

# **Data Engineering and Privacy Challenges in Loyalty Card Programs: Insights from Retail, Banking, and Hospitality**

## **ABSTRACT**

This paper explores the data engineering practices surrounding loyalty card programs in the retail banking, travel, and hospitality industries. It examines how businesses use personal and transactional data to enhance customer retention and gain competitive advantages. The research highlights industry-specific practices, regulatory frameworks (such as GDPR and CCPA), and the security measures adopted to protect Personally Identifiable Information (PII). Key insights into how data is leveraged for targeted marketing and the persistent vulnerabilities in data protection systems are also discussed. The study provides a comprehensive view of the ethical, economic, and strategic implications of loyalty card data on modern businesses and consumers.

With practical and theoretical methods applied, we discover and share our findings from how consumers' loyalty is segmented across industries, their preferred payment methods, the bridge between points acquired and utilized by customers across various tiers, as well as churn analysis with a focus on clients who left the business and factors that motivated the exist. Recent surveys (in the USA and UK) also help us understand users' preferences among different age groups, therefore providing insights for businesses to leverage. The invaluable nature of data begs the need for protection laws to be in place, hence, we explore measures that are set in place for all parties involved to protect consumer data from collection to usage. The research also proposes a framework that could be modelled by companies hoping to create a loyalty program.

*Keywords:* Loyalty card, loyalty program, data, personal data, customer, consumer, security, industry.

## **1. INTRODUCTION**

On a Monday morning, a well-dressed lady by the name of Adele walks into a new branch of a coffee shop, hoping to get a cappuccino to boost her focus and productivity at work. Upon being greeted and welcomed by the attendant across the counter, she makes her order; after the order collection and processing, the attendant asks if she has their coffee shop card. Adele replies, "No," and she is informed of a 50% discount based on a first-time purchase and more promotional offers if she signs up by scanning the QR code seated on the counter. The QR code hyperlinks to the mobile application store to download the app, complete a form to create an account, and then apply a discount on her purchase.

The form collects information like her full name, email address, phone number, gender, date of birth, and home address. Upon successfully signing up for this loyalty program, Adele receives the coffee and pays for it at a discount. This is a supposed win-win for both parties. Where Adele gets her coffee, hoping to get more eventually based on continuous promotions, and loyalty to the brand. The coffee company leverages her data to provide targeted marketing campaigns rendering promotional offers from her coffee purchasing patterns as told by the data and the underlying ones collected too. Due to a

high coffee-drinking population around this business area, all other customers' patronage of their services is analysed as well to effectively sell to them, hoping that this analysis is effective with a single result. More sales.

Customers are majorly focused on getting their desired product alongside a discount (if applicable), and "due to improvements in technology and decreasing costs of processing and storage, data are exchanged without ever providing notice or transparency about the handling and storage of such information to those individuals whose personal data are involved." (Kline, 2007). This end goal of consumption leaves the consumer in the dark, unbeknownst to them about how their data is handled.

This paper aims to dive deep into exploring the collection, analysis, use and misuse of data which is deemed as Personally Identifiable Information (PII) and more, of customers with a focal point on how this data flows in loyalty cards and is utilised by businesses in various industries. It shines a light on the security, economic, social and strategic impacts it has on customers and businesses (both positively and negatively).

The following comprehensive and descriptive questions guide this research and support the topic to achieve its goal:

1. What types of personal data are collected through loyalty cards, and how is this data categorized?
2. How does the flow of personal data within loyalty card programs differ across industries (e.g. retail, hospitality, grocery chains) in terms of collection, usage and sharing?
3. What competitive advantages do companies gain by leveraging loyalty card data, and how does this impact their market position?
4. What are the key regulations governing the collection, use and transfer of personal data through loyalty card programs, and how are companies complying with them?
5. What security measures are in place to protect the personal data collected by loyalty cards, and how vulnerable are they to data breaches or cyber-attacks?

Accompanied by datasets, extensive research, methodologies used, Key Intelligence Questions (KIQs) asked, diagrams, images and other necessary visuals to traverse, we will present significant findings and discuss the limitations encountered.

## 1.1. Objectives

To develop a comprehensive understanding of our aim, we explore the various KIQs and provide a broader insight with some examples. We look at the:

### 1. Data Collection Practices

This data spans from name, email, phone number, purchase history, payment information, and demographic details (like gender and age). More in-depth data dives like shopping patterns, preferred products, and frequency of visits, which are classified as behavioural data, are also explored. PII, transactional, behavioural and demographic data are the major categories that will be looked at.

The objective of this is to understand how companies utilize them as a source for carrying out operational and marketing campaigns to achieve the goals of their loyalty program.

### 2. Industry-Specific Insights

A holistic view of how different industries collect, analyze and share personal data. We shine a light from the top down to provide an understanding of their various practices.

A common difference is the discrepancy in the collection of data types between the hospitality industry and grocery chains; whereby airlines collect more sensitive data such as travel history, dietary preferences and medical needs (e.g., allergies). For airlines with a wider range of services like car rentals, there is an integration of third-party services such as Google Maps, Apple Maps, etc. This integration involves data sharing across industries. Grocery chains differ from the former by collecting and leveraging data like customers' purchase history and location to track consumer preferences and provide geo-targeted promotions around where they frequently shop.

### 3. Advantages and Performance

Alongside other common customer knowledge, the cost of finding a new client is 5 to 6 times more expensive than keeping an existing one. (Herbert, 2020). A restaurant that notices the regular patronage of a client to savour his/her favourite meal and offers them a discount on that meal after multiple purchases has a high probability of retaining them as patrons, as well as others who patronise them.

Market demographics, buying patterns, and the effectiveness of marketing campaigns could be understood by companies from loyalty card data. The awareness created can help them course the path of uncharted markets, create niche offerings or optimize their pricing strategies.

This aims to help companies benchmark their performance against competitors by truly understanding their market advantage, therefore improving on it to obtain a larger market share and position.

### 4. Compliance and Security

Across continents, various data protection laws exist with common examples like the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA), the Health Insurance Portability and Accountability Act (HIPAA), the Children's Online Privacy Protection Act (COPPA) and the Gramm-Leach-Bliley Act (GLBA) in the United States of America, which place stringent requirements on the collection, storage, and sharing of personal data. These laws ensure that companies fully adhere to data minimization, inform customers about their data collection practices with the choice to opt-out, and enforce that third-party vendors or service providers adhere to the same data protection standards and maintain customer data security properly.

Our viewpoint in this section fixates on how companies who engage in loyalty card programs comply with key data protection regulations as well as the steps they take to ensure personal data is handled responsibly and lawfully.

### 5. Security measures and Vulnerabilities

Britain's largest airline, British Airways (BA), suffered a data breach whereby their loyalty program was hacked into. The hackers had access to over 400,000 clients' log-ins, payment details, travel details and addresses. This cybersecurity breach resulted from vulnerabilities in the system, and it took two months for BA to realise this breach; in which they eventually got fined \$26 million (Tidy, 2020).

