An Analysis of Canadian Marine Accident Fatality Factors

Literature Review, Data Description, and Methodology

Name: Rares Finatan
Student Number: 501140875
Supervisor: Uzair Ahmad, Ph.D
Week of Oct 24, 2022

Table of Contents

Ryerson University

Defining the Research Question

The main topic of research in this study is the analysis of marine accidents occurring in Canadian waters for the years 1995 to 2022. More specifically, the research question is aimed at identifying the key factors involved in Canadian marine accident fatalities and quantifying the respective factors' contributions to serious marine incidents.

A formalized version of the research question reads as follows:

For the years 1995 to 2022, what key factors involved in Canadian marine accidents have quantifiably shown a statistically significant contribution to marine accident fatalities?

Approach of Research

All literature articles that are selected for the purpose of this literature review were found using the ProQuest Ebook Central in collaboration with Ryerson University's library, Google Scholar searches, as well as the Transportation Safety Board of Canada's database on marine transportation safety.

Literature Analysis

Summary of Literatures

A large majority of available studies pertain to marine accident investigations conducted in bodies of water not belonging to Canada. Given the scope of the research question being limited to Canadian waters, only investigations originating from the Transportation Safety Board of Canada are directly applicable. Despite this, international or cross-national studies may prove useful in determining research methods and areas of focus not explored by Canadian investigators.

Investigations conducted on Canadian marine transportation safety incidents are typically conducted on a per-occurrence basis[1]. This means that investigations are conducted on specific maritime incidents, and studies are not often formed to analyze marine accidents on aggregate. Given that the scope of the research questions looks to identify summarized variables influencing marine accidents, the literatures under review will also be considered only if they take on a macro-level approach to the research question.

Canadian-Specific Literatures

1. Transportation Safety Board Investigation M09Z0001: Safety Issues

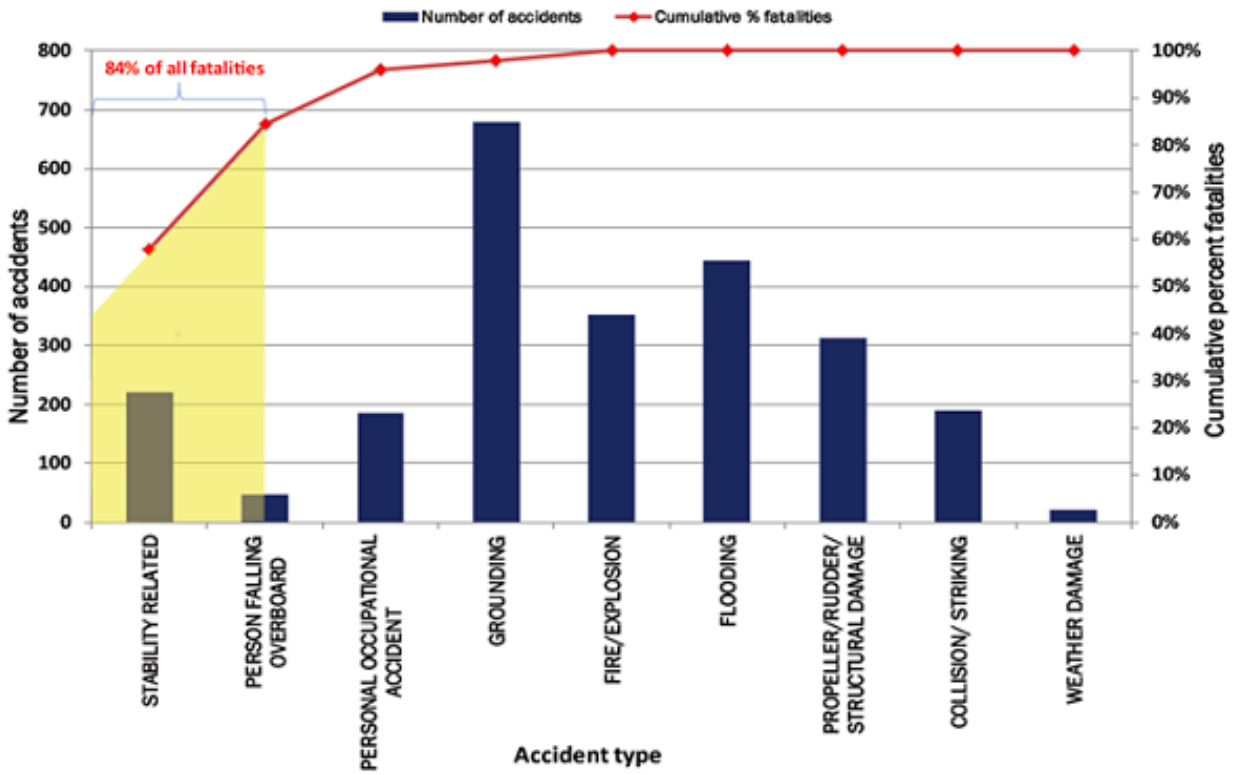
Investigation into Fishing Safety in Canada[2]

