Credit Card Fraud Analytics

Shreyash Kondakindi

April 15, 2025

1 Data Cleaning Analysis

1.1 Outlier Removal:

The dataset was filtered to exclude non-purchase transactions and extremely large transaction amounts. All records with `Transtype` not equal to `'P'` (purchase) were removed, as these were adjustment or non-purchase entries not relevant to fraud analysis. Additionally, transactions with `Amount > 3,000,000` were dropped as outliers (unrealistically high values). This eliminated 356 records and ensured the amount distribution is reasonable for modeling.

1.2 Missing Merchant Number (Merchnum):

Merchant IDs (`Merchnum`) had placeholders `'0'` and `'unknown'` indicating missing values. These were first replaced with `NaN`. A mapping was then created from **Merchant Description** to **Merchant ID** using records where both were present, under the assumption that the same merchant description corresponds to the same ID. This dictionary was used to fill in missing `Merchnum` for transactions that had a known description. After this step, missing `Merchnum` count dropped from 3,306 to 2,136.

1.3 Adjustment Transactions:

Certain transactions labeled as "RETAIL CREDIT ADJUSTMENT" or "RETAIL DEBIT ADJUSTMENT" (system adjustments) lacked valid merchant info. These were handled separately: their `Merchnum` was explicitly set to `'unknown'` for identification. This reduced missing `Merchnum` further (to 1,437). The remaining missing cases (about 1,403) represented unique merchants with no ID available. For each unique merchant description among these, a **new merchant ID** was created (using the next available numeric ID). These new IDs were assigned to all transactions sharing that description. After assignment, `Merchnum` had **no missing values**.

1.4 Merchant State Imputation:

The merchant state field had 1,037 missing entries after the above steps. To fill these, a multi-step strategy was used:

- **Infer from ZIP:** A dictionary mapping from merchant ZIP code to state was built using transactions where state was known. For special cases (e.g., ZIP codes in Puerto Rico like `00926` appearing as `926.0` due to numeric conversion), entries were manually mapped to the correct state (e.g. `'PR'` for Puerto Rico). Using this map, any missing `Merch state` with a known `Merch zip` was imputed with the corresponding state.
- **Foreign vs Unknown:** Any state values not recognized as valid U.S. state codes were labeled as `'foreign'` (assuming those transactions occurred outside the US). After that, any remaining missing states were filled with `'unknown'`. This ensured that `Merch state` had no nulls and that non-U.S. merchants are distinctly marked.

1.5 Merchant ZIP Imputation:

Merchant ZIP had the most missing values (4,373 initially). The cleaning logic mirrored the state process:

- **Infer from Merchant & Description:** A mapping of Merchant (ID) to a known ZIP and another of Merchant Description to ZIP were created from available data. Missing `Merch zip` were first filled by looking up the merchant's ID, and if still missing, by the merchant's description.
- **Adjustment Transactions:** ZIP for adjustment records was set to `'unknown'` (similar to `Merchnum`).
- **Infer from State:** For cases with a known state but still missing ZIP (e.g., a merchant with state info but no ZIP recorded), the ZIP was imputed with the **most populous ZIP code in that state**. A prepared dictionary of state → largest ZIP (by population) was used for this purpose.
- **Finalize Unknowns:** After these steps, the remaining 533 missing ZIPs (which generally corresponded to merchants with state unknown or foreign) were filled with `'unknown'`. This eliminated all missing values in ZIP, while flagging unresolvable cases as `'unknown'`.

