

# Package: fishstat (via r-universe)

February 18, 2025

**Version** 2024.1.0.1

**Date** 2025-02-14

**Title** Global Fishery and Aquaculture Statistics

**Depends** R (>= 3.5.0)

**Suggests** areaplot

**LazyData** yes

**Description** The Food and Agriculture Organization of the United Nations (FAO) FishStat database is the leading source of global fishery and aquaculture statistics and provides unique information for sector analysis and monitoring. This package provides the global production data from all fisheries and aquaculture in R format, ready for analysis.

**License** CC BY-NC-SA 4.0

**URL** <https://www.fao.org/fishery/en/fishstat>,  
<https://github.com/sofia-taf/fishstat>

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**Repository** <https://sofia-taf.r-universe.dev>

**RemoteUrl** <https://github.com/sofia-taf/fishstat>

**RemoteRef** HEAD

**RemoteSha** 942b1cabdd08ee150b205bc654ffc801695b86c7

## Contents

fishstat-package . . . . .	2
aquaculture . . . . .	4
area . . . . .	5
capture . . . . .	7
country . . . . .	8
environment . . . . .	9

measure . . . . .	10
production . . . . .	11
source . . . . .	13
species . . . . .	14
status . . . . .	15

<b>Index</b>	<b>17</b>
--------------	-----------

---

fishstat-package	<i>Global Fishery and Aquaculture Statistics</i>
------------------	--

---

## Description

The Food and Agriculture Organization of the United Nations (FAO) **FishStat** database is the leading source of global fishery and aquaculture statistics and provides unique information for sector analysis and monitoring.

This package provides the global production data from all fisheries and aquaculture in R format, ready for analysis.

## Details

*Production tables:*

<a href="#">aquaculture</a>	aquaculture production
<a href="#">capture</a>	capture production
<a href="#">production</a>	aquaculture and capture production

*Lookup tables:*

<a href="#">area</a>	fishing areas
<a href="#">country</a>	countries and territories
<a href="#">environment</a>	aquaculture environments
<a href="#">measure</a>	units of measurement
<a href="#">source</a>	sources of production
<a href="#">species</a>	taxonomic groups
<a href="#">status</a>	status of data entries

## Note

The data in the package were downloaded from the FAO **data server** and imported into R. The R package version indicates the version of FishStat data it includes. Column names have been simplified to facilitate quick exploration and plotting in R.

Production tables can be joined with lookup tables using the [merge](#) function, as demonstrated in the help page examples for production tables. The column names in the R package have been designed to allow automatic inference of which columns to join, and the resulting table will have unique column names.

To achieve a small memory footprint and fast computations, one can filter or aggregate rows and select columns of interest before joining tables, as demonstrated in the capture and production examples. On the other hand, for the sake of convenience, one can also construct a full table with all data records and all columns. The resulting table will have many rows and columns, but most computers will handle these without problems:

```
prod.all <- merge(merge(merge(merge(merge(merge(production,
  area), country), measure), source), species), status)
cap.all <- merge(merge(merge(merge(merge(capture,
  area), country), measure), species), status)
aqua.all <- merge(merge(merge(merge(merge(merge(aquaculture,
  area), country), environment), measure), species), status)
```

When aggregating data, simple utility functions can be handy for tasks like rounding the sum of thousands/millions of tonnes:

```
sum3 <- function(x, digits=0) round(sum(x)/1e3, digits=digits)
sum6 <- function(x, digits=0) round(sum(x)/1e6, digits=digits)
aggregate(value~environment, aquaculture, sum6)
aggregate(value~area, capture, sum6, subset=measure=="Q_tlw", digits=1)
```

An effort has been made to describe each table in the corresponding R help page. However, the official and authoritative documentation of the FishStat database is found on the FAO [FishStat](#) website.

The example below demonstrates how the FishStat data can be used to produce an overview of global fisheries and aquaculture. The combination of FishStat and the R environment can also be very efficient for analyses that focus on selected areas, countries, species, and/or taxonomic groups.

