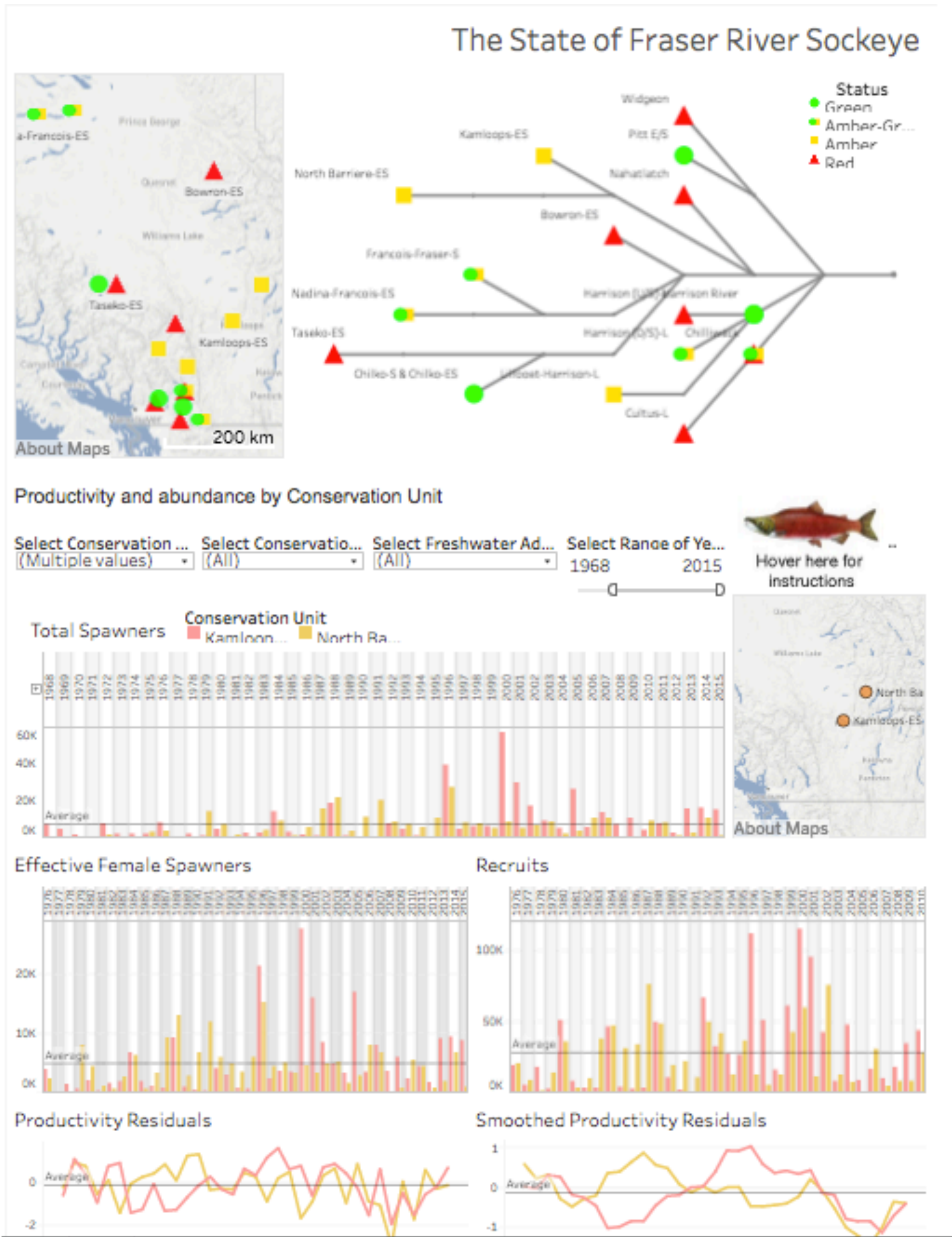
# The State of the Salmon:

# Visualizing population data from Fraser River Sockeye

Michael Barrus




**Abstract** In this project, I integrated data about Fraser River salmon populations and develop an interactive data visualization tool that allows fisheries scientists and managers to explore this data at different spatial and temporal scales. The objective of this project is to provide fisheries scientists and managers with the tools to easily and comprehensively explore trends in salmon data, and discover features and attributes of that data.

## Introduction

Fisheries managers for the federal Department of Fisheries and Oceans are tasked with determining when and where to conduct fisheries, as well as how many fish to harvest at a time. These decisions are informed by advice from fisheries scientists, who in turn are responsible for understanding the factors that affect fish population dynamics. The Fraser river sockeye salmon fishery is the most economically and culturally important fishery in British Columbia. Salmon are born in dozens of locations throughout the Fraser river, and these different populations show variation across their life histories. The primary form of data collection on these populations takes place when field biologists with the DFO perform annual fish counts of both adults and juveniles for these populations. Those counts are analysed by fisheries scientists in order to measure the size and health of the populations.

