Know Your Clients' Behaviours: A Cluster Analysis of Financial Transactions

Quantitative Finance Seminar

Fields Institute

January 27, 2021

Acknowledgements

- Funding
  - Fields-CQAM Financial Data Analytics Lab
  - MITACS Elevate PDF
  - NSERC
  - Laurier, Western
- Anonymous Data Donor
  - Team members

Acknowledgements

- People
  - Chuck Grace, Ivey Business School, Western
  - John Thompson, Postdoctoral Fellow, Laurier and Western
  - Leon (Longlong) Feng, Research Associate, Laurier and Western
  - Matt Davison, Western
  - Adam Metzler, Laurier
  - Jet (Yuhao) Zhou, Western
  - Lori Weir, Foureyes Financial
  - Kendall McMenamon, Foureyes Financial
  - Philip Patterson, Foureyes Financial

Executive Summary

- Clustered clients using features from a Recency, Frequency, Monetary, Profile behavioural finance model
  - Subject and objective KYC information
  - Trading data
- Client trading behavior not informed by
  - KYC information
  - KYC-inferred Risk Tolerances
- Challenges assumption that risk tolerance and trading behavior are related
- Regulators/advisors should use more advanced techniques to understand clients

Client Onboarding and Suitability

- Retail investors hire financial advisors
  - Select and manage investments
- Financial advisors must provide advice/products that are suitable for each client
- Ontario Securities Commission (2014) describes suitability as a
  - meaningful dialogue with the client to obtain a solid understanding of the client's investment needs and objectives, and to explain how a proposed investment strategy is suitable for the client in light of the client's investment needs and objectives
- Know Your Client (KYC)

Client Onboarding and Suitability

- KYC Information typically collected (questionnaire)

1. Demographic and identifying information
   - Age, gender, occupation, income, etc
2. Subjective financial situation information
   - Questions to asses risk preferences/tolerance
   - Investment goals, time horizon

- Information is used to determine client's risk tolerance
  - Used to determine the suitability of investments for the client

Hypothesis: KYC

- Groups of investors with similar KYC attributes will have similar
  - KYC-inferred Risk tolerances
  - Trading behaviours
- We find
  - Similar KYC gives similar risk tolerance
  - Similar KYC/risk tolerance have different trading behaviours

(Brief) Literature Review

- Suitability and KYC
  - Required in many jurisdictions (US, UK, Europe, AUS,…)
  - Existing research uses
    - objective KYC information to study things besides trading behavior (eg, Moyano and Ross 2017)
  - Little research on using subjective KYC to study trading behavior
    - Picard and de Palma (2010) --- risk tolerance through surveys

(Brief) Literature Review

- Trading Behaviour
  - Zahera and Bansal (2018), Chen et al 2007, Barber and Odean 2008
  - Use different tools/datasets and/or have different focus
- Machine Learning in Finance (Rundo et al 2019)
  - Portfolio construction (Emerson et al 2019)
  - Financial distress prediction (Huang and Yen 2019)
  - Banking risk management (Leo et al 2019)
  - Credit card fraud, money laundering (Van Liebergen 2017)
  - None study retail client suitability and trading behaviour

Data Description

- Dealer
  - 30+ years experience
  - 300 advisors, over $5 Billion in assets
- Dataset
  - 23,970 clients with 52,025 accounts
  - KYC information
  - Trade and transaction details
  - January 1 to August 12, 2019

