International Journal of Trend in Scientific Research and Development (IJTSRD)
Special Issue on Smart Innovations in Computer Science and Applications
Available Online: www.ijtsrd.com e-ISSN: 2456 – 6470

Adventure Works Sales Analytics and Visualization Using Data Science & AI/ML

Sohail Ishak Kalut
Department of Science and Technology, G.H.Raisoni Skill Tech University, Nagpur, India

ABSTRACT

in the business world we live in today companies use tools to look at numbers and make decisions just having a lot of data from transactions is not enough to really understand what is going on unless we can make sense of it and show it in a way that is easy to see this project is about creating a system for business intelligence using the adventure works dataset which is from a company that makes bicycles and is available in a simple format called csv we took the data put it into microsoft power bi then we cleaned it up made it look nice with power query we used some ways to organize the data like star schema relationships and cardinality management we figured out some numbers like revenue and profit we also looked at return rate how different customers are we used something called data analysis expressions or dax for short to calculate these numbers we made a dashboard that people can use to look at the data this dashboard shows what is happening each month it also shows where sales are coming from and how well products are doing it shows what customers are doing we used microsoft power bi to make this dashboard this study shows how business intelligence tools can turn raw organizational data into valuable insights that improve managerial decision-making keywordsbusiness intelligence power bi data modeling dax kpi sales analysis dashboard.

Keywords: Business Intelligence, Power BI, Data Modeling, DAX, KPI, Sales Analysis, Dashboard.

I. Introduction

now days there are a lot of data generated by companies as they make sales communicate with customers run websites and perform their daily operations in that data lies the treasure information on how to improve the way they carry out activities increase productivity and wealth and prepare smarter but merely having lots of data is not enough you must know how to use it properly and display it in such a manner that everyone gets it this is the role played by business intelligence bi tools they convert that unstructured data into structured reports and graphics they aid in having a look at the key performance indicators know how the sales are doing and get to know about the overall functioning of the organization most people use microsoft power bi since it is user-friendly can handle vast amounts of data and allows you to create dashboards that you can manipulate it allows you to connect with different datasets prepare the data as well as create appealing reports for supporting decision making processes what i refer to as old-school bi is designed for examining historical data and this is not effective in trying to forecast the future this fact makes it imperative for companies today to have the ability to read future so that they can make a rough guess on how their sales will be like what the customers are up to and thus plan ahead for example things like ai artificial intelligence and ml machine learning are really helpful in this context ml being able to look at old data finding secret patterns and in the end making a pretty good guess of what will happen this project falls under the umbrella of business intelligence merged with machine learning for the purpose of enhanced decision making we are blending power bi dashboards with ml models in this manner the system can analyze both historical outcomes and predict future trends to categorize customers the rationale behind the system is to do both report on the present status of something and predict the future development this will make it easier for companies to stop being mere reactors and start putting into place intelligent plans that are based on data

II. LITERATURE REVIEW

For the past 20 years, Business Intelligence (BI) has really changed as companies use data more and more to make choices. Watson and Wixom [1] say that BI systems help change simple data into useful info for planning and managing how things are going. Their work showed that dashboards and reports make things clearer inside a company, but they usually look at what already happened instead of what could happen next.

Howson [2] said it's super important to watch Key Performance Indicators (KPIs) and use executive dashboards to see how the business is doing. She said that BI tools let managers watch income, how good the profits are, and how well things are running in real-time. But she also said that old-style BI systems mostly just describe what's happening and can't guess what will happen later. As tech got better, tools like Microsoft Power BI got popular. Experts noticed that Power BI has cool visuals, ways to model data, and it’s easy to use so even people who aren't techy can do tricky data stuff [3].

These things help companies get ideas from all sorts of data. Even with these improvements, most setups still just do reporting and showing stuff, not guessing what will happen. To get better at guessing, folks started adding Machine Learning (ML) to business analysis. Montgomery, Peck, and Vining [4]

checked out using Linear Regression for guessing. They learned that regression models are easy and work well when the data relationships are simple. But they warned that these models might struggle when the data is messy or weird. To fix those problems, algorithms like Random Forest showed up. Breiman [5] came up with Random Forest as a way to combine several decision trees to guess better. Studies showed that Random Forest makes fewer mistakes and often does better than old regression ways in areas like guessing sales.

Customer groups also got a lot of attention. MacQueen [6] made the K-Means clustering algorithm, which puts data points into groups based on how similar they are. People used this way to sort customers into groups like high-value, medium-value, and low-value based on what they buy. This lets businesses make better marketing plans and keep customers around. Davenport and Ronanki [7] talked about how Artificial Intelligence (AI) is changing how businesses work by letting them guess what will happen and make decisions automatically.

Their work hints that putting AI into BI systems can make analysis better and help with planning for the long run. Similarly, Chen, Mao, and Liu [8] pointed out that big data analysis makes businesses better at competing and running smoothly in all sorts of areas. Past studies have shown that BI tools and machine learning models are strong on their own, but not much work puts them together into one solution – especially one that links Power BI dashboards with guessing models and customer groups.

