State University of Trade and Economics

Department of Digital Economy and System Analysis

QUALIFICATION WORK

On the topic:

«System Analysis of Volkswagen Group Sales»

Student of the 2nd year, 7 am group,
of Master's degree (Second Cycle)
Subject area
124 «System analysis»
educational programme
«Information technologies and business analytics (DataScience)» _____ Tatiana Panchenko
signature of student

Scientific supervisor
PhD in Economics,
Associate Professor _____ Volodymyr Kulazhenko
signature of supervisor

Director of the educational programme
Doctor of Sciences (Economics),
Professor _____ Andrii Roskladka
signature of director

Kyiv 2024

State University of Trade and Economics

Faculty of Information Technologies
Department of Digital Economy and System Analysis
Educational degree Master's degree
Subject area 124 «System analysis»
Educational programme «Information technologies and business analytics (DataScience)»

Approved by

Head of the department _____ A.Roskladka
«15» December 2023

Task
For the qualification work for

Panchenko Tatiana
(full name)

1.Topic of the qualification work
«System Analysis of Volkswagen Group Sales»
Approved by the SUTE decree dated on «27» November 2023 № 4143
2. Deadline for submitting the finished work by the student «05» December 2024
3. Goal setting and initial data to work
Goal of the work: is to conduct a system analysis of Volkswagen Group's car sales to identify key trends and build forecast model
Object of the work: Volkswagen Group car sales
Subject of the work: financial indicators of sales of the Volkswagen Group for 2006-2023.

4. Contents of the qualification work (list of questions for each chapter)

5. Calendar plan of work performance

№	Name of stages of graduation qualification work	Term of work performance stages	
		By plan	In fact
1	2	3	4
1	*Choosing thetopic of the qualification work*	*20.11.2023*	*20.11.2023*
2	*Development and approval of the task for the qualification work*	*15.12.2023*	*15.12.2023*
3	*Introduction*	*01.03.2024*	
4	*Chapter1. The Volkswagen Group*	*25.06.2024*	
5	*Chapter2. Tools for data analysis and visualizations*	*01.09.2024*	
6	*Publication of an article in the journal of scientific articles of the masters' degree students*	*15.09.2024*	
7	*Chapter 3. Sales forecasting using machine learning*	*18.10.2024*	
8	*Conclusions and recommendations*	*01.11.2024*	
9	*Submission of the qualification work to the department to the scientific supervisor*	*15.11.2024*	
10	*Preliminary presentation of the qualification work*	*02.12.2024*	
11	*Correction of remarks, external review of the qualification work*	*04.12.2024*	
12	*Presentation of the finished stitched qualification work to the department*	*05.12.2024*	
13	*Public presentation of the qualification work*	*According to the work schedule of the EB*	

6. Date of the task issue «15» December 2023

7. Scientific supervisor of the qualification work _____ V. Kulazhenko
(signature) *(name)*

8. Director of the educational programme _____ A. Roskladka
(signature) *(name)*

9. The task was accepted by the student _____ T. Panchenko
(signature) *(name)*

10. Review of the scientific supervisor of the qualification work

Scientific supervisor of the qualification work_____ _____
(signature) *(date)*

Preliminary defence_____ _____
(signatures of th emembers of the preliminary defense board) *(date)*

11. Conclusion on the qualification work

The qualification work of the student T. Panchenko
(name)
can be admitted to the defense in the examination board.

Director of the educational programme _____ A.Roskladka.
(signature) *(name)*

Head of the department_____A.Roskladka
(signature) *(name)*

«_____»_____2024

Abstract

This thesis investigates the process of selling Volkswagen Group cars and develops a system for analysing and forecasting their sales using modern data analysis methods. The main tools for the study were the Python programming language and the Tableau software. Based on the collected data, analytical reports were created to assess sales trends, and machine learning models were implemented to predict future sales volumes. Particular attention was paid to visualising the results, which allows for effective data interpretation and informed decision-making. Python programming language libraries such as pandas, scikit-learn, and matplotlib were used to provide flexibility and power in data analysis.

Keywords: Volkswagen Group, car sales, data analysis, machine learning, Python, Tableau, analytical reports, forecasting.

Анотація

У даній роботі досліджується процес продажу автомобілів Volkswagen Group та розробляється система для аналізу і прогнозування їх продажів з використанням сучасних методів аналізу даних. Основними інструментами для реалізації дослідження стали мови програмування «Python», а також програмне забезпечення «Tableau». На основі зібраних даних було створено аналітичні звіти, що дозволяють оцінити тенденції продажу, а також реалізовано моделі машинного навчання для прогнозування майбутніх обсягів продажів. Особливу увагу приділено візуалізації результатів, що дозволяє ефективно інтерпретувати дані та приймати обґрунтовані рішення. Використані бібліотеки мови програмування «Python», такі як «pandas», «scikit-learn» та «matplotlib», забезпечили гнучкість та потужність у аналізі даних.

Ключові слова: Volkswagen Group, продажі автомобілів, аналіз даних, машинне навчання, Python, Tableau, аналітичні звіти, прогнозування.

CONTENT

INTRODUCTION

Over the past decades, rapid advances in technology have had a significant impact on the business environment, opening up opportunities for deeper data analysis and more accurate forecasting. This is especially true in sales, where effective management and planning depend on the ability to anticipate market changes. The automotive industry, including the Volkswagen Group, needs innovative approaches to accurately forecast demand for its products, which allows it to optimise production and adapt business strategies to rapidly changing market conditions.

Using statistical methods such as ARIMA (AutoRegressive Integrated Moving Average), a reliable sales forecasting model can be created that ensures high accuracy and allows the company to make informed decisions based on the forecast data. The ARIMA method, which is based on time series analysis, is particularly effective for processing large amounts of data, allowing to take into account both trends and seasonal fluctuations.

Relevance of the work: in the current conditions of dynamic development of the automotive industry, the Volkswagen Group faces challenges related to changes in consumer priorities, increased competition, and the introduction of environmental standards. Sales analysis and forecasting are critical to the effective management of production, logistics and marketing. A system approach to sales analysis allows a company to adapt its strategies to a rapidly changing market, and the introduction of modern data analysis tools ensures improved accuracy of management decision-making.

The purpose of the study is to conduct a system analysis of Volkswagen Group's car sales to identify key trends and build forecast model.

To achieve this goal, the following tasks need to be completed:

1. Study modern methods of system analysis, including statistical modeling techniques, for application to the analysis of Volkswagen Group car sales.

2. Develop an analytical report based on the collected data.

3. Analyze the current situation in the car sales market, identifying key trends and features of demand for Volkswagen Group products.

4. Study a sales forecasting model to analyze and build reliable forecasts.

5. Collect the necessary data on Volkswagen Group car sales and explore the application of machine learning algorithms for sales forecasting.

Object of research: Volkswagen Group car sales.

Subject of research: financial indicators of sales of the Volkswagen Group for 2006 - 2023.

The structure of the final qualification work:

The work consists of an introduction, three chapters, conclusions, and a list of references. The total volume of the work is 55 pages, contains 17 figures illustrating the main stages of the analysis, and a list of references, which includes 20 items.

CHAPTER 1. THE VOLKSWAGEN GROUP

1.1 Development and structure of the Group

The Volkswagen Group, commonly referred to as the VW Group, is a conglomerate of automotive companies with Volkswagen AG as its parent company. The group is headquartered in Wolfsburg, Germany. Ownership of the VW Group was clarified in 2012. Before that, Porsche SE held a 50.73% stake in Volkswagen AG, despite Volkswagen AG owning 100% of Porsche GmbH. Currently, Porsche is wholly owned by the Volkswagen Group.

Herbert Diess has been Chairman of Volkswagen AG and concurrently Chairman of the Management Board of Porsche SE since April 2018.

Herbert Diess has served as Chairman of Volkswagen AG and Chairman of the Management Board of Porsche SE since April 2018.

The Volkswagen Group consists of 342 companies, many of which provide services related to car production rather than manufacturing vehicles directly. The VW Group has frequently been the world's largest automaker, competing with General Motors, Toyota, and Renault-Nissan.

In December 2009, the Volkswagen Group entered into a partnership with Japan's Suzuki to develop environmentally friendly vehicles. As part of this collaboration, Volkswagen acquired a 20% stake in Suzuki. However, the alliance dissolved in late 2011.The corporate structure of the VW Group:

- Volkswagen. Focuses on passenger car production and operates under the direct management of Volkswagen AG.

- Audi. Acquired in 1964 from Daimler and is the last remaining member of the former Auto Union group.

- NSU Motorenwerke. Owned by the VW Group since 1969, it operates under the Audi division. It ceased functioning as an independent brand in 1977.

- SEAT. Volkswagen acquired a 53% controlling stake in SEAT in 1986 through a deal with the Spanish government. By 1990, the VW Group became the sole owner, holding 99.99% of the Spanish automaker's shares.

- Škoda. VW Group has had exclusive rights to manage this Czech automaker since 1991.

- Volkswagen Commercial Vehicles. The company produces commercial vehicles: vans, buses, etc.[1]

Fig. 1.1. Corporate structure of the VW Group. Source: [11]

1.2 Global markets and production centres

The Volkswagen Group's production sites are strategically located around the world to ensure efficient production and delivery of vehicles. The international innovation centres serve as catalysts for pioneering research and new technologies, working closely with the Group brands to turn new ideas into reality - shaping the vehicles of the future.

In addition, battery factories are being set up to meet the growing demand for electric vehicles. PowerCo, the battery company of the Volkswagen Group, has already identified three locations for battery cell production: Salzgitter (Germany), Valencia (Spain) and St Thomas (Canada).

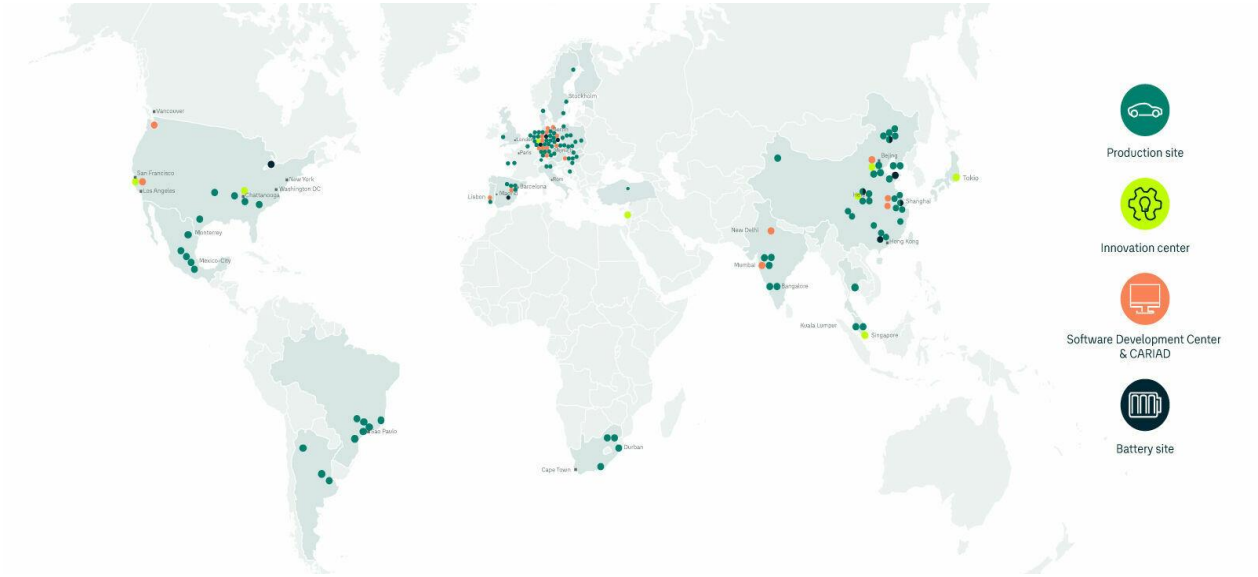


Figure 1.2. Global markets and production centres of the VW Group in the world.

