Package: nhscancerwaits (via r-universe)

June 20, 2026

Type Package

Title NHS Cancer Waiting-Time Analysis, Benchmarking and Multilevel Modelling

Version 1.0.1

Description Provides tools for importing, harmonising, cleaning, analysing, benchmarking and visualising National Health Service (NHS) England Cancer Waiting Times data. The package supports national performance monitoring, provider-level benchmarking and cancer pathway comparisons through key performance indicator summaries, provider filtering, clustering analyses, mixed-effects regression models, variance decomposition, intraclass correlation coefficient estimation, adjusted provider performance estimation and sensitivity analyses. Functions are included for exploratory analysis, publication-ready visualisations and spreadsheet exports, supporting reproducible health services research, cancer services evaluation, quality improvement and assessment of waiting-time performance across healthcare organisations.

License MIT + file LICENSE

Encoding UTF-8

Language en-GB

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 4.2.0)

Imports dplyr, tidyr, readr, readxl, stringr, lubridate, rlang, ggplot2, scales, lme4, broom.mixed, performance, cluster, stats, writexl

Suggests testthat (>= 3.0.0), knitr, rmarkdown, quarto, covr

VignetteBuilder knitr

Config/testthat/edition 3

URL https://github.com/zerish12/nhscancerwaits

BugReports https://github.com/zerish12/nhscancerwaits/issues

Config/pak/sysreqs cmake make libicu-dev libx11-dev zlib1g-dev

Repository https://zerish12.r-universe.dev

Date/Publication 2026-05-31 01:28:56 UTC

RemoteUrl https://github.com/zerish12/nhscancerwaits

RemoteRef HEAD

RemoteSha 3a6bcd5963ae1f3ff72550bce75698a06370483d

Contents

calculate_icc *Calculate Intraclass Correlation Coefficient*

Description

Calculates the intraclass correlation coefficient from a fitted mixed-effects model.

Usage

```
calculate_icc(model)
```

Arguments

model	A fitted mixed-effects model from `lme4::lmer()`.

Value

A data frame with provider variance, residual variance, total variance and ICC.

calculate_silhouette_score *Calculate Provider Clustering Silhouette Score*

Description

Calculates the average silhouette score for provider clusters.

Usage

```
calculate_silhouette_score(clustered_data)
```

Arguments

clustered_data	Data frame from `cluster_providers()`.

Value

Average silhouette score.

check_cwt_model *Check Mixed-Model Performance Diagnostics*

Description

Returns standard model diagnostics using the performance package.

Usage

```
check_cwt_model(model)
```

Arguments

model	A fitted mixed-effects model.

Value

A model diagnostic object.

clean_cwt_data *Clean NHS Cancer Waiting-Time Data*

Description

Standardises column names, provider fields, reporting dates, performance values and activity denominators.

Usage

```
clean_cwt_data(data)
```

Arguments

data	A data frame of NHS Cancer Waiting-Time records.

Value

A cleaned data frame.

cluster_providers *Cluster Providers by Performance and Activity*

Description

Groups providers into performance clusters using provider-level summaries.

Usage

```
cluster_providers(
  data,
  provider_var = "provider_code",
  performance_var = "performance_percent",
  activity_var = "total_treated",
  k = 3
)
```

Arguments

data	A cleaned and filtered cancer waiting-time data frame.
provider_var	Provider column. Default is "provider_code".
performance_var	Performance percentage column. Default is "performance_percent".
activity_var	Activity denominator column. Default is "total_treated".
k	Number of clusters. Default is 3.

Value

A data frame containing provider summaries and cluster labels.

export_excel_tables *Export Analysis Tables to Excel*

Description

Exports one or more analysis tables to an Excel workbook.

Usage

```
export_excel_tables(tables, path = "nhscancerwaits_outputs.xlsx")
```

Arguments

tables	A named list of data frames.
path	Output Excel file path. Default is "nhscancerwaits_outputs.xlsx".

Value

The output file path, invisibly.

`export_provider_rankings` *Export Provider Rankings to Excel*

Description

Exports adjusted provider effects or provider ranking tables to Excel.

