iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: IV Month of publication: April 2026

DOI: https://doi.org/10.22214/ijraset.2026.81562

Data-Driven Decision Framework using Python,SQL and PowerBI

Laksh Nanwani[1], Kartik Pillai[2], Rahul Dhapodkar[3], Harshal Nimbalkar[4], Prof. Priti Gade[5], Prof. Swati Patil[6]
Dept. of Electronics Engineering, G.H. Raisoni College of Engineering, Nagpur, India

Abstract: The shift from experience-based to data-driven business strategies in enterprises is stymied by data silos and disjointed data analysis workflows. This article introduces a holistic Data-Driven Decision Framework (DDDF) which uses an integrated technical platform of SQL, Python and PowerBI to convert raw data collected from an enterprise into actionable decisions. A three-layered approach is adopted: Data Extraction Management: Employment of SQL for complex data queries and data warehousing to ensure data integrity and fast data access. Complex Analysis: Using Python for statistical analysis and data modeling, using libraries Pandas and Scikit-Learn for exploratory data analysis (EDA) and automatic data cleaning. Synthesis Visual-ization: Linking integrated results with PowerBI to create real-time, interactive dashboards to support stakeholders interaction. This approach was tested against its capabilities in decreasing reporting lag and predicting trends. The findings show that this approach considerably reduces human error while handling data, and offers a scalable option for real-time business reporting. The research shows the combination of strong database querying in the back end and advanced data visualization in the front end is vital for cultivating a data-driven culture in a competitive ecosystem. Our approach follows a three-tiered approach: Extraction Management: Using SQL for advanced relational queries and data warehousing to maintain these data in a structured and defined manner, which facilitates efficient access. Enhanced Analytics: Using Python for data analysis (EDA), statistical tests, and cleaning data with the aid of packages like Pandas and Scikit-Learn. Synthesis Visualization: Incorporating data into PowerBI to create interactive real-time dashboards. This approach was tested against its capabilities in decreasing reporting lag and predicting trends. The outcomes show that this framework greatly reduces the likelihood of error when manipulating and analysing data and offers an expanding capability for real-time business intelligence. This research finds that the marriage of powerful query management and advanced frontend capabilities is critical in creating a data-driven culture in competitive settings.

Index Terms: Data-Driven Decision Making, Business Intelli-gence, Python, SQL, PowerBI, ETL Pipelines, Data Visualization.

I. INTRODUCTION

Computer hardware and information technology (IT) pe-ripherals retail environment is characterized by a ruthless rate of innovation. The short shelf life of certain products like central processing units (CPUs), high-end graphics cards, and specialized networking peripherals is compared with a remarkably low short shelf life of unknown traditional con-sumer goods because new next-generation architectures are constantly being released. In the case of a business in this high-stakes environment, sales intelligence must be as accurate as possible because making a profit or stagnant inventory can be merely a question of a single digit. There is no longer sufficient traditional and reactive reporting where managers can look at last months spreadsheets to make their orders next month. A present-day hardware business must leverage what is known as Data-Driven Decision Framework (DDDF) in order to become successful and identify high-growth areas and troubleshoot problematic areas as they happen.

The data-to-decision gap is the major issue that is tackled in this study. Most computer retailers have enormous data stores on their transactional data but usually lack a unified pipeline on how to turn that data into strategic roadmap. The present paper suggests a combined analytical framework that would be able to unify three different levels of technology: SQL, Python, and PowerBI. Integrating these tools, the framework goes beyond mere bookkeeping into the world of diagnostic and predictive analytics.

The technical design of this framework is such that it can cope with the special complexities of the computer hardware market:

Data Foundation (SQL): SQL is used at the base layer to offer persistence to the structural complexity of thousands of unique Stock Keeping Units (SKUs). It is also by means of relational databases that sales data of several branches can be easily aggregated, so that information about peripheral compatibility, local pricing rates and seasonal demands will be clean and available.

