DataDeficiency\_analysis

Alex

8/29/2019

### This is an R Markdown file that will show the annotated analysis of our data on factors contributing to data deficiency in all known shark species.

library(xtable); library(MASS); library(tidyverse);library(lme4);library(sjPlot)

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

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

## ── 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()  
## ✖ dplyr::select() masks MASS::select()

## Loading required package: Matrix

##   
## Attaching package: 'Matrix'

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

## Warning in checkMatrixPackageVersion(): Package version inconsistency detected.  
## TMB was built with Matrix version 1.2.15  
## Current Matrix version is 1.2.14  
## Please re-install 'TMB' from source using install.packages('TMB', type = 'source') or ask CRAN for a binary version of 'TMB' matching CRAN's 'Matrix' package

# Read in data  
dd <- read.csv("DD\_analysis\_datafile\_3.csv", stringsAsFactors = FALSE)  
dd = dd[dd$Order != "",]  
  
# Set NAs to 0  
dd[is.na(dd)] = 0  
table(dd$Data.Deficient)

##   
## 0 1   
## 313 188

#head(dd)  
#summary(dd)  
# Rename some columns  
dd = dd %>% rename("Rep\_Strategy" = "X.Reproductive.Strategy",  
 "Vulnerability" = "Vulnerability..",  
 "Antarctic" = "Antartic")  
  
class(dd$Epipelagic)

## [1] "numeric"

#Cleaning up depth data  
# assigning values to the 5 layers of the ocean  
# epipelagic - 0-200 m  
# mesopelagic - 200 - 1000 m  
# bathypelagic - 1,000 - 4,000 m   
# abyssopelagic - 4,000-6,000 m   
  
dd[dd$Epipelagic == ""] <- "0"  
dd$Epipelagic <- factor(dd$Epipelagic, levels = c("0", "1"))  
class(dd$Epipelagic)

## [1] "factor"

dd[dd$Mesopelagic == ""] <- "0"  
dd$Mesopelagic <- factor(dd$Mesopelagic, levels = c("0", "1"))  
class(dd$Mesopelagic)

## [1] "factor"

dd[dd$Bathypelagic == ""] <- "0"  
dd$Bathypelagic <- factor(dd$Bathypelagic, levels = c("0", "1"))  
class(dd$Bathypelagic)

## [1] "factor"

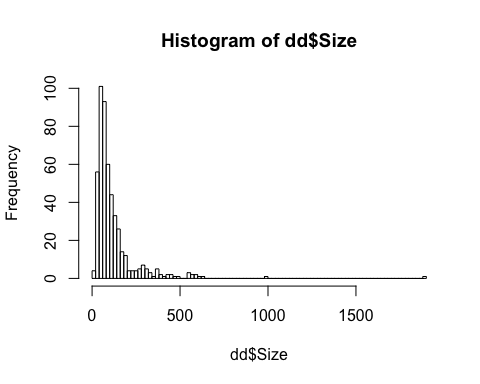
dd[dd$Abyssopelagic == ""] <- "0"  
dd$Abyssopelagic <- factor(dd$Abyssopelagic, levels = c("0", "1"))  
class(dd$Abyssopelagic)

## [1] "factor"

# Cleaning up fisheries data  
dd$Fisheries[dd$Fisheries == "NA"] <- "Unknown"  
dd$Fisheries[dd$Fisheries == ""] <- "Unknown"  
dd$Fisheries <- factor(dd$Fisheries, levels = c("Unknown", "0", "1"))  
  
# Cleaning up reproductive strategy data  
dd$Rep\_Strategy[dd$Rep\_Strategy == ""] = "Unknown"  
dd$Rep\_Strategy = factor(dd$Rep\_Strategy, levels = c("Unknown", "Ovoviviparous", "Viviparous", "Oviparous"))  
  
# Removing columns unused in data  
dd$Benthic <- as.numeric(dd$Benthic) %>% replace\_na(0)

## Warning in eval(lhs, parent, parent): NAs introduced by coercion

dd$Vulnerability <- as.numeric(dd$Vulnerability) %>% replace\_na(0)  
  
## Check normality of size variable; needs to be log transformed  
hist(dd$Size,100)



dd$log\_size<- log(dd$Size) #added a column with log transformed values, which we will use in the GLM  
  
#simple look:   
#Out of 501 shark species assessed: 37.5% are classified as DD, with ~ 4% not evaluated  
sum(dd$Data.Deficient==1, na.rm=T)/nrow(dd) #.375

## [1] 0.3752495

sum(dd$Data.Deficient==0, na.rm=T)/nrow(dd) # .585

## [1] 0.6247505

### After we have our superficial values describing the data, time to do some stats and visualization. What potential predictive factors are we interested in?