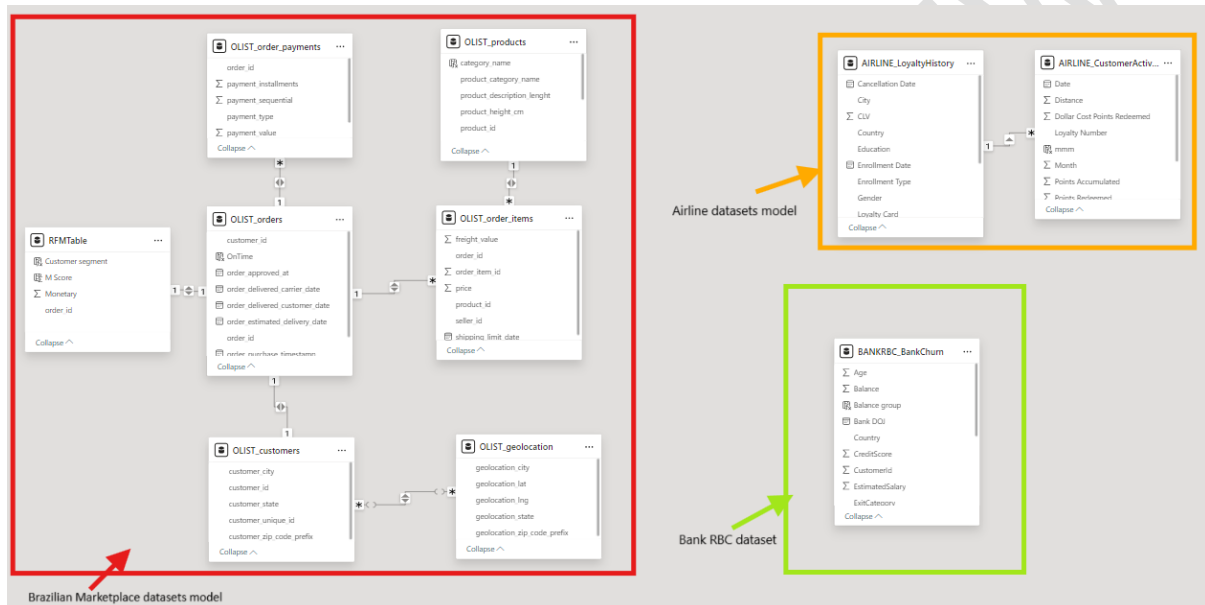
Most companies grant access to sensitive information like PII only to authorized personnel, and they use encryption technologies to protect customer data. Despite the numerous security measures, loyalty programs are still vulnerable to data breaches and cyberattacks. Therefore, companies must regularly update their cybersecurity protocols and conduct penetration testing to identify and mitigate vulnerabilities. This is the objective pursued here.

## 2. METHODOLOGY

Our applied methodology involved four steps. We started by picking our Key Important Topic (KIT) and then meticulously posed intelligent questions that are key to our topics while voyaging through the USA,

Europe, Canada and Brazil as they constitute our Key Important Territories in understanding the impact of loyalty card data on customers and businesses. Secondly, we conducted an extensive literature review around the surface and deep web by querying for useful materials through various intelligence techniques. We browsed for open-source intelligent materials from websites like Google Scholar, Statista (for surveys and statistical information), Web of Science, and the World Intellectual Property Organization (WIPO). The querying techniques coupled with streamlined prompting helped us achieve the next step, data collection; we collected 9 datasets from websites like Kaggle, Google Dataset Search, and GitHub, then reviewed some surveys relevant to our goals.

Finally, we used 3 datasets, namely: “Real data of RBC bank Canada”, “Airline Loyalty Dataset”, and “Brazilian E-Commerce Public Dataset by Olist”. These datasets contained huge amounts of real customer data that were made publicly available. We chose to use them because they did not have simulated inputs, leakage, or a high amount of missing data. With the Power BI software, cleaning was done as the first part of the data preprocessing methods; afterwards, filtration of columns relevant to our analysis took place (see Figure 1). Then the data underwent some formatting to output visuals used within the paper to visually back up our findings.



**Fig. 1. Power Query cleaning and transformation of data**

The premiere limitation is the unavailability of a vast number of actual industry-related datasets for analysis, findings and supplying solutions. The datasets provided by companies are mostly used for hackathons within a limited period, thus barring the possibility of interested persons carrying out research at any time. Most available datasets are also simulated or anonymized with abstract names. Anonymization aims to protect actual user data as per data protection laws. Simulated datasets provide (limited) insights on how businesses can engineer data for profit and bring about better service provision to their customers; they lack the real-world factor of understanding customer purchasing patterns, the quantity of data collected, and active service patrons as opposed to the total onboarded ones.

### 3. RESULTS – DATA VISUALIZATION AND STORYTELLING

#### 3.1. Key Intelligence Question 1

"Big data" is what defines the retail industry in the modern day. Because of the need to adjust to the ways that modern customers shop—in-store, online through mobile channels, and via many other connected devices—retailers in this period face a fast-shifting market landscape. Using this data, retailers can make the shopping experience convenient for their customers even better by introducing new business strategies like omnichannel and personalized recommendations.

##### 1. Traditional Enterprise Data Capture.

Traditional sales data from Universal Product Code (UPC) scanners and inventory data from Enterprise Resource Planning (ERP) or Supply Chain Management (SCM) software comprise the initial set of data sources. The four components of marketing: product, price, promotion, and place—as well as product performance at the store, aisle, shelf, and other levels are covered in these sources. Retailers can see product relationships and cross-category dependencies by using this data to look at market baskets as Figure 2 exemplifies using Olist market data.

product_id	product_category_name	product_name_length	product_description_length	product_photos_qty	product_weight_g	product_length_cm	product_height_cm	product_width_cm
5d2754abcd85429bc7a4c37d52fba43	fashion_bolsas_e_acessorios	53	770	1	200	16	2	11
49f9eeab25d653114bbd4e9e7c926c36	fashion_bolsas_e_acessorios	48	731	1	200	16	2	11
8f4735c09292c695473969af57466365	fashion_bolsas_e_acessorios	60	83	1	200	16	2	11
9892435c317b92530812ef9fe8cfc6	fashion_bolsas_e_acessorios	51	731	1	200	16	2	11
165948bc2c1e846d7d09b56826a747f	fashion_bolsas_e_acessorios	55	89	1	200	16	2	11
07d56eb519d315a2d8fb298f095e6bf	fashion_bolsas_e_acessorios	56	1447	1	200	16	2	11
fe09eb57bfcb79d05df7c4e83ec204e	consoles_games	60	574	1	200	16	2	11
da93a362382f7e8e659cfea14aa9b918	fashion_bolsas_e_acessorios	49	54	1	200	16	2	11
05216a143a65ab59c48fb996e7c12d7	fashion_bolsas_e_acessorios	53	353	1	200	16	2	11
683487a7a9ed71b5776c9b9cad37eed	fashion_bolsas_e_acessorios	52	23	1	200	16	2	11
447b7437663c07b241787482238e1d58	fashion_bolsas_e_acessorios	58	1359	1	200	16	2	11
9f1fbb90131eec1287c7ee32bb32a31	fashion_bolsas_e_acessorios	48	796	1	200	16	2	11
059d3c4de5bc9099c4c10ee7e84d5	automotivo	58	824	1	200	16	2	11
6c04a069e5ab37749c980c42a036b9e3	automotivo	52	1367	1	200	16	2	11
3ca3e63474ec7f039c2935cd25d202	fashion_bolsas_e_acessorios	55	1434	1	200	16	2	11
b842a6c9c2c51a699729d378d8d674c	fashion_bolsas_e_acessorios	51	48	1	200	16	2	11
62f68be290c2e9c8403e306bd023041c	fashion_bolsas_e_acessorios	41	52	1	200	16	2	11
1632189d4c3d9e64c3fd1110b93f1a19	fashion_bolsas_e_acessorios	56	1474	1	200	16	2	11

Fig. 2. Products table of Brazilian marketplace Olist

### 2. Customer Characteristics from Social Media.

The second set of data sources is about customers and gathering data about them, including basic contact info like email addresses, purchase history, preferences, responses to promotions and demographics. A single customer ID is utilized to connect all relevant information about a customer from several sources. Inquiries, complaints, email click-throughs, and data acquired from syndicated sources regarding customers' online searches and products viewed but not necessarily purchased are among the other ways retailers can obtain information about their online behaviour. It is possible to have a better understanding of consumers' purchasing contexts and promotion responses by utilizing this data.