### Author(s)

Arni Magnusson and Rishi Sharma created this R package.

All credit for the FishStat database goes to the Statistics Team of the FAO Fisheries and Aquaculture Division, as well as national data submitters. The database [terms of use](#) are based on the [CC BY-NC-SA 3.0 IGO](#) license. The R package is released under a similar [CC BY-NC-SA 4.0](#) license.

*To cite the use of FishStat data:*

FAO. [Year]. FishStat data. Fisheries and Aquaculture Division. Rome. <https://www.fao.org/fishery/en/fishstat>.

*To cite the use of this R package to access the data, cite FishStat (above) as well as:*

Magnusson, A. and R. Sharma. [Year]. fishstat: Global Fishery and Aquaculture Statistics. R package version [Version]. <https://cran.r-project.org/package=fishstat>.

### See Also

Useful links:

- <https://www.fao.org/fishery/en/fishstat>
- <https://github.com/sofia-taf/fishstat>

**Examples**

```

head(production)

# Analyze production measured in tonnes
prod <- production[production$measure == "Q_tlw" & production$value > 0,]
prod <- merge(prod, species[c("species", "yearbook")])

# Select SOFIA species, excluding mammals, reptiles, and plants
prod <- prod[prod$yearbook == "Fish, crustaceans and molluscs, etc.",]

# Determine origin
prod <- merge(prod, area[c("area", "inlandmarine")])
prod$origin <- ifelse(prod$source == "CAPTURE", "Capture", "Aquaculture")
prod$w <- ifelse(prod$inlandmarine == "Marine areas", "marine", "inland")
prod$origin <- paste0(prod$origin, " (", prod$w, ")")

# World capture fisheries and aquaculture production
x <- xtabs(value~year+origin, aggregate(value~year+origin, prod, sum))
x <- x[,c(2,1,4,3)] / 1e6
library(areaplot)
areaplot(x, legend=TRUE, args.legend=list(x="topleft"), ylab="million tonnes")

```

---

aquaculture

*Global Aquaculture Production*


---

**Description**

Aquaculture production quantity by species, area, country, and aquatic environment for the years 1950-2022, compiled and published by FAO (2024).

**Usage**

```

aquaculture

```

**Format**

Data frame containing eight columns:

species	<a href="#">species</a> code
year	year
area	<a href="#">area</a> code
country	<a href="#">country</a> code
value	quantity in tonnes
measure	<a href="#">measure</a> code
status	<a href="#">status</a> code
environment	<a href="#">environment</a> code

## Details

This data frame contains the full set of 102,342 data records from the FishStat *Aquaculture Quantity* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Aquaculture Production. Fisheries and Aquaculture Division. Rome.

<https://www.fao.org/fishery/en/collection/aquaculture>

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

## Examples

```
head(aquaculture)

# Add species columns
aqua <- merge(aquaculture, species)

# Top 10 aquaculture species in 2022, production in tonnes
x <- aggregate(value~species_name, aqua, sum, subset=year==2022)
x$value <- round(x$value)
head(x[order(-x$value),], 10)

# Total aquaculture production by major taxa since 1950, in million tonnes
aggregate(value~tolower(major), aqua, function(x) round(sum(x/1e6)))

# Annual aquaculture production of all animals
x <- aggregate(value~year, aqua, sum, subset=yearbook != "Aquatic plants")
plot(value/1e6~year, x, ylim=c(0,105), ylab="million tonnes", type="l")
title(main="Aquaculture production: All animals")
```

---

area

*Areas*

---

## Description

Major inland and marine fishing areas, defined by FAO (2024).