*Biology*  
Size  
Reproductive strategy

*Human Use*  
Human use in general (any use)  
Type of use (fisheries, gamefish, aquaculture) Fisheries (any fishery)

*Ecology*  
Geographic range (oceans, trans-oceanic) Latitude/temperature (tropical/temperate)  
Habitat (benthic, deepwater, pelagic, coastal, freshwater)  
Depth range (epipelagic, mesopelagic, bathypelagic, abyssopelagic)

# Potential models (both first- and second-order interactions included)

#First, we will look at the coefficients of every single first order variable and extract p-values to see what we are dealing with  
firstorder.mod <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater +  
 Tropical + Temperate +   
 Pacific + Atlantic + Indian + Arctic + Antarctic + Trans.oceanic +  
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic + Abyssopelagic+  
 log\_size + Rep\_Strategy, data = dd, family = "binomial")  
  
firstorder.mod

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Tropical + Temperate + Pacific +   
## Atlantic + Indian + Arctic + Antarctic + Trans.oceanic +   
## Fisheries + commercial + subsistence + Gamefish + Aquarium +   
## Epipelagic + Mesopelagic + Bathypelagic + Abyssopelagic +   
## log\_size + Rep\_Strategy, family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 2.01426 0.36476   
## Coastal Pelagic   
## -0.31384 -0.16998   
## Benthic BrackishFreshwater   
## 0.65385 -1.50380   
## Tropical Temperate   
## 0.35459 0.07011   
## Pacific Atlantic   
## 0.30989 0.31155   
## Indian Arctic   
## -0.79531 1.98724   
## Antarctic Trans.oceanic   
## NA 0.01491   
## Fisheries1 commercial   
## -0.38310 0.11039   
## subsistence Gamefish   
## -0.01720 -0.47231   
## Aquarium Epipelagic1   
## 0.96691 -0.46342   
## Mesopelagic1 Bathypelagic1   
## -0.25317 -0.80966   
## Abyssopelagic1 log\_size   
## 14.09947 -0.58724   
## Rep\_StrategyOvoviviparous Rep\_StrategyViviparous   
## 0.06466 0.45305   
## Rep\_StrategyOviparous   
## -0.01807   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 475 Residual  
## Null Deviance: 663   
## Residual Deviance: 554.9 AIC: 606.9

# considering removing abyssopelagic - n=1  
firstorder.mod2 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater +  
 Tropical + Temperate +   
 Pacific + Atlantic + Indian + Arctic + Antarctic + Trans.oceanic +  
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +  
 log\_size + Rep\_Strategy, data = dd, family = "binomial")  
  
firstorder.mod2

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Tropical + Temperate + Pacific +   
## Atlantic + Indian + Arctic + Antarctic + Trans.oceanic +   
## Fisheries + commercial + subsistence + Gamefish + Aquarium +   
## Epipelagic + Mesopelagic + Bathypelagic + log\_size + Rep\_Strategy,   
## family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.95925 0.36748   
## Coastal Pelagic   
## -0.29071 -0.17184   
## Benthic BrackishFreshwater   
## 0.66975 -1.51682   
## Tropical Temperate   
## 0.33800 0.06358   
## Pacific Atlantic   
## 0.30143 0.33019   
## Indian Arctic   
## -0.79701 1.94814   
## Antarctic Trans.oceanic   
## NA 0.00991   
## Fisheries1 commercial   
## -0.40738 0.11258   
## subsistence Gamefish   
## -0.00787 -0.48129   
## Aquarium Epipelagic1   
## 0.97677 -0.47704   
## Mesopelagic1 Bathypelagic1   
## -0.25370 -0.77269   
## log\_size Rep\_StrategyOvoviviparous   
## -0.57498 0.08402   
## Rep\_StrategyViviparous Rep\_StrategyOviparous   
## 0.47387 -0.02513   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 476 Residual  
## Null Deviance: 663   
## Residual Deviance: 556.8 AIC: 606.8

### note that we did try to include "Family" as a random variable; however, it was "rank deficient", which means that it was correlated with multiple other variables and therefore redundant. However, code is below  
  
#firstorder.mod.re <- glmer(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater +   
 # Tropical + Temperate +   
 # Global + Pacific + Atlantic + Indian + Arctic + Antarctic + Trans.oceanic +  
 # Fisheries + commercial + subsistence + Gamefish + Aquarium + Vulnerability +   
 #log\_size + Rep\_Strategy +  
 # (1|Order), data = dd, family = "binomial")  
  
  
### Second order model with all potential interactions  
secondorder.mod <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic +   
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian + Arctic + Trans.oceanic + Trans.oceanic\*Pacific + Trans.oceanic\*Atlantic + Trans.oceanic\*Indian + Trans.oceanic\*Arctic +  
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic + Abyssopelagic+  
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Trans.oceanic\*Fisheries, data = dd, family = "binomial")  
  