customer_id	customer_unique_id	customer_zip_code_prefix	customer_city	customer_state
4e7b3e00288586ebd08712fdd0374a03	060e732b5b29e8181a18229c7b0b2b5e	1151	sao paulo	SP
fd826e7cf63160e536e0908c76c3f441	addec96d2e059c80c30fe6871d30d177	4534	sao paulo	SP
eabebad39a88bb6f5b52376faec28612	295c05e81917928d76245e842748184d	5704	sao paulo	SP
9b8ce803689b3562defaad4613ef426f	7f3a72e8f988c6e735ba118d54f47458	5416	sao paulo	SP
2938121a40a20953c43caa8c98787fcb	482441ea6a06b1f72fe9784756c0ea75	5713	sao paulo	SP
cb721d7b4f271fd87011c4c83462c076	a5844ba4bfc8d0cc61d13027c7e63bcc	8225	sao paulo	SP
6e359a57a91f84095cc64e1b351aef8c	2e6a42a9b5cbb0da62988694f18ee295	4571	sao paulo	SP
e3109970a3fe8021d5ff82c577ce5606	a8654e2af5da6bb72f52c22b164855e1	5528	sao paulo	SP
a3b0fda37b7ae14cf754877bed475e80c	c9158d089637ab443c78984d20da7fc0	5727	sao paulo	SP
d3b6830d18c7de943d1e707d1f061d40	27cf4b153010911a0957150255a6c6db	5351	sao paulo	SP
a9b0d1c26105279e1b8edc63d06bd668	3d49f4455a3947c8dd7e972b3ad8cb76	5017	sao paulo	SP
03f846ad03437d864a8d2a22976dcafe	7677c213007e9a6ec9267ea50b5ce5bc	2075	sao paulo	SP
0aae2862f8eac77f10a34f44860720ac	cd076285a12f40041b32f5ad8c98699f	5890	sao paulo	SP
6c9a5923526346cb0bd7bbd92269c01	cf6d4152d758efc43910e0141ae5b912	3733	sao paulo	SP
8264e3518163dd09211870b24a5d741d	67d21c8bea9d6017d1b124d3879dd815	5415	sao paulo	SP
38d1cd89306128348ffdf4cc23f3a50a	d491a65a6ef3c04e145d37395996bad7	4548	sao paulo	SP

Fig. 3. Customers table of Brazilian marketplace Olist

### 3. Customer location-based data.

Utilizing consumers' locations to learn about their preferences and buying patterns to create marketing campaigns that are specifically targeted towards them is the third category of data sources. The way location-based data is received and used has been changed by smartphones. These devices have mobile shopping apps that collect data (Jonathan Zhang et al., 2023). Figure 4 shows a small amount of customers' geolocation collected across Brazil with their zip code, latitude, longitude, city, and state.

geolocation_zip_code_prefix	geolocation_lat	geolocation_lng	geolocation_city	geolocation_state
4363	-23.646845934214372	-46.6660547231519	sao paulo	SP
4363	-23.646845934214372	-46.6660547231519	sao paulo	SP
4363	-23.647623590339908	-46.66684794962203	sao paulo	SP
4363	-23.64789871412087	-46.669586894656454	sao paulo	SP
4363	-23.64793806578562	-46.66826871856427	sao paulo	SP
4363	-23.648279532056094	-46.66647444962204	sao paulo	SP
4363	-23.648613296150906	-46.66888515528881	sao paulo	SP
4363	-23.648170861503843	-46.66887473036527	sao paulo	SP
4363	-23.64772359616828	-46.66575475353455	sao paulo	SP
4363	-23.647247591724888	-46.666924501529216	sao paulo	SP

Fig. 4. Geolocation table of Brazilian marketplace Olist

### 3.2. Key Intelligence Question 2

The loyalty programs are segmented into different programs, namely, points programs, tiered programs, and paid programs, amongst others. These three listed are the most popular, with the points program having the largest slice of the pie based on market share as of 2023, as these programs are simple and flexible to use (Sharma, 2024). The author proffers that the points-based program ranks top, largely because of the instantaneous benefits from accumulating points per purchase.

From the global market, we narrow our research to the retail, BFSI (Banking, Financial Services, and Insurance) and travel & hospitality industries to answer our second KIQ. This reason stems from those industries being the big key players over time, coupled with a forecast of dominance in the future as they possess a high customer acquisition and retention rate and a Compounded Annual Growth Rate (CAGR) as customers continuously make use of their designed reward initiatives.

We analysed data provided by the largest department store in the Brazilian marketplace called Olist for the retail industry with a focus on e-commerce; the dataset contained 100,000 orders in 2 years (2016 to 2018) from multiple marketplaces in Brazil.

Customer Segment	auto	bed_bath_table	computers_accessories	furniture_decor	health_beauty	housewares	sports_leisure	telephony	toys	watches_gifts
Loyal	825	1,494	1,221	1,234	1,856	905	1,361	468	661	1,931
Potential Loyalist	902	1,985	1,523	1,282	1,463	1,087	1,722	205	754	1,484
Promising	575	2,478	1,372	1,552	2,053	1,074	1,747	251	974	1,005
Need attention	688	2,151	1,127	1,497	1,693	1,344	1,546	687	856	737
At Risk	896	1,104	1,428	832	1,726	1,465	1,325	2,578	616	462

Fig. 5. Order Frequency by Product Category

The heat map displayed in Figure 5 provides information on various product categories with a total of 62,182 orders across 10 product segments and the purchase frequency by customer segment. According to this customer segmentation, “Promising” customers with 13,235 orders across various product types, those “At Risk” as well as “Potential Loyalist” (with over 12,000 orders), are the top 3 in the segment. The “bed\_bath\_table” category occupies the first position with 9,412 total orders, with

8,791 and 7,701 orders for “health\_beauty” and “sports\_leisure”, respectively, placing them second and third, which is also illustrated in Figure 6 in the form of a Bar Chart.

### | Top 10 Products by Orders

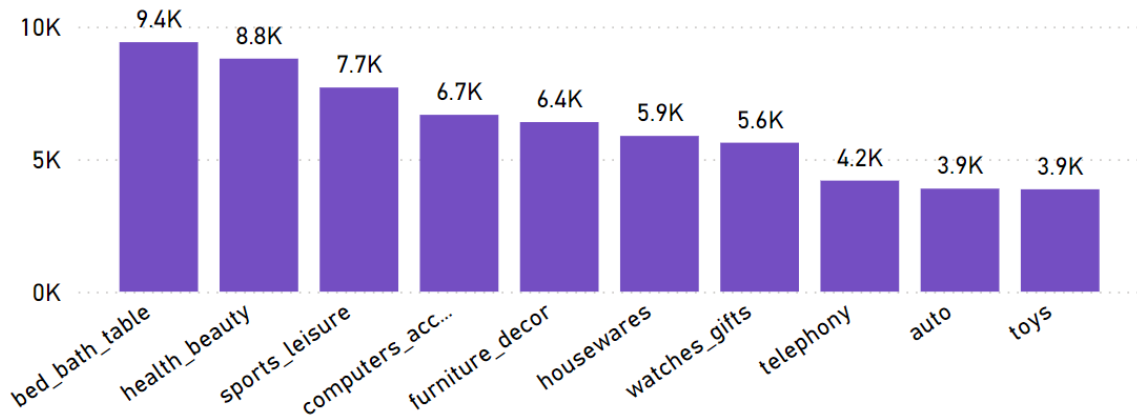


Fig. 6. Top 10 Products by Orders

Companies can focus on product categories frequently purchased by consumers and tailor loyalty programs towards those, as well as orchestrate payment-related promotions or partnerships with preferred payment providers for increased loyalty engagement (Idugboe, 2014). Figure 7 indicates the ratio of payment methods used, therefore providing the companies with an opportunity to attenuate and strategize their customer engagement by collaborating with their third-party payment providers to optimise swift payment solutions and cashback benefits. The result of this analysis offers the competitive and business intelligence department of a company some sage advice on how to better engage customers in purchasing lower-performing products and using the less frequented payment methods.

