Capstone Project (30.01 – 02.02.2024)

Lecture Business Intelligence Systems (BIS)

Prof. Dr. Alexander Mädche

Jonas Gunklach (KIT)

Dr. Dominik Jung, Dr. Sebastian Seifen,
Patrick Gärtner, Cindy Tröger-Hegedusch (Porsche AG)


PORSCHE

Patrick Gärtner

Head Data Science & Governance
Sales & Marketing

Cindy Tröger-Hegedusch

Specialist Sales Planning &
Steering

Dr. Dominik Jung

Senior Data Scientist

Dr. Sebastian Seifen

Senior Data Scientist,
Product Owner Data Science

Mission

"In the beginning,
I looked around and could not
find the car I dreamed of.
So I decided to build it myself."

Ferry Porsche

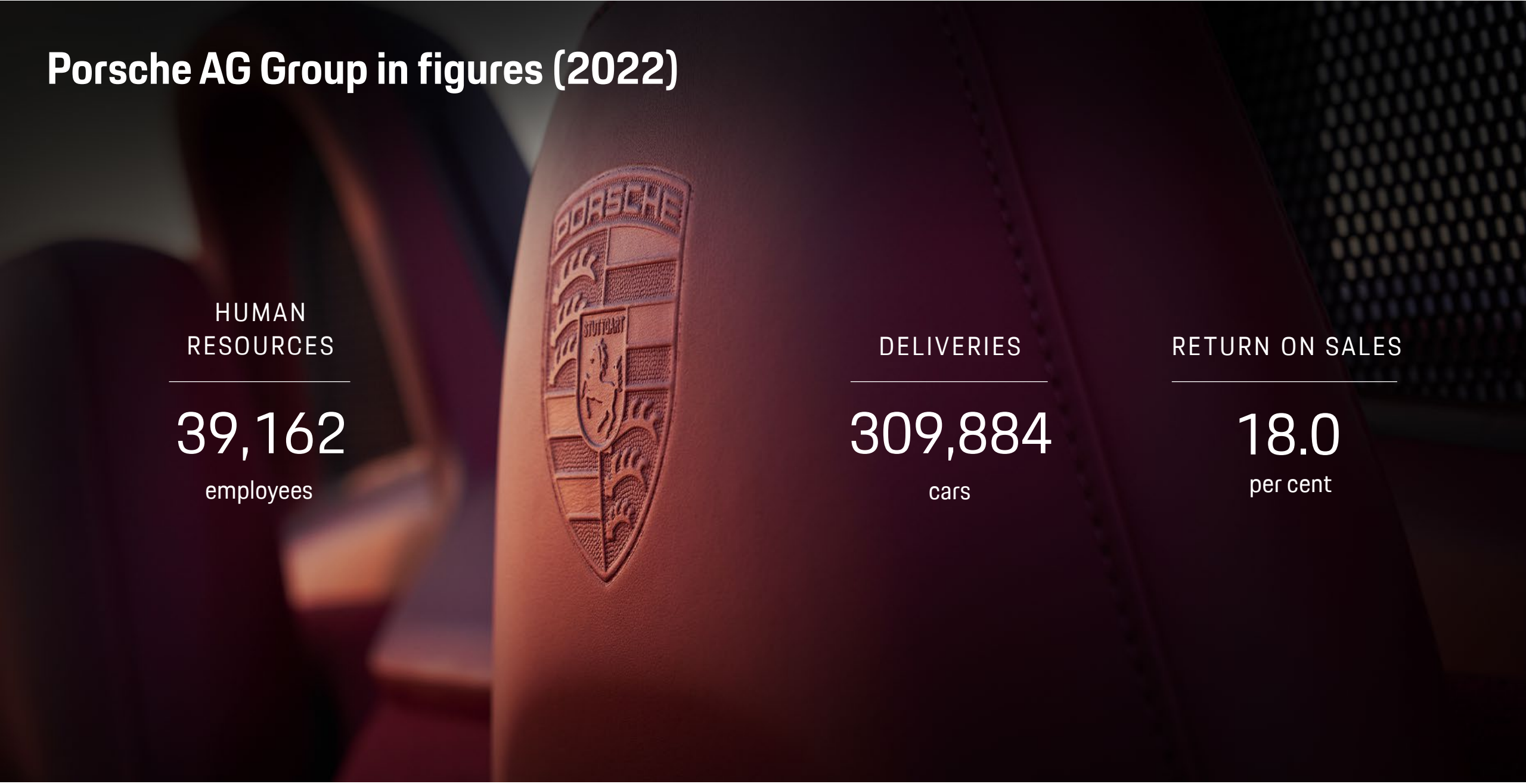
AND THIS IS HOW WE AIM TO SHAPE
THE FUTURE OF THE SPORTS CAR.

