

*ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ*



*ΣΧΟΛΗ ΝΑΥΠΗΓΩΝ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ*

*<<Στατιστική διερεύνηση και εφαρμογή προσέγγισης  
εξόρυξης δεδομένων για τα ναυτικά ατυχήματα πλοίων  
Γενικού Φορτίου>>*

*Διπλωματική Εργασία*

*Νάστας Εμμανουήλ*

*nm11551*

*Αθήνα 2018*

Επιβλέπων :

Νικόλαος Π. Βεντίκος, Αναπληρωτής Καθηγητής Ε.Μ.Π.

Τριμελής Επιτροπή :

Νικόλαος Π. Βεντίκος, Αναπληρωτής Καθηγητής Ε.Μ.Π.

Γεώργιος Ζαραφωνίτης, Καθηγητής Ε.Μ.Π.

Κώστας Ι. Σπύρου, Καθηγητής Ε.Μ.Π.

## ΕΥΧΑΡΙΣΤΙΕΣ :

Ένα μεγάλο ευχαριστώ στους καθηγητές αυτής της σχολής που με τις γνώσεις τους μας κέρδιζαν καθημερινά και μας έκαναν να θέλουμε να ασχοληθούμε με τη ναυπηγική κάθε μέρα και περισσότερο. Μια ιδιαίτερη αναφορά στον κ.Βεντίκο και τον Παναγιώτη Σωτήραλη που με τις επεξηγήσεις και τις συμβουλές με βοήθησαν να ολοκληρωθεί αυτή η διπλωματική. Τέλος, θέλω να ευχαριστήσω την οικογένειά μου, τους φίλους και τους συμφοιτητές για τη στήριξη αλλά και την υπομονή που έκαναν στα φοιτητικά μου χρόνια.

## ΠΕΡΙΛΗΨΗ :

Η παρούσα διπλωματική εργασία ασχολείται με τα ναυτικά ατυχήματα πλοίων τύπου Γενικού Φορτίου που διαδραματίστηκαν την περίοδο 1995-2016. Στόχος της εργασίας είναι η διερεύνηση της μεταβολής των ατυχημάτων, τόσο αριθμητικά όσο και σε σχέση με τη σοβαρότητα των συνεπειών τους, αλλά και η αποσαφήνιση πιθανών παραμέτρων που μπορεί να συνέβαλαν στο εκάστοτε γεγονός ναυτικού ατυχήματος.

Από τη μελέτη που ολοκληρώθηκε, βγήκε το συμπέρασμα πως η αύξηση των ατυχημάτων την τελευταία 25ετία ήταν ραγδαία και το γεγονός αυτό πρέπει να συγκεντρώσει την προσοχή όλων των εμπλεκόμενων. Η ανάλυση των ατυχημάτων χωρίστηκε σε δύο άξονες. Σύμφωνα με το μέγεθος του πλοίου, μετρούμενο σε κόρους και σύμφωνα με τον τύπο του ατυχήματος. Η συντριπτική πλειοψηφία των πλοίων ανήκουν στην κατηγορία μεγέθους  $1000 \leq GT \leq 20000$  ενώ ο μεγαλύτερος αριθμός ατυχημάτων ήταν με ζημιά στη γάστρα ή το μηχανοστάσιο. Επιπλέον εξετάστηκε η σημασία της ηλικίας του πλοίου τη στιγμή που συνέβη το ατύχημα. Επίσης τα δεδομένα υποβλήθηκαν σε κατάλληλα στατιστικά τεστ από τα οποία διαπιστώθηκε πως το μέγεθος του πλοίου παίζει ρόλο σε ένα ναυτικό ατύχημα για τα πλοία Γ/Φ.

Στη συνέχεια το δείγμα χωρίστηκε με βάση τον τύπο ατυχήματος και χρησιμοποιώντας δέντρα παλινδρόμησης (Regression Trees) προέκυψαν σημαντικά αποτελέσματα με την αλληλεπίδραση των παραγόντων που επηρεάζουν τον κάθε τύπο ατυχήματος και συμβάλουν στο να αναπτυχθεί το ατύχημα.

Μέσω των αποτελεσμάτων που προέκυψαν, θα ήταν συνετό η διπλωματική αυτή και τα συμπεράσματά της να αποτελέσουν μια βάση περαιτέρω ερευνών με σκοπό την ελαχιστοποίηση του αριθμού των ναυτικών ατυχημάτων πλοίων τύπου Γ/Φ και των ανεπιθύμητων συνεπειών αυτών.

## Περιεχόμενα

<<Στατιστική διερεύνηση και εφαρμογή προσέγγισης εξόρυξης δεδομένων για τα ναυτικά ατυχήματα πλοίων Γενικού Φορτίου>> .....	1
ΕΥΧΑΡΙΣΤΙΕΣ : .....	2
ΠΕΡΙΛΗΨΗ : .....	3
Περιεχόμενα.....	4
ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ: .....	6
ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ : .....	8
ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ : .....	8
ΚΕΦΑΛΑΙΟ 1 <sup>ο</sup> : ΕΙΣΑΓΩΓΗ.....	9
ΚΕΦΑΛΑΙΟ 2 <sup>ο</sup> : ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ.....	11
ΚΕΦΑΛΑΙΟ 3 <sup>ο</sup> : ΑΠΟΤΥΠΩΣΗ ΤΟΥ ΣΤΟΛΟΥ ΠΛΟΙΩΝ ΓΕΝΙΚΟΥ ΦΟΡΤΙΟΥ .....	13
3.1 Εισαγωγή .....	13
3.2 Η Βάση Δεδομένων .....	15
3.3 Αναφορές Ατυχημάτων .....	20
3.4 Στόλος πλοίων Γ/Φ ανά έτος.....	30
ΚΕΦΑΛΑΙΟ 4 <sup>ο</sup> : ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΤΩΝ ΑΤΥΧΗΜΑΤΩΝ .....	32
4.1 Σοβαρά ατυχήματα ανά πλοία-έτη για την περίοδο 01/01/1995-31/12/2016.....	32
4.1.1 Confidence Intervals (c.i).....	35
4.2 Συχνότητες Ατυχημάτων ανά κατηγορία Ατυχήματος.....	38
4.3 Ολικές Απώλειες ανά Πλοία×Έτη .....	42
4.4 Ανθρώπινες Απώλειες ανά Πλοία×Έτη.....	43
4.5 «Σοβαρά Ατυχήματα» ανά Έτος Κατασκευής των πλοίων .....	46
4.6 Ηλικία των πλοίων που είχαν ‘Σοβαρά’ ατυχήματα την περίοδο 1995-2016. ....	47
4.7 Στατιστικά Τεστ.....	49
ΚΕΦΑΛΑΙΟ 5 <sup>ο</sup> : ΑΝΑΠΤΥΞΗ ΔΕΝΤΡΩΝ ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗΣ & ΠΑΛΙΝΔΡΟΜΗΣΗΣ.....	62
5.1 R & R-Studio.....	62
5.2 Εισαγωγή Δεδομένων στο πρόγραμμα R.....	63
5.2.1 Το Συνολικό Δείγμα.....	64
5.2.2 Το Δείγμα για ατυχήματα σύγκρουσης (Collision).....	65
5.2.3 Το Δείγμα για ατυχήματα επαφής (Contact) .....	66
5.2.4 Το Δείγμα για ατυχήματα πυρκαγιάς/έκρηξης (Fire/Explosion) .....	66
5.2.5 Το Δείγμα για ατυχήματα από βλάβη στη γάστρα ή στο μηχανοστάσιο (Hull/Mchy Dam.) .....	67
5.2.6 Το Δείγμα για ατυχήματα πλοίων που καταστράφηκαν ή ακινητοποιήθηκαν (Wrecked/Stranded).....	68
5.2.7 Το Δείγμα για ατυχήματα βύθισης (Foundered) .....	69

ΚΕΦΑΛΑΙΟ 6 <sup>ο</sup> : ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ .....	70
ΒΙΒΛΙΟΓΡΑΦΙΑ .....	71
ΠΑΡΑΡΤΗΜΑ Α : ΣΥΝΟΛΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ.....	73
ΠΑΡΑΡΤΗΜΑ Β : ΣΥΝΟΛΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΔΕΝΤΡΩΝ .....	76
ΠΑΡΑΡΤΗΜΑ Γ : ΣΗΜΑΙΕΣ ΠΛΟΙΩΝ.....	89

## ΠΕΡΙΕΧΟΜΕΝΑ ΣΧΗΜΑΤΩΝ:

<a href="#">Σχήμα 1 Ποσοστά ατυχημάτων για όλους τους τύπους G.C. με GT&gt;499 και για την περίοδο 01/01/1995 έως 31/12/2016.</a>	20
<a href="#">Σχήμα 2 Γραφική αναπαράσταση της ποσοστιαίας κατανομής ατυχημάτων πλοίων τύπου Γ/Φ με βάση το GT για την περίοδο 01/01/1995-31/12/2016</a>	21
<a href="#">Σχήμα 3 Γραφική αναπαράσταση της ποσοστιαίας κατανομής του συνολικού αριθμού των ατυχημάτων στις διάφορες κατηγορίες ατυχημάτων.</a>	23
<a href="#">Σχήμα 4 Γραφική αναπαράσταση του αριθμού των ατυχημάτων για κάθε τύπο ατυχήματος για την περίοδο 01/01/1995-31/12/2016.</a>	23
<a href="#">Σχήμα 5 Ετήσια μεταβολή του αριθμού ατυχημάτων για κάθε τύπο ατυχήματος, για την περίοδο 01/01/1995-31/12/2016.</a>	24
<a href="#">Σχήμα 6 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με 500≤GT≤1000 για την περίοδο 01/01/1995-31/12/2016.</a>	24
<a href="#">Σχήμα 7 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με 1000&lt;GT≤20000 για την περίοδο 01/01/1995-31/12/2016.</a>	25
<a href="#">Σχήμα 8 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με GT&gt;20000 για την περίοδο 01/01/1995-31/12/2016.</a>	25
<a href="#">Σχήμα 9 Ετήσια μεταβολή του αριθμού των ατυχημάτων για κάθε κατηγορία μεγέθους.</a>	26
<a href="#">Σχήμα 10 Ο αριθμός των νεκρών που προκύπτουν κάθε χρονιά</a>	26
<a href="#">Σχήμα 11 Ατυχήματα με total loss του πλοίου μαζί με αναφορές για νεκρούς</a>	27
<a href="#">Σχήμα 12 Ατυχήματα με total loss του πλοίου μαζί με τον αριθμό των νεκρών</a>	27
<a href="#">Σχήμα 13 Ετήσιος μέσος όρος της ηλικίας των πλοίων που είχαν ατυχήματα την περίοδο 1995-2016.</a>	28
<a href="#">Σχήμα 14 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία 500≤GT&lt;1000</a>	29
<a href="#">Σχήμα 15 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία 1000≤GT≤20000</a>	29
<a href="#">Σχήμα 16 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία GT&gt;20000</a>	30
<a href="#">Σχήμα 17 Ετήσια συχνότητα ατυχημάτων για όλες τις κατηγορίες μεγέθους.</a>	32
<a href="#">Σχήμα 18 Ετήσια συχνότητα ατυχημάτων για πλοία με 500≤GT&lt;1000.</a>	33
<a href="#">Σχήμα 19 Ετήσια συχνότητα ατυχημάτων για πλοία με 1000≤GT≤20000.</a>	33
<a href="#">Σχήμα 20 Ετήσια συχνότητα ατυχημάτων για πλοία με GT&gt;20000.</a>	33
<a href="#">Σχήμα 21 Ετήσια μεταβολή συχνοτήτων ατυχημάτων για όλες τις κατηγορίες μεγέθους την περίοδο 1995-2016</a>	38
<a href="#">Σχήμα 22 Ετήσια Accident Rates για ατυχήματα Collision, για την περίοδο 1995-2016.</a>	39
<a href="#">Σχήμα 23 Ετήσια Accident Rates για ατυχήματα Hull/Mchy Dam, για την περίοδο 1995-2016.</a>	39
<a href="#">Σχήμα 24 Ετήσια Accident Rates για ατυχήματα Wrecked/Strand, για την περίοδο 1995-2016.</a>	40
<a href="#">Σχήμα 25 Ετήσια Accident Rates για ατυχήματα Contact, για την περίοδο 1995-2016.</a>	40
<a href="#">Σχήμα 26 Ετήσια Accident Rates για ατυχήματα Fire/Explosion, για την περίοδο 1995-2016.</a>	41
<a href="#">Σχήμα 27 Ετήσια Accident Rates για ατυχήματα Foundered, για την περίοδο 1995-2016.</a>	41
<a href="#">Σχήμα 28 Ετήσια Accident Rates για ατυχήματα που δεν προσδιορίζεται το είδος τους, για την περίοδο 1995-2016.</a>	42
<a href="#">Σχήμα 29 Κατανομή των ατυχημάτων που οδήγησαν σε ολική απώλεια του πλοίου, για την περίοδο 1995-2016.</a>	43

<a href="#">Σχήμα 30 Ετήσια συχνότητα ατυχημάτων που οδήγησαν σε Ολική Απώλεια πλοίου για την περίοδο 1995-2016. Συμπεριλαμβάνονται και τα c.i 95%.....</a>	43
<a href="#">Σχήμα 31 Ετήσια συχνότητα ατυχημάτων που οδήγησαν σε Ολική Απώλεια πλοίου, για τις τρεις κατηγορίες μεγέθους, για την περίοδο 1995-2016. ....</a>	44
<a href="#">Σχήμα 32 Κατανομή των ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες, για την περίοδο 1995-2016. ....</a>	45
<a href="#">Σχήμα 33 Ετήσιες συχνότητες ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες, για την περίοδο 1995-2016. Μαζί παριστάνεται και η γραμμή τάσης. Συμπεριλαμβάνονται και τα c.i 95%.....</a>	46
<a href="#">Σχήμα 34 Αριθμός ατυχημάτων ανά πλοία®έτη για κάθε έτος κατασκευής, εντός της περιόδου 1990-2016. ....</a>	47
<a href="#">Σχήμα 35 Συχνότητα ατυχημάτων ανά πλοία®έτη ανά ηλικιακή ομάδα τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.....</a>	48
<a href="#">Σχήμα 36 Συχνότητα ατυχημάτων ανά πλοία®έτη για την ηλικιακή ομάδα 0-4 τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.....</a>	49
<a href="#">Σχήμα 37 Συχνότητα ατυχημάτων ανά πλοία®έτη για την ηλικιακή ομάδα 5-9 τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.....</a>	49
<a href="#">Σχήμα 38 Regression Tree με αρχική μεταβλητή Casualty Type. ....</a>	65
<a href="#">Σχήμα 39 Regression Tree με αρχική μεταβλητή Flag Color.....</a>	66
<a href="#">Σχήμα 40 Regression Tree με αρχική μεταβλητή Age Group. ....</a>	66
<a href="#">Σχήμα 41 Regression Tree με αρχική μεταβλητή Fatalities.....</a>	67
<a href="#">Σχήμα 42 Regression Tree με αρχική μεταβλητή Flag Color.....</a>	68
<a href="#">Σχήμα 43 Regression Tree με αρχική μεταβλητή Flag Color.....</a>	69
<a href="#">Σχήμα 44 Regression Tree με αρχική μεταβλητή Flag Color.....</a>	69
<a href="#">Σχήμα 45 Regression Tree με αρχική μεταβλητή Age Group. ....</a>	70

## ΠΕΡΙΕΧΟΜΕΝΑ ΠΙΝΑΚΩΝ :

<a href="#">Πίνακας 1 Κατηγορίες δεδομένων για τις οποίες παρείχε πληροφορίες η Sea-Web.</a>	13
<a href="#">Πίνακας 2 Αλφαβητική λίστα Νηογνομώνων που αναφέρεται στη βάση δεδομένων της Sea-Web.</a>	18
<a href="#">Πίνακας 3 Αναφορές ατυχημάτων για κάθε κατηγορία πλοίων G.C. με GT&gt;499 και για την περίοδο 01/01/1995 έως 31/12/2016.</a>	19
<a href="#">Πίνακας 4 Αριθμός ατυχημάτων για την κάθε κατηγορία μεγέθους για την περίοδο 01/01/1995 έως 31/12/2016.</a>	20
<a href="#">Πίνακας 5 Αριθμός ατυχημάτων ανά κατηγορία ατυχήματος για πλοία Γ/Φ GT&gt;499 και για την περίοδο 01/01/1995-31/12/2016.</a>	22
<a href="#">Πίνακας 6 Αριθμός ατυχημάτων που προκάλεσαν μόλυνση του περιβάλλοντος για πλοία Γ/Φ GT&gt;499 και για την περίοδο 01/01/1995-31/12/2016.</a>	27
<a href="#">Πίνακας 7 Ετήσιος υπό Μελέτη Στόλος για την περίοδο 1995-2016</a>	31
<a href="#">Πίνακας 8 Αριθμός ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες και μέσος όρος απωλειών/ατυχημάτων, για την περίοδο 1995-2016.</a>	46
<a href="#">Πίνακας 9 Κατανομή ατυχημάτων σε ηλικιακές ομάδες πενταετίας για κάθε κατηγορία μεγέθους, για την περίοδο 1995-2016</a>	48
<a href="#">Πίνακας 10 Συγκεντρωτικά αποτελέσματα των accident rates για κάθε κατηγορία μεγέθους, για την περίοδο 1995-2016.</a>	53
<a href="#">Πίνακας 11 Αποτελέσματα ANOVA όπως προέκυψαν από το πρόγραμμα MS EXCEL.</a>	54
<a href="#">Πίνακας 12 Αποτελέσματα t-Test όπως προέκυψαν από το πρόγραμμα MS EXCEL.</a>	59
<a href="#">Πίνακας 13 Αποτελέσματα F-Test όπως προέκυψαν από το πρόγραμμα MS EXCEL.</a>	60
<a href="#">Πίνακας 14 Δεδομένα μεταβλητής Age Group με Τύπο Ατυχήματος</a>	61
<a href="#">Πίνακας 15 Δεδομένα μεταβλητής Flag Color με Τύπο Ατυχήματος</a>	62

## ΠΕΡΙΕΧΟΜΕΝΑ ΕΙΚΟΝΩΝ :

<a href="#">Εικόνα 1 Απεικόνιση του επιπέδου εμπιστοσύνης C και της τιμής z* στην περίπτωση της τυπικής κανονικής κατανομής.</a>	36
<a href="#">Εικόνα 2 Κατανομή των ομάδων ελέγχου και θεραπείας μιας μελέτης.</a>	56
<a href="#">Εικόνα 3 Τρία διαφορετικά σενάρια κατανομών με ίση διασπορά μέσων τιμών.</a>	57
<a href="#">Εικόνα 4 Συνάρτηση στο t-Test.</a>	57
<a href="#">Εικόνα 5 Standard error of difference.</a>	58
<a href="#">Εικόνα 6 Τελική συνάρτηση του t-Test.</a>	58
<a href="#">Εικόνα 7 Λογότυπο προγράμματος R.</a>	63



## ΚΕΦΑΛΑΙΟ 1<sup>ο</sup> : ΕΙΣΑΓΩΓΗ

Ένας σημαντικός παράγοντας που επηρεάζει την κοινωνία είναι τα ναυτικά ατυχήματα, ειδικά στην Ελλάδα που θεωρείται μια ναυτική χώρα. Τα ναυτικά ατυχήματα θα βρίσκονται πάντα στις ζωές μας και δε μπορεί αυτό να αλλάξει μιας και πάντα θα υπάρχουν ανεξαρτήτως τι μπορεί να αλλάξει. Το σημαντικότερο είναι να ελαχιστοποιηθούν όσο είναι ανθρωπίνως δυνατό μέσω της συμβολής όλων των εμπλεκόμενων στη ναυτιλία. Για να γίνει αυτό εφικτό πρέπει να μελετώνται όλα τα ατυχήματα μαζί αλλά και το καθένα ξεχωριστά ώστε να ληφθούν μέτρα πρόληψης και αποφυγής. Επίσης όλα τα συμπεράσματα των μελετών θα πρέπει να μοιράζονται και να κοινοποιούνται ώστε όλοι οι εμπλεκόμενοι να μπορούν να βοηθήσουν και να βοηθηθούν.

Η παρούσα διπλωματική εργασία διαπραγματεύεται ναυτικά ατυχήματα πλοίων τύπου Γενικού Φορτίου (Γ/Φ) για την περίοδο 1995-2016 και ως στόχο έχει αρχικά την εξακρίβωση της σοβαρότητας των συγκεκριμένων ατυχημάτων και στη συνέχεια την ανάδειξη των παραμέτρων εκείνων που πιθανότατα παίζουν κυρίαρχο ρόλο σε ένα ναυτικό ατύχημα με σκοπό τη μελλοντική έρευνα αυτών για αποφυγή ή σημαντική μείωση των ατυχημάτων αυτών.

Το κεφάλαιο 3 της διπλωματικής αυτής εργασίας περιέχει μια αναφορά στα ιστορικά δεδομένα για ναυτικά ατυχήματα πλοίων τύπου Γ/Φ με βάση τις αναφορές ατυχημάτων που υπάρχουν στη βάση δεδομένων της Sea-Web. Το σημαντικό που προέκυψε από την εξόρυξη των δεδομένων είναι η μεγάλη αύξηση ατυχημάτων που παρατηρείται την περασμένη 25ετία για πλοία τύπου Γ/Φ γεγονός που φαίνεται καλύτερα αν τα συγκρίνουμε με πλοία διαφορετικού τύπου όπως τα Tankers, των οποίων η ετήσια μεταβολή του αριθμού ατυχημάτων ακολουθεί πολύ καλύτερη πορεία, μιας και βαίνει συνεχώς μειούμενη. Προκύπτει δηλαδή ότι οι αυστηρότεροι κανονισμοί αποδίδουν έμπρακτα αποτελέσματα και ίσως θα έπρεπε να εφαρμοστούν παρόμοιοι κανονισμοί και στα πλοία Γ/Φ (General Cargo Ships). Στη συνέχεια και αφού υπολογίστηκαν τα πλοία<sup>®</sup> έτη της μελέτης, αριθμός απαραίτητος για τον υπολογισμό των συχνοτήτων, ακολούθησε μια στατιστική ανάλυση των αποτελεσμάτων. Το σύνολο των δεδομένων χωρίστηκε σε δύο άξονες :

1. Με βάση το μέγεθος σε 3 κατηγορίες
  - $500 \leq GT \leq 1000$
  - $1000 < GT \leq 20000$
  - $GT > 20000$
2. Με βάση τον τύπο του ατυχήματος σε 6 κατηγορίες
  - Σύγκρουση (Collision)
  - Επαφή (Contact)
  - Βύθιση / Ημιβύθιση (Foundered)
  - Φωτιά / Έκρηξη (Fire / Explosion)
  - Προσάραξη (Wrecked / Stranded)
  - Αστοχία Γάστρας / Μηχανική Βλάβη (Hull / Machinery Damage)

Το κεφάλαιο 4 περιλαμβάνει τη στατιστική ανάλυση των δεδομένων με τον υπολογισμό συχνοτήτων ατυχημάτων και ειδικότερα συχνοτήτων ατυχημάτων που οδήγησαν σε ολική απώλεια του πλοίου (total loss) και σε ανθρώπινες απώλειες (fatalities). Παράλληλα εφαρμόστηκαν τα κατάλληλα στατιστικά τεστ και το συμπέρασμα ήταν πως υπάρχει στατιστική διαφορά τόσο ως προς τις κατηγορίες μεγέθους, όσο και ως προς το έτος του ατυχήματος. Πρέπει όμως να σημειωθεί πως η μεσαία κατηγορία μεγέθους στην οποία ανήκουν πλοία με  $1000 < GT \leq 20000$ , συγκεντρώνει τη μεγάλη πλειοψηφία των πλοίων.