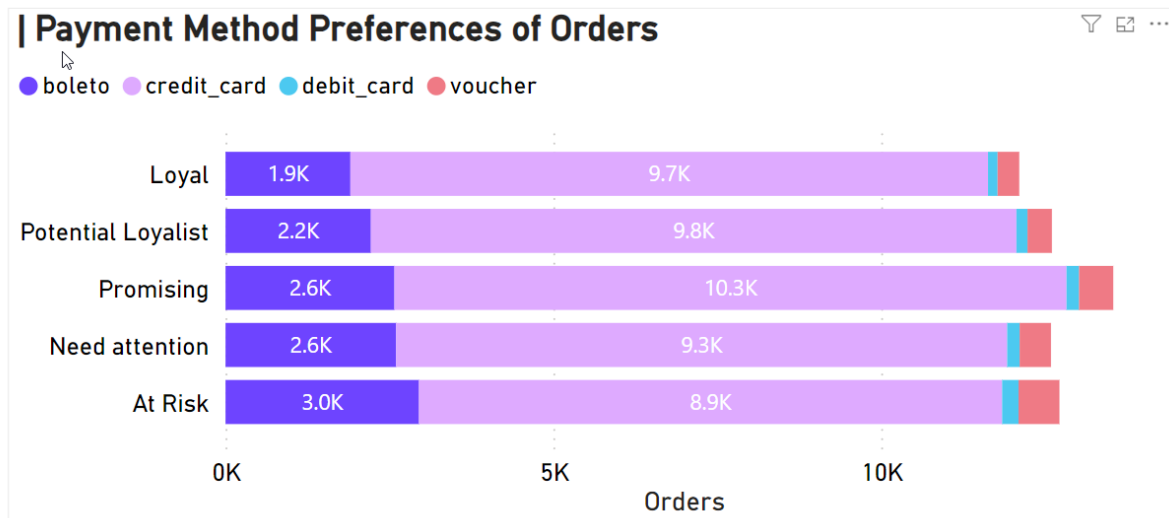


Fig. 7. Payment Method Preferences of Orders

Figure 8 demonstrates sales made within two years of the data collection with a focus on various customer segments categorising loyalty. An average of 370 customers and above are loyal across

various months; although this is good, those who fall into the “Potential Loyalist”, “Promising”, “Need attention”, and “At Risk” segments need a new dynamic in customer service provision to fully convert them because they constitute 46.58% in total as seen in the Pie chart on Figure 9.

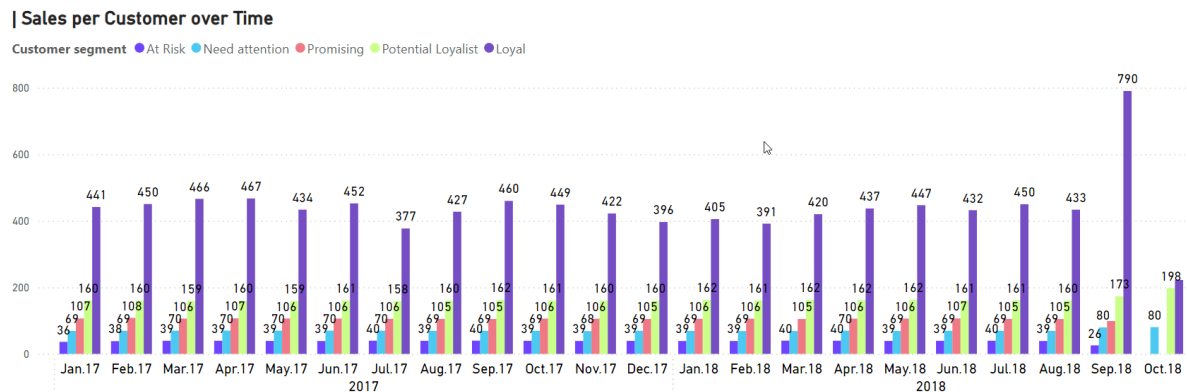


Fig. 8. Sale per Customer over Time

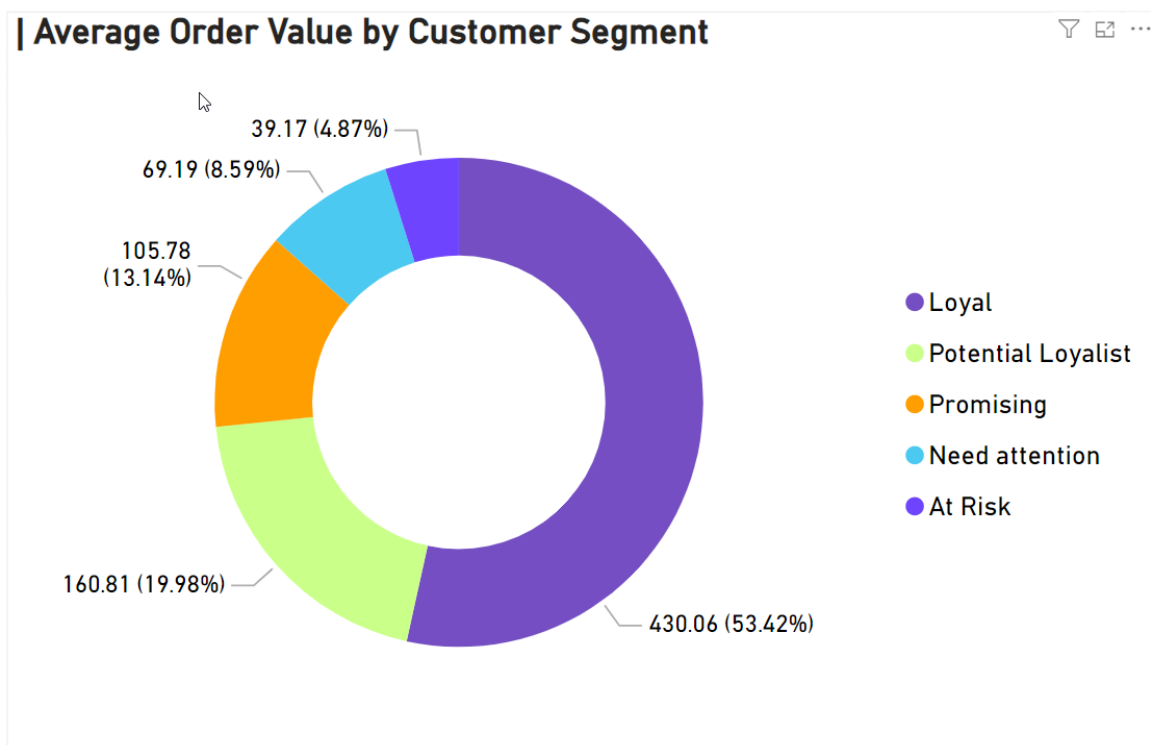
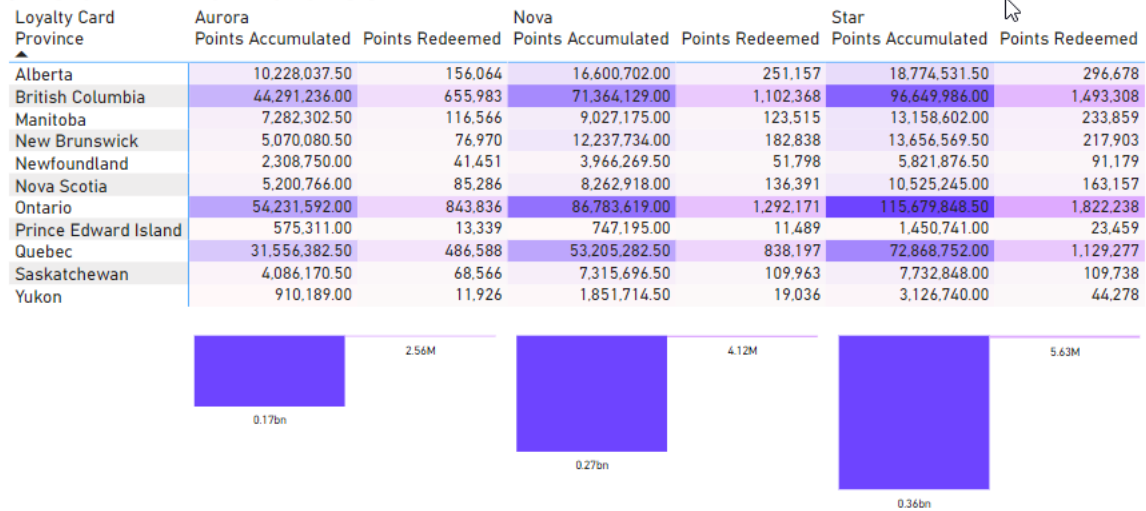


