Sherlock AI: Enhancing Financial Reporting for LinkedIn Marketing Solutions

Fall 2024 Practicum Report

By Xiaofan Jiao

Abstract

This report presents Sherlock AI, an AI-driven system developed to enhance the efficiency of analyzing business outcomes within LinkedIn Marketing Solutions (LMS), with the aim of improving productivity and precision in financial assessments. Sherlock AI is tailored to accommodate LinkedIn's specific business structure and the interrelationships among various data metrics, thereby facilitating effective detection of anomalies across both short-term and long-term metrics. By automating the labor-intensive processes inherent in LinkedIn's Financial Planning and Analysis (FP&A) function, Sherlock AI identifies key drivers underlying anomalies in business performance and generates insightful, data-driven narratives.

The system employs a combination of machine learning techniques, advanced data analytics, and prompt engineering to deliver insights. Its comprehensive workflow incorporates data preprocessing, anomaly detection, and automated report generation, ensuring both accuracy and scalability in financial analysis. The implementation of Sherlock AI provides LinkedIn with actionable insights that support strategic financial decision-making, enabling a data-driven approach to managing business outcomes and driving organizational growth.

1. Introduction

1.1 Motivation

LinkedIn, recognized as the world's largest professional networking platform, empowers users to form meaningful professional connections, discover new career opportunities, and develop essential skills. Within LinkedIn's Financial Planning and Analysis (FP&A) division, the Strategic Finance Data & Analytics (FP&A) team combines business expertise with advanced analytical methodologies to facilitate data-driven decision-making in evaluating financial performance. The FP&A team plays a critical role in monitoring LinkedIn Marketing Solutions (LMS), which serves as a key driver of revenue by enabling B2C advertising (Smith & Johnson, 2021).

The process of monitoring business performance requires the evaluation of both demand-side and supply-side factors, including metrics such as ad engagement, customer expenditure, and website visitation rates (Davis et al., 2022). It is also important to identify the underlying drivers responsible for metric variances, which enables a deeper understanding of business outcome. The analysis of LinkedIn financial reporting data typically involves the synthesis of multiple data tables, metrics, and time periods—a process that is complex and time-consuming. For instance, a decline in LMS revenue may be attributed to reduced advertiser engagement or a decrease in user visits on LinkedIn, resulting in fewer ad impressions.

The objective of this project is to address these challenges by introducing automation to streamline financial reporting, data analysis, and outcome summarization processes. Automating these functions will allow the FP&A team to focus more on strategic decision-making rather than repetitive data handling, thus enhancing efficiency and effectiveness in financial assessments.

1.2 Problem Statement

Conventional methods for assessing business performance within Financial Planning and Analysis (FP&A) are often labor-intensive and challenging to scale. LinkedIn's FP&A and Strategic Finance teams currently engage in manual analysis across multiple dimensions, including products, geographic regions, and customer segments. Such manual processes face scalability limitations given the complexity of LinkedIn's Learning Management System (LMS), which requires the consideration of 15 metrics, 6 time periods, and 5 dimensions. Each of these elements must be contextualized within the broader business framework to discern performance drivers effectively (Smith et al., 2020).

The LMS business at LinkedIn involves both demand-side and supply-side factors, which further complicates the assessment of business performance. For instance, lower-than-anticipated revenue may be attributed to a decrease in advertisers' willingness to allocate budget towards ads (demand-side) or a reduction in user traffic that limits ad views (supply-side). Additionally, shifts in consumer behavior, such as changes in engagement and interaction patterns, introduce real-time complexities that must be understood to maintain an accurate assessment of performance (Johnson & Lee, 2021). These dynamic and multifaceted influences necessitate the adoption of an automated, data-driven approach to effectively assess and contextualize the impact of these variables on LinkedIn's reporting process.

Therefore, the objective of this project is to develop an AI agent capable of automating the analysis of these business metrics. By leveraging artificial intelligence, the proposed system aims to deliver timely, accurate, and actionable insights that address the limitations of manual analysis, thus improving overall efficiency and supporting data-driven strategic decisions (Davis et al., 2022).

1.3 Objectives

The primary objective of this project is to create an AI system named 'Sherlock' that automates the process of identifying anomalies in LinkedIn's LMS business performance data and generates meaningful narratives about these anomalies. Specifically, Sherlock is designed to:

- Automate the data collection, transformation, and analysis phases of business performance reporting.
- Detect anomalies in key performance metrics across multiple dimensions (e.g., product, region).
- Generate meaningful narratives and insights to support decision-making processes, reduce human intervention, and improve the quality of financial insights provided to LinkedIn's executives.
- Understand both the business model (metric relationships) and data model (data relationships) to generate comprehensive narratives.

1.4 Scope

The scope of this project includes automating data analytics functions for LinkedIn's LMS business data, anomaly detection techniques, systematic report generation, visualization, and narrative construction. This report also discusses various approaches for anomaly detection and the usage of a multi-agent system to enhance the capabilities of Sherlock.