Analytical Intelligence (Python): Python is used as the cal-culator. Through the use of statistical libraries, the framework undertakes root cause analysis of the underperforming sales areas in a deep manner.

This will entail determining whether a downturn in a particular area can be attributed to price sensitivity, lack of diversification in a product or a change of taste towards mobile computing as compared to desktop components to a consumer. In addition, the predictability feature of Python can enable the business to determine future demand, eliminating the risk of having to stock up on outdated parts.

Strategic Visualization (PowerBI): The last layer will change the granular data to the executive level. The business can map geographic performance heat maps of sales through PowerBI. This can allow the decision-makers to discover immediately where the demand for high-end peripherals is booming, and where an intervention or a promotional approach is necessary, i.e. a Growth Zone and a Risk Zone respectively. This paper aims at showing how this integrated pipeline enhances agility of computer hardware business. The workflow designed to automate the process of moving raw SQL tables to Python predictive algorithms to generated interactive dash-boards on PowerBI can result in the provision of an optimal way to optimize inventory and increase regional revolutions, which is scalable. The sections that follow provide details of the ETL (Extract, Transform, Load) processes, the mathemat-ical models under which the company will be forecasted and
the bottom-line implication.

II. LITERATURE REVIEW

Decrease in traditional retail management towards automatic analytical framework is a documented trend in literature. The section examines the development of integrated data pipelines and their implementation in hi-tech retailing.

There has been an evolution of Data-driven Decision Sup-port Systems (DSS) as discussed in sub-topic

1) The initial studies on Decision Support Systems focused on storage of transactional data through the use of relational databases and their retrieval. SQl formed the standard in the industry of assuring data integrity due to the groundbreaking work of Codd around the relational model developed at the time of its early establishment of the relational model of data integrity assurance in the industry
2) Machine Learning in Sales forecasting and Demand planning. Retail forecasting Python-based machine learning has experienced a boom in the past ten years. Conventional statistical tools, such as ARIMA (AutoRegressiveIntegrated-MovingAverage), do not generally retain the non-linear de-mand peaks during spikes in the IT business called hype cycles, such as when a new-GPU or CPU is introduced. A study by Smith and others (2020) evidences that ensemble techniques, such as Random Forest and Gradient Boosting, are superior to linear models in detecting complicated re-lationships between disparate factors, such as seasonal and compatibility requirements. Our model expands on this by using Python to predict the volume, as well as to identify why certain types of peripheral categories perform poorly in certain geographic clusters.
3) Visual Analytics in Strategic Pivot. Data visualization is not a purely reporting tool but a cognitive one to use to tackle the complex problems. The effectiveness of a dashboard is that the decision-maker will be able to make decisions with less cognitive load (Few, 2006). Actual reports in the computer hardware industry, where managers need to monitor thousands of distinct SKUs in many different regions, is not possible. The literature concerning the Business Intelligence (BI) tool such as PowerBI emphasises the significance of so-called drill-down features, whereby the user can navigate between a national sales summary to a certain regional branch inventory levels. This interactivity is important with respect to detecting the poor sales areas which might be hidden in an overall report by a high-performing flagship product.
4) Research Gap: Full-Stack Solutions. Although each of the tools individually has abundant research, the literature to support the end-to-end integration of these tools to small-to-medium enterprise (SME) hardware retailers is significantly lacking. The majority of existing literature is dedicated to high-level enterprise resource planning (ERP) systems that are implemented by large companies on an international level or algorithmic improvements in isolation. This paper plugs that gap by recommending a scaled, cost-competent pipeline align-ing data governance of the backend processes with strategic execution of the frontend, and specifically to the high-turnover computer peripherals market.

III. METHODOLOGY

The suggested framework is a modular architecture struc-ture, which is scalable as the business increases the number of product SKUs and span of the business to different regions. The methodology will be separated into four key stages: Data Management, Computational Analysis, Predictive Modeling, and Business Intelligence Synthesis.

