

# Analytics based Seamless Automation of Key Performance Indicator Setting Process

Lavanya Naik

M. Tech in Computer Network Engineering  
Dept of CSE, RV College of Engineering, Bangalore

Dr Nagaraja G S

Professor & Assoc. Dean  
Dept of CSE, RV College of Engineering, Bangalore

**Abstract**—Business Relationship with the stakeholders is important for any organization to grow in their business. This is evaluated based on Key Performance Indicators (KPIs) that is set in the beginning of the financial year. Hence it is important to track the KPIs in regular intervals so that organization do not end up in trouble last moment. Manual processes for tracking and analyzing KPIs are often time-consuming and prone to errors, which can lead to inaccurate data and poor decision-making. The proposed approach is the automating the process of calculating actuals and forecast for the current month. This solution is achieved using ETL process of relevant data and writing stored procedures in PostgreSQL and thereby coming up with automated view for KPI setting process. The major factor is to finalize the models in order to calculate the forecast values for the remaining day of the month. Forecasts provide valuable insights into future trends, allowing individuals and organizations to make informed and strategic decisions. This in turn gives streamlined solution proving efficient solution both in terms of reduction in error and time.

**Keywords**— Database, Key Performance Indicator, Machine Learning, Structured Query Language, Forecast, Power Query.

## I. INTRODUCTION

In today's data-driven world, businesses are constantly looking for ways to optimize their performance and achieve their goals. One way to do this is through the use of key performance indicators (KPIs) as described in Fig 1, which are metrics that are used to measure and track progress towards specific business objectives. However, manually tracking and analysing KPIs can be a time-consuming and error-prone process, which is why many businesses are turning to analytics-based seamless automation to streamline the process. Analytics-based seamless automation of KPIs involves a multi-step process that includes identifying relevant KPIs, defining metrics, gathering data, analysing it for insights, automating the KPIs through reports or dashboards, monitoring them, and refining the process as needed. This approach there by compares actual values with that of targets set for that month. For developing this system, it is important to have the relevant data in database By understanding potential outcomes, businesses can allocate resources, plan production, manage inventory, and set realistic goals.

The building blocks of analytics-based seamless automation of key performance indicators (KPIs) are the essential elements that work together to create an effective and efficient process. These building blocks include:

**Data sources:** Data sources are the building blocks of the analytics-based seamless automation of KPIs. They include all the systems and tools used to gather data, such as website analytics, CRM systems, financial software, and other relevant data sources. The data collected from these sources should be accurate, up-to-date, and consistent.

**Metrics and KPIs:** Metrics and KPIs are the measures that are used to track and analyze business performance. They ought to be in sync with the business objectives. The selection of metrics and KPIs should be based on the business needs and should be continually reviewed and updated.



Fig 1: Key Performance Indicator

**Data visualization and reporting:** Data visualization and reporting tools are used to turn data into actionable insights. They provide a clear and concise picture of business performance and help identify trends and areas for improvement. Data visualization and reporting tools should be customizable to meet the specific needs of the business.

**Automation tools:** Automation tools are used to streamline the process of tracking and analyzing KPIs. They include dashboards, reports, and alerts that provide real-time updates on business performance. Automation tools should be customized to meet the business needs and should be easy to use.

**Data analysis and refinement:** Data analysis and refinement are ongoing processes that involve identifying trends and patterns in business performance and making adjustments as needed. This requires ongoing data analysis to ensure that the

KPIs are relevant and accurate, and that they provide valuable insights that can inform decision-making.

Overall, these building blocks work together to create an effective and efficient analytics-based seamless automation of KPIs. By leveraging these building blocks, businesses can optimize their performance, achieve their goals, and gain a competitive advantage in their industry.



Fig 2: Steps followed in Business Intelligence

The problem statement is businesses' challenge in swiftly and accurately tracking and analyzing performance due to manual KPI processes, disconnected data systems, and inflexible approaches. This impedes informed decisions, misses optimization opportunities, and slows market response, leading to a competitive lag. Analytics-based KPI automation aims to streamline processes, utilize advanced tools for real-time insights, and enhance overall performance and competitiveness. This can be achieved by using following all steps mentioned in Fig 2.

The objectives of analytics-driven KPI automation are: enhancing data accuracy and consistency, increasing efficiency and real-time insights, and aligning KPIs with business goals for informed decision-making and improved collaboration, ultimately optimizing performance and gaining a competitive edge.

