Type or paste the title of the work here

**by**  
**[Volodymyr Kozyr]**

[Previous degree, institution, YEAR]  
[Previous degree, institution, YEAR – if more than one list in descending order]

[Thesis/Project/Extended Essays] Submitted in Partial Fulfillment of the  
Requirements for the Degree of  
[Master of…/Doctor of…]

in the  
[School/Department of …]  
[Faculty of …]

© [Volodymyr Kozyr] 2021  
SIMON FRASER UNIVERSITY  
[Term YEAR]

Copyright in this work rests with the author. Please ensure that any reproduction   
or re-use is done in accordance with the relevant national copyright legislation.

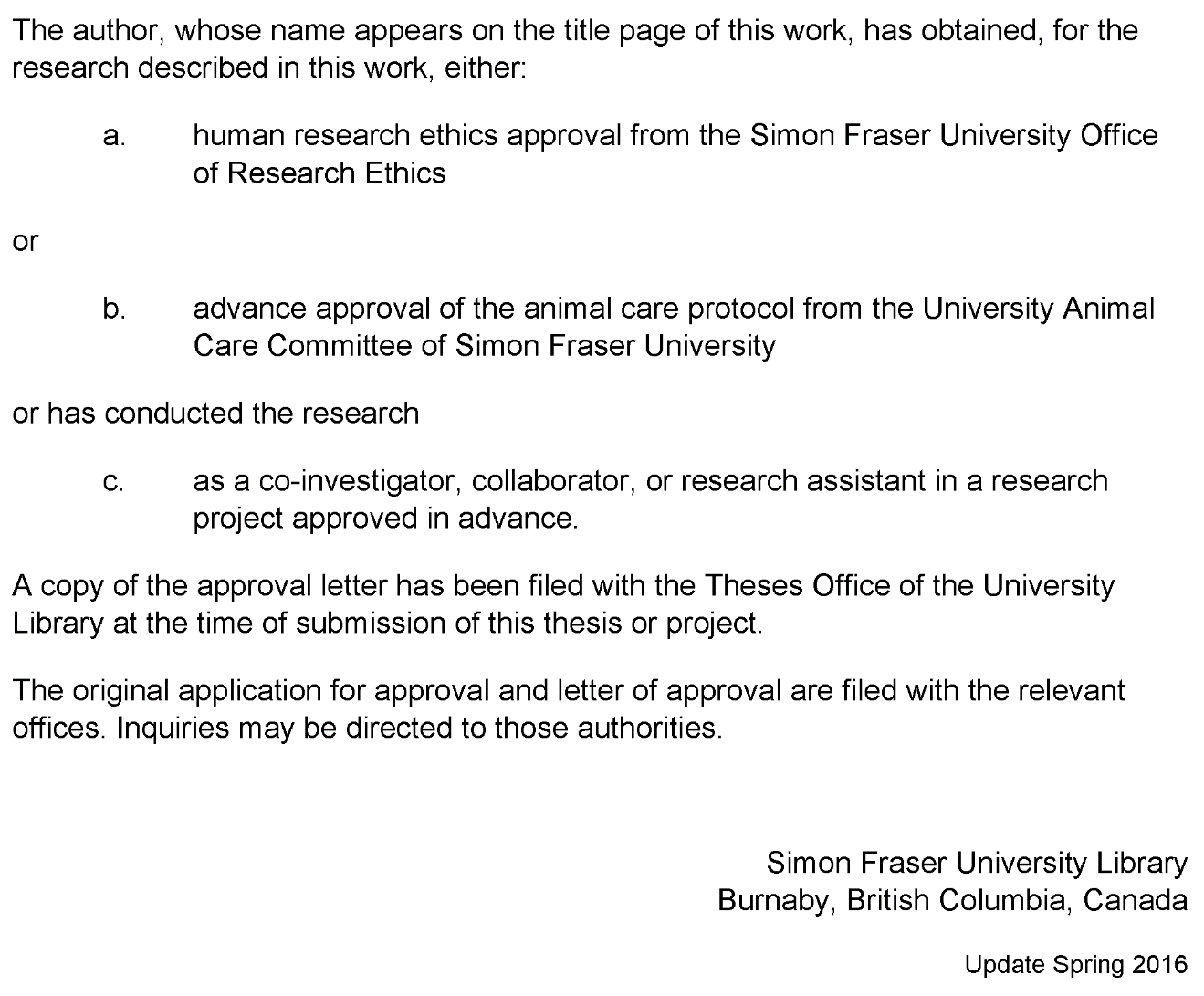
Declaration of Committee

|  |  |  |
| --- | --- | --- |
| **Name:** | **[Firstname Lastname]** | |
| **Degree:** | **[Master of …/Doctor of …]** | |
| **Thesis title:** | **Type or paste the title of the work here** | |
| **Committee:** | **Chair:** | [Firstname Lastname] [Academic Role, Academic Unit] |
|  | **[Firstname Lastname]** Supervisor [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** Committee Member [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** Committee Member [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** Committee Member [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** Committee Member [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** Examiner [Academic Role, Academic Unit] | |
|  | **[Firstname Lastname]** External Examiner  [Academic or Professional Role, Academic Unit] [Institution] | |

Please review the Declaration of Committee page instructions on the library’s thesis website before completing this page:   
[**https://www.lib.sfu.ca/help/publish/thesis/format/declaration-committee**](https://www.lib.sfu.ca/help/publish/thesis/format/declaration-committee)

Remember to delete this note before submission.

Ethics Statement



This page must appear as p. iii if your research required ethics approval. Remember to delete this note before submission.

If an ethics exemption was granted, this page must be removed.

If ethics approval was not required, this page must be removed.

Remember to update the Table of Contents after removing this page.

Abstract

Fishery data visualization plays a vital role for companies in countries with developed fishery infrastructure.

This paper will show trends, anomalies, year-by-year comparison (drill-down) of fishing catches and money value.

Showing correlation between these two parameters for each Canadian province and fish species is a prominent part of the work.

This research will provide a user-friendly way to show all necessary data for non-domain expert users in a web browser, using the novel js library amCharts.

The tool will help users see/identify fish amounts in specific regions and help fishery managers add or remove restrictions on fish quotas. It will allow answering questions about the number of caught species or possible threat to biological diversity.

Having the visualization tool in a web browser is very handy for users because they do not have to install software on their computers that may not be compatible with OS.