This study on marine accidents in Canada is published by the Transportation Safety Board of Canada (TSB), with analysis completed for both public and private sector vessels involved in a marine incident between 1999 and 2010. It is arguably one of the most comprehensive studies of its type, identifying the context behind the impact of commercial vessels and their involvement in marine incident rates, the accident and fatality rates for the respective period, as well as the most pertinent safety issues contributing to marine accident fatalities. The study's results are a compilation of analyses conducted by investigators with expertise across multiple marine disciplines, with a focus on observation and interpretation of 370 investigation reports, 42 safety recommendations, and 100 safety advisory letters published prior to the report release date.

---

[1] Government of Canada, T. S. B. of C. (2019, May 6). *Marine Transportation Safety Investigations and reports*. Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/index.html

[2] Government of Canada, T. S. B. of C. (2012, August 10). *Marine investigation report M09Z0001*. Marine Investigation Report M09Z0001 - Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/m09z0001/m09z0001.html



A visual analysis of marine accident types conducted as part of the study.[3]

While the TSB study is highly exhaustive in preparing future safety documentation for marine safety, it has a few limitations, and thus presents several information gaps that can be attenuated in the currently proposed study.

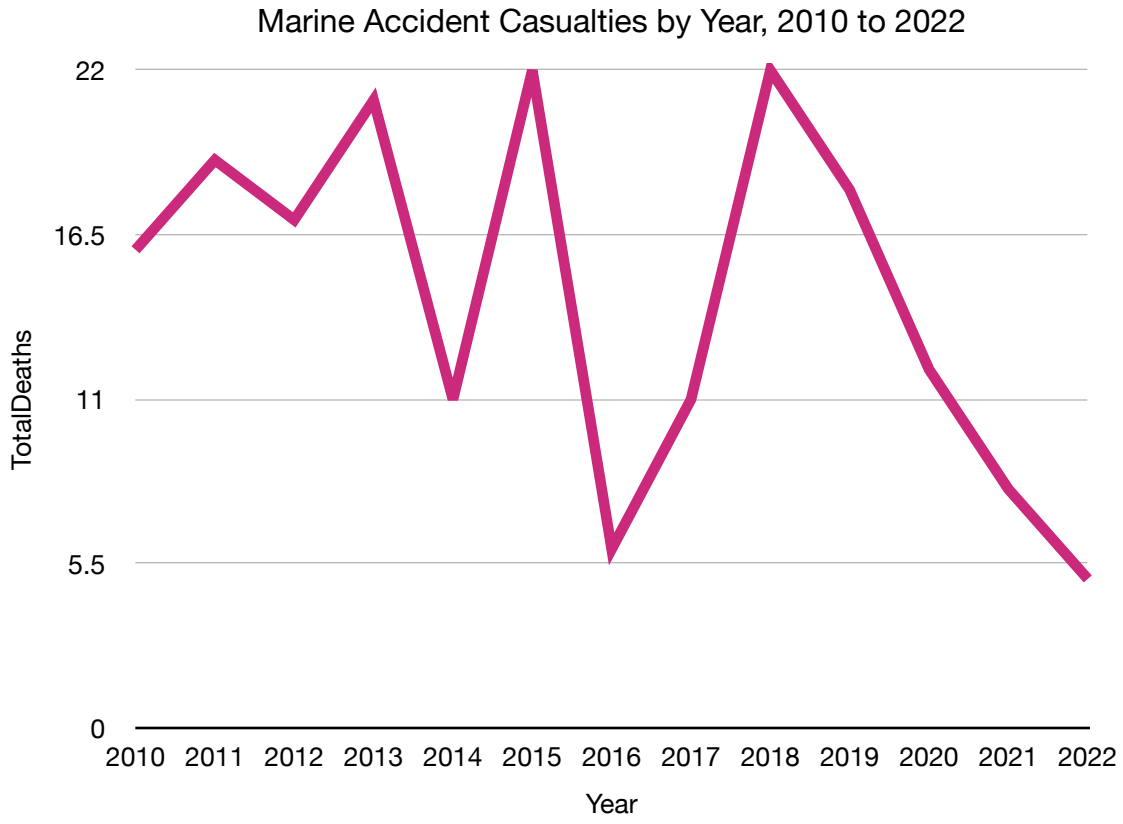
i. The dataset under analysis is strictly constrained to marine vessels operating under the classification of a fishing vessel and does not take into account marine accidents as holistically as possible. More specifically, the investigation does not analyze any marine incidents for the respective time period unless the incident involved a commercial fishing vessel under 24.4 metres in length, or 150 gross tons.