## Usage

```
area
```

## Format

Data frame containing five columns:

area	area code
area_name	area name
inlandmarine	inland or marine
faregion	northern, central, or southern (marine fishing areas)
ocean	Atlantic, Indian, Pacific, or Southern Ocean (marine fishing areas)

## Details

This data frame contains the full set of 29 data records from the FishStat *Water Area Groups* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

## Examples

```
head(area)

# 28 'area' entries have non-zero production, 1 has no production
nonzero <- unique(production$area[production$value > 0])
area.nz <- area[area$area %in% nonzero,]
area$area_name[!(area$area %in% nonzero)] # Antarctica - Inland waters

# Areas 98 and 99 are only used for whale catches
species_98_99 <- unique(production$species[production$area %in% 98:99])
cbind(unique(species$isscaap[species$species %in% species_98_99]))

# Marine fishing areas in northern, central, and southern regions
area$area[area$faregion == "Northern regions"]
area$area[area$faregion == "Central regions"]
area$area[area$faregion == "Southern regions"]

# Inland waters and marine areas
area[area$inlandmarine == "Inland waters", c("area", "area_name")]
area[area$inlandmarine == "Marine areas", c("area", "area_name")]

# Examine one area
print.simple.list(area[area$area == 71,])
```

---

capture	<i>Global Capture Production</i>
---------	----------------------------------

---

## Description

Capture production quantity by species, area, and country for the years 1950-2022, compiled and published by FAO (2024).

## Usage

```
capture
```

## Format

Data frame containing seven columns:

species	species code
year	year
area	area code
country	country code
value	quantity in tonnes or number of individuals
measure	measure code
status	status code

## Details

This data frame contains the full set of 1,050,119 data records from the FishStat *Capture Quantity* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Capture Production. Fisheries and Aquaculture Division. Rome.  
<https://www.fao.org/fishery/en/collection/capture>

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.  
[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.  
[fishstat-package](#) gives an overview of the package.

## Examples

```
head(capture)

# Analyze catches measured in tonnes
cap <- aggregate(value~species+year, capture, sum, subset=measure=="Q_tlw")
cap <- merge(cap, species[c("species", "species_name", "major")])
```

```
# Top 10 capture species in 2022
x <- aggregate(value~species_name, cap, sum, subset=year==2022)
x$value <- round(x$value)
head(x[order(-x$value),], 10)

# Total capture production by major taxa since 1950, in million tonnes
x <- aggregate(value~tolower(major), cap, function(x) round(sum(x/1e6)))
x[x$value > 0,]

# Annual capture production of all aquatic animals
x <- aggregate(value~year, cap, sum)
plot(value/1e6~year, x, ylim=c(0,105), ylab="million tonnes", type="l")
title(main="Capture production: All")
```

---

country

*Countries*


---

### Description

Countries and various territories, defined by FAO (2024).

### Usage

```
country
```

### Format

Data frame containing eight columns:

country	country code
country_name	country name
iso2	ISO 2-alpha code
iso3	ISO 3-alpha code
continent	continent
georegion	geographic region
ecoclass	economic class
official	official country name

### Details

This data frame contains the full set of 275 data records from the FishStat *Country Groups* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

### Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)



**See Also**

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.  
[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.  
[fishstat-package](#) gives an overview of the package.

**Examples**

```
head(country)

# Regions within continents
table(country$georegion, country$continent)

# 240 'country' entries have non-zero production, 35 have no production
nonzero <- unique(production$country[production$value > 0])
country.nz <- country[country$country %in% nonzero,]
country$country_name[!(country$country %in% nonzero)]

# Only 'country' and 'country_name' are always defined
cbind(sapply(country, function(x) all(x!="")))

# Plus 'ecoclass' and 'official' for non-zero production
cbind(sapply(country.nz, function(x) all(x!="")))

# Economic class levels
sort(unique(country$ecoclass))

# Examine individual countries
print.simple.list(country[country$iso2 == "IS",])
print.simple.list(country[country$country_name == "Samoa",])
```

---

environment

*Environments*


---

**Description**

Aquatic environments for aquaculture, defined by FAO (2024).

**Usage**

```
environment
```

**Format**

Data frame containing two columns:

environment	environment code
environment_name	environment name

**Details**

This data frame contains the full set of 4 data records from the FishStat *Production Environment* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

**Source**

FAO (2024). Aquaculture Production. Fisheries and Aquaculture Division. Rome.