2. Data Analytics

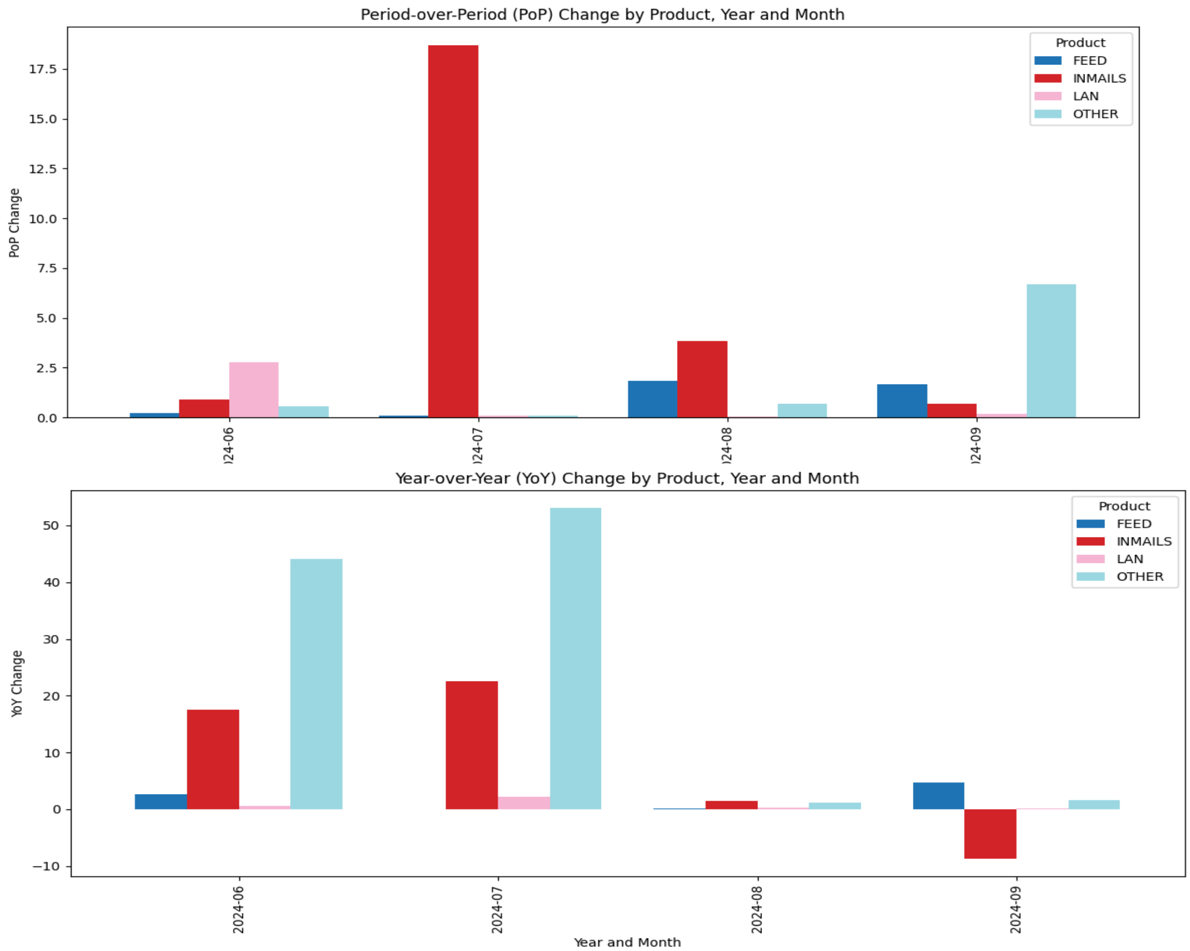
2.1 Data Preparation

The data cleaning and transformation process ensures consistency, completeness, and proper formatting of metrics for business reporting and automation. Below are examples of some data cleaning procedures that we have done.

- Replaced missing period_label values for 'L7D' with '2024-09-22' to ensure valid entries.
- Calculated percentage metrics as the ratio of numerator divided by denominator. Set value to None if denominator is missing or zero.
- Removed raw value columns to streamline the dataset.
- Converted period_label to proper date formats for chronological sorting.

2.2 Exploratory Data Analysis (EDA)

Utilizing techniques such as summary statistics, correlation analysis, and time-series analysis, the agent efficiently detects Year-over-Year (YoY) and Period-over-Period (PoP) trends, while variance detection highlights discrepancies between actual performance and outlook. The charts below show some examples of EDA that we have conducted.



Year over Year analysis:

- "OTHER" grew over 40% YoY in June and July 2024 due to marketing or new features, but growth stabilized in August.
- "INMAILS" showed YoY growth but declined in September, possibly due to competition or reduced engagement.
- "FEED" and "LAN" had flat YoY growth, indicating stagnation needing possibly strategic action.

Period over Period analysis:

- "INMAILS" spiked in July but lost momentum afterwards.
- "LAN" saw minor increases but remained mostly static, indicating limited growth without intervention.
- "FEED" had steady PoP growth in August and September, suggesting potential for targeted marketing.