secondorder.mod

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Tropical + Temperate + Tropical \* Temperate + Pacific +   
## Atlantic + Indian + Arctic + Trans.oceanic + Trans.oceanic \*   
## Pacific + Trans.oceanic \* Atlantic + Trans.oceanic \* Indian +   
## Trans.oceanic \* Arctic + Fisheries + commercial + subsistence +   
## Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
## Abyssopelagic + log\_size + Rep\_Strategy + Coastal \* Fisheries +   
## Pelagic \* Fisheries + Benthic \* Fisheries + Trans.oceanic \*   
## Fisheries, family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.62416 0.33125   
## Coastal Pelagic   
## 0.97044 -0.26737   
## Benthic BrackishFreshwater   
## 1.69196 -1.57920   
## Tropical Temperate   
## 0.97310 0.71100   
## Pacific Atlantic   
## 0.12594 0.14524   
## Indian Arctic   
## -1.00555 1.98539   
## Trans.oceanic Fisheries1   
## -25.85270 0.33731   
## commercial subsistence   
## 0.12718 -0.22087   
## Gamefish Aquarium   
## -0.44971 1.06295   
## Epipelagic1 Mesopelagic1   
## -0.46649 -0.15508   
## Bathypelagic1 Abyssopelagic1   
## -0.84728 16.33379   
## log\_size Rep\_StrategyOvoviviparous   
## -0.74458 0.01359   
## Rep\_StrategyViviparous Rep\_StrategyOviparous   
## 0.50669 -0.03093   
## Coastal:Benthic Deepwater:Benthic   
## -1.56859 -0.34225   
## Tropical:Temperate Pacific:Trans.oceanic   
## -0.88272 13.56734   
## Atlantic:Trans.oceanic Indian:Trans.oceanic   
## 13.02195 -0.83433   
## Arctic:Trans.oceanic Coastal:Fisheries1   
## NA -1.04445   
## Pelagic:Fisheries1 Benthic:Fisheries1   
## -0.15849 0.12491   
## Trans.oceanic:Fisheries1   
## 0.72373   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 465 Residual  
## Null Deviance: 663   
## Residual Deviance: 535.2 AIC: 607.2

#removing trans-oceanic as a variable  
secondorder.mod2 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic +   
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian + Arctic +   
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic + Abyssopelagic+  
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Fisheries, data = dd, family = "binomial")  
  
secondorder.mod2

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Tropical + Temperate + Tropical \* Temperate + Pacific +   
## Atlantic + Indian + Arctic + Fisheries + commercial + subsistence +   
## Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
## Abyssopelagic + log\_size + Rep\_Strategy + Coastal \* Fisheries +   
## Pelagic \* Fisheries + Benthic \* Fisheries + Fisheries, family = "binomial",   
## data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.43548 0.30136   
## Coastal Pelagic   
## 0.98862 -0.23499   
## Benthic BrackishFreshwater   
## 1.69212 -1.62311   
## Tropical Temperate   
## 0.95723 0.69202   
## Pacific Atlantic   
## 0.25373 0.30266   
## Indian Arctic   
## -0.90967 1.84471   
## Fisheries1 commercial   
## 0.54085 0.14035   
## subsistence Gamefish   
## -0.26366 -0.39664   
## Aquarium Epipelagic1   
## 1.10634 -0.45815   
## Mesopelagic1 Bathypelagic1   
## -0.17786 -0.81097   
## Abyssopelagic1 log\_size   
## 14.28824 -0.72779   
## Rep\_StrategyOvoviviparous Rep\_StrategyViviparous   
## -0.01900 0.41247   
## Rep\_StrategyOviparous Coastal:Benthic   
## -0.07415 -1.56503   
## Deepwater:Benthic Tropical:Temperate   
## -0.29480 -0.80215   
## Coastal:Fisheries1 Pelagic:Fisheries1   
## -1.20770 -0.21423   
## Benthic:Fisheries1   
## 0.05544   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 470 Residual  
## Null Deviance: 663   
## Residual Deviance: 537.8 AIC: 599.8

#removing abyssopelagic as a variable, including first order variable transoceanic  
secondorder.mod3 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic + Trans.oceanic+  
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian + Arctic +   
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Fisheries, data = dd, family = "binomial")  
  
secondorder.mod3

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Trans.oceanic + Tropical + Temperate + Tropical \*   
## Temperate + Pacific + Atlantic + Indian + Arctic + Fisheries +   
## commercial + subsistence + Gamefish + Aquarium + Epipelagic +   
## Mesopelagic + Bathypelagic + log\_size + Rep\_Strategy + Coastal \*   
## Fisheries + Pelagic \* Fisheries + Benthic \* Fisheries + Fisheries,   
## family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.442089 0.304890   
## Coastal Pelagic   
## 1.021824 -0.228864   
## Benthic BrackishFreshwater   
## 1.714197 -1.653228   
## Trans.oceanic Tropical   
## 0.244396 0.943214   
## Temperate Pacific   
## 0.695713 0.186597   
## Atlantic Indian   
## 0.244279 -0.976869   
## Arctic Fisheries1   
## 1.835066 0.507676   
## commercial subsistence   
## 0.140497 -0.241929   
## Gamefish Aquarium   
## -0.408926 1.119525   
## Epipelagic1 Mesopelagic1   
## -0.472627 -0.186623   
## Bathypelagic1 log\_size   
## -0.772855 -0.720038   
## Rep\_StrategyOvoviviparous Rep\_StrategyViviparous   
## 0.006118 0.460724   
## Rep\_StrategyOviparous Coastal:Benthic   
## -0.075654 -1.551644   
## Deepwater:Benthic Tropical:Temperate   
## -0.282970 -0.830103   
## Coastal:Fisheries1 Pelagic:Fisheries1   
## -1.199796 -0.225219   
## Benthic:Fisheries1   
## 0.045595   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 470 Residual  
## Null Deviance: 663   
## Residual Deviance: 539.8 AIC: 601.8

