Package 'AIGovernance'

November 12, 2026

Type Package

Title Statistical Auditing, Drift Monitoring and Governance Reporting for AI Decision Systems

Version 0.2.0

Description Provides statistical auditing, risk documentation, and reporting tools to support AI governance workflows for employment, hiring and other AI-assisted decision systems. Implements the EEOC four-fifths adverse impact rule (Equal Employment Opportunity Commission, 1978, <https://www.ecfr.gov/current/title-29/subtitle-B/chapter-XIV/part-1607>), NYC Local Law 144 bias audit requirements (New York City, 2023, <https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page>), and the AI Risk Management Framework checklist items from the National Institute of Standards and Technology (2023, <doi:10.6028/NIST.AI.100-1>). Optionally supports EU AI Act high-risk classification (European Parliament and Council, 2024, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>). It also audits decision drift over time: the change in a decision rate is decomposed into case-mix, allocation (routing of cases to decision units) and rule-change components in the spirit of Kitagawa (1955) <doi:10.1080/01621459.1955.10501299>; net and absolute drift are estimated over a fixed reference case mix with overlap trimming and cluster bootstrap uncertainty; categorical decisions and unit-level drift with noise-corrected heterogeneity are supported; and a simulator with known drift mechanisms is included. The package does not provide legal advice or certify legal compliance; it is a statistical and documentation support tool.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 4.1.0)

Imports cli, rlang, stats, utils, tibble, dplyr, nnet

Suggests ggplot2, knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Language en-US

Config/testthat/edition 3

URL https://github.com/subirhait/AIGovernance

BugReports https://github.com/subirhait/AIGovernance/issues

NeedsCompilation no

Author Subir Hait [aut, cre] (ORCID: <https://orcid.org/0009-0004-9871-9677>)

Maintainer Subir Hait <haitsubi@msu.edu>

Config/roxygen2/version 8.1.1

Repository CRAN

Date/Publication 2026-11-12 14:50:34 UTC

Contents

aigov_adverse_impact 2
aigov_audit_nist 4
aigov_audit_nyc 5
aigov_build 6
aigov_checklist 8
aigov_classify 8
aigov_drift 10
aigov_monitor 13
aigov_report 14
aigov_scope 15
aigov_simulate_drift 16
hiring_panel 17
hiring_sim 18

Index **20**

`aigov_adverse_impact` *EEOC Adverse Impact Analysis (4/5ths Rule)*

Description

Computes selection rates by group and applies the EEOC Uniform Guidelines 4/5ths (80%) rule to assess adverse impact in employment selection procedures (EEOC, 1978).

The **Adverse Impact Ratio (AIR)** for each group is:

$$AIR_g = \frac{\text{selection rate}_g}{\text{selection rate}_{\text{ref}}}$$

A group is flagged for adverse impact when $AIR < 0.80$. The function also reports the two-proportion Z-test and Fisher's exact test p-values as supplementary statistics.

Note: The 4/5ths rule is a rule of thumb, not a bright-line legal standard. Small sample sizes reduce reliability. This function does not provide legal advice.

Usage

```
aigov_adverse_impact(gov, min_n = 30L)
```

Arguments

gov	An aigov object from aigov_build.
min_n	Integer. Minimum group sample size for a reliable AIR estimate. Groups below this threshold are flagged with a warning. Default 30.

Value

The input gov object with gov$results$adverse_impact appended, a tibble containing:

group Group label.

n Total applicants in group.

n_selected Number selected.

selection_rate Proportion selected.

AIR Adverse Impact Ratio vs reference group.

fourfifths_flag Logical: TRUE if AIR < 0.80.

z_stat Two-proportion Z statistic.

p_value Two-sided p-value from Z test.

fisher_p Fisher's exact test p-value.

small_n_flag Logical: TRUE if n < min_n.