Understanding both local and global salmon population trends is essential to effective fisheries management. However, routine analysis of this information by fisheries scientists within the DFO is impeded by decentralized datasets and the lack of user-friendly tools with which to explore the available information. There is presently no interface with which scientists can review and examine existing salmon data across species and watersheds. The development of such an interface would provide salmon scientists with a more thorough understanding of salmon in the Fraser river.

I was recruited by the DFO in the summer of 2017 to join a project called “The State of the Salmon” (SoS). The project team (henceforth referred to as the SoS team) is composed of four members—a senior biologist, a junior biologist, a senior analyst, and myself. The other members of the team expressed a desire to consolidate available data from Fraser salmon populations and provide a simple, interactive interface through which managers and researchers can explore that data, identify patterns, and discover how these patterns change over time. My role with the group would be to develop a visualization tool that department scientists could use to explore and examine the existing data on Fraser river sockeye populations.

## Related Work

Forms of data visualization are used by fisheries scientists within the DFO, but these are usually limited to static bar, line and dot plots that encode population data. Typically, these are designed by an individual user and distributed to other users of their choice as PDFs or printed documents. Figure 1 shows a representative document. The document **[1]** summarizes and presents information relating to various measures of population health for subgroups salmon within the larger Fraser river watershed. These reports juxtapose line plots of different metrics across a common timescale.

The encoding choices employed here are largely defensible; the line plots are effective encodings for a continuous variable plotted over a time series. However, there are several potential problems with this visualization that are related to their ability to . These static plots make it difficult for users to make comparisons between populations, as the user would have to navigate through the document to find the desired charts and then physically juxtapose them against each other (placing sheets side-by-side) or flipping back and forth between them. Furthermore, these figures do not give users the ability to drill down, nor is it possible to get details from the underlying data from the plot.

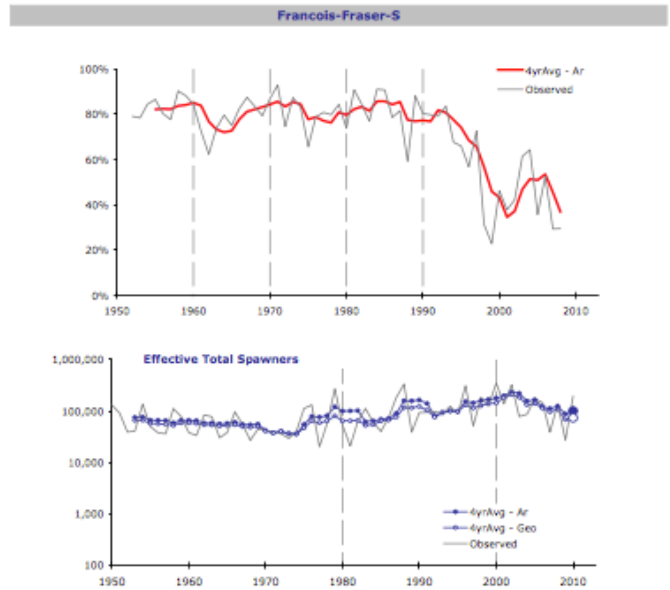



*Figure 1. Line plots from the annual status assessment reports issued within the DFO. These reports are representative of current visualizations within the department. Using static images spread across multiple pages makes it difficult to compare information about different salmon populations.*

Another existing solution is the Pacific Salmon Explorer tool, shown as figure 2. The Salmon Explorer visualizes salmon population and environmental data from the Skeena river in northern BC. This interface has advantages over the static interface shown in figure 1. The use of small multiple line plots juxtaposes the time series data from many populations against each other, which might make it easier for users to compare trends and features between the populations. However, this solution is imperfect because the user does not have the ability to zoom, filter across items or attributes, or superimpose the line plots from each system in order to more precisely compare attributes such as trends, outliers or similarities.

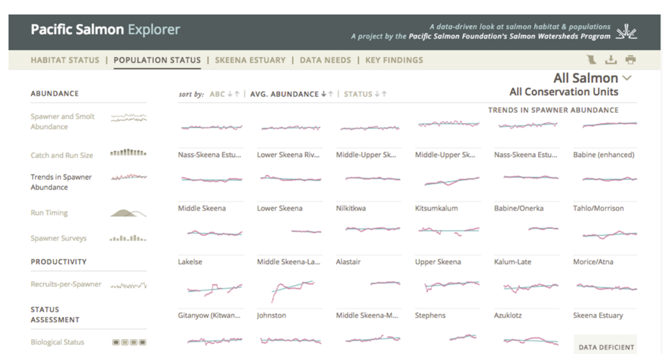



*Figure 2: The Pacific Salmon Explorer.*

Fisheries scientist within the department have also incorporated geographic information into these visualizations. In [1], stoplight marks representing the health of individual salmon populations were superimposed over map of the Fraser River, in an attempt to represent the relationship between geography and population health. This visualization communicates information about the health of salmon populations and the location of those populations, allowing scientists to explore the topology of statuses throughout the watershed. However, information density is not evenly distributed throughout the figure, and potentially important information about the connectivity between populations is not emphasized.

