Package: atsalibrary (via r-universe)

August 10, 2024

Title Packages, data and scripts for ATSA course and lab book

Version 1.5

Maintainer Elizabeth E. Holmes <eli.holmes@noaa.gov>

Description This package will load the needed packages and data files for the ATSA course material when students install from GitHub.

Depends R (>= 3.5.0), broom (>= 0.5.1), corrplot, datasets (>= 3.5.3), dplyr (>= 0.7.5), forecast (>= 8.4), fpp2 (>= 2.4), ggmap (>= 3.0.0), ggplot2 (>= 3.0.0), grid (>= 3.4.1), gridExtra (>= 2.3), Hmisc, kableExtra (>= 0.9.0), knitr (>= 1.20), loo (>= 2.1.0), magrittr (>= 1.5), MARSS (>= 3.11.2), quantmod (>= 0.3-13), RCurl (>= 1.95-4.12), readr, reshape2 (>= 1.4.3), rmarkdown (>= 1.11), stats (>= 3.5.3), stringr (>= 1.4.0), tidyr (>= 1.1.2), tseries (>= 0.10-46), urca (>= 1.3-0), zoo

Suggests atsar, coda (>= 0.19-2), devtools (>= 2.0.1), learnr, R2jags (>= 0.5-7), rjags (>= 4-8), rstan (>= 2.18.2)

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.2.3

Roxygen list(markdown = TRUE)

Remotes atsa-es/atsar

URL https://atsa-es.github.io/atsalibrary/

Repository https://atsa-es.r-universe.dev

RemoteUrl https://github.com/atsa-es/atsalibrary

RemoteRef HEAD

RemoteSha 950df5a4c779998c9dab3673806de25be2282fa0

Contents

`atsalibrary-package` *atsalibrary*

Description

This package will load the needed packages and data files for the ATSA course material when students install from GitHub.

Data

- chinook
- ECNNO2
- greeklandings
- lakeWA
- snotel
- sockeye
- neon_barc
- MLCO2
- NHTemp

chinook *Monthly and annual Chinook salmon commercial landings*

Description

Monthly (WA, OR and CA) and yearly (AK, CA, MI, OR, PA, and WA) data on commercial landings and value of Chinook salmon. Yearly data were downloaded from the NOAA Fisheries, Fisheries Statistics Division and the monthly data were downloaded from PacFIN.

Usage

```
data(chinook)
```

Format

Objects of class `"data.frame"`. Columns are Year, Month (if monthly), Species, State, log.metric.tons, metric.tons, and value.usd (USD)

Details

There are two datasets included: `chinook.month` and `chinook.year`. The monthly data are available from 1990 and the annual data are available from 1950. The raw data and R script to process the data are in the `inst/original_data` folder.

From the NOAA Fisheries Statistics Divison: "Collecting these data is a joint state and federal responsibility. State-federal systems gather landings data from state-mandated fishery trip-tickets, landing weighout reports from seafood dealers, federal logbooks of fishery catch and effort, and shipboard and portside interviews and biological sampling of catches. State fishery agencies are usually the main collectors of these data, though they and NOAA Fisheries gather data jointly in some states. Surveys are done differently in different states; NOAA Fisheries takes supplemental surveys to ensure that the data from different states and years are comparable."

In addition from NOAA Fisheries Statistics Division: "Statistics for each state represent a census of the volume and value of finfish and shellfish landed and sold at the dock, not an expanded estimate of landings based on sampling data. The main statistics collected are the pounds and ex-vessel dollar value of landings identified by species, year, month, state, county, port, water, and fishing gear. Most states get their landings data from seafood dealers who submit monthly reports of the weight and value of landings by vessel. Increasingly, though, states are getting landings data from mandatory trip-tickets—filled out by seafood dealers and fishermen at the end of every fishing trip, indicating their landings by species."