Fig. 9. Average Order Value by Customer Segment

To improve enrolment and usage of their loyalty program, we engaged in some analysis of a Canadian-based airline as our second industry of focus. Flight activity of various residents in Canada between February and April 2018 was monitored by looking at their loyalty tiers and correlating it with other data points.

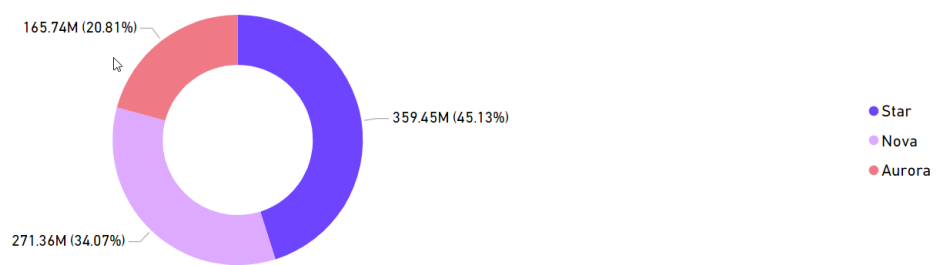
**Geographical Heatmap of Loyalty engagement**



**Fig. 10. Geographical heatmap of loyalty engagement**

Figure 10 illustrates that this airline collects and analyses data based on the various provinces in which its customers reside. With a total of over 796 million points accumulated from all customers, only a total of 12.3 million points were redeemed, with the star points tier grossing the most accumulated and redeemed points from Ontario, British Columbia and Quebec provinces. **The huge deficit of over 780 million unredeemed points creates a concern. Businesses within this industry are advised to approach each customer tier differently; because there are several factors that might affect client retention and accumulated points usage, these factors may include better service provision from competitors which clients like and migrated to, increased earnings of individual clients with current benefits within the tier unsuitable for their current financial status.**

**CUSTOMER TIER DISTRIBUTION BY ACCUMULATED POINTS**



**Fig. 11. Customer tier distribution by accumulated points**

To further strengthen Star tier as the top distribution among the 3 categories, Figure 11 displays (in percentages) its popularity in usage by 45.13%, Nova with 34.07% and Aurora with 20.81%. This demonstrates that the least tier category is the most used, and despite a higher salary earned by some of their clients, it doesn't guarantee allegiance to the loyalty program or the airline (exhibited in Figure 12).

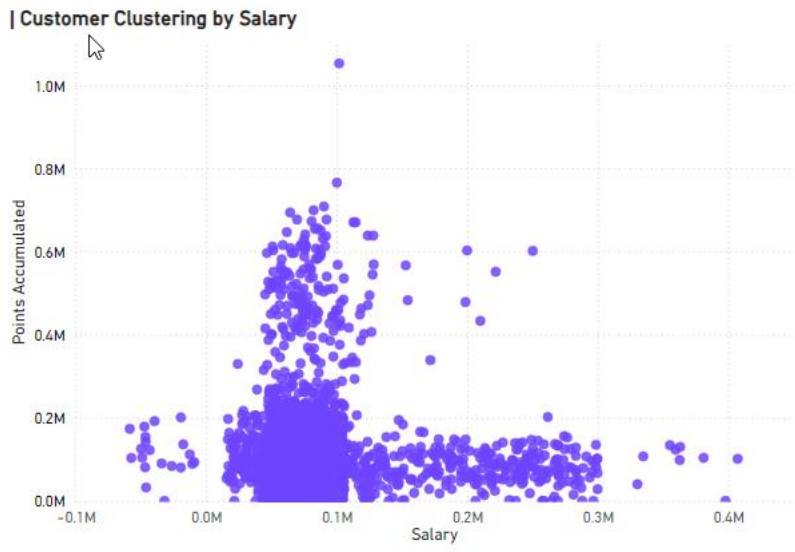


Fig. 12. Customer clustering by salary

Bank RBC Overview

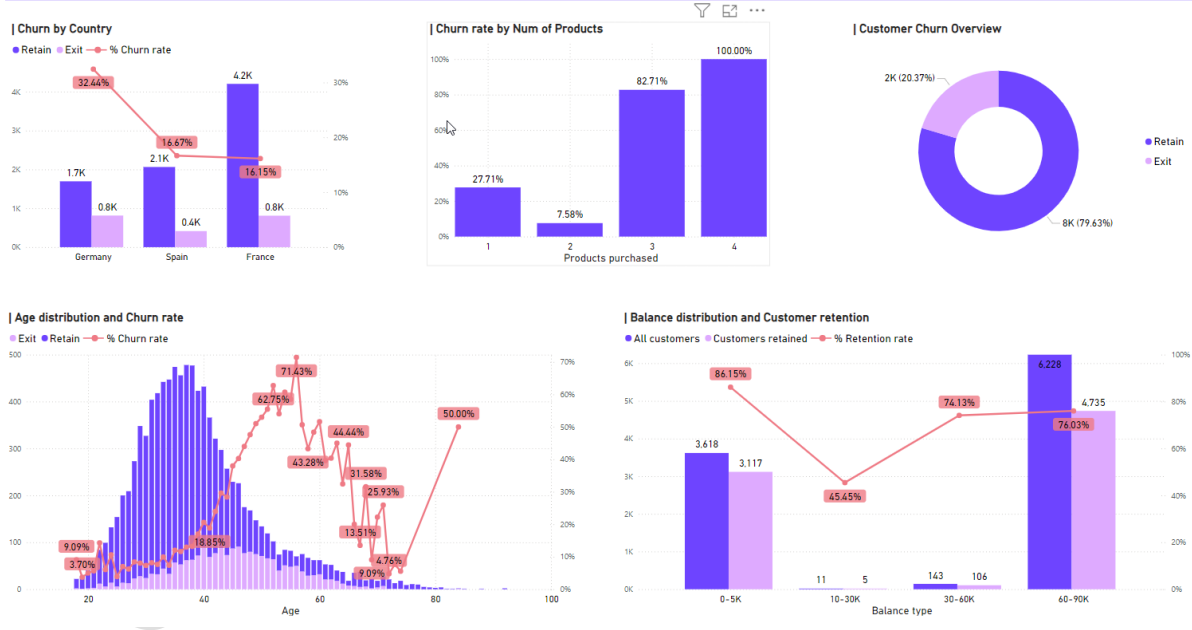


Figure 13: Royal Bank of Canada (RBC) bank overview

In the BFSI sector, a client's credit score, geography, tenure, and estimated salary are part of the data utilized to create marketing campaigns for their clients. We focus on churn analysis to get some insights into how certain factors can improve targeted campaigns to onboard new customers and retain the current ones. We find out that the age and location of clients are the main churn factors; clients around the age of 30 to 50 exit more from the banking service; this indicates that older people are more loyal, maybe due to having a more stable income, contentment with the banking services like mortgages, loans or retirement plans etc. The location is the other factor, which might be due to the user experience

provided by the bank in one location and how it differs from another – the efficacy of their services and how the end user enjoys them.