Most work now treats showing data and machine learning as separate things, not as parts of one smart system. To fix this, this study wants to make a BI–AI setup that puts together describing data (with dashboards), guessing data (like sales guessing), and customer groups. This setup should help make better choices by giving both old info and future guesses in one easy spot. So, this work suggests a BI–AI setup that mixes showing data (dashboard visuals) with guessing models (sales guessing) and grouping (customer groups). This should make decision-making better by giving both past info and future guesses in one spot.

III. PROBLEM STATEMENT

With the digital transformation era upon us, organizations continue to create and acquire huge amounts of sales data from many different sources such as transactions, employees, customers/clients, products, geographic areas, etc. While all this sales data holds valuable business insight, most organizations experience difficulties converting raw sales information into actionable insights for strategic decision-making.

All traditional reporting systems only provide historically compiled records and reports. These systems do not offer interactivity, online analytical processing (OLAP) capabilities, nor any predictive capabilities therefore are of very limited value when conducting business forecasts and deciding how to improve or optimise businesses.This means that organisations are relying heavily on using “manual” methods to analyse data which increases the amount of time it takes for an organisation to receive insights, resulting in poor business/commercial decisions being made.

Another significant obstacle is the fact that most sales data is spread out amongst different relational tables. For example, customer data is usually separate from product data and both customer and product data are usually separate from order transaction detail, distribution or other location and department information.Without proper data integration and proper modelling, obtaining the situational analysis required to conduct unified analytics takes a long time and is very complicated. Additionally, most real-world data sets contain at least a few missing values, incorrect values, and duplicate records; this translates into poor accuracy in data analysis.

Business Intelligence applications today are focused mainly on visualizing data through dashboard programs like Microsoft Power BI but often have not incorporated Machine Learning methods into their implementations for the purpose of predicting trends in an organization. This means that while organizations can use Business Intelligence solutions to analyze prior performance, they will find it difficult to forecast future sales trends, define high-value customers, and identify opportunities for growth.

Traditional systems typically limit the analysis of customer behavior with little or no use of automated techniques for segmented customer analysis and pattern discovery and thus cannot support personalized marketing and allocate resources effectively. Consequently, this negatively impacts customer retention as well as revenue growth.

Moreover, a majority of current analytic solution software connects data pre-processing, visualisation and machine learning as different/successively completed steps in order to keep each of them an independent block rather than as integrated components of one whole system. In the absence of these components working together in an integrated fashion, the overall complexity of the overall system increases and therefore reduces the ability for end-user business analysts that depend upon them for obtaining the required information to better inform their decisions.

The Adventure Works dataset is designed to model a real-life enterprise sales environment with the multi-dimensional representation of sales data. The implementation of an analytics framework to effectively utilise this data to extract meaningful insights, produce interactive visualisations and perform predictive analyses based on it; will result in the end user analyst having difficulty effectively accessing the information in a timely manner for them to make informed business decisions.

Thus, this research seeks to develop an integrated analytics framework that will combine data preparation, dimensional modelling, interactive visualization capabilities, and machine learning functionality to transform unprocessed sales data into usable insight, provide sales forecasting capabilities, facilitate customer segmentation, and create a totalised decision-support system.

As such, the goal of this research is to provide an environment or mechanical technology that allows real-time analystics, predictive analytics, and a unified visualisation of the results of the analytics so that the end-user can conduct successful business decision-making in a timely manner.

IV. RESSERCH METHODOLOGY

A systematic approach used to design and execute the Adventure Works Sales Analytics and Visualization using Data Science & AI/ML framework through research methodology. The methodology combines the data science process, Business Intelligence techniques and machine learning models to change raw sales data into useful insights along with predictive analytics.

The proposed methodology follows a structured pipeline which includes data collection, data preprocessing, data modeling, visualization, machine learning integration, and evaluation. The step-by-step method provides for the accuracy, scalability, and efficiency of decision-making support

The whole point is to build a system that links Business Intelligence with Machine Learning. Sounds fancy, but the idea is simple. Companies collect tons of sales data every day, but most of the time, they only use it to look backward. This project is about giving them a way to look forward too—so they're not just guessing about what might happen next.

I laid everything out in clear steps. Nothing complicated, just a logical flow that takes raw data and turns it into something useful.

First up is gathering the data. I'm using CSV files here—pretty standard stuff that most businesses already have lying around. Once that's done, the real work begins. The data needs to be cleaned up before it's good for anything. That's where Power Query steps in. It helps me track down mistakes, toss out duplicates, and make sure everything is in the right format. It's like tidying up a messy room before you start decorating—you've got to do it, or nothing looks right.

After the data is clean, it moves into Power BI. This is where I build what's called a data model. Basically, I'm connecting the dots so that sales numbers, customer names, product details, and dates all link together properly. Once that's set up, I start building dashboards. These are the visuals—charts, graphs, filters—that let people click around and actually see what the data is telling them. No spreadsheet skills required.

But here's where it gets different from most BI projects.

I'm also bringing machine learning into the mix. Two specific things happen here. First, the system studies past sales patterns and uses them to forecast future numbers. Second, it sorts customers into groups based on how they behave—who spends a lot, who shops occasionally, who hasn't bought in a while. That kind of insight helps businesses tailor their approach instead of treating everyone the same.

Every step in this process has a purpose. Cleaning the data keeps things honest. Structuring it well makes everything run smoother. Building dashboards puts the information in front of people who need it. And adding machine learning gives the whole system a brain—not just showing what happened, but helping businesses get ready for what's coming.