ii. The study does not quantify incident exposure risk from variables that may interact with the 100+ safety actions common to Canadian fishing vessel accidents involved in this

[3] Transportation Safety Board of Canada. (2010). *Figure 4*. Safety Issues Investigation into Fishing Safety in Canada. Government of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/M09Z0001/images/m09z0001-figure-04.png.

study. Despite this, there is a qualitative attempt made by the study to consolidate these findings as part of Table 3, Section 7.

iii. The study only contains data from 1999 to 2010, and may suffer from data freshness. Despite updated marine safety standards since the report's release date, holistic marine accident rates per annum have not decreased significantly since 2010. There may be influencing factors not under consideration in the study that have since become sizeable for observation since the report's release date.



Observed marine accident casualties over time in Canada, from 2010 to 2022.[4]

[4] [4] Transportation Safety Board of Canada. (2019, June 12). *Marine occurrence data from January 1995 to present*. Open Government Portal. Retrieved October 23, 2022, from https://open.canada.ca/data/en/dataset/ad8d1b73-df09-4521-9bdb-61c529328218

2. Transportation Safety Board Investigation SM9501: A Safety Study of The Operational Relationship Between Ship Masters/Watch-keeping Officers and Marine Pilots[5]

This study is the second safety issue investigation released by the Transportation Safety Board of Canada, pertaining to the operational relationship between ship masters/watch-keeping officers and marine pilots. The dataset of the study involved 273 marine accident occurrences between February 1981 and May 1992, whereby collisions, groundings, strikings, contacts, and sinking of vessels were reported. Unlike Transportation Safety Board Investigation M09Z0001, there was no self imposed limitation on type, size, or activity of vessel in Canadian waters. Of the 273 incidents considered as part of the study, the research question of the investigation focused on a subset of 200 accidents involving erroneous human factors. This included vessel accident occurrences primarily influenced by misunderstandings, inattention, lack of communication, misjudgement, and miscellaneous human factors.

Of the 200 accident subset wherein human error was involved, statistical analysis was conducted on a manufactured dataset from responses originating from a standardized questionnaire. As part of data collection, groups of respondents (pilots, ship masters, and officers) provided information on the attitudes, behaviours, and interactions among marine personnel involved in the respective accidents.

The study concluded that improved standards of communication, reduction of language barriers, increased monitoring of vessel movements, and greater cooperation amongst crew members were all necessary in order to improve marine safety standards.

---

[5] Government of Canada, T. S. B. of C. (1995, January 1). *A SAFETY STUDY OF THE OPERATIONAL RELATIONSHIP BETWEEN SHIP MASTERS/ WATCHKEEPING OFFICERS AND MARINE PILOTS*. Marine Investigation Report SM9501 - Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/SM9501/SM9501.html