Source: [12]

Europe

The company operates ten core brands from five European countries, organised into brand groups:

Brand Group Core with Volkswagen (Wolfsburg, Germany), Volkswagen Commercial Vehicles (Hanover, Germany), SKODA (Mladá Boleslav, Czech Republic), SEAT and CUPRA (Barcelona, Spain).

Brand Group Progressive with Audi (Ingolstadt, Germany), Lamborghini (Sant'Agata Bolognese, Italy), Bentley (Crewe, UK), Ducati (Bologna, Italy).

Brand Group Sport Luxury with Porsche (Stuttgart-Zuffenhausen, Germany).

Europe plays a central role in the global R&D strategy of the automotive industry. Numerous innovation centres and production facilities have been established here, making a significant contribution to the development of new technologies and vehicles worldwide.[12]

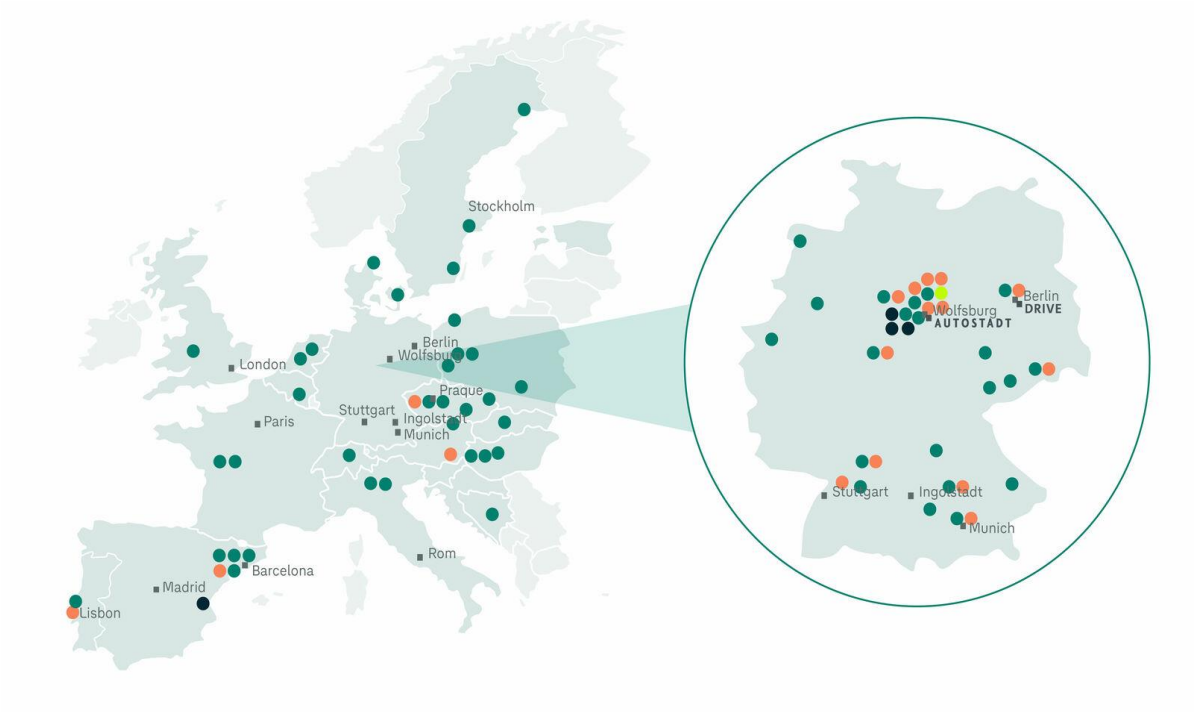
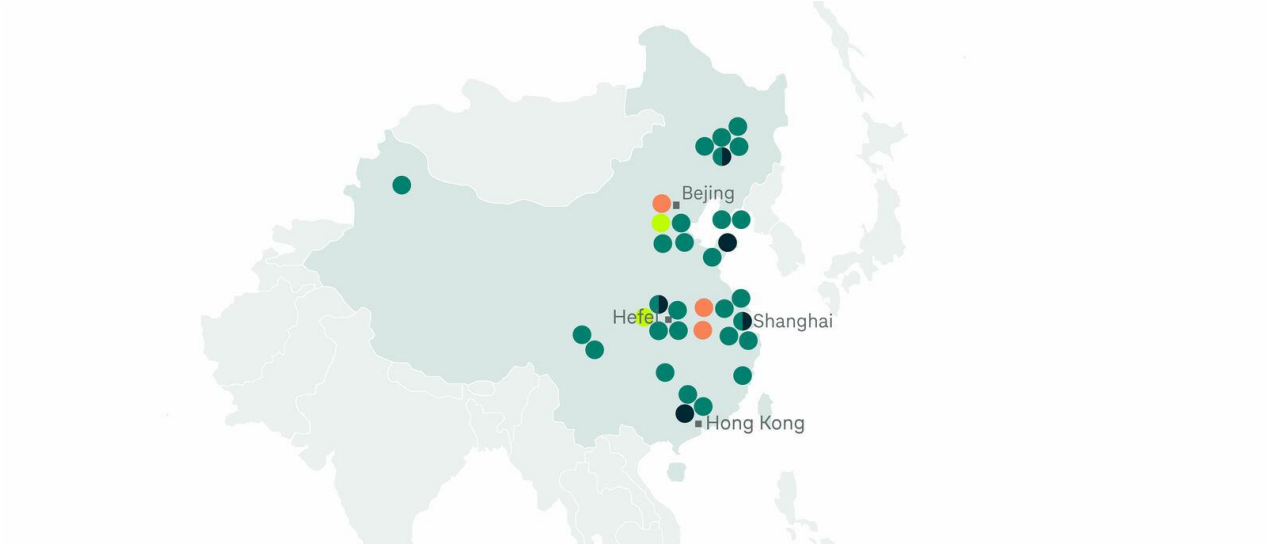


Figure 1.2. Global markets and production centres of the VW Group in Europe. Source: [12]

For the Volkswagen Group, China is a key emerging market and a driving force for innovation in the areas of electric mobility and connected automated driving, which is crucial to the company's transformation in the automotive industry. The Group takes the ‘In China for China’ approach, emphasising accelerated innovation and promoting technological localisation to meet customer needs. In line with the strategy, the Group is establishing the Group's largest procurement and development centre outside Germany in China together with Volkswagen China Technology Company (VCTC) in Hefei to adapt products to the needs of Chinese customers more quickly. Flexible working methods, an efficient network of different areas and the use of state-of-the-art technology will reduce the development time for new products by 30 per cent. Thus,adapting to market dynamics and strengthening competitiveness.[12]



.

Figure 1.3. Global markets and production centres of the VW Group in China. Source: [12]

The Volkswagen Group has initiated a special development programme for North America with a commitment to invest heavily in the region. This includes the expansion of PowerCo cell production from Europe to Canada. The first battery cell factory in North America will begin operations in St Thomas, Ontario. In addition, the company is presenting one of the broadest portfolios of all-electric vehicles in the United States. An electric assembly line has already been opened in Chattanooga, Tennessee, and there are plans to significantly expand the charging infrastructure across the continent with the help of Electrify America. The Group is also electrifying the iconic American brand Scout by setting up an electric vehicle plant near Columbia, South Carolina.[12]

Figure 1.4. Global markets and production centres of the VW Group in North America. Source: [12]

The overall strategy of the Volkswagen Group is focused on profitability, electrification and a more balanced global footprint with an emphasis on innovation, sustainability and customer focus to shape the mobility of the future.

Volkswagen has no production facilities in Ukraine. However, the company actively cooperates with local businesses to ensure the sale and service of its vehicles.[12]

The Volkswagen Group is represented in Ukraine through a network of authorised dealers that sell new cars. Among them:

- Avtodom Atlant: an authorised Volkswagen dealer in Kyiv offering a wide range of models, including T-Roc, Tiguan, Touareg, Golf, Arteon, Caddy, T6.1, Multivan, Caravelle and Crafter.[19]

- Jerman Centre: an authorised Volkswagen dealer offering new models and special offers for vehicles older than 4 years.

- AvtoSoyuz: the first Volkswagen concept car centre in Kyiv, which consistently ranks at the top of the Volkswagen sales ranking in Ukraine. [18]

To ensure high-quality service for Volkswagen owners, official service centres in Ukraine meet the brand's high standards. For example, Avtodom Atlant on Dniprovska embankment provides a full range of services, including warranty service, painting and bodywork, and offers original spare parts.

In addition, Volkswagen in Ukraine offers the NORA Service Centre programme, which involves cooperation with independent service companies that service Volkswagen vehicles, providing high quality service and access to original spare parts.[20]

Volkswagen also provides car owners in Ukraine with access to various services, such as the Car-Net programme, which provides convenient access to car functions via smartphone or computer.

Thus, although Volkswagen does not have its own production facilities in Ukraine, its presence in the market is ensured through a network of authorised dealers and service centres, which contributes to the development of the country's automotive industry and meets the needs of Ukrainian consumers.[29]

1.3 Strategic structure of brands and technologies

The Volkswagen Group comprises ten powerful brands organised into three key groups: Core, Progressive and Sport Luxury. Each of the brands has a unique focus, which ensures coverage of different market segments and a high level of synergy in achieving the company's strategic goals.

The Core Group focuses on the mass market and affordability. The core Volkswagen brand offers cars for everyday use with a focus on reliability and innovation. SEAT attracts a younger audience by offering stylish and affordable models, while Škoda is known for its economy and quality. Volkswagen Commercial Vehicles specialises in commercial vehicles, while CUPRA, as SEAT's sports brand, focuses on performance models for speed enthusiasts.

The Progressive Group comprises premium brands that use advanced technologies for demanding consumers. Audi is a leader in the premium segment, known for its elegant design and technological innovations in electric and autonomous vehicles. Bentley offers luxury cars for very wealthy customers, while Lamborghini focuses on supercars with impressive design and performance. Ducati, an Italian motorcycle brand, is known for its sporty models.

The Sport & Luxury Group is represented by the Porsche brand, which is synonymous with luxury sports cars and is actively introducing innovative solutions in electric vehicles, such as the Taycan.

In addition to its core brands, the Volkswagen Group is investing in strategic areas to strengthen its influence in global mobility.

