

## Executive Summary

### Objective of the EDA

- The objective of this Exploratory Data Analysis (EDA) is to discover insights into the *performance* and *adoption* rates for membership of various stores within a retail chain. This analysis aims to describe historical data that can help the business define acceptable KPI metrics for conversion and adoption rates in different stores across the country, in the same line, the goal is to identify the correlation between the conversion and adoption rates and understand the big picture of the general performances of the stores across the country

### Data Overview

- The dataset comprises **300** records, with five key columns:
  - store\_id: Identifier for each store.
  - week: Week number ranging from **1** to **5**.
  - conversion\_rate: The rate at which users are converted ranges from **0.011** to **0.419**.
  - adoption: The adoption rate of features within the store ranges from **0.003** to **0.966**.
  - province: The province where the store is located.

### Data Sanity Test

- The dataset consists of 300 rows and 5 columns.
- There are null or missing values present in the dataset.
- There are 3 unique values for provinces (**ON**, **BC**, **AB**).
- There are 60 unique stores in the dataset.
- The methodology applied to find outliers was 1.5 IQR.

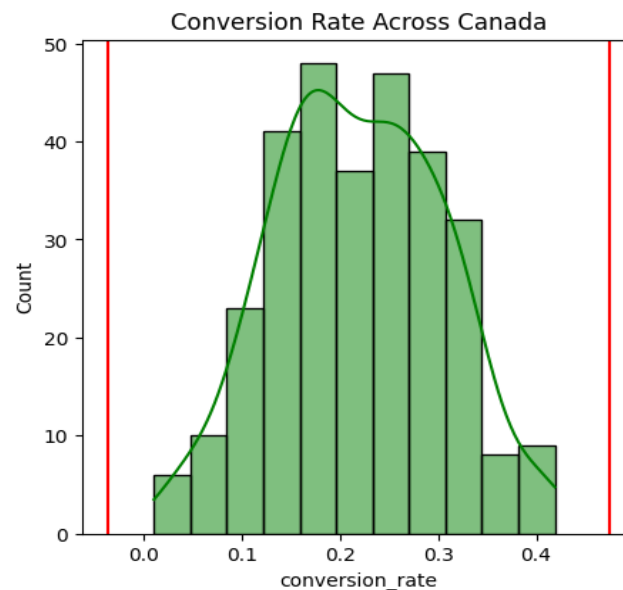
### Descriptive Statistic

	conversion_rate	adoption
count	300.000000	300.000000
mean	0.218540	0.27596
std	0.085436	0.20992
min	0.011000	0.00300
25%	0.155000	0.11375
50%	0.215000	0.22750
75%	0.282750	0.38450
max	0.419000	0.96600