Στο πέμπτο κεφάλαιο γίνεται χρήση του προγράμματος R και των βοηθητικών του προγραμμάτων γραμμένων σε γλώσσα προγραμματισμού R με σκοπό την παραγωγή δέντρων παλινδρόμησης (Regression Trees) για να αναδειχτεί η αλληλεπίδραση των παραγόντων που οδηγούν στους κυριότερους τύπους ατυχημάτων για General Cargo Ships. Με αυτό τον τρόπο ίσως γίνει πιο ξεκάθαρη η αιτία που οδήγησε σε καθένα από τους τύπους ατυχήματος και στη συνέχεια δοθεί η κατεύθυνση προς τη δραστική μείωση των ατυχημάτων αυτών.

## ΚΕΦΑΛΑΙΟ 2<sup>ο</sup> : ΒΙΒΛΙΟΓΡΑΦΙΚΗ ΕΠΙΣΚΟΠΗΣΗ

Το εύρος των παραμέτρων που απαιτείται για να χαρακτηρίσουν τα ναυτικά ατυχήματα είναι τεράστιο. Έτσι, η μελέτη και ανάλυση των ναυτικών ατυχημάτων είναι μια δύσκολη υπόθεση αφού χρειάζεται να συνυπολογιστούν παράγοντες όπως η λειτουργική κατάσταση στην οποία βρισκόταν το εκάστοτε πλοίο που είχε ατύχημα, το φορτίο που μετέφερε, το μέγεθός του, ακόμα και ο καιρός που συνάντησε στο ταξίδι του. Ωστόσο η συλλογή αυτού του όγκου δεδομένων γίνεται κυρίως από ιδιωτικές βάσεις, ασφαλιστικές εταιρίες, με αποτέλεσμα η επεξεργασία και η μελέτη να χρειαστεί να ολοκληρωθεί από Πανεπιστημιακά Ιδρύματα σε προπτυχιακό ή σε μεταπτυχιακό επίπεδο.

Η βιβλιογραφία που χρησιμοποιήθηκε στην παρούσα έρευνα βασίζεται κυρίως σε παρόμοιες διπλωματικές εργασίες του τομέα Μελέτης Πλοίου και Θαλάσσιων Μεταφορών της σχολής Ναυπηγών Μηχανολόγων Μηχανικών του Εθνικού Μετσόβιου Πολυτεχνείου. Με τον τρόπο αυτό μπορούν να παραχθούν συγκρίσεις αποτελεσμάτων και παράλληλα να βγουν αξιόπιστα συμπεράσματα για την εξέλιξη και τη βελτίωση της κατάστασης όσων αφορά τη μείωση των ατυχημάτων. Το 2011, στη διπλωματική εργασία με τίτλο «Μελέτη και Ανάλυση Ναυτικών Ατυχημάτων Πλοίων τύπου Γενικού Φορτίου», ο Χρήστος Χρυσουγής μελέτησε για την περίοδο 1995-2016 τα ναυτικά ατυχήματα πλοίων Γ/Φ με  $GT \geq 500$ , από όπου βγήκαν αποτελέσματα για τον καταγεγραμμένο υπό μελέτη στόλο της βάσης δεδομένων Sea-Web εκείνης της περιόδου αλλά και οι τιμές συχνοτήτων ατυχημάτων, συχνοτήτων ολικής απώλειας πλοίου και συχνότητες ανθρώπινων απωλειών οι οποίες κρίθηκε χρήσιμο να συγκριθούν με τα αποτελέσματα έως σήμερα ως προς την ακρίβεια και την τάξη μεγέθους.

Το έτος 2012 η κ. Κ. Μπίθα στη διπλωματική της με τίτλο «Στατιστική Ανάλυση Ναυτικών Ατυχημάτων κατά την περίοδο 1980-2009» μελέτησε τα ναυτικά ατυχήματα για όλους τους τύπους πλοίων, συμπεριλαμβανομένου και των πλοίων Γ/Φ. Η ανάλυση περιελάμβανε τα αίτια και τη σοβαρότητα των ατυχημάτων, τις τοποθεσίες στις οποίες αυτά συνέβησαν, τη σημαία που έφεραν τα πλοία τη στιγμή του ατυχήματος καθώς και τα πλοία που χάθηκαν από κάθε κατηγορία, και τις απώλειες σε ανθρώπινες ζωές που προκλήθηκαν από τα ατυχήματα αυτά.

Χρήσιμη πηγή πληροφοριών αποτέλεσε και η διεθνής ένωση Νηογνομόνων (IACS) με τη δημοσίευση το 2008 (Formal Safety Assessment of General Cargo Ships) όπου μελέτησε για την περίοδο 1997-2007 τα ναυτικά ατυχήματα πλοίων Γ/Φ με  $GT \geq 500$ , από όπου βγήκαν αποτελέσματα για τον καταγεγραμμένο υπό μελέτη στόλο εκείνης της περιόδου αλλά και οι τιμές συχνοτήτων ατυχημάτων, συχνοτήτων ολικής απώλειας πλοίου και συχνότητες ανθρώπινων απωλειών, με τα αποτελέσματα να αποτελούν εργαλείο σύγκρισης όσων αφορά την παρούσα διπλωματική εργασία.

Επίσης πολύ σημαντική βοήθεια και γνώση παρείχαν τα papers τα οποία συμπεριλήφθηκαν για την ανάλυση : M. Bogalecka & M. Popek «Analysis of Sea Accidents in 2006», όπου καταγράφηκαν αναλυτικά τα ατυχήματα εκείνης της χρονιάς με πληροφορίες για τις τοποθεσίες, τις συνέπειες στα πλοία και τις ανθρώπινες απώλειες. Eleftheria Elioroulou & Apostolos Papanikolaou «Statistical analysis of ship accidents and review of safety level» 2006, όπου περιγράφεται αναλυτικά το πλαίσιο των νέων κανονισμών καθώς και οι επιπτώσεις αυτών στη σχεδίαση εμπορικών πλοίων.

Τέλος, έχουν συμπεριληφθεί και έχουν χρησιμοποιηθεί οδηγίες των papers που αφορούν την κατασκευή δέντρων παλινδρόμησης (Regression Trees) και αναφέρονται ακολούθως :

Ο κ. Τζαννάτος Ερνέστος και ο κ. Κοκοτός Δημήτρης το 2009 δημοσίευσαν ένα Paper με τίτλο «Analysis of Accidents in Greek shipping during the pre and post-ISM period», όπου συμπεριέλαβαν όλα τα πλοία που κατείχαν Ελληνική σημαία την περίοδο 1993-2006 και μέσω των δέντρων αποφάσεων όπου περνάει τα δεδομένα, κατηγοριοποιεί τους παράγοντες που οδηγούν στο εκάστοτε ατύχημα.

Επίσης χρήσιμη πηγή οδηγιών αποτέλεσαν οι δημοσιεύσεις «Evtree-Evolutionary Learning of Globally Optimal Classification and Regression Trees in R» 2014. «Classification and Regression Trees» 2017. « Party –A Laboratory for Recursive Part(y)itioning» 2006. «Classification and Regression Trees-A Practical Guide for Describing a Dataset» 2013.

## ΚΕΦΑΛΑΙΟ 3<sup>ο</sup> : ΑΠΟΤΥΠΩΣΗ ΤΟΥ ΣΤΟΛΟΥ ΠΛΟΙΩΝ ΓΕΝΙΚΟΥ ΦΟΡΤΙΟΥ

### 3.1 Εισαγωγή

Ύστερα από πρωτοβουλία της Ρωσίας στα πλαίσια διάσκεψης του IMO εν έτη 2006 (Maritime Safety Committee 82/21/19,2006) συζητήθηκε έντονα το θέμα της ασφάλειας πλοίων τύπου Γενικού Φορτίου. Αφορμή στάθηκε το γεγονός ότι παρά το χαμηλό ποσοστό παγκόσμιου στόλου που αποτελούν (17%), καταλαμβάνουν το 42% των συνολικών ατυχημάτων που καταλήγουν σε total loss του πλοίου και το 27% των fatalities για το διάστημα 1999-2004. Αυτά τα ποσοστά αντιστοιχούν σε 73 κατεστραμμένα πλοία για κάθε έτος της παραπάνω περιόδου. Επίσης με βάση τις στατιστικές αναλύσεις του ParisMOU (Paris Memorandum of Understanding on Port State Control), προκύπτουν για τα πλοία Γ/Φ τα εξής ενδιαφέροντα στοιχεία για το έτος 2004:

- Εμφανίζουν το δεύτερο μεγαλύτερο αριθμό επιθεωρήσεων που επιβάλλεται κράτηση (detentions), περίπου 8% ενώ γενικά για όλα τα πλοία είναι 6%.
- Χαρακτηρίζονται με το δεύτερο μεγαλύτερο αριθμό επιθεωρήσεων που υπάρχουν ελλείψεις (deficiencies), περίπου 60% ενώ γενικά για όλα τα πλοία ο μέσος όρος είναι στο 54%.

Στον οργανισμό ParisMOU συμμετέχουν 26 Ευρωπαϊκές χώρες και ο Καναδάς ενώ προστατεύει τα Ευρωπαϊκά παράλια και τον Β. Ατλαντικό Κόλπο από τη Β. Αμερική έως την Ευρώπη. Στόχος του είναι η εξάλειψη υπό-πρότυπων πλοίων μέσω του περιβάλλοντος εναρμονισμένων port State control. Έτσι διενεργούνται πάνω από 24000 έλεγχοι στα λιμάνια υποδοχής για την εξασφάλιση υψηλών προτύπων ασφαλείας και μέριμνας, τόσο για τα πληρώματα των πλοίων όσο και για το περιβάλλον.

Αμέσως μετά τη διάσκεψη του 2006 έγιναν πολλές έρευνες που παραδόθηκαν στον IMO με στόχο τη διαλεύκανση των αιτιών που οδηγούν σε total loss και fatalities για πλοία Γ/Φ.

Η διπλωματική αυτή εργασία θα αποτυπώσει το στόλο πλοίων Γ/Φ που είχαν ατυχήματα από το 1995 έως το 2016 και θα αναλύσει τα δεδομένα με στόχο να εντοπιστούν συσχετίσεις και τάσεις χαρακτηριστικών μεγεθών της ταυτότητας ενός πλοίου που πιθανόν να μας δώσουν μια πιο ολοκληρωμένη εικόνα για τα υψηλά ποσοστά που αναφέρθηκαν παραπάνω.

Για την ανάλυση που πραγματοποιήθηκε, αντλήθηκαν στοιχεία από τις βάσεις δεδομένων της Clarksons και της Sea-Web. Θα πρέπει να επισημάνουμε μερικούς παράγοντες σχετικά με τις βάσεις δεδομένων όταν χρησιμοποιούνται για τον καθορισμό στατιστικών δεδομένων. Ορισμένα στοιχεία σε μερικά πλοία δεν είναι δηλωμένα. Επίσης η αλλαγή ορισμένων στοιχείων δεν έχει καταχωρηθεί (π.χ. ο τύπος του πλοίου ή η σημαία). Τέτοιες αναφορές βρίσκουμε σε όλες τις βάσεις δεδομένων, πάντως αυτές που χρησιμοποιήθηκαν χαρακτηρίζονται από μεγάλη πληρότητα.

Πίνακας 1 Κατηγορίες δεδομένων για τις οποίες παρείχε πληροφορίες η Sea-Web.

LR/IMO_Ship_No. & Incident_No.	Pollution	Demolition_Commenced
LR/IMO_Ship_No	GT	Demolition_Complete
Ship_Name	Flag	Demolition_Date
Cargo	Incident_No.	DWT
Casualty_Date	Killed	Severity
Casualty_Type	Loss_or_Disposal	Ship_type
Date_of_Build	Missing	Zone

Η ανάλυση ιστορικών στοιχείων έχει ως στόχο να αποδώσει ένα στιγμιότυπο της περιόδου που μελετάται. Η ακρίβεια που αποτυπώνεται μεταξύ πραγματικότητας και στατιστικών επηρεάζεται σε περίπτωση ελλείψεων ή λανθασμένων καταγραφών στις βάσεις δεδομένων. Οι ελλείψεις αυτές οδηγούν γενικά σε πιο αισιόδοξες εκτιμήσεις. Για να εξεταστούν εξ ολοκλήρου όμως οι αναφορές ατυχημάτων απαιτούνται πιστή καταγραφή των στοιχείων του κάθε ατυχήματος και συλλογή όσο το δυνατό περισσότερων στοιχείων για αυτό. Με γνώμονα όμως ότι η λειτουργία των πλοίων είναι σε ολόκληρο τον κόσμο, απαιτείται ένα εξειδικευμένο και συνεχώς εξελίξιμο σύστημα για την απόκτηση των δεδομένων των ατυχημάτων συμπεριλαμβανομένου των αιτιών και συνεπειών τους. Μέσω της ανάλυσης που ακολουθήσαμε προέκυψε πως ακόμα δεν υπάρχει συνεπής απόκτηση και ενημέρωση στοιχείων που αφορούν τη σημαία και την κλάση του πλοίου. Ο λόγος που συμβαίνει αυτό δεν είναι ξεκάθαρος, όμως έχει σχέση σίγουρα με το underreporting, δηλαδή με το γεγονός ότι δεν υπάρχει ακριβής δήλωση στις αναφορές.

Ένας επιπλέον σημαντικός παράγοντας για τη συχνότητα ατυχημάτων είναι ο αριθμός πλοία-έτη δηλαδή ο αριθμός των πλοίων επί τα έτη. Για αυτό και απαιτείται ο ακριβής καθορισμός του. Στη μελέτη αυτή ο αριθμός πλοία-έτη (shipyears) προκύπτει από τις ημερομηνίες due or delivered και λαμβάνοντας υπόψιν την ημερομηνία λειτουργίας (και όχι μόνο το έτος) ενώ θα ήταν πιο ακριβές σε περίπτωση που δινόταν το πλήθος των scrapped or lost και οι ημερομηνίες τους, κάτι που τελικά δε κατέστη δυνατό να βρεθεί. Τα δεδομένα που αντλήσαμε μας επιτρέπουν τον υπολογισμό της συχνότητας ατυχημάτων μόνο για έναν τύπο πλοίου. Σε περίπτωση που θέλαμε να συγκρίνουμε πλοία μιας κλάσης (ενός νηογνώμονα) δεν είναι εφικτό διότι τα πλοία καταγράφονται βάσει μίας κλάσης και αυτή δεν ενημερώνεται το ιστορικό σε περίπτωση αλλαγής.

Στην διπλωματική αυτή, η ανάλυση που πραγματοποιήθηκε αφορά όλα τα πλοία Γ.Φ. ανεξαρτήτως κλάσης. Οι παράμετροι που χρησιμοποιήθηκαν για την αναζήτηση στις βάσεις δεδομένων ήταν :

- Πλοία που παραδόθηκαν μετά από την 01/01/1990 και πριν την 01/01/2016 (πράγμα που σημαίνει ότι η μέγιστη ηλικία πλοίου που παρατηρείται στην έρευνά μας είναι τα 26 έτη)
- Πλοία με GT>499 κόρους
- Οι αναφορές αφορούν σοβαρά ατυχήματα

Δε χρησιμοποιήθηκαν η κλάση και η σημαία του πλοίου ως κριτήρια αναζήτησης. Άρα τα στατιστικά στοιχεία και οι συχνότητες ατυχημάτων δημιουργούνται για τις εξής κατηγορίες:

1. Σε σχέση με το μέγεθος του πλοίου

- $500 \leq GT \leq 1000$

- $1000 \leq GT \leq 20000$
- $20000 < GT$

## 2. Σε σχέση με τις κατηγορίες ατυχημάτων

- Όλα τα σοβαρά ατυχήματα, δηλαδή τα ατυχήματα που οδηγούν σε δομική βλάβη που καθιστά αναξιόπιστο το πλοίο για πλεύση (π.χ. ρήγμα πυθμένα), σε σταμάτημα της κύριας μηχανής, σε εκτενή βλάβη και γενικά σε οποιαδήποτε βλάβη ή οικονομική ζημιά που θεωρείται σημαντική.
- Ατυχήματα που οδηγούν σε ολική απώλεια του πλοίου
- Ατυχήματα που έχουν θύματα ή αγνοούμενους ως αποτέλεσμα

## 3.2 Η Βάση Δεδομένων

Η βάση δεδομένων από την οποία αντλήθηκαν τα ιστορικά στοιχεία είναι η Sea-Web, στην οποία χρησιμοποιείται ο κώδικας 'StatCode5v' για κατηγοριοποίηση των διάφορων τύπων πλοίου.

Με βάση αυτόν τον κώδικα κατάταξης τα πλοία General Cargo χωρίζονται σε τρεις κατηγορίες :

- General Cargo(A31A)
- Palletized Cargo Ship(A31B)
- Deck Cargo Ship(A31C)

Για αυτές τις υποκατηγορίες, προκύπτουν επόμενες υποκατηγορίες με βάση το 'StatCode5v'. Εν τέλει οι σημαντικότερες υποκατηγορίες 'General Cargo Ship', στις οποίες θα γίνει εστίαση για την αρχική αναζήτηση ιστορικών στοιχείων είναι οι ακόλουθες :

- General Cargo Ship (with Ro/Ro facilities), A31A2GA
- Open Hatch Cargo Ship, A31A2GO
- General Cargo/Tanker (Container/Oil/Bulk-COB Ships), A31A2GS
- General cargo/tanker, A31A2GT
- General Cargo Ship, A31A2GX
- Palletized Cargo Ship, A31B2GP
- Deck Cargo Ship, A31C2GD

Ενδεικτικά, αναφέρουμε μερικά χαρακτηριστικά του 'StatCode5v'. Οι κωδικοποιημένες ονομασίες που αναφέρονται μετά τον κάθε τύπο πλοίου προκύπτουν από το διαχωρισμό των πλοίων σε πέντε επίπεδα.

Για αρχή, στο πρώτο επίπεδο ορίζεται ο γενικός τύπος του πλοίου-κατασκευαστή :

- A:Cargo Carrying Ship
- B:Work Vessel
- W:Non Seagoing Merchant Ships

- X:Non Merchant
- Y:Non propelled
- Z:Non Ship Structures

Στο δεύτερο επίπεδο ορίζεται η γενική κατηγορία που ανήκει το πλοίο όπως :

- A1:Tankers
- A2:Bulk Carriers
- A3:Dry Cargo/Passenger
- B1:Fishing
- B2:Offshore
- B3:Miscellaneous

Το τρίτο επίπεδο διαχωρίζει τα πλοία με βάση τον τύπο του φορτίου που κατασκευάστηκαν να μεταφέρουν ορισμένο κατά μια γενική έννοια (κατηγορία A) ή με βάση το γενικευμένο όρο που χρησιμοποιείται για το πεδίο λειτουργίας τους (κατηγορία B). Παραδείγματος χάρη :

- A11:Liquified Gas
- A12:Chemical
- A13:Oil
- A14:Other Liquids
- A21:Bulk Dry
- A22:Bulk Dry/Oil
- A23:Self Discharging Bulk Dry
- A24:Other Bulk Dry
- A31:General Cargo
- A32:Passenger/General Cargo
- A33:Container
- A34:Refrigerated Cargo
- A35:Ro-Ro Cargo
- A36:Passenger/ Ro-Ro Cargo
- A37:Passenger
- A38:Other Dry Cargo
- B11:Fish Catching
- B12:Other Fishing
- B21:Offshore Supply
- B22:Other Offshore
- B31:Research
- B32:Towing/Pushing
- B33:Dredging
- B34:Other Activities



Το τέταρτο επίπεδο ορίζεται ακριβώς το μεταφερόμενο φορτίο. Όμως επειδή οι κατηγορίες έχουν αυξηθεί σημαντικά θα δοθούν μόνο μερικές ενδεικτικά.

- A11A:LNG Tanker
- A11B:LPG Tanker
- A13A:Crude Oil Tanker
- A21A:Bulk Carrier
- A21B:Ore Carrier
- A31A:General Cargo
- A31B:Palletised Cargo Ship
- A31C:Deck Cargo Ship

Τέλος, στο επίπεδο πέντε ορίζεται αν το πλοίο έχει διπλά τοιχώματα (double hull, κωδικός 2) και συμπληρώνεται από δύο γράμματα που αποτελούν την κωδική ονομασία του συγκεκριμένου τύπου πλοίου που μεταφέρει το συγκεκριμένο φορτίο και επιτελεί τη συγκεκριμένη λειτουργία. Για παράδειγμα :

- A13B2TP:Product Tanker
- A12A2TC:Chemical Tanker
- A11C2LC:CO2 Tanker
- A22A2BB:Bulk/Oil Carrier(OBO)
- A24B2BW:Woods Chips Carrier
- A24H2BZ:Powder Carrier

Δίνονται οι ορισμοί για τους τύπους πλοίων General Cargo :

1. **General Cargo Ship with Ro-Ro facility** StatCode5v: A13A2GA , είναι πλοία τύπου Γ/Φ με δυνατότητα φορτοεκφόρτωσης οχημάτων σε ένα περιορισμένο ποσοστό του διαθέσιμου χώρου φορτίου.
2. **Open Hatch Cargo Ship** StatCode5v: A13A2GO , είναι μεγάλο μεγέθους πλοία Γ/Φ ενός καταστρώματος με ανοίγματα πλήρους πλάτους στο κατάστρωμα και τετραγωνισμένα αμπάρια για τη μεταφορά μοναδοποιημένου ξηρού φορτίου όπως δασικά προϊόντα και *containers*. Συνήθως εφοδιάζονται με ατσάλινο γερανό.
3. **General Cargo/Tanker/(Container/Oil/Bulk-COB Ship)** StatCode5v: A31A2GS , είναι πλοία τύπου Γ/Φ με αναστρέψιμες μπουκαπόρτες. Ενώ η μια πλευρά είναι ισόπεδη, η άλλη έχει διαφράγματα ώστε να μεταφέρονται υγρά φορτία. Επίσης όταν οι μπουκαπόρτες είναι σε λειτουργία για ξηρό φορτίο, μπορούν να μεταφερθούν και *containers*.
4. **General Cargo/Tanker** StatCode5v: A31A2GT, είναι πλοία τύπου Γ/Φ με δυνατότητα για επιπρόσθετη μεταφορά υγρού φορτίου λόγω των δεξαμενών που διαθέτει.
5. **General Cargo Ship** StatCode5v: A31A2GX, είναι πλοίο με ένα ή περισσότερα καταστρώματα και μπορεί να μεταφέρει πολλών ειδών φορτία ξηρού τύπου. Στις περιπτώσεις που υπάρχει ένα μόνο κατάστρωμα, μεταφέρονται φορτία σε μορφή κιβωτίων. Η φορτοεκφόρτωση γίνεται από τα ανοίγματα του καταστρώματος καιρού.