Usage

```
export_provider_rankings(provider_rankings, path = "provider_rankings.xlsx")
```

Arguments

`provider_rankings`	A provider ranking data frame.
`path`	Output Excel file path. Default is `"provider_rankings.xlsx"`.

Value

The output file path, invisibly.

`extract_model_effects` *Extract Fixed Effects from Mixed Model*

Description

Extract Fixed Effects from Mixed Model

Usage

```
extract_model_effects(model)
```

Arguments

`model`	A fitted mixed-effects model.

Value

A data frame of fixed-effect estimates.

extract_provider_effects *Extract Adjusted Provider Effects*

Description

Extracts provider-specific random effects from a fitted mixed-effects model.

Usage

```
extract_provider_effects(model, provider_name = "provider_code")
```

Arguments

model	A fitted model from fit_cwt_mixed_model().
provider_name	Name of provider grouping variable.

Value

A data frame of adjusted provider effects.

filter_providers *Filter Providers for Reliable Analysis*

Description

Keeps providers with sufficient activity, repeated observations and stable performance values.

Usage

```
filter_providers(
  data,
  provider_var = "provider_code",
  activity_var = "total_treated",
  performance_var = "performance_percent",
  min_mean_activity = 20,
  min_observations = 5,
  max_cv = 0.5
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
provider_var	Provider code column. Default is "provider_code".
activity_var	Activity denominator column. Default is "total_treated".
performance_var	Performance percentage column. Default is "performance_percent".
min_mean_activity	Minimum mean activity per provider.
min_observations	Minimum number of observations per provider.
max_cv	Maximum coefficient of variation.

Value

A filtered data frame.

---

fit_cwt_mixed_model *Fit Mixed-Effects Model for Cancer Waiting-Time Performance*

Description

Fits a random-intercept mixed-effects model with provider as the grouping variable.

Usage

```
fit_cwt_mixed_model(
  data,
  performance_var = "performance_percent",
  month_var = "month_index",
  pathway_var = "cancer_type",
  provider_var = "provider_code"
)
```

Arguments

data	A cleaned and filtered cancer waiting-time data frame.
performance_var	Outcome column. Default is "performance_percent".
month_var	Month index column. Default is "month_index".
pathway_var	Cancer pathway column. Default is "cancer_type".
provider_var	Provider column. Default is "provider_code".

Value

A fitted lmerMod model object.

load_cwt_data *Load NHS Cancer Waiting-Time CSV Files*

Description

Reads one or more NHS Cancer Waiting-Time CSV files and combines them into a single data frame.

Usage

```
load_cwt_data(files)
```

Arguments

files	Character vector of CSV file paths.

Value

A data frame containing combined cancer waiting-time records.

load_cwt_excel *Load NHS Cancer Waiting-Time Excel Files*

Description

Reads one or more NHS Cancer Waiting-Time Excel files and combines them into a single data frame.

Usage

```
load_cwt_excel(files, sheet = 1, .name_repair = "unique")
```

Arguments

files	Character vector of Excel file paths.
sheet	Sheet name or sheet number. Default is 1.
.name_repair	Passed to readxl::read_excel(). Default is "unique".

Value

A combined tibble with a source_file column.

pivot_provider_months *Reshape Monthly Provider Performance to Wide Format*

Description

Creates a provider-by-month performance table.

Usage

```
pivot_provider_months(
  data,
  provider_var = "provider_code",
  month_var = "reporting_date",
  performance_var = "performance_percent"
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
provider_var	Provider column.
month_var	Month column.
performance_var	Performance column.

Value

A wide provider-by-month table.

plot_national_trends *Plot National Cancer Waiting-Time Trends*

Description

Creates a monthly performance trend plot.

Usage

```
plot_national_trends(
  data,
  month_var = "reporting_date",
  performance_var = "performance_percent",
  group_var = "standard"
)
```

Arguments

`data`	A cleaned cancer waiting-time data frame.
`month_var`	Reporting date column. Default is `"reporting_date"`.
`performance_var`	Performance percentage column. Default is `"performance_percent"`.
`group_var`	Grouping column. Default is `"standard"`.

Value

A ggplot object.

`plot_pathway_predictions` *Plot Adjusted Pathway Performance*

Description

Creates a horizontal plot of adjusted predicted performance by cancer pathway.