Volkswagen Financial Services offers a wide range of services, including lending, leasing and insurance, which makes cars more affordable for consumers and contributes to the company's business growth.[25]

Volkswagen Group Technology is responsible for technological innovation, particularly in electric mobility, autonomous driving systems and intelligent digital solutions for traffic and safety optimisation.

TRATON SE specialises in the production of heavy commercial vehicles such as trucks and buses and is actively developing electric models through the MAN and Scania brands to reduce environmental impact and increase logistics efficiency. MOIA focuses on new forms of urban mobility, such as electric car sharing, contributing to the sustainable development of urban transport infrastructure.[26]

NEW AUTO's strategy defines electric mobility as a priority area. The Technology Group is responsible for the production of batteries and the expansion of charging infrastructure. Through PowerCo SE, the company is setting up global production of battery cells and is developing a network of fast-charging stations through the Elli, Electrify America and CAMS brands. The Volkswagen Group's technologies are based on four main areas: Battery, Charging & Energy, Components, and Platform Business.

1. Group Components manufactures key technical components (electric motors, transmissions) to improve efficiency and sustainability.

2. Platform Business works with external partners such as Ford and Mahindra to share platforms and components.

3. Charging & Energy, through Elli, Electrify America and CAMS, develops charging infrastructure, provides green energy and charging services.

4. Battery, through PowerCo SE, is responsible for global battery production, the development of standard cells for electric vehicles and the support of innovative technologies.

This structure enables the Volkswagen Group to take a holistic approach to technology development and integrate innovations to create environmentally sustainable and efficient transport solutions for the future.[20]

Conclusions to Chapter 1

The Volkswagen Group is one of the largest automotive corporations in the world, uniting 10 major brands, each of which has its own unique focus. Volkswagen is engaged in the production of mass-produced cars, Audi specializes in premium models, Škoda is known as an economy-class brand, and Bentley, Lamborghini and Porsche represent the luxury and sports car segment. Ducati, in turn, focuses on the production of motorcycles, and Volkswagen Commercial Vehicles manufactures commercial vehicles.

The group's production and research facilities are located around the world. Europe is a key region with the largest concentration of factories and innovation centers. In China, the company is actively developing localized technologies for electric vehicles, and in North America, attention is focused on the production of batteries and expanding the range of electric vehicles. Thanks to this geography, Volkswagen ensures its global presence and effectively adapts to the requirements of regional markets.

Although Volkswagen does not have its own production facilities in Ukraine, the company is represented by a network of official dealers and service centers. This allows us to ensure high-quality car service and the availability of new models on the Ukrainian market.

CHAPTER 2. TOOLS FOR SYSTEM DATA ANALYSIS AND VISUALISATION

2.1 Tools for data processing

Data processing is an essential step in the information analysis process, as it prepares data for further analysis, modeling, and visualization. There are many tools used for data processing, but among the most commonly used in business intelligence are Python, R, and SQL. Each of them has its own features and benefits that make them useful in different data analysis contexts.

Python is one of the most popular programming languages in the data analytics industry due to its simplicity, versatility, and a large number of libraries that facilitate data processing. The main libraries used for data processing in Python are:

- Pandas: This is the main tool for working with data in table format. Pandas makes it easy to read, clean, transform, and aggregate data. It provides powerful methods for manipulating data, such as joining tables, filtering, grouping, and working with time series.

- NumPy: A library for processing multidimensional arrays and matrices, which also includes a large number of high-level math functions. NumPy is the basis for many other libraries in Python, such as Pandas, and is used to process large amounts of data quickly and efficiently.

- SciPy: This library extends the functionality of NumPy by adding additional tools for numerical calculations such as integration, optimization, and signal processing. SciPy is especially useful for complex math problems and scientific calculations.

- Dask: This is a library for parallel processing of large data sets that allows you to work with data that does not fit into computer memory. Dask provides distributed data processing and integrates with Pandas and NumPy.

Features of Python:

- Simplicity and readability of the code: Python is known for its simplicity and clarity of syntax. The language is designed to make code easy to read and write. It uses a clear and understandable structure with indentation to define blocks of code instead of curly brackets, which makes it easier to understand the structure of the program. This is

suitable for both beginners and experienced developers, as it reduces syntax errors and makes debugging easier.

- Versatility: Python is a versatile programming language widely applied across fields such as web development, task automation, data science, and machine learning. It supports multiple programming paradigms, including object-oriented, functional, and procedural programming. This allows Python to be an effective tool in many contexts and for different types of projects.

- A large number of libraries and frameworks: One of the main advantages of Python is the huge number of available libraries and frameworks. The Python Package Index (PyPI) contains hundreds of thousands of libraries that cover almost any aspect of programming: data processing (Pandas, NumPy), visualisation (Matplotlib, Seaborn), web development (Django, Flask), machine learning (Scikit-learn, TensorFlow, PyTorch), and many others. This provides a wide range of tools for quick problem solving.

- Interpreted language: Python is an interpreted language, which means that the code is executed line by line without the need to compile to machine code. This allows developers to develop and test code faster because there is no need for the compilation process. Interpretability also simplifies debugging, as errors can be detected and corrected at runtime.

- Support for diverse programming paradigms: Python accommodates various programming paradigms, enabling developers to select the most suitable approach for a given task. These include object-oriented programming (OOP) for organizing code into classes and objects, functional programming for writing pure functions without side effects, and procedural programming for structuring code into functions and procedures.

- Platform independence: Python is a platform-independent programming language, allowing the same application to run on multiple operating systems without code modification. This flexibility is achieved because Python interprets code instead of compiling it into machine-specific instructions. As a result, Python programs are compatible with Windows, macOS, Linux, and other platforms.

- Active community and rich documentation: Python has a large and active developer community that is constantly working on new libraries, frameworks, and tools. This provides access to a wealth of resources and support. Python's documentation is detailed and well-structured, helping developers quickly find the information they need and solve problems they may encounter.

- Dynamic typing: In Python, variable types are assigned at runtime, rather than during compile time. This makes development more flexible because you don't have to explicitly define variable types. However, dynamic typing can also lead to type errors that become apparent only at runtime.

- Built-in testing tools: Python has built-in tools for writing and executing tests, such as the unittest module, which is part of the standard Python library. This allows developers to write automated tests to check the correctness of their code. There are also other tools, such as pytest, that provide additional testing capabilities and facilitate the debugging process.

- Widespread use in science and education: Python is widely regarded as one of the most popular programming languages in academia and research, thanks to its simplicity and versatility. Many universities and educational institutions use Python for teaching programming and conducting scientific research, largely due to its extensive collection of libraries for data analysis and mathematical computing.

- Supports integration with other languages: Python integrates easily with other programming languages such as C, C++, and Java. This allows you to use Python for high-level programming and at the same time use high-speed libraries written in other languages for computing and productivity. Python also supports integration with libraries and APIs written in other languages, which extends its capabilities.[27]

R is a programming language and environment created for statistical computing and graphical visualization, specifically tailored for data analysis. It serves as a robust tool for conducting statistical analyses, processing data, and generating graphs. Developed by Ross Ihaka and Robert Gentleman in the 1990s, R was introduced as a free alternative to commercial statistical software. Its primary objective is to offer users efficient and accessible tools for statistical analysis and data visualization.[24]

R has a large set of built-in functions and capabilities that allow you to perform complex statistical calculations, modelling and data visualisation. In addition, it supports extensions through additional libraries and packages, making it very flexible and adaptable to different needs.

The main R packages and libraries:

- stats: This is the main package included in the basic installation of R. It provides functions for performing statistical analyses such as regression analysis (lm() for linear models), hypothesis testing (t.test() for t-tests), and time series analysis (functions for decomposing and predicting time series).

- ggplot2: One of the most popular packages for graphical data visualisation. It implements the principles of graphing grammar, which allows you to create complex graphs with high quality visualisation. It can be used to create a variety of graph types, including histograms, scatter plots, line graphs, and more.

- dplyr: A data manipulation package that provides a convenient and efficient way to filter, sort, aggregate and transform data. dplyr offers a clear syntax for performing operations on data tables.

- tidyr: Complements dplyr by specialising in data organisation. The package offers functions for ‘unpacking’ and ‘assembling’ data, which is useful for transforming data into formats that are easy to analyse.

- MASS: A package for applied statistics that contains tools for working with various statistical models and methods. It includes functions for regression, clustering, and analysis of variance, as well as datasets for practical problems.

- xts: A package designed for processing and analyzing time series data. It offers tools for storing, handling, and aggregating data in time series format, which is essential for financial and econometric analysis.

- Caret: A machine learning package that provides a user-friendly interface for building and evaluating machine learning models. It supports many classification and regression algorithms and includes tools for parameter fitting and model evaluation.

- sf: A package for working with geospatial data. It implements the concept of spatial features and provides tools for analysing and visualising geographic data.

- shiny: A package for creating interactive web applications in R. With shiny, you can develop web applications to visualise data and provide an interactive user interface.

SQL (Structured Query Language) is a standard language for interacting with relational databases. SQL allows you to create, modify, retrieve and manage data in relational database management systems (DBMS). SQL provides the ability to:

- Create and modify the database structure: define tables, their columns, data types, and constraints.

- Data manipulation: adding new records, updating existing records, deleting data, and executing queries to retrieve information.

- Access control: control access rights for different users and groups.

- Ensure data integrity: implement constraints to ensure data is correct and consistent, and manage transactions to ensure the integrity of transactions.

The main database management systems (DBMS) that support SQL:

- MySQL is an open source relational database widely used for data management in web applications and medium-sized projects.

- PostgreSQL is a powerful open-source relational database that supports advanced data manipulation functions and provides high flexibility.

- Microsoft SQL Server is a commercial relational database from Microsoft that provides extensive business intelligence capabilities and integrates with other Microsoft products.

- Oracle Database - a commercial relational database from Oracle, known for its scalability and reliability, often used in large corporate systems.

- SQLite is a lightweight relational database that can be built into applications without the need for a separate server environment. It is used in mobile applications and embedded systems.

- MariaDB is a fork of MySQL created to provide open source and performance improvements, compatible with MySQL and often used as an alternative.

2.2. Data visualization tools (Tableau, Power BI, Looker Studio)

Google Looker Studio is a powerful data analysis and visualization platform that makes it easy to create and understand reports. The main advantage of the service is that it allows you to connect to more than 800 data sources, such as databases, spreadsheets, Google Analytics, Google Ads, and more, and then use the web editor to drag and drop objects, customize properties, and create charts and graphs that meet your needs. Most of the features of the service are easy to use, and you can share them with colleagues, work in real time, or embed your reports on any web page. You can also access a library of report templates that will help you visualize your data in minutes.

Looker Studio provides powerful data visualization tools that make it easy to create beautiful interactive graphs, charts, and tables from any data by simply dragging and dropping objects onto the canvas. The service's multifunctional visual editor allows you to customize the appearance of graphs, choose colors, fonts, and styles to create unique visualizations that best suit your needs. And a large library of templates allows you to visualize data in minutes. You can add up to 10 widgets to one page at a time and create a report for any number of pages. To prevent graphs or charts from being too static and boring, you can make them interactive.

