

GENESIS-Online

Die Datenbank des Statistischen Bundesamtes



Bundeskriminalamt

Big Data Projekt
Daten-Realität statt Mythen: Kriminalität & Herkunft im 40-Jahre-Vergleich

Justin Bergmann | 28.01.2026

Inhaltsverzeichnis

- Motivation
- Forschungsfragen
- Data Acquisition
- Data Pre-processing
- Data Analysis
- Result Validation
- Einordnung
- Fazit

Forschungsfragen

Vier Forschungsfragen welche ich mithilfe der Daten analysiert habe:

- **Regionale Verteilung:** Welches Delikt dominiert in welchem Bundesland?
- **Der Racial Profiling-Mythos:** Sind Nichtdeutsche nur überrepräsentiert, weil sie von der Polizei öfters kontrolliert werden?
- **Der Armut-Mythos:** Ist diese Überrepräsentation durch die Soziale Lage erklärbar?
- **Der Bagatell-Mythos:** Ist die Statistik verzerrt durch Bagatell-Delikte?

Motivation

Saskia Esken (SPD-Vorsitzende), 2020:

"Auch in Deutschland gibt es latenten Rassismus in den Reihen der Sicherheitskräfte."

Amnesty International / Gängiges Narrativ:

"Die Polizeistatistik bildet nicht die Kriminalität ab, sondern das Kontrollverhalten der Polizei. Wer öfter kontrolliert wird, taucht öfter in der Statistik auf."

Thomas Kutschaty (SPD), NRW-Wahlkampf:

"Kriminalität ist keine Frage des Passes, sondern der sozialen Lage."

Data Acquisition



*1



Bundeskriminalamt

*2

Quellen waren hauptsächlich:

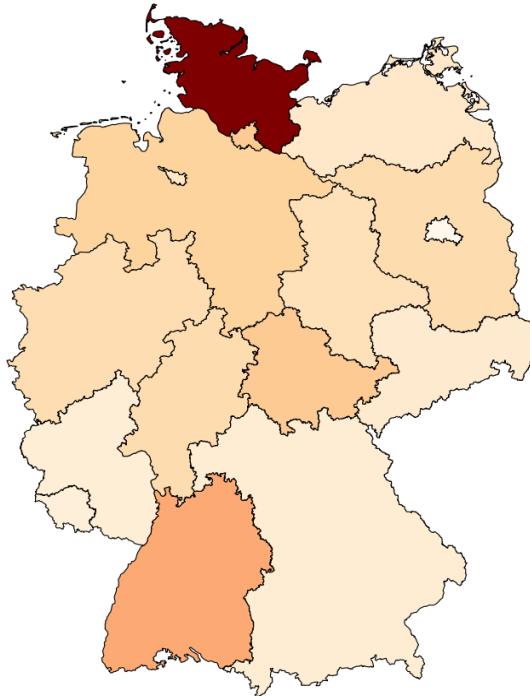
- Genesis-Online vom Statistischen Bundesamt
- PKS vom BKA

Data Pre-processing **Medallion Architecture**



Data Analysis Heatmap häufigste Straftat Deutschland pro Bundesland

Dominanz des Top-Delikts (Intensität)