KYC data

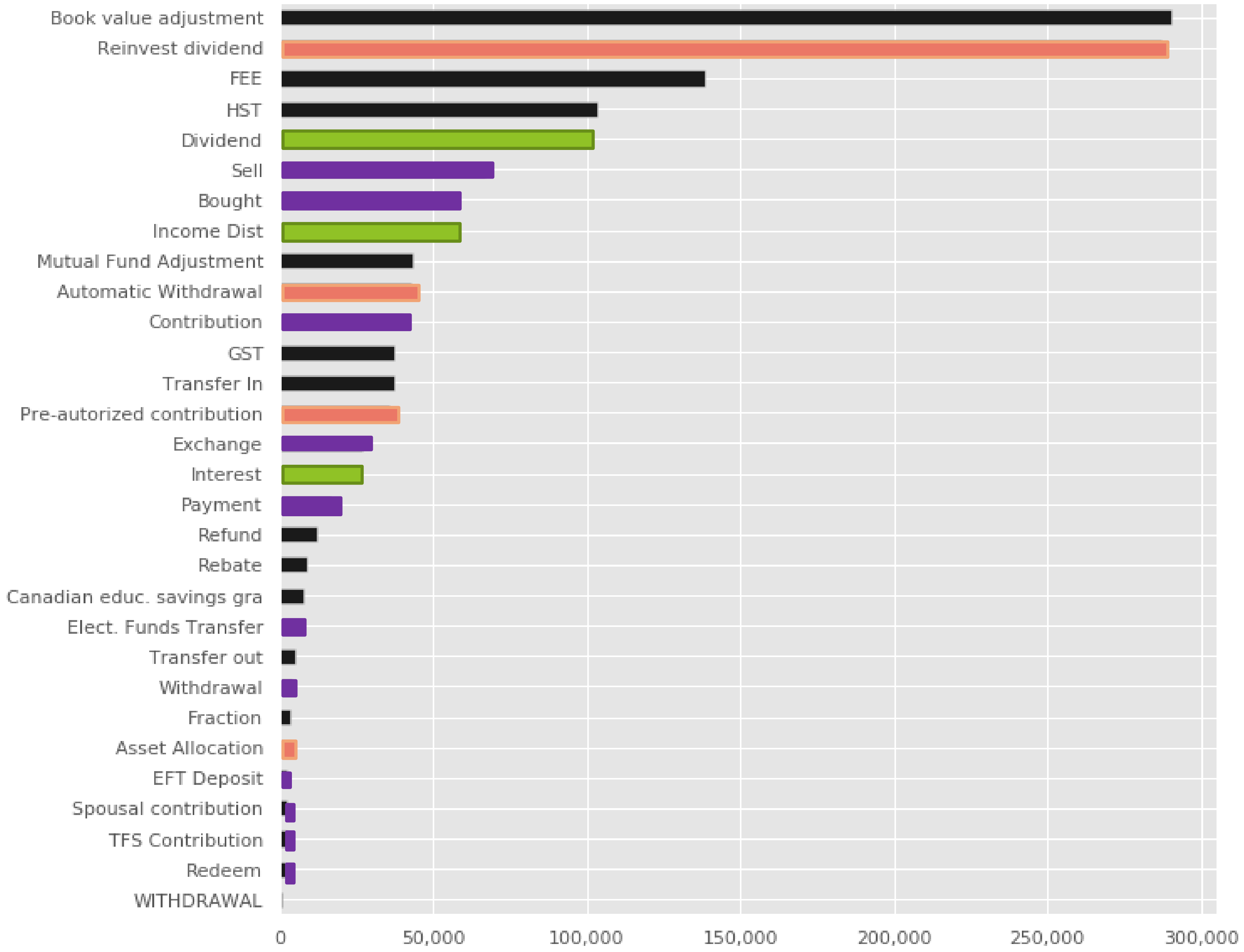
- “Tombstone” Data
  - Age, Gender, Residency, Marital Status
- Retirement Indicator
- Annual Income
- Investment Knowledge (self-reported)
- Number of accounts
- “Profile” features category in RFMP model
- Subjective KYC data not used for clustering features

Feature Engineering

- RFMP Behavioural Finance Model
  - Recency
  - Frequency
  - Monetary
  - Profile (KYC)

- Monetary
  - Types of transactions
    - *Third party initiated*
    - *Systematic*
    - *Periodic*

Types of transactions (Top 30)



Third Party Initiated Trade

- Dividend
- Income dist.
- Interest

Systematic Trade

- Automatic withdrawal
- Pre-authorized contribution
- Asset allocation
- Reinvest dividend

Periodic Trade

- Buy
- Sell
- Exchange
- Contribution
- Payment
- Electronic funds transfer
- Withdrawal
- EFT deposit
- TFS contribution
- Spousal contribution
- Redeem

Features extracted for clustering

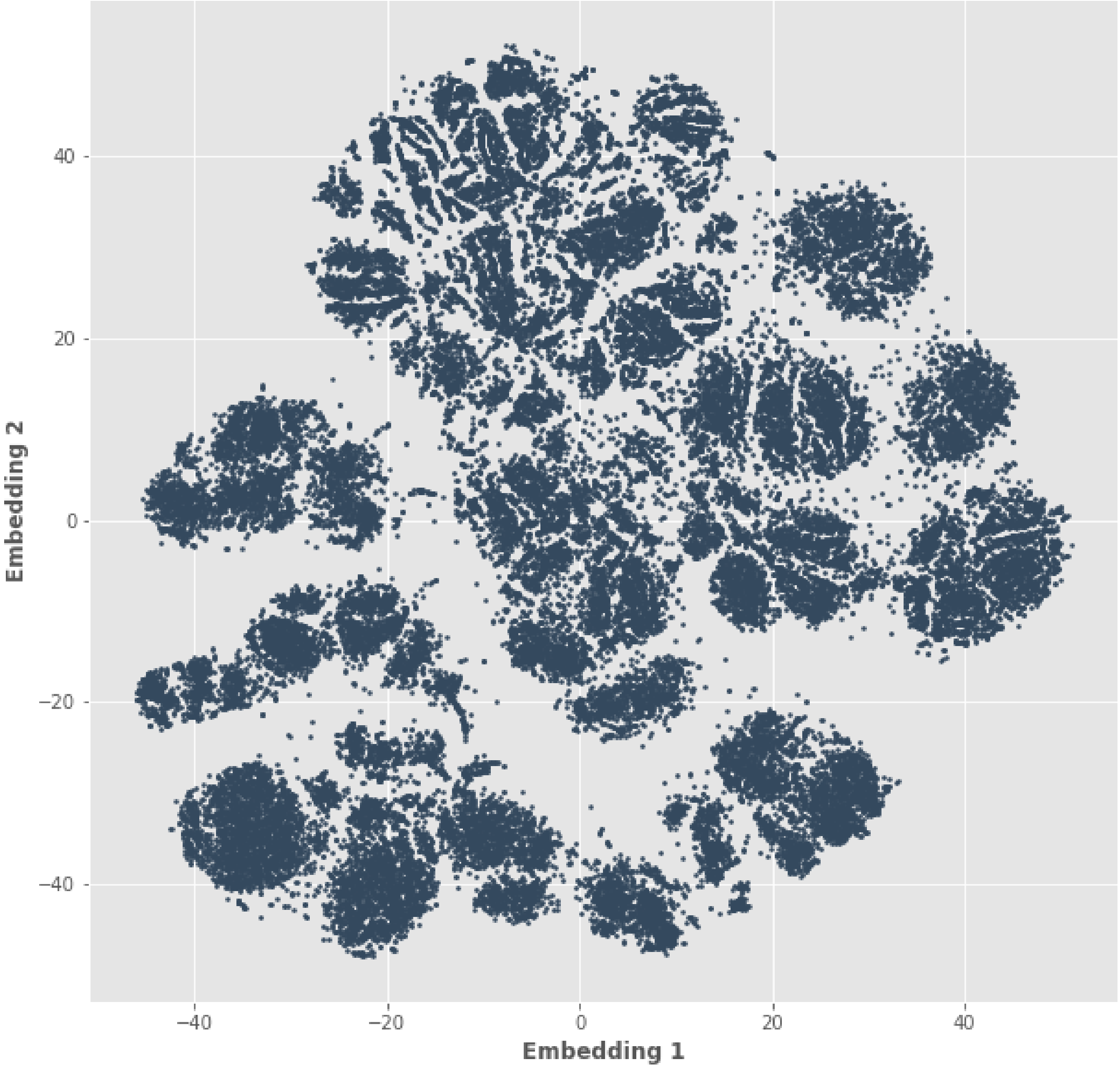
Feature Type	Feature
Recency	Number of days since last trade on record
Frequency	Total number of trades Average number of days between trades
Monetary	Buy and Sell size totals Buy and Sell size minimum and maximum *For each trade type* • Third party initiated • Systematic • Periodic Trade size Variability of trade size
Profile	Age / Gender / Residency / Annual Income / investment knowledge level / number of accounts / marital status / retirement indicator