The main phases of working with Looker Studio:

- Import or connect data to Looker Studio.
- Prepare data for use and configure the source.
- Create an empty dashboard (report).
- Adding a connected data source.
- Creating visualisations / parameters / calculation fields based on this data (i.e. filling the dashboard).
- Checking the finished dashboard from the user's point of view.
- Providing access to the dashboard to other users.[3]

Tableau is a powerful data visualization tool that is rapidly growing and used in the business intelligence industry. It helps simplify raw data into a understandable format.

Tableau has more than one product, and the real wow effect can be achieved by combining them in your work:

- Tableau Desktop is a tool for creating analytical dashboards and visualisations in a few clicks;

- Tableau Prep is a tool for collecting and processing data. It allows you to prepare data in a format convenient for analysis;

- Tableau Server - a solution that allows you to place analytics in a browser tab and make it available to any user, such as your colleagues;

- Tableau Cloud - duplicates the capabilities of Tableau Server, without the need to maintain your own server. It does not require installation, special software or additional settings;

- Tableau Public is a free, shortened version of Tableau Desktop.

The Tableau concept is based on three principles:

- Speed and simplicity of the interface. The software uses the drag-and-drop method of working with elements. They are easy and convenient to move, making this tool intuitive.

- Versatility. Tableau can connect to more than 60 data sources - from file systems, PDF, JSON to databases and web interfaces. In four years, I haven't encountered a situation where Tableau didn't support a format. You can also mix different formats at the same time. For example, you can download and combine a database, a Google Sheets table, and JSON in one report.

- Constant access. All data is quickly updated and the team has access to real-time analytics.

Working in Tableau is based on interacting with data in a workbook (Book) using sheets, similar to Microsoft Excel. A workbook contains three types of tabs:

- Worksheet. This is where you create visualisations (charts, graphs). It can contain only one visualisation.

- Dashboard is a collection of several worksheets. It allows you to visually collect several charts and graphs in one place at once, adding context, legends, filters, and building a work logic. You need it if you want to see the dynamics, collect data on different segments, and compare them with each other.

- Story is a sequence of worksheets or dashboards in the form of a presentation. You can add captions and comments.[6]

Microsoft Power BI is a powerful tool for data analysis, information visualization, and team collaboration. With Power BI, you can connect to multiple data sources, create interactive reports and dashboards, use built-in artificial intelligence capabilities, integrate with Excel and other Microsoft applications, and access your data from any device. Here are some of the benefits and capabilities of Microsoft Power BI:

- Quick and easy connection to data. Power BI has many built-in connectors for various services and databases, so you can pull in the right data set from different sources, link them together, and build consolidated reports and charts in minutes1. You can also use Power Query to transform and cleanse your data to suit your needs.

- Flexible and powerful data visualisation. Power BI lets you choose from hundreds of visualisation types to suit different purposes and audiences. You can customise the appearance, colours, filters, axes, legends, and other settings of your graphs, tables, maps, and other elements. You can also create your own visualisations or use pre-built solutions from the Power BI Visualisation Store.

- Interactive and dynamic reports and dashboards. Power BI allows you to create reports and dashboards that respond to user actions, such as selecting, filtering, sorting, zooming, and more. You can add segmenters, buttons, scroll bars, and other controls to make your reports and dashboards more interactive and intuitive. You can also use the Q&A feature to provide natural language visualisations, such as ‘show sales by region as a map’.

- Built-in artificial intelligence capabilities. Power BI has a number of features that allow you to use artificial intelligence to analyse your data, such as key influencers, value distribution, anomaly detection, forecasting, time series decomposition, and more. You can also integrate Power BI with Azure Machine Learning to build and use your own machine learning models.

- Tight integration with Excel and other Microsoft applications. Power BI is highly compatible with Excel, making it easy to import and export data, reports, and visualisations between the two applications. You can also use Power BI to analyse data

from other Microsoft applications, such as SharePoint, Dynamics 365, Teams, Outlook, and more.

- Accessibility and mobility. Power BI makes it easy and fast to provide access to analytical information at various levels to employees, partners, and customers. You can publish your reports and dashboards online, embed them in your websites and apps, or send them via email. You can also use the Power BI mobile app to view and edit your data from any device.[4]

2.3 Creating an analytical report using Tableau

This chapter explores the process of generating an analytical report using contemporary data analysis tools like Tableau.

It focuses on the following steps:

- Identifying and preparing data sources for analysis.

- Loading and processing data in Tableau, including cleaning and transformation.

- Creating interactive visualisations and generating a report that provides in-depth analysis and helps with decision-making.

The input for the work is a dataset ready for analysis, which is provided from a specialised resource. It includes data on the Volkswagen Group's car sales revenue from fiscal years 2015 to 2023 by major subsidiaries or brands (in billions of euros).

The data was obtained from the official website [5] of the Volkswagen Group and covers the period from 2015 to 2023. It includes information on sales volumes for various brands within the group, including Volkswagen, Audi, Porsche, Škoda and others. This data allows us to analyse sales dynamics over a nine-year period. The data is organized in a table, with each row representing the sales of a specific brand in a given year. The key columns include:

Brand: The name of the brand (e.g. Volkswagen, Audi, Porsche).

Year: The period for which the data is displayed (2015-2023).

Sales: Total number of units sold or total sales for the year.

The data contains information about the following brands belonging to the Volkswagen Group:

- Volkswagen Passenger Cars (VW Passenger Cars): The main brand that produces passenger cars for the general market.

- MAN Commercial Vehicles: Truck and bus manufacturer specialising in commercial vehicles.

- Audi Progressive Brand Group: Premium brand known for its technological and design innovations.

- Porsche: Luxury sports car manufacturer known for its high performance and design.

- Škoda: A brand focused on affordable and reliable cars for the mass market.

- Scania: A truck and bus manufacturer known for its reliability and durability.

- TRATON Commercial Vehicles: A group that includes commercial vehicle brands such as MAN and Scania.

- Volkswagen Commercial Vehicles (VW Commercial Vehicles): Manufacturer of light commercial vehicles such as vans and pick-up trucks.

- SEAT: A Spanish brand that produces passenger cars with a focus on youthful design and dynamic performance.

- Bentley: British luxury brand specialising in the production of exclusive cars with a focus on comfort and luxury.

Before starting the analysis, you should make sure that all data has been transferred correctly and open the field (Fig. 2.1) on the corresponding tab.

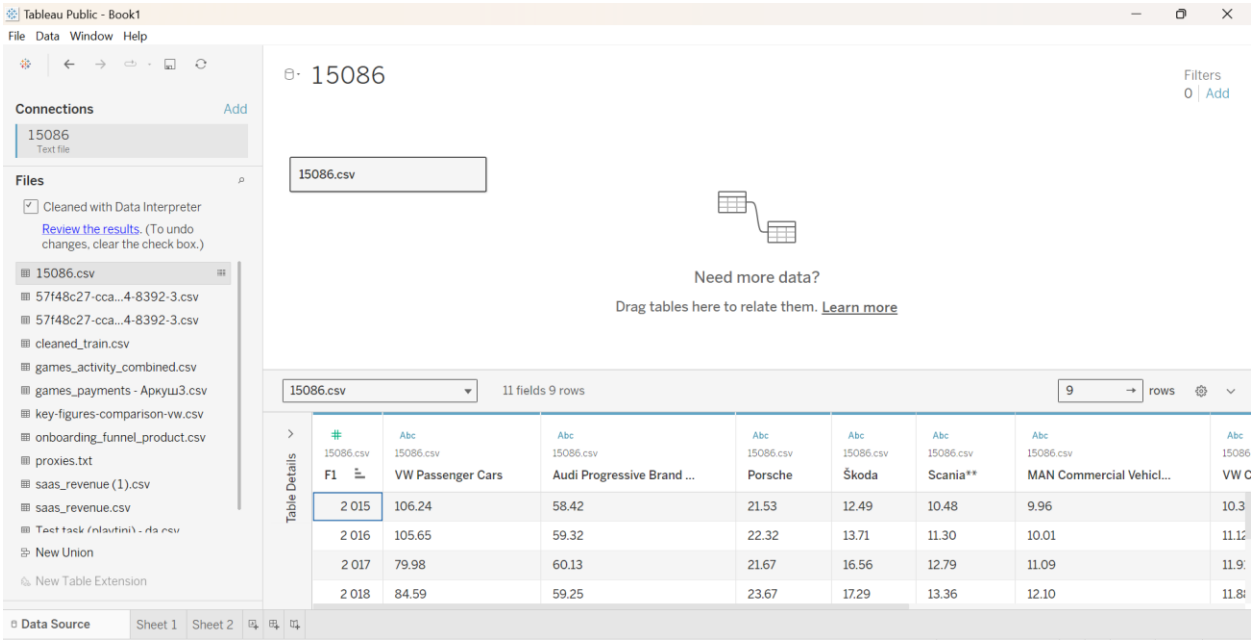


Fig.2.1 Data verification [Source: author's development]

Before you start working with your data, Tableau offers the Data Interpreter feature. It helps you automatically prepare your data for analysis.

Data Interpreter is a tool built into Tableau that automatically cleans and structures data, especially when importing from spreadsheets (such as Excel). It helps you extract the information you need from such sources, which often contain redundant or incorrectly structured data.

Data Interpreter makes data preparation much easier by reducing the time spent on manual cleaning and formatting. As a result, you get a clean, well-structured data set ready for further analysis and reporting in Tableau.

This action is especially useful when working with large or complex spreadsheets, where data may be arranged in a chaotic manner or contain unnecessary information.[28]

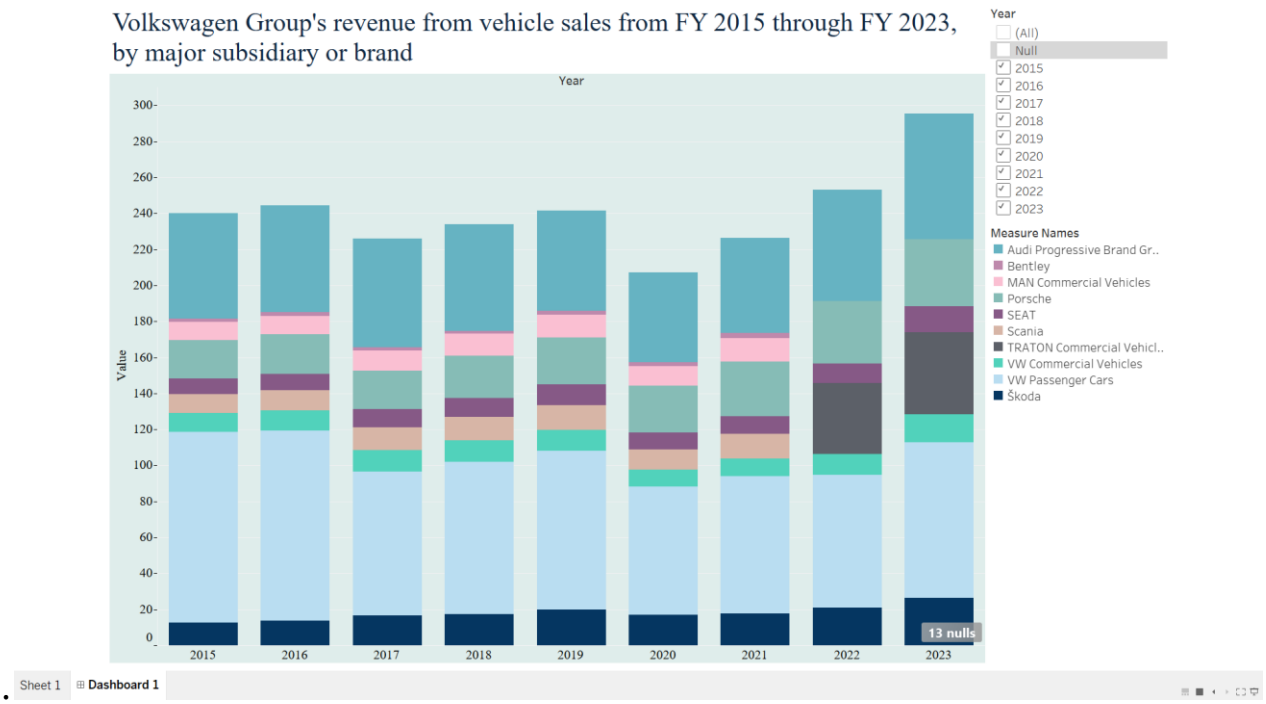


Fig. 2.2 “Dynamics of financial indicators” [Source: author's development]

The figure shows a diagram of the financial performance of car brands by year. The graph shows that the largest sales volume is observed in the brands that dominate the market. These are Audi and VW passenger cars.