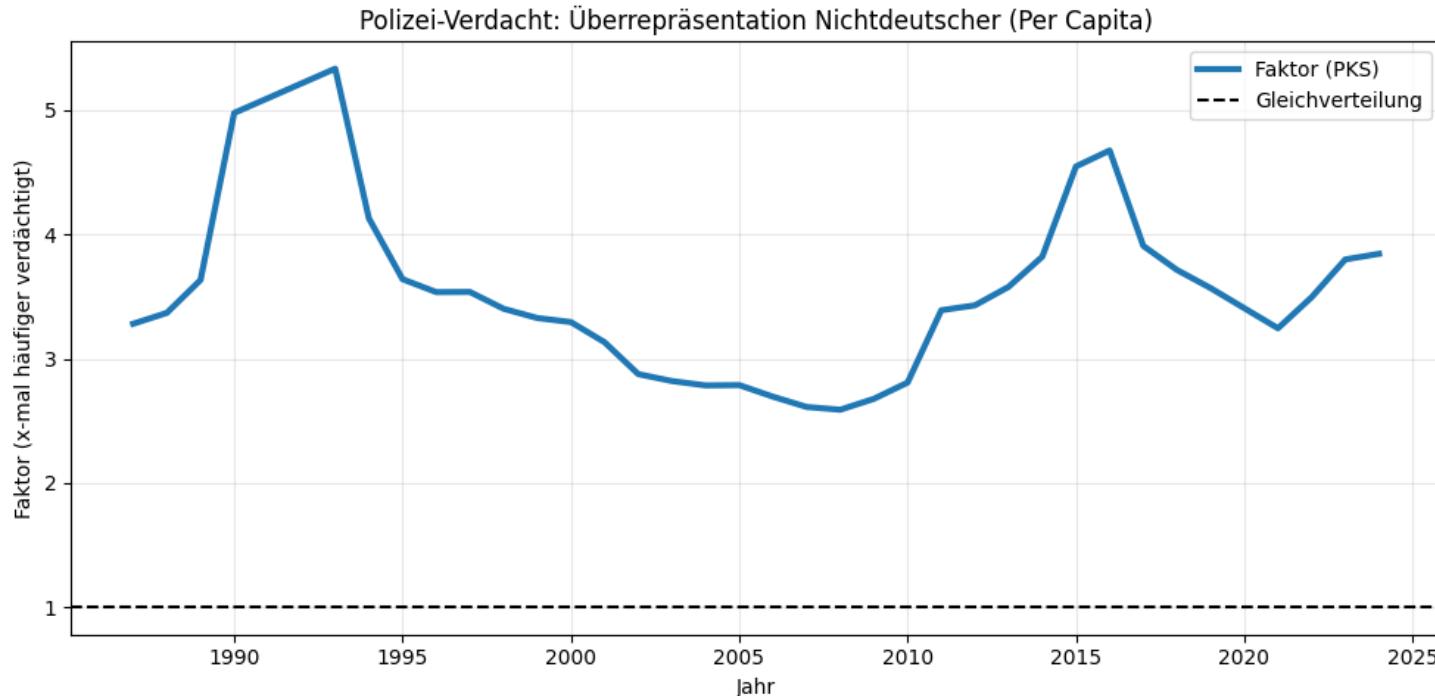


Rangliste: Welches Delikt dominiert?