References

Equal Employment Opportunity Commission (1978). Uniform guidelines on employee selection procedures. *Federal Register*, 43(166), 38295–38309.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_adverse_impact(gov)
gov$results$adverse_impact
```

`aigov_audit_nist` *NIST AI Risk Management Framework (AI RMF 1.0) Audit*

Description

Implements a structured checklist audit aligned with the NIST AI Risk Management Framework (AI RMF 1.0, NIST, 2023). The RMF organises AI risk management into four core functions: **GOVERN**, **MAP**, **MEASURE**, and **MANAGE**.

This function presents the checklist items most relevant to employment AI systems and records user-supplied responses (or defaults to `NA` for items that cannot be verified from data alone).

Usage

```
aigov_audit_nist(gov, responses = list())
```

Arguments

`gov`	An `aigov` object from `aigov_build`.
`responses`	An optional named list of logical values (`TRUE`/ `FALSE`) for checklist items. Use `aigov_checklist(gov,"NIST_RMF")` to see item names. Items not supplied default to `NA`. Some items are filled automatically from modules already run: `MEASURE_1_2` after `aigov_adverse_impact`, and `MEASURE_3_1` after `aigov_drift` or `aigov_monitor`. A value you supply always takes precedence. `MANAGE_4_1` (periodic review of monitoring results) requires a human review and is never set automatically.

Value

The input `gov` object with `gov$results$nist_rmf` appended, containing:

`checklist` Named list with `TRUE`/`FALSE`/`NA` per item.

`scores` Per-function completion scores (proportion of confirmed items).

`overall_score` Overall confirmed proportion.

`verdict` One of `"GREEN"` (>= 0.75), `"AMBER"` (0.50–0.74), or `"RED"` (< 0.50).

References

National Institute of Standards and Technology (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1. doi:10.6028/NIST.AI.1001

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_audit_nist(gov, responses = list(
  GOVERN_1_1 = TRUE,
  GOVERN_1_2 = TRUE,
```

```
  MAP_1_1     = TRUE
))
```

`aigov_audit_nyc` *NYC Local Law 144 Bias Audit Module*

Description

Implements the statistical components of a NYC Local Law 144 (2023) bias audit for Automated Employment Decision Tools (AEDTs). The law requires employers using AEDTs to:

1. Conduct an annual independent bias audit.
2. Publish a summary of audit results on their website.
3. Provide advance notice to candidates/employees.

This function computes the required **impact ratio** statistics for each race/ethnicity and sex category and generates a publication-ready summary table matching the format expected in public disclosures.

The NYC LL144 impact ratio is:

$$IR_g = \frac{\text{selection rate}_g}{\text{selection rate of most-selected category}}$$

Note: LL144 uses the *most-selected category* (not a user-specified reference group) as the denominator – this differs from the EEOC approach. This function implements both.

Usage

```
aigov_audit_nyc(gov, use_most_selected = TRUE)
```

Arguments

`gov`	An `aigov` object from `aigov_build`.
`use_most_selected`	Logical. If `TRUE` (default), the denominator is the group with the highest selection rate (NYC LL144 standard). If `FALSE`, uses `gov$ref_group` (EEOC standard).

Value

The input `gov` object with `gov$results$nyc_ll144` appended, containing:

`impact_table` Tibble with selection rates and impact ratios.

`most_selected_group` The reference category used.

`disclosure_table` Formatted table for public disclosure.

`checklist` LL144 procedural checklist (named logical vector).

`verdict` Statistical verdict: `"PASS"` or `"REVIEW"`.

References

New York City Local Law 144 of 2021 (effective January 1, 2023). NYC Department of Consumer and Worker Protection (DCWP). https://www.nyc.gov/site/dca/about/automated-employment-decision-tools.page

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White",
                   frameworks = c("NYC_LL144"))
gov <- aigov_audit_nyc(gov)
```

`aigov_build` *Build an AIGovernance audit object*

Description

Constructs an `aigov` object from employment decision data. This is the entry point for all auditing, classification, and reporting functions in the **AIGovernance** package.

Disclaimer: AIGovernance provides statistical and documentation support tools only. It does not provide legal advice and does not certify compliance with any law or regulation.

Usage

