Price disaggregation protocol: from spatial ICES cells to individual ports.

Sbrana, Sabatella

2025-07-30

Table of Contents

## Introduction

The protocol is part of the SURIMI project and is designed to extract the prices of individual species over time for each landing port.

To implement this methodology effectively, two key datasets are required:

1. The [FDI landing dataset](https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/FAD/fdi/FDI_spatial_data_EU28.zip), which provides detailed records of landings over time by year, country, GSA, gear, vessel length and ices-cell.
2. Dataset containing the number of vessels operating during the same time period, divided by port, type of fishing gear used, and vessel length:

| year | port\_code | gear | vlength | n |
| --- | --- | --- | --- | --- |
| 2013 | ESSCR | DTS | VL1218 | 11 |
| 2014 | ESSCR | DTS | VL1218 | 11 |
| 2015 | ESSCR | DTS | VL1218 | 10 |
| 2016 | ESSCR | DTS | VL1218 | 10 |

1. Ports coordinates:

| port\_code | lon | lat |
| --- | --- | --- |
| ESSCR | 40.582276 | 0.598944 |

1. The [FAO Geographical Sub-Areas](https://gfcmsitestorage.blob.core.windows.net/website/5.Data/ArcGIS/GSAs_simplified_updated_division%20(2).zip): Area of application, comprised of the Mediterranean and the Black Sea, as Major Fishing Area 37.
2. The [FAO ASFIS List of Species for Fisheries](https://www.fao.org/fishery/static/ASFIS/ASFIS_sp.zip) which represents the standard taxonomic reference system for the FAO Statistics Team.

This information is essential to contextualize the landing data and to link fishing effort with species availability and pricing trends at the port level.

## Disaggregation procedure

Save the data input to a folder and set the folder as the data location in the R environment:

wd = "SET YOUR DATA FOLDER DIRECTORY"

### Data manipulation for a case study area

Users could establish parameters for their case study, which will subsequently inform the procedure.

This study examines the activity of bottom otter trawlers (OTB) and purse seiners (PS) within the Spanish waters of GSA 6.

CS\_name = "FAO GSA06 "  
Gear = c("OTB", "PS")  
Year = c(2013:2022)  
Country = c("ESP")  
GSAa\_CS = c("GSA6")  
GSAs\_CS = c("GSA06")

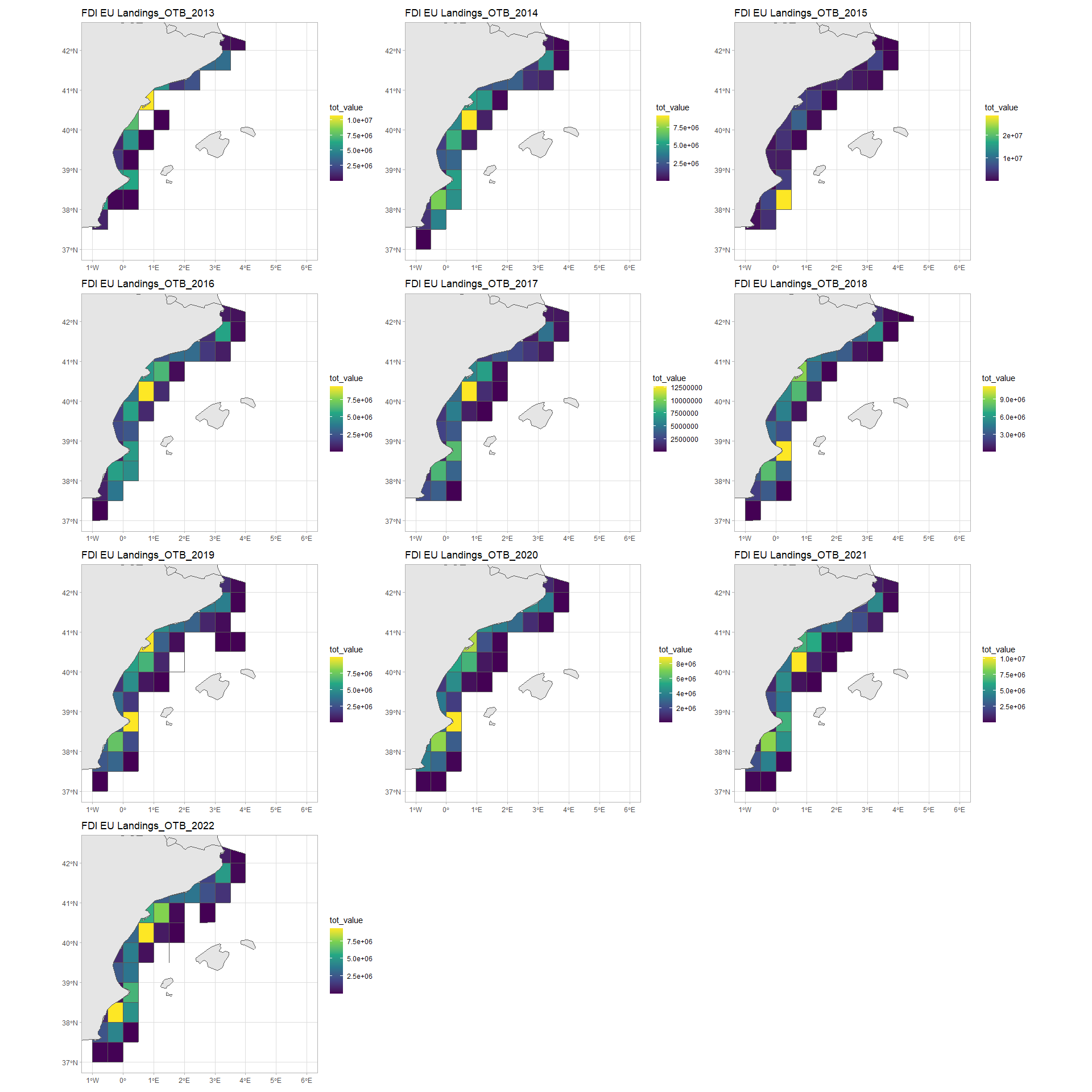
library(dplyr)  
library(ggplot2)  
library(ggrepel)  
library(ggpmisc)  
library(openxlsx)  
library(paletteer)  
library(patchwork)  
library(reshape2)  
library(rnaturalearth)  
library(rnaturalearthdata)  
library(sf)  
library(tidyverse)  
library(tidytext)  
library(webr)  
library(webshot2)

