

The background features four dark silhouettes of people walking away from the viewer along a glowing blue path. The path is set against a vibrant, swirling blue and orange background that resembles a nebula or a digital tunnel.

FinApprenticeship

2025.06.23

MEET THE TEAM



Tetiana

Data scientist

Back-end developer



Jonny

Data Scientist

Virologist



Alexander

Data Scientist

Computer Scientist
Programmer



Delia

Data Scientist

Japanese studies and
Statistics
Director

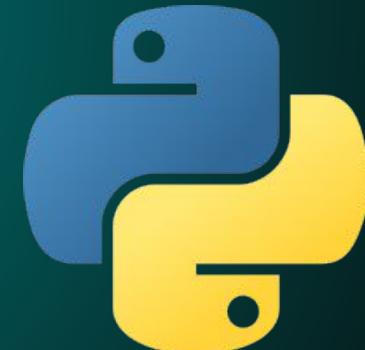
TECHNOLOGY STACK



Streamlit



OPTUNA



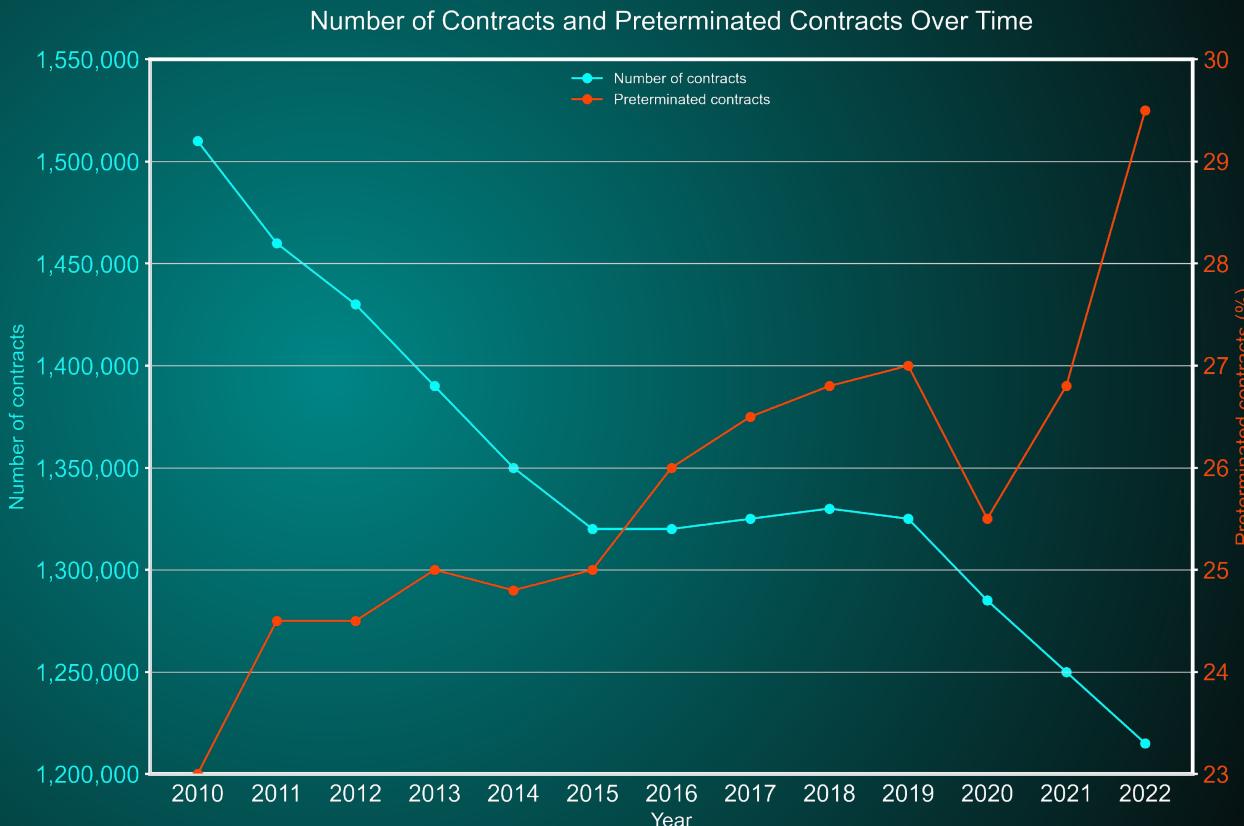
BUSINESS PROBLEM

What's the issue?