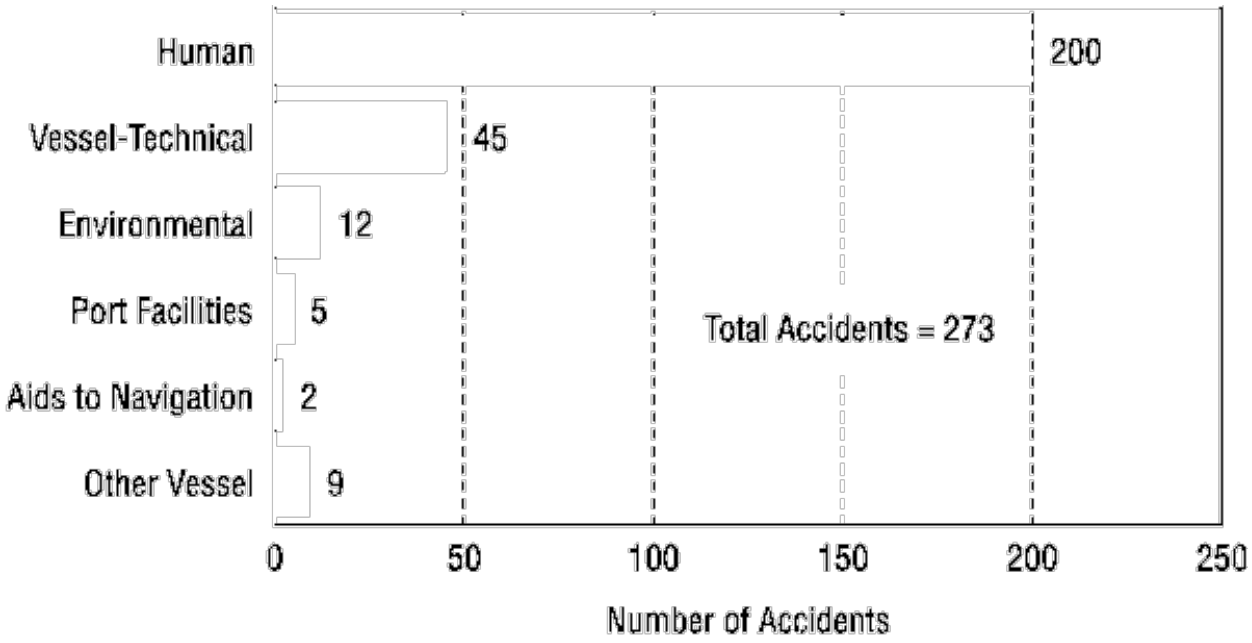


Figure 1, Section 1.1, showcasing the propensity of human error in marine accidents between 1981 and 1992.[6]

The study presents multiple findings on the importance of minimizing human error within marine safety procedures that had not been explored thoroughly prior to the report's release date. While it investigates the scope of marine safety on a similar basis when compared to Transportation Safety Board Investigation M09Z0001, SM9501 limits its overall ability to derive insights due to several research limitations:

i. The study presents an inherent bias in the types of vessels under investigation. The dataset primarily focuses on marine accident occurrences where 87% of the 273 occurrences were involving vessels greater than 5,000 gross registered tons. This impacts the scope of data that could have been captured in occurrences involving small registered vessels, which could possibly experience similar human error scenarios. This presents an opportunity for the current study proposal which aims to consider all vessel classes, weights, and sizes to provide an overall understanding of human and non-human factors affecting marine accident fatality rates.

---

[6] Transportation Safety Board of Canada. (1995). *Figure 1*. Transportation Safety Board of Canada. Government of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/SM9501/images/ems9501a.gif.

ii. The contributing factors under study aimed to assess the impact of human error on marine accidents and fatalities. However, of the attributes under consideration, no direct interaction calculation was attempted between human and non-human factors. It is possible that non-human factors, such as technical failures of the vessels, environmental conditions, port facilities conditions, and interference from other vessels could impact the contributing human factors under study. Of the variables under consideration in the proposed study, if attributes pass the stage whereby feature importance is calculated, interactions between the selected attributes will be calculated - regardless of their classification (human, technical, environmental , etc.).

iii. As briefly mentioned in the critical review of Transportation Safety Board Investigation M09Z0001, SM9501 also suffers from data freshness given the dataset was collected between 1981 and 1992. The survey data gathered may also be subject to survey question bias given that "*[…]each question began with the prefix "In my experience.…"[…]*", and did not offer any other form of questionnaire data input.

The dataset of the proposed study does not need to contend with an issue of bias given that it is factual data of marine accident occurrences simply logged monthly. However, the data under proposal may be susceptible to accident occurrence validation given its high-level of data freshness[7].

[7] Transportation Safety Board of Canada. (2019, June 12). *Marine occurrence data from January 1995 to present*. Open Government Portal. Retrieved October 23, 2022, from https://open.canada.ca/data/en/dataset/ad8d1b73-df09-4521-9bdb-61c529328218

International Literatures

3. An Analysis of Factors Affecting the Severity of Marine Accidents[8]

This study comprises of a statistical analysis of 1,207 marine accidents occurring across the globe from 2010 to 2019. The data collected for this study are from 7 marine safety boards, inclusive of the Transportation Safety Board of Canada, the Marine Accident Investigation Branch of the United Kingdom, the Australian Transport Safety Bureau, the United States Office of Marine Safety, the German Federal Bureau of Maritime Casualty Investigation, the China Maritime Safety Administration, and the Japan Transport Safety Board. Given various severities of marine accident categories, Wang et al. explore the influencing factors of the severity of marine accidents using an ordered logistic regression model.

The results of the study are able to provide a probability of the positive associative factors resulting in a higher marine accident severity. In addition, the same results are individually documented for types of vessels, as well as seven accident categories: collision, grounding, fire/explosion, contact, sinking, equipment failure, and others. Relativity of accident severity ratios are provided for different influencing factors as part of the study, inclusive of accident type, human element, ship type, ship condition, and environmental interference. The study aims to address some of the limitations of the aforementioned Canadian literature by including human, technical, and environmental factors as part of the interactions leading up to marine accident occurrences.