### Step 1 - Open and subset FDI Landing data

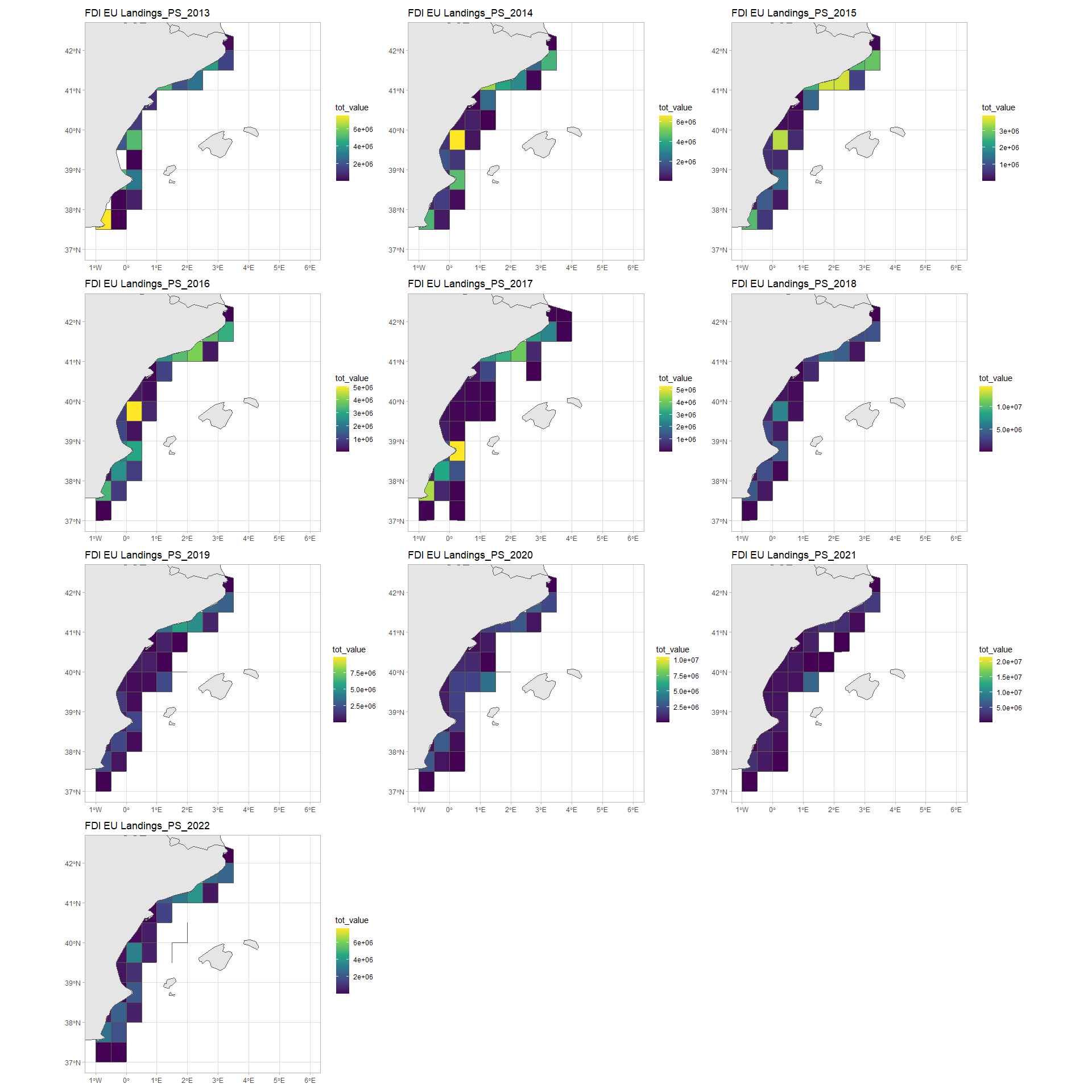
Open and subset data from FDI by: *Gear type*, *Year*, and *Country* The GSA polygon was used to subset spatial landings.

Total landings coverage for the case study area - resulting from FDI data

#Set parameter for the map  
world <- ne\_countries(scale = "medium", returnclass = "sf", continent = "europe")  
world = st\_transform(world, crs = st\_crs(GSA))  
  
xmin = as.numeric(st\_bbox(GSA)[1])-0.0001  
xmax = as.numeric(st\_bbox(GSA)[3])+0.0001  
ymin = as.numeric(st\_bbox(GSA)[2])-0.0001  
ymax = as.numeric(st\_bbox(GSA)[4])+0.0001  
  
