**EXPLORATION OF DATA SKILLS AND ISSUES**

*CASE STUDY ON A MARKETING CAMPAIGN DATASET FROM A TELECOMMUNICATION COMPANY*

Contents

[List of figures 2](#_Toc144034290)

[I. Data preprocessing: 3](#_Toc144034291)

[II. Univariate analysis: 4](#_Toc144034292)

[1. Continuous variables: 4](#_Toc144034293)

[2. Categorical variables: 8](#_Toc144034294)

[III. Relationships Between Predictors: 14](#_Toc144034295)

[IV. Variable relationships over time: 15](#_Toc144034296)

[V. Key findings: 29](#_Toc144034301)

[REFERENCES 32](#_Toc144034302)

# **List of figures**

[***Figure 1. Summary of Dataset Columns*** *3*](#_Toc144028296)

[***Figure 2. number employed - quarterly indicator (numeric)*** *4*](#_Toc144028297)

[***Figure 3. euribor 3 month rate - daily indicator (numeric)*** *4*](#_Toc144028298)

[***Figure 4. Age*** *5*](#_Toc144028299)

[***Figure 5. Number of contacts performed during this campaign.*** *5*](#_Toc144028300)

[***Figure 6. Number of days that passed by after the client was last contacted from a previous campaign)*** *6*](#_Toc144028301)

[***Figure 7. Last contact duration, in seconds (numeric).*** *6*](#_Toc144028302)

[***Figure 8. consumer price index - monthly indicator (numeric)*** *7*](#_Toc144028303)

[***Figure 9. consumer confidence index monthly indicator (numeric)*** *7*](#_Toc144028304)

[***Figure 10. employment variation rate - quarterly indicator (numeric)*** *8*](#_Toc144028305)

[***Figure 11. Type of job*** *8*](#_Toc144028306)

[***Figure 12. Marital status*** *9*](#_Toc144028307)

[***Figure 13. Level of education*** *9*](#_Toc144028308)

[***Figure 14. Has credit in default?*** *10*](#_Toc144028309)

[***Figure 15. Has a housing loan?*** *10*](#_Toc144028310)

[***Figure 16. Has a personal loan?*** *10*](#_Toc144028311)

[***Figure 17. Month of contact*** *11*](#_Toc144028312)

[***Figure 18. Contact communication types.*** *11*](#_Toc144028313)

[***Figure 19. weekdays of contact*** *12*](#_Toc144028314)

[***Figure 20. Did the client subscribe to a Telecom plan?*** *12*](#_Toc144028315)

[***Figure 21. Number of contacts performed before this campaign and for this client.*** *13*](#_Toc144028316)

[***Figure 22. Outcome of the previous marketing campaign*** *13*](#_Toc144028317)

[***Figure 23. Correlation matrix of numerical variables*** *14*](#_Toc144028318)

[***Figure 24. Percentage Distribution of Education Level on a Monthly Basis*** *15*](#_Toc144028319)

[***Figure 25. Percentage Distribution of Jobs on a Monthly Basis*** *15*](#_Toc144028320)

[***Figure 26. Percentage Distribution of Education Level for Individuals with Credit Default on a Monthly Basis*** *15*](#_Toc144028321)

[***Figure 27. Percentage Distribution of Job for Individuals with Credit Default on a Monthly Basis*** *15*](#_Toc144028322)

[***Figure 28. Percentage Distribution of telecom subscriptions on a monthly basis*** *16*](#_Toc144028323)

***Figure 29. Average number of CPI, EURIBOR3M, Employee Variation Rate, Subscription Rate, and rate of default*** *16*

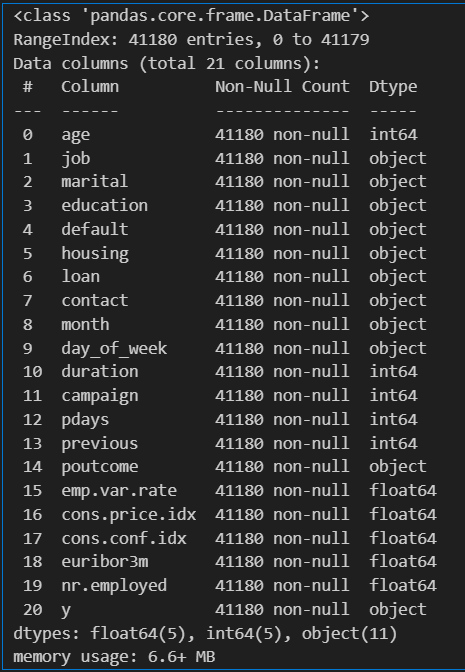
***Figure 30. Phases of economic cycle*** *17*

