REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

REVA Academy for Corporate Excellence (RACE)

Credit card Segmentation and Recommendation system

Surendra Tanniru

SRN: R20MBA09

Date: 20/08/2022

PGDM/MBA in Business Analytics

Capstone Project Presentation
Year: I

race.reva.edu.in

REVA UNIVERSITY
Bengaluru, India
Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Agenda

01 Indian Credit Card Market
Back Ground | Current status | Why this study

02 Previous Study on Credit card
Seminal works | Summary | Research Gap

03 Problem Statement
Business Problem | Analytics Solution

04 Objectives of this Study
Primary & Secondary Objectives | Expected Outcome

05 Project Methodology
Conceptual Framework | Research Design

06 Data Collection
Data Collection | Variables

07 Data Preparation
Pre-processing | Process \| Techniques

08 Data Analysis
Univariate | Bivariate | Hypothesis

09 Modeling
Machine Learning | Model Evaluation | Insights

10 Results and Insights
Applications | Demo

11 Conclusions and Suggestions
Insights | Next Step \| Future Scope

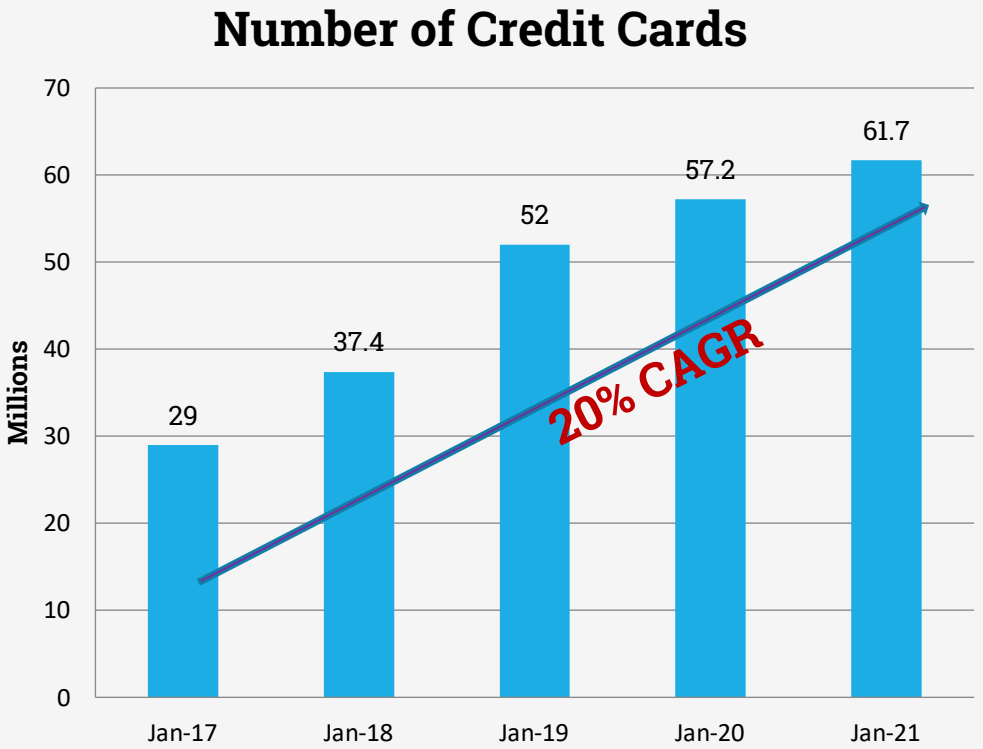
12 Annexure
References | Publications | Plagiarism Score

Indian Credit Card Market

Traditionally, India is a debit card market.

