OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

MAGDEBURG RESEARCH AND COMPETENCE CLUSTER VLBA

ARBEITSGRUPPE WIRTSCHAFTSINFORMATIK

AG WI SAP UCC VLBA Lab

MPD II
Project VI: Manufacturing & Product Distribution

VLBAII – System Architectures

Magdeburg Research and Competence Cluster
Workgroup Business Informatics I

Faculty of Computer Science
Department of Technical and Business Information Systems

Otto-von-Guericke-University Magdeburg

Dustin Herrmann 227954
Jerry Jiang 218873
Loris Guba 227127
Nitin Bharadwaj Nataraj 248993
Ranjith Mahesh 249303

AGENDA

- Introduction
- Objective
- Integration Architecture - Architecture, BSP, Component diagram
- Clustering
- Product category selection
- Machine learning forecasts
- Kubeflow pipeline with Vertex AI
- Evaluation and Visualisation past and future developments
- Interpretation

INTRODUCTION

- Mobility Worldwide (MWW) is a global leader headquartered in Magdeburg, specializing in mobility innovation, production, and services
- Product portfolio includes electric scooters, segways, bicycles, e-bikes, and accessories.
- MWW operates under four strategic departments: Consulting & Customer Retention (CCR), **Manufacturing & Product Distribution (MPD)**, Research & Solution Development (RSD), and Maintenance & Service (MS)
- They have hired an Analyst team to analyse past historical data on customers, orders, and materials from SAP Datasphere and forecast future sales, and profits of mobility products across diverse markets

OBJECTIVE

- **Market Clustering**

  Group customer markets by geographical location using BigQuery machine learning algorithm

  Visualize clusters with Google Cloud visualization tools

- **Product Category Analysis**

  Select two key product categories

  Apply data analytics and machine learning to predict future product sales and material demand

- **Profit Trend Comparison**

  Compare past and future profit trends in these markets

  Provide insights to support management decisions

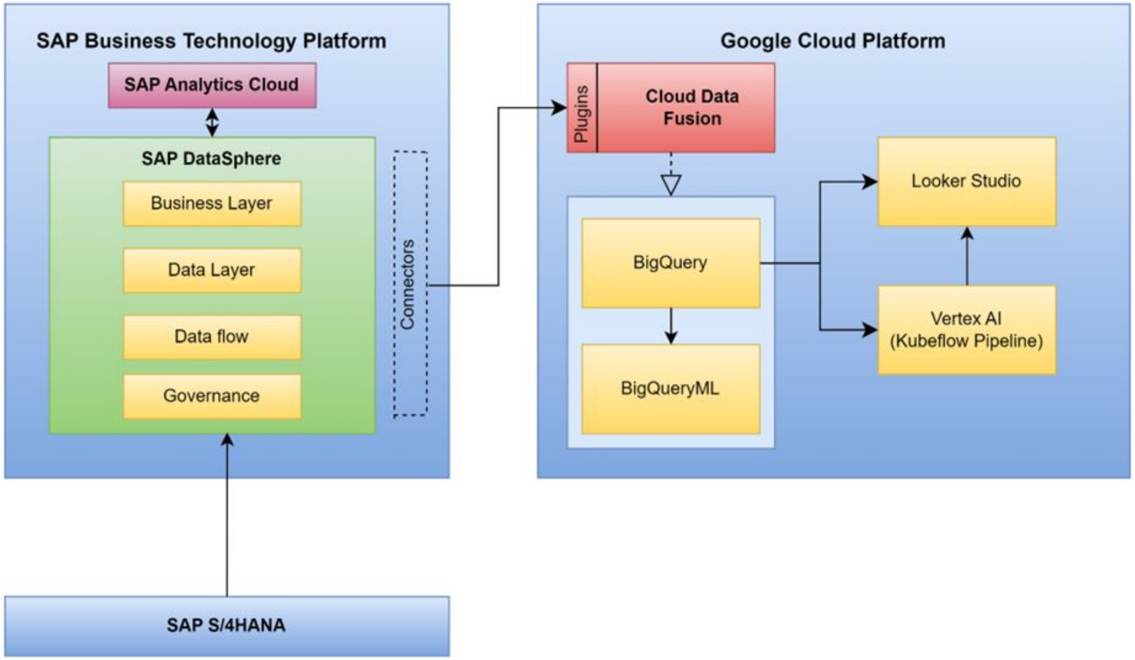
- **Results Summary**

  Summarize and interpret the results and implications for the company in a comprehensive report

INTEGRATION ARCHITECTURE

Data Extraction from SAP Datasphere:

- Uses Cloud Data Fusion's predefined cloud Data Integration patterns.
- Custom workflows for non-standard data extraction

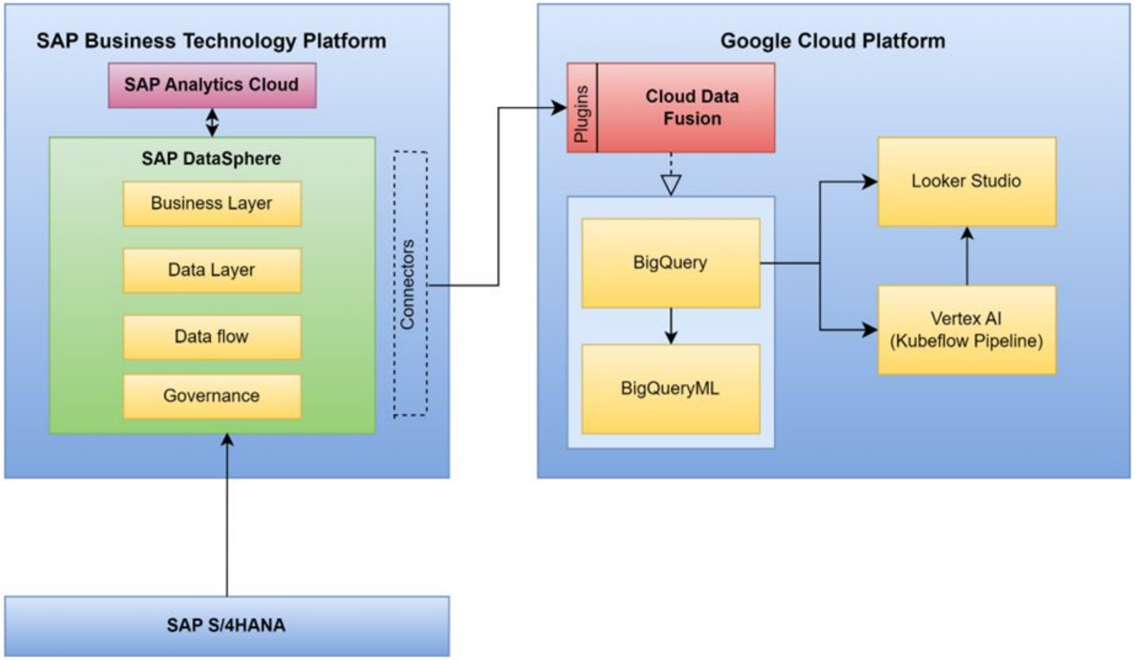


Architecture Diagram

INTEGRATION ARCHITECTURE

Data Ingestion with Cloud Data Fusion:

- Visual tool in the UI for data integration
- Ingests data into BigQuery as tables

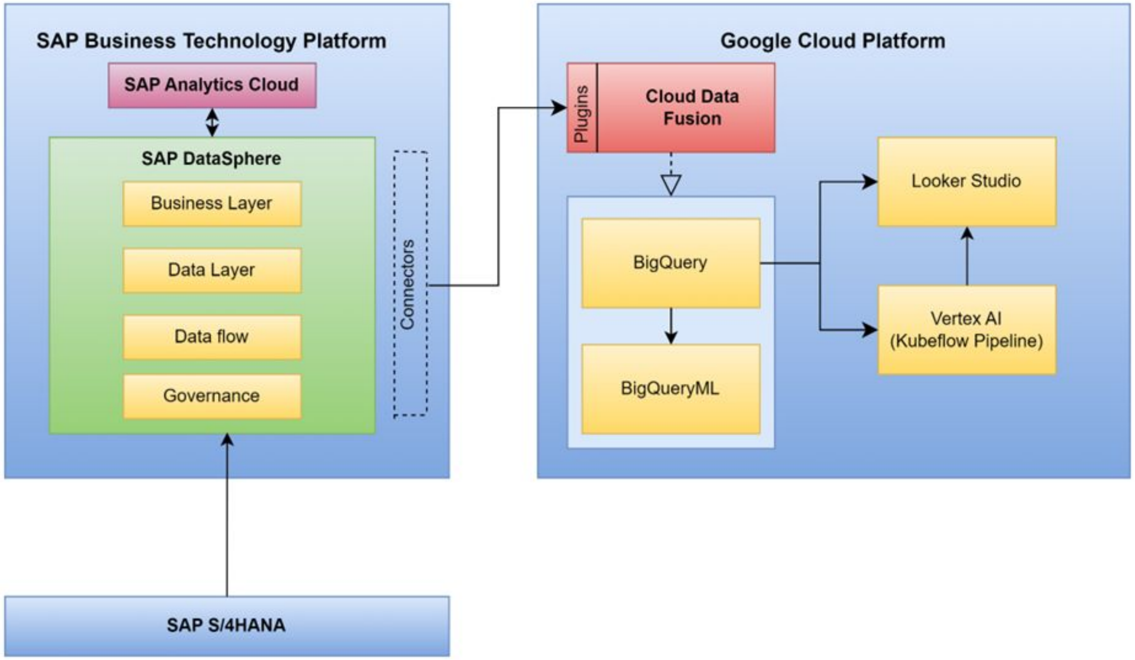


Architecture Diagram

INTEGRATION ARCHITECTURE

Data Storage and Processing in BigQuery:

- Scalable, serverless data warehouse
- Data cleaning, processing and transformation

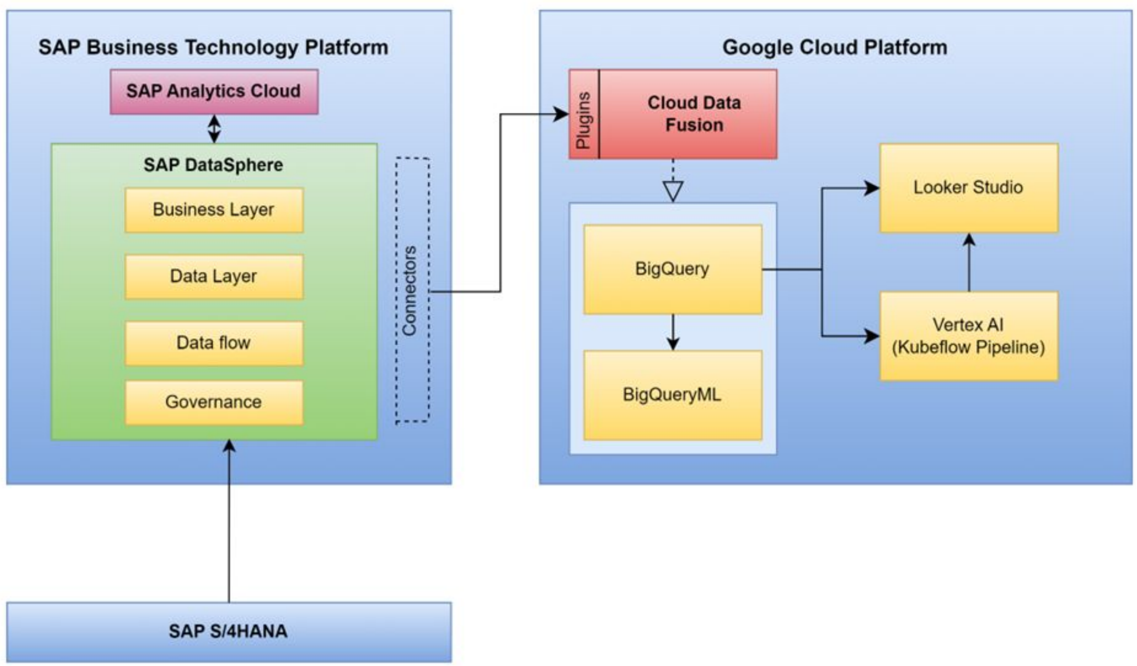


Architecture Diagram

INTEGRATION ARCHITECTURE

Machine Learning with BigQuery ML:

- Build and evaluate ML models using SQL

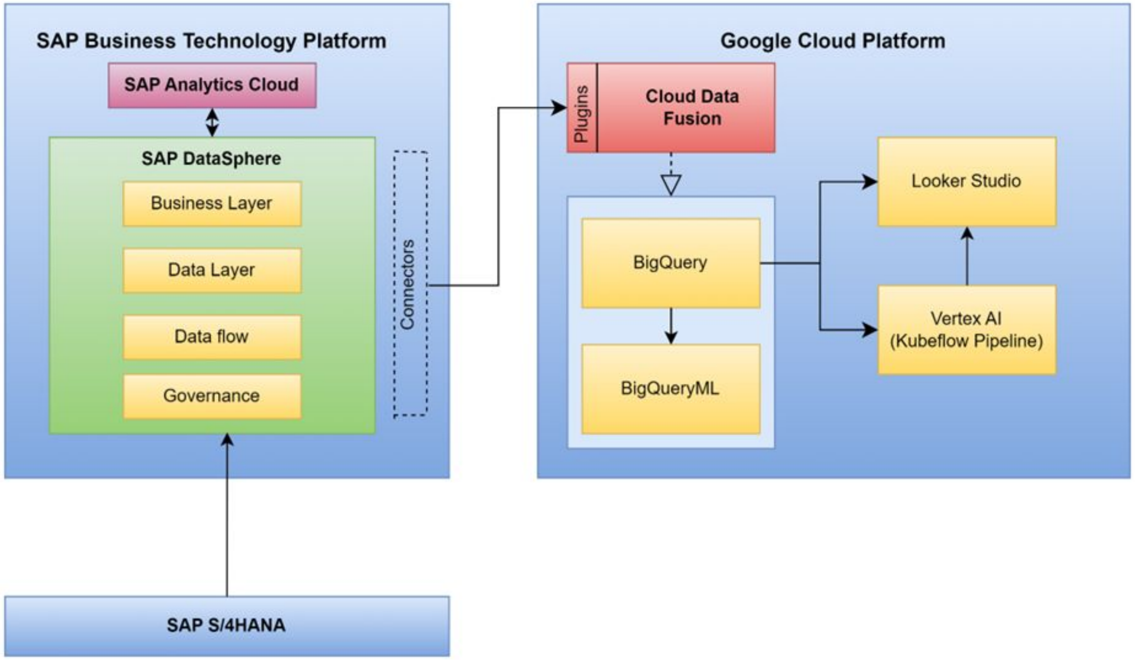


Architecture Diagram

INTEGRATION ARCHITECTURE

Kubeflow Pipeline in Vertex AI:

- End-to-end ML orchestration workflows (preprocessing, training, evaluation, deployment)

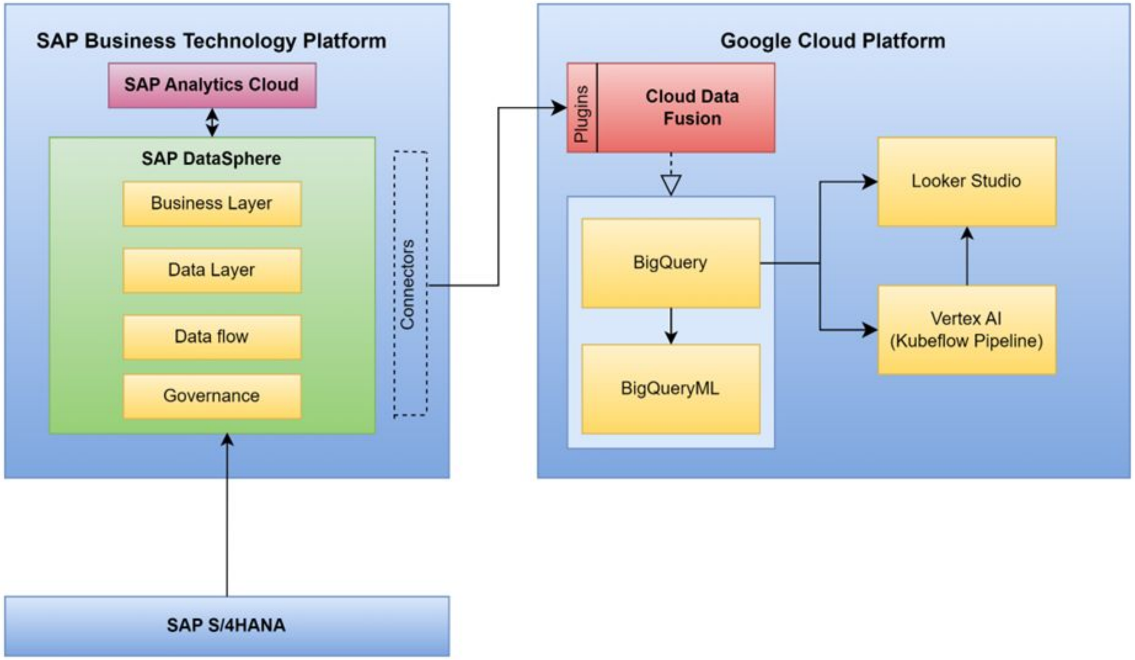


Architecture Diagram

INTEGRATION ARCHITECTURE

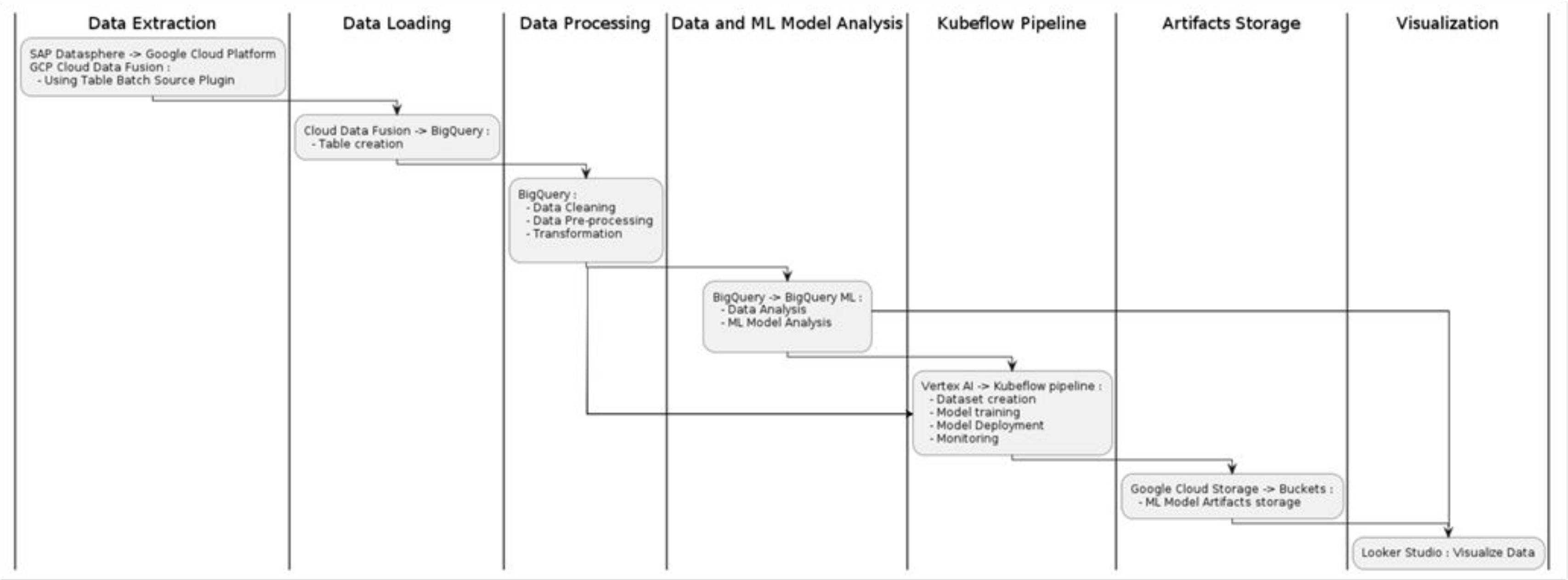
Data Visualization with Looker Studio:

- Interactive dashboards and reports for ML insights



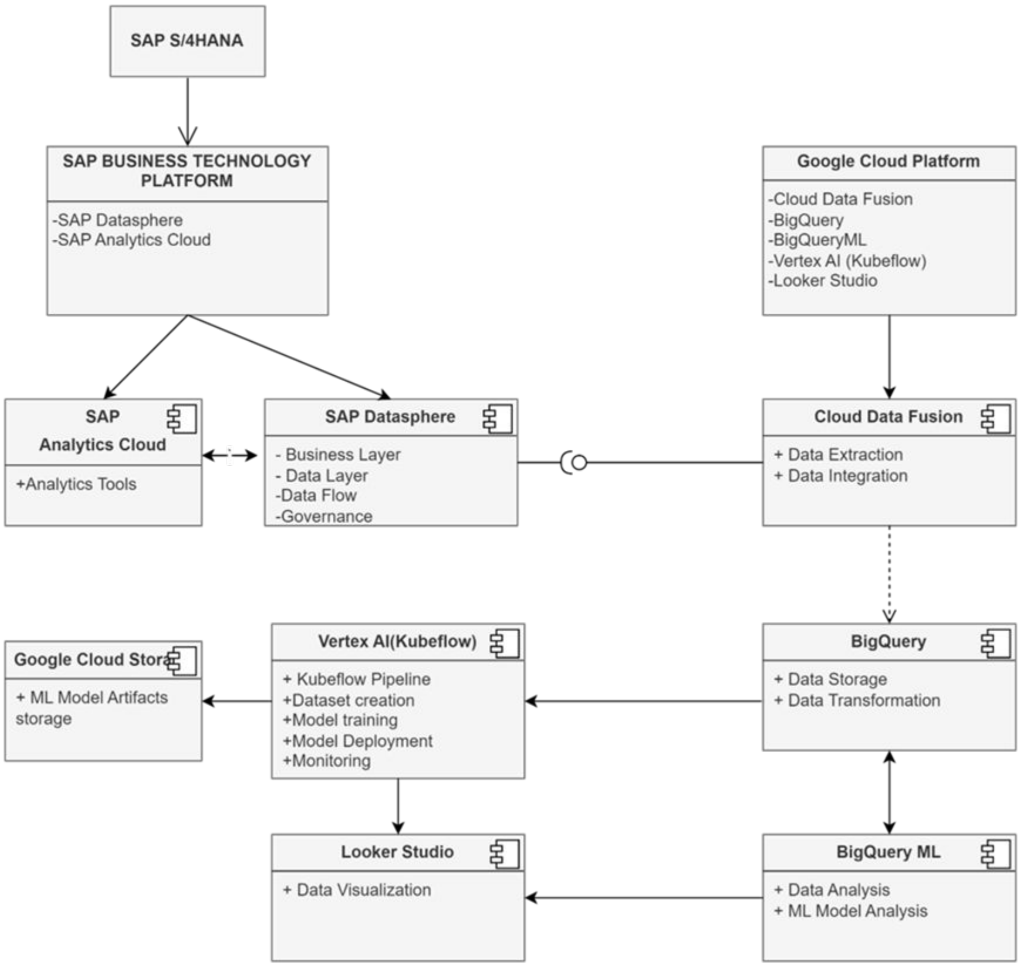
Architecture Diagram

BUSINESS SYSTEM PLANNING DIAGRAM



BSP Diagram

ARCHITECTURE COMPONENT DIAGRAM



UML Component Diagram

CLUSTERING

- Different regions in the existing customer markets need to be clustered based on their geographical location
  - Input table creation for clustering
    - Created a table MPD_CustomerData_ML which is used for clustering based on Countries.

```
Create Table mpd_g6_data.MPD_CustomerData_ML AS
SELECT MPD_CustomerAttr.CustomerId, CustDescr,City, Country,
CASE Country
  WHEN 'BR' THEN 1
  WHEN 'CH' THEN 2
  WHEN 'DE' THEN 3
  WHEN 'FR' THEN 4
  WHEN 'US' THEN 5
  ELSE NULL
END AS Country_Num
FROM mpd_g6_data.MPD_CustomerAttr
JOIN mpd_g6_data.MPD_CustomerText
ON MPD_CustomerAttr.CustomerId = MPD_CustomerText.CustomerId
order by Country;
```

MPD_CustomerData_ML QUERY SHARE COPY SNAPSHOT DELETE EXPORT

SCHEMA DETAILS PREVIEW LINEAGE DATA PROFILE DATA QUALITY

Row	CustomerId	CustDescr	City	Country	Country_Num
1	9000	Brazil Bikes	Rio de Janeiro	BR	1
2	12000	Belo Bikes	Belo Horizonte	BR	1
3	5000	Beantown Bikes	Sao Paulo	BR	1
4	4000	Peachtree Bikes	Shanghai	CH	2
5	7000	Furniture City Bikes	Peking	CH	2
6	10000	Valley Bikes	Wuhan	CH	2
7	20000	Neckarrad	Heidelberg	DE	3
8	18000	Drahtesel	Leipzig	DE	3
9	15000	Bavaria Bikes	München	DE	3
10	24000	Velodrom	Magdeburg	DE	3
11	14000	Alster Cycling	Hamburg	DE	3
12	17000	Cruiser Bikes	Hannover	DE	3
13	23000	Red Light Bikes	Hamburg	DE	3
14	13000	Airport Bikes	Frankfurt	DE	3
15	16000	Capital Bikes	Berlin	DE	3
16	22000	Velo Seine	Paris	FR	4

CLUSTERING

- Different regions in the existing customer markets need to be clustered based on their geographical location
  - Creation of K-means model for clustering based on geographical location
    - Used K Means clustering ML algorithm using BigQueryML to cluster markets based on geographical location

```
CREATE OR REPLACE MODEL `vlba-2024-mpd-group-6.mpd_g6_data.mpd_customer_clusters`
OPTIONS(
  MODEL_TYPE = 'KMEANS',
  NUM_CLUSTERS = 5,
  STANDARDIZE_FEATURES = FALSE
) AS
SELECT
  Country_Num
FROM
  `vlba-2024-mpd-group-6.mpd_g6_data.MPD_CustomerData_ML`;
```

CLUSTERING

- Different regions in the existing customer markets need to be clustered based on their geographical location
    - Create a table to get the output of the clustering model

```
Create table `mpd_g6_data.MPD_Clustering_Output`
AS

SELECT
  Centroid_id as Cluster,
  CustomerId,
  CustDescr as `Customer_name`,
  City,
  Country

FROM
  ML.PREDICT(MODEL `vlba-2024-mpd-group-6.mpd_g6_data.mpd_customer_clusters`, TABLE `vlba-2024-mpd-group-6.
mpd_g6_data.MPD_CustomerData_ML`)
```

MPD_Clustering_Output   QUERY   SHARE   COPY   SNAPSHOT   DELETE   EXPORT

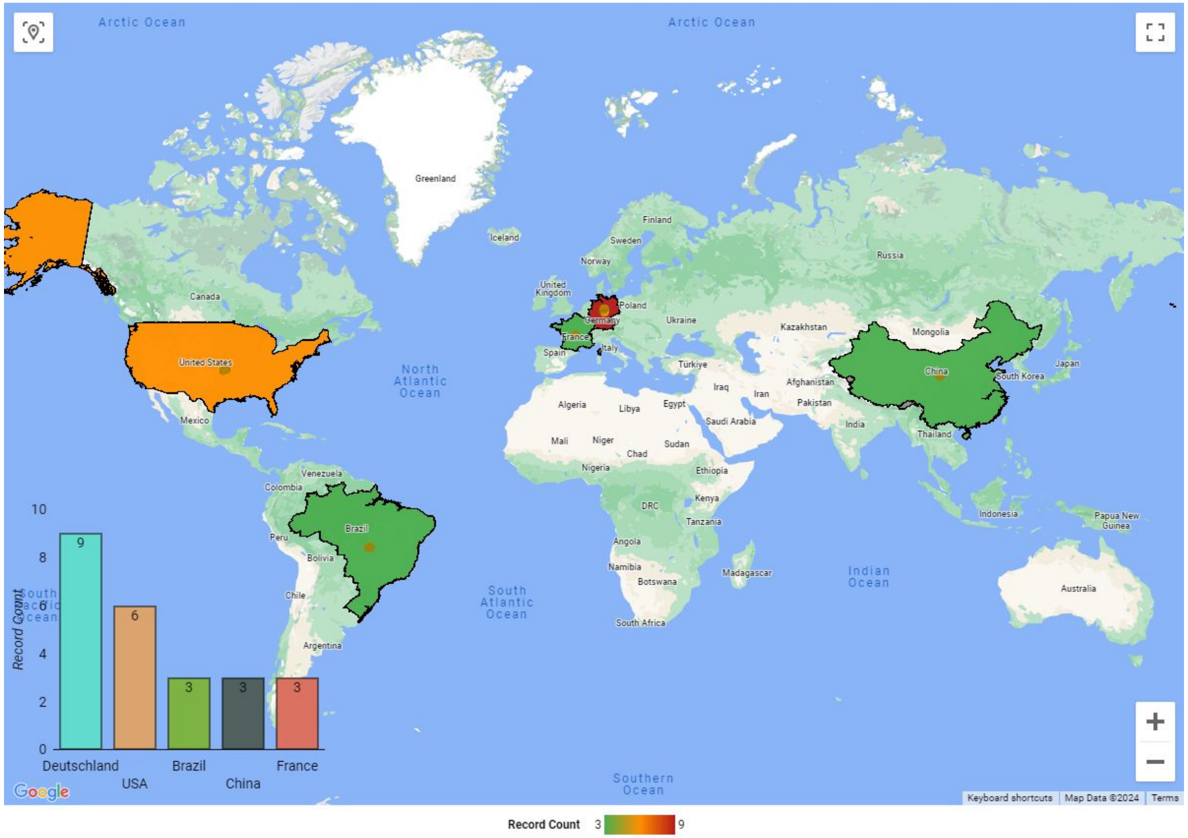
SCHEMA   DETAILS   PREVIEW   LINEAGE   DATA PROFILE   DATA QUALITY