6. **Palletized General Cargo** StatCode5n: A31B2GP, είναι πλοίο τύπου Γ/Φ με ένα ή περισσότερα καταστρώματα στο οποίο η φορτοεκφόρτωση πραγματοποιείται με χρήση ανελκυστήρα φορτωμένο με παλέτες. Επίσης δεν υπάρχουν ανοίγματα στο κατάστρωμα καιρού σε αυτόν τον τύπο πλοίου.
7. **Deck Cargo Ship** StatCode5n: A31C2GD, είναι πλοία τα οποία φορτώνονται μόνο στο κατάστρωμα με μοναδοποιημένο φορτίο. Για την πρόσβαση σε αυτά χρειάζεται ράμπα ανάλογη των Ro-Ro.

Στα πλοία που πραγματοποιήθηκε ανάλυση δε συμπεριλαμβάνονται αυτά που :

1. Δόθηκαν για scrapping ή χάθηκαν πριν από τις 01/01/1995
2. Ήρθαν σε λειτουργία μετά τις 31/12/2015
3. Ήταν υπό παραγγελία
4. Ήταν σε αναμονή παράδοσης ή ακυρώθηκε η παραγγελία

Το μεγαλύτερο ποσοστό των πλοίων Γ/Φ που βρίσκονται σε λειτουργία αλλά και των ατυχημάτων που σημειώθηκαν την εξεταζόμενη περίοδο 1995-2016, ανήκουν στην κατηγορία General Cargo Ship με κωδικό A31A2GX. Η εκτίμησή μου είναι πως τα πλοία που δεν αντιστοιχούν σε κάποια άλλη κατηγορία εντάσσονται σε αυτήν.

Ένα ακόμη ενδιαφέρον στοιχείο είναι η απουσία του ονόματος της κλάσης του πλοίου στα δεδομένα της Sea-web. Παρατίθεται όμως ο πίνακας με τη λίστα των νηογνωμόνων που βρέθηκε στη σελίδα της Sea-Web.

Πίνακας 2 Αλφαβητική λίστα Νηογνωμόνων που αναφέρεται στη βάση δεδομένων της Sea-Web.

American Bureau of Shipping	Iranian Classification Society	Panama Register
Soc Andina de Certification	Indian Register of Shipping	Panama Marine Documentation
Asia Classification Society	International Register	Panama Marine Survey
Bulgarski Koraben Register	Intermaritime	Panama Shipping Registrar
Black Sea Bureau of Shipping	International Maritime Bureau	Phoenix Register of Shipping
Bureau Veritas	International Naval Surveys	Polski Rejestr Statkow
Belize Maritime Bureau	International Ship Classification	Registro Brasileiro
Columbus American Register	Jugoslavenski Registar Brodova	Registro Cubano de Buques
China Classification Society	Korea Classification Society	Registro Italiano Navale
CNRIN	Biro Klasifikasi Indonesia	Registrul Naval Roman
China Corporation Register	Korean Register of Shipping	Rihane Portuguesa
Croatian Register of Shipping	Korean Ship Safety Technology	Russian River Register
Cyprus Bureau of Shipping	Lloyd's Register	Russian Maritime Register
R J Del Pan	Maritime Bureau of Shipping	Sing-Lloyd
Dromon Bureau of Shipping	Macosnar Corp	Turk Lloyd
Deutsche Schiffs-Revision	Maritime Lloyd Georgia	Ukraine Shipping Register
Germanischer Lloyd	Maritime Technical Systems	Union Bureau of Shipping
Global Marine Bureau	Ships Classification Malaysia	Universal Maritime Bureau
Hellenic Register of Shipping	National Shipping Adjuster	Unknown
HORINSIB	Nippon Kaiji Kyokai	Universal Shipping Bureau
Isthmus Bureau of Shipping	Det Norske Veritas	Vietnam Register
INCLAMAR	Overseas Marine Certification	Zianlian Chuen
Isthmus Maritime Classification	Panama Bureau of Shipping	Joson Classification Society

### 3.3 Αναφορές Ατυχημάτων

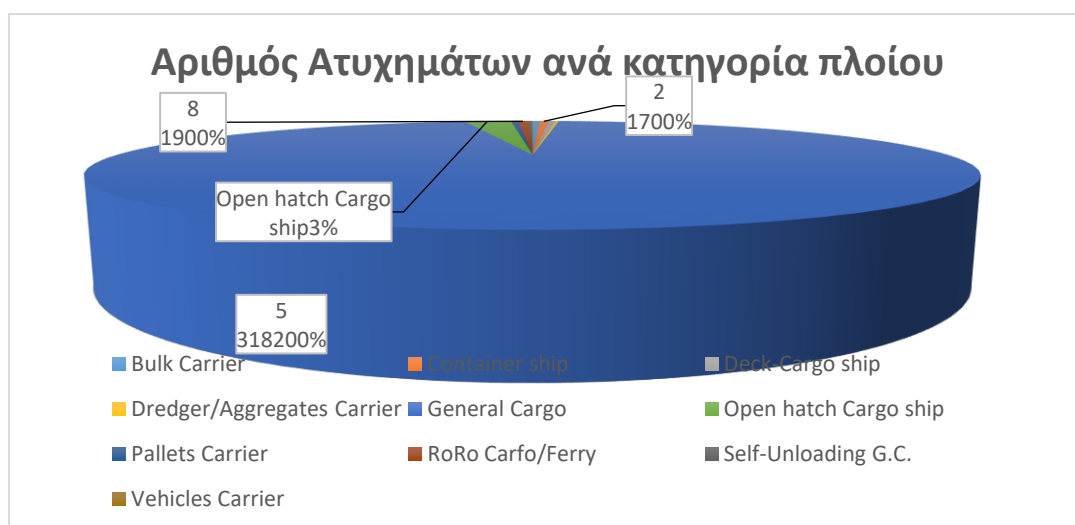
Ο αριθμός των ατυχημάτων προκύπτει εξ ολοκλήρου από τη βάση δεδομένων της Sea-Web και επειδή δεν επέτρεψαν πρόσβαση σε άλλη μεγάλη βάση, η σύγκριση για να διευκρινιστεί το επίπεδο πληρότητας της Sea-Web δε καταφέραμε να γίνει. Τα ατυχήματα αφορούν σοβαρές περιπτώσεις και όλα τα πλοία μεγέθους μικρότερου των 500 κόρων αποκλείστηκαν από την ανάλυση αφού η ολοκληρωμένη SOLAS συμπεριλαμβάνει μόνο τα μεγαλύτερα πλοία.

Αρχικά όσων αφορά στις διάφορες κατηγορίες πλοίων τύπου Γ/Φ βρέθηκαν τα εξής στοιχεία για αναφορές ατυχημάτων από τη βάση δεδομένων της Sea-Web.

Πίνακας 3 Αναφορές ατυχημάτων για κάθε κατηγορία πλοίων G.C. με GT>499 και για την περίοδο 01/01/1995 έως 31/12/2016.

Type of ship	Number	%
Bulk Carrier	12	0.36
Container ship	17	0.51
Deck-Cargo ship	12	0.36
Dredger/Aggregates Carrier	5	0.15
General Cargo	3182	95.24
Open hatch Cargo ship	82	2.45
Pallets Carrier	13	0.39
RoRo Carfo/Ferry	19	0.57
Self-Unloading G.C.	5	0.15
Vehicles Carrier	2	0.06
Total	3341	100

Από τον πίνακα 3 γίνεται ξεκάθαρο πόσο μεγάλο ποσοστό ατυχημάτων καλύπτουν τα General Cargo ships από το συνολικό αριθμό των πλοίων Γενικού Φορτίου. Παρατίθενται γραφικά στο επόμενο σχήμα (Σχήμα 1) τα αποτελέσματα του πίνακα 3.



Σχήμα 1 Ποσοστά ατυχημάτων για όλους τους τύπους G.C. με GT>499 και για την περίοδο 01/01/1995 έως 31/12/2016.

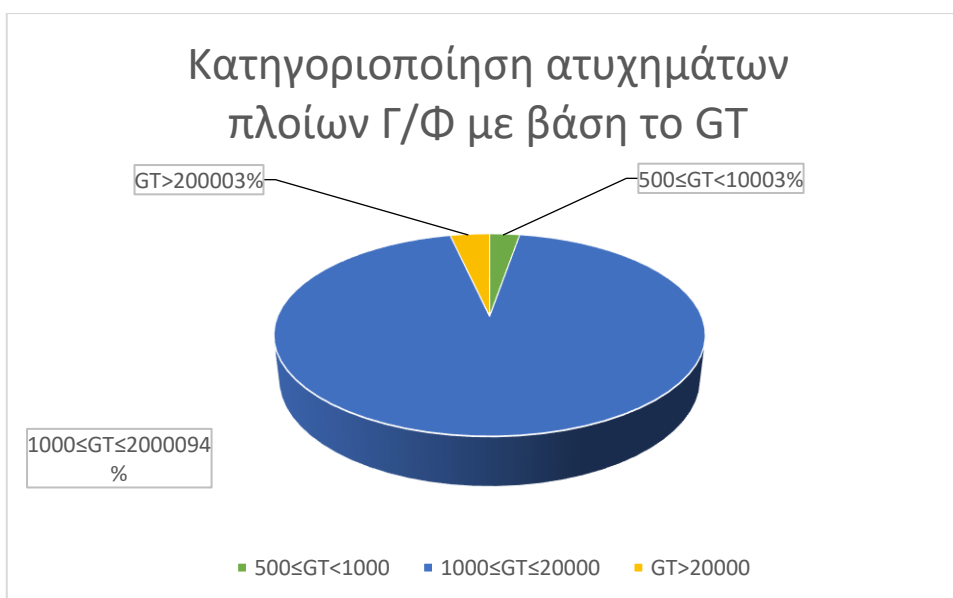
Ακολουθούν τα στοιχεία ταξινόμησης των περιπτώσεων ατυχημάτων που κατέχουμε από τη βάση της Sea-Web χωρισμένα κατά:

1. Το μέγεθος του πλοίου
2. Τον τύπο του ατυχήματος

Ο πίνακας που ακολουθεί (**Πίνακας 4**) παρουσιάζει, σύμφωνα με την κατηγοριοποίηση των πλοίων τύπου Γ/Φ που έγινε με βάση το μέγεθός τους, το συνολικό αριθμό ατυχημάτων για τα οποία υπάρχουν καταχωρημένα στοιχεία.

Πίνακας 4 Αριθμός ατυχημάτων για την κάθε κατηγορία μεγέθους για την περίοδο 01/01/1995 έως 31/12/2016.

$500 \leq GT \leq 1000$	$1000 < GT \leq 20000$	$GT > 20000$
84	2928	110



Σχήμα 2 Γραφική αναπαράσταση της ποσοστιαίας κατανομής ατυχημάτων πλοίων τύπου Γ/Φ με βάση το GT για την περίοδο 01/01/1995-31/12/2016

Είναι εμφανές από το **Σχήμα 2** ότι τα περισσότερα ατυχήματα συμβαίνουν σε πλοία της δεύτερης σε μέγεθος κατηγορίας, δηλαδή  $1000 < GT \leq 20000$ , και αυτό συμβαίνει διότι είναι και η μεγαλύτερη κατηγορία σε αριθμό πλοίων. Παραπάνω παρατηρήσεις θα γίνουν στη συνέχεια όπου θα δοθούν στοιχεία του υπό μελέτη στόλου της περιόδου που μας απασχολεί.

Τα είδη των ατυχημάτων που μπορούν να αναζητηθούν από τη βάση δεδομένων είναι τα εξής:

- Collision
- Contact
- Fire/Explosion
- Foundered
- Hull/Mchy Damage

- Wrecked/Stranded
- War Loss/Hostilities
- Missing
- Miscellaneous

Από τα παραπάνω είδη ατυχήματος εξαιρέθηκαν αυτά που οφείλονται σε πόλεμο/εχθροπραξίες και αυτά που αγνοούνται διότι δεν υπάρχει κάποια χρήσιμη πληροφορία που να μπορεί να χρησιμοποιηθεί στην παρούσα διπλωματική.

Στο σημείο αυτό θα δοθούν οι ορισμοί για την κατάσταση που αντιπροσωπεύει ο καθένας από τους παραπάνω τύπους ατυχημάτων.

➤ **Σύγκρουση (Collision)**

Στα ατυχήματα σύγκρουσης υπάρχει απρόσμενη επαφή μεταξύ δύο πλοίων. Στην κατηγορία αυτή το πλοίο Γ/Φ χτυπά ένα άλλο ή το αντίστροφο, δηλαδή δέχεται το χτύπημα.

➤ **Επαφή (Contact)**

Ατυχήματα επαφής είναι αυτά στα οποία το πλοίο ήρθε σε επαφή με σταθερά ή επιπλέοντα στοιχεία.

➤ **Ημιβύθιση/Βύθιση (Foundered)**

Ατυχήματα βύθισης είναι αυτά που οδηγούν σε βύθιση του πλοίου.

➤ **Φωτιά/Έκρηξη (Fire explosion)**

Ατυχήματα φωτιάς/έκρηξης είναι αυτά στα οποία το κύριο και αρχικό γεγονός (main event) ήταν φωτιά/έκρηξη.

➤ **Αστοχία Γάστρας/Μηχανική Βλάβη (Hull/machinery Damage)**

Ατυχήματα αστοχίας της γάστρας είναι αυτά στα οποία παρουσιάστηκε ρήγμα/ρωγμή στη γάστρα του πλοίου με άμεση συνέπεια στην πλευστότητα και την ακεραιότητα του περιβλήματος.

Στα ατυχήματα μηχανικής βλάβης ανήκουν αυτά που το κύριο και αρχικό γεγονός αποτελεί βλάβη σε κάποιο από τα εξαρτήματα της μηχανής ή κάποιας αντλίας. Στην κατηγορία αυτή εντάσσονται επίσης και βλάβες στο σύστημα πηδαλίου.

➤ **Προσάραξη (Wrecked/Stranded)**

Είναι ατυχήματα όπου το πλοίο έμεινε ακινητοποιημένο είτε λόγω προσάραξης είτε λόγω κακοκαιρίας.

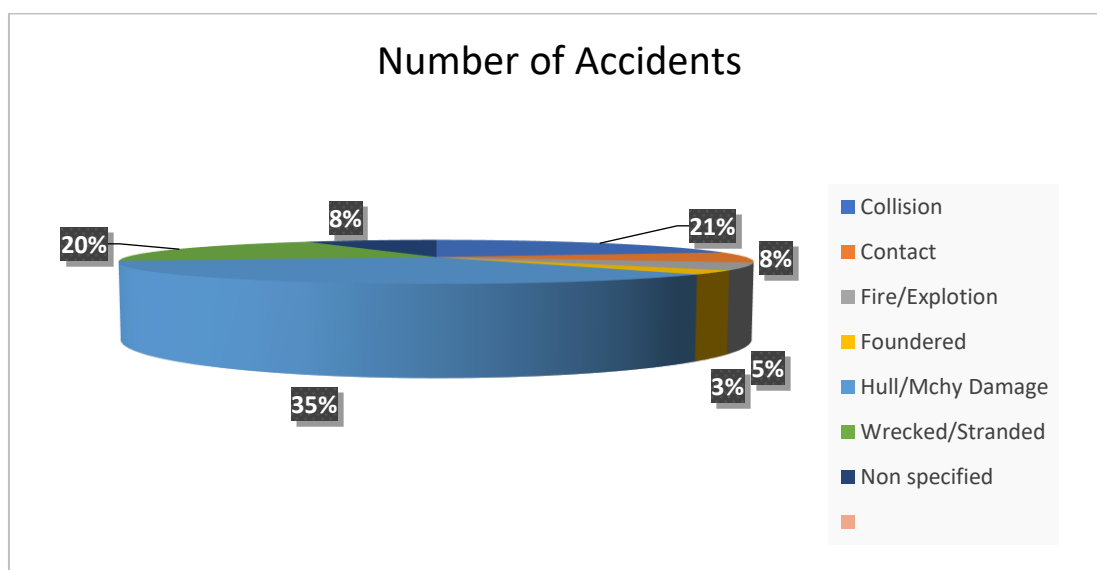
➤ **Υπόλοιπα (Miscellaneous)**

Στην κατηγορία αυτή εντάσσονται ατυχήματα που δε μπορούν να αντιστοιχηθούν με κάποια από τις παραπάνω κατηγορίες. Στη συνέχεια η κατηγορία αυτή θα χρησιμοποιηθεί με όνομα Non Specified διότι υπήρχαν και αρκετά πλοία που τα δεδομένα που βρήκαμε για αυτά ήταν ελάχιστα και δε μπορούσαν να χρησιμεύσουν στην ανάλυση. Έτσι, συνδυάστηκαν με την μικρή σε αριθμό κατηγορία Miscellaneous.

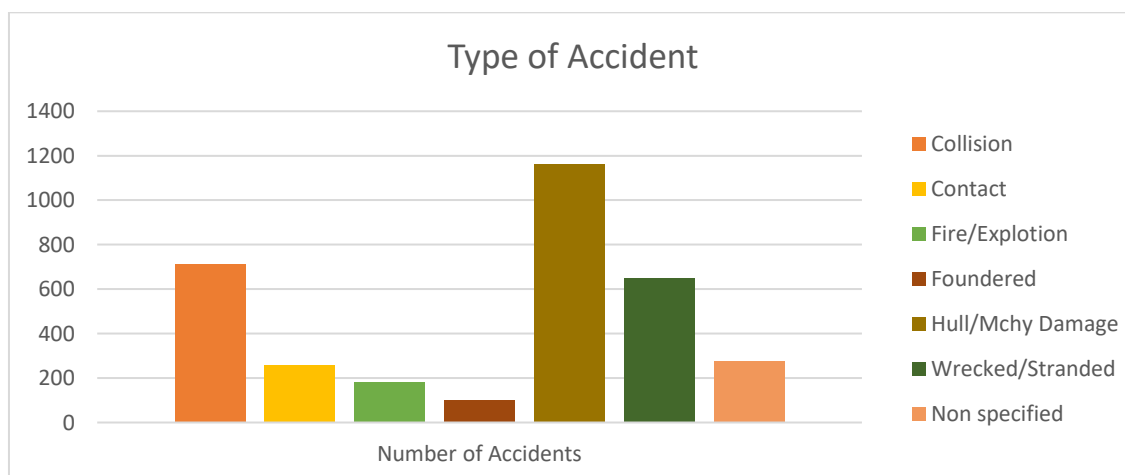
Αρχικά παρουσιάζεται μια πρώτη εικόνα για το αριθμητικό μέγεθος της κάθε κατηγορίας ατυχήματος για την υπό μελέτη περίοδο. Συγκεκριμένα από τα ιστορικά στοιχεία που εξορύχθηκαν από τη βάση δεδομένων της Sea-Web προκύπτουν τα εξής:

Πίνακας 5 Αριθμός ατυχημάτων ανά κατηγορία ατυχήματος για πλοία Γ/Φ GT>499 και για την περίοδο 01/01/1995-31/12/2016.

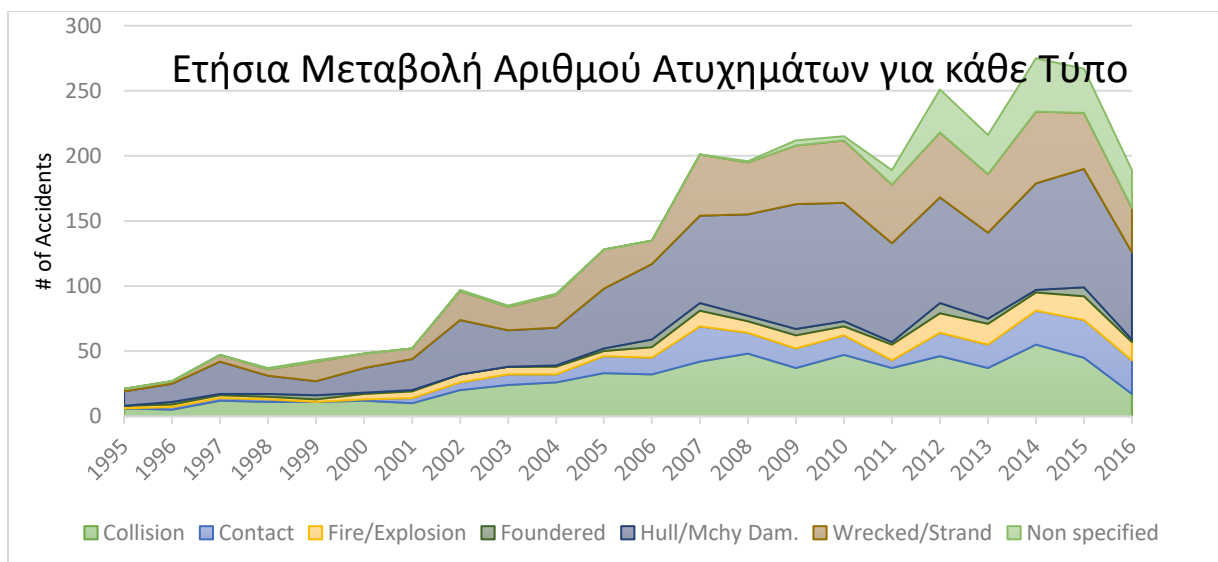
Casualty type	Number of Accidents	% of Total
Collision	711	21.28
Contact	258	7.72
Fire/Explosion	182	5.45
Foundered	99	2.96
Hull/Mchy Damage	1162	34.78
Wrecked/Stranded	651	19.49
Non specified	278	8.32
Total	3341	100.00



Σχήμα 3 Γραφική αναπαράσταση της ποσοστιαίας κατανομής του συνολικού αριθμού των ατυχημάτων στις διάφορες κατηγορίες ατυχημάτων.



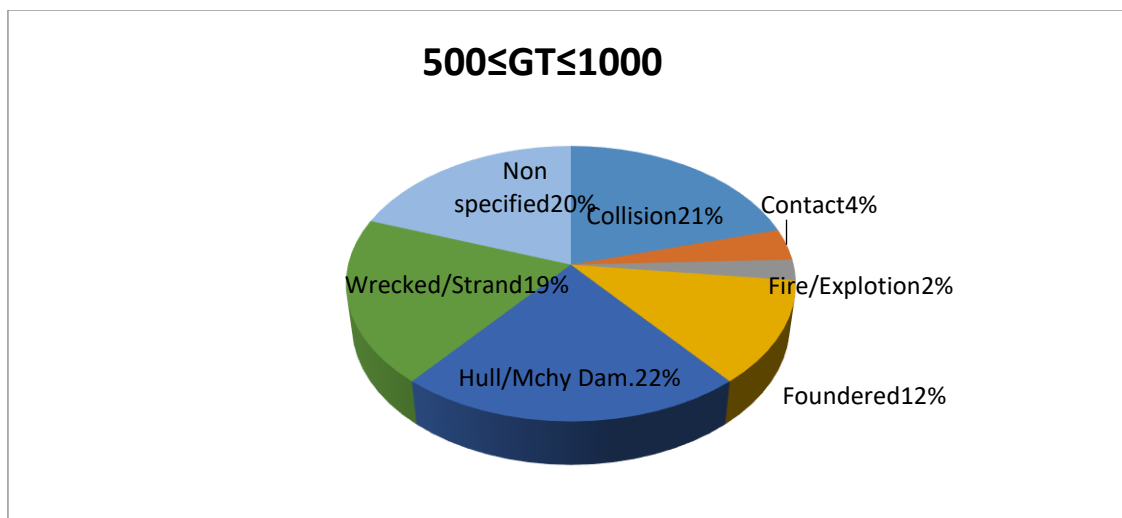
Σχήμα 4 Γραφική αναπαράσταση του αριθμού των ατυχημάτων για κάθε τύπο ατυχήματος για την περίοδο 01/01/1995-31/12/2016.



