Project Report

On

GOOGLE vs AI

Submitted

In partial fulfilment of the requirements for the degree of

BACHELOR OF SCIENCE

in

MATHEMATICS AND STATISTICS

By

MARIYA XAVIER

(Register no: AB21AMAT019)

(2021-2024)

Under the Supervision of

ANU MARY JOHN

DEPARTMENT OF MATHEMATICS AND STATISTICS

ST. TERESA'S COLLEGE (AUTONOMOUS)

ERNAKULAM, KOCHI - 682011

APRIL 2024

ST. TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

CERTIFICATE

This is to certify that the dissertation entitled, **GOOGLE vs. AI** is a bonafide record of the work done by MARIYA XAVIER under my guidance as partial fulfilment of the award of the degree of **Bachelor of Science in Mathematics and Statistics** at St. Teresa's College (Autonomous), Ernakulam affiliated to Mahatma Gandhi University, Kottayam. No part of this work has been submitted for any other degree elsewhere.

Date:

Place: Ernakulam

ANUMARY JOHN
Assistant Professor,
Department of Mathematics and Statistics
St. Teresa's College (Autonomous),
Ernakulam.

Dr. URSALA PAUL
Assistant Professor and HOD
Department of Mathematics and Statistics
St. Teresa's College (Autonomous)
Ernakulam

External Examiners

1.....29.04.2024.....

2.....

ST. TERESA'S COLLEGE (AUTONOMOUS) DEPARTMENT OF MATHEMATICS AND STATISTICS ERNAKULAM

DECLARATION

I hereby declare that the work presented in this project is based on the original work done by me under the guidance of **ANU MARY JOHN,** Assistant Professor, Department of Mathematics and Statistics, St. Teresa's College (Autonomous), Ernakulam and has not been included in any other project submitted previously for the award of any degree.

Ernakulam
Date: 29/04/2024

MARIYA XAVIER
AB21AMAT019

ACKNOWLEDGEMENT

I express our deep gratitude to Miss. **ANU MARY JOHN,** Assistant Professor, Department of Mathematics and Statistics, St. Teresa's College, Ernakulam for always motivating and encouraging me to learn and research everything independently.

I am also very thankful to **Dr. URSALA PAUL,** Assistant Professor and HOD, Department of Mathematics and Statistics, St. Teresa's College (Autonomous), Ernakulam for their valuable suggestions, critical examination of work during the progress.

Above all, I thank God Almighty and my parents for giving me the blessings to take over this project.

Ernakulam
Date: 29/04/2024

MARIYA XAVIER
AB21AMAT019

GOOGLE Vs. AI

[1*]Julia Beril Sharon, [2]Mary Riya C. R, [3]Lakshmi Umesh, [4]Mary Niya N.J, [5]MariyaXavier

[12345]B.sc mathematics, St. Teresa's College (Autonomous), Ernakulam, India.

[6]Anu Mary John

[6]Assistant Professor, Department of Mathematics and Statistics, St. Teresa's College, Ernakulam, India.

ABSTRACT

This paper explores the nexus of technology, finance, and information retrieval, focusing on a comprehensive analysis of the stock market. It predicts trends in Google and AI stocks from 2024 to 2028, utilizing advanced analytics to uncover insights into their financial trajectories and potential market impact. The study states forecasting and some performance measure of time series are discussed. Our report outlines findings from an online survey gauging public perceptions on the comparative merits, exploring public opinions regarding the comparative efficacy of Google and artificial intelligence (AI).

Keywords: ARIMA, Forecast, Box- Jenkins methodology, Short-term prediction, Chi-square Test, Correlation Analysis, Z-Test.

INDEX

CHAPTER 1

INTRODUCTION

In the rapidly evolving landscape of artificial intelligence, natural language processing (NLP) models have emerged as transformative tools for understanding and generating human-like text. Among these models, ChatGPT, developed by OpenAI, and Google's suite of NLP technologies, represent two of the most prominent offerings. As organizations and individuals seek advanced AI solutions for tasks ranging from customer support to content generation, the question arises: which of these powerful tools is the better choice?

1.1 AI (ARTIFICIAL INTELLIGENCE)

Artificial intelligence - Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert systems, natural language processing, speech recognition and machine vision. AI systems are developed using a combination of techniques such as machine learning, deep learning, natural language processing, and computer vision. Machine learning algorithms enable machines to learn patterns from data and make predictions or decisions without

explicit programming. Deep learning, a subset of machine learning, involves training neural networks with multiple layers to process complex data and extract meaningful insights.

WHY AI IS USED -

India sees AI as a tool for social empowerment and inclusion and focuses on developing AI tools for the global south. AI makes our lives easier by automating tasks and providing us with information and recommendations tailored to our individual needs. The goal of AI is to create intelligent machines that can perform tasks without human intervention. AI is used in many areas of our lives. You might have seen AI in action when you use voice assistants like Siri or Alexa. They can answer your questions and help you find information. AI is also used in games, self-driving cars, and even in medical research to help find new treatments for diseases.

OPEN AI

OpenAI is an artificial intelligence research laboratory and company. It was founded with the goal of developing AI technologies that are safe, beneficial, and accessible to all of humanity. OpenAI conducts research, builds AI models, and collaborates with organizations to promote the responsible and ethical use of AI.

OpenAI has contributed significantly to the field of AI with ground-breaking models such as GPT (Generative Pre-Trained Transformer). These models, like GPT-3, have demonstrated remarkable abilities in natural language processing, text generation, and other AI applications.

ChatGPT

ChatGPT is an advanced language model developed by Open AI. It is part of the GPT-3.5 series, which stands for "Generative Pre-Trained Transformer 3.5." ChatGPT is designed to understand and generate human-like text responses in a conversational manner. Using a large amount of data from the internet, ChatGPT has been trained to understand and generate text in multiple languages. It can answer questions, provide explanations, engage in discussions, and assist with a wide range of topics. ChatGPT can be seen as a virtual assistant that utilizes natural language processing and machine learning techniques to communicate with users.

The underlying technology behind ChatGPT is based on deep learning, specifically utilizing transformer neural networks. These networks allow the model to analyses and understand context, generate coherent responses, and provide meaningful information based on the input it receives.

ChatGPT has been trained on a diverse range of topics, but it is important to note that it has limitations. While it strives to provide accurate and helpful responses, it may occasionally produce incorrect or nonsensical answers. Users should exercise critical thinking and fact-check information obtained from ChatGPT.

OpenAI continues to refine and improve ChatGPT by gathering feedback from users and making iterative updates to enhance its capabilities. The goal is to create a more reliable and useful tool for various applications such as customer support, content creation, and general conversational assistance. ChatGPT represents a significant advancement in natural language processing and AI, enabling more interactive and dynamic interactions between humans and machines.

1.2 GOOGLE

Google is a search engine that started development in 1996 by Sergey Brin and Larry Page as a research project at Stanford University to find files on the Internet. It is one of the largest companies in the world. Google analyses the text, images, and video files on the page, and stores the information in the Google index, which is a large database. Apart from search, Google offers many other helpful tools and services. One of the most popular is

Gmail, which is an email service that lets people send and receive messages online.
Google also has Google Maps, which helps us find our way and discover places around the world.
It shows us maps, directions, and even real-time traffic information.
Another amazing service by Google is YouTube. It is a website where people can watch and share videos. From funny clips to educational tutorials and music videos, YouTube has millions of videos on almost every topic you can imagine!
How Google is useful - Google's mission is to organize the world's information and make it universally accessible and useful. Google provides students the opportunity.

1.3 STOCK MARKET

The stock market is where investors buy and sell shares of companies. It is a set of exchanges where companies issue shares and other securities for trading. It also includes over the counter (OTC) marketplaces where investors trade securities directly with each other (rather than through an exchange). This project delves into the transformative potential of integrating Google's expansive resources with the capabilities of Artificial Intelligence (AI). The exchange between Google and AI in this context creates a symbiotic relationship that has the potential to revolutionize how we understand and navigate the complexities of the stock market. Stock market prediction is a complex and challenging task that involves forecasting the future movements of stock prices based on various factors.

1.4 OVERVIEW OF GOOGLE Vs AI

The growth of AI and Google are interconnected, as Google has been a major player in advancing and leveraging AI technologies. Google's growth has been significantly fuelled by integrating AI into its products and services. The company uses AI algorithms for search engine optimization, personalized recommendations, language processing, and more, enhancing user experience and maintaining a competitive edge.

On the other hand, the growth of AI as a field is not solely attributed to Google, but Google's contributions and innovations have played a substantial role. Google's deep involvement in

AI research, through projects like DeepMind, has contributed to the broader development of AI technologies globally.

1.5 OVERVIEW OF GOOGLE Vs. CHATGPT

ChatGPT and Google are two different technologies designed for different purposes. The former is designed to generate text and answer questions using natural language, while the latter is designed to search the internet and provide relevant results for a user's query.

ChatGPT's accuracy depends on its training data and the quality of prompts it receives. However, since it relies on learned patterns from data, there is a chance it might generate responses that sound plausible but are not entirely accurate, especially in complex or nuanced queries. OpenAI has stated clearly that ChatGPT has limited knowledge of the world and events after 2021. Therefore, OpenAI has made efforts to reduce the likelihood of misinformation, but there's still room for improvement.

Google's accuracy is rooted in its continuous data collection. Google's accuracy is high for well-established facts and widely recognized information. However, it might not filter out misinformation. So, the accuracy of search results can vary depending on the query and the sources it pulls from.

Google has an edge in providing accurate information due to its sophisticated ranking algorithms and access to a wide array of reputable sources. ChatGPT's accuracy heavily depends on the quality of its training data and the precision of user prompts.

1.6 OBJECTIVES OF THE STUDY.

In this study, we apply a time series model to forecast future stock indices that have significant impact on Google and AI.

And, by applying primary data from the survey on Google Vs. ChatGPT.

•To Analyse and report on the percentage shift in trends or statistics during the year 2028.

•To inquire about the user's personal experience with adopting and AI-oriented lifestyle, explore their views on the potential transformative impact of AI on the world, discuss the necessity of learning more about AI, and address the question of whether AI contributes to human cognitive laziness.

•To provide accurate and reliable information comparable to Google search results.

•To determine and quantify the rate of preference among users for ChatGPT versus Google in various scenarios.

•To investigate whether people are more likely to use chatGPT or Google to gather primary information.

1.7 LITERATURE REVIEW

1. A comprehensive dataset derived from two secondary sources: the "marcrotens of alphabet" (Google stock market data) covering the period from 2004 to July 2023, and AI stocks data obtained from the "Statista" website covering the period from 2021 to July 2023. The dataset comprises four fundamental elements—open price, closed price, low price, and high price—forming the basis for a thorough exploration of market dynamics. This analytical approach plays a pivotal role in understanding the temporal evolution of stock prices, facilitating the identification of potential forecasting opportunities. The four key elements mentioned earlier—open price, closed price, low price, and high price constitute the cornerstone for the conducted analysis.

   The application of advanced forecasting models to the dataset aims to offer insights into the possible trajectory of the market, extending the analysis to predict data trends up to the year 2028.

2. Understanding users' experiences in adopting an AI-oriented lifestyle is vital, as highlighted by studies from Smith et al. (2020) and Chen & Lee (2019), emphasizing the exploration of comfort, challenges, and benefits associated with integrating AI technologies into daily routines. Investigating the necessity of learning more about AI, as argued by scholars Brown & Smith (2017) and Li & Zhang (2020), is critical for educational initiatives, underscoring the pivotal role of user understanding and awareness in fostering responsible interactions with AI technologies.

   The debated question of whether AI contributes to cognitive laziness, as suggested by studies from Lee et al. (2018) and Wang & Patel (2019), adds nuance, acknowledging the potential efficiency gains of AI while addressing concerns about over-reliance and cognitive disengagement. Aligning with research on information quality by Xu & Li (2020), the objective to provide accurate and reliable information comparable to Google search results Acknowledges that users' trust in AI-generated content hinges on the accuracy and relevance of information.