```
aigov_build(
  data,
  outcome,
  group,
  ref_group,
  frameworks = c("EEOC", "NYC_LL144", "NIST_RMF"),
  org_name = NULL,
  system_name = NULL,
  audit_date = NULL
)
```

Arguments

`data`	A data frame containing the employment decision records.
`outcome`	Unquoted column name of the binary decision variable (1 = selected / hired / advanced; 0 = not selected).
`group`	Unquoted column name of the protected-class variable (e.g., race/ethnicity, gender).
`ref_group`	Character string identifying the reference group (typically the highest-selection-rate group, e.g. `"White"` or `"Male"`).

frameworks	Character vector of governance frameworks to activate. One or more of "EEOC", "NYC_LL144", "NIST_RMF", "EU_AI_Act". Default is c("EEOC", "NYC_LL144", "NIST_RMF").
org_name	Optional character string — organisation name for reports.
system_name	Optional character string — name of the AI system being audited (e.g., "Resume screening tool v2.1").
audit_date	Optional Date or character string (ISO format). Defaults to Sys.Date().

Value

An object of class "aigov" containing:

data The input data frame.

outcome Name of the outcome column (character).

group Name of the group column (character).

ref_group Reference group label.

frameworks Active frameworks.

org_name Organisation name.

system_name AI system name.

audit_date Audit date.

group_levels All observed group labels.

n_total Total number of records.

Examples

```
data(hiring_sim)
gov <- aigov_build(
  data        = hiring_sim,
  outcome     = selected,
  group       = race_ethnicity,
  ref_group   = "White",
  frameworks  = c("EEOC", "NYC_LL144", "NIST_RMF"),
  org_name    = "Acme Corp",
  system_name = "Resume Screening Tool v1.0"
)
print(gov)
```

aigov_checklist *Display Governance Checklist for a Framework*

Description

Returns all checklist item names and descriptions for a given framework. Use the returned item names as keys in the `responses` argument of `aigov_audit_nist`.

Usage

```
aigov_checklist(gov, framework)
```

Arguments

`gov`	An `aigov` object (used only for class checking).
`framework`	Character. One of `"NYC_LL144"`, `"EEOC"`, `"NIST_RMF"`, or `"EU_AI_Act"`.

Value

A tibble with columns `item_id`, `function_area`, and `description`.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
aigov_checklist(gov, "NYC_LL144")
aigov_checklist(gov, "NIST_RMF")
```

aigov_classify *Classify AI System Risk Level*

Description

Classifies the AI system under two frameworks:

1. **EU AI Act (2024)** – Assigns one of four risk tiers: Unacceptable, High, Limited, or Minimal risk. Employment/worker-management AI is listed in Annex III as *High Risk*.
2. **NIST AI RMF** – Assigns a risk tier (1–4) based on impact on individuals' rights and opportunities.

Usage

```
aigov_classify(
  gov,
  domain = "employment",
  makes_final_decision = TRUE,
  human_oversight = NA
)
```

Arguments

gov	An aigov object from aigov_build.
domain	Character. Application domain. One of "employment", "education", "credit", "housing", "healthcare", "law_enforcement", "other". Default "employment".
makes_final_decision	Logical. Does the AI system make or substantially influence a final employment decision? Default TRUE.
human_oversight	Logical. Is meaningful human review in place before the AI decision takes effect? Default NA (unknown).

Value

The input gov object with gov$results$risk_class appended, containing EU AI Act and NIST risk classifications with explanatory text.

References

European Parliament and Council (2024). Regulation (EU) 2024/1689 (EU AI Act). https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

