Lowcountry Shrimp Dataset Deep Dive

2025-06-10

Table of contents

Summary

About the Project

Funded by the NERRS Science Collaborative, link to page

Collaborators

Name people here!

Part I

Shrimp Data Exploration

Here, we go through each file to see the distribution of data and to make some graphics illustrating what is present in each dataset.

1 Postlarval

1.1 Dataset

The only dataset representing postlarvae is from North Inlet-Winyah Bay. Two files are used: `Penaeus_PostLarvae_NInlet_1981_2017_wide_kac.csv` and `Penaeus_Postlarval_Lengths_NInlet_1981 - 2017.xlsx`. The latter file contains sizes, which were explored in this portion of the project but not expanded upon during later compiling stages.

1.2 Tabular summary of abundance

Table 1.1: Abundance data frame summary

Table 1.1: Data summary

Name	postlarv_abund
Number of rows	1830
Number of columns	21
Column type frequency:	
character	6
Date	1
numeric	14
Group variables	None

Variable type: character

Table 1.2: Abundance data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
replicate	0	1	1	10	0	13	0
date	0	1	0	10	6	915	0
season	0	1	4	6	0	4	0

Table 1.2: Abundance data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
ol_sal	0	1	0	7	2	181	0
ol_temp	0	1	0	7	2	241	0
notes	0	1	0	29	1746	22	0

Variable type: Date

Table 1.3: Abundance data frame summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
date2	6	1	1981-01-20	2017-12-30	1999-06-17	914

Variable type: numeric

Table 1.4: Abundance data frame summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
sample	0	1	458.00	264.21	1.0	229.25	458.00	686.75	915.00
year	0	1	1998.67	12.07	1900.0	1990.00	1999.00	2008.00	2017.00
month	0	1	6.48	3.45	1.0	3.00	6.00	9.00	12.00
day	0	1	15.58	8.83	0.0	8.00	16.00	23.00	31.00
dayofyear	0	1	181.63	105.57	0.0	90.25	181.00	271.75	365.00
week	0	1	26.96	15.08	1.0	14.00	27.00	40.00	53.00
bb_surface_salinity	0	1	31.59	5.08	5.9	30.83	33.20	34.50	38.60
bb_surface_temp	0	1	19.28	7.15	3.3	13.20	19.50	26.37	33.00
total_153_zoop_ind_m0.3	0	1	9946.45	10692.28	0.0	3440.03	6531.50	12306.58	119401.90
vol.filt.0.13	7	1	32.36	4.76	0.0	29.36	32.47	35.74	45.68
total_ppl_density	9	1	0.20	0.63	0.0	0.00	0.00	0.13	8.40
brown_density	9	1	0.05	0.19	0.0	0.00	0.00	0.03	3.77
white_density	9	1	0.05	0.29	0.0	0.00	0.00	0.00	8.19
pink_density	9	1	0.09	0.42	0.0	0.00	0.00	0.00	6.71

1.3 Tabular summaries of size

1.3.1 Brown Shrimp

Table 1.5: Brown shrimp size summary

Table 1.5: Data summary

Name	postlarv_size_brown
Number of rows	1952
Number of columns	16
_____	
Column type frequency:	
character	5
numeric	10
POSIXct	1
_____	
Group variables	None

Variable type: character

Table 1.6: Brown shrimp size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Shrimp_ID	0	1.00	6	8	0	1949	0
TowID	0	1.00	3	4	0	414	0
Replicate	0	1.00	1	1	0	2	0
Species	0	1.00	5	5	0	1	0
Notes	1830	0.06	1	34	0	33	0

Variable type: numeric

Table 1.7: Brown shrimp size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Cruise	0	1.00	510.58	192.75	32.00	388.00	521.00	606.00	909.00
Shrimp_ID_number	0	1.00	10.36	13.36	1.00	2.00	5.00	14.00	73.00
Surface_Salinity	2	1.00	26.58	7.84	5.90	23.10	29.60	32.38	38.60
Surface_Temperature	2	1.00	17.30	5.44	7.00	13.60	16.50	19.50	30.10
Rostrum_teeth	40	0.98	2.68	0.84	0.00	2.00	3.00	3.00	6.00
Rostrum_length	61	0.97	3141.90	415.89	1391.45	3018.81	3278.33	3415.37	4266.98
Telson_length	70	0.96	1516.31	179.36	749.62	1441.84	1567.86	1635.95	1872.39
Uropod_length	38	0.98	1768.37	214.82	799.73	1675.06	1822.97	1914.16	2324.15
Year	0	1.00	2001.33	7.81	1982.00	1996.00	2001.50	2005.00	2017.00
Month	0	1.00	4.26	2.31	1.00	3.00	4.00	4.00	12.00

Variable type: POSIXct

Table 1.8: Brown shrimp size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1982-04-20	2017-10-03	2002-01-03	302

1.3.2 White Shrimp

Table 1.9: White shrimp size summary

Table 1.9: Data summary

Name	postlarv_size_white
Number of rows	1096
Number of columns	16
_____	
Column type frequency:	
character	5
numeric	10
POSIXct	1
_____	
Group variables	None

Variable type: character

Table 1.10: White shrimp size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Shrimp_ID	0	1.00	7	8	0	1087	0
TowID	0	1.00	3	4	0	249	0
Replicate	0	1.00	1	1	0	2	0
Species	0	1.00	5	5	0	1	0
Notes	1029	0.06	5	35	0	29	0

Variable type: numeric

Table 1.11: White shrimp size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Cruise	0	1.00	487.77	228.39	15.00	356.00	482.00	632.50	907.00
Shrimp_ID_number	0	1.00	7.28	8.19	1.00	2.00	4.00	10.00	47.00
Surface_Salinity	0	1.00	31.57	4.82	6.90	31.40	33.20	34.40	38.60
Surface_Temperature	0	1.00	26.03	2.43	13.90	25.20	26.70	27.50	30.10
Rostrum_teeth	16	0.99	1.65	0.79	0.00	1.00	2.00	2.00	6.00
Rostrum_length	27	0.98	1940.35	321.57	1107.64	1710.19	1831.93	2179.20	3514.92
Telson_length	28	0.97	1033.78	155.01	593.65	928.25	990.63	1144.55	1708.58
Uropod_length	18	0.98	1142.02	197.06	654.82	995.32	1081.15	1281.64	2075.20
Year	0	1.00	2000.24	9.24	1981.00	1995.00	2000.00	2006.00	2017.00
Month	0	1.00	6.65	1.16	3.00	6.00	6.00	7.00	10.00

Variable type: POSIXct

Table 1.12: White shrimp size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1981-08-13	2017-09-01	2000-06-29	184

1.4 Graphics - Size Distributions

In these plots, there is a panel for each shrimp species. Size is represented on the y-axis, and a cloud of points called a "beeswarm" has representation for every data point. Each point is colored by the numeric month in which the individual was caught, allowing us to see differences in size at different parts of the year and life-cycle.

