# Package: amazonasdatahub (via r-universe)

November 10, 2026

**Title** Integrating Data, Research, and Education

**Version** 1.2.1

**Description** Various datasets from the state of Amazonas (AM), Brazil, to apply Deep Learning and statistical methods.

**License** MIT + file LICENSE

**Encoding** UTF-8

**RoxygenNote** 7.3.3

**LazyData** TRUE

**LazyDataCompression** xz

**Depends** R (>= 3.5)

**Imports** dplyr, ggplot2

**Suggests** rmarkdown, forcats

**URL** https://onelsoncarvalho.github.io/amazonasdatahub/

**NeedsCompilation** no

**Author** Nelson Carvalho [aut, cre], Leonardo Nascimento [aut]

**Maintainer** Nelson Carvalho <nelsoncarvalho.negocios@gmail.com>

**Repository** https://cran.r-universe.dev

**Date/Publication** 2026-11-10 17:58:44 UTC

**RemoteUrl** https://github.com/cran/amazonasdatahub

**RemoteRef** HEAD

**RemoteSha** 4edc3cd911416dc871ff390b3d65d4a25956d4a5

## Contents

agriculture_amazonas . . . . . 2
aids_amazonas . . . . . 3
gdp_amazonas . . . . . 4
humidity_manaus . . . . . 6
malaria_amazonas . . . . . 7

prf_amazonas . . . . . 9
rionegro_amazonas . . . . . 10
srl_muni . . . . . 11

**Index** **13**

---

`agriculture_amazonas` *agriculture_amazonas - Plant production data from the Institute of Sustainable Agricultural and Forestry Development of Amazonas - IDAM (2023).*

---

## Description

This database, which is the consolidation of data from IDAM's activity reports, provides crop production data for crops that this institute considers a priority.

## Usage

```
agriculture_amazonas
```

## Format

**'agriculture_amazonas' A data frame with 1072 rows and 13 columns::**

**municipality** Municipality

**local_unit** Local Unit of the Institute

**n_beneficiary** Number of assisted beneficiaries

**planted** Planted area in hectares (ha)

**harvested** Harvested area in hectares (ha)

**production** Production (in the unit of measurement - measure)

**estimated_n_beneficiary** Estimate of the number of beneficiaries

**estimated_planted** Estimate of the planted area (ha)

**estimated_harvested** Estimate of the harvested area (ha)

**estimated_production** Estimate of production (in the unit of measurement - measure)

**measure** Unit of Measurement of production

**cultivation** Cultivation

**year** Year of cultivation

## Source

INSTITUTO DE DESENVOLVIMENTO AGROPECUÁRIO E FLORESTAL SUSTENTÁVEL DO ESTADO DO AMAZONAS (IDAM). Relatório de Atividades (RAT 23-2). Manaus: IDAM, 2023. Disponível em: https://www.idam.am.gov.br/biblioteca/relatorio-de-atividades-rat-23-2/.

## Examples

```
# Cassava production plot (Gráfico de Produção de Mandioca)
mandioca_prod <- agriculture_amazonas[agriculture_amazonas$cultivation == "Mandioca", ]
plot(
  mandioca_prod$planted,
  mandioca_prod$harvested,
  xlab = "Planted area (hectare)",
  ylab = "Harvested area (hectare)",
  main = "Cassava production in Amazonas",
  sub = "Planted area x Harvested Area"
)
```

---

# aids_amazonas *aids_amazonas - A dataset of AIDS occurrences from 2011 to 2023 in Amazonas*

---

## Description

This dataset contains data of the AIDS occurrences in each municipality from Amazonas.

## Usage

```
aids_amazonas
```

## Format

**'aids_amazonas' A data frame with 1612 rows and 5 columns::**

**code_muni** Municipality code

**name_muni** Municipality name

**cases** Case count

**gender** Gender of the infected individual - M for Male and F for female

**year** Year of observation

## Source

BRASIL. Ministério da Saúde. Departamento de HIV, Aids, Tuberculose, Hepatites Virais e Infecções Sexualmente Transmissíveis. Indicadores HIV, Aids. Available in: https://indicadores.aids.gov.br/.

## Examples

```
# Aids - case count in Manaus
# Loading dplyr and ggplot to structure the data
require(dplyr)
require(ggplot2)

# Filtering by municipality and plotting case count by gender
```

```
aids_amazonas %>%
  filter(name_muni == "Manaus") %>%
  group_by(gender) %>%
  ggplot(aes(x = year, y = cases, group = gender, color = gender)) +
  geom_line() +
  scale_color_manual(values = c("blue", "red")) +
  theme_minimal() +
  labs(
    title = "AIDS ocurrences in Manaus (2011-2023)",
    x = "Year",
    y = "Case count",
    color = "Gender"
  )
```

---

gdp_amazonas *gdp_amazonas - Gross Domestic Product - GDP of Amazonas*

---

## Description

This database contains information about the Gross Domestic Product (GDP) of the state of Amazonas over the years (from 2010 to 2021). It includes values in R$ 1,000,000.