The proposed web-interface will be a highly accessible multi-platform tool that will help users interested in the fishery domain analyze table data quickly and effectively.

**Keywords**: fishery; web-interface; visualization; online platform

Dedication

This is an optional page. Use your choice of paragraph style for text on this page (**1\_Para\_FlushLeft** shown here).

To hide the heading at the top of this page, select the text and change the text colour to white.

Acknowledgements

This is an optional page. Use your choice of paragraph style for text on this page (**1\_Para** shown here).

Table of Contents

[Declaration of Committee ii](#_Toc67830726)

[Ethics Statement iii](#_Toc67830727)

[Abstract iv](#_Toc67830728)

[Dedication v](#_Toc67830729)

[Acknowledgements vi](#_Toc67830730)

[Table of Contents vii](#_Toc67830731)

[List of Tables ix](#_Toc67830732)

[List of Figures x](#_Toc67830733)

[List of Acronyms xi](#_Toc67830734)

[Glossary xii](#_Toc67830735)

[Preface/Executive Summary/Image xiii](#_Toc67830736)

[Chapter 1. Introduction 1](#_Toc67830737)

[Chapter 2. Related Work 2](#_Toc67830738)

[2.1. Role of the Fishery Visualization 2](#_Toc67830739)

[2.2. Marine Environmental Management 2](#_Toc67830743)

[2.3. Approaches to Visualization 4](#_Toc67830744)

[Chapter 3. Design and Use Cases 6](#_Toc67830745)

3.1. [Fishery Reports 6](#_Toc67830746)

[3.2. Data Sources 12](#_Toc67830749)

[3.3. Fishery Domain Problems 13](#_Toc67830750)

[3.4.](#_Toc67830751) [Visualization Motivation 15](#_Toc67830752)

[3.5. Task Abstraction 16](#_Toc67830753)

[Task 1. Exploring Relationships for Fish Amount and Price 16](#_Toc67830754)

[3.5.1. 16](#_Toc67830755)

[3.5.2. Task 2. Paired Time Series for Fish Amount and Price 16](#_Toc67830756)

[Task 3. 17](#_Toc67830758)

[3.5.3. Identifying Top Fish Species by Catch Amount or Price 17](#_Toc67830759)

[Task 4. 18](#_Toc67830760)

[3.5.4. Consequent Years Fishery Data Comparison 18](#_Toc67830761)

[Chapter 4. Implementation 19](#_Toc67830762)

[4.1. Overview of the Tool’s User Interface 19](#_Toc67830763)

[4.2. Data Processing 19](#_Toc67830764)

[4.3. Visualizations Overview 20](#_Toc67830765)

[4.3.1. Task 1 20](#_Toc67830766)

[4.3.2. Task 2 21](#_Toc67830767)

[4.3.3. Task 3 21](#_Toc67830769)

[4.3.4. Task 4 22](#_Toc67830770)

[Chapter 5. Evaluation 24](#_Toc67830774)

[Chapter 6. Conclusions and Future Work 25](#_Toc67830775)

[Chapter 7. Bibliography 26](#_Toc67830776)

[New References (not used yet) 28](#_Toc67830777)

[Appendix A. An Example of an Appendix 30](#_Toc67830778)

**Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.**Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.**Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.**Error! Hyperlink reference not valid.**Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.**Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.Error! Hyperlink reference not valid.

List of Tables

**No table of figures entries found.**

Use References>Insert Caption to create caption labels and numbers. Right-click on the text above and select Update Field to update this list. See the Thesis Template Instructions for information on creating tables, figures, and captions.

Remember to delete this note before submission.

If there are no tables in the document, remove this page.

List of Figures

**No table of figures entries found.**

Use References>Insert Caption to create caption labels and numbers. Right-click on the text above and select Update Field to update this list. See the Thesis Template Instructions for information on creating tables, figures, and captions.

Remember to delete this note before submission.

If there are no figures in the document, remove this page.

List of Acronyms

|  |  |
| --- | --- |
| SFU | Simon Fraser University |
| LAC | Library and Archives Canada |
|  | This is an optional page and can be removed if not used. |
|  | Use one table row for each item to allow sorting using Word’s table tools. |
|  | Apply the style **1\_Para\_NoSpace** to table rows as shown here. |
|  |  |
|  |  |

Glossary

|  |  |
| --- | --- |
| Thesis | An extended research paper that is part of the final exam process for a graduate degree. The document may also be classified as a project or collection of extended essays. |
| Glossary | An alphabetical list of key terms |
|  | This is an optional page and can be removed if not used. |
|  | Use one table row for each item to allow sorting using Word’s table tools. |
|  | Apply the style **1\_Para\_NoSpace** to table rows as shown here. |
|  |  |

Preface/Executive Summary/Image

This page can be used for a Preface, Executive Summary, or introductory image. This is an optional page and can be deleted if not used.

To hide the heading at the top of this page – e.g., if using an introductory image – select the text and change the text colour to white.

Do not delete the section break that follows this paragraph. If the section break is not visible, turn on non-printing characters using the Show/Hide icon (¶) on the Home ribbon.

# Introduction

The primary motivation for this research is to help users to make their work faster and more productive. Thus, a software development company I am working for frequently meets requests for giving access to raw table data and data processing, analysis, and making some reports.

Of course, digging into raw data might give the expected results, but it usually takes much time to analyze a significant amount of data, especially if users are not familiar with computer software.

Visual data representation plays a crucial role in data analysis. It can condense vast amounts of data into several plots and labels, giving information about trends; it is also much easier to compare pictures than data rows for sure.

This research thesis will help people who are domain experts but not very familiar with data analysis tools. The tool can help fishery management regulate fishery catching in certain Canadian provinces and help decide which policies or fishing quotas for specific fish types should be applied.

# Related Work

## Role of the Fishery Visualization

A visualization is an essential tool for decision support in fisheries information systems. It can give a person who is working with fishery information more insights about data. It usually saves time for making correct decisions about the fishing company's business logic because it is easier to see trends and outliers while using charts and interactive diagrams rather than just looking through a spreadsheet, sometimes located even in multiple files.

There are different groups of users in the fishery domain. Papers discussed in this chapter are oriented primarily for fishery management. However, the content is complicated for readers who are not data analysts or/and experienced computer users.

So, the work's primary goal is to create a tool that will be easily accessible for both types of users: data scientists and fishery management, making decisions based on conclusions of the data presented in a tool.