3. Sherlock AI: An AI-Powered Approach

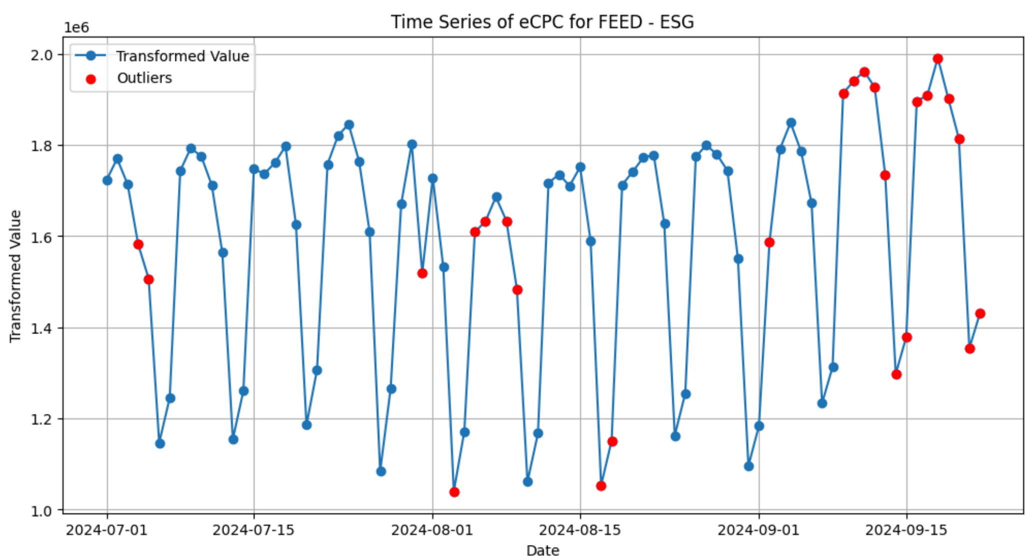
3.1 Anomaly Detection Methodology Overview

Sherlock AI combines time-series and aggregated data analysis to identify both short and long-term performance anomalies, offering a comprehensive view for decision-makers. Its two main components – outlook anomaly detection and short- & long-term anomaly detection – enable easier monitoring of budget utilization, bookings, and related metrics. Outlook anomaly detection uses statistical benchmarks to detect significant daily deviations in budget spending, while a regression-based approach predicts expected growth in bookings, flagging under- or over-performance relative to targets. Short and long-term anomaly detection goes further by capturing fluctuations over various time frames: short-term anomalies highlight unexpected shifts in the current period and long-term anomalies predict year-over-year variances. Through this hybrid approach, Sherlock delivers insights into both immediate issues and future performance trends, allowing stakeholders to make timely, strategic decisions.

3.2 Metric Categorization in Anomaly Detection

We analyzed anomalies in supply-side and demand-side metrics to identify deviations that could indicate operational inefficiencies or unusual behavior. Metrics were categorized into "supply-side," reflecting resource availability (e.g., Budgets, Inventory, eCPM), and "demand-side," highlighting engagement or performance (e.g., CTR, eCPC, Impressions). This classification was used to better understand the root causes of anomalies.

To detect outliers, we calculated the average metric value for each {product, segment, objective, metric_category_group, day_of_week} grouping and measured deviations as the absolute difference from the average. Outliers were flagged if deviations exceeded three standard deviations from the group mean. For visualization, we plotted time series data with high-deviation points marked, enabling us to link anomalies to either supply or demand-side factors. Chart below shows an example of eCPC time-series data, with anomalies that are flagged as red points. As observed from the chart, data points post 2024-09-05 have the most anomalies that need investigations.



3.3 Outlook Anomaly Detection

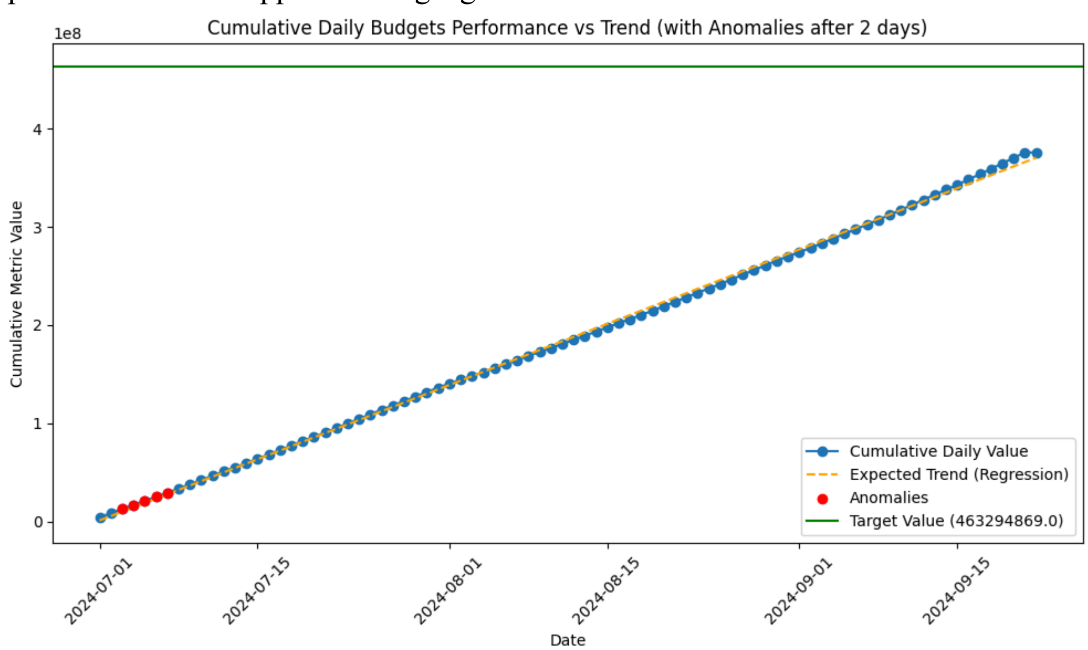
The Outlook Anomaly Detection component assesses real-time performance against business benchmarks using a time-series dataset from July to September 2024. It focuses on metrics of budget utilization, overall budget, and net bookings across different categories. By leveraging statistical and regression techniques, it identifies deviations from expected trends, enabling proactive adjustments and timely business interventions to support informed decision-making.

