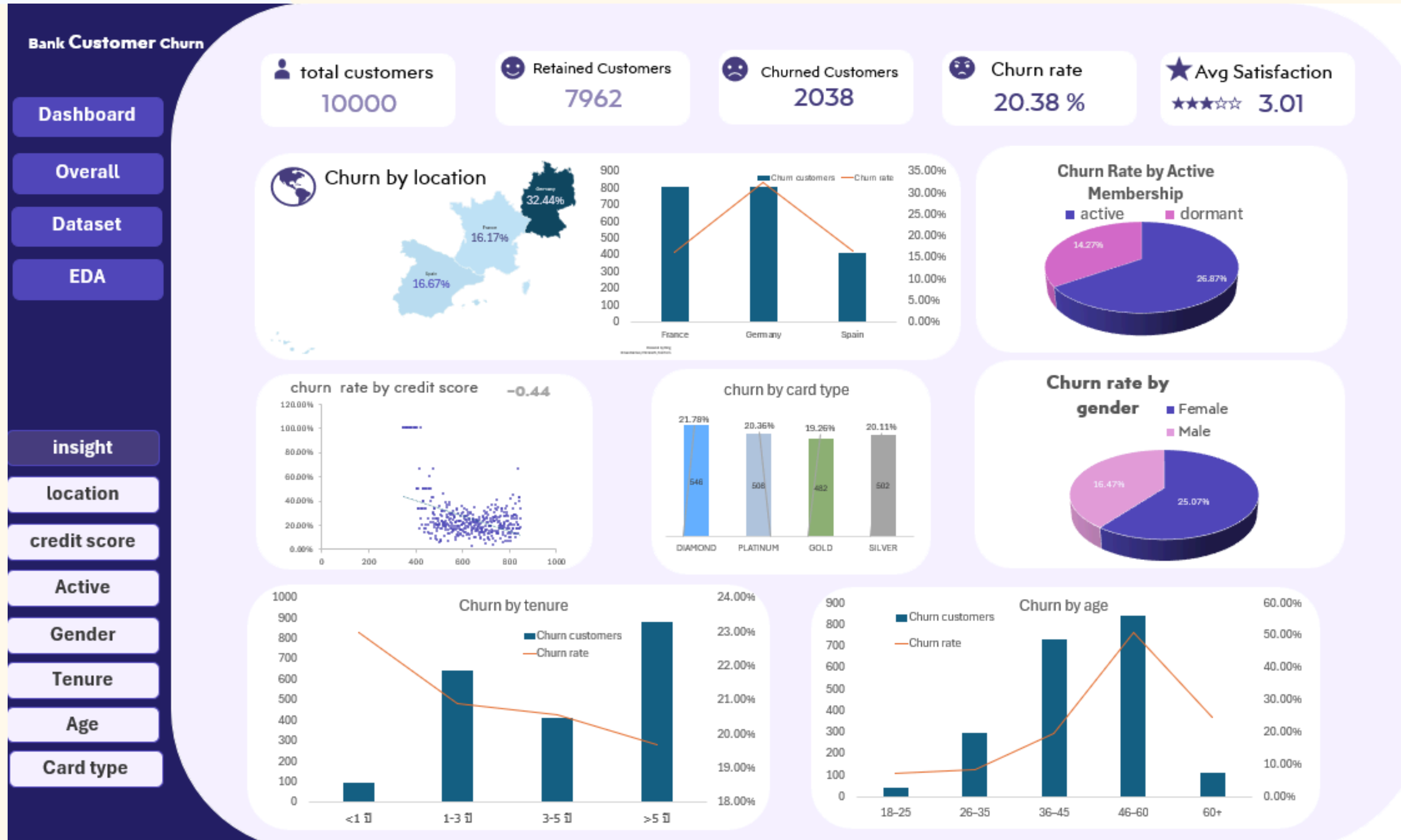
	Churned	retained	
	Obs Suc	Obs Fail	Total
Pred Suc	1432	555	1987
Pred Fail	606	1506	2112
Total	2038	2061	4099
Accuracy	0.70265	0.730713	0.71676
Cutoff	0.5		
AUC	0.787856		

ใช้ Random Undersampling ด้วยฟังก์ชัน RAND() เพื่อปรับสัดส่วนข้อมูลให้สมดุล (50:50) ทำให้ accuracy ของกลุ่มยกเลิกบริการเพิ่มขึ้นอย่างชัดเจน แม้ Accuracy รวมจะลดลงเล็กน้อย ส่งผลให้โมเดลเข้าใจพฤติกรรมลูกค้าเสี่ยงลูกค้าที่ยกเลิกได้ดีขึ้น