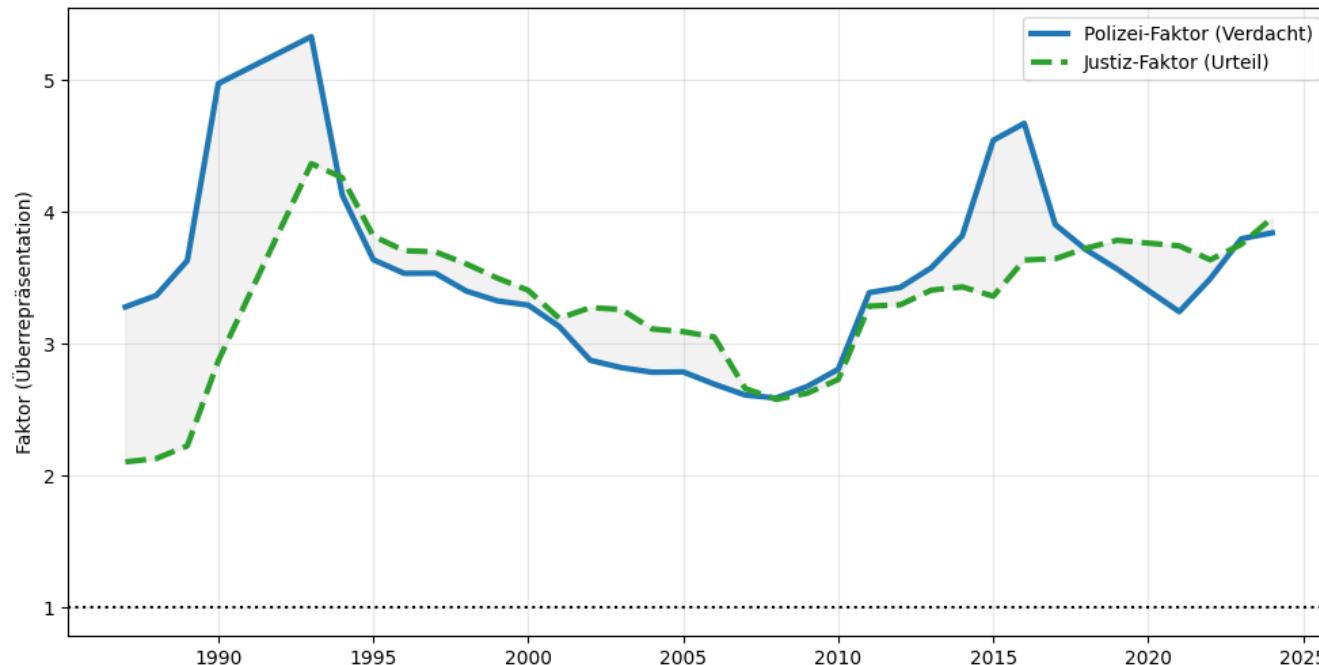
(Sortiert nach Anteil an der Gesamtkriminalität)

Schleswig-Holstein:	Betrug (Gesamt)	17.5%
Baden-Württemberg:	Betrug (Gesamt)	14.3%
Hamburg:	Betrug (Gesamt)	13.6%
Thüringen:	Betrug (Gesamt)	13.6%
Niedersachsen:	Betrug (Gesamt)	13.4%
Hessen:	Betrug (Gesamt)	13.1%
Brandenburg:	Nebengesetze (Asyl/Drogen)	13.1%
Nordrhein-Westfalen:	Betrug (Gesamt)	13.0%
Sachsen-Anhalt:	Betrug (Gesamt)	12.9%
Bremen:	Betrug (Gesamt)	12.7%
Mecklenburg-Vorpommern:	Sachbeschädigung	12.6%
Sachsen:	Nebengesetze (Asyl/Drogen)	12.6%
Bayern:	Betrug (Gesamt)	12.4%
Rheinland-Pfalz:	Körperverletzung	12.4%
Saarland:	Körperverletzung	12.3%
Berlin:	Betrug (Gesamt)	11.9%