Lastly, investigating user behavior and preferences to understand reliance on Google or ChatGPT, drawing from studies by Liu & Wang (2018) and Zhang & Chen (2022), emphasizes the significant roles of user habits, trust, and perceived utility in shaping information-seeking behavior. This integrated exploration forms a comprehensive foundation for understanding user attitudes and behaviours in the AI and information technology landscape.

Having conducted experiments utilizing statistical tools such as Chi-square distribution, correlation, regression, and Z-test, we have quantified and validated our survey data. Our analysis contributes robust empirical evidence to the nuanced insights provided by the literature, offering statistical significance to user preferences for ChatGPT versus Google in various scenarios. The experimentally supported findings shed light on the intricate relationship between user behaviours, preferences, and the transformative impact of AI on information retrieval practices.

CHAPTER 2
TIME SERIES ANALYSIS

SECTION 1
MATERIALS AND METHODS.

1.1 INTRODUCTION

First and foremost, step in analysis is to study the data in detail. Since the data is time dependent, time series analysis was utilized for forecasting. Forecasting uses the developed models to predict future values of the time series. This involves extrapolating trends and considering seasonality. This chapter describes the material and techniques used for the study.

1.2 TIME SERIES

Time series analysis is a specific way of analysing a sequence of data points collected over an interval of time. In time series analysis, analysts record data points at consistent intervals over a set period of time rather than just recording the data points intermittently or randomly. However, this type of analysis is not merely the act of collecting data over time. Time is a crucial variable because it shows how the data adjusts over the course of the data points as well as the results. It provides an additional source of information and a set order of dependencies between the data. Additionally, time series data can be used for forecasting—predicting future data based on historical data.

Time series are used for,

i) Obtain an understanding of the underlying forces and structure that produced the observed data.

ii) Fit a model and proceed to forecasting, monitoring. Time series analysis can be useful to see how X given asset, security or economic variables change over time. A time series can be

- Stationary
- Non stationary

1.2.1 Stationary.

A time series is considered stationary if its statistical properties, such as mean and variance, remain constant over time. This implies that the underlying data-generating process does not change with time.

1.2.2 Non-Stationary

Non-stationary A time series is non-stationary when its statistical properties change over time. This often involves trends, seasonality, or other patterns that can affect the mean and variance, making it more challenging to analyse and predict.

1.3 VARIOUS COMPONENTS OF THE TIME SERIES

1.3.1 Trend analysis

Trend analysis is a technique used in technical analysis that attempts to predict future stock price movements based on recently observed trend data. Trend analysis uses historical data, such as price movements and trade volume, to forecast the long-term direction of market sentiment.

1.3.2 Seasonal variations

Seasonal variations refer to the changes that take place due to the rhythmic forces which operate in a regular and periodic manner. These forces usually have the same or most similar pattern year after year. When we record data weekly, monthly or quarterly, we can see and calculate seasonal variations. Thus, when a time series consists of data only based on annual figures, there will be seen no seasonal variations. These variations may be due to seasons, weather conditions, habits, customs or traditions.

1.3.3 Cyclical variations

Cyclical variations are due to the ups and downs recurring after a period from time to time. These are due to the business cycle and every organization must phase all the four phases of a business cycle some time or the other. Prosperity or boom, recession, depression, and recovery are the four phases of a business cycle.

1.3.4 Random or irregular variations

Random variations are fluctuations which are a result of unforeseen and unpredictable forces.

These forces operate in an absolutely random or erratic manner and do not have any definite pattern. Thus, these variations may be due to floods, famines, earthquakes, strikes, etc.

1.4 FORECASTING TECHNIQUES

There are several forecasting methods that can be broadly segmented as either qualitative or quantitative. Within each category, there are several techniques at one's disposal.

- Under **qualitative** methods, techniques may involve interviews, on-site visits, and the Delphi method of pooling experts' opinions, focus groups, and text analysis of financial documents, news items, and so forth.
- Under **quantitative** methods, techniques generally employ statistical models that look at time series or cross-sectional data, such as econometric regression analysis or causal inference (when available).

1.5 CHOOSING THE RIGHT FORECASTING METHOD

The right forecasting method will depend on the type and scope of the forecast. Qualitative methods are more time-consuming and costly but can make very accurate forecasts given a limited scope. For instance, they might be used to predict how well a company's new product launch might be received by the public.

For quicker analyses that can encompass a larger scope, quantitative methods are often more useful. Looking at big data sets, statistical software packages today can crunch the numbers in a matter of minutes or seconds. However, the larger the data set and the more complex the analysis, the pricier it can be.

Thus, forecasters often make a sort of cost-benefit analysis to determine which method maximizes the chances of an accurate forecast in the most efficient way. Furthermore, combining techniques can be synergistic and improve the forecast's reliability.

1.6 STEPS IN FORECASTING

1.6.1 Understanding why changes in the past have occurred.

One of the basic principles of statistical forecasting indeed of all forecasting when historical data are available is that the forecaster should use the data on past performance

to get a "speedometer reading" of the current rate (of sales, say) and of how fast that rate is increasing or decreasing. The current rate and changes in the – "acceleration" and "deceleration" constitute the basis of forecasting. Once they are known, various mathematical techniques can develop projections from them. If an attempt is made to forecast business fluctuations without understanding why past changes have taken place, the forecast will be purely mechanical, based solely upon the application of mathematical formulae and subject to serious error.

1.6.2 Determining which phases of business activity must be measured.

After it is known why business fluctuations have occurred, or if there is a reasonable supposition it is necessary to measure certain phases of business activity to predict what changes will probably follow the present level of activity.

1.6.3 Selecting and compiling data to be used as measuring devices.

There is an interdependent relationship between the selection of statistical data and determination of why business fluctuations occur. Statistical data cannot be selected and compiled in an intelligent manner unless there is a sufficient understanding of business fluctuations; likewise, it is important that reasons for business fluctuations be stated in such a manner that it is possible to secure data that are related to the reasons.

1.6.4 Analysis of data.

In this last step, the data are analysed in the light of one is understanding of the reason why changes occur. For example, if it is reasoned that a certain combination of forces will result in each change, the statistical part of the problem is to measure these forces and from the data available, to draw conclusions on the future course of action. The methods of drawing conclusions may be called Forecasting techniques, which represent any one of many analytical devices for summarizing data and drawing inferences from the summaries.

1.7 REQUIREMENTS GOOD FORECASTING SYSTEM

A forecasting system to be instrumental in contributing to better management decision-making needs certain conditions;

1. It must involve the managers whose decisions are affected.
2. Individual forecasts and group of forecasts must be specifically relevant to the decisions being taken.
3. The forecasts must not claim too much validity or authority.
4. Implications of the various probable errors in the predictions for the organizations need to be thoroughly worked through so that management can evaluate the consequences of the probable range of likely outcomes.
5. Management must at least know how badly things could go wrong if all the guesses turned out wrong.

1.8 METHODS OF FORECASTING

1.8.1 Percent of Sales

Internal pro forma statements are often created using percent of sales forecasting. This method calculates future metrics of financial line items as a percentage of sales. For example, the cost of goods sold is likely to increase proportionally with sales; therefore, it is logical to apply the same growth rate estimate to each.

To forecast the percent of sales, examine the percentage of each account's historical profits related to sales. To calculate this, divide each account by its sales, assuming the numbers will remain steady. For example, if the cost of goods sold has historically been 30 percent of sales, assume that trend will continue.

1.8.2Straight Line

The straight-line method assumes a company's historical growth rate will remain constant. Forecasting future revenue involves multiplying a company's previous year's revenue by its growth rate. For example, if the previous year's growth rate was 12 percent, straight-line forecasting assumes it will continue to grow by 12 percent next year.

Although straight-line forecasting is an excellent starting point, it does not account for market fluctuations or supply chain issues.

1.8.3Moving Average

Moving average involves taking the average—or weighted average—of previous periods to forecast the future. This method involves more closely examining a business's high or low demands, so it is often beneficial for short-term forecasting. For example, you can use it to forecast next month's sales by averaging the previous quarter.

Moving average forecasting can help estimate several metrics. While it is most applied to future stock prices, it is also used to estimate future revenue.

To calculate a moving average, use the following formula:

A1 + A2 + A3 … / N

Formula breakdown:

A = Average for a period

N = Total number of periods

Using weighted averages to emphasize recent periods can increase the accuracy of moving average forecasts.

1.8.4 Simple Linear Regression

Simple linear regression forecasts metrics based on a relationship between two variables: dependent and independent. The dependent variable represents the forecasted amount, while the independent variable is the factor that influences the dependent variable.

The equation for simple linear regression is:

Y= BX + A

Formula breakdown:

Y = Dependent variable (the forecasted number)

B = Regression line's slope

X = Independent variable

A = Y-intercept

1.8.5 Multiple Linear Regression

If two or more variables directly impact a company's performance, business leaders might turn to multiple linear regression. This allows for a more accurate forecast, as it accounts for several variables that ultimately influence performance.

To forecast using multiple linear regression, a linear relationship must exist between the dependent and independent variables. Additionally, the independent variables cannot be so closely correlated that it is impossible to tell which impacts the dependent variable.

1.8.6 Exponential Smoothing Model

Like the moving average, exponential smoothing is another time-series forecasting model which can be used to predict new values by using a set of weighted averages based on past observations. Exponential smoothing helps to predict the future by using past company data. The weights start declining exponentially with past observations to predict the upcoming period. To put it simply, if the observation is a more recent one — the associated weight is higher. This means that more weight is given to recent values instead of past values.

New forecasts are predicted by including the past forecast and the percentage of value (the difference between the current and the past forecast). The idea behind this model is to attribute importance to more recent values in the series — when observations become older, past values get exponentially smaller.

1.8.7 Trend Projection Model

The trend projection model works best in situations where you could work out the future influence of certain variables (dependent or independent) based on its Past behavior. The model examines past events to identify patterns and trends that could recur frequently.

Trend projection can be used to forecast future activity since it considers that all factors involved in past trends will continue in the future as well. The model requires long and

reliable time-series data which is arranged in chronological order for evaluation. By identifying the patterns of trends, the company will be able to get a vision of the future. Consequently, once the trend has been identified, it will be able to predict future demand.

1.8.8 Bayesian Information Criterion (BIC)

The model that minimizes the Bayesian Information Criterion (BIC) is selected as the final model. The BIC criterion attempts to balance model complexity with goodness-of-fit over the historical data period (between history start date and forecast start date). The BIC criterion rewards a model for goodness-of-fit and penalizes a model for its complexity. The complexity penalty is necessary to avoid over fitting.

There are various equivalent versions of the Bayesian Information Criterion, but RDF minimizes the following:

$$BIC = s \cdot n^{k/2n}$$

where *n* is the number of periods in the available data history, *k* is the number of parameters to be estimated in the model (a measure of model complexity), and *so* is the root mean squared error computed with one-step-ahead forecast errors resulting from the fitted model (a measure of goodness-of-fit). Note that since each member of the model candidate list is a family of models, an optimization routine to select optimal smoothing parameters is required to minimize *s* for each model form (that is, to select the best model).

1.8.9 Box-Jenkins Methodology

The Box-Jenkins Model may be one of several, time series analysis models a forecaster will encounter when using programmed forecasting software. In many cases, the software will be programmed to automatically use the best fitting forecasting methodology based on the time series data to be forecasted. Box-Jenkins is reported to be a top choice for data sets that are mostly stable and have low volatility. Autoregressive integrated moving average (ARIMA) models are a form of Box-Jenkins model. The terms ARIMA and Box Jenkins are sometimes used interchangeably.

1.8.10 Autoregressive Integrated Moving Average (ARIMA)

An autoregressive integrated moving average model is a form of regression analysis that gauges the strength of one dependent variable relative to other changing variables. The model's goal is to predict future securities or financial market moves by examining the differences between values in the series instead of through actual values. A statistical model is autoregressive if it predicts future values based on past values. For example, an ARIMA model might seek to predict a stock's future prices based on its past performance or forecast a company's earnings based on past periods.