---

[8] Wang, H., Liu, Z., Wang, X., Graham, T., & Wang, J. (2021). An analysis of factors affecting the severity of marine accidents. *Reliability Engineering & System Safety*, *210*, 107513. doi:10.1016/j.ress.2021.107513

Data Description

The dataset under study is titled "*Marine occurrence data from January 1995 to present"* and is published in CSV format by the Government of Canada in cooperation with the Transportation Safety Board of Canada. The data is directly sourced from the Marine Safety Information Systems (MARSIS) operated by the TSB. Notably, the data set is refreshed on or soon after the 15th of every month. As such, there are some small changes that may occur between the time of writing this data description and the final version of the data set used in model production. The full data set ranges from January 1995 to the present day, up until the latest cycle of 15 days occurring at the time of writing. The full data set operates under an Open Government Licence by the Government of Canada and can be found at this URL:

https://www.tsb.gc.ca/includes/stats/csv/Marine/MARSISdb_MDOTW_VW_OCCURRENCE_PUBLIC.csv

For the respective marine occurrences under study, there are also records on the relevant vessels, navigation equipment, lifesaving appliances equipment, recording equipment, and injuries descriptions. These are all distributed as separate CSV files, but will not be under study as part of the core hypothesis analysis. There may be additional academic links made to the CSVs for explanatory or descriptive purposes, but not statistical or quantifiable analysis.

The raw dataset contains 83,797 entries across 160 attributes. For the analysis under study, not all 160 attributes will be considered. To view all 160 attributes, a data dictionary has been compiled by the TSB and is accessible here.

To understand the impact of attributes contributing to marine accident fatalities, a subset of the data must be taken, whereby only severe accidents resulting in fatalities have occurrence records. The dataset by default classifies occurrences by International Maritime Organization class levels, as well as TSB investigation classification levels of the occurrence. These two classifications differ in the way they numerically attribute fatalities to the

occurrences - subsequently, it is most prudent to simply look at occurrences which have non-zero fatality records.

```
import pandas as pd

marsis_raw =
pd.read_csv("MARSISdb_MDOTW_VW_OCCURRENCE_PUBLIC.csv")
marsis_deadly = marsis_raw[marsis_raw["TotalDeaths"] > 0]
```

Of interest, there are 3,235 records of fatalities. These records ought to be further subdivided by the severity of the accident leading to the fatality occurrence. Not all fatalities are the result of extreme circumstances, and observing the breakout of accident severity may give an indication of the distribution of accident severity respective to fatality occurrences.

For the purposes of enabling this study to be universally accessible, accident severity is best subcategorized under the International Maritime Organization's standards[9]. These include three levels of severity, identified as the following:

- **Level 1 - Very Serious Marine Casualty:** A marine incident which involves a loss of life, severe pollution, or a total loss of the vessel involved.
- **Level 2 - Marine Casualty:** An occurrence not classified as a very serious casualty, typically involving a fire, explosion, collision, grounding, contact, heavy weather damage, ice damage, hull cracking, or suspected hull defects, resulting in:
    - immobilization of main engines, extensive accommodation damage, severe structural damage, such as penetration of the hull under water, etc., rendering the ship unfit to proceed

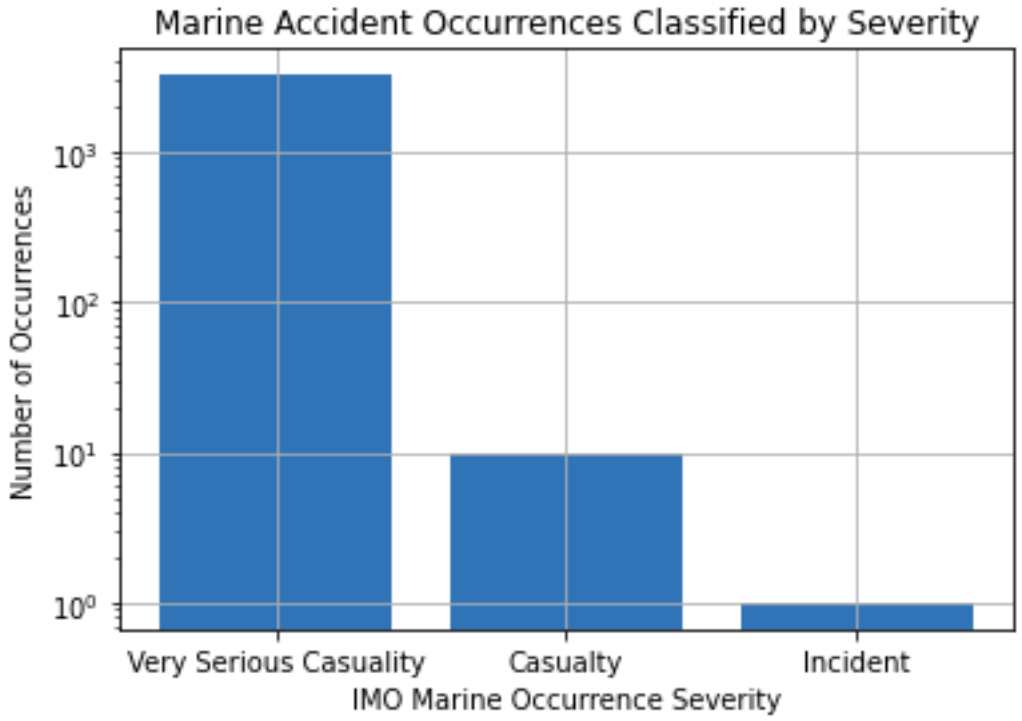
[9] International Maritime Organization. (2014, November 18). *CASUALTY-RELATED MATTERS* REPORTS ON MARINE CASUALTIES AND INCIDENTS Revised harmonized reporting procedures – Reports required under SOLAS regulations I/21 and XI-1/6, and MARPOL, articles 8 and 12*. International Maritime Organization. Retrieved October 23, 2022, from https://wwwcdn.imo.org/localresources/en/OurWork/MSAS/Documents/MSC-MEPC3/MSC-MEPC.3-Circ.4%20Rev%201%20%20Revised%20harmonized%20reporting%20procedures%20-%20Reports%20required%20under%20SOLAS%20regulations%20I21.pdf