The dominance of the Audi and Volkswagen Passenger Cars brands in the Volkswagen Group is explained by a number of factors that provide them with leading positions in the automotive industry. First of all, these are different market positioning and a wide target audience. Volkswagen has a significant presence in Europe, China and the USA, which makes it the largest in terms of sales among the group's brands. Audi, for its part, operates in the premium segment and offers cars for wealthy customers who value technology, design and status. This brand is one of the market leaders in Europe and North America and is actively introducing electric cars, which adds weight to the growing popularity of electric cars.

In the 2023 financial year, Volkswagen Passenger Cars introduced the updated Tiguan to the market and launched the fully electric ID.7 sedan. The operating performance significantly improved compared to the previous year. The main highlights was the debut of the third-generation Tiguan, which features a modernized digital cockpit, an infotainment screen, a head-up display, and intuitive controls inspired by customer feedback. The new plug-in hybrid engines offer an electric range of up to 100 kilometers

and support fast DC charging. Additionally, the advanced DCC Pro adaptive chassis system, typically found in premium vehicles, enhances driving comfort and agility. New generations of the Passat and Touareg were also launched during the year.

In April 2023, Volkswagen unveiled the ID.7, its first global all-electric model in the upper middle-class segment. This sedan combines long range, rapid charging, a spacious interior, and intuitive operation, setting a new benchmark for premium electric vehicles within the brand. Among its innovative features is a panoramic sunroof with smart glass technology that can switch between opaque and transparent with touch controls. Volkswagen also revealed two new electric concept cars: the compact ID. 2all and the sporty ID. GTI Concept.

During the year, Volkswagen Passenger Cars sold 4.9 million vehicles globally, a 6.7% increase compared to 2022. The strongest growth was recorded in Western Europe (+14.7%), North America (+17.0%), and South America (+15.9%). The brand produced 4.8 million vehicles worldwide in 2023, with Volkswagen do Brasil achieving a milestone of 25 million vehicles manufactured in its 70 years of operation.

The brand's sales revenue increased by 17.1% year-on-year to €86.4 billion, with operating profit rising by €2.6 billion. Higher sales volumes, better pricing, and reduced fixed costs positively impacted financial results, despite rising product costs and the deconsolidation of Volkswagen Group Rus.

The Progressive brand group, which includes Audi, Bentley, Lamborghini, and Ducati, performed strongly in a challenging economic environment. Audi introduced its first fully electric model based on the Premium Platform Electric (PPE), the Q6 e-tron, alongside the launch of the Q8 e-tron, the successor to the pioneering Audi e-tron. Updates were also introduced to the A6, A7, and Q8 models, enhancing their design and technology. Bentley celebrated the 20th anniversary of the Continental GT with the exclusive Le Mans Collection, while Lamborghini debuted its first mass-produced hybrid model, the Revuelto, marking its move towards electrification. Ducati also reported robust performance.

The Progressive brand group delivered 1.6 million vehicles globally in 2023, with significant growth in Western Europe (+19.0%), Central and Eastern Europe (+25.3%),

and North America (+21.7%). Audi sold 1.1 million units, driven by strong demand for the Q5, A3, and fully electric Q4 e-tron, while Chinese joint ventures added 599,000 locally produced Audi vehicles. Lamborghini delivered 10,600 cars (+19.4%), Bentley sold 15,500 vehicles, and Ducati recorded sales of 58,400 motorcycles.

In terms of production, Audi manufactured 1.7 million vehicles worldwide, Bentley produced 16,400 cars, Lamborghini assembled 9,900 units, and Ducati built 70,400 motorcycles. The Progressive brand group's revenue grew to €61.8 billion in 2023, driven by increased volumes at Audi and strong contributions from Bentley, Lamborghini, and Ducati.

Compared to other brands of the group such as Škoda, SEAT, Porsche or Lamborghini, Volkswagen and Audi benefit from the breadth of the model range, greater investment in technology and a global presence. Škoda and SEAT, for example, are focused mainly on the European market and offer more budget-friendly cars, which limits their global market share. Porsche and Lamborghini, on the other hand, represent luxury segments with limited production volumes, which also limits their overall sales.

One of the key strengths of the Volkswagen Group is its innovative approach to production and electrification. Audi and Volkswagen are actively developing the direction of electric cars, which helps them meet the requirements of environmental standards and new consumer trends. In addition, the Volkswagen Group has a strong production base and ample opportunities for global scaling of its products. However, a weakness may be the slower pace of innovation in lower-end brands such as SEAT and Škoda, which reduces their position compared to competitors.

The Volkswagen Group's competitors, especially for the Audi and Volkswagen brands, are automotive giants such as Toyota, General Motors, Ford, BMW and Mercedes-Benz. Toyota, for example, is a leader in the mass-market segment due to its reliability, environmental technologies (hybrids and electric vehicles), as well as strong positions in the Asian and US markets. In the premium segment, Audi's main competitors are BMW and Mercedes-Benz. These brands also focus on innovation, electrification and luxury and often outpace Audi in the pace of implementation of new technologies, especially in the electric and autonomous driving segments.[10]

Table 2.1

Top 100 Brands by Car Sales as of 2023[10]

Brand	**2023 Car Sales**	**2023 Market Share %**
Toyota	8,569,973	11.07%
Volkswagen	4,965,621	6.41%
Honda	3,774,161	4.87%
Ford	3,729,817	4.82%
Hyundai	3,540,787	4.57%
Nissan	2,975,048	3.84%
Suzuki	2,916,149	3.77%
Kia	2,731,370	3.53%
Chevrolet	2,692,391	3.48%
BYD	2,683,743	3.47%
BMW	2,101,487	2.71%
Mercedes-Benz	2,063,820	2.67%
Audi	1,772,892	2.29%
Tesla	1,770,354	2.29%

Volkswagen and Audi, despite their strong positions, face challenges in the form of competition from electric car manufacturers such as Tesla, which is a leader in electrification and innovation. Tesla has a competitive advantage in high-tech

developments such as autopilot and efficient batteries, which puts pressure on traditional automakers. Despite Volkswagen investing heavily in electrification, Tesla continues to dominate this segment, especially in the North American and European markets.

The figure contains filters by year and color, which allow you to select data for analysis for the period from 2015 to 2023, as well as segment visualizations by specific categories, such as brands or types of vehicles, which makes analysis more convenient.

Filters in Tableau are a tool that allows users to select and display only the part of the data that they need for analysis. Filters help focus on certain aspects of the data, hiding data that is not relevant to the current context. They can be applied to different types of fields, such as numeric, categorical, date, and others.

Basic types of filters:

- Category Filters: Category filters enable you to choose data according to the values in text or category fields.

- Range Filters: Range filters allow you to select a range of values for numeric fields. Date Filters: Date filters allow you to select data for specific periods, such as a year, month, or day. Measure Filters: These are filters that are based on numeric fields, allowing you to display data only for certain aggregate values, such as sums, averages, and more.

- Context Filters: Context filters create a subset of the data on which other filters are applied. They are used to optimise work with large data sets.

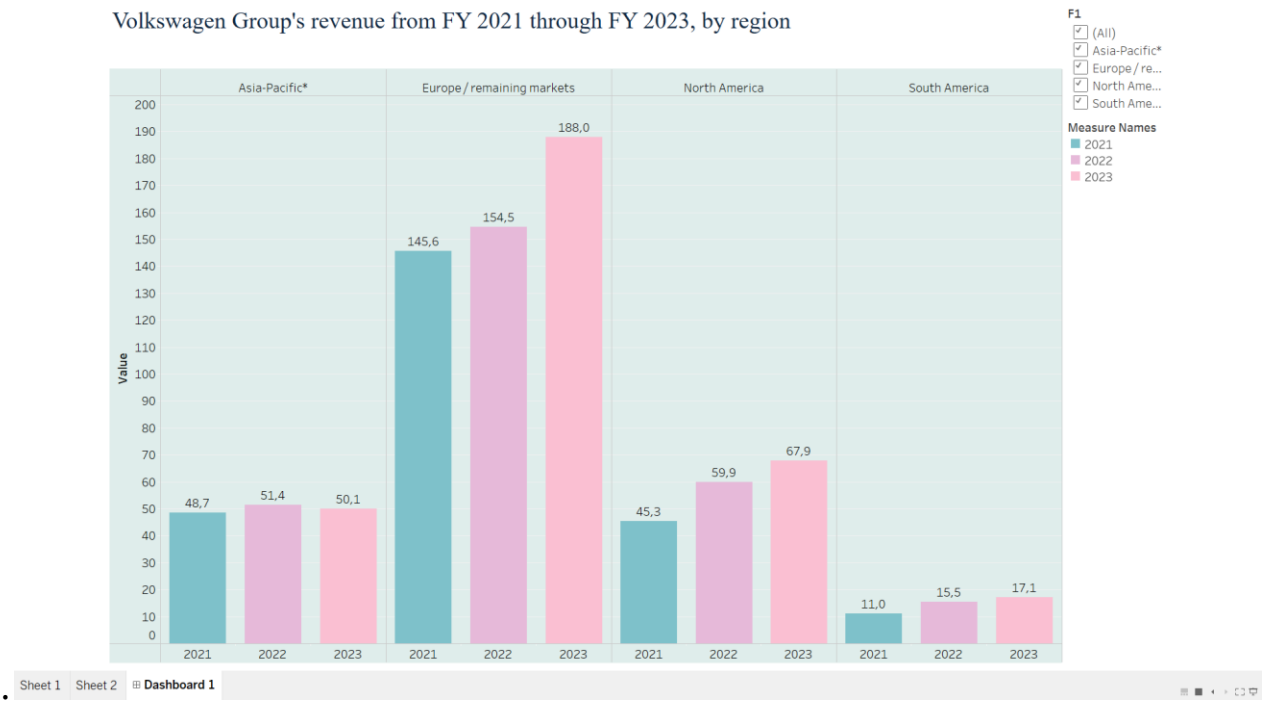


Fig.2.2 Volkswagen Group revenue from fiscal year 2021 to 2023 by region (in billion euros) [Source: author's development]

This data shows the Volkswagen Group's revenue in billions of euros for the financial years 2021 to 2023, broken down by region. In particular, the data contains information about the company's revenues in the following regions: Europe/Other Markets, North America, Asia Pacific, and South America.[15]

In the 2023 reporting year, Volkswagen Group achieved total sales of 9.4 million vehicles, up from 8.5 million in the previous year. Sales revenue increased by 15.5% to €322.3 billion, driven by higher volumes, improved pricing strategies, and synergistic effects.

