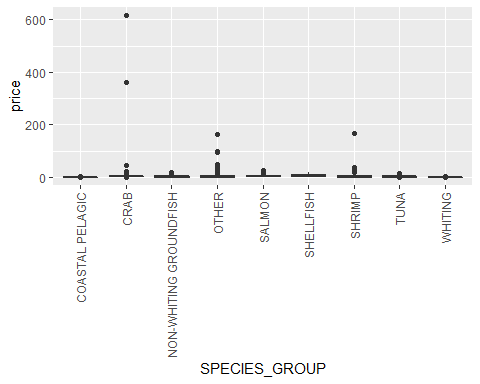
Outlier Analysis

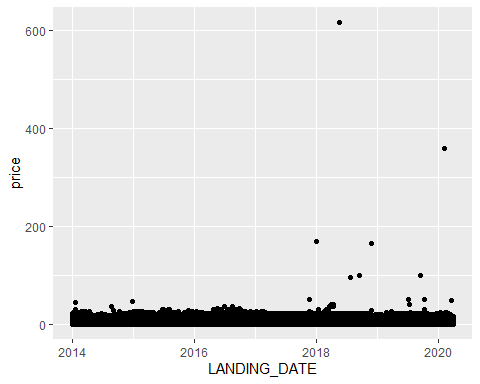
## Overview

There are a four data points from 2017-2020 that look like obvious errors (one instance of crab in 2018, crab in 2020, shrimp in 2017, and other in 2018). Removing prices > $150/lb from comp\_dat\_raw, or bounding them would deal with them. See below for outlier analysis.