## Marine Environmental Management

FishCAM2000 (FC) [1] is a computer-based integrated information system for fisheries management and marine environmental monitoring. It illustrates a visualization of the fishery activity over the same spatial area of interest for a particular type of fish. It has a user-friendly and straightforward interface implemented in Windows Forms.

Figure 2.1.1 below shows the output after the user passes seven screens of settings of the query wizard. It takes too many steps for a user to get a result, but it must be as generic as possible. The excellent point is that it presents complex geodata on the map instead of the data table, which is a considerable time-saving for users to understand the query output.



Figure 2.1.1

Subsequent work, which is to mention, is a tool that gets and analyzes data that is directly coming from vessels [2]. It also depicts geographical data and the amount of fish caught by a particular vessel in a specific region. Authors show not only the geographical spread of fish but also year-to-year comparison charts (Figure 2.2.2).

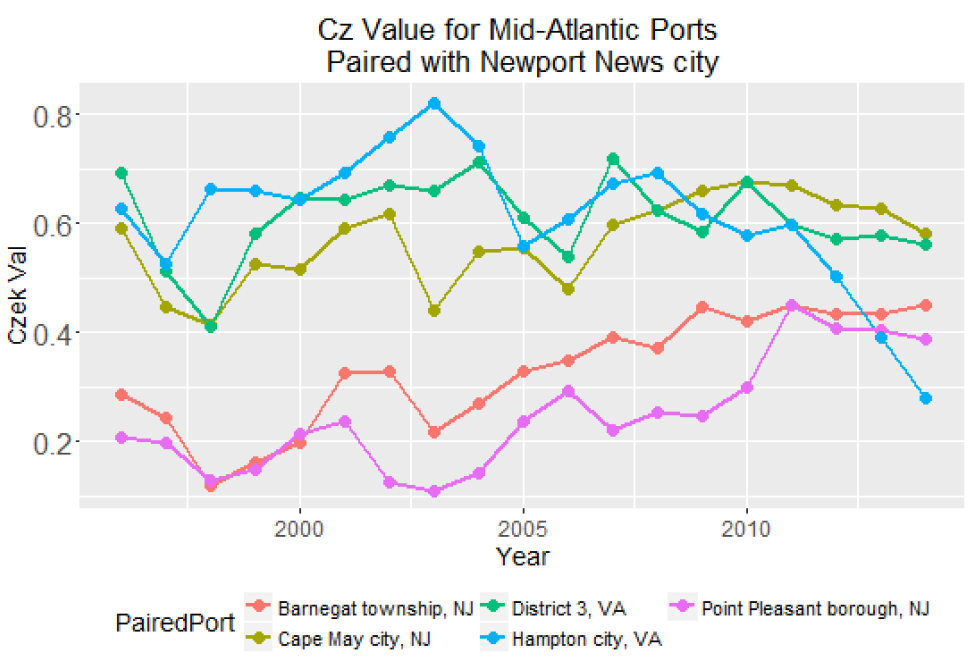


Figure 2.2.2

Whereas the works we have discussed so far focussed on the visualization of geographical and vessel-related data, the proposal of da Silva, Charles Fulcher [3] allows the investigation of land-sea connections. It shows the reader the human impact on the sea from land and vice versa. The maps show the connection between vessels and ports and depict the distribution of gear types used in different regions using descriptive labels, lines, and appropriate legends that any person can easily understand.

At the same time, the work of Barrus about the state of the salmon [4] provides a reasonable basis for a platform to support analysis and interaction for fishery data through visualization. It uses DFO data related to salmon in BC rivers, and the author discusses and tries to produce visualizations for questions that marine experts asked him to solve. In particular, in his work, he states that all data sources for DFO are decentralized and there is practically no interface which allows user to analyze data. As a result, he combined several data sources and produced visualizations that answer questions about the state of salmon in Canadian rivers.

## Approaches to Visualization

The next paper to be discussed with time series prediction on stocks [5] mostly shows the difference between prediction algorithms, but the visualization part requires special attention. Authors used different colour coding to show results, and there is no extra information on the chart which distracts or confuses readers.

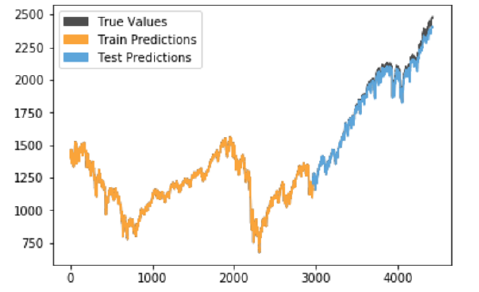


Figure 2.3.1

IDMVis [6]: a visualization tool for a patient with diabetes that shows multidimensional interrelated data during the day. IDMVis includes a novel technique for folding and aligning records by dual sentinel events and scaling the intermediate timeline. The designed tool helps doctors track the state of patients' important parameters and detect anomalies. After that, doctors use it as a decision support tool for the treatment of diabetes. Six clinicians evaluated design decisions positively.

Papers discussed in this subsection mainly explain one visualization technique, but Sofia Semikina, in her thesis work, Stress Data Visualization [7], compares methods of visualizing the same data in various amounts of charts and diagrams of different types. She uses bar charts, line charts, pie charts, spiral charts. Sofia Semikina also includes users' studies in her work, showing which particular visualization users understand better than others.

# Design and Use Cases

The environment and nature change over time, so scientists and environmentalists interested and responsible for the sustainability of different fish populations are using data analysis tools to determine if there are problems with biodiversity in a particular region. For instance, at the species level, fishing can reduce the abundance and alter physiology and lifestyles, thereby affecting species' role in the biological community. Besides, fishing can cause changes in the open ocean community's trophodynamics and reduce the biodiversity and resilience of these ecosystems. That is why scientists work to detect any unexpected fish population trends and design methods and steps to address the issue if one exists. For getting to such conclusions, they are going through an enormous amount of data. In these cases, visualization tools help them to understand the problems faster. Policymakers and fish companies' stakeholders are also involved in discussing the fishery data state. Policymakers need to issue quotas for the fishing season in a particular water zone for a specific fish type to avoid such a problem known as overfishing.

Visualizations play an essential role in communicating for fisheries scientists, environmentalists and policymakers. In the article “On the Role of Visualization in Fisheries Management” [8], Polina Levonin discusses how policymakers, stakeholders, and fishery management communicate with each other and provides a visualization tool that makes the process easier on all sides. From the articles and papers discussed in the following paragraphs, it is becoming clear that more and more specialists are preoccupied with the fish population in the world ocean. There are groups of users interested in fishery data analysis for different purposes, but they need visualization tools to make their work more productive.