Around 30% of apprenticeship contracts are terminated prematurely.

Termination rates vary by region, age, and education level.

This worsens the skilled labor shortage and strains the vocational system.



WHO BENEFITS FROM THIS PROJECT

- Policymakers
- Students
- Job Center Staff

⇒ Supports data-driven decisions to strengthen vocational education and reduce dropouts.



Bundesagentur für Arbeit
bringt weiter.



DATA COLLECTION + EDA

DATA SOURCES

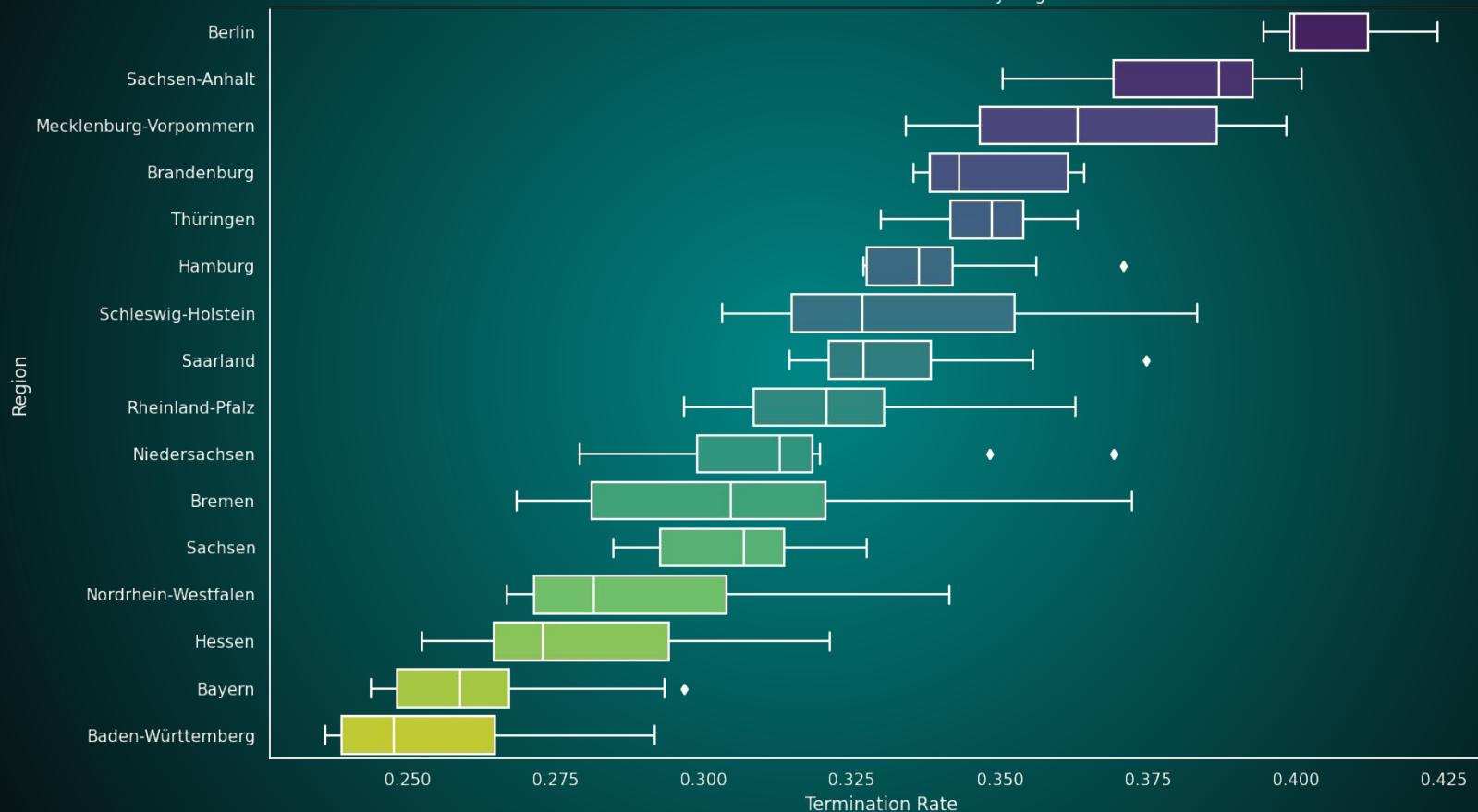
Datensystem Auszubildende -
Zeitreihen (DAZUBI)