Σχήμα 5 Ετήσια μεταβολή του αριθμού ατυχημάτων για κάθε τύπο ατυχήματος, για την περίοδο 01/01/1995-31/12/2016.

Στη συνέχεια εξετάζονται οι τρεις κατηγορίες μεγέθους GT όπως αυτές ορίστηκαν νωρίτερα με σκοπό να προσδιοριστούν από μια ακόμη σκοπιά τα αποτελέσματα των ατυχημάτων.

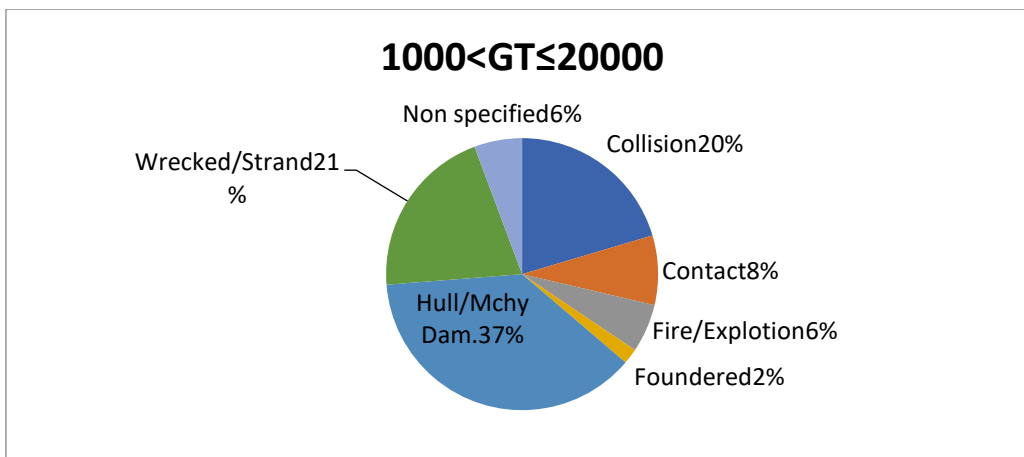
Ξεκινάμε την παρουσίαση των αποτελεσμάτων με την κατηγορία πλοίων με  $500 \leq GT \leq 1000$  όπου στο σχήμα που ακολουθεί (Σχήμα 6) απεικονίζεται η ποσοστιαία κατανομή των συνολικών ατυχημάτων με βάση τον τύπο του ατυχήματος. Βλέπουμε ότι το μεγαλύτερο ποσοστό ανήκει στον τύπο ατυχήματος Hull/Machinery damage (22%). Ακολουθούν τα ατυχήματα Collision (21%) και Wrecked/Strand (19%). Μικρότερα ποσοστά καταλαμβάνουν τα ατυχήματα Foundered (12%) και Contact (4%), Fire/Explosion (2%). Σημαντικό όμως είναι και το ποσοστό των πλοίων που είχαν ατύχημα αλλά δεν προσδιορίζεται το είδος στη βάση δεδομένων.



Σχήμα 6 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με  $500 \leq GT \leq 1000$  για την περίοδο 01/01/1995-31/12/2016).

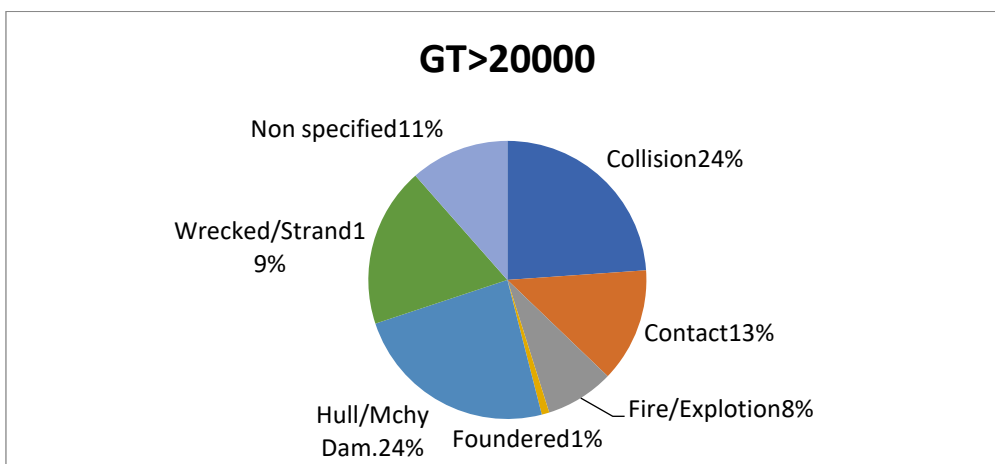
Για την κατηγορία πλοίων με  $1000 < GT \leq 20000$ , όπως παρουσιάζεται και στο ακόλουθο σχήμα (Σχήμα 7), το κυρίαρχο ποσοστό των ατυχημάτων με 37% καταλαμβάνουν τα ατυχήματα Hull/Machinery damage. Ακολουθούν τα ατυχήματα Wrecked/Strand (21%) Collision (20%). Μικρότερα ποσοστά καταλαμβάνουν τα ατυχήματα Contact (8%), Fire/Explosion (6%) και Foundered (2%).





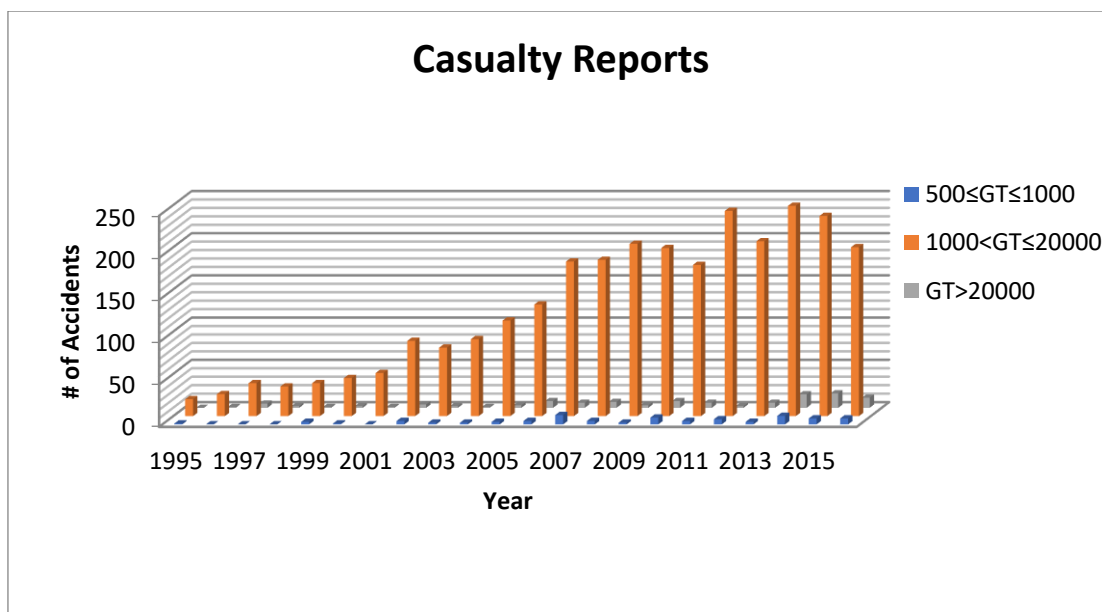
Σχήμα 7 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με 1000<GT≤20000 για την περίοδο 01/01/1995-31/12/2016.

Το **Σχήμα 8** παρουσιάζει την κατανομή των ατυχημάτων στους διάφορους τύπους για τα πλοία Γ/Φ με GT>20000 κόρων. Προκύπτει ότι το μεγαλύτερο ποσοστό ανήκει στους τύπους ατυχήματος Hull/Machinery damage και Collision (24%). Ακολουθούν με μικρότερα ποσοστά οι τύποι Wrecked/Strand (19%), Contact (13%), Fire/Explosion (8%) και Foundered (1%).



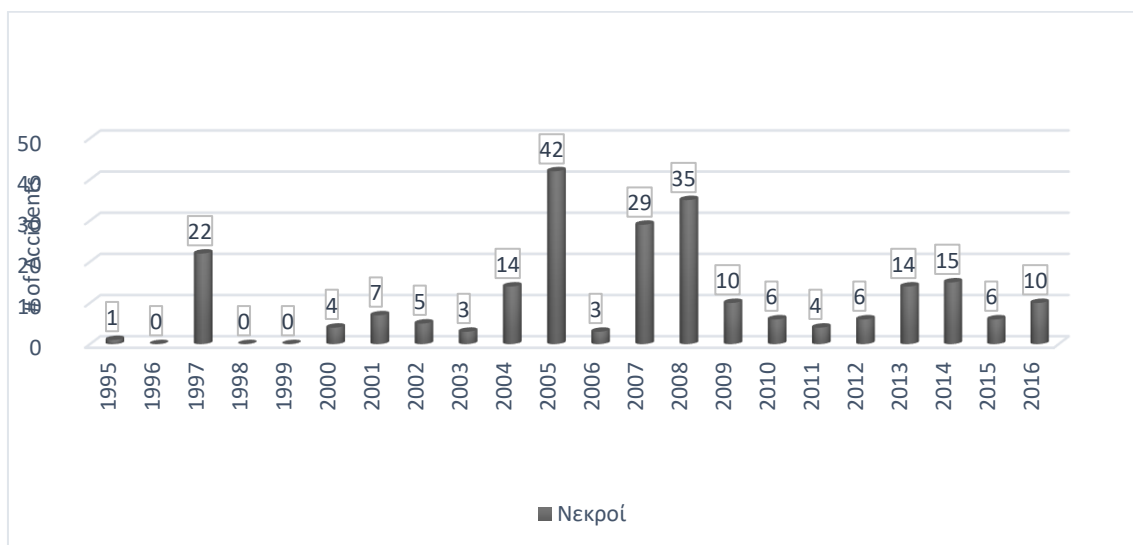
Σχήμα 8 Κατανομή των ατυχημάτων με βάση τον τύπο ατυχήματος (πλοία τύπου Γ/Φ με GT>20000 για την περίοδο 01/01/1995-31/12/2016.

Στο επόμενο σχήμα (**Σχήμα 9**), αναδεικνύεται για κάθε μία από τις τρεις κατηγορίες μεγέθους, η ετήσια μεταβολή του αριθμού των ατυχημάτων για την περίοδο 1995-2016. Παρατηρούμε ότι για την μεσαία κατηγορία μεγέθους, η οποία αποτελείται και από τα περισσότερα πλοία, υπάρχει μια σταθερή αύξηση με την πάροδο των χρόνων. Συγκρίνοντας τις πρώτες χρονιές καταγραφής με τις τελευταίες, καταγράφεται αύξηση ατυχημάτων σχεδόν 1000%. Στις άλλες δύο κατηγορίες δε μπορούν να βγουν ξεκάθαρα συμπεράσματα επειδή οι στόλοι είναι κατά πολύ μικρότερη από τη μεσαία κατηγορία.

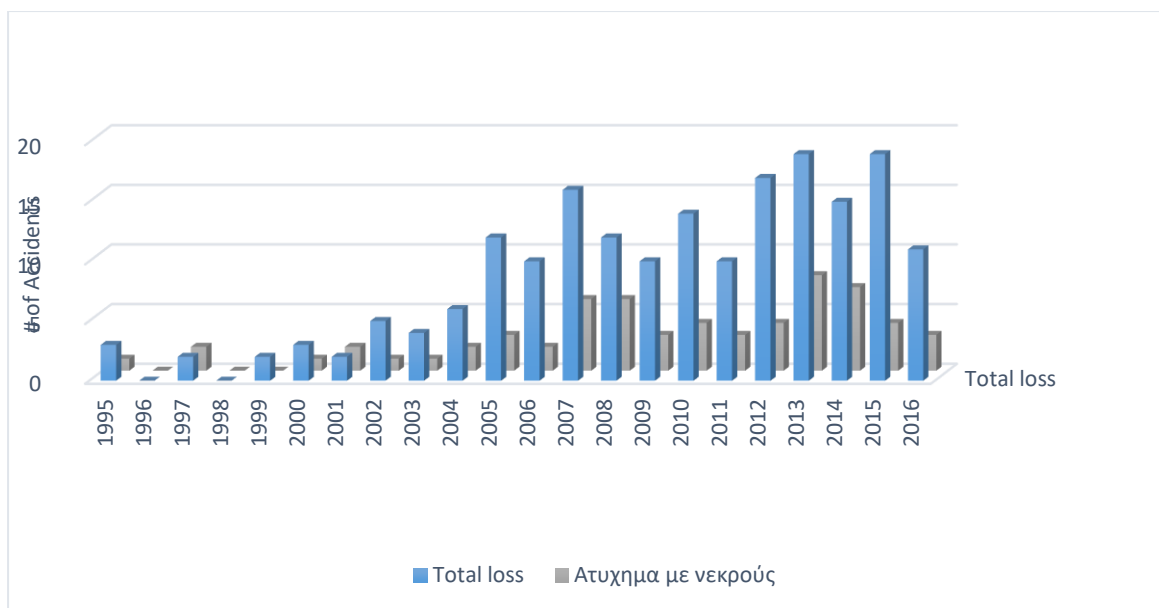


Σχήμα 9 Ετήσια μεταβολή του αριθμού των ατυχημάτων για κάθε κατηγορία μεγέθους.

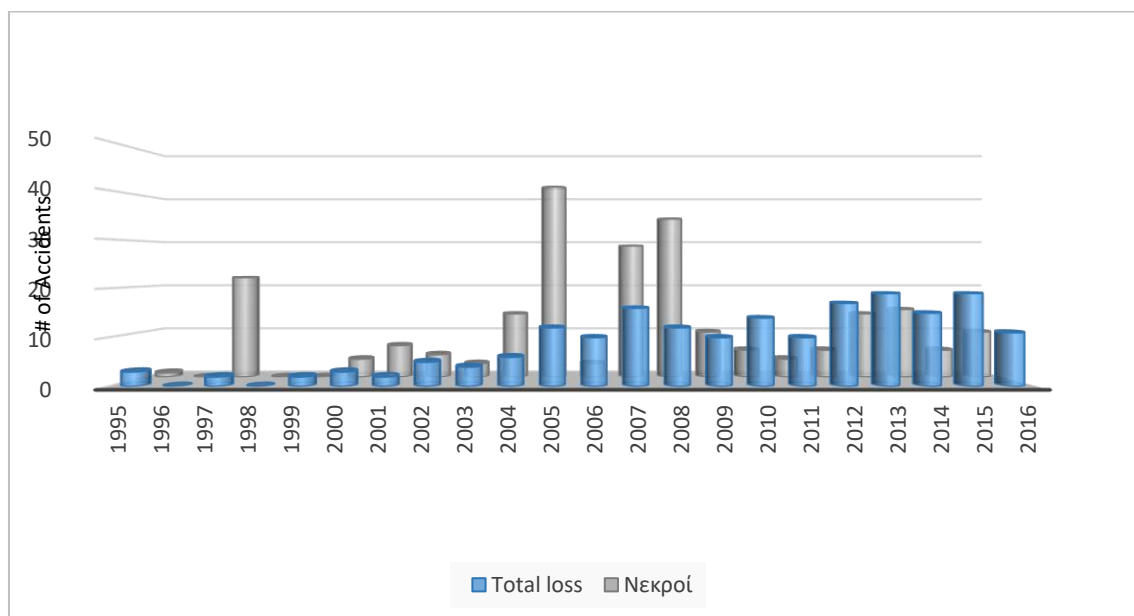
Στο **Σχήμα 10** απεικονίζεται ο αριθμός των νεκρών που προκύπτουν κάθε χρονιά από τα ατυχήματα στα πλοία Γ/Φ. Σε συνδυασμό με αυτό προκύπτουν τα σχήματα 11 όπου συνυπάρχουν ατυχήματα με total loss του πλοίου μαζί με αναφορές για νεκρούς και 12 όπου συνυπάρχουν ατυχήματα με total loss του πλοίου μαζί με τον αριθμό των νεκρών. Κοιτάζοντας τα διαγράμματα ταυτόχρονα και συγκρίνοντας τα αποτελέσματα βλέπουμε πως το 2005 και το 2008 ενώ οι αναφορές για ατυχήματα με νεκρούς ήταν λίγες, τα ατυχήματα ήταν πολύνεκρα.



Σχήμα 10 Ο αριθμός των νεκρών που προκύπτουν κάθε χρονιά



Σχήμα 11 Ατυχήματα με ολική απώλεια του πλοίου (total loss) του πλοίου μαζί με αναφορές για νεκρούς

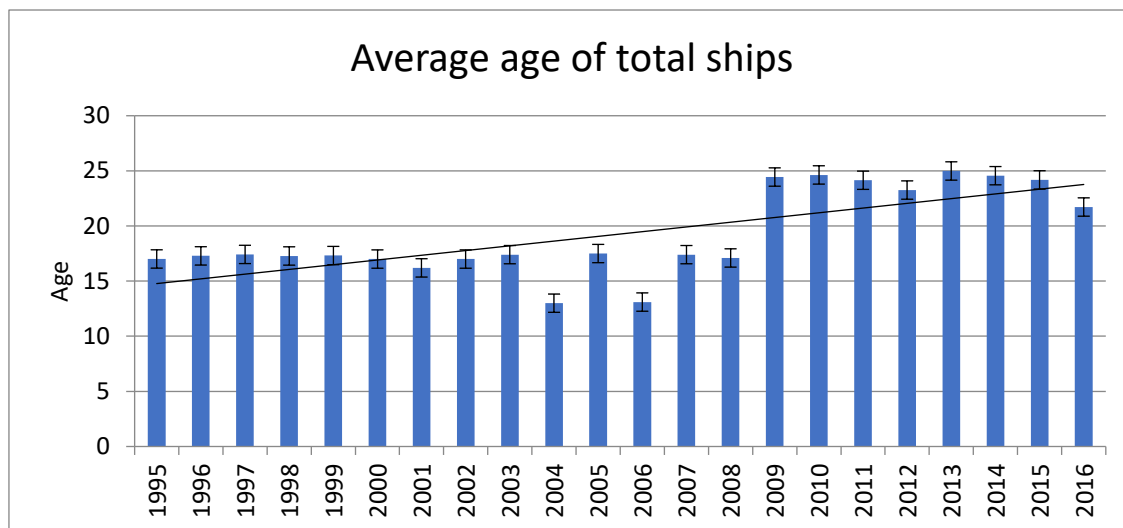


Σχήμα 12 Ατυχήματα με ολική απώλεια του πλοίου (total loss) του πλοίου μαζί με τον αριθμό των νεκρών

Πίνακας 6 Αριθμός ατυχημάτων που προκάλεσαν μόλυνση του περιβάλλοντος για πλοία Γ/Φ GT>499 και για την περίοδο 01/01/1995-31/12/2016.

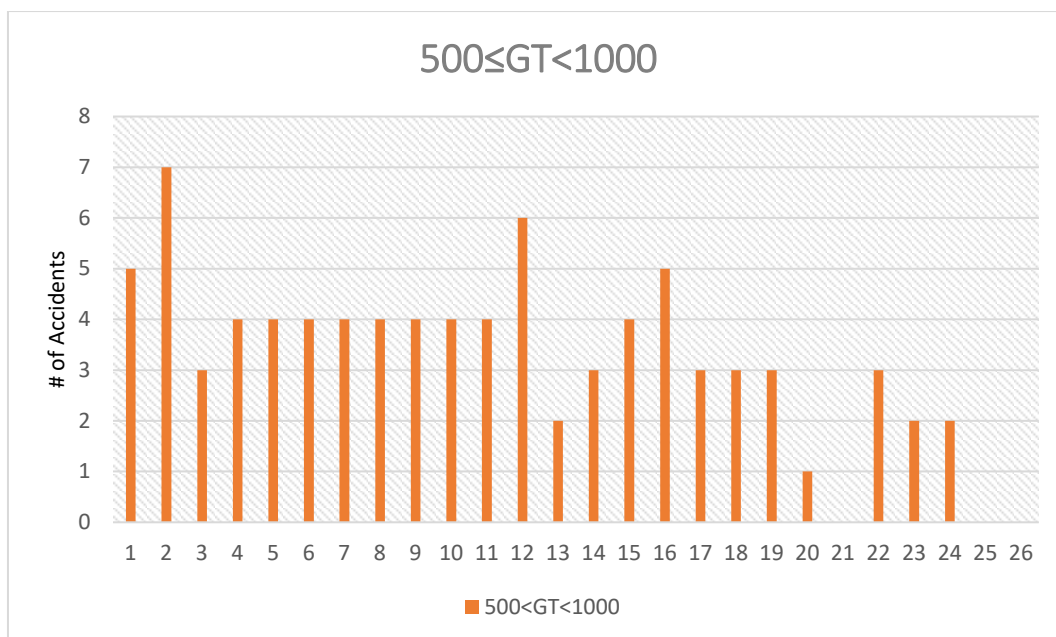
Pollution 1995-2016	
Yes	70
Unknown	143
No	2846
Total	3059

Ένα πολύ σημαντικό στοιχείο που θα χρησιμοποιηθεί για την αξιολόγηση των ιστορικών δεδομένων είναι η ηλικία του πλοίου τη στιγμή που πραγματοποιήθηκε το ατύχημα. Στο **Σχήμα 13** απεικονίζεται ο ετήσιος μέσος όρος της ηλικίας των πλοίων που είχαν ατυχήματα την περίοδο 1995-2016.

