FinApprenticeship

Predicting the Future of Apprenticeships

Our Team

Tetiana

Data Scientist

Back-end developer

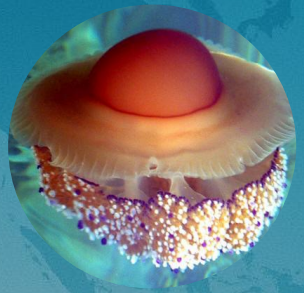
Delia

Data Scientist
Statistics and japanese studies
Director

Jonny

Data Scientist

Virologist



Alexander

Data Scientist
Computer Scientist
Programmer

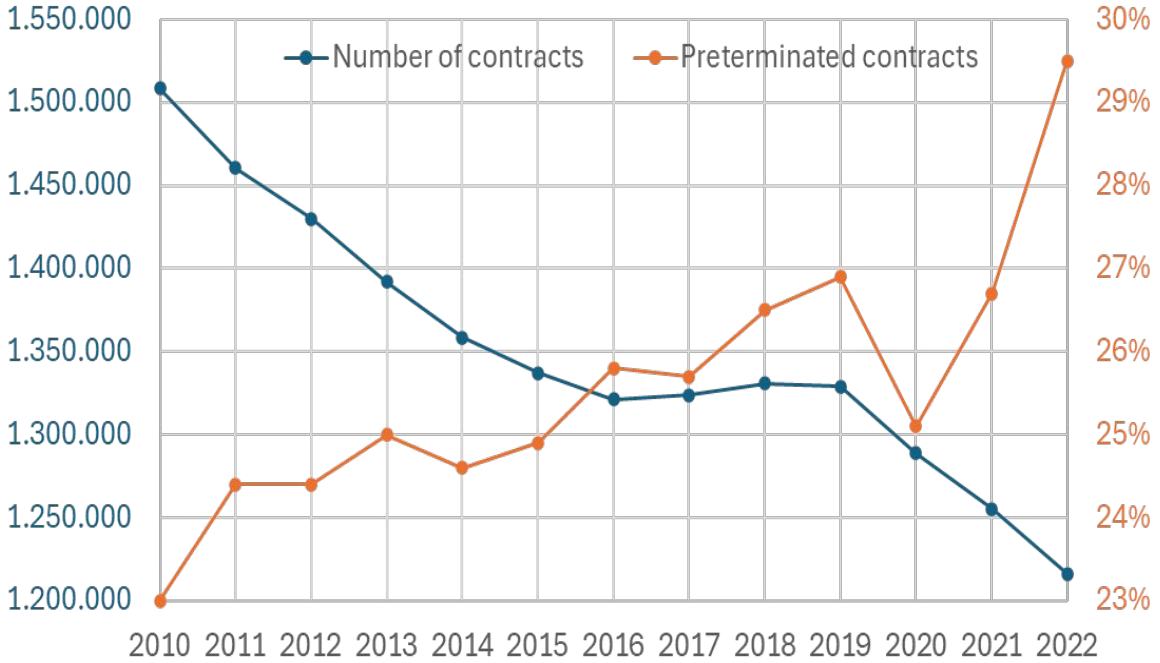
Business Problem

What's the issue?

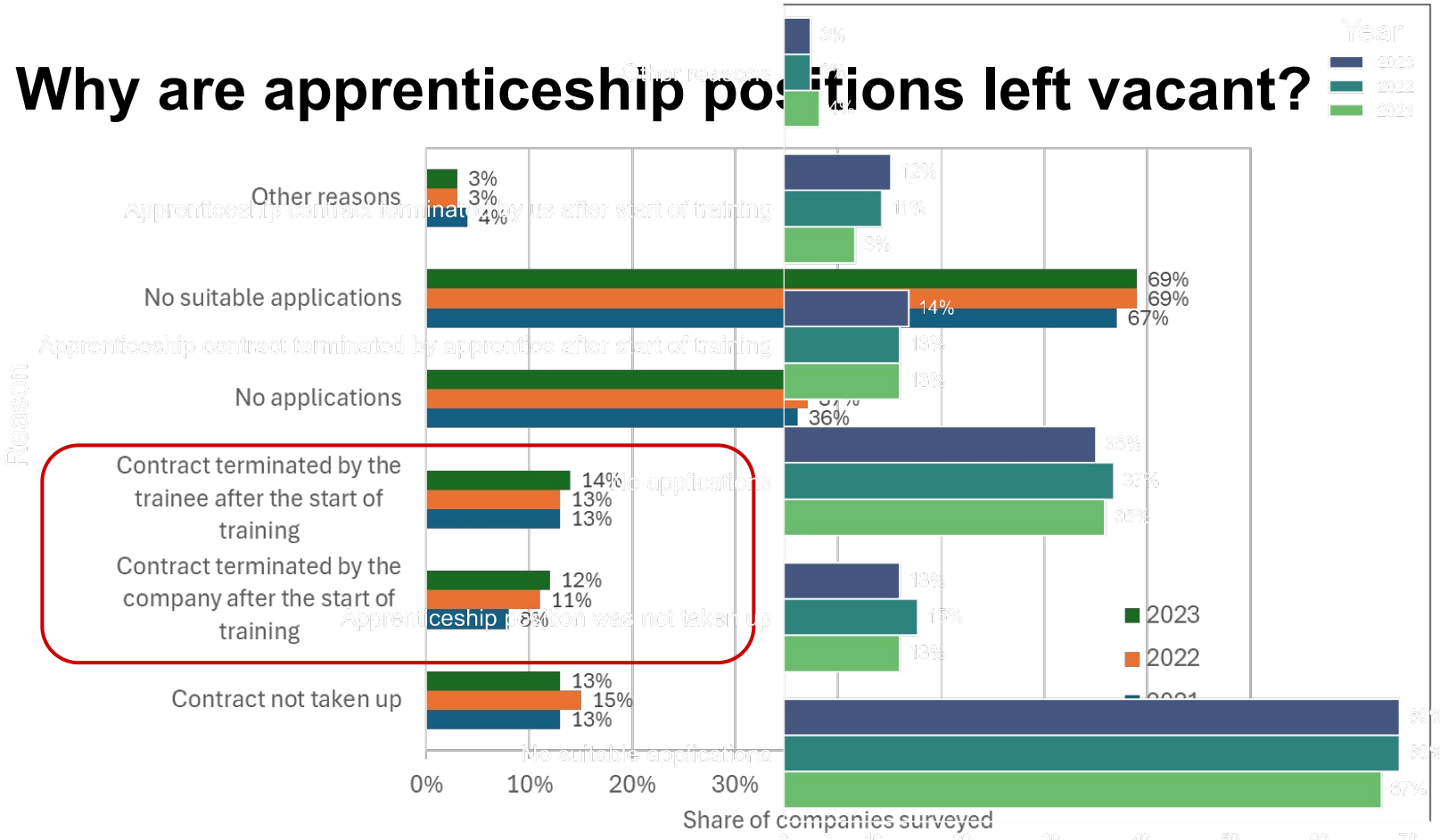
Around **30%** of apprenticeship contracts are terminated prematurely.