  
#landing  
  
for(g in 1 : length(Gear)){  
  
l\_data = purrr::map(landing\_sf, ~ .x %>%  
 filter(gear\_type == Gear[g]))  
   
plot\_list <- list()   
   
 for(y in 1: length(Year)){  
l\_plot = l\_data[[y]] %>%   
 group\_by(year,gear\_type,sub\_region,cscode) %>%   
 summarise(tot\_kg = sum(totwghtlandg), tot\_value = sum(totvallandg)) %>%   
 ggplot()+  
 geom\_sf(aes(fill = tot\_value))+  
 geom\_sf(data = world)+  
 coord\_sf(xlim = c(xmin, xmax), ylim = c(ymin, ymax))+  
 scale\_fill\_viridis\_c(option = "D")+   
 ggtitle(paste0("FDI EU Landings\_",Gear[g],"\_",Year[y]))+  
 theme\_light()  
  
plot\_list[[y]] <- l\_plot  
 }  
 combined\_plot <- wrap\_plots(plotlist = plot\_list, ncol = 3)  
 print(combined\_plot)  
   
 cat("-----")  
 }



## -----



## -----

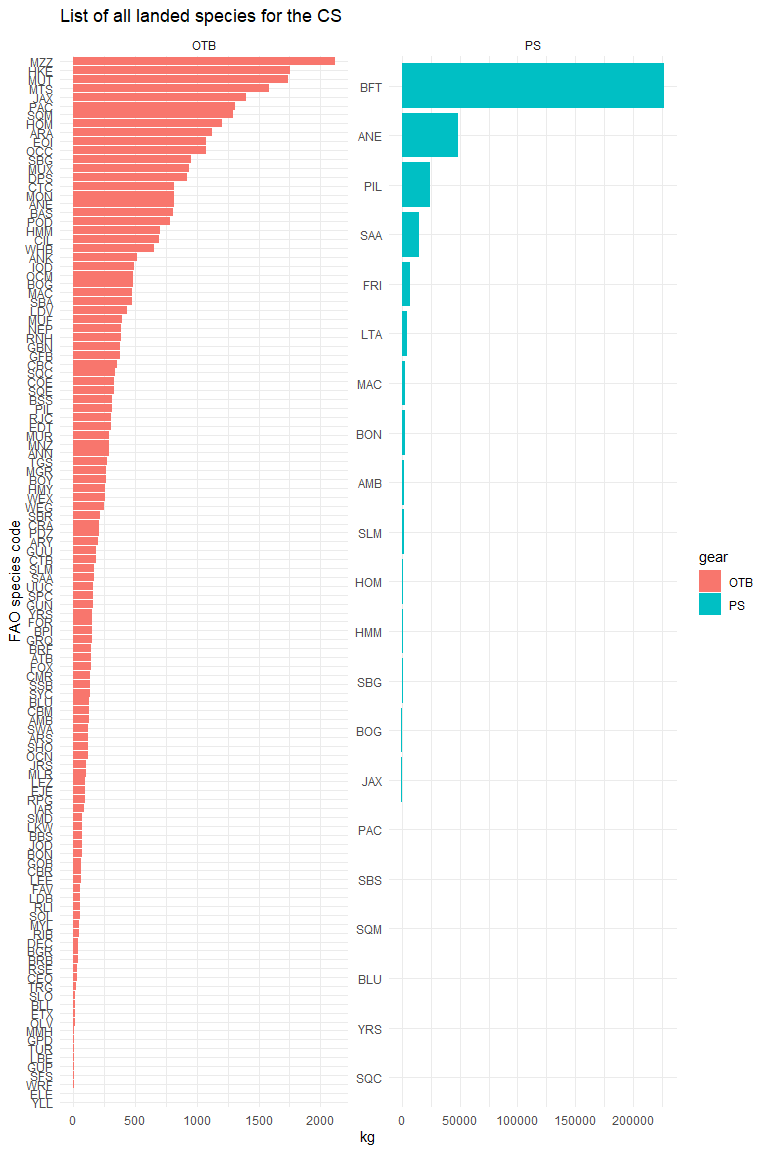
Save all the filtered data in a specific folder fd = “CaseStudy/”

fd = "CaseStudy/"  
  
saveRDS(landing\_sf, paste0(fd,"Landing\_sf.RData"))  
  
saveRDS(  
 purrr::map(landing\_sf, ~ .x %>%   
 st\_drop\_geometry() %>%   
 rename(id = cscode, gear = gear\_type, vlength = vessel\_length, tot\_fish\_weight = totwghtlandg, tot\_fish\_value = totvallandg)),   
 paste0(fd,"landing\_CS.rData"))

### Step 2 - Filter and clean landing data

This R script processes landing data to identify key species consistently landed over time, and calculates average weights, values, and prices. It focuses on analyzing catch data by gear type and vessel length, and select top 15 species by landed weight for each gear.