***Figure 31. Summarized five phases of a business cycle.*** *17*

***Figure 32. Education distribution of people having credit default (percentage by month)*** *19*

# **Data preprocessing:**

In the preparation of EDA on a telecommunication company’s marketing campaign dataset, preprocessing steps are performed. Double quotes will be eliminated, then cleaned data will be loaded into a panda’s data frame by using String IO.



**Figure 1. Summary of Dataset Columns**

1. **Univariate analysis:**
2. Continuous variables:

|  |  |
| --- | --- |
|  | **A screenshot of a computer  Description automatically generated** |

**Figure 2. number employed - quarterly indicator (numeric)**

50% of employed individuals fall within 5191 to 5228. The data is highly skewed

|  |  |
| --- | --- |
| **A screenshot of a computer  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 3. euribor 3 month rate - daily indicator (numeric)**

%50 of euribor3m range from 4.857 to 5.045 with a solid right skewed.

|  |  |
| --- | --- |
| **A screenshot of a graph  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 4. Age**

Age has 468 outliers above 69, suggesting a distinct subgroup.

|  |  |
| --- | --- |
|  | **A screenshot of a computer  Description automatically generated** |

**Figure 5. Number of contacts performed during this campaign.**

Most of the number of contacts falls below 6, even though the range is 1 to 37, showing numerous outliers.

|  |  |
| --- | --- |
|  | **A screenshot of a computer  Description automatically generated** |

**Figure 6. Number of days that passed by after the client was last contacted from a previous campaign**

‘999’ (96% of pdays) is often used for missing values rather than an actual value, so it is recommended that we should not use pdays for analysis.

|  |  |
| --- | --- |
|  |  |

**Figure 7. Last contact duration, in seconds (numeric).**

Duration has many outliers. Most values are below 644 seconds, with a median of 180 seconds.

|  |  |
| --- | --- |
| **A screenshot of a computer  Description automatically generated** | **A screenshot of a computer** |

**Figure 8. consumer price index - monthly indicator (numeric)**

|  |  |
| --- | --- |
|  | **A screenshot of a computer** |

**Figure 9. consumer confidence index monthly indicator (numeric)**

***cons.conf.idx*** are negative, indicating decreased confidence, while all ***cons.price.idx*** are positive.

|  |  |
| --- | --- |
| **A screenshot of a graph  Description automatically generated** | **A screenshot of a computer** |

**Figure 10. employment variation rate - quarterly indicator (numeric)**

The ***emp.var.rate*** spread suggests economic fluctuations. Median (1.1) and upper quartile (1.4) are positive, implying employment increase, possibly tied to economic growth.

1. Categorical variables:

|  |  |
| --- | --- |
| **A pie chart with different colors  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 11. Type of job**

Most surveyed individuals are administrators, blue-collar, or technicians, totaling around 64%.

|  |  |
| --- | --- |
| **A screenshot of a graph** | **A screenshot of a computer  Description automatically generated** |

**Figure 12. Marital status**

The analysis reveals a substantial 63% married, 28.1% unmarried, and less than 10% divorced.

|  |  |
| --- | --- |
| **A screenshot of a graph  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 13. Level of education**

Over 50% surveyed have high school degrees or higher, with a small minority being illiterate.

|  |  |
| --- | --- |
| **A screenshot of a graph  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 14. Has credit in default?**

|  |  |
| --- | --- |
| **A screenshot of a pie chart  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 15. Has a housing loan?**

|  |  |
| --- | --- |
| **A screenshot of a graph  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 16. Has a personal loan?**

While 82.4% lack personal loans and 79.1% have no credit defaults, over 50% hold home loans.

|  |  |
| --- | --- |
| **A pie chart with different colors  Description automatically generated** | **A screenshot of a computer** |

**Figure 17. Month of contact**

Most surveys were conducted between May and August, totaling over 78%.

|  |  |
| --- | --- |
| **A screenshot of a pie chart  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 18. Contact communication types.**