## II. LITERATURE SURVEY

The use of key performance indicators (KPIs) is crucial for evaluating manufacturing plant performance [1]. Fuzzy cognitive maps (FCMs) have been used to analyze and model complex systems, including manufacturing plants. Studies have shown that FCMs can accurately model the relationships between different KPIs, identify the most important KPIs and their interdependencies, and allow for targeted improvement strategies. Therefore, FCMs can help manufacturing plants improve their performance and achieve their strategic goals.

Business intelligence (BI) is a critical tool for data-driven decision-making in organizations.[2] With the increasing amount of data generated by businesses, machine learning (ML) and data mining (DM) Approaches have gained increased prominence in BI. Studies have shown that ML and DM can be used to develop BI systems for predicting customer behaviour, analyzing customer feedback, and improving marketing strategies. Additionally, techniques have been used to analyse financial data, customer segmentation, and fraud detection.

Overall, the use of ML and DM in BI has the ability to provide insights and improve business performance.

The SMART KPI Management System Framework [3] is a tool that aids organizations in meaningful KPI setting and management. It emphasizes specific, measurable, achievable, relevant, and time-bound KPIs. Research shows its effectiveness in improving KPI management, aligning with objectives, and benefiting industries like healthcare. Overall, it's valuable for enhancing KPI management processes.

The KPI for Managing and Controlling a Demand Response System [4] is a testing framework that helps end users manage and control their demand response systems. The framework includes several KPIs that are used to assess system performance and pinpoint areas for enhancement. One study evaluated a demand response system for a university campus, using the KPI framework to evaluate the performance of the system and identify areas for improvement. Another study proposed a KPI framework for demand response systems in smart grids, including KPIs related to system reliability, efficiency, and customer satisfaction. Overall, the KPI framework is a useful tool for evaluating and improving the performance of demand response systems in various industries.

"A Method for Modelling and Organizing ETL Processes" [5] is a research paper that presents a framework for designing and managing ETL (extract, transform, load) processes. The framework includes a graphical notation for representing ETL processes and a set of guidelines for organizing and managing ETL processes. One of the main contributions of the framework is its ability to handle complex ETL processes that involve multiple data sources, transformations, and data quality checks. The paper also includes a case study that demonstrates the use of the framework in the context of a real-world ETL project. Overall, the framework provides a useful tool for data professionals who are involved in designing and managing ETL processes.

"Design of a Shared Memory Mechanism for Efficient Parallel Processing in PostgreSQL" [6] is a research paper that proposes a new shared memory mechanism for efficient parallel processing in PostgreSQL, a popular open-source relational database management system. The paper discusses the limitations of the existing shared memory mechanism in PostgreSQL and introduces a new design that aims to improve parallel processing performance. The paper offers experimental findings that demonstrate the performance benefits of the new shared memory mechanism compared to the existing mechanism in PostgreSQL. Overall, the paper offers valuable insights into the design of efficient parallel processing mechanisms in database management

"The Challenges of Extract, Transform and Loading (ETL) System [7] Implementation For Near Real-Time Environment" is a research paper that discusses the challenges of implementing ETL systems for near real-time environments, where data needs to be processed and analyzed in a timely manner. The paper highlights the importance of ETL systems in business intelligence and analytics and identifies the challenges in designing and implementing these systems for near real-time environments. The challenges discussed in the paper include data volume, data quality, data velocity, and data variety. The paper also discusses the key considerations in designing ETL systems for near real-time environments, such as the use of parallel processing, data streaming, and data

validation techniques. Overall, the paper provides insights into the complexities of implementing ETL systems in near real-time environments and offers recommendations for addressing these challenges.

In this study [8], the authors aimed to design and improve a KPI system for materials management in a power group enterprise. They conducted a comprehensive review of the literature on KPI systems and materials management and identified several key performance indicators relevant to the enterprise. The authors then developed a KPI system that incorporated these indicators and tested the system using data from the enterprise. They found that the system was effective in providing meaningful insights into materials management performance and could be used to drive continuous improvement in the enterprise's operations. The authors concluded that the KPI system they developed could serve as a useful model for other organizations seeking to improve their materials management practices.

This literature survey [9] explores the ETL function realization of a data warehouse system using the SSIS platform. The authors examine the primary constituents of the ETL process and deliberate on the significance of data quality and integration. They additionally furnish an intricate account of employing the SSIS platform to create and execute a proficient ETL process. The paper underscores the advantages of SSIS for ETL, encompassing its capability to manage substantial data volumes and its versatility with different data origins and endpoints. In conclusion, the authors assert that SSIS stands as a dependable and effective platform for integrating ETL processes into a data warehousing system.

