Longhurst\_figure\_cleanup\_ALX.R

test

Tue Jul 2 10:57:09 2019

## Alex prepping data for longhurst province figure ##  
# load packages and data ----  
library(tidyverse)

## ── Attaching packages ─────────────────────────── tidyverse 1.2.1 ──

## ✔ ggplot2 3.2.0 ✔ purrr 0.2.4  
## ✔ tibble 1.4.2 ✔ dplyr 0.7.8  
## ✔ tidyr 0.8.0 ✔ stringr 1.3.1  
## ✔ readr 1.1.1 ✔ forcats 0.3.0

## Warning: package 'ggplot2' was built under R version 3.5.2

## ── Conflicts ────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(gbm)

## Loading required package: survival

## Loading required package: lattice

## Loading required package: splines

## Loading required package: parallel

## Loaded gbm 2.1.3

library(dismo)

## Loading required package: raster

## Loading required package: sp

##   
## Attaching package: 'raster'

## The following object is masked from 'package:dplyr':  
##   
## select

## The following object is masked from 'package:tidyr':  
##   
## extract

library(mgcv)

## Loading required package: nlme

##   
## Attaching package: 'nlme'

## The following object is masked from 'package:raster':  
##   
## getData

## The following object is masked from 'package:dplyr':  
##   
## collapse

## This is mgcv 1.8-23. For overview type 'help("mgcv-package")'.

library(lme4)

## Loading required package: Matrix

##   
## Attaching package: 'Matrix'

## The following object is masked from 'package:tidyr':  
##   
## expand

##   
## Attaching package: 'lme4'

## The following object is masked from 'package:nlme':  
##   
## lmList

library(gamm4)

## This is gamm4 0.2-5

library(readxl)  
library(readr)  
library(ggplot2)  
  
data <- read.csv("Plastics ingestion records fish master\_UDPATED\_AGM-MSS2.csv")  
  