## Broad Look

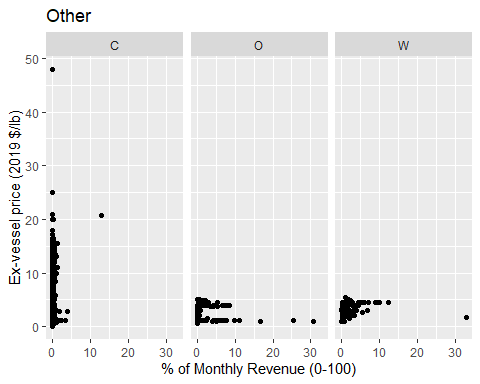
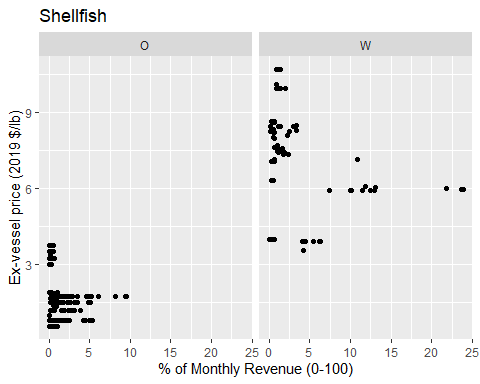
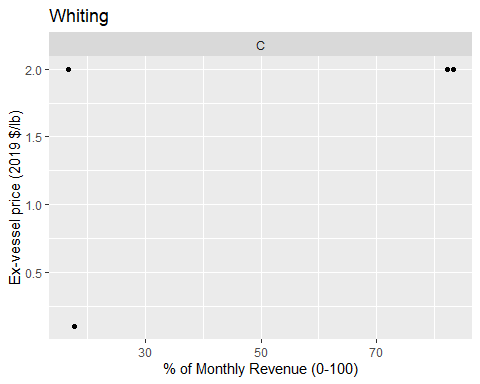
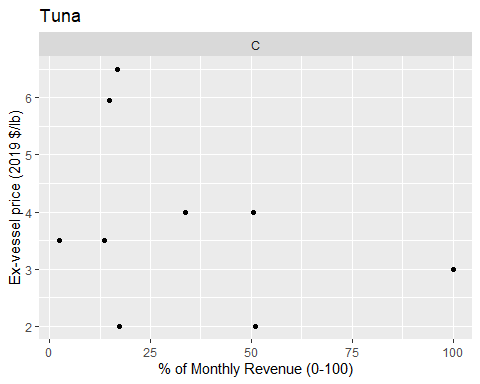
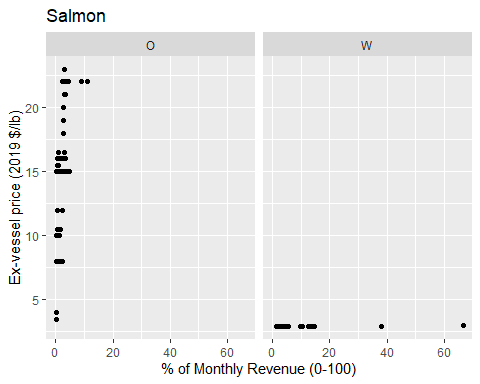
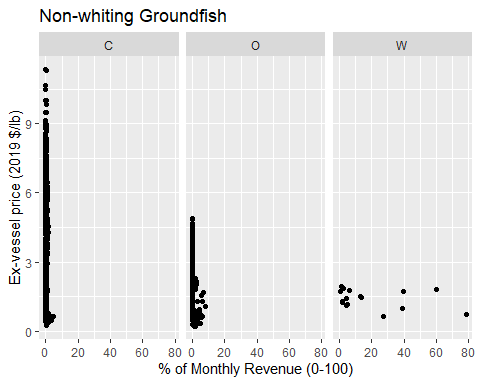
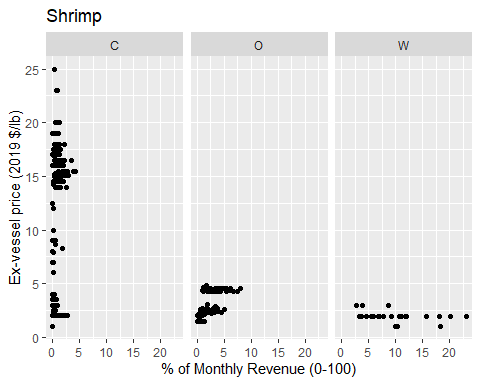
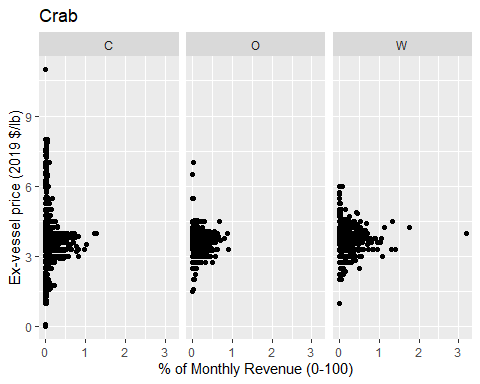
These graphs take a very high level look at all the data together. As you can see, there are several fairly obvious outliers in the crab in 2018 and 2020, and maybe some in shrimp and other over the years. So I start by removing prices > $200/lb.