Institut für Arbeitsmarkt- und
Berufsforschung (IAB)

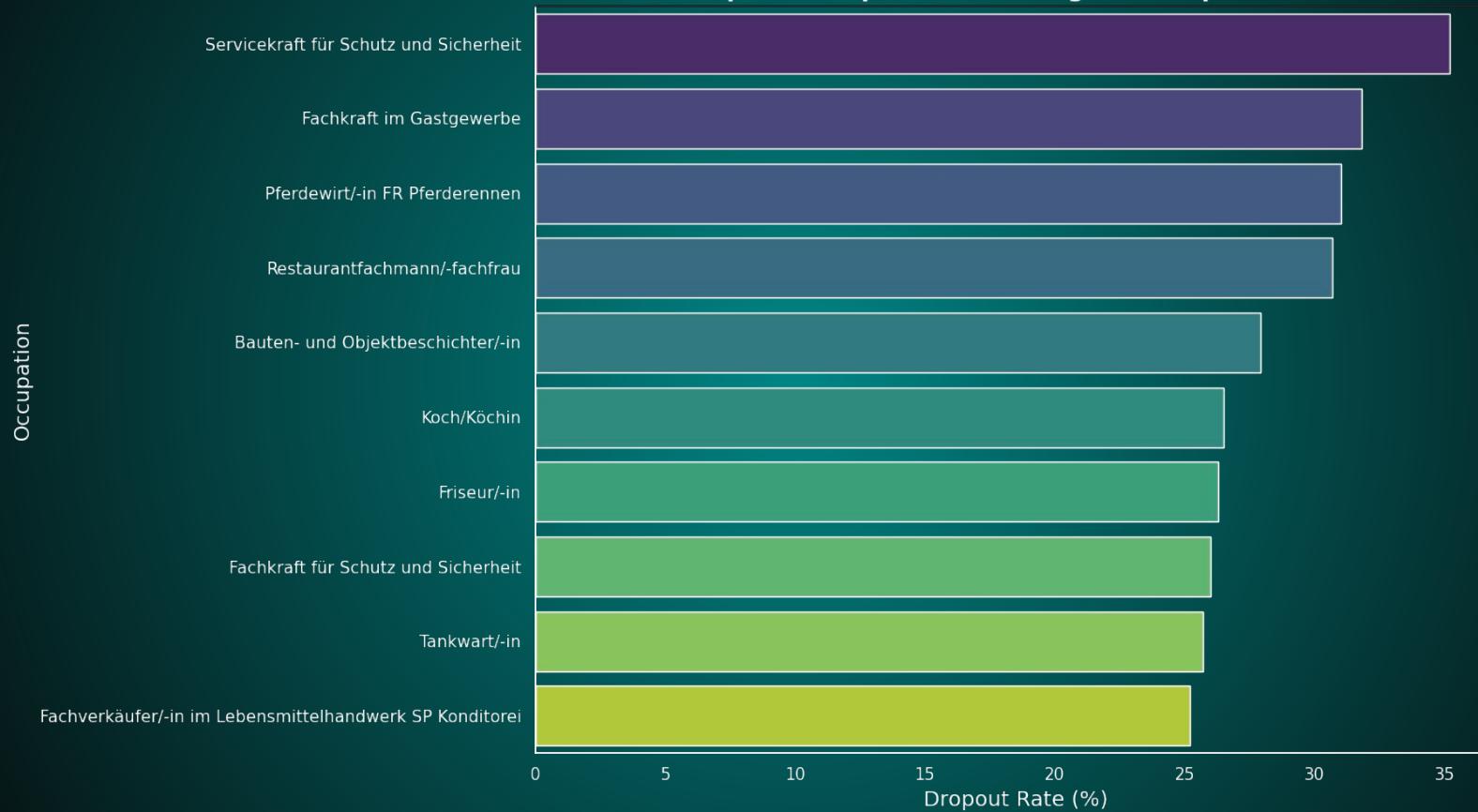
German Federal Statistical Office



Distribution of Termination Rate by Region



Top 10 Occupations with Highest Dropout Rate



APPRENTICESHIP DROPOUTS (Dazubi)