23 features in total used, appended to client records

Clustering Algorthim

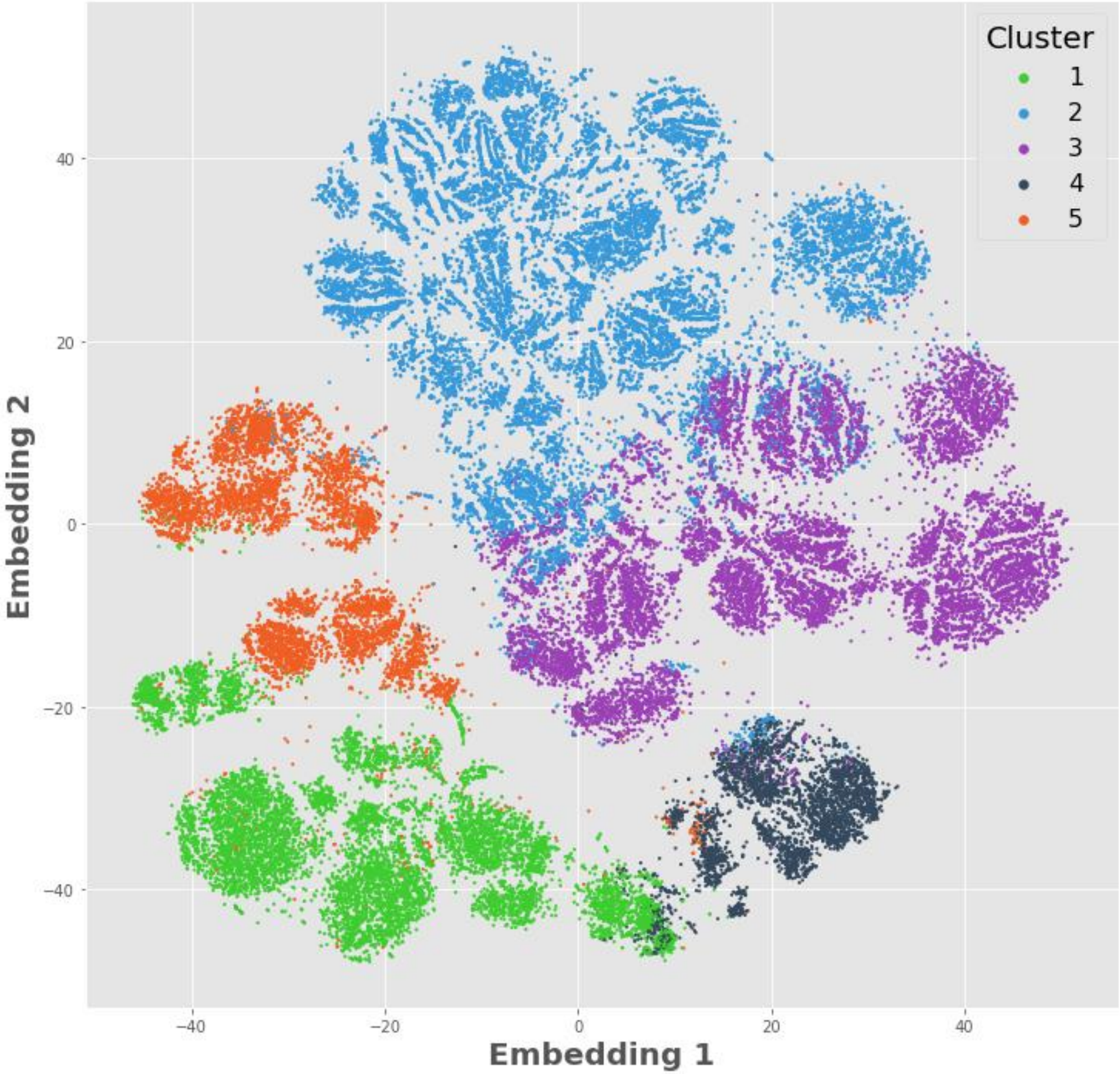
- K-Prototypes
  - Allows for mixed variable types
- Goal: Put clients into groups/clusters such that
  - Each client belongs to one group
  - Clients within clusters are similar
  - Clients in different clusters are dissimilar
- Choose optimal cluster number
  - Silhouette Coefficient
  - DB Index