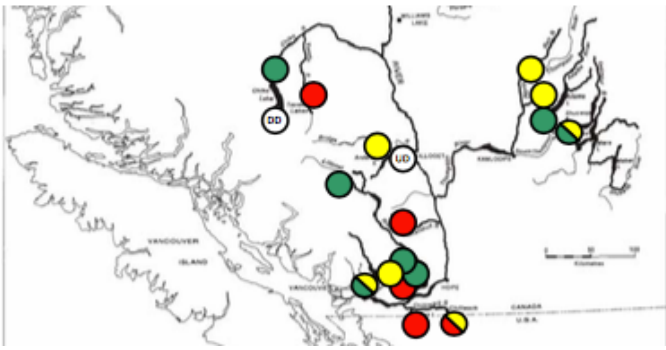


*Figure 3. Stoplight marks encode information about the health of salmon populations, and are plotted over the location of those populations on a map of the Fraser river.*

## Data and Task Abstractions

### 1.1 Data

The data I used to develop this visualization was provided to me by a field biologist from the SoS team. The DFO has performed annual salmon surveys continuously since 1951, with the most recent year of available data being 2015. During these surveys, they count the number of fish within a survey area. These counts are performed separately for both juveniles and adults in each survey area. From these raw counts and geospatial positions, they derive several attributes that describe the population size, health, spatial position.

Fish found within particular survey areas are aggregated into larger attributes, called Conservation Units (CUs). CUs are aggregations based on spatial proximity and genetic similarity, and each CU used in this dataset was first defined in [2]. These CUs were the primary categorical attribute used to group the population data. Each CU is assigned a status by a panel of DFO scientists, which is a categorical measure of overall health ranging from best ("Green" status), to intermediate ("Amber"), to worst ("Red" status). If there is sufficient variation across the sub-populations within the CU, the CU can be assigned a "mixed" status of two or more primary statuses (resulting in a status like Amber/Red). () also included the definitions of "Freshwater Adaptive Zones", which are categorical groupings of CUs based on shared geography.

Estimates of population size ("Total Returns"), number of spawning fish and number of spawning females ("Spawner Estimate" and "Female Spawner Estimate", respectively) are derived from the raw adult fish counts. Productivity ("Productivity residuals") is calculated by comparing adult counts and juvenile counts, estimating how many offspring each spawning fish produces, then comparing that to a fixed baseline rate. Smoothed productivity residuals are a running average calculated across the previous four years; as sockeye have four-year life cycles, this reduces the variability across years that is driven by different spawning classes. These values were calculated by Gottfried Pedstal, the analyst for the SoS team, and provided to me as a series of excel tables. I joined these excel tables using Trifacta Wrangler, and order to create an appropriate .csv file for tableau.

I conducted a separate process to gather the data that was necessary to construct the tree plot. Topological stream order is a method of classifying streams based on their position within a larger watershed. This hierarchical numbering system starts at the mouth of the river (a stream order of one) and increases at each upstream confluence. A schematic is provided as figure 4. Each CU was assigned a stream order number according to this measure, and I determined this number by manually going through a map of the Fraser river watershed and counting the confluences between the CU and the mouth of the Fraser River. During this process, I also recorded the path from the CU to the mouth, which was used to construct the tree diagram.

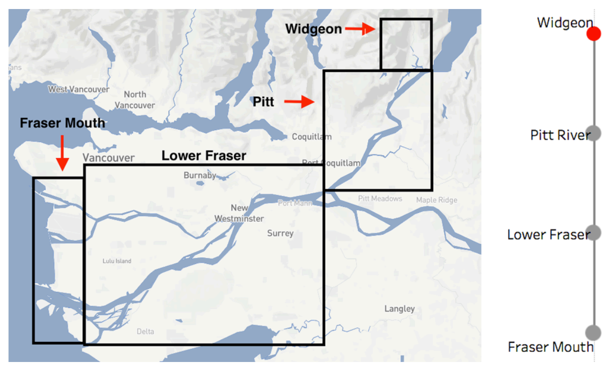



*Figure 4. Topological stream order. Each node depicts a unique subsection of river that is bounded box a box on the map. Links between nodes indicate direct connectivity.*

### 1.2 Tasks

Initially, the other members of the SoS team had difficulty articulating the tasks they wished to achieve with a visualization tool. Sue Grant, the senior biologist on the SoS team, expressed a desire to use visualizations to help users browsing the dataset to "understand the state of the salmon in British Columbia". After further discussion, we decided to narrow our focus on one of several subsets of user types within the department in order to characterize the tasks a user might. There were several end user groups that we were interested in developing a visualization tool for, but ultimately decided our target users were senior science managers. These individuals are mostly concerned with high-level tasks and are broad rather than deep consumers of information. More specifically, these users would be interested in the following types of tasks:

1. Exploring the existing dataset.
2. Discover trends and features in measures of sockeye population size and health over time.
3. Compare trends and features across different Conservation Units (CUs).
4. Discover similarities between CUs.