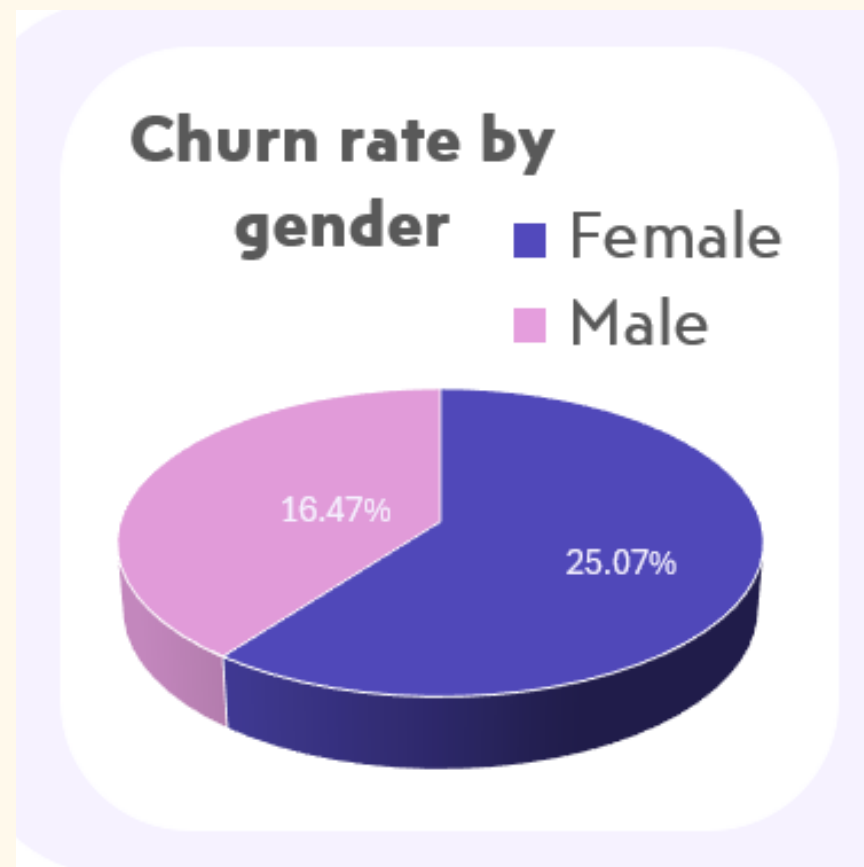
Balanace class (random undersampling)

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

Dashboard



Insight



Gender				
Count of Gender	Exited			
Geography	Retained (0)	Churned (1)	Grand Total	Churn rate
Female	3404	1139	4543	25.07%
Male	4558	899	5457	16.47%
Grand Total	7962	2038	10000	