Vision

Porsche model range

911 Carrera models: Fuel consumption combined: 11.4 - 10.1 l/100 km (WLTP) | CO_2 emissions combined 259 - 229 g/km (WLTP)
Taycan Sport limousine models: Electricity consumption combined: 24.1 - 19.6 kWh/100 km (WLTP) | CO2 emissions combined: 0 g/km (WLTP) | Electric range: 370 - 512 km; Electric range city: 440 - 630 km
Macan GTS: Fuel consumption* combined (WLTP) 11.7 - 11.3 l/100 km | CO_2 emissions* combined (WLTP) 265 - 255 g/km
718 Cayman models: Fuel combined: 13.2 -8.9 l/100 km (WLTP) | CO2 emissions combined: 299 -201 g/km (WLTP)
Cayenne GTS: Fuel consumption combined: 14.1 - 13.3 l/100km (WLTP) | CO2 emissions combined: 319 - 301 g/km (WLTP)
Panamera GTS: Fuel consumption combined: 13.1 - 12.1 l/100km (WLTP) | CO2 emissions combined: 296 - 275 g/km (WLTP)

Porsche AG Group in figures (2022)



HUMAN RESOURCES

39,162

employees

DELIVERIES

309,884

cars

RETURN ON SALES

18.0

per cent

Porsche markets in 2023

> 900

POINTS OF SALE

> 120

MARKETS WORLDWIDE

Sales subsidiary/
regional office

Product strategy

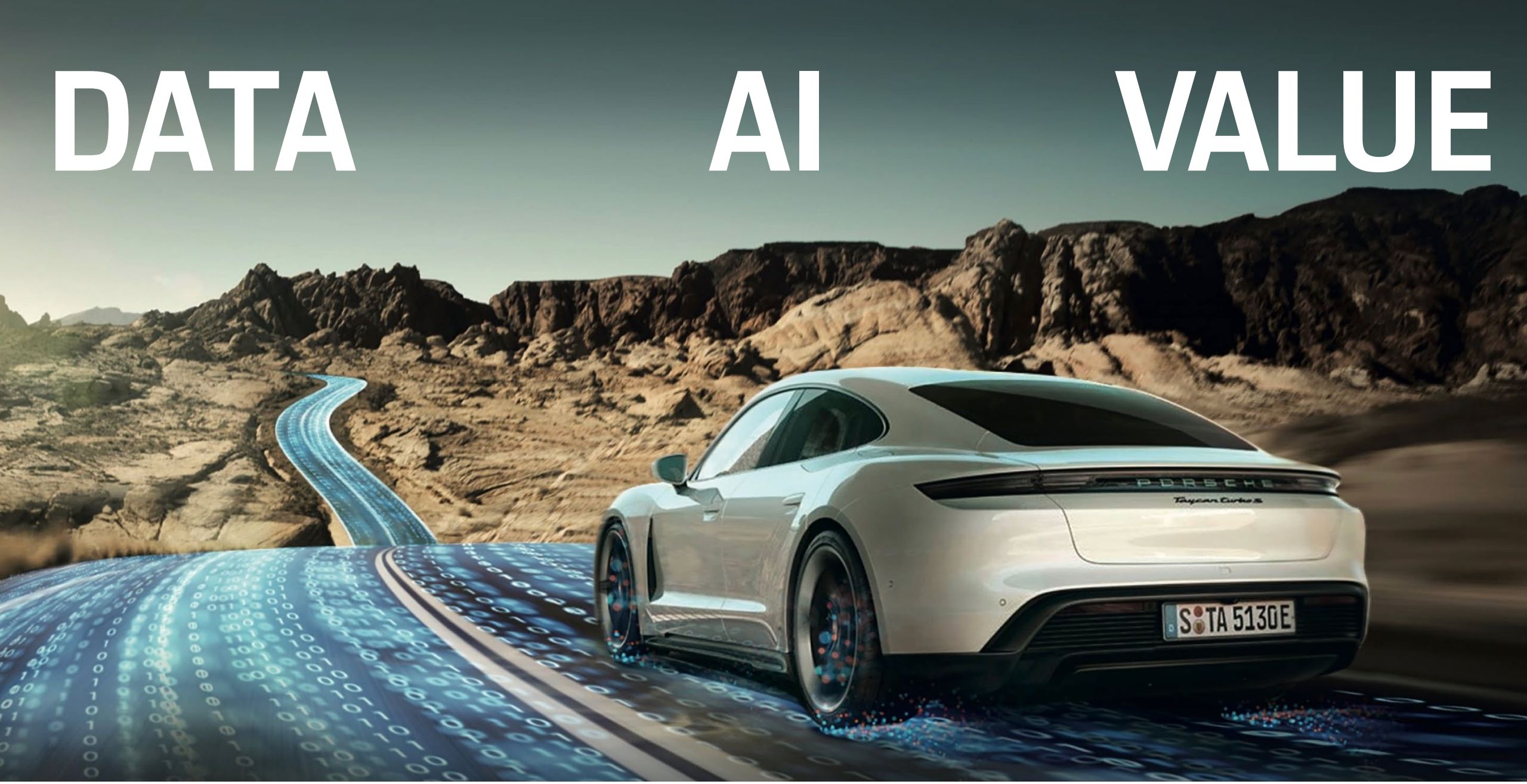
SPORTLICHKEIT
Image

HERZBLUT
Heritage

FAMILY
Core

PIONIERGEIST
Future

911 GT3 models: Fuel consumption combined: 13.4 - 12.9 l/100 km (WLTP); CO2 emissions combined: 305 - 292 g/km (WLTP)


DATA
AI
VALUE
PORSCHE
Taycan Turbo S
S TA 5130E

Customer Centricity

Customer Experiences, Smart Products & Services

Operational Excellence

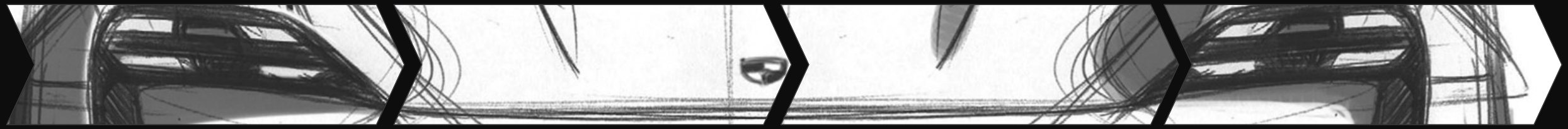
Optimized processes and decisions

Our sales vision | BALANCED DEMAND AND SUPPLY

Customer places **intent to purchase a specific Porsche** in Porsche Centre.