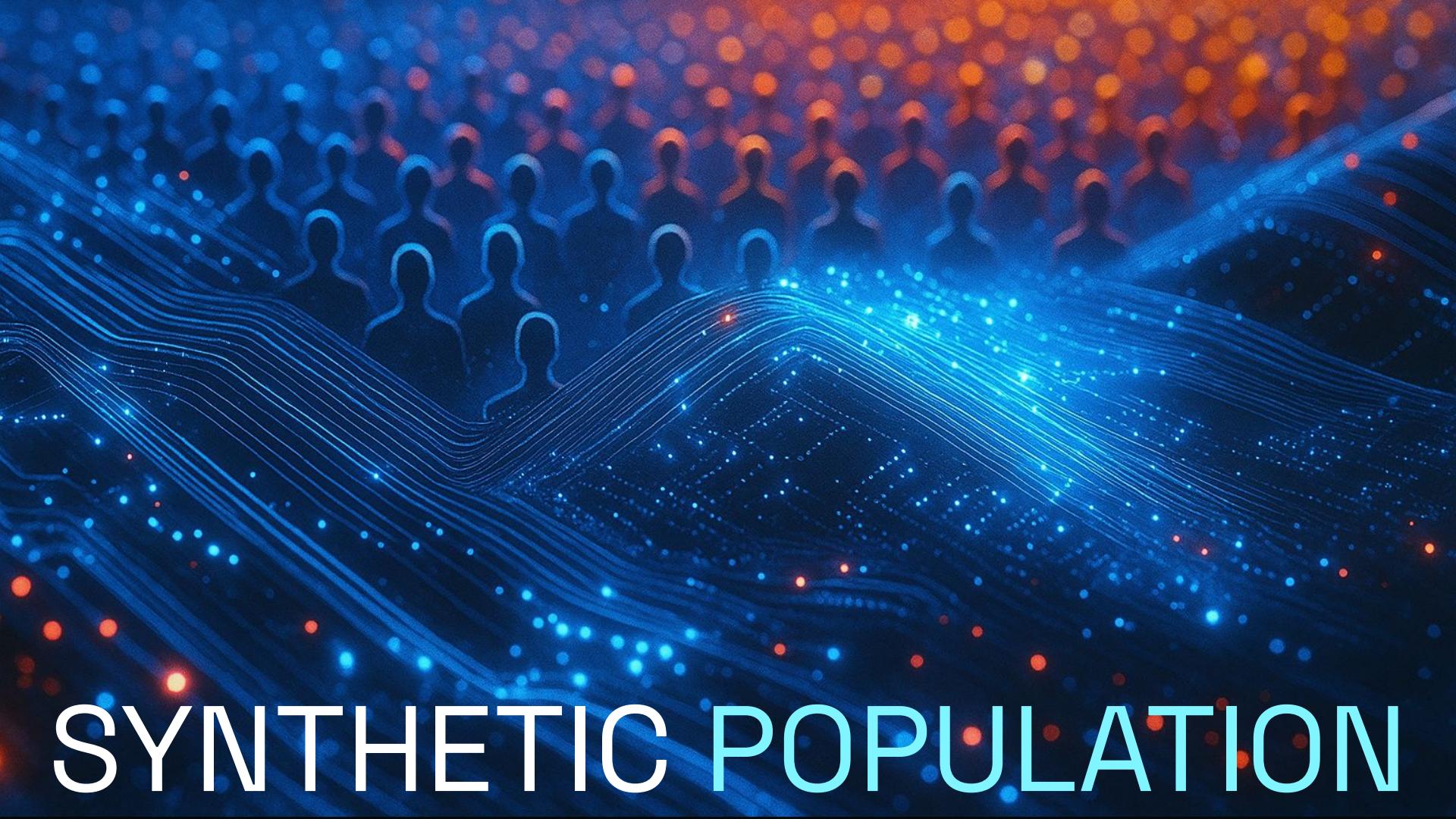
Key Findings:

- Apprenticeship Dropouts (2010–2023)
- Total dropouts stable until 2019, sharp dip in 2020 (COVID), then recovery
- Dropouts among non-German apprentices more than doubled since 2010
- East–West divide: Highest rates in Berlin & eastern states; lowest in southern Germany
- Gap between German and non-German dropouts is narrowing, but still exists