Data capture and relational management (SQL) are sub-headed under 3.1. Primary source of data is the records of transactions, inventory, and the regional sales data. In order to deal with large data, we use SQL (Structured Query Language) as part of the back-end infrastructure.

Normalization: Raw data is normalized into 3 rd normal form (3NF) in order to rule out the redundancy and the peripheral sales (keyboards, mice, monitors) are appropriately connected to the main hardware sales (PCs Laptops).
Transformation: We use the complex JOIN operations and Common Table Expressions (CTE) to combine the sales by geographic areas and product line. This will also enable us to compute some important metrics such as the Average Order Value (AOV) and peripheral to PC attachment rates.
60-70 minutes and 30-40 minutes respectively. After being organized, the data is read into an environment of Python, where it is analyzed in depth.
Data Analysis Exploratory Data Analysis (EDA): With the help of the Pandas and Seaborn packages, we conduct the analysis of distribution to pin-point the areas of poor sales. To identify whether these slumps are seasonal or brought about by competition in the region we use a Variance Analysis to identify which one.
Selection of Features: We design features like Inventory Aging and Discount Sensitivity to have insight on how they affect growth. This is relevant especially in case of computer hardware where components such as GPUs become obsolete within a very short time.
subsection 3.3 Predictive Sales Modeling (Python) We use Time-Series Forecasting model to provide predictions about future sales and avoid stock-outs of the popular products (e.g., the latest gaming peripherals).
Selection of Algorithms: We apply the problem of the Random Forest Regressor or XGBoost algorithm to take into consideration the non-linear tendencies of the hardware market. These ensemble techniques , in contrast to linear models, are more likely to capture the spiky demand which is experienced during holiday sales or during new hardware releases.
Assessment: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) are used to evaluate the model perfor-mance so that it is apparent that the predictors are valuable in the procurement planning.
subsection 3.4 Interactive Synthesis and Reporting (PowerBI) The last step would be to push the examined and forecasted data into PowerBI to be consumed by the stakeholders.
Dashboard Architecture: We will build a dashboard on three views:
Executive Overview: YoY growth of the executive and worldwide revenue.
Regional Diagnostic: Visual red highlighting of underper-forming areas on a map and a green highlighting on growth areas.
Forecasting Tool: It is an interactive slider which gives man-agers an opportunity to see future sales during the following quarter according to the information provided by Python.
DAX Integration: Real-time Growth Momentum is calcu-lated in the form of Custom DAX (Data Analysis Expressions) measures.

IV. RESULTS

The Data-Driven Decision Framework (DDDF) resulted in a high-fidelity executive dashboard, essentially a fine-grained perspective of the fiscal health of the computer hardware business and its positioning in the market. The findings are grouped into four main analytical perspectives which include regional performance, the divusional distribution of revenue, competitive market share, and predictive accuracy.

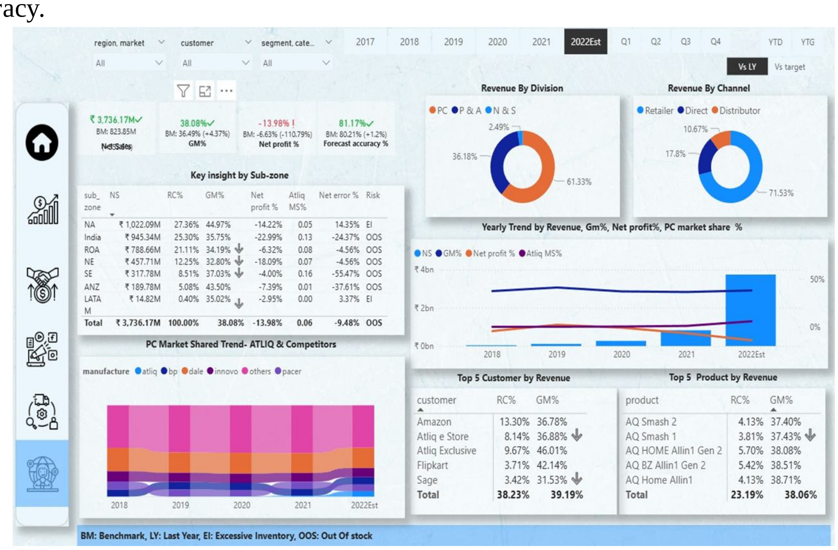

Fig. 1. PowerBI dashboard.