Customers picks up his/ her **new vehicle** at a Porsche Center.

Sales planning processes for order intakes and retails at Porsche AG



I

Strategic Planning

planning horizon:

NEXT 10 YEARS

frequency:

YEARLY

II

Forecast Planning

planning horizon:

CURRENT YEAR

frequency:

MONTHLY

III

Weekly Estimation

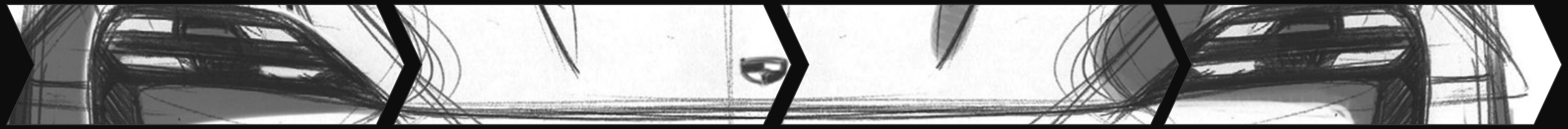
planning horizon:

CURRENT MONTH

frequency:

WEEKLY

Sales planning processes for order intakes and retails at Porsche AG



I

Strategic Planning

planning horizon:

NEXT 10 YEARS

frequency:

YEARLY

II

Forecast Planning

planning horizon:

CURRENT YEAR

frequency:

MONTHLY

III

Weekly Estimation

planning horizon:

CURRENT MONTH

frequency:

WEEKLY

01

Availability of actual data

02

Forecast planning for upcoming months

03

Communication to Management & Provision to other tools and processes

01

02

03

Availability of actual data

Forecast planning for upcoming months

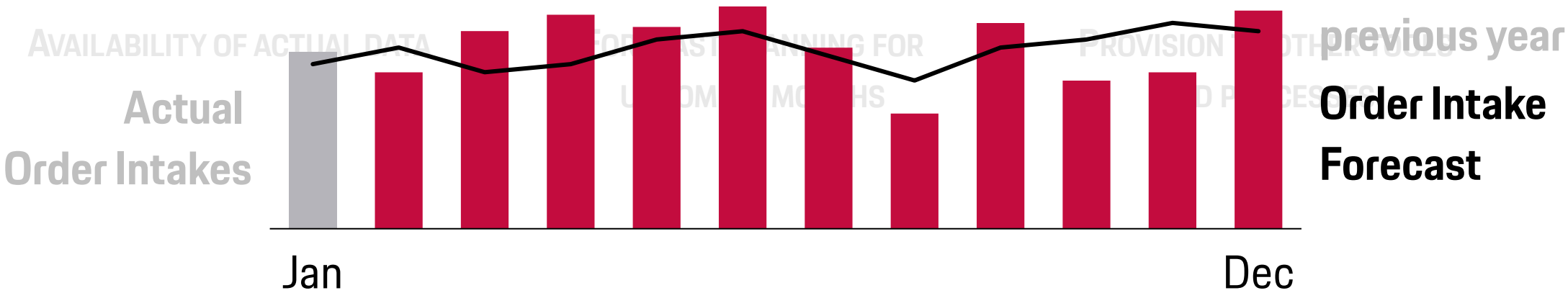
Provision to other users and processes

fictitious figures for schematic visualization

01

02

03



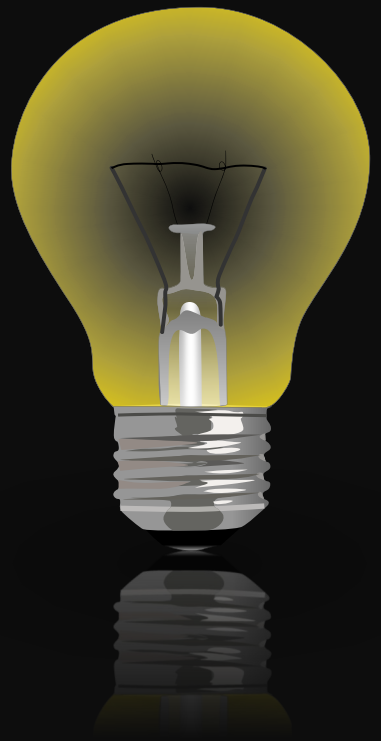
fictitious figures for schematic visualization

Current objectives in sales planning

Process improvement driven by ...