Row	Cluster	CustomerId	Customer_name	City	Country
1	1	12000	Belo Bikes	Belo Horizonte	Brazil
2	1	9000	Brazil Bikes	Rio de Janeiro	Brazil
3	1	5000	Beantown Bikes	Sao Paulo	Brazil
4	2	11000	DC Bikes	Washington DC	USA
5	2	2000	Big Apple Bikes	New York City	USA
6	2	8000	Motown Bikes	Detroit	USA
7	2	6000	Windy City Bikes	Chicago	USA
8	2	3000	Philly Bikes	Philadelphia	USA
9	2	1000	Rocky Mountain Bikes	Denver	USA
10	3	21000	Velo d'Atlantique	Brest	France
11	3	19000	Velo de la Mer	Marseille	France
12	3	22000	Velo Seine	Paris	France
13	4	24000	Velodrom	Magdeburg	Deutschland
14	4	14000	Alster Cycling	Hamburg	Deutschland
15	4	17000	Cruiser Bikes	Hannover	Deutschland
16	4	16000	Capital Bikes	Berlin	Deutschland

CLUSTERING

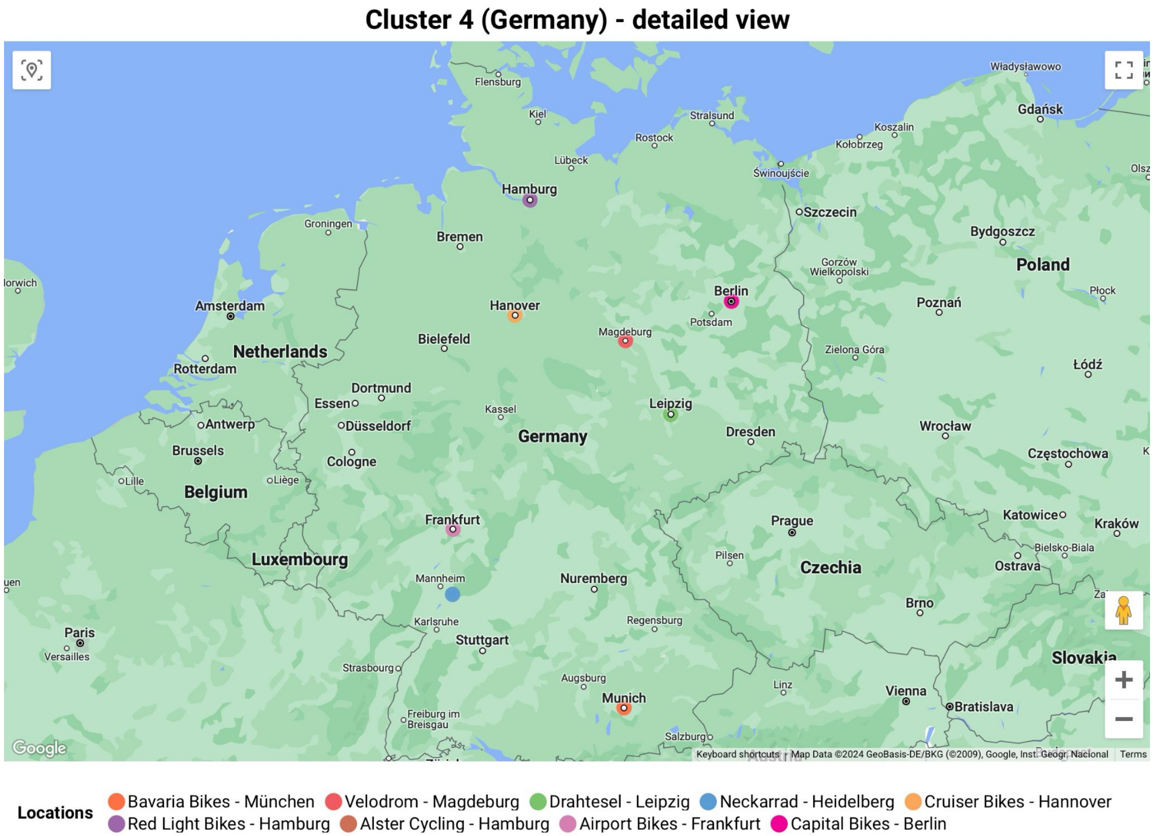
- Different regions in the existing customer markets need to be clustered based on their geographical location
  - Visualisation of the clusters

Customers clustered by country



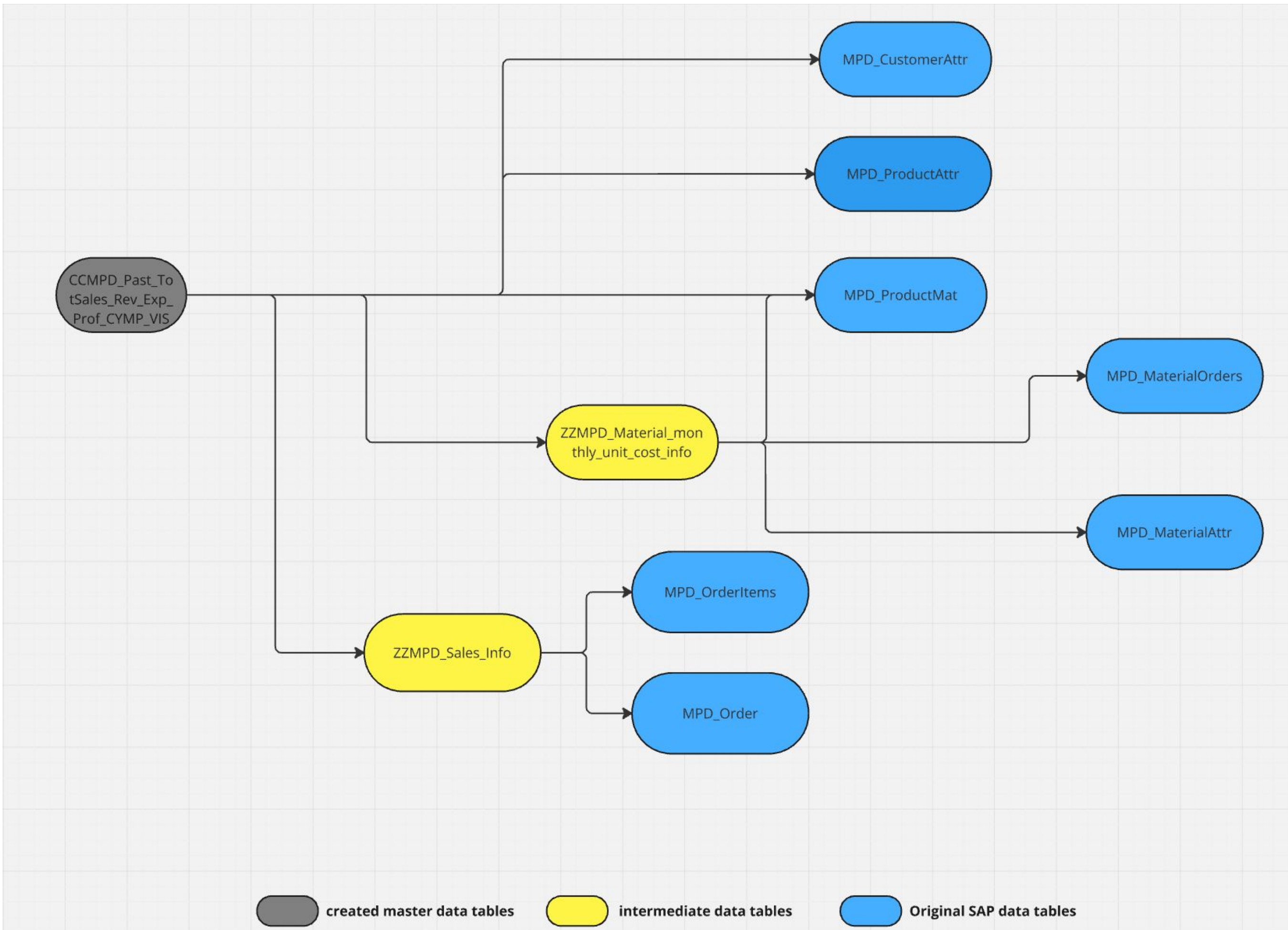
CLUSTERING

- Different regions in the existing customer markets need to be clustered based on their geographical location
  - Visualisation of the clusters