In [10] the paper titled "KPI Anomaly Detection Based on LSTM with Phase Space," the authors propose a method for detecting anomalies in key performance indicators (KPIs) using long short term memory (LSTM) neural networks and phase space analysis. The method involves transforming KPI data into a phase space representation, which allows for the detection of anomalies that may not be visible in the original data. The authors then apply an LSTM neural network to the phase space data to detect anomalies in real-time. The proposed method is tested on a real-world dataset, and the results demonstrate its effectiveness in detecting anomalies in KPIs.

### III. DESIGN

Project design is a crucial phase in any undertaking, laying the foundation for successful implementation and achievement of goals.

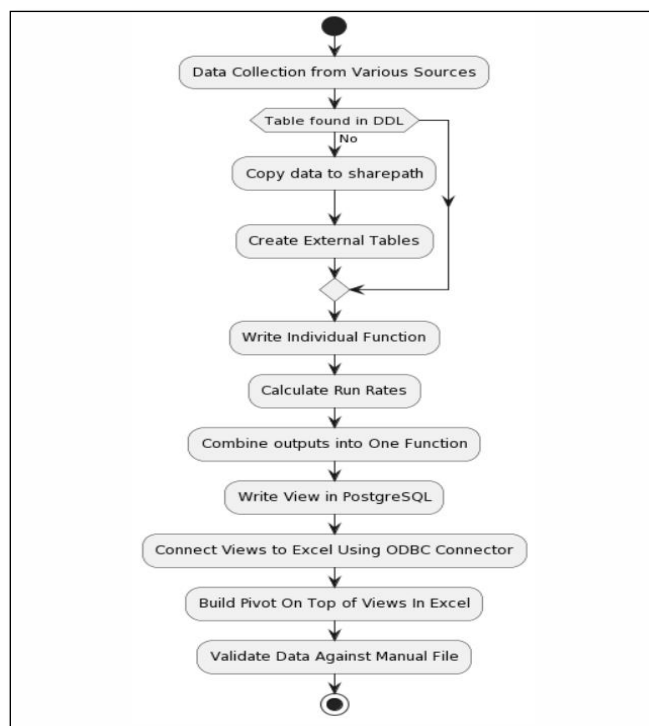


Fig 4: Complete flow of Project

It involves outlining the project's scope, objectives, activities, and resources required to bring the idea to fruition. The complete flow of implementing the given problem statement is shown in fig 3. First and foremost, step is to collect the data from various sources which is crucial for automation. Once data is collected, check on the availability of those table. If not available as it is in database, then create an external table to query the path of flat file having the data. Creating external tables allows you to access and manage data stored outside the database without moving it internally. It simplifies data management by providing a virtual view of external data. Next step is to proceed with individual functions for all the model, to evaluate run rates, etc. Once the functions are completely running, create view on top of it. Creating a view on top of a function can encapsulate the functionality of the function and provide a simplified way to query its results. This abstraction helps hide the complexity of the underlying function, making it easier to use and maintain in your queries and applications. This view is connected to excel via ODBC connector and using power query get the data in excel. Next step is to build pivot and have the data in required format. Last step is to do the data validation and monitor the forecast values at regular intervals.

### IV. METHODOLOGY

The methodology of analytics-based seamless automation of KPI involves the following steps:

Identifying which all attributes should be considered for calculating the forecast is major step. Once that is completed, collected the relevant data required for writing functions. Functions are stored procedures in PostgreSQL. This is written in DBeaver which offers several useful features that make the process easier and more efficient.

SQL Server Integration Services (SSIS) is a component of Microsoft's SQL Server database management system. It is a powerful data integration and ETL (Extract, Transform, Load) tool used to extract data from various sources, transform it into a desired format, and load it into a target destination, typically a SQL Server database.

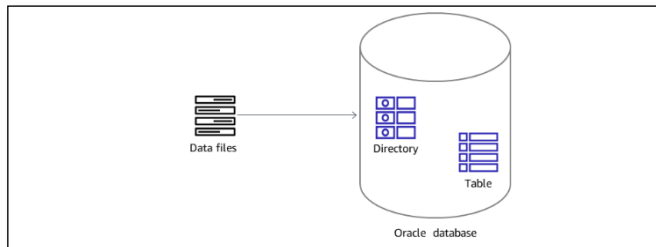


Fig 5: External table creation

If the relevant data is not found in database, it is important to create external tables that extracts data from csv file in local or shared path and gives in database. This is explained in figure 5 Once all the data is ready functions are built for calculating run rates using the actuals in terms of purchases of every priority group. Based on this data, forecast values are evaluated for remaining days of ongoing month.