#keeping abyssopelagic as a variable, including first order variable transoceanic  
secondorder.mod4 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic + Trans.oceanic+  
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian + Arctic +   
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic + Abyssopelagic +  
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Fisheries, data = dd, family = "binomial")  
  
secondorder.mod4

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Trans.oceanic + Tropical + Temperate + Tropical \*   
## Temperate + Pacific + Atlantic + Indian + Arctic + Fisheries +   
## commercial + subsistence + Gamefish + Aquarium + Epipelagic +   
## Mesopelagic + Bathypelagic + Abyssopelagic + log\_size + Rep\_Strategy +   
## Coastal \* Fisheries + Pelagic \* Fisheries + Benthic \* Fisheries +   
## Fisheries, family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.49661 0.31032   
## Coastal Pelagic   
## 1.02035 -0.23882   
## Benthic BrackishFreshwater   
## 1.72145 -1.63940   
## Trans.oceanic Tropical   
## 0.25760 0.95637   
## Temperate Pacific   
## 0.69733 0.19609   
## Atlantic Indian   
## 0.22364 -0.97482   
## Arctic Fisheries1   
## 1.88841 0.51407   
## commercial subsistence   
## 0.13472 -0.25590   
## Gamefish Aquarium   
## -0.39732 1.10982   
## Epipelagic1 Mesopelagic1   
## -0.45726 -0.18345   
## Bathypelagic1 Abyssopelagic1   
## -0.81579 14.30001   
## log\_size Rep\_StrategyOvoviviparous   
## -0.73487 -0.01668   
## Rep\_StrategyViviparous Rep\_StrategyOviparous   
## 0.43769 -0.06552   
## Coastal:Benthic Deepwater:Benthic   
## -1.59096 -0.29820   
## Tropical:Temperate Coastal:Fisheries1   
## -0.82281 -1.18696   
## Pelagic:Fisheries1 Benthic:Fisheries1   
## -0.20323 0.06250   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 469 Residual  
## Null Deviance: 663   
## Residual Deviance: 537.7 AIC: 601.7

#removing abyssopelagic and transoceanic as a variable  
secondorder.mod5 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic +  
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian + Arctic +   
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Fisheries, data = dd, family = "binomial")  
  
secondorder.mod5

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Tropical + Temperate + Tropical \* Temperate + Pacific +   
## Atlantic + Indian + Arctic + Fisheries + commercial + subsistence +   
## Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
## log\_size + Rep\_Strategy + Coastal \* Fisheries + Pelagic \*   
## Fisheries + Benthic \* Fisheries + Fisheries, family = "binomial",   
## data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.383883 0.296403   
## Coastal Pelagic   
## 0.991663 -0.225270   
## Benthic BrackishFreshwater   
## 1.686177 -1.638300   
## Tropical Temperate   
## 0.944115 0.690626   
## Pacific Atlantic   
## 0.241312 0.319154   
## Indian Arctic   
## -0.914958 1.793604   
## Fisheries1 commercial   
## 0.532970 0.145835   
## subsistence Gamefish   
## -0.249369 -0.408247   
## Aquarium Epipelagic1   
## 1.116002 -0.473415   
## Mesopelagic1 Bathypelagic1   
## -0.181329 -0.768531   
## log\_size Rep\_StrategyOvoviviparous   
## -0.713234 0.003757   
## Rep\_StrategyViviparous Rep\_StrategyOviparous   
## 0.436582 -0.083734   
## Coastal:Benthic Deepwater:Benthic   
## -1.527132 -0.279674   
## Tropical:Temperate Coastal:Fisheries1   
## -0.810423 -1.219435   
## Pelagic:Fisheries1 Benthic:Fisheries1   
## -0.235337 0.039156   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 471 Residual  
## Null Deviance: 663   
## Residual Deviance: 540 AIC: 600

### removing arctic, abyssopelagic, transoceanic  
secondorder.mod6 <- glm(Data.Deficient ~ Deepwater + Coastal + Pelagic + Benthic + BrackishFreshwater + Coastal\*Benthic + Deepwater\*Benthic + Deepwater\*Fisheries +  
 Tropical + Temperate +   
 Tropical\*Temperate + Pacific + Atlantic + Indian +   
 Fisheries + commercial + subsistence + Gamefish + Aquarium + Epipelagic + Mesopelagic + Bathypelagic +   
 log\_size + Rep\_Strategy + Coastal\*Fisheries + Pelagic\*Fisheries + Benthic\*Fisheries + Fisheries, data = dd, family = "binomial")  
  