20% CAGR

Credit cards issuance in last 4 years



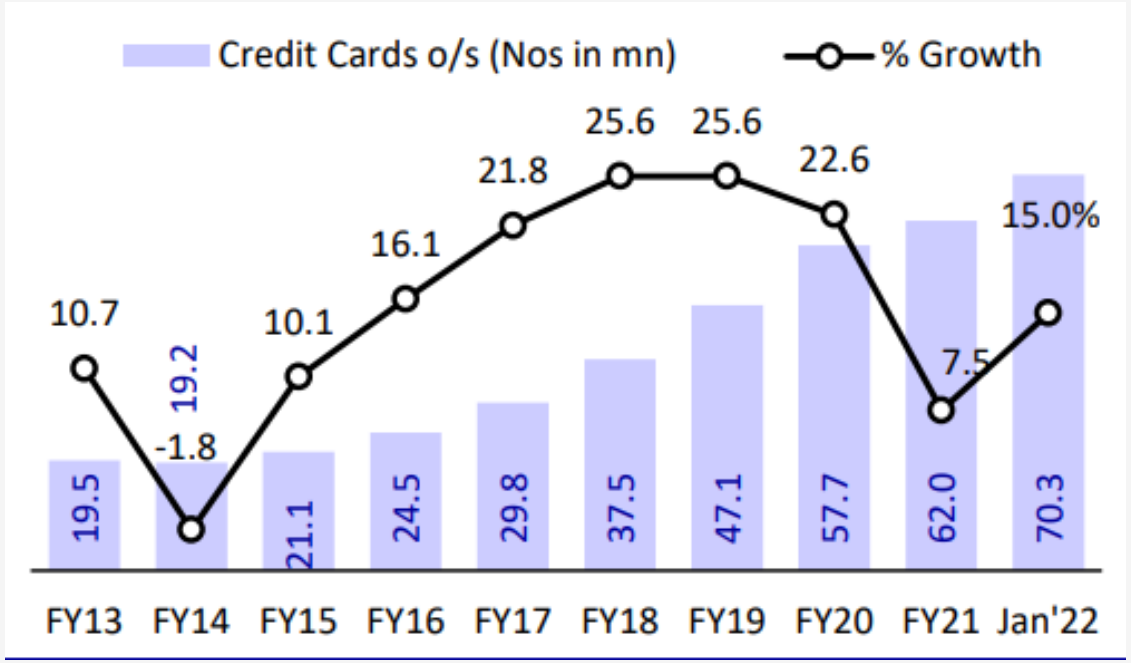
Credit cards during pandemic

15% ↑

Number of Credit cards - YOY

1.3m ↑

New Credit cards in Jan'22



+86% YoY when compared to Jan'21

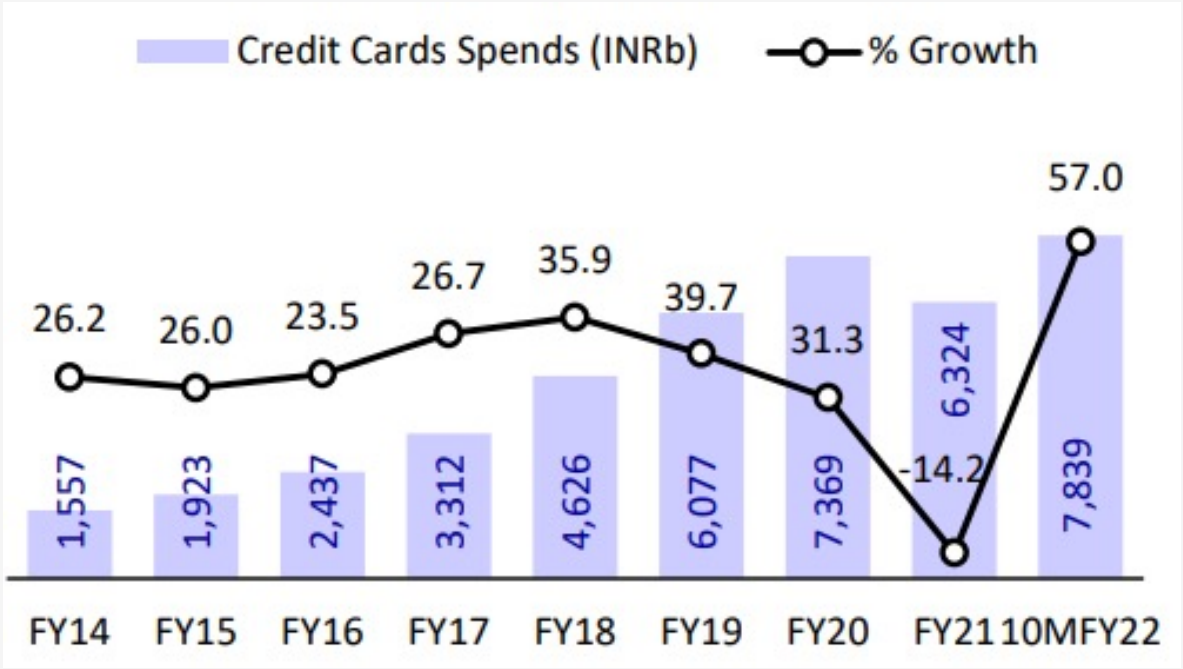
Spends during pandemic

7% ↓
Credit card Spends MOM

35% ↑
Credit card Spends YOY

9% ↓
number of transactions per card reduced to 2.8 MOM from 3.1

Rs.12500
Average monthly spends per card
Increased from Rs.11700



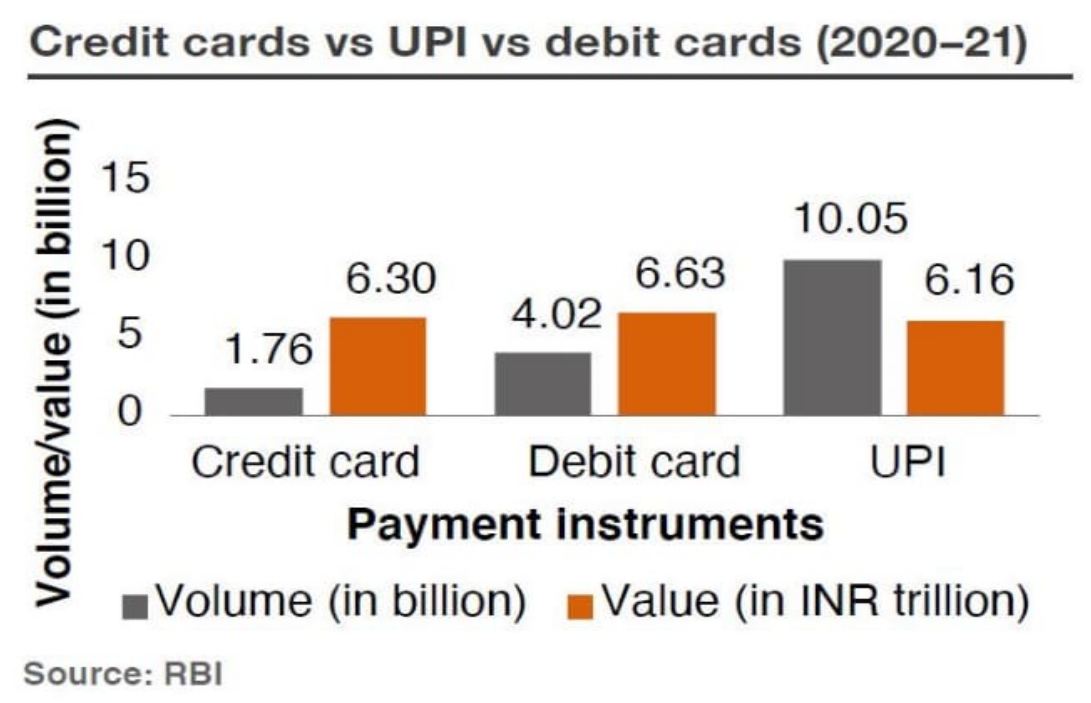
Compare with other payment modes

26.43% ↑

CAGR
Volume of Credit card transactions in next 5 years

39.22% ↑

CAGR
Value of Credit card transactions in next 5 years



Credit cards provide almost same value to that of other payment modes with less volume.

Market Share of Banks

31
Card Issuers

70m
Total number of Cards

80%
6 Card Issuers

Market share (%)	FY18	FY19	FY20	FY21	10MY22
HDFCB	28.5	26.5	25.1	24.2	22.8
SBICARD	16.7	17.6	18.3	19.1	18.9
ICICIBC	13.3	14.1	15.8	17.1	17.9
AXSB	12.0	12.7	12.1	11.5	11.9
CITI	7.1	5.8	4.8	4.2	3.6
RBK	2.1	3.6	4.6	4.8	5.1
KMB	3.9	4.3	4.0	3.9	4.1
AMEX	3.2	3.1	2.9	2.5	2.0
SCB	3.3	2.7	2.5	2.4	1.9
IIB	2.1	2.3	2.3	2.5	2.5
Others	7.7	7.4	7.5	8.0	9.1

Why this study

- Vast availability of credit cards
- Not every card fulfil our needs
- Difficult to select and lack of knowledge
- Save on your spending

What we do

- Analyze the offers and benefits.
- Group the cards with similar offers
- Provide better selection of a card that fulfills most of the customer's needs.
- Recommend a card that suits the customer's needs.
- Help banks to identify and analyze where the user spends and what he looks for. Helps in implementing a target marketing.

Literature Review

Title	Author & Year	Source	Insights	Research gap
Best Offer Recommendation Service	Web Services Group Samsung R&D Institute India, Bangalore, 2016	2016 Intl. Conference on Advances in Computing, Communications and Informatics (ICACCI), Sept. 21-24, 2016, Jaipur, India	Recommend which is the best online aggregator coupons and offers can be used for a payment.	Recommends based on offers provided by online aggregators rather than what credit card provides.
Credit Card Customer Segmentation and Target Marketing Based on Data Mining	Wei Li, Xuemei Wu 2010	2010 International Conference on Computational Intelligence and Security	Based on the real data of a Chinese commercial bank's credit card, the credit card customers are grouped into four classifications by K-means	The data is about customers transactions rather than offers and benefits provided by the credit cards. The clustering is based on customer's income and consumption habits.
Mining and Exploration of Credit Cards Data in UAE	Sarween Zaza, Mostafa Al-Emran 2015	2015 Fifth International Conference on e-Learning	Results indicated how people are grouped based on their income which will help in making region based targeted marketing. Moreover, results revealed how different work sectors use different types of credit cards with regard to their income.	The study is about credit card-holder's behavior in order to predict the market segmentation based on their income

Literature Review

Title	Author & Year	Source	Insights	Research gap
Using data mining predictive models to classify credit card applicants.	Yap Bee Wah, Irma Rohaiza Ibrahim 2010	2010 6th International Conference on Advanced Information Management and Service	Using historical data on payments, demographic characteristics and statistical techniques, credit scoring models can help identify the important demographic characteristics related to credit risk and provide a score for each customer.	The study is to provide a score to the credit card holder based on his historical loans and applications
A Data Mining Approach to Classify Credit Cardholders' Behavior	Aihua Li , Yong Shi , Meihong Zhu 2006	Sixth IEEE International Conference on Data Mining - Workshops (ICDMW'06)	The dataset used is from a major US bank with 65 attributes such as over limit fee, over charge fee and other information etc.,. Used PCA for dimensionality reduction and MCLP for classifying the card holders into Good or Bad customer that identify defaulter.	The study talks about whether a card-holder is a defaulter or not based on his credit card transaction usage.

Literature Review

- The most important techniques used in previous studies are K-means, Decision Tree, and Logistic Regression.
- Some other techniques like Multi-criteria linear programming (MCLP) and Principal Component Analysis (PCA) are used for large datasets.
- Most of the studies on credit cards are based on customer demographics, income, credit risk identification but not on Offers, benefits and terms & conditions that a credit card provides.

Problem Statement

Business Problem | Analytics Solution

- There are huge number of credit cards available in the market which caters different needs of the customers
- Need to consider many factors like Annual Fee, redeem points, cashback. Fuel surcharge, discounts and terms & conditions etc.,
- Check the services it offers and the credit limit.
- Difficult to acquire a best card to fulfill most of our needs.