Clustering (2D Visualization using t-SNE)



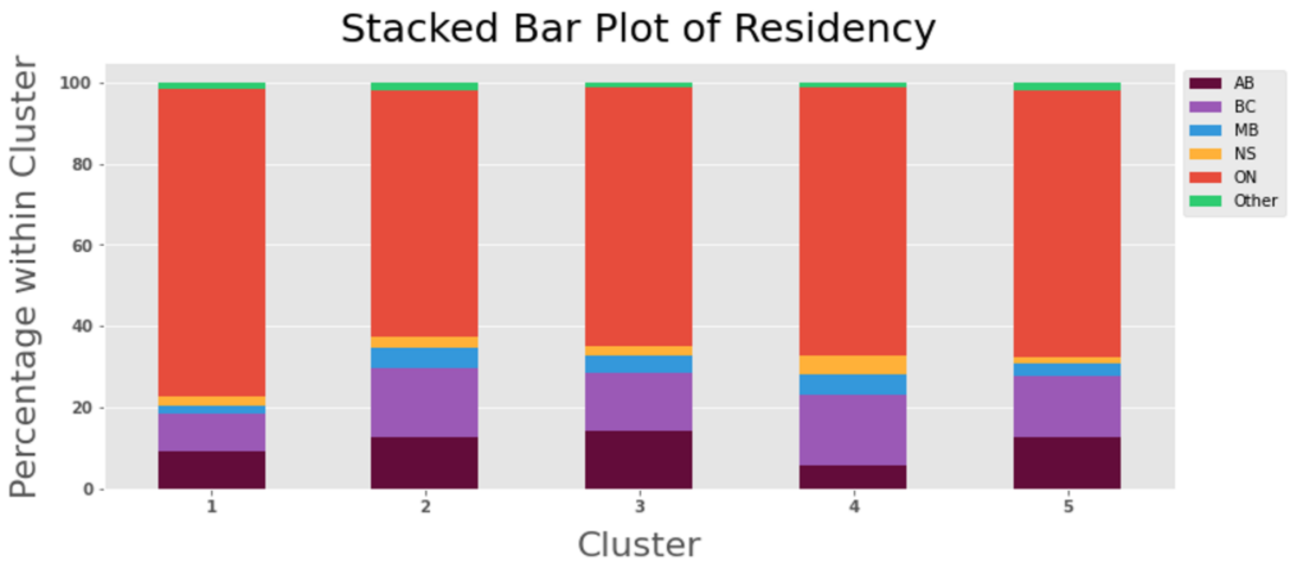
Client Data Projected onto two Embeddings

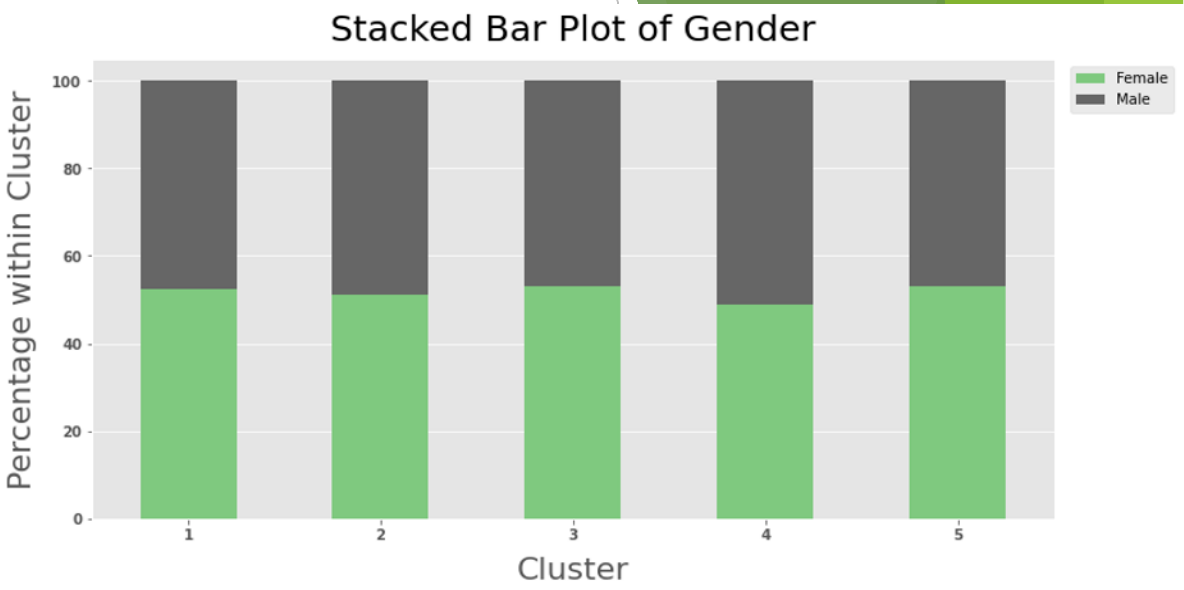
Clustering (2D Visualization using t-SNE)

