Franco Lorenzino - ANLT 214 Customer Analytics - Fall 2025

Capstone Project

Customer Segmentation for Proactive Risk Management

The Problem: Our “One-Size-Fits-All” Strategy is Failing

Two core challenges:

1. **Inefficient Marketing** → We are spending money on the wrong customers
2. **Reactive Risk** → We are failing to stop high-risk players until it’s too late

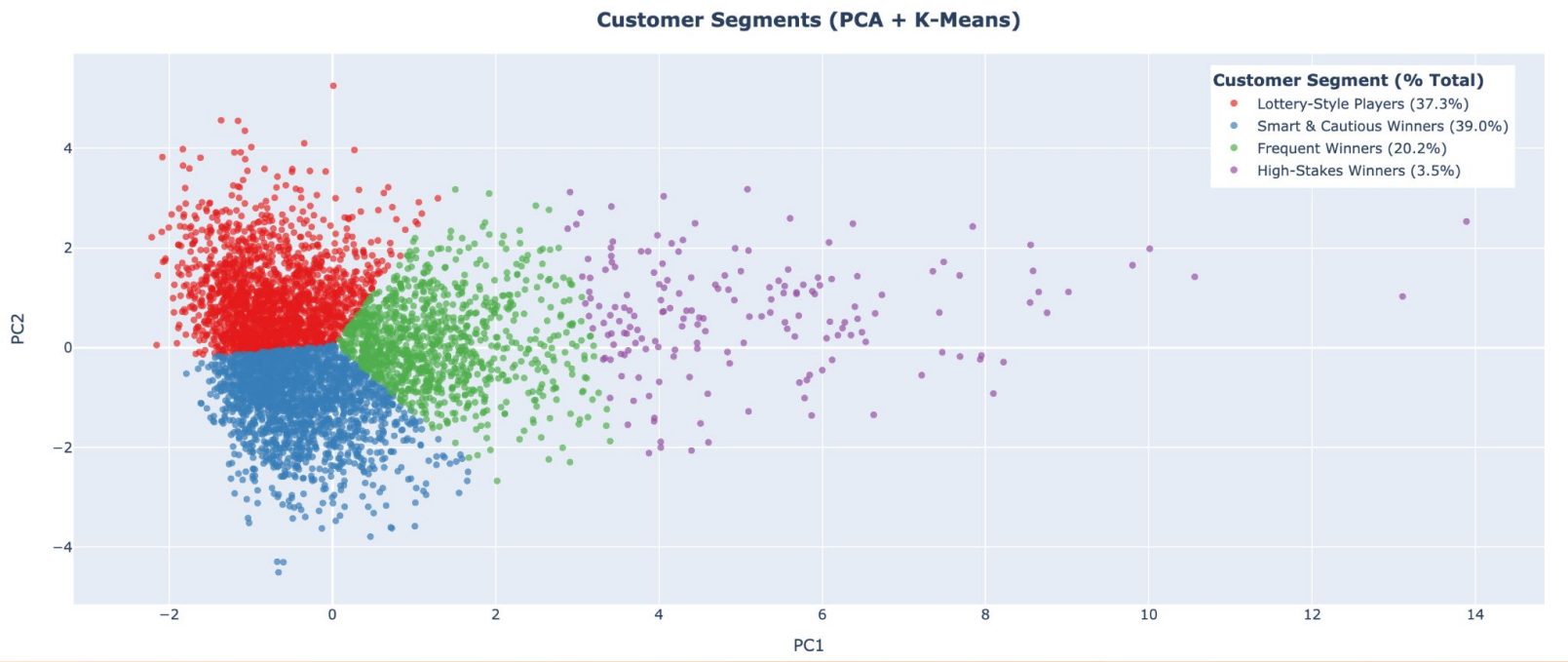
The Goal
Move from a reactive, generic strategy to a **proactive, segment-based** one.