In Europe and other markets, sales grew by 19.3% year-on-year, reaching 4.2 million vehicles. Improved price positioning contributed to a significant rise in sales revenue, which totaled €187.9 billion compared to €154.5 billion in 2022.

North American markets saw a 24.2% increase in sales, with 1.1 million vehicles sold during the year. This growth, primarily driven by increased sales volumes, resulted in revenue of €67.9 billion, up from €59.9 billion in the previous year.

In South America, Volkswagen Group sold 513,000 vehicles in 2023, a 5.3% increase compared to 2022. This growth translated into a 10.7% increase in sales revenue, reaching €17.1 billion.

In the Asia-Pacific region, unit sales, including contributions from Chinese joint ventures, remained stable at 3.6 million vehicles. Sales revenue for the region amounted to €50.1 billion, slightly down from €51.4 billion in 2022. It's important to note that this revenue figure excludes contributions from Chinese joint ventures, which are accounted for using the equity method.[17]

In South America and Asia, the Volkswagen Group encounters several challenges that contribute to weak sales performance.

In South America, economic instability and political crises are the primary factors driving the reduced demand for new cars, as these issues diminish consumer purchasing power and interest in new vehicles. High inflation and unstable currency exchange rates are making consumers' financial situation more difficult, reducing demand for imported cars. In addition, high customs duties and taxes increase the final cost of cars, making them less affordable for consumers.

In Asia, the Volkswagen Group faces lower sales due to intense competition from local manufacturers like Toyota, Honda, Hyundai, and BYD. These companies dominate the market with locally tailored models and lower production costs. BYD, in particular, stands out as a leader in electric vehicle production in China, leveraging innovative technologies that exert significant pressure on traditional automakers such as Volkswagen.

Conclusions to Chapter 2

As a result of the analysis of Volkswagen Group's competitors, the main players in the market are Toyota, BMW, Mercedes-Benz, Tesla and BYD. Toyota holds a leading position due to its reliability and development of hybrid and electric technologies, especially in the Asian and US markets. BMW and Mercedes-Benz actively compete with Audi in the premium segment, in particular due to innovation and development of electric models. Tesla stands out due to its advanced technologies in the field of electric cars and autonomous driving, creating serious competition for traditional automakers. BYD holds a strong position in the Chinese market, where it offers affordable electric cars, in particular competing with Volkswagen in the electric car segment.

CHAPTER 3. SALES FORECASTING USING MACHINE LEARNING

3.1 Theoretical foundations of sales forecasting using machine learning

Machine Learning (ML) is a subset of artificial intelligence (AI) that employs methods and technologies to develop algorithms capable of learning from data. Unlike traditional approaches where systems are explicitly programmed for specific tasks, ML models improve their performance over time by analyzing data and identifying patterns. This enables ML systems to tackle complex tasks that are challenging or impractical to solve through manual programming.[8]

One important type of data that requires specific approaches to analysis and forecasting is time series. A time series is a sequence of values collected at regular intervals of time that reflects how certain indicators or events change over time. For example, data on monthly car sales of the Volkswagen Group is a time series where changes in sales volumes are recorded over a certain period. Time series analysis enables the identification of fundamental patterns and the prediction of future values using historical data.

The ARIMA (AutoRegressive Integrated Moving Average) model is commonly used for forecasting time series data. It enables predictions based on past observations, making it effective for series with a clear relationship between prior and future values. Combining autoregression, integration, and moving average components, ARIMA captures various characteristics of a time series, including trends, seasonal patterns, and noise.

It is important to note that the ARIMA model is usually attributed to statistical methods, and not to classical machine learning methods. However, in the context of data analysis, it is often used together with ML methods, especially when integrated into analytical systems where accurate forecasting is required. For example, ARIMA can complement ML algorithms in hybrid models where the time series is a significant part of the data. ARIMA is also often used in time series analysis before preparing data for ML, helping to isolate key components (trends and seasonality) to improve the forecast.

The ARIMA model is defined by three parameters: p, representing the order of the autoregressive component that models the relationship between the current value

and past observations; d, indicating the degree of differencing required to make the time series stationary; and q, specifying the order of the moving average component, which addresses errors from previous predictions. The model is formulated as follows:

$$Y_t = \alpha + \sum_{\{i=1\}}^{p} \phi_i Y_{\{t-i\}} + \sum_{\{j=1\}}^{q} \theta_j \epsilon_{\{t-j\}} + \epsilon_t, \quad (3.1)$$

де Y_t— value at a point in time *t*;

α — constant, ϕ_i — coefficients of the autoregressive component;

θ_j — moving average coefficients;

a ϵ_t — error at the time *t*.

Building an ARIMA model involves several steps. The process begins with a preliminary analysis of the data to detect trends, seasonal patterns, or structural breaks in the time series. If the series is non-stationary, differencing is applied to stabilize it. Optimal values for the parameters p, d and q are determined through graphical analysis, using autocorrelation and partial autocorrelation functions (ACF and PACF), alongside criteria such as the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Once the parameters are selected, they are estimated using methods like maximum likelihood estimation or least squares..

Another important step is to check the model residuals for autocorrelation, normality, and homoscedasticity. Despite its numerous advantages, such as the ability to account for autocorrelation and trends over time, the ARIMA model has its limitations. It requires the time series to be stationary, which may require additional transformations. In addition, the ARIMA model may not be effective for forecasting data with strong seasonal fluctuations or in cases where significant structural changes are expected in the future.

Machine learning methods:

- Supervised learning. The algorithm is trained on a dataset in which each example has a label indicating its class.

- Unsupervised learning. The algorithm is trained on a dataset that has no labels.

- Semi-supervised learning. The algorithm is trained on a data set that contains both labelled and unlabelled examples.

- Reinforcement Learning (RL). The algorithm learns to perform tasks by receiving a ‘reward’ or ‘penalty’ for its actions.

Supervised learning is the most widely used form of machine learning. The ML algorithm is trained on a dataset where each example has a label. The label describes the desired outcome for this example. Based on this data, the algorithm learns to associate the input data with the output labels.

There are many types of supervised learning algorithms, including:

- Linear regression. It is used to predict a numerical value based on one or more inputs.

- Logistic regression. Used to classify data into categories.

- Decision trees. Used to make decisions based on a set of rules.

- K-nearest neighbours (KNN) method. Used to classify data based on similarity to other data in the set.

- Support vector machine (SVM). Used for data classification and regression.

Unsupervised learning. This type of machine learning involves working with an unlabeled dataset. The algorithm's objective is to detect patterns, identify structures within the data, group objects based on similar characteristics, and, where possible, predict how the situation may evolve in the future.

Unsupervised learning is often used to divide data into groups with similar characteristics (clustering), reduce the number of data features without losing information (dimensionality reduction), and detect unusual or abnormal data (anomaly detection).

The algorithms for learning without a teacher include:

- The k-means method is a clustering algorithm that divides data into a predetermined number of groups.

- Approximate k-means algorithm. A more efficient version of the previous algorithm that can handle large data sets.

- Principal component analysis. A dimensionality reduction method that selects the most informative features in data sets.

A hybrid approach combining supervised and unsupervised methods is known as semi-supervised learning (or mixed learning). In this approach, the algorithm is trained on a dataset that includes both labeled and unlabeled data.

The labeled data guides the algorithm during the supervised learning phase, enabling it to learn specific mappings or predictions. Meanwhile, the unlabeled data is leveraged for unsupervised learning, helping the model uncover hidden patterns and structures within the dataset. This combination allows the algorithm to make better use of available data, especially in cases where labeled data is limited or expensive to obtain.

"Semi-supervised learning" can be effective in situations where there is a small set of labeled and a large set of unlabeled data. For example, during self-learning or co-training of ML models.

Reinforcement Learning (RL) is a machine learning paradigm where an algorithm learns by interacting with its environment. The algorithm, known as an agent, takes actions in the environment, receives feedback in the form of rewards or penalties, and adjusts its strategy to maximize long-term rewards. Unlike other ML methods, where the algorithm is trained on a labeled dataset, an RL algorithm receives "reinforcement" — a "reward" or "penalty" — for certain outcomes and over time learns to choose actions that maximize the reward.

Advantages of RL:

- Self-learning capability. Such models do not require a large set of labelled data.

- Efficiency in complex tasks. The RL model can be used for tasks that are difficult or impossible to solve using other machine learning methods.

- Adaptability. Algorithms can adapt to new conditions and tasks.

Deep learning is a specialized branch of machine learning that employs artificial neural networks to tackle complex problems. These networks are inspired by the structure and function of biological neural networks, such as those in the human brain.

Artificial neural networks consist of multiple layers of artificial neurons, which are computational units that perform basic calculations. Each neuron receives input

signals from other neurons, processes these signals, and generates an output signal. As information moves through the network's layers, it is progressively transformed and analyzed, enabling deep learning models to extract patterns and insights from the data. Thanks to its multilayer architecture, Deep Learning allows you to detect complex patterns in huge data sets and can solve AI tasks that previously seemed impossible.

There are several main types of deep learning:

- Convolutional neural networks (CNNs). Such networks analyse input data (images, videos) in parts using small ‘filters’. Each filter focuses on certain features, such as shape, colour, or texture. Convolutional neural networks (CNNs) process data by applying filters to extract a hierarchy of features, enabling them to recognize and classify objects in images as well as detect patterns in visual data. They are highly effective for tasks such as pattern recognition, image analysis, and video processing.

- Recurrent neural networks (RNNs) are designed to handle sequential data, such as text or speech, by leveraging the context of previous elements in the sequence to inform the processing of current inputs. Information from previous elements of the sequence influences how the RNN interprets the current element (this is generally similar to the model of dialogue between people). RNNs are ideal for tasks where order and context are important. They are well suited for natural language processing, speech understanding, and machine translation.

- Transformer is a new neural network architecture (introduced in 2017) that has brought natural language processing to a whole new level of efficiency. Unlike RNNs, which process data sequentially, transformer models can simultaneously ‘see’ all elements of a sequence, taking into account the relationships between them. This makes transformers efficient and fast in solving tasks related to long texts.

Deep Learning makes it easy to solve tasks that were previously impossible (or difficult) for computing systems. For example, it is image and speech recognition, “understanding” human language and generating adequate answers to questions asked by a person.