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

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



Why are apprenticeship positions left vacant?



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.

EDA


120
100
80
60
40
20
0
50
40
30
20
10
0
1/6/2002
1/8/2002
1/7/2002
1/9/2002

Data sources

Datensystem Auszubildende - Zeitreihen (DAZUBI)

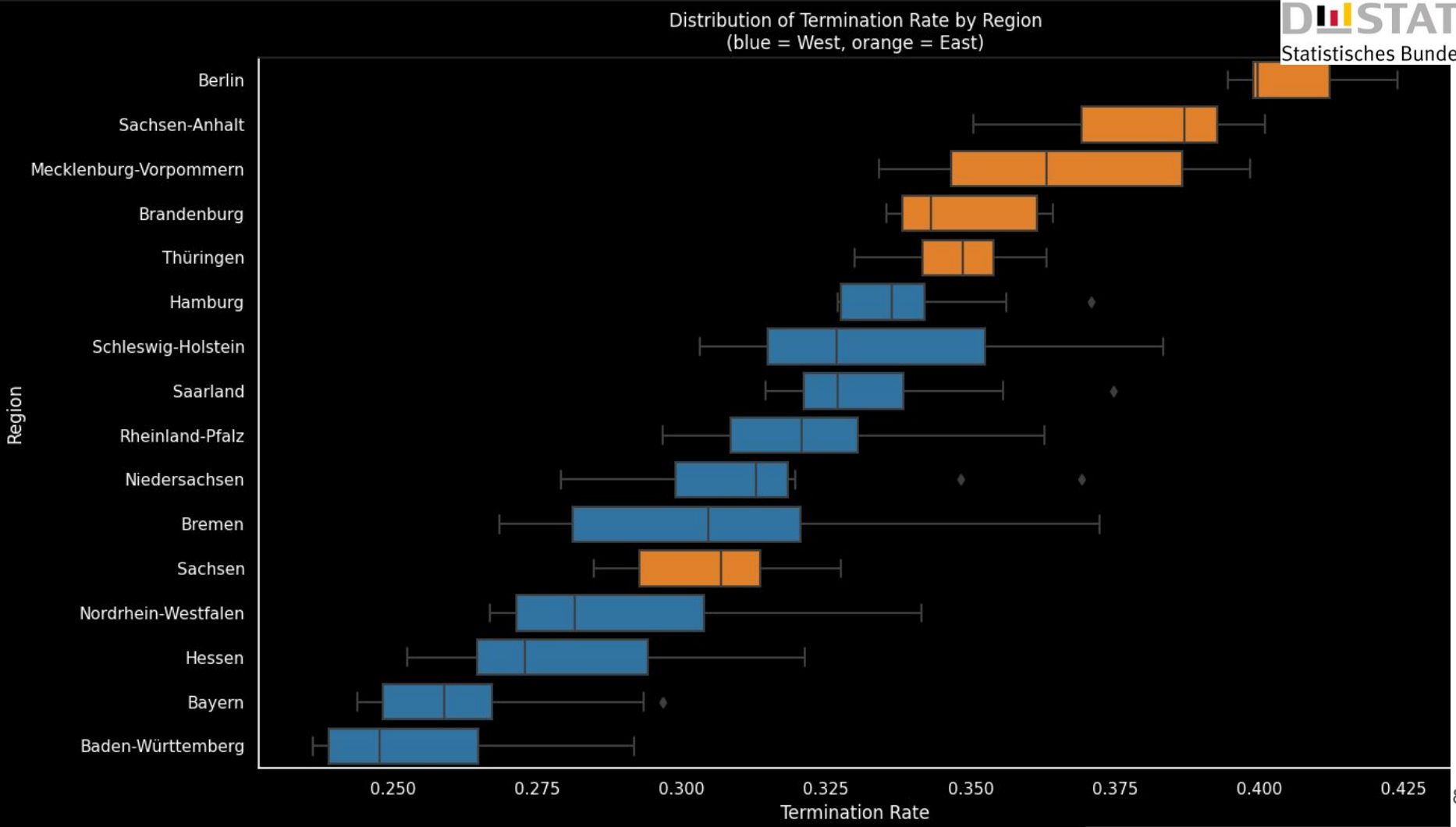
Bundesinstitut für Berufsbildung

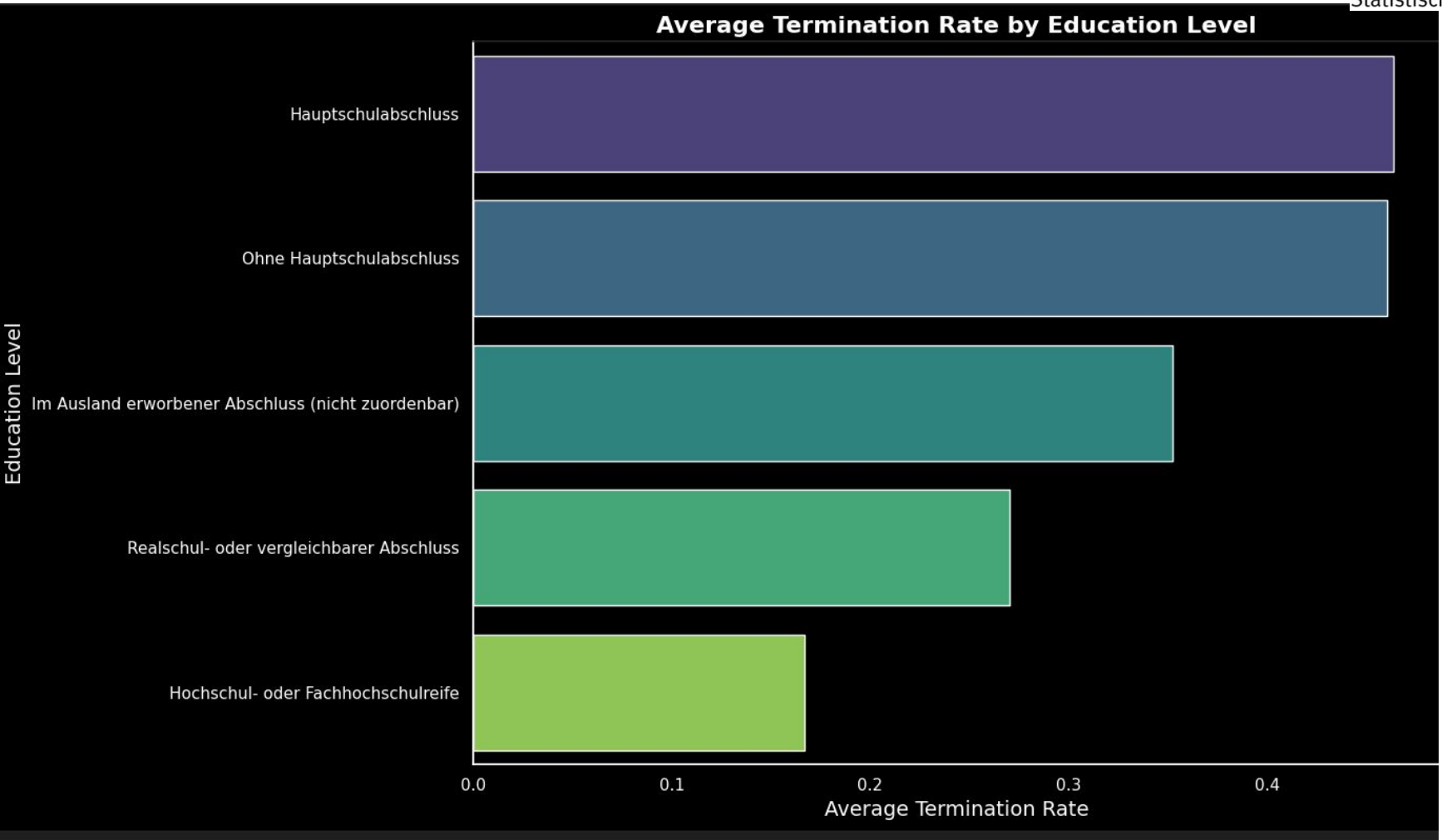
Institut für Arbeitsmarkt- und Berufsforschung (IAB)

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

German Federal Statistical Office

DESTATIS
Statistisches Bundesamt




Average Termination Rate by Education Level
Hauptschulabschluss
Ohne Hauptschulabschluss
Im Ausland erworbener Abschluss (nicht zuordenbar)
Realschul- oder vergleichbarer Abschluss
Hochschul- oder Fachhochschulreife
Education Level
0.0
0.1
0.2
0.3
0.4
Average Termination Rate