- pollution (regardless of quantity)
- a breakdown necessitating towage or shore assistance

- **Level 3 - Marine Incident:** Includes occurrences not classified as very serious or serious casualties. Includes records in which hazardous occurrences or “near misses” took place, and if no corrective action were taken, would have endangered the safety of the vessel, its occupants, or any other person or the environment.

```
names = ["Very Serious Casuality", "Casualty", "Incident"]
values = [
    len(marsis_deadly[marsis_deadly["ImoClassLevelID"] == 1]),
    len(marsis_deadly[marsis_deadly["ImoClassLevelID"] == 2]),
    len(marsis_deadly[marsis_deadly["ImoClassLevelID"] == 3])
]

plt.bar(names, values)
plt.yscale("log", base=10)
plt.xlabel("IMO Marine Occurrence Severity")
plt.ylabel("Number of Occurrences")
plt.title("Marine Accident Occurrences Classified by Severity")
plt.grid(True)
plt.show()
```



In addition to analyzing the severity classification of each individual fatality, the cause of the accident or incident will be an important variable in feature engineering. Each accident / incident type is provided in the dataset by a presumed attribute named AccIncTypeDisplayEng, containing more granular classifications of each incident under study. To understand how many deaths are attributed to marine occurrence severity and their respective impact on total death counts, the following grouping is made to better illustrate the data:

Accident/Incident Type	Severity Class	Total Deaths
ABANDONED	1	131
BOTTOM CONTACT	1	2
CAPSIZES	1	754
	2	2
CARGO SHIFT/CARGO LOSS - Cargo shifted	1	2
COLLISION - Struck by vessel	1	4
COLLISION - With another vessel or other floating object	1	362
EXPLOSION	1	88
	3	1
FIRE	1	165
	2	1
GROUNDING - Not under power (includes drifting) (non-intentional)	1	123
GROUNDING - Under power (non-intentional)	1	79
MISSING (the ship is)	1	239
PERSON SERIOUSLY INJURED OR KILLED - Boarding, being on board, falling overboard from the ship	1	265
	2	6
PERSON SERIOUSLY INJURED OR KILLED - In contact with any part of the ship or its contents	1	1517
RISK OF SINKING	1	2
SANK - Flooding	1	163
SANK - Founders (taking on water above the waterline)	1	1265
	2	1
STRIKING - Allision with a fixed object (striking - includes berthed/docked vessels)	1	40
SUSTAINS DAMAGE RENDER UNSEAWORTHY/UNFIT FOR PURPOSE - Unfit for purpose - ice, weather, etc.	1	14

One assumption that needed also to be tested early on for our hypothesis was the correlation found between injuries sustained in an occurrence and the total deaths in an incident depending on the severity class of the incident. A correlation matrix was created to primarily analyze the directional relationship between total deaths and several attributes commonly occurring in highly severe classes of incidents.

The correlation matrix below analyses the following attributes commonly seen in highly severe classes of incidents:

- **TotalMinorInjuries:** The total number of persons who received minor injuries as a result of the occurrence.
- **TotalSeriousInjuries:** The total number of persons seriously injured as a result of the occurrence.
- **TotalMissingIndividuals**: The total number of persons who are missing as a result of the occurrence.
- **TotalPeopleInTheWater:** The total number of persons in the water as a result of the occurrence.