Recommendations:

- More targeted support for foreign apprentices
- Regional programs for high-risk states





SYNTHETIC POPULATION

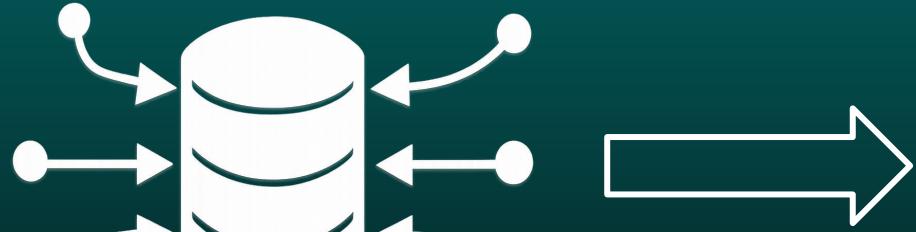
SYNTHETIC POPULATION

Aggregated Data



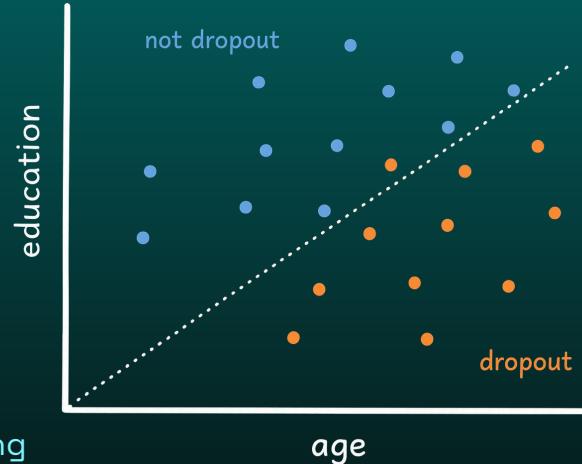
SYNTHETIC POPULATION

Aggregated Data



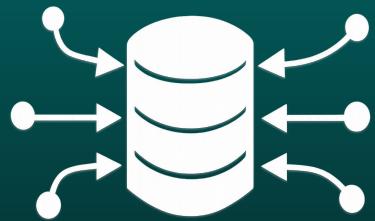
De-aggregation by
Iterative Proportional Fitting