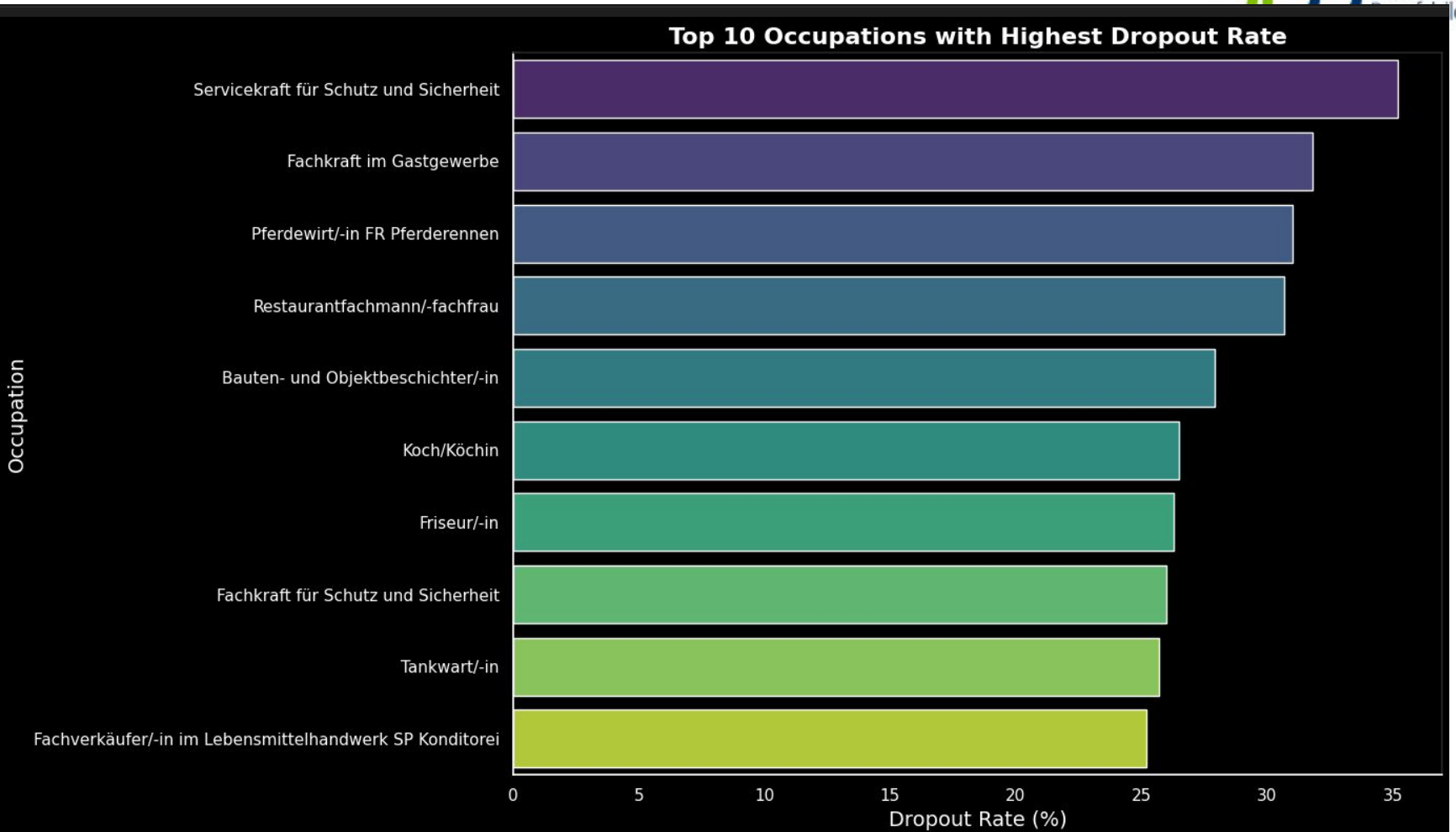
Apprenticeship Dropouts (Dazubi)

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:

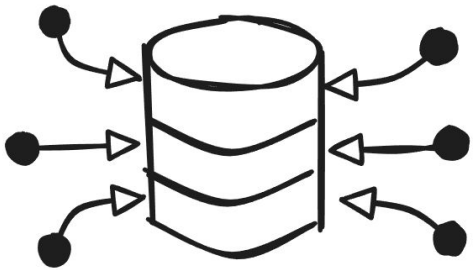
- More targeted support for foreign apprentices
- Regional programs for high-risk states


Top 10 Occupations with Highest Dropout Rate
Servicekraft für Schutz und Sicherheit
Fachkraft im Gastgewerbe
Pferdewirt/-in FR Pferderennen
Restaurantfachmann/-fachfrau
Bauten- und Objektbeschichter/-in
Koch/Köchin
Friseur/-in
Fachkraft für Schutz und Sicherheit
Tankwart/-in
Fachverkäufer/-in im Lebensmittelhandwerk SP Konditorei
Occupation
0
5
10
15
20
25
30
35
Dropout Rate (%)