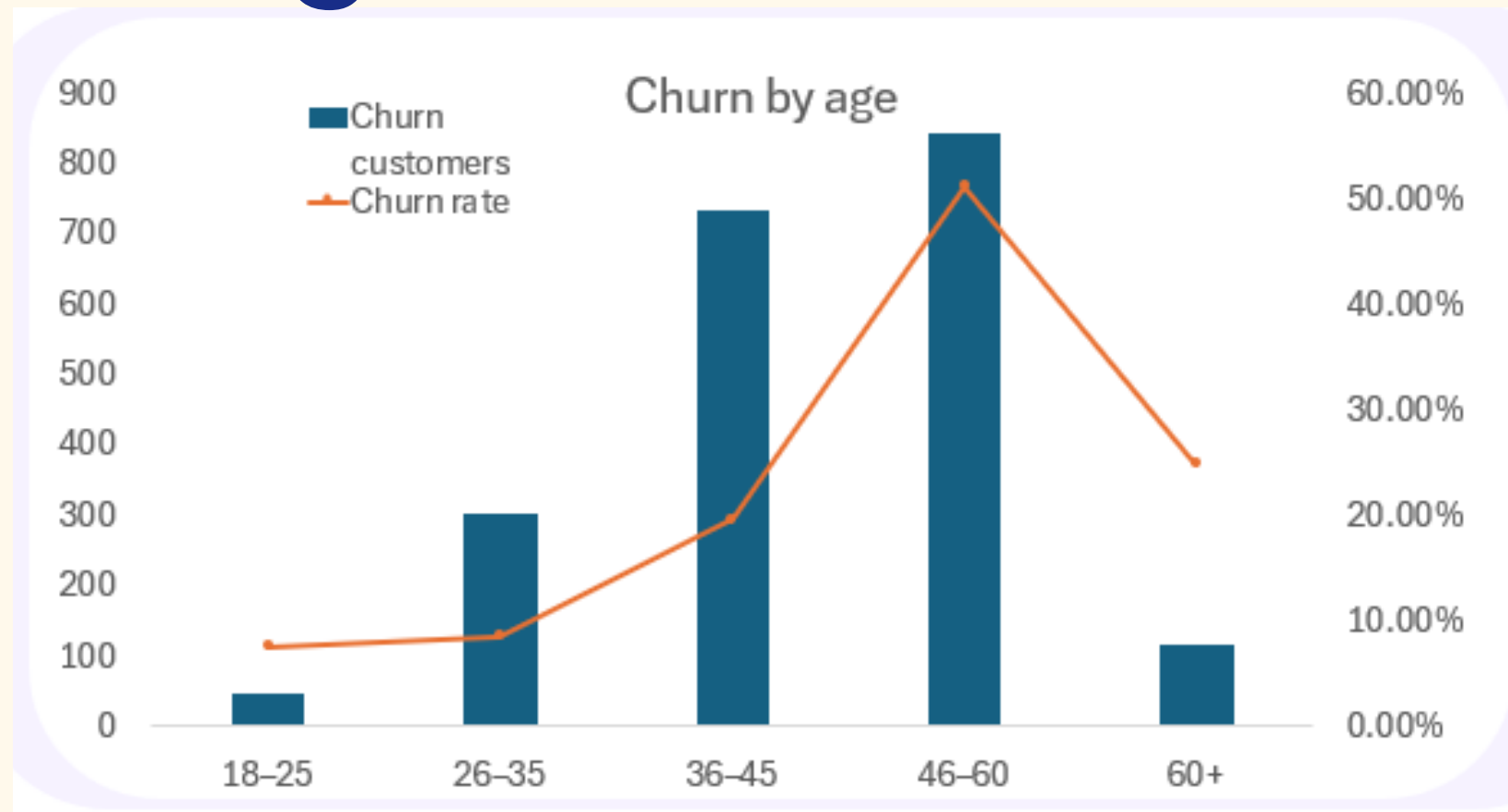
insight

ผู้หญิงมีอัตราการยกเลิกบริการสูงกว่าผู้ชายอย่างชัดเจน แนะนำให้เพิ่มการดูแลด้านประสบการณ์ลูกค้าและมอบสิทธิประโยชน์ที่ตอบโจทย์เพื่อช่วยลด Churn ในกลุ่มลูกค้าผู้หญิง

solution

- 1.เพิ่มช่องทางบริการที่รวดเร็วและเป็นมิตร เช่น แชทบอท / สายด่วนเฉพาะกลุ่มลูกค้าผู้หญิง
- 2.จัดแคมเปญสิทธิพิเศษที่ตรงกับไลฟ์สไตล์ เช่น ส่วนลดร้านอาหาร สุขภาพ และความงาม
- 3.ทำคอนเทนต์สั้น ๆ เกี่ยวกับ “วางแผนการเงินง่าย ๆ สำหรับผู้หญิง” เพื่อสร้างความผูกพัน

Insight



Age				
Count of Age Group	exited			
Row Labels	Retained (0)	Churned (1)	Grand Total	Churn rate
18-25	565	46	611	7.53%
26-35	3241	301	3542	8.50%
36-45	3002	734	3736	19.65%
46-60	805	842	1647	51.12%
60+	349	115	464	24.78%
Grand Total	7962	2038	10000	