3.3.1 Approach One: Statistical Analysis of Budget Utilization

This statistical approach provides a focused list of anomalies by region, vertical, and product, identifying where spending patterns significantly deviate from recent trends. For example, Sherlock AI can perform daily monitoring of budget utilization, leveraging a two-sigma rule to identify statistical anomalies. Specifically, this method calculates a rolling mean and standard deviation for budget utilization within each category (e.g., region, vertical, objective) to establish a baseline. Any deviation that exceeds two standard deviations from this mean is flagged as an anomaly, indicating either an over- or under-utilization of the budget. This threshold-based detection method is particularly effective for capturing extreme fluctuations that could indicate significant spending inefficiencies.

3.3.2 Approach Two: Regression-Based Analysis of Budgets and Net Bookings

The second approach employs regression-based analysis to monitor cumulative daily values of both budgets and net bookings. In this approach, we used a time-series regression model to establish an expected growth trend based on historical cumulative budget and booking data. This model projects future values, and thus provides predictive insights into whether performance is on track to meet quarterly or monthly targets. As shown in the chart below, deviations from the trend line are measured daily, with a specified threshold applied to flag significant deviations as anomalies.



An additional enhancement in this method is the incorporation of a business outlook benchmark. By comparing cumulative performance to a pre-established target, the system identifies not only statistical outliers but also data points where budget utilization or net bookings significantly exceed or below the expected quarterly monthly or quarterly target. This dual-layered approach, considering both statistical trends and business benchmarks, enables Sherlock to contextualize anomalies, offering a nuanced view of potential over- or under-performance in relation to the organization's strategic goals.

3.4 Short- and Long-Term Anomaly Detection

In addition to Outlook Anomaly Detection, the Sherlock AI system evaluates both short- and long-term performance trends, utilizing historical comparisons to detect deviations in key metrics over different time frames. By combining statistical trends with business benchmarks, Sherlock can analyze anomalies in a way that provides a deeper understanding of how they relate to the organization's goals, helping identify areas of potential over- or under-performance.

3.4.1 Short-Term Anomaly Detection

For short-term anomaly detection, Sherlock assesses metric changes in the current period relative to the previous period. By calculating the percentage change of metrics and comparing it to the historical mean growth rate, the system identifies deviations that exceed two standard deviation as anomalies. This threshold-based approach, applied across different groupings such as product, segment, region, and vertical, allows Sherlock to distinguish between growth fluctuations from significant performance shifts that may signal underlying issues or opportunities.

For percentage-based metrics, the system employs a modified z-score method to enhance the detection of anomalies with higher sensitivity, particularly in cases where distributions may not follow a normal curve. This approach measures the deviation from the median growth rate for each grouping, flagging outliers based on a threshold of modified z-scores exceeding an absolute value of 2. This layer of analysis enables Sherlock to highlight both positive and negative anomalies in percentage growth, allowing decision-makers to understand where performance has exceeded or fallen short of expectations over the current period.

3.4.2 Long-Term Anomaly Detection

For long-term anomaly detection, Sherlock evaluates year-over-year (YoY) metric changes, providing insights into broader strategic trends. Like short-term detection, this process involves calculating percentage changes relative to the same period in the previous year, with an emphasis on detecting significant shifts that exceed two times the standard deviation of historical YoY changes. This method accounts for seasonal or cyclical variations, identifying areas where growth patterns have diverged from historical norms. By employing a modified z-score on percentage-based YoY metrics, Sherlock ensures a high level of accuracy, particularly when distributions are skewed or have a high degree of variability.