5. Explore the topology of CUs within the Fraser river. Managers want to be able to explore the relationship between population health and the location of those populations relative to other populations within the watershed. Similar attributes (such as “red” status) amongst CUs that were grouped together by their position within the watershed could indicate that local factors were affecting sockeye populations.

Science managers in the DFO are distinct from research scientists in that they consume information, rather than produce it. They do not usually perform analysis or generate hypotheses about salmon populations; instead, they expect to have a high-level understanding of the short and long term trends in population health and abundance.

## Solution

The solution I developed is a Tableau dashboard with geospatial and temporal representations of salmon population data. The top half of the dashboard depicts CU status and their location in both geographic (map) and quasi-geographic (tree plot) representations. CU status is encoded as a glyph that combines shape and hue. The bottom half of the dashboard is a series of three bar plots and two line plots that depict five measures of population size and health. These plots can be manipulated using a series of drop-down filters that The subcomponents of the dashboard are described below.

### 1.1 Status glyphs

The CU statuses were encoded as glyphs that combined hue and shape to represent each status attribute in the dataset. The DFO uses a red-yellow-green “stoplight” encoding for population health, as described previously in the derived data section. These encodings are potentially problematic, as it would be difficult for individuals with red/green colorblindness to differentiate between them. Adopting a colorblind-friendly color plot (such as a categorical diverging blue-orange) is an obvious solution, but an issue here is that the attributes being encoded take the names of the hues they are encoded by. Using a blue hue to encode the attribute “green” seems undesirably confusing.


Given this constraint, adopting a dual hue/shape encoding seemed to be an appropriate solution. In this encoding, “Green” status is represented by a green circle, “Amber” by an amber square, and “Red” by a red triangle. Using both hue and shape channels allowed me to maintain the color-attribute relationship DFO users are familiar with, while also making the visualization more colourblind friendly. Another solution would be to vary the luminance of each hue in order to make them distinguishable along that channel. I chose to use shape rather than luminance as I expected to use these statuses as nodes in a tree plot, where they would necessarily assume a shape. Furthermore, these channels do not exclude the use of gradations in luminance, although the use of three separate channels to encode a single attribute seems unnecessarily redundant.

### 1.2 Map

The visualization features two maps—one that is juxtaposed with the tree plot, and another that is situated near the filters for the lower half of the dashboard. Both were designed to emphasize bodies of water, while de-emphasizing elements such as human development, place names and topological features. Including this information would potentially distract from the purpose of the map, which was to provide the location of the CUs both within the province and in relation to one another. Including extraneous details could detract from that. Furthering this simplicity, the map uses only two hues, one for water and the other for land. On the upper map, status glyphs (values determined by the 2017 status assessment) are plotted over the location of CUs. On the lower map, the mark is identical across all CUs. The name of the CU is displayed beneath the glyph or mark, and the user can pan and zoom the map by using a toolbar.

I included maps to provide information about the spatial position of CUs within the province to users. The upper map encoded information about the spatial location of CUs as well as their status, as users reported that they wanted to be able to identify the physical location of each CU. The marks displayed on each map are linked to the other plots they are juxtaposed against. As the upper section of the dashboard (map and tree plot) is meant to give a broad overview of the Fraser system in its entirety, every CU in the dataset has been plotted. In contrast, the lower map marks only the CUs that have been selected by the user through the drop-down menus, as a linked view here helps users rapidly identify the location of the chosen CUs.

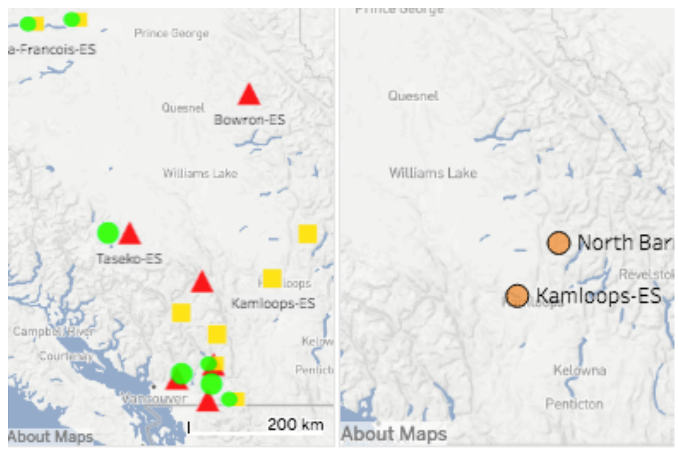



*Figure 5. The map to the left is part of the upper dashboard, and displays the status glyphs of the CUs in their geospatial locations. The map to the right is the smaller of the two, and shows only those CUs selected from the drop down filters.*