Source

Yearly NOAA Commercial Landings Statistics

Monthly PacFIN

References

NOAA Fisheries Office of Science and Technology, Commercial Landings Query, Available at: www.fisheries.noaa.gov/foss, Accessed 14 April 2023

Pacific Fisheries Information Network (PacFIN) retrieval dated 14 April 2023, Pacific States Marine Fisheries Commission, Portland, Oregon (www.psmfc.org). Available at https://reports.psmfc.org/pacfin

Examples

```
data(chinook)
dat <- subset(chinook.month, state="WA")$log.metric.tons
datts <- ts(dat, start=c(1990,1), frequency=12)
plot(datts)

ggplot(
  chinook.month %>%
    mutate(t = zoo::as.yearmon(paste(Year, Month), "%Y %b")),
 aes(x=t, y=log.metric.tons)) +
  geom_line() +
  facet_wrap(~State)
```

ECNNO2 *UK Atmospheric NO2*

Description

UK Environmental Change Network (ECN) atmospheric nitrogen chemistry data: 1993-2015

Usage

```
data(ECNNO2)

ECNNO2

ECNmeta
```

Format

ECNNO2 is the data and is an object of class `"data.frame"`. Columns are Year, Month, TubeID, SiteCode, and Value. ECNmeta is the site locations.

An object of class `data.frame` with 9108 rows and 7 columns.

An object of class `data.frame` with 12 rows and 4 columns.

Details

From https://catalogue.ceh.ac.uk/documents/baf51776-c2d0-4e57-9cd3-30cd6336d9cf: "Atmospheric Nitrogen Dioxide data from the UK Environmental Change Network (ECN) terrestrial sites. These data are collected by diffusion tubes at all of ECN's terrestrial sites using a standard protocol. They represent continuous fortnightly records from 1993 to 2015." The data in ECNNO2 are the NO2 (micrograms/m3). Note although the dataset information do not state this, I believe that the data are averaged iver the time exposure (minutes), i.e. are micrograms/m3-minute, because the exposure time does not affect the N02 values in the dataset, i.e. longer exposure does not equal higher NO2 values. The link above has all the background information on the dataset. The original data are taken every 7-31 days, with 14 days apart being the target and most common time difference. ECNNO2 is the monthly average of of all the readings for the experimental tubes. The average for the E1 tube will be average of the 1-3 E1 tubes from that month. Same for E2, E3 etc.

These data are part of a long-term monitoring program in the UK. N02 is just one component that is monitored. The sites locations are in ECNmeta.

Source

UK Centre for Ecolgoy & Hydrology

References

Rennie, S.; Adamson, J.; Anderson, R.; Andrews, C.; Bater, J.; Bayfield, N.; Beaton, K.; Beaumont, D.; Benham, S.; Bowmaker, V.; Britt, C.; Brooker, R.; Brooks, D.; Brunt, J.; Common, G.; Cooper, R.; Corbett, S.; Critchley, N.; Dennis, P.; Dick, J.; Dodd, B.; Dodd, N.; Donovan, N.; Easter, J.; Flexen, M.; Gardiner, A.; Hamilton, D.; Hargreaves, P.; Hatton-Ellis, M.; Howe, M.; Kahl, J.; Lane, M.; Langan, S.; Lloyd, D.; McCarney, B.; McElarney, Y.; McKenna, C.; McMillan, S.; Milne, F.; Milne, L.; Morecroft, M.; Murphy, M.; Nelson, A.; Nicholson, H.; Pallett, D.; Parry, D.; Pearce, I.; Pozsgai, G.; Rose, R.; Schafer, S.; Scott, T.; Sherrin, L.; Shortall, C.; Smith, R.; Smith, P.; Tait, R.; Taylor, C.; Taylor, M.; Thurlow, M.; Turner, A.; Tyson, K.; Watson, H.; Whittaker, M.; Wood, C. (2017). UK Environmental Change Network (ECN) atmospheric nitrogen chemistry data: 1993-2015. NERC Environmental Information Data Centre. https://doi.org/10.5285/baf51776-c2d0-4e57-9cd3-30cd6336d9cf

