Package ‘NHSRdatasets’

Type Package
Title NHS and Healthcare Related Data for Education and Training
Version 0.1.2
Maintainer Chris Mainey <chris.mainey@uhb.nhs.uk>
Description Free United Kingdom National Health Service (NHS) and other healthcare, or population health-related data for education and training purposes. This package currently contains a single simulated hospital dataset for teaching regression methods, with the addition of more datasets planned for future releases. This package exists to support skills development in the NHS-R community: <https://nhsrcommunity.com/>.
License CC0
Encoding UTF-8
LazyData true
RoxygenNote 6.1.1
Depends R (>= 3.5.0)
BugReports https://github.com/nhs-r-community/NHSRdatasets/issues
Suggests dplyr, ggplot2, lme4, MASS, ModelMetrics, lmtest, rmarkdown, scales, ggrepel, lubridate, tidyr, forcats, knitr
VignetteBuilder knitr
NeedsCompilation no
Author Chris Mainey [aut, cre] (<https://orcid.org/0000-0002-3018-6171>), Tom Jemmett [aut] (<https://orcid.org/0000-0002-6943-2990>), NHS-R community [cph]
Repository CRAN
Date/Publication 2019-09-27 12:20:02 UTC

R topics documented:

ae_attendances .......................................................... 2
LOS_model ...................................................................... 3

Index 5
ae_attendances

NHS England Accident & Emergency Attendances and Admissions

Description

Reported attendances, 4 hour breaches and admissions for all A&E departments in England for the years 2016/17 through 2018/19 (Apr-Mar). The data has been tidied to be easily usable within the tidyverse of packages.

Usage

data(ae_attendances)

Format

Tibble with six columns

- **period** The month that this data relates to
- **org_code** The ODS code for this provider
- **type** The department type, either 1, 2 or other
- **attendances** the number of patients who attended this department in this month
- **breaches** the number of patients who breaches the 4 hour target in this month
- **admissions** the number of patients admitted from A&E to the hospital in this month

Details

Data sourced from NHS England Statistical Work Areas which is available under the Open Government Licence v3.0

Source

NHS England Statistical Work Areas

Examples

data(ae_attendances)
library(dplyr)
library(ggplot2)
library(scales)

# Create a plot of the performance for England over time
ae_attendances %>%
group_by(period) %>%
summarise_at(vars(attendances, breaches), sum) %>%
mutate(performance = 1 - breaches / attendances) %>%
ggplot(aes(period, performance)) +
geom_hline(yintercept = 0.95, linetype = "dashed") +
geom_line() +
geom_point() +
scale_y_continuous(labels = percent) +
labs(title = "4 Hour performance over time")

# Now produce a plot showing the performance of each trust
ae_attendances %>%
group_by(org_code) %>%
# select organisations that have a type 1 department
filter(any(type == "1")) %>%
summarise_at(vars(attendances, breaches), sum) %>%
array(desc(attendances)) %>%
mutilate(performance = 1 - breaches / attendances,
overall_performance = 1 - sum(breaches) / sum(attendances),
rank = rank(-performance, ties.method = "first") / n()) %>%
ggplot(aes(rank, performance)) +
geom_vline(xintercept = c(0.25, 0.5, 0.75), linetype = "dotted") +
geom_hline(yintercept = 0.95, colour = "red") +
geom_hline(aes(yintercept = overall_performance), linetype = "dotted") +
geom_point() +
scale_y_continuous(labels = percent) +
theme_minimal() +
theme(panel.grid = element_blank(),
axis.text.x = element_blank()) +
labs(title = "4 Hour performance by trust",
subtitle = "Apr-16 through Mar-19",
x = "", y = "")

---

**LOS_model**

*Hospital Length of Stay (LOS) Data*

**Description**

Artificially generated hospital data. Fictional patients at 10 fictional hospitals, with LOS, Age and Date status data. Data were generate to learn Generalized Linear Models (GLM) concepts, modelling either Death or LOS.

**Usage**

data(LOS_model)

**Format**

Data frame with five columns

- **ID** A fictional patient ID number
- **Organisation** A factor representing one of ten fictional hospital trusts, e.g. Trust1
- **Age** Age in years of each fictional patient
LOS  In-hospital length of stay in days. The difference between admission and discharge date in dates

Death  Binary for death status: 0 = survived, 1 = died in hospital

Source

Generated by Chris Mainey <chris.mainey@uhb.nhs.uk>, Feb-2019

Examples

data(LOS_model)

model1 <- glm(Death ~ Age + LOS, data=LOS_model, family="binomial")
summary(model1)

# Now with an Age, LOS, and Age*LOS interaction.
model2 <- glm(Death ~ Age * LOS, data=LOS_model, family="binomial")
summary(model2)
Index

*Topic a&e
   ae_attendances, 2
*Topic datasets
   ae_attendances, 2
   LOS_model, 3
*Topic hospital
   ae_attendances, 2
   LOS_model, 3
*Topic regression
   LOS_model, 3

ae_attendances, 2

LOS_model, 3