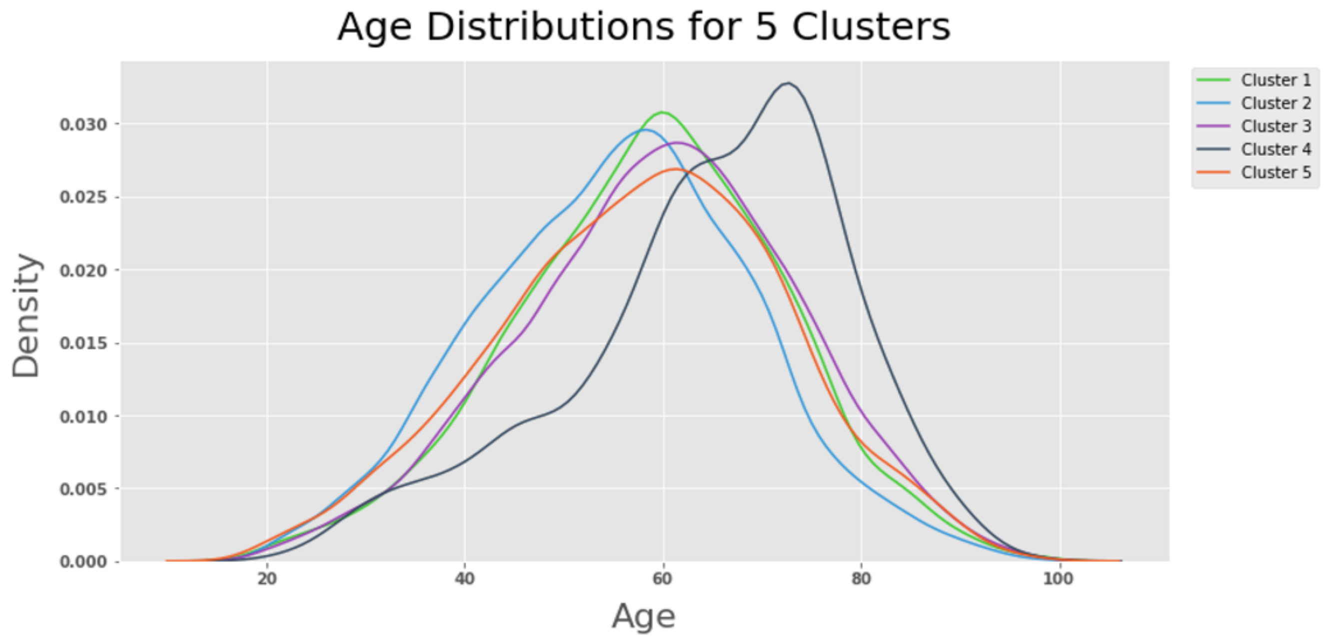


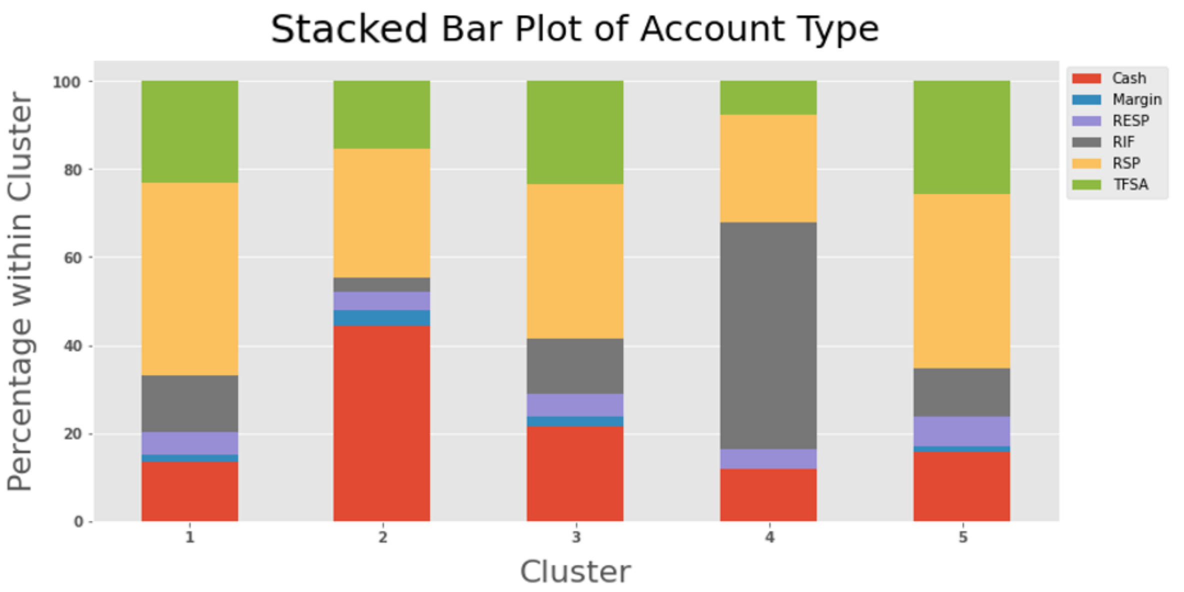
Cluster percentages range from 7% to 36%