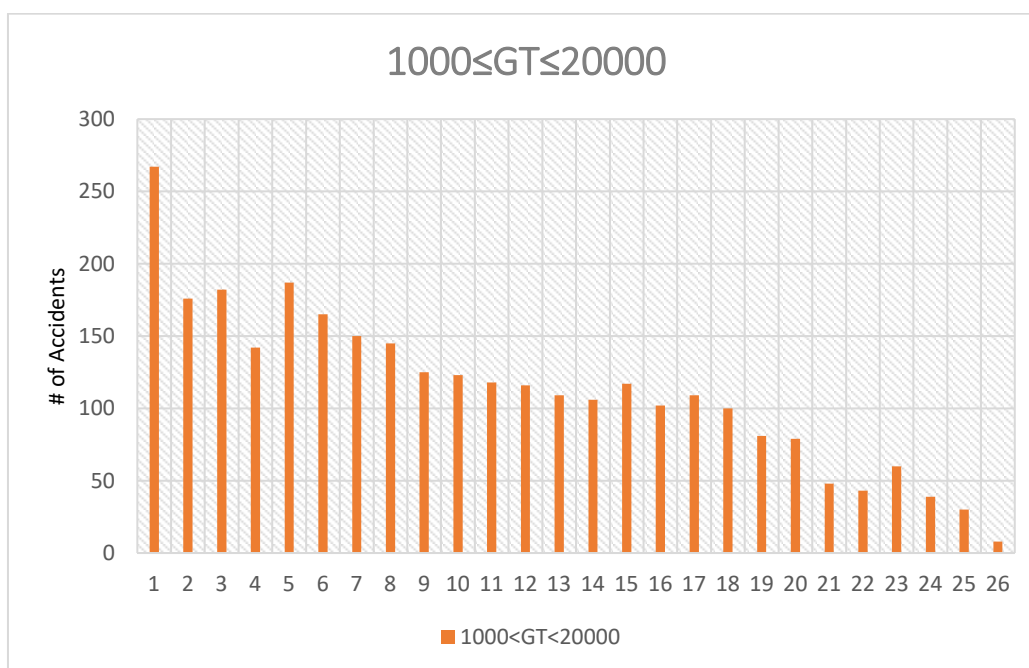


Σχήμα 13 Ετήσιος μέσος όρος της ηλικίας των πλοίων που είχαν ατυχήματα την περίοδο 1995-2016.

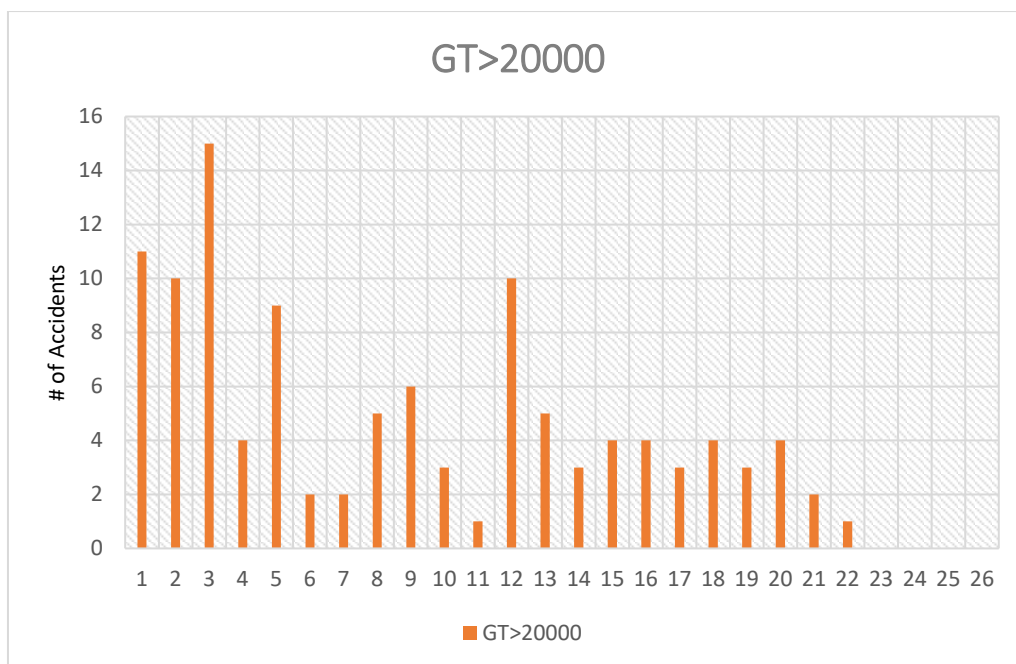
Επίσης σημαντικό στοιχείο αποτελεί η αναλυτική ηλικία των πλοίων στην οποία προκύπτει το ατύχημα. Σε συνδυασμό με το μέγεθος του πλοίου προκύπτουν τα παρακάτω διαγράμματα. Για παράδειγμα στην κατηγορία μεγέθους  $500 \leq GT \leq 1000$  στο **Σχήμα 14** δεν προέκυψε ατύχημα με πλοίο 21 ετών, ενώ σημειώθηκαν 7 ατυχήματα σε πλοία που είχαν ηλικία 2 ετών. Όσον αφορά την κατηγορία μεγέθους  $1000 < GT \leq 20000$ , το **Σχήμα 15** παρουσιάζει μια ομοιόμορφα πτωτική τάση για τον αριθμό των ατυχημάτων καθώς η ηλικία του πλοίου αυξάνεται. Αυτό προκύπτει πρώτον επειδή είναι η ομάδα με τα περισσότερα σε αριθμό πλοία και δεύτερον επειδή τα δεδομένα της Sea-Web είναι εμπλουτισμένα για τα σχετικά πρόσφατα χρόνια του παρελθόντος άλλα παρουσιάζουν ελλείψεις για δεδομένα που αφορούν δύο δεκαετίες πριν.



Σχήμα 14 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία 500≤GT<1000



Σχήμα 15 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία 1000≤GT≤20000



Σχήμα 16 Αριθμός ατυχημάτων πλοίων ηλικίας 1-26 ετών τα οποία ανήκουν στην κατηγορία GT>20000

### 3.4 Στόλος πλοίων Γ/Φ ανά έτος

Για τη συνέχεια της ανάλυσης των στοιχείων με στόχο τον υπολογισμό συχνοτήτων διαφόρων μεγεθών για την περίοδο που εξετάζουμε και με βάση τα ιστορικά δεδομένα που εξορύχθηκαν από τη Sea-Web είναι απαραίτητο να δηλώσουμε τον ετήσιο υπό μελέτη στόλο πλοίων Γ/Φ που βρίσκονταν σε λειτουργία για τη δεδομένη περίοδο (1995-2016). Όμως θα φανεί παρακάτω πως χρειαζόμαστε τον στόλο και για μερικά χρόνια πριν το 1995, γι' αυτό και ο πίνακας περιλαμβάνει το στόλο που μας ενδιαφέρει αλλά με δεδομένα από το 1990 έως το 2016.

Στον υπό μελέτη στόλο λοιπόν ανήκουν πλοία με  $GT \geq 500$  και Date of Build 01/01/1990 έως 31/12/2016.

Πίνακας 7 Ετήσιος υπό Μελέτη Στόλος για την περίοδο 1990-2016

Fleet	Year
4153	1990
4348	1991
4734	1992
5074	1993
5411	1994
5807	1995
6211	1996
6585	1997
6896	1998
7162	1999
7375	2000
7615	2001
7802	2002
8004	2003
8225	2004
8527	2005
8876	2006
9231	2007
9658	2008
10127	2009
10568	2010
10988	2011
11329	2012
11670	2013
11650	2014
11729	2015
11683	2016

## ΚΕΦΑΛΑΙΟ 4<sup>ο</sup> : ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΤΩΝ ΑΤΥΧΗΜΑΤΩΝ

Σε αυτό το κεφάλαιο γίνεται η ανάλυση όλων των αναφορών ατυχημάτων, που εξετάστηκαν μέσω της βάσης δεδομένων της Sea-Web, σχετικά με τις συνέπειές τους, όπως ανθρώπινες απώλειες (fatalities), ολικές απώλειες πλοίων (total losses), αν υπήρχαν αγνοούμενοι (missing) κ.α. Περιβαλλοντικές συνέπειες δεν θα εξεταστούν διότι ο αριθμός των ατυχημάτων στα οποία υπήρξε διαρροή πετρελαίου ή απώλεια φορτίου είναι ελάχιστος. Σημαντική παράμετρος του κεφαλαίου αυτού είναι ο υπολογισμός συχνότητας των παραπάνω ατυχημάτων και η γραφική τους απεικόνιση.

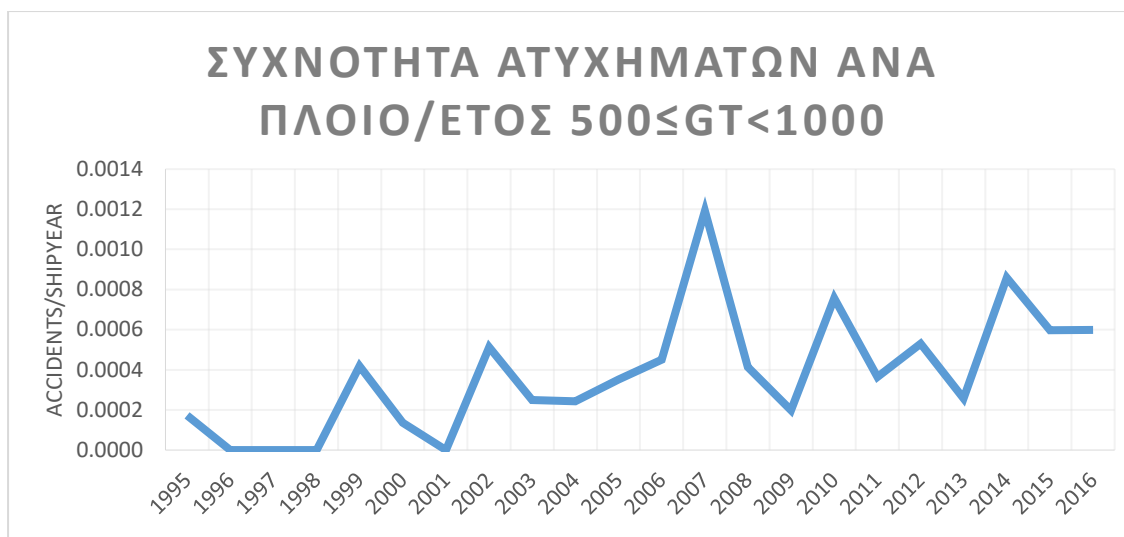
### 4.1 Σοβαρά ατυχήματα ανά πλοία-έτη για την περίοδο 01/01/1995-31/12/2016

Στο πρώτο από τα διαγράμματα που ακολουθούν (**Σχήμα 17**) παρουσιάζεται η ετήσια μεταβολή της συχνότητας ατυχημάτων για το σύνολο των υπό μελέτη πλοίων της εξεταζόμενης περιόδου. Παρατηρείται ανοδική πορεία των accident rates με υψηλότερη τιμή τη χρονιά 2014.



Σχήμα 17 Ετήσια συχνότητα ατυχημάτων για όλες τις κατηγορίες μεγέθους.





Σχήμα 18 Ετήσια συχνότητα ατυχημάτων για πλοία με 500≤GT<1000.



Σχήμα 19 Ετήσια συχνότητα ατυχημάτων για πλοία με 1000≤GT≤20000.



Σχήμα 20 Ετήσια συχνότητα ατυχημάτων για πλοία με GT>20000.

Αν παρουσιάσουμε τον Μέσο Όρο για την καθεμία από τις τρεις κατηγορίες μεγέθους οι τιμές θα είναι οι ακόλουθες:

- $0,38 \times 10^{-3}$  για πλοία με  $500 \leq GT \leq 1000$
- $13 \times 10^{-3}$  για πλοία με  $1000 < GT \leq 20000$
- $0,52 \times 10^{-3}$  για πλοία με  $GT > 20000$

Η δεύτερη κατηγορία μεγέθους φαίνεται ξεκάθαρα ότι αντιπροσωπεύει το σύνολο του δείγματος με μέση τιμή πολλαπλάσια των άλλων κατηγοριών. Επίσης παρουσιάζει ομαλότητα και ομοιομορφία στα διαγράμματα αφού το δείγμα της είναι μεγαλύτερο σε αριθμό. Οι άλλες κατηγορίες με μικρό στον αριθμό δείγμα παρουσιάζουν αυξομειώσεις στις ετήσιες συχνότητες και τα συμπεράσματα στο συγκεκριμένο πεδίο δε μπορούν να είναι ακριβή.

#### 4.1.1 Confidence Intervals (c.i)

Στη Στατιστική, ένα confidence interval αποτελεί ένα συγκεκριμένο είδος διαστήματος εκτίμησης μιας παραμέτρου ενός πληθυσμού και χρησιμοποιείται ως ένδειξη της αξιοπιστίας της εκάστοτε εκτίμησης. Είναι ένα διάστημα που υπολογίζεται από τις παρατηρήσεις και είναι διαφορετικό από δείγμα σε δείγμα. Η τιμή της παραμέτρου μπορεί να περιλαμβάνεται στο διάστημα εκτίμησης. Το πόσο συχνά περιλαμβάνεται η τιμή της παραμέτρου στο διάστημα εκτίμησης εξαρτάται από το επίπεδο ή σταθερά εμπιστοσύνης (confidence level or confidence coefficient).

Ένα confidence interval με ένα συγκεκριμένο επίπεδο εμπιστοσύνης έχει σκοπό να δώσει τη βεβαιότητα πως αν το στατιστικό μοντέλο είναι σωστό τότε το ποσοστό των δεδομένων που χρησιμοποιήθηκαν για την κατασκευή του c.i και περιέχονται μέσα σε αυτό θα είναι ίσο με το επίπεδο εμπιστοσύνης. Πιο συγκεκριμένα, η έννοια του όρου «επίπεδο εμπιστοσύνης» είναι ότι αν c.i κατασκευάζονταν για αναλύσεις δεδομένων ενός επαναλαμβανόμενου πειράματος, το ποσοστό αυτών των c.i που θα περιείχαν την αληθινή τιμή της παραμέτρου θα ήταν περίπου ίσο με το «επίπεδο εμπιστοσύνης». Αυτή είναι η αρχή στην οποία βασίζεται ο υπολογισμός των c.i.

Ένα confidence interval δεν προβλέπει ότι η αληθής τιμή της παραμέτρου έχει συγκεκριμένη πιθανότητα να είναι εντός του c.i με βάση τα δεδομένα που ο μελετητής έχει στα χέρια του. Ένα interval με αυτή την ιδιότητα, καλείται credible interval και υπολογίζεται με Bayesian μεθόδους ενώ έχει τα δικά του πλεονεκτήματα και μειονεκτήματα.

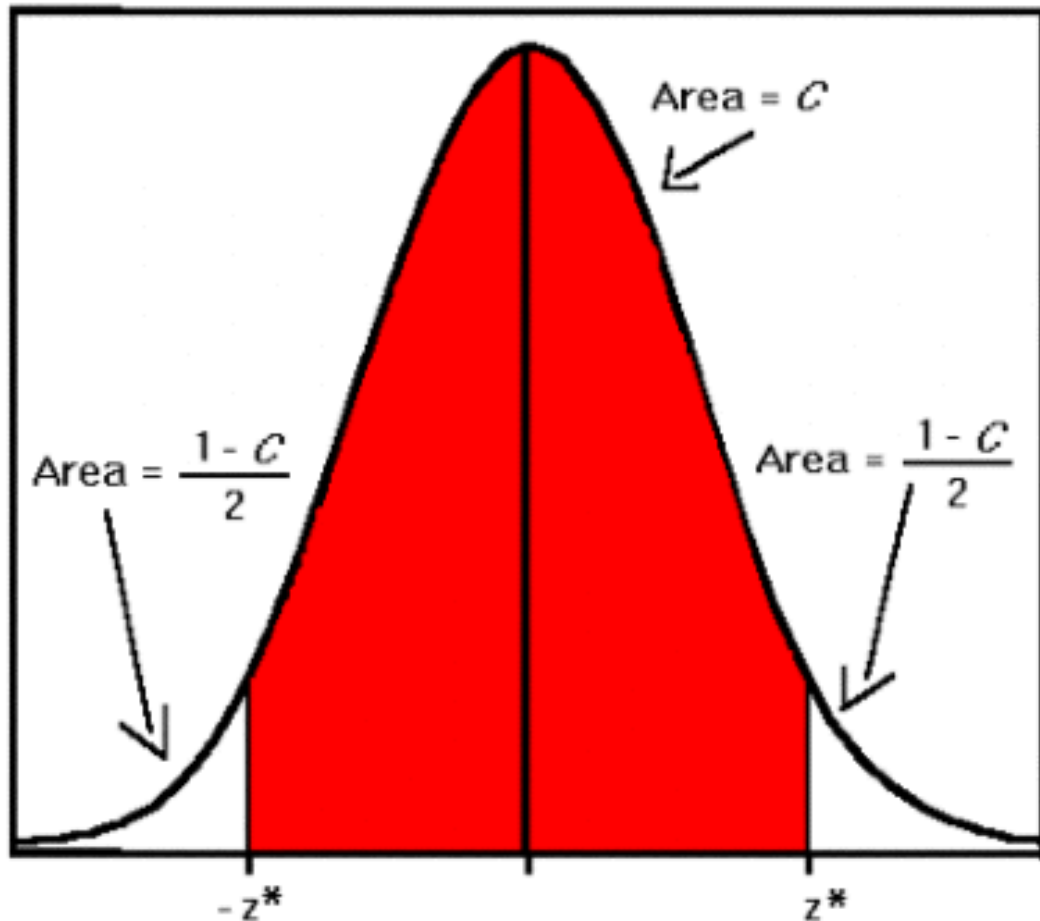
Η επιλογή του επιπέδου εμπιστοσύνης καθορίζει την πιθανότητα του να περιέχεται η αληθής τιμή της παραμέτρου στο παραγόμενο c.i. Οι πιο κοινές επιλογές επιπέδου εμπιστοσύνης είναι 0.90,0.95,0.99. Οι τιμές αυτές αντιπροσωπεύουν το ποσοστό των αντίστοιχων περιοχών σε μια καμπύλη τυπικής κανονικής κατανομής, όπως φαίνεται στο επόμενο σχήμα. Για παράδειγμα, ένα 95% confidence interval καλύπτει το 95% της περιοχής κάτω από την καμπύλη υποδηλώνοντας πως υπάρχει πιθανότητα μόλις 5% η τιμή της παραμέτρου να είναι εκτός αυτής της περιοχής. Στην περίπτωση της τυπικής κανονικής κατανομής η καμπύλη είναι συμμετρική οπότε το 5% ισομοιράζεται στις δύο ακραίες περιοχές με ποσοστό 2.5% η κάθε μια. Στο επόμενο σχήμα με C συμβολίζεται το επίπεδο εμπιστοσύνης.

Η τιμή  $z^*$  παριστά αυτό το σημείο της καμπύλης κανονικής κατανομής για το οποίο η πιθανότητα μιας παράστασης με τιμή μεγαλύτερη από  $z^*$  είναι ίση με  $p$ , γνωστή ως p-κρίσιμη τιμή της τυπικής κανονικής κατανομής. Για παράδειγμα, αν  $p=0.025$ , τότε η τιμή  $z^*$  ώστε να ισχύει  $P(Z>z^*)=0.975$  είναι 1.96. Για ένα c.i με επίπεδο εμπιστοσύνης C, η τιμή  $p$  ισούται με  $(1-C)/2$ . Επομένως ένα 95% confidence interval για την τυπική κανονική κατανομή είναι το διάστημα (-1.96,1.96), αφού το 95% της περιοχής κάτω από την καμπύλη ανήκει σε αυτό το διάστημα.

Για ένα πληθυσμό με άγνωστη μέση τιμή  $m$  και γνωστή τυπική απόκλιση  $\sigma$ , ένα c.i για τη μέση τιμή ενός τυχαίου δείγματος μεγέθους  $n$  του πληθυσμού υπολογίζεται από τη σχέση :

$$c.i =$$

Η παραπάνω σχέση ισχύει μόνο όταν ο πληθυσμός ακολουθεί τη κανονική κατανομή. Για μεγάλα δείγματα από πληθυσμούς με διαφορετική κατανομή αυτή η τιμή διορθώνεται με βάση το **Κεντρικό Οριακό Θεώρημα**.



Εικόνα 1 Απεικόνιση του επιπέδου εμπιστοσύνης  $C$  και της τιμής  $z^*$  στην περίπτωση της τυπικής κανονικής κατανομής.

Ένας τυπικός ορισμός του Κ.Ο.Θ είναι ο παρακάτω:

<<Εάν  $\bar{x}$  είναι η μέση τιμή ενός τυχαίου δείγματος  $x_1, x_2, \dots, x_n$  μεγέθους που ακολουθεί μια κατανομή με πεπερασμένη μέση τιμή  $\mu$  και πεπερασμένη διακύμανση  $\sigma^2$ , τότε η κατανομή  $W$  είναι  $N(0,1)$  καθώς το  $n$  πλησιάζει το άπειρο>>.

Αυτό σημαίνει ότι η μεταβλητή  $\bar{x}$  ακολουθεί την κατανομή  $N(\mu, \sigma^2/n)$ .

Μια ευρέως γνωστή εφαρμογή του Κ.Ο.Θ είναι η προσέγγιση της διωνυμικής κατανομής με την κανονική. Στην περίπτωση αυτή η σχέση που χρησιμοποιείται είναι η εξής :

$$c.i = \hat{g} \pm (z_{1-\alpha/2}^* \cdot \sigma_{\hat{g}})$$

όπου,

- $\hat{g}$ , είναι το ποσοστό επιτυχιών σε μια διαδικασία δοκιμών Bernoulli όπως υπολογίζεται από το στατιστικό δείγμα.

- $Z_{1-\alpha/2}$ , είναι το  $1-\alpha/2$  εκατοστημόριο μιας τυπικής κανονικής κατανομής.
- $\alpha$ , είναι το επίπεδο εμπιστοσύνης.
- $n$ , το μέγεθος του δείγματος.

Επειδή τα παραπάνω επικεντρώνονται στην περίπτωση της τυπικής κανονικής κατανομής θα προχωρήσουμε σε ένα πιο γενικευμένο ορισμό του c.i.

Έστω  $X$  ένα τυχαίο δείγμα μιας πιθανοθεωρητικής κατανομής με  $\theta$  την παράμετρο που μας ενδιαφέρει και  $\phi$  μια παράμετρο εκτός ενδιαφέροντος. Ένα c.i για την παράμετρο  $\theta$  με επίπεδο εμπιστοσύνης  $\gamma$  είναι ένα διάστημα με τυχαία άκρα  $(u(X), v(X))$  με την ιδιότητα :

$$\Gamma = \Pr_{\theta, \phi} (u(X) < \theta < v(X)), \text{ για κάθε } (\theta, \phi).$$

Οι ποσότητες  $\phi$  καλούνται nuisance-parameters και αποτελούν μελλοντικό αντικείμενο μελέτης. Ο αριθμός  $\gamma$  λαμβάνει τιμές κοντά στο 1, όπως ήδη αναφέρθηκε. Η  $\Pr_{\theta, \phi}$  είναι η συνάρτηση πιθανότητας της τυχαίας μεταβλητής  $X$  που χαρακτηρίζεται από τις παραμέτρους  $(\theta, \phi)$ .

