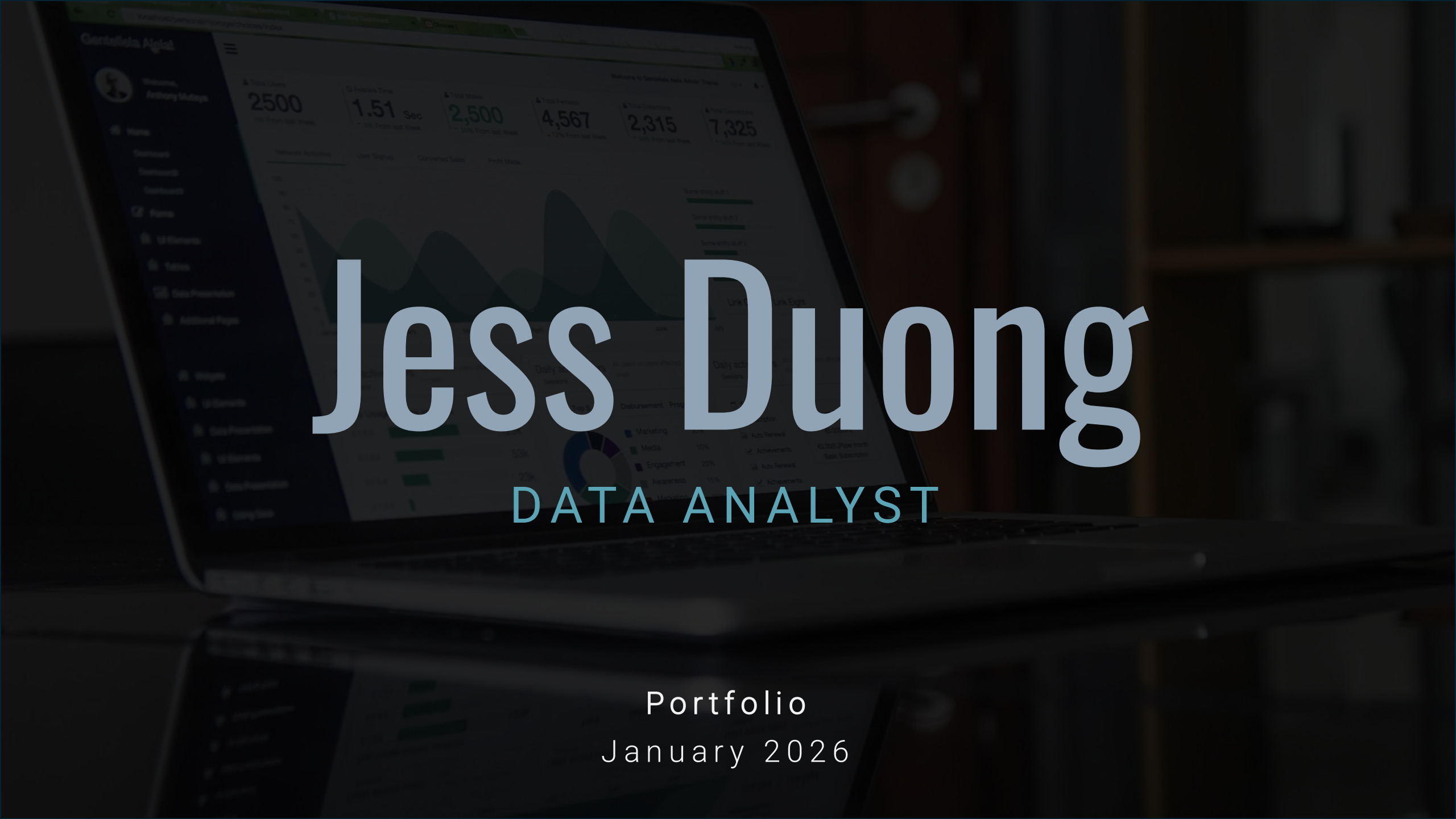


# Jess Duong

## DATA ANALYST

Portfolio

January 2026

# PROJECT 5: Data Ethics & Analytical Foundations

## Excel I Statistical Analysis I Time Series Forecasting I Data Ethics Framework I CRISP-DM I Decision Trees

### Business Problem

Pig E. Bank's compliance analytics department faced critical data ethics violations and needed foundational analytical capabilities across multiple initiatives.