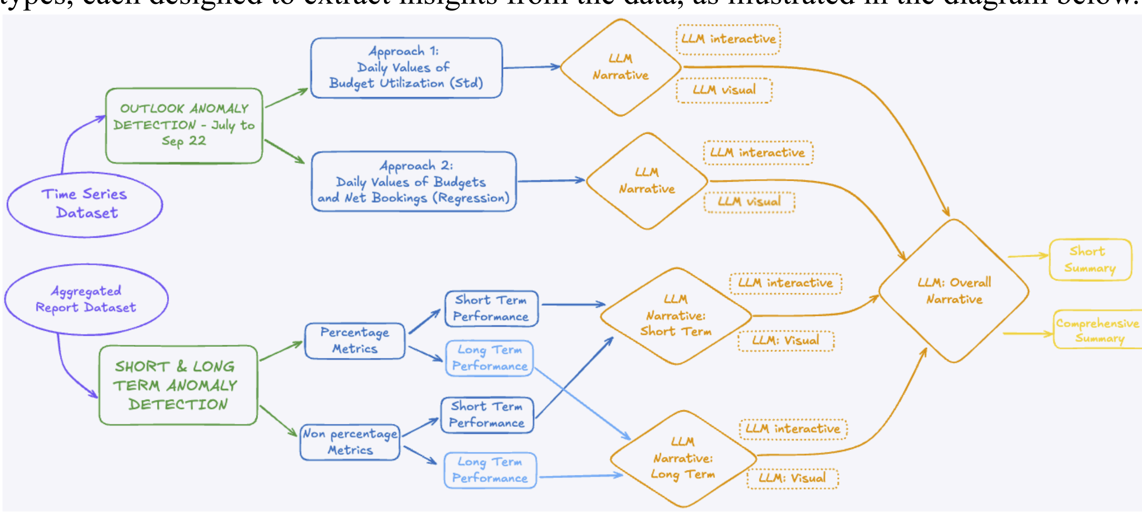
4. Automatic Narrative Generation: Multi-Agent Systems (MAS)

4.1 Background

Multi-Agent Systems (MAS) are an AI architecture where multiple autonomous agents work together, each specializing in distinct tasks within a larger, complex system. MAS are particularly effective for handling big data processing and analysis by distributing tasks across each specialized agents, which can function either collaboratively or independently. This architecture enhances scalability and efficiency, allowing MAS to handle complex applications like anomaly detection and decision support (Wooldridge, 2009). Each agent's unique capability, or "competence," enables MAS to solve specific components of a larger problem, providing focused insights and contributing to the system's overall analysis (Russell & Norvig, 2020).

MAS is especially beneficial in applications requiring real-time, layered insights, where both high-level summaries and detailed analysis are needed. Sherlock AI leverages specialized Narrative, Visual, and Interactive Agents that work collaboratively to generate contextually enriched insights and deliver tailored, interactive reports, enhancing the analytical depth and user engagement of its outputs. The multi-agent approach not only streamlines data processing but also enhances flexibility, accessibility, and depth in the insights generated, making MAS an effective framework for complex, user-driven AI systems (Jennings et al., 1998).

Sherlock AI is built as a Multi-Agent System (MAS) that uses multiple specialized large language model (LLM) agents to analyze data in a detailed and context-rich way. This setup allows users to explore insights in different formats and levels of detail, making it flexible and customizable. The MAS is made up of key agent types, each designed to extract insights from the data, as illustrated in the diagram below.



4.2 Individual Narrative Agents

Sherlock's MAS assigns a dedicated Primary Narrative Agent to each core analytical output, focusing on generating concise, accessible narratives that summarize key findings. These agents create text-based summaries from various analyses, such as time-series and aggregated report data, highlighting critical aspects of anomaly detection and performance trends. For instance, in the case of Outlook Anomaly Detection, a Narrative Agent will contextualize deviations by referencing business benchmarks.

To provide a more granular analysis, Sherlock AI employs Short- and Long-Term Narrative Agents that specialize in analyzing data across different time scales. These agents independently process data for short- (period-over-period) and long-term (year-over-year) anomalies, broken down by data type (percentage or non-percentage metrics). By isolating short-term trends and sustained long-term patterns, these agents enable Sherlock to present a balanced view of both immediate fluctuations and broader strategic changes. This division also ensures that the analysis of transient and structural trends can be understood independently, allowing for clear and targeted insights.

4.3 Optional Visual and Interactive Agents

Each Narrative Agent in Sherlock's MAS is supported by optional Visual and Interactive Agents, enhancing the system's flexibility and engagement capabilities.

- **Visual Agents:** These agents create graphical representations, including line charts, bar graphs, and trend analyses, to help users quickly understand complex data patterns. The visual summaries bring attention to key anomalies and deviations, making it easier for users to interpret Sherlock's findings briefly.
- **Interactive Agents:** Users can dive deeper into Sherlock's outputs through Interactive Agents, which enable real-time interaction with the data. Users may ask questions to clarify specific anomalies, request additional context on findings, or even ask the Interactive Agent to create custom visualizations by generating and executing code. This feature is particularly beneficial for decision-makers who need customized insights or wish to explore trends further.