1.6 Logic & Rationale:

Each cleaning decision was made to either preserve useful information or clearly mark the data:

- Placeholder values like `'0'` or blank were converted to proper NA and then imputed rather than dropping records, retaining transactions by assigning them plausible or flag values.
- Merchant identities were preserved by using descriptions and creating new IDs instead of dropping those transactions, ensuring subsequent feature engineering could treat them as distinct entities.
- Geographic consistency was enforced by cross-using state and ZIP data: if one was present it informed the other. This maintains realistic location data.
- Special categories (`'foreign'`, `'unknown'`) were used to capture inherently missing or out-of-scope values rather than imputing incorrect data.
- An anomaly in the amount distribution was noted: a **spike at $3.62**. This was investigated separately. As discussed in class, these correspond to the shipping transactions through FEDEX. Histograms were plotted to see the difference in frequency of amounts before and after these FEDEX transactions were filtered from the dataset. As predicted, the spike around $3.62 has disappeared.

2 Entity Creation and Target Encoding

Before engineering new variables, the dataset was structured around a series of defined **entities** to support feature generation. These entities served as grouping keys for behavioral and temporal aggregations, including:

- `Cardnum`
- `Merchnum`
- `Merch description`
- `Dow` (day-of-week)
- `Cardnum + Merchnum`
- `Cardnum + Merch description`
- `Cardnum + Dow`
- `Merchnum + Dow`
- `Merch description + Dow`
- `Cardnum + Merchnum + Merch description`
- `Cardnum + Merchnum + Merch zip`
- `Cardnum + Merch description + Merch zip`
- `Cardnum + Merchnum + Merch state`
- `Cardnum + Merch description + Merch state`
- `Merchnum + Merch description + Merch state`
- `Merchnum + Merch description + Merch zip`
- `Merchnum + Merch zip`
- `Merchnum + Merch state`
- `Merch description + Merch zip`
- `Merch description + Merch state`

Each of these combinations was filtered to retain only those entity pairs with **at least 10 prior transactions** before a given point in time. This threshold helps prevent overfitting and reduces the influence of sparse or unrepresentative history in later feature calculations.

In parallel, the dataset underwent **target encoding** to convert high-cardinality categorical variables into smoothed numerical indicators of fraud risk. A `TargetEncoder` class was implemented to compute:

- The mean fraud rate per category (e.g., per ZIP or state)
- Smoothed estimates using a weighted average of the global fraud rate and the local (category-specific) rate
- Out-of-fold encoding logic to ensure no row uses its own target value for encoding (avoiding leakage)

This resulted in the following target-encoded features:

- `Merch state_TE` – smoothed fraud probability for merchant's state
- `Merch zip_TE` – smoothed fraud probability for merchant's ZIP code
- `Dow_TE` – smoothed fraud probability for the day-of-week of the transaction

These encoded variables capture population-level risk signals and serve as informative, low-dimensional predictors for the model.

3 Variable Creation Summary

Feature Family	Description and Purpose	# Variables
Card Transaction History	`Card-level velocity features:` For each credit card, transaction counts and amount statistics (average, max, median, total) were computed over rolling time windows (0, 1, 3, 7, 14, 30, 60 days prior). Includes recency features (*day_since* last transaction). These metrics capture the spending frequency, magnitude, and recent activity of the cardholder, highlighting anomalies or sudden spikes indicative of fraudulent behavior.	215
Merchant Transaction History	Merchant-level aggregates including the number of transactions, transaction amount statistics (average, max, median, total), and recency metrics computed over rolling windows (0-60 days). This identifies typical merchant activity and size patterns, crucial for spotting anomalous merchant behavior or transactions that deviate from historical norms.	208
Card-Merchant Interaction History	Historical interaction metrics between a specific card and merchant, including transaction frequency, amounts, and recency. Designed to highlight first-time interactions or unusual activity patterns, these variables serve as strong fraud indicators by identifying irregularities in card-merchant relationships.	126
Card-ZIP Transaction History	Aggregated historical transaction data of cards within specific ZIP codes, detailing counts and amounts over multiple time windows. Useful for detecting geographic transaction anomalies or unusual local spending behaviors of cardholders.	126
Card-State Transaction History	Similar to ZIP-level but aggregated at the state level, these features track transaction counts and amounts, assisting in identifying uncommon or anomalous state-level transaction activity for cards.	126
Merchant-ZIP Transaction History	Transaction aggregates for merchants within specific ZIP codes, capturing typical location-specific transaction volume and amounts. Helps flag transactions deviating significantly from a merchant's usual geographic activity patterns.	154
Merchant-State Transaction History	Aggregation of merchant transactions by state, summarizing typical regional transaction behaviors and facilitating the detection of anomalies related to a merchant's state-level activity.	138
State-Descriptor Aggregate History	Aggregates transaction data by merchant type (descriptor) within states. These features capture overall regional transaction behaviors for merchant categories, assisting in identifying unusual spending patterns or activity anomalies within specific states.	156
Card-Descriptor Interaction History	Historical spending behavior of cards with specific merchant types. These aggregates highlight whether merchant category transactions are consistent or anomalous for each card, flagging unusual spending patterns indicative of potential fraud.	138
Merchant-Descriptor Combination History	Aggregated metrics for merchants, explicitly including merchant type context, reinforcing typical merchant behaviors and transaction patterns, aiding in anomaly detection.	132
Merchant Descriptor Aggregate Patterns	Broad transaction patterns of merchant types across various times and locations, identifying category-based anomalies or deviations from typical merchant descriptor behaviors.	235
Card, Merchant, ZIP Interaction	Granular features capturing historical card interactions with specific merchant locations (ZIP code), designed to flag unusual geographic card-merchant interactions or anomalies at localized levels.	132