There are three different measures of size in this dataset: Rostrum length, Telson length, and Uropod length. A graph is below for each.

With brown shrimp, looks like we've got a couple of size classes in here - bigger shrimp earlier in the year, and smaller ones that are mostly later months.

The plot below shows the entire time series, with date (as year-month) along the x-axis and uropod length on the y-axis. Points are again colored by month.

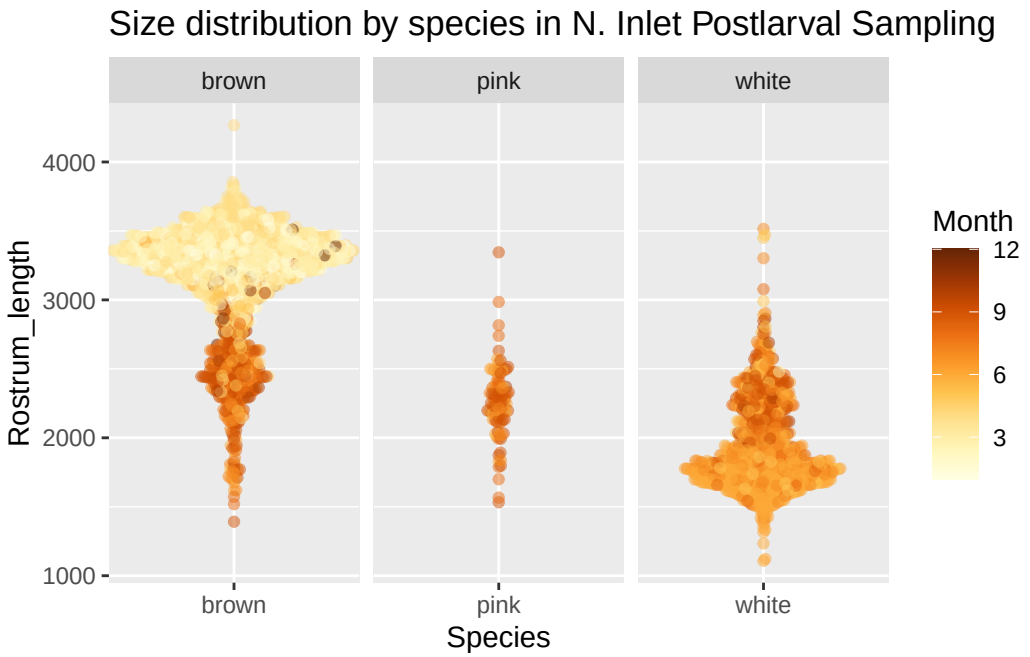

Figure 1.1: Beeswarm plot of rostrum length in postlarval shrimp. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

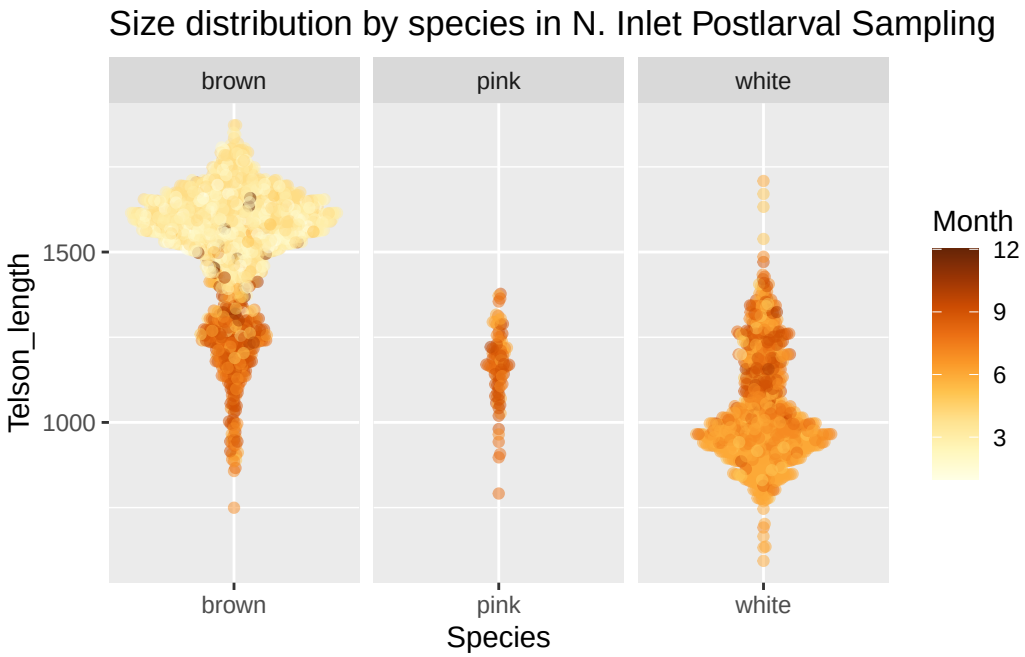

Figure 1.2: Beeswarm plot of telson length in postlarval shrimp. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

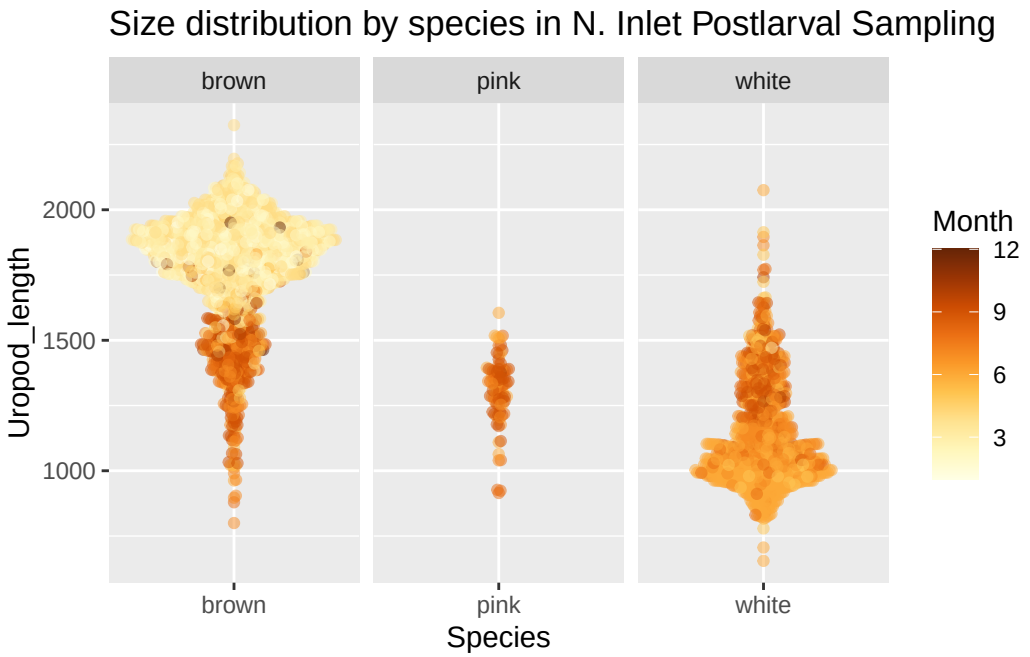

Figure 1.3: Beeswarm plot of uropod length in postlarval shrimp. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