Synthetic Population

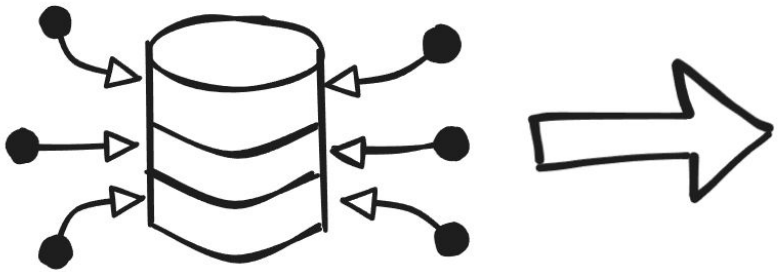
Synthetic Population

Aggregated Data

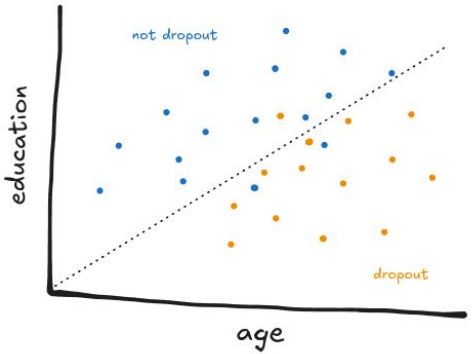


Synthetic Population

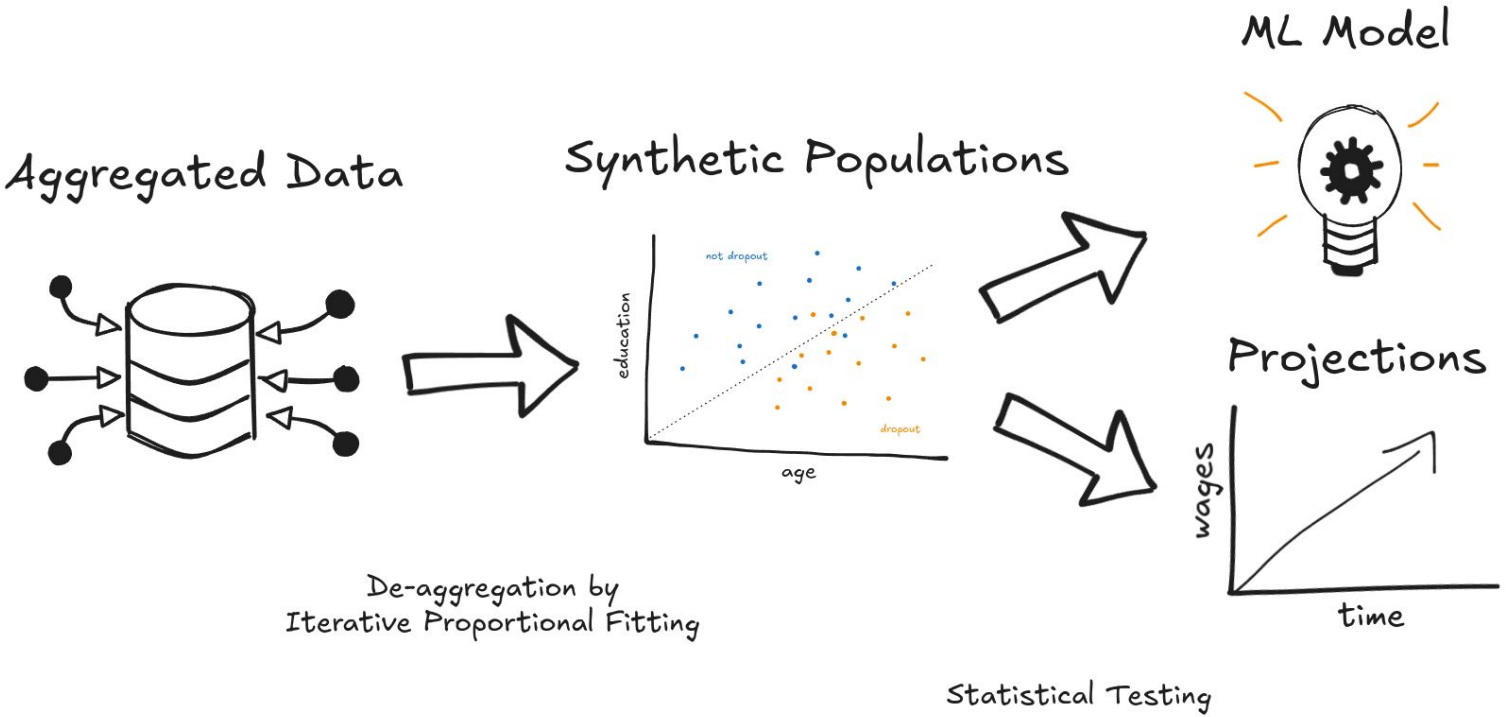
Aggregated Data



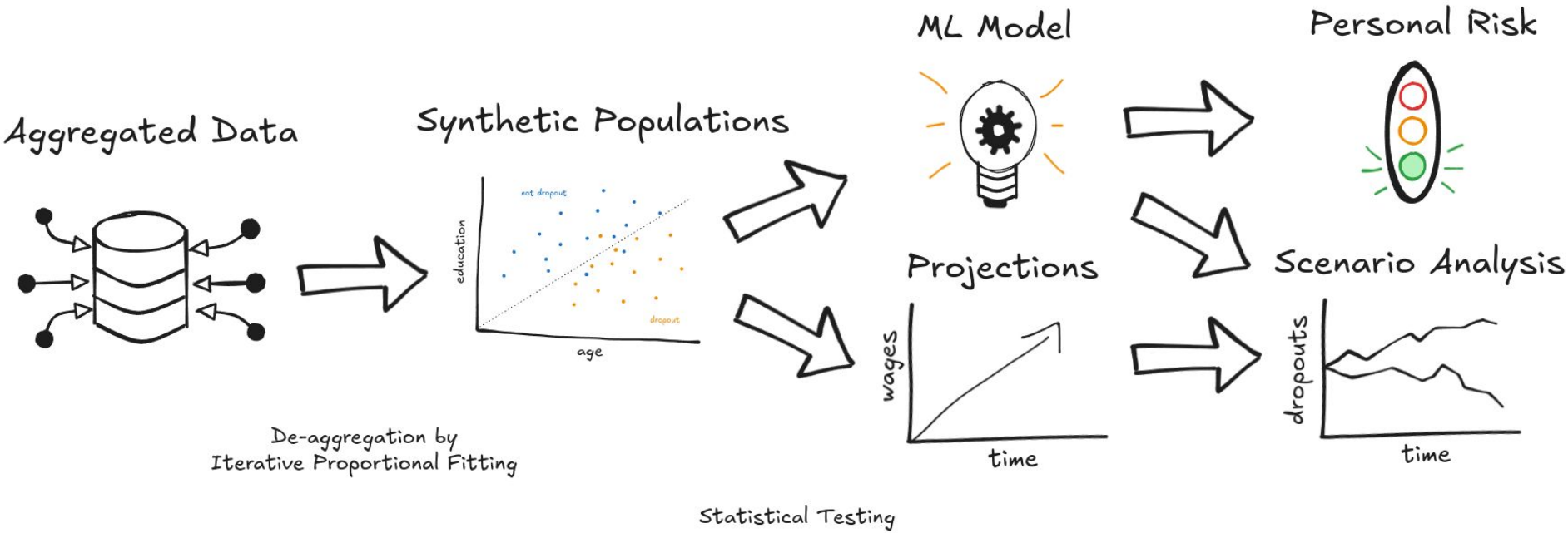
Synthetic Populations



Synthetic Population