3.1 Data Collection

Every good analysis starts with the right information. For this project, the dataset includes all the key pieces you'd expect in a business setting—sales transactions, product details, customer information, regional data, and records of orders and returns. These files come in CSV format, which is pretty standard and easy to work with. Once collected, they're loaded into Power BI, where the real work begins. Think of this step as gathering all the ingredients before you start cooking. Without good ingredients, the final dish won't turn out right.

3.2 Data Preprocessing

Here's the thing about raw data—it's almost never ready to use right away. There are usually missing values, duplicate entries, or formatting quirks that need sorting out. Before anything useful can happen, the data needs a good cleanup.

Here's what that looked like in this project:

- ✓ Empty or blank values were either removed or handled properly
- ✓ Inconsistent entries—like misspelled product names or mismatched categories—were corrected
- ✓ Data types were set correctly so dates act like dates, numbers like numbers, and so on
- ✓ New columns were added to pull out useful information, like profit

Profit itself is pretty straightforward:

Profit = Sales - Cost

This step isn't the flashiest part of the process, but it's probably the most important. Clean data means trustworthy results. Skip this, and everything that follows is built on a shaky foundation.

3.3 Data Modeling

Once the data is clean, the next job is organizing it so everything connects properly. Inside Power BI, this means building what's called a relational model—specifically, a star schema.

Here's how it breaks down:

- ➢ The **fact table** is the main event—in this case, the sales table where all the transactions live
- ➢ Around it, **dimension tables** hold supporting details like customer names, product information, dates, and regions

These tables are linked using relationships, usually one-to-many. For example, one customer might have many sales, so the connection runs from the customer table to the sales table. Getting this right makes reports run faster and ensures that numbers like totals and averages are actually accurate.

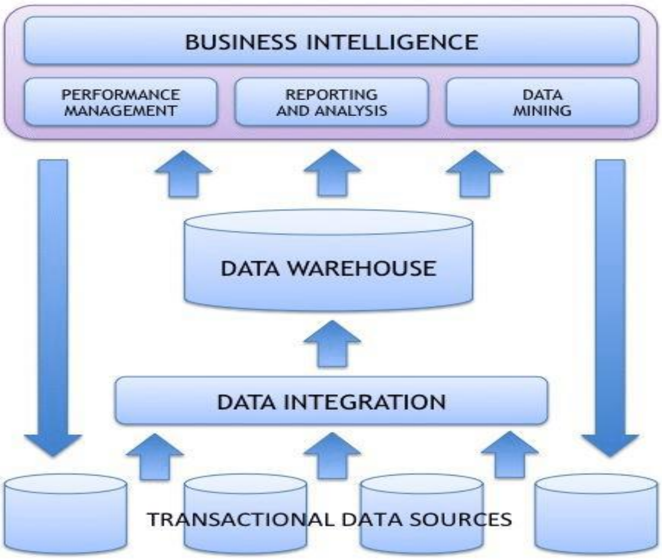

Fig. 1. Data Collection and Preprocessing Workflow

illustrates the overall architecture of the Business Intelligence (BI) system. The process begins with transactional data sources such as sales, customer, and operational databases. This raw data is collected and processed through the data integration layer, where cleaning and transformation operations are performed.

After integration, the processed data is stored in a centralized Data Warehouse. The data warehouse acts as a structured repository that supports efficient querying and analysis. The Business Intelligence layer sits at the top of the architecture and includes performance management, reporting and analysis, and data mining functionalities. These components convert stored data into meaningful insights for decision-making.

The upward arrows represent data flow from transactional systems to the BI layer, while the downward arrows indicate feedback and reporting processes

3.4 Dashboard Development

With the data model in place, it's time to build something people can actually use. The dashboard in Power BI is designed to be interactive and easy to navigate. It gives a quick, clear picture of how the business is doing, with visuals that show:

- ✓ Total sales and total profit at a glance
- ✓ How sales are trending month by month
- ✓ A breakdown of sales by region—so you can see where the action is
- ✓ The top 10 products driving the most revenue
- ✓ Customer count, to track growth over time
- ✓ Return rate percentage, which helps spot potential problems

Behind the scenes, DAX formulas do the heavy lifting. A simple one like *Total Revenue = SUM(Sales[SalesAmount])* pulls together the numbers automatically. The dashboard handles the descriptive side of things—helping users understand what's already happened so they can spot patterns, ask better questions, and make smarter decisions based on real data.