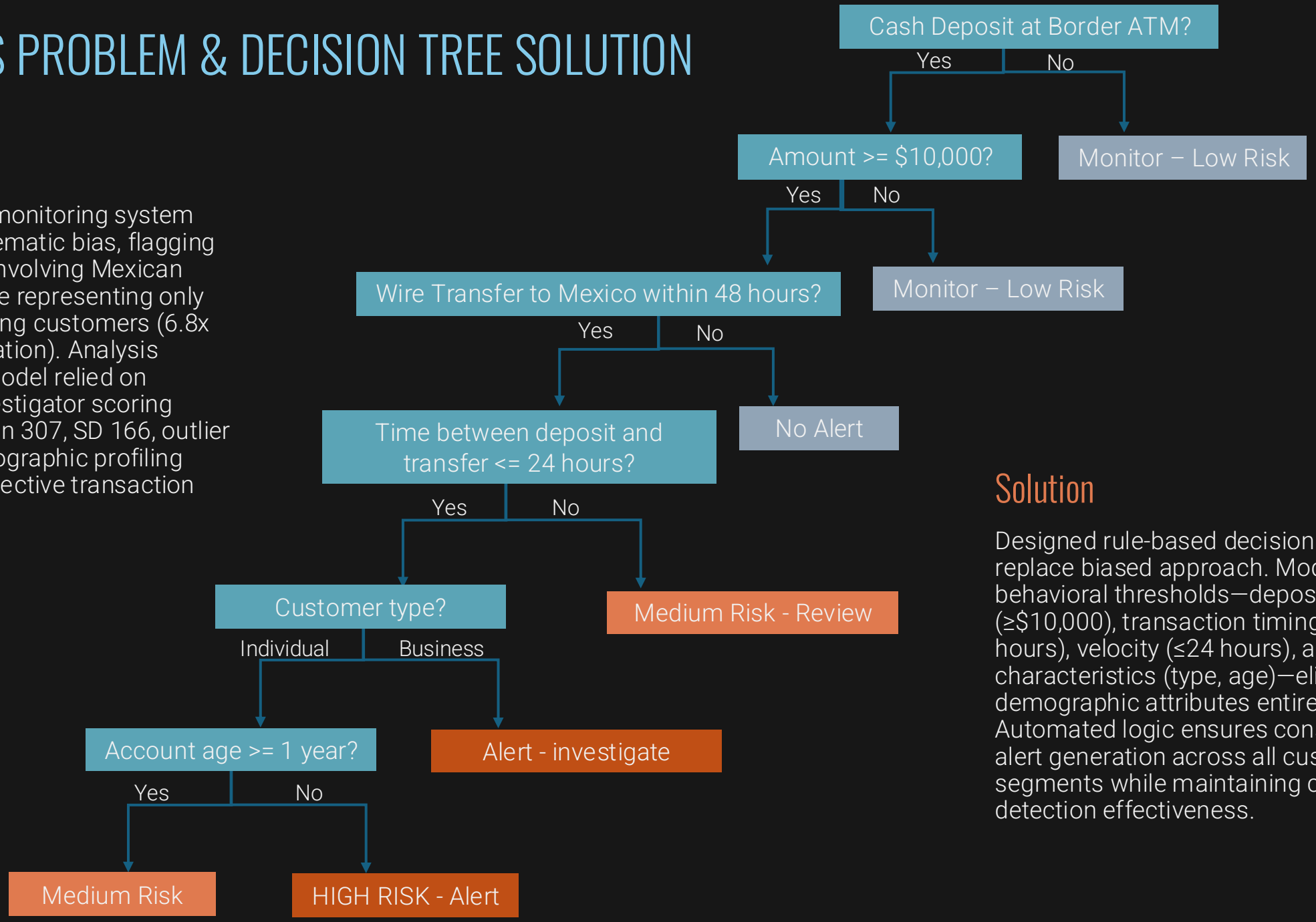
### Key Problems

- **Systematic Bias in AML Monitoring:**
  Existing model flagged 75% alerts involving Mexican citizens despite representing only 11% of qualifying customers (6.8x over representation). Root causes: geographic collection bias (100-mile border restriction), unvalidated sample data, investigator scoring variance (mean 307, SD 166, outlier 759).
- **Data Security & Privacy Gaps:**
  Investigators photographing PII screens; foreign outsourcing proposal involving military customer data; inadequate device policies.
- **Analytical Needs:**
  Customer churn risk analysis for sales team; time series forecasting for financial planning and inventory optimization.