# want to get average proportion of plastic per province  
summary(data)

## Species.name Common.name   
## :1048105 :1048122   
## Sardina pilchardus : 6 European pilchard : 6   
## Merluccius merluccius: 5 Atlantic horse mackerel: 5   
## Mullus barbatus : 5 Chub mackerel : 5   
## Mullus surmuletus : 5 European hake : 5   
## Platichthys flesus : 5 Lesser spoted dogfish : 5   
## (Other) : 444 (Other) : 427   
## Order Family   
## :1048114 :1048114   
## Perciformes : 200 Myctophidae: 27   
## Gadiformes : 38 Scombridae : 27   
## Scorpaeniformes: 32 Sparidae : 24   
## Clupeiformes : 29 Gadidae : 21   
## Myctophiformes : 27 Clupeidae : 20   
## (Other) : 135 (Other) : 342   
## Diet Trophic.level.via.fishbase  
## :1048358 Min. :2.0   
## predatory : 46 1st Qu.:3.3   
## no data : 22 Median :3.7   
## carnivore : 17 Mean :3.6   
## invertebrates, piscivorous: 13 3rd Qu.:4.1   
## crustaceans : 6 Max. :4.7   
## (Other) : 113 NA's :1048116   
## NwP N Prop.w.plastic   
## Min. : 0.0 Min. : 1.0 Min. :0.0   
## 1st Qu.: 0.0 1st Qu.: 7.0 1st Qu.:0.0   
## Median : 2.0 Median : 20.0 Median :0.1   
## Mean : 9.8 Mean : 324.5 Mean :0.2   
## 3rd Qu.: 9.0 3rd Qu.: 83.0 3rd Qu.:0.3   
## Max. :506.0 Max. :25914.0 Max. :1.0   
## NA's :1048105 NA's :1048105 NA's :1048105   
## average.debris.proportion....mass.  
## :1048400   
## 0 : 55   
## n/a : 31   
## 1 : 3   
## <1% ingested plastic: 2   
## 0.3±0.6 : 2   
## (Other) : 82   
## Type.of.debris   
## :1048401   
## n/a : 61   
## plastics 86.5% in 25 individuals, metals 8.1% 3 individuals, wood 2.7% 1 individual: 26   
## no data : 18   
## microplastics, fibres : 7   
## \*see note : 6   
## (Other) : 56   
## Color.of.debris   
## :1048401   
## n/a : 67   
## no data : 50   
## blue and black (bags), brown (hard plastic), transparent green (fishing gear): 26   
## \*only shows % by color/type : 10   
## \*gives % of colored debris, see note : 3   
## (Other) : 18   
## Average.size.of.plastic  
## :1048486   
## 2.11±1.67 mm and 1 particle 9.432 mm above average size range: 26   
## 5-60 mm : 26   
## n/a : 23   
## no data : 6   
## 0.03±0.1 mg average : 1   
## (Other) : 7   
## Size.of.fish Habitat Average.depth   
## :1048542 :1048118 Min. : 1.0   
## \*see note : 13 demersal : 151 1st Qu.: 50.0   
## n/a : 2 reef-associated: 72 Median : 150.0   
## no data : 2 benthopelagic : 66 Mean : 261.0   
## not specified: 2 pelagic-neritic: 61 3rd Qu.: 300.4   
## 102.29 cm : 1 bathypelagic : 51 Max. :2999.0   
## (Other) : 13 (Other) : 56 NA's :1048149   
## All\_forage   
## :1048107   
## Active predation, benthic foraging : 96   
## Active predation, benthic foraging, particulate feeder: 85   
## Filter-feeder: facultative, particulate feeder : 47   
## Benthic foraging : 38   
## Particulate feeder : 33   
## (Other) : 169   
## Prime\_forage Second\_forage   
## :1048106 :1048246   
## Benthic foraging : 188 Active predation : 109   
## Active predation : 117 Benthic foraging : 80   
## Particulate feeder: 102 Particulate feeder : 46   
## not listed : 25 Filter-feeder: facultative: 38   
## Grazing : 20 not listed : 25   
## (Other) : 17 (Other) : 31   
## Primary\_aggregation Secondary\_aggregation  
## :1048375 :1048534   
## Pairing : 1 Pairing : 2   
## Schooling : 168 Schooling: 5   
## Schooling (juv): 4 Solitary : 14   
## Solitary : 20 Spawning : 20   
## Spawning : 7   
##   
## Commercial Aquaculture Recreational  
## :1048113 :1048113 :1048113   
## commercial : 175 experimental: 15 N: 286   
## highly commercial: 106 N : 395 Y: 176   
## minor commercial : 97 Y : 52   
## none : 81   
## subsistence : 3   
##   
## IUCN.status   
## :1048113   
## LC : 296   
## Not assessed: 99   
## NT : 20   
## DD : 19   
## VU : 17   
## (Other) : 11   
## Vulnerability.score..via.fishbase.from.Cheug.et.al.2005.  
## Min. :10.0   
## 1st Qu.:31.0   
## Median :40.0   
## Mean :44.1   
## 3rd Qu.:59.0   
## Max. :88.0   
## NA's :1048113   
## Oceanographic.province..from.Longhurst.2007. Basin   
## :1048105 :1048193   
## NECS : 104 North Atlantic: 151   
## MEDI : 98 Mediterranean : 86   
## NPTG : 40 South Pacific : 53   
## SPSG : 37 North Pacific : 50   
## NASE : 29 Red Sea : 26   
## (Other): 162 (Other) : 16   
## Gyre Year.of.collection   
## :1048131 :1048139   
## Mediterranean Sea : 58 2014 : 61   
## Iberian Sea : 39 2015 : 56   
## South Pacific Islands: 34 1999-2016 : 39   
## North Sea : 33 Sept 2015-Oct 2016: 34   
## Portuguese Coast : 29 (Other) : 245   
## (Other) : 251 NA's : 1   
## Publicaton.year Source   
## Min. :1972 :1048105   
## 1st Qu.:2013 Lucia et al. 2018 : 39   
## Median :2016 Markic et al. 2018 : 34   
## Mean :2014 Guven et al. 2017 : 28   
## 3rd Qu.:2018 Davison and Asch 2011 : 27   
## Max. :2018 Anastasopoulou et al. 2013: 26   
## NA's :1048105 (Other) : 316   
## Notes   
## :1048515   
## \*Only considered meso and macroplastics, may exclude from analyses\* : 39   
## \*\*River Thames : 2   
## \*noted to eat garbaage from fish base : 2   
## \* Particles: 33 total, 25 plastic particles, 4 black aluminium silicate, 3 white calcium carbonate, 1 unknown: 1   
## \*\* Celtic Sea; amount consumed positively correlated with weight and velocity : 1   
## (Other) : 15

head(data)