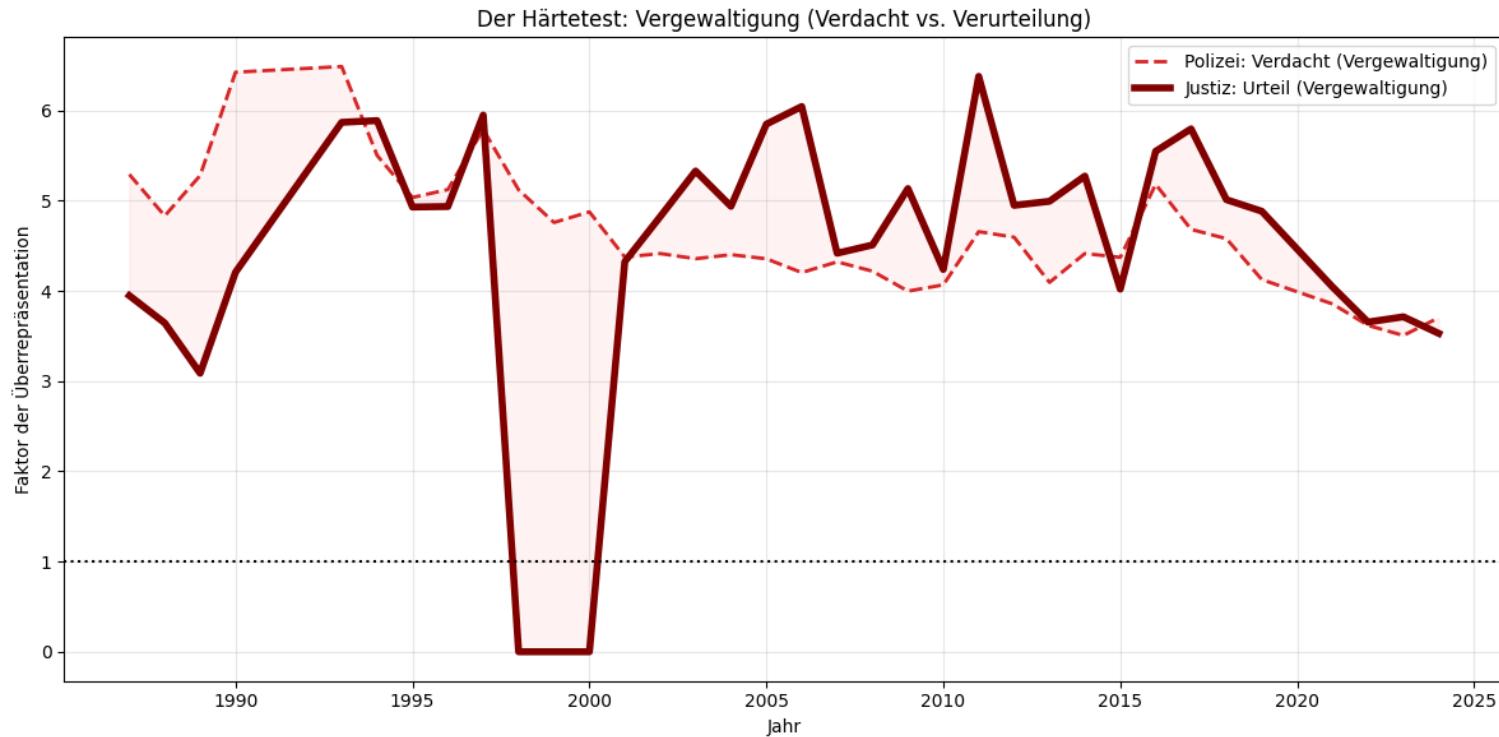
Data Analysis Per Capita Ausländer Repräsentation