# BUSINESS PROBLEM & DECISION TREE SOLUTION

## Problem

Existing AML monitoring system exhibited systematic bias, flagging 75% of alerts involving Mexican citizens despite representing only 11% of qualifying customers (6.8x overrepresentation). Analysis revealed the model relied on subjective investigator scoring (variance: mean 307, SD 166, outlier 759) and demographic profiling rather than objective transaction behavior.




## Solution

Designed rule-based decision tree to replace biased approach. Model uses behavioral thresholds—deposit amount (≥$10,000), transaction timing (≤48 hours), velocity (≤24 hours), and account characteristics (type, age)—eliminating demographic attributes entirely. Automated logic ensures consistent, fair alert generation across all customer segments while maintaining compliance detection effectiveness.

# STRATEGIC SOLUTIONS & OUTCOMES

## Key Findings

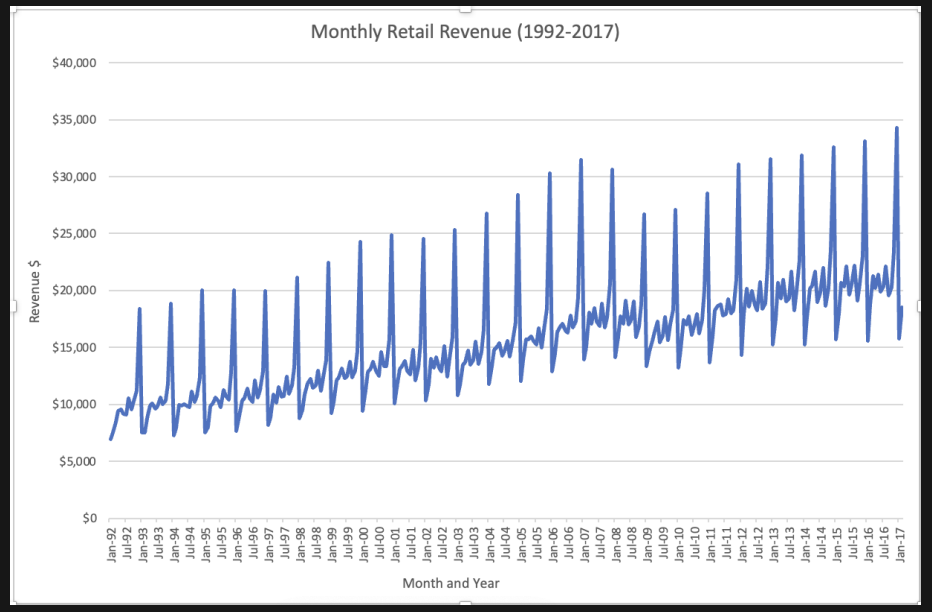
- **Biased Eliminated Through Behavioral Logic:** Decision tree removes demographic profiling; automated thresholds eliminate investigator variance (SD 166 → 0) while maintaining detection effectiveness.
- **Security Vulnerabilities Identified:** PII exposure from screen photography and outsourcing created regulatory violation risk requiring immediate policy changes.
- **Predictable Patterns Enable Action:** Customer churn follows identifiable risk factors (low balances, <1 year tenure); retail value revenue shows seasonal peaks; oil prices smooth ed via 5-year moving average reveal long-term trends.

## Recommendations

- Deploy behavioral decision tree; expand geographic sampling; eliminate outlier score (>2SD); standardize investigator training
- Enforce no-device policies; implement screen-capture prevention; require compliance sign-off for data sharing
- Target high-risk customer segments; use seasonal patterns for inventory/staffing planning

## Business Impact

- **Compliance:** Prevented discriminatory model employment; avoided Fair Lending Act violations
- **Efficiency:** Automated logic ensures 100% consistent alerts; time series enables data-driven planning
- **Security:** Framework prevents PII exposure incidents and regulatory penalties




*Retail sales time series (1992-2017) displays non-stationary pattern with upward trend and annual seasonal peaks, enabling inventory optimization and staffing allocation.*

## Project Deliverables

Full stakeholder presentation available on GitHub

# CONTACT

## Jess Duong


duong.t.jess@gmail.com

inkedin.com/in/jessica-duong-35690847/

github.com/jess-duong

public.tableau.com/app/profile/jess.duong