Card, Merchant, State Interaction	Metrics aggregating card transactions with merchants within specific states. These detailed interactions assist in identifying transactions where cards engage with known merchants but in unfamiliar or unusual geographic contexts.	107
Card, Merchant, Descriptor Interaction	Detailed interaction history combining cards, specific merchants, and merchant categories, providing nuanced detection of anomalies in transaction patterns, strengthening fraud detection capabilities.	126
Card, Descriptor, ZIP Interaction	Card interaction history with merchant categories within ZIP codes, used to identify localized and category-specific anomalies or uncommon spending behaviors.	126
Card, Descriptor, State Interaction	Card spending behavior with merchant types aggregated by state, highlighting atypical or unusual state-level category interactions for individual cards.	132
Merchant, Descriptor, State Interaction	Transaction history aggregation for merchant types within states, useful in identifying state-specific anomalies in merchant transaction behaviors and activities.	132
Merchant, Descriptor, ZIP Interaction	Merchant-specific transaction metrics aggregated within ZIP codes, aimed at spotting unusual local merchant transaction behaviors or anomalies.	127
Risk Encoding Features	Target-encoded historical fraud rates for categorical attributes (state, ZIP, day-of-week). These variables provide smoothed estimates of fraud likelihood based on historical patterns, enhancing predictive modeling accuracy.	3
Transaction-Level Features	Fundamental details of transactions including transaction amount and weekend indicators, offering basic contextual information critical for initial fraud risk assessment.	4
Distance-Based Features	Geographic metrics measuring distances between sequential transactions and flags indicating unusual distances, highlighting potential improbable travel scenarios or geographic anomalies indicative of fraud.	2
Unsupervised Anomaly Features	Unsupervised machine learning-derived anomaly scores (latent factors) for cards and merchants. These features identify atypical transaction behaviors or patterns without using label information directly.	2
Heuristic Anomaly Flags	Simple rule-based indicators flagging high-value transaction anomalies or inconsistent geographic transaction information, enabling rapid identification of clearly suspicious activities.	2
Benford's Law Features	Statistical deviation metrics from Benford's law expected distribution for transaction amounts. These features serve as forensic indicators, highlighting potential numerical manipulation or fraud-related anomalies.	$\sim$40

Notes: The *temporal and velocity features* detect surges or lulls in activity and repeat behaviors. The *amount features* capture monetary deviations. *Entity-based aggregates* establish baseline behaviors. *Day-of-week and Benford features* uncover cyclical or mathematical anomalies. *Target encoding* introduces broader population-level patterns.