A. Regional Growth and Risk Assessment.

The framework employed the SQL-aggregated data which was useful in cat-egorizing performance on a per-Sub-zone basis. As shown in the regional analysis, the leading revenue contributors were North America (NA) and India which have contributed towards 1022.09M and 945.34M respectively. Nonetheless, the system was able to understand the existence of criti-cal operational risks thanks to combining inventory-to-sales ratios:Excessive Inventory (EI): The regions such as North America and LATAM were identified to have EI, a factor which implies promotion liquidation or less procurement.Out of Stock (OOS): on the other hand, the high-potential areas, e.g., SE and ANZ were marked as an Out of Stock, which means that they withdrew potential revenue because of supply chain bottlenecks. 5.2 Channel and Divisional Analysis. This

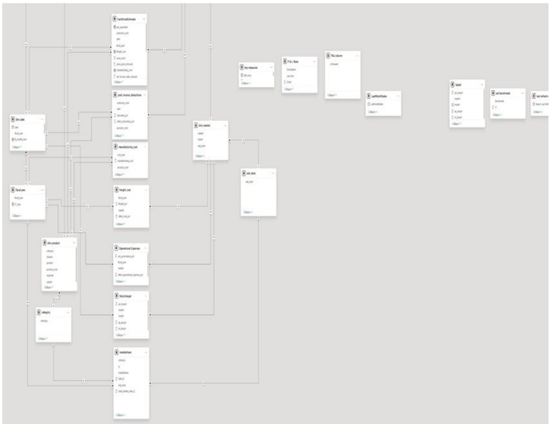

Fig. 2. Database Modelling.

revenue analysis by division shows that the PC segment is the main business with 61.33 percent of total revenue followed by the Peripheral and Accessories (P and A) with 36.18 percent. This dissection, which PowerBI achieves by its donut charts, enables the management to realize that, although the peripherals have a smaller revenue base, they have a high-yield opportunity as an attach-rate. More so, the channel analysis has indicated that Retailers hold the largest portion of sales (71.53.percent) implying that regional growth initiatives must focus more on physical retail collaboration rather than direct-to-consumer (DTC) channel within the present cycle.

B. Long-term Trends and Market Share.

An analysis of Net Sales (NS) conducted over a period from 2018 to 2022 showed that the trend of Net Sales (NS) was steadily rising, even though Gross Margin (GM%)), fluctuated. The framework monitored the market share of the company against the main competitors such as Dale, Innovo and Pacer. The PC Market Share Trend chart validated that the market is very fragmented although the business has had a consistent footprint of “Atliq. This is a longitudinal perspective that could differentiate between short-time seasonal peaks and long-term sustainable growth.

C. Profitability Metrics and Forecasting.

One of the major technical successes of the pipeline that was Python-integrated was the achievement of a Forecast Accuracy of 81.17%. This level of accuracy in the ”2022Est” projections allowed the business to have a Gross Margin of 38.08%. Although the Net Profit as a percentage displayed a loss of 13.98 percental, a diagnostic drill-down could be made by the insights table on the dashboard in understanding whether this has been driven by the high costs of expansion in the specific regions or by high operation overhead in certain zones such as India (-22.99 percental Net Profit).