Cluster Germany detailed view

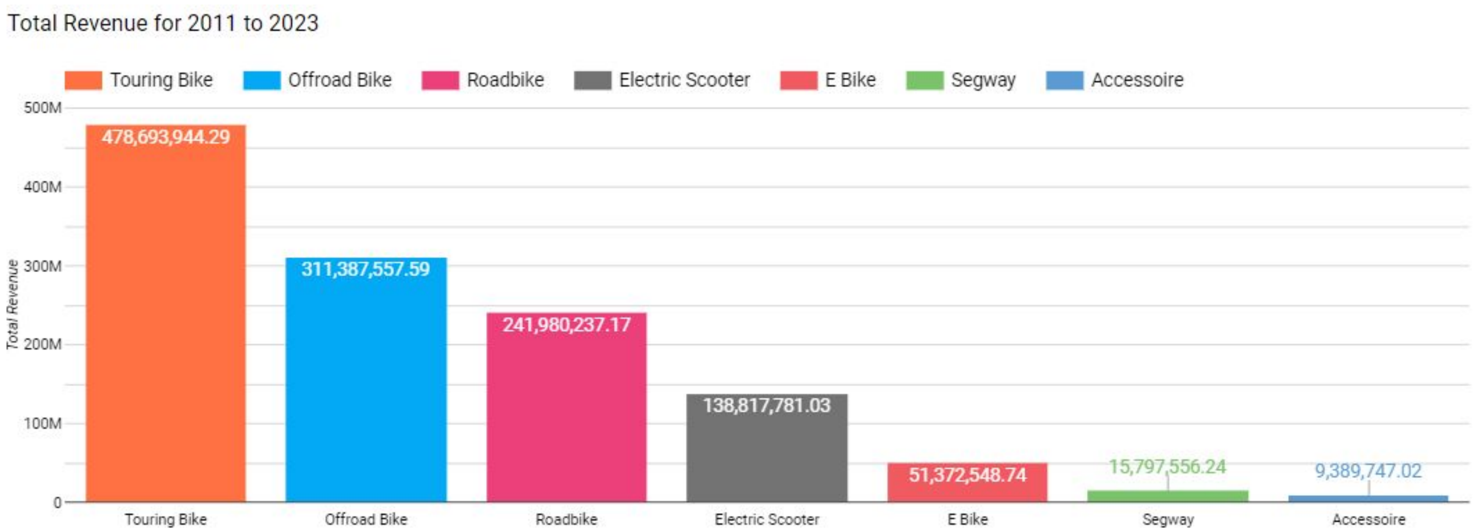
PRODUCT CATEGORY SELECTION



Overview of table creation for past data

PRODUCT CATEGORY SELECTION

- Identify the two sensible product categories
  - Analyse the total revenue of each product category from 2011 to 2023

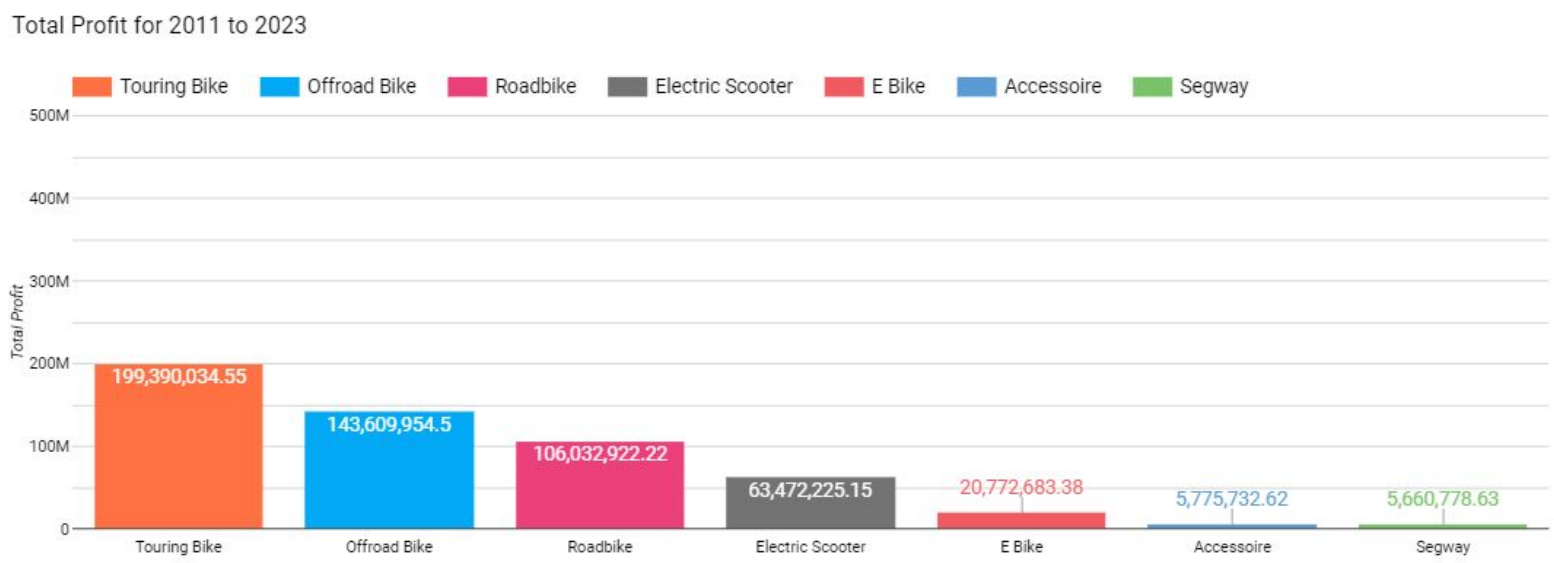


Total revenue from 2011 to 2023 by product category

- Touring Bike and Offroad Bike achieved the highest revenue in all product categories

PRODUCT CATEGORY SELECTION

- Identify the two sensible product categories
  - Analyse the profit of each product category from 2011 to 2023



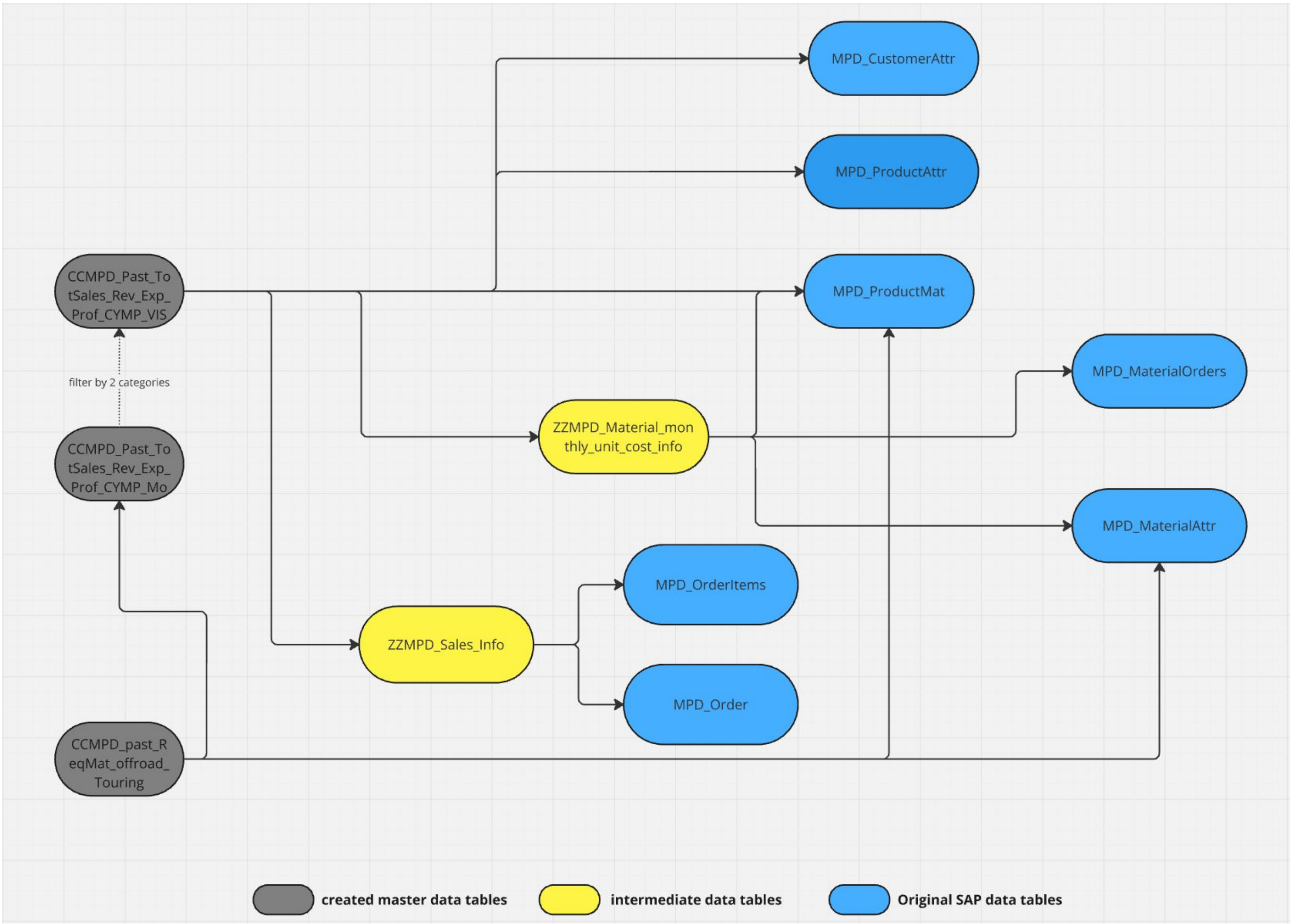
Total profit from 2011 to 2023 by product category

- The total profit over the same period Touring Bike and Offroad Bike have also the highest profit

PRODUCT CATEGORY SELECTION

- Considering total revenue and total profit together was a strategic decision
- High revenue does not necessarily correlate with high profitability
- A product category with substantial revenue but low profit margins may not be sustainable in the long term
- Comprehensive assessment by considering both revenue and profit
- Strategy for stability and sustainability: Touring Bike and Offroad Bike