4.4 Narrative Agents

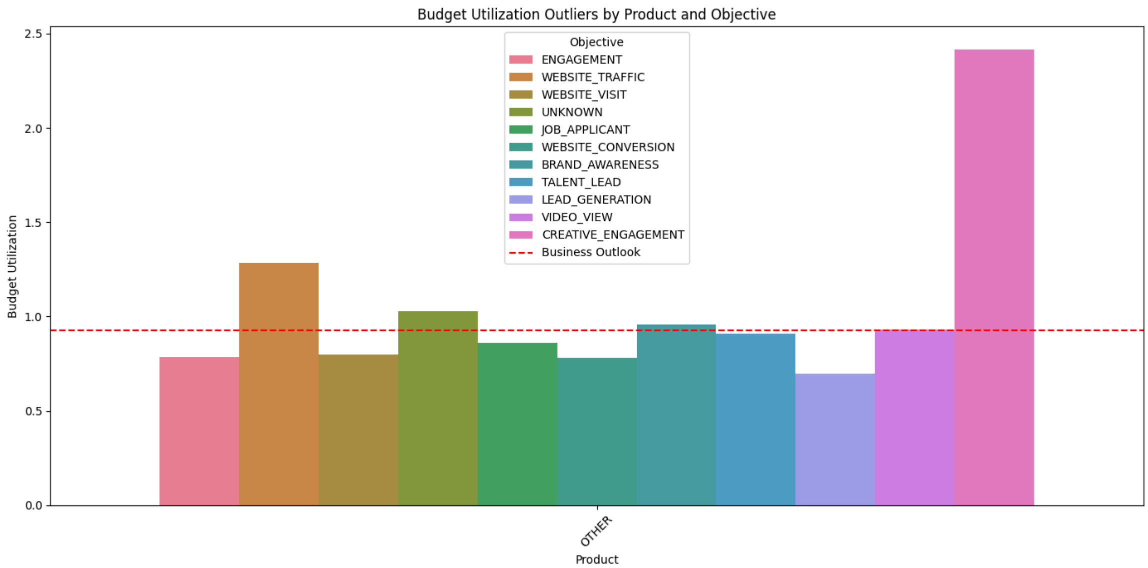
The Narrative Agents act as central aggregators, synthesizing outputs from all primary Narrative Agents into cohesive reports. This comprehensive agent generates two types of high-level summaries:

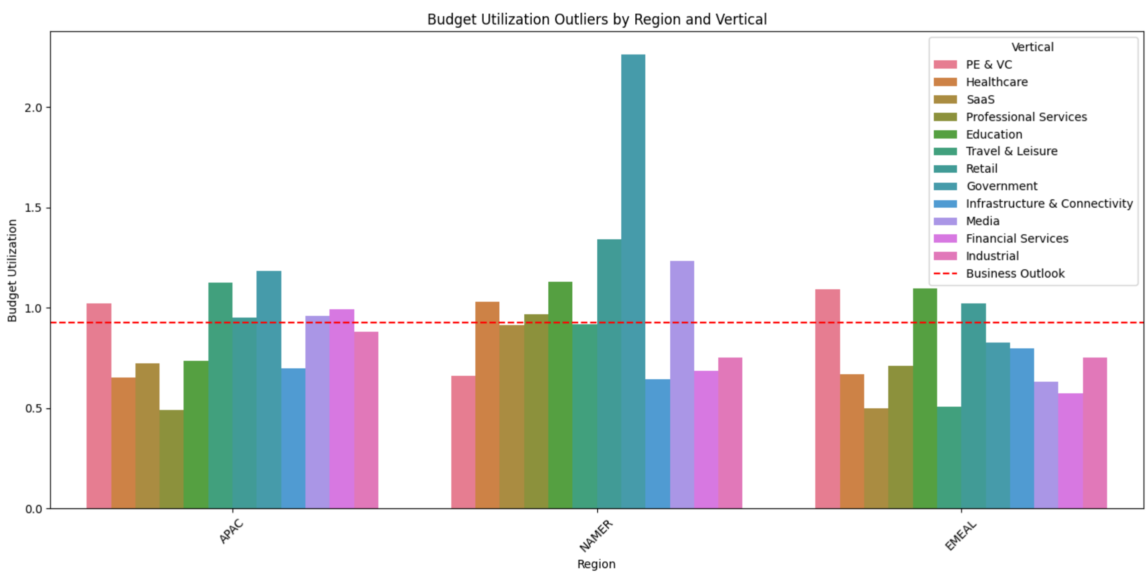
- **Short Summary**: A brief, high-level overview suitable for executives or stakeholders who require quick access to the most critical insights.
- **Comprehensive Summary:** A detailed report that encompasses both short-term and long-term trends, cross-analyzing data across multiple dimensions. This in-depth summary is ideal for strategic planning, providing stakeholders with a complete view of Sherlock's findings.

5. Output Analysis

5.1. Weekly Performance vs Outlooks

Chart below shows an example of analysis for budget utilization anomalies. It indicates key deviations across regions, verticals, and objectives compared to a benchmark of 0.9287, as indicated by the red dotted line. Key findings include under-utilization in the Travel & Leisure (0.53) and Healthcare (0.65) verticals within EMEAL, as well as in Retail (0.58) in APAC, primarily driven by low impressions and CTR. Over-utilization was observed in the Government vertical (1.64), indicating potential budget inefficiencies due to high CPC. Root causes included low engagement metrics for under-utilized budgets and high cost-per-click and CPM for over-utilized budgets. Recommendations include improving CTR and impressions for underperforming verticals, optimizing CPC strategies in over-utilized areas, and enhancing real-time monitoring for anomalies to improve budget efficiency.