Ο αντίστοιχος ορισμός για τα Bayesian intervals καλούμενα credible intervals είναι ο ακόλουθος :

$$\Pr(u(x) < \theta < v(x), | X=x) = 1-\alpha$$

Στην περίπτωση αυτή η άγνωστη τιμή της παραμέτρου  $\theta$  λαμβάνεται ως τυχαία μεταβλητή  $\Theta$ . Συγκρίνοντας τους 2 παραπάνω ορισμούς μπορούμε να πούμε ότι :

1. Ο ορισμός ενός confidence interval εμπεριέχει πιθανότητες υπολογισμένες από την κατανομή της μεταβλητής  $X$  για δεδομένα  $(\theta, \phi)$  και τη συνθήκη που πρέπει να ισχύει για όλα τα ζεύγη  $(\theta, \phi)$ .
2. Ο ορισμός ενός credible interval εμπεριέχει πιθανότητες υπολογισμένες από την κατανομή της μεταβλητής  $\Theta$  για τις τιμές  $X=x$  των παρατηρήσεων και περιορισμένες από τις τιμές της μεταβλητής  $\Phi$ , η οποία είναι η τυχαία μεταβλητή που αντιπροσωπεύει την αβεβαιότητα γύρω από τις τιμές της nuisance-παραμέτρου  $\phi$ .

Σε κάθε περίπτωση, ο υπολογισμός ενός confidence interval οφείλει να διαθέτει τις τρεις παρακάτω ιδιότητες :

- Εγκυρότητα (validity): πρέπει δηλαδή το επίπεδο εμπιστοσύνης που έχει οριστεί πρέπει να απαντάται ακριβώς ή το πολύ κοντά στις επικείμενες μετρήσεις-υπολογισμούς επί του δείγματος.
- Βέλτιστη απόδοση (optimality) : πρέπει δηλαδή να γίνεται χρήση του μέγιστου δυνατού αριθμού δεδομένων από αυτά που διαθέτει ο μελετητής. Ένα μέσο κρίσης του κατά πόσο βέλτιστο είναι ένα c.i αποτελεί το μήκος του. Γενικά, c.i μικρότερου μήκους θεωρούνται πιο αποδοτικά.
- Μη-διακύμανση (invariance) : υπάρχουν αρκετές περιπτώσεις εφαρμογών που η εκτιμώμενη ποσότητα δεν είναι σαφώς ορισμένη. Για παράδειγμα, μια έρευνα που μελετά το μέσο εισόδημα ενός πληθυσμού, είναι πιθανό να εκτιμά το λογάριθμο του μέσου εισοδήματος, διότι είναι ένας συχνός τρόπος παρουσίασης αριθμητικών αποτελεσμάτων. Θα ήταν επομένως επιθυμητό η μέθοδος κατασκευής των c.i να δίνει ισοδύναμα αποτελέσματα και για τους 2 τρόπους εκτίμησης του μέσου εισοδήματος. Πιο συγκεκριμένα τα άκρα του διαστήματος για εκτίμηση του

λογαρίθμου του μέσου εισοδήματος πρέπει να είναι οι λογάριθμοι των τιμών των άκρων του διαστήματος για απλή εκτίμηση του μέσου διαστήματος.

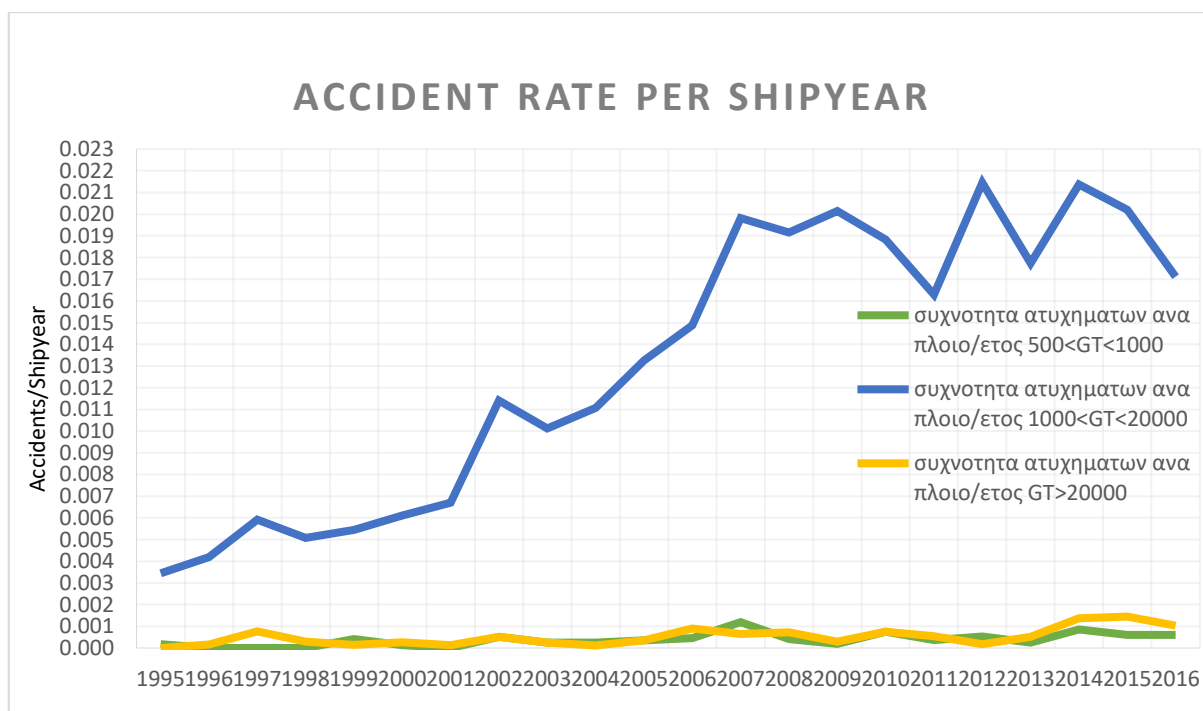
Από τις παραπάνω ιδιότητες, η πιο σημαντική είναι η validity και την ακολουθεί η optimality. Η invariance είναι μια ιδιότητα που χαρακτηρίζει περισσότερο την παραγωγή των c.i και λιγότερο την κατασκευή τους.

Για την παρούσα ανάλυση λοιπόν, λόγω της φύσης των ιστορικών δεδομένων που χρησιμοποιήθηκαν, η διωνυμική ανάλυση εμπιστοσύνης (binomial confidence analysis) κρίθηκε προτιμητέα για να οριστούν τα περιθώρια αβεβαιότητας επί των υπολογισμένων μέσων τιμών των διαφόρων μεγεθών. Επίσης, έγινε αντιληπτό πως σε μερικές περιπτώσεις τα confidence intervals έχουν μεγάλο μέγεθος σε σχέση με τους αντίστοιχους μέσους όρους και σε αυτές τις περιπτώσεις πρέπει να δίνεται βάση όταν χρησιμοποιούνται οι τιμές αυτές για θέματα λήψεων αποφάσεων decision making purposes).

#### 4.2 Συχνότητες Ατυχημάτων ανά κατηγορία Ατυχήματος

Σε αυτή την ενότητα εξετάζεται κατά πόσον ένας τύπος ατυχήματος από αυτούς που ορίστηκαν νωρίτερα εμφανίζεται πιο συχνά ή πιο σπάνια για την περίοδο 01/01/1995-31/12/2016.

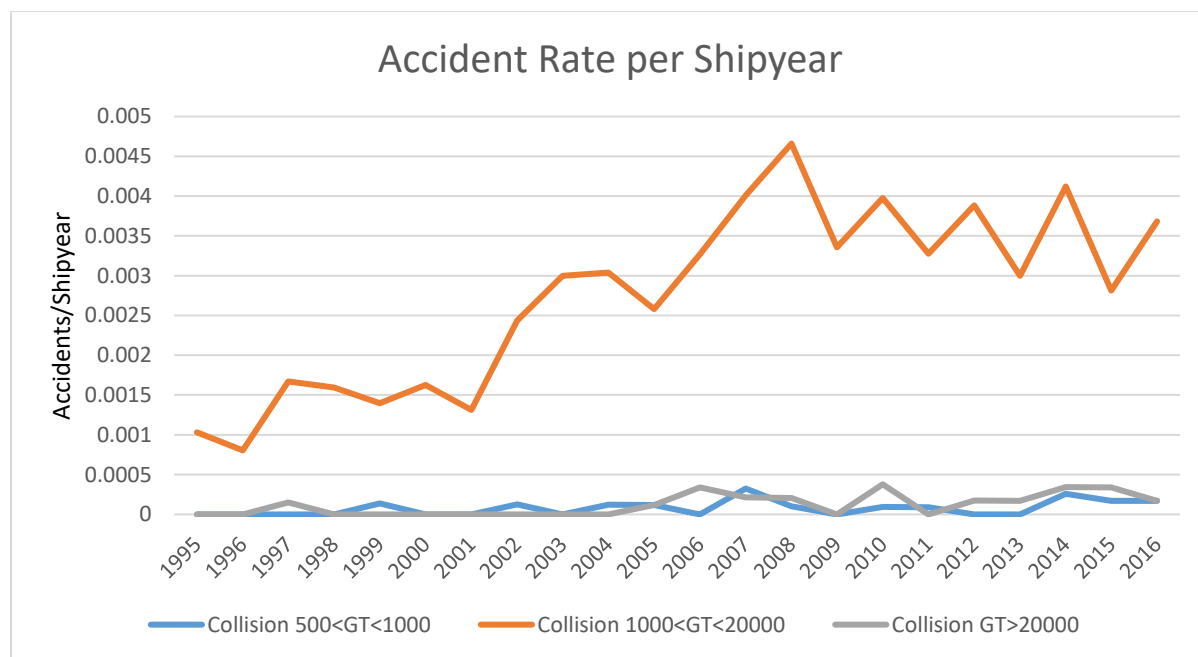
Αρχικά παρουσιάζεται μια συγκεντρωτική εικόνα των μεταβολών των ετήσιων συχνοτήτων για όλες τις κατηγορίες μεγέθους (Σχήμα 21), με την κατηγορία  $1000 \leq GT \leq 2000$  να υπερτερεί ξεκάθαρα σε σχέση με τις άλλες δύο σε συχνότητα ατυχημάτων.



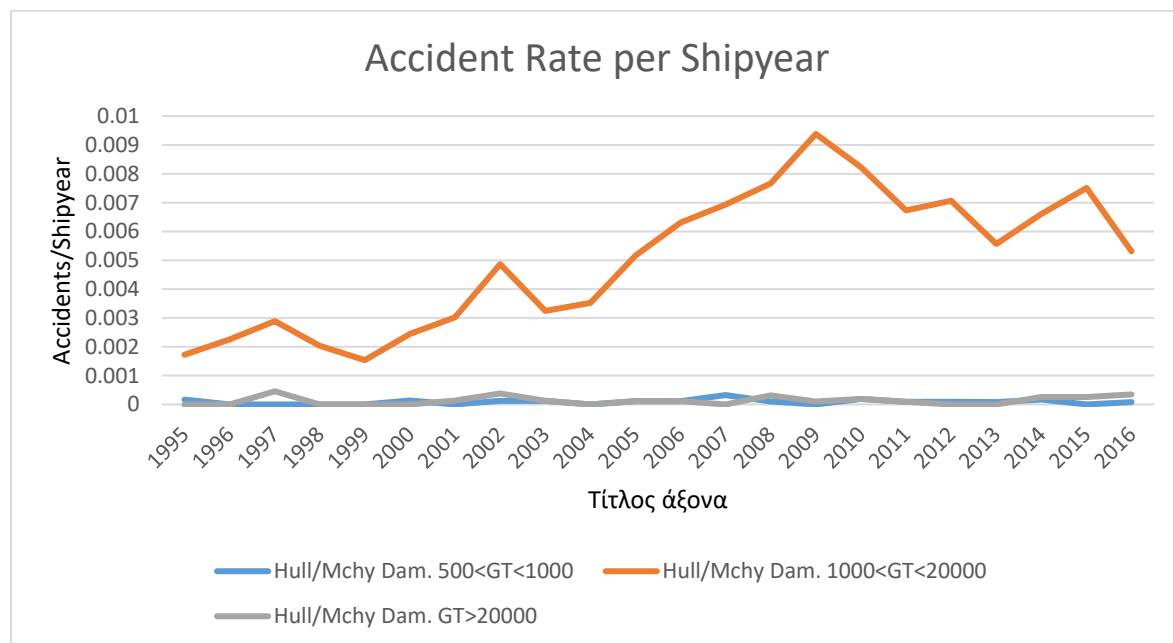
Σχήμα 21 Ετήσια μεταβολή συχνοτήτων ατυχημάτων για όλες τις κατηγορίες μεγέθους την περίοδο 1995-2016

Στη συνέχεια παρουσιάζονται 6 διαγράμματα το καθένα από τα οποία αφορά μια κατηγορία ατυχήματος. Σε κάθε διάγραμμα υπάρχουν τρεις καμπύλες, μια για κάθε κατηγορία μεγέθους, ενώ οι τιμές αποτελούνται από τον αριθμό των ατυχημάτων της κάθε κατηγορίας μεγέθους διαιρεμένη με το συνολικό στόλο ανά έτος. Οι συχνότητες ατυχημάτων για την μεσαία

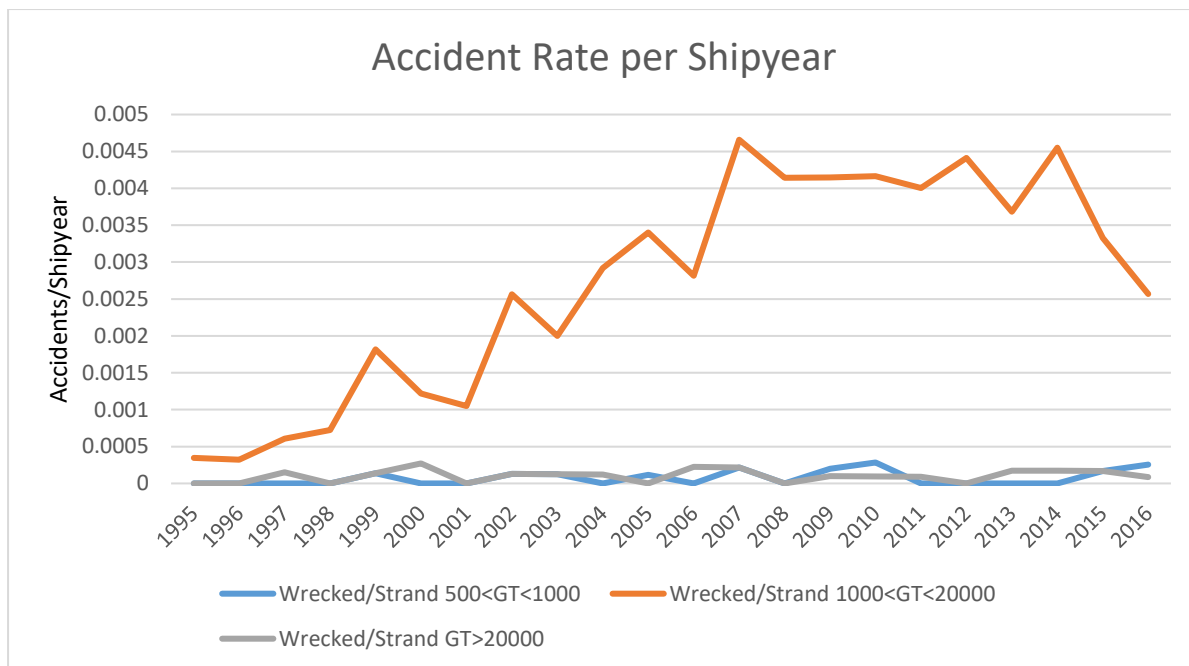
κατηγορία μεγέθους είναι μεγάλες διότι περιέχουν τα περισσότερα ατυχήματα. Οι άλλες κατηγορίες μεγέθους δείχνουν μικρές συχνότητες διότι τα ατυχήματα που περιέχουν είναι ελάχιστα σε σχέση με το συνολικό στόλο.



Σχήμα 22 Ετήσια Accident Rates για ατυχήματα Collision, για την περίοδο 1995-2016.

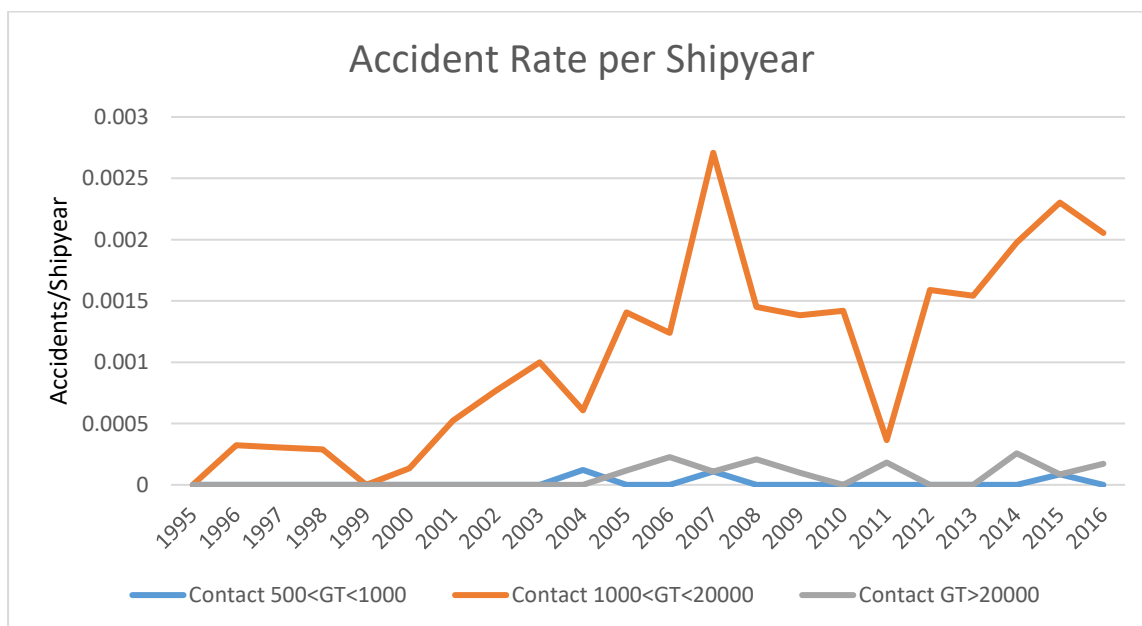


Σχήμα 23 Ετήσια Accident Rates για ατυχήματα Hull/Mchy Dam, για την περίοδο 1995-2016.



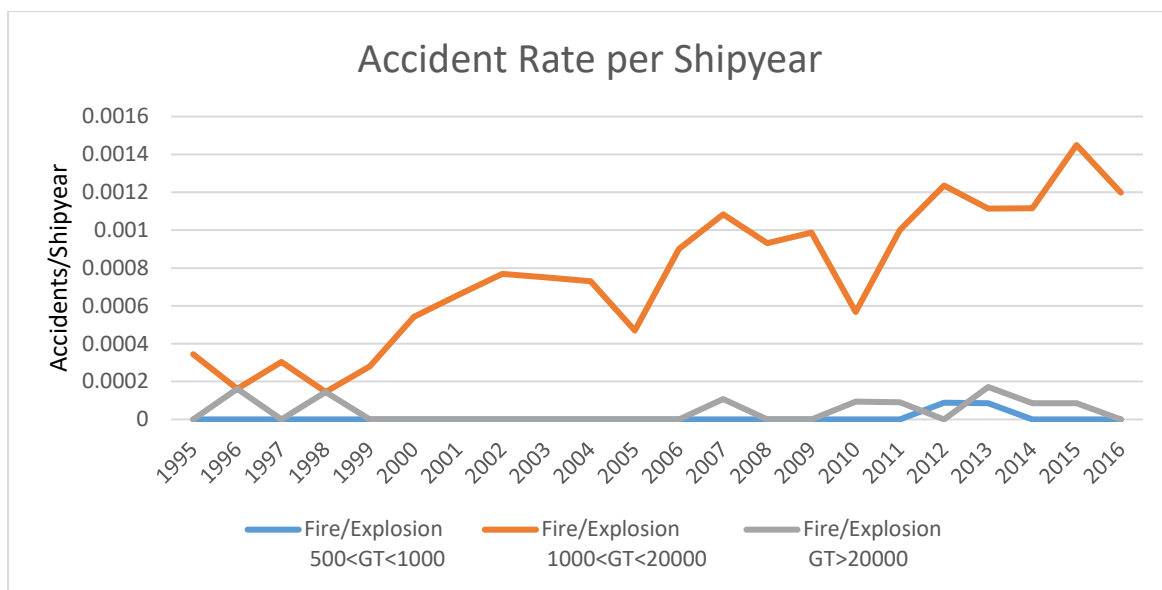
Σχήμα 24 Ετήσια Accident Rates για ατυχήματα Wrecked/Strand, για την περίοδο 1995-2016.

Οι τρεις παραπάνω κατηγορίες ατυχημάτων δείχνουν μια ανοδική πορεία και υψηλές συχνότητες διότι τα ατυχήματα στις κατηγορίες Collision, Hull/Mchy Dam, και Wrecked/Strand είναι αυτά που παρατηρούνται πιο συχνά.

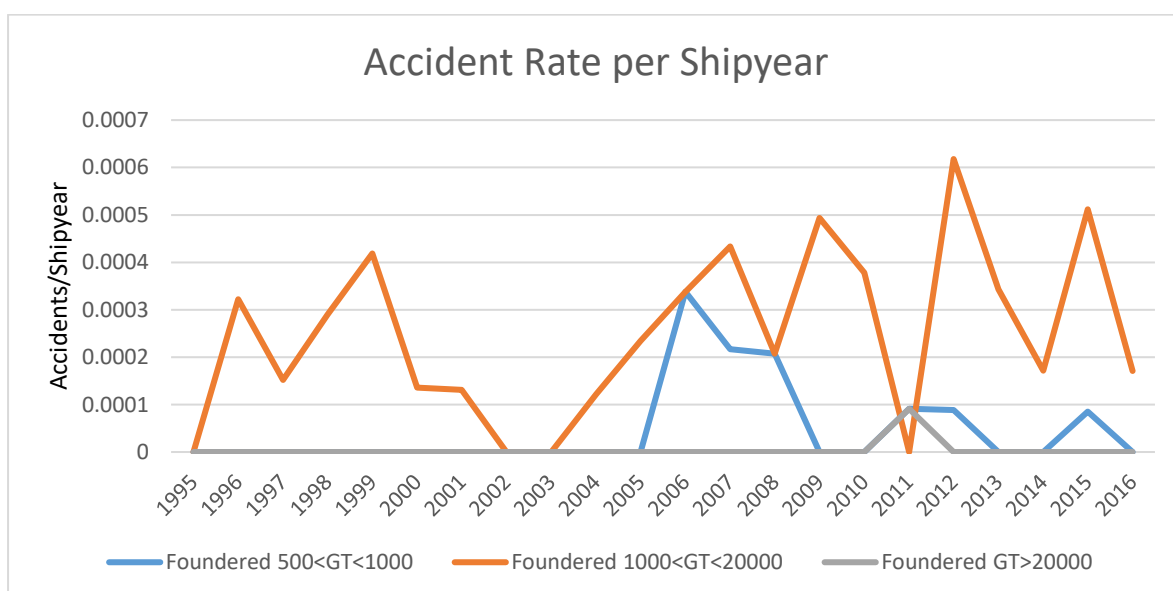


Σχήμα 25 Ετήσια Accident Rates για ατυχήματα Contact, για την περίοδο 1995-2016.