PRODUCT CATEGORY SELECTION



Overview of table creation for past data

MACHINE LEARNING FORECAST

Sales and Profit Forecasting:

- Utilized an Exponential Smoothing model to accurately forecast product sales and profits
- Data for the model was sourced from the BigQuery table **"CCMPD_Past_TotSales_Rev_Exp_Prof_CYMP_Model"**
- The forecasted results were stored back in BigQuery in the table **"CCMPD_Sal_Prof_Seasonal_Model_Output"**

IA DETAILS PREVIEW LINEAGE DATA PROFILE DATA QUALITY

year	month	ProductId	ProdCat	Country	total_sales_qty	total_revenue	total_expense	monthly_profit
2011	1	DXTR2000	Touring Bike	BR	9	28503.18	18452.29	10050.89
2011	1	ORHT1000	Offroad Bike	BR	8	12416.0	7984.88	4431.12
2011	1	ORHT2000	Offroad Bike	BR	27	49478.43	29239.65	20238.78
2011	2	DXTR1000	Touring Bike	BR	25	75847.46	51289.94	24557.52
2011	2	DXTR2000	Touring Bike	BR	27	83724.5	55356.89	28367.61
2011	2	DXTR3000	Touring Bike	BR	12	37662.36	24619.17	13043.19
2011	2	ORHT1000	Offroad Bike	BR	22	37219.24	21590.8	15628.44
2011	2	ORHT2000	Offroad Bike	BR	12	21173.14	12364.92	8808.22
2011	2	ORMN1000	Offroad Bike	BR	22	52799.16	34332.32	18466.84
2011	2	PRTR1000	Touring Bike	BR	10	31621.83	21825.6	9796.23
2011	2	PRTR2000	Touring Bike	BR	4	13413.8	8723.08	4690.72
2011	2	PRTR3000	Touring Bike	BR	1	3094.15	2182.56	911.59
2011	3	DXTR1000	Touring Bike	BR	6	18260.34	12532.88	5727.46
2011	3	DXTR2000	Touring Bike	BR	8	25243.64	16710.5	8533.14
2011	3	DXTR3000	Touring Bike	BR	7	20370.0	14619.9	5750.1
2011	3	ORHT1000	Offroad Bike	BR	15	24848.52	14863.95	9984.57

Input table for the ML model

MACHINE LEARNING FORECAST

CCMPD_Sal_Prof_Seasonal_Model_Output

QUERY SHARE COPY SNAPSHOT DELETE EXPORT

SCHEMA DETAILS PREVIEW LINEAGE DATA PROFILE DATA QUALITY

Row	Country	ProductId	ProdCat	Year	Month	Forecasted_Sales_Quantity	Forecasted_Revenue	Forecasted_Expense	Forecasted_Profit
1	BR	ORHT1000	Offroad Bike	2025	1	20	43876.355235813...	20675.631855531...	23831.551791813457
2	BR	ORHT1000	Offroad Bike	2025	2	17	38849.142345215...	18852.061377271...	19627.195594003395
3	BR	ORHT1000	Offroad Bike	2025	3	11	28379.46762772584	13422.749769586...	14157.381918784027
4	BR	ORHT1000	Offroad Bike	2025	4	29	66399.064007893...	30879.927694764...	36163.738207194583
5	BR	ORHT1000	Offroad Bike	2025	5	37	83382.454803793...	40743.859107377...	42160.384093082896
6	BR	ORHT1000	Offroad Bike	2025	6	53	118077.29528626...	55579.481222989...	63406.28059599195
7	BR	ORHT1000	Offroad Bike	2025	7	40	92035.9292751949	46488.245870281...	43774.888045212734
8	BR	ORHT1000	Offroad Bike	2025	8	38	86573.563500521...	41215.804441727...	45422.969791244024
9	BR	ORHT1000	Offroad Bike	2025	9	43	96750.161465466...	48017.488862693...	48480.745929675992
10	BR	ORHT1000	Offroad Bike	2025	10	12	35004.149845802...	17334.866097639...	16800.414918002971
11	BR	ORHT1000	Offroad Bike	2025	11	14	39208.366738867...	19650.380580542...	18782.141579674771
12	BR	ORHT1000	Offroad Bike	2025	12	17	47163.2194476976	23117.480441369...	23987.422037276949
13	BR	ORHT2000	Offroad Bike	2025	1	9	24444.85528934616	10383.175740080...	14718.468118564451
14	BR	ORHT2000	Offroad Bike	2025	2	18	45504.127827135...	20116.136678369...	25446.385030566169
15	BR	ORHT2000	Offroad Bike	2025	3	17	43772.6440017827	18689.433728476...	24962.795609138226
16	BR	ORHT2000	Offroad Bike	2025	4	23	53808.078701681...	24222.022427795...	29563.008150860733

Output table of the ML model

MACHINE LEARNING FORECAST

Material Requirements Forecasting:

- Projections were based on the forecasted sales data generated by the machine learning model
- Created the table **"CCMPD_future_ReqMat_forecast"** to project future material requirements

CCMPD_future_ReqMat_forecast QUERY SHARE COPY SNAPSHOT DELETE

SCHEMA DETAILS PREVIEW LINEAGE DATA PROFILE DATA QUALITY

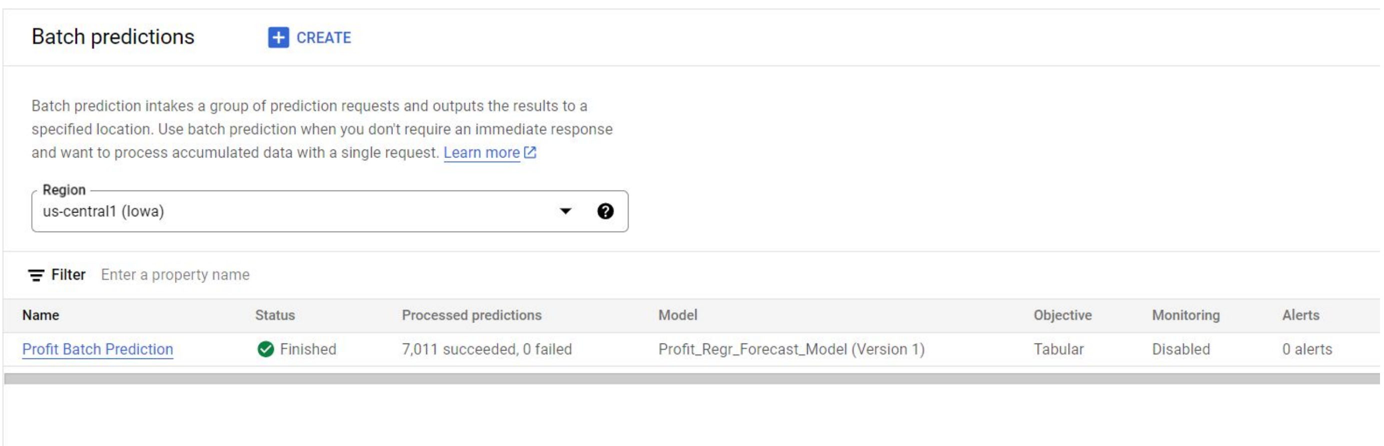
Row	Country	Year	Month	MatId	Mat_group	Forecasted_Material_Qty
1	BR	2025	1	CBLA1000	Paint	27.5
2	CH	2025	1	CBLA1000	Paint	34.5
3	DE	2025	1	CBLA1000	Paint	38.5
4	FR	2025	1	CBLA1000	Paint	33.75
5	US	2025	1	CBLA1000	Paint	21.25
6	BR	2025	1	CRED1000	Paint	19.75
7	CH	2025	1	CRED1000	Paint	20.0
8	DE	2025	1	CRED1000	Paint	20.75
9	FR	2025	1	CRED1000	Paint	14.75
10	US	2025	1	CRED1000	Paint	18.0
11	BR	2025	1	CSIL1000	Paint	18.75
12	CH	2025	1	CSIL1000	Paint	30.75
13	DE	2025	1	CSIL1000	Paint	35.75
14	FR	2025	1	CSIL1000	Paint	29.25
15	US	2025	1	CSIL1000	Paint	37.25
16	BR	2025	1	OFRA1000	Frames	20.0

Material Required Forecast table

MACHINE LEARNING FORECAST

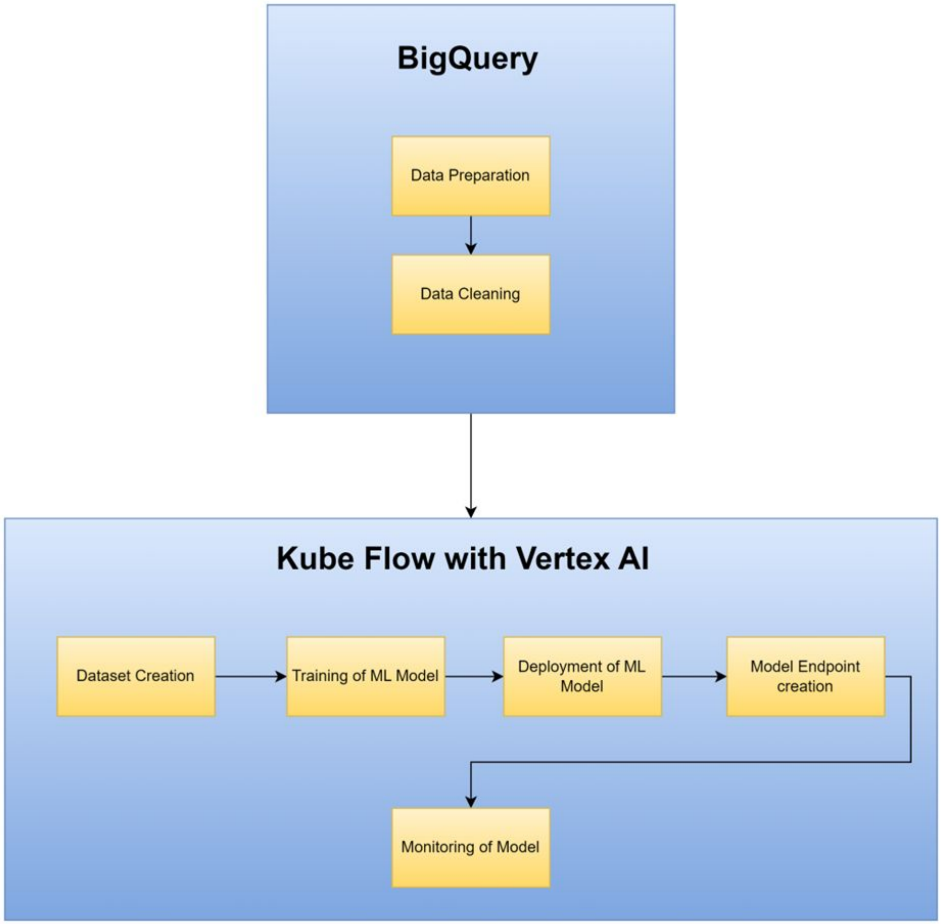
Kubeflow pipeline with Vertex AI:

- Developed a Kubeflow pipeline using the Kubeflow SDK
- Integrated the pipeline with Vertex AI to deploy an AutoML regression model for profit prediction
- The deployed AutoML regression model enables further profit forecasting through batch predictions by providing necessary input data to the deployed model



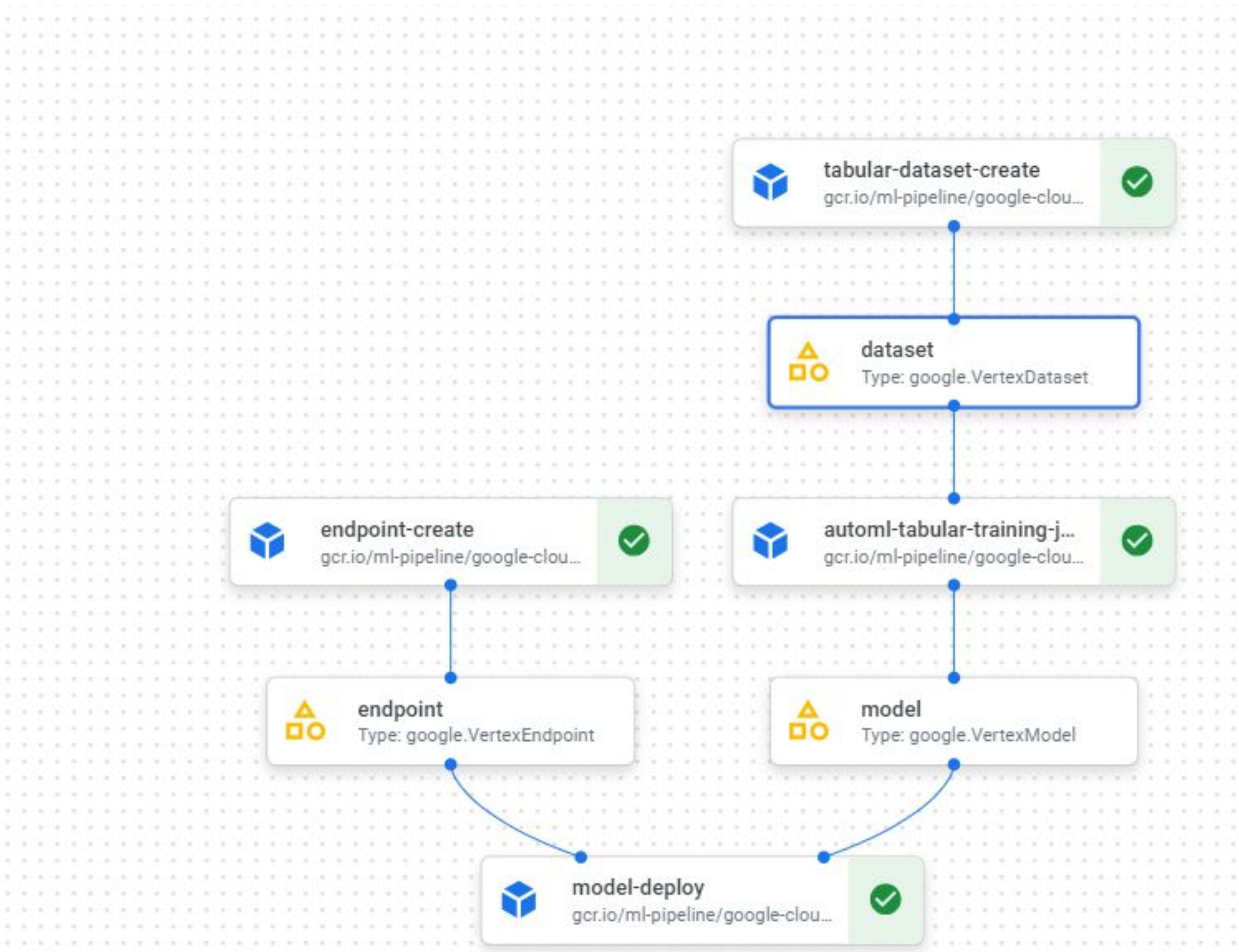
Batch prediction in Vertex AI

KUBEFLOW PIPELINE WITH VERTEX AI



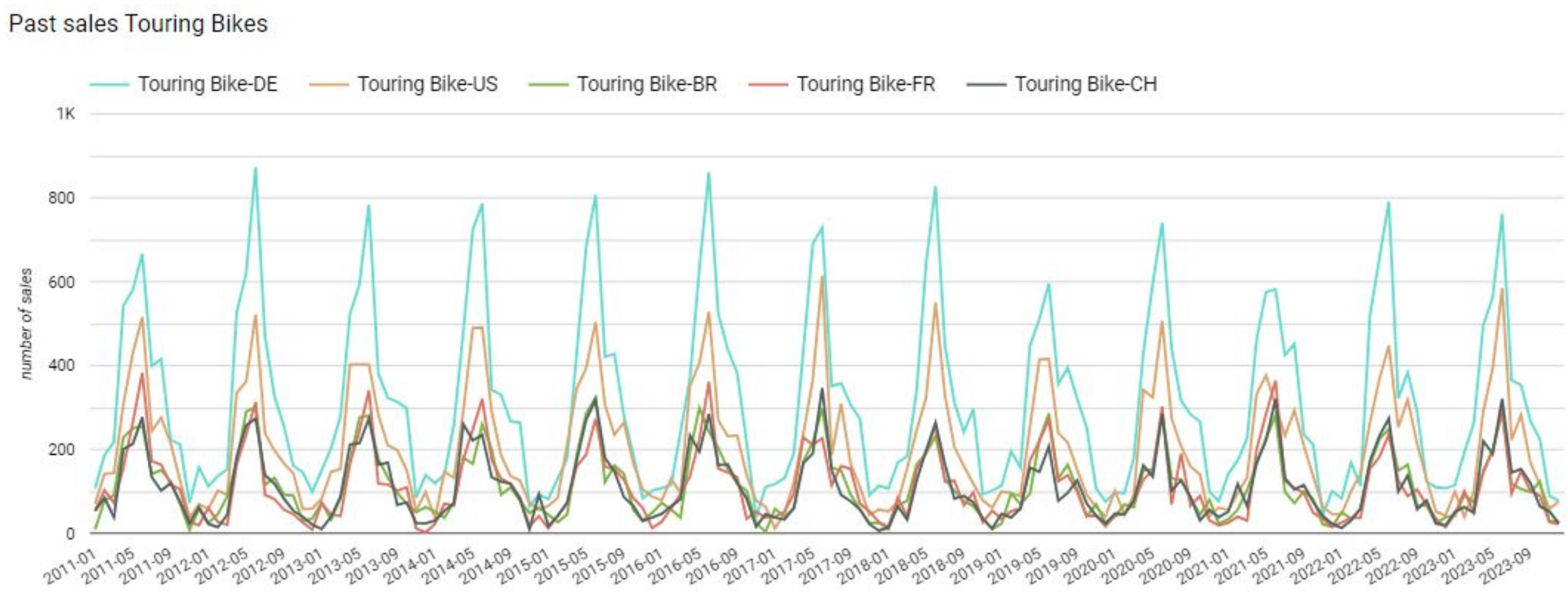
- This process involves creating and deploying a machine learning pipeline using Kubeflow SDK integrated with Vertex AI, starting with data extraction from a BigQuery table for preparation and cleaning
- A Kubeflow pipeline is defined and compiled, including steps for dataset creation, model training using AutoML, model deployment, endpoint creation for predictions using Kubeflow SDK along with monitoring of the model