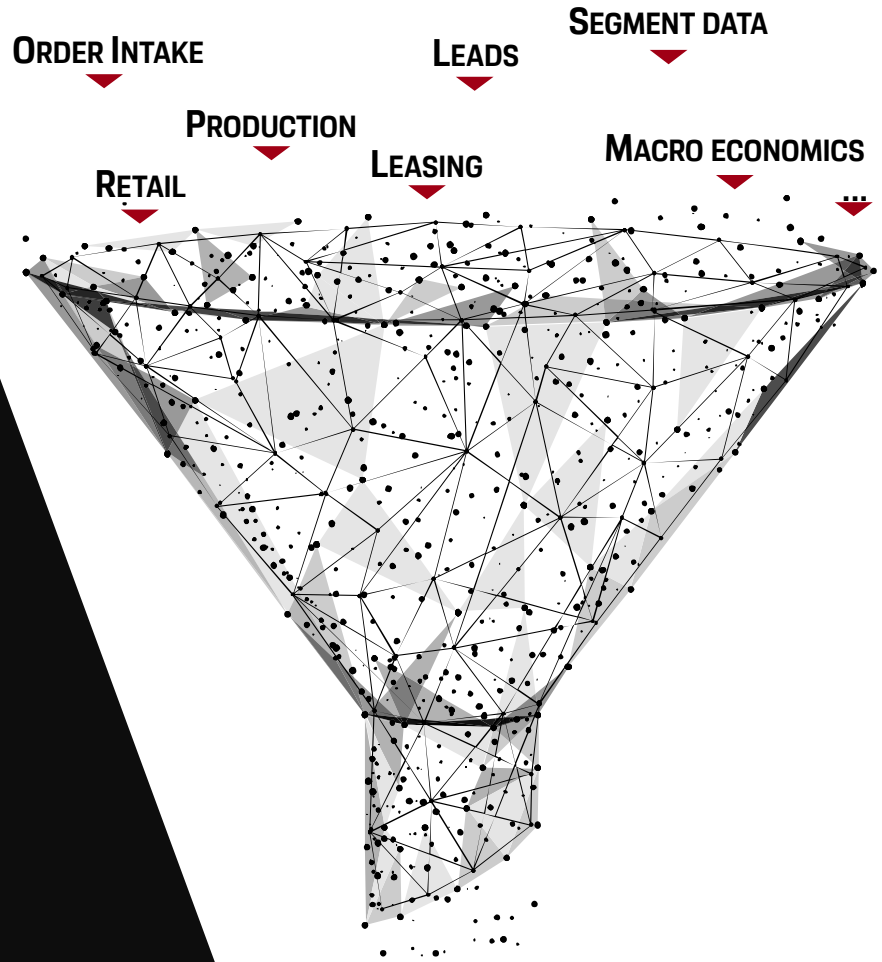
- **efficiency** in **capacity** commitment
- **enhancement** of forecast **quality** through integration of the entire available data set
- **elimination of risks** through human error in manual processes

Project Idea | AI-BASED DEMAND PROGNOSIS

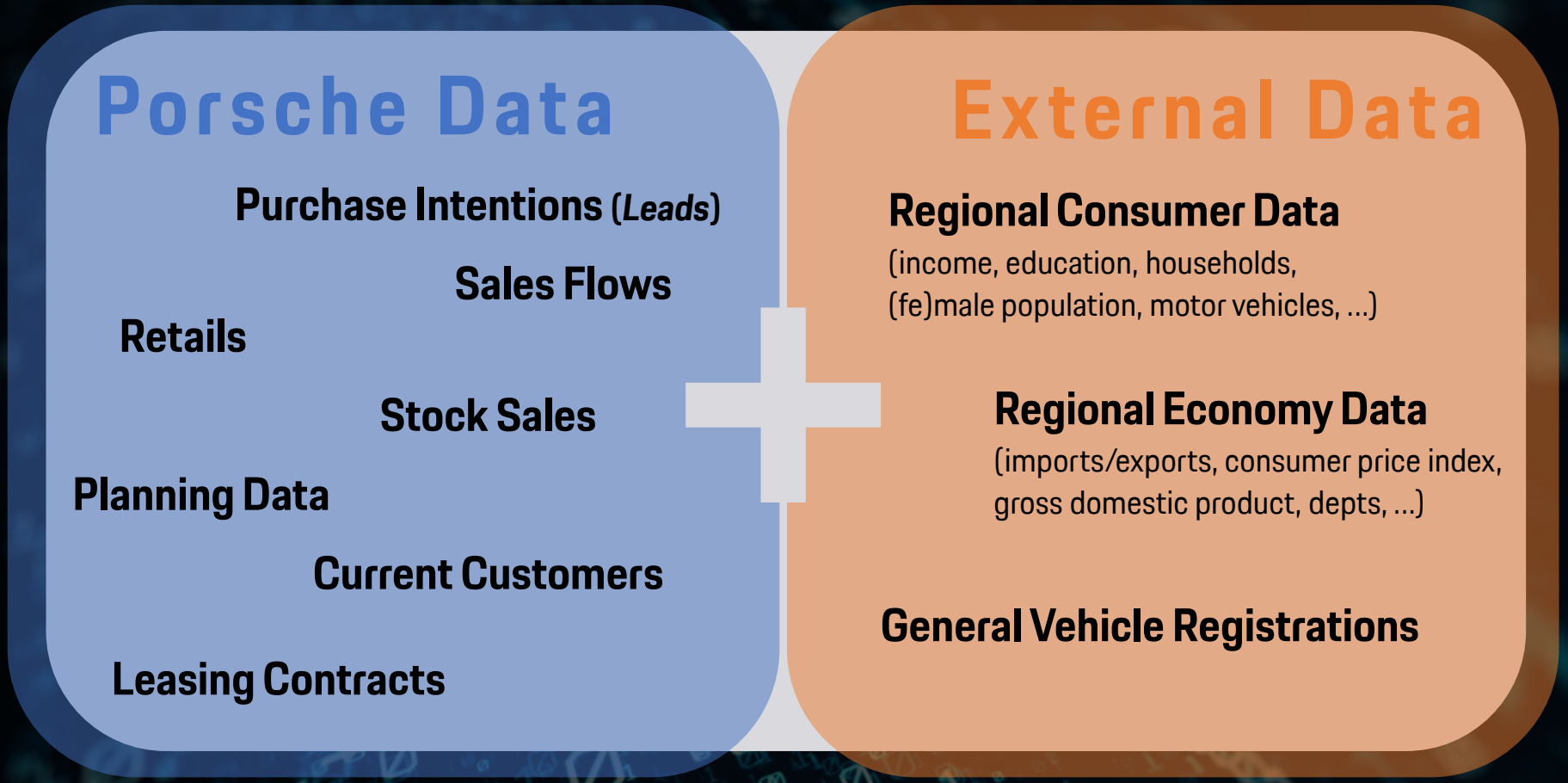


Improved forecast process by **providing markets** an automatically generated, **AI-based demand forecast** of Order Intakes for the upcoming **18 months** on market & derivative level as baseline for subsequent planning processes.