1.9 TIME SERIES FORECASTING

Time series forecasting uses information regarding historical values and associated patterns to predict future activity. Most often, this relates to trend analysis, cyclical fluctuation analysis, and issues of seasonality. As with all forecasting methods, success is not guaranteed.

The Box-Jenkins Model, for instance, is a technique designed to forecast data ranges based on inputs from a specified time series. It forecasts data using three principles: auto regression, differencing, and moving averages. These three Principles are known as p, d, and q, respectively. Each principle is used in the Box-Jenkins analysis, and together they are collectively shown as an autoregressive integrated moving average, or ARIMA (p, d, and q). ARIMA can be used, for instance, to forecast stock prices or earnings growth. Another method, known as rescaled range analysis, can be used to detect and evaluate the amount of persistence, randomness, or mean reversion in time series data. The rescaled range can be used to extrapolate a future value or average for the data to see if a trend is stable or likely to reverse.

1.10 FORECASTING STOCK PRICES

One use for Box-Jenkins Model analysis is to forecast stock prices. This analysis is typically built out and coded through Excel. The analysis results in a logarithmic outcome, which can be applied to the data set to generate the forecasted prices for a specified period of time in the future.

ARIMA models are based on the assumption that past values have some residual effect on current or future values. For example, an investor using an ARIMA model to forecast stock prices would assume that new buyers and sellers of that stock are influenced by recent market transactions when deciding how much to offer or accept for the security

SECTION 2

DATA ANALYSIS AND INTERPRETATION

2.1 INTRODUCTION

This chapter comprises the statistical analysis of the data and the results and interpretation of the data. For analysing the stock prices of Google stock market data covering the period from 2004 to July 2023 and AI stock data obtained covering the period from 2021 to July 2023, we analyse the volume of both stocks for forecasting the prices. We follow the procedure for forecasting after plotting the time series data and decomposing the time series into different components.

2.2 DATA SUMMARIZATION

The collected data from Google stock market performance spanning from 2004 to July 2023 reveals trends and patterns in the company's financial trajectory over nearly two decades. Simultaneously, the AI stock data covering the period from 2021 to July 2023, offers insights into the broader landscape of artificial intelligence investments during this specific timeframe. Examining these datasets enables a comprehensive understanding of the evolving dynamics in both the technology and financial sectors.

2.2.1 VOLUME OF AI AND GOOGLE

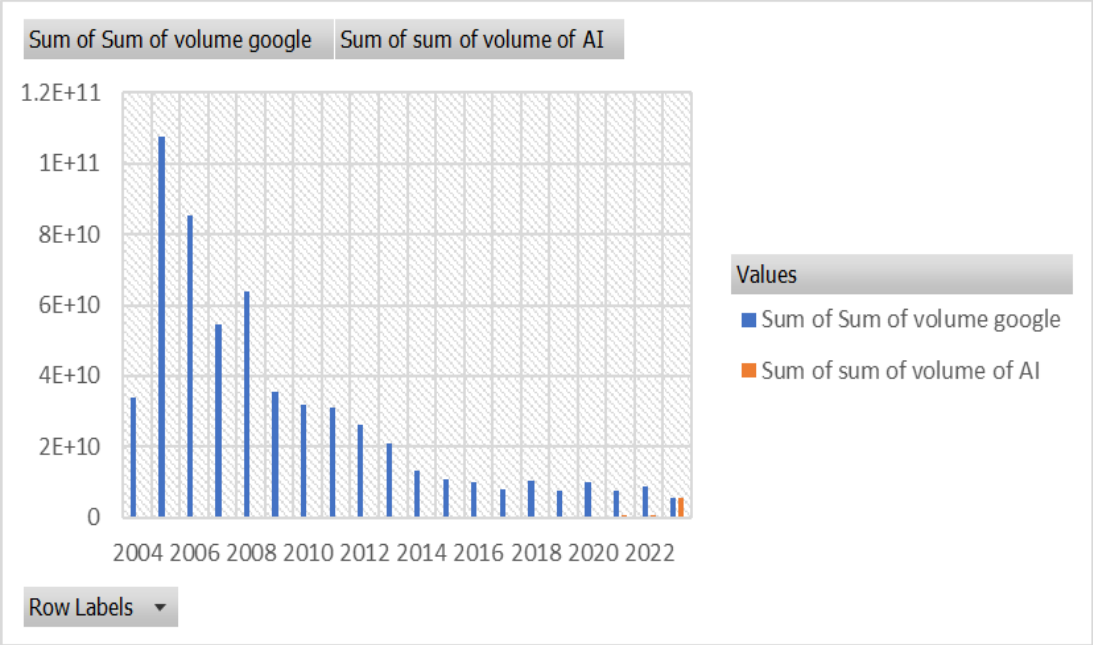
Row Labels	Sum of volume Google	Sum of volume of AI
2004	34094928000	0

2005	1.07639E+11	0
2006	85179590000	0
2007	54552820000	0
2008	63910472000	0
2009	35336700000	0
2010	32021128000	0
2011	30997184000	0
2012	26017452000	0
2013	21019706000	0
2014	13394592980	0
2015	10934456820	0

2016	9923281360	0
2017	8116826520	0
2018	10271638720	0
2019	7653541340	0
2020	10045411580	0
2021	7594192080	815928900
2022	8595186324	643872900
2023	5400345577	5482158849

Table 1; volume of AI and google from 2004- 2023

This table represents the historical stock market volumes for Google and AI-related stocks from 2004 to August 2023. The data showcases the annual trading volumes for Google and AI stocks, providing insights into their market activity over the years. As of the latest available data in August 2023, it sets the foundation for predicting future trends and making informed decisions in anticipation of the stock market volumes for 2028. Analysing this historical context, along with the partial data from 2023, can aid in forecasting potential patterns and behaviours in the stock market for both Google and AI-related investments.



Figures 2.1

2.3 FORECASTING OF GOOGLE

To effectively forecast future values, the process begins by meticulously gathering historical data pertaining to the variable under consideration, such as the volume of Google stock over the years. Following data collection, a comprehensive analysis ensues, delving into discernible trends, patterns, and potential seasonality within the historical dataset. This thorough understanding of the data's characteristics serves as a foundation for judiciously selecting an appropriate forecasting method.

Given the diverse landscape of forecasting methodologies, the choice of a specific approach is contingent upon the inherent nature of the data. Time Series Analysis offers avenues like auto regression (ARIMA), while Regression Analysis, Machine Learning Models, and Exponential Smoothing represent additional common techniques.
Subsequently, the historical data is partitioned into training and testing sets to facilitate the model-building process. The training set becomes the crucible for constructing the forecasting model, wherein parameters, such as coefficients or weights, are meticulously estimated to encapsulate the intricate relationships within the data.

Once the model is crafted, the validation phase ensues, scrutinizing its efficacy using the testing set. Upon successful validation, the model stands ready to be deployed for predicting future values. This application extends beyond the confines of the training and testing periods, offering insights into forthcoming trends. Vital to the forecasting process is the ongoing vigilance of the model's performance, especially in the face of newly available data. Regular monitoring and judicious updates ensure the model's sustained accuracy over time, reflecting a dynamic approach to forecasting in a constantly evolving landscape.

Row Labels	Sum of volume GOOGLE	Forecast (Sum of volume)	Lower Confidence Bound (Sum of volume)	Upper Confidence Bound (Sum of volume)
2004	34094928000			
2005	1.07639E+11			
2006	85179590000			
2007	54552820000			
2008	63910472000			

2009	35336700000
2010	32021128000
2011	30997184000
2012	26017452000
2013	21019706000
2014	13394592980
2015	10934456820
2016	9923281360
2017	8116826520
2018	10271638720
2019	7653541340

2020	10045411580			
2021	7594192080			
2022	8595186324			
2023	5400345577	5400345577	5.40E+09	5.40E+09
2024		-588975673.7	-3.89E+10	3.77E+10
2025		-4371689761	-4.72E+10	3.85E+10
2026		-8154403849	-5.51E+10	3.88E+10
2027		-11937117936	-6.27E+10	3.88E+10
2028		-15719832024	-7.00E+10	3.86E+10

Table 2: Forecasting of Google

2.4 FORECASTING OF AI

Embarking on the forecasting journey for AI, our starting point involves the assembly of historical data, spanning from 2021 to August 2023. A notable aspect of this scenario is the temporal constraint of three years, compelling the adoption of a monthly-wise approach to enhance forecasting precision within the limited timeframe.

The process unfolds with a meticulous analysis, scrutinizing trends, patterns, and potential seasonality present in the historical monthly data. This discerning examination of the data's characteristics lays the groundwork for the selection of an appropriate forecasting method tailored to the nuanced dynamics of the AI domain.

Within the expansive realm of forecasting methodologies, the choice is guided by the intrinsic nature of the data. Time Series Analysis, encompassing techniques such as auto regression (ARIMA), takes a central role, while Regression Analysis, Machine Learning Models, and Exponential Smoothing contribute additional tools for unravelling the intricacies of AI-related trends.

The subsequent phase involves the systematic partitioning of the monthly historical data into training and testing sets. The training set becomes the crucible for constructing the forecasting model, emphasizing the monthly intervals. Parameters like coefficients or weights are diligently estimated to capture the intricate relationships within the data, recognizing the nuances revealed through the monthly granularity.

Post-model construction, the validation phase rigorously assesses its efficacy using the testing set, ensuring adaptability to the dynamic nature of AI data. With successful validation, the model stands poised to predict future values monthly, effectively extending insights beyond the observed period.

In the context of a limited three-year timeframe, the monthly-wise approach enriches the granularity of insights, contributing to a more nuanced and accurate forecasting analysis for AI. The forecasting journey, dynamic in nature, necessitates continual vigilance of the model's performance. As the data evolves, regular monitoring and judicious updates become imperative, fortifying the model's resilience and adaptability in the face of emerging trends within the dynamic AI landscape.

Date	Volume AI	Forecast (Volume)	Lower Confidence Bound (Volume)	Upper Confidence Bound (Volume)
01-02-2021	51568500			
01-03-2021	125732400			
01-04-2021	77656500			
01-05-2021	45006500			
01-06-2021	113381600			
01-07-2021	48154500			
01-08-2021	46652700			

01-09-2021	82000200
01-10-2021	33260600
01-11-2021	68434800
01-12-2021	124080600
01-01-2022	65686200
01-02-2022	56424000
01-03-2022	92294100
01-04-2022	35813900

01-05-2022	59884100
01-06-2022	98972500
01-07-2022	40847900
01-08-2022	35496700
01-09-2022	56556800
01-10-2022	27774600
01-11-2022	25532800
01-12-2022	48589300

01-01-2023	135437600
01-02-2023	620668000
01-03-2023	477531300
01-04-2023	495238200
01-05-2023	607401600
01-06-2023	787467500
01-07-2023	424612700
01-08-2023	159677300

01-09-2023	413166760.5
01-10-2023	430680459.9
01-11-2023	456639850.1
01-12-2023	473637578
01-01-2024	487212932.5
01-02-2024	506738045.4
01-03-2024	537880257.8
01-04-2024	558425479.2

01-05-2024	632689616.5			
01-06-2024	656196731.4			
01-07-2024	663300044.4			
01-08-2024	687469268.1			
01-09-2024	711174947.7			
01-10-2024	734714616.2			
01-11-2024	758839699.6			
01-12-2024	782726014.6	782726014.6	782726014.6	782726014.6

01-01-2025	797727505.7	575687268.6	1019767743
01-02-2025	814957449.7	537273872.4	1092641027
01-03-2025	832187393.7	508167992.9	1156206794
01-04-2025	849417337.6	484803604.6	1214031071
01-05-2025	866647281.6	465434028.5	1267860535
01-06-2025	883877225.6	449049246.8	1318705204
01-07-2025	901107169.5	435002929.1	1367211410
01-08-2025	918337113.5	422851948.8	1413822278

01-09-2025	935567057.5	412276864.3	1458857251
01-10-2025	952797001.5	403038422.6	1502555580
01-11-2025	970026945.4	394951947.4	1545101943
01-12-2025	987256889.4	387871370.9	1586642408
01-01-2026	1004486833	381678810.8	1627294856
01-02-2026	1021716777	376277507.1	1667156048
01-03-2026	1038946721	371586882.8	1706306560
01-04-2026	1056176665	367538996.3	1744814334