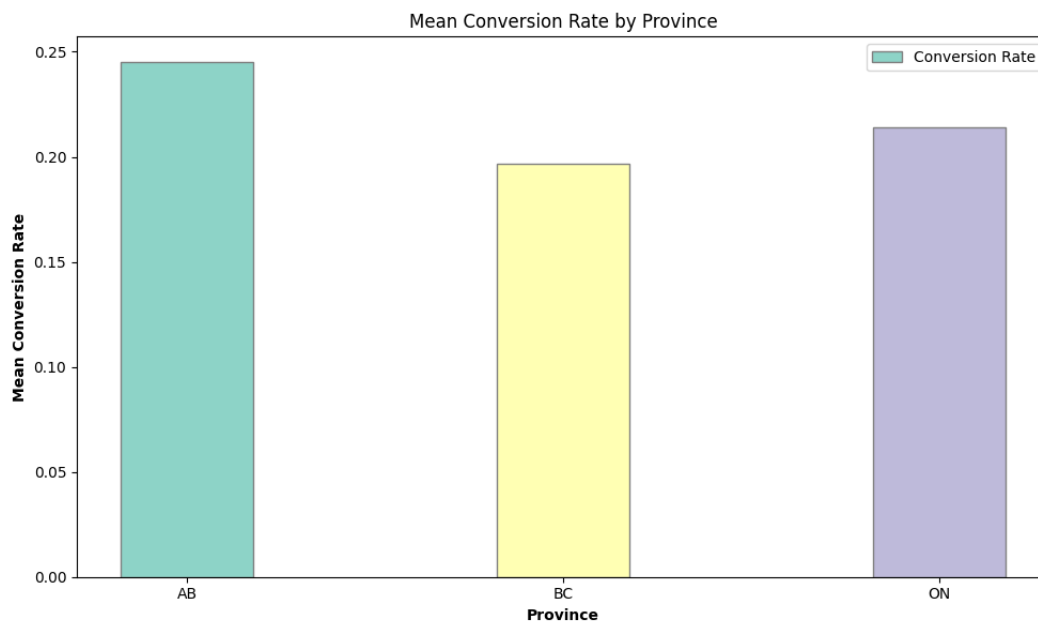
## Key Findings

### Conversion Rate Distribution:

- The overall conversion rate across all stores shows normal distribution. The conversion rates mostly fall between 0.1 and 0.3, as shown in the following histogram:



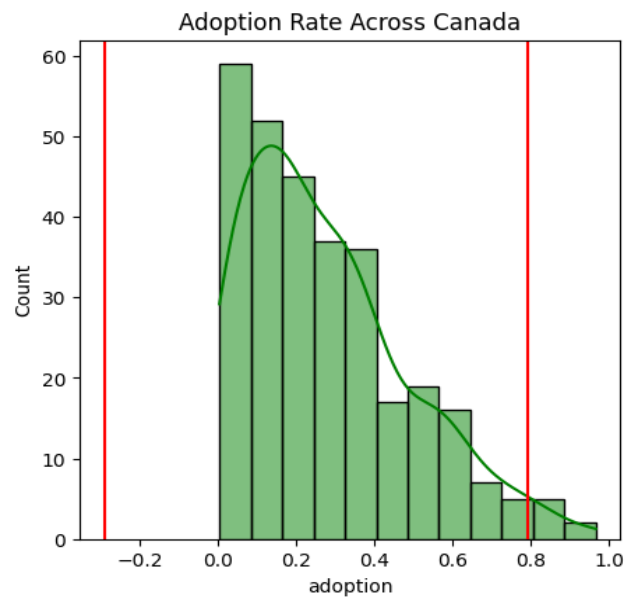
- There are no outliers presented on the conversion rate distribution.



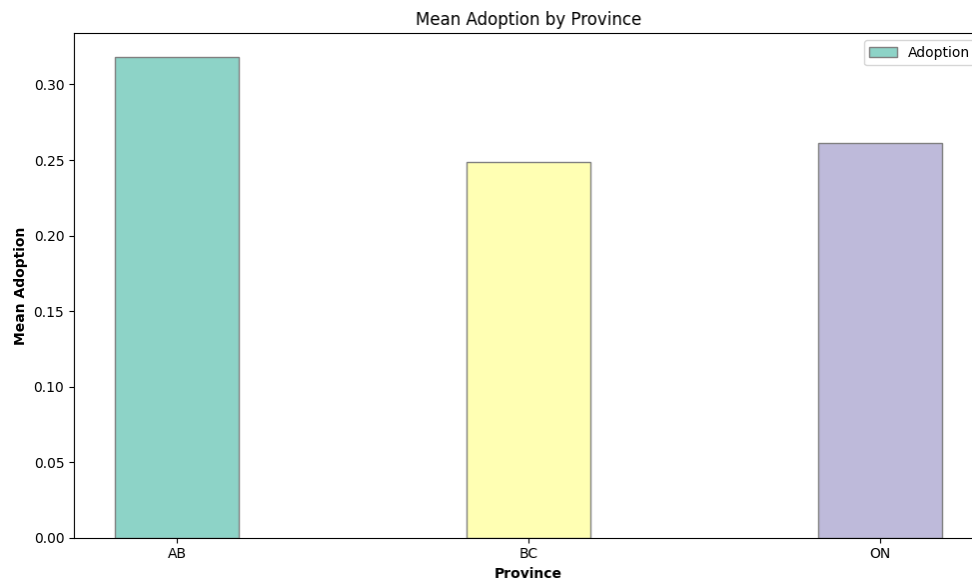
- Distribution by province reveals that Alberta (AB) has the highest average conversion rate of 0.24497, followed by Ontario (ON) with 0.21413, and British Columbia (BC) with 0.19652.