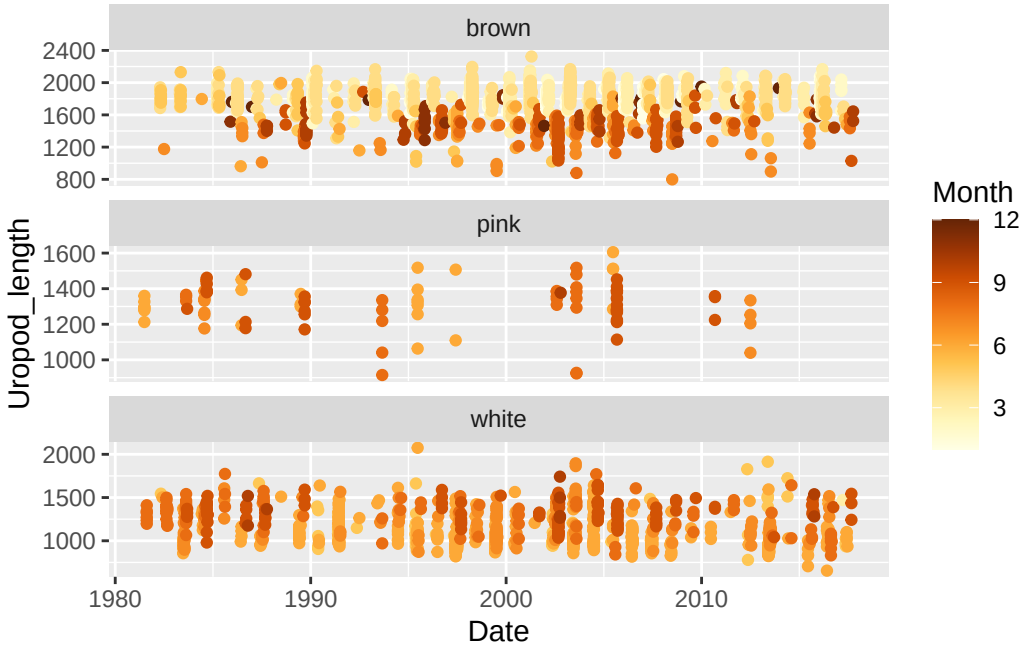

Figure 1.4: Time series plot of uropod length in postlarval shrimp. Points represent measurements of individual shrimp, and are colored by the month in which they were captured.

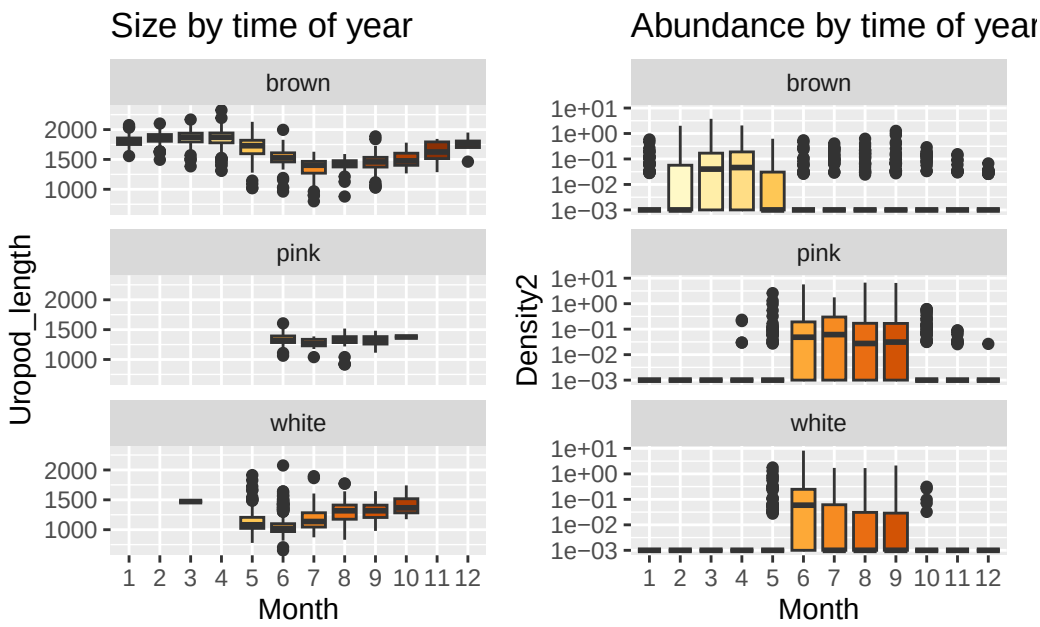

Figure 1.5: Boxplots representing shrimp size and abundance Note the log10-scaled y-axes.

1.5 Graphics - Abundance and Size Boxplots

So even though we're seeing that pattern in brown shrimp of being so much bigger early in the year and smaller later, they're just really not all that common later in the year.

2 Juveniles

2.1 Datasets

There are two datasets that represent juveniles: SCDNR's Creek Trawls, and NIW NERR's Oyster Landing seines. We will explore each.

Pull out date characteristics and deal with column types if necessary; change Species columns to common names.

(OL's temp and salinity columns are character, due to entries of "DataGap". Weight has an entry that is only a decimal point. I'll turn these to NAs manually before forcing to numeric so that if anything else is odd (e.g. two decimal points in a number) I'll get a warning about the coercion and can look into it.)

Count and size are in the same file for Oyster Landing - up to 100 lengths taken, in columns. Probably should split this and pivot the lengths to long.