01-05-2026	1073406609	364075935.1	1782737283
01-06-2026	1090636553	361147861.6	1820125245
01-07-2026	1107866497	358711522.8	1857021472
01-08-2026	1125096441	356729096.1	1893463786
01-09-2026	1142326385	355167281.9	1929485488
01-10-2026	1159556329	353996582.8	1965116075
01-11-2026	1176786273	353190723.9	2000381822
01-12-2026	1194016217	352726181.5	2035306253

01-01-2027	1211246161	352581797.5	2069910525
01-02-2027	1228476105	352738460	2104213750
01-03-2027	1245706049	353178837.1	2138233261
01-04-2027	1262935993	353887154.5	2171984831
01-05-2027	1280165937	354849006.4	2205482867
01-06-2027	1297395881	356051195.8	2238740566
01-07-2027	1314625825	357481597.5	2271770052
01-08-2027	1331855769	359129040.5	2304582497

01-09-2027	1349085713	360983206.6	2337188219
01-10-2027	1366315657	363034542.1	2369596771
01-11-2027	1383545601	365274181.4	2401817020
01-12-2027	1400775545	367693879.9	2433857210
01-01-2028	1418005489	370285954.4	2465725023
01-02-2028	1435235433	373043232.3	2497427633
01-03-2028	1452465377	375959004.7	2528971749
01-04-2028	1469695321	379026985.9	2560363655

01-05-2028	1486925265	382241277.4	2591609252
01-06-2028	1504155209	385596335.3	2622714082
01-07-2028	1521385153	389086941.3	2653683364
01-08-2028	1538615097	392708176.7	2684522016
01-09-2028	1555845040	396455399.1	2715234682
01-10-2028	1573074984	400324221.1	2745825748
01-11-2028	1590304928	404310491.4	2776299365
01-12-2028	1607534872	408410277.4	2806659467

Table 3; Forecasting of AI

2.4.1 FORECASTING OF AI IN YEARS

In the ongoing analysis, we are simplifying the presentation of monthly AI data by aggregating it into a yearly format. This involves adding up the monthly values to provide a clearer overview of trends on an annual basis, making it easier to interpret and analyse.

Additionally, we are incorporating predicted values for AI from 2021 to 2028 into the dataset. These predictions, derived from rigorous modelling techniques, offer insights into the anticipated trends and trajectories for AI during this period. By merging historical monthly data with these forecasts, we are creating a comprehensive dataset that facilitates a holistic understanding of AI trends over the specified timeframe. This combined approach enhances the dataset's utility for making informed decisions and strategic plans in the dynamic landscape of artificial intelligence.

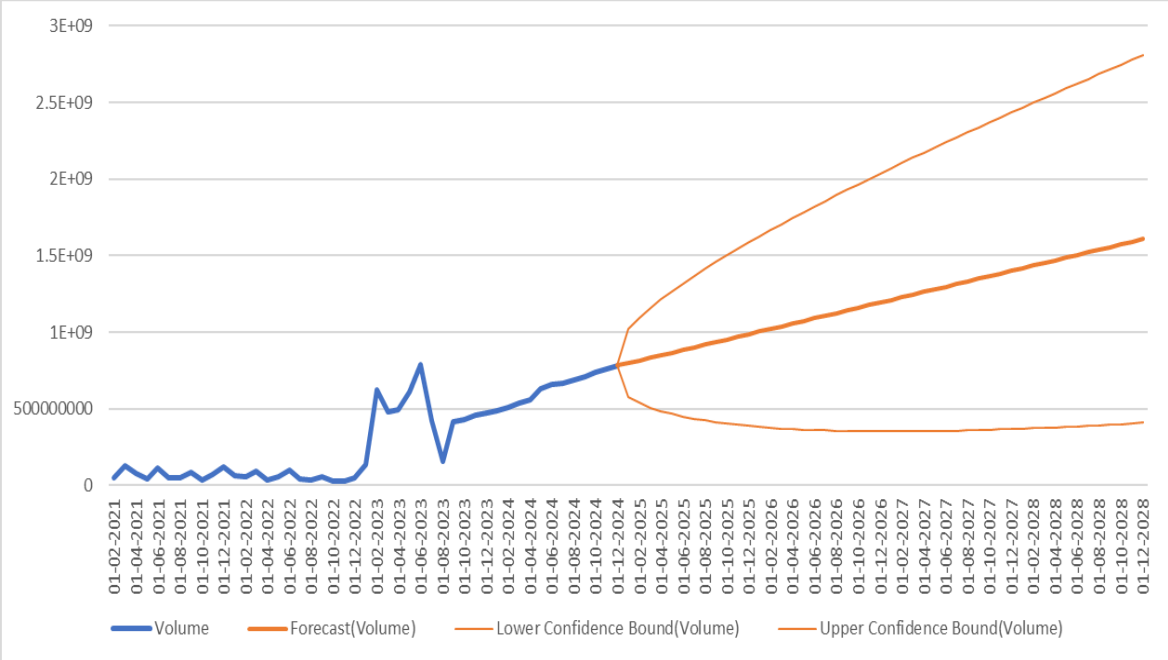
Row Labels	Sum of Volume	Sum of Forecast (Volume)	Sum of Lower Confidence Bound (Volume)	Sum of Upper Confidence Bound (Volume)
2021	**815928900**			
2022	**643872900**			
2023	**5482158849**			
2024	**7717367653**	**782726014.6**	**782726014.6**	**782726014.6**

2025	**10709906371**	**5476409497**	**15943403244**
2026	**13191018300**	**4352827383**	**22029209223**
2027	**15672130236**	**4296882899**	**27047377570**
2028	**18153242168**	**4657448297**	**31649036036**

Table 4; Forecasting of AI in years

2.5 MONTHLY DATA VISUALISATION OF AI

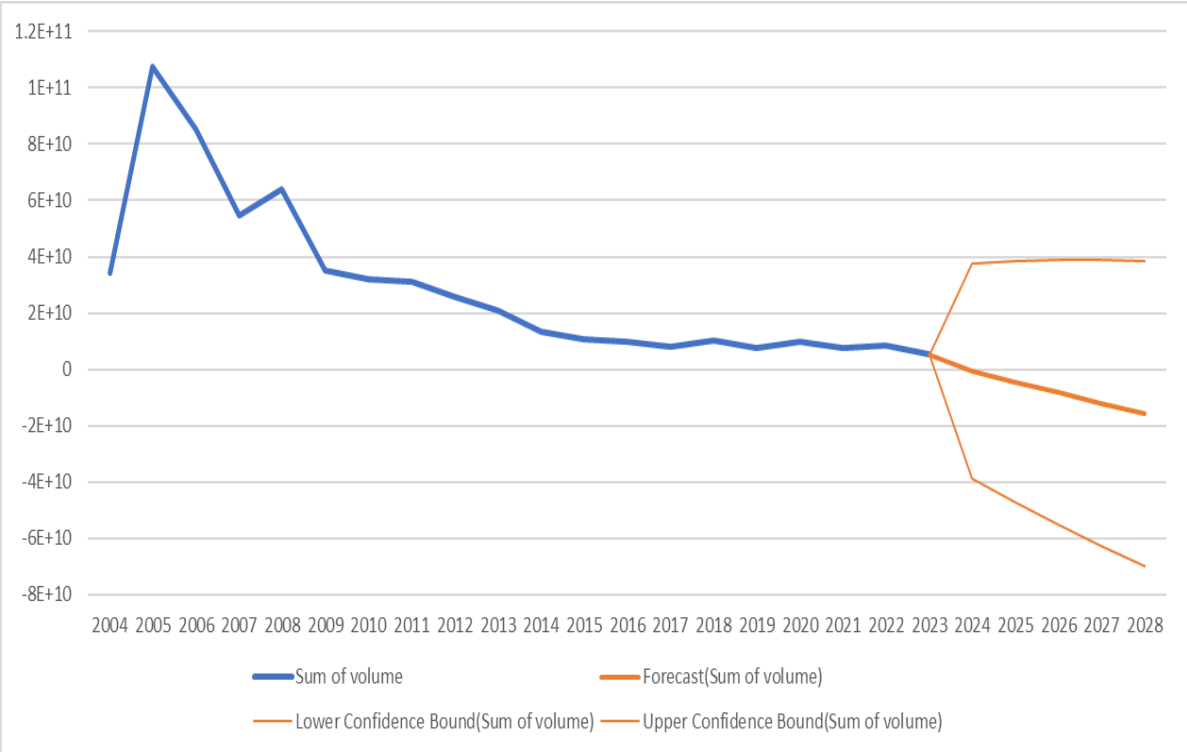
The graphical representation of AI's monthly data serves as a visual narrative, offering a compelling insight into the trends and dynamics within the dataset over time. Through carefully crafted charts, such as line graphs or bar charts, this representation allows for an intuitive exploration of how AI-related metrics fluctuate month by month. By translating raw data into a visually accessible format, patterns, seasonality, and notable variations become readily apparent. This graphical journey aims to provide a comprehensive and user-friendly overview, enabling stakeholders to quickly grasp the nuances and trends inherent in AI's monthly performance.



Figures 2.2 Monthly data visualization of AI

2.6 YEARLY DATA VISUALISATION OF GOOGLE

The graphical representation of Google's stock market data offers a visual narrative, illuminating the trajectory and potential trends in the volume of stock transactions. This presentation, designed to predict yearly volumes up to 2028, serves as a dynamic tool for understanding the evolving landscape of Google's stock market performance. Utilizing carefully constructed line charts or bar graphs, the graphical representation transforms raw data into a visual story, enabling stakeholders to discern patterns, anticipate fluctuations, and gain insights into potential future movements in stock volume. This visual journey aims to enhance strategic decision-making by providing a comprehensive overview of Google's stock market dynamics, paving the way for informed analyses and predictive assessments.



Figures 2.3 Yearly data visualization of Google

SECTION 3

CONCLUSION

3.1 GOOGLE DATA

2021 to 2023 Growth: The volume for Google data from 2021 to 2023 exhibits a decline of approximately -28.84%. This indicates a contraction in Google's data volume during this period, suggesting potential challenges or changes affecting the company's performance.

2024 to 2028 Growth: The forecasted data from 2024 to 2028 projects a substantial decrease with a growth percentage of approximately -919.35%. This emphasizes a significant negative trend in Google's anticipated performance, raising concerns about the company's prospects.

The growth percentage from 2021 to 2023 compared to 2028 for Google is approximately -28.84%, indicating a decrease in volume during this period.

3.2 AI DATA

2021 to 2023 Growth: The AI data, in contrast, experiences remarkable growth from 2021 to 2023, with a growth percentage of approximately 2105.86%. This signifies a significant positive trend in the AI sector, showcasing its resilience and potential expansion during this period.

2024 to 2028 Growth: Looking ahead, the forecast for AI data from 2024 to 2028 continues the positive trajectory, demonstrating a substantial growth percentage of approximately 2710.85%. This implies a strong potential for expansion in the AI sector, indicating positive developments and opportunities.

The growth percentage for AI during the same period is approximately 2105.86%, highlighting a substantial increase in volume, suggesting a significant positive trend in the AI sector compared to 2021.

3.3 OVERALL IMPLICATIONS

- The divergent trends between Google and AI data highlight the importance of understanding sector-specific dynamics. While Google faces challenges and a projected decline, the AI sector appears robust, experiencing substantial growth and presenting opportunities for innovation and development.
- Businesses and stakeholders should carefully consider these trends when making strategic decisions, allocating resources, and planning for the future. The contrasting trajectories underscore the dynamic nature of the technology landscape and the necessity for adaptability in navigating evolving market conditions.