## Fishery Reports

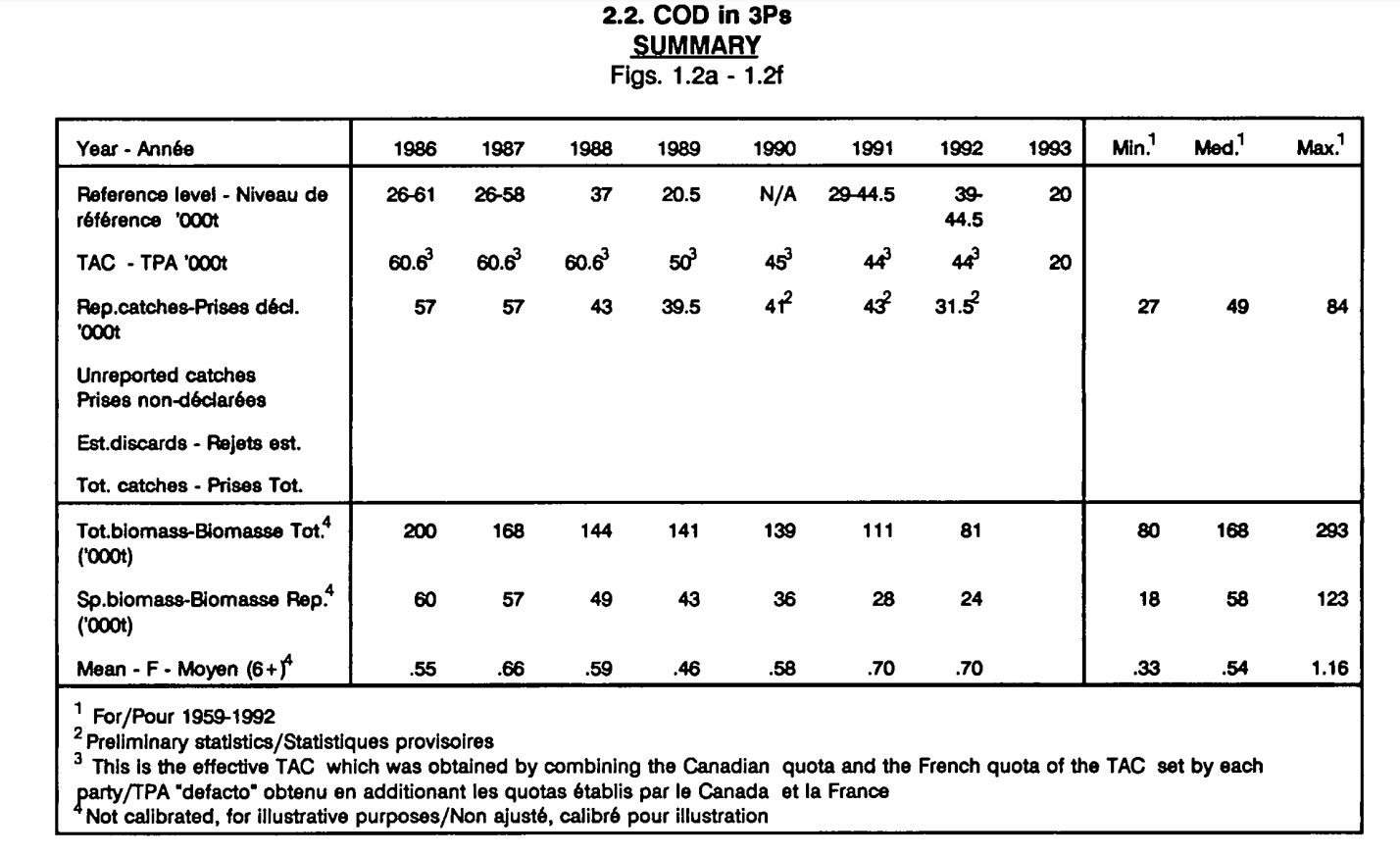
Reporting on the status of aquatic resources and how they are changing over time, along with understanding the impacts of large-scale disturbance and human activities on aquatic resources, are among the most critical processes for assessing the state of ecosystems.

To display the human influence on an ecosystem or a specific type of fish, data on the amount of catch of this species by region, the number of the species in comparison with previous years, data from the SARA and COSEWIC organization about species at risk can be helpful [9].

For example, the 1993 Report on the status of groundfish stocks in the Canadian Northwest Atlantic [9] reported cod catch data. The authors used multiplicative analyzes of catch-at-age from research surveys and the commercial fishery and, on average, fishing mortalities from preliminary SPA.

Based on these data, it was concluded that fishing mortality increased according to multiple reasons (page 17) and that there has been a “gradual decline in stock size through the second half of the 1980s”.

This conclusion can be drawn from the following tabular presentation of the data, where in the total biomass section user can see that for every next year amount of cod is decreasing.



“Table Data Presentation” (Figure 3.1.1)

Similarly, this tabular format allows the reader to see, in the same report, data on the amount of cod catches from 1960 to 1994. The user can compare these data with the data on the population of different age groups of this species, fish mortality, recruitment, supply and demand and other data. For a person who uses such kind of data representation, some additional steps should be done to come up to colslusions. User has to look for the value of interest and compare each column year by year to figure out if there is a trend pattern (increase/decrease) in the data.

Instead of providing information in tabular format, line charts presented in different colours give the reader a clear and quicker understanding of trends and processes ongoing. The Stock Status Report [10] for the 2004 year provided abundance data for Northern Abalone for each year (see Figure 3.1.2).

This figure shows that since 1998, this species' abundance has fallen below the permissible short-term recovery objective line. This can be inferred from a figure 3.1.2 listed below. After year of 1998, data points are located below the line which is located on the level of 40%.