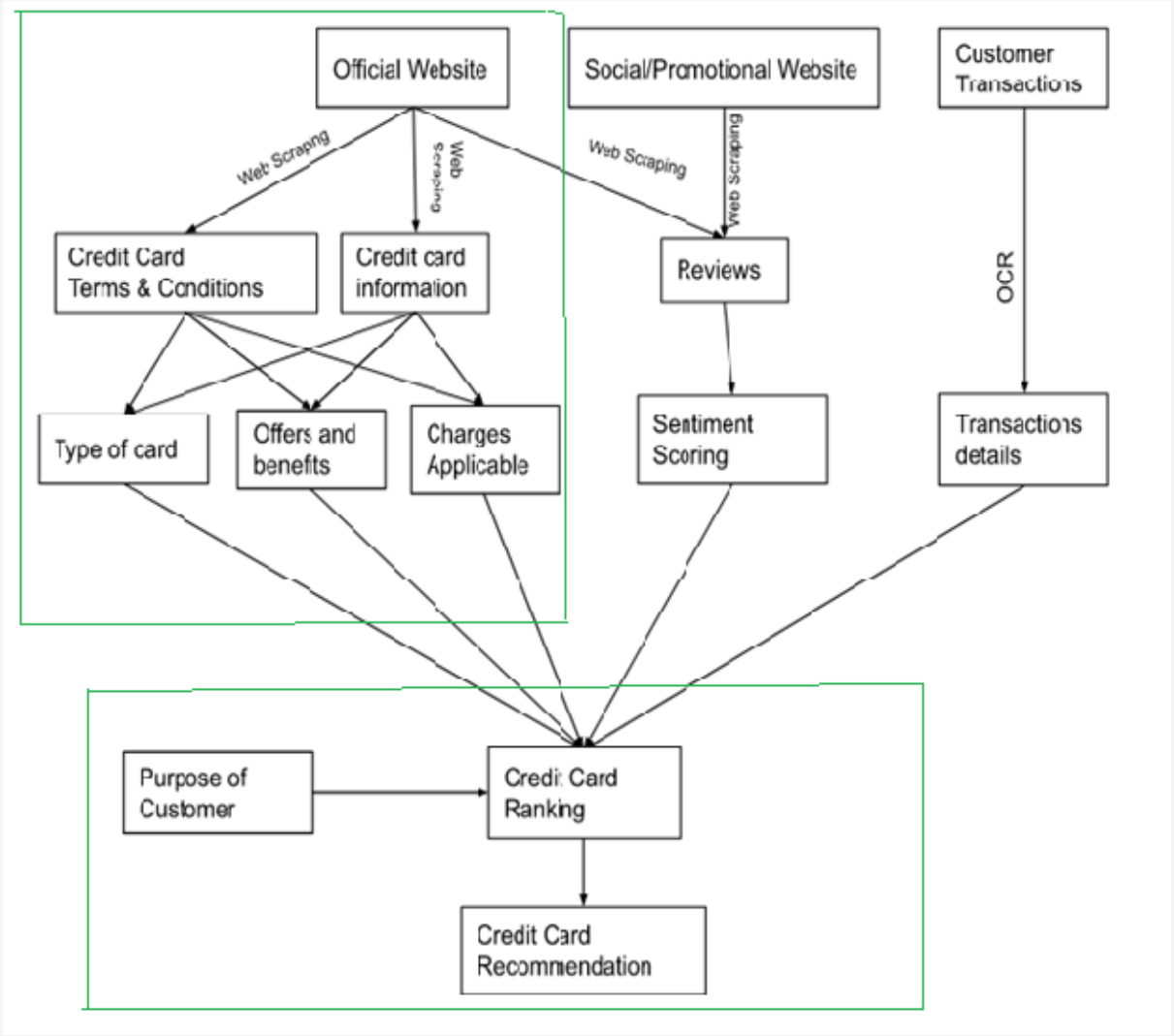
Project Objectives

- Create a corpus of credit card data by Web scraping the Offers and Benefits and terms and conditions of the credit card from official websites.
- Analyze different credit card characteristics.
- Summarize the Offers and Benefits based on different categories/needs.
- Cluster the cards that provide similar set of offers and benefits.
- Build a Credit card recommendation system by identifying the most similar card using KNN.

Project Methodology

The scope of this study is limited to web scraping of the bank websites to collect the offers and benefits of credit cards and group them into clusters.

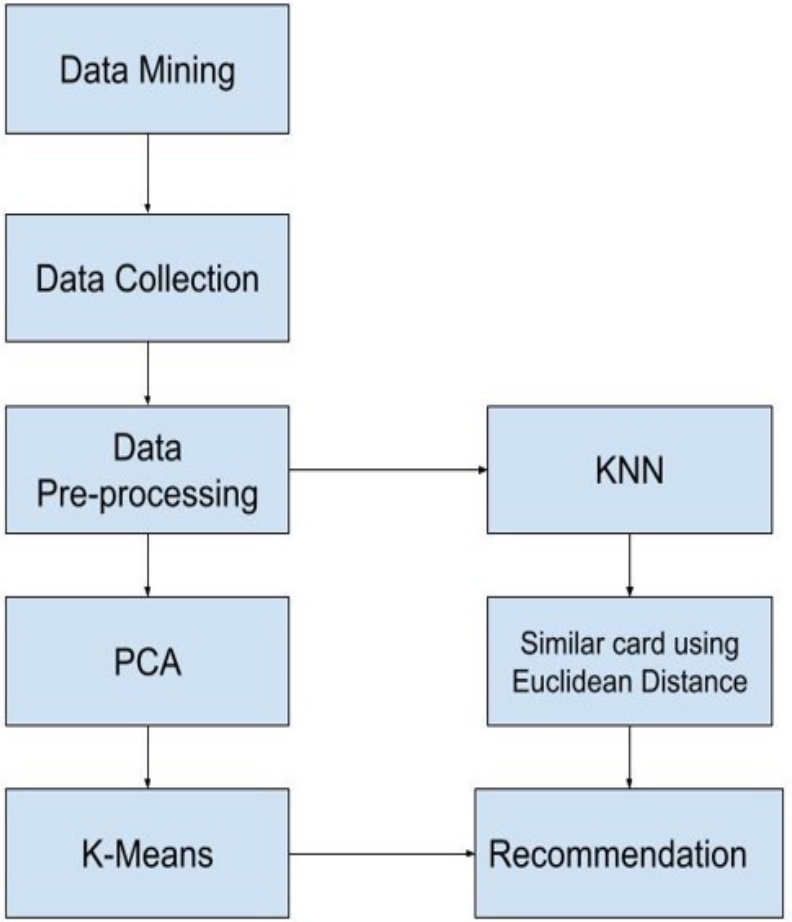
Further, the similarity in cards are used to recommend the most similar card.



System Design

Project Methodology

- Data is collected by scraping card details on different bank websites.
- Data is pre-processed for different features like offers, cashback, terms and conditions etc.,
- **Principal component Analysis (PCA)** is used for dimensionality reductions which reduces the number of features that covers most of the variance in the data.
- **K-means** is applied on the data with the features obtained from PCA to cluster them based on the offers they provide.
- **K-Nearest Neighbor (KNN)** is used to identify the similar card for recommendation. We used Euclidean distance to identify the similar instances between two cards.



Flow chart

Data Collection

Data Extraction Techniques:

- Selenium Web Driver

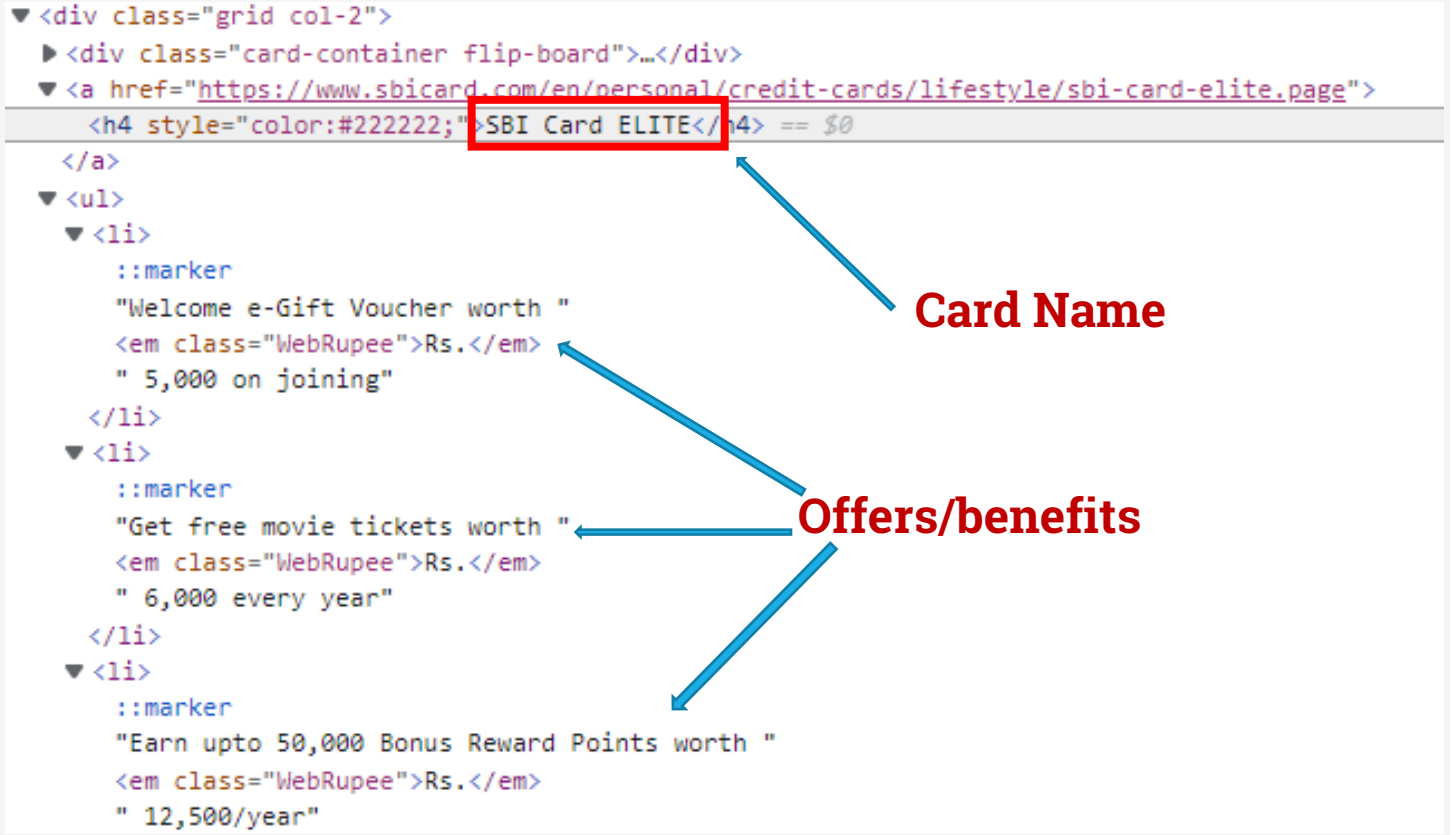
```
import numpy as np
import pandas as pd
import requests
from bs4 import BeautifulSoup

from selenium import webdriver
from selenium.webdriver.common.by import By
from selenium.webdriver.support.ui import WebDriverWait
from selenium.webdriver.support import expected_conditions as EC
from selenium.common.exceptions import TimeoutException
```

```
['https://www.icicibank.com/card/credit-cards/coral-card.page?ITM=nli_cms_CC_index_coral_explore_btn',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/rubyx-card/key-privileges.page?ITM=nli_cms_CC_index_rubyx_knowmore_btn',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/sapphiro-card/key-privileges.page?ITM=nli_cms_CC_index_sapphiro_knowmore_btn',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/Platinum-chip/key-privileges.page?ITM=nli_cms_CC_index_platinum_knowmore_btn',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/makemytrip-cc-index.page?ITM=nli_cms_btn_cc_index_signaturecc_knowmore',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/manchester-index.page?ITM=nli_cms_btn_cc_index_mucc_knowmore',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/fuel-index.page?ITM=nli_cms_CC_index_hpcl_knowmore_btn',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/accelero-card/index.page?ITM=nli_cms_btn_cc_index_accelero_knowmore',
 'https://www.icicibank.com/Personal-Banking/cards/Consumer-Cards/Credit-Card/fuel-index.page?ITM=nli_cms_CC_index_hpcl_knowmore_btn']
```