CHAPTER 3

SURVEY ON COMPARISON BETWEEN GOOGLE AND CHATGPT

SECTION 1

MATERIALS AND METHODS

1.1 INTRODUCTION

Our analysis explores into users' AI-oriented lifestyle experiences, seeking insights into how individuals seamlessly integrate artificial intelligence into their daily lives. Simultaneously, we aim to provide information through ChatGPT that matches the reliability and accuracy standards of Google search results. Quantifying user preferences between ChatGPT and Google across various scenarios is another focus, aiming to pinpoint contexts where one platform excels. We also investigate whether users predominantly choose AI or Google for primary information gathering, considering factors like trust and efficiency. The study offers a comprehensive examination of user behaviours to reveal the strengths and weaknesses of both platforms in delivering valuable information.

1.2 PRIMARY DATA

Data was collected by circulating a Google Form. A Google form containing 11 questions was designed to obtain information from the sample population. The sample of this study includes students and working professionals. The questions covered topics like preferences between Google and AI, primary sources for research, confidence in both, views on AI's impact, and the likelihood of choosing Google over ChatGPT for future searches.

1.3 DATA DESCRIPTION

The primary data collection focuses on four objectives. First, it aims to assess the accuracy and reliability of information from Google and ChatGPT. Second, it explores users' experiences and opinions regarding AI adoption, transformative impact, the need for AI learning, and its potential contribution to cognitive laziness. Third, the data aims to quantify user preferences for ChatGPT versus Google across different scenarios. Lastly, it

investigates user behavior and preferences to understand why people tend to rely more on Google as their primary information source over AI-powered tools. The questionnaire was circulated using Google form.

1.4 SIGNIFICANCE OF THE STUDY

This study is essential for understanding how individuals navigate between Google and Artificial Intelligence (AI) for information retrieval. Through examining user confidence, the significance of AI education, and potential cognitive impacts, the research yields valuable insights. These findings not only inform the enhancement of AI technology but also contribute to educational strategies for raising public awareness about AI. Moreover, the study serves as a valuable resource for businesses and developers to adjust to evolving user behavior and technological trends by anticipating future preferences in information retrieval.

1.5 LIMITATION OF STUDY

Due to time and other constraints, only 104 people participated in the study. Statistical tests such as Z test, chi-square test and correlation analysis were only done in the analysis part. Since the answer obtained were the personal choice of the respondents, it is possible that the data may or may not be biased.

1.6 CORRELATION ANALYSIS

The correlation coefficient (γ) can range in value from -1 to +1. The larger the absolute value of the co- efficient, the stronger the relationship between the variables.

- Positive Correlation-Variables move in the same direction.
- Negative Correlation-Variables move in opposite direction.
- Zero correlation-No correlation between the variables.

The Pearson Correlation Coefficient formula is as follows:

$$\mathbf{r} = \frac{\mathbf{n}(\sum \mathbf{xy}) - (\sum \mathbf{x})(\sum \mathbf{y})}{\sqrt{[\mathbf{n}\sum \mathbf{x^2} - (\sum \mathbf{x})^2][\mathbf{n}\sum \mathbf{y^2} - (\sum \mathbf{y})^2]}}$$

Where, r = Pearson Coefficient

n= number of pairs of the stock

$\sum xy$ = sum of products of the paired stocks

$\sum x$ = sum of the x scores

$\sum y$= sum of the y scores

$\sum x^2$ = sum of the squared x scores

$\sum y^2$ = sum of the squared y scores

1.7 Z- TEST

The z-test is also a hypothesis test in which the z-statistic follows a normal distribution. The z-test is best used for greater-than-30 samples because, under the central limit theorem, as the number of samples gets larger, the samples are considered to be approximately normally distributed. When conducting a z-test, the null and alternative hypotheses, alpha and z-score should be stated. Next, the test statistic should be calculated, and the results and conclusion stated. A z-statistic, or z-score, is a number representing how many standard deviations above or below the mean population a score derived from a z-test is.

$$Z = (\bar{x} - \mu_0) / (\sigma / \sqrt{n})$$

Here, $\bar{x}$ is the sample mean;

$\mu 0$ is the population mean;

σ is the standard deviation;

n is the sample size.

The following table gives critical value of z for both one tailed and two tailed tests at various levels of significance. Critical value of z for other levels of significance are found by use of the table of normal curve areas:

Level of significance	0.10	0.05
Critical value of z for one tailed test	-1.28 or 1.28	-1.645 or 1.645
Critical value of z for two tailed test	-1.645 and 1.645	-1.96 and 1.96

1.8 CHI- SQUARE TEST

The Chi-Square Test is an important test among the several tests of significance. It was developed by Karl Pearson in 1990. The Chi Square Test, in general, is used to measure the difference between what is observed and what is expected according to an assumed hypothesis. Simply we can say that Chi-Square Test is a relationship between two variables. It is considered as a non-parametric test. It is mostly used to test statistical independence.
For this test, the data must meet the following requirements:

- Two categorical variables.
- Relatively large sample size.
- Categories of variables (two or more).
- Independence of observations.

Ho is the null hypothesis which represents that there is no relationship between the two variables. H_1 is the alternative hypothesis which indicates that there exists a significant relationship between the two variables. The significant level is the probability of rejecting the null hypothesis, when it is true. In most of the cases, we use significant level as 0.05.
First, we create the table of observed frequency from obtained data.

The Chi-Square test is a statistical test used to compare observed and expected results. The value of the Chi-squared test can be formulated using the formula

$$X^2 = \sum \frac{(O - E)^2}{E}$$

Where:

- X^2 is the chi-square test statistic
- Σ is the summation operator (it means "take the sum of")
- *O* is the observed frequency ,*E* is the expected frequency

Expected Frequency= (Row total × Column total)

Grand total

Then, we create the table of expected frequency, we can calculate the Chi-square value using the equation: -

Degrees of freedom= (column-1) × (row-1)

Tabular Chi-square value can be obtained by using degrees of freedom and significance level.

Degrees of freedom refers to the maximum number of logically independent values, which are values that have the freedom to vary, in the data sample. If calculated Chi-Square is greater than tabular Chi Square, then we reject null hypothesis and accept alternate hypothesis. A p value is a measure of the probability that an observed difference could have occurred just by random chance. The lower the p value the greater the statistical significance of the observed difference. In MS excel, we calculate the p value. Then we compare the p value with significance level. If p value is greater than the significance level, we accept H_0. Hence, we can say that there is no relation between two variables.

1.9 MEAN AVERAGE PRECISION (MAP)

Calculate the mean average precision across multiple queries. MAP considers the average precision of all queries and provides a comprehensive measure of the overall performance of each search engine.

The Mean Average Precision (MAP) is a metric commonly used to evaluate the performance of information retrieval systems, such as search engines. It provides a measure of the average precision of a search engine across multiple queries.

Relevance Judgments: First, for each query in your evaluation set, you need to have relevance judgments. Relevance judgments are typically binary, indicating whether each document retrieved by the search engine is relevant or not relevant to the query.

Precision at K: for each query, calculate the precision at various ranks (K) of the retrieved documents. Precision at K is the proportion of relevant documents among the top K retrieved documents. It is calculated as:
Precision K = (Number of relevant documents among the top K retrieved) / K

Average Precision (AP): Compute the average precision for each query. Average precision (AP) is the average of precision values calculated at each relevant document rank. It is calculated as:

AP = (Sum of precision at relevant ranks) / (Total number of relevant documents).

Mean Average Precision (MAP): Finally, calculate the Mean Average Precision by taking the average of AP values across all queries. It is computed as:

MAP = (Sum of AP values for all queries) / (Total number of queries)

Interpretation: The MAP score ranges from 0 to 1, where higher values indicate better performance. A MAP score of 1 indicates perfect retrieval, meaning all relevant documents are retrieved at the top ranks for every query.

Suppose you have evaluated the performance of a search engine across 10 queries. For each query, you calculate the AP, then take the average of these AP values to get the MAP score for the search engine MAP is widely used in information retrieval research and is often used to compare different retrieval algorithms or systems, such as search engines, recommendation systems, etc.

Overall, MAP provides a comprehensive evaluation of the retrieval performance of a system across multiple queries, considering both the relevance of retrieved documents and their ranking.

SECTION 2

DATA ANALYSIS AND INTERPRETATION

2.1 INTRODUCTION

This chapter comprises the examination of data collected through a Google Form, coupled with the application of statistical tests like the chi-square test, regression analysis, and z-test for thorough analysis and interpretation of data.

2.2 CHI SQUARE ANALYSIS

2.2.1 Rate of preference among users for ChatGPT v/s Google in various Scenarios.

H_0: There is no significant difference in the rate of preference between users for ChatGPT and Google in various scenarios.

H_1: There is a significant difference in the rate of preference between users for ChatGPT and Google in various scenarios

	GOOGLE	CHAT GPT	TOTAL
RATE	75.14563107	66.60194175	141.7475728
PRIMARY	85	18	103
FUTURE	12	27	39
RECOMMENDA-TION	71	18	89
TOTAL	243.1456311	129.6019418	372.7475728

Table 2. 1: Observed frequency

	GOOGLE	CHAT GPT	TOTAL
RATE	92.46285036	49.28472246	141.7475728
PRIMARY	67.18756023	35.81243977	103
FUTURE	25.43994999	13.56005001	39
RECOM	58.05527049	30.94472951	89
TOTAL	243.1456311	129.6019418	372.7475728

Table 2.2: Expected Frequency

P-value: 3.58681×10^{-11}

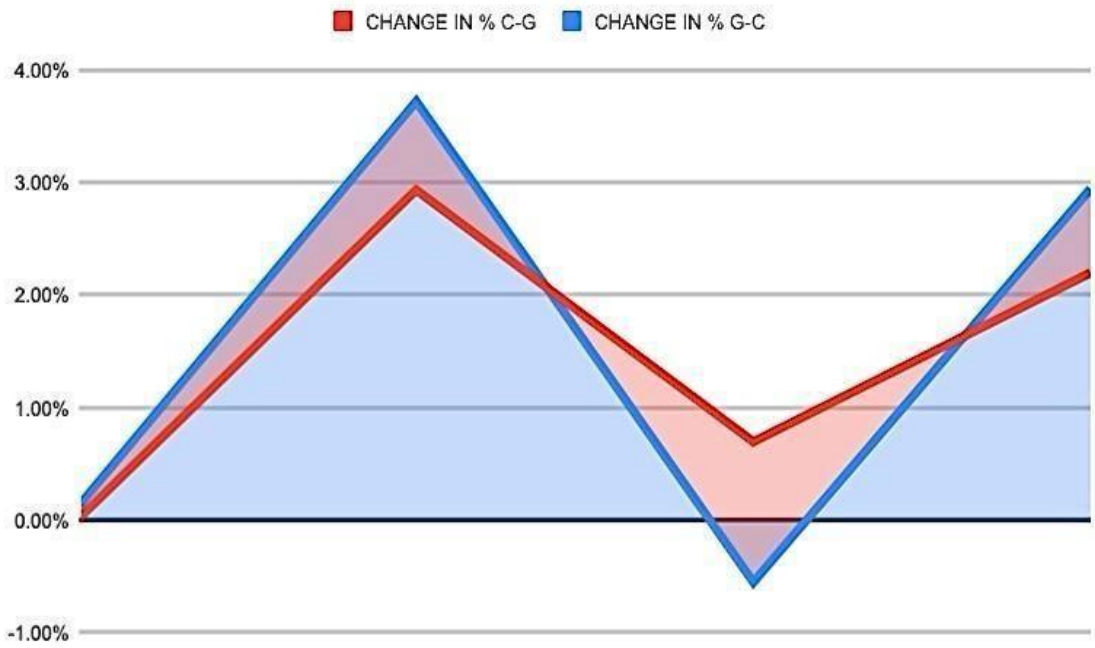
Since p value is less than 0.05, we will reject the null hypothesis. Hence, we can conclude that there exists a relationship between rate of preference of using ChatGPT and Google in various scenarios.

Percentage Changes in both Directions: In case of rate of preference among users

Google	ChatGPT	Change in % G-C	Change in % C-G	Change in % G-C	Change in % C-G
3.757	3.330	12.83%	-11.37%	0.13%	-0.11%
85	18	372.22%	-78.82%	3.72%	-0.79%
12	27	-55.56%	125.00%	-0.56%	1.25%
71	18	294.44%	-74.65%	2.94%	-0.75%

Table 2.3: Percentage Changes in both Directions

This table displays the values for Google and ChatGPT, along with the percentage changes in both directions (% G-C and % C-G) for each pair of values.