## Usage

```
gdp_amazonas
```

## Format

**'gdp_amazonas' A data frame with 48 rows and 21 columns::**

**year** Indicates the year in which the GDP info was collected

**quarter** Specifies the quarter of the year

**agriculture** Refers to the economic output from agricultural activities, including cultivation, harvesting, and logistical support after harvest

**livestock** Represents the economic contribution from livestock farming, covering meat, dairy production, and related support services

**forest_fishing_prod** Outputs from forestry (timber and non-timber products), fishing activities and aquaculture (fish farming)

**extractive_ind** Extractive industries of natural resources as minearls and oil

**manufacturing_ind** Refers to the transformation of raw materials into finished goods industries

**basic_infrastructure** Encompasses services related to electricity and gas generation, water supply, sewage treatment, and waste management.

**construction** Represents economic activities related to building construction.

**vehicle_retail_repair** Includes all commercial activities related to selling goods and services as well as vehicle repair services.

**transport_storage_post** Covers services related to transporting people and goods, storage facilities, and postal services.

**accomodation_food** Refers to economic activities in tourism-related sectors such as hotels and restaurants.

**information_communication** Encompasses sectors dealing with information technology and telecommunications.

**financial_activities** Includes banks, financial institutions, insurance companies, and various financial services.

**real_state_activities** Refers to transactions involving buying, selling, or renting properties.

**professional_activities** Covers a wide range of specialized professional services including consulting and administrative support.

**public_admin_defense_health_edu** Represents public sector spending on government administration, national defense, public education, health care services, and social security programs.

**private_edu_health** Refers to educational institutions and health care providers operating in the private sector.

**art_culture_sport** Encompasses activities related to entertainment, cultural events, sports facilities, and other service-oriented sectors.

**domestic_services** Includes personal services provided within households such as cleaning or caregiving.

**taxes** Represents taxes levied on goods and services included in GDP calculations.

## Source

Scientific Journal of Applied Social and Clinical Science - TIME SERIES ANALYSIS FOR THE QUARTERLY GROSS DOMESTIC PRODUCT OF AMAZONAS

## Examples

```
# Time series - Gross Domestic Product of Amazonas
# Loading dplyr and ggplot2
require(dplyr)
require(ggplot2)

# Selecting only year and taxes and ploting
gdp_amazonas %>%
  select(year, taxes) %>%
  ggplot(., aes(x = year, y = taxes)) +
  geom_line(linewidth = 1L, colour = "#cb181d") +
  geom_hline(
    yintercept = mean(gdp_amazonas$taxes),
    linetype = "dashed",
    size = 1
  ) +
  theme_light() +
  theme(
    plot.title = element_text(face = "bold", size = 16)
  ) +
  labs(
    x = "Year",
```

```
  y = "Taxes",
  title = "Quarterly GDP of Amazonas",
  subtitle = "2010 - 2021"
)
```

---

humidity_manaus *humidity_manaus - Time series with Extreme Values of Limited Range*

---

## Description

A time series dataset consisting of the minimum relative humidity observed in the city of Manaus from January 2009 to December 2020. This dataset is a sample of size n = 144 monthly blocks.

## Usage

```
humidity_manaus
```

## Format

**'humidity_manaus' A tibble with 144 rows and 5 columns:**

**year** The year in which this data was observed

**rh** Relative Humidity

**dbt** Dry Bulb Temperature

**wbt** Wet Bulb Temperature

**month** The month in which this data was observed

## Source

NASCIMENTO, Leonardo Brandão Freitas; LIMA, Max Sousa; DUCZMAL, Luiz H. P-min-stable regression models for time series with extreme values of limited range. Environmetrics, Issue 2, v. 36, 2025.

## Examples

```
# Relative humidity time series of Amazonas
# Loading dplyr and ggplot to structure the data
require(dplyr)
require(ggplot2)

# Creating date column and plotting the time series
humidity_manaus %>%
  mutate(date = as.Date(paste0(year, "-", month, "-","01"))) %>%
  ggplot(aes(x = date, y = rh)) +
  geom_line() +
  theme_minimal() +
  labs(
```

```
  title = "Relative Humidity of Amazonas (2010 - 2020)",
  x = "Date",
  y = "Relative Humidity"
)
```

---

`malaria_amazonas` *malaria_amazonas - An integrated dataset of malaria notifications in the state of Amazonas*

---

## Description

This dataset contains medical records of patients who underwent malaria screening tests in the Legal Amazon. The data in this dataset were filtered to include only the cases reported in the state of Amazonas and for patients who reside specifically in the state of Amazonas between 2014 and 2019.