2.2 Tabular summaries of abundance

2.2.0.1 SCDNR Creek Trawls

Table 2.1: Abundance (SC) data frame summary

Table 2.1: Data summary

Name	juv_sc_cpue
Number of rows	5619
Number of columns	13
_____	
Column type frequency:	
character	5
numeric	8
_____	
Group variables	None

Variable type: character

Table 2.2: Abundance (SC) data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
StationCode	0	1	4	4	0	7	0
EstuaryCode	0	1	2	2	0	1	0
DTStart	0	1	8	10	0	621	0
SpCode	0	1	4	4	0	3	0
Species	0	1	4	5	0	3	0

Variable type: numeric

Table 2.3: Abundance (SC) data frame summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1	2810.00	1622.21	1.00	1405.50	2810.00	4214.50	5619.00
Coll	0	1	20030686.47	3365.28	19801003.00	19911032.00	20031007.00	20161081.00	20231084.00
Year	0	1	2002.97	13.34	1980.00	1991.00	2003.00	2016.00	2023.00
Month	0	1	6.62	1.91	1.00	5.00	7.00	8.00	12.00
Day	0	1	15.29	8.38	1.00	9.00	15.00	22.00	31.00
CPUE	3	1	546.00	2690.49	0.00	0.00	0.00	82.00	62976.00
Lat	0	1	32.84	0.07	32.75	32.80	32.86	32.86	32.95
Long	0	1	-79.89	0.08	-79.99	-79.98	-79.88	-79.85	-79.76

2.2.0.2 Oyster Landing Seines

Table 2.4: Abundance (Oyster Landing) data frame summary

Table 2.4: Data summary

Name	juv_ol_count
Number of rows	1796
Number of columns	12
_____	
Column type frequency:	
character	4
Date	1
numeric	7
_____	

Group variables	None

Variable type: character

Table 2.5: Abundance (Oyster Landing) data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Protocol	0	1	7	8	0	3	0
WQ.Source	2	1	8	19	0	4	0
WQ.Time	2	1	0	5	36	74	0
Species	0	1	5	5	0	2	0

Variable type: Date

Table 2.6: Abundance (Oyster Landing) data frame summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1984-01-04	2023-12-21	2002-02-18	898

Variable type: numeric

Table 2.7: Abundance (Oyster Landing) data frame summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Sample	0	1.00	449.76	259.63	1.0	225.0	449.5	675.00	899.0
Year	0	1.00	2002.63	11.71	1984.0	1993.0	2002.0	2012.00	2023.0
Month	0	1.00	6.60	3.26	1.0	4.0	7.0	9.00	12.0
Temp	94	0.95	20.47	6.91	3.8	14.5	21.6	26.90	31.4
Sal	106	0.94	31.82	4.78	2.3	30.0	33.3	35.00	38.7
Weight	489	0.73	625.27	2355.88	0.0	0.0	0.0	77.05	39196.1
Count	240	0.87	380.32	1326.23	0.0	0.0	0.0	62.00	22560.0

2.3 Tabular summaries of size

2.3.1 Brown Shrimp

2.3.1.1 SCDNR Creek Trawls

Table 2.8: Brown shrimp (SC) size summary

Table 2.8: Data summary

Name	sz_sc_brn
Number of rows	34818
Number of columns	14
_____	
Column type frequency:	
character	4
Date	1
numeric	9
_____	
Group variables	None

Variable type: character

Table 2.9: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
DTStart	0	1	8	10	0	489	0
StationCode	0	1	4	4	0	7	0
SpCode	0	1	4	4	0	1	0
Species	0	1	5	5	0	1	0

Variable type: Date

Table 2.10: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1980-05-08	2023-12-11	1996-06-12	489

Variable type: numeric

Table 2.11: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1.00	35073.55	22862.74	1.00	15867.25	32821.50	52650.75	81763.00
Coll	0	1.00	19981663.41	109955.91	19801003.00	19891005.00	19961014.00	20061026.00	20231080.00
Length	0	1.00	69.54	22.67	12.00	52.00	71.00	86.00	151.00

Table 2.11: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Lat	0	1.00	32.83	0.07	32.75	32.80	32.81	32.86	32.95
Long	0	1.00	-79.91	0.08	-79.99	-79.98	-79.96	-79.85	-79.76
TempS	372	0.99	27.61	2.92	8.40	25.90	28.00	29.70	36.70
SalinityS	913	0.97	17.58	5.21	0.00	14.00	18.00	21.00	28.00
Year	0	1.00	1998.07	12.00	1980.00	1989.00	1996.00	2006.00	2023.00
Month	0	1.00	5.81	0.86	1.00	5.00	6.00	6.00	12.00

2.3.1.2 Oyster Landing Seines

Table 2.12: Brown shrimp (Oyster Landing) size summary

Table 2.12: Data summary

Name	sz_ol_brn
Number of rows	15212
Number of columns	12
Column type frequency:	
character	5
Date	1
numeric	6
Group variables	None

Variable type: character

Table 2.13: Brown shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Protocol	0	1	7	8	0	3	0
WQ.Source	0	1	8	19	0	4	0
WQ.Time	0	1	0	5	409	48	0
Species	0	1	5	5	0	1	0
Rep	0	1	4	6	0	100	0

Variable type: Date

Table 2.14: Brown shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1984-05-14	2023-08-29	1999-06-11	410

Variable type: numeric

Table 2.15: Brown shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Sample	0	1.00	384.06	230.34	10.0	185	383.0	543	891.0
Year	0	1.00	1999.55	10.18	1984.0	1991	1999.0	2006	2023.0
Month	0	1.00	6.27	1.50	1.0	5	6.0	7	12.0
Temp	895	0.94	25.64	2.82	6.6	24	26.0	28	31.4
Sal	903	0.94	32.54	3.53	7.7	31	33.8	35	38.7
Length	0	1.00	21.20	6.70	0.0	16	21.0	26	71.0

2.3.2 White Shrimp

2.3.2.1 SCDNR Creek Trawls

Table 2.16: White shrimp (SC) size summary

Table 2.16: Data summary

Name	sz_sc_wht
Number of rows	46792
Number of columns	14
_____	
Column type frequency:	
character	4
Date	1
numeric	9
_____	
Group variables	None

Variable type: character

Table 2.17: White shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
DTStart	0	1	8	10	0	467	0
StationCode	0	1	4	4	0	7	0
SpCode	0	1	4	4	0	1	0
Species	0	1	5	5	0	1	0

Variable type: Date

Table 2.18: White shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1980-05-08	2023-12-11	2003-07-29	467

Variable type: numeric

Table 2.19: White shrimp (SC) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1.00	45299.88	23170.58	7.00	25467.75	48030.50	65043.25	81764.00
Coll	0	1.00	20031152.32	4812.02	19801003.00	19921067.00	20031027.00	20151040.00	20231080.00
Length	1	1.00	53.26	21.32	8.00	37.00	51.00	66.00	152.00
Lat	0	1.00	32.84	0.07	32.75	32.80	32.81	32.86	32.95
Long	0	1.00	-79.90	0.08	-79.99	-79.98	-79.96	-79.85	-79.76
TempS	215	1.00	29.22	2.98	11.00	28.30	29.60	31.00	36.70
SalinityS	822	0.98	15.99	5.81	0.00	12.00	17.00	20.00	30.00
Year	0	1.00	2003.02	12.48	1980.00	1992.00	2003.00	2015.00	2023.00
Month	0	1.00	7.43	1.11	1.00	7.00	7.00	8.00	12.00