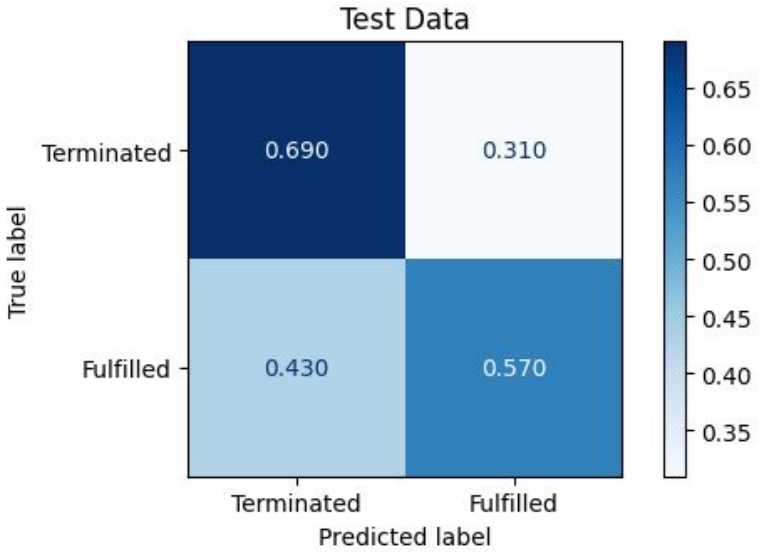
Synthetic Population

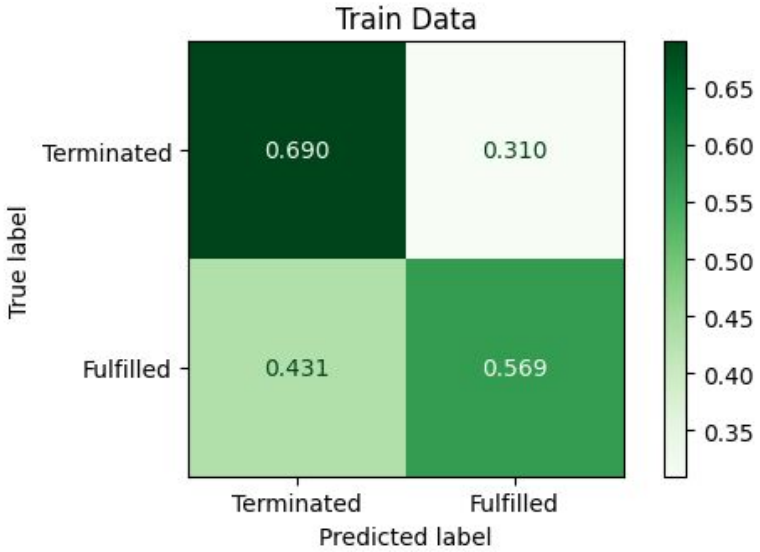


Model + Dashboard

Modelling

- Different classification models give the same results → Accuracy: 0.63





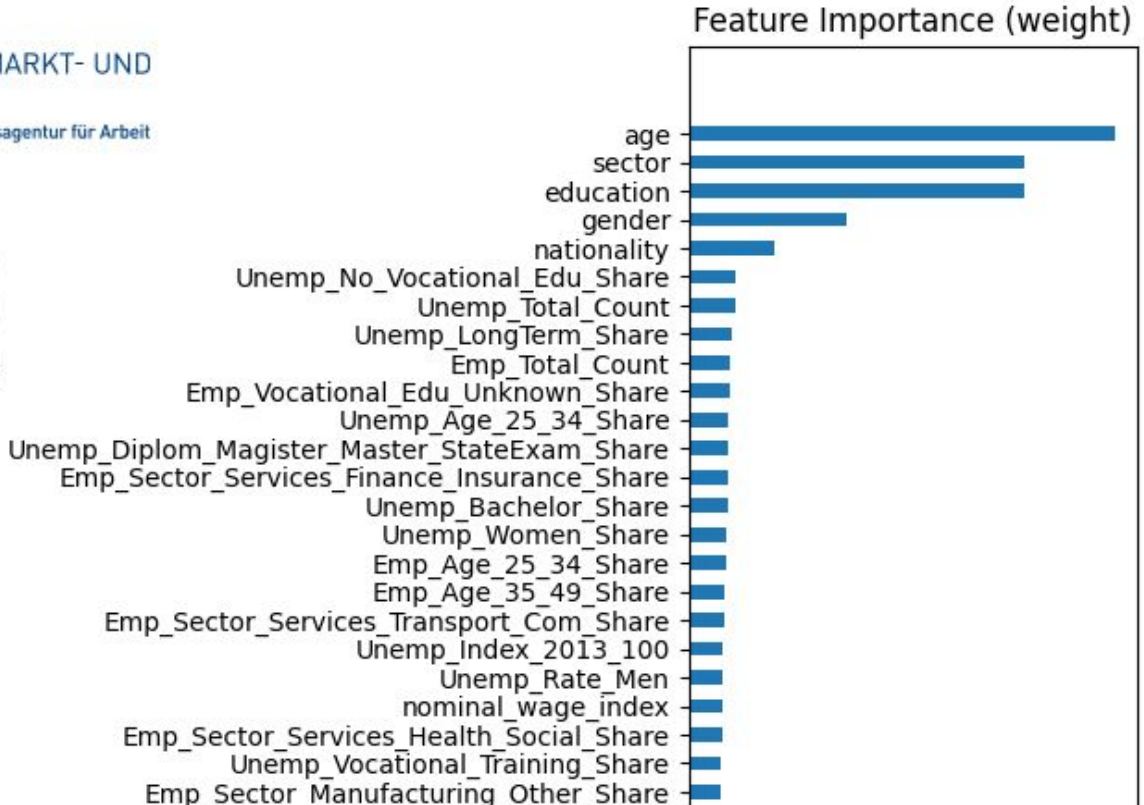
- We need **more features**! → back to data collection