Data in websites

Data resides in the form of HTML in the browser



Data Preparation

Pre-processing | Techniques

Beautiful Soup is used to retrieve the Card Names and its associated offers and benefits from HTML Tags.

```
def cards_list(card_names):
    for i in range(len(card_names)):
        for c in card_names[i].find_all('img',class_='lazy'):
            card = c.get('title')
            cards.append(card)
    return cards

def offers(offers_ul):
    card_offers = [[] for _ in range(len(offers_ul))]
    for j in range(len(offers_ul)):
        for offer in offers_ul[j].find_all("li",class_='',recursive=False):
            card_offers[j].append(offer.text.strip())
    return card_offers
```

Data Preparation

Pre-processing | Techniques

```
['SBI Card ELITE',
 'SBI Card PULSE',
 "Doctor's SBI Card (in association with IMA)",
 'SBI Card ELITE Advantage',
 "Doctor's SBI Card",
 'SBI Card PRIME',
 'OLA Money SBI Card',
 'Apollo SBI Card',
 'SBI Card PRIME Advantage',
 'Tata Platinum Card',
 'Tata Titanium Card',
 'Paytm SBI Card SELECT',
 'Paytm SBI Card',
 'FABINDIA SBI CARD SELECT',
 'FABINDIA SBI CARD',
 'Lifestyle Home Centre SBI Card PRIME',
 'Max SBI Card PRIME',
 'Spar SBI Card PRIME',
 'Lifestyle Home Centre SBI Card SELECT',
 'Max SBI Card SELECT',
 'Spar SBI Card SELECT',
 'Lifestyle Home Centre SBI Card',
```

List of Cards

List of Offers

```
['Welcome e-Gift Voucher worth Rs. 5,000 on joining',
 'Get free movie tickets worth Rs. 6,000 every year',
 'Earn upto 50,000 Bonus Reward Points worth Rs. 12,500/year',
 'Complimentary membership to Club Vistara and Trident Privilege program'],
['Noise Color Fit Pulse Smartwatch worth Rs. 4,999 on payment of joining fees',
 '12 Month Membership for FITPASS and Netmeds First Pass on payment of joining fee and card activation',
 '5X Reward Points on Chemist, Pharmacy, Dining and Movie Spends',
 'E-Voucher worth Rs.1,500 on achieving Retail spends of Rs. 4 Lakh in one annual year'],
['Professional Indemnity Insurance cover of Rs. 20 Lakhs',
 'e-Gift Voucher worth Rs. 1,500 on joining',
 '5X Reward Points on Medical Supplies, Travel Bookings, International Spends and Doctors' Day',
 'e-Gift Voucher worth Rs. 5,000 on annual spends of Rs. 5 Lakhs'],
['Welcome e-Gift Voucher worth Rs. 5,000 on joining',
 'Get free movie tickets worth Rs. 6,000 every year',
 'Earn upto 50,000 Bonus Reward Points worth Rs. 12,500/year',
 'Complimentary membership to Club Vistara and Trident Privilege program'],
['Professional Indemnity Insurance cover of Rs. 20 Lakhs',
 'e-Gift Voucher worth Rs. 1,500 on joining',
```

Data Preparation

Pre-processing | Techniques

The list of Cards and the lists of Offers and benefits are mapped together into a dataframe

	Cards	Offer1	Offer2	Offer3	Offer4	Offer5	Offer6	Offer7
0	SBI Card ELITE	Pay Now	Report Lost Card	Manage Pin	Book Flexipay	View Card Statement	Utility Bill Payment	Blog
1	SBI Card PULSE	Welcome e-Gift Voucher worth Rs. 5,000 on joining	Get free movie tickets worth Rs. 6,000 every year	Earn upto 50,000 Bonus Reward Points worth Rs....	Complimentary membership to Club Vistara and T...	None	None	None
2	Doctor's SBI Card (in association with IMA)	Noise Color Fit Pulse Smartwatch worth Rs. 4,9...	12 Month Membership for FITPASS and Netmeds Fi...	5X Reward Points on Chemist, Pharmacy, Dining ...	E-Voucher worth Rs.1,500 on achieving Retail s...	None	None	None
3	SBI Card ELITE Advantage	Professional Indemnity Insurance cover of Rs. ...	e-Gift Voucher worth Rs. 1,500 on joining	5X Reward Points on Medical Supplies, Travel B...	e-Gift Voucher worth Rs. 5,000 on annual spend...	None	None	None
4	Doctor's SBI Card	Welcome e-Gift Voucher worth Rs. 5,000 on joining	Get free movie tickets worth Rs. 6,000 every year	Earn upto 50,000 Bonus Reward Points worth Rs....	Complimentary membership to Club Vistara and T...	None	None	None

Data Preparation

Pre-processing | Techniques

- Scraped credit cards data from top 6 banks.
- The data set contains 189 unique cards and a total of 65 different offers and benefits like Cashback, Fuel, Annual Fees, welcome benefits etc.
- These 65 offers, benefits and terms and conditions are represented as variables and its associated data as observations.

CardVariant	WelcomeBenefits	AnnualBenefits	AirportLoungeBenefits	Cashback
Flipkart Axis Bank Credit Card	Rs. 1000 worth of joining and activation benefits on your Flipkart Axis Bank		Enjoy 4 complimentary lounge visits	5% cashback on Flipkart &
Axis Bank Privilege Credit Card	Get 12500 EDGE REWARD Points, re	Annual fees reversal on achieving spe	2 Complimentary access per calendar quarter to select domes	
AXIS Bank SELECT Credit Card	Get Amazon voucher worth Rs 2000	Priority Pass Membership renewal on	Get complimentary Priority Pass me	Earn 10 Axis EDGE points
AXIS Bank MY ZONE Credit Card			Enjoy 1 complimentary access to select airport lounges within	
AXIS Bank Magnus Credit Card	Choose between one complimentary	Annual fee of Rs 10,000 + Taxes waive	Enjoy 8 complimentary end-to-end VIP services at the airport	
IndianOil Axis Bank Credit Card	Earn 100% cashback up to INR 250 c	Spend more than INR 50,000 in a year and you will be eligible for annual fee waiver		