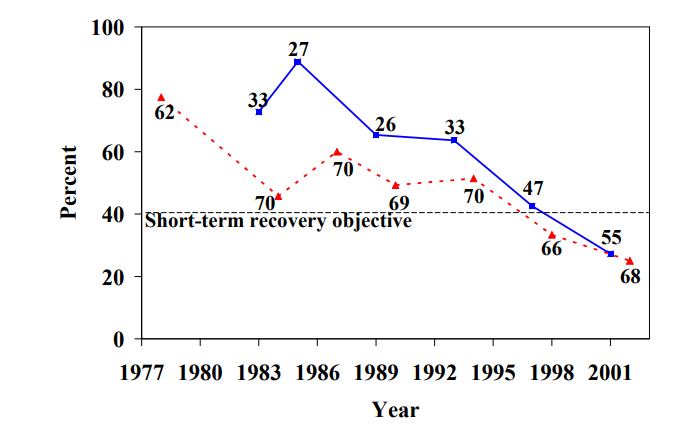
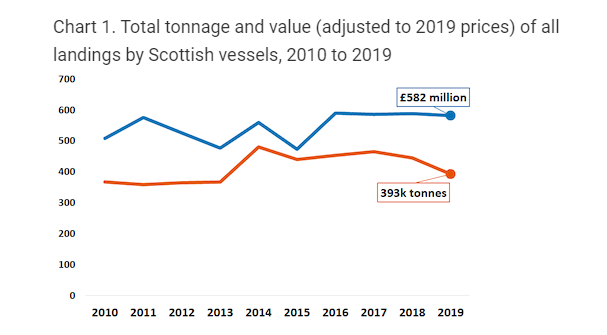


Figure 3.1.2

The same approach of using line graph visualizations can be seen in a more recent annual environmental report on this kind of data. For example, Scottish Sea Fisheries Statistics [11] for 2019 from the Cabinet Secretary for Rural Economic and Tourism provides summary data and more detailed statistics by region.

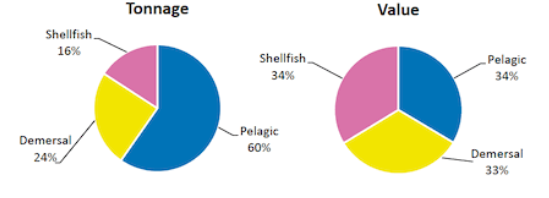
From the information related to this topic, the following can be distinguished:

The total amount of fish caught (tonnage) is shown in blue, and the value of all landings by Scottish vessels is shown in red. By comparing the blue and red graphs, the user can easily and quickly determine the connection (will be discussed in the following chapter) between the catch's value and the amount of fish caught in a specific year. For instance, the user can see that despite the tonnage falling since 2017, the value of landings remains constant.



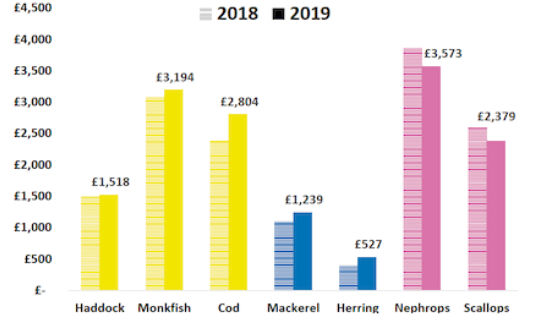
* **Percentage of Scottish vessels’ landings by species type in 2019**

We can see that marine industry requires to see overall reports on such parameters as “tonnage” and “value”. From the pie charts users can determine which type of fish gives which revenue according to tonnage.



* **Real terms price per tonne for key species (value of £20 million or over landed by Scottish vessels) 2018 and 2019**

On the next report there are bar charts for two consequtive years presented. This is done mainly to see trends and then to decide if fishery industry doing better or worse than previous year. After comparing values for two years some adjustments in fishery policies could potentially be implemented by ecologists or fishery companies.



The Scottish Sea Fisheries Statistics report [11] is an excellent example of data visualization to assess both the environmental and economic side of a problem. The reports described above cover different periods of time and fish types.

Summarizing the importance of data vizualization in comparing a larger number of data it can be seemed that the data provided is very similar and differs only in the methods of their visualization. Some provide tabular information, others provide graphs, line plot, bar chart, pie chart, scatter plot or bubble chart. Moreover, each of the visualizations is aimed at displaying specific data for a particular user's request. By the fact that this report is used on the government level it can be concluded that these data visualizations are relevant and will remain relevant for a long time for specific users in fishery domain.

However, there are several disadvantages contain in all the examples discussed above.

For instance, in tables, data for information analysis is presented exclusively in vertical and horizontal numbers, and there are no visible accents, like, colour. In this case, it slows down and complicates practical visual assessment and analysis of information.

In turn, when in other examples the information is presented using colour graphics, there are no labels that explain the missing data for analysis.

Consequently, even with these examples, it is difficult for a consumer to comprehensively evaluate and analyze information.

Accordingly, in all the described work examples, information is presented in static images, which does not allow interactive intercommunication with information and work with data in real-time.

## Data Sources

According to National Research Council. 2000. Improving the Collection, Management, and Use of Marine Fisheries Data. Washington, DC: The National Academies Press. “The phrase “fisheries data” is a general way of referring to data that may be of use in the management of a fishery as well as for commercial, recreational, cultural, and scientific purposes.”

### Such data includes biological information on operational fish and related species, economic information from fishermen and markets for catch and environmental information affecting the species' productivity.

### A primary source of information is fishery-dependent data. So-called trip tickets or logbooks contain the timeliest information on current fishery conditions.

### Secondly, catch sampling programs are an essential source of information provided by observers placed on commercial fishing vessels to collect data based on the species composition, sex ratio, and age composition of the catch.

### And finally, scientific surveys are the primary source of fishery-independent data, including estimates of fish populations' age structure and relative abundance of stocks.

The scope of fisheries data use is very wide. It ranges from stock assessment by scientists to strategic planning by industry, and fishery monitoring and allocation decisions by managers.

Wherein, as mentioned in “Improving the Collection, Management, and Use of Marine Fisheries Data” report, “each use implies a set of users and a suite of requirements that the data must satisfy, including timeliness, level of detail, accuracy, accessibility to users, coverage or completeness, and credibility of the data collection process and the management process that uses the data.”

We can emphasize that fisheries data are vital to strategic planning activities in coastal communities that rely on fisheries. At the same time, civil authorities use fisheries data to site marinas, underwater pipes and cables, and other maritime facilities and develop infrastructure for the fishing industry. Besides, bankers use fisheries data to plan economic development and loan packages to fishermen, fish processors, and ship suppliers.

And of course, fishermen themselves, besides their own logbooks and observations and what they learn from other fishermen and buyers, use government fisheries data to plan shifts to new fishing grounds, changes in fishing gear, and changes in species targeted. Environmental and other interest groups also have become increasingly involved in monitoring fishing activities. Monitoring often requires data with great detail in both time and space as well as frequent updates, often within a fishing season.