## Usage

```
malaria_amazonas
```

## Format

**'malaria_amazonas' A data frame with 551633 rows and 35 columns::**

**notification_hr** Health region of notification
**notification_county** County of notification
**infection_state** Probable state where the patient was infected
**infection_hr** Probable health region where the patient was infected
**infection_county** Probable county where the patient was infected
**home_hr** Health region of residence of the patient
**home_county** County of residence of the patient
**exam_interval** Time interval between notification and examination
**treatment_interval** Time interval between examination and beginning of treatment
**notification_interval** Time interval between symptom and notification
**notification_month** Month in which the notification was recorded
**notification_year** Year in which the notification was recorded
**symptom_month** Month in which the patient felt the first symptoms of malaria
**symptom_year** Year in which the patient felt the first symptoms of malaria
**exam_month** Month in which the examination was performed
**exam_year** Year in which the examination was performed
**treatment_month** Month in which the treatment started
**treatment_year** Year in which the treatment started
**migration** Health region of residence different than that of notification
**exam_result** Result of examination

**detection_type** Type of detection

**exam_type** Type of examination

**symptom** Indicates if the patient felt a symptom

**hemiparasite** The result of the examination for other hemiparasites

**previous_treatment** Previous treatment for P. vivax or P.falciparum

**occupation** Main activity in the last 15 days

**education_level** Level of education of the patient

**age** Interval of the age of the patient

**race** Race/color of the patient

**gender** Gender of the patient

**pregnancy** Indicates pregnancy and the gestational age

**crosses** Indicates the amount of parasitemia in crosses

**cvl_case** Indicates the existence of cases of canine visceral leishmaniasis

**scheme** Indicates the treatment scheme employed

**qty_parasites** Indicates the number of parasites per mm^3

## Source

Lais Baroni, M. P. (2020). An Integrated Dataset of Malaria Notifications in the Legal Amazon (Dataset). Synapse. https://doi.org/10.7303/SYN21552203

## Examples

```
# Bar plot of malaria tests results (2019)
library(dplyr)
library(ggplot2)

malaria_amazonas %>%
  filter(exam_year == 2019) %>%
  count(exam_result) %>%
  ggplot(aes(x = reorder(exam_result, -n), y = n)) +
  geom_bar(stat = "identity") +
  geom_text(
    aes(label = n),
    vjust = -0.5
  ) +
  theme_minimal() +
  labs(
    title = "Results of malaria tests conducted in 2019",
    subtitle = "Documented cases in Amazonas and of patients residing in the state",
    x = "Result",
    y = "Count"
  )
```

`prf_amazonas` *prf_amazonas - Traffic accident data from Amazonas recorded by Federal Highway Police (2012-2025)*

## Description

Traffic accident data from Amazonas (2012-2025)

## Usage

```
prf_amazonas
```

## Format

**'prf_amazonas' A data frame with 116 rows and 30 columns::**

**id** Record ID
**date** Date of the record - yyyy-mm-dd
**days_of_week** Day of the week
**time** Hour of the record
**federative_unit** Federative unit (UF)
**br** Federal Highway
**km** Kilometer of BR
**municipality** Municipality
**accident_cause** Accident cause
**accident_type** Accident type
**accident_classification** Accident classification: victims, no victims, unharmed victims, fatal victims
**road_direction** Road direction
**weather** Weather conditions at the time of the accident
**track_type** Track type: Simples, Múltipla ou Dupla
**track_layout** Track layout
**land_use** Land use
**people** People involved in the accident
**deceased** Death toll
**minor_injuries** Number of victims with minor injuries
**serious_injuries** Number of victims with serious injuries
**unharmed** Number of unharmed victims
**ignored** Ignored
**harmed** Number of harmed victims
**latitude** Latitude
**longitude** Longitude
**regional** Superintendency of the Federal Highway Police of Brazil
**police_station** Police station where the report was filed
**operating_unit** Operating unit and police station

## Source

Brasil. Polícia Rodoviária Federal (PRF). Dados Abertos da PRF: Acidentes de Trânsito. https://www.gov.br/prf/pt-br/acesso-a-informacao/dados-abertos/dados-abertos-da-prf.