Modelling - Feature enrichment

Modelling - Feature enrichment

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

DESTATIS
Statistisches Bundesamt



Deploy

🌈 Ausbildungsabbrüche im bunten Überblick

Beruf: Anlagenmechaniker/-in in allen Regionen

🎛️ Filteroptionen

🔧 Beruf auswählen

Anlagenmechaniker/-in

📍 Region wählen (optional)

Alle

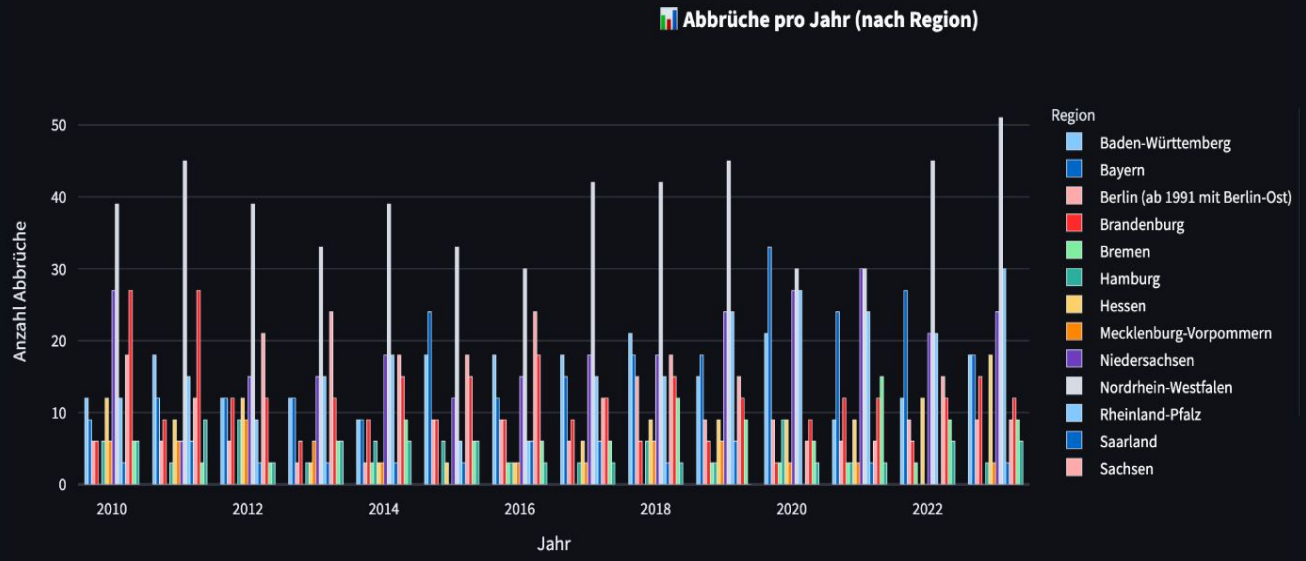
📅 Zeitraum

2010 — 2023

2010 — 2023

☐ 🗂️ Rohdaten anzeigen

☐ 🔮 Forecast anzeigen



FinApprenticeship

Let’s connect on linkedin!

Tetiana

Alexander

Deila

Jonny

Thank you all for your time and attention

Agenda

1. Definition of the Business Problem
2. Motivation & Stakeholder Benefits
3. Data Sources & Challenges
4. Methodology & Tech Stack
5. Key Insights from Data Analysis
6. Modeling & Prototyping
7. Future Goals & Next Steps

A Perfect Fit for Our Data App

What is Streamlit?

- An open-source Python framework for building interactive web apps
- Designed specifically for **data scientists** and **analysts**
- Requires **no frontend knowledge** (HTML/CSS/JS)

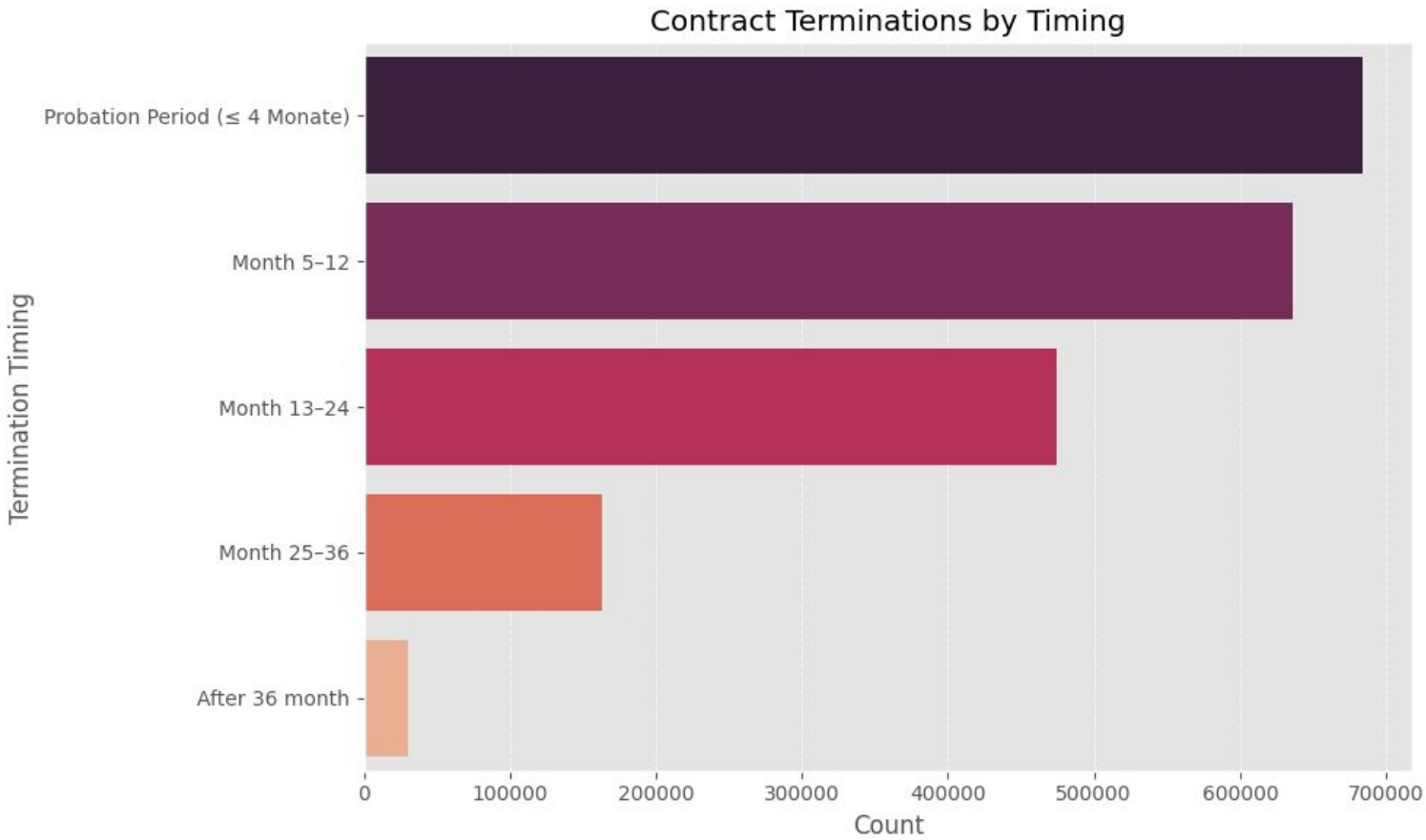
Why We Chose Streamlit for Our Product:

- **Fast development** – UI and logic in a single Python file
- **Interactive filters & visualizations** for real-time data exploration
- **Beautiful dashboards** with minimal effort (custom styling possible)
- **Scalable** for internal demos or public product deployment

Our Prototype Demo:

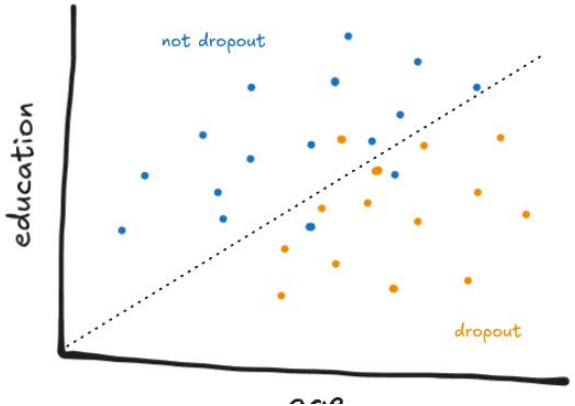
👉 Open in Browser

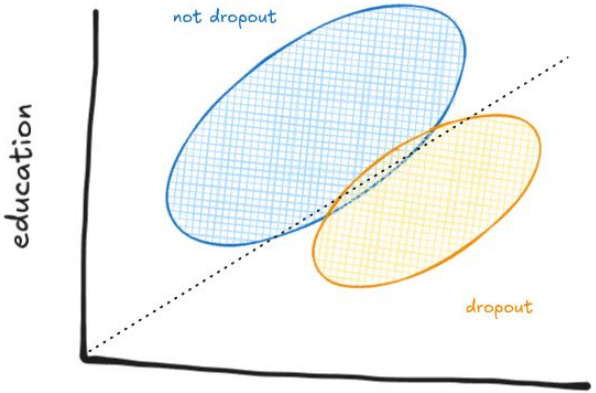
(Shows dropout trends by job, region, and includes forecasts)

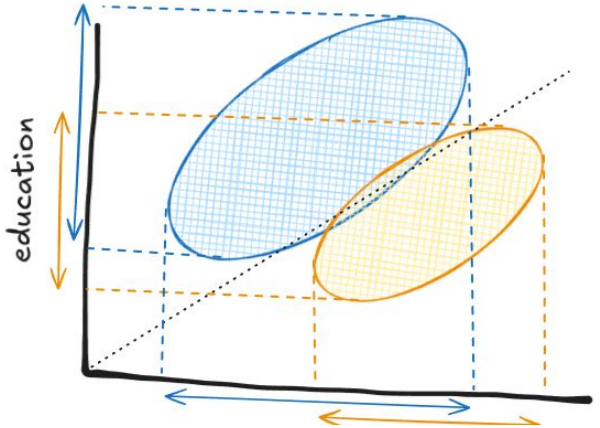

Contract Terminations by Timing
Termination Timing
Probation Period (≤ 4 Monate)
Month 5–12
Month 13–24
Month 25–36
After 36 month
0
100000
200000
300000
400000
500000
600000
700000
Count

Synthetic population: method

Known distributions for Age, Gender,
Education Level, Nationality, Sector of







education \ age	17 yo	18 yo	19 yo	total	target
basic					12
intermediate					14
high					8
total					
target	6	12	8		

education \ age	17 yo	18 yo	19 yo	total	target
basic	4	4	4	12	12
intermediate	4	5	5	14	14
high	2	3	3	8	8
total					
target	6	12	8		

education \ age	17 yo	18 yo	19 yo	total	target
basic	2	4	4	10	12
intermediate	2	5	5	12	14
high	2	3	3	8	8
total	6				
target	6	12	8		

Project Benefits for Stakeholders

Policymakers:

- Insights into key risk factors
- Enable targeted interventions

Students:

- Identify risk groups
- Support informed choices

Job Center Staff:

- Improve counseling with data
- Reduce early terminations

Overall:

- Boost retention
- Strengthen vocational training
- Address skills shortages

Bundesagentur für Arbeit
bringt weiter.