Data Preparation

Pre-processing | Techniques

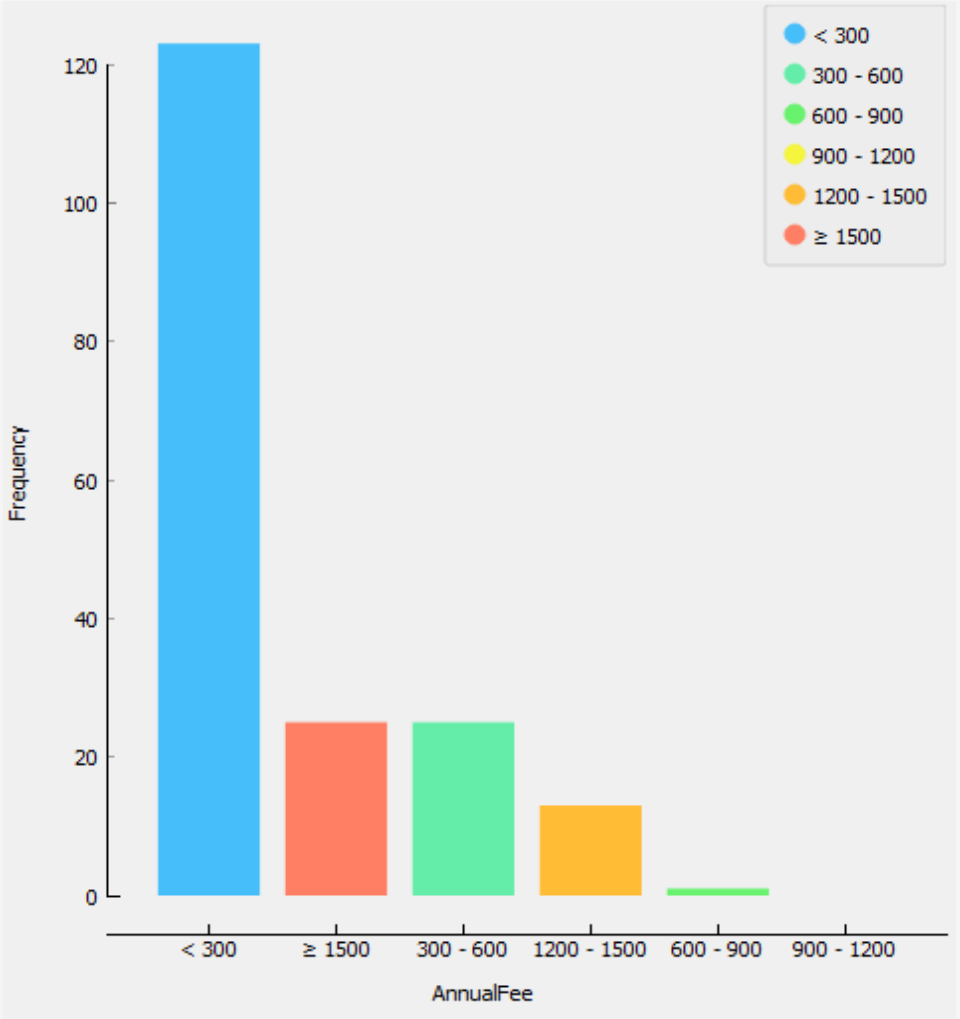
The data is then converted into Numerical format for further analysis.

BankName-CardVariant	AnnualFee2ndYear	MinimumSpend	CashWithdrawalFe	JoiningFee	AnnualMinSpend	LoungeBenefits	FuelSurcharge	FuelDiscount/C
HDFC-6E Rewards XL-IndiGo HDFC Bank	2500	0	2.5	1	0	0	0	0
HDFC-6E Rewards-IndiGo HDFC Bank	700	0	2.5	1	0	0	0	0
ICICI-Accelero ICICI Bank Credit Card	499	125000	0	0	1	1	1	2.5
SBI-Air India SBI Platinum Card	1499	0	2.5	1	0	8	1	0
SBI-Air India SBI Signature Card	4999	0	2.5	1	0	0	1	0
SBI-Allahabad Bank SBI Card ELITE	4999	0	2.5	1	0	8	1	0
SBI-Allahabad Bank SBI Card PRIME	2999	0	2.5	1	0	8	0	0
SBI-Allahabad Bank SimplySAVE SBI Car	499	100000	2.5	0	1	0	1	0

Cards Vs Annual Fee

~65 % of the cards have Annual fee < Rs.300

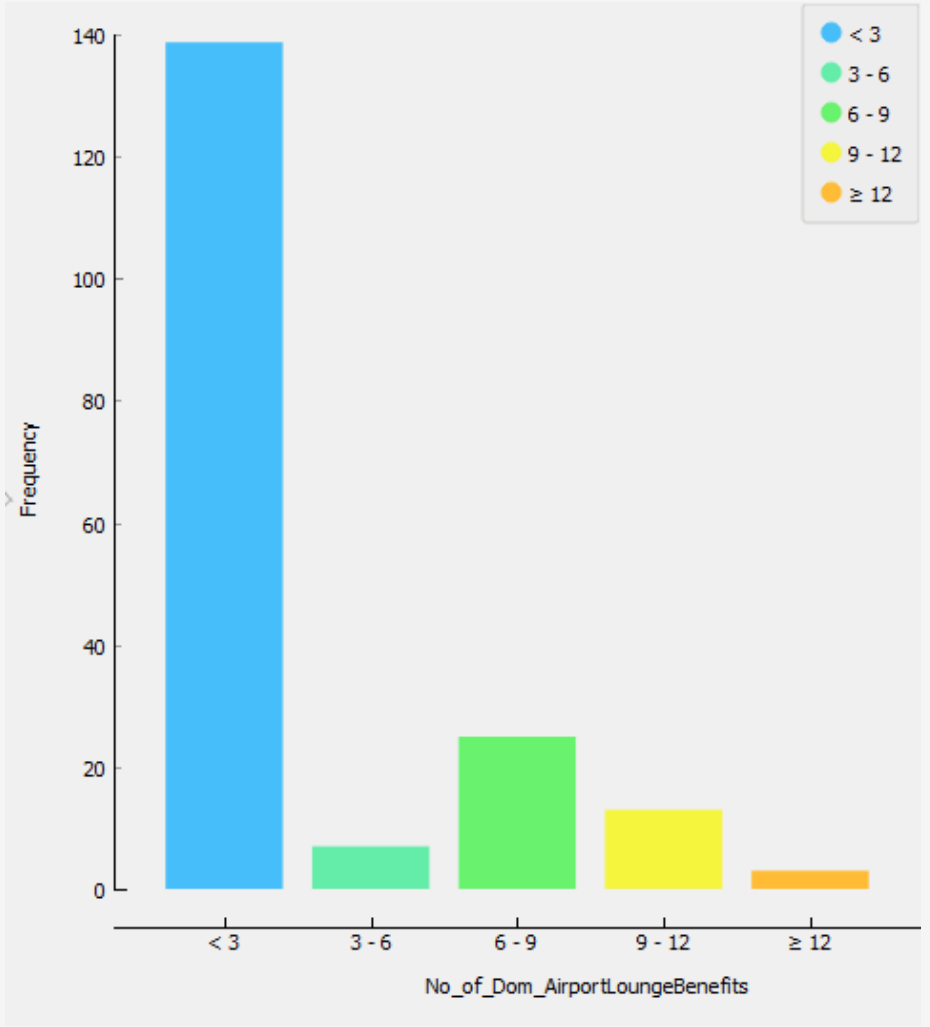
12% of cards charge Annual Fee > Rs.1500



Cards Vs Lounge Benefits

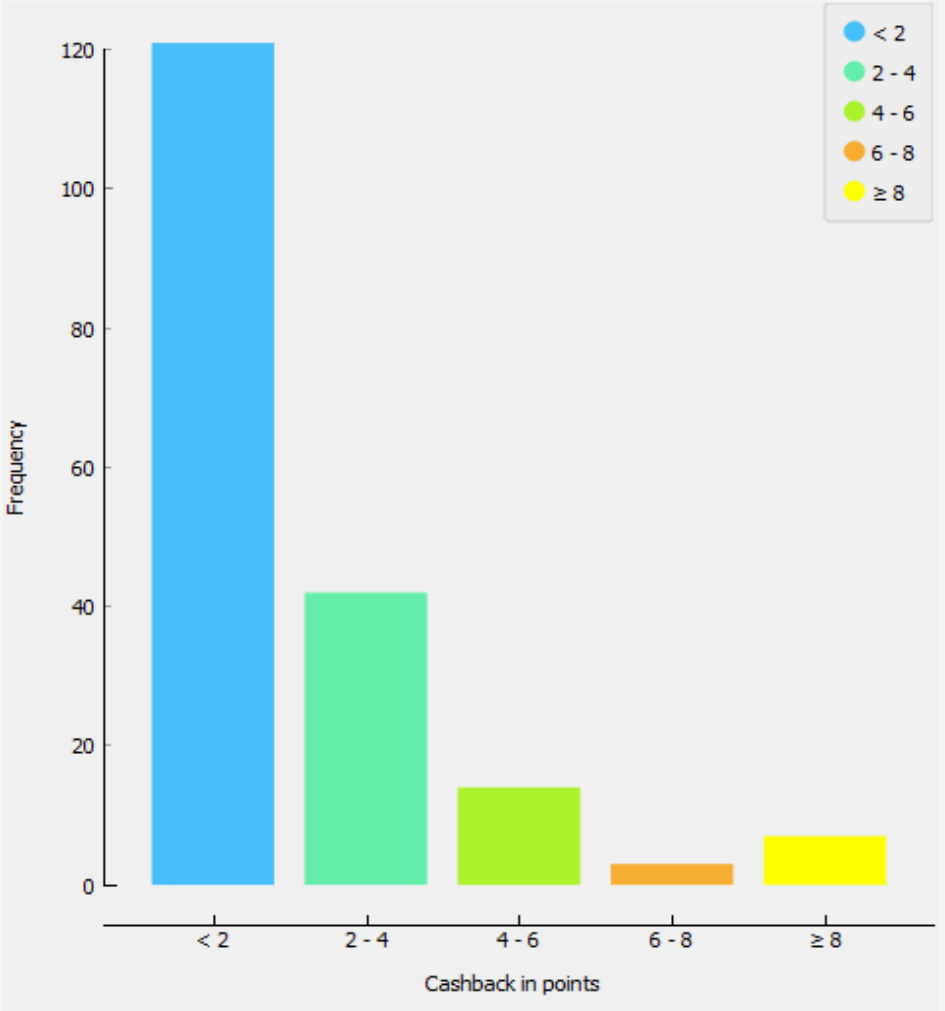
~75 % of the cards provide only 2 types of Lounge benefits