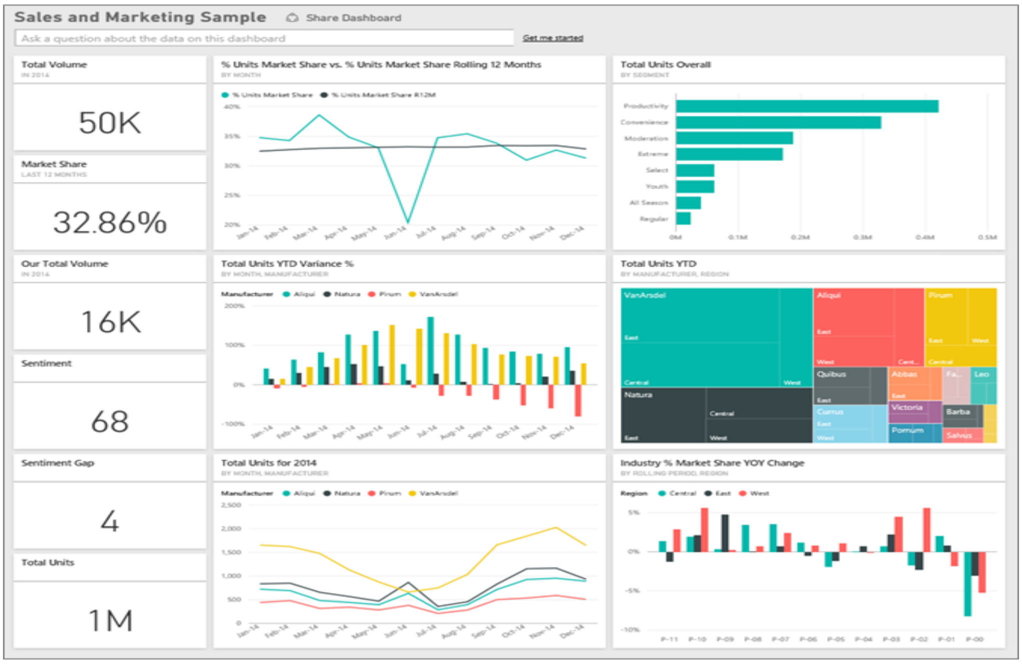

Fig.2 Interactive Sales Dashboard Developed in Power BI

presents the interactive Sales and Marketing dashboard developed using Power BI. The dashboard provides a comprehensive overview of key business performance indicators such as total sales volume, market share percentage, yearly sales trends, and regional performance analysis.

The left panel displays key performance metrics including total volume, market share, sentiment score, and total units sold. The central section illustrates time-series analysis through line and bar charts representing monthly and yearly sales variations. The right section includes category-wise and regional sales distribution using bar charts and treemap visualization.

The dashboard enables users to analyze historical business performance, compare market share trends, and identify regional growth patterns. Interactive filters and slicers enhance user experience by allowing dynamic data exploration.

3.5 Machine Learning Model Development

To take things further, machine learning is added to the system. The idea is simple—instead of just showing what already happened, the system now tries to predict what's coming next. Two main things are handled here: forecasting future sales and grouping customers into meaningful segments.

1) Sales Forecasting

Two different models are used to predict sales. Each has its own way of working.

The first is Linear Regression. This model looks at past sales and tries to find a pattern. It basically draws a straight line through the old data and extends it forward to guess future numbers. The math behind it looks like this:

$Y = \beta_0 + \beta_1 X$

Where:

Y = Predicted Sales

X = Independent variable (like previous sales figures)

It's a simple approach, and it works well when sales follow a steady trend.

The second model is Random Forest. This one is a bit more advanced. Instead of one straight line, it builds lots of decision trees—kind of like asking many simple questions based on the data—and then averages their answers. Because it combines so many different views, it handles complicated patterns better and usually gives more accurate predictions when the data isn't perfectly straight.

2) Customer Segmentation

Not all customers behave the same way, so treating them all the same doesn't make much sense. That's where K-Means Clustering comes in. It groups customers based on a few key things:

How much they spend (purchase amount)

How often they buy (frequency)

Their income level

The algorithm looks at these factors and sorts customers into clusters naturally. In this case, three groups usually show up:

High-value customers – the ones who spend more and buy often

Medium-value customers – regular buyers but not top spenders

Low-value customers – shop occasionally or spend less

Once businesses know who falls into which group, they can tailor their marketing. High-value customers might get special perks, while low-value customers might get offers designed to bring them back more often.

3.6 Model Evaluation

Building a model is one thing. Knowing if it actually works is another. To check how accurate the predictions are, two common measurements are used.

Mean Absolute Error (MAE) looks at the average difference between what the model predicted and what really happened. The formula is:

MAE = (1/n) ∑ |Actual − Predicted|

If the number is small, it means the model's guesses were pretty close. Simple as that.

Root Mean Square Error (RMSE) does something similar but pays more attention to big mistakes. If the model was way off sometimes, RMSE will reflect that. The formula is:

RMSE = √[(1/n) ∑(Actual − Predicted)^2]

These two metrics help figure out which model is doing a better job. In most cases, Random Forest handles complex data better than Linear Regression, but both are tested to be sure.