V. DISCUSSION

The Data-Driven Decision Framework (DDDF) led to a high-fidelity executive dashboard which is simply a fine-grained view of the fiscal health of the business of computer hardware, and its position within the market. The findings have been categorized into four broad analytical lenses that are region performance, devolution of revenue by the divusional, market share based on competitors, as well as predictive accuracy.

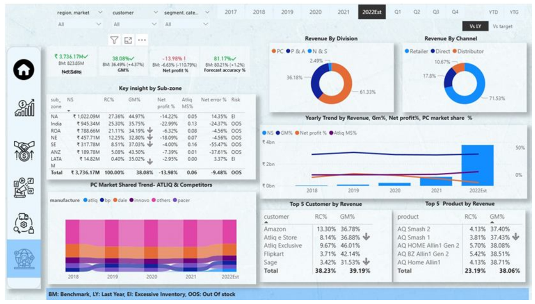

Fig. 3. PowerBI dashboard.

A. Regional Growth and Risk Assessment

The framework used the SQL-aggregated data which proved handy in per-formance classification on a per-Sub-zone scale. As shown in the regional analysis, the leading revenue contributors were North America (NA) and India which have contributed towards 1022.09M and 945.34M respectively. Nonetheless, the system was able to understand the existence of critical operational risks thanks to combining inventory-to-sales ratios:Excessive Inventory (EI): The regions such as North America and LATAM were identified to have EI, a factor which implies pro-motion liquidation or less procurement.Out of Stock (OOS): in its turn, high-potential areas, e.g., SE and ANZ were marked as an Out of Stock, i.e., they took away the possible revenue due to bottlenecks in supply chains. 5.2 Channel and Divisional Analysis. This revenue analysis by division shows that the PC segment is the main business with 61.33 percent of total

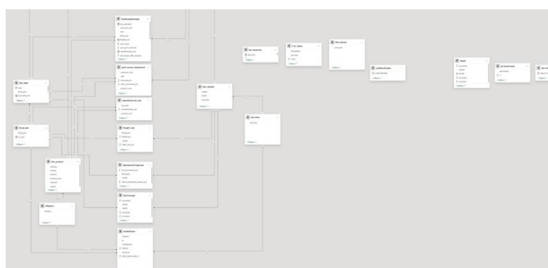

Fig. 4. Database Modelling.

revenue followed by the Peripheral and Accessories (P and A) with 36.18 percent. The dissection that PowerBI is able to fulfill with the help of its donut charts allows the management to recognize that even though the peripherals have lesser revenue base, it has a great-yield opportunity as an attach-rate. Greater, the channel analysis has shown that Retailers have the highest share of sales (71.53.percent) suggesting that the development of the regional growth initiatives should concentrate more on the physical retail collaboration than on direct-to-consumer (DTC) channel in the current cycle.

B. Long-term Trends and Market Share

A discussion of Net Sales (NS) done between 2018 and 2022 revealed that Net Sales (NS) trend was consistently increasing albeit with Gross Margin (GM%), which varied. The framework kept track of the companies market share as compared to the core competition of the company like Dale, Innovo and Pacer. The PC Market Share Trend chart confirmed the fact that the market is highly fragmented whilst the business has been enjoying a stable presence of “Atliq. This is a perspective of the long-range and might distinguish long-term sustainable growth and short-time seasonal maximization.

C. Metrics and Forecasting Profitability

Achieving a Fore-cast Accuracy of 81.17 was one of the most significant technical success stories of the pipeline which was Python-integrated. The projection accuracy of the references made on the 2022Est resulted in the Business having a Gross Margin of 38.08. The Net Profit as a percentage although indicated a loss of 13.98 percental, could be diagnostically drilled down by the insights table shown in the dashboard to determine whether it has been caused by the high cost by expansion in the areas or high overhead of the operations in certain areas like India (-22.99 percental Net Profit).