Only 10% of cards provide more than 10 Lounge benefits



Cards Vs Cashback Points

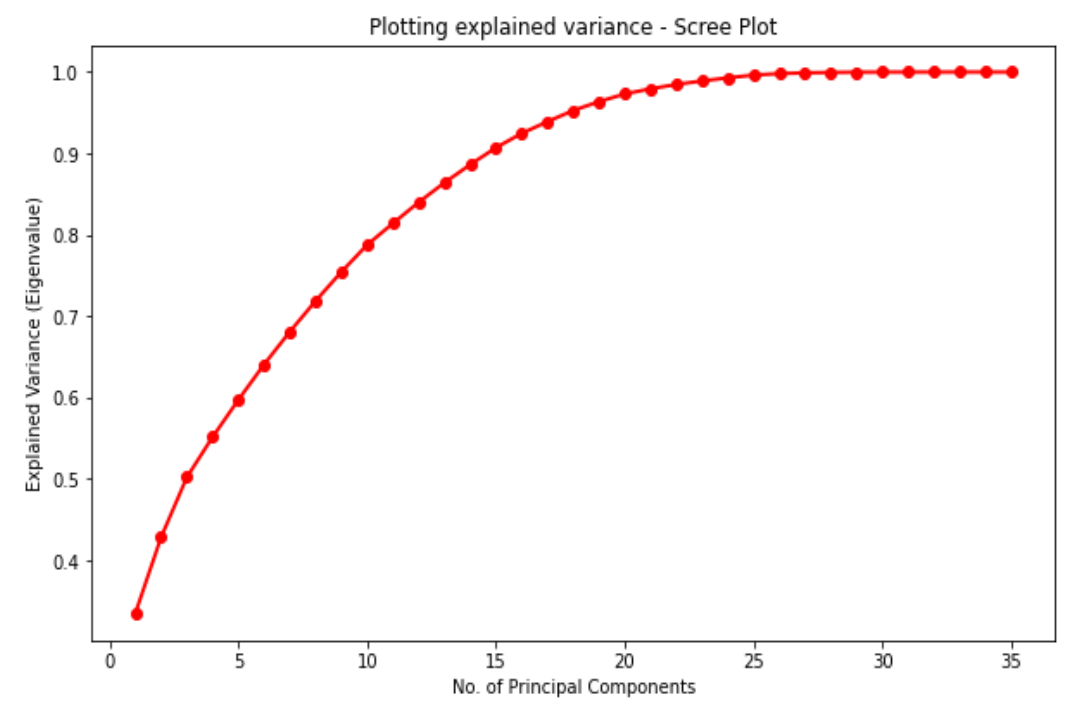
~65 % of the cards provide only 1 Cashback point for every Rs.100 spent(1% Cashback)

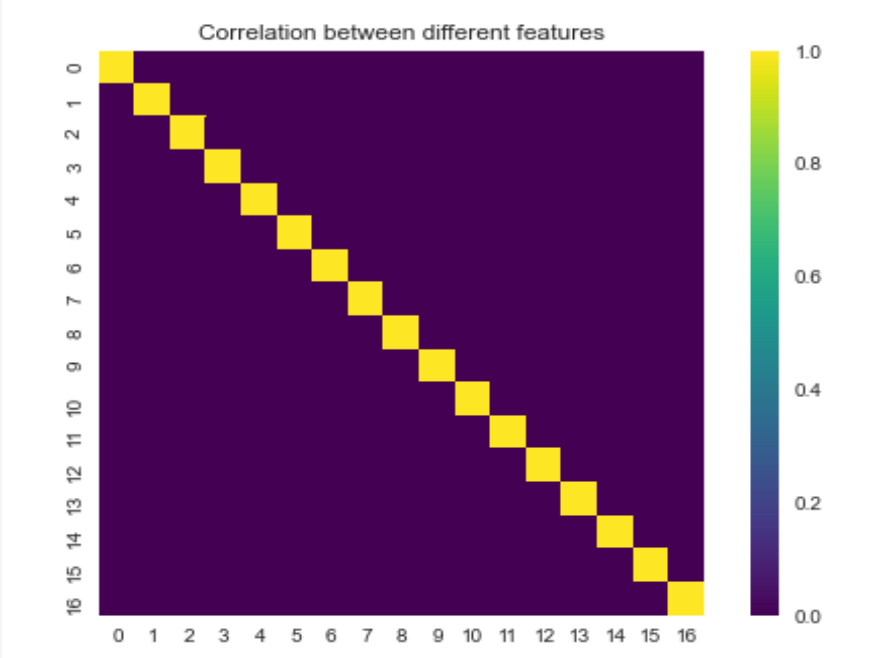


Modeling

Principal Component Analysis (PCA)

- PCA is used for dimensionality reduction.
- 17 Features covered 90% of variance
- Correlation is Zero.



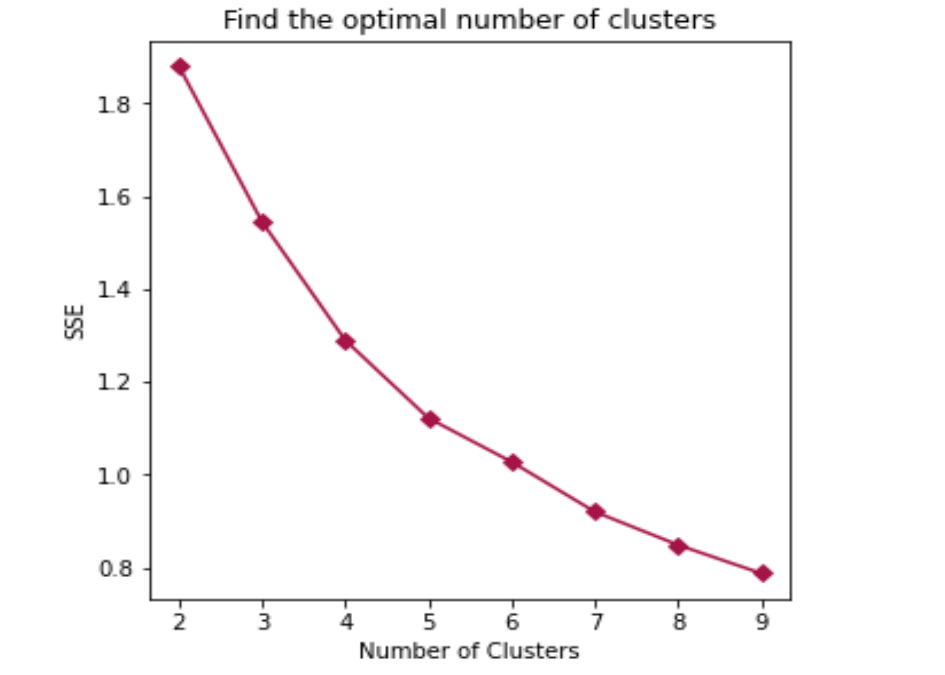


Modeling

K-Means Clustering

- The idea of clustering is to define number of clusters where the total intra-cluster variation is minimal and the inter-cluster distance within the samples of the same cluster is less.

- Cluster the Credit-cards based on similar offers and benefits.