<https://www.fao.org/fishery/en/collection/aquaculture>

**See Also**

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

**Examples**

```
environment

# Aquaculture production by environment in 2022, in million tonnes
x <- merge(aquaculture, environment)
x <- aggregate(value~environment_name, x, sum, subset=year==2022)
transform(x, value=round(value/1e6))
```

---

measure

*Measures*

---

**Description**

Units of measurement, defined by FAO (2024).

**Usage**

```
measure
```

**Format**

Data frame containing two columns:

measure	measure code
measure_name	measure name
short	short name
multiplier	unit multiplier
unit	unit symbol
measure_description	measure description
sws	SWS code

## Details

This data frame contains the full set of 11 data records from the FishStat *Units* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

## Examples

```
head(measure, 3)

# Aquaculture production is measured in tonnes live weight
table(aquaculture$measure)

# Capture production is measured in tonnes or number of individuals
table(capture$measure)

# When number of individuals is used, it is for mammals and reptiles
x <- merge(capture[capture$measure=="Q_no_1",], species)
aggregate(value~isscaap, x, sum)
aggregate(value~isscaap, x, sum, subset=year==2022)

# Examine one measure
print.simple.list(measure[measure$measure=="Q_tlw",])
```

---

production

*Global Aquaculture and Capture Production*

---

## Description

Aquaculture and capture production quantity by species, area, and country for the years 1950-2022, compiled and published by FAO (2024).

## Usage

```
production
```

**Format**

Data frame containing eight columns:

species	species code
year	year
area	area code
country	country code
value	quantity in tonnes or number of individuals
measure	measure code
status	status code
source	source code

**Details**

This data frame contains the full set of 1,152,461 data records from the FishStat *Production Quantity* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

**Source**

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

**See Also**

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

**Examples**

```
head(production)

# Analyze production measured in tonnes
prod <- production[production$measure == "Q_tlw" & production$value > 0,]
prod <- merge(prod, species[c("species", "yearbook")])

# Select SOFIA species, excluding mammals, reptiles, and plants
prod <- prod[prod$yearbook == "Fish, crustaceans and molluscs, etc.",]

# Determine origin
prod <- merge(prod, area[c("area", "inlandmarine")])
prod$origin <- ifelse(prod$source == "CAPTURE", "Capture", "Aquaculture")
prod$w <- ifelse(prod$inlandmarine == "Marine areas", "marine", "inland")
prod$origin <- paste0(prod$origin, " (", prod$w, ")")
cbind(sort(unique(prod$origin)))

# World capture fisheries and aquaculture production
x <- xtabs(value~year+origin, aggregate(value~year+origin, prod, sum))
```

```
x <- x[,c(2,1,4,3)] / 1e6
library(areaplot)
areaplot(x, legend=TRUE, args.legend=list(x="topleft"), ylab="million tonnes")
```

---

source

*Sources*


---

## Description

Sources of aquaculture and capture production, defined by FAO (2024).

## Usage

```
source
```

## Format

Data frame containing two columns:

source	source code
source_name	source name

## Details

This data frame contains the full set of 4 data records from the FishStat *Production Source* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format. [area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables. [fishstat-package](#) gives an overview of the package.

## Examples

```
source

# Analyze production measured in tonnes
prod <- production[production$measure == "Q_tlw" & production$value > 0,]
prod <- merge(prod, source)

# Production by source in 2022, in million tonnes
x <- aggregate(value~source_name, prod, sum, subset=year==2022)
transform(x, value=round(value/1e6))
```

---

species	<i>Species</i>
---------	----------------

---

### Description

Aquatic species and taxonomic groups, defined by FAO (2024).

### Usage

species

### Format

Data frame containing ten columns:

species	species code
species_name	species name
scientific	scientific name
isscaap	ISSCAAP group
major	major taxa
cpc_class	CPC class
cpc_group	CPC group
yearbook	yearbook category
author	author of scientific name
taxonomic	taxonomic code

### Details

This data frame contains the full set of 13,615 data records from the FishStat *Species Groups* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

### Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

### See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.

[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.

[fishstat-package](#) gives an overview of the package.