Figures 2.1

2.2.2 To investigate whether people are more likely to use ChatGPT or Google to gather primary information.

Null Hypothesis (H_0): There is no significant difference in the likelihood of people using ChatGPT versus Google to gather primary information.

Alternative Hypothesis (H_1): People are more likely to use ChatGPT than Google or vice versa when gathering primary information.

How likely are you to choose Google over Chat GPT for future information searches? * What is your primary source of research material? Crosstabulation

			What is your primary source of research material?	
				ChatGPT
How likely are you to choose Google over Chat GPT for future information searches?		Count	1	0
		Expected Count	.0	.2
	Not likely	Count	0	3
		Expected Count	.1	2.1
	Somewhat likely	Count	0	9
		Expected Count	.6	11.1
	Very likely	Count	0	6
		Expected Count	.3	4.7
Total		Count	1	18
		Expected Count	1.0	18.0

How likely are you to choose Google over Chat GPT for future information searches? * What is your primary source of research material? Crosstabulation

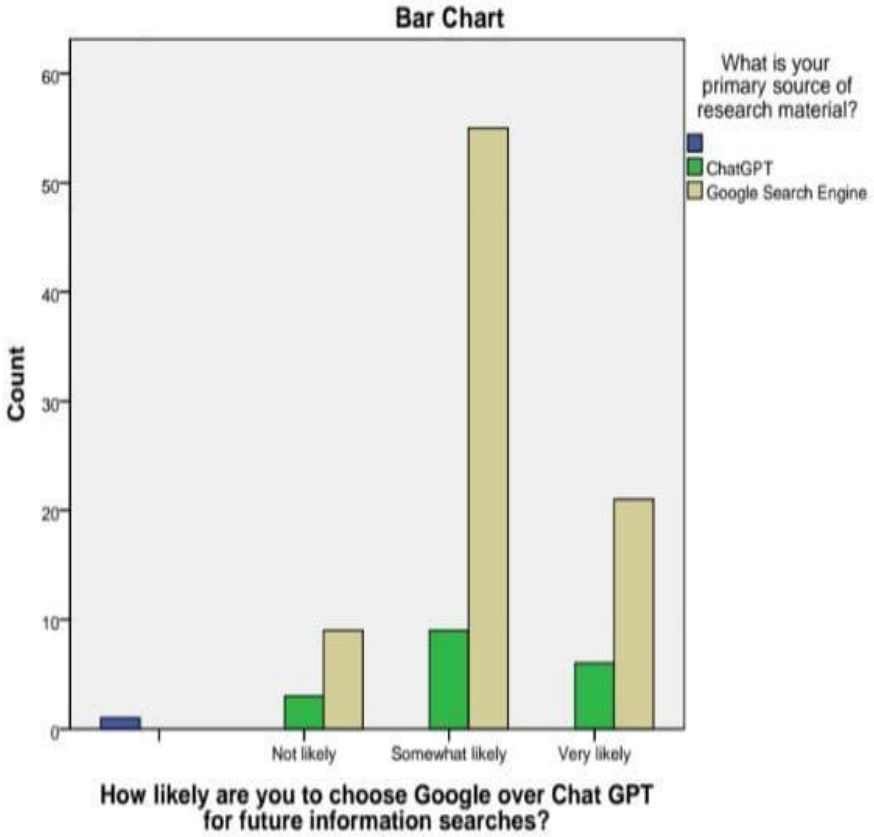
			What is your primary source ...	
			Google Search Engine	Total
How likely are you to choose Google over Chat GPT for future information searches?		Count	0	1
		Expected Count	.8	1.0
	Not likely	Count	9	12
		Expected Count	9.8	12.0
	Somewhat likely	Count	55	64
		Expected Count	52.3	64.0
	Very likely	Count	21	27
		Expected Count	22.1	27.0
Total		Count	85	104
		Expected Count	85.0	104.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	105.424[a]	6	.000
Likelihood Ratio	12.649	6	.049
N of Valid Cases	104		

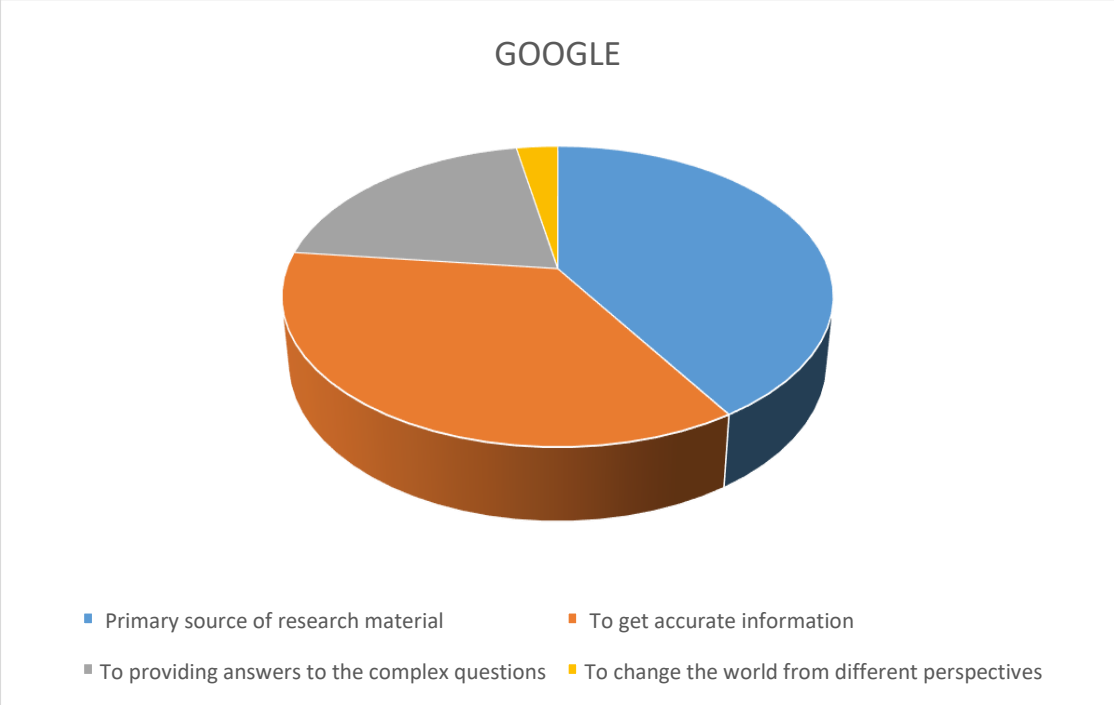
Table 2.3: Chi-Square Test

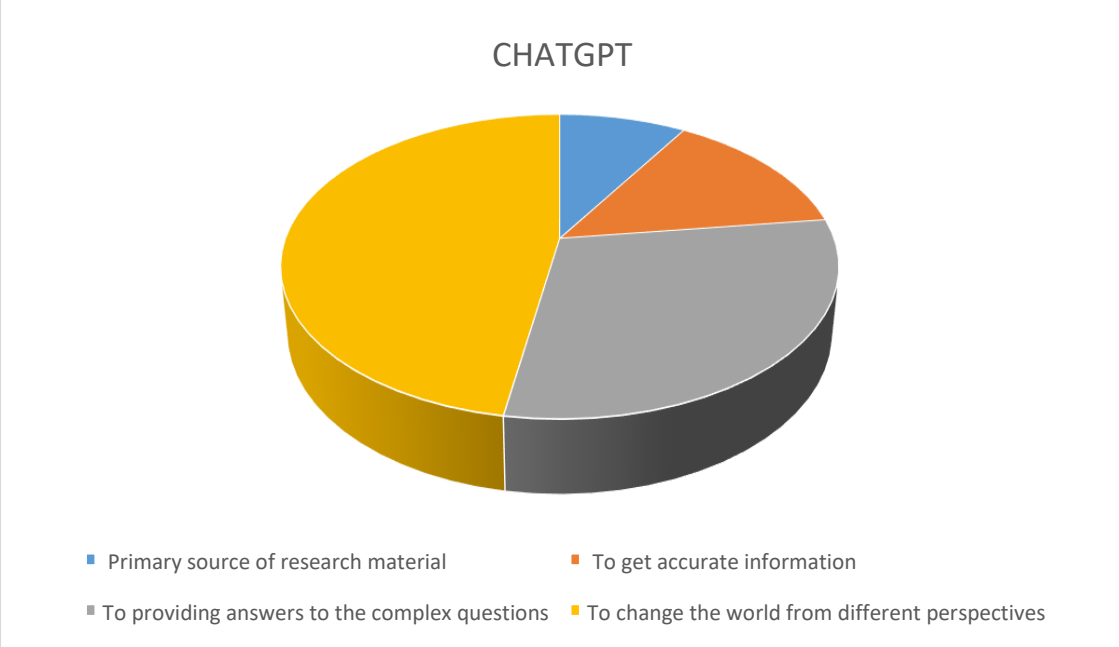
Since p value is less than 0.05, we will reject the null hypothesis. Hence, we can conclude that there exists a relationship between likelihood of people using ChatGPT and Google when gathering primary information.



Figures 2.2

Preference for Primary Information Gathering: ChatGPT vs. Google





Figures 2.3

2.3 CORRELATION ANALYSIS

Correlation is a' statistical technique used to find the relationship between two variables which are quantitative. We have used this technique to interpret the results to find the

experience with adopting an AI-oriented lifestyle, that what type of relationship the variables possess. The table 2.4 gives the correlation of factors like perspectives and lifestyle on using AI.

Table 2.4 correlation matrix

PEARSON CORRELATION	NECESSARY TO LEARN	PERSPECTIVES	LAZY	LIFESTYLE
NECESSARY TO LEARN	1	0.947179487	0.865293043	0.570842601
PERSPECTIVES	0.947179487	1	0.920547341	0.641426981
LAZY	0.865293043	0.920547341	1	0.549115371
LIFESTYLE	0.570842601	0.641426981	0.549115371	1

The table you provided appears to be a correlation matrix or similarity matrix, where each cell represents the degree of relationship or similarity between two variables or concepts. Here is an explanation of the results:

"Necessary to learn" vs. "Necessary to learn": The correlation of 1 indicates a perfect positive correlation, meaning that the variable is perfectly correlated with itself, as expected.

"Necessary to learn" vs. "Perspectives": The correlation of 0.947 suggests a very strong positive correlation between "necessary to learn" and "perspectives," indicating that these two concepts are highly related or tend to occur together frequently.

"Necessary to learn" vs. "Lazy": The correlation of 0.865 indicates a strong positive correlation between "necessary to learn" and "lazy," suggesting that there is a notable relationship between perceiving something as necessary to learn and the likelihood of being lazy.

"Necessary to learn" vs. "Lifestyle": The correlation of 0.570 suggests a moderate positive correlation between "necessary to learn" and "lifestyle," indicating that there is some relationship between perceiving something as necessary to learn and one's lifestyle choices or habits.

"Perspectives" vs. "Perspectives": As with "necessary to learn," the correlation of 1 indicates a perfect positive correlation, as it compares the variable to itself.