We need to note that among most vital users of fisheries data stand stock assessment, government and private sectors’ scientists, national, and international fishery agencies and university.

From the above, we must consider that when creating software for analyzing fisheries data, we must assume that users' requirements will vary. The system should help users who require data with various spatial resolutions and different levels of timelines.

Data source for the visualization is taken from DFO Canada website for provincial seafisheries for several reasons. At first, data is provided on a Canadian government website, so users can rely on the data with high level of trust. Secondly, its publically available data, so users can check the source data if they want. And thirdly, data is updated on regular basis, so it will be available for the following years as well (not only histarical data).

Format: Data tables provide the volume and value of seafisheries landings. Data is organized by species-groups, by main species and by province.

Period covered: Data is available from 1990 to 2018 year before the current fishing season, data prior to 1990 is available upon request.

Sources: Data collected by DFO regional offices.

Data is separated yearly (one Microsoft Excel file for each year) and grouped inside by Canadian provinces for each fish type. In total this dataset consists of 56 files (28 for fish amount in tonnes and 28 for fish total value in thousand of Canadian dollars.

To present such kind of data, it should go through several processes which includes data extraction, cleaning/filtering and massaging (converting into suitable format) which will be discussed in implementation part of the work (chapter 4).

3.3. Fishery Domain Problems

It’s also worth mentioning problems on which marine scientist are eager to work on accordind to papers and researches.

1) determination of the optimal amount of catch for each type of fish to reduce environmental damage in a specific region. The visualization of the problem analysis can be presented in the form of a bar chart growth/decrease in the catch of a specific type of fish for a particular period in different regions.

2) understanding how the catch affects the ecosystem and other species. The visualization of the problem can be presented in the form of graphs of two colours. One of them shows the increase/decrease in the catch, and the graph of the other colour shows the diversity of the ecosystem in the number of species

3) find out what species of fish and other animals are on the verge of extinction. Visualization can be depicted using pie charts showing endangered species by year.

4) identifying the regions that need attention in the first place and introducing measures to restore or prevent species' extinction or changes in the ecosystem. The problem can be visualized in a geo map, where endangered species in a specific region are shown in different colours.

5) predicting which species may also be subject to negative or positive effects (trends). A line chart would be appropriate here, defining the trend in the development of species over some time.

6) the establishment of quotas (the proper amount of unloading of fish) which will minimize the negative impact on the environment. The visualization could be a bar graph showing the dependence between the allowable catch and the accepted catch quotas.

7) determining the recovery rate of the ecosystem after taking appropriate measures can be displayed using two line graphs of different colours.

8) analysis of the safety of methods for catching a particular type of fish in each region. A comparative bar chart can be used to visualize the relationship between safe fishing practices and the fish population in a given area.

Based on the results of the analysis, scientists adjust methods that have adverse effects on the state of the environment and suggest alternative methods.

In a 2008 publication of Bedford Institute of Oceanography [12], which described many factors affecting different ecosystems, aquaculture research, various data collection methods and technologies, the authors mentioned Recovery Objectives for North Atlantic Right Whales, among which the following issues were described:

- reduction of death and injury from vessel strikes

- reduction of death and injury from fishing gear entanglement

- reduction of injury and disturbance from vessel noise, exposure to contaminants, and other forms of habitat degradation

- monitoring populations and threats

Based on these data (amount of injured or accidentally killed animals and detailed reports on each case) it can be concluded that the environment is adversely affected by the rising amount of fishing vessels and equipment for fishing in the ocean.

The Federal Register - The Daily Journal of the United States Government [13] provides data on Marine Mammal Species and Stocks Incidentally Killed or Injured during the 2018 fishing catch.

By obtaining the average annual data on the number of accidentally killed or damaged fish during the catch for Canada and comparing it with the data about the number of fish caught in the same year, it is possible to estimate the environmental damage. Depending on the amount and detail of the data, it is possible to estimate the damage for a specific type of fish and/or for each region. Also, fish species can be combined into groups according to their catching methods, and ecological influence can be analyzed for each group.

## 3.4. Visualization Motivation

Analysis of the data presented in table or text format may take significant amount of time [cite report paper]. For example, it is hard to see trends, how values change through the years, as well as comparing data for different provinces and fish type. Another thing which is hard to capture is correlation [report paper] between price and quantities of a specific types of fish for a certain period.

The tool itself and it’s implementation will be discussed in Chapter 4, however, it is worth mentioning that it is developed for people who may not be data scientists. The main goal is to make it usable for people with average knowledge about computers. It also will not require any installation steps, because it is a web application which can be accessible just by typing url in any of the modern browsers. Another feature of the tool is that it will allow user to select range, provinces and any fish type from dropdowns, zoom into details etc. This type of UI experience if not available if it is done through Excel charts or Python library PyPlot. Visualizations in these cases are static and should be re-rendered if some parameters of visualization are changed. Also it requires advanced knowledge of Microsoft Excel or programming.

Design requirements for visualization were set based on the informal task abstractions discussed above and literature research (real reports which people use for marine operations from section 3.1).

**Requirement 1. Interactivity**

The main difference between existing static reports and this tool is that tool allows user to interact with data before producing visualization (filtering, zooming etc.)

**Requirement 2. Data Scaling for Further Analysis**

Tool should allow users to discover patterns, trends and anomalies (Task 1 and Task 2)

**Requirement 3. Summary and Overall Statistics**

Visualizations for Task 3 and Task 4 are designed to show global summarized data.

## Task Abstraction

### Task 1. Exploring Relationships for Fish Amount and Price

In section 3.1, we introduced a report [11] shows that charts for the fish amounts and price are relevant for people who work in marine industry. This work contains plots for fish amount and prices separately on different figures. They also have different visualizations (bar charts, line chart, pie charts) for the prices and tonnage of the fish of specific type.

Improvement of that report would be placing them on one line chart with two axes to help user see both amount at the same time [report paper]. This helps to understand correlation between the values and quantities without a need for looking into two different charts (they may be in different scale etc.) or comparing row table data.

Visualization (combined with external data sources and/or user’s knowledge and experience in the domain) may be used by users for solving range of issues such as listed in chapter 3.1.