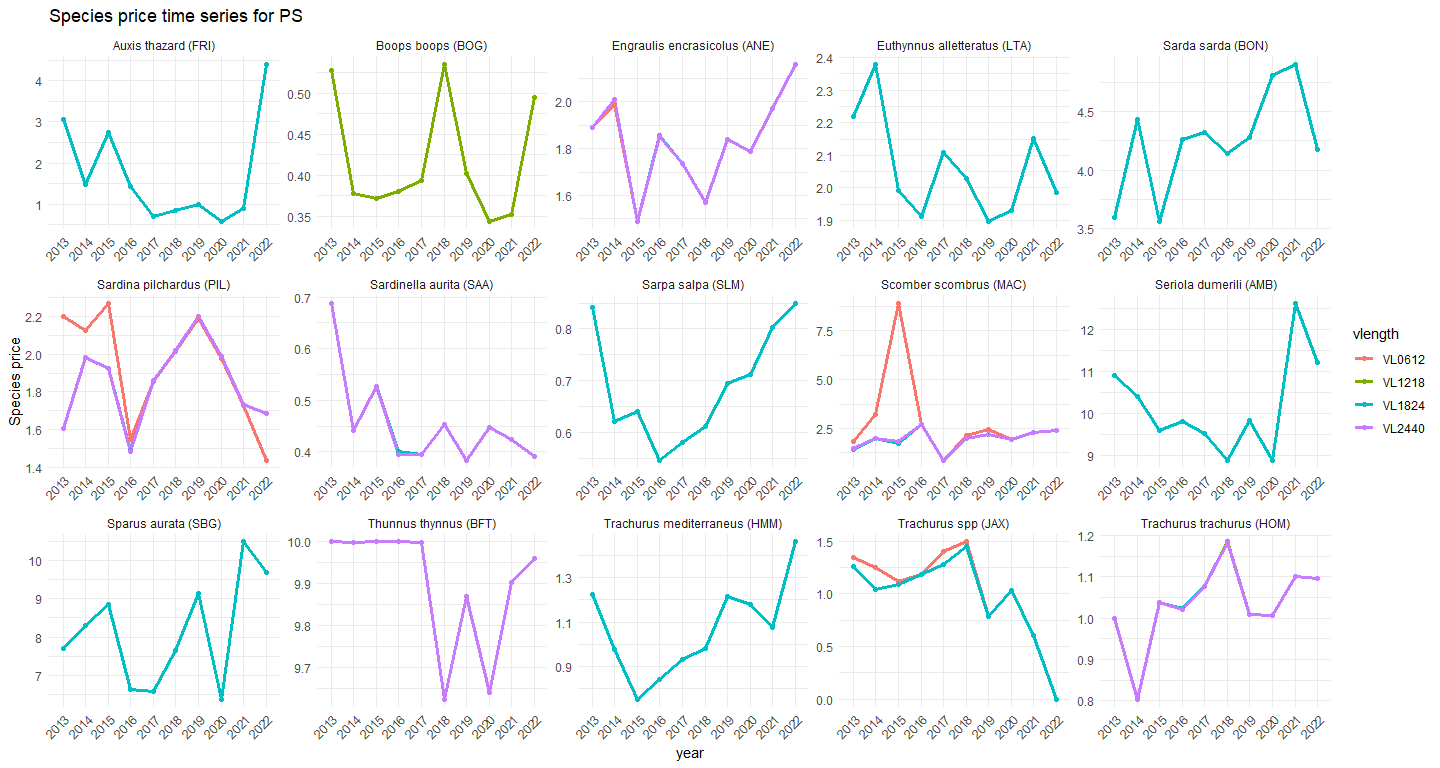
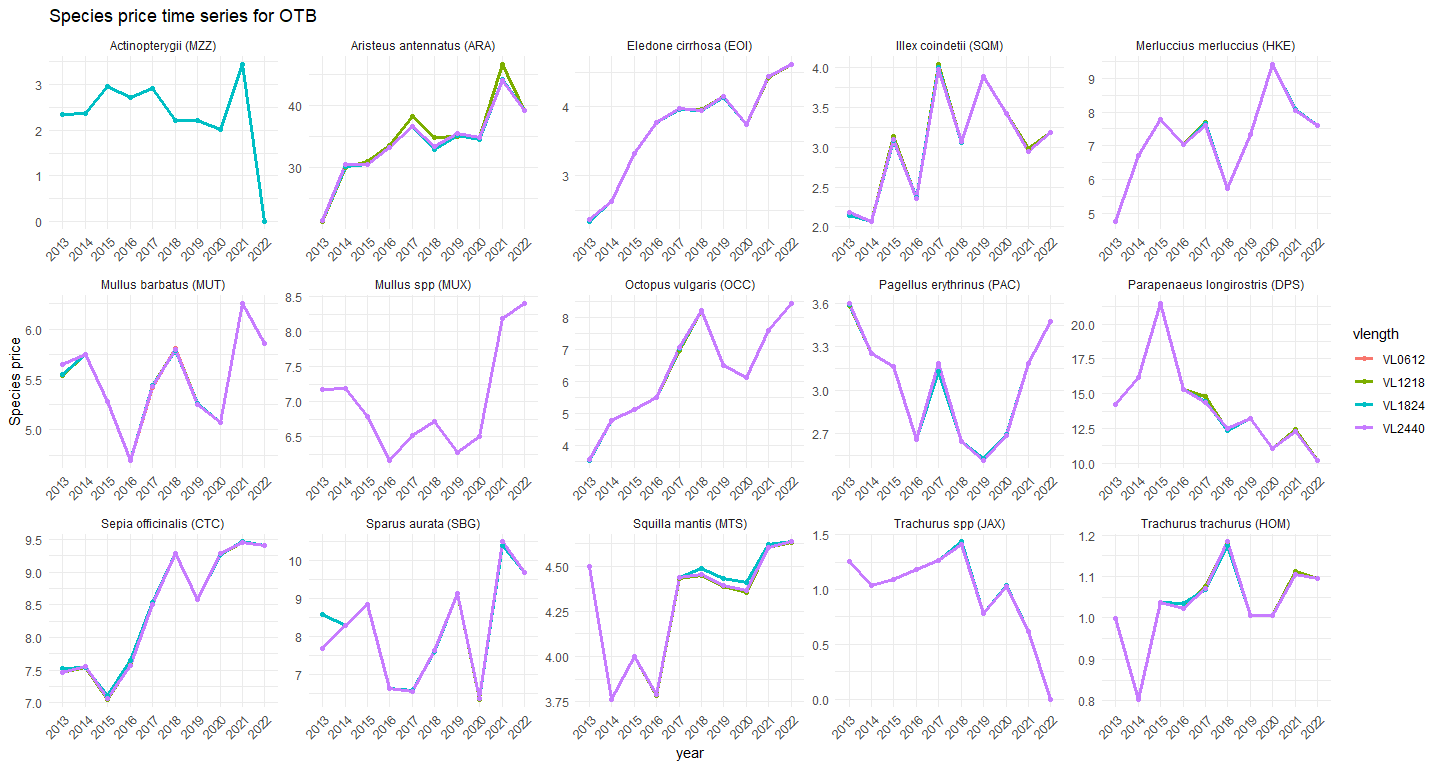
landing\_CS = readRDS(paste0(fd,"Landing\_CS.RData"))  
  