VI. CONCLUSION

The creation and application of this Data-Driven Decision Framework (DDDF) illustrate the impact of bringing domain-focused data analytics tools to computer retail hardware stores. In transitioning from fragmented data management to a consolidated framework, taking advantage of SQL for effective data governance, Python for advanced predictive analytics and PowerBI for compelling data visualisation, the company has created unprecedented operational transparency.

From the results, it is apparent that the system is capable of dealing with the special problems of the IT peripheral marketplace, including short product lifespan and regional inventory spikes. PoS level insights into ”Excessive Inventory” and ”Out of Stock" have enabled flexible supply chain strategies and has its impact on the bottom line. In addition, the planning accuracy of 81.17% plays a valuable role in future procurement planning - preventing valuable capital being lost on obsolete equipment.

In summary, this project has demonstrated that ”Full-Stack” data analytics not only delivers value but is a strategic imperative. By connecting database rows to C-level board rooms, the framework helps stakeholders to shift from a reactive to a win-win, growth-focused business strategy.

Future work Although the present framework is secure, future research will investigate the inclusion of more unstructured data, like customer feedback from social networks and tech forums, into the Python software. An additional Natural Language Processing (NLP) layer could potentially forecast the demand for peripherals based on "viral" trends, or product reviews, prior to them showing up in the sales figures. Finally, future work will explore extending the model to include real-time data scraping of competitor’s prices to give a further contextual understanding of the market, and to ensure the business remains viable in an increasingly price-driven online environment.’

REFERENCES

[1] E. F. Codd, “A Relational Model of Data for Large Shared Data Banks,” Comm. ACM, 1970. Link
[2] R. Kimball and M. Ross, The Data Warehouse Toolkit, 3rd ed. Wiley, 2013. Link
[3] P. Vassiliadis, “A Survey of ETL Technology,” Int. J. Data Warehousing, 2009. Link
[4] J. Celko, SQL for Smarties, 5th ed. Morgan Kaufmann, 2014. Link
[5] C. J. Date, An Introduction to Database Systems, 8th ed. Pearson, 2006. Link
[6] W. McKinney, “Data Structures for Statistical Computing in Python,” Proc. 9th Python in Science Conf., 2010. Link
[7] F. Pedregosa et al., “Scikit-learn: Machine Learning in Python,” J. Mach. Learn. Res., 2011. Link
[8] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” Proc. 22nd ACM SIGKDD, 2016. Link
[9] S. Raschka, Python Machine Learning, Packt, 2019. Link
[10] L. Breiman, “Random Forests,” Mach. Learn., 2001. Link J. Brownlee, Time Series Forecasting with Python, ML Mastery, 2017. Link
[11] C. R. Harris et al., “Array programming with NumPy,” Nature, 2020. Link
[12] S. Few, Information Dashboard Design, O’Reilly, 2006. Link
[13] E. Tufte, The Visual Display of Quantitative Information, Graphics Press, 2001. Link
[14] M. Russo and A. Ferrari, The Definitive Guide to DAX, Microsoft Press, 2019. Link
[15] B. Larson, Data Analysis with Microsoft Power BI, McGraw-Hill, 2020. Link
[16] C. Ware, Information Visualization, Morgan Kaufmann, 2012. Link
[17] F. Provost and T. Fawcett, Data Science for Business, O’Reilly, 2013. Link
[18] J. Wang, “Business Intelligence in Retail,” IEEE Int. Conf. Big Data, 2019. Link
[19] V. Oliveira et al., “Data-Driven Decision Making,” IEEE Access, 2021. Link
[20] R. Sharda, Business Intelligence and Data Science, Pearson, 2017. Link
[21] T. Davenport, “Competing on Analytics,” Harvard Bus. Rev., 2006. Link
[22] S. Chopra, Supply Chain Management, Pearson, 2016. Link
[23] G. Waller, “Data Science in Supply Chain Design,” J. Bus. Logistics, 2013. Link
[24] M. Berry, Data Mining Techniques for Sales, Wiley, 2011. Link