Synthetic Population

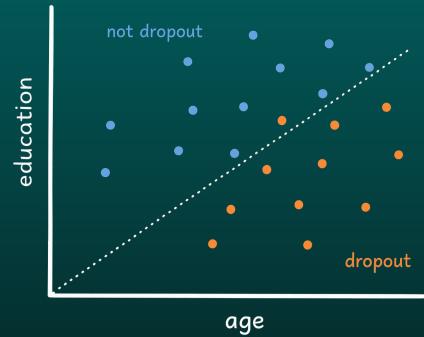


SYNTHETIC POPULATION

Aggregated Data



Synthetic Population



De-aggregation by
Iterative Proportional Fitting

ML Model

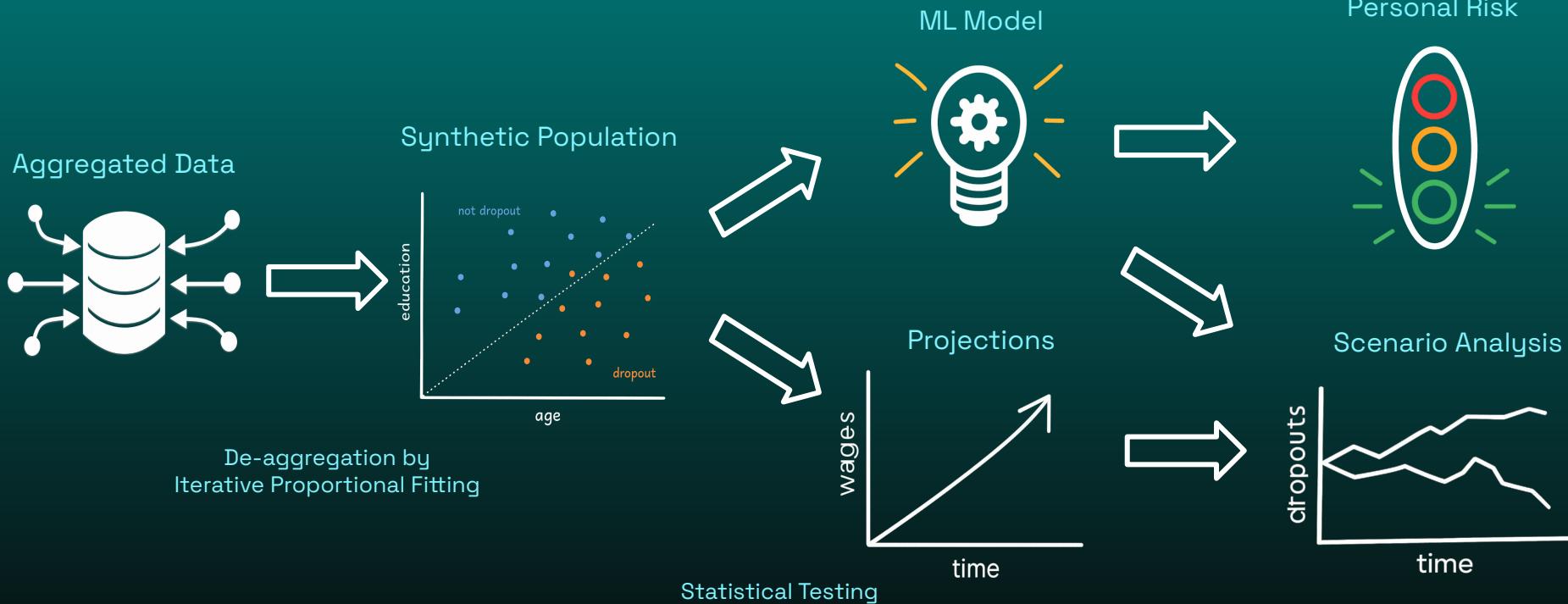


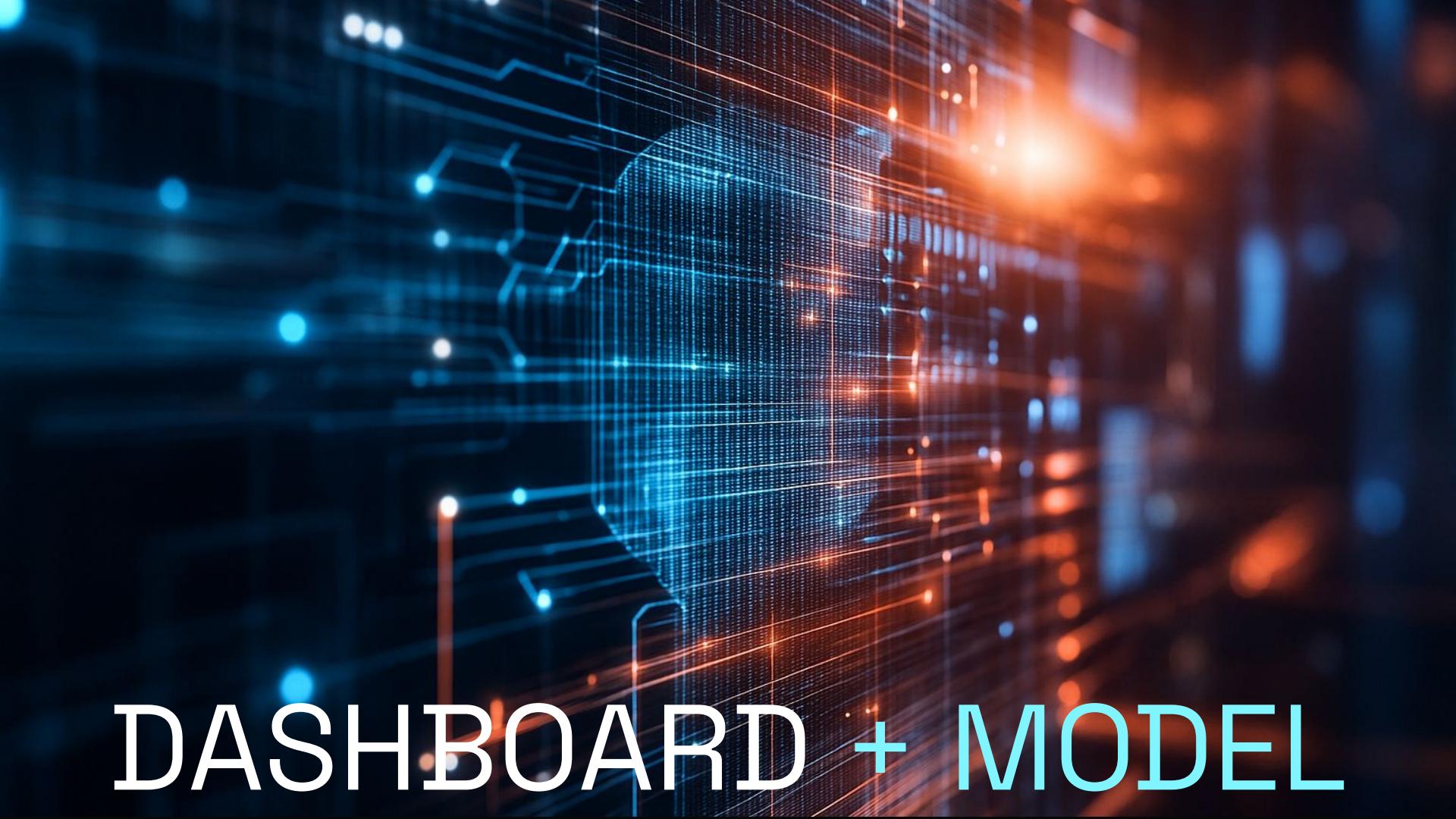
Projections