2.3.2.2 Oyster Landing Seines

Table 2.20: White shrimp (Oyster Landing) size summary

Table 2.20: Data summary

Name	sz_ol_wht
Number of rows	24387
Number of columns	12

Column type frequency:	
character	5
Date	1
numeric	6
Group variables	None

Variable type: character

Table 2.21: White shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
Protocol	0	1	7	8	0	3	0
WQ.Source	0	1	8	19	0	4	0
WQ.Time	0	1	0	5	874	51	0
Species	0	1	5	5	0	1	0
Rep	0	1	4	6	0	100	0

Variable type: Date

Table 2.22: White shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
Date	0	1	1984-07-12	2023-12-11	2000-09-11	436

Variable type: numeric

Table 2.23: White shrimp (Oyster Landing) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Sample	0	1.00	421.82	222.99	14.0	236.0	414	592.0	898.0
Year	0	1.00	2001.01	9.94	1984.0	1993.0	2000	2008.0	2023.0
Month	0	1.00	8.45	1.39	1.0	7.0	8	10.0	12.0
Temp	2200	0.91	25.85	3.78	10.6	23.8	27	28.8	31.4
Sal	2305	0.91	32.81	3.96	5.4	31.8	34	35.1	38.7
Length	0	1.00	23.84	7.09	3.0	19.0	23	29.0	71.0

2.4 Graphics - Size Distributions

Beeswarm plots can be slow to render so I have subsetted both data frames here to 15,000 rows.

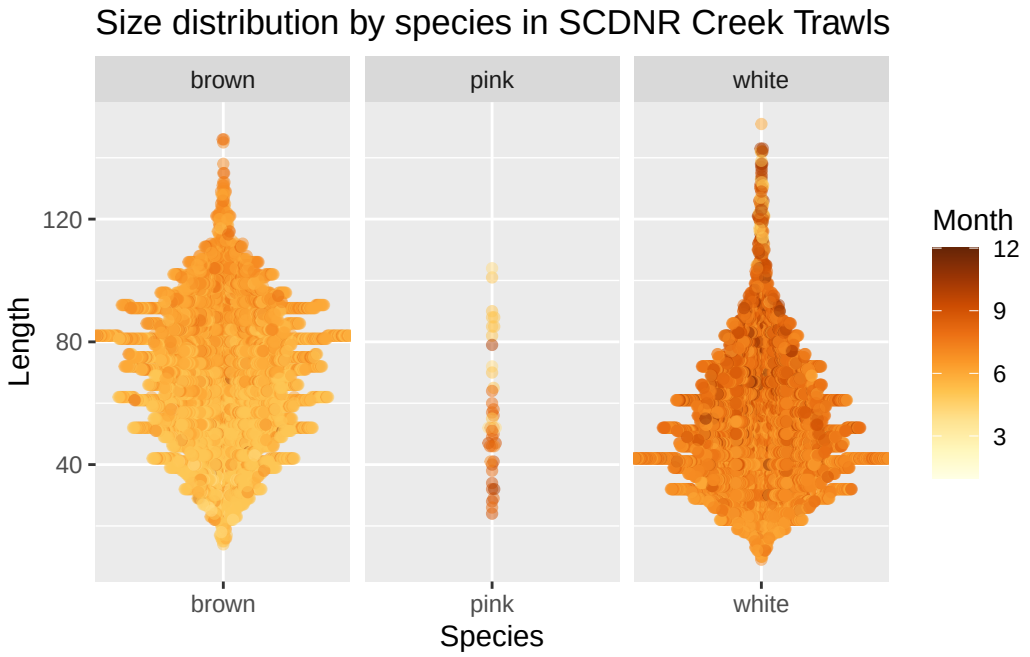

Figure 2.1: Beeswarm plot of juvenile shrimp length for the SCDNR data. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

There was a typo in the original files - Sample 748, from 2017-07-20, LEN21, value of 221. [confirmed from reserve that this was a typo; updated to '21' in '_corrected' file, september 2024 and then file with (kac 2025-06-02) appended to name]

Looks like OL seines are generally catching smaller individuals of both species than the SCDNR creek trawls.

2.5 Graphics - Abundance/CPUE Boxplots

Both surveys (unsurprisingly) show the same temporal pattern of abundance - aztecus in May/June, with setiferus later in the year.

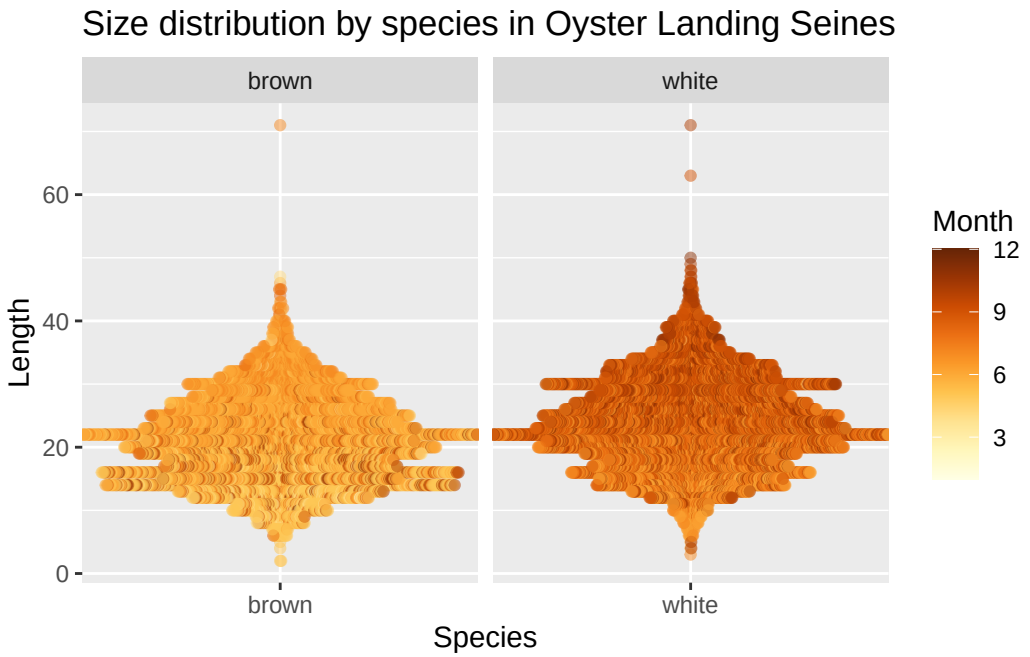

Figure 2.2: Beeswarm plot of juvenile shrimp length for the Oyster Landing seine data. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