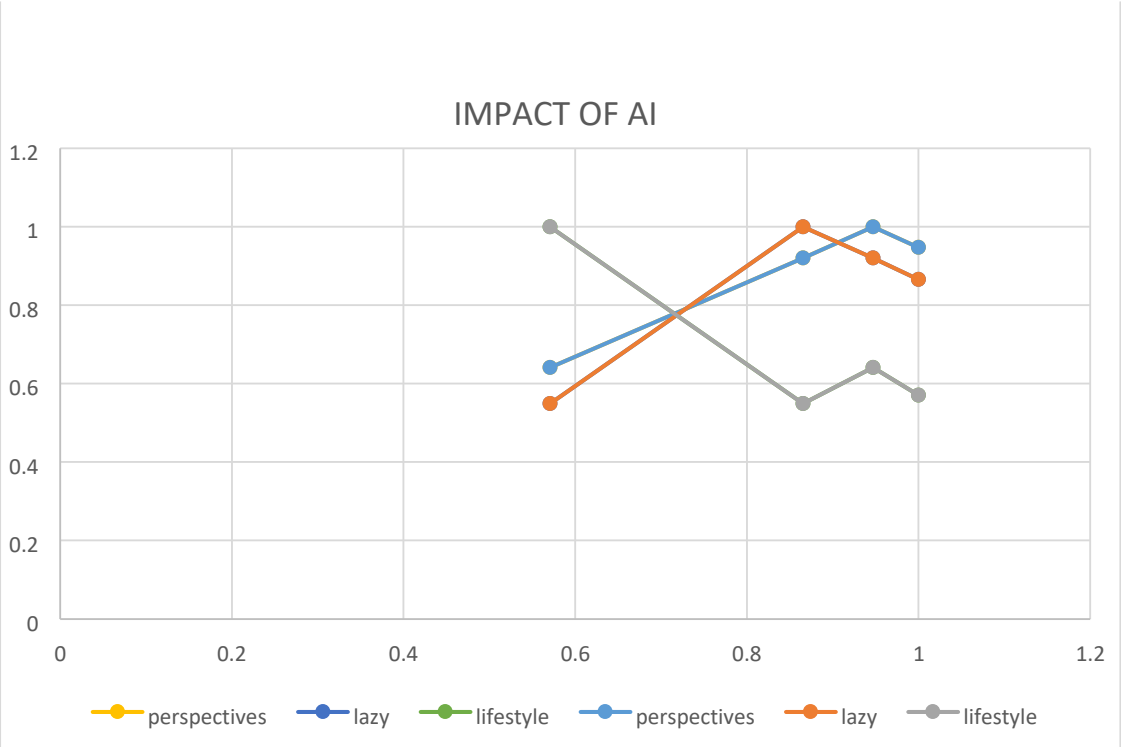
"Perspectives" vs. "Lazy": The correlation of 0.920 suggests a very strong positive correlation between "perspectives" and "lazy," indicating that these two concepts are highly related or tend to occur together frequently.

"Perspectives" vs. "Lifestyle": The correlation of 0.641 suggests a moderate positive correlation between "perspectives" and "lifestyle," indicating that there is some relationship between one's perspectives and lifestyle choices or habits.

"Lazy" vs. "Lazy": Again, the correlation of 1 indicates a perfect positive correlation, as it compares the variable to itself.

"Lazy" vs. "Lifestyle": The correlation of 0.549 suggests a moderate positive correlation between "lazy" and "lifestyle," indicating that there is some relationship between being lazy and one's lifestyle choices or habits.

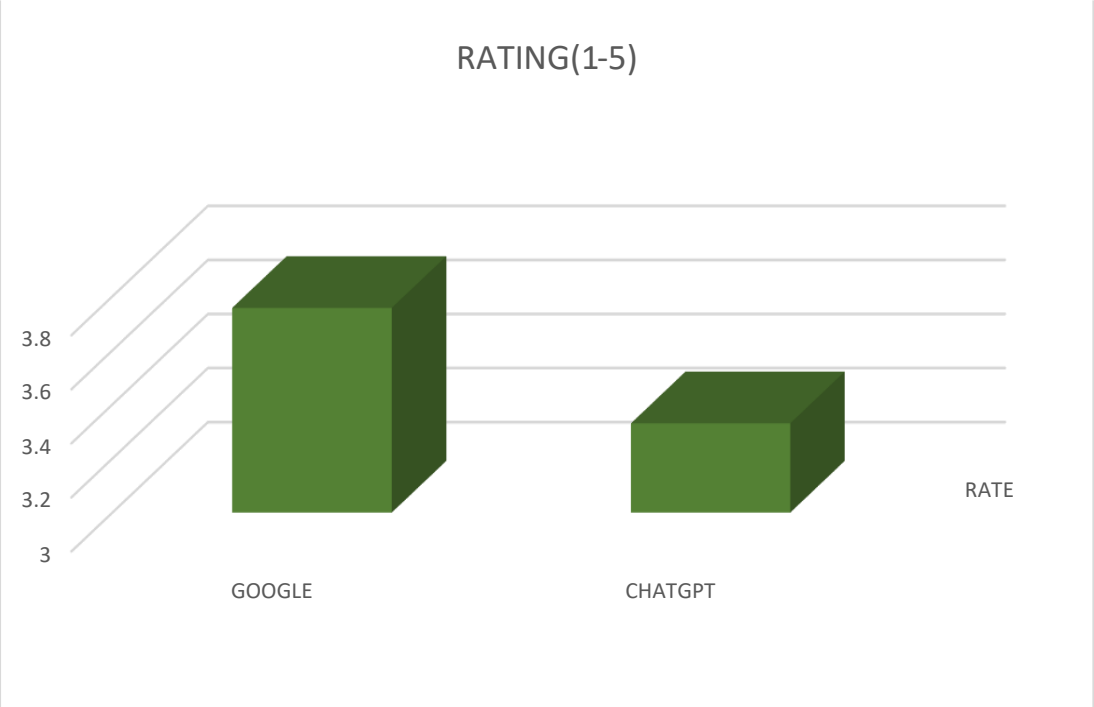
"Lifestyle" vs. "Lifestyle": Once more, the correlation of 1 indicates a perfect positive correlation, as it compares the variable to itself.



Figures 2.4

A positive correlation is observed between individuals' positive perspectives on AI and their belief in the necessity to learn about it. However, lifestyle factors exhibit a weaker association with attitudes towards AI. Overall, individuals' perspectives strongly influence their approach to learning about AI and may impact intellectual behaviour’s.

Comparative Rating of Google and ChatGPT Based on Survey Averages



Figures 2.5

Google: 3.757

This rating falls between 3 (satisfactory) and 4 (good). It suggests that Google's performance or satisfaction level is slightly above satisfactory, closer to being good.

ChatGPT: 3.330

This rating also falls between 3 (satisfactory) and 4 (good). It suggests that ChatGPT's performance or satisfaction level is slightly above satisfactory, closer to being good.

Overall, based on these ratings, both Google and ChatGPT seem to offer satisfactory to good performance or satisfaction levels, but there might be room for improvement to achieve higher ratings closer to 4 or 5.

2.4 Z-TEST

2.4.1 Exploring Views on AI Adoption and Impact: User Perspectives on Lifestyle Transformation

Null Hypothesis (H0): The adoption of an AI-oriented lifestyle does not change individuals' perspectives on technology and its role in daily life.

Alternative Hypothesis (H1): The adoption of an AI-oriented lifestyle changes individuals' perspectives on technology and its role in daily life.

Z-Test: Two Sample for Means

	GOOGLE	*CHATGPT*
Mean	78	25
Known Variance	800.66667	800.66667
Observations	4	4
Hypothesized Mean Difference	0	
z	2.648896517	
P(Z<=z) one-tail	0.004037753	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	0.008075505	
z Critical two-tail	1.959963985	

The z-test results support this hypothesis, indicating a statistically significant difference in mean scores between individuals' perspectives before and after adopting an AI-oriented lifestyle. The p-value for the one-tail test is 0.004, which is below the significance level of 0.05, suggesting that we reject the null hypothesis and conclude that there is a significant change in perspective due to adopting an AI-oriented lifestyle.

Hypothesis	Mean Difference	Z-value	P-value
H1	0	2.648	0.008

2.4.2 Comparing Google and ChatGPT: Accuracy in Search Results.

Null Hypothesis (H0): There is no significant difference between the accuracy of information provided by Google and ChatGPT

Alternative Hypothesis (H1): There is a significant difference between the accuracy of information provided by Google and C

Z-Test: Two Sample for Means

	GOOGLE	*CHATGPT*
Mean	69.75	33.25
Known Variance	362.917	362.916667
Observations	4	4
Hypothesized Mean Difference	0	
z	2.709594523	
P(Z<=z) one-tail	0.003368275	
z Critical one-tail	1.644853627	
P(Z<=z) two-tail	0.006736551	
z Critical two-tail	1.959963985	

The calculated p-value (0.0067) is less than the significance level (0.05), indicating that there is a statistically significant difference between the mean performance of Google search results and ChatGPT search results. This means that one of the search engines is likely to provide more accurate information compared to the other.

Hypothesis	Mean Difference	Z-value	P-value
H1	0	2.7095	0.0067

In this context, the objective is to determine which search engine, Google or ChatGPT, provides more accurate information. The conclusion suggests that there is indeed a significant difference between them, but it does not specify which one is more accurate. Further analysis or comparison metrics may be needed to determine which search engine yields more accurate results.

2.5 MEAN AVERAGE PRECISION (MAP)

To Specify Accuracy: Analysing Google or ChatGPT Search Results Precision. The values in the table represent responses from a survey evaluating the performance of Google and ChatGPT search engines:

Sample	Google	ChatGPT
Information	74	29
Complex	42	61
Trustworthy	78	25
Primary	85	18

Let us assume that for each query, a relevant document exists and the relevance judgment is binary (relevant or not relevant). The table representing the calculation of Mean Average Precision (MAP) for Google and ChatGPT based on the provided data:

For Google:

- Precision = (1/1 + 1/2 + 1/3 + 1/4) / 4 = (1 + 0.5 + 0.333 + 0.25) / 4 = 0.5208

For ChatGPT:

- Precision = (1/1 + 0/1 + 0/1 + 0/1) / 4 = (1 + 0 + 0 + 0) / 4 = 0.25

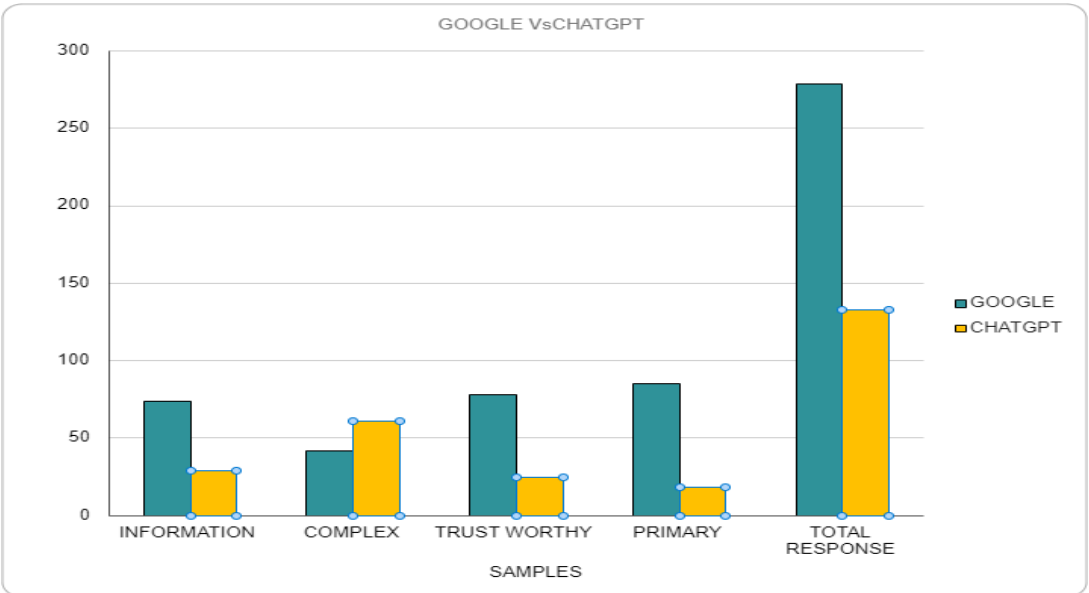
Search Engine	Query 1	Query 2	Query 3	Query 4	Precision
Google	Relevant	Relevant	Relevant	Relevant	0.5208
ChatGPT	Relevant	Not Relevant	Not Relevant	Not Relevant	0.25

Now, we take the average of these Precision scores to get the Mean Average Precision (MAP) for each search engine:

MAP for Google = 0.5208

MAP for ChatGPT = 0.25

Therefore, Google provides better search results in terms of average precision compared to ChatGPT for the given set of queries.