To engage customers in using the loyalty program to close the frontier between points accumulation and redemption to ensure customer retention, elements like improving customer experience should be employed where the focus should not be on ensuring that customers amass and redeem points but on improving their experience when they use their services (Idugboe, 2014). Idugboe (2014, p. 57) also suggests that profitable customers should be recognised and targeted specially, using Southwest Airlines as an example where air mileage points are rewarded based on the amount per ticket purchased instead of the conventional usage of travel miles post a long accumulation period.

### 3.3. Key Intelligence Question 3

In the USA, 60 percent of the average consumer prefers using digital coupons, with the remaining 40 percent taking paper coupons as their preferred form of reward. Furthermore, in 2023, a study conducted on consumer preferences (from 18 years and older) in ways of accessing loyalty rewards divulged that the younger generation (Gen X, Millennials and Gen Z) preferred using digital coupons while the older ones preferred paper coupons, as shown in Figure 14. From a sample size of 988 respondents, 57% represent those in favour of digital coupons and 43% for paper coupons. (Researchscape, 2023).

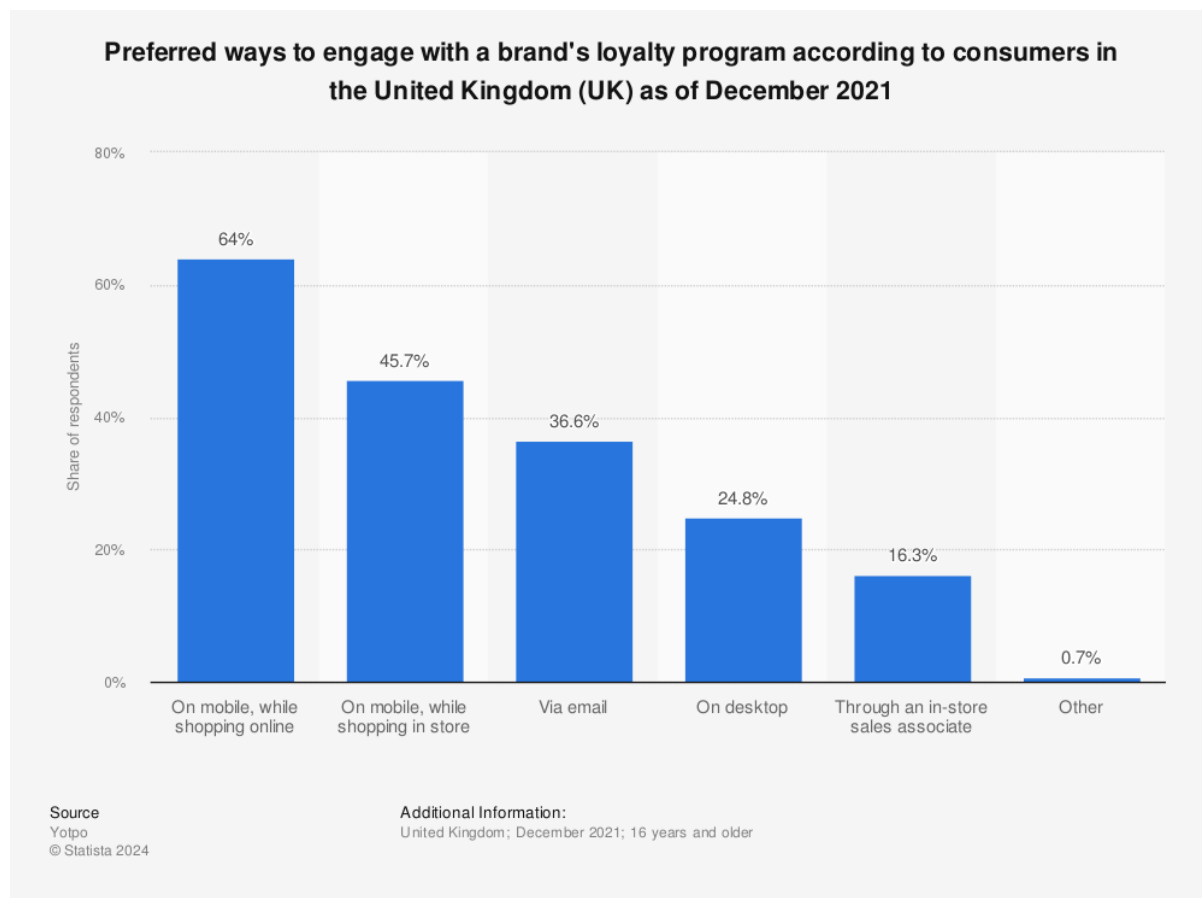
**Note. From Statista Share of shoppers using digital vs. paper coupons in the United States as of 2023, by generational cohort (Researchscape, 2023)**



**Fig. 14. Share of shoppers using digital vs. paper coupons in the United States as of 2023, by generational cohort**

In the United Kingdom, retailers offer reward programs for loyalty differently: A mobile application has the upper hand over physical reward cards or coupons; this is because all of a customer's data can be stored, accessed and analysed centrally on the business side and consumers largely prefer to use their mobile devices while shopping in-store or online as seen in Figure 15 below.

**Note. From Statista Preferred ways to engage with a brand’s loyalty program according to consumers in the United Kingdom (UK) as of December 2021 (Yetpo, 2022)**



**Fig. 15. Preferred ways to engage with a brand’s loyalty program according to consumers in the United Kingdom (UK) as of December 2021**

Businesses like Tesco, Sainsbury’s and COOP explore their competitive advantage in loyalty schemes by developing and releasing mobile applications for their customers to find new offers and personalised discounts in real time. Environmentally, they reduce the use of paper and plastic materials and save costs from postage vis-a-vis saving the customers’ time to receive the posted item(s).

With the rise of Artificial Intelligence (AI), “AI can also improve loyalty programs with innovations like voice activation to use the loyalty app hands-free, or facial recognition, used by Kentucky Fried Chicken (KFC) with their “Smile to Pay” feature.” (Statista Research Department, 2023). Businesses can leverage on these especially AI, by using it in ways targeted towards their business goals to achieve the desired outcomes.

Some businesses across industries (like fitness, retail etc) use gamification to engage their customers. While the customer has fun completing missions **or engaging in fitness exercises to earn** points from playing a game, they also engage with the brand through steady and frequent participation and competition. (Mastercard, 2023).

By replicating these methods leveraged by some businesses, a broader knowledge about the customers is acquired. Customers playing a game that is informative and fun, provided by a company,

helps them to understand how to better deliver services and improve upon their customer experience; brand loyalty through retainment would not only be achieved but referral via word of mouth is spread around too, thus increasing acquisition and growing market share and position.

The competitive advantages deployed by a business also enable successful data monetization. These insights are derived from the company's internal departments to optimize things like cost to enable the business to understand where to channel funds increment or decrement on costs related to production, orders, deliveries, inventory, etc. Profits can also be forecasted, and trends regarding what consumers want more could be analysed. (Masternak, n.d).