Examples

```
# Data were prepared from the downloaded data set as follows:
## Not run:
library(tidyr)
library(reshape2)
dat <- read.csv("ECN_AN1.csv", stringsAsFactors = FALSE)
dat$Date <- as.Date(dat$SDATE, "%d-%B-%y")
dat$Year <- format(dat$Date, "%Y")
dat$Mon <- format(dat$Date, "%b")
a <- subset(dat, TUBEID %in% c("E1","E2","E3"))
b <- spread(a, FIELDNAME, VALUE)
b <- b[order(b$Date), ]
bad <- c(103, 104, 107:120) # bad quality codes. Burning nearby or contaminant in tube.
b$NO2[b$Q1 %in% bad | b$Q2 %in% bad | b$Q3 %in% bad] <- NA
val <- tapply(b$NO2, list(b$Year, b$Mon, b$TUBEID, b$SITECODE), mean, na.rm=TRUE)
```

```
dat.mon <- melt(data=val, value.name="NO2")
colnames(dat.mon) <- c("Year", "Month", "TubeID", "SiteCode", "Value")
dat.mon$Date <- as.Date(paste(1,dat.mon$Mon, dat.mon$Year), "%d %b %Y")
dat.mon <- dat.mon[order(dat.mon$Date),]
rownames(dat.mon) <- NULL
ECNNO2 <- dat.mon

## End(Not run)

data(ECNNO2)
```

---

greeklandings *Annual Anchovy, Sardine and Mackerel landings in Hellenic Waters 1964-2007*

Description

Annual commercial landings of anchovy, sardine and mackerel from Greek fisheries compiled by the Hellenic Statistical Authority. Also included are covariates for the catch.

Usage

```
data(greeklandings)

greeklandings

covs

covsmean.year

covsmean.mon
```

Format

Objects of class "data.frame". Columns are Year, Species, log.metric.tons, metric.tons

An object of class `data.frame` with 264 rows and 4 columns.

An object of class `list` of length 3.

An object of class `data.frame` with 55 rows and 6 columns.

An object of class `data.frame` with 656 rows and 7 columns.

Details

Data are from Table IV in the "Sea Fishery by Motor Vessels" statistical reports published by the Hellenic Statisitcal Authority. The reports are available in Digital Library (ELSTAT), Special Publications, Agriculture-Livestock-Fisheries, Fisheries. In Table IV, the landings data were taken from the total column, units are metric tons. In the table, sardine is denoted ' Pilchard'. The data were assembled manually for the Fish Forecast eBook.

The covariates are Year, air.degC (air temperature), slp.millibars, sst.degC (sea surface temperature), vwnd.m/s (N/S wind speed), wspd3.m3/s3 (overall wind speed cubed). `covsmean.mon` and `covsmean.year` are averages over the 3 regions in the `covs` list. See the the Fish Forecast eBook for details.

Source

Hellenic Statisitcal Authority Digital Library

Examples

```
data(greeklandings)
anchovy = ts(subset(greeklandings, Species=="Anchovy")$log.metric.tons, start=1964)
plot(anchovy)

library(ggplot2)
ggplot(greeklandings, aes(x=Year, y=log.metric.tons)) +
      geom_line() + facet_wrap(~Species)
```

`hourlyphyto` *Hourly phytoplankton*

Description

no information about these data

Usage

```
data(hourlyphyto)
```

Format

Objects of class `"data.frame"` with column V1

Examples

```
data(hourlyphyto)
```

`KvichakSockeye` *Kvichak River SR Data*

Description

Spawner-recruit data for sockeye salmon from the Kvichak River.

Usage

```
data(KvichakSockeye)
```

Format

Object of class `"data.frame"`

- **brood_year** Brood year.
- **spawners** Spawners (count) in thousands
- **recruits** Recruits (count) in thousands
- **pdo_summer_t2** Pacific Decadal Oscillation (PDO) from Apr-Sep of brood year $t+2$
- **pdo_winter_ts** Pacific Decadal Oscillation (PDO) from Oct (brood year $t+2$) to Mar (brood year $t+3$)

Details

Spawner-recruit data for sockeye salmon (*Oncorhynchus nerka*) from the Kvichak River in SW Alaska that span the years 1952-1989. The data come from a large public database begun by Ransom Myers many years ago. The database is now maintained as the RAM Legacy Stock Assessment Database.