Cluster Analysis

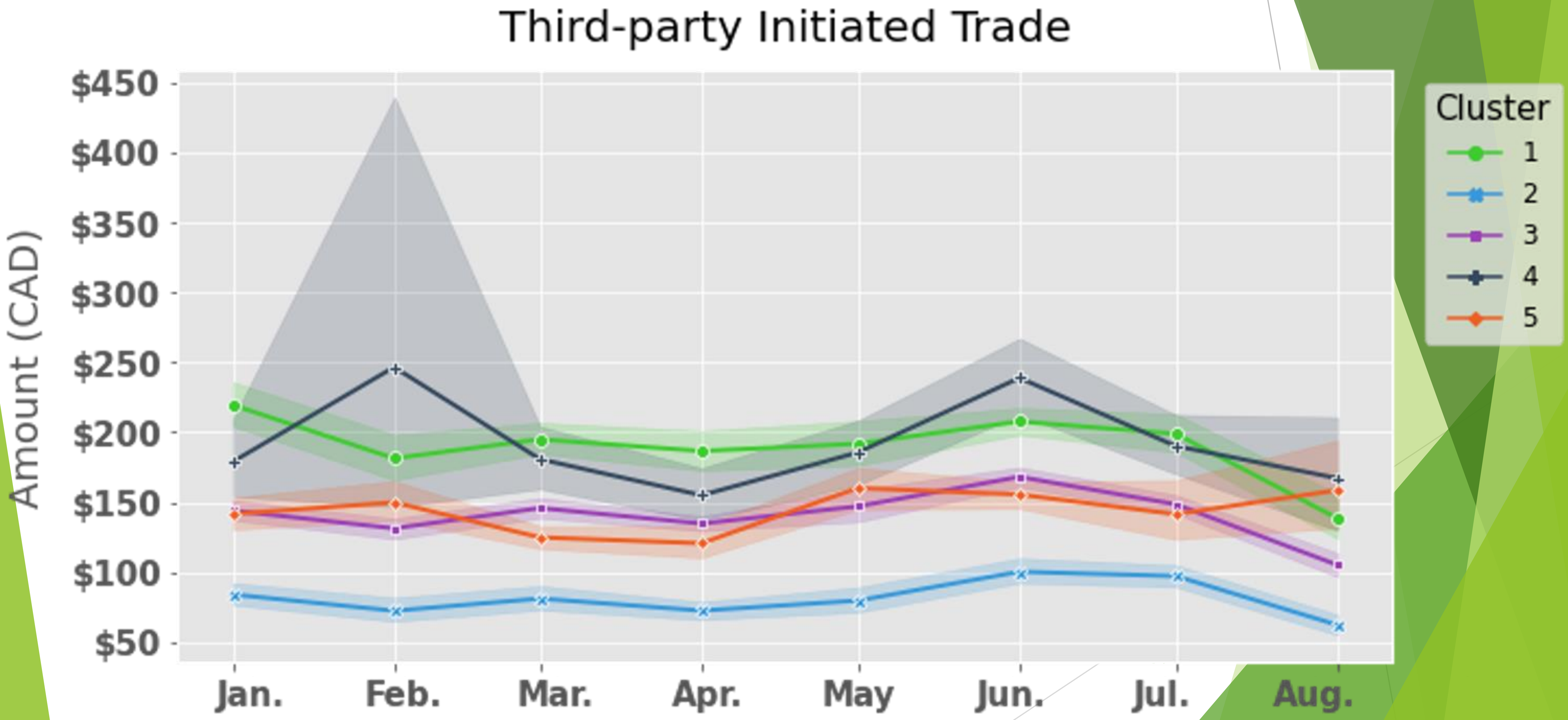




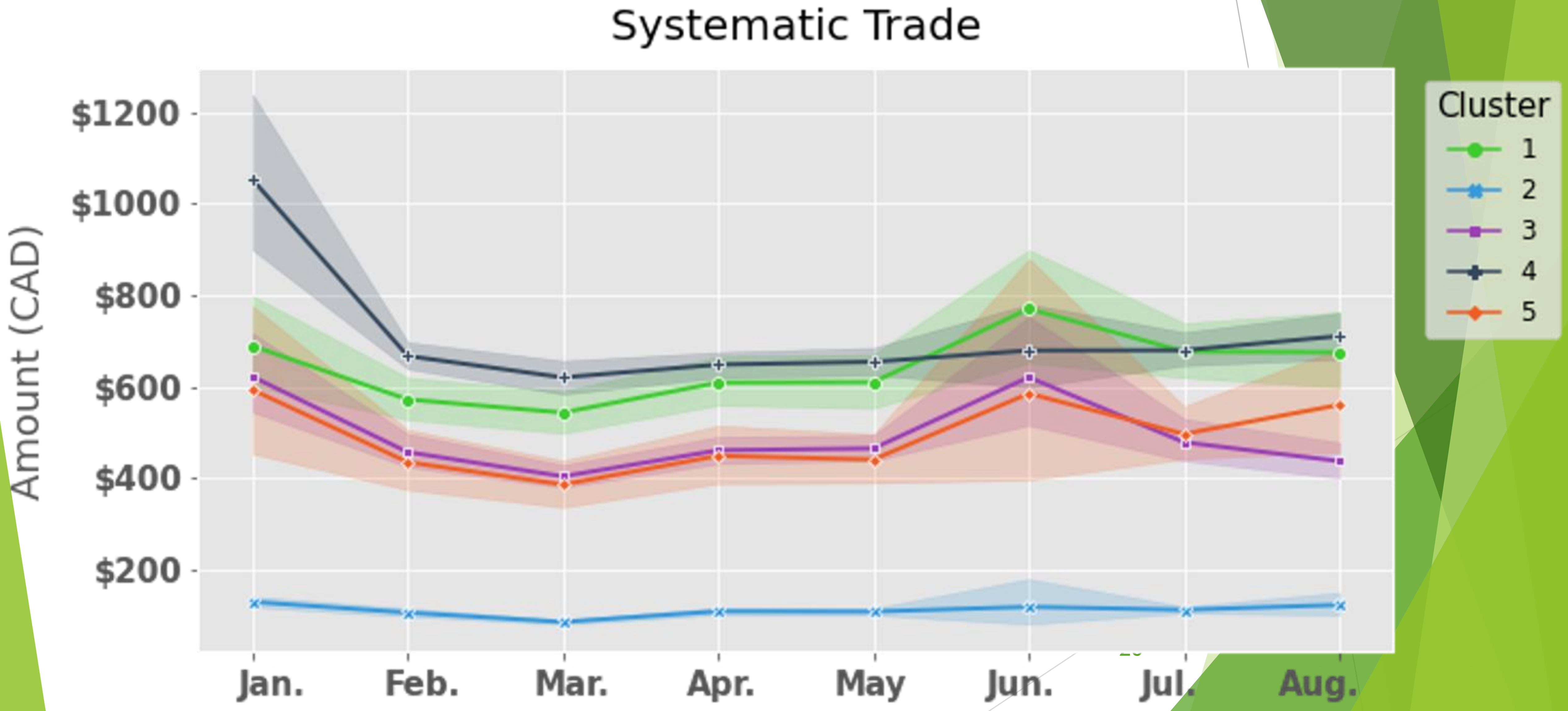