The majority, 63.5%, were contacted through cellular communication, while the rest were reached by traditional phone methods.

|  |  |
| --- | --- |
| **A pie chart with different colored circles  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 19. weekdays of contact**

Customers were contacted evenly on weekdays.

|  |  |
| --- | --- |
| **A screenshot of a pie chart  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 20. Did the client subscribe to a Telecom plan?**

|  |  |
| --- | --- |
| **A screenshot of a pie chart  Description automatically generated** | **A screenshot of a computer  Description automatically generated** |

**Figure 21. Number of contacts performed before this campaign and for this client.**

|  |  |
| --- | --- |
| **A screenshot of a pie chart  Description automatically generated** | **A screenshot of a computer** |

**Figure 22. Outcome of the previous marketing campaign**

New customers prevailed, lacking prior campaign data with existing ones. Hence, outcomes might unveil novel patterns.

1. **Relationships Between Predictors:**

**A screenshot of a graph

Description automatically generated**

**Figure 23. Correlation matrix of numerical variables**

The correlation matrix reveals significant correlations between **emp.var.rate**, **nr.employed, EURIBOR3M**, and **cons.price.idx**. These indicators, sensitive to economic changes, offer insights into customer behaviors.

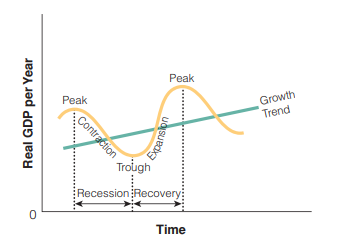
1. **Variable relationships over time:**

|  |  |
| --- | --- |
| A graph of a number of numbers  Description automatically generated with medium confidence  **Figure 24. Percentage Distribution of Education Level on a Monthly Basis** | A graph with different colored lines  Description automatically generated  **Figure 25. Percentage Distribution of Jobs on a Monthly Basis** |

|  |  |
| --- | --- |
| A graph of credit failure  Description automatically generated with medium confidence  ***Figure 26. Percentage Distribution of Education Level for Individuals with Credit Default on a Monthly Basis*** | ***A graph with a number of months  Description automatically generated with medium confidence***  ***Figure 27. Percentage Distribution of Job for Individuals with Credit Default on a Monthly Basis*** |

|  |
| --- |
| A graph with numbers and a number of months  Description automatically generated with medium confidence  ***Figure 28. Percentage Distribution of telecom subscriptions monthly*** |
| ***Figure 29. Average number of CPI, EURIBOR3M, Employee Variation Rate, Subscription Rate,***  ***and rate of default monthly*** |

Monthly averages show the business cycle’s impact on many aspects. Despite timing differences, core principles persist, revealing economic effects on consumer behavior. The analysis uncovers economic factors' influence on consumer behavior during phases.



**Figure 30. Phases of economic cycle**

(Sexton, 2010)

**Figure 31. Summarized five phases of a business cycle.**

1. March-April: **Expansion Phase: Emerging Optimism**
2. Characteristics:

* Economic Indicators: CPI, CCI, EURIBOR3M, and employee variation rate increased. This suggests an overall positive economic environment, potentially indicating increased consumer spending and business confidence.
* Default Rate: The percentage number of people having default increased sharply. This might be due to various factors, such as changes in employment stability or financial difficulties for borrowers.
* Occupation and Education Changes: The decrease in the percentage of people in admin, technician, retired, and university degree categories, along with the sharp increase in the percentage of blue-collar workers, could indicate shifts in the job market, potentially influenced by economic conditions or industry changes.
* Subscribe: The average number of people subscribing to a Telecom plan decreased considerably.

1. Analysis:

During this period, the increase in economic indicators like the Consumer Price Index (CPI) and Consumer Confidence Index (CCI) is evidence of a surge in consumer optimism and spending, possibly pointing towards increased consumer spending and greater business confidence. However, the simultaneous increase in the default rate could be attributed to underlying financial difficulties faced by certain segments of the population despite the overall economic positivity.

The occupation and education changes, including a decrease in admin, technician, retired, and university degree categories, along with a spike in blue-collar workers, may reflect a realignment of the job market due to economic influences. The slight decrease in telecom subscriptions could be linked to financial caution among consumers, despite the overall positive economic environment.