NIST (2023). *AI RMF 1.0*. doi:10.6028/NIST.AI.1001

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_classify(gov, domain = "employment",
                      makes_final_decision = TRUE,
                      human_oversight = FALSE)
```

`aigov_drift` *Audit Decision Drift Between Two Periods*

Description

Separates a change in decision rates between two periods into changes in the cases being decided and changes in how comparable cases are decided. Single-period audits such as the four-fifths rule or the NYC Local Law 144 impact ratio cannot make this distinction.

Let D be the decision, X the recorded case characteristics (`covariates`), G the group and t the period. The function estimates the system-level decision rule $m_t(x, g) = P_t(D = k \mid X = x, G = g)$ in each period and reports:

Decomposition The raw change in the rate of each decision category is split into a *case-mix* component and a *rule-change* component evaluated at the current case mix. When `unit` is given (for example recruiter, office, judge or court), the change is split three ways: case mix, *allocation* of cases to units, and *within-unit* rule change. The components add to the raw change exactly in the analysis sample.

Net drift $\Delta_g = E_{Q_g}[m_1(X, g) - m_0(X, g)]$ for the focal category, over a fixed reference case mix Q_g.

Absolute drift $A_g = E_{Q_g}[\frac{1}{2}\sum_k |m_1(X, g)_k - m_0(X, g)_k|]$, the mean total-variation change in the decision rule. For a binary decision this is $E_{Q_g}|m_1 - m_0|$. Increases and decreases for different cases do not cancel, so A_g can be large when Δ_g is near zero ("masked drift").

Unit drift With `unit`, the drift of each unit over a common reference case mix, with delta-method standard errors, a noise-corrected between-unit variance τ^2, empirical-Bayes shrunken estimates and Benjamini-Hochberg q-values.

Cases outside the region where both periods have comparable cases are removed before estimation (`overlap_trim`). Uncertainty is estimated with a bootstrap that resamples units (or `cluster`) when given.

Interpretation. Drift is a change in decisions *conditional on the recorded characteristics*. It does not identify the causal effect of an AI system or of human discretion, and a change in an unrecorded decision-relevant characteristic appears as drift. The absolute-drift estimate is biased upward by estimation noise; `abs_drift_bc` reports a bootstrap bias-corrected value, and alerts for absolute drift compare the lower confidence bound with `tolerance` rather than with zero. The decomposition and drift estimates are model-based standardizations: they are accurate when the decision models are adequately specified. When the true rule is nonlinear in the covariates (for example because cases are routed to units according to their characteristics), add flexible terms with `extra_terms` and compare. These procedures are approximate and should be reviewed with the overlap and sample-size diagnostics. This function does not provide legal advice.

Usage

```
aigov_drift(
  gov,
  period,
  covariates,
```

```
  extra_terms = NULL,
  unit = NULL,
  decision = NULL,
  baseline = NULL,
  current = NULL,
  reference = c("pooled", "baseline", "current"),
  overlap_trim = 0.02,
  tolerance = 0.05,
  focal = NULL,
  cluster = NULL,
  B = 200L,
  conf_level = 0.95,
  min_unit_n = 20L,
  max_ref = 2000L,
  seed = NULL,
  verbose = TRUE
)
```

Arguments

gov	An aigov object from aigov_build.
period	Column identifying the period (unquoted name or string).
covariates	Character vector of recorded case characteristics available before the decision.
extra_terms	Optional character vector of additional model terms built from covariates, for example "I(qualification^2)", "qualification:degree" or "splines::ns(qualification, 3)". Used in the decision and overlap models. Comparing results with and without such terms is a check of functional-form sensitivity.
unit	Optional column identifying the decision unit. Enables the three-way decomposition and unit-level drift.
decision	Optional column with a categorical decision (for example reject / review / advance). Defaults to the binary outcome given to aigov_build().
baseline, current	Values of period to compare. Default: the first and last sorted values.
reference	Reference case mix for net and absolute drift: "pooled" (default; both periods), "baseline" or "current".
overlap_trim	Cases with an estimated period-membership probability outside [overlap_trim, 1 - overlap_trim] are removed. Use 0 to keep all cases.
tolerance	Practical tolerance on the probability scale (default 0.05). An alert requires the confidence interval to lie beyond this value, so changes that are plausibly within tolerance never trigger an alert.
focal	Decision category for net and unit drift. Default: the last category (1 for a binary outcome).
cluster	Optional column for cluster bootstrap resampling. Defaults to unit when given.
B	Number of bootstrap replicates (default 200). Use 0 to skip uncertainty estimation.

conf_level	Confidence level for intervals (default 0.95).
min_unit_n	Minimum decisions per unit in each period for the unit to enter the multilevel analysis (default 20).
max_ref	Maximum number of reference cases used for unit-level drift (default 2000).
seed	Optional random seed for the bootstrap.
verbose	Print a summary (default TRUE).

Value

The input gov object with gov$results$drift, a list containing:

decomposition Tibble by group and decision category: rates, raw change, case-mix, allocation (multilevel only), rule-change and net drift, with bootstrap percentile intervals.

drift Tibble by group: net drift for the focal category, absolute drift, bias-corrected absolute drift, intervals, p-values and alert. net_p tests $\Delta_g = 0$; net_p_tol is a one-sided test of $|\Delta_g| \leq$ tolerance (normal approximation). Alerts use interval nulls: "net drift" when the whole interval for Δ_g lies beyond $\pm$ tolerance; "masked drift" when the lower bound for A_g exceeds tolerance while $|\hat{\Delta}_g| \leq$ tolerance; "conditional drift" when the lower bound for A_g exceeds tolerance otherwise; "none" otherwise.

overlap Cases kept and trimmed by group.

units Unit-level drift table and summary (multilevel only).

settings Analysis settings.

verdict "PASS" if no alert and no flagged unit, otherwise "REVIEW".

References

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, 50(272), 1168–1194. doi:10.1080/01621459.1955.10501299

Cai, T. T., Namkoong, H., & Yadlowsky, S. (2023). Diagnosing model performance under distribution shift. *arXiv*. doi:10.48550/arXiv.2303.02011

See Also

aigov_monitor, aigov_simulate_drift

Examples

