Clustering Project

Connor & Jagvir

01

Project Flow

Steps Taken

1. Exploratory Data Analysis
    a. Clean & Wrangle
2. Feature Engineering
3. Building Radar & PCA Charts
4. Finding Trends

EDA & Feature Selection

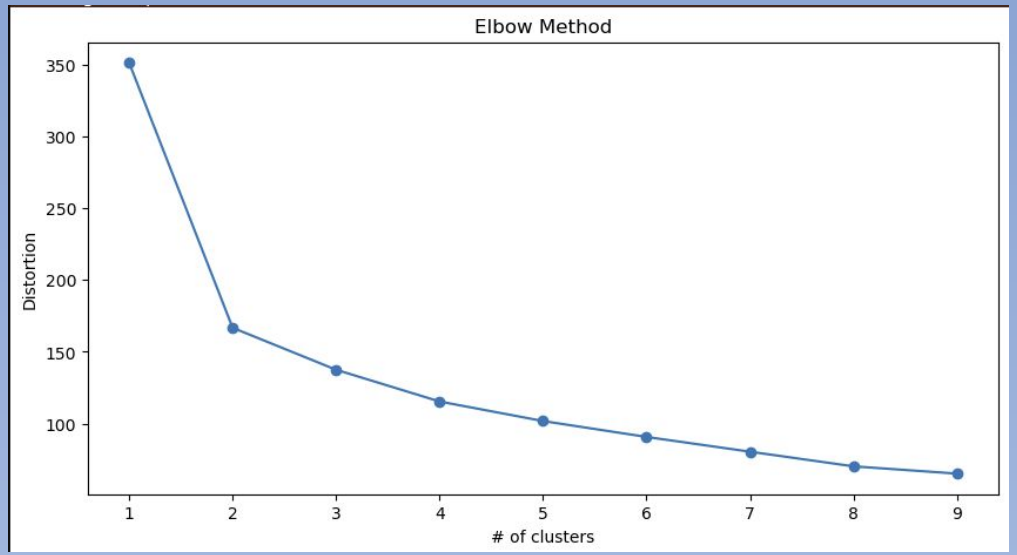
Customer Demographics

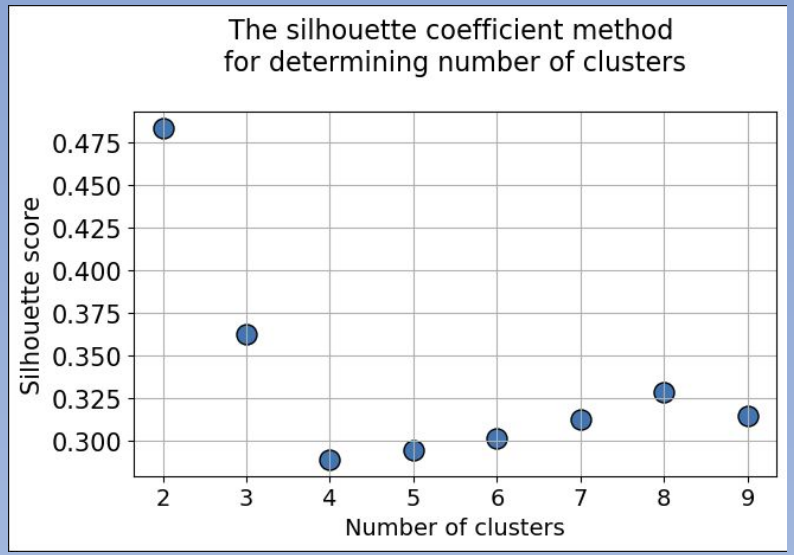
- No null values
- Features:
  - Income, age, years with bank, children, & gender
- Dummy df to show categorical data as numerical
- Used MinMaxScaler()
- 3 clusters from elbow/silhouette method*

Customer Banking Behaviour

- Only Active accounts
- Features:
  - count, subtotal, transactions, est median
- Dummy df
- Used MinMaxScaler()
- 3 clusters from elbow/silhouette method*