### Task 2. Paired Time Series for Fish Amount and Price

Visualization of the paired time series plays an important role in the marine fishery industry, such as in work retaled to parrotfish population by Valle and Oxenford [14]. In this paper there are scatter-plots showing relationships between human population size and fish density for selected fish groups across the Caribbean (fig. 3.5.2.1).

Chart

Description automatically generated

fig. 3.5.2.1

For DFO data visualization was used similar approach but with different time series (fish amount and quantities). One axis is fish quantity, second is fish price and dots represent years which is a third dimension in that case. Visualization shows user the trend of how these values change over time together. Improvements of the visualization comparing to the paper [14] are that it has more advanced filtering, zooming and used interaction. Details of implementation are discussed in the following chapter 4.

### Task 3. Identifying Top Fish Species by Catch Amount or Price

Knowing which types of fish give the biggest income is valuable for the fishermen and management. In British reports [1] there are pie charts which briefly describe the data for the year.

Pie charts presented in this thesis are more advanced. This visualization can help to find recordmens and outliars for the selected year.

### Task 4. Consequent Years Fishery Data Comparison

Similarly to the previous section 3.5.3, there is a visualization for two consequtive years DFO data.

The main point for this kind of visualization is to show fishery management and policy makers trend for the quantities and fish price.

Policy makers and environment workers or biologists are mostly interested in seeing fish quantities trend to determine if the decisions made in previous year lead do the desired results in the current (such as establishing fishing quotas, studying how some kinds of pollutions affect fishing population).

Howewer, fishery management are more looking into fish price to decide which kind of fish has more sense to aim for catching for the following year.

This chart also helps environmentalists and fishery management to communicate with each other. In a way that based on these values fishery management requests which type and how much fish are they planning to catch. And policy makers based on the same plot and internal data and knowledge approve or deny a request.

# Implementation

In this chapter there will be discussions related to the impementetion of the tool, it’s structure, architecture levels and explanations of visualization methods for each use case (task).

The goal of the current thesis project is to create a visualization tool to help people understand/analyze table data in a more suitable format as charts, comparison diagrams, etc.

The tool itself is implemented in TypeScript (wrapper for JavaScript) from Microsoft. The front-end framework is Angular 10, one of the most powerful and highly used web engines. The back end is not needed for now as it is a test project. If the data source is changed, there will be minimum code modifications to get/process data.

The main reason for choosing a web solution for implementing the visualization is that it will be accessible for users without needing the installation of any additional software.

Having everything in a web browser is a modern way of viewing and sharing visualizations. The user doesn’t have to do any extra manipulations with a computer system to access visualizations.

Next sections will be describing tool architechure and design decisions in detail.

## System Overview

1. Data layer

As it was written in the previous chapters, the data for visualization tool uses public data provided by DFO Canada. Data is converted from Excel to JSON format, which is suitable for using it inside web browser environment.

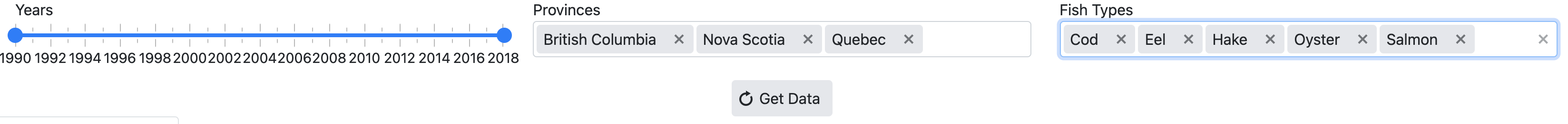
Web browser gets data by using HTTP REST request, therefore if dataset is changed, it can point to any url with file which contains JSON file with data in suitable format for visualization tool to use. So, for the future use any web API can be used as datasource, which is very convenient if several people work with fishery data (one user uploads file with data, second user analyzes it with visualization tool etc.)

1. Logic Engine

As a framework engine, the tool uses TypeScript framework Angular version 10. It is a popular JavaScript based web framework. For presenting data there is amCharts4 library included in the project, which allows to generate sophisticated interactive data visualizations using JavaScript inside web browser.

1. User interface overview

On the top of the page, there are controls which allows users to work with data which includes filtering, selecting, etc. (See picture below)



The date range slider (from 1990 to 2018) allows users to select particular years of interest that users would like to study and create visualizations. This type of control allows user to select start and end year of the time range without a need to select each year individually, which possibly could be 28 actions (28 click for each year presented in the dataset) for user to do. Instead, it is minimized to 2 (just selecting first and the last year in the dataset).

There are two multiple selection pickers for provinces and fish types which are also filtering data. With these controls in place, user has an ability to generate visualizations which are not restricted to a particular province or fish type. This feature makes the tool more flexible than reports which are discussed in the section 1 and 2.

Based on the user input, there may be 4 different visualizations generated.

## Data Processing

The data source comes from DFO (Fisheries and Oceans Canada). It reflects fishing amount catches and money profit for years from 1990 until 2018.

1. Data extraction

At the stage of data extraction phase Excel files are downloading from the official DFO website. It does not require any additional permissions. This data is publicly available. However, the process of downloading is manual. It can be improved if DFO could provide API up-to-date data for use.

1. Data cleaning

Excel files contain some information not relevant for visualization. For example, these are table borders, text colour, other metadata, etc.

So, the irrelevant information is removed at this step.

1. Data converting

The last step of data processing is converting. As an intermediate step, the Excel file converts to CSV. Then CSV converts into JSON file, which can be easily interpreted by visualization library in any modern browser.

Having all these prosessing steps as automated process would be a good improvement of the tool, but DFO data is not that structured (especially for the 1990s), so manual evaluation is required.

But processing of the data is not the goal of the research, so it is enough to mension this brief bullet points.

## Visualizations Overview

In this particular subchapter there will be given visualization solutions for each of the 4 problems and discussion of that how it improved current visualizations (fishery reports and papers discussed in the previous sections 2 and 3).

### Task 1

If the date range selected is more than 2 years, there will be visualization generated as on the figure below.

Chart

Description automatically generated

(sample screenshot)

This visualization will help user to figure out relationships between fish quantities and prices for the particular provinces and / or fish type.

is

### Task 2

Second chart is a scatter plot on which x axes is price and y axes is quantities of particular fish types and province (picture is below). Bullet points and labels represent years.

This visualization helps users to clearly see correlation between price and quantities throughout the selected years.

Chart, line chart, scatter chart

Description automatically generated

### Task 3

If the user selection is only one year, then pie charts for fishing quantities and prices for province will be created.

