# BMW Worldwide Sales — Analysis Report

Mini Project | Big Data Analytics

**Dataset (summary)**
- **Source:** Kaggle — BMW Worldwide Sales Dataset
- **Period:** Covers worldwide BMW car sales across multiple years.
- **Granularity:** Reported by year, country, and model.
- **Typical columns:** Year, Country, Model, Units_Sold, Revenue, Segment.

**Project Objectives**
- Perform exploratory data analysis to understand BMW global sales trends.
- Identify top markets, top-selling models, seasonal patterns, and growth/decline periods.
- Compute key performance metrics (total units, CAGR, market share by region/model).
- Provide actionable recommendations and visualizations for stakeholders.

**Methodology**
- Data ingestion using PySpark or pandas.
- Data cleaning: handling missing values, correcting model names, normalizing currencies.
- Aggregation: summing units and revenue by year, country, model; computing YoY % changes.
- Visualization: line charts for sales trends, bar charts for top countries/models.
- Optional: simple forecasting (time series models) on annual totals.

**Key Findings (summary)**
- Overall global sales show steady growth with some dips during economic slowdowns.
- Top countries contribute the majority of sales, with strong dominance in Europe, USA, and China.
- Certain models consistently drive revenue, while new releases create temporary sales spikes.
- Seasonality observed in some regions, with higher sales in year-end quarters.
- Strong positive correlation found between units sold and revenue.

**Representative PySpark / pandas snippets used**

```
# PySpark read CSV spark = SparkSession.builder.appName("BMWSales").getOrCreate() df =
spark.read.option("header",True).option("inferSchema",True).csv("data/bmw_sales.csv") #
Annual sales aggregation annual =
df.groupBy("Year").agg(F.sum("Units_Sold").alias("TotalUnits")).orderBy("Year") # Top 10
models top_models =
df.groupBy("Model").agg(F.sum("Units_Sold").alias("Units")).orderBy(F.desc("Units")).limit(10)
```

**Recommendations**
- Focus resources on top markets and best-selling models.
- Adjust inventory and marketing strategies based on seasonal peaks.
- Explore low-performing regions for growth opportunities.
- Consider competitive pricing strategies in price-sensitive regions.

**Conclusion**

BMW's worldwide sales show resilience and growth with strong regional dominance. This project identifies opportunities to strengthen global strategy and supports future planning with data-driven insights.