## Examples

```
# Accident count by weather condition
# Loading dplyr and ggplot
require(dplyr)
require(ggplot2)
require(forcats)

# Counting accidents by weather condition and ploting
prf_amazonas %>%
  count(weather) %>%
  mutate(
    weather = fct_reorder(weather, -n)
  ) %>%
  ggplot(aes(x = weather, y = n)) +
  geom_bar(stat = "identity", fill = "steelblue") +
  geom_text(aes(label = n), vjust = -0.5) +
  theme_minimal() +
  labs(
    title = "Accident count by weather condition",
    x = "Weather",
    y = "Count"
  )
```

---

# rionegro_amazonas *rionegro_amazonas - A dataset of the Rio Negro River (Amazonas) level from 2003 to 2023*

---

## Description

This dataset contains data of the Rio Negro River level in the state of Amazonas (Brazil), compiling information on years, semesters and days, along with the river's level in meters and whether it filled or dropped (in centimeters).

## Usage

```
rionegro_amazonas
```

## Format

**'rionegro_amazonas' A data frame with 8003 rows and 7 columns::**

**level_m** River level in meters

**increase_decrease_cm** Wheter the river level filled or dropped (in centimeters)

**day** Measurement day

**month** Measurement month

**year** Measurement year

**semester** Measurement semester

## Source

Porto de Manaus. Nível do Rio Negro. https://portodemanaus.com.br/nivel-do-rio-negro/.

## Examples

```
# Rio Negro - Boxplot (water level by year)

# Loading ggplot
require(ggplot2)
require(dplyr)

rionegro_amazonas %>%
  ggplot(aes(x = year, y = level_m, group = year)) +
  stat_boxplot(geom = "errorbar") +
  geom_boxplot() +
  theme_minimal() +
  labs(
    x = "Year",
    y = "Water level (m)"
  )
```

---

`srl_muni` *srl_muni - School Read Levels - Physical literacy and reading performance of amazonian schoolchildren*

---

## Description

This is a dataset from an association, cross-sectional, quantitative study carried out with Amazonian schoolchildren (n=100; aged between 10 and 13 years) from 3 public schools in the municipality of Borba (Amazonas).

## Usage

```
srl_muni
```

## Format

**'srl_muni' A data frame with 100 rows and 17 columns::**

**school** School of observation

**sex** Sex - 1 for Masculine and 2 for Feminine

**age** Age in years

**weight_kg** Weight in kilograms (kg)

**height_m** Height in meters (m)

**bmi** Body Mass Index (BMI)

**bmi_status** Body Mass Index Status - A categorical description of the BMI

**phv** Peak Height Velocity (PHV) is a continuous variable, measured by age (in years) relative to the peak height in the period where maximum rate of growth occurs and this is an indication of maturity

**phv_state** Categorical description for Peak Height Velocity (PHV)

**eses** Exercise Self-Efficacy Scale measures how confident the students are with regard to carrying out regular physical activities and exercise, measured by a psychometric scale

**pmc** Perceived Motor Competence (PMC) assesses the child's self-perception regarding their ability to perform motor movements. Measured by a psychometric scale

**kupa** Knowledge and Understanding of Physical Activity (KUPA) - Evaluates knowledge regarding the benefits, safety, and engagement in physical activity. Measured by scores obtained in objective questions

**sb** Sedentary Behavior - Evalutes the exposition time to sedentary behaviors. Measured by psychometric scale

**pa** Pyshical Activity (PA) - Level of physical activity ian children, based on the frequency and duration of routine daily activities. Measured by score on the physical activity frequency questions

**mc** Motor Competence - Supine To Stand (in seconds) - Measures the time (in seconds) it takes for the child to stand up from the floor, from a supine position to a standing position

**tde** Total Score on the TDE test - Measures reading level based on time and accuracy in word reading

**tde_categorical** Categorical description of the TDE Score

## Source

ALMEIDA, Thiago da Cruz de. Physical Literacy e desempenho em leitura de escolares amazônicos: um estudo de associação. 2024. 104 f. Dissertação (Mestrado em Educação) - Universidade Federal do Amazonas, Manaus, 2024. Disponível em: https://tede.ufam.edu.br/handle/tede/10627.

# Index

∗ **datasets**
agriculture_amazonas, 2
aids_amazonas, 3
gdp_amazonas, 4
humidity_manaus, 6
malaria_amazonas, 7
prf_amazonas, 9
rionegro_amazonas, 10
srl_muni, 11

agriculture_amazonas, 2
aids_amazonas, 3

gdp_amazonas, 4

humidity_manaus, 6

malaria_amazonas, 7

prf_amazonas, 9

rionegro_amazonas, 10

srl_muni, 11