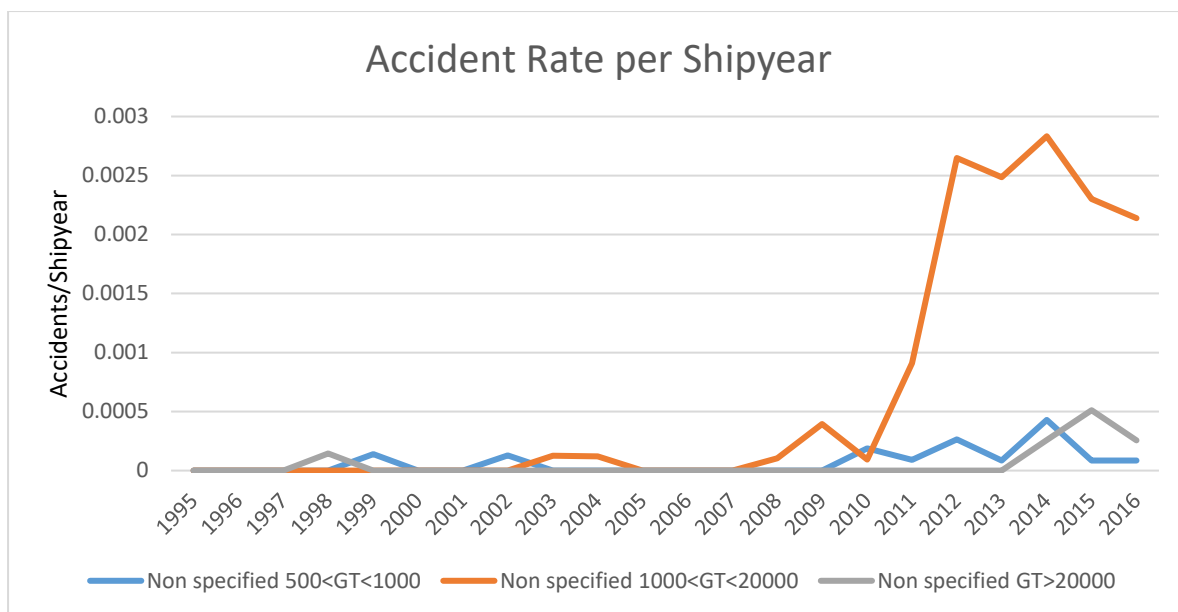


Σχήμα 26 Ετήσια Accident Rates για ατυχήματα Fire/Explosion, για την περίοδο 1995-2016.



Σχήμα 27 Ετήσια Accident Rates για ατυχήματα Foundered, για την περίοδο 1995-2016.

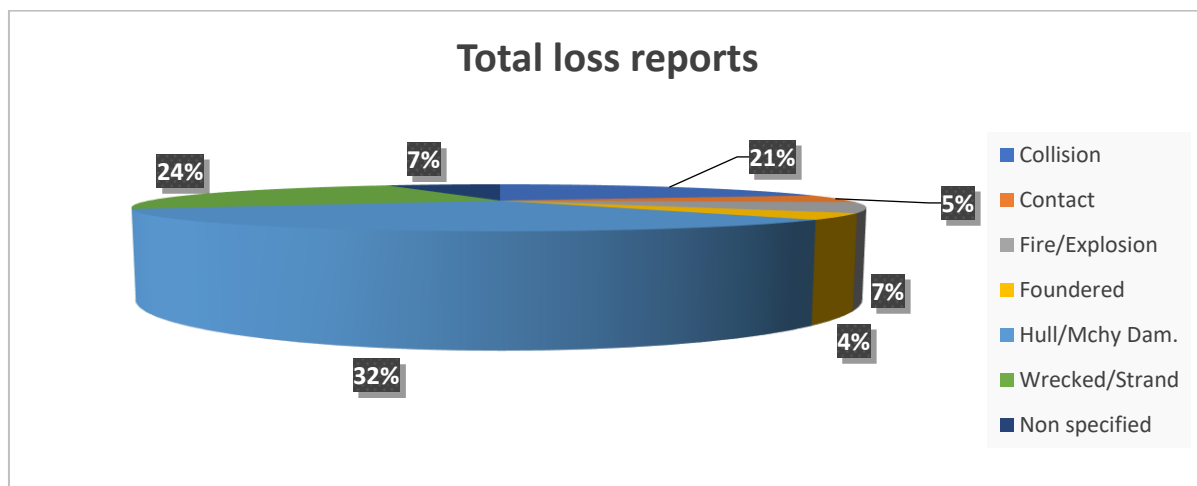
Το επόμενο διάγραμμα (**Σχήμα 28**) είναι αντίστοιχο με τα παραπάνω, άλλα περιέχει τα ατυχήματα που στη βάση δεδομένων δε προσδιορίζονται το είδος τους. Αυξάνεται ο αριθμός τους στην περίοδο 2011-2016 και αυτό οφείλεται στην αυξημένη και πιο απαραίτητη καταγραφή ατυχημάτων που παρατηρείται γενικά στη βάση από το 2010 και μετά.



Σχήμα 28 Ετήσια Accident Rates για ατυχήματα που δεν προσδιορίζεται το είδος τους, για την περίοδο 1995-2016.

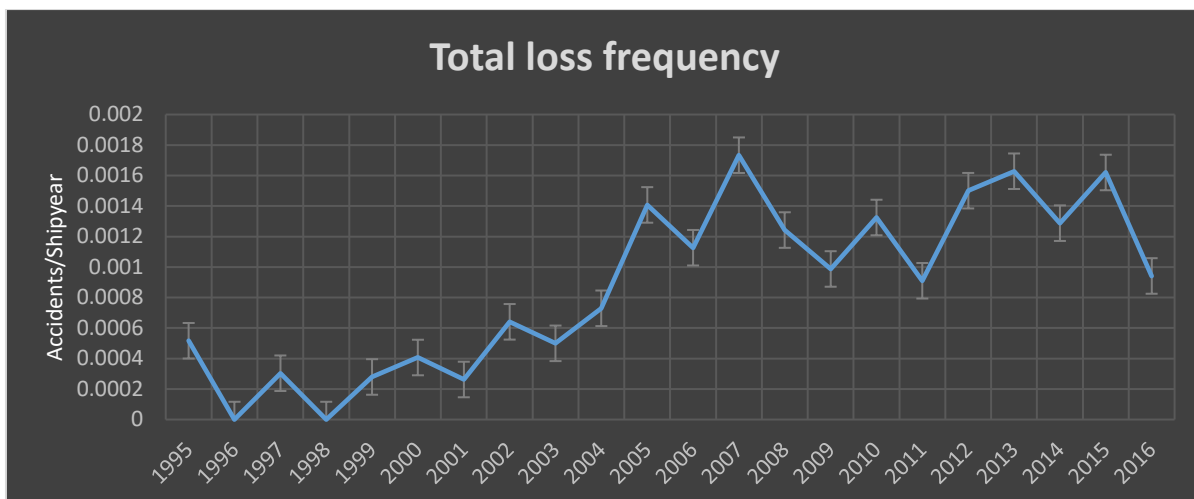
#### 4.3 Ολικές Απώλειες ανά Πλοία×Ετη

Τα ατυχήματα που κατέληξαν σε Ολική Απώλεια πλοίου την περίοδο 1995-2016 είναι 192 στον αριθμό. Στο επόμενο σχήμα (**Σχήμα 29**) φαίνεται ποσοστιαία ποια ατυχήματα ξεκίνησαν τη ζημιά από την οποία τα πλοία κατέληξαν σε Total Loss.

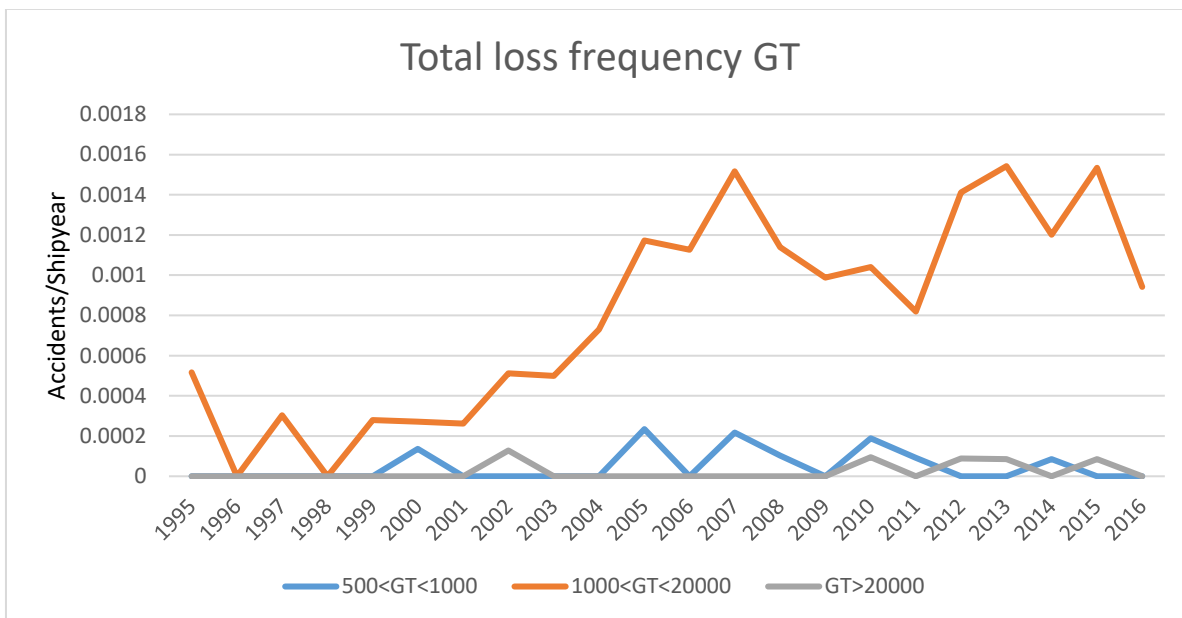


Σχήμα 29 Κατανομή των ατυχημάτων που οδήγησαν σε ολική απώλεια του πλοίου, για την περίοδο 1995-2016.

Το μεγαλύτερο ποσοστό κατέχει η κατηγορία Hull/Mchy Dam. με ποσοστό 32%, κάτι που υποδεικνύει τη σοβαρότητα που μπορεί να προκαλέσει αυτό το είδος, ενώ σημαντικό ποσοστό κατέχουν τα ατυχήματα Wrecked/Strand 24% και Collision 21%. Τα διαγράμματα που ακολουθούν παρουσιάζουν τις συχνότητες ατυχήματος με Ολική Απώλεια πλοίου για κάθε χρονιά συνολικά αλλά και χωρισμένο με βάση το GT.



Σχήμα 30 Ετήσια συχνότητα ατυχημάτων που οδήγησαν σε Ολική Απώλεια πλοίου για την περίοδο 1995-2016. Συμπεριλαμβάνονται και τα c.i 95%.



Σχήμα 31 Ετήσια συχνότητα ατυχημάτων που οδήγησαν σε Ολική Απώλεια πλοίου, για τις τρεις κατηγορίες μεγέθους, για την περίοδο 1995-2016.

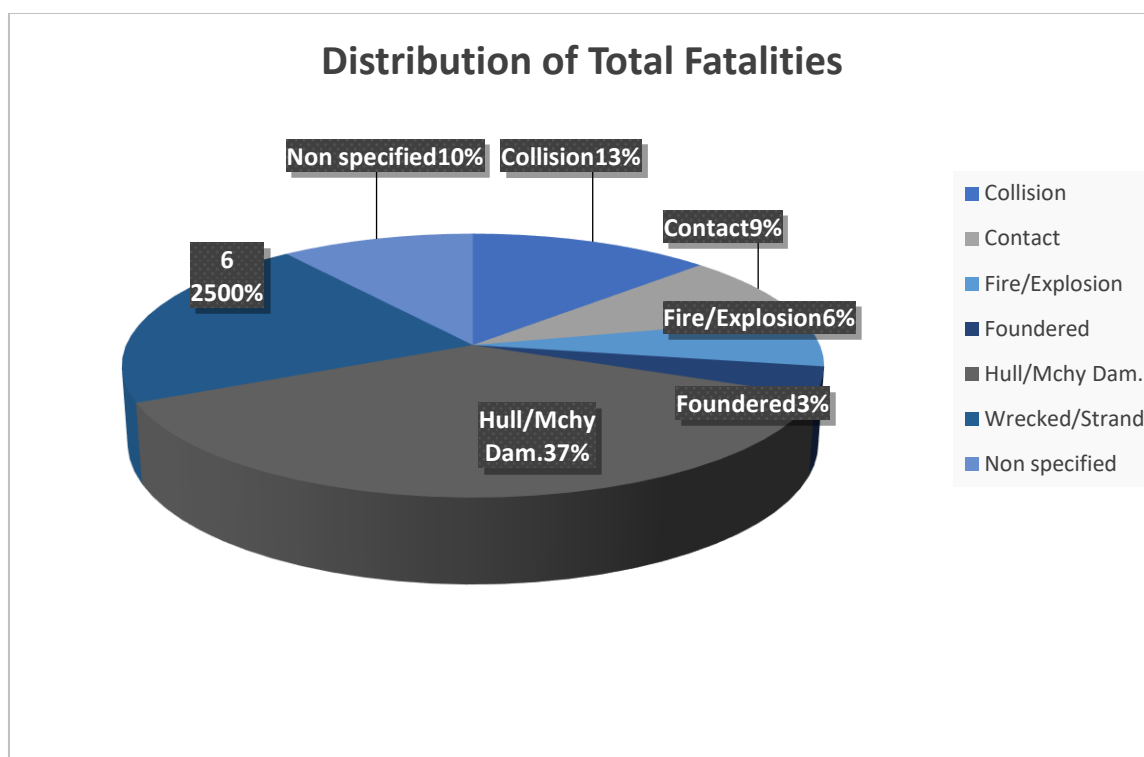
Από το σύνολο των 192 Total Loss, στην μεσαία κατηγορία μεγέθους ανήκουν τα 177 και για αυτό το λόγο διαιρεμένα με το σύνολο των shipyears παρουσιάζουν πολύ μεγαλύτερη συχνότητα από τις άλλες δύο κατηγορίες μεγέθους.

#### 4.4 Ανθρώπινες Απώλειες ανά Πλοία×Ετη

Συνδυάζοντας την κατηγορία “killed” με την κατηγορία “missing” προκύπτει ο αριθμός ανθρώπινων απωλειών ή αλλιώς “fatalities”. Όλα τα αποτελέσματα προέκυψαν από τη βάση δεδομένων Sea-Web και τα διαγράμματα που ακολουθούν βασίζονται σε αυτά τα δεδομένα.

Για την περίοδο 01/01/1995 έως 31/12/2016 ο απολογισμός ήταν 236 νεκροί και 321 αγνοούμενοι. Άρα ως fatalities έχουμε το άθροισμα αυτών των ατόμων, δηλαδή 557 άτομα. Ο αριθμός των ατυχημάτων για την κατηγορία “killed” φτάνει τα 63 ενώ τα ατυχήματα για την κατηγορία “missing” φτάνουν τα 54. Όμως 34 από τα “missing” ήταν ατυχήματα που περιείχαν και “killed” άρα ο αριθμός ατυχημάτων “fatalities” είναι  $63+54-34=83$ .

Στο ακόλουθο διάγραμμα (Σχήμα 32) παριστάνεται η κατανομή των fatalities στους διάφορους τύπους ατυχημάτων. Το μεγαλύτερο ποσοστό καταλαμβάνει με διαφορά η κατηγορία Hull/Mchy Dam. με ποσοστό 37%, όπως ακριβώς και στην προηγούμενη ενότητα με total losses. Ακολουθούν με ποσοστά 22% και 13% τα ατυχήματα Wrecked/Strand και Collision αντίστοιχα, ενώ υπάρχει σημαντικό ποσοστό ατυχημάτων που οδήγησαν σε fatalities τα οποία δεν έχει γίνει ξεκάθαρο το που βασιζόταν η ζημιά (10%).



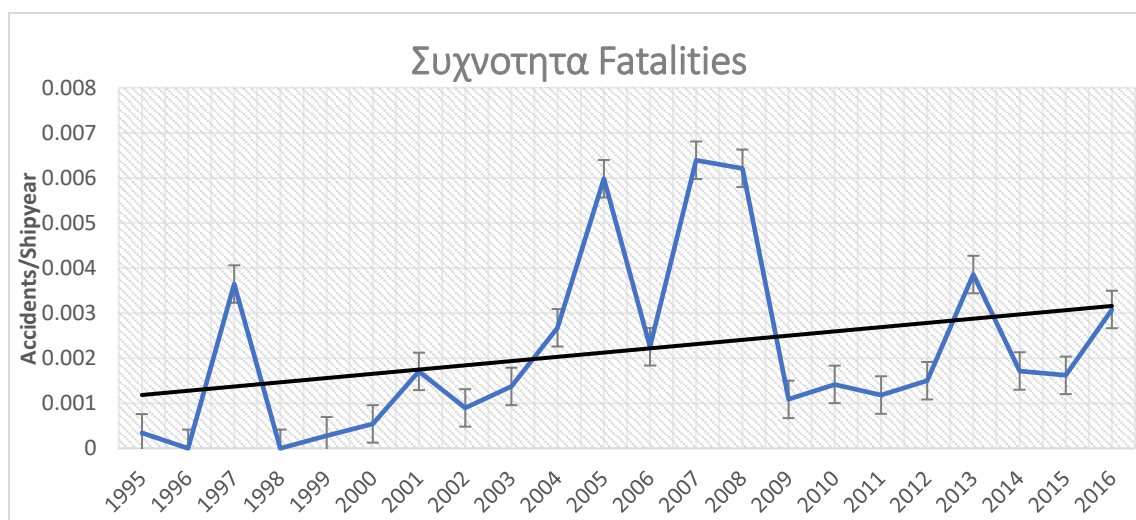
Σχήμα 32 Κατανομή των ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες, για την περίοδο 1995-2016.

Στον πίνακα που ακολουθεί (Πίνακας 8) παρουσιάζεται ο αριθμός των ατυχημάτων κάθε τύπου που είχαν ανθρώπινες απώλειες και επιβεβαιώνεται ότι οι κατηγορίες ατυχημάτων Hull/Mchy Dam. και Wrecked/Strand κατέχουν τα υψηλότερα νούμερα και προκαλούν κατά μέσο όρο τις περισσότερες ανθρώπινες απώλειες.

Πίνακας 8 Αριθμός ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες και μέσος όρος απωλειών/ατυχημάτων, για την περίοδο 1995-2016.

Κατηγορία Ατυχήματος	Αριθμός Ατυχημάτων	Αριθμός ατυχημάτων που είχαν Fatalities	Αριθμός Fatalities	Fatalities/ Ατύχημα
Collision	711	15	36	0.051
Contact	258	10	35	0.136
Fire/Explosion	182	7	18	0.099
Foundered	99	4	26	0.263
Hull/Mchy Dam.	1162	43	160	0.138
Wrecked/Strand	651	25	156	0.239
Non specified	278	12	20	0.072

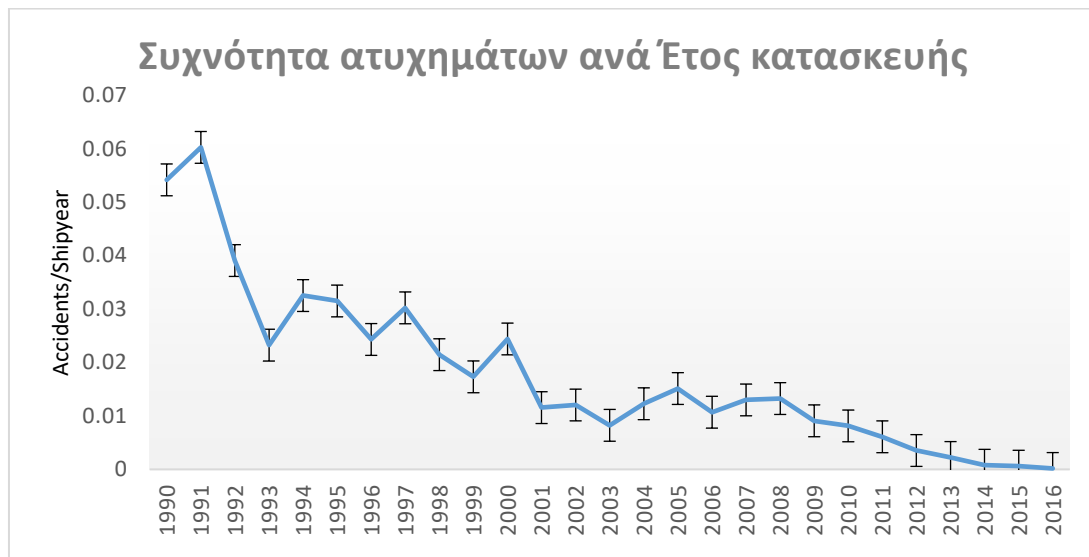
Στη συνέχεια παρατίθεται διάγραμμα ετήσιας μεταβολής των fatalities για το σύνολο των πλοίων της παρούσας μελέτης στο οποίο συνυπάρχει η γραμμή τάσης της καμπύλης (**Σχήμα 33**). Πέραν της διετίας 2007-2008, οι συχνότητες δεν απέχουν πολύ μεταξύ τους. Τη χρονιά 2007 σημειώθηκε ατύχημα Wrecked/Strand του πλοίου Kocatere S κατασκευής 1990 με απολογισμό 15 killed. Το πλοίο είχε σημαία Τουρκίας ενώ ανήκει στην κατηγορία total loss.



Σχήμα 33 Ετήσιες συχνότητες ατυχημάτων που οδήγησαν σε ανθρώπινες απώλειες, για την περίοδο 1995-2016. Μαζί παριστάνεται και η γραμμή τάσης. Συμπεριλαμβάνονται και τα c.i 95%.

#### 4.5 «Σοβαρά Ατυχήματα» ανά Έτος Κατασκευής των πλοίων

Μια ακόμα παράμετρος που αξίζει να ερευνηθεί αποτελεί είναι το έτος κατασκευής-κτίσης του πλοίου. Η συσχέτιση αυτής της παραμέτρου με τη συχνότητα των ατυχημάτων για κάθε έτος απεικονίζεται στο επόμενο διάγραμμα (**Σχήμα 34**) μαζί με τα 95% confidence intervals.



Σχήμα 34 Αριθμός ατυχημάτων ανά πλοία®έτη για κάθε έτος κατασκευής, εντός της περιόδου 1990-2016.

Το συμπέρασμα που βγαίνει από το παραπάνω διάγραμμα είναι πως τα πλοία που κατασκευάστηκαν μεταξύ 1990-1992 έχουν μεγαλύτερη συχνότητα ατυχημάτων, όμως οι πληροφορίες που έχουμε για τα ατυχήματα της περιόδου 1990-1994 είναι ελάχιστα διότι η παρούσα διπλωματική αναλώνεται κυρίως στην περίοδο 1995-2016.