### 1.3 Tree Plot

The upper half of the dashboard also features a quasi-geographic tree plot that represents connectivity between CUs within the larger Fraser watershed. The tree plot features links and nodes to encode this connectivity. The status glyphs stand in as the nodes, and as described in “Status glyphs”, the shape and color of these nodes encodes the status of that CU

The structure of the tree plot was derived from the geography of the Fraser river. The tree diagram was hand coded by assigning each CU and stream directly downriver of those CUs an X-Y position according to that topological stream order, and their position to river right or to river left of the Fraser mainstem. In order to avoid overlap across CUs that occupied the same stream order, the CUs with the highest stream order value were plotted nearest to the center, and then systems with lower stream order values were plotted immediately outside of those. Edges were drawn from each node to the one directly downstream of it and so on until the terminal node was reached. This formulation maintained characteristics of the geographic space (such as a CU's right/left positioning relative to the Fraser mainstem) while obscuring details that were not directly related to connectivity, such as absolute distance between systems.

Using a tree plot idiom enabled viewers to explore the topology of the Fraser in a way that the map alone did not. The map and the tree were juxtaposed against each other because each idiom encoded different aspects of the geo-spatial data of the Fraser river. While the map showed the lat-long coordinates of each CU and the layout of those CUs within the watershed as a whole, the tree plot showed the connectivity across the system. Furthermore, the tree plot is a simple, intuitive design that users have probably encountered before. Tracing paths through the plot is simple and reveals information about the relationships between systems that is not readily apparent from the map.

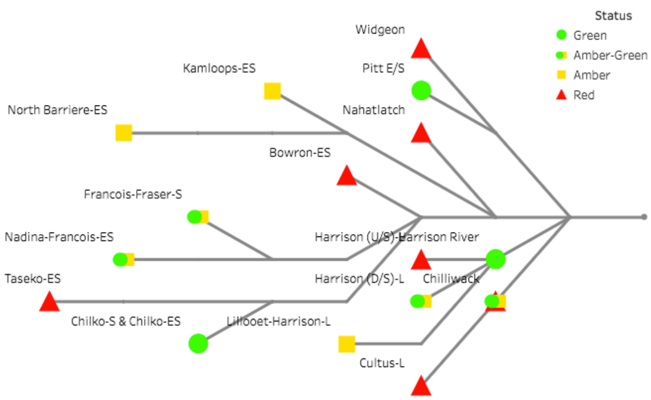



*Figure 6. Tree plot.*

### 1.4 Bar and line plots

Below the geo-spatial encodings are a series of five bar and line plots. Filters at the top of the plots allow users to select CUs by several attributes (Name, status, freshwater adaptive zone), and then select a range of time for which to display data. The plots are juxtaposed and linked, such that using any of the filters above the plots changes all of them simultaneously. Hovering above any point on the plot results in a pop-up tooltip that includes the name of the CU, the year of the selected item, and its exact value. The selected CUs are displayed on the small map to the right of the filters. Each plot displays an average line that is calculated across all items on the plot.

Attributes relating to yearly population size are encoded as bars, whereas attributes relating to productivity are encoded as line. This design choice was dictated by the nature of the variables. Salmon exhibit roughly 4-year lifespans, so the populations within a CU exist in four discrete cycles. The size of a return is not predictive of the size of the return the next year, or two or three years after, only the fourth year after that. Comparing year to year population sizes is somewhat misleading, and encoding the data with bars for each year emphasizes that these are discreet populations. In contrast, productivity is more dependent on local environmental conditions, which are continuous across years and independent of the total size of the population. Line plots are thus a more appropriate encoding for these continuous variables.

Each CU is associated with a hue, such that all lines and bars associated with that CU are plotted in that hue. These hues are categorical, and the pallet was a built-in Tableau pallet . While using 16 unique hues to encode categorical attributes could make the visualization indecipherably complex, the small scale of the plots already prohibits users from making meaningful comparisons across more than 40-50 items. Given that this is time-series data, this limits users to looking at only 3-4 CUs at a time, for which categorical hue encodings are appropriate.

The plots are ordered according to SoS team comments about what information fisheries scientists look to when assessing salmon populations. The top plot shows total returns, or the sum of all fish returning to a CU in a given year. According to the SoS team, scientists use this to get a broad overview of the health of the system, and often consult it first before more detailed metrics. Below this are plots of the other two population attributes, and below these are two measures of productivity.

## Implementation

All derived variables were calculated by the other members of the SoS team. These variables were sent to me as excel workbooks, with each CU as a separate workbook, and each year of data on a separate sheet within that workbook. I performed data cleaning in Trifacta Wrangler, using pivot functions to arrange the information into a format that was compatible with Tableau. The data was then stored as a csv file and imported to Tableau.

I designed the background map using Mapbox, and imported it using Mapbox studio. CU status shapes were designed in Adobe Photoshop and imported to Tableau.