What business tasks can machine learning and artificial intelligence solve?

Machine learning (ML) and artificial intelligence (AI) are powerful tools that can help businesses in many aspects. For example, automate many routine tasks, increase process efficiency, improve customer service, develop new products and services, and make useful recommendations based on the analysis of huge data sets.

Here are some examples of how Machine Learning and artificial intelligence can be used to solve business tasks:

- Automate repetitive tasks such as data entry, document processing, and customer service. This frees up employees' time for more complex work.

- Optimisation of business processes. ML/AI can analyse data to identify inefficiencies and optimise business processes, increasing the company's productivity.

- Personalisation of services. ML/AI can analyse customer data to provide personalised offers, recommendations, and support.

- Automate answers to questions and anticipate customer needs.

- Fraudulent transaction detection and cybersecurity.

- Risk management.

Machine learning is now widely used in various areas of our lives, from medicine and finance to marketing and manufacturing. Let's look at some examples of ML applications that will help you better understand how this technology can be used to solve real-world problems.

- Amazon, Netflix, Spotify. They use ML to recommend their services to customers based on personal purchase history, browsing, preferences, and other data.

- Google, Facebook, X (formerly Twitter). They use ML to show users the most relevant ads based on their interests, search history, demographic data, and social media behaviour.

- Virtual assistants Apple Siri, Amazon Alexa, Google Assistant. They use ML to understand natural language and fulfil user requests.

- IBM Watson. Uses ML to analyse medical images such as X-rays and MRIs to help doctors diagnose diseases.

- Verily. Uses ML to develop algorithms that can diagnose diseases based on patient data such as symptoms, medical history, and test results.

- Walmart. Uses ML to predict demand for goods, which helps the company optimise its inventory and production.

- Uber. Uses ML to predict the demand for rides, which helps the company optimise the distribution of its drivers.

These are just a few examples of how machine learning is used in business.

Today, ML technologies and algorithms are becoming increasingly powerful and productive, which will allow you to solve even more complex tasks. Companies that can effectively use machine learning will have a tangible competitive advantage in the market. However, it is worth remembering that Machine Learning is not a magic wand. In order for the technologies to be effective, you need to have high-quality data, competent specialists and a well-thought-out implementation strategy. The combination of these factors significantly increases the possibility of effective use of ML technologies.[7]

3.2. Practical implementation of sales forecasting

In this section, we will perform a practical implementation of sales forecasting based on historical revenue data for Volkswagen Group. We will use the company's revenue time series from 2006 to 2023 and build a model to forecast future values based on the ARIMA model.

Historical revenue data[14] (in billion euros):

2006: 104.88

2007: 108.90

2008: 113.81

2009: 105.19

2010: 126.88

2011: 159.34

2012: 192.68

2013: 197.01

2014: 202.46

2015: 213.29

2016: 217.27

2017: 229.55

2018: 235.85

2019: 252.63

2020: 222.88

2021: 250.20

2022: 279.05

2023: 322.28

To build the forecast, we used the ARIMA (AutoRegressive Integrated Moving Average) model. This model is well suited for time series analysis, where there is a relationship between past values and future forecasts.

First, we import the pandas library, which will be used to work with data tables. Then we load an Excel file containing data on the revenue of the Volkswagen Group for the years 2006-2023. The skiprows=4 parameter allows you to skip the first four rows of the file, which do not contain useful information.

```
[ ]  import pandas as pd

     data = pd.read_excel('/content/statistic_id264349_volkswagen-groups-sales-revenue-2006-2023 (1).xlsx', sheet_name='Data', skiprows=4)
```

Fig.3.1. Installing libraries and reading a file. [Source: author's development]

Next, we output the first few rows of the table to make sure the data is loaded correctly.

```
[ ]  print(data.head())

     Unnamed: 0  Unnamed: 1  Unnamed: 2
0         NaN        2006      104.88
1         NaN        2007      108.90
2         NaN        2008      113.81
3         NaN        2009      105.19
4         NaN        2010      126.88
```

Fig.3.2. Output of lines. [Source: author's development]

We use the isnull().sum() function to check if there are any missing values in the table that could affect further analysis.

```
[ ]  print(data.isnull().sum())

Unnamed: 0    18
Unnamed: 1     0
Unnamed: 2     0
dtype: int64
```

Fig.3.3. The isnull().sum() function. [Source: author's development)]

We remove the extra column Unnamed: 0, which appears due to the specifics of the Excel file structure. We also rename the columns to make them more informative. We output the data to make sure that everything looks correct after cleaning.

The Year column can be stored in floating point format. To avoid problems with display and processing, we convert it to an integer.

```
[ ]  data['Year'] = data['Year'].astype(int)
```

Fig.3.4. Conversion to an integer. [Source: author's development]

After preparing the data, we use the matplotlib.pyplot library to plot the revenue by year. The plt.plot() function creates a line graph with the years on the X-axis and the company's revenue on the Y-axis. We use plt.xticks() to add labels for each year on the X-axis and rotate them 45 degrees for better readability. We add a title, axis labels, and a grid. We add a title, axis labels, and a grid and display.

```
import matplotlib.pyplot as plt

plt.plot(data['Year'], data['Revenue'], marker='o')

plt.xticks(data['Year'], rotation=45)

plt.title('Volkswagen Group revenues by year')
plt.xlabel('Year')
plt.ylabel('Revenue (billion euros)')
plt.grid(True)
plt.show()
```

Fig.3.5. Creating a graph. [Source: author's development]

Fig. 3.6. Dynamics of Volkswagen Group revenues from 2006 to 2023. [Source: author's development]

The graph shows the dynamics of Volkswagen Group's revenue from 2006 to 2023, showing a general upward trend, despite some fluctuations. Overall, the company's revenue is increasing, which indicates stable development and growing demand for its products. However, there are years when revenue declines are observed. In particular, in 2009, there was a noticeable drop, which is likely related to the global economic crisis of 2008, which negatively affected the automotive industry. Another decline occurred in 2020, due to the COVID-19 pandemic, which led to a decrease in demand for cars and disruptions in production.

In contrast, from 2010 to 2018, a gradual increase in revenue is visible. This can be attributed to the development of new models, the expansion of sales markets and innovations in production. Since 2021, after the pandemic, there has been significant growth, which may be a result of economic recovery and increased demand for cars,

including electric vehicles. It is also worth noting that during this period there may be minor fluctuations in revenues, which may be due to seasonal factors or the impact of competition in the market.

```
from statsmodels.tsa.arima.model import ARIMA

revenue_data = data['Revenue']

model = ARIMA(revenue_data, order=(1, 1, 1))
model_fit = model.fit()

forecast = model_fit.forecast(steps=5)

future_years = pd.DataFrame({'Year': range(2024, 2029), 'Forecast': forecast})

print(future_years)
```

```
    Year    Forecast
18  2024  333.646285
19  2025  345.012472
20  2026  356.378561
21  2027  367.744553
22  2028  379.110448
```

Fig. 3.6. Forecasting the income of the Volkswagen Group for the next 5 years. [Source: author's development]

In this code, we use the ARIMA model to forecast the Volkswagen Group's revenue for the next 5 years based on available historical data. First, the ARIMA class is imported from the statsmodels library, which specializes in statistical time series modeling. We select only the Revenue column with revenue data, since it is on its basis that we will build the forecast. The ARIMA model works with time series and consists of three components: autocorrelation (AR), integration (I), which is responsible for the difference transformations to achieve stationarity of the series, and moving average (MA). In our case, the model is initialized with parameters (1, 1, 1), where the first value "1" means that the model uses one previous period to forecast the current value, which corresponds to the autoregression parameter (p). This parameter indicates that the current value of the time series depends on one previous value. In cases where the time series exhibits significant autocorrelation at longer lags, the value of p can be increased, but for simple series, choosing p = 1 is often sufficient, since a single prior value significantly influences the forecast. The second value, "1", is the difference transformation (d = 1), which determines how many times the difference should be applied to make the series stationary. Often, time series, especially financial or economic, have a trend, so a single difference transformation helps to make the series

stationary by removing the effect of the trend. If after one difference, the series still does not become stationary, it may be necessary to increase the value of d. The initial value of d = 1 is chosen based on the assumption that a single difference should be enough to eliminate the trend in most series. The third value, the moving average (q = 1), takes into account one prior value of random fluctuations to correct the forecast, smoothing out short-term fluctuations in the series. This allows the model to account for short-term changes that depend on a single previous period. If the residuals (after accounting for autoregression and differencing) show significant correlation at longer lags, it is worth considering increasing q. Thus, the initial choice of (1, 1, 1) for ARIMA provides a simple baseline for data analysis and is a starting point that can be adjusted according to the model's results, especially if the accuracy estimate turns out to be insufficient. After initializing the model, it is trained on historical revenue data using the .fit() method. This allows the model to adjust its parameters to best describe the dynamics of the company's revenue.

After setting up the model, we use the .forecast(steps=5) method to forecast values 5 years ahead. This method calculates future values based on patterns in the historical data and the model we created. The result of the forecast is stored in the forecast variable, which contains the forecasted revenue values for the period from 2024 to 2028.

Next, we create a new DataFrame future_years where we combine the forecast years (2024 to 2028) and the corresponding revenue values obtained from the forecast. Finally, using print(future_years), we output this DataFrame containing the forecasted revenues of Volkswagen Group for each of the next five years.

```
plt.plot(data['Year'], data['Revenue'], marker='o', label="Historical data")
plt.plot(future_years['Year'], future_years ['Forecast'], marker='o', linestyle='--', color='orange', label='Forecast')
plt.xticks(list(data['Year']) + list(future_years['Year']), rotation=45)
plt.title('Volkswagen Group revenue forecast')
plt.xlabel('Year')
plt.ylabel('Revenue (billion euros)')
plt.legend()
plt.grid(True)
plt.show()
```

Fig.3.7. Creating a graph. [Source: author's development]