Source

RAM Legacy Stock Assessment Database

References

RAM Legacy Stock Assessment Database. 2018. Version 4.44-assessment-only. Released 2018-12-22. Retrieved from DOI:10.5281/zenodo.2542919.

Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. (2012) Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries 13 (4) 380-398. DOI: 10.1111/j.1467-2979.2011.00435.x

Spawner escapement estimates are originally from: 1997-2017 Appendix Table A12.

See Also

sockeye

Examples

```
data(KvichakSockeye)
```

---

lakeWA *Lake Washington Plankton Data*

Description

Monthly Lake Washington (WA, USA) plankton, temperature, total phosphorous, and pH data 1962 to 1994.

Usage

```
data(lakeWA)
```

Format

Object of class "data.frame". #'

Details

The lakeWA is a 32-year time series (1962-1994) of monthly plankton counts from Lake Washington, Washington, USA. lakeWA is a transformed version of the raw data (available in the MARSS package data(lakeWAplanktonRaw, package="MARSS")). Zeros have been replaced with NAs (missing). The plankton counts are logged (natural log) and standardized to a mean of zero and variance of 1 (so logged and then z-scored). Temperature, TP & pH were also z-scored but not logged (so z-score of the untransformed values for these covariates). The single missing temperature value was replaced with -1 and the single missing TP value was replaced with -0.3. The two missing pH values were interpolated. Monthly anomalies for temperature, TP and pH were computed by removing the monthly means (computed over the 1962-1994 period). The anomalies were then z-scored to remove mean and standardize variance to 1.

References

Adapted from the Lake Washington database of Dr. W. T. Edmondson, as funded by the Andrew Mellon Foundation; data courtesy of Dr. Daniel Schindler, University of Washington, Seattle, WA.

Hampton, S. E. Scheuerell, M. D. Schindler, D. E. (2006) Coalescence in the Lake Washington story: Interaction strengths in a planktonic food web. Limnology and Oceanography, 51, 2042-2051.

Examples

```
# The lakeWA data frame was created with the following code:
## Not run:
data(lakeWAplankton, package = "MARSS")
lakeWA <- data.frame(lakeWAplanktonTrans)
# add on month and date columns
```

```
lakeWA$Month.abb <- month.abb[lakeWA$Month]
lakeWA$Date <- as.Date(paste0(lakeWA$Year, "-", lakeWA$Month,"-01"))
# interpolate 2 missing values in pH
lakeWA$pH[is.na(lakeWA$pH)] <- MARSS::MARSS(lakeWA$pH)$states[1, is.na(lakeWA$pH)]
# create monthly anomalies
lakeWA$Temp.anom <- residuals(lm(Temp ~ Month.abb, data=lakeWA))
lakeWA$TP.anom <- residuals(lm(TP ~ Month.abb, data=lakeWA))
lakeWA$pH.anom <- residuals(lm(pH ~ Month.abb, data=lakeWA))
# resort the columns
lakeWA <- lakeWA[,c(22,1:2,21,3:5,23:25,6:20)]
# zscore everything
for (i in 5:25) lakeWA[[i]] <- MARSS::zscore(lakeWA[[i]])
save( lakeWA, file="data/lakeWA.RData" )

## End(Not run)

library(ggplot2)
library(tidyr)
df <- lakeWA %>%
     pivot_longer(
         cols = Temp:Non.colonial.rotifers,
         names_to = "variable",
         values_to = "value"
         )
ggplot(df, aes(x=Date, y=value)) +
     geom_line() +
     facet_wrap(~variable)

data(lakeWA)
```

MLCO2 *Mauna Loa C02 measurements*

Description

Monthly Snow Water Equivalent (SWE) at Stations in Washington State

Usage

```
data(MLCO2)
```

Format

Objects of class "`data.frame`". Columns are year, month, ppm (part per million)

Details

These are monthly average C02 values recorded on Mauna Loa in Hawaii since 1958. Data are provided by NOAA Earth System Rearch Laboratory, Global Monitoring Division.