### 3.4. Key Intelligence Question 4

The purpose of data protection laws is to ensure that private data about individuals is used correctly. The law gives individuals rights regarding what the data is about, the data subjects, and the obligations of those who possess that personal data.

Recognizing the value of preserving data for historical purposes, the Data Protection Act 2018 (DPA) and the General Data Protection Regulations, which went into effect in March 2018, include provisions for doing so within the bounds of the law.

According to the DPA, organizations that hold personal data are required to register it with the Information Commissioner, adhere to data protection standards, and provide persons with access to, and, in certain cases, the ability to correct and remove, personal data about them. The Data Protection Act of 1998 was expanded to cover all information about live people stored by public agencies, regardless of the format or structure of the records, with the implementation of the Freedom of Information Act (FOIA) in 2005. This extension of the act's scope is maintained under the 2018 Act (Kew et al., 2019).

The General Data Protection Regulation is based on the notion of accountability. GDPR mandates that businesses and organizations adhere to all data protection standards and provide proof of compliance. Several tools are available to corporations and organizations under the GDPR to aid in demonstrating accountability, and they must be legally implemented.

A business is allowed to follow a code of conduct created by a DPA-approved business association. An implementing act of the Commission may confer European Union (EU)-wide legitimacy onto a code of conduct.

They may adhere to a certification mechanism administered by one of the certifying bodies that have acquired accreditation from a DPA or a national accrediting authority or both, as decided in each EU Member State.

Since certification and codes of conduct are optional tools, it is up to the business or organization to choose whether to apply for certification or follow a particular code of conduct. While adhering to such instruments may be taken into account in the event that you are subject to an enforcement action for violating the GDPR, your company or organization must nonetheless respect and comply with the GDPR (GDPR, 2016).

Leveraging predictive AI (also known as Artificial Intelligence and Machine Learning) is a security measure that can help build the trust of consumers not only in the loyalty program but in the company itself. The use of predictive AI to monitor and detect unusual patterns and behaviours could be insightful for misuse or fraud detection; AI can identify hackers with the aim of tricking the loyalty system or employees of a company/business who, as a gesture, offer their points to family and friends to use. Developing an alert system for an unusually high volume of points collections or transactions at certain times of the day etc, is a good step towards achieving this. A good example of this is how the popular ride-hailing company, Lyft, uses AI to detect suspicious behaviours like inconsistencies with the user's regular patterns such as the use of multiple accounts to earn rewards or a spike in the

number of rides taken; and when the data team confirms such activity, the account is restricted off the platform or the discount is denied.

Such security measures enable the customers to understand that their data is not only collected for profit-based mining but care is given to it too.

Also, to ensure a company achieve their compliance objectives around data privacy and information security, the implementation of ISO 27001 (the information security management standard) should be in place. This standard is beneficial because it demonstrates information security best practices, helps organizations meet the various data protection law requirements around the world, and helps avoid any fines associated with a data breach, it is also considered the 'gold standard' for designing a comprehensive set of security controls by many organizations (Vigilant software, n.d).

A business/company is advised to enact the use of this standard as it sets out guidelines for an optimal and holistic approach towards information security management.

### 3.5. Key Intelligence Question 5

Businesses need to take action to protect customer data and make their loyalty programs open, transparent and compliant with the law to ease customer concerns about data security. One way to do that is to have strong data security measures in place like monitoring, access controls and encryption. Also, businesses need to educate customers clearly and simply about how their data will be shared and used. Government and consumer protection agencies need to push for stronger laws with bigger penalties for data privacy and personal data (Eduina Maksuti et al., 2019). Companies in the EU (or those intending to venture in) that offer loyalty programs are required to take data privacy rules into account when designing both new and ongoing projects.

#### Data protection principles

1. Lawfulness, fairness and transparency: Any member's data must be processed fairly, lawfully and transparently.
2. Limitation of use: Only use data for the purposes it was collected for and made clear to the member at the time of collection.
3. Collect and process only what you have to and only when you have to.
4. Accuracy: Keep personal information up to date and correct.
5. Limitation of storage: Keep personally identifiable information only as long as you need, to achieve the purpose.
6. Integrity and confidentiality: Process in a way that ensures the necessary level of security, integrity and confidentiality (e.g. using encryption).
7. Accountability: The "data controller" or **data protection officer** of **an** organisation is responsible for proving it complies with all GDPR rules.

Businesses should implement technical and organizational measures for data protection. There are technical measures such as employing a platform provider using end-to-end encryption or requiring two-factor authentication for all staff on systems storing personal data. Organizational measures include staff training, **inputting** a data privacy policy in your employee handbook, or allowing only the personnel who require it to access personal data. Data Protection Impact Assessments (DPIA) and Data Protection Officers (DPO) are extra conditions that must be fulfilled for the loyalty program **on a solo or group level**. DPIAs are required of all

organizations for any processing that poses a high risk to a person's rights and freedoms. It is mandatory for some organizations to designate a Data Protection Officer to supervise compliance with data protection laws. The standard is a 72-hour notification period for a data breach, or else a risk of huge fines is posed (Hunter Murray, 2024). The notification must include details such as the nature of the breach, the likely consequences, and the measures taken or proposed to address it (Kambourakis et al., 2021, Victor-Mgbachi, 2024).

## 4. CONCLUSION

### 4.1. Implications

Personal information that loads through loyalty cards has crucial insights not only about clients' preferences but also sensitive data about their health conditions when we focus on data collected in certain industries. For businesses, the use of loyalty card information consists of customer preferences and patterns. This is beneficial as companies solicit consumers with focused marketing plans, meeting the customers' expectations, which in turn increases the retention of those clients, hence achieving global competitiveness. Forecasting and analyzing consumer behaviour makes any firm appreciate their client's needs, therefore enabling the firm to spend resources on effective product development and customer service.

From the viewpoint of users, loyalty cards promote the availability of preferences in terms of the implementation of promotional programs for them and their satisfaction with the final product. On the other hand, the user will have to exchange it for a fair amount of their information, the ethics of which is questionable at least.

Companies should be less rigid and more transparent by explicitly informing their users on how their PII is collected, processed and used. Most companies/businesses vaguely provide this information, and they are often wrapped in ambiguity and lengthy words, which discourage customers/readers. They should oblige to accurate processing of personal data with the provision to adequately rectify inaccuracies if demanded by a user (within a Service Level Agreement period or earlier), to ensure the possession of accurate and up-to-date data, which categorizes as some of the characteristics of qualitative data. This creates a high level of accountability and a demonstration of compliance with data protection principles, which includes creating and updating detailed records of data collection and usage activities and conducting data protection impact assessments (Bakare et al., 2024).

The allocation of resources towards implementing technological solutions, data protection officers and legal counsel should be put in place to ensure that business activities that permeate the use of customers' data strictly adhere to regulatory requirements.

### 4.2. Limitations

There are some limitations to analyzing loyalty card data flow. First, while the data collected is useful in decision-making, it is mostly narrow and sometimes ignored with regard to system integration and synchronization across industries. This hinders the ability to effectively understand decisions made or the entire customer journey taken at different levels.

The risk of bias in the collection and/or analysis of the data is another downside. For instance, the program may be biased for or against certain customers (like loyal card users) hence, there is a distortion of the insights that do not cover the other customers.

Also, the degree of data protection is likely to differ from one industry to another; therefore, the adherence to privacy laws is likely to vary. Some industries may not have the capacity or the realization for such measures, posing data as more dangerous.