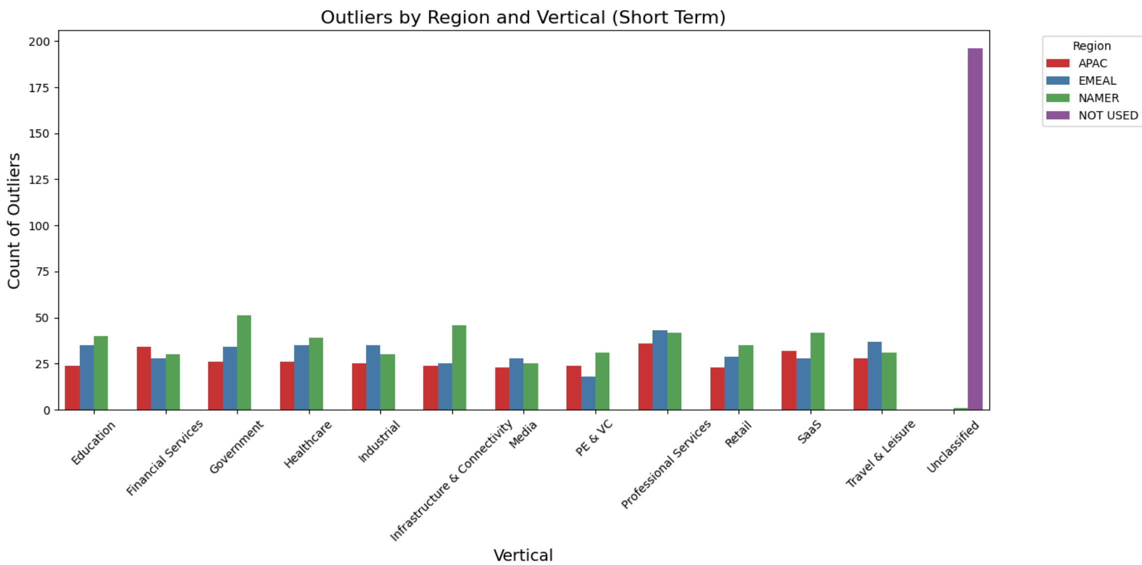




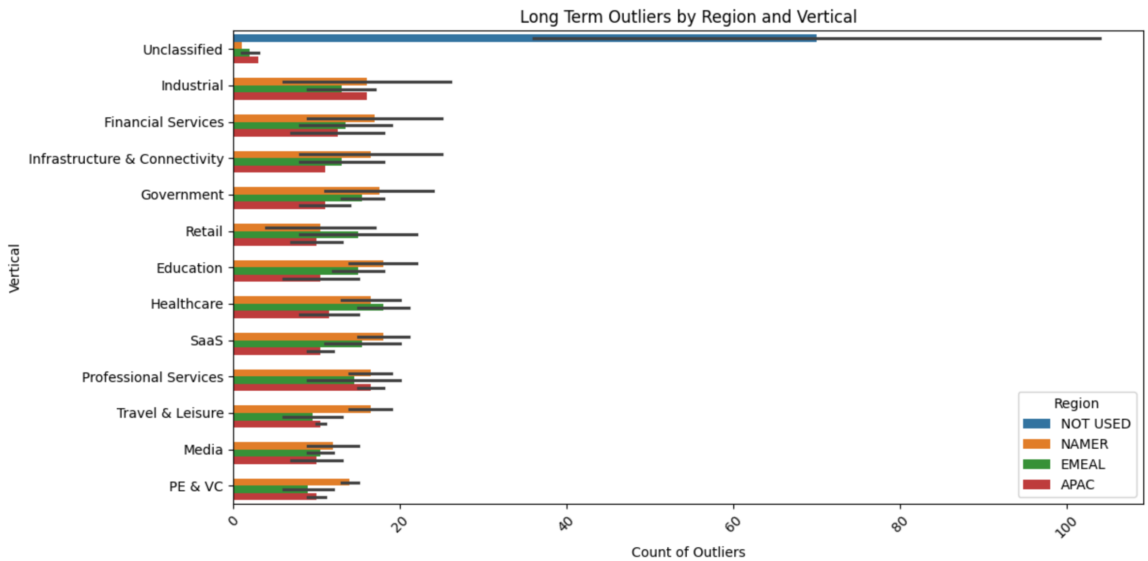
The analysis of budgets and net bookings highlighted regional and vertical-specific deviations. APAC's Healthcare vertical showed a -37.63% anomaly, driven by low CTR (-15%), while EMEAL's Travel & Leisure vertical had a +45% positive anomaly supported by high engagement and impressions (+30%). Conversely, NAMER's Education vertical faced underperformance (-50%) due to high CPC (+20%). Recommendations include targeted strategies to enhance engagement in Healthcare, reallocating budgets to successful verticals like Travel & Leisure, and optimizing ad placements in Education. Supply-side inefficiencies, such as high CPC, were noted in Education, while demand-side success in ESG segments demonstrated effective targeting and engagement strategies, supporting lead generation and brand awareness.

5.2. Weekly Short-Term and Long-Term Performance

The short-term analysis highlighted key anomalies in budget utilization across regions and verticals. Notable findings include a -37.63% decline in budget utilization for Professional Services in NAMER, driven by high eCPM (+45%), and a -25% drop for SaaS in EMEAL due to high CPC (+30%). Similarly, APAC's Financial Services showed a -18.63% decline, linked to low impressions and CTR (-20%). Positive trends were observed in NAMER's Media with a +3.92% increase, driven by improved engagement. Recommendations include optimizing bidding strategies, enhancing targeting for low CTR verticals, and maintaining successful strategies in regions showing growth.