Elbow Method

Cluster	Silhouette Score
Cluster 3	0.770
Cluster 4	0.787
Cluster 5	0.798
Cluster 6	0.851
Cluster 7	0.872
Cluster 8	0.881
Cluster 9	0.880

Silhouette Method

Clusters

Cluster 1: Low Annual Fee Low Cashback

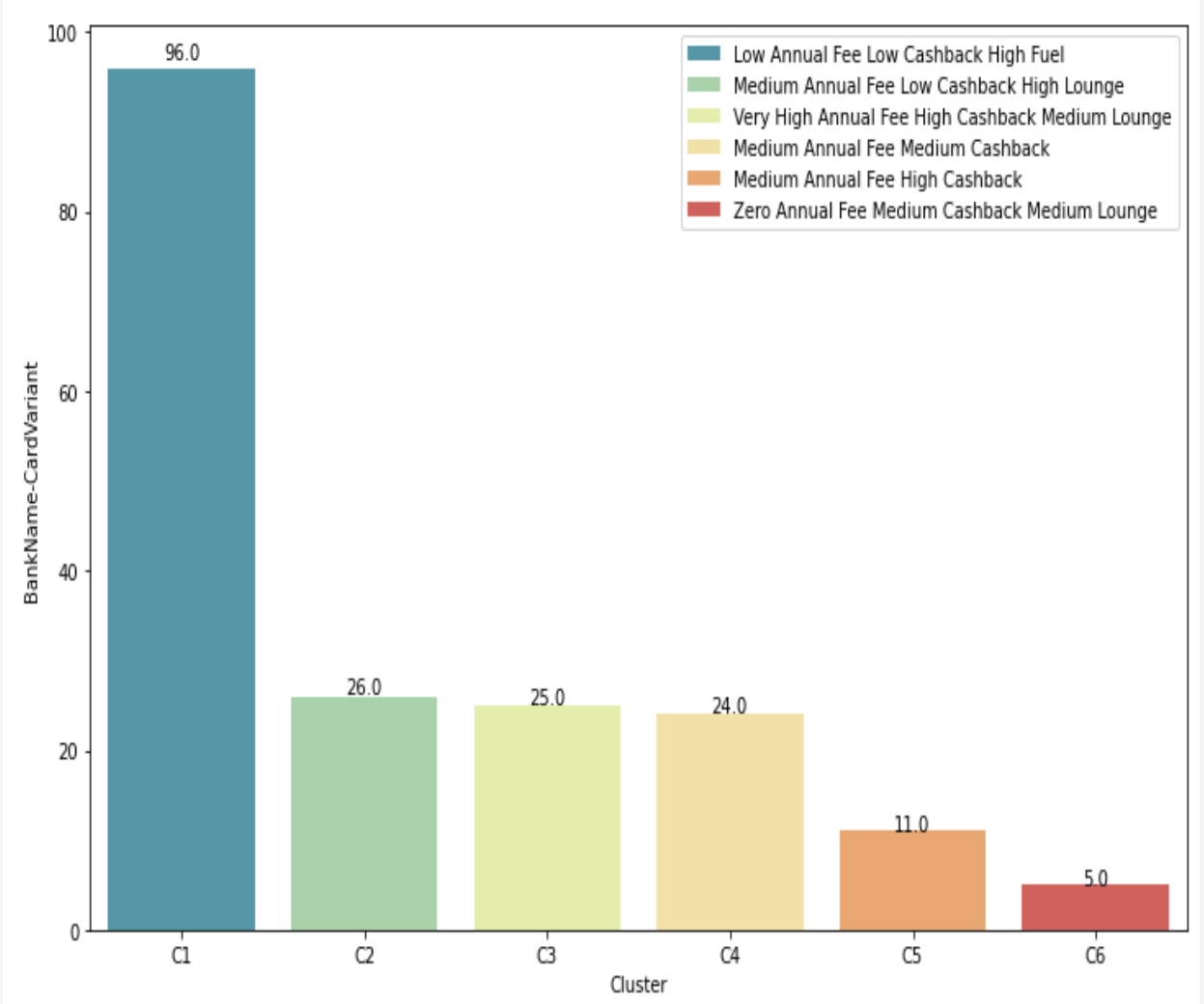
Cluster 2: High Annual Fee High Cashback Medium Lounge

Cluster 3: Medium Annual Fee Medium Cashback

Cluster 4: Medium Annual Fee High Cashback

Cluster 5: Zero Annual Fees Medium Cashback Medium Lounge

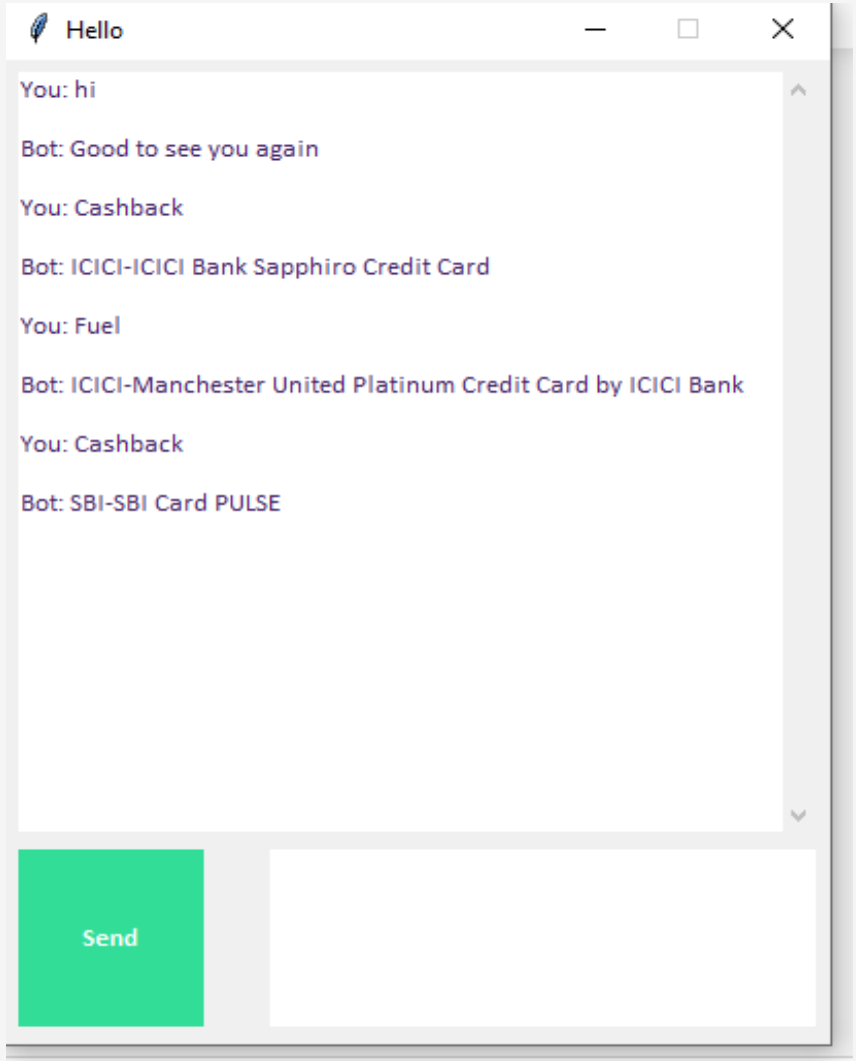
Cluster 6: Medium Annual Fee Low Cashback High Lounge



Chatbot based on Clusters

Chatbot for recommending credit card based on user input for his/her needs

For eg: If Customer is looking for a card with cashback benefits or Fuel or payback points etc.,



Recommendation with KNN

K-Nearest Neighbor (KNN)

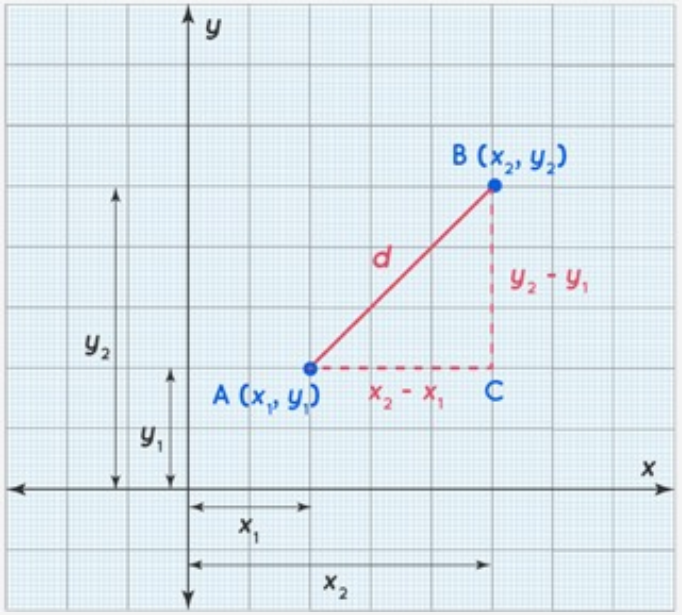
Identify K most similar Cards

Considered K=1

1-Nearest Neighbor classifier points the predicted credit card to its closest neighbor in the feature space

One Credit-card will be mapped to each and every card in the dataset based on its features which are similar to each other.