## Species.name Common.name Order  
## 1 Alepisaurus ferox Longsnouted lancetfish Aulopiformes  
## 2 Makaira indica Black marlin Perciformes  
## 3 Boreogadus saida Polar cod Gadiformes  
## 4 Boreogadus saida Polar cod Gadiformes  
## 5 Triglops nybelini Bigeye sculpin Scorpaeniformes  
## 6 Priacanthus arenatus Atlantic bigeye Perciformes  
## Family Diet  
## 1 Alepisauridae piscivorous, cephalopods, crustaceans  
## 2 Istiophoridae piscivorous, cephalopods, crustaceans  
## 3 Gadidae   
## 4 Gadidae   
## 5 Cottidae   
## 6 Priacanthidae   
## Trophic.level.via.fishbase NwP N Prop.w.plastic  
## 1 4.0 1 1 1.00000000  
## 2 4.5 1 1 1.00000000  
## 3 3.1 2 72 0.02777778  
## 4 3.1 15 85 0.17647059  
## 5 3.3 24 71 0.33802817  
## 6 4.0 60 122 0.49180328  
## average.debris.proportion....mass. Type.of.debris Color.of.debris  
## 1 100% rope green, orange  
## 2 100% plastic blue and transparent  
## 3   
## 4   
## 5   
## 6   
## Average.size.of.plastic Size.of.fish Habitat Average.depth  
## 1 2.0 mm 77 cm bathypelagic 915  
## 2 no data 128 cm pelagic-oceanic 100  
## 3 demersal 200  
## 4 demersal 200  
## 5 demersal 400  
## 6 reef-associated 25  
## All\_forage  
## 1 Active predation, benthic foraging  
## 2 Active predation, benthic foraging  
## 3 Benthic foraging, particulate feeder  
## 4 Active predation, benthic foraging, particulate feeder  
## 5 Benthic foraging  
## 6 Active predation, benthic foraging, particulate feeder  
## Prime\_forage Second\_forage Primary\_aggregation  
## 1 Active predation Benthic foraging   
## 2 Active predation Benthic foraging   
## 3 Benthic foraging Particulate feeder   
## 4 Particulate feeder Benthic foraging   
## 5 Benthic foraging   
## 6 Particulate feeder Benthic foraging   
## Secondary\_aggregation Commercial Aquaculture Recreational  
## 1 none N N  
## 2 commercial N Y  
## 3 highly commercial N N  
## 4 highly commercial N N  
## 5 none N N  
## 6 minor commercial N Y  
## IUCN.status Vulnerability.score..via.fishbase.from.Cheug.et.al.2005.  
## 1 LC 83  
## 2 DD 78  
## 3 Not assessed 46  
## 4 Not assessed 46  
## 5 Not assessed 21  
## 6 LC 25  
## Oceanographic.province..from.Longhurst.2007. Basin Gyre  
## 1 AUSW North Atlantic Indian Gyre  
## 2 AUSW North Atlantic Indian Gyre  
## 3 BPLR Arctic Ocean  
## 4 BPLR Arctic Ocean Arctic Ocean  
## 5 BPLR Arctic Ocean Arctic Ocean  
## 6 BRAZ Brazil coast  
## Year.of.collection Publicaton.year Source Notes  
## 1 August-September 2005 2008 Fujieda et al. 2008   
## 2 August-September 2005 2008 Fujieda et al. 2008   
## 3 2012-2015 2018 Kuhn et al. 2018   
## 4 2015 2018 Morgana et al. 2018   
## 5 2015 2018 Morgana et al. 2018   
## 6 May 2015-Nov 2016 2018 Cardozo et al. 2018