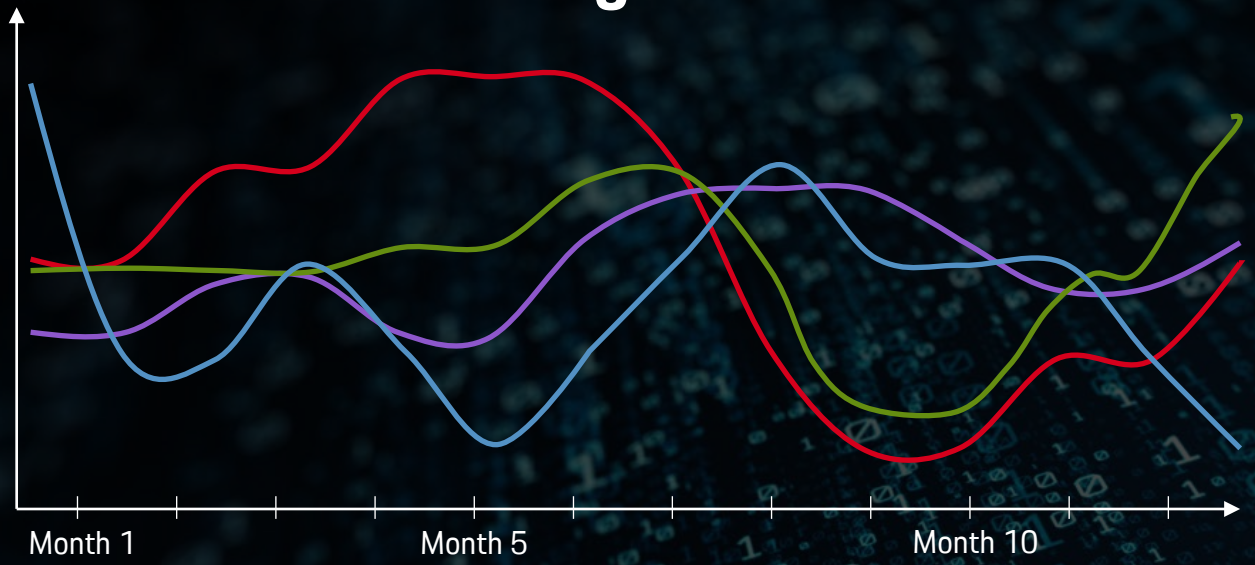
E.g.: How many customers will order a 911 Cabriolet in Germany in August this year?



The database comes from two sources: (a) Porsche-internal sales and customer data and (b) external macro-economical data



The whole data set consists of time series with values on monthly basis for four different countries/regions

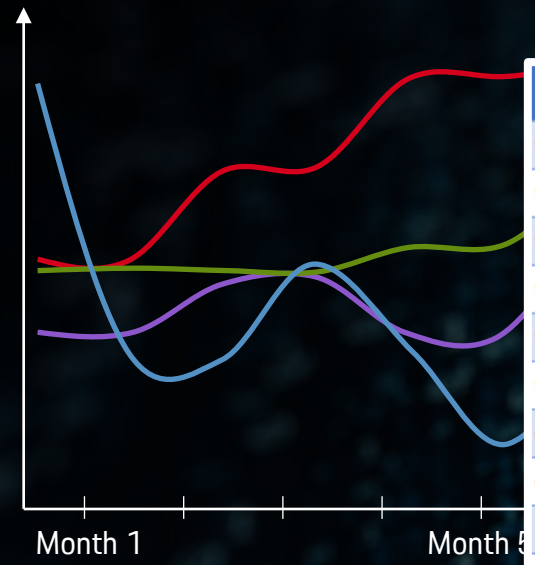


- All fields are numerical time series – except *month* (= time variable) and *countryname* (= categorical filter criterion).
- The time series begin with *Month 1* at the earliest and extend to *Month 258*.
- Some fields also contain values for *Month 259* to *Month 360*, e.g. about known leasing contract ends.
- Some fields do not have data stored for all points in time.
- The internal Porsche data are masked, i.e. the numerical values are manipulated so that no conclusions can be drawn about the actual values.

month	countryname
1	Country A
2	Country A
3	Country A
...	...
359	Country A
360	Country A
1	Country B
2	Country B
3	Country B
...	...
359	Country B
360	Country B
1	Country C
2	Country C
3	Country C
...	...
359	Country C
360	Country C
1	Country D
2	Country D
3	Country D
...	...
359	Country D
360	Country D