insight

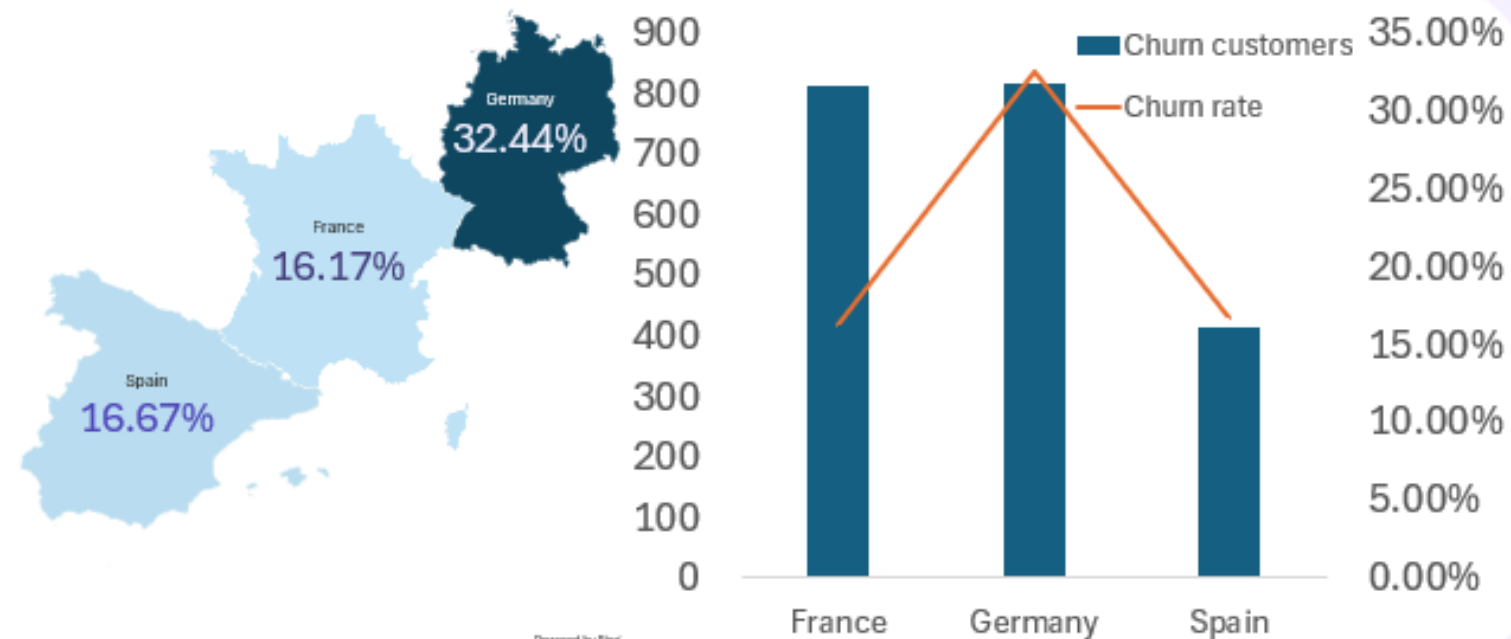
อัตราการลาออกสูงที่สุดในกลุ่มอายุ 46-60 ปี ซึ่งเป็นกลุ่มลูกค้าที่มีความคาดหวังสูงและมีผลต่อรายได้รวมมาก ดังนั้นควรให้ความสำคัญกับการดูแลลูกค้ากลุ่มนี้แบบเฉพาะบุคคล เช่น การให้คำแนะนำด้านการเงินและสิทธิพิเศษเพิ่มเติม

solution

- 1.สร้างโปรแกรม “วางแผนการเงินวัยทำงานปลาย” เช่น การลงทุน / ประกันชีวิต
- 2.เพิ่มช่องทางบริการแบบมีที่ปรึกษา (Financial Advisor)
- 3.สื่อสารผ่านช่องทางที่กลุ่มนี้เข้าถึงง่าย เช่น อีเมล หรือ SMS พร้อมข้อความสั้น ชัด เข้าใจง่าย

Insight

Churn By Location



Location				
Count of Location	Exited			
Card Type	retained (0)	Churned (1)	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	