secondorder.mod6

##   
## Call: glm(formula = Data.Deficient ~ Deepwater + Coastal + Pelagic +   
## Benthic + BrackishFreshwater + Coastal \* Benthic + Deepwater \*   
## Benthic + Deepwater \* Fisheries + Tropical + Temperate +   
## Tropical \* Temperate + Pacific + Atlantic + Indian + Fisheries +   
## commercial + subsistence + Gamefish + Aquarium + Epipelagic +   
## Mesopelagic + Bathypelagic + log\_size + Rep\_Strategy + Coastal \*   
## Fisheries + Pelagic \* Fisheries + Benthic \* Fisheries + Fisheries,   
## family = "binomial", data = dd)  
##   
## Coefficients:  
## (Intercept) Deepwater   
## 1.314401 0.160176   
## Coastal Pelagic   
## 0.915408 -0.355342   
## Benthic BrackishFreshwater   
## 1.701033 -1.592057   
## Fisheries1 Tropical   
## 0.063759 0.999919   
## Temperate Pacific   
## 0.744911 0.185000   
## Atlantic Indian   
## 0.300599 -0.977038   
## commercial subsistence   
## 0.125116 -0.223500   
## Gamefish Aquarium   
## -0.416465 1.140461   
## Epipelagic1 Mesopelagic1   
## -0.505385 -0.216696   
## Bathypelagic1 log\_size   
## -0.809479 -0.650059   
## Rep\_StrategyOvoviviparous Rep\_StrategyViviparous   
## 0.032899 0.537690   
## Rep\_StrategyOviparous Coastal:Benthic   
## -0.007323 -1.512872   
## Deepwater:Benthic Deepwater:Fisheries1   
## -0.361663 0.727963   
## Tropical:Temperate Coastal:Fisheries1   
## -0.901199 -1.059778   
## Pelagic:Fisheries1 Benthic:Fisheries1   
## 0.031379 -0.035568   
##   
## Degrees of Freedom: 500 Total (i.e. Null); 471 Residual  
## Null Deviance: 663   
## Residual Deviance: 539.2 AIC: 599.2

### Now to visualize a coefficient plots to see what is significant!

# # Pull out plot data  
# coefs <- plot\_model(firstorder.mod, transform = NULL)  
# # Making custom figure  
# coefs$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# # Color by sign of change, transparency by p-value  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# # Add CIs around point  
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# # Adding labelled y axis  
# scale\_y\_continuous(labels = as.character(coefs$data$term),  
# breaks = seq(1, 25, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables") + ggtitle("First order model")  
#   
# coefs1.2 <- plot\_model(firstorder.mod2, transform = NULL)  
#   
# # Making custom figure  
# coefs1.2$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs1.2$data$term),  
# breaks = seq(1, 24, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables") + ggtitle("First order model w/o Abyssopelagic")  
#   
#   
# ############# REPEAT FOR SECOND ORDER MODEL #############  
#   
# # Pull out plot data  
# coefs2 <- plot\_model(secondorder.mod, transform = NULL)   
#   
# # Making custom figure  
# coefs2$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2$data$term),  
# breaks = seq(1, 35, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables") + ggtitle("Second Order Model - w/Transoceanic")  
#   
# ###### without trans-oceanic ######  
#   
# # Pull out plot data  
# coefs2.2 <- plot\_model(secondorder.mod2, transform = NULL)   
# coefs2.2$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2.2$data$term),  
# breaks = seq(1, 30, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables")+ ggtitle("Second Order Model - w/o Transoceanic")  
#   
# ##### Remove abyssopelagic, keep transoceanic first order  
# # Pull out plot data  
# coefs2.3 <- plot\_model(secondorder.mod3, transform = NULL)   
# # Making custom figure  
# coefs2.3$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2.3$data$term),  
# breaks = seq(1, 30, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables")+ ggtitle("Second Order Model - no abyssopelagic + 1st order transoceanic")  
#   
# ##### Keep abyssopelagic, keep transoceanic first order  
# # Pull out plot data  
# coefs2.4 <- plot\_model(secondorder.mod4, transform = NULL)   
# # Making custom figure  
# coefs2.4$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2.4$data$term),  
# breaks = seq(1, 31, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables")+ ggtitle("Second Order Model - w/ abyssopelagic + 1st order transoceanic")  
#   
# ##### Remove abyssopelagic, remove transoceanic  
# # Pull out plot data  
# coefs2.5 <- plot\_model(secondorder.mod5, transform = NULL)   
# # Making custom figure  
# coefs2.5$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2.5$data$term),  
# breaks = seq(1, 29, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables")+ ggtitle("Second Order Model - no abyssopelagic or TO")  
#   
# ##### Remove abyssopelagic, remove transoceanic and Arctic  
# # Pull out plot data  
# coefs2.6 <- plot\_model(secondorder.mod6, transform = NULL)   
# # Making custom figure  
# coefs2.6$data %>%  
# mutate(term = factor(term, levels = term)) %>%  
# ggplot(aes(x = estimate,  
# y = as.numeric(term),  
# color = as.factor(sign(estimate)),  
# alpha = as.factor(p.value < 0.05))) +  
# geom\_point() +   
# geom\_segment(aes(x = conf.low,  
# xend = conf.high,  
# yend = as.numeric(term))) +  
# geom\_vline(xintercept = 0) +  
# scale\_y\_continuous(labels = as.character(coefs2.6$data$term),  
# breaks = seq(1, 28, 1)) +  
# scale\_alpha\_discrete(range = c(.5, 1)) +  
# xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
# ylab("Variables")+ ggtitle("Second Order Model - remove TO and all n<3")  
#