1. April to May: **Peak Phase: “Sentiment Shift”**
2. Characteristics:

* Economic Indicators: CCI and Subscribe variation rate increased sharply, while employee variation rate, CPI, EURIBOR3M collapsed.
* Default Rate: The percentage number of people having default plunged to a low.
* Occupation and Education Changes: The percentage of blue-collar workers peaked and the percentage of people holding university degrees bottomed out could reflect changes in employment patterns, possibly due to shifts in demand for certain skills or industries.
* Subscribe: The average number of people subscribing to Telecom shot.

1. Analysis:

The peak in the CCI along with the sharp decline in economic indicators like employee variation rate, CPI, and EURIBOR3M signifies a shift in sentiment. This change could be attributed to external factors causing uncertainty. The contrasting trends in default rates (a plunge) and occupation/education percentages (a peak in blue-collar workers and a bottoming out of university degree holders) might suggest a complex interaction between economic shifts and individual financial decisions. It seem that the economy is going to boom again.

The increase in telecom subscriptions during this phase could be a result of consumers’ confidence about the economy.

1. May-June: **Contraction Phase**: “**Volatility Unleashed**”
2. Characteristics:

* Economic Indicators: CCI and Subscribe variation rate went into free fall, while employee variation rate, CPI and EURIBOR3M bounced back.
* Default Rate: The percentage number of people having default remained at a low level.
* Occupation and Education Changes: The percentage of blue collar decreased slightly but it still maintained a high ratio. The percentage of admins increased moderately. The percentage of people holding university degrees went up slightly.
* Subscribe: The average number of people subscribing to Telecom collapse.

Reduced **CCI**, varying **CPI** and **EURIBOR3M** signal economic instability, prompting spending uncertainty. Rising university degrees and admin roles reflect job outlook and structure shifts. Subscriptions decline, emphasizing vital spending.

1. Analysis:

The plummeting CCI and rebounding indicators like employee variation rate and CPI highlight increased economic volatility. This period might be marked by uncertainty, leading to fluctuations in consumer spending and investment decisions. The increase in admin roles suggests a potential shift towards administrative or managerial positions, possibly driven by businesses restructuring in response to the economic environment. The rise in the percentage of people holding university degrees could reflect individuals seeking higher education to improve their job prospects in the uncertain economic climate.

In such times, consumers could prioritize essential expenses over discretionary ones, possibly leading to a decline in new subscriptions or upgrades for telecom services.

1. June-July: **Trough Phase**: “**Inflationary Respite**”
2. Characteristics:

* Economic Indicators: CPI peaked, and CCI increased slightly.
* Default Rate: Number of people having default remained a low level.
* Occupation and Education Changes: There was a slight drop in the percentage of blue-collar workers and a moderate increase in the percentage of admin people.
* The percentage of people holding a university degree was unchanged.
* Subscribe: The average number of people subscribing increased moderately.

1. Analysis:

The peak in CPI suggests a brief inflationary period, possibly driven by increased demand or supply chain disruptions. The decline in defaults and the increase in the CCI might indicate improved financial conditions and consumer sentiment. The shift in job categories might reflect a temporary rebound in certain sectors, while the overall decrease in blue-collar workers could hint at the effects of changing economic conditions.

The moderate increase in subscriptions despite the peak in CPI could be a result of individuals re-evaluating their priorities and realizing the value of the service after a period of reduced subscriptions. This might indicate that these services are viewed as essential or offering convenience and utility that consumers are unwilling to forego. These are products or services that consumers consider vital and necessary for their daily lives, and they are often prioritized even during challenging economic times.

1. July-August: **Trough Phase**: “**Stability and Rebound**”
2. Characteristics:

* Economic Indicators: The significant decline in CPI, EURIBOR3M, employee variation rate. It suggests potential economic contraction or stability.
* Default rate: The number of people having default remains at a low level.
* Occupation and Education Changes: The peak in admin, technician, and university degree holders might indicate certain sectors rebounding or stabilizing. This could be linked to the economic conditions.
* Subscribe: The average number of people subscribing increased rapidly.

1. Analysis:

In august all people having credit default get professional course and being

Technician

There are many reasons for it. First, technicians might have experienced income instability due to fluctuations in demand for their specific technical skills. This could have led to challenges in meeting debt obligations during certain periods. Second, it is possible that individuals who pursued professional courses and became technicians faced a mismatch between their education and the job market's demands. This mismatch could have led to lower income and financial strain, contributing to credit defaults. Third, people who pursued professional courses might have incurred educational debt. If this debt burden was coupled with difficulties in finding suitable employment, it could have led to financial stress and credit defaults.