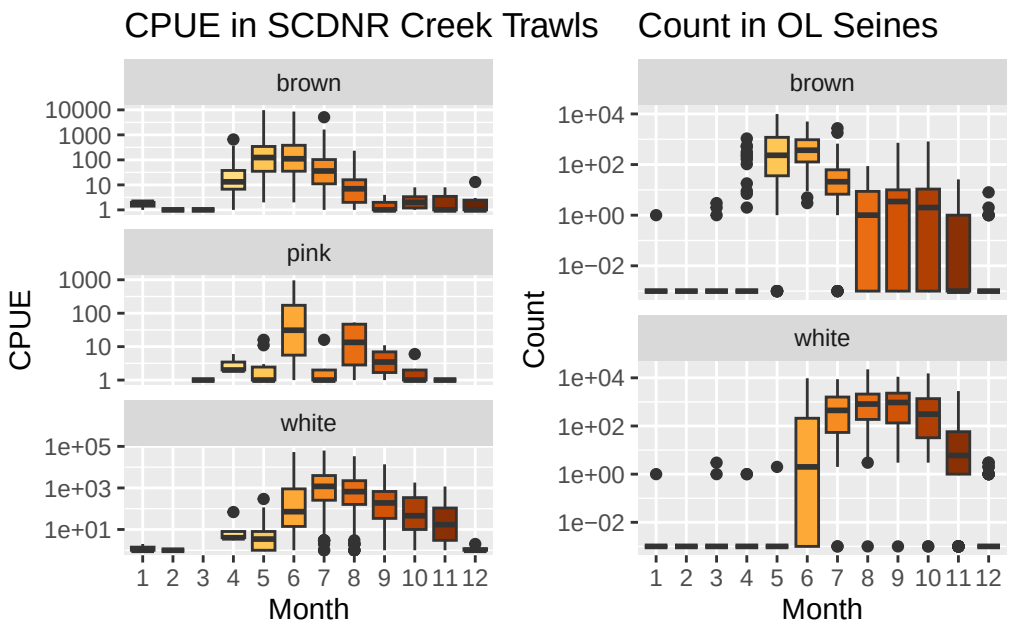

Figure 2.3: Boxplots representing shrimp abundance (as either CPUE or count, depending on the survey) by month. Note the log10-scaled y-axes.

3 Subadults

3.1 Datasets

There are two datasets that represent subadults: GADNR's EMTS sampling and SCDNR's Estuarine Trawls. We will explore each.

Combine species in GA data frames; pull out date characteristics and deal with column types if necessary; change shrimp species to common names.

3.2 Tabular summaries of abundance

3.2.0.1 GADNR EMTS

Table 3.1: Abundance (GA) data frame summary

Table 3.1: Data summary

Name	subad_ga_count
Number of rows	39669
Number of columns	14
_____	
Column type frequency:	
character	2
Date	1
numeric	11
_____	
Group variables	None

Variable type: character

Table 3.2: Abundance (GA) data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
RefNum	0	1	12	12	0	19828	0
Species	0	1	5	5	0	2	0

Variable type: Date

Table 3.3: Abundance (GA) data frame summary

skim_variable	n_missing	complete_rate	min	max	median	n_unique
GuessedDate	0	1	1975-12-19	2021-10-26	2000-01-10	3254

Variable type: numeric

Table 3.4: Abundance (GA) data frame summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
TotWt	3	1.00	1491.00	8839.01	0	0.00	27.22	680.39	1394789.25
TotNum	83	1.00	115.10	610.81	0	0.00	2.00	43.00	73800.00
SampleWt	13	1.00	382.37	716.10	0	0.00	27.22	680.39	65770.89
SampleNum	120	1.00	28.03	47.17	0	0.00	2.00	42.00	1220.00
NumMeas	21154	0.47	12.23	16.40	0	0.00	2.00	30.00	207.00
LbsperHr	12269	0.69	17.83	75.69	0	0.16	2.08	12.88	7060.50
NumperLb	12359	0.69	33.89	253.35	0	15.27	25.60	40.00	31751.50
NumFemales	2395	0.94	7.63	10.59	0	0.00	2.00	15.00	122.00
NumMales	2382	0.94	6.27	9.46	0	0.00	1.00	11.00	97.00
Year	0	1.00	1998.94	13.00	1975	1988.00	2000.00	2010.00	2021.00
Month	0	1.00	6.49	3.42	1	4.00	6.00	9.00	12.00

3.2.0.2 SCDNR Estuarine Trawls

Table 3.5: Abundance (SC) data frame summary

Table 3.5: Data summary

Name	subad_sc_cpue
Number of rows	21771
Number of columns	13

Column type frequency:	
character	5
numeric	8
Group variables	None

Variable type: character

Table 3.6: Abundance (SC) data frame summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
estuary	0	1	2	2	0	6	0
StationCode	0	1	4	4	0	24	0
DTStart	0	1	8	10	0	1647	0
SpCode	0	1	4	4	0	3	0
Species	0	1	4	5	0	3	0

Variable type: numeric

Table 3.7: Abundance (SC) data frame summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1	10886.00	6284.89	1.000e+00	5443.50	10886.00	16328.50	21771.00
Coll	0	1	19986283.92	2808.61	1.979e+07	19890034.00	19970150.00	20090016.00	20230114.00
Year	0	1	1998.61	12.29	1.979e+03	1989.00	1997.00	2009.00	2023.00
Month	0	1	6.60	3.37	1.000e+00	4.00	6.00	9.00	12.00
Day	0	1	17.04	8.02	1.000e+00	11.00	18.00	23.00	31.00
CPUE	4	1	119.11	525.17	0.000e+00	0.00	0.50	24.00	20052.00
Latitude	0	1	32.57	0.23	3.215e+01	32.32	32.67	32.77	32.83
Longitude	0	1	-80.30	0.36	-8.085e+01	-80.65	-80.29	-79.92	-79.89

3.3 Tabular summaries of size

3.3.1 Brown Shrimp

3.3.1.1 GADNR EMTS

Table 3.8: Brown shrimp (GA) size summary

Table 3.8: Data summary

Name	subad_ga_size_brown
Number of rows	11029
Number of columns	4
_____	
Column type frequency:	
character	3
numeric	1
_____	
Group variables	None

Variable type: character

Table 3.9: Brown shrimp (GA) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
RefNum	0	1	12	12	0	745	0
TowDate	0	1	8	10	0	308	0
Species	0	1	5	5	0	1	0

Variable type: numeric

Table 3.10: Brown shrimp (GA) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Size	0	1	113.84	17.45	8	103	114	125	182