	TotalDeaths	TotalMinorInjuries	TotalSeriousInjuries	TotalMissingIndividuals	TotalPeopleInTheWater
TotalDeaths	1.00	0.05	0.01	0.05	0.05
TotalMinorInjuries	0.05	1.00	0.09	0.03	-0.00
TotalSeriousInjuries	0.01	0.09	1.00	-0.02	0.07
TotalMissingIndividuals	0.05	0.03	-0.02	1.00	-0.01
TotalPeopleInTheWater	0.05	-0.00	0.07	-0.01	1.00

Given the per occurrence correlation between total deaths and the aforementioned attributes, there appears to be no correlation (or an extremely weak positive correlation). As there is little directionality in the correlation matrix to indicate a relationship between the suspected attributes, additional exploratory analysis will need to be conducted during feature engineering. There is also additional intra-attribute analysis that needs to be conducted in order to determine deeper relationships between suspected (and not already considered) variables.

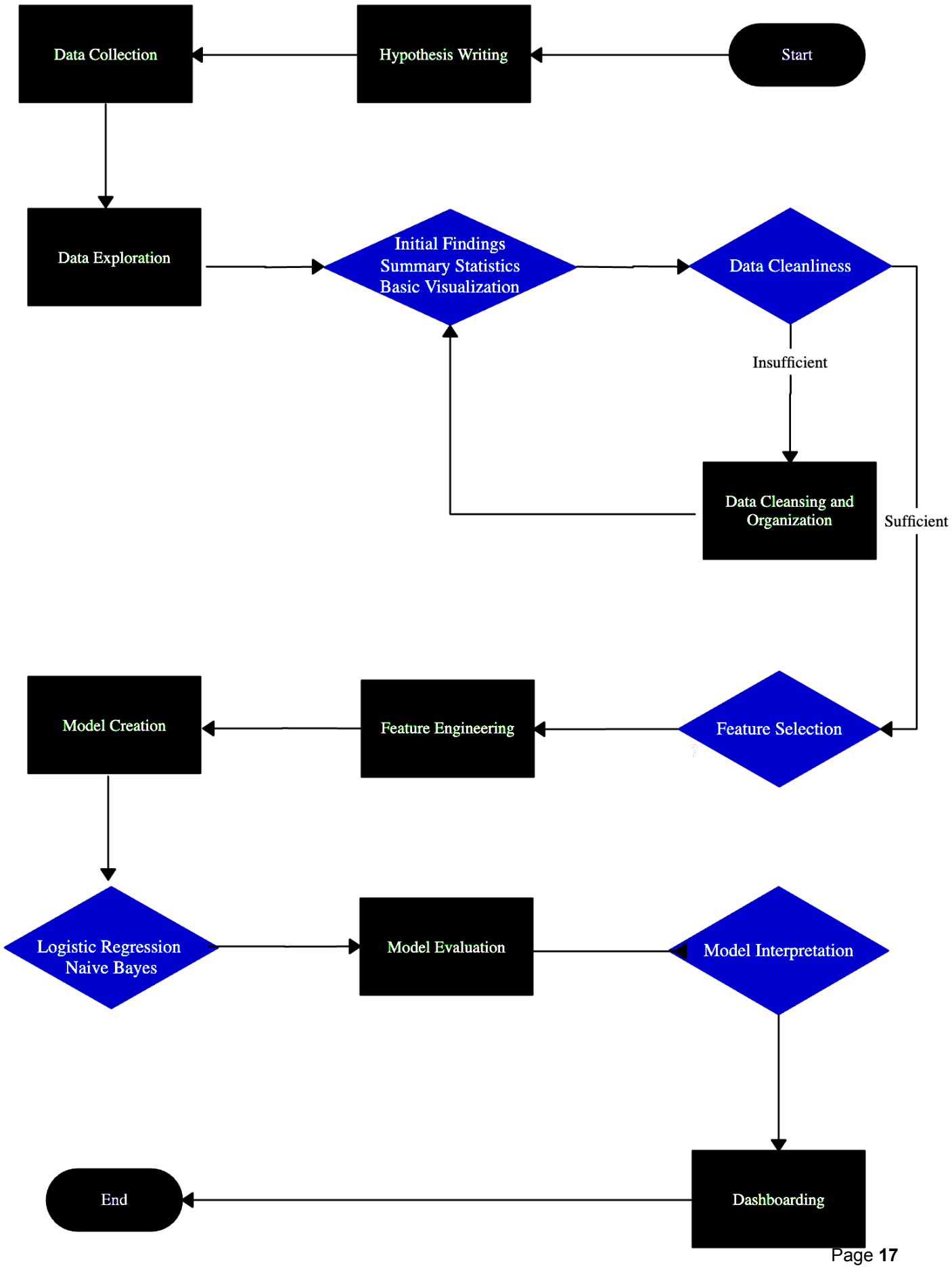
Lastly, some summary statistics on fatality occurrences provide us with an understanding of some of the dataset's aforementioned attributes and the range of values that are appropriate for further analysis and interpretation.