euro\_species = purrr::map(  
 landing\_CS, ~ .x %>%   
 filter(tot\_fish\_weight != 0)   
 )  
  
euro\_species\_df = do.call("rbind", euro\_species)  
  
species\_all\_years <- euro\_species\_df %>%  
 group\_by(gear,vlength,species) %>%   
 distinct(year, species) %>%  
 summarise(n\_years = n\_distinct(year), .groups = "drop") %>%  
 filter(n\_years == length(unique(euro\_species\_df$year)))  
  
euro\_species\_filtered <- euro\_species\_df %>%  
 inner\_join(species\_all\_years, by = c("species", "vlength", "gear"))  
  
# remove quarter  
  
euro\_species\_filtered <- euro\_species\_filtered %>%   
 group\_by(year,vlength, gear, id, species) %>%   
 summarise(yearly\_kg = mean(tot\_fish\_weight)\*1000, yearly\_value = mean(tot\_fish\_value))  
   
euro\_species\_filtered %>%  
 group\_by(species, gear) %>%  
 summarise(mvalue = mean(yearly\_value), kg = (mean(yearly\_kg))) %>%  
 ggplot() +  
 geom\_bar(aes(x = kg, y = reorder\_within(species, kg, gear), fill = gear), stat = "identity")+  
 ylab("FAO species code")+  
 facet\_wrap(~ gear, scales = "free")+  
 xlab("kg")+  
 ggtitle("List of all landed species for the CS")+  
 scale\_y\_reordered()+  
 theme\_minimal()



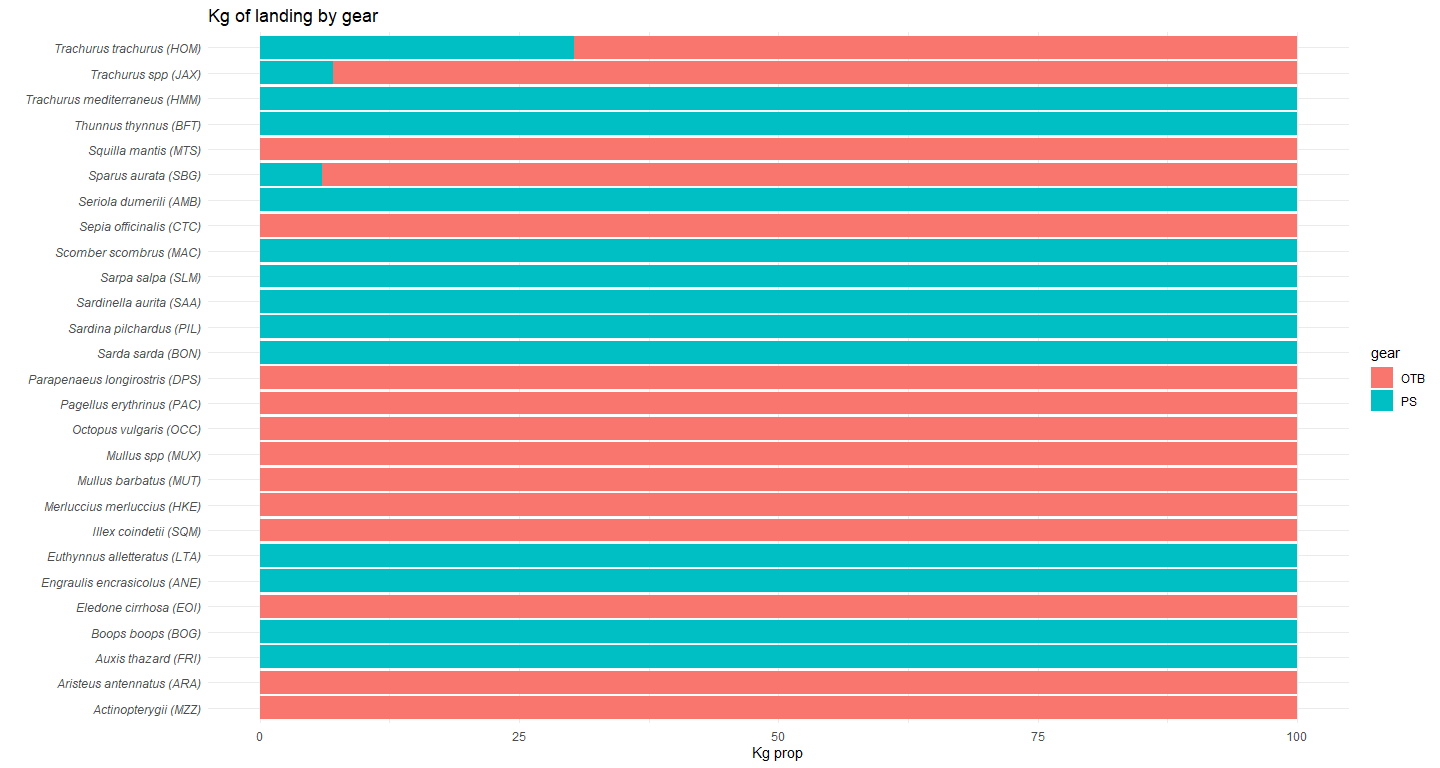
species\_sel <- euro\_species\_filtered %>%  
 group\_by(species, gear) %>%  
 summarise(m\_kg = mean(yearly\_kg), .groups = "drop") %>%  
 group\_by(gear) %>%  
 slice\_max(m\_kg, n = 15) %>%  
 ungroup()   
  
