Project Overview

This project demonstrates an automated, LLM-driven workflow for generating structured analytical reports from structured datasets, exemplified here using synthetic BMW sales data. The goal is to design a prototype where a language model functions as the central orchestrator, autonomously performing data analysis, visualization, statistical modeling, and report composition. The workflow leverages modern GenAI capabilities combined with modular Python tools to produce a cohesive, reproducible output in DOCX format.

By using this approach, we illustrate how LLMs can act as AI data scientists: reasoning, summarizing, and producing actionable insights without manual intervention. While the case study focuses on sales data, the underlying architecture and tool suite are designed to generalize to other datasets and business reporting tasks.

System Architecture

The system is structured around two core components:

1. **LLM Orchestrator**
   A single LLM, **Gemini 2.5 Flash**, serves as the central orchestrator. This model was chosen for its lightweight footprint, fast inference, and robust support for tool calling. The orchestrator receives a high-level prompt describing the report requirements and autonomously determines which tools to invoke, in what order, and with which parameters. Using **auto tool calling**, the LLM can execute multiple tools in sequence or in parallel, retry operations if results are unsatisfactory, and refine outputs until the report reaches a desired quality—following principles of the ReAct framework for reasoning and acting.
2. **Reusable Python Tool Suite**
   The tool suite performs all concrete data operations and report-building tasks. Key design principles include:
   - **Generality:** Tools are not hard-coded for sales data. They can operate on any structured dataset, enabling broader application across domains.
   - **Clarity for LLM:** Each tool provides structured input/output definitions with clear examples, allowing the LLM to use them autonomously.

The combination of the orchestrator and tools ensures that the system behaves as an intelligent agent, capable of complex, multi-step analytical reasoning.

Tooling and Functionality

Automated Visualization

Visualization is a core capability of this system. The LLM has complete autonomy to:

- Generate **line charts**, including multi-line charts to display group-level breakdowns.
- Generate **bar charts**, aggregated or grouped by category.
- Retrieve the **underlying dataset used for the plot**, enabling the LLM to summarize trends, compare groups, and extract insights directly from the numerical data.

By integrating the underlying data with visual outputs, the system ensures that narrative commentary aligns with quantitative evidence, supporting more accurate and actionable insights. Examples include regional sales trends over time, model-level performance comparisons, and aggregated market analyses.

OLS Statistical Analysis

OLS regression is included to provide a simple yet structured statistical analysis of key drivers. The LLM selects relevant numerical and categorical variables based on the dataset and research objectives. The tool performs:

- Automatic encoding of categorical variables (one-hot encoding)
- Regression modeling
- Output of coefficients, p-values, and R-squared metrics

This allows the report to quantify relationships between features (e.g., price, region, model type) and sales outcomes, providing data-driven explanations for observed trends.

DOCX Report Assembly

The final stage of the workflow is automated report generation:

- A structured JSON schema guides the LLM on how to populate sections.
- Sections can include **text and plots**, with plots embedded in-line in DOCX format.
- The tool ensures consistent formatting and structure while allowing the LLM to provide nuanced, natural-language explanations.

This structured input-output approach reduces errors and helps the LLM maintain a coherent narrative across multiple sections, from the executive summary to recommendations.

LLM Orchestration and Autonomy

The LLM's ability to autonomously call tools is central to the system:

- The model can **invoke multiple tools in parallel** or in sequence.
- Results from one tool can feed into subsequent analyses.
- The LLM can **retry or refine outputs**, ensuring high-quality results before producing the final report.

This design follows the ReAct paradigm, enabling reasoning and action in tandem. The approach allows for iterative and dynamic report generation that mimics the workflow of a human data scientist.

Extensibility and Creative Analysis

A key advantage of this architecture is flexibility. Beyond the pre-defined analytical sections, the LLM can leverage the generic tools to:

- Explore additional hypotheses
- Conduct further statistical or visual analyses
- Generate novel recommendations

This capability ensures that the system can adapt to new business questions or datasets without requiring additional code changes.

Business Insights and Impact

The generated reports provide multiple levels of insights:

- **Performance trends**: Clear identification of top and underperforming models, regions, and markets.
- **Sales drivers**: Quantified relationships between features and outcomes via OLS regression.
- **Actionable recommendations**: Strategic guidance based on analysis, allowing decision-makers to adjust pricing, product focus, or marketing strategy.

By automating reporting, this workflow reduces manual effort, accelerates decision-making, and provides a consistent and reproducible framework for data-driven analysis.

Conclusion

This project demonstrates how LLMs can be integrated with modular, reusable tools to automate complex business reporting tasks. By combining autonomous orchestration, statistical modeling, visualization, and structured document generation, the system behaves as an AI data scientist. The approach is highly extensible, reproducible, and applicable across multiple domains, offering a practical blueprint for leveraging GenAI to accelerate and standardize data-driven decision-making.