Data Analysis Per Capita Vergleich Polizei gegenüber Justiz

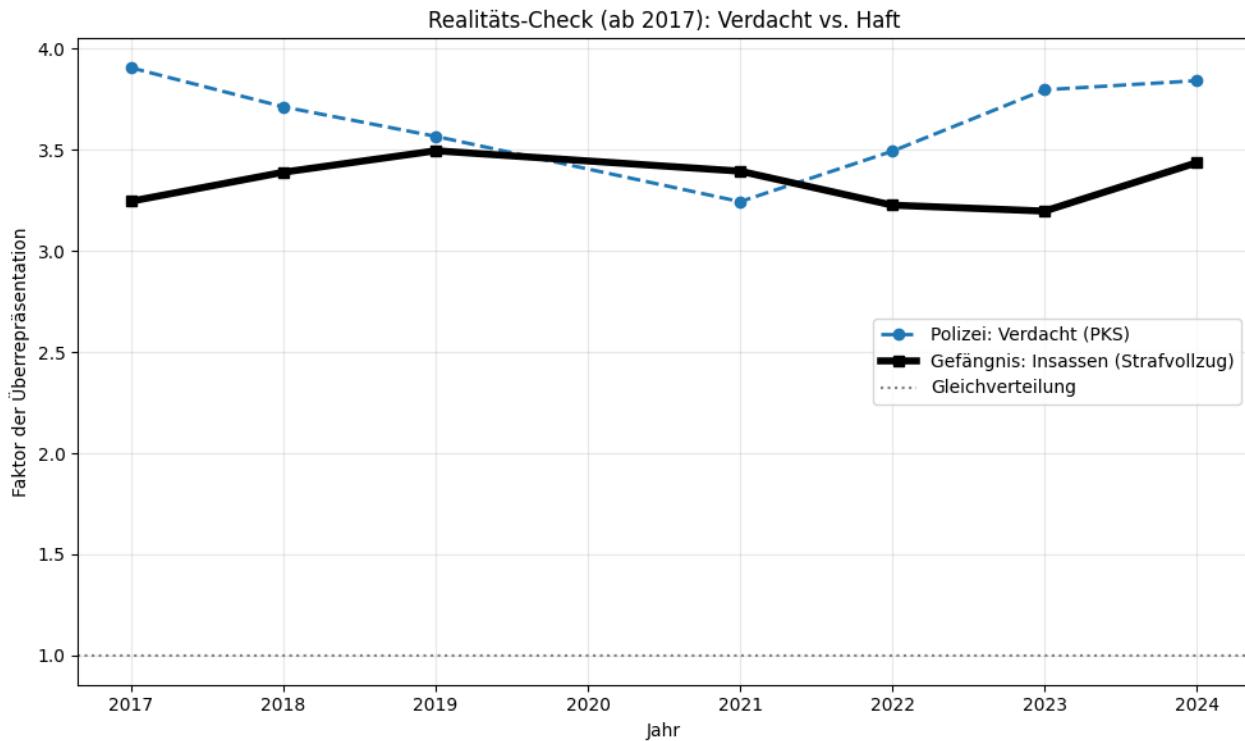


Data Analysis Überrepräsentation bei Vergewaltigungen



Data Analysis Faktische Überrepräsentation im Strafvollzug

- Überrepresentation bleibt auch in den JVAs hoch
- Bagatelldelikte scheinen nicht für die Verzerrung verantwortlich



Die Einordnung



*4

Das Nenner-Problem verzerrung durch:

- Touristen
- Pendler
- stationierte Streitkräfte
- Dunkelziffer

Das Fazit

- **Validierung durch Justiz:** Haft- und Urteilsdaten bestätigen PKS-Trends
- **Statistische Evidenz:** Konsistente Überrepräsentation auch bei Schwerkriminalität
- **Technologie:** Apache Spark ermöglicht performante ETL-Pipeline und Skalierbarkeit.
- **Datenqualität:** Exzellenze Struktur der PKS und Genesis Datenbank

“One of the great mistakes is to judge policies and programs by their intentions rather than their results.”

– Milton Friedman



Vielen Dank für Ihre Aufmerksamkeit!

Quellenverzeichnis

- *1 https://de.linkedin.com/posts/statistisches-bundesamt-destatis_statistischesbundesamt-datenbank-activity-7346424773578551296-ygUL
- *2 https://www.bka.de/DE/AktuelleInformationen/StatistikenLagebilder/PolizeilicheKriminalstatistik/pks_node.html
- *3 <https://www.c-sharpcorner.com/article/understanding-medallion-architecture-in-microsoft-fabric-lakehouse/>
- *4 https://www.reichelt.de/de/de/shop/produkt/warnaufkleber_gefahrenstelle-54447?PROVID=2788&gad_source=1&gad_campaignid=19009429406&gbraid=0AAAAADwnxtag9GDIUDYpfCv-8toWd5dz&gclid=Cj0KCQiA4eHLBhCzARIsAJ2NZoIXn2JWv2fDcBeqL82x7N23CGFMrCVNsXzYMoIDSKT0P7pxMTpBzOsaAkC2EALw_wcB