 FDI\_land\_spe\_filter <- euro\_species\_filtered %>%  
 inner\_join(species\_sel, by = c("species", "gear")) %>%   
 select(-m\_kg) %>%   
 mutate(price = yearly\_value/yearly\_kg)

Some results:

FAO\_sp = read.csv(paste0(wd,"ASFIS\_sp\_2025.csv")) %>% select("Alpha3\_Code","Scientific\_Name") %>% rename("species" = "Alpha3\_Code")  
  
FDI\_land\_spe\_filter$year = as.character(FDI\_land\_spe\_filter$year)  
  
gear\_filter = unique(FDI\_land\_spe\_filter$gear)  
  
for (g in seq\_along(gear\_filter)) {  
  
 print(FDI\_land\_spe\_filter %>%   
 filter(gear == gear\_filter[g]) %>%   
 group\_by(year, vlength, species) %>%   
 summarise(m\_price = mean(price)) %>%   
 left\_join(FAO\_sp) %>%  
 mutate(sp\_id = paste0(Scientific\_Name," (",species,")")) %>%   
   
 ggplot(aes(x = year, y = m\_price, color = vlength,  
 group = interaction(species, vlength)  
   
 )) +  
 geom\_line(size = 1.1) +  
 geom\_point() +  
 facet\_wrap(~ sp\_id, scales = "free", ncol = 5) +  
 theme\_minimal() +  
 ggtitle(paste0("Species price time series for ", gear\_filter[g])) +  
 ylab("Species price")+  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))  
)  
}



FDI\_land\_spe\_filter %>%  
 group\_by(gear,species) %>%  
 summarise(kg\_gear = sum(yearly\_kg)) %>%  
 group\_by(species) %>%   
 mutate(kg\_tot = sum(kg\_gear)) %>%   
 mutate(kg\_prop = (kg\_gear/kg\_tot)\*100) %>%  
 left\_join(FAO\_sp) %>%  
 mutate(sp\_id = paste0(Scientific\_Name," (",species,")")) %>%   
 ggplot() +  
 geom\_bar(aes(y = sp\_id, x = kg\_prop, fill = gear),stat = "identity") +  
 labs(title = "Kg of landing by gear", y = "", x = "Kg prop") +  
 theme\_minimal()+  
 theme(axis.text.y = element\_text(face = "italic"))