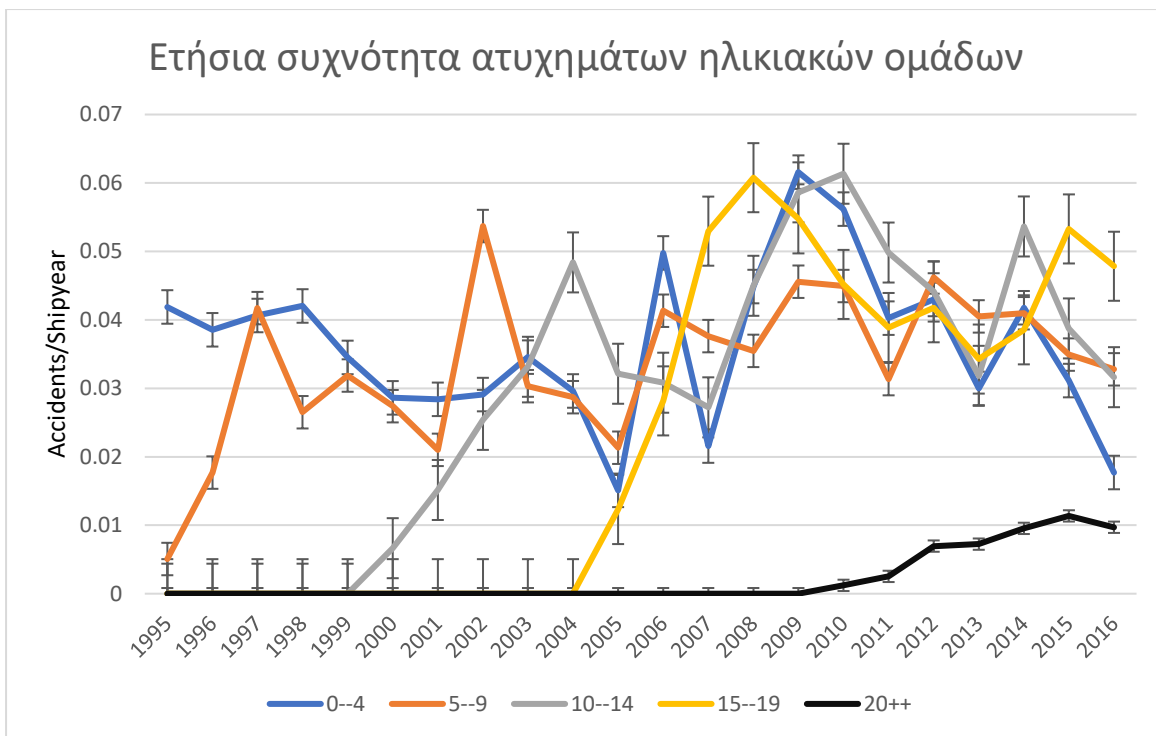
#### 4.6 Ηλικία των πλοίων που είχαν ‘Σοβαρά’ ατυχήματα την περίοδο 1995-2016.

Ένας ακόμα σημαντικός παράγοντας για την ανάλυση των ναυτικών ατυχημάτων είναι η ηλικία που είχε το κάθε πλοίο τη στιγμή του ατυχήματος και η συγκεντρωτική εικόνα αυτού του παράγοντα δίδεται στον ακόλουθο πίνακα χωρισμένη ανάλογα με το μέγεθος του πλοίου με βάση τις τρεις κατηγορίες μεγέθους.

Πίνακας 9 Κατανομή ατυχημάτων σε ηλικιακές ομάδες πενταετίας για κάθε κατηγορία μεγέθους, για την περίοδο 1995-2016

Age	500<GT<1000	1000<GT<20000	GT>20000
0--4	19.00	767.00	40.00
5--9	20.00	772.00	24.00
10--14	19.00	572.00	22.00
15--19	18.00	509.00	18.00
20++	8.00	307.00	7.00

Στο επόμενο διάγραμμα παρουσιάζεται η σχέση μεταξύ της συχνότητας ατυχημάτων ανά πλοία®έτη και της ηλικίας του πλοίου τη στιγμή του ατυχήματος σε ηλικιακές ομάδες πενταετίας (**Σχήμα 35**) μαζί με τα 95% confidence intervals.



Σχήμα 35 Συχνότητα ατυχημάτων ανά πλοία\*Έτη ανά ηλικιακή ομάδα τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.

Ενδιαφέρον στοιχείο αποτελεί η υψηλή συχνότητα ατυχημάτων στα νέα πλοία που ανήκουν δηλαδή στην ηλικιακή ομάδα 0-4. Επίσης, αποτελέσματα της ηλικιακής ομάδας 20+ εμφανίζονται μετά το 2010 διότι δεν υπάρχουν πληροφορίες στη συγκεκριμένη διπλωματική για πλοία που κτίστηκαν πριν το 1990.



Σχήμα 36 Συχνότητα ατυχημάτων ανά πλοία\*Έτη για την ηλικιακή ομάδα 0-4 τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.





Σχήμα 37 Συχνότητα ατυχημάτων ανά πλοία®έτη για την ηλικιακή ομάδα 5-9 τη στιγμή του ατυχήματος, εντός της περιόδου 1990-2016.

Η γραμμή τάσης για τη συχνότητα ατυχημάτων της ηλικιακής ομάδας 0-4 (Σχήμα 36) φαίνεται σταθερή. Την τελευταία όμως πενταετία δείχνουν ότι τα νέα πλοία μειώνουν αρκετά τα ατυχήματά τους. Λίγο μικρότερη συχνότητα ατυχημάτων φαίνεται να έχει η ηλικιακή ομάδα 5-9 (Σχήμα 37). Αυτές είναι οι ηλικίες που είχαν τα πλοία όταν συνέβη το ατύχημα.

#### 4.7 Στατιστικά Τεστ

##### • 4.7.1 Ανάλυση Διακύμανσης (ANOVA)

Η ανάλυση της διακύμανσης (Analysis Of Variance – ANOVA) είναι μια στατιστική μέθοδος με την οποία η μεταβλητότητα που υπάρχει σε ένα σύνολο δεδομένων διασπάται στις επιμέρους συνιστώσες της με στόχο την κατανόηση της σημαντικότητας των διαφορετικών πηγών προέλευσής της. Η ανάπτυξη της μεθοδολογίας οφείλεται στον θεμελιωτή της σύγχρονης στατιστικής επιστήμης, Άγγλο στατιστικό Sir Ronald Aylmer Fisher (1890-1962). Στην πραγματικότητα η ANOVA περιλαμβάνει μια ομάδα στατιστικών μεθόδων κατάλληλων για την ανάλυση δεδομένων που προκύπτουν από πειραματικούς σχεδιασμούς.

Τα δεδομένα ενός δείγματος ανάλογα με την προέλευσή τους διακρίνονται σε **παρατηρήσεις (observational sampling)** ή σε **πειραματικά (designed sampling)**. Στην πρώτη κατηγορία ο στατιστικός ερευνητής απλά παρατηρεί τις τιμές που εμφανίζονται χωρίς να έχει δυνατότητα επέμβασης στις αντίστοιχες μεταβλητές. Αντίθετα, στη δεύτερη κατηγορία ο στατιστικός ερευνητής προσπαθεί να ελέγξει τα επίπεδα μιας ή περισσότερων **ανεξάρτητων (independent)** μεταβλητών προκειμένου να προσδιορίσει την επίδραση που έχουν πάνω στην υπό μελέτη μεταβλητή που καλείται **εξαρτημένη (dependent)** η **απόκριση (response)**. Για παράδειγμα, απόκριση μπορεί να είναι η βαθμολογία στην εξέταση του μαθήματος της στατιστικής, ο όγκος των πωλήσεων μιας επιχείρησης ή το συνολικό εισόδημα μιας οικογένειας κατά τη διάρκεια του έτους.

Στόχος κάθε στατιστικού πειράματος είναι ο προσδιορισμός της επίδρασης μιας ή περισσότερων ανεξάρτητων μεταβλητών πάνω στην απόκριση. Οι μεταβλητές αυτές αναφέρονται συνήθως σαν **παράγοντες (factors)** και μπορεί να είναι είτε ποσοτικές είτε ποιοτικές. Για παράδειγμα θα ήταν ενδιαφέρον να διερευνήσουμε την επίδραση που έχει ο ποιοτικός παράγοντας «φύλο» στη βαθμολογία της στατιστικής ή ο ποσοτικός παράγοντας «πλήθος καταστημάτων πώλησης» στον όγκο των πωλήσεων. Σε άλλες πάλι περιπτώσεις ενδεχομένως να ενδιαφερόμαστε για την επίδραση που έχουν πάνω στην απόκριση περισσότερες της μιας ανεξάρτητες μεταβλητές όπως ο ποσοτικός παράγοντας «πλήθος εργαζομένων» και ο ποιοτικός παράγοντας «πόλη διαμονής» πάνω στο οικογενειακό εισόδημα.

Οι τιμές του παράγοντα που προσδιορίζονται στο πείραμα λέγονται **επίπεδα (levels)**. Για παράδειγμα τα επίπεδα για τον ποιοτικό παράγοντα φύλο είναι «αρσενικό» - «θηλυκό» ενώ για τον ποσοτικό παράγοντα πλήθος καταστημάτων πώλησης είναι «θετικός ακέραιος». Σε ένα πείραμα με έναν παράγοντα οι **μεταχειρίσεις (treatments)** του πειράματος είναι τα επίπεδα του παράγοντα. Για παράδειγμα αν στο πείραμα βαθμολογίας της στατιστικής μας ενδιαφέρει η επίδραση του παράγοντα «φύλο» τότε οι μεταχειρίσεις του πειράματος είναι «αρσενικό» - «θηλυκό». Σε ένα πείραμα με δύο ή περισσότερους παράγοντες οι μεταχειρίσεις είναι συνδυασμοί παραγόντων – επιπέδων. Για παράδειγμα αν μας ενδιαφέρει η επίδραση των παραγόντων «φύλο», «ηλικία» στη βαθμολογία της στατιστικής, τότε οι μεταχειρίσεις είναι οι συνδυασμοί των επιπέδων φύλου και ηλικίας π.χ. (αρσενικό,20), (θηλυκό,18).

Η ANOVA βασίζεται στον έλεγχο μιας στατιστικής υπόθεσης. Στους στατιστικούς ελέγχους υποθέσεων επιθυμούμε να ελέγξουμε αν μια ή περισσότερες παράμετροι ενός πληθυσμού ικανοποιούν μια βασική υπόθεση. Στην πράξη, επειδή σχεδόν πάντα δεν είμαστε σε θέση να καταγράψουμε τις τιμές των παραμέτρων για ολόκληρο τον πληθυσμό, αρκούμαστε σε ένα τυχαίο δείγμα αυτού. Σκοπός του ελέγχου είναι να επιβεβαιωθεί ή να απορριφθεί η αρχική υπόθεση. Για το σκοπό αυτό αρχικά κατασκευάζεται μια στατιστική συνάρτηση η οποία ακολουθεί μια συγκεκριμένη κατανομή όταν η αρχική υπόθεση επιβεβαιώνεται ή λαμβάνει κάποιες ακραίες τιμές όταν απορρίπτεται. Το επίπεδο εμπιστοσύνης που ο μελετητής ορίζει, καθορίζει ποιες τιμές θεωρούνται κάθε φορά ακραίες και ποιες επιτρεπτές. Πιο συγκεκριμένα :

Έστω  $X_1, X_2, \dots, X_n$  ένα τυχαίο δείγμα από έναν πληθυσμό με κατανομή  $F_\theta$ , η οποία εξαρτάται από τις παραμέτρους  $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ . Θέλουμε να ελεγχθεί η υπόθεση :

**$H_0: \theta \in \Theta_0$**  (βασική ή μηδενική υπόθεση)

**$H_1: \theta \in \Theta_1$**  (εναλλακτική υπόθεση)

Όπου  $\Theta_0, \Theta_1$  υποσύνολα του χώρου  $\Theta$  που περιλαμβάνει όλες τις επιτρεπτές τιμές των παραμέτρων  $\theta$ , τα οποία φυσικά είναι μεταξύ τους ξένα.

Εν συνεχεία, επιλέγεται στατιστική συνάρτηση  $T(X) = T(X_1, X_2, \dots, X_n)$  με τις δύο επόμενες ιδιότητες:

1. Όταν ισχύει η  $H_0$ , λαμβάνει τις τιμές μέσα σε μια περιοχή του  $R$  (σύμφωνα με μια γνωστή κατανομή  $F_T$  χωρίς άγνωστες παραμέτρους)
2. Όταν ισχύει η  $H_1$ , λαμβάνει ακραίες τιμές εκτός της περιοχής αυτής (συνήθως μεγαλύτερες).

Με βάση τη συνάρτηση αυτή, αν  $x_1, x_2, \dots, x_n$  είναι οι τιμές του δείγματος (πραγματοποίηση των  $X_1, X_2, \dots, X_n$ ) τότε συνήθως

- Αν  $T(x) > c$ , απορρίπτουμε την  **$H_0: \theta \in \Theta_0$**
- Αν  $T(x) \leq c$ , δεν απορρίπτουμε την  **$H_0: \theta \in \Theta_0$**

Για κάποιο προεπιλεγμένο  $c$ . Ανάλογα με την απόφαση που θα παρθεί ελλοχεύουν τα εξής δύο πιθανά σφάλματα :

- 1) **Σφάλμα τύπου I** , απόρριψη της  $H_0$ , ενώ ισχύει η  $H_0$ .
- 2) **Σφάλμα τύπου II** , αποδοχή της  $H_0$ , ενώ ισχύει η  $H_1$ .

Από τους δύο παραπάνω τύπους σφαλμάτων, μας ενδιαφέρει πρωτίστως να αποφευχθεί ο πρώτος. Για το λόγο αυτό επιλέγουμε  $c$  τέτοιο ώστε :

$$P(\text{σφάλμα τύπου I}) = P(T(x) > c / H_0) \leq \alpha$$

Όπου  $\alpha$  το επίπεδο εμπιστοσύνης το οποίο προεπιλέγεται με τιμή 0,05 ή 0,01 συνήθως. Δηλαδή η μηδενική υπόθεση  $H_0$  απορρίπτεται όταν :

$$T(x) > c = F_T^{-1}(1-\alpha)$$

Η παραπάνω εξίσωση περιγράφει το άνω  $\alpha$ -σημείο της κατανομής  $T(X)$  όταν ισχύει η  $H_0$ .

Πριν την ευρεία χρήση των  $H/Y$  υπήρχαν πίνακες των άνω  $\alpha$ -σημείων για συγκεκριμένες κατανομές και συγκεκριμένες τιμές του  $\alpha$ . Σήμερα ο έλεγχος γίνεται μέσω των στατιστικών πακέτων χρησιμοποιώντας το **p-value**, η έννοια του οποίου περιγράφεται αναλυτικά στη συνέχεια.

Αν η περιοχή απόρριψης της μηδενικής υπόθεσης  $H_0$  έχει τη μορφή  $T(x) > c$  , τότε το p-value των τιμών  $x$  του δείγματος ορίζεται ως η πιθανότητα

$$p\text{-value} = P(T(X) > T(x) / H_0) = 1 - F_T^{-1}(T(x))$$

η οποία εκφράζει την πιθανότητα να εμφανιστεί ένα τόσο ή ακόμη πιο ακραίο δείγμα από αυτό που εμφανίστηκε, δεδομένου ότι ισχύει η  $H_0$  (αν η  $T(X)$  θεωρηθεί συνεχής, που ισχύει συνήθως, τότε η πιθανότητα περιλαμβάνει και την ισότητα, δηλαδή είναι

$$p\text{-value} = P(T(X) \geq T(x) / H_0) = 1 - F_T^{-1}(T(x))$$

Αν η τιμή του p-value είναι κοντά στο μηδέν τότε είναι απίθανο να εμφανιστεί αυτό το δείγμα, ενώ ισχύει η  $H_0$ . Συμπεραίνουμε επομένως ότι μάλλον δεν πρέπει να ισχύει η  $H_0$ . Πράγματι αν  $p\text{-value} < \alpha$ , απορρίπτουμε την  $H_0$  διότι

$$T(X) > c = F_T^{-1}(1-\alpha) \Leftrightarrow F_{T|H_0}(T(x)) > 1-\alpha \Leftrightarrow P(T(X) > T(x) | H_0) < \alpha$$

Συνεπώς, αντί να εξετάζεται αν  $T(x) > c$  , ισοδύναμα εξετάζεται :

-Αν  $p\text{-value} < \alpha$  , απορρίπτουμε την  $H_0$

- Αν  $p\text{-value} \geq \alpha$  , δεχόμαστε την  $H_0$ .

Αν η τιμή του p-value είναι πάρα πολύ μικρή (π.χ. 0,0001), τότε απορρίπτεται η μηδενική υπόθεση  $H_0$  αμέσως, ενώ αν η τιμή του είναι σχετικά μικρή (π.χ. 0,05), τότε διατηρούνται κάποιες επιφυλάξεις.

Στα στατιστικά πακέτα αρχικά εισάγονται οι τιμές  $x$  του δείγματος, έπειτα επιλέγεται το επίπεδο εμπιστοσύνης  $\alpha$ , και μετά την εφαρμογή του τεστ ελέγχου εμφανίζεται το p-value. Αν η τιμή αυτή είναι μικρότερη του  $\alpha$ , απορρίπτεται η μηδενική υπόθεση  $H_0$ .

Το βασικό πλεονέκτημα της χρήσης του p-value είναι ότι πέρα από την απόρριψη ή την αποδοχή της  $H_0$  είναι δυνατό να διευκρινιστεί πόσο πιθανή είναι η εμφάνιση του τυχαίου δείγματος  $x$  που επιλέχθηκε, ενώ μπορούν να πραγματοποιηθούν συγκρίσεις για διάφορες τιμές του  $\alpha$ . Η χρήση  $H/Y$  είναι απαραίτητη για μεθόδους με p-value διότι δεν είναι εύκολο να υπολογιστεί ή να πινακοποιηθεί για κάθε τιμή του  $T(x)$ .

Οι προϋποθέσεις για την εφαρμογή της ανάλυσης διακύμανσης είναι οι παράμετροι των δειγμάτων να είναι ανεξάρτητοι και τα δείγματα να ακολουθούν την κανονική κατανομή. Όσον αφορά τις τυπικές αποκλίσεις, οι τιμές στη μεσαία κατηγορία είναι μια τάξη μεγέθους μεγαλύτερες. Στην περίπτωση αυτής της διπλωματικής εργασίας, ανεξάρτητες μεταβλητές είναι το μέγεθος του πλοίου και το έτος που συνέβη το ατύχημα, ενώ εξαρτημένη μεταβλητή είναι τα accident rates. Επομένως θα πραγματοποιηθεί τεστ δύο παραγόντων. Επιλέχθηκε  $\alpha=0,05$ . Ο πίνακας που ακολουθεί περιέχει τα δεδομένα προς ανάλυση.

Πίνακας 10 Συγκεντρωτικά αποτελέσματα των accident rates για κάθε κατηγορία μεγέθους, για την περίοδο 1995-2016.

Statistical tests	συχνότητα ατυχημάτων πλοιο/έτος $500 \leq GT \leq 1000$	συχνότητα ατυχημάτων πλοιο/έτος $1000 < GT \leq 20000$	συχνότητα ατυχημάτων πλοιο/έτος $GT > 20000$
1995	0.00035	0.00692	0.00000
1996	0.00000	0.00826	0.00032
1997	0.00000	0.01147	0.00147
1998	0.00000	0.00957	0.00055
1999	0.00077	0.00997	0.00026
2000	0.00024	0.01080	0.00048
2001	0.00000	0.01154	0.00023
2002	0.00086	0.01903	0.00086
2003	0.00041	0.01643	0.00041
2004	0.00039	0.01755	0.00019
2005	0.00055	0.02077	0.00055
2006	0.00070	0.02317	0.00140
2007	0.00185	0.03075	0.00101
2008	0.00064	0.02981	0.00113
2009	0.00030	0.03087	0.00045
2010	0.00122	0.03037	0.00122
2011	0.00064	0.02886	0.00097
2012	0.00097	0.03917	0.00032
2013	0.00048	0.03337	0.00097
2014	0.00161	0.04014	0.00258
2015	0.00113	0.03821	0.00274
2016	0.00113	0.03224	0.00193

Η διαδικασία υλοποιήθηκε με χρήση της ανάλυσης δεδομένων του MS EXCEL και τα αποτελέσματα που προέκυψαν παρουσιάζονται στους επόμενους πίνακες.

Πίνακας 11 Αποτελέσματα ANOVA όπως προέκυψαν από το πρόγραμμα MS EXCEL.

Ανάλυση διακύμανσης δύο παραγόντων χωρίς αλληλεπίδραση				
ΣΥΜΠΕΡΑΣΜΑ	Πλήθος	Άθροισμα	Μέσος όρος	Διακύμανση
1995	3	0.007263923	0.002421308	1.51952E-05
1996	3	0.008582327	0.002860776	2.19251E-05
1997	3	0.012937371	0.004312457	3.89333E-05
1998	3	0.01012035	0.00337345	2.89035E-05
1999	3	0.010994631	0.003664877	2.9899E-05
2000	3	0.011521843	0.003840614	3.63572E-05
2001	3	0.011762045	0.003920682	4.35059E-05
2002	3	0.020744226	0.006914742	0.000110146
2003	3	0.017237883	0.005745961	8.55585E-05
2004	3	0.018125723	0.006041908	9.92888E-05
2005	3	0.02187098	0.007290327	0.000136241
2006	3	0.025280899	0.008426966	0.000163233
2007	3	0.033607797	0.011202599	0.000286785
2008	3	0.03158234	0.010527447	0.000278917
2009	3	0.031628329	0.010542776	0.000309954
2010	3	0.0328094	0.010936467	0.000283182
2011	3	0.030469128	0.010156376	0.000262311
2012	3	0.040464291	0.013488097	0.000494951
2013	3	0.03482186	0.011607287	0.000355301
2014	3	0.044333387	0.014777796	0.000482736
2015	3	0.042076415	0.014025472	0.000439221
2016	3	0.035305497	0.011768499	0.00031455
500≤GT≤1000	22	0.014234606	0.000647028	2.63953E-07
1000<GT≤20000	22	0.499276782	0.022694399	0.000122255
GT>20000	22	0.020029256	0.000910421	5.59565E-07

ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ	Στήλη1	Στήλη2	Στήλη3	Στήλη4	Στήλη5	Στήλη6
Προέλευση διακύμανσης	SS	βαθμοί ελευθερίας	MS	F	τιμή-P	κριτήριο F
Γραμμές	0.000996	21	4.74E-05	1.253018	0.260795	1.812816807
Στήλες	0.007045	2	0.003523	93.10307	3.66E-16	3.219942293
Σφάλμα	0.001589	42	3.78E-05			
Σύνολο	0.00963	65				

Στο πρώτο μέρος των αποτελεσμάτων υπολογίζονται οι διακυμάνσεις ανά γραμμή και ανά στήλη, δηλαδή ανά κατηγορία μεγέθους για όλη την περίοδο 1995-2016 και ανά έτος για όλες τις κατηγορίες μεγέθους.

Στο δεύτερο τμήμα των αποτελεσμάτων που αποτελεί και τον πίνακα ANOVA, εμφανίζονται τα εξής μεγέθη :