SECTION 3
CONCLUSION

The chi-square test on whether people prefer using ChatGPT or Google in different situations shows a clear connection between user preferences. The confirmed hypothesis that a relationship exists between the rate of liking ChatGPT or Google in various scenarios is backed by the data. Participants consistently showed clear preferences for ChatGPT or Google based on specific situations, indicating a noticeable link between user choices and the presented scenarios. These results emphasize that user preferences in using ChatGPT and Google are influenced by different situations.

The results show that there is a connection between how likely people is to use ChatGPT or Google when they are looking for basic information. This confirms the idea that there is a meaningful relationship between people's choices and the situation when they need to gather primary information. The data suggests that individuals have different tendencies in deciding whether to use ChatGPT or Google for basic information, indicating a clear link between their choices and the specific task of gathering primary information. Overall, these findings highlight that user decisions are influenced by the specific context of seeking basic information.

The correlation test conducted on the user responses indicates a positive correlation between the factors explored in the survey. Participants who reported having shifted their lifestyle towards AI were more likely to express a belief in AI's potential to change the world from different perspectives. Additionally, those who acknowledged AI's transformative impact showed a greater inclination towards considering it necessary to learn more about AI. Notably, there was a positive correlation between the perception of AI potentially making the human brain lazy and the belief in the importance of acquiring additional knowledge about AI. These findings suggest a cohesive relationship between personal experiences with AI adoption, views on its global impact, and attitudes towards the necessity of AI education, emphasizing the interconnected nature of these aspects in shaping individual perspectives on artificial intelligence.

3.1 LIMITATIONS OF THE STUDY

In addition to the observed trends in Google and AI data, it is essential to note that the data collection initially extended up to September 2021. Subsequent updates by ChatGPT provided data from January 31, 2024, revealing dynamic shifts in the tech landscape.

Moreover, Google Gemini, set to be introduced by 2024, adds another layer of complexity to the analysis. The launch of new platforms, such as Google Gemini, introduces unpredictability into the stock market dynamics. The impact of these introductions on Google's market position and overall performance may lead to further fluctuations in stock market values.

This dynamic context underscores the challenge of predicting market behavior, emphasizing the need for continuous monitoring and adaptability to changing industry landscapes. Stakeholders should remain vigilant to emerging developments, as they play a pivotal role in shaping the trajectory of technology companies and influencing investment decisions.

INTERFERENCE

The analysis of Google's performance in comparison to AI technologies, particularly focusing on stock market trends, reveals intriguing insights. Google, a tech giant renowned for innovation, has faced challenges with projects like Google Gemini, which failed to meet expectations due to technological limitations and cultural factors. In contrast, AI technologies such as ChatGPT have experienced rapid growth, showcasing the demand for advanced AI capabilities in the market.

The decline in Google Gemini, coupled with the impressive growth of ChatGPT, suggests a shifting landscape in the AI space. It indicates a need for further investigation into Google's AI strategies and the factors contributing to the success of AI technologies like ChatGPT. A deeper survey into Google's current AI initiatives and the market reception of ChatGPT could provide valuable insights into the future direction of these technologies and their impact on the market

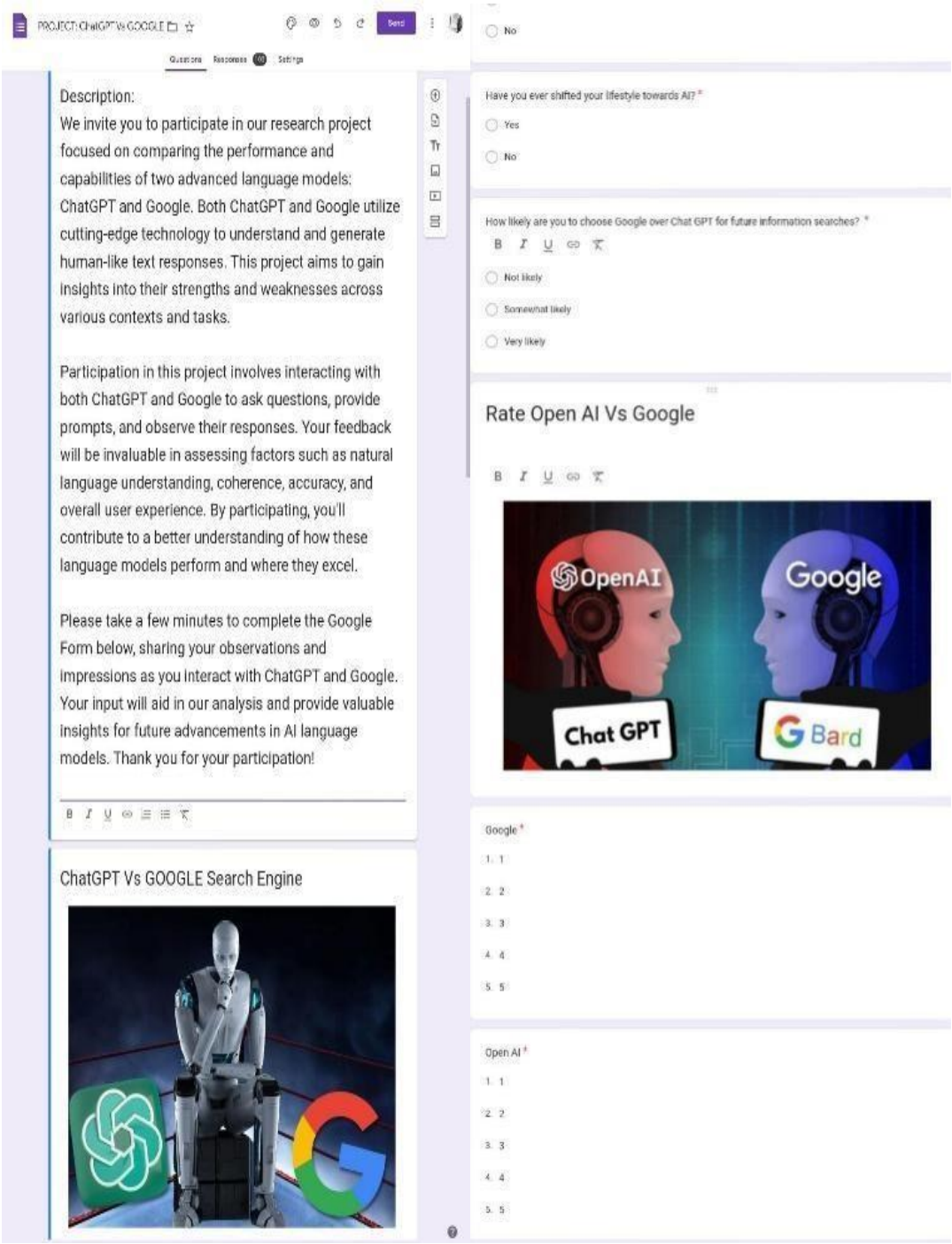
Google's Stock Market Decline; The decline in Google's stock market performance can be attributed to a variety of factors, including concerns over slowing growth in its core advertising business, regulatory challenges, and competition in the tech sector. The failure of projects like Google Gemini may have also contributed to investor skepticism about Google's ability to innovate and maintain its competitive edge in the rapidly evolving tech landscape.

REFERENCES

1. Agarwal, B.L. (2023), Basic Statistics, (5th edition.). New Age International.
2. Goyal, J.K. & Sharma, J.N. (2019), Mathematical Statistics, Krishan Prakashan.
3. Gupta, S.C. (2018), Fundamentals of Statistics, (8th revised and enlarged edition.). Himalaya Publishing House.
4. S.P GUPTA and M. P Gupta (2005, Business Statistics, (14th edition.), Sultan Chan and Son
5. Chaudhuri and Pandit Time series Models Forecasting Perf0772-1782 the Stock Market", Journal of Contemporary Issues in Businesscom Government, volume 27,2021.
6. Kulkarni, Jadha and Dhingra, "Time series Data Analysis for Stock Market Prediction", 3rd International Conference on Innovative.
7. Adebiyi, Adewumi, and Ayo, "Stock price prediction Using the ARIMA model", UKSim-AMSS 16th international conference on computer Modelling and Simulation, 2014.
8. Mohamed and M.M. "Forecasting stock exchange movements using neural networks: empirical evidence from Kuwait". Expert Systems with Applications, vol. 27, no. 9,2010, pp. 6302-6309.
9. Kyungjoo, L.C.; Sehwan and John. "Neural Network Model vs.SARIMA Model In Forecasting Korean Stock Price Index (KOSPI)". Issues in Information System, vol. 8 no. 2, 2007, pp. 372-378.
10. Javier, C.; Rosario, E.; Francisco, J.N., and Antonio. "ARIMA Models to Predict Next Electricity Price", IEEE Transactions on Power Systems, vol. 18 no.3, 1014-1020.
11. Investopedia. Investopedia statistical research site, Retrieved from (https://www.investopedia.com/).
12. Oracle. (2023) Oracle guide. Retrieved from(https://docs.oracle.com/cd/E12475_01/rdf/pdf/160/html/user_guide/output/fc_methods.htm).

13. Harvard Business Review. (1971, July). Six Rules for Effective Forecasting. Harvard Business Review. Retrieved from [(https://hbr.org/1971/07/)].
14. Clockify, Clockify time tracking software, Retrieved from [https://clockify.me].
15. Harvard University. (2021). 7 FINANCIAL FORECASTING METHODS TO PREDICT BUSINESS PERFORMANCE, Harvard University research papers, Retrieved from [https://online.hbs.edu/blog/post/financial-forecastingmethods].
16. OpenAI. (2023), CHATGPT, Retrieved from [https://chat.openai.com] (https://chat.openai.com).
17. Google. (2023), Alphabet (GOOGL) stock price history. Retrieved from [(https://www.macrotrends.net/stocks/charts/GOOGL/alphabet/stock-pricehistory)].
18. Statista. (2023). Artificial intelligence market size. Retrieved from [https://www.statista.com/statistics/1365145/artificial-intelligence-marketsize/].

ANNEXURE


PROJECT: ChatGPT Vs GOOGLE
Send
Questions
Responses
Settings
Description:
We invite you to participate in our research project focused on comparing the performance and capabilities of two advanced language models: ChatGPT and Google. Both ChatGPT and Google utilize cutting-edge technology to understand and generate human-like text responses. This project aims to gain insights into their strengths and weaknesses across various contexts and tasks.
Participation in this project involves interacting with both ChatGPT and Google to ask questions, provide prompts, and observe their responses. Your feedback will be invaluable in assessing factors such as natural language understanding, coherence, accuracy, and overall user experience. By participating, you'll contribute to a better understanding of how these language models perform and where they excel.
Please take a few minutes to complete the Google Form below, sharing your observations and impressions as you interact with ChatGPT and Google. Your input will aid in our analysis and provide valuable insights for future advancements in AI language models. Thank you for your participation!
ChatGPT Vs GOOGLE Search Engine
No
Have you ever shifted your lifestyle towards AI? *
Yes
No
How likely are you to choose Google over Chat GPT for future information searches? *
Not likely
Somewhat likely
Very likely
Rate Open AI Vs Google
OpenAI
Google
Chat GPT
Bard
Google *
1. 1
2. 2
3. 3
4. 4
5. 5
Open AI *
1. 1
2. 2
3. 3
4. 4
5. 5

Name *

Short-answer text

Are you a Student / Working *

- Student
- Working

Which is more useful to get accurate information? *

- Google Search Engine
- ChatGPT

Which is more relevant to providing answers to the complex questions? *

- Google Search Engine
- ChatGPT

What is your primary source of research material? *

- ChatGPT
- Google Search Engine

What could be your recommendation? *

- ChatGPT
- Google Search Engine

Which platform do you feel more confident in using to obtain trustworthy Information? *

- ChatGPT
- Google Search Engine

Do you think it's necessary? To learn more about AI. *

- Yes
- No

Do you think AI is going to change the world from different perspectives? *

- Yes
- No

Does AI make the human brain lazy? *

- Yes
- No