Finding #1: Not All Bettors Are the Same

Even a simple analysis shows 4 distinct betting styles.

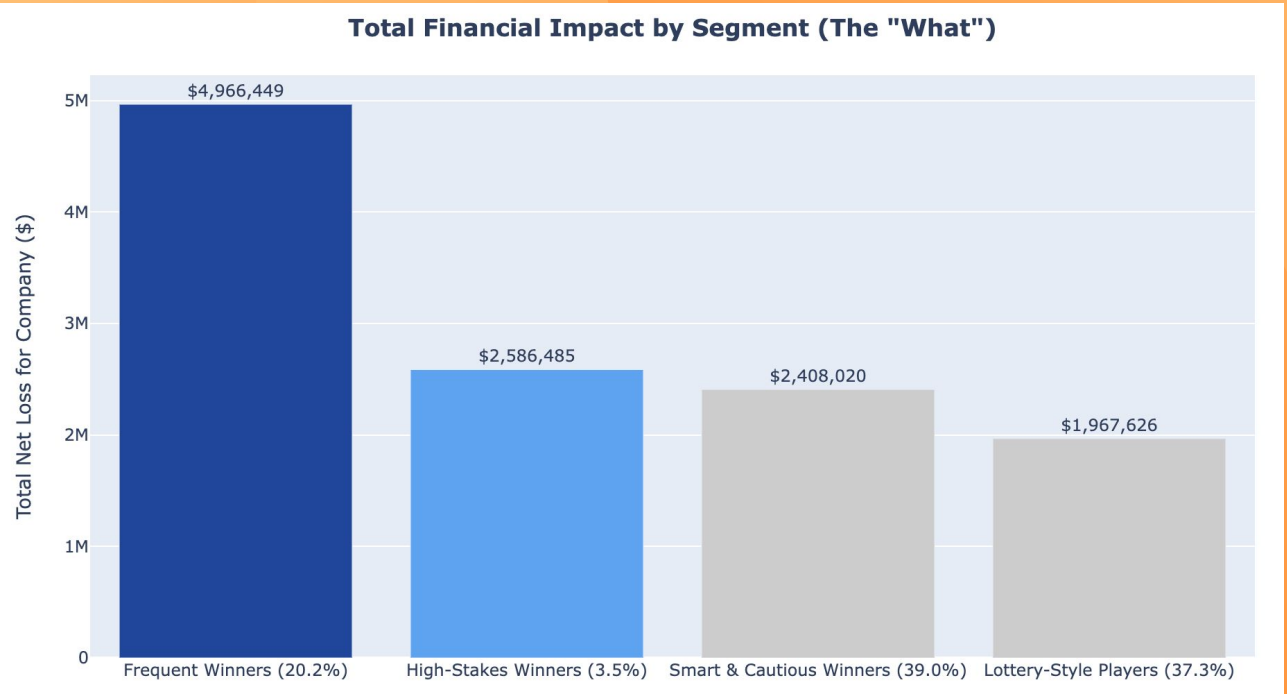
Finding #2: Our Model Identified 4 Actionable Segments

We implemented a K-Means Clustering model to analyze 6 key behavioral features.