Save data filtered

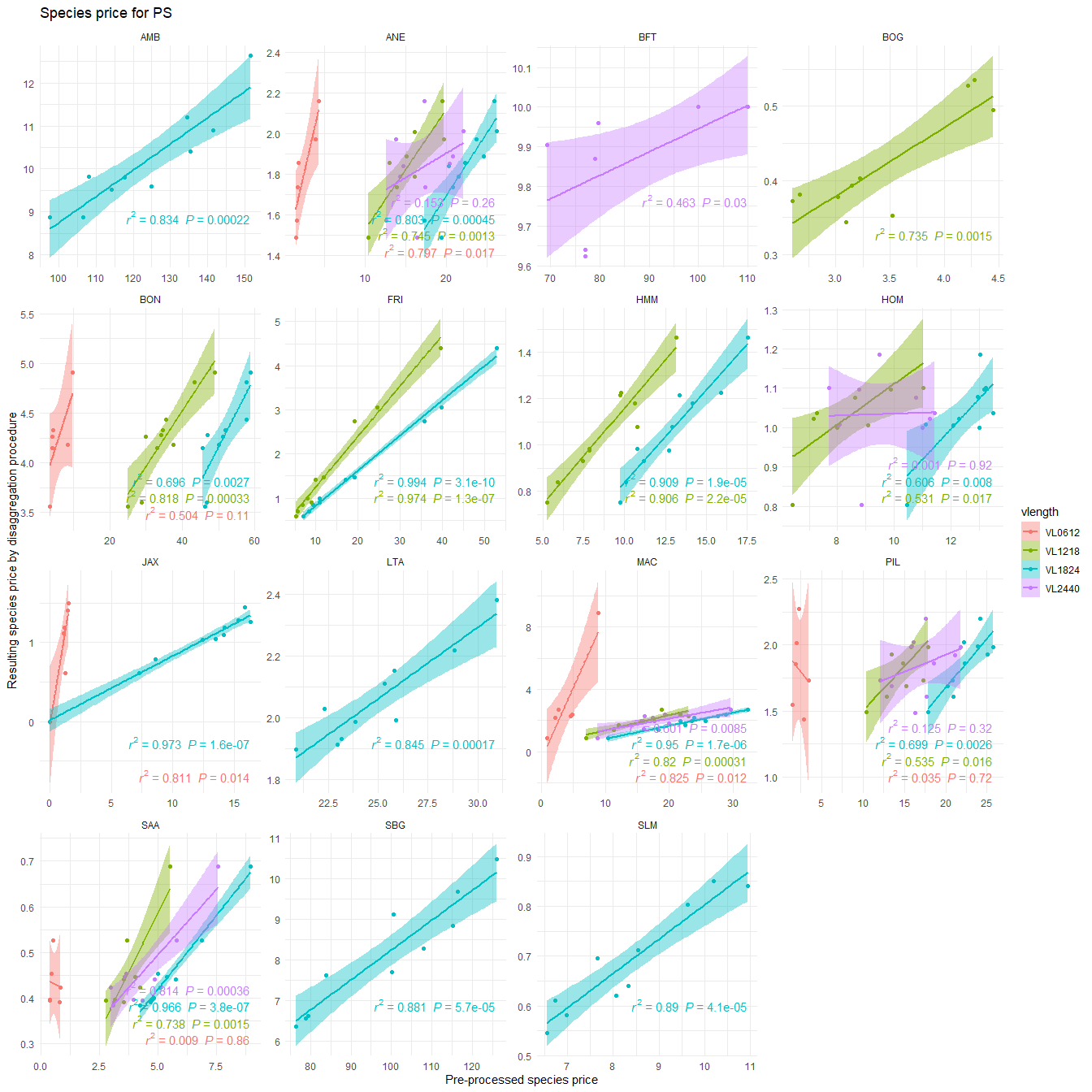
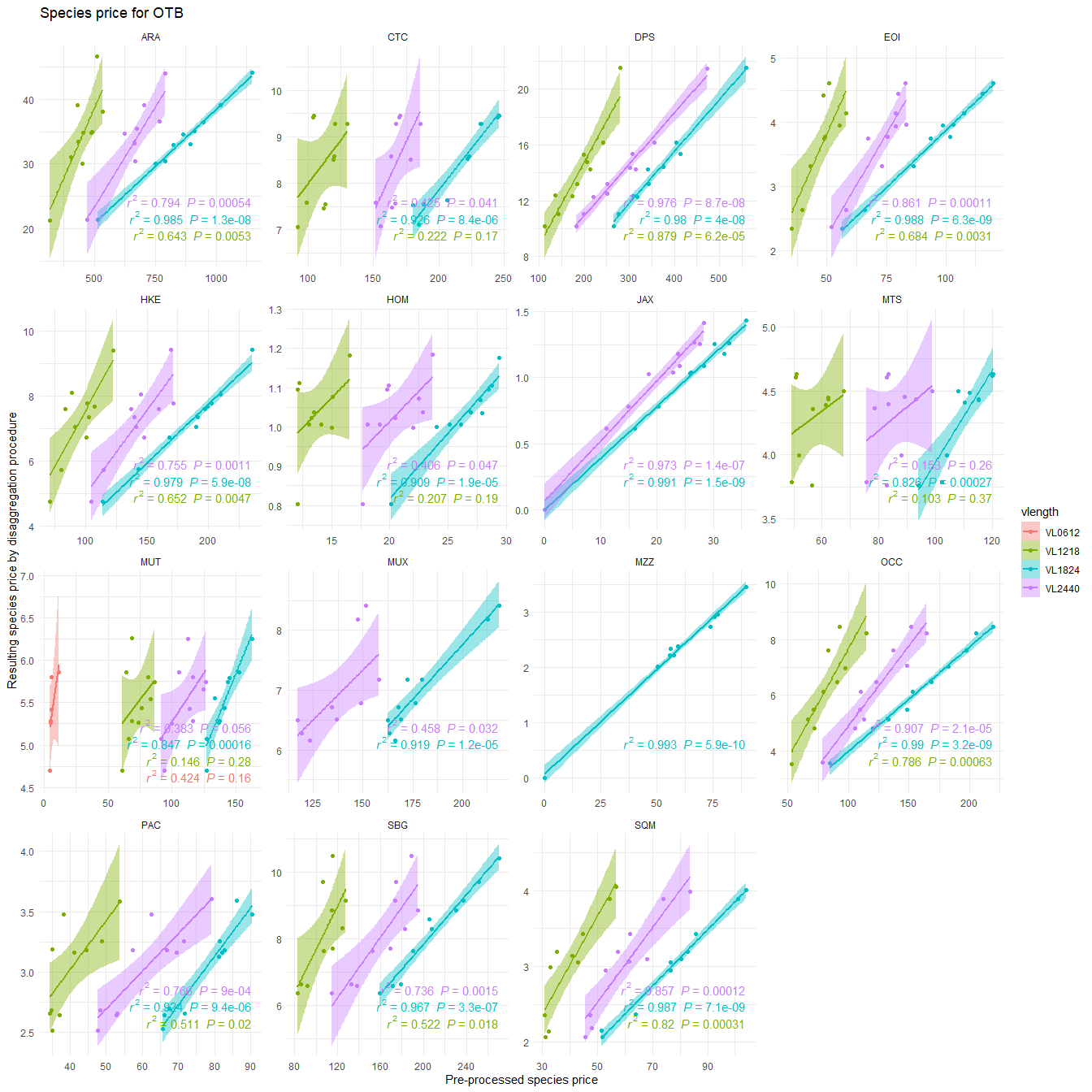
saveRDS(FDI\_land\_spe\_filter, paste0(fd,"FDI\_land\_spe\_filter.RData"))

### Step 3 - Calculate and add price by species

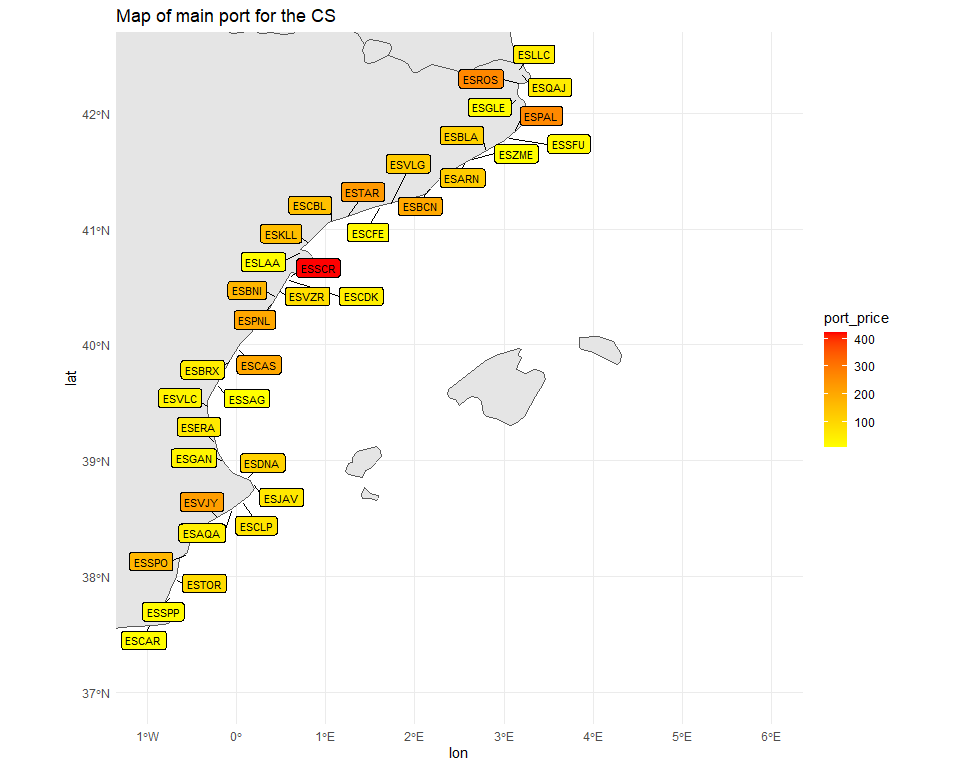
Load dataset containing the number of vessels by port, type of fishing gear used, and vessel length. Load port coordinates.