```
data(hiring_panel)
gov <- aigov_build(hiring_panel, selected, group, ref_group = "A")
gov <- aigov_drift(gov, period = period,
                   covariates = c("qualification", "degree"),
                   baseline = 2024, current = 2026, B = 30, seed = 1)
gov$results$drift$drift


# three-way decomposition with recruiters as decision units
gov <- aigov_drift(gov, period = period,
```

```
                    covariates = c("qualification", "degree"),
                    unit = recruiter, baseline = 2024, current = 2026,
                    B = 100, seed = 1)
gov$results$drift$units$summary
```

---

aigov_monitor *Monitor Decision Drift Across Several Periods*

Description

Runs aigov_drift for a sequence of prespecified comparisons (each period against the baseline, or each period against the previous one) and adjusts the net-drift p-values for the number of comparisons. A net-drift alert is kept only if the adjusted one-sided p-value for $|\Delta| \leq$ tolerance is below $(1-$ conf_level$)/2$. This supports the periodic monitoring of deployed decision systems.

Note: the adjustment covers the comparisons made in one call. It does not make repeated, data-dependent looks at accumulating data valid; that requires sequential (anytime-valid) methods.

Usage

```
aigov_monitor(
  gov,
  period,
  covariates,
  unit = NULL,
  decision = NULL,
  compare = c("baseline", "previous"),
  adjust = "holm",
  tolerance = 0.05,
  B = 200L,
  seed = NULL,
  verbose = TRUE,
  ...
)
```

Arguments

gov	An aigov object from aigov_build.
period	Column identifying the period (unquoted name or string).
covariates	Character vector of recorded case characteristics available before the decision.
unit	Optional column identifying the decision unit. Enables the three-way decomposition and unit-level drift.
decision	Optional column with a categorical decision (for example reject / review / advance). Defaults to the binary outcome given to aigov_build().

compare	"baseline" (default) compares every later period with the first; "previous" compares consecutive periods.
adjust	Multiplicity adjustment passed to p.adjust (default "holm").
tolerance	Practical tolerance on the probability scale (default 0.05). An alert requires the confidence interval to lie beyond this value, so changes that are plausibly within tolerance never trigger an alert.
B	Number of bootstrap replicates (default 200). Use 0 to skip uncertainty estimation.
seed	Optional random seed for the bootstrap.
verbose	Print a summary (default TRUE).
...	Further arguments passed to aigov_drift, such as reference, overlap_trim, focal, cluster or min_unit_n.

Value

The input gov object with gov$results$monitor, a list with table (one row per comparison and group, with adjusted p-values and alerts), fits (the individual drift results) and verdict.

Examples