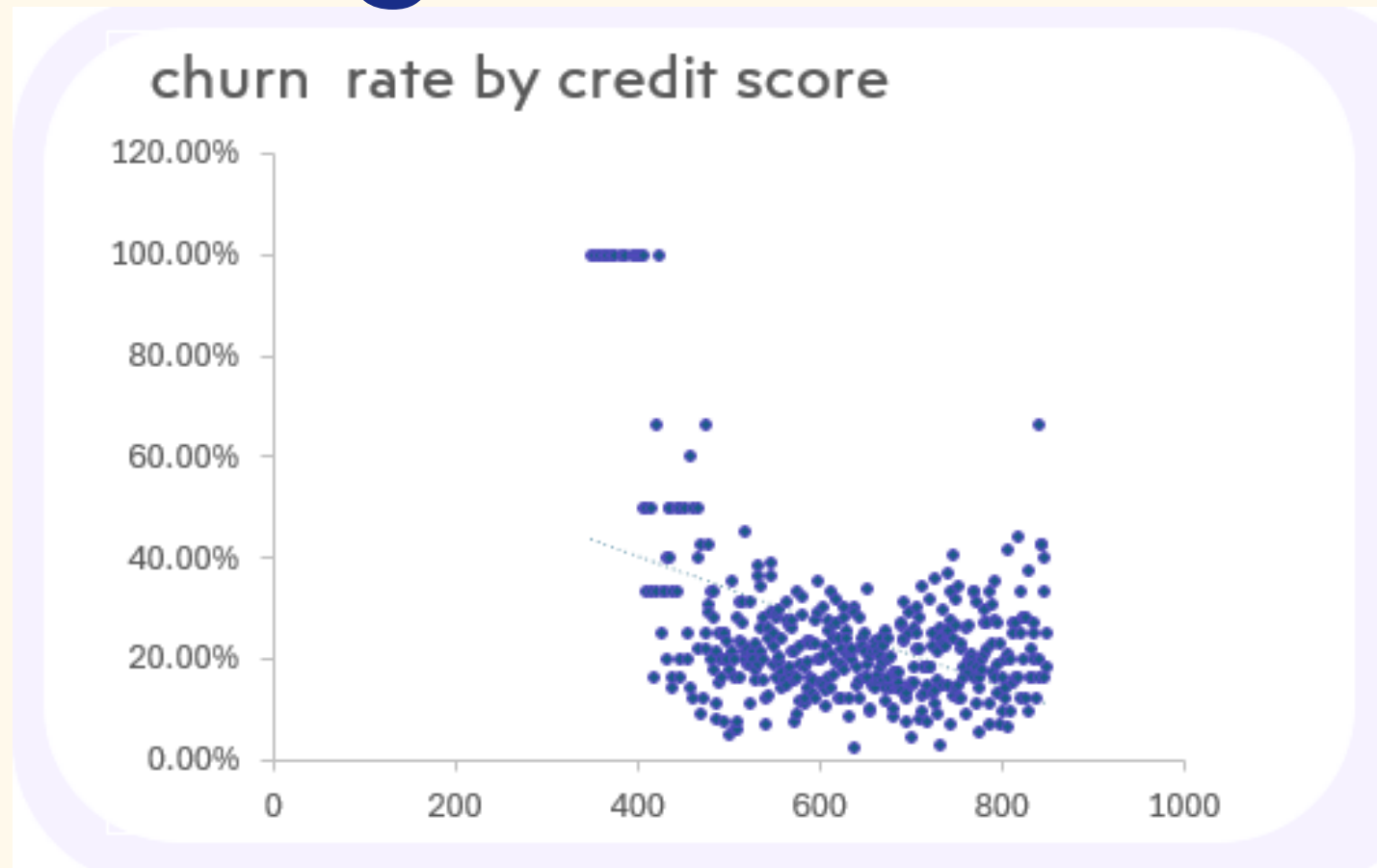
insight

เยอรมนีเป็นตลาดที่มีปัญหา Churn สูงที่สุด จึงต้องวิเคราะห์ประสบการณ์การให้บริการเชิงพื้นที่อย่างเร่งด่วน ขณะที่ฝรั่งเศสเป็นตลาดสำคัญที่ต้องรักษาไว้เพราะมีฐานลูกค้าขนาดใหญ่ แม้อัตราการลาออกจะต่ำก็ตาม

solution

1. วิเคราะห์ประสบการณ์การใช้งานของลูกค้าในเยอรมนี (เช่น ค่าธรรมเนียม ภาษา หรือการบริการหลังการขาย)
2. ปรับบริการให้เหมาะกับพฤติกรรมของแต่ละประเทศ
3. สำหรับฝรั่งเศส ควรเน้นรักษาฐานลูกค้าด้วยโปรแกรมสะสมแต้มระยะยาว

Insight



✕ ✓ fx $=CORREL(B5:B424,C5:C424)$						
B	C	D	E	F	G	H
ดูความสัมพันธ์ระหว่าง อัตราการยกเลิกบริการ กับ คะแนนความน่าเชื่อถือ						
Label	Churn rate	credit score	churn rate			
	100.00%	350	100.00%		correlation	-0.44482
	100.00%	351	100.00%			