Time series for Country A with frequency 1/month

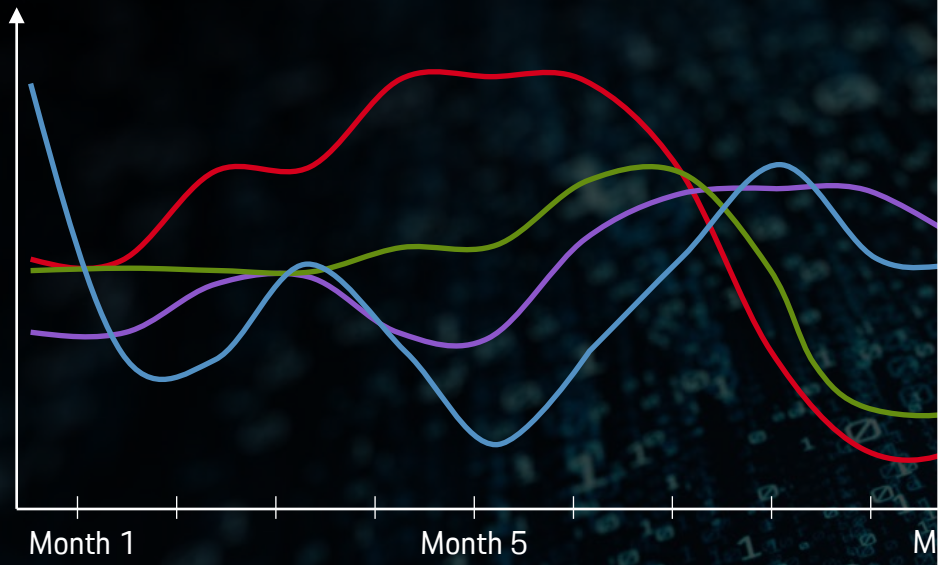
The internal Porsche data contain sales and customer data



Field Name	Field Description
sales_actuals_monthly__[...]__orderintake	Number of order intakes for vehicles
sales_actuals_monthly__[...]__retail	Number of actually sold vehicles
sales_flow_monthly__[...]__cp8	Number of vehicles that reached end of production
sales_stock_monthly__[...]__physical	Number of cars that are physically in stock
sales_stock_monthly__[...]__grossall	Number of cars in stock among all sales levels (Porsche AG, importers, dealers)
sales_stock_monthly__[...]__netimporteranddealer	Net stock of importers and dealers (stock cars without customer relationship)
customer_vehicle_relationships__[...]__terminated_cars	Number of terminated customer-vehicle relationships (customers sold their vehicle)
customer_vehicle_relationships__[...]__active_cars	Number of active customer-vehicle relationships (current customers)
leads__[...]__won	Number of known purchase interests that have resulted in an order
leads__[...]__first_touch_point	Number of known purchase interests with pending decision for/against purchase
leads__[...]__lost	Number of known purchase interests that have not resulted in an order
leasing_contracts__[...]__end	Number of leasing contracts commenced
leasing_contracts__[...]__start	Number of leasing contracts ending
premises__[...]__flag_order_start	Date at which new model can be officially ordered (Start of Sale)
premises__[...]__in_sales	Current model which can be ordered by customers
premises__[...]__flag_production_start	Start of production of a new model
premises__[...]__in_production	Current model in production
quota__[...]__aq	Number of cars which are produced within a specific month

Each field is available for various groups of Porsche models

The external data contain macro-economical data for each of the specific countries/area



Field Name	
global_consumer__alcoholic_beverages	consumer indices
global_consumer__apparel	
global_consumer__apparel_footwear	
global_consumer__appliances	
global_consumer__automobiles	
global_consumer__beer	
...	
global_economy__capital_consumption_total_us__u_s_dollar	economic indices
global_economy__change_in_international_reserves__u_s_dollar	
global_economy__change_in_international_reserves_excluding_gold__u_s_dollar	
global_economy__consumer_price_index__index	
global_economy__consumer_price_index_core__index	
global_economy__consumer_price_index_core_dollar_basis__index	
...	
registrations__vwm_premium_mini	official vehicle registrations
registrations__vwm_premium_acura	
registrations__vwm_premium_infiniti	
registrations__vwm_luxus_maserati	
registrations__vwm_luxus_bentley	
registrations__vwm_luxus_ferrari	

global means „for the entire country/area" (related to field *countryname*)

The Challenge

Challenge

You are selected to assist the planning team in Country A. With your AI and analytics skills, you are exactly the right people to support in the following ways:

1. For **Country A*** and **Vehicle Group 01****, make an **18 months***** forecast with an algorithm of your choice.
2. Find out **which of the used time series are particularly helpful** for forecasting and try to identify correlations. Find further interesting insights.

month	countryname	sales_actuals_monthly__vehiclegroup01__orderintake
...	...	...
255	Country A	621
256	Country A	451
257	Country A	420
258	Country A	544
259	Country A	???
260	Country A	???
261	Country A	???
262	Country A	???
263	Country A	???
264	Country A	???
265	Country A	???
266	Country A	???
267	Country A	???
268	Country A	???
269	Country A	???
270	Country A	???
271	Country A	???
272	Country A	???
273	Country A	???
274	Country A	???
275	Country A	???
276	Country A	???

These 18 values are to be predicted

To accomplish the task, you can use all time series for all countries in the data set or other publicly available data. Choose different training and test periods to make the forecast model robust. The quality of the forecast is evaluated at the end based on the *mean absolute percentage error* (MAPE)****. Choose a suitable type of presentation for your results, for example a dashboard, interactive visualizations or a PowerPoint.