## Results

### 1.1 Use scenario

A potential use scenario would be that of a science manager facing questions from non-scientists, from either within or outside of the department. Science managers are often the point of contact for media inquiries and governmental officials who are seeking general information about salmon stocks. A common query is for comment on the most recent year of salmon returns, or to describe the health of a population aggregate (such as all populations within the Fraser, all populations located in the greater Vancouver area, etc).

If a science manager received a call from a member of the press who had a general inquiry about the current state of Fraser river salmon, they could pull up the dashboard and refer first to the map and tree plot, which display status for individual populations across the entire river. The manager might browse these plots and answer that the health of populations varied throughout the watershed, as indicated by the distribution of different status glyphs across these plots, but CUs in the upper watershed tended to have lower status than those in the lower watershed. If the reporter asked if there were areas of concern, the manager could identify the

CUs assigned a "red" status on the tree plot, and report those CUs to the reporter.

If the reporter asked for follow up on these regions or for specific details about these populations, the manager could then select "Red" under the "Select Conservation Status" filter, and see plots of population measures. The manager could then browse the bar and line plots to identify trends, outliers or features across these populations. Examining these plots suggests that population size is highly variable across these CUs, but all CUs showed a reduction in total recruits beginning in the late 1990s, and these have not recovered.

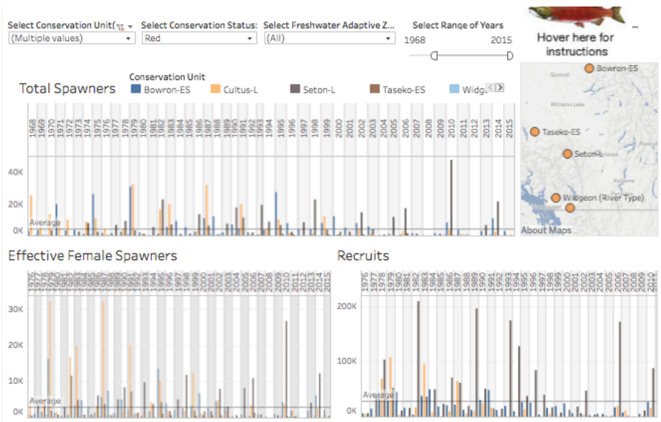



*Figure 7. Dashboard showing all "Red" CUs selected.*

However, the smoothed productivity residual plot shows that productivity appears to be exhibiting a positive trend across most CUs and returning to the long-term average. The manager could also identify Cultus Lake as an outlier, where this CU has continued to have sharp declines in productivity and population size.

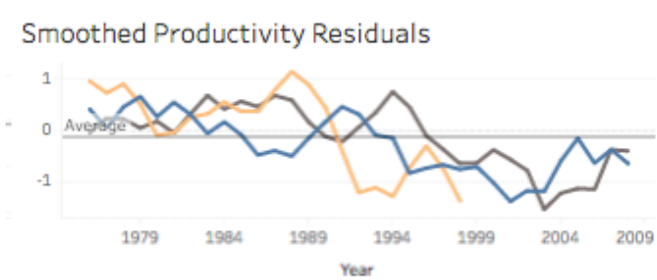



*Figure 8. Smoothed productivity residuals for the selected CUs.*

In this scenario, the manager is able to rapidly browse the statuses of CUs across the Fraser river and identify features of that topology. Furthermore, they can filter items and compare them to discover trends or features in the data, and identify outliers. The visualization enables rapid exploration of the existing dataset and comparisons across CUs in a way that was not possible using the department's existing visualizations. While there are ways to improve the visualization, this represents a significant improvement over current techniques.

### 1.2 Evaluation

I showed iterations of the visualization to the SoS team during three meetings, as well as providing them with a link to a live demo on Tableau Public where they could evaluate it at any time.

An early comment was that they wanted the ability to filter plots by attribute, and thus the lower half of the dashboard incorporated a series of drop-down filters that allowed users to select data by CU, status and year. This interactivity was an important advantage over the traditional static visualizations used in the department. A suggestion for future visualizations was to provide additional attributes to filter by; however, this would involve the expansion of the existing dataset to include the data by which they wanted to filter, so we decided it was beyond the scope of this project.

Initial iterations of the dashboard featured small multiples, with a series of juxtaposed plots showing the same attribute and same time scale plotted for each CU in the dataset. Because salmon population trends are noisy and exhibit high year-to-year variability, it was difficult for them to compare trends and features across CUs in this format. In response, I redesigned the interface to allow users to superimpose the data from multiple CUs into a single plot. Similarly, users reported that seeing the entire time series was sometimes distracting, as the high variability across years resulted in information-dense plots. Implementing filters that allowed users to select the range of years for which to display data resolved these issues.