Moreover, the decline in CPI and EURIBOR3M, along with increased confidence, suggests a period of relative stability or even slight contraction. The peak in admin, technician, and university degree holders might point toward specific sectors rebounding or stabilizing, potentially driven by an improved economic outlook. As economic indicators stabilize and show signs of recovery, consumer confidence could rebound. This might result in increased willingness to invest in additional services like telecom subscriptions, as people feel more secure in their financial situations. So, the subscription numbers increase slightly

1. August-September: **Recovery phase: “Bouncing Back”**
2. Characteristics:

* Economic Indicators: The rapid recovery of CPI, EURIBOR3M, employee variation rate. These indicators suggest improved consumer confidence, stable borrowing conditions, and potentially increased economic activity.
* Default rate: The number of people having default remains at a low level.
* Occupation and Education Changes: The drop in the percentage of people in the technician category could be a response to changes in demand for specific technical skills.
* Subscribe: The average number of people subscribing dropped dramatically.

1. Analysis:

The drop in technician roles might be due to a shift in demand for specific technical skills. The overall trends could signify a rapid adaptation to the changing economic landscape.

During periods of economic uncertainty or downturns, people might cut back on non-essential spending. As the economy starts recovering, there's often a release of pent-up demand for goods and services that were deferred during tougher times. Subscriptions to services like telecom plans might fall into this category. People who postponed upgrading their plans or signing up for new services due to financial concerns might now feel more comfortable spending on these non-essential activities.

1. September-October: **Recovery phase: “Cautious Optimism”**
2. Characteristics:

Economic Indicators: The drop in CPI and employee variations might signify an improving economic situation, potentially indicating a decrease in financial stress and inflation. However, the default rate increases dramatically.

Occupation and Education Changes: Stable job and education distribution might reflect a period of relative equilibrium in the job market and educational pursuits.

Subscribe: The average number of people subscribing increases slightly.

1. Analysis:

This behavior is in line with the notion that during times of economic uncertainty or recovery, consumers often prioritize essential expenditures over discretionary ones. As individuals emerge from challenging economic periods, they may be cautious about additional spending, particularly on non-essential services like telecom subscriptions.

There are two main reasons for the caution. First, it is the consequence of lagging effect. Even though the recession might officially be over, its effects can linger. Some individuals, households, and businesses might still be dealing with the financial repercussions of the recession, which could result in delayed defaults or financial difficulties. Second, debt accumulation could pose a threat to people’s financial health. During an economic downturn, people and businesses might accumulate debt to manage through the tough times. These debts could become problematic after the immediate crisis has passed, leading to an increase in defaults.

The intersection of lagging effects and debt accumulation illustrates that economic recovery is not a uniform experience for everyone. The broader economic trends can instill hope and optimism, but personal circumstances and financial burdens can introduce complexities that might slow down the process of regaining confidence.

1. October-November: **Recovery phase: “Consolidating Stability”**
2. Characteristics:

Economic Indicators: The continued decrease in the number of people having default, along with the collapsed EURIBOR3M and employee variation rate, suggests ongoing economic stability and improved financial conditions. These trends suggest that individuals are experiencing fewer financial difficulties and that credit markets are functioning more smoothly.

Occupation Changes: The increase in the percentage of people in management roles could indicate growth or expansion in sectors requiring management positions. It's likely that businesses are regaining confidence and are investing in management capacities to drive growth and navigate the stabilized economic environment.

Subscribe: The average number of people subscribing increases dramatically

1. Analysis:

The continued decrease in the default rate and indicators like EURIBOR3M and employee variation rate indicates ongoing economic stability. The increase in management positions could signify growth in sectors requiring leadership roles, potentially indicating further expansion.

As economic stability continues, consumers may be more receptive to long-term commitments like subscription plans. The company could emphasize the benefits of bundled services or loyalty programs to encourage customer retention and discourage churn. However, the subscriptions surprisingly dropped. The reason for it is the sudden increase of people having credit default. In term of total number of people being default, default incidents clustered in only August and November. Mature debt obligation may be the main reason.**A graph of credit failure

Description automatically generated with medium confidence**

**Figure 32. Education distribution of people having credit default (percentage by month)**

1. November-December: **Recovery phase: “Year-End Resilience”:**
2. Characteristics:

Economic Indicators: Slight increases in CPI and Employee Variation Rate, along with the highest CCI level, could suggest strong economic sentiment and potential growth.