# data of interest  
data2 <- data[,c("Species.name", "Oceanographic.province..from.Longhurst.2007.", "Prop.w.plastic", "NwP", "N", "Source")]  
head(data2)

## Species.name Oceanographic.province..from.Longhurst.2007.  
## 1 Alepisaurus ferox AUSW  
## 2 Makaira indica AUSW  
## 3 Boreogadus saida BPLR  
## 4 Boreogadus saida BPLR  
## 5 Triglops nybelini BPLR  
## 6 Priacanthus arenatus BRAZ  
## Prop.w.plastic NwP N Source  
## 1 1.00000000 1 1 Fujieda et al. 2008  
## 2 1.00000000 1 1 Fujieda et al. 2008  
## 3 0.02777778 2 72 Kuhn et al. 2018  
## 4 0.17647059 15 85 Morgana et al. 2018  
## 5 0.33802817 24 71 Morgana et al. 2018  
## 6 0.49180328 60 122 Cardozo et al. 2018

colnames(data2) <- c("Species", "OceanProv", "PropPlastic", "NwP", "N", "Source")  
head(data2)

## Species OceanProv PropPlastic NwP N Source  
## 1 Alepisaurus ferox AUSW 1.00000000 1 1 Fujieda et al. 2008  
## 2 Makaira indica AUSW 1.00000000 1 1 Fujieda et al. 2008  
## 3 Boreogadus saida BPLR 0.02777778 2 72 Kuhn et al. 2018  
## 4 Boreogadus saida BPLR 0.17647059 15 85 Morgana et al. 2018  
## 5 Triglops nybelini BPLR 0.33802817 24 71 Morgana et al. 2018  
## 6 Priacanthus arenatus BRAZ 0.49180328 60 122 Cardozo et al. 2018

length(table(data2$OceanProv))

## [1] 22

data3<- data2[order(data2$OceanProv),]  
head(data3)

## Species OceanProv PropPlastic NwP N Source  
## 471 NA NA NA   
## 472 NA NA NA   
## 473 NA NA NA   
## 474 NA NA NA   
## 475 NA NA NA   
## 476 NA NA NA

### set up new dataframe   
prov <- unique(data3$OceanProv)[-1]  
prov

## [1] AUSW BPLR BRAZ CARB CCAL CHIL CHIN EAFR FKLD GFST GUIA INDE KURO MEDI  
## [15] NASE NECS NPTG REDS SPSG SUND TASM  
## 22 Levels: AUSW BPLR BRAZ CARB CCAL CHIL CHIN EAFR FKLD GFST ... TASM

length(prov)

## [1] 21

aveplast <- rep(NA, length(prov))  
numfish <- rep(NA, length(prov))  
numstudies <- rep(NA, length(prov))  
numspecies <- rep(NA, length(prov))  
  
## using this to double check the produced averages  
sub <- data3[data3$OceanProv=="CHIL",]  
sub

## Species OceanProv PropPlastic NwP N Source  
## 41 Odontesthes regia CHIL 0.11111111 1 9 Ory et al. 2018  
## 42 Strangomera bentincki CHIL 0.00000000 0 10 Ory et al. 2018  
## 43 Sardinops sagax CHIL 0.00000000 0 7 Ory et al. 2018  
## 44 Opisthonema libertate CHIL 0.02500000 2 80 Ory et al. 2018  
## 45 Cetengraulis mysticetus CHIL 0.02500000 1 40 Ory et al. 2018  
## 46 Engraulis ringens CHIL 0.00862069 1 116 Ory et al. 2018  
## 47 Scomber japonicus CHIL 0.03333333 1 30 Ory et al. 2018

unique(sub$Source)

## [1] Ory et al. 2018  
## 67 Levels: Alomar and Deudero 2017 ... Wieczorek et al. 2018

length(unique(sub$Source))

## [1] 1

sum(sub$NwP)/sum(sub$N)

## [1] 0.02054795

mean(sub$PropPlastic) #3 = .1077803

## [1] 0.0290093

sum(sub$N)

## [1] 292

nrow(sub)

## [1] 7

length(unique(sub$Species))