	TotalDeaths	TotalMinorInjuries	TotalSeriousInjuries	TotalMissingIndividuals	TotalPeopleInTheWater
mean	1.46	0.19	0.10	0.16	0.11
std	2.90	0.78	0.52	1.11	0.83
min	1.00	0.00	0.00	0.00	0.00
25%	1.00	0.00	0.00	0.00	0.00
50%	1.00	0.00	0.00	0.00	0.00
75%	1.00	0.00	0.00	0.00	0.00
max	84.00	12.00	8.00	28.00	21.00

Methodology

As part of the iterative process of this study, the illustration seen below will act as a primary flow of work to be completed given the established approach thus far.

1. Hypothesis writing and initial problem framing
2. Data collection and data imports
3. Data exploration
    - 3.1. Initial findings, summary statistics
    - 3.2 Data cleanliness and data logic
4. Feature selection
5. Feature engineering
6. Model creation
    - 6.1. Logistic regression
    - 6.2. Naive Bayes
7. Model evaluation
    - 7.1. Model interpretation
8. Dashboarding and final presentation


Start
Hypothesis Writing
Data Collection
Data Exploration
Initial Findings
Summary Statistics
Basic Visualization
Data Cleanliness
Insufficient
Data Cleansing and
Organization
Sufficient
Feature Selection
Feature Engineering
Model Creation
Model Interpretation
Model Evaluation
Logistic Regression
Naive Bayes
Dashboarding
End

References

Government of Canada, T. S. B. of C. (1995, January 1). A SAFETY STUDY OF THE OPERATIONAL RELATIONSHIP BETWEEN SHIP MASTERS/ WATCHKEEPING OFFICERS AND MARINE PILOTS. Marine Investigation Report SM9501 - Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/SM9501/SM9501.html

Government of Canada, T. S. B. of C. (2012, August 10). Marine investigation report M09Z0001. Marine Investigation Report M09Z0001 - Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/m09z0001/m09z0001.html

Government of Canada, T. S. B. of C. (2019, May 6). Marine Transportation Safety Investigations and reports. Transportation Safety Board of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/index.html

International Maritime Organization. (2014, November 18). CASUALTY-RELATED MATTERS* REPORTS ON MARINE CASUALTIES AND INCIDENTS Revised harmonized reporting procedures – Reports required under SOLAS regulations I/21 and XI-1/6, and MARPOL, articles 8 and 12. International Maritime Organization. Retrieved October 23, 2022, from https://wwwcdn.imo.org/localresources/en/OurWork/MSAS/Documents/MSC-MEPC3/MSC-MEPC.3-Circ.4%20Rev%201%20%20Revised%20harmonized%20reporting%20procedures%20-%20Reports%20required%20under%20SOLAS%20regulations%20I21.pdf

Transportation Safety Board of Canada. (1995). Figure 1. Transportation Safety Board of Canada. Government of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/SM9501/images/ems9501a.gif.

Transportation Safety Board of Canada. (2010). Figure 4. Safety Issues Investigation into Fishing Safety in Canada. Government of Canada. Retrieved October 23, 2022, from https://www.tsb.gc.ca/eng/rapports-reports/marine/etudes-studies/M09Z0001/images/m09z0001-figure-04.png.

Transportation Safety Board of Canada. (2019, June 12). Marine occurrence data from January 1995 to present. Open Government Portal. Retrieved October 23, 2022, from https://open.canada.ca/data/en/dataset/ad8d1b73-df09-4521-9bdb-61c529328218

Transportation Safety Board of Canada. (2019, June 12). Marine occurrence data from January 1995 to present. Open Government Portal. Retrieved October 23, 2022, from https://open.canada.ca/data/en/dataset/ad8d1b73-df09-4521-9bdb-61c529328218

Wang, H., Liu, Z., Wang, X., Graham, T., & Wang, J. (2021). An analysis of factors affecting the severity of marine accidents. Reliability Engineering & System Safety, 210, 107513. doi:10.1016/j.ress.2021.107513 [Original source: https://studycrumb.com/alphabetizer]

Github Repository

All relevant project files can be found at https://github.com/rfinatan/CIND-820-Big-Data-Analytics-Project.