Statistical Testing

SYNTHETIC POPULATION





DASHBOARD + MODEL

CREATING THE MODEL

Different classification models give the same results → Accuracy: 0.63



→ We need **more features!** → back to data collection

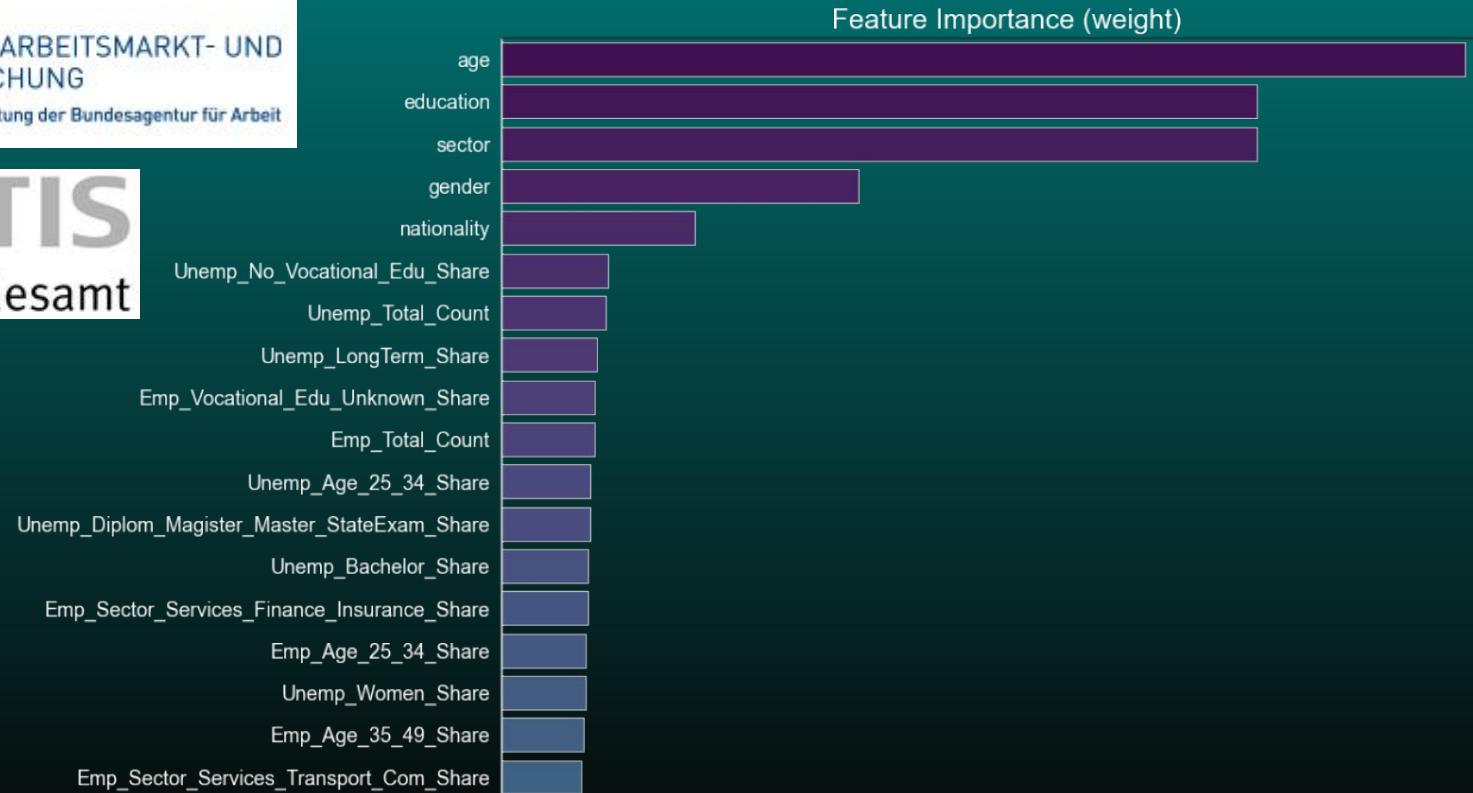
MODELING - FEATURE ENRICHMENT