KUBEFLOW PIPELINE WITH VERTEX AI



Kubeflow Pipeline components flow

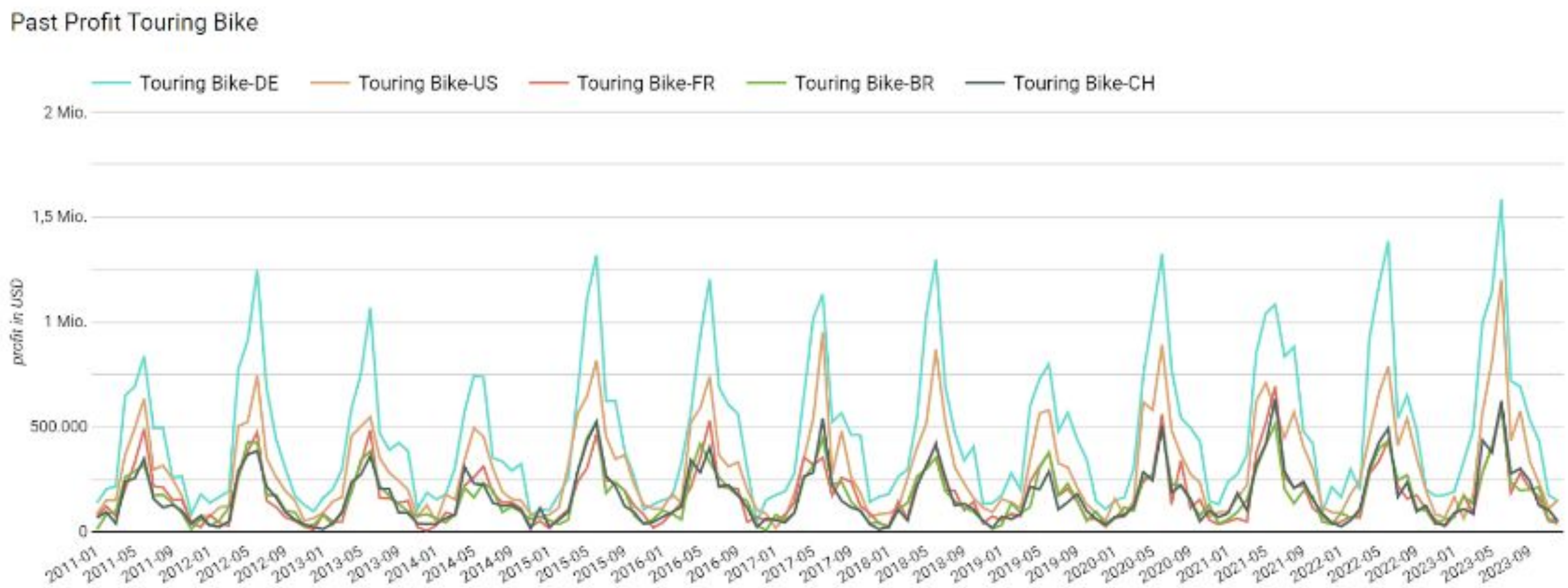
EVALUATION AND VISUALISATION



Past sales of Touring Bikes

- Steady but slow decline in sales volumes for Offroad Bikes and Touring Bikes
- Decline observed across all clustered customer markets
- Corresponding decrease in required materials

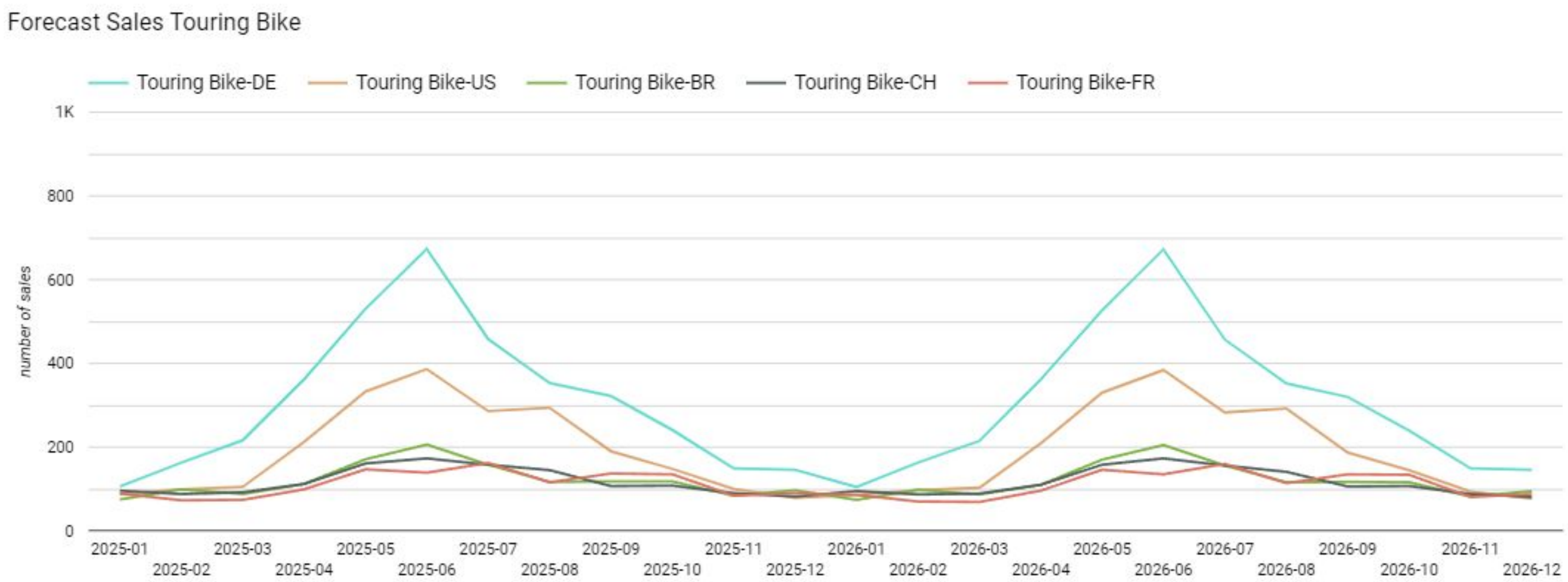
EVALUATION AND VISUALISATION



Past profit of Touring Bikes

- Slight increase in profit margins despite declining sales
- Indicates improved cost management or higher profit per unit sold

EVALUATION AND VISUALISATION



Forecast sales of Touring Bikes

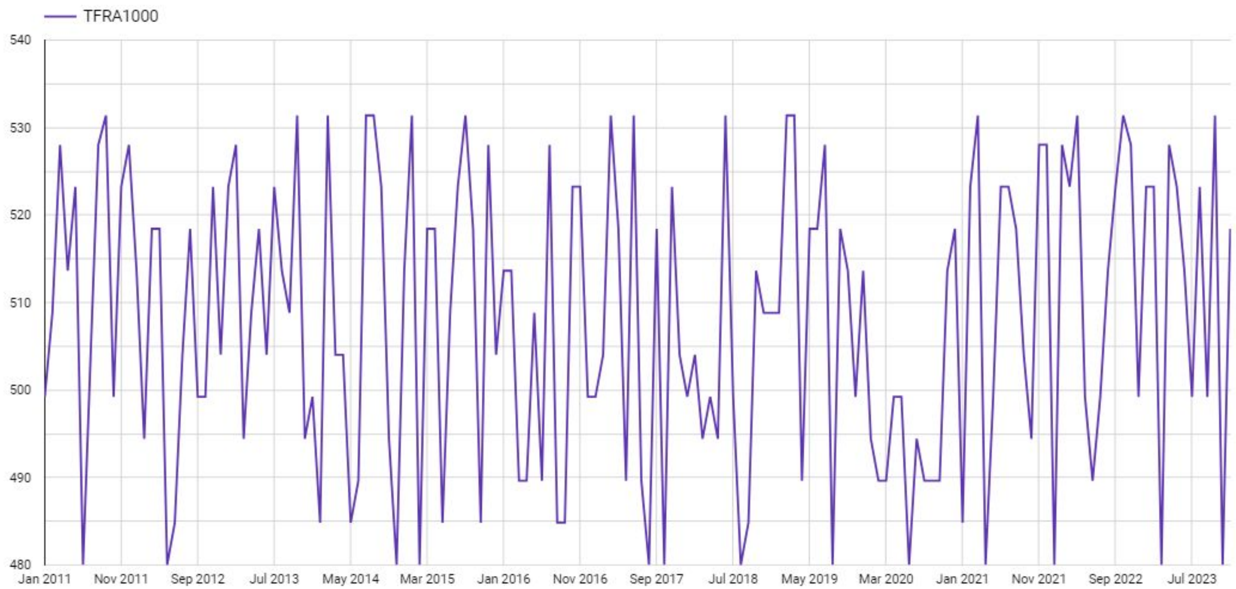
- Projected decline in both sales and profits for Offroad Bikes and Touring Bikes

EVALUATION AND VISUALISATION

Forecast profit of Touring Bikes

- Indicates potential challenges in maintaining profitability with shrinking sales volume

INTERPRETATION - PROCUREMENT



Price fluctuations for the material TFRA1000

- Material Order and Price Fluctuations
  - Price constantly changing for every order for every material
- Strategy to avoid Price Fluctuations
  - collaboration with suppliers to negotiate bulk purchasing agreements for non perishable materials
  - Benefit: significantly lower unit cost due to economies of scale
  - prevent overstocking and ensure continuous material availability

INTERPRETATION - PRODUCTION PLANNING

- Align production with bulk purchasing and seasonal demand
  - Maximize efficiency and reduce costs
- Increase production capacity with the availability of large quantities of material
  - Adjust allocations such as worker to match increased production need
  - Analyze production flow to identify bottlenecks and optimize processes
- Optimizing our machine learning model for more accurate and long term demand forecast
  - Improve production planning with precise data
  - Prevent material shortages or excesses

INTERPRETATION - MARKETING

- Market Research
  - Conduct research to understand regional preferences, regulations and competitors
  - Adjust localized strategies base on market insights
- Pre Season Promotion
  - Launch country custom campaigns from December to March
  - Benefits: Build awareness and anticipation
- Event Marketing
  - Organize or participate in events, trade shows and corporate expos in different countries to promote our mobility devices

INTERPRETATION- BUSINESS DECISIONS

- Market opportunity
  - Growing market for sustainable and eco friendly transportation solutions
  - Increasing demand for bike rental driven by environmental awareness etxc
- Bike rental system
  - Leverage existing infrastructure and expertise in the mobility sector
  - Utilize existing logistics, maintenance facilities, and customer service infrastructure
  - Reduce dependency on existing product lines

Thank You!