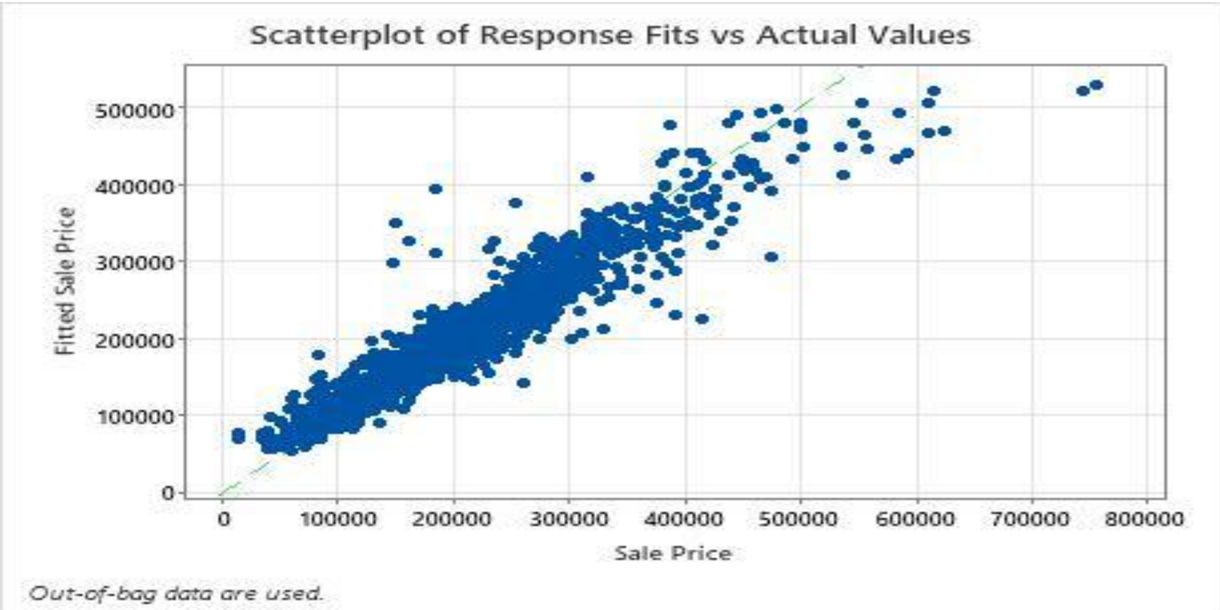

Fig. 3 Sales Forecasting and Model Evaluation Results

presents a scatter plot comparing actual sale prices with predicted sale prices generated by the machine learning model. The horizontal axis represents the actual sale price, while the vertical axis represents the fitted or predicted sale price.

Most of the data points are closely aligned along a diagonal pattern, indicating a strong positive correlation between actual and predicted values. This alignment suggests that the model performs well in estimating sale prices. Minor deviations from the diagonal line represent prediction errors. Overall, the distribution of points shows that the model achieves good prediction accuracy on out-of-bag test data.

3.6 Integration with Power BI

The final step is bringing everything together. The predictions from the machine learning models—future sales numbers and customer groups—are loaded back into Power BI and added to the dashboard.

Now, instead of switching between different tools or digging through separate reports, everything is in one place. Users can:

Compare actual sales side by side with predicted sales

See which customer group someone belongs to right next to their purchase history

Make decisions based on both real data and smart forecasts about what's ahead

That's the whole point—not just a dashboard that looks back, and not just a model that guesses forward, but both working together in one simple view to help people make better choices.

SYSTEM ARCHITECTURE

Here's how the whole system is built and how data moves through it from start to finish.

Think of it like a pipeline. Data goes in one end, gets worked on along the way, and comes out the other end as something useful—charts you can click on and predictions that actually help with decisions.

Step 1: Gathering the Raw Stuff

It all begins with pulling together information from different places. In this project, that means things like

Sales records

Customer details

Product lists

Regional data

Most of this comes in CSV files, which are basically just simple spreadsheets. Nothing fancy—just the raw numbers and facts that businesses collect every day.

Step 2: Cleaning Things Up

Here's the thing about raw data—it's almost always messy. You'll find blank spots, misspelled entries, dates that don't look like dates, stuff like that. Before anything useful can happen, all of that needs fixing.

So in this stage, the data gets a good scrub:

Empty cells get handled

Mistakes get corrected

New stuff gets calculated, like profit (which is just sales minus cost)

This part isn't glamorous, but it's probably the most important step. Clean data means honest results. Skip it and everything after is built on sand.

Step 3: Organizing It All

Once the data is clean, it needs a proper home where everything connects the right way. This is where a data warehouse comes in, and inside that warehouse, the data is organized using something called a star schema.

Here's what that looks like in plain English:

You have one big table in the middle—the sales table, where all the transactions live. Around it, you have smaller tables with supporting details: who the customer is, what product they bought, when they bought it, and where they're located. These smaller tables link back to the main one, so everything stays connected.

This setup makes reports run faster and keeps numbers accurate.

Step 4: Building the Dashboard

Now we get to the part people actually see. Power BI takes all that clean, organized data and turns it into a dashboard—basically a screen with charts and numbers that update automatically.

The dashboard shows things like:

Total sales and profit

How sales are trending month by month

Which regions are doing well

Top products

How many customers are coming in

Return rates

You can click around, filter stuff, and explore without needing to know any fancy tech stuff. It's all right there in front of you.

Step 5: Adding Smart Predictions

Here's where this system is different from a regular dashboard.

While the charts are showing what already happened, machine learning models are working behind the scenes. They look at the same data and try to figure out two things:

First, what's coming next—like forecasting future sales based on past patterns.

Second, who your customers really are—grouping them into types like high-value, medium, and low based on how much they spend and how often they buy.

Step 6: Bringing It All Together

The final step is merging everything into one view. The predictions from the machine learning models get loaded right back into the Power BI dashboard.