# Model selection

ic <- data.frame(Model = c("secondorder.mod", "secondorder.mod2", "secondorder.mod3", "secondorder.mod4", "firstorder.mod", "firstorder.mod2", "secondorder.mod5", "secondorder.mod6"),  
AIC = c(AIC(secondorder.mod), AIC(secondorder.mod4), AIC(secondorder.mod2), AIC(secondorder.mod3), AIC(firstorder.mod), AIC(firstorder.mod2), AIC(secondorder.mod5), AIC(secondorder.mod6)),  
stringsAsFactors = FALSE  
)  
  
ic <- ic[order(ic$AIC),]  
ic

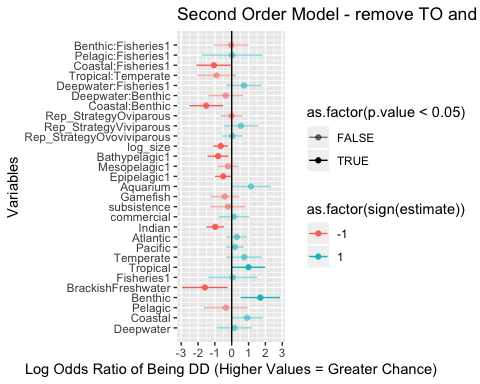
## Model AIC  
## 8 secondorder.mod6 599.2491  
## 3 secondorder.mod3 599.8086  
## 7 secondorder.mod5 599.9532  
## 2 secondorder.mod2 601.6607  
## 4 secondorder.mod4 601.8200  
## 6 firstorder.mod2 606.8285  
## 5 firstorder.mod 606.9077  
## 1 secondorder.mod 607.2367

# AIC selection

*Generally seeing three top models, all second order -*  
1) no abyssopelagic, first order transoceanic (note that transoceanic is redundant, so not using this one)  
2) no abyssopelagic or transoceanic  
**3) no abyssopelagic, Arctic, or transoceanic**  
*only going to use variables where we have n>3*

##### Chosen model  
coefs2.6 <- plot\_model(secondorder.mod6, transform = NULL)   
coefs2.6$data %>%  
 mutate(term = factor(term, levels = term)) %>%  
 ggplot(aes(x = estimate,  
 y = as.numeric(term),  
 color = as.factor(sign(estimate)),  
 alpha = as.factor(p.value < 0.05))) +  
 geom\_point() +   
 geom\_segment(aes(x = conf.low,  
 xend = conf.high,  
 yend = as.numeric(term))) +  
 geom\_vline(xintercept = 0) +  
 scale\_y\_continuous(labels = as.character(coefs2.6$data$term),  
 breaks = seq(1, 29, 1)) +  
 scale\_alpha\_discrete(range = c(.5, 1)) +  
 xlab("Log Odds Ratio of Being DD (Higher Values = Greater Chance)") +  
 ylab("Variables")+ ggtitle("Second Order Model - remove TO and all n<3")

## Warning: Using alpha for a discrete variable is not advised.



### Let’s visualize some of our our significant results

Start with making a table of the emmeans, which give us our predicted, backtransformed values of probability of DD based on the glm

## subset coefficients by category  
fullcoeffs <- summary(secondorder.mod3)$coefficients %>% as.data.frame(.) %>% rownames\_to\_column()  
coef\_conf <- confint(secondorder.mod3) %>% as.data.frame(.) %>% rownames\_to\_column() #error here

## Waiting for profiling to be done...

coef\_table <- left\_join(fullcoeffs, coef\_conf) %>%  
 rename("Coef" = "rowname")

## Joining, by = "rowname"

########### Generating Estimated Marginal Mean ##########  
# estimates the average value for that combination of parameters, and predicts the resulting mean for a set of specified values #  
  
library(emmeans)  
  
# size  
logsize\_means <- emmeans(secondorder.mod6, ~ log\_size,   
 var = "log\_size",   
 at = list(log\_size = seq(2, 8, by = .25)),  
 type = "response")  
  
#habitat  
Brackish\_means <- emmeans(secondorder.mod6, ~ BrackishFreshwater,   
 var = "BrackishFreshwater",   
 at = list(BrackishFreshwater = c(0,1)),  
 type = "response")  
  
benthicfish\_means <- emmeans(secondorder.mod6, ~ Benthic\*Fisheries,   
 var = "Benthic\*Fisheries",   
 at = list(Benthic = c(0,1),   
 Fisheries = factor(c(0,1))),  
 type = "response") #giving probability that the animal is data deficient if benthic = 0,1 and fished  
  
pelagicfish\_means <- emmeans(secondorder.mod6, ~ Pelagic\*Fisheries,   
 var = "Pelagic\*Fisheries",   
 at = list(Pelagic = c(0,1),  
 Fisheries = factor(c(0,1))),  
 type = "response")  
  
coastalfish\_means <- emmeans(secondorder.mod6, ~ Coastal\*Fisheries,   
 var = "Coastal\*Fisheries",   
 at = list(Coastal = c(0,1)),  
 Fisheries = factor(c(0,1)),  
 type = "response")  
  
coastalbenthic\_means <- emmeans(secondorder.mod6, ~ Coastal\*Benthic,   
 var = "Coastal\*Benthic",   
 at = list(Coastal = c(0,1),  
 Benthic = c(0,1)),  
 type = "response")  
  
dwbenthic\_means <- emmeans(secondorder.mod6, ~ Deepwater\*Benthic,   
 var = "Deepwater\*Benthic",   
 at = list(Deepwater = c(0,1),  
 Benthic = c(0,1)),  
 type = "response")  
  
dwfish\_means <- emmeans(secondorder.mod6, ~ Deepwater\*Fisheries,   
 var = "Deepwater\*Fisheries",   
 at = list(Deepwater = c(0,1),  
 Benthic = c(0,1)),  
 type = "response")  
  