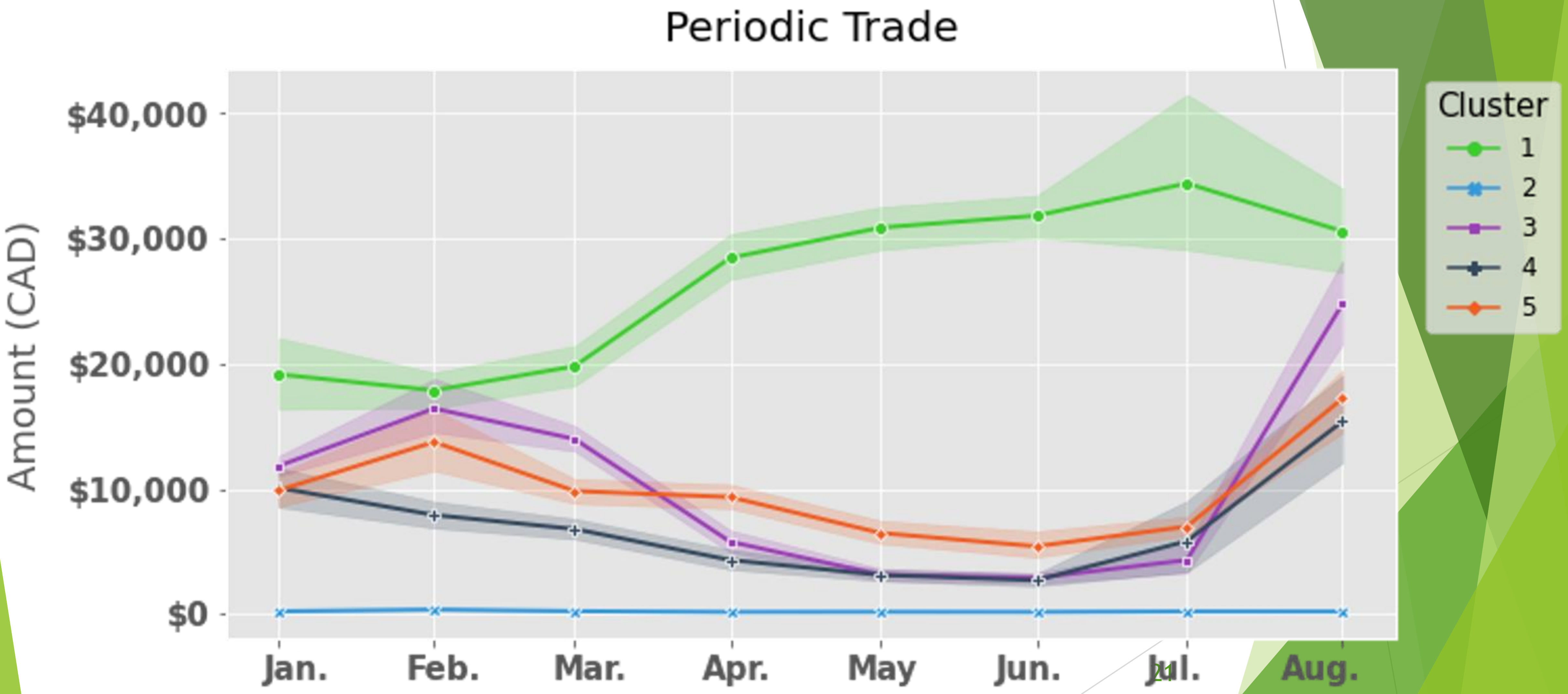
Third Party Initiated Average Monthly Trade Size