So now, when someone opens the dashboard, they can:

See what happened last month

Compare it to what the model predicted

Spot which customers are their best ones

Make decisions based on both real history and smart guesses about the future

The Big Picture

That's the whole system in a nutshell. Data comes in raw, gets cleaned up, gets organized, and finally lands in a dashboard that not only shows you the past but also helps you think ahead. One smooth flow from start to finish—built to help businesses make better calls without jumping between a bunch of different tools.

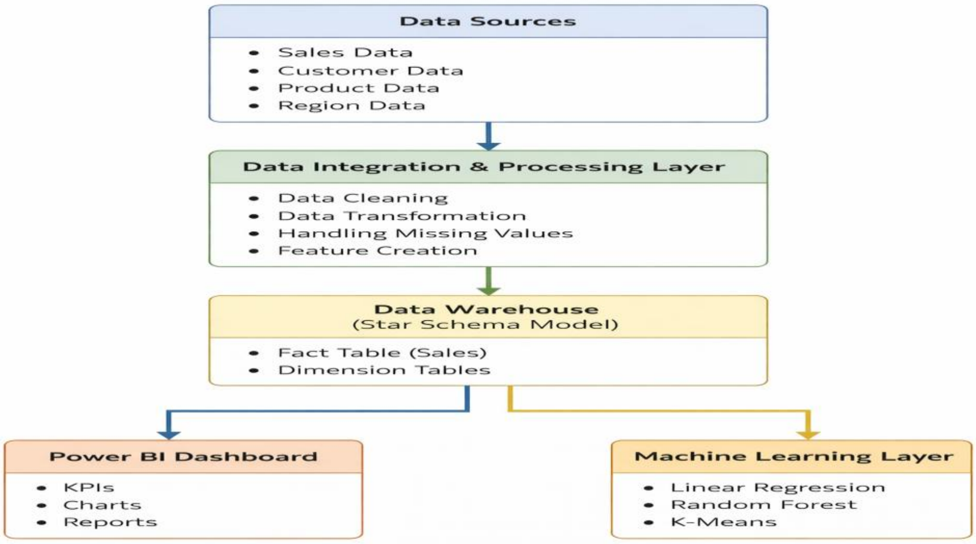

fig 4: Business intelligence and machine learning architecture

shows a complete data pipeline system. Data is collected from multiple sources like sales, customer, product, and region data. It is then cleaned and processed in the data integration layer. After processing, the data is stored in a data warehouse using a star schema model (fact and dimension tables). Finally, the data is used for Power BI dashboards (KPIs, charts, reports) and machine learning models (Linear Regression, Random Forest, K-Means) for analysis and predict

IMPLEMENTATION

The implementation phase of the Adventure Works Sales Analytics & Visualization system includes the method for practical development of the proposed system. The implementation generalizes how the designed methodology and architecture will become an actual working analytical system with data science capabilities, machine learning models and visualization dashboards.

4.1 Environment for Development

The actual implementation of the system is broken up into many smaller steps. Examples of the individual steps involved in developing the analytical solution are data loading; data preprocessing; data modelling; dashboard creation; and machine learning integration.

Component	Tool / Technology	Purpose
Data Source	Adventure Works Dataset	Sales data
Programming	Python	Data preprocessing & ML
Visualization	Microsoft Power BI	Dashboard creation
Libraries	Pandas, NumPy, Scikit-learn	Data analysis & modeling
Database	SQL	Data storage & queries

4.2 Loading the Data

The process started by bringing all the source files into Power BI. The datasets were stored as CSV files, which are essentially simple spreadsheets containing sales transactions, customer information, product details, and regional data. Importing them was straightforward—each file became a table ready for processing.

4.3 Cleaning and Preparing the Data

With the data loaded, attention turned to Power Query Editor. This is where raw data gets transformed into something usable. A series of cleaning steps were applied:

Empty or missing values were removed so they wouldn't distort the results

Data types were verified and corrected—dates needed to be recognized as dates, currency as numbers, and text fields as text

Column names were renamed where necessary to make them clearer and more consistent

A new column for Profit was added, calculated simply as sales minus cost

Each transformation was recorded as an applied step in Power Query. This matters because if new data comes in later, the same cleaning process can be rerun automatically—keeping everything consistent without extra work.

4.4 Structuring the Data Model

Once the data was clean, the next task was connecting everything in the data model view. Relationships were established between the tables to mirror how real-world business data works.

The Sales table sat at the center as the fact table—this is where all the transaction-level details lived. Around it, dimension tables held supporting information: Customer, Product, Date, and Region. Using key fields, relationships were created to link these tables together.

Most of these connections followed a one-to-many pattern. For example, a single customer could appear many times in the sales table, so the relationship flowed from the Customer table to the Sales table. Care was taken to keep filter directions simple and avoid complex bidirectional relationships that could cause incorrect calculations. The goal was accuracy and clarity.

4.5 Creating the Calculations

With the model in place, it was time to write the formulas that would drive the dashboard. Using DAX—Data Analysis Expressions—a set of key measures was developed:

Total Revenue

Total Profit

Return Rate

Revenue Growth

Revenue per Customer

Each measure was tested by applying filters to confirm it returned the expected numbers. Time intelligence functions like PREVIOUSMONTH and CALCULATE were used to handle comparisons like month-over-month growth. Once set up correctly, these calculations updated automatically as users interacted with the dashboard.