Elbow & Silhouette Chart

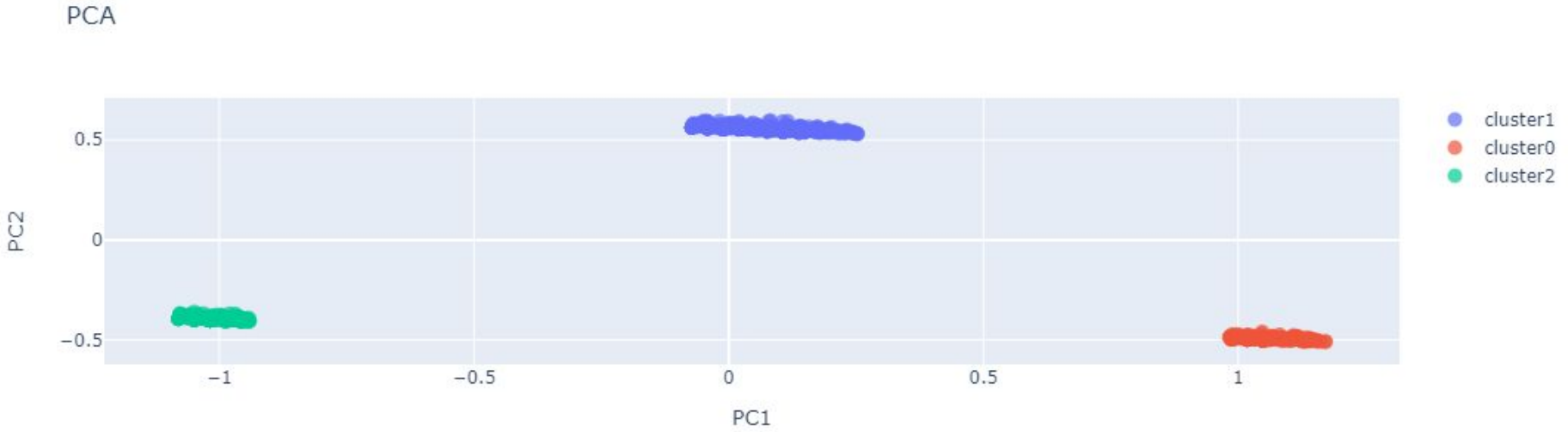




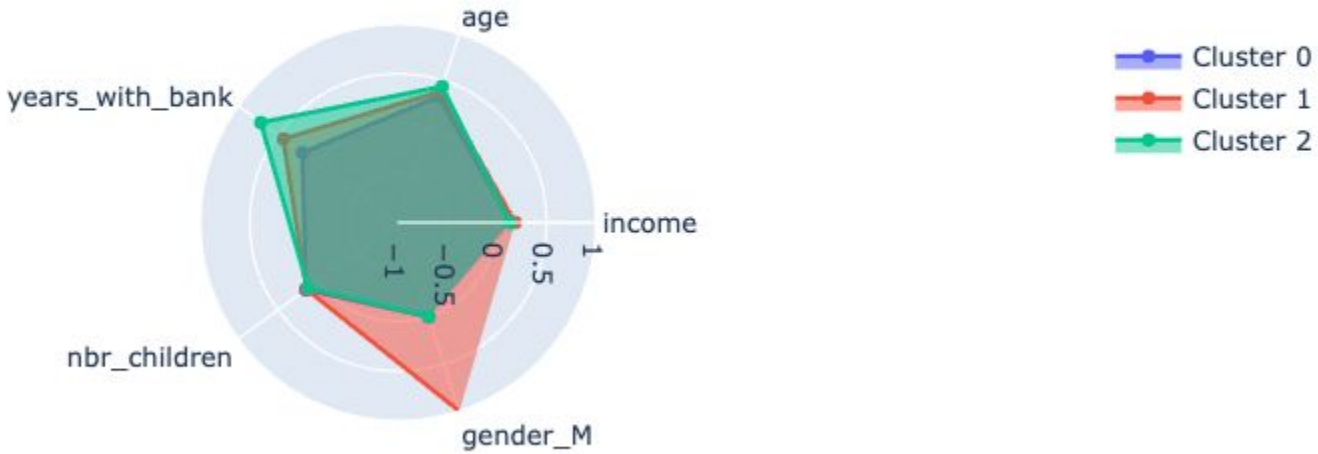
02

Results

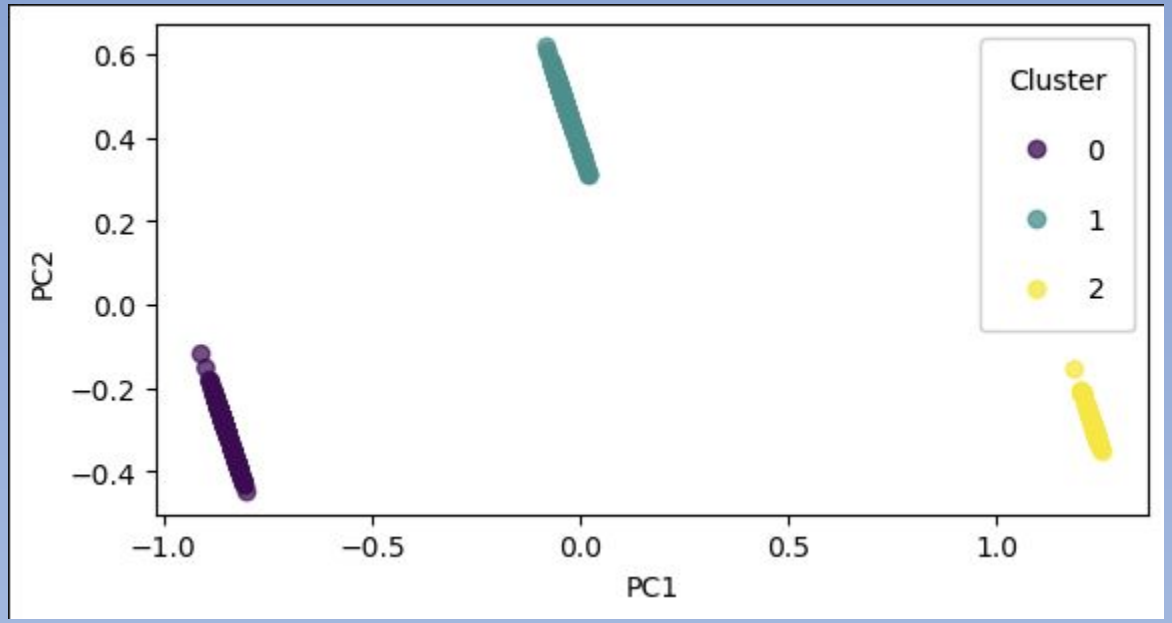
Customer Demo PCA Chart



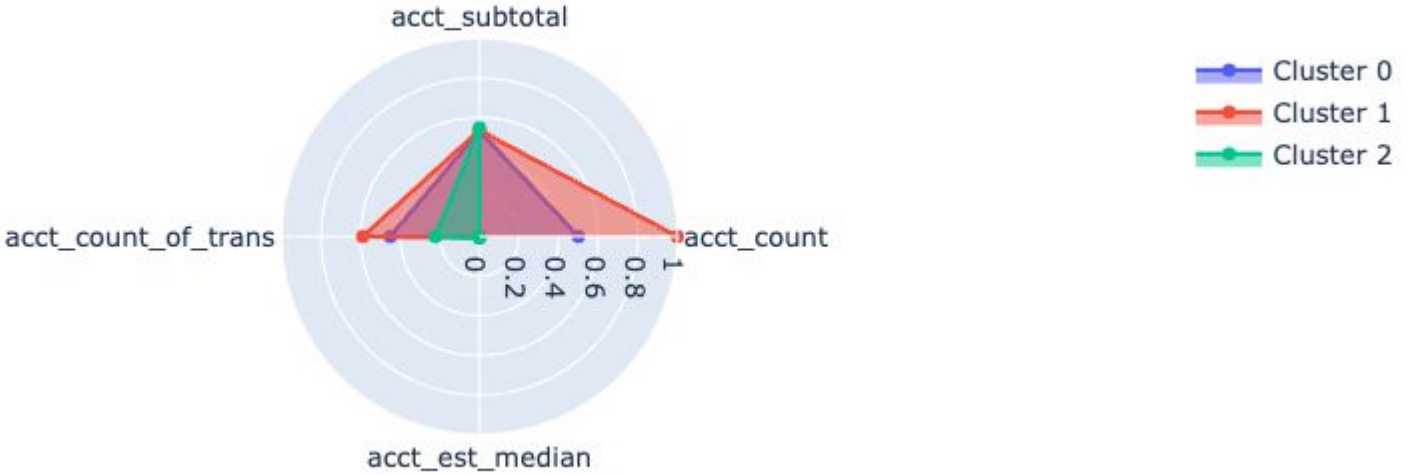
Customer Demo Radar Chart



Banking Behaviour PCA Chart



Banking Behaviour Radar Chart



03 Challenges & Future

EDA & Feature Selection

Challenges

- Wrangling with interest
- Outliers within variables, esp transaction amounts.
- Interpreting the PCA charts

Future Goals

Technical:

- # of clusters: Elbow vs Silhouette method
- Understanding PCA - additional charts
- Understand linearity within charts
- Investigate t-SNE

Practical:

- How customers use their accounts

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**