Data from March 1958 through April 1974 have been obtained by C. David Keeling of the Scripps Institution of Oceanography (SIO) and were obtained from the Scripps website (scrippsco2.ucsd.edu). The ppm column contains the monthly mean CO2 mole fraction determined from daily averages. The mole fraction of CO2, expressed as parts per million (ppm) is the number of molecules of CO2 in every one million molecules of dried air (water vapor removed). If there are missing days, the value is the interpolated value. See www.esrl.noaa.gov/gmd/ccgg/trends/ for additional details.

Source

NOAA ESRL Trends in Atmospheic Carbon Dioxide

References

Dr. Pieter Tans, NOAA/ESRL (www.esrl.noaa.gov/gmd/ccgg/trends/) and Dr. Ralph Keeling, Scripps Institution of Oceanography (scrippsco2.ucsd.edu/).

Examples

```
# We downloaded this data and created the `MLCO2` dataframe with the following code:
## Not run:
library(RCurl)
## get CO2 data from Mauna Loa observatory
ww1 <- "ftp://aftp.cmdl.noaa.gov/products/"
ww2 <- "trends/co2/co2_mm_mlo.txt"
CO2fulltext <- getURL(paste0(ww1,ww2))
MLCO2 <- read.table(text=CO2fulltext)[,c(1,2,5)]
## assign better column names
colnames(MLCO2) <- c("year","month","ppm")
save(MLCO2, file="MLCO2.RData")

## End(Not run)

data(MLCO2)
```

---

neon_barc *Lake Barco Aquatic Data*

Description

Lake Barco data for the EFI 2021 Forecasting Challenge

Usage

```
data(neon_barc)
```

Format

neon_barc is a tibble with columns are date, siteID, oxygen, temperature, oxygen_sd, temperature_sd, depth_oxygen, depth_temperature, neon_product_ids, year, and cal_day.

Details

The data for the aquatics challenge comes from a NEON site at Lake Barco (Florida). More about the data and challenge is here and the Github repository for getting all the necessary data is eco4cast.

The neon_barc data set was created with the aquatics-targets.csv.gz file produced in the neon4cast-aquatics GitHub repository and saved in the inst folder in the atsalibrary package. From that file, neon_barc is created with

```
library(tidyverse)
library(lubridate)
# This taken from code on NEON aquatics challenge
targets <- readr::read_csv("aquatics-targets.csv.gz", guess_max = 10000)
site_data_var <- targets %>%
  filter(siteID == "BARC")
# This is key here - I added the forecast horizon on the end of the data for the forecast period
full_time <- tibble(time = seq(min(site_data_var$time), max(site_data_var$time), by = "1 day"))
# Join the full time with the site_data_var so there aren't gaps in the time column
site_data_var <- left_join(full_time, site_data_var) %>%
  dplyr::rename(date = time)
site_data_var$year <- year(site_data_var$date)
site_data_var$cal_day <- yday(site_data_var$date)
neon_barc <- site_data_var
```

Source

https://github.com/eco4cast/neon4cast-aquatics

References

Ecological Forecasting Initiative https://ecoforecast.org/

Neon4Cast Aquatics GitHub repository https://github.com/eco4cast/neon4cast-aquatics

Examples

```
data(neon_barc)
```

NHTemp *Northern Hemisphere land and ocean temperature anomalies*

Description

Northern Hemisphere land and ocean temperature anomalies from NOAA

Usage

```
data(NHTemp)
```

Format

Object of class "data.frame". Columns are Year and Value. #'

Details

From https://www.ncdc.noaa.gov/cag/global/data-info: "Global temperature anomaly data come from the Global Historical Climatology Network-Monthly (GHCN-M) data set and International Comprehensive Ocean-Atmosphere Data Set (ICOADS), which have data from 1880 to the present. These two datasets are blended into a single product to produce the combined global land and ocean temperature anomalies. The available timeseries of global-scale temperature anomalies are calculated with respect to the 20th century average."