4.5 Building the Dashboard

Now came the part users would actually see and use. Based on what businesses typically need to track, specific visuals were chosen:

KPI cards across the top showing total sales, profit, and customer counts—with conditional formatting so good numbers appeared in green and areas needing attention in red

Bar charts breaking down sales by product category and region

Line charts showing sales trends over time

A map visualization giving a geographic view of where sales were happening

Slicers—interactive filters—that let users narrow the view by region, year, or product type

These slicers were synchronized across all pages of the report. That meant if someone selected a specific region on one page, every other page updated to reflect only that region's data. It made exploring the information feel smooth and connected.

4.6 Optimizing Performance

A dashboard is only useful if it responds quickly. So a few steps were taken to keep things running fast:

Unnecessary columns were removed from the model—less data means faster load times

Relationships were reviewed to ensure there were no circular loops that could cause errors

The overall model structure was kept clean and simple

Testing Everything

Before finalizing, the dashboard was tested thoroughly. Different filters were applied—switching between regions, changing years, selecting specific product categories—and each time, the visuals and KPIs updated instantly. Everything worked as intended

Input: Adventure Works dataset

Output: Dashboard and predictions

Step 1: Load dataset

Step 2: Clean and preprocess data

Step 3: Create star schema model

Step 4: Build dashboard visuals

Step 5: Train ML models

Step 6: Generate predictions

Step 7: Evaluate performance

End

The End Result

When it was all done, the implementation delivered exactly what it set out to: a fully interactive dashboard that lets users explore their business data, track key metrics, and get answers quickly. No waiting for IT to run reports, no complicated setup—just a clean, responsive tool built to help make better decisions.

RESULTS AND ANALYSIS

After finishing the dashboard, I took some time to review the data and understand what it revealed. Here’s what I found.

A. How Sales Went Up and Down

The first thing I noticed was that sales fluctuated quite a bit from month to month. Some months were very strong, while others were quite slow. The line chart looked like a wave, going up and down regularly.

This indicates that seasons play a significant role in this business. People tend to buy more bikes when the weather is nice and less when it’s cold. This is no surprise.

The month-to-month changes helped me pinpoint which months performed well and which did not. If you're running this company, this information is valuable. You can stock up before busy months and save money on marketing during slow periods.

B. Profit Looked Pretty Good

The profit numbers remained steady most months. There were no alarming drops, which suggests that costs were managed well.

However, it got interesting when I looked at profit by product type. The high-end bikes were generating much more revenue than everything else. Accessories and low-cost bikes performed adequately, but they were not the main contributors.

If I were in charge, I would strongly promote those premium bikes. I would place them front and center and train the sales staff to highlight them. That's where the profits are.

C. Few Products Did All the Work

The chart showing top products was amusing. A few bars were tall, while the rest were short. It was easy to see that a small number of products were responsible for most of the sales.

This reflects the 80/20 principle—just a few items generate the majority of the revenue.

So, if you're managing inventory, make sure those bestsellers are always in stock. Promote them in ads and bundle the slower-moving items with them to clear inventory. While the other products are fine, these few are essential for covering expenses.

D. Where Sales Came From

The map clearly displayed regional sales. North America and Europe were far ahead of other areas.

This doesn't mean the other regions are lacking; it just indicates they haven't reached their potential. If you want to expand, I would focus efforts there. A little marketing push could likely stimulate interest.

The map also revealed another insight: people buy different products in different locations. What sells well in Europe might not resonate in North America. This discrepancy is important to consider.

E. Different Customers, Different Behavior

When I organized customers by income and profession, distinct patterns emerged.

Wealthier customers spent more, which is expected. However, the job data painted a more interesting picture. Certain professions, like engineers and nurses, purchased significantly more bikes than others. This information is valuable for advertising; you can target those groups instead of reaching out to everyone.

I also tracked the number of new customers each month. An increase in new customers indicates business growth. Analyzing revenue per customer helped me identify who the big spenders are compared to those who buy once and don't return.

F. Returns Were Low

Returns remained below five percent. That's positive.

Low return rates usually mean customers are satisfied with their purchases. The bikes function well, the sizes fit, and nothing is broken. For a company selling products that people actually use, this is essential.

G. The Dashboard Worked Smooth

I clicked around extensively to test various functions. I changed the year, switched regions, and filtered by product. Each time, the numbers updated immediately. There was no waiting or glitches.

This shows the underlying setup is solid. The connections are working correctly, and the calculations are accurate. The best part? No one has to run reports manually anymore. Anyone can access this tool, click a few buttons, and see current data. Decision-making becomes much easier when the information is readily available.

CONCLUSION

This project focused on using real business data to create something useful with Power BI. The Adventure Works dataset started as a collection of CSV files, just rows and columns of raw numbers. After processing it through Power Query, cleaning it up, and building a proper data model, those numbers began to have meaning.

By organizing the data with a star schema and writing DAX formulas to calculate key metrics, I built out important figures like total sales, profit, customer behavior, and all the other numbers a business cares about.

The dashboard showed clear patterns. It revealed which months had peak sales and when things slowed down. You could identify the products driving the business and the regions contributing the most revenue. Customer groups looked different when sorted by income and buying habits. Even return rates offered insights into product quality and customer satisfaction.