3.3.1.2 SCDNR Estuarine Trawls

Table 3.11: Brown shrimp (SC) size summary

Table 3.11: Data summary

Name	sz_sc_brn
Number of rows	66499
Number of columns	15
_____	

Column type frequency:	
character	5
numeric	10
Group variables	None

Variable type: character

Table 3.12: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
SpCode	0	1	4	4	0	1	0
Sex	0	1	0	1	61324	4	0
estuary	0	1	2	2	0	6	0
StationCode	0	1	4	4	0	24	0
Species	0	1	5	5	0	1	0

Variable type: numeric

Table 3.13: Brown shrimp (SC) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1.00	129622.34	80691.21	1.000e+05	59387.50	126106.00	187401.50	304572.00
Coll	0	1.00	19955732.85	2429.92	1.979e+07	19870163.00	19940120.00	20010143.00	20230106.00
Length	0	1.00	102.12	19.99	2.100e+01	90.00	103.00	115.00	177.00
Year	0	1.00	1995.56	11.25	1.979e+03	1987.00	1994.00	2001.00	2023.00
Month	0	1.00	6.66	1.22	1.000e+00	6.00	6.00	7.00	12.00
Day	0	1.00	17.19	7.95	1.000e+00	10.00	19.00	24.00	31.00
Latitude	0	1.00	32.68	0.19	3.215e+01	32.67	32.77	32.80	32.83
Longitude	0	1.00	-80.12	0.30	-8.085e+01	-80.29	-79.97	-79.92	-79.89
TempB	1640	0.98	27.81	2.84	0.000e+00	27.00	28.40	29.40	32.40
SalinityB	2759	0.96	24.83	6.13	0.000e+00	21.00	26.00	30.00	39.00

3.3.2 White Shrimp

3.3.2.1 GADNR EMTS

Table 3.14: White shrimp (GA) size summary

Table 3.14: Data summary

Name	subad_ga_size_white
Number of rows	54505
Number of columns	4
Column type frequency:	
character	3
numeric	1
Group variables	None

Variable type: character

Table 3.15: White shrimp (GA) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
RefNum	0	1	12	12	0	2231	0
TowDate	0	1	8	10	0	726	0
Species	0	1	5	5	0	1	0

Variable type: numeric

Table 3.16: White shrimp (GA) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
Size	0	1	124.02	21.61	2	108	124	140	212

3.3.2.2 SCDNR Estuarine Trawls

Table 3.17: White shrimp (SC) size summary

Table 3.17: Data summary

Name	sz_sc_wht
Number of rows	224834
Number of columns	15

Column type frequency:	
character	5
numeric	10
Group variables	None

Variable type: character

Table 3.18: White shrimp (SC) size summary

skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
SpCode	0	1	4	4	0	1	0
Sex	0	1	0	1	173246	4	0
estuary	0	1	2	2	0	6	0
StationCode	0	1	4	4	0	24	0
Species	0	1	5	5	0	1	0

Variable type: numeric

Table 3.19: White shrimp (SC) size summary

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
X	0	1.00	165616.57	89820.94	7.000e+00	89692.25	172903.50	244450.75	307960.00
Coll	0	1.00	19991381.22	3497.97	7.979e+07	19890293.00	19970176.00	20090085.00	20230114.00
Length	0	1.00	113.38	22.51	1.100e+01	98.00	114.00	130.00	207.00
Year	0	1.00	1999.12	12.36	1.979e+03	1989.00	1997.00	2009.00	2023.00
Month	0	1.00	7.49	3.49	1.000e+00	4.00	8.00	10.00	12.00
Day	0	1.00	16.44	8.29	1.000e+00	10.00	17.00	23.00	31.00
Latitude	0	1.00	32.59	0.22	3.215e+01	32.46	32.67	32.77	32.83
Longitude	0	1.00	-80.27	0.35	-8.085e+01	-80.54	-80.24	-79.92	-79.89
TempB	4077	0.98	20.86	6.76	0.000e+00	15.00	20.60	27.90	32.40
SalinityB	4665	0.98	25.93	7.06	0.000e+00	22.00	27.00	31.00	209.00

3.4 Graphics - Size Distributions

South Carolina's file has over 300,000 points, which makes for a beeswarm plot that is very large and slow to render. Georgia's file is "only" 65k and even it is too slow to render beeswarm

plots. So rather than graphing the full datasets, I have randomly sampled 10,000 rows from GA and 15,000 rows from SC (because SC has three species).

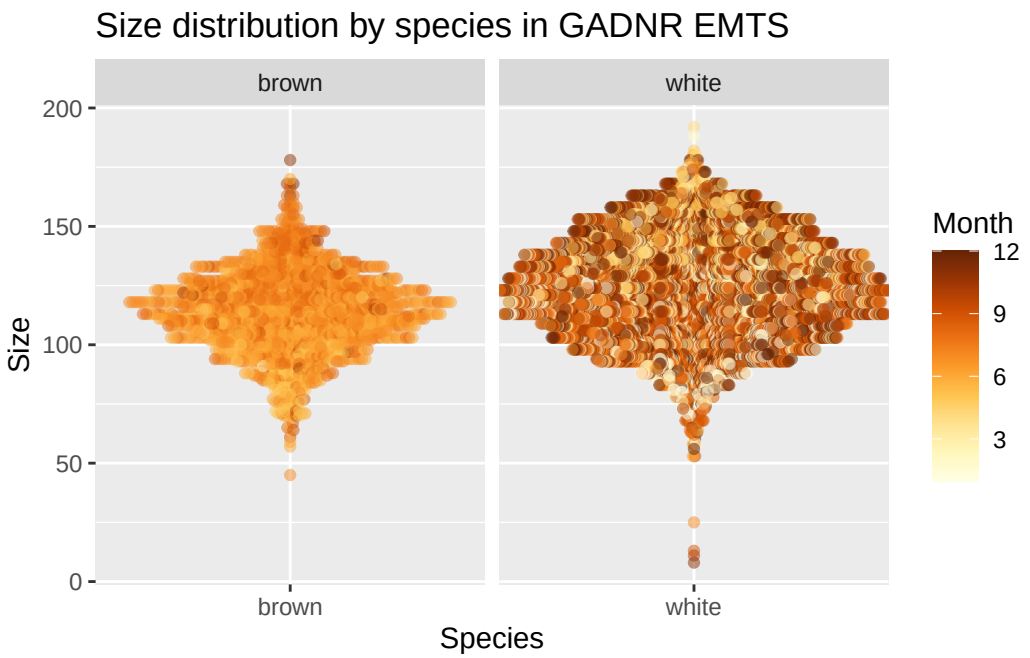

Figure 3.1: Beeswarm plot of subadult shrimp length for the GADNR data. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

3.5 Graphics - Abundance/CPUE Boxplots

Both surveys (unsurprisingly) show the same temporal pattern of abundance - brown shrimp most abundant in June and July, also high in August; and May in SC. For white shrimp, we see a dip in June and (less so) July.