MODELING - FEATURE ENRICHMENT

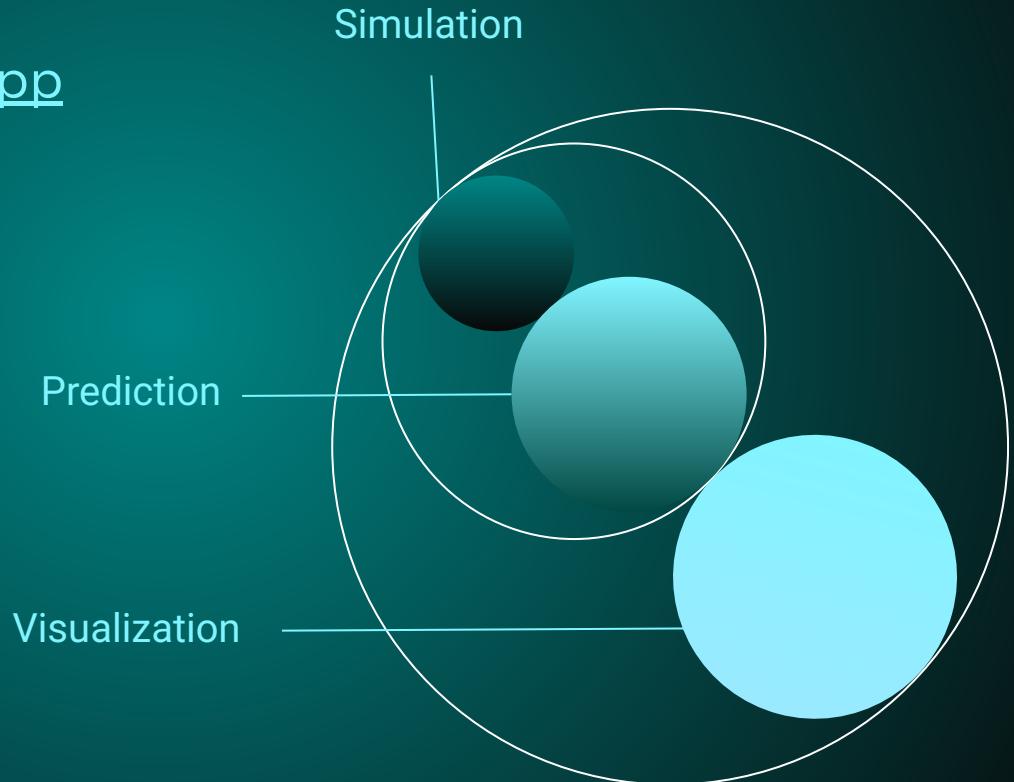


INSTITUT FÜR ARBEITSMARKT- UND
BERUFSFORSCHUNG
Die Forschungseinrichtung der Bundesagentur für Arbeit



DASHBOARD - WEB-APP

finapprenticeship.streamlit.app



FinApprenticeship

LET'S DISCUSS!

Tetiana



Alexander



Delia



Jonny