Source

NOAA Climate Monitoring

Examples

```
# We downloaded this data and created the `NHTemp` dataframe with the following code:
## Not run:
library(RCurl)
ww1 <- "https://www.ncdc.noaa.gov/cag/time-series/"
ww2 <- "global/nhem/land_ocean/p12/12/1880-2014.csv"
NHTemp <- read.csv(text=getURL(paste0(ww1,ww2)), skip=4)
save(NHTemp, file="NHTemp.RData")

## End(Not run)

data(NHTemp)
```

---

snotel *Washington State SNOTEL measurements*

Description

Monthly Snow Water Equivalent (SWE) at Stations in Washington State

Usage

```
data(snotel)

snotel

snotelmeta
```

Format

- snotel: object of class "`data.frame`". Columns are Station, Station.Id, Year, Month, SWE, and Date
- snotelmeta: Station_Name, Station.Id, State.Code, Network.Code, Network.Name, Elevation, Latitude, Longitude"

An object of class `data.frame` with 27720 rows and 6 columns.

An object of class `data.frame` with 70 rows and 8 columns.

Details

These data are the snow water equivalent percent of normal. This represents the snow water equivalent (SWE) compared to the average value for that site on the same day. The data were downloaded from the USDA Natural Resources Conservation Service website for the Washington Snow Survey Program. The `snotel` object is the SWE data and the `snotelmeta` is the station metadata (lat/lon and elevation).

Source

USDA WA Snow Survey Program

Examples

```
data(snotel)
```

`sockeye` *Bristol Bay Sockeye Spawner-Recruit Data*

Description

Spawner-recruit data for sockeye salmon from the KVICHAK, WOOD, IGUSHIK, NAKNEK, EGEGIK, NUSHAGAK, UGASHIK, TOGIAK Rivers.

Usage

```
data(sockeye)
```

Format

Object of class "`data.frame`"

- **brood_year** Brood year.
- **region** River.
- **spawners** Spawners (count) in thousands
- **recruits** Recruits (count) in thousands
- **pdo_summer_t2** Pacific Decadal Oscillation (PDO) from Apr-Sep of brood year $t+2$
- **pdo_winter_ts** Pacific Decadal Oscillation (PDO) from Oct (brood year $t+2$) to Mar (brood year $t+3$)

Details

Spawner-recruit data for sockeye salmon (*Oncorhynchus nerka*) from the Bristol Bay region in SW Alaska that span the years 1952-2005. The data come from a large public database begun by Ransom Myers many years ago. The database is now maintained as the RAM Legacy Stock Assessment Database.

The R script which created the data can be found in the atsalibrary GitHub site in the inst/orginal_data/sockeye folder.

For convenience in the lab book, KvichakSockeye is also created which is just the Kvichak data.

Source

RAM Legacy Stock Assessment Database v4.491

References

RAM Legacy Stock Assessment Database. 2018. Version 4.44-assessment-only. Released 2018-12-22. Retrieved from DOI:10.5281/zenodo.2542919.

Ricard, D., Minto, C., Jensen, O.P. and Baum, J.K. (2012) Evaluating the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. Fish and Fisheries 13 (4) 380-398. DOI: 10.1111/j.1467-2979.2011.00435.x

Spawner escapement estimates are originally from: Bristol Bay Area Annual Management Reports from the Alaska Department of Fish and Game and are based on in river spawner counts (tower counts and aerial surveys).

See an example application of these data see: Rogers, L.A. and Schindler, D.E. 2008. Asynchrony in Population Dynamics of Sockeye Salmon in Southwest Alaska. Oikos 117: 1578-1586

Examples

```
data(sockeye)

a <- pivot_longer(sockeye,
     cols=spawners:recruits,
     names_to="value",
     values_to="count")
ggplot(a,aes(brood_year, count, color=value, fill=value)) +
   geom_area() +
   theme(axis.text.x = element_text(angle = 90)) +
     ylab("Count (thousands)") +
     xlab("Brood Year") +
   facet_wrap(~region, scales="free_y")
```

Index