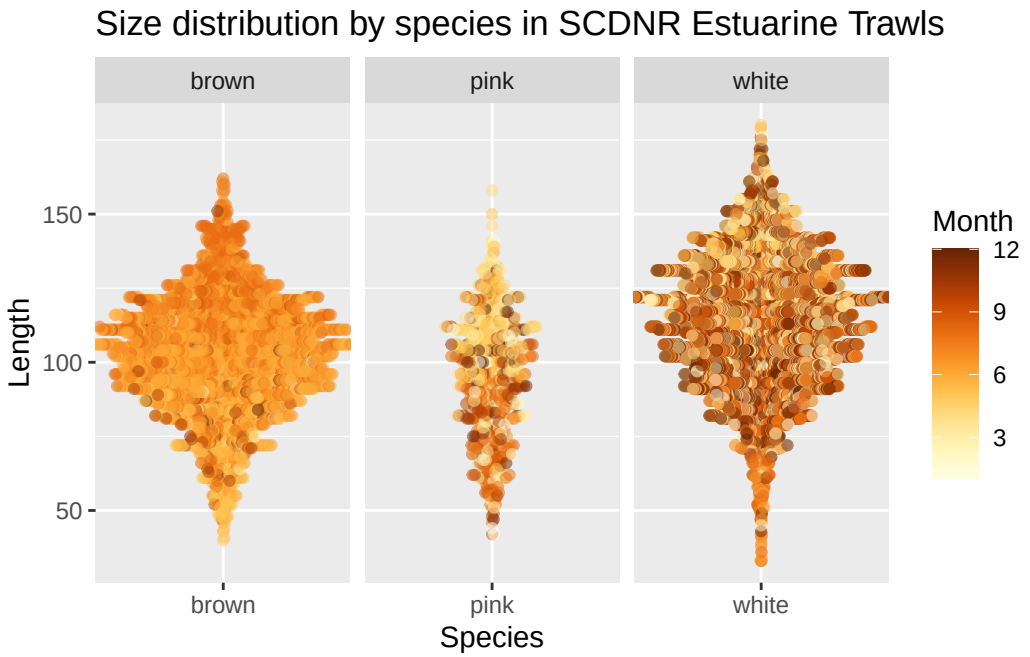

Figure 3.2: Beeswarm plot of subadult shrimp length for the SCDNR data. Points represent measurements of individual shrimp, and are colored by the month in which they were captured. The y-axis represents length. More points spread out along the x-axis for a given length means that there were more individuals of that length captured than regions where points remain closer to the center.

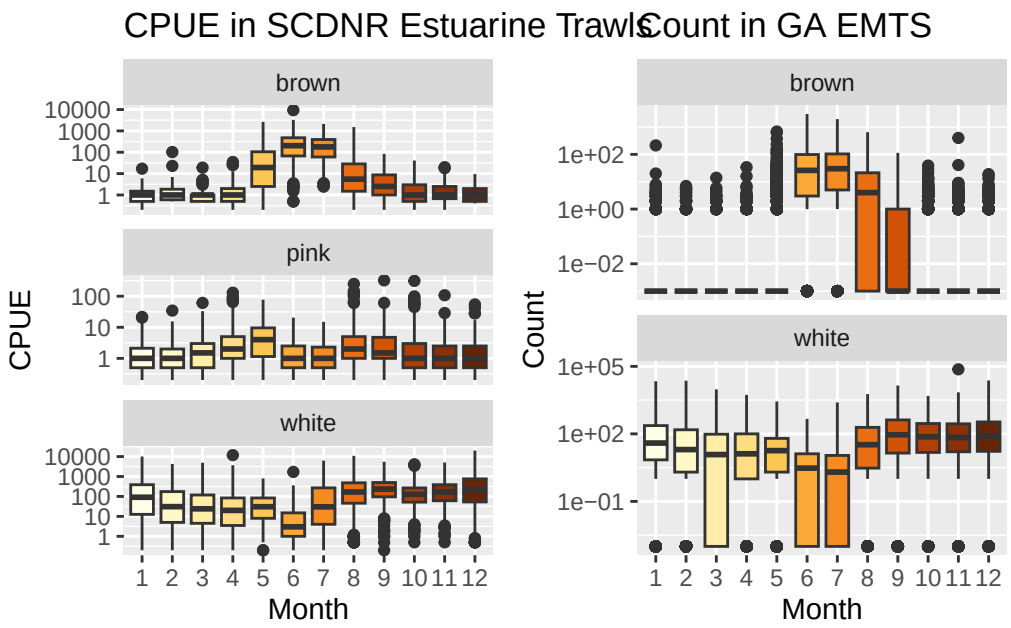

Figure 3.3: Boxplots representing shrimp abundance (as either CPUE or count, depending on the survey) by month. Note the log10-scaled y-axes.

4 Adults - fishery-independent

SEAMAP info here

5 Adults - landings and trip tickets

Commercial data here

6 Food sources (benthic cores)

Stuff here

Part II

Environmental Data Exploration

7 Water Temperature at Charleston Harbor

stuff

8 Salinity

from SCDNR trawls

Part III

Shrimp Year Definition

9 Shrimp Year Explanations

Here we show how we defined 'shrimp year' for each species, and what that actually means in the context of winter temperatures.

Part IV

Summarizing to Abundance Index

10 Postlarval Stage

Datasets - .....

11 Juvenile

Datasets .....

Multiple datasets here; abundances and sizes also present

12 Subadults

Stuff here ....

13 Adults - fishery-independent

SEAMAP info here

13.1 test

rendering with freeze: auto

13.2 another test

```
[1] 4
```

14 Adults - landings and trip tickets

Commercial data here

14.1 test

code:

```
[1] 2
```

The end

15 Food sources (benthic cores)

Stuff here

Part V

Summarizing Environmental Data

16 Water Temperature at Charleston Harbor

stuff

17 Salinity

from SCDNR trawls

Part VI

Relationships between datasets

18 Temperature thresholds

some stuff

19 Nursery Period Salinity

stuff

20 Regression Models

Start to explore what happens when we do more than just correlations.

A Data Dictionaries - raw datasets

Data dictionaries from data files explored in parts I & II.

A.1 Postlarval data

description

A.2 Juvenile data

description

... etc.

B Data Dictionaries - processed data

Explanations of files after summarizing/compiling. Shrimp abundance index files will all follow the same column convention. Environmental data files will probably take more explanation.

B.1 Shrimp Data

details

B.2 Environmental Data

etc.