### Adoption Rate Distribution:

- Adoption rates tend to be right-skewed, meaning that most stores have lower adoption rates, but a few stores have much higher adoption rates. This pattern is clearly shown in the adoption rate histogram.

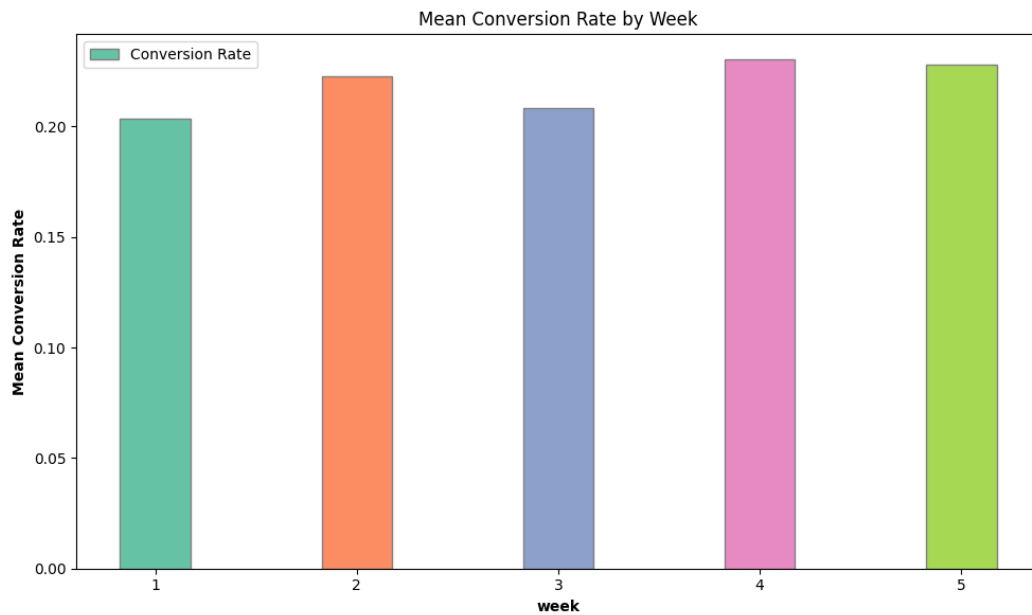


- Outliers are presented in the adoption rate distribution. **It is pending confirmation from the team if we should exclude those from the analysis.**