Default Rate: The stable number of people having default at a low level might indicate that the worst of the financial difficulties has passed.

Occupation and Education Changes: it is stable. There is a considerable decrease in the percentage of people being management and blue collar. This shift might imply changes in industry demand or a strategic workforce restructuring in response to evolving economic factors.

Subscribe: The average number of people subscribing increases slightly.

1. Analysis:

The slight increases in CPI and Employee Variation Rate, coupled with the highest CCI level, suggest a strong year-end sentiment and potential economic growth. The stable default rate might indicate that the worst financial challenges have passed. The changes in management positions could be reflective of ongoing shifts in demographics and professional roles.

As economic conditions continue to improve, individuals may feel more confident in making non-essential expenditures, such as subscribing to telecom services. The slight increase in subscriptions suggests a growing willingness among consumers to invest in additional services.

1. **Conclusions:**
2. Interest Rates and Economic Conditions:

Interest rates, such as the EURIBOR3M, serve as a pivotal lever in economic policy, influencing borrowing costs and consumer behavior. The observed trends in your analysis demonstrate how changes in these rates can have cascading effects on various economic aspects. When interest rates rise, borrowing becomes more expensive, leading to reduced consumer spending and business investment. This could explain the collapse in economic indicators like the employee variation rate and the CPI during certain periods. Conversely, when interest rates are low, businesses are more likely to invest in expansion, potentially driving job creation and stimulating economic growth.

1. Impact on Job Choices:

Interest rates can greatly influence the demand for certain job categories. A surge in blue-collar workers suggests industries like manufacturing, construction, or logistics might be expanding during periods of economic positivity. This aligns with the notion that these sectors could be taking advantage of lower borrowing costs to fuel their growth. The decrease in admin and technician roles might be reflective of cost-cutting measures in response to tighter financial conditions, as well as automation or changes in business models.

1. Impact on Education Choices:

Interest rates, as a component of broader economic conditions, can shape individuals' decisions regarding education. During economic upswings, individuals might be more willing to invest in higher education to capitalize on growth prospects. This trend aligns with the decrease in university degree holders during periods of economic positivity, suggesting that people may prioritize immediate job opportunities over extended education. Conversely, during economic downturns or uncertainty, there might be a preference for shorter-term vocational training to quickly enter the workforce.

1. Default Rates and Financial Stability:

Fluctuations in default rates can provide insights into the financial health of borrowers. The increase in the number of people having default during certain periods might reflect the sensitivity of borrowers to changes in interest rates and economic stability. Higher interest rates could strain borrowers' ability to repay loans, contributing to a rise in default rates. This underscores the crucial role that interest rates play in determining the affordability of debt and its impact on individuals' financial well-being.

1. Occupation and Education Changes:

The patterns of occupation and education changes mirror the ebb and flow of economic conditions. When economic indicators peak, signaling confidence and growth, there's a corresponding increase in management and university degree holders, indicating opportunities for skilled roles. During economic volatility or decline, the emphasis on blue-collar workers might reflect industries prioritizing cost-effectiveness and immediate labor needs over higher-skilled roles. Similarly, reductions in technician roles could reflect adjustments made due to fluctuations in demand for technical expertise.

1. Economic Sentiment and Decision-Making:

The Consumer Confidence Index (CCI) provides insight into consumer sentiment, which is instrumental in shaping economic activity. Positive sentiment fosters spending, investment, and borrowing, all of which contribute to economic growth. When CCI peaks, as observed in some instances, it indicates that consumers are optimistic about the economic outlook and are more likely to engage in economic activities that drive growth.

1. Final thoughts:

In conclusion, the interplay between interest rates, economic indicators, job choices, education decisions, default rates, and economic sentiment illustrates the complex web of relationships within an economy. These factors not only influence each other but also impact individuals' decisions and behavior. Understanding these dynamics is crucial for policymakers, businesses, and individuals alike, as they navigate the ever-changing landscape of economic conditions and their far-reaching effects.

# **REFERENCES**

Sexton, R. L. (2010). *Exploring Economics* *(5th ed.).* Mason, OH: South Western Educational Publishing.