#oceans  
Atlantic\_means <- emmeans(secondorder.mod6, ~ Atlantic,   
 var = "Atlantic",   
 at = list(Atlantic = c(0,1)),  
 type = "response")  
  
Pacific\_means <- emmeans(secondorder.mod6, ~ Pacific,   
 var = "Pacific",   
 at = list(Pacific = c(0,1)),  
 type = "response")  
  
Indian\_means <- emmeans(secondorder.mod6, ~ Indian,   
 var = "Indian",   
 at = list(Indian = c(0,1)),  
 type = "response")  
  
# depth  
bathy\_means <- emmeans(secondorder.mod6, ~ Bathypelagic,   
 var = "Bathypelagic",   
 at = list(Bathypelagic = factor(c(0,1))),  
 type = "response")  
  
meso\_means <- emmeans(secondorder.mod6, ~ Mesopelagic,   
 var = "Mesopelagic",   
 at = list(Mesopelagic = factor(c(0,1))),  
 type = "response")  
  
epipelagic\_means <- emmeans(secondorder.mod6, ~ Epipelagic,   
 var = "Epipelagic",   
 at = list(Epipelagic = factor(c(0,1))),  
 type = "response")  
  
  
  
logsize\_means

## log\_size prob SE df asymp.LCL asymp.UCL  
## 2.00 0.6674290 0.12901343 Inf 0.3911208 0.8624468  
## 2.25 0.6304307 0.12298826 Inf 0.3774116 0.8275952  
## 2.50 0.5918335 0.11480119 Inf 0.3635716 0.7863388  
## 2.75 0.5520694 0.10474634 Inf 0.3495167 0.7387024  
## 3.00 0.5116281 0.09327706 Inf 0.3351121 0.6852910  
## 3.25 0.4710340 0.08099580 Inf 0.3201402 0.6274143  
## 3.50 0.4308195 0.06864131 Inf 0.3042482 0.5671255  
## 3.75 0.3914972 0.05708172 Inf 0.2868667 0.5071507  
## 4.00 0.3535347 0.04730467 Inf 0.2671346 0.4506943  
## 4.25 0.3173337 0.04031882 Inf 0.2440075 0.4010067  
## 4.50 0.2832155 0.03677086 Inf 0.2169412 0.3604175  
## 4.75 0.2514147 0.03637329 Inf 0.1869621 0.3290927  
## 5.00 0.2220782 0.03795367 Inf 0.1565424 0.3051251  
## 5.25 0.1952720 0.04023195 Inf 0.1280948 0.2861184  
## 5.50 0.1709902 0.04233748 Inf 0.1030291 0.2702726  
## 5.75 0.1491682 0.04382450 Inf 0.0818166 0.2564764  
## 6.00 0.1296954 0.04452916 Inf 0.0643525 0.2440788  
## 6.25 0.1124287 0.04444813 Inf 0.0502502 0.2326950  
## 6.50 0.0972040 0.04366166 Inf 0.0390214 0.2220901  
## 6.75 0.0838463 0.04228777 Inf 0.0301729 0.2121147  
## 7.00 0.0721774 0.04045559 Inf 0.0232543 0.2026700  
## 7.25 0.0620225 0.03829003 Inf 0.0178764 0.1936881  
## 7.50 0.0532144 0.03590369 Inf 0.0137150 0.1851205  
## 7.75 0.0455964 0.03339308 Inf 0.0105061 0.1769313  
## 8.00 0.0390240 0.03083765 Inf 0.0080382 0.1690929  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

Brackish\_means

## BrackishFreshwater prob SE df asymp.LCL asymp.UCL  
## 0 0.31610096 0.03947029 Inf 0.24423523 0.3979776  
## 1 0.08597468 0.05341963 Inf 0.02421757 0.2628031  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

benthicfish\_means

## Benthic Fisheries prob SE df asymp.LCL asymp.UCL  
## 0 0 0.2318397 0.05710313 Inf 0.1386679 0.3613509  
## 1 0 0.3581289 0.06299398 Inf 0.2458999 0.4884049  
## 0 1 0.2026310 0.06571267 Inf 0.1027472 0.3605920  
## 1 1 0.3119452 0.06305536 Inf 0.2031313 0.4463956  
##   
## Results are averaged over the levels of: Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

pelagicfish\_means

## Pelagic Fisheries prob SE df asymp.LCL asymp.UCL  
## 0 0 0.3142605 0.05288666 Inf 0.22075431 0.4257346  
## 1 0 0.2431258 0.12574062 Inf 0.07763729 0.5507384  
## 0 1 0.2734332 0.05274805 Inf 0.18277453 0.3877251  
## 1 1 0.2139570 0.11262161 Inf 0.06826001 0.5028141  
##   
## Results are averaged over the levels of: Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

coastalfish\_means

## Coastal Fisheries prob SE df asymp.LCL asymp.UCL  
## 0 0 0.3140109 0.06325307 Inf 0.2047342 0.4487050  
## 1 0 0.3044296 0.05961754 Inf 0.2013130 0.4318085  
## 0 1 0.4134182 0.09823672 Inf 0.2416194 0.6092396  
## 1 1 0.1893117 0.04721202 Inf 0.1133046 0.2991067  
##   
## Results are averaged over the levels of: Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