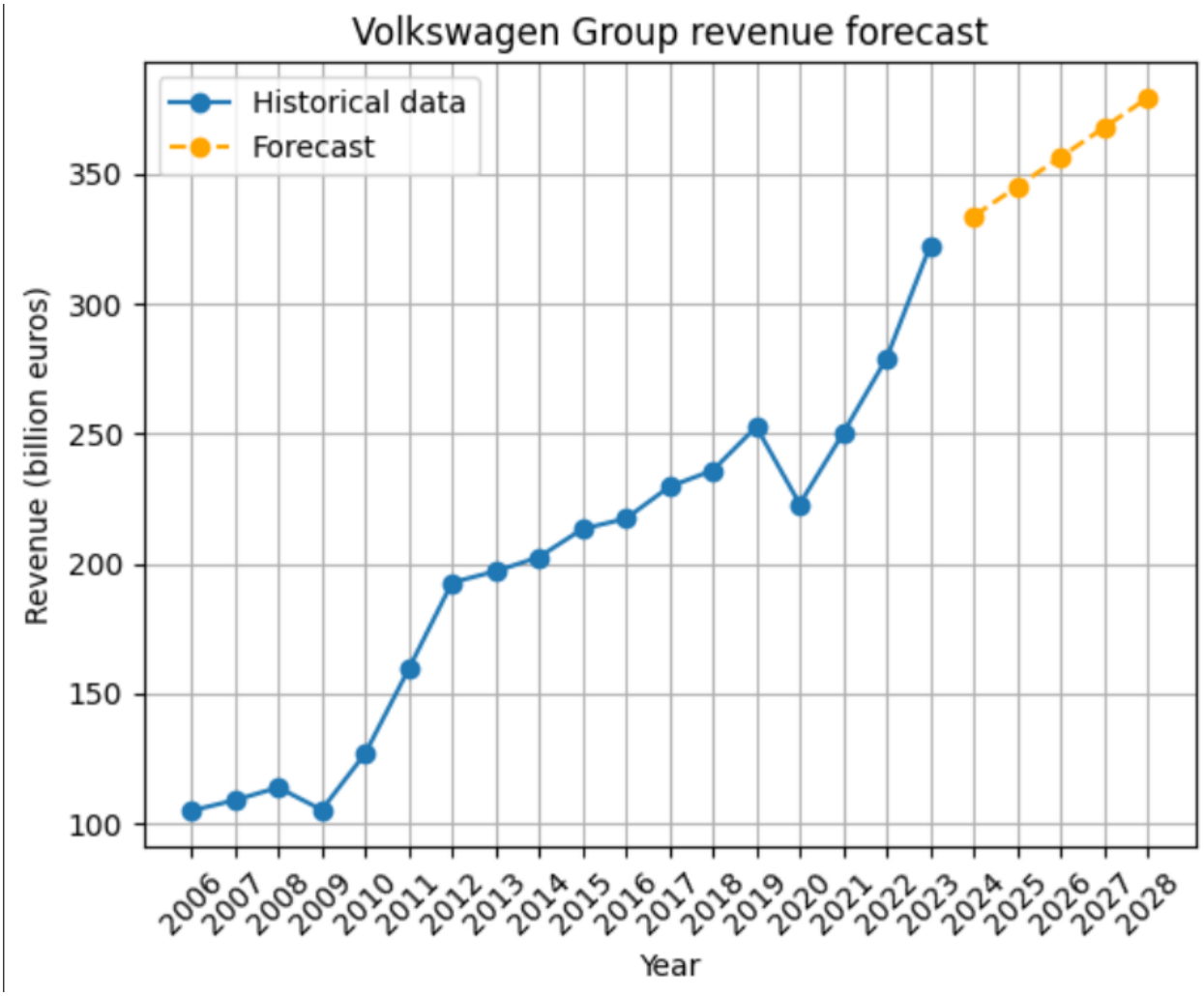


Fig. 3.6. Volkswagen Group revenue forecast from 2023 to 2028. [Source: author's development]

The graph shows that Volkswagen Group's revenue has had a steady upward trend from 2006 to 2023, although there have been some fluctuations. For example, there is a decline in 2009, which is likely related to the economic crisis, and another decline in 2020, possibly due to the impact of the COVID-19 pandemic. The projected portion of the graph (from 2024 to 2028) indicates further growth in the company's revenue, although the pace may fluctuate.

Volkswagen Group's sales and revenue growth depends on a number of factors, such as demand for cars, the economic situation, innovation and environmental friendliness of products, as well as external global crises. For example, a stable economy and growing demand for environmentally friendly vehicles can have a positive impact on revenue. At the same time, events such as economic recessions or political crises can lead to a decline in sales, as can be seen in historical data.

Thus, the graph shows an optimistic forecast for the Volkswagen Group, but its implementation will depend on stability and favorable market conditions, as well as on the company's ability to adapt to changing conditions and implement innovative solutions.

Conclusions to Chapter 3

An analysis of Volkswagen Group’s revenue from 2006 to 2023 showed an overall increase in the company’s revenue, with revenue reaching €322.28 billion in 2023, up 13% from the previous year. This growth can be attributed to the economic recovery following the COVID-19 pandemic, increased demand for cars, particularly electric models, as well as improved production processes and the launch of new models.

The forecast for 2024-2028 shows steady revenue growth, but with a gradual decline in the growth rate. Revenue growth is forecast at 3.53% in 2024, followed by a gradual decline to 3.41% in 2025, 3.30% in 2026, 3.19% in 2027 and 3.09% in 2028.

Modeling and forecasting of these indicators was carried out using the ARIMA model, which has demonstrated the ability to accurately predict future values based on historical data.

CONCLUSIONS

As a result of the analysis of the activities of the Volkswagen Group, several important conclusions can be drawn. The Volkswagen Group is one of the largest automotive corporations in the world, which unites 10 main brands, each of which has its own unique focus. The company is actively developing in various segments of the automotive market, from mass production to luxury and sports models, and is also engaged in the production of motorcycles through Ducati. An important feature is the strategy of global presence, with production and research facilities located around the world, with the main focus on Europe, China and North America. However, despite the lack of its own production facilities in Ukraine, Volkswagen maintains a strong presence in the Ukrainian market through a network of official dealers and service centers.

Regarding the competitive environment, Volkswagen's main competitors in the global market are Toyota, BMW, Mercedes-Benz, Tesla and BYD. Each of these brands has its own strengths, including Toyota’s reliability and development of hybrid and electric technologies, BMW and Mercedes-Benz’s innovation in the premium segment, Tesla’s advanced electric vehicle and autonomous driving technologies, and BYD’s affordable electric vehicles, especially in the Chinese market.

An analysis of the Volkswagen Group’s financial results from 2006 to 2023 shows a steady increase in the company’s revenue, reaching €322.28 billion in 2023, up 13% from the previous year. This growth can be attributed to the economic recovery following the COVID-19 pandemic, increased demand for cars, particularly electric models, as well as improved production processes and the launch of new models. The forecast for the period 2024-2028 predicts a steady increase in revenue, but with a gradual decrease in the growth rate. Thus, in 2024, growth is forecasted at 3.53%, with a further decrease to 3.09% in 2028.

Modeling and forecasting of these indicators was carried out using the ARIMA model, which has proven its effectiveness in predicting future values based on historical data.

REFERENCES

1. About Volkswagen Group concern. History, structure of VAG. URL: https://www.vwdrive.com.ua/o-volkswagen-group/ (Date of application: 04.09.2024)

2. Volkswagen Group Technology. URL: https://www.volkswagen-group.com/en/volkswagen-group-technology-16016 (Date of application: 13.09.2024)

3. Basics of working with Looker Studio. URL: https://www.edu.goit.global/uk/learn/18173459/20561705/20561857/training?blockId=20599322 (Date of application: 11.10.2024)

4. Opportunities and advantage of Microsoft Power BI. URL: https://danco.com.ua/useful/vozmozhnosti-i-preimushhestva-microsoft-power-bi/ (Date of application: 08.08.2024)

5. Volkswagen Group's revenue from vehicle sales from FY 2015 through FY 2023, by major subsidiary or brand URL: https://www.statista.com/statistics/275870/revenue-of-volkswagen-group-by-brand/ (Date of application: 09.09.2024)

6. Guide for working with Tableau.URL: https://dou.ua/forums/topic/39355/ (Дата звернення: 18.09.2024)

7. What is Machine Learning? URL: https://denovo.ua/resources/what-is-machine-learning (Date of application: 23.09.2024)

8. What is a machine learning model? URL: https://learn.microsoft.com/ru-ru/windows/ai/windows-ml/what-is-a-machine-learning-model (Date of application: 27.09.2024)

9. Machine Learning: What It Is, How It Works, and What It's Used For. URL: https://goit.global/ua/articles/mashynne-navchannia-shcho-tse-take-iak-pratsiuie-i-dlia-choho-vykorystovuietsia/ (Date of application: 23.09.2024)

10. Worldwide Car Sales: Top-Selling Manufacturers. URL: https://roadgenius.com/cars/statistics/sales-by-manufacturer/ (Date of application: 30.09.2024)

11. VAG Group Concern. URL: https://vag.ua/blog-vw/kontsern-vag-group/ (Date of application: 29.09.2024)

12. Volkswagen Group sites. URL: https://www.volkswagen-group.com/en/volkswagen-group-sites-17481 (Date of application: 29.10.2024)

13. Brands and brand groups. URL: https://www.volkswagen-group.com/en/brands-and-brand-groups-15811 (Date of application: 11.11.2024)

14. Volkswagen Group's sales revenue from FY 2006 to FY 2023. URL: https://www.statista.com/statistics/264349/sales-revenue-of-volkswagen-ag-since-2006/ (Date of application: 22.11.2024)

15. Volkswagen Group's revenue from FY 2021 through FY 2023, by region. URL: https://www.statista.com/statistics/260805/revenue-of-volkswagen/ (Date of application: 12.11.2024)

16. Brand Group Core. URL: https://www.volkswagen-newsroom.com/en/brand-group-core-17663 (Date of application: 15.11.2024)

17. Annual Report 2023. URL: https://www.volkswagen-group.com/en/publications/more/annual-report-2023-2671 (Date of application: 01.12.2024)

18. Auto Union. URL: https://www.avtosojuz.ua/kontakty/istoriya-nashoho-rozvytku?utm_source (Date of application: 24.11.2024)

19. Auto centers "Avtodim Atlant". URL: https://www.infocar.ua/avtosalon_avtocentri_atlant-m.html?utm_source (Date of application: 24.11.2024)

20. NORA Service Center. URL: https://www.volkswagen.ua/servis-ta-aksesuari/nora-service-center?utm_source (Date of application: 26.11.2024)

21. Key trends in machine learning and AI URL: https://techcrunch.com/2016/07/06/key-trends-in-machine-learning-and-ai/

(Дата звернення: 29.11.2024)

22. Digital Workspace Concepts. URL: https://link.springer.com/chapter/10.1057/9781137554123_3 (Date of application: 30.11.2024)

23. The Volkswagen Group. URL: https://www.volkswagen-group.com/en/about-us-16013 (Date of application: 08.11.2024)

24. R (programming language). URL: https://uk.wikipedia.org/wiki/R_(%D0%BC%D0%BE%D0%B2%D0%B0_%D0%BF%D1%80%D0%BE%D0%B3%D1%80%D0%B0%D0%BC%D1%83%D0%B2%D0%B0%D0%BD%D0%BD%D1%8F) (Date of application: 08.08.2024)

25. Volkswagen Financial Services. URL: https://www.volkswagen-group.com/en/volkswagen-financial-services-16056 (Дата звернення: 01.08.2024)

26. TRATON SE URL: https://www.volkswagen-group.com/en/traton-se-16021 (Date of application: 22.08.2024)

27. Most popular Python libraries. URL: https://foxminded.ua/biblioteky-python/ (Date of application: 20.09.2024)

28. What is Tableau and why do analysts need it? URL: https://profit.store/uk/blog/expert/scho-take-tableau-i-navischo-vona-potribna-analitykam (Date of application: 30.09.2024)

29. Current news from the world of Volkswagen in Ukraine. URL: https://www.volkswagen.ua/ (Date of application: 30.11.2024)

30. Basic statistical tools in research and data analysis. URL: https://pmc.ncbi.nlm.nih.gov/articles/PMC5037948/ (Date of application: 18.11.2024)