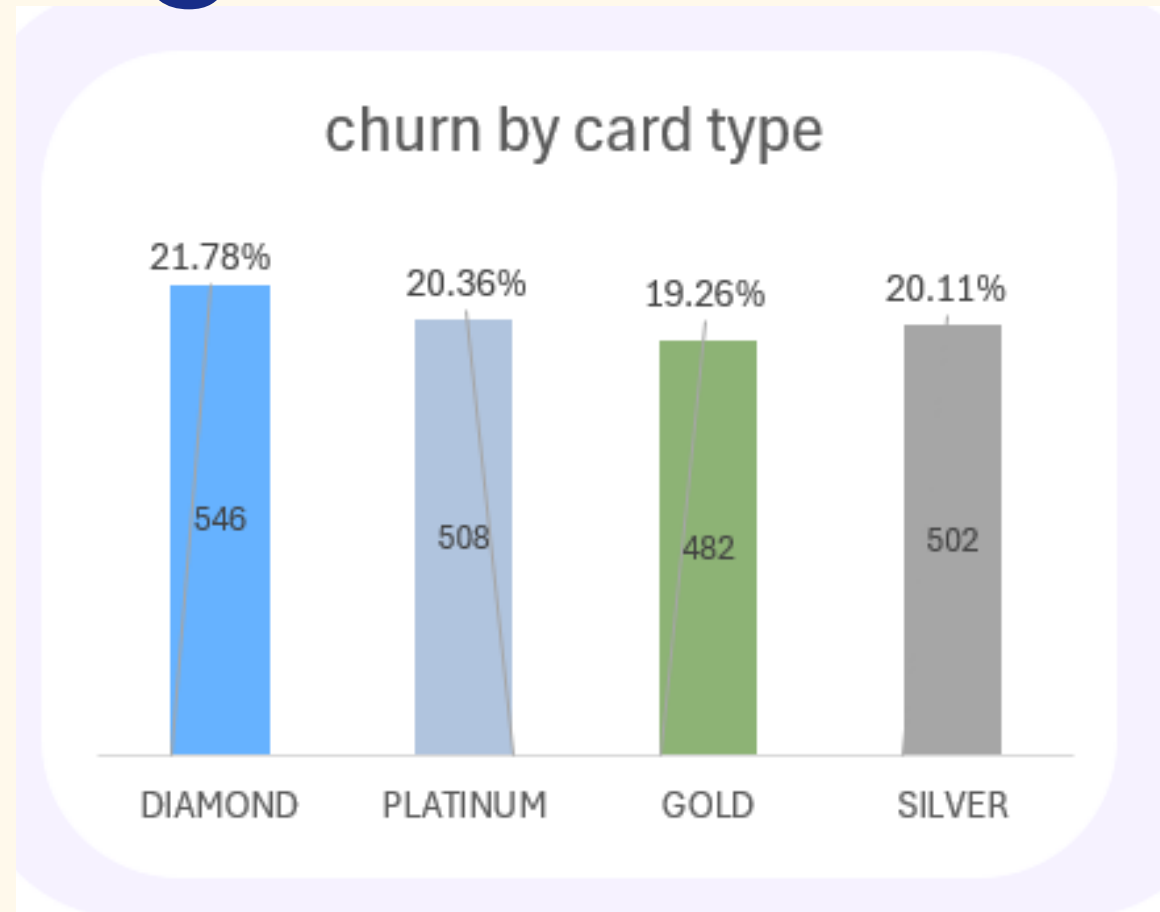
insight

ลูกค้าที่มีคะแนนเครดิตต่ำมีอัตราการยกเลิกสูงกว่ากลุ่มอื่นอย่างเห็นได้ชัด ซึ่งสะท้อนว่าสถานะทางการเงินและต้นทุนที่ลูกค้ารับรู้เป็นปัจจัยสำคัญในการตัดสินใจเลิกใช้บริการ ดังนั้นกลยุทธ์ควรเน้นช่วยลดภาระและให้การสนับสนุนมากกว่าการขายเพิ่ม

solution

1. จัดโปรแกรม “ช่วยเหลือลูกค้าการเงิน” เช่น ผ่อนจ่าย / พักหนี้ระยะสั้น
2. ให้คำปรึกษาด้านการวางแผนการเงินฟรีสำหรับกลุ่มเครดิตต่ำ
3. ลดการเสนอขายสินค้าซ้ำในกลุ่มนี้ และเน้นการช่วยให้กลับมาใช้บริการได้มั่นคง