* countryname == `Country A` ** sales_actuals_monthly__vehiclegroup01__orderintake *** Month 259 – Month 276 **** https://en.wikipedia.org/wiki/Mean_absolute_percentage_error

The Awards

	Criteria	Tips
 **Best BI-Artefact or Prototype**	**Could you use your analytical skills to handle the identified business case?** Please include the code and documentation of your data product or analytics/AI/BI artefact or prototype	• Take a look at the lecture and the different analytics systems and how they are used • Simple solutions or prototypes are also welcomed
 **Best Prediction**	**Can you predict the next 18 months order intakes?** Please submit the 18 months forecast of order intakes for Country A and Vehicle Group 01 by email to Dr. Sebastian Seifen <sebastian.seifen@porsche.de> until Thursday 01. February evening at 18:00 MEZ!	• Simpler models are also valuable • What are the relevant features • Data preprocessing might help you
 **Best Presentation/ Data Insights**	**Can you tell a coherent and exciting story about the analysis and your insights?** PowerPoint presentation, live demo or comparable form.	• Build an interesting and clear story • Remember that there are also some practioneers and non-experts in the jury • Less is more: get to the point and pack the details into the backup

Capstone Project

30.01.2024, 9:30 – 18:00

- Greetings & Welcome
- Guest Lecture
- Use Case Introduction

30.01. – 01.02.2024

- Teamwork
- Use Case Q&A (online via MS Teams)
- Expert sessions with PAG employees
- Lecture Evaluation (15:00)
- Submission of the 18 predictions (until 01. Feb. 18:00 to sebastian.seifen@porsche.de)

02.02.2024, 9:30-12:00

- Final pitches
- Nomination of the three winner teams

Award


PORSCHE MUSEUM
SAVE-THE-DATE!
Porsche Experience Zuffenhause
21.03.2024
9:00-15:00 Uhr

Final Presentations

Day 4 – Friday (preliminary)

- 09:30 Greedings and Welcome
- 10:00 Start
- 10:10 Presentation Team 1
- 10:20 Presentation Team 8
- 10:30 Presentation Team 2
- 10:40 Presentation Team 7
- 10:50-11:15 Break
- 11:15 Presentation Team 3
- 11:25 Presentation Team 6
- 11:35 Presentation Team 4
- 11:45 Presentation Team 5
- 11:45-12:15 Break
- 12:30 Awards and Results

Modalities

- 10 min presentation
- Please stick to the given amount of time
- Not every team member has to present

Presentation & Content

- Ideas
- Approach
- Results
  - Code Run-Through
  - Demonstration of dashboard, data model, application
  - Evaluation, e.g. of classifier
- Discussion
- Next Steps

Work Environment

Tools

- Data Science Environment
  - Python Anaconda
  - R + RStudio
  - KNIME
- Microsoft Power BI
- GITlab KIT-SCC
- H-lab Jupyter Hub

Data Access

- Every team member has access to a file exchange
- For downloading data and for uploading your final submission
- Link and password via E-Mail

Work Environment – Jupyter Hub

Human-Centered Systems Lab:

Open link once (KIT VPN required):
https://auth.k8s.iism.kit.edu/auth/realms/default/account/#/

Email your user to issd-researchserver@kit.edu to get access after you decided about the technology

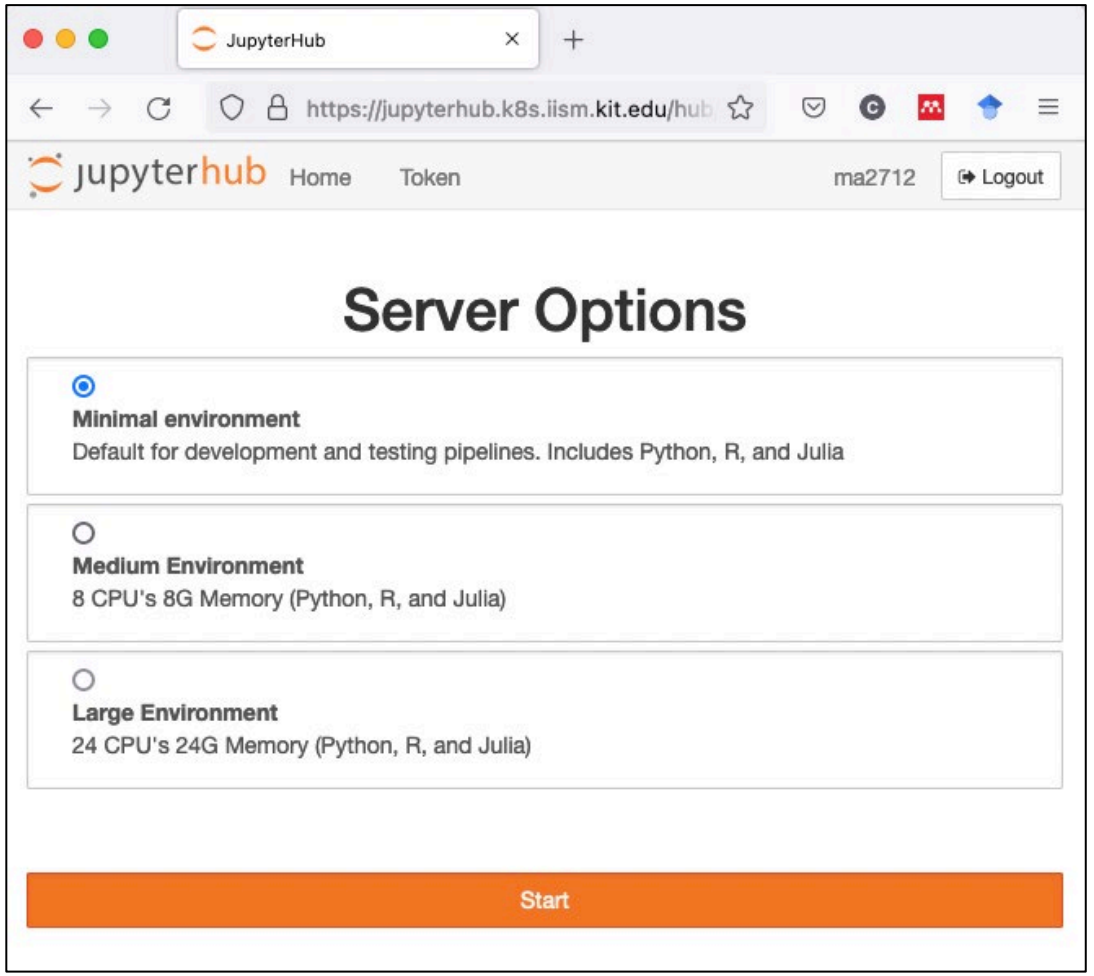
Choose minimal environment for defining your models

For training, you can choose medium and large environments.