Usage

```
plot_pathway_predictions(
  pathway_predictions,
  pathway_var = "cancer_type",
  prediction_var = "predicted_performance"
)
```

Arguments

`pathway_predictions`	Data frame from `predict_pathway_performance()`.
`pathway_var`	Cancer pathway column. Default is `"cancer_type"`.
`prediction_var`	Predicted performance column. Default is `"predicted_performance"`.

Value

A ggplot object.

plot_provider_clusters *Plot Provider Clusters*

Description

Creates a scatter plot of provider clusters.

Usage

```
plot_provider_clusters(clustered_data)
```

Arguments

clustered_data	Data frame from cluster_providers().

Value

A ggplot object.

plot_provider_effects *Plot Adjusted Provider Effects*

Description

Creates a ranked plot of adjusted provider effects.

Usage

```
plot_provider_effects(
  provider_effects,
  provider_var = "provider_code",
  effect_var = "adjusted_effect"
)
```

Arguments

provider_effects	Data frame from extract_provider_effects().
provider_var	Provider column. Default is "provider_code".
effect_var	Adjusted effect column. Default is "adjusted_effect".

Value

A ggplot object.

predict_pathway_performance

Predict Adjusted Performance by Cancer Pathway

Description

Generates adjusted predicted performance for each cancer pathway from a fitted mixed-effects model.

Usage

```
predict_pathway_performance(
  model,
  data,
  pathway_var = "cancer_type",
  month_var = "month_index",
  provider_var = "provider_code"
)
```

Arguments

`model`	A fitted model from `fit_cwt_mixed_model()`.
`data`	The data frame used to fit the model.
`pathway_var`	Cancer pathway column. Default is `"cancer_type"`.
`month_var`	Month index column. Default is `"month_index"`.
`provider_var`	Provider column. Default is `"provider_code"`.

Value

A data frame with adjusted predicted pathway performance.

run_sensitivity_analysis

Run Sensitivity Analyses

Description

Runs mixed-effects models across alternative provider-filtering thresholds. If a cohort is too small or a model fails, the function returns a row with status information rather than stopping.

Usage

```
run_sensitivity_analysis(
  data,
  cohorts = data.frame(cohort = c("Main cohort", "No outlier removal",
   "Stricter cohort"), min_mean_activity = c(20, 20, 30), min_observations = c(5, 5, 5),
    max_cv = c(0.5, 0.75, 0.5)),
  provider_var = "provider_code",
  activity_var = "total_treated",
  performance_var = "performance_percent",
  month_var = "month_index",
  pathway_var = "cancer_type",
  min_providers = 3,
  min_rows = 20
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
cohorts	A data frame defining sensitivity cohorts.
provider_var	Provider column. Default is "provider_code".
activity_var	Activity denominator column. Default is "total_treated".
performance_var	Performance column. Default is "performance_percent".
month_var	Month index column. Default is "month_index".
pathway_var	Cancer pathway column. Default is "cancer_type".
min_providers	Minimum number of providers required to fit a model.
min_rows	Minimum number of rows required to fit a model.

Value

A data frame summarising model results across cohorts.

---

summarise_kpis *Summarise Cancer Waiting-Time KPIs*

Description

Calculates summary statistics for cancer waiting-time performance.

Usage

```
summarise_kpis(
  data,
  group_var = "standard",
  performance_var = "performance_percent"
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
group_var	Column used to group results. Default is "standard".
performance_var	Performance percentage column. Default is "performance_percent".

Value

A summary data frame.

---

summarise_pathways *Summarise Cancer Pathway Performance*

Description

Summarise Cancer Pathway Performance

Usage

```
summarise_pathways(
  data,
  pathway_var = "cancer_type",
  performance_var = "performance_percent"
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
pathway_var	Cancer pathway column.
performance_var	Performance column.

Value

Pathway-level summary table.

summarise_providers *Summarise Provider Performance*

Description

Summarise Provider Performance

Usage

```
summarise_providers(
  data,
  provider_var = "provider_code",
  performance_var = "performance_percent",
  activity_var = "total_treated"
)
```

Arguments

data	A cleaned cancer waiting-time data frame.
provider_var	Provider column.
performance_var	Performance column.
activity_var	Activity column.

Value

Provider-level summary table.

Index