## [1] 7

for (i in 1:length(prov)){  
 sub <- data3[data3$OceanProv==prov[i],]  
 aveplast[i] <- sum(sub$NwP)/sum(sub$N)  
 numfish[i] <- sum(sub$N)  
 numstudies[i] <- length(unique(sub$Source))  
 numspecies[i] <- length(unique(sub$Species))  
}  
  
aveplast

## [1] 1.000000000 0.179824561 0.129587156 0.251890359 0.029850746  
## [6] 0.020547945 1.000000000 0.001298701 0.144927536 0.068020930  
## [11] 0.179571664 0.214285714 0.765625000 0.267062767 0.252604167  
## [16] 0.006479657 0.184549356 0.146067416 0.232535885 0.231884058  
## [21] 0.002923977

length(aveplast)

## [1] 21

numfish

## [1] 2 228 872 2116 1005 292 378 10780 69 3249  
## [11] 607 28 64 6564 384 123309 932 178 1045 69  
## [21] 342

table(numfish)

## numfish  
## 2 28 64 69 178 228 292 342 378 384   
## 1 1 1 2 1 1 1 1 1 1   
## 607 872 932 1005 1045 2116 3249 6564 10780 123309   
## 1 1 1 1 1 1 1 1 1 1

numstudies

## [1] 1 2 4 2 2 1 1 1 1 6 2 1 1 16 2 15 4 1 3 1 1

head(data3)

## Species OceanProv PropPlastic NwP N Source  
## 471 NA NA NA   
## 472 NA NA NA   
## 473 NA NA NA   
## 474 NA NA NA   
## 475 NA NA NA   
## 476 NA NA NA

numspecies

## [1] 2 2 15 6 14 7 21 7 1 19 6 2 1 70 29 74 37 26 35 10 21

newdat <- data.frame(prov, aveplast, numfish, numstudies, numspecies)  
newdat

## prov aveplast numfish numstudies numspecies  
## 1 AUSW 1.000000000 2 1 2  
## 2 BPLR 0.179824561 228 2 2  
## 3 BRAZ 0.129587156 872 4 15  
## 4 CARB 0.251890359 2116 2 6  
## 5 CCAL 0.029850746 1005 2 14  
## 6 CHIL 0.020547945 292 1 7  
## 7 CHIN 1.000000000 378 1 21  
## 8 EAFR 0.001298701 10780 1 7  
## 9 FKLD 0.144927536 69 1 1  
## 10 GFST 0.068020930 3249 6 19  
## 11 GUIA 0.179571664 607 2 6  
## 12 INDE 0.214285714 28 1 2  
## 13 KURO 0.765625000 64 1 1  
## 14 MEDI 0.267062767 6564 16 70  
## 15 NASE 0.252604167 384 2 29  
## 16 NECS 0.006479657 123309 15 74  
## 17 NPTG 0.184549356 932 4 37  
## 18 REDS 0.146067416 178 1 26  
## 19 SPSG 0.232535885 1045 3 35  
## 20 SUND 0.231884058 69 1 10  
## 21 TASM 0.002923977 342 1 21

## binning data  
# ave plastic  
newdat$aveplastbin <- rep(NA, length(prov))  
newdat$aveplastbin[newdat$aveplast<=.10]= "<.10"  
newdat$aveplastbin[newdat$aveplast >.10 & newdat$aveplast<=20]= ".11-.20"  
newdat$aveplastbin[newdat$aveplast >.20 & newdat$aveplast<=.30]= ".21-.30"  
newdat$aveplastbin[newdat$aveplast >.30 & newdat$aveplast<=.40]= ".31-.40"  
newdat$aveplastbin[newdat$aveplast >.40 & newdat$aveplast<=.50]= ".41-.50"  
newdat$aveplastbin[newdat$aveplast >.50 & newdat$aveplast<=.60]= ".51-.60"  
newdat$aveplastbin[newdat$aveplast >.60 & newdat$aveplast<=.70]= ".61-.70"  
newdat$aveplastbin[newdat$aveplast >.70 & newdat$aveplast<=.80]= ".71-.80"  
newdat$aveplastbin[newdat$aveplast >.80 & newdat$aveplast<=.90]= ".81-.90"  
newdat$aveplastbin[newdat$aveplast >.90 & newdat$aveplast<=.99]= ".90-.99"  
newdat$aveplastbin[newdat$aveplast ==1]= "1"  
  
head(newdat$aveplastbin, 50)