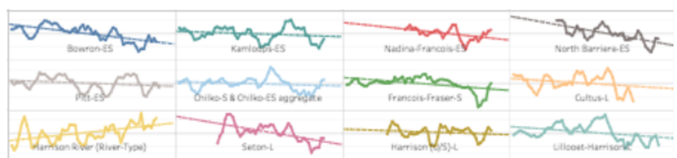


*Figure 9: Early iteration of the dashboard featuring small multiples.*

In a similar vein, these users also reported that seeing a single attribute from each system (such as "Total returns") was not satisfactory, as fisheries scientists within the department tended to consider all five of these measures at the same time when making status assessments or even exploring the dataset. My early iterations of the dashboard featured a single plot which displayed a single attribute, but in response to this discussion, I changed the layout to incorporate five juxtaposed plots so that each attribute could be displayed simultaneously.

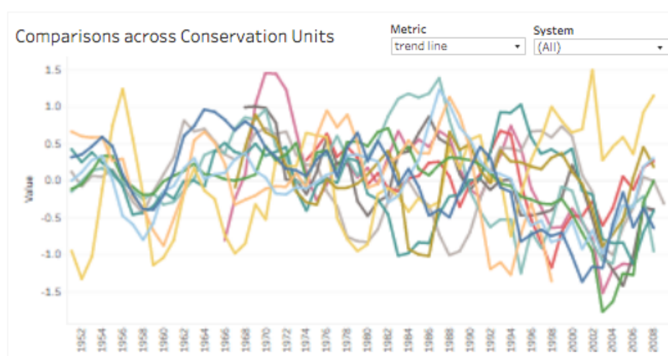



*Figure 10. Early iteration of the dashboard with a single plot featuring a drop down menu for attribute selection.*

When I initially presented the tree plot to the SoS team, the dashboard displayed either the map or tree plot but not both.

A drop down menu at the top gave users the option of displaying one or the other. However, users reported that it was useful to have both juxtaposed against each other. The task that these idioms were designed to assist with (namely exploring connections and topology) was performed more easily if users were able to see both where the CUs were located and how they were connected within the Fraser river.

## DISCUSSION

### 1.1 Critique

The State of the Salmon dashboard makes several contributions that add to or improve on existing visualizations that are in use in the department. Perhaps the most important contribution is that it consolidates important data about Fraser river sockeye; this data had previously been distributed throughout the department and was not available on-demand. Furthermore, the dashboard allows researchers to explore relationships between the geography and connectivity across systems and the health of those systems in ways they were not able to before. This interface also allows researchers to superimpose attributes and items from multiple systems on the same plot, making it easier for them to compare and identify trends, outliers and features across these systems.

There are a few weakness of the present tool that could be improved upon. First, the tree plot is only capable of displaying status, which is only one attribute of a CU; this simplification belies the complexity of the underlying data and does not allow the user to explore the topology and connectivity of the Fraser across time. Ways to improve the tree plot and address these issues are discussed in the "Future Work" section.

Another weakness is due to the scale of the bar plots. When more than ~40-50 items are displayed on each plot, the bar marks decrease in size and become difficult to see. A potential solution to this would be using stacked, rather than juxtaposed bars, but it is not clear to me that this would make comparisons across CUs easier, as moving the marks from a common scale to an unaligned scale might increase the cognitive effort of making comparisons.

A key limitation of the visualization is that the network was handbuilt by providing x,y coordinates for each mark; this may make it difficult to scale as it essentially must be redesigned each time nodes or links are added. However, this is ultimately a limitation on future work, and does not affect the efficacy of the existing visualization.

The SoS team also described an interest in incorporating some analytic tools into the visualization interface, such as time-series clustering capabilities. In its present form, the visualization lacks these. This could be an area for further work.

### 1.2 Future work

The SoS team is interested in continuing to develop this visualization and others based on this dataset, and they have provided a stipend for me to continue to work with them from Febuarary 2018 to August 2018. There are several ways I want to develop this project beyond its current form.

I am interested in developing more complex quasi-geographic encodings that are able to display more than one value per CU. While the tree plot provides important information about connectivity and one measure of population health, it is an oversimplified view of a complex system. Displaying more complex data such as a uni- or multivariate time series in place of the status glyphs would provide users with more granular information and could improve their ability to discover trends and explore the topology of the system. In [3], the authors designed a quasi-spatial representation of arteries called **HemoVis** that encoded both endothelial shear stress (a marker of disease) and the topology of the arterial tree. This idiom presented dense information in an abstraction that removed less important data such as the three-dimensional shape of the arteries while still maintaining meaningful spatial and topological relationships. I would like to employ a similar idiom to improve or supplement the tree plot, perhaps by replacing the CU marks with plots of the key metrics (see figure 11). While this is somewhat similar to a small multiple view (which users reported being less effective than other representations), grouping them based on geospatial similarity and connectivity might enable users to recognize patterns or similarities more readily.