Here we first chech gear code used and change as FDI code.

nvessel\_port = read.csv("Data Input/num\_vessels\_per\_year\_port\_gear\_length.csv") %>%   
 rename( gear = gear\_code,  
 vlength = vessel\_length) %>%   
 mutate(year = as.character(year))  
  
nvessel\_port$gear[which(nvessel\_port$gear == "DTS")] <- "OTB"  
   
total\_n <- nvessel\_port %>%  
 group\_by(year, gear, vlength) %>%  
 summarise(total\_n = sum(n))  
  
nvessel\_weighted <- nvessel\_port %>%  
 left\_join(total\_n, by = c("year", "gear", "vlength")) %>%  
 mutate(weight = n / total\_n)  
  
FDI\_land\_spe\_filter = readRDS(paste0(fd,"FDI\_land\_spe\_filter.RData")) %>%   
 group\_by(year, gear, vlength, species) %>%   
 summarise(mval = mean(yearly\_value), mkg = mean(yearly\_kg), mprice = mean(price))   
   
FDI\_land\_spe\_nvessel <- FDI\_land\_spe\_filter %>%  
 inner\_join(nvessel\_weighted, by = c("year", "gear", "vlength"))  
  
FDI\_land\_spe\_nvessel <- FDI\_land\_spe\_nvessel %>%  
 mutate(  
 mval\_by\_port = mval \* weight,  
 mkg\_by\_port = mkg \* weight,  
 mprice\_by\_port = mprice \* weight)  
  
df\_port\_species <- FDI\_land\_spe\_nvessel %>%  
 group\_by(year, gear, vlength, port\_code, species) %>%  
 select(mval\_by\_port, mkg\_by\_port, mprice\_by\_port, mval, mkg, mprice)  
  
  
for (g in seq\_along(gear\_filter)) {  
  
 print(df\_port\_species %>%   
 group\_by(year, gear, vlength, species) %>%   
 summarise(mprice\_by\_port\_tot = sum(mprice\_by\_port), mprice = sum(mprice)) %>%   
 filter(gear == gear\_filter[g]) %>%   
  
 ggplot(aes(x = mprice, y = mprice\_by\_port\_tot, fill = vlength, color = vlength))+  
 geom\_point()+  
 geom\_smooth(method = "lm")+  
 facet\_wrap(~ species, scale = "free")+  
 theme\_minimal() +  
 ggtitle(paste0("Species price for ", gear\_filter[g]))+  
 stat\_fit\_glance(method = "lm",  
 label.x = "right",  
 label.y = "bottom",  
 method.args = list(formula = y ~ x),  
 mapping = aes(label = sprintf('italic(r)^2~"="~%.3f~~italic(P)~"="~%.2g',  
 after\_stat(r.squared), after\_stat(p.value))),  
 parse = TRUE,  
 show.legend = FALSE) +  
 ylab("Resulting species price by disaggregation procedure")+  
 xlab("Pre-processed species price")  
 )  
   
 }



port\_coord = read.csv("Data Input/ports.csv") %>%   
 st\_as\_sf(coords = c("lon", "lat"), crs = st\_crs(GSA))  
  
df\_port\_species %>%   
 group\_by(port\_code) %>%   
 summarise(port\_price = sum(mprice\_by\_port), port\_kg = sum(mkg\_by\_port)) %>%  
 left\_join(port\_coord) %>%   
 st\_as\_sf() %>%   
 mutate(lon = st\_coordinates(.)[,1],   
 lat = st\_coordinates(.)[,2]) %>%   
  
 ggplot() +  
 geom\_sf(data = world)+  
 geom\_label\_repel(  
 aes(x = lon, y = lat, label = port\_code, fill = port\_price),  
 size = 3,  
 min.segment.length = 0  
 ) +  
 scale\_fill\_gradient(low = "yellow", high = "red", na.value = NA)+  
 coord\_sf(xlim = c(xmin, xmax), ylim = c(ymin, ymax))+  
 theme\_minimal()+  
 ggtitle("Map of main port for the CS")



df\_port\_species %>%   
 group\_by(species, gear, vlength) %>%   
 summarise(mean\_price = mean(mprice)) %>%   
   
ggplot( aes(y = reorder\_within(species, mean\_price, interaction(gear, vlength)), x = mean\_price, fill = species)) +  
 geom\_bar(stat = "identity", position = "dodge") +  
 facet\_wrap(~ gear + vlength, scales = "free", ncol = 4) +  
 theme\_minimal() +  
 scale\_y\_reordered() +  
 theme(legend.position = "none")+  
 labs(title = "Average price by species and port", x = "Mean price (€)", y = "Species")