## [1] "1" ".11-.20" ".11-.20" ".21-.30" "<.10" "<.10" "1"   
## [8] "<.10" ".11-.20" "<.10" ".11-.20" ".21-.30" ".71-.80" ".21-.30"  
## [15] ".21-.30" "<.10" ".11-.20" ".11-.20" ".21-.30" ".21-.30" "<.10"

sort(newdat$aveplastbin)

## [1] ".11-.20" ".11-.20" ".11-.20" ".11-.20" ".11-.20" ".11-.20" ".21-.30"  
## [8] ".21-.30" ".21-.30" ".21-.30" ".21-.30" ".21-.30" ".71-.80" "<.10"   
## [15] "<.10" "<.10" "<.10" "<.10" "<.10" "1" "1"

# number of studies  
summary(newdat$numstudies)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.000 1.000 2.000 3.238 3.000 16.000

newdat$numstudiesbin <- rep(NA, length(prov))  
newdat$numstudiesbin[newdat$numstudies == 1] = "1"  
newdat$numstudiesbin[newdat$numstudies >1 & newdat$numstudies<=5]= "2-5"  
newdat$numstudiesbin[newdat$numstudies >5 & newdat$numstudies<=10]= "6-10"  
newdat$numstudiesbin[newdat$numstudies >=10]= ">10"  
  
head(newdat, 50)

## prov aveplast numfish numstudies numspecies aveplastbin  
## 1 AUSW 1.000000000 2 1 2 1  
## 2 BPLR 0.179824561 228 2 2 .11-.20  
## 3 BRAZ 0.129587156 872 4 15 .11-.20  
## 4 CARB 0.251890359 2116 2 6 .21-.30  
## 5 CCAL 0.029850746 1005 2 14 <.10  
## 6 CHIL 0.020547945 292 1 7 <.10  
## 7 CHIN 1.000000000 378 1 21 1  
## 8 EAFR 0.001298701 10780 1 7 <.10  
## 9 FKLD 0.144927536 69 1 1 .11-.20  
## 10 GFST 0.068020930 3249 6 19 <.10  
## 11 GUIA 0.179571664 607 2 6 .11-.20  
## 12 INDE 0.214285714 28 1 2 .21-.30  
## 13 KURO 0.765625000 64 1 1 .71-.80  
## 14 MEDI 0.267062767 6564 16 70 .21-.30  
## 15 NASE 0.252604167 384 2 29 .21-.30  
## 16 NECS 0.006479657 123309 15 74 <.10  
## 17 NPTG 0.184549356 932 4 37 .11-.20  
## 18 REDS 0.146067416 178 1 26 .11-.20  
## 19 SPSG 0.232535885 1045 3 35 .21-.30  
## 20 SUND 0.231884058 69 1 10 .21-.30  
## 21 TASM 0.002923977 342 1 21 <.10  
## numstudiesbin  
## 1 1  
## 2 2-5  
## 3 2-5  
## 4 2-5  
## 5 2-5  
## 6 1  
## 7 1  
## 8 1  
## 9 1  
## 10 6-10  
## 11 2-5  
## 12 1  
## 13 1  
## 14 >10  
## 15 2-5  
## 16 >10  
## 17 2-5  
## 18 1  
## 19 2-5  
## 20 1  
## 21 1

# number of fish/study  
summary(newdat$numfish)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 2 178 384 7263 1045 123309

newdat$numfishbin <- rep(NA, length(prov))  
newdat$numfishbin[newdat$numfish >1 & newdat$numfish<10]= "< 10"  
newdat$numfishbin[newdat$numfish >=10 & newdat$numfish<=50]= "10-50"  
newdat$numfishbin[newdat$numfish >50 & newdat$numfish<=100]= "51-100"  
newdat$numfishbin[newdat$numfish >100 & newdat$numfish<500]= "101-500"  
newdat$numfishbin[newdat$numfish >500 & newdat$numfish<=1000]= "501-1000"  
newdat$numfishbin[newdat$numfish >1000 & newdat$numfish<=1500]= "1001-1500"  
newdat$numfishbin[newdat$numfish >1500]= ">1500"  
  
head(newdat, 50)

