

MBAn Student Team:

Eishna Ranganathan & Julia Schwartz

Visa Company Team:

Abby Srivicharnkul, John Taylor & Scott Abernathy

MIT Faculty Advisor:

Jose Pacheco



Problem Statement

Help identify underserved Small Businesses (SMB) via creating a look-alike model and visualizing key trends

SMB look-alike model: a model that identifies "consumer" accounts that exhibit SMB-like behaviors

- This project carries out Visa’s mission to accelerate SMB digitization and growth
- This model’s outputs can be commercialized as a product offered to Visa’s bank clients

Key Questions

Using **Visa transactional data, US Census demographic data**, and public data sources we want to investigate:



How can we identify underserved SMBs?



Where do these SMBs exist?



What demographic trends correlate with these SMBs?

Impact

We have identified **1.33%** of consumer accounts as SMB lookalikes; scaled to Visa’s entire **2023 H2** book this amounts to:

1.5M+
Accounts

\$20B+
Attributable Purchase Volume

\$12M+
Est. Cash Rewards Back to SMBs

Our analysis affirms the hypothesis that these SMB owners are **currently underserved**

High SMB-lookalike rates are correlated with **lower income, lower college graduation rate, and older populations**

Methodology

Data Preparation

- Created scripts to filter and sample accounts
- Aggregated transactions data into 65+ account-level features

Build Model

- Feature selection with RF baseline model
- Ran XGB and CATBoost models
- Wanted errors identifying Consumers as SMBs

Apply Model

- Ran model on consumer-only data
- Reviewed results and demographic correlations and trends

Create Dashboard

- Created a dashboard showing where these SMB look-alikes exist and identified trends

Our Model



Our model is significantly better than random at identifying SMBs, the model had a 100% accuracy rate for 40% of all SMB accounts, which have the highest SMB propensity scores

Key Takeaway
Our model has learned what distinguishes SMB behavior

Output Insights

Our model uses only 5 features, 3 of which are novel and designed by us

Top 5 Features:

- Average transaction size
- Percent of purchase volume (PV) on business-to-business services
- Percent of PV in the SMB merchant class bucket (*new*)
- Percent of transactions in the SMB merchant class bucket (*new*)
- Percent of PV on business-to-business services, home improvement or supply, automotive, or fuel (*new*)

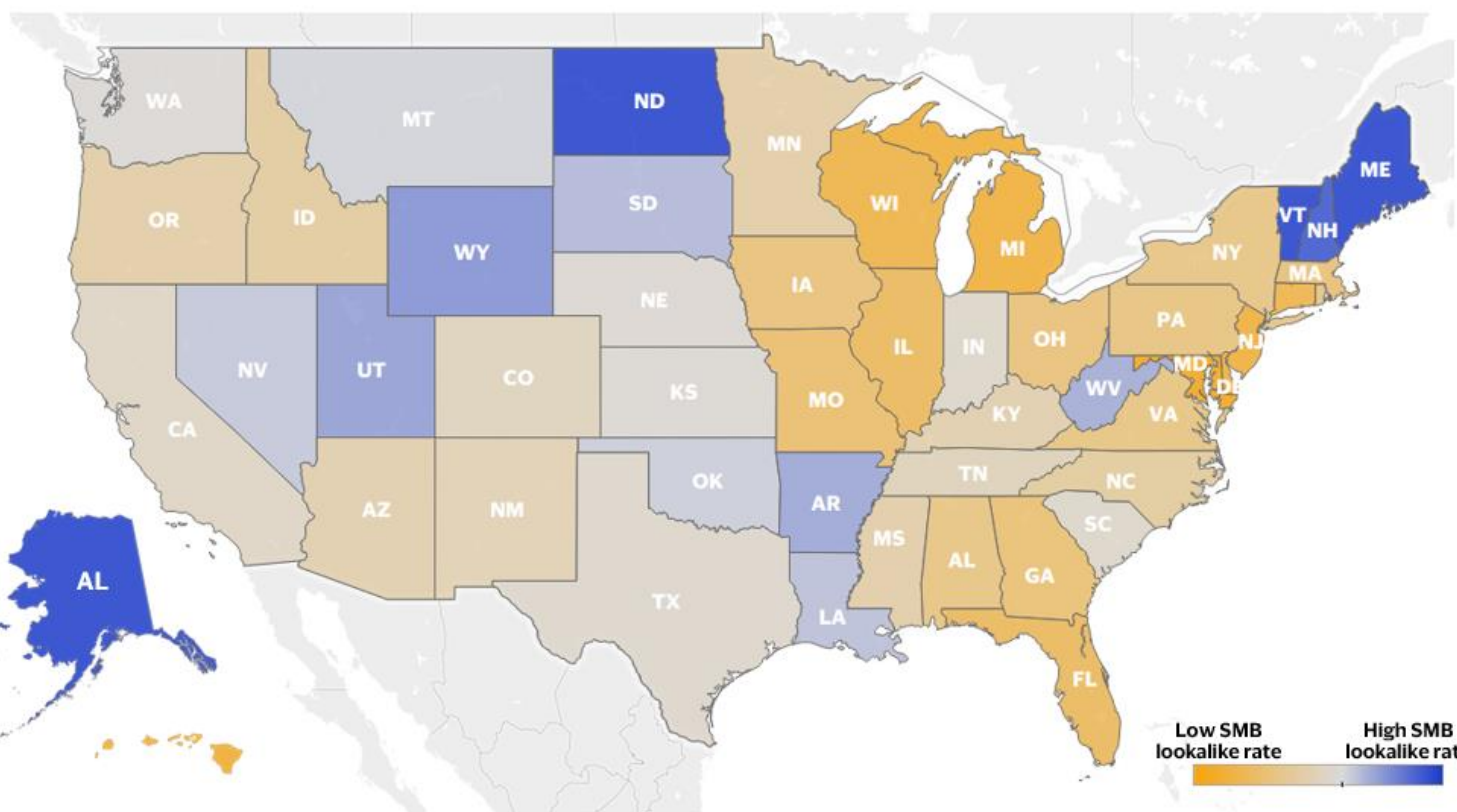
Key Takeaways

- SMBs tend to make bigger transactions than consumers
- SMBs are more likely to spend at businesses geared towards serving other businesses and similar, related services

Geographical Insights

SMB look-alike rates from our model are correlated with low population density and certain industries at the state level

(population density Pearson correlation coefficient (PCC) -10%)



Correlated with high rates of mining, logging and construction employment (46%)

Correlated with high rates of agriculture employment (30%)

Correlated with low rates of professional and business services employment (-45%)

Demographic Insights

High SMB look-alike rates are correlated with lower incomes, lower college graduation rates, and older populations

Core Based Statistical Area (CBSAs) with higher SMB look-alike rates have:

- Older populations (8% PCC)
- Lower college graduation rates (7% PCC)
- Higher percentage of white people (9% PCC)
- Lower mean incomes (5% PCC)