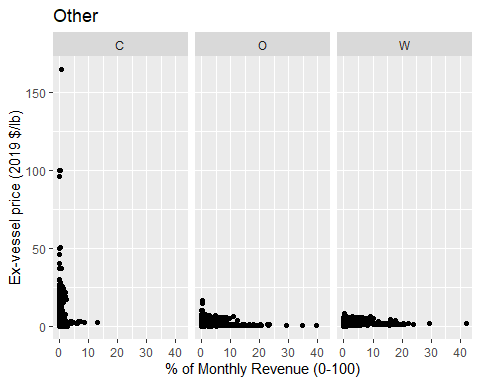
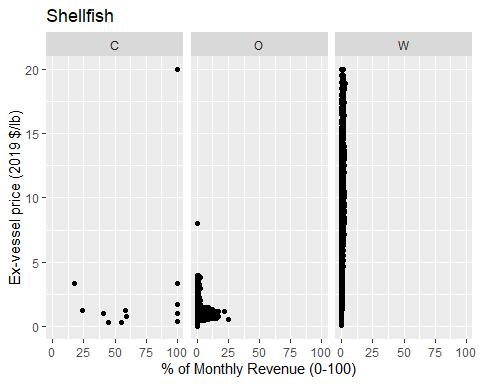
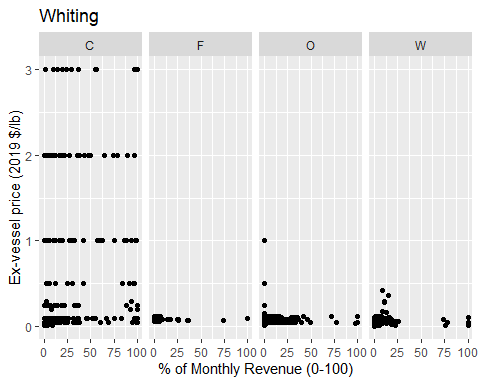
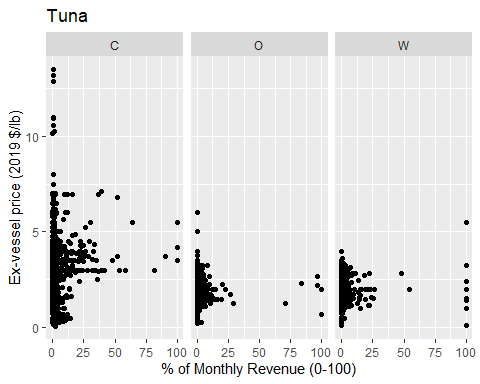
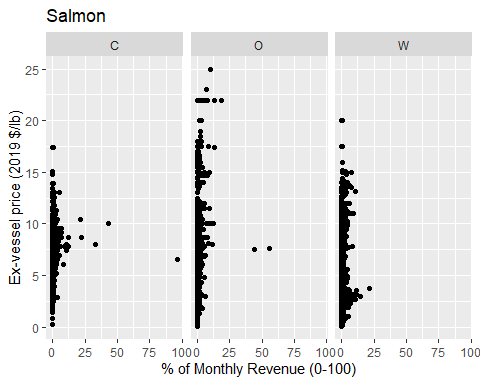
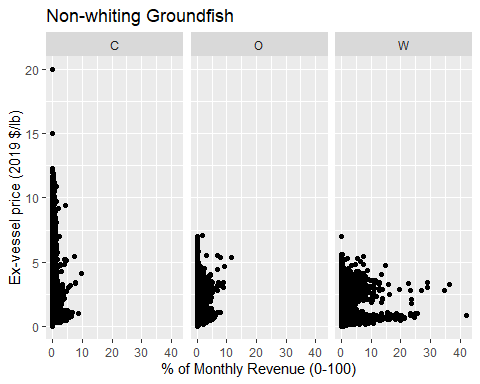
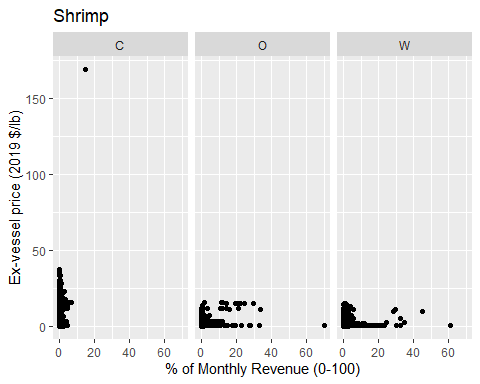
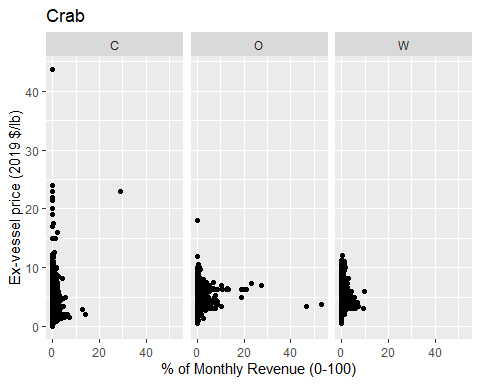


## 2020

I take a look at 2020 data to see if we observe any outliers in terms of ex-vessel price and compare that with how “important” each observation is to monthly ex-vessel revenue by state. Once the prices > $200/lb are removed - I don’t see any 2020 observations that look like obvious errors that would have large impact on our data.



## All Years

Doing the same exercise for all years, the only observation that looks potentially problematic is the shrimp price over $150/lb. 

## By Fishery over Time