The long-term analysis identified significant deviations from benchmarks, with notable negative anomalies in NAMER budget utilization (-2.30%), driven by high eCPM, and engagement metrics in EMEAL (-4.60%), linked to low CTR (-15%). Positive trends were observed in APAC's Website Conversion (+46.53%), driven by effective engagement strategies. Supply-side inefficiencies, such as high eCPM in FEED campaigns, and demand-side challenges, including low impressions in LAN segments, were highlighted. Recommendations include optimizing ad placements to reduce supply costs, improving CTR with targeted campaigns, and prioritizing regions with strong performance, such as APAC, for resource allocation.



5.3 Potential Root Causes

The possible root causes indicate that low impressions and CTR are significant demand-side factors contributing to underperformance across several verticals and regions, directly impacting budget utilization and net bookings. On the supply side, high eCPM and CPC values are driving inefficient budget spending, particularly in the Professional Services vertical in North America and the SaaS vertical in EMEAL, highlighting the need for cost optimization and improved engagement strategies.

5.4. Strategic Recommendations

To address inefficiencies and enhance performance, the following strategies are recommended:

-Optimize Cost Efficiency: Renegotiate supplier contracts to reduce elevated eCPM and CPC values, ensuring better cost control. Additionally, refine ad placement and bidding strategies to maximize the value derived from ad spend.

-Improve Engagement in Underperforming Areas: Focus on increasing CTR and impressions in sectors with low engagement, such as Healthcare in the APAC region and Professional Services in North America. Implement targeted campaigns tailored to these verticals to address specific audience needs and behaviors.

-Refine Campaign Effectiveness: Conduct A/B testing to evaluate and improve ad creatives, ensuring they resonate with target audiences and drive better engagement metrics.

-Leverage Proven Strategies: Apply successful engagement tactics from high-performing regions and verticals, such as the Travel & Leisure sector in EMEAL and APAC's strong Website Conversion metrics, to underperforming areas. By adapting these approaches, positive outcomes can be replicated and scaled across different regions and industries.

6. Conclusion

This research has demonstrated the effectiveness of Sherlock AI as a powerful analytical tool for assessing business performance analysis. By leveraging a multi-agent system (MAS) architecture and employing both statistical and regression-based methods, Sherlock AI can identify and contextualize anomalies across various time horizons and business dimensions. The system's ability to process and analyze both time-series and aggregated report data allows it to capture immediate fluctuations and long-term performance trends, providing a robust foundation for data-driven decision-making in LinkedIn Marketing Solutions.

Sherlock AI's integration of layered outputs, including concise weekly summaries, comprehensive reports, real-life feedback, and visualizations, enables stakeholders to engage with the data in a way that aligns with their specific needs. By focusing on both granular and high-level perspectives, Sherlock AI delivers actionable insights that help identify underutilization and capitalize on high-performing areas.

7. Future Steps and Considerations

As advancements in AI and large language models (LLMs) continue, future developments in Sherlock AI have the potential to significantly enhance its capabilities and expand its utility for addressing diverse business objectives. Improvements in predictive modeling, such as advanced time-series forecasting, could refine the accuracy of trend prediction and anomaly detection, particularly for long-term analyses. Incorporating factors like seasonality and external indicators, including market sentiment and economic data, would provide a more comprehensive contextual framework for Sherlock AI's insights. This would enable a deeper understanding of how macroeconomic trends influence key metrics such as budget utilization and engagement, factors that invariably affect businesses advertising on platforms like LinkedIn.

Furthermore, the extension of real-time monitoring functionalities to include automated alerts could facilitate faster issue resolution by providing immediate feedback on performance anomalies. Enhancements such as natural language explanations, scenario analysis, and actionable recommendations would augment the user-focused features of Interactive Agents, improving both usability and user engagement. Lastly, expanding the application of Sherlock AI's methodology to additional organizational domains, such as sales, operations, and customer support, could foster a more holistic and integrated perspective on overall business performance, driving strategic decision-making and operational alignment.

References

Davis, S., Green, E., & Patel, L. (2022). *Integrating AI for Automated Business Insights: Opportunities and Challenges*. Journal of Business Intelligence Research, 15(4), 99-116.

Davis, R., Brown, P., & Walker, H. (2022). *Analyzing Demand and Supply Factors in Digital Advertising*. Journal of Business Analytics, 24(1), 101-115.

Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). *"A Roadmap of Agent Research and Development."* Autonomous Agents and Multi-Agent Systems, 1(1), 7-38.

Johnson, T., & Lee, M. (2021). *The Role of Demand-Side and Supply-Side Factors in Business Performance Assessments*. Journal of Business Analytics, 19(3), 112-127.

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach. Pearson. Wooldridge, M. (2009).* An Introduction to MultiAgent Systems. Wiley.

Smith, J., & Johnson, M. (2021). Business Analytics in Professional Networking Platforms. Journal of Strategic Finance, 29(3), 78-93.

Smith, J., Brown, P., & Walker, R. (2020). *Challenges in Scaling FP&A Processes in Large Tech Firms*. Journal of Strategic Finance, 28(2), 45-58.