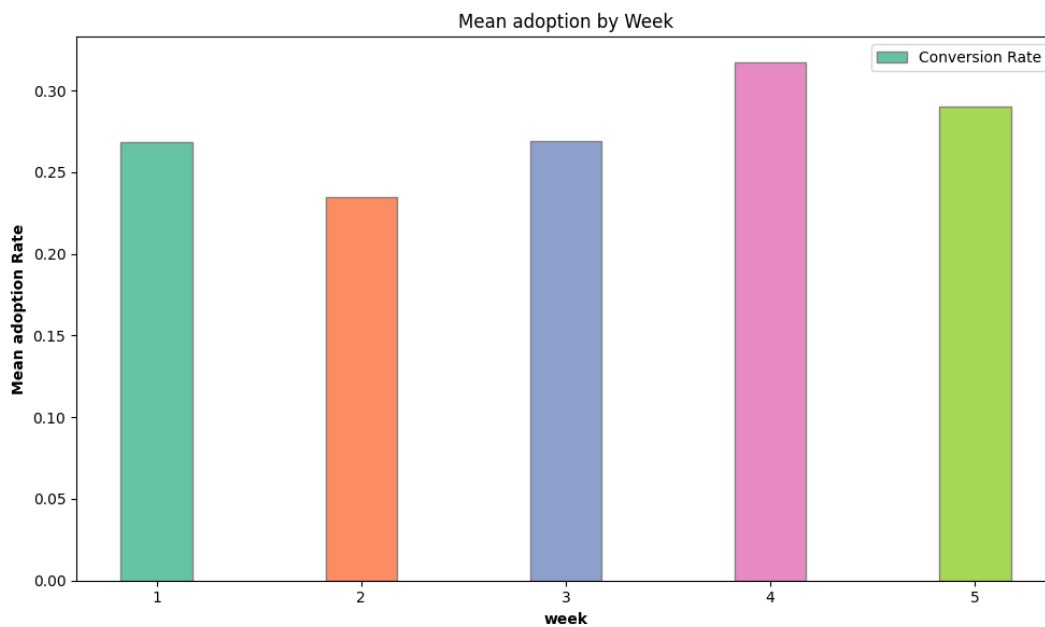


- Similar to conversion rates, Alberta (AB) leads with an average adoption rate of 0.31796, followed by Ontario (ON) with 