Used Euclidean distance to find the similarity between the features.



$$EuclideanDistance(x, xi) = sqrt(sum\left((xj - xij)^2\right))$$

1-NN

1-nearest neighbor for each card using KNN

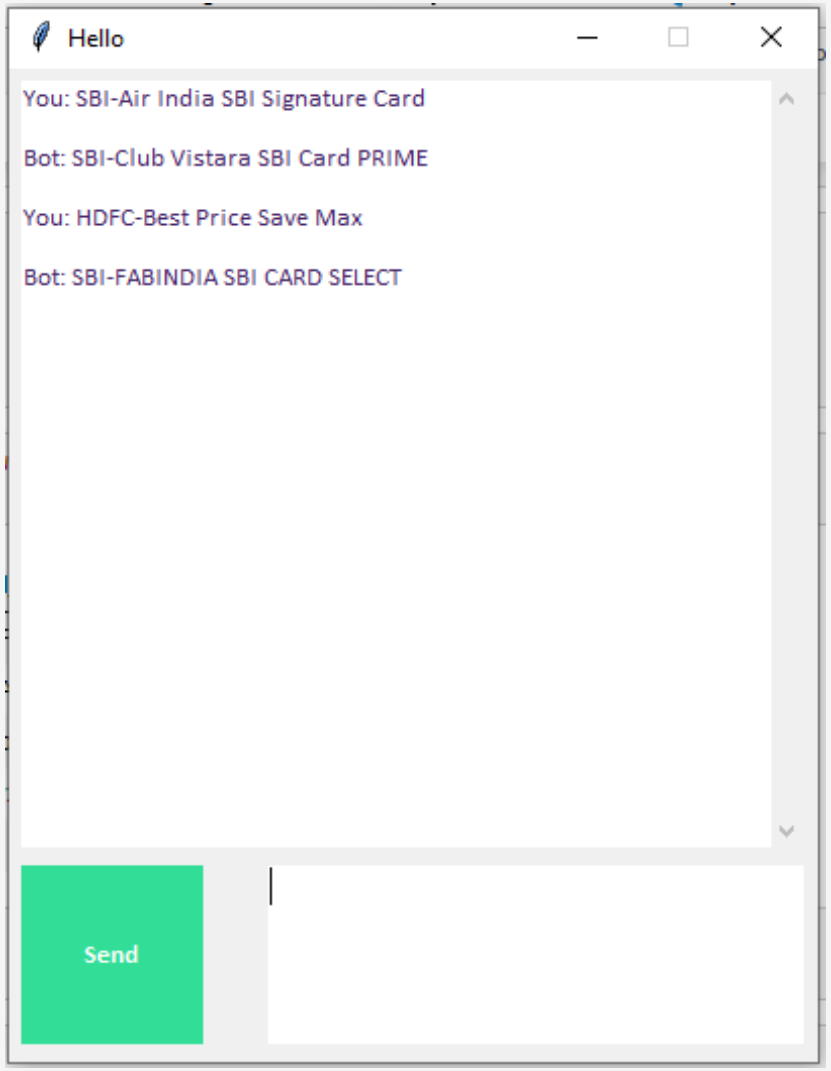
	idx	Bank-card	pred_id	pred_Bank_Card
0	0	HDFC-6E Rewards XL-IndiGo HDFC Bank	24	SBI-BPCL SBI Card
1	1	HDFC-6E Rewards-IndiGo HDFC Bank	55	SBI-FABINDIA SBI CARD SELECT
2	2	ICICI-Accelero ICICI Bank Credit Card	59	ICICI-ICICI Bank Coral American Express Credit...
3	3	SBI-Air India SBI Platinum Card	64	ICICI-ICICI Bank Rubyx Credit Card
4	4	SBI-Air India SBI Signature Card	40	SBI-Club Vistara SBI Card PRIME
...	...	...	...	...
182	182	HDFC - Regalia ForexPlus Card	167	HDFC - MakeMyTrip Hdfc Bank ForexPlus Card
183	183	HDFC - Reward Card	167	HDFC - MakeMyTrip Hdfc Bank ForexPlus Card
184	184	HDFC - Titanium Edge Credit Card	176	HDFC - Multicurrency Platinum ForexPlus Chip Card
185	185	HDFC - Visa Signature Credit Card	179	HDFC - Platinum Plus Credit Card
186	186	HDFC - World MasterCard Credit Card	176	HDFC - Multicurrency Platinum ForexPlus Chip Card

Chatbot for Similar card Recommendation

Chatbot for recommending credit card which is similar to user input card

If Customer is looking for a card which is most similar to existing card (similar card recommendation)

High level recommendation



Conclusion and Future Work

Proposed solutions | Scope for future work

- The credit cards from different banks are grouped to form 6 different clusters.
- Dimensionality reduction using PCA gives us that 17 features can be considered for analysis that has a variance of about 90%.
- These features are mostly focused on Annual Fee, Cashback, Fuel and Lounge benefits.
- 1-nearest neighbor is used to identify the most similar card for an existing credit card.
- The features can be extended to cover a higher range of benefits. This helps in identifying more clusters and recommend better suitable cards that cater different needs.
- 1-nearest neighbor can further be extended to 2 and more neighbors which will help in identifying competitive cards among different financial institutions.

References

Bibliography | Webliography

- Aggarwal, N. (2022). Digital Payments Tracker Technology. Mumbai: Motilal Oswal.
- Chadha, S. (2021, December). What is fuelling India's credit card splurge despite the rise of UPI payments? Retrieved 2022, from https://timesofindia.indiatimes.com/business/india-business/decoded-in-charts-what-is-fuelling-indias-credit-card-splurge-despite-the-rise-of-upi-payments/articleshow/88612339.cms
- Gandhi, M. (2021, August). The credit industry in India. Retrieved 2021 2021, 2021, from The changing landscape of India’s credit industry: https://www.pwc.in/industries/financial-services/fintech/dp/the-changing-landscape-of-indias-credit-industry.html
- J. MacQueen, e. a. (n.d.). Some methods for classification and analysis of multivariate observations, in: Proceedings of the fifth Berkeley symposium.
- Kiran Gajanan Javkar, S. H. (2016). Best Offer Recommendation Service. 2016 Intl. Conference on Advances in computing, Communications and Informatics(ICACCI) (p. 7). Jaipur: IEEE.

References

Bibliography | Webliography

- Lorran Santos Rodrigues, M. d. (2022). Application of DEA and Group Analysis using K-means; compliance in the context of the performance evaluation of school networks. Sciencedirect, 10.
- Ram Sharan Chaulagain, S. P. (2017). Cloud Based Web Scraping for Big Data. 2017 IEEE International Conference on Smart Cloud, 6.
- S. Wold, K. E. (n.d.). Principal component analysis, Chemometrics and intelligent laboratory systems 2 (1-3) (1987) 37–52.
- Sumit Agarwal, J. C. (2008). Learning in the Credit Card Market. NBER Working Paper No. 13822, 37.
- Wei Li, X. W. (2010). Credit Card Customer Segmentation and Target Marketing Based on Data Mining . 2010 International Conference on Computational Intelligence and Security, 4.
- Zaza, S. (2015). Mining and Exploration of Credit Cards Data in UAE. 2015 Fifth International Conference on e-Learning, 5.

Annexure

Credit Card Segmentation and Recommendation System

ORIGINALITY REPORT

8%	6%	2%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	global.oup.com Internet Source	2%
2	www.financialexpress.com Internet Source	1%
3	eprints.maynoothuniversity.ie Internet Source	1%
4	Lorran Santos Rodrigues, Marcos dos Santos, Claudio de Souza Rocha Junior. "Application of DEA and Group Analysis using K-means; compliance in the context of the performance evaluation of school networks", Procedia Computer Science, 2022 Publication	1%
5	Submitted to University College London Student Paper	1%
6	Submitted to University of Nottingham Student Paper	<1%
7	www.coursehero.com Internet Source	<1%
8	Submitted to Georgia Institute of Technology Main Campus Student Paper	<1%
9	www.communitylawcentre.org.za Internet Source	<1%
10	rickreagan.com Internet Source	<1%
11	en.wikipedia.org Internet Source	<1%
12	onlineasdma.assam.gov.in Internet Source	<1%
13	link.springer.com Internet Source	<1%
14	iiespace.iie.ac.za Internet Source	<1%

REVA UNIVERSITY

Bengaluru, India

Established as per the section 2(f) of the UGC Act, 1956,
Approved by AICTE, New Delhi

Thank You!