There is also grouping for fish types which have a small percentage of catch or value compariong to the others which is an improvement (for more than 20 legends pie charts usually look overcrowded and not readable.

The main feature of the chart is that it catches of some fish types are smaller that particular threshold, it will be grouped into “Other” category. This solves a well known problem with having too many unnessessary labels (legends). But on click on the “Other” category user can drill down to see more details for the fish types which are the “outliers” in the dataset.

Chart, sunburst chart

Description automatically generated

### Task 4

If only two years are selected, bar chart report as in “British Reports” will be created. This kind of data presentation is particularly useful for determining trends for the current and previous year. For fishery management it is important to know if, for ex. New policies and laws or fishing quotas are working into intended direction.

Chart itself in an improvement of the bar charts, called categorized bar chart. This visualization will show price and fish amounts for the selected two consequtive years based on province and grouped by fish type.

Timeline

Description automatically generated

# Evaluation

# Conclusions and Future Work

* Bringing different datasources
* Remambering (serialization) and sharing visualizations

Chapter 7. Bibliography

|  |  |
| --- | --- |
| [1] | Scottish Government Riaghaltas na h-Alba, "Scottish Sea Fisheries Statistics," 2019. |
| [2] | Henri Valle, Hazel A. Oxenford, "Parrotfish Size: A Simple yet Useful Alternative Indicator of Fishing Effects on Caribbean Reefs?," 2014. |
| [3] | Z. Kemp and G. Meaden, "Visualization for Fisheries Management from a Spatiotemporal Perspective," 2002. |
| [4] | Sharon Benjamin, Min-Yang Lee, Geret DePiper, "Visualizing Fishing Data as Rasters," 2018. |
| [5] | Hiroshi Ashida, "Spatial and Temporal Differences in the Reproductive Traits of Skipjack Tuna," 2020. |
| [6] | Marlene A. Bellman, Scott A. Heppell, Chris Goldfinger, "Evaluation of a US West Coast Groundfish Habitat Conservation Regulation via Analysis of Spatial and Temporal Patterns of Trawl Fishing Effort," 2005. |
| [7] | Aaron Elliot, Cheng Hua Hsu, "Time Series Prediction: Predicting Stock Price," 2017. |
| [8] | Patricia Pinto da Silva, Charles Fulcher, "Using GIS to Visualize Land/Sea Connections: Case Study - NE Herring Fishery," 2007. |
| [9] | M. Barrus, "The State of the Salmon," 2017. |
| [10] | Yixuan Zhang, Kartik Chanana, Cody Dunne, "IDMVis: Temporal Event Sequence Visualization for Type 1 Diabetes Treatment Decision Support," 2018. |
| [11] | S. Semikina, "Stress Data Visualization," 2014. |
| [12] | Ministry of Natural Resources, "Methods for Monitoring Fish Populations," 2014. [Online]. Available: https://www.ontario.ca/page/methods-monitoring-fish-populations. |
| [13] | Fisheries and Oceans Canada, "Report on the Status of Groundfish Stocks in the Canadian Northwest Atlantic," 1993. |
| [14] | Fisheries and Oceans Canada , "Northern Abalone," 2004. |
| [15] | Bedford Institute of Oceanography, "Fishery Report," 2008. |
| [16] | Federal Register, "List of Fisheries," 2018. [Online]. Available: https://www.federalregister.gov/documents/2018/02/07/2018-02442/list-of-fisheries-for-2018. |
| [17] | Polina Levontin, Paul Baranovski, Adrian W Leach, Alexandria Bailey, "On the Role of Visualization in Fisheries Management," 2017. |
| [18] | Commission on Geosciences, Environment, and Resources, IMPROVING THE COLLECTION, MANAGEMENT, AND USE OF MARINE FISHERIES DATA, 2000. |







New References (not used yet)

Aquatic Species at Risk found in Canadian waters

- <https://www.dfo-mpo.gc.ca/species-especes/sara-lep/identify-eng.html>

Aquatic species at risk map

- <https://www.dfo-mpo.gc.ca/species-especes/sara-lep/map-carte/index-eng.html>

Annual (2008) report about influence on ocean, ecosystems; data collection methods and technologies; species at risk;

RECOVERY OBJECTIVES FOR NORTH ATLANTIC RIGHT WHALES (page 44, maybe 37 – 40…)

<https://waves-vagues.dfo-mpo.gc.ca/Library/353989.pdf>

Report on the status of groundfish stocks in the Canadian northwest Atlantic

pages 17 – 48 => total amount of catches is comparing with other values/indexes

has abundance and mortality rates legends

Page 59 : price/catches legend

<https://waves-vagues.dfo-mpo.gc.ca/Library/40622745.pdf>

Habitat Status Reports (2003-2004)  
<https://waves-vagues.dfo-mpo.gc.ca/Library/281840.pdf>

Competition between Marine Mammals and Fisheries: Food for Thought

Marine mammal bycatch in gillnet and other entangling net fisheries, 1990 to 2011

<https://www.bmis-bycatch.org/system/files/zotero_attachments/library_1/DL6ER5VA%20-%20Reeves%20et%20al.%20-%202013%20-%20Marine%20mammal%20bycatch%20in%20gillnet%20and%20other%20entangl.pdf>

US Government site: list of fisheries and incidentally killed mammals:  
<https://www.federalregister.gov/documents/2018/02/07/2018-02442/list-of-fisheries-for-2018>

BEDFORD INSTITUTE of OCEANOGRAPHY (2000 report) Mortality: page 31  
<https://waves-vagues.dfo-mpo.gc.ca/Library/254420-00.pdf>

Factors Affecting the Responses of Marine Mammals to Acoustic Disturbance  
<https://www.researchgate.net/publication/233638567_Factors_Affecting_the_Responses_of_Marine_Mammals_to_Acoustic_Disturbance>

Sea mammals population decline  
<https://www.pnas.org/content/114/44/11781>

Marine mammal population decline linked to obscured by-catch

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0077908>

Marine Mammal Impacts in Exploited Ecosystems: Would Large Scale Culling Benefit Fisheries?

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0043966>

<https://www.un.org/Depts/los/global_reporting/WOA_RPROC/Chapter_37.pdf>

Farms influence

<https://waves-vagues.dfo-mpo.gc.ca/Library/40654345.pdf>

Appendix A.  
  
An Example of an Appendix