0.26125, and British Columbia (BC) with 0.24867.

## Store and Week Analysis:

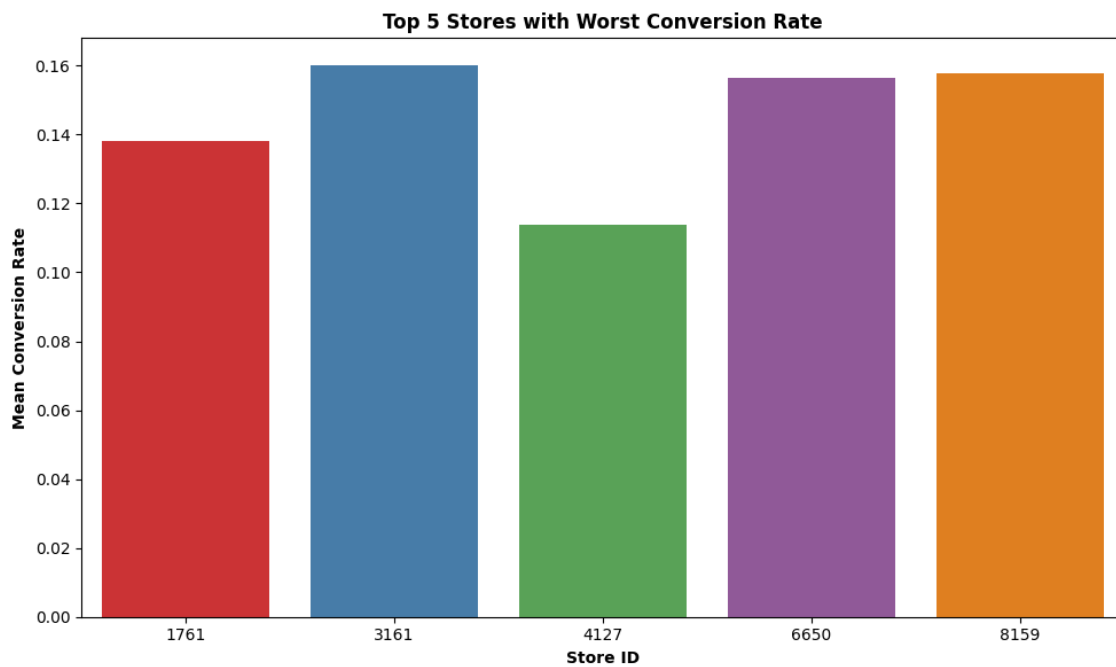
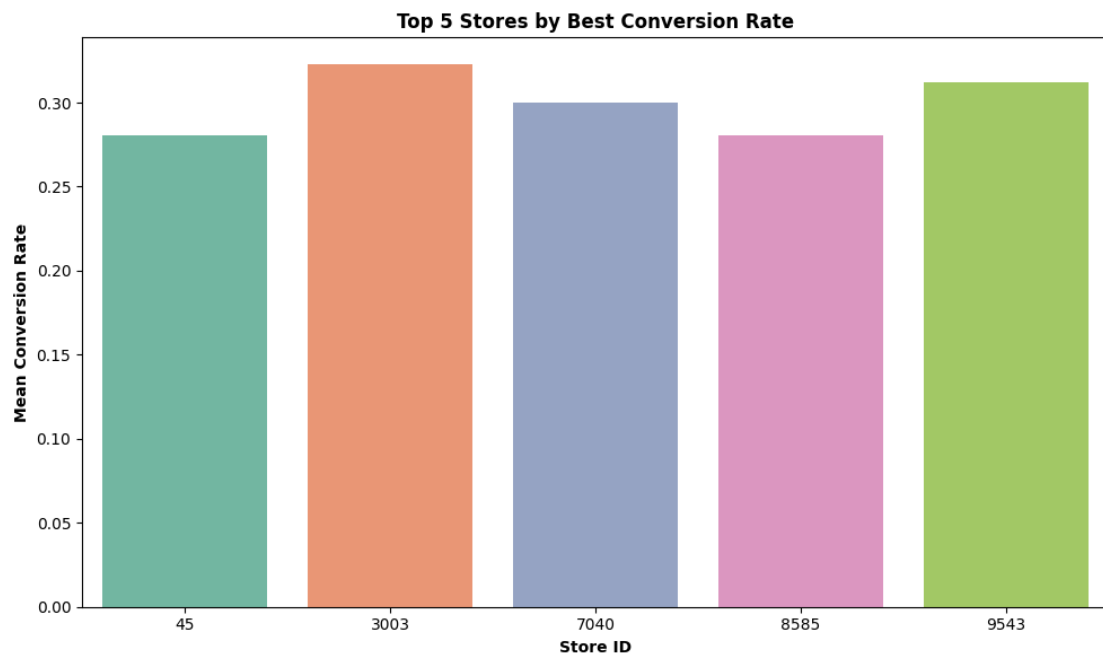