Insight



card type				
Count of Complain exited				
Row Labels	Retained = 0	Churned = 1	Grand Total	Churn rate
DIAMOND	1961	546	2507	21.78%
GOLD	2020	482	2502	19.26%
PLATINUM	1987	508	2495	20.36%
SILVER	1994	502	2496	20.11%
Grand Total	7962	2038	10000	

insight

กลุ่มลูกค้าบัตรระดับสูงอย่าง Diamond มีอัตราการยกเลิกสูงที่สุด
แม้เป็นกลุ่มที่มีมูลค่าสูง ดังนั้นควรเพิ่มสิทธิประโยชน์และดูแลแบบเฉพาะ
บุคคล เพื่อป้องกันไม่ให้ย้ายไปใช้บริการของคู่แข่ง

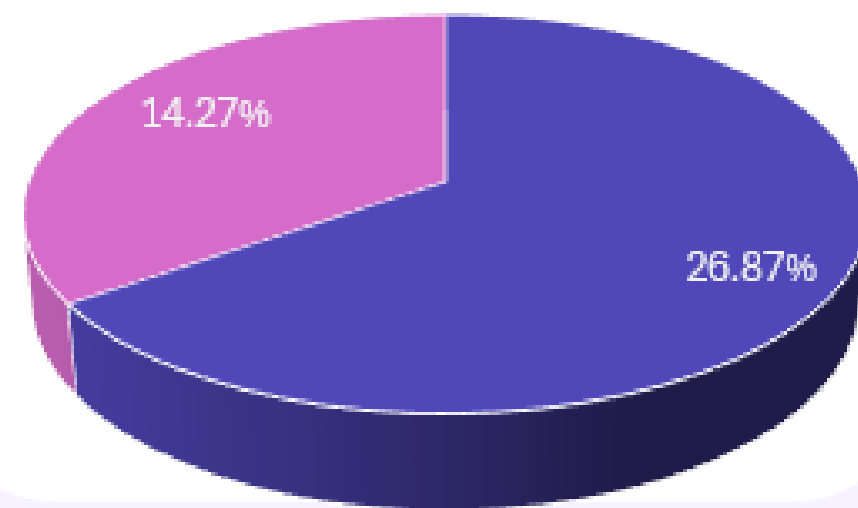
solution

1. ตั้งทีมดูแลลูกค้า Diamond เฉพาะ (Personal Banker)
2. จัดสิทธิพิเศษเฉพาะ เช่น ห้องรับรอง, กิจกรรมสุดพิเศษ หรือ Cashback เพิ่ม
3. วิเคราะห์สิทธิคู่แข่งและปรับสิทธิประโยชน์ให้เหนือกว่า