**Examples**

```

head(species, 3)

# 3524 'species' entries have non-zero production, 9893 have no production
nonzero <- unique(production$species[production$value > 0])
species.nz <- species[species$species %in% nonzero,]
length(unique(species.nz$species))

# Only 'species', 'scientific', 'major', and 'taxonomic' are always defined
cbind(sapply(species, function(x) all(x!="")))

# Plus 'isscaap' and 'yearbook' for non-zero production
cbind(sapply(species.nz, function(x) all(x!="")))

# A variety of species are missing 'species_name', 'cpc_class', 'cpc_group'
cbind(table(species.nz$major[species.nz$species_name == ""]))
cbind(table(species.nz$major[species.nz$cpc_class == ""]))
cbind(table(species.nz$major[species.nz$cpc_group == ""]))

# 3524 non-zero production species by major taxa
cbind(table(species.nz$major))

# 3524 species in 7 major taxa, 50 isscaap groups, and 3 yearbook categories
cbind(sapply(species.nz, function(x) length(unique(x))))

# 'scientific' and 'species_name' entries are not unique
table(species.nz$scientific)[table(species.nz$scientific) > 1]
table(species.nz$species_name)[table(species.nz$species_name) > 1]

# A closer look at the yearbook categories
cbind(table(species.nz$yearbook)) # now inspect the 149 "other" species
other <- species.nz[species.nz$yearbook == "Other aq. animals & products",]
cbind(table(other$major)) # majority is mammals, what about the rest
other <- other[other$major != "MAMMALIA",]
cbind(table(other$isscaap)) # reptiles and inedible ornamental animals

# Examine one species
print.simple.list(species[species$species_name == "Atlantic cod",])
print.simple.list(species[species$species == "YFT",])

```

---

status

*Status*


---

**Description**

Status of data entries, defined by FAO (2024).

**Usage**

```
status
```

## Format

Data frame containing four columns:

status	status code
status_name	status name
status_description	short name
alternate	unit multiplier

## Details

This data frame contains the full set of 16 data records from the FishStat *Symbols* data table. Column names have been simplified to facilitate quick exploration and plotting in R.

## Source

FAO (2024). Global Production. Fisheries and Aquaculture Division. Rome.

[https://www.fao.org/fishery/en/collection/global\\_production](https://www.fao.org/fishery/en/collection/global_production)

## See Also

[aquaculture](#) and [capture](#) data are also available in a combined [production](#) format.  
[area](#), [country](#), [environment](#), [measure](#), [source](#), [species](#), and [status](#) are lookup tables.  
[fishstat-package](#) gives an overview of the package.

## Examples

```
head(status, 3)

# Aquaculture data entries
# 80% have official status, 17% are estimated, and 3% are negligible
100 * proportions(table(aquaculture$status))

# Capture data entries
# 94% have official status, 3% are estimated, and 3% are negligible
100 * proportions(table(capture$status))

# Examine one status definition
print.simple.list(status[status$status=="N",])
```



# Index

aquaculture, [2](#), [4](#), [5–7](#), [9–14](#), [16](#)  
area, [2](#), [4](#), [5](#), [5](#), [6](#), [7](#), [9–14](#), [16](#)  
capture, [2](#), [5–7](#), [7](#), [9–14](#), [16](#)  
country, [2](#), [4–7](#), [8](#), [9–14](#), [16](#)  
environment, [2](#), [4–7](#), [9](#), [9](#), [10–14](#), [16](#)  
fishstat (fishstat-package), [2](#)  
fishstat-package, [2](#)  
measure, [2](#), [4–7](#), [9](#), [10](#), [10](#), [11–14](#), [16](#)  
merge, [2](#)  
production, [2](#), [5–7](#), [9–11](#), [11](#), [12–14](#), [16](#)  
source, [2](#), [5–7](#), [9–13](#), [13](#), [14](#), [16](#)  
species, [2](#), [4–7](#), [9–14](#), [14](#), [16](#)  
status, [2](#), [4–7](#), [9–14](#), [15](#), [16](#)