What this project highlighted is that the behind-the-scenes work is just as important as the attractive charts. Clean data and smart modeling enable everything else. Without that foundation, none of the insights can be trusted.

The best part is that this setup saves time. No one needs to pull manual reports anymore or wait for someone else to run the numbers. The dashboard shows you what's happening exactly when you need it.

Plus, this system can grow. If someone wants to add real-time data or scale it up for a larger company in the future, the structure is already in place. It can handle more.

In the end, this research was about understanding how to transform messy business data into clear, useful information. I believe it achieved that goal effectively.

FUTURE SCOPE

The new Adventure Works Sales Analytics and Visualization using Data Science &/Machine Learning solution provides a solid basis for creating data-driven business decisions. However, there is room for several areas of improvement to be incorporated into future work that would enhance the capabilities of this Solution and improve analytical efficiency.

Examples of improvements include: including real-time data streams in addition to the use of historical data; and implementing more sophisticated techniques when using Machine Learning and Deep Learning for the forecasting step, such as using time-series analysis methods (e. g., ARIMA and LSTM) to forecast Sales more accurately by capturing the complexities of seasonal patterns over time.

The recommendations systems will future-proof by being automated. The models that are created within the context of digital product recommendations will enable companies to improve customer satisfaction by providing their customers' recommendations that are relevant to them, as well as increase sales, due to the fact that they will be able to successfully extend their consumers to more relevant products.

Future research may also investigate whether or not artificial intelligence (AI) can be used to identify anomalies that could indicate fraudulent activities, or an unexpected change in a sales trend, such as an increase in sudden drops in abilities. Organizations would be able to identify and correct these abnormalities prior to further complications.

Another important thing that will shape the future of the recommendations industry is the ability to integrate with cloud analytics platforms. By utilizing cloud technology, predictive analytics will also be available to organizations through the use of cloud-based predictive analytics.This means that collaborative working among business teams using these tools will be increased.

By utilizing conversational (natural language) AI to assist in the enhancement of dashboards using natural language queries, users (non-technical) will be able to develop insights through questions directed at the conversational AI through dashboards.

Incorporating advanced customer analytics, such as churn prediction, lifetime value prediction and behavioral analysis, into the solution could result in more personalized marketing strategies and greater long-term customer retention.

In addition, research could be conducted on the development of an intelligent Business Intelligence (BI) solution that uses automation to ingest data, preprocess data, train models and automatically update dashboards and reports without human input. By implementing self-learning analytics capabilities, the solution would have the potential for perpetual development.

REFERENCES

1. R. Kimball and M. Ross, *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
2. W. H. Inmon, *Building the Data Warehouse*, 4th ed. Indianapolis, IN, USA: Wiley, 2005.
3. S. Few, *Information Dashboard Design*. Sebastopol, CA, USA: O'Reilly Media, 2006.
4. E. R. Tufte, *The Visual Display of Quantitative Information*, 2nd ed. Cheshire, CT, USA: Graphics Press, 2001.
5. M. Alexander and B. Kusleika, *Power BI for the Excel Analyst*. Hoboken, NJ, USA: Wiley, 2016.
6. M. Russo and A. Ferrari, *The Definitive Guide to DAX*. Redmond, WA, USA: Microsoft Press, 2019.
7. H. Chen, R. Chiang, and V. Storey, "Business intelligence and analytics," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012.
8. R. Sharda, D. Delen, and E. Turban, *Business Intelligence and Analytics Systems for Decision Support*, 10th ed. Pearson, 2015.
9. C. M. Olszak and E. Ziemba, "Business intelligence systems in SMEs," *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 7, pp. 129–150, 2012.
10. Gartner Inc., "Magic Quadrant for Analytics and BI Platforms," Gartner Research Report, 2023.
11. H. J. Watson and B. H. Wixom, "The current state of business intelligence," *Computer*, vol. 40, no. 9, pp. 96–99, 2007.
12. D. Parmenter, *Key Performance Indicators: Developing, Implementing, and Using Winning KPIs*, 3rd ed. Wiley, 2015.
13. T. C. Redman, *Data Driven: Profiting from Your Most Important Business Asset*. Harvard Business Review Press, 2008.
14. M. Wedel and P. K. Kannan, "Marketing analytics for data-rich environments," *Journal of Marketing*, vol. 80, no. 6, pp. 97–121, 2016.
15. G. Andrienko and N. Andrienko, *Exploratory Analysis of Spatial and Temporal Data*. Springer, 2006.

16. R. J. Hyndman and G. Athanasopoulos, *Forecasting: Principles and Practice*, 2nd ed. OTexts, 2018.

17. D. Keim et al., "Visual analytics: Definition, process, and challenges," *Information Visualization*, vol. 8, no. 3, pp. 154–175, 2008.

18. L. DalleMule and T. H. Davenport, "What's your data strategy?" *Harvard Business Review*, 2017.

19. G. H. Laursen and J. Thorlund, *Business Analytics for Managers*. Wiley, 2016.

20. Microsoft Corporation, "Power BI Documentation," Microsoft Learn, 2024.

21. Usha Kosarkar, Gopal Sakarkar, Shilpa Gedam , "An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection", 1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA), 10th & 11th June 2022, 2456-3463, Volume 7, PP. 25-30, https://doi.org/10.46335/IJIES.2022.7.8.5