

# Package ‘statprograms’

July 23, 2025

**Title** Graduate Statistics Program Datasets

**Version** 0.3.0

**Description** A small collection of data on graduate statistics programs from the United States.

**URL** <https://brettklamer.com/work/statprograms/>

**License** MIT + file LICENSE

**Depends** R (>= 2.10.0)

**LazyData** TRUE

**Encoding** UTF-8

**RoxygenNote** 7.3.2

**NeedsCompilation** no

**Author** Brett Klamer [aut, cre]

**Maintainer** Brett Klamer <code@brettklamer.com>

**Repository** CRAN

**Date/Publication** 2025-05-01 00:40:02 UTC

## Contents

degreesawarded . . . . .	1
statprograms . . . . .	2

<b>Index</b>	<b>5</b>
--------------	----------

---

degreesawarded	<i>Degrees Awarded by Year</i>
----------------	--------------------------------

---

## Description

This dataset contains the number of degrees awarded per year. It's based on data from the National Center for Education Statistics as retrieved by Steve Pierson. See <https://community.amstat.org/blogs/steve-pierson/2014/07/28/categorization-of-statistics-degrees> for more information.

**Usage**

degreesawarded

**Format**

A data.frame with 4606 observations and 5 columns. The columns are defined as follows:

school The college

program\_category The program type categorized as either "Statistics" or "Biostatistics"

degree\_category The degree categorized as either "Master" or "Doctorate"

year The year the degrees were awarded

count The number of degrees awarded

**Source**

"Statistics and Biostatistics Degree Data.", <https://www.amstat.org/education/statistics-and-biostatistics-deg>

**Examples**

```
## Not run:
data(degreesawarded)
summary(degreesawarded)

# In wide format as provided by Steve Pierson
library(tidyr)
spread(degreesawarded, key = year, value = count)

## End(Not run)
```

---

statprograms

*Graduate Statistics Program Data*


---

**Description**

This dataset contains various information from the majority of graduate statistics programs in the United States.

**Usage**

statprograms

**Format**

A data.frame with 490 observations and 16 columns. The columns are defined as follows:

school The college

program The program type as advertised by the department

program\_category The program type categorized as either "Statistics" or "Biostatistics"

degree The degree given by the department

degree\_category The degree categorized as either "Master" or "Doctorate"

state The state

city The city

square\_miles The square miles of the city (or region) from <https://www.wikipedia.org/>

population The population of the city (or region) from <https://www.wikipedia.org/> or <https://www.census.gov/programs-surveys/popest/data/data-sets.html>. Most are estimates from 2010 to 2014.

density The population density

average\_winter The average winter temperature

average\_summer The average summer temperature

latitude The latitude of the department's building (or as close as possible) from <https://www.gps-coordinates.net/>

longitude The longitude of the department's building (or as close as possible) from <https://www.gps-coordinates.net/>

link The URL of the department's website

date\_collected The date the information was recorded

**Author(s)**

Brett Klammer

**Examples**

```
## Not run:
data(statprograms)
summary(statprograms)

#-----
# Plot locations on a map
#-----
library(maps)
library(ggplot2)
library(mapproj)

us_states <- map_data("state")

ggplot(
  data = statprograms[statprograms$state != "Alaska", ],
```

```
mapping = aes(x = longitude, y = latitude)
) +
geom_polygon(
  data = us_states,
  aes(x = long, y = lat, group = group),
  fill = "white",
  color = "gray50"
) +
geom_point() +
guides(fill = "none") +
coord_map(
  projection = "albers",
  lat0 = 39,
  lat1 = 45
) +
theme_bw()

## End(Not run)
```

# Index

## \* **datasets**

degreesawarded, [1](#)

statprograms, [2](#)

degreesawarded, [1](#)

statprograms, [2](#)