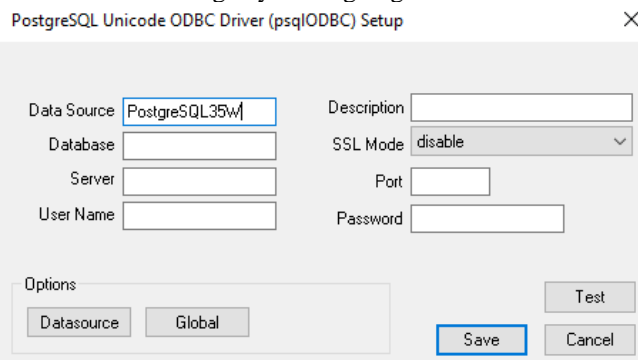


Fig 6: ODBC connector to connect to PostgreSQL function

On top this function views are built which is then connected to excel using ODBC connector and power query This views are being rearranged or formatted using pivot table and given to stakeholders to check on how much target is met. This is shown in Fig 6. Finally by completing all this steps and doing the data validation against manual file, analytics based automation of KPI setting process gets completed

### V. RESULTS

The proposed system offers numerous benefits that can significantly enhance organizational performance and decision-making. Model validation is done using MAPE as the metrics. Mean Absolute Percentage Error is metrics that is useful for analyzing time series data and identify error in percentage form between actual and forecast values. This metrics has helped in selecting the model that has least MAPE and thereby giving accurate results. Fig 7 shows how MAPE is used for selecting the model out of 4 models for every month.

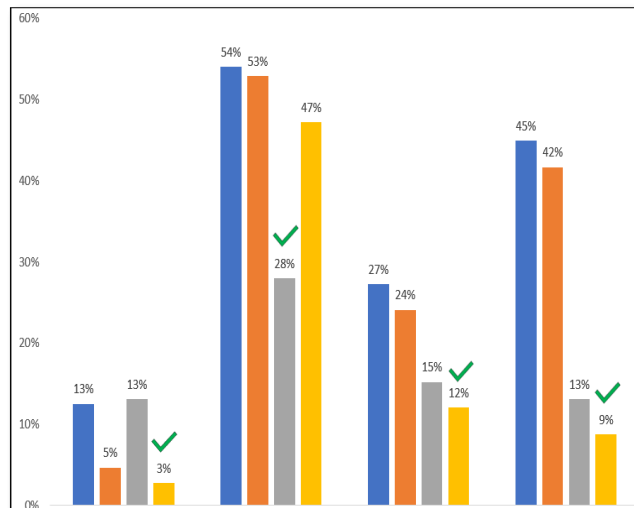


Fig 7: MAPE comparison

Proposed output would be Autonomous BI system capable of extracting the relevant data from various sources and giving out the KPI outlook number in the form of dashboard. By leveraging data analytics and automation technologies, organizations can streamline and optimize the KPI setting process in the following ways:

Automation eliminates manual interventions and reduces the potential for human errors in defining and tracking KPIs. Leveraging analytics allows organizations to identify relevant and meaningful KPIs based on historical data, industry benchmarks, and real-time insights.

### XI CONCLUSION

Major significance is Streamlines the monitoring process by automating data collection, calculation, and comparison of metrics against target values. Automated data integration system that collects real-time data from relevant sources that contain the necessary information. Provides real-time visibility and proactive management of business performance through automated monitoring and alerts. Eliminates manual formulation, input, Excel formulas, and manual errors in the process of monitoring actuals meeting target values. The proposed approach does automate the process of calculating actuals and forecast for the current month of the quarter.

By leveraging data analytics and automation technologies, organizations can streamline and optimize the KPI setting process in the following ways: Automation eliminates manual interventions and reduces the potential for human errors in defining and tracking KPIs. Finally, this approach allows organizations to identify relevant and meaningful KPIs based on historical data, industry benchmarks, and real-time insights.

Future enhancement of this automation can be done by building a dashboard in data visualization tools like Power BI, Tableau. This will give fair understanding to stakeholders with no technical background and there by not only avoids any loss but also will keep the relationship with stakeholders intact.