## prov aveplast numfish numstudies numspecies aveplastbin  
## 1 AUSW 1.000000000 2 1 2 1  
## 2 BPLR 0.179824561 228 2 2 .11-.20  
## 3 BRAZ 0.129587156 872 4 15 .11-.20  
## 4 CARB 0.251890359 2116 2 6 .21-.30  
## 5 CCAL 0.029850746 1005 2 14 <.10  
## 6 CHIL 0.020547945 292 1 7 <.10  
## 7 CHIN 1.000000000 378 1 21 1  
## 8 EAFR 0.001298701 10780 1 7 <.10  
## 9 FKLD 0.144927536 69 1 1 .11-.20  
## 10 GFST 0.068020930 3249 6 19 <.10  
## 11 GUIA 0.179571664 607 2 6 .11-.20  
## 12 INDE 0.214285714 28 1 2 .21-.30  
## 13 KURO 0.765625000 64 1 1 .71-.80  
## 14 MEDI 0.267062767 6564 16 70 .21-.30  
## 15 NASE 0.252604167 384 2 29 .21-.30  
## 16 NECS 0.006479657 123309 15 74 <.10  
## 17 NPTG 0.184549356 932 4 37 .11-.20  
## 18 REDS 0.146067416 178 1 26 .11-.20  
## 19 SPSG 0.232535885 1045 3 35 .21-.30  
## 20 SUND 0.231884058 69 1 10 .21-.30  
## 21 TASM 0.002923977 342 1 21 <.10  
## numstudiesbin numfishbin  
## 1 1 < 10  
## 2 2-5 101-500  
## 3 2-5 501-1000  
## 4 2-5 >1500  
## 5 2-5 1001-1500  
## 6 1 101-500  
## 7 1 101-500  
## 8 1 >1500  
## 9 1 51-100  
## 10 6-10 >1500  
## 11 2-5 501-1000  
## 12 1 10-50  
## 13 1 51-100  
## 14 >10 >1500  
## 15 2-5 101-500  
## 16 >10 >1500  
## 17 2-5 501-1000  
## 18 1 101-500  
## 19 2-5 1001-1500  
## 20 1 51-100  
## 21 1 101-500

head(newdat)

## prov aveplast numfish numstudies numspecies aveplastbin numstudiesbin  
## 1 AUSW 1.00000000 2 1 2 1 1  
## 2 BPLR 0.17982456 228 2 2 .11-.20 2-5  
## 3 BRAZ 0.12958716 872 4 15 .11-.20 2-5  
## 4 CARB 0.25189036 2116 2 6 .21-.30 2-5  
## 5 CCAL 0.02985075 1005 2 14 <.10 2-5  
## 6 CHIL 0.02054795 292 1 7 <.10 1  
## numfishbin  
## 1 < 10  
## 2 101-500  
## 3 501-1000  
## 4 >1500  
## 5 1001-1500  
## 6 101-500

table(newdat$prov)

##   
## AUSW BPLR BRAZ CARB CCAL CHIL CHIN EAFR FKLD GFST GUIA INDE KURO MEDI   
## 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1   
## NASE NECS NPTG REDS SPSG SUND TASM   
## 1 1 1 1 1 1 1

head(newdat)

## prov aveplast numfish numstudies numspecies aveplastbin numstudiesbin  
## 1 AUSW 1.00000000 2 1 2 1 1  
## 2 BPLR 0.17982456 228 2 2 .11-.20 2-5  
## 3 BRAZ 0.12958716 872 4 15 .11-.20 2-5  
## 4 CARB 0.25189036 2116 2 6 .21-.30 2-5  
## 5 CCAL 0.02985075 1005 2 14 <.10 2-5  
## 6 CHIL 0.02054795 292 1 7 <.10 1  
## numfishbin  
## 1 < 10  
## 2 101-500  
## 3 501-1000  
## 4 >1500  
## 5 1001-1500  
## 6 101-500

newdat$aveplast

## [1] 1.000000000 0.179824561 0.129587156 0.251890359 0.029850746  
## [6] 0.020547945 1.000000000 0.001298701 0.144927536 0.068020930  
## [11] 0.179571664 0.214285714 0.765625000 0.267062767 0.252604167  
## [16] 0.006479657 0.184549356 0.146067416 0.232535885 0.231884058  
## [21] 0.002923977

sort(newdat$numstudies)