dwfish\_means

## Deepwater Fisheries prob SE df asymp.LCL asymp.UCL  
## 0 0 0.2932367 0.06905657 Inf 0.1775863 0.4435803  
## 1 0 0.2889743 0.06471146 Inf 0.1798072 0.4296971  
## 0 1 0.1888600 0.05556858 Inf 0.1026254 0.3215887  
## 1 1 0.3207937 0.08236259 Inf 0.1837700 0.4976895  
##   
## Results are averaged over the levels of: Benthic, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

coastalbenthic\_means

## Coastal Benthic prob SE df asymp.LCL asymp.UCL  
## 0 0 0.1806850 0.05590732 Inf 0.0951853 0.3161499  
## 1 0 0.2448641 0.05737978 Inf 0.1500245 0.3733247  
## 0 1 0.4946985 0.06942011 Inf 0.3622731 0.6278719  
## 1 1 0.2407522 0.04799110 Inf 0.1593406 0.3466092  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

dwbenthic\_means

## Deepwater Benthic prob SE df asymp.LCL asymp.UCL  
## 0 0 0.1731760 0.05209927 Inf 0.09309394 0.2994042  
## 1 0 0.2613188 0.06792392 Inf 0.15074135 0.4135152  
## 0 1 0.3156432 0.06405820 Inf 0.20503929 0.4519878  
## 1 1 0.3517463 0.06541837 Inf 0.23619068 0.4877361  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

Atlantic\_means

## Atlantic prob SE df asymp.LCL asymp.UCL  
## 0 0.267643 0.04242288 Inf 0.192972 0.3583770  
## 1 0.330480 0.05770038 Inf 0.228442 0.4514285  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

Pacific\_means

## Pacific prob SE df asymp.LCL asymp.UCL  
## 0 0.2660970 0.04852623 Inf 0.1821967 0.3711009  
## 1 0.3037474 0.04350285 Inf 0.2257120 0.3949988  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

Indian\_means

## Indian prob SE df asymp.LCL asymp.UCL  
## 0 0.3694258 0.04882768 Inf 0.2797906 0.4690741  
## 1 0.1806842 0.03579885 Inf 0.1207143 0.2615835  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

bathy\_means

## Bathypelagic prob SE df asymp.LCL asymp.UCL  
## 0 0.3774291 0.03476129 Inf 0.3120779 0.4475618  
## 1 0.2124945 0.05133077 Inf 0.1288504 0.3298759  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Mesopelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

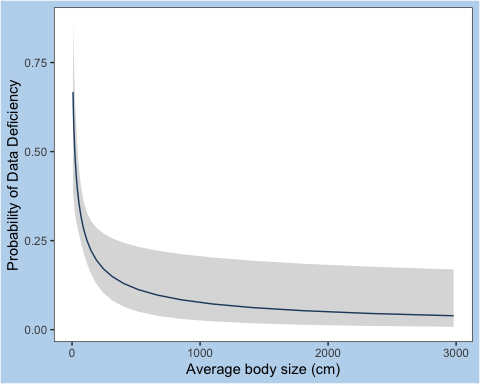
meso\_means

## Mesopelagic prob SE df asymp.LCL asymp.UCL  
## 0 0.3106962 0.05444436 Inf 0.2149867 0.4258977  
## 1 0.2662832 0.04434406 Inf 0.1887106 0.3615334  
##   
## Results are averaged over the levels of: Fisheries, Epipelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

epipelagic\_means

## Epipelagic prob SE df asymp.LCL asymp.UCL  
## 0 0.3424214 0.04872345 Inf 0.2541441 0.4431437  
## 1 0.2390478 0.04013945 Inf 0.1693279 0.3262016  
##   
## Results are averaged over the levels of: Fisheries, Mesopelagic, Bathypelagic, Rep\_Strategy   
## Confidence level used: 0.95   
## Intervals are back-transformed from the logit scale

# Plotting the size figure  
data.frame(logsize\_means) %>%  
 ggplot(aes(x = exp(log\_size), # Exponentiating log size  
 y = prob)) + geom\_line(color="#1F4E79")+  
 geom\_ribbon(aes(ymin = asymp.LCL,  
 ymax = asymp.UCL),  
 alpha = .2) + theme\_bw() + theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank()) + theme(plot.background=element\_rect(fill = "#BDD7EE"),  
 panel.background = element\_rect(fill = "white"))+  
xlab("Average body size (cm)") +  
 ylab("Probability of Data Deficiency")



### try with new color  
data.frame(logsize\_means) %>%  
 ggplot(aes(x = exp(log\_size), # Exponentiating log size  
 y = prob)) + geom\_line(color="#1F4E79")+  
 geom\_ribbon(aes(ymin = asymp.LCL,  
 ymax = asymp.UCL),  
 alpha = .2) + theme\_bw() + theme(panel.grid.major = element\_blank(), panel.grid.minor = element\_blank()) + theme(plot.background=element\_rect(fill = "#EDD58B"),  
 panel.background = element\_rect(fill = "white"))+   
xlab(" ") +  
 ylab(" ")