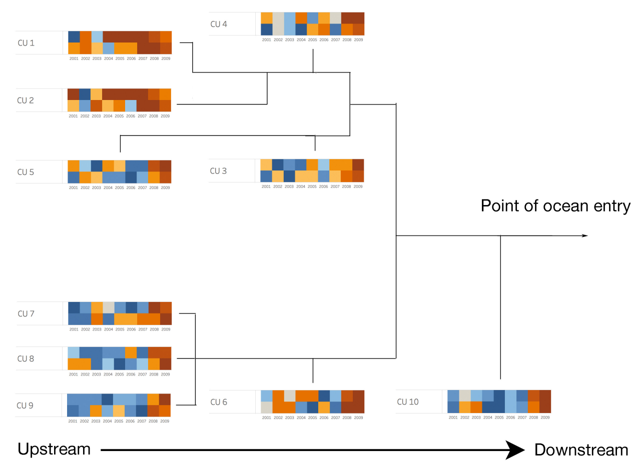



*Figure 11: A mock-up of a quasi-geographic idiom. The function of this idiom is similar to that of the tree plot, but this plot encodes 20 items per CU rather than the single attribute encoded in the current tree plot. Using heat maps or small multiple line plots as the mark for each CU would allow much greater amounts of data to be represented than is currently possible in the tree plot.*

I believe the work could also be improved by incorporating more user research into the design process. I intend to perform a series of semi-structured interviews with DFO scientists and science managers in order to better understand the tasks they perform related to salmon population datasets. The tasks listed previously in this paper were generated through conversation with the two biologists on the SoS team; this is a small subset of the potential end users within the department and speaking with more end users would probably allow us to characterize more tasks that visualization could aid.

In order to develop a more comprehensive list of these tasks, Sue Grant has recruited ~15 additional scientists within the department who are willing to participate in a semi-structured user interview (see supplemental materials: "State of the Salmon Questionnaire"). This questionnaire asks users to describe the tasks they perform, what data they need to perform those tasks, what tasks they would like to perform but are unable to, what barriers they encounter that prevent them from completing those tasks, and asks them to rate the value of various attributes by which salmon data could be aggregated. There are also biographical questions (education, position, length of tenure with the department) that will help us identify potential subgroups within the users we survey. Identifying these subgroups could help us categorize the tasks by user groups and thus develop visualizations that have greater specificity for certain types of end users.

In a similar vein, I am interested in performing user testing with these same scientists on future iterations of the visualizations. One merit of well-designed information visualization is that it allows users to rapidly explore large datasets and perform abstract tasks such as developing ideas and insights about those datasets. However, the concepts of 'understanding', 'ideas' and 'insights' are somewhat vague, and it was not immediately obvious to me how to measure a visualization's ability to promote these things among our end users. Given that we are designing the State of the Salmon visualization in part to encourage 'understanding', 'ideas' and 'insights', it seemed important to operationalize these terms and develop methods to evaluate how well our visualization accomplishes those goals.

In [4], Saraiya et al developed methods to evaluate the ability of visualization to achieve goals similar to those listed above. I have adopted and built on their user testing methodology so that it is relevant to the SoS project (see supplemental materials). Having quantitative measures of "insight", etc., will help us determine how well our tool is performing.

## Conclusions

This dashboard improves the ability of fisheries scientists within the DFO to explore information about salmon within the Fraser river. Existing techniques limited the ability of fisheries scientists to make comparisons across CUs, which is a task of central importance to their jobs. Providing an interactive visual interface should improve their ability to explore and consume the data. The development of quasi-geographic encodings such as the tree plot allow users to explore the topology of the Fraser river in ways they were not able to before, and provide an important emphasis on connectivity between attributes which has not been addressed by previous visualizations in the department. The development of more quasi-geographic plots could further the ability of scientists within the department to explore these relationships, and appears to be a promising area for further work. The use of more extensive user surveys and interviews also promises to improve my understanding of the tasks fisheries scientists need to complete as part of their job, and will in turn lead to the development of more effective visualizations.

## References

[1] S.C.H. Grant and G. Pestal. (2012). Integrated Biological Status Assessments Under the Wild Salmon Policy Using Standardized Metrics and Expert Judgement: Fraser River Sockeye Salmon (Oncorhynchus nerka) Case Studies. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2012/106. http://waves-vagues.dfo-mpo.gc.ca/Library/349637.pdf

[2] L.B. Holtby and K.A. Ciruna. (2007). Conservation Units for Pacific Salmon under the Wild Salmon Policy. DFO Can. Sci. Advis. Sec. Proceed. Ser. 2007/070. http://waves-vagues.dfo-mpo.gc.ca/Library/334860.pdf

[3] M. Borkin, K. Gajos, A. Peters, D. Mitsouras, S. Melchionna, F. Rybicki, C. Feldman, and H. Pfister. Evaluation of artery visualizations for heart disease diagnosis. IEEE TVCG. 2011;17(12):2479–2488.

[4] P. Saraiya, C. North, and K. Duca. An insight-based methodology for evaluating bioinformatics visualizations. IEEE Trans. Visualization and Computer Graphics (TVCG), 11(4):443-456, 2005