

Skeena River Sockeye Salmon Data Set Skeena_River.csv

The Skeena River Sockeye Salmon fishery is in British Columbia. The independent variable is X = number of Spawners (in thousands) = number of mother fish, and Y = number of Recruits (in thousands) = number of teenage fish migrating into the ocean. For a multivariate regression, there is also a variable called Year, ranging from 1940 to 1967. The idea that the numbers depend linearly on Year may be far-fetched, since the Salmon spawn and the Recruits come back 4 years later. The data after 1967 tend to be dreadful because of factory fishing for salmon.

The fishery is in a beautiful place, and is also, like all west coast salmon fisheries, in dire danger. In 1951 there was a rock slide that limited the number of recruits who actually got to the ocean. The 4-year cycle of this fishery meant that one would expect the number of Spawners to be low in 1955, which it distinctly is.

I live part-time in the North Olympic Peninsula of Washington State, along the Strait of Juan de Fuca separating the U.S. from Vancouver Island in Canada. The loss of the fishery has also caused trouble for the resident orca pod.

Finally, these data are a classic case of heteroscedastic regression, i.e., non-constant variation, and much of the analysis is improved by transforming both Recruits and Spawners, e.g., by logs.