Insight

Churn Rate by Active Membership

■ dormant ■ active



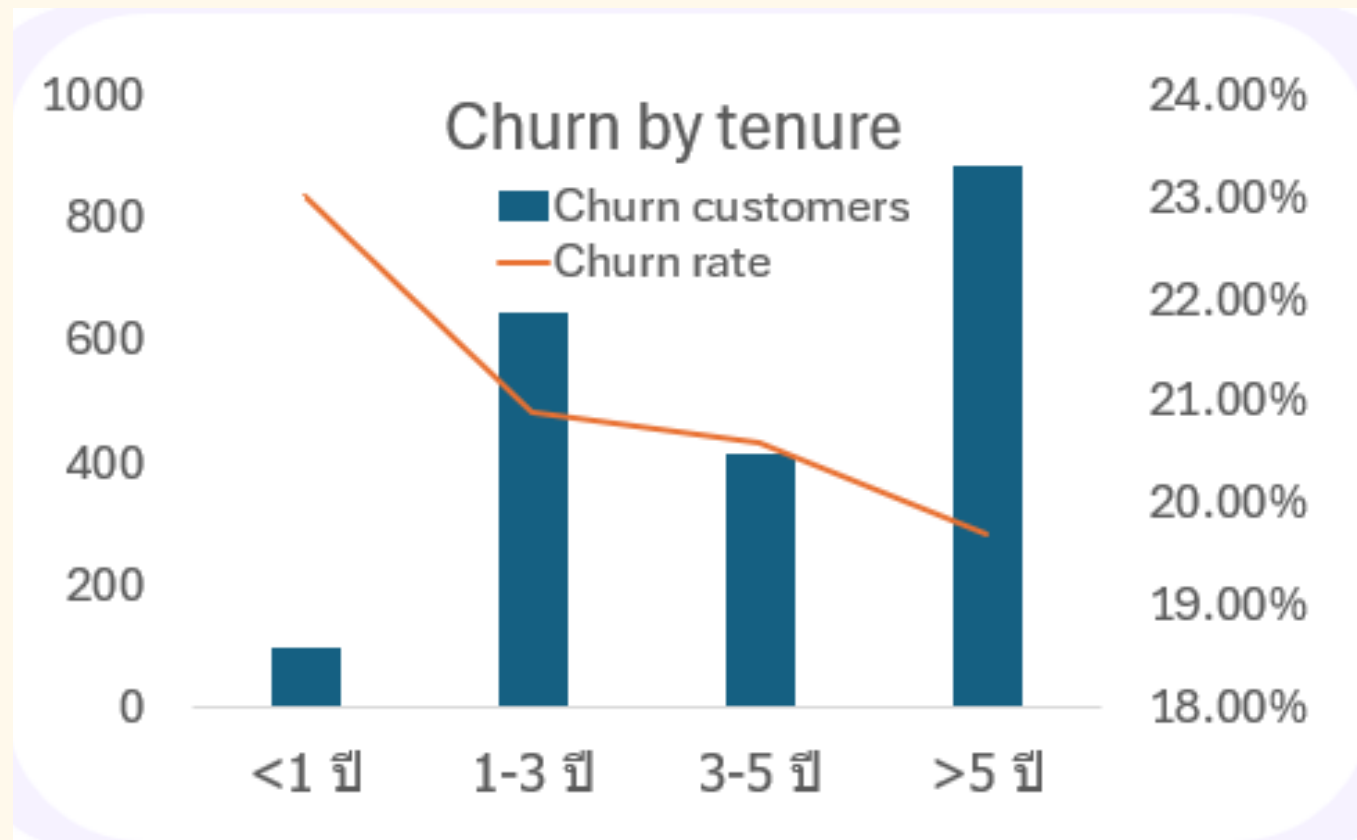
insight

ลูกค้าที่ไม่ค่อยใช้งานมีโอกาสยกเลิกบริการมากกว่าลูกค้าที่ใช้งานสม่ำเสมอเกือบสองเท่า ดังนั้นกลยุทธ์ต้องเน้น กระตุ้นการใช้งาน (Re-engagement) สำหรับกลุ่ม Dormant และ สร้างรางวัล/คะแนนสะสม เพื่อรักษากลุ่ม Active

solution

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

Insight



Tenure				
Count of Tenure group	exited			
Row Labels	Retained (0)	Churned (1)	Grand Total	Churn rate
<1 ปี	318	95	413	23.00%
>5 ปี	4412	1094	5506	19.87%
1-3 ปี	1650	433	2083	20.79%
3-5 ปี	1582	416	1998	20.82%
Grand Total	7962	2038	10000	

insight

แม้ลูกค้าที่อยู่กับธนาคาร มากกว่า 5 ปี จะมี Churn Rate ต่ำที่สุด แต่เป็นกลุ่มที่มีจำนวนลูกค้าที่ลาออกมากที่สุด เพราะเป็นกลุ่มลูกค้าฐานใหญ่ ส่งผลกระทบต่อรายได้รวมอย่างมีนัยสำคัญ จึงเป็นกลุ่มที่ควรให้ความสำคัญที่สุดในการรักษา และ ลูกค้าที่เป็นสมาชิก น้อยกว่า 1 ปี มี Churn Rate สูงที่สุด แสดงว่าช่วง Onboarding / เริ่มต้นใช้งาน มีปัญหา หรือ ลูกค้ายังไม่รู้สึกถึง Value ของบริการ

solution

ลูกค้าเก่า: ให้ลูกค้าเก่าที่มีการใช้งานต่อเนื่องเลื่อนระดับสถานะ

เช่น จาก Silver → Gold → Diamond พร้อมสิทธิพิเศษเพิ่ม เช่น ดอกเบี้ยพิเศษ, ฟรีค่าธรรมเนียมบางรายการ, หรือสิทธิลดเงิน/รูบ์ครในต่างประเทศโดยไม่เสียค่าใช้จ่าย

ลูกค้าใหม่: เมื่อเป็นสมาชิกครบทุก 1 ปี จะได้รับคะแนนพิเศษเพิ่ม หรือของขวัญขอบคุณประจำปี (Gift Card, ส่วนลดร้านค้า, สิทธิพิเศษต่าง ๆ)

ส่งแบบสอบถามหลัง 1 เดือน เพื่อเก็บ Feedback ปรับปรุงทันที

ผู้จัดทำ

66050172 ธนภฤติ อักษรเจริญสุข

66050191 ธราเทพ จูเจริญ

66050606 จิรายุต์ม์ อยู่อ่วม