```
data(hiring_panel)
gov <- aigov_build(hiring_panel, selected, group, ref_group = "A")
gov <- aigov_monitor(gov, period = period,
                     covariates = c("qualification", "degree"),
                     B = 50, seed = 1)
gov$results$monitor$table
```

---

aigov_report *Generate an AI Governance Audit Report*

Description

Produces a self-contained HTML (or plain-text) governance audit report from a completed aigov object. The report is suitable for internal documentation, legal review, or public disclosure (e.g., NYC LL144 website posting requirement).

Usage

```
aigov_report(gov, format = "html", output_file = NULL, open = TRUE)
```

Arguments

gov	A completed aigov object (at least one audit module run).
format	Character. Output format: "html" (default) or "text".
output_file	Optional character. File path for the output. If NULL, a temporary file is used and opened in the browser (HTML) or printed to console (text).
open	Logical. If TRUE (default) and format = "html", attempt to open the report in the default browser.

Value

Invisibly returns the path to the generated file.

Examples

```
## Not run:
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
gov <- aigov_adverse_impact(gov)
gov <- aigov_audit_nyc(gov)
gov <- aigov_audit_nist(gov)
aigov_report(gov, format = "html")

## End(Not run)
```

---

aigov_scope *Determine Applicable Governance Frameworks*

Description

Returns a summary of which governance frameworks apply to the AI system described in the aigov object, based on domain and jurisdiction.

Usage

```
aigov_scope(gov, domain = "employment", us_state = NULL)
```

Arguments

gov	An aigov object from aigov_build.
domain	Character. Application domain (default "employment").
us_state	Optional character. US state, e.g. "NY", "CO", "IL" for state-specific law flags.

Value

A tibble with columns framework, applies, jurisdiction, and note.

Examples

```
data(hiring_sim)
gov <- aigov_build(hiring_sim, selected, race_ethnicity, ref_group = "White")
aigov_scope(gov, domain = "employment", us_state = "NY")
```

`aigov_simulate_drift` *Simulate Decision Data With Known Drift Mechanisms*

Description

Generates multi-period decision records from a known data-generating process so that the targets of `aigov_drift` are known exactly. Used for testing, teaching and simulation studies of audit procedures.

Cases have a continuous `qualification` score, a binary `degree`, a group (`"A"`, `"B"`, `"C"`) and an unrecorded binary characteristic z. Each case is routed to one of `n_units` decision units (`recruiter`). The decision follows a logistic (binary) or cumulative-logistic (three-category) model:

$$\eta = -0.2 + qualification + 0.6\,degree - z + \gamma_g + \alpha_j + v_j s_t + \delta_t(x, g),$$

where s_t rises from 0 in the first period to 1 in the last.

Mechanisms (`scenario`; several may be combined):

`"stable"` Nothing changes.

`"case_mix"` Mean qualification falls by 0.5 SD.

`"rule"` All groups face a stricter rule ($\delta = -$`effect`).

`"offsetting_groups"` Group B gains and group A loses by `effect`.

`"offsetting_cases"` Within every group, cases with a degree gain and cases without lose by `effect`.

`"unobserved"` The share of cases with the unrecorded z rises from 0.3 to 0.6; the full-information rule does not change.

`"reallocation"` High-qualification cases are routed to more lenient units; no unit changes its rule.

`"unit_offsetting"` Half of the units become more lenient and half stricter by `effect`; the average unit does not change.

`"category_shift"` (Requires `categorical = TRUE`.) The threshold between "review" and "advance" rises by `effect` while the share passing the first screen is unchanged.

Usage

```
aigov_simulate_drift(
  n = 2000L,
  n_periods = 2L,
  n_units = 30L,
  scenario = "stable",
```