## [1] 1 1 1 1 1 1 1 1 1 1 2 2 2 2 2 3 4 4 6 15 16

sort(newdat$numfish)

## [1] 2 28 64 69 69 178 228 292 342 378  
## [11] 384 607 872 932 1005 1045 2116 3249 6564 10780  
## [21] 123309

write.csv(newdat, "Longhurst\_FishSummaryData\_fullbinned.csv")  
  
  
newdat

## prov aveplast numfish numstudies numspecies aveplastbin  
## 1 AUSW 1.000000000 2 1 2 1  
## 2 BPLR 0.179824561 228 2 2 .11-.20  
## 3 BRAZ 0.129587156 872 4 15 .11-.20  
## 4 CARB 0.251890359 2116 2 6 .21-.30  
## 5 CCAL 0.029850746 1005 2 14 <.10  
## 6 CHIL 0.020547945 292 1 7 <.10  
## 7 CHIN 1.000000000 378 1 21 1  
## 8 EAFR 0.001298701 10780 1 7 <.10  
## 9 FKLD 0.144927536 69 1 1 .11-.20  
## 10 GFST 0.068020930 3249 6 19 <.10  
## 11 GUIA 0.179571664 607 2 6 .11-.20  
## 12 INDE 0.214285714 28 1 2 .21-.30  
## 13 KURO 0.765625000 64 1 1 .71-.80  
## 14 MEDI 0.267062767 6564 16 70 .21-.30  
## 15 NASE 0.252604167 384 2 29 .21-.30  
## 16 NECS 0.006479657 123309 15 74 <.10  
## 17 NPTG 0.184549356 932 4 37 .11-.20  
## 18 REDS 0.146067416 178 1 26 .11-.20  
## 19 SPSG 0.232535885 1045 3 35 .21-.30  
## 20 SUND 0.231884058 69 1 10 .21-.30  
## 21 TASM 0.002923977 342 1 21 <.10  
## numstudiesbin numfishbin  
## 1 1 < 10  
## 2 2-5 101-500  
## 3 2-5 501-1000  
## 4 2-5 >1500  
## 5 2-5 1001-1500  
## 6 1 101-500  
## 7 1 101-500  
## 8 1 >1500  
## 9 1 51-100  
## 10 6-10 >1500  
## 11 2-5 501-1000  
## 12 1 10-50  
## 13 1 51-100  
## 14 >10 >1500  
## 15 2-5 101-500  
## 16 >10 >1500  
## 17 2-5 501-1000  
## 18 1 101-500  
## 19 2-5 1001-1500  
## 20 1 51-100  
## 21 1 101-500

### Summary table:  
# geographic summary of data  
Geo\_summ\_pt2 <- newdat %>%   
 group\_by(`prov`) %>%   
 summarize(num\_studies = numstudies,  
 num\_sp = numspecies,  
 num\_ind\_studied = numfish)  
  
Geo\_summ\_pt2

## # A tibble: 21 x 4  
## prov num\_studies num\_sp num\_ind\_studied  
## <fct> <int> <int> <int>  
## 1 AUSW 1 2 2  
## 2 BPLR 2 2 228  
## 3 BRAZ 4 15 872  
## 4 CARB 2 6 2116  
## 5 CCAL 2 14 1005  
## 6 CHIL 1 7 292  
## 7 CHIN 1 21 378  
## 8 EAFR 1 7 10780  
## 9 FKLD 1 1 69  
## 10 GFST 6 19 3249  
## # ... with 11 more rows

write.table(Geo\_summ\_pt2, file = "Geo\_summ\_pt2.txt", sep = ",", quote = FALSE, row.names = F)