- Weekly analysis indicates that weeks 2 and 4 show the highest conversion rates. **It is pending the team's confirmation if certain promotional activities or external factors influenced user behavior during these weeks.**



- The analysis shows that weeks 3 and 4 have the highest adoption rates. Based on both graphs, we can infer that there is not a strong correlation between adoption and conversion rates.

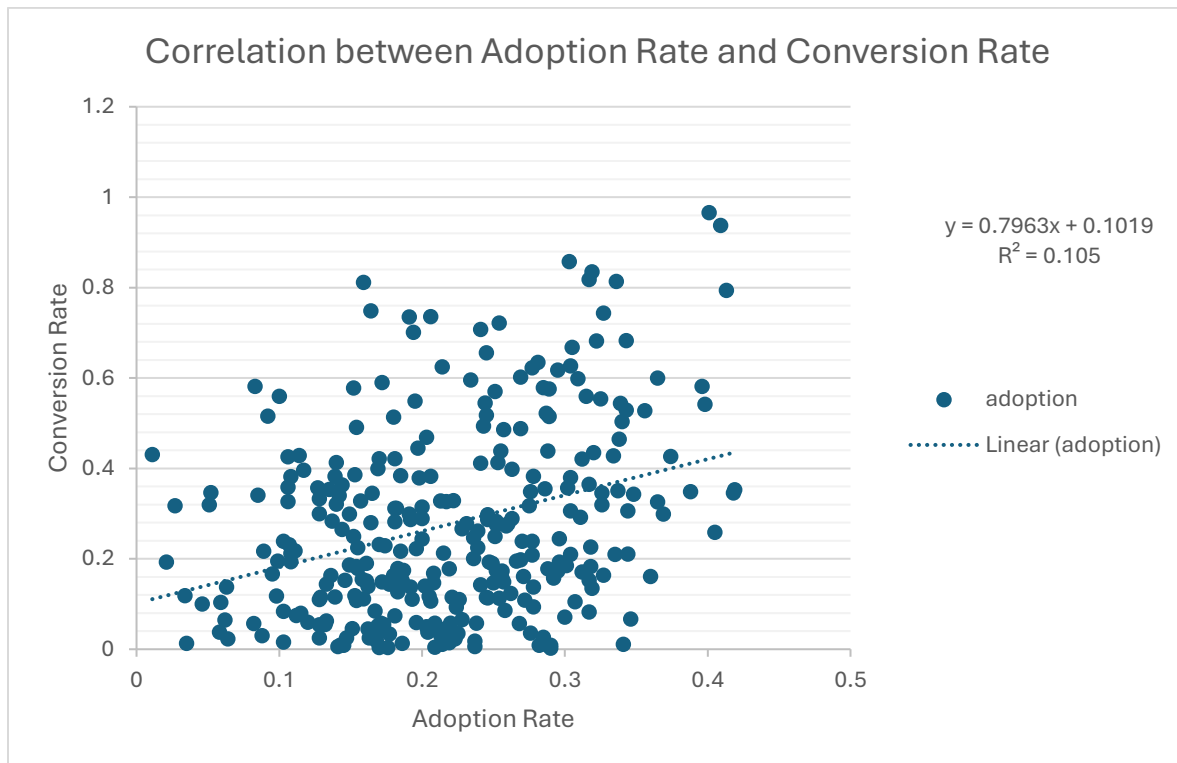
## Top Performing Stores:



- The top 5 stores by best conversion rates and adoption rates have been identified. Store ID 3003 has the highest conversion rate, while Store ID 251 leads in adoption rates.

## Correlation Analysis:

There is a positive correlation between the adoption and conversion rate, however, the dispersion of the data indicates that the precision and the accuracy of the model is weak.



The  $R^2$  value is as low as 0.105, which indicates a very low correlation between the variables. In other words, the behavior of one variable cannot explain the behavior of the other.

The linear regression equation for this dataset is  $y=0.7963x+0.1019$ , where the conversion rate is the dependent variable and the adoption rate is the independent variable.

However, **this model should not be used for predictions**, as the dispersion of the data points and the low  $R^2$  value suggest that the relationship between the variables is weak and unreliable, leading to potential errors.

## Conclusions

- **Comparison of Conversion Rates by Province:**
  - Based on the comparison with Alberta's conversion rate of 0.24497, Ontario's conversion rate is approximately 12.59% lower, and British Columbia's conversion rate is about 19.78% lower. This significant difference suggests that Alberta's strategies are more effective, and by analyzing strategies implemented in Alberta, we may identify key practices that could improve conversion rates in Ontario and British Columbia.
- **Adoption Rate Analysis:**
  - Similar to conversion rates, Alberta (AB) leads with an average adoption rate of 0.31796. Ontario's adoption rate is approximately 17.83% lower, and British Columbia's adoption rate is about 21.78% lower.
- **Weekly Analysis:**
  - Weekly analysis indicates that weeks 2 and 4 exhibit the highest conversion rates.
  - The analysis shows that weeks 3 and 4 have the highest adoption rates. Based on both graphs, we can infer that **there is not a strong correlation between adoption and conversion rates.**
- **Performance Improvement:**
  - To improve overall performance, it is ideal to select metrics of the top-performing stores and set them as targets for other stores.
  - Additionally, reviewing the performance of the lowest-performing stores will help identify issues and areas for improvement.
- **Correlation and Model Reliability:**
  - The dispersion of the data points and the low  $R^2$  value of 0.105 suggest that the relationship between the conversion and adoption rates is weak and unreliable, leading to potential errors. The linear regression equation  $y=0.7963x+0.1019$  should not be used for predictions due to this low level of correlation.