```
  effect = 0.8,
  categorical = FALSE,
  first_period = 2024L,
  seed = NULL
)
```

Arguments

n	Number of cases per period (default 2000).
n_periods	Number of periods (default 2).
n_units	Number of decision units (default 30).
scenario	Character vector of mechanisms (see Description).
effect	Size of rule changes on the logit scale (default 0.8).
categorical	If TRUE, adds a three-category decision (reject / review / advance) and defines selected as passing the first screen (review or advance).
first_period	Label of the first period (default 2024).
seed	Optional random seed.

Value

A data frame with columns case_id, period, group, qualification, degree, recruiter, selected and, if categorical = TRUE, decision. The attribute "truth" is a list with drift (true net and absolute drift of selected, by group, between the first and last period, over the pooled cases conditional on the recorded characteristics; with categorical = TRUE also the true net drift of "advance" and the true total-variation drift of decision), decomposition (true case-mix, allocation and within-unit rule-change components of the change in selected, evaluated on the realized cases) and units (true unit drift over the pooled cases).

Examples

```
sim <- aigov_simulate_drift(n = 500, scenario = "offsetting_cases", seed = 1)
attr(sim, "truth")$drift
```

---

hiring_panel *Simulated Multi-Period Hiring Decisions With Known Drift*

Description

A synthetic panel of screening decisions made by 24 recruiters over three annual periods, generated with aigov_simulate_drift. It is built so that a single-period audit and a drift audit tell different stories: applicant qualifications fall over time (case-mix change), the rule changes in opposite directions for applicants with and without a degree (masked drift), stronger applicants are increasingly routed to more lenient recruiters (allocation change), and the threshold between "review" and "advance" rises (category shift). All records are fictional.

Usage

```
hiring_panel
```

Format

A data frame with 4,500 rows and 8 variables:

`case_id` Integer case identifier.

`period` Integer period (2024, 2025, 2026).

`group` Character group label (`"A"`, `"B"`, `"C"`).

`qualification` Numeric standardized qualification score.

`degree` Integer (0/1) degree indicator.

`recruiter` Character recruiter (decision unit) identifier.

`selected` Integer (0/1): passed the first screen (`decision` is "review" or "advance").

`decision` Factor: "reject", "review" or "advance".

Details

The attribute `"truth"` holds the true net and absolute drift between 2024 and 2026 by group and the true drift of each recruiter. See `data-raw/hiring_panel.R` in the package sources for the generating call.

Source

Simulated with `aigov_simulate_drift(n = 1500, n_periods = 3, n_units = 24, scenario = c("case_mix", "offsetting_cases", "reallocation", "category_shift"), categorical = TRUE, seed = 2026)`.

Examples

```
data(hiring_panel)
with(hiring_panel, tapply(selected, list(group, period), mean))
attr(hiring_panel, "truth")$drift
```

`hiring_sim` *Simulated Employment Screening Dataset*

Description

A synthetic dataset of 500 job applicants processed by a hypothetical automated resume-screening tool. Generated for illustrative and testing purposes only. All individuals are fictional.

Usage

```
hiring_sim
```

Format

A data frame with 500 rows and 6 variables:

`applicant_id` Integer applicant identifier.

`race_ethnicity` Character. One of `"White"`, `"Black"`, `"Hispanic"`, `"Asian"`, `"Other"`.

`gender` Character. `"Male"` or `"Female"`.

`years_experience` Numeric. Years of relevant experience.

`score` Numeric. AI screening score (0–100).

`selected` Integer (0/1). Whether the applicant was advanced to the next stage (1 = selected, 0 = not selected).

Details

Selection probabilities were set to produce a realistic adverse impact pattern across race/ethnicity groups, consistent with published empirical ranges. The data are purely synthetic and do not represent any real organisation or hiring process.

Source

Simulated by the package authors for illustration purposes.

Examples

```
data(hiring_sim)
head(hiring_sim)
table(hiring_sim$race_ethnicity, hiring_sim$selected)
```

Index

∗ **datasets**
hiring_panel, 17
hiring_sim, 18

aigov_adverse_impact, 2, *4*
aigov_audit_nist, 4, *8*
aigov_audit_nyc, 5
aigov_build, *3–5*, 6, *9*, *11*, *13*, *15*
aigov_checklist, 8
aigov_classify, 8
aigov_drift, *4*, 10, *13, 14*, *16*
aigov_monitor, *4*, *12*, 13
aigov_report, 14
aigov_scope, 15
aigov_simulate_drift, *12*, 16, *17*

hiring_panel, 17
hiring_sim, 18

p.adjust, *14*