Systematic Average Monthly Trade Size



Periodic Average Monthly Trade Size



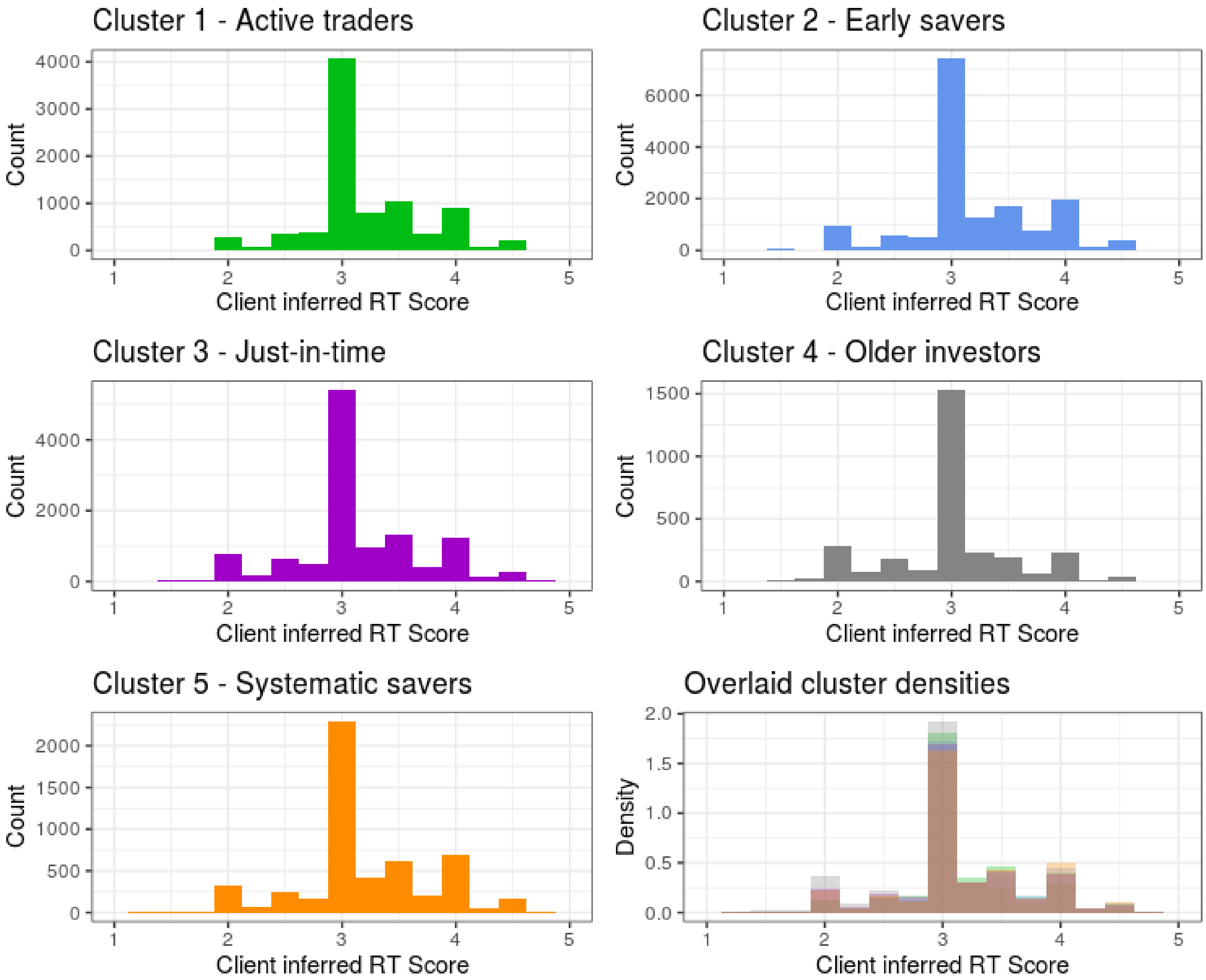
Cluster Analysis Summary

- Features most important in determining cluster membership
  - Monetary
  - Frequency
  - Recency
- Profile features less important

Cluster Personas

Cluster	Nickname	Key Attributes
1	Active Traders 19%	Trade frequently Large amounts Random trading pattern
2	Early Savers 36%	Systematic transactions No active trading Younger
3	Just-in-Time 27%	Trades manually Less frequently, smaller amounts than Active Traders
4	Older Investors 7%	Infrequent trades Systematic or 3rd-party initiated Older, slightly lower KYC-inferred risk tolerance
5	Systematic Savers 12%	Recurrent small systematic trades Large periodic transactions

KYC-inferred Risk Tolerance by Cluster



Distribution of Risk Tolerance same for each cluster

Summary

- Clustered clients using an RFMP model
- Client trading behavior not informed by
  - Subjective and objective KYC information
  - KYC-inferred Risk Tolerances
- Challenges assumption that risk tolerance and trading behavior are related
- Current work looks at trading behavior and asset mix
  - Uses security risk ratings
  - Ties actual risk to KYC-prescribed risk