Finding #3.1: Our Financial Risk is Dangerously Concentrated

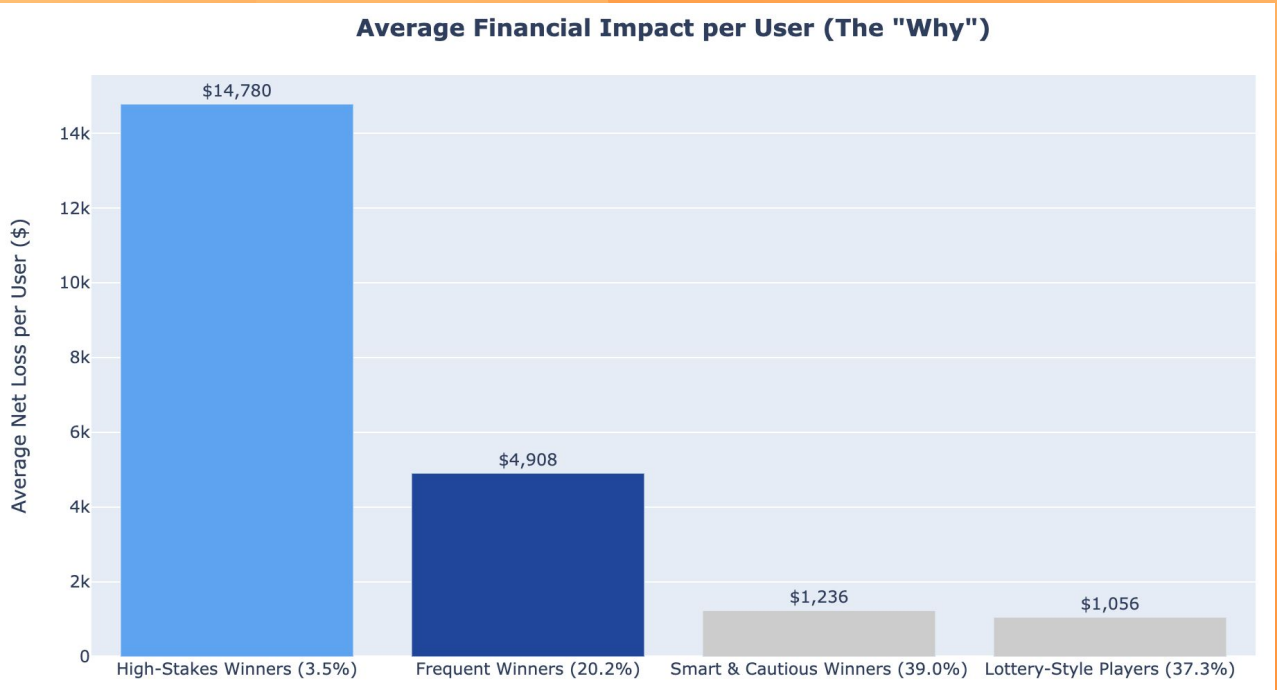
Just two of our four segments drive 63% of all company costs.



The Volume Problem
The “Frequent Winners” segment is our single largest drain, costing us **$4.97** Million in total.

Finding #3.2: The Risk Has Two Different Shapes

The total cost is only half of the story.



The "Risk" Problem
The "High-Stakes Winners" are a time bomb. Each one costs us **$14,780** on average.

From Insight to Action: A Targeted Strategy

High-Stakes Winners

Manage Risk

Apply strict limits and exclude from all marketing lists.

Frequent Winners

Monitor

Shift marketing from bonuses to higher-margin bets (e.g., parlays).

Smart & Cautious

Engage

Use “Bet-Get” campaigns to increase volume without increasing risk.

Lottery-Style Winners

Encourage

Target with all jackpot and parlay promotions.

The Next Step: From Proactive to Predictive

Build a Predictive Model
Use these segments to train a classification model that predicts a new user's segment at sign-up.

Create Real-Time Alerts
Flag "Potential High-Stakes Winners" for the Risk Team before they can cost us $14k, not after.

A/B Test
Prove the dollar value of this strategy by testing it on a control group.

Franco Lorenzino - ANLT 214 Customer Analytics - Fall 2025

Thank You!

Questions?

Appendix: Technical Information

Dataset: https://www.kaggle.com/datasets/emiliencoicaud/sports-betting-profiling-dataset (100,000 transactions → 5,000 customers)

Features Used: total_bets, total_staked, net_gain_loss, avg_odds, win_rate, multiple_bet_ratio

Model: StandardScaler → PCA (2 Components) → K-Means (K=4)

Validation: Elbow Method & Silhouette Score (0.384)