Don't forget to shut down your environment afterwards

SCC:

- https://www.scc.kit.edu/dienste/JupyterHPC.php



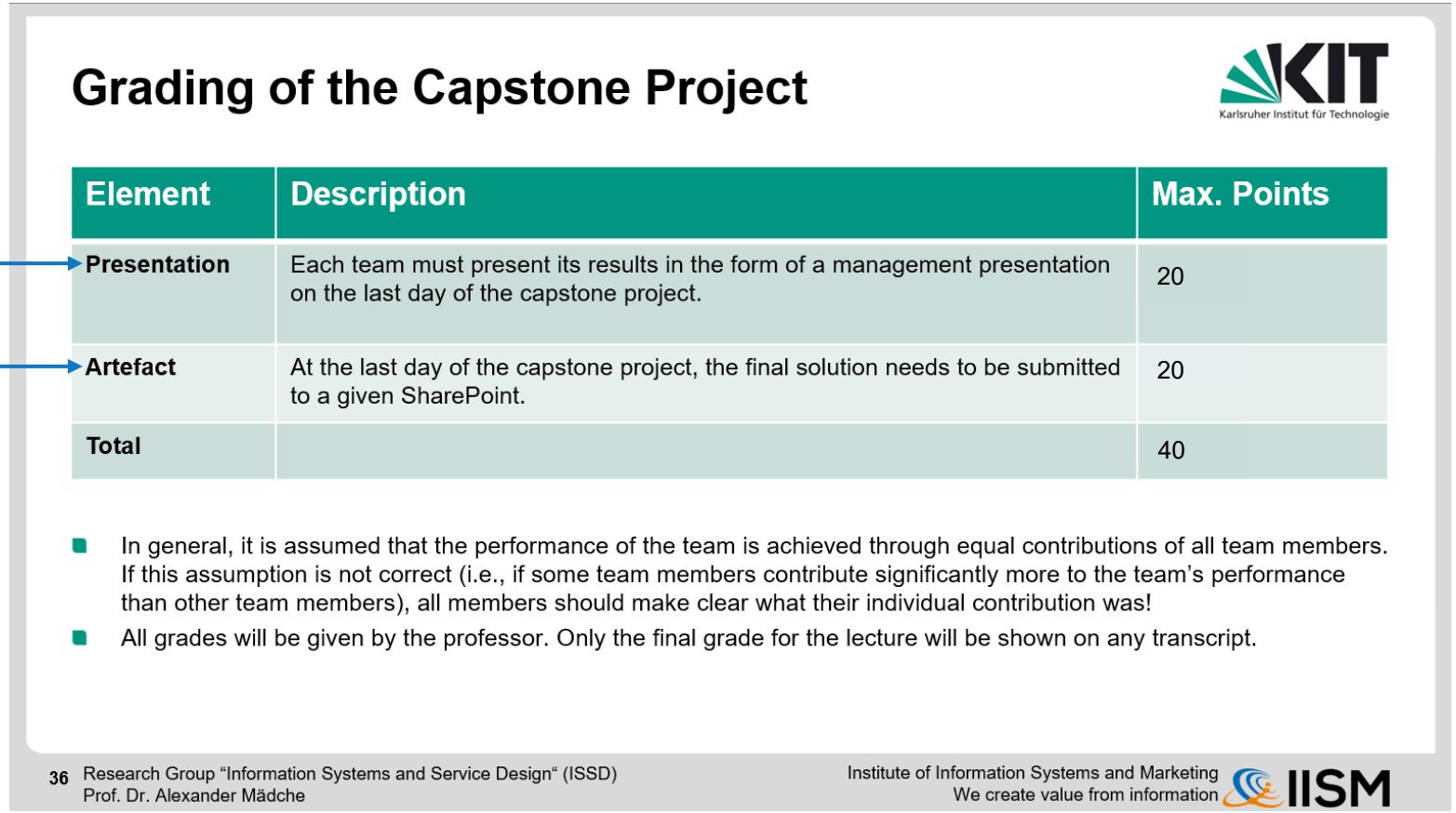
Required Submissions on Friday by the EOD

(1) PDF – File

(2) Code and Power BI File as ZIP Archive

Target Folder:

`File Exchange`



Grading of the Capstone Project

Element	Description	Max. Points
Presentation	Each team must present its results in the form of a management presentation on the last day of the capstone project.	20
Artefact	At the last day of the capstone project, the final solution needs to be submitted to a given SharePoint.	20
Total		40

- In general, it is assumed that the performance of the team is achieved through equal contributions of all team members. If this assumption is not correct (i.e., if some team members contribute significantly more to the team's performance than other team members), all members should make clear what their individual contribution was!
- All grades will be given by the professor. Only the final grade for the lecture will be shown on any transcript.

NDA for Capstone Project Required

Submit the NDA to Ilias (in Folder: Capstone Project) until 21.01.2024 23:55


#Let‘s accelerate