### REFERENCES

[1] H. J. Moon, S. H. Lee, S. J. Yoo, E. J. Yu and C. S. Leem, "A KPI-Based Performance Assessment Framework for Korean e-Government," 2008

- Second International Conference on Future Generation Communication and Networking Symposia, Hinan, China, 2018, pp. 71-76, doi: 10.1109/FGCNS.2018.96..
- [2] A. Ferreira and I. Pedrosa, "Data-driven management using Business Analytics: the case study of data sets for new business in tourism," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), Madrid, Spain, 2022, pp. 1-5, doi: 10.23919/CISTI54924.2022.9819995.
- [3] A. Baykasoglu, Z. N. Atalay and İ. Gölcük, "Analysis of key performance indicators in a manufacturing plant via fuzzy cognitive maps," 2019 Innovations in Intelligent Systems and Applications Conference (ASYU), Izmir, Turkey, 2019, pp. 1-4, doi: 10.1109/ASYU48272.2019.8946387.
- [4] A. P. Pereira, B. P. Cardoso and R. M. S. Laureano, "Business intelligence: Performance and sustainability measures in an ETL process," 2018 13th Iberian Conference on Information Systems and Technologies (CISTI), Caceres, Spain, 2018, pp. 1-7, doi: 10.23919/CISTI.2018.8399473.
- [5] J. Sreemathy, R. Brindha, M. Selva Nagalakshmi, N. Suvkha, N. Karthick Ragul and M. Praveennandha, "Overview of ETL Tools and Talend-Data Integration," 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2021, pp. 1650-1654, doi: 10.1109/ICACCS51430.2021.9441984.
- [6] A. Kabiri and D. Chiadmi, "A method for modelling and organazing ETL processes," Second International Conference on the Innovative Computing Technology (INTECH 2018), Casablanca, Morocco, 2018, pp. 138-143, doi: 10.1109/INTECH.2012.6457795.
- [7] R. Sharma and P. Srinath, "Business Intelligence using Machine Learning and Data Mining techniques - An analysis," 2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2018, pp. 1473-1478, doi: 10.1109/ICECA.2018.8474847.
- [8] Z. Ishak, S. L. Fong and S. C. Shin, "SMART KPI Management System Framework," 2019 IEEE 9th International Conference on System Engineering and Technology (ICSET), Shah Alam, Malaysia, 2019, pp. 172-177, doi: 10.1109/ICSEngT.2019.8906478.
- [9] P. Triantafillou, "Data-Less Big Data Analytics (Towards Intelligent Data Analytics Systems)," 2018 IEEE 34th International Conference on Data Engineering (ICDE), Paris, France, 2018, pp. 1666-1667, doi: 10.1109/ICDE.2018.00205.
- [10] B. Pan, G. Zhang and X. Qin, "Design and realization of an ETL method in business intelligence project," 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), Chengdu, China, 2018, pp. 275-279, doi: 10.1109/ICCCBDA.2018.8386526.
- [11] A. Sabtu et al., "The challenges of Extract, Transform and Loading (ETL) system implementation for near real-time environment," 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), Langkawi, Malaysia, 2017, pp. 1-5, doi: 10.1109/ICRIIS.2017.8002467.
- [12] L. Hu, H. Su, R. Cui, H. Han and L. Sun, "KPI Anomaly Detection Based on LSTM with Phase Space," 2022 IEEE 5th International Conference on Big Data and Artificial Intelligence (BDAl), Fuzhou, China, 2022, pp. 130-135, doi: 10.1109/BDAl56143.2022.9862720.
- [13] S. Subbalakshmi and C. Prabhu, "Protagonist of Big Data and Predictive Analytics using data analytics," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 2018, pp. 276-279, doi: 10.1109/CTEMS.2018.8769141.
- [14] R. Wijaya and B. Pudjoatmodjo, "An overview and implementation of extractiontransformation-loading (ETL) process in data warehouse (Case study: Department of agriculture)," 2015 3rd International Conference on Information and Communication Technology (ICoICT), Nusa Dua, Bali, Indonesia, 2015, pp. 70-74, doi:10.1109/ICoICT.2015.7231399.
- [15] L. J. Tozin and A. C. S. Amaro, "Business intelligence on Supply Chain Management," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), Madrid, Spain, 2022, pp. 1-4, doi: 10.23919/CISTI54924.2022.9820567.s