With the EU's GDPR and US data privacy regulations being the two strongest pillars and most popular that uphold the protection of user data, the diversity in the legal traditions and cultural perspectives in shaping data protection frameworks serves as a limitation to businesses. While the EU GDPR is a well-grounded and tested framework which serves its citizens within the European Union and other European states, the USA fuses federal and state laws whereby jurisdictions and enactments vary across the various states. "This raises challenges for businesses operating across multiple states, requiring them to navigate diverse legal frameworks." (Bakare et al., 2024). The need to create a

comprehensive framework that minimizes the use of various laws is paramount as it effectively enables the bodies responsible to ensure adherence to the set laws and makes it easier for companies to implement and use them, especially multinational corporations.

Regulators must also ensure that constant updates are made to currently existing regulations to keep up with rapid trends like AI and ML, Internet of Things (IoT), and Web3.0 to guarantee that as data collection and its engineering quickly grows, these laws protect those to whom it belongs to, and are used appropriately.

### 4.3. Significance

Being aware of how data flows from loyalty cards is important because it helps predict the future of consumer-business interaction and provides an understanding of the need for data regulation. Knowing very well the cycle of loyalty card data collection, usage, and distribution enables companies to come up with more efficient, responsible, yet compliant advertising plans. Such not only improves customer satisfaction and loyalty but also helps maintain the customer's privacy.

From a regulatory standpoint, emphasis has shifted towards the recognition of the essentials for tougher legislation on data protection e.g., the General Data Protection Regulation awareness. There is no need to elaborate on the adherence to these presented regulations because their breach would entail unimaginable fines and a loss of confidence from clients. Sectors, where customers' data are used, are expanding; firms need to learn how to handle such information effectively without compromising the privacy and security of customer data.

Understanding your client's needs and preferences from their behavioural (transactional and demographic) data is a cheat code that startups and businesses can use due to this data being an internal possession. SMEs and even larger companies can leverage these large volumes of data by analyzing and incorporating them with predictive AI to engage in the right Research and Development (RnD) to innovate and develop new products or services therefore increasing their profit and gaining a solid market position. (Bharadiya, 2023). Data analytics of loyalty program data can effectively control how resources should also be allocated to areas within the business where they are needed the most, which would, in turn, reduce business costs and improve operational efficiency.

Also, a growing small business or startup would profit from running an optimal inventory where available stocks align with the predicted demand. The use of machine learning algorithms can help achieve this through demand forecasting and dynamic pricing (leveraging real-time market data and customer behaviours).

Gamification massively reduces customer churn with added benefits. Games such as trivia, puzzles, treasure hunts, and quizzes about the company and its products or services for points (on correctly answered questions) not only engage and teach your customers but create a memorable experience, making them come back to engage with your business and all it offers.

While the business is envisioned to thrive, due diligence in the form of setting up and respecting data protection and governance practices within the company and all third parties involved should be planned and executed, this shows an understanding of the need for transparency and mindfulness for the use of customer's personal data.

Ultimately, this paper proposes a framework for businesses (small or large scale) to adopt in order to create a loyalty program as part of their competitive and business intelligence strategy to gain market share and a vantage point over their competitors. The use of predictive AI coupled with data engineering, analytics processes and gamification as technology models are suggested and their benefits, as well as data privacy laws they must follow to secure trust and avoid regulatory problems.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

## Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc. have been used during the writing or editing of manuscripts. This explanation will include the name, version, model, and source of the generative AI technology and as well as all input prompts provided to the generative AI technology

Details of the AI usage are given below:

- 1.
- 2.
- 3.

## References

1. Candice L. Kline. (2007). Security Theater and Database-Driven Information Markets: A Case for an Omnibus U.S. Data Privacy Statute, 39 U. TOL. L. REV. 443, 448.
2. Jonathan Zhang, Fletcher Mangum, David Zorn. "Know Your Customer: How Retailers Have Used Data Throughout History". 2023.06.21.
3. Raksha Sharma. (2024). Loyalty Program Market. DATAINTELO. <https://dataintelo.com/report/loyalty-program-market>.
4. Kew, Richmond, Surrey, TW9 4DU. (2019) . The National Archives. "Data Protection and Personal Information."
5. GDPR. (2016). Article 24, Articles 40 to 43 and Article 83 and Recitals (98), (99), (100), (148), (150) and (151). "How can I demonstrate that my organisation is compliant with the GDPR?"
6. Sandra Idugboe. (2014). Loyalty programs and their effectiveness. [https://vskp.vse.cz/english/58407\\_loyalty-programs-and-their-effectiveness-from-the-company-point-of-view?title=loyalty&page=1](https://vskp.vse.cz/english/58407_loyalty-programs-and-their-effectiveness-from-the-company-point-of-view?title=loyalty&page=1).
7. Eduina Maksuti, Bledar Maksuti. (2023). "Application of Customer loyalty programs and issues of data protection."
8. Hunter Murray. (2024). "Unpacking the Data Dilemma: A Guide to GDPR and Privacy in Loyalty Programs."
9. Researchscape. (February 5, 2023). Share of shoppers using digital vs. paper coupons in the United States as of 2023, by generational cohort [Graph]. In *Statista*. Retrieved October 02, 2024, from <https://www.statista.com/statistics/1389319/digital-paper-coupons-usage-us-generation/>.
10. Statista Research Department. (December 19, 2023). Loyalty programs & marketing in the United Kingdom – statistics & facts. Statista. <https://www.statista.com/topics/3961/coupons-and-loyalty-schemes-united-kingdom-uk/#topicOverview>.
11. Yotpo. (November 13, 2022). Preferred ways to engage with a brand's loyalty program according to consumers in the United Kingdom (UK) as of December 2021 [Graph]. In *Statista*. Retrieved October 02, 2024, from <https://www.statista.com/statistics/1399792/ways-engage-loyalty-program-uk/>.
12. Mastercard. (2023). The impact of gamification on loyalty strategies. Mastercard. <https://www.mastercardservices.com/en/advisors/consumer-engagement-loyalty-consulting/insights/impact-gamification-loyalty-strategies>.
13. Weronika Masternak. (n.d). Loyalty program benefits: How a loyalty program can grow your business. OPEN LOYALTY. <https://www.openloyalty.io/insider/loyalty-program-benefits-how-a-loyalty-program-can-grow-your-business>.

14. Mark Herbert. (2020). Important Loyalty Program Statistics for 2024. Extu. <https://extu.com/blog/loyalty-program-statistics/>.
15. Joe Tidy. (2020). British Airways fined £20m over data breach. BBC. <https://www.bbc.com/news/technology-54568784>.
16. Mary Pilecki. (2024). Can AI Improve Your Loyalty Program? FORRESTER. <https://www.forrester.com/blogs/can-ai-improve-your-loyalty-program/>.
17. Tamara Thurman. (2023). How AI Will Shape the Future of Loyalty Programs. stamp Me. <https://www.stampme.com/blog/ai-loyalty-programs>.
18. Vigilant software. (n.d). ISO 27001 – the Information Security Management Standard. <https://www.vigilantsoftware.co.uk/topic/whatisiso27001>.
19. Kambourakis, G., Neisse, R., & Nai-Fovino, I. (2021). Information security in the age of EU-Institutions digitalisation, a landscape analysis.
20. Victor-Mgbachi, T. O. Y. I. N. (2024). Navigating cybersecurity beyond compliance: understanding your threat landscape and vulnerabilities.
21. Bakare, S. S., Adeniyi, A. O., Akpuokwe, C. U., & Eneh, N. E. (2024). Data privacy laws and compliance: a comparative review of the EU GDPR and USA regulations. *Computer Science & IT Research Journal*, 5(3), 528-543.
22. Bharadiya, Jasmin. (2023). Machine Learning and AI in Business Intelligence: Trends and Opportunities. *International Journal of Computer (IJC)*. 123-134.

UNDER PEER REVIEW