Extract to create a csv file used to feed FGP/OGSL/ETC

Pablo Vergara

January 27 2026

This document is the narrative of how an Excel file called "sGSL-September-RV-FGP.csv" was generated.

The DFO Gulf Region September RV survey follows a stratified random sampling design and covers Division 4T of the Northwest Atlantic Fisheries Organisation (NAFO).

```r
`%nin%` = Negate(`%in%`) ## a useful operator
suppressPackageStartupMessages(library(gulf))
```

SETS

Using the function "read.card" from the DFO Gulf Region's R package "gulf", read in the set cards and exclude hydrographic stations and null sets, so as to keep only representative tows in strata 401 to 403 and 415 to 439.

```r
yrs <- 1971:2025

projectBase = here::here()

x <-  rv.good.sets(year=yrs, source="oracle")
```

```
## Data are being retrieved from Oracle

## Data channel is being created for the PTRAN database

## we want to keep only representative tows, within the historical context that some tows were identifi
uu <- sort(unique(x$experiment))

knitr::kable(data.frame(uu,experiment.str(uu)))## what are the experiment codes
```

uu	experiment.str.uu.
1	stratified random survey set
2	regular survey set (fixed)
5	comparative fishing experiment
8	exploratory fishing
9	hydrography

```
#Need to remove data that is of no use for analysis

## keep only the strata in NAFO 4T,
logical4 <- x$stratum %in% c(401,402,403,415:439)

#Subset set cards with conditions
x <- x[logical4,]
```

Add a few useful columns to the data frame containing the set card information (including depth and swept area as requested).

```
## add useful columns
x[x$vessel.code=="T" & x$year==2003,"vessel.code"] <- "TE" ## CCGS Templeman used in 2003
x$unique.id <- paste(x$year, x$cruise.number, x$vessel.code, x$set.number, sep="-")
x$experiment.str <- experiment.str(x$experiment)
x$vessel.str <- vessel.str(x$vessel.code)
x$gear.str <- gear.str(x$gear)
```

```
## Data channel is being created for the PTRAN database
```

```
x$longitude <- longitude(x)
x$latitude <- latitude(x)
x$mission <- paste(x$vessel.code, x$cruise.number, sep="")

ox <- order(x$year, x$month, x$day, x$start.hour)
x <- x[ox,] # reorder chronologically

## remove missions N176 and H245 which were comparative missions conducted in August 1992
x <- x[which(x$mission !="N176"),]
x <- x[which(x$mission !="H245"),]
```

To deal with "repeat tows" (fishing locations that were purposefully sampled more than once in a given year) an additional column called station.number is added to identify tows that were conducted at the same location within a survey and that should be treated differently when estimating species density. These repeating tows are not independent observations and should be treated accordingly in analyses.

```
x$station.number <- unlist(sapply(yrs, function(y){x.t <- x[x$year==y,]; station.number(x.t, method="ob
```

Catch cards contain the total catch information for the species of interest. Here they are adjusted for distance towed, estimated diurnal effects and estimated vessel-gear effects.

```
### Catch card for all years requested (1970 - 2025)
y <- read.card(card.type="catch", year = yrs)
```

```
## Data channel is being created for the PTRAN database
```

CATCH

```r
y[y$vessel.code=="T" & y$year==2003,"vessel.code"] <- "TE" ## CCGS Templeman used in 2003


y$unique.id <- paste(y$year, y$cruise.number, y$vessel.code, y$set.number, sep="-")

uni.spec = unique(y$species)
columns = c("species","english.name", "french.name", "latin.name")



#Create a Empty DataFrame with length(uni.spec) rows and length(columns) columns
df = data.frame(matrix(nrow = length(uni.spec), ncol = length(columns)))
colnames(df) = columns
df$species = uni.spec
df$english.name <- species.str( uni.spec, "english")
```

Data channel is being created for the PTRAN database

```r
df$latin.name    <- species.str( uni.spec, "latin")
```

Data channel is being created for the PTRAN database

```r
df$french.name  <- species.str( uni.spec, "french")
```

Data channel is being created for the PTRAN database

```r
y <- merge(y, df, by = "species", names = c("english.name","french.name", "latin.name"))
z <- merge(y,x, by = "unique.id", names = c("longitude","latitude", "gear.str"))

##  CSV
## write catch cards to file
fn2 <- "sGSL-September-RV-FGP.csv"
fp  <-paste0(projectBase,"/stock_assessment_surveys/1989de32 (rv_survey)/")

csv.fn2 <- file.path(fp,fn2)

ooz <- order(z$year, z$month, z$day, z$start.hour, z$start.minute, z$species)

#columns to keep
fvars2 <- c("year","month","day","start.hour","start.minute","latitude","longitude","gear.str","species
zz=  z[ooz,fvars2]

#header for FGP must be in both languages
names(zz) <- c("year__annee","month__mois","day__jour","start_hour__heure_de_depart","start_minute__min

#fixes encoding issue
con<-file(csv.fn2,encoding="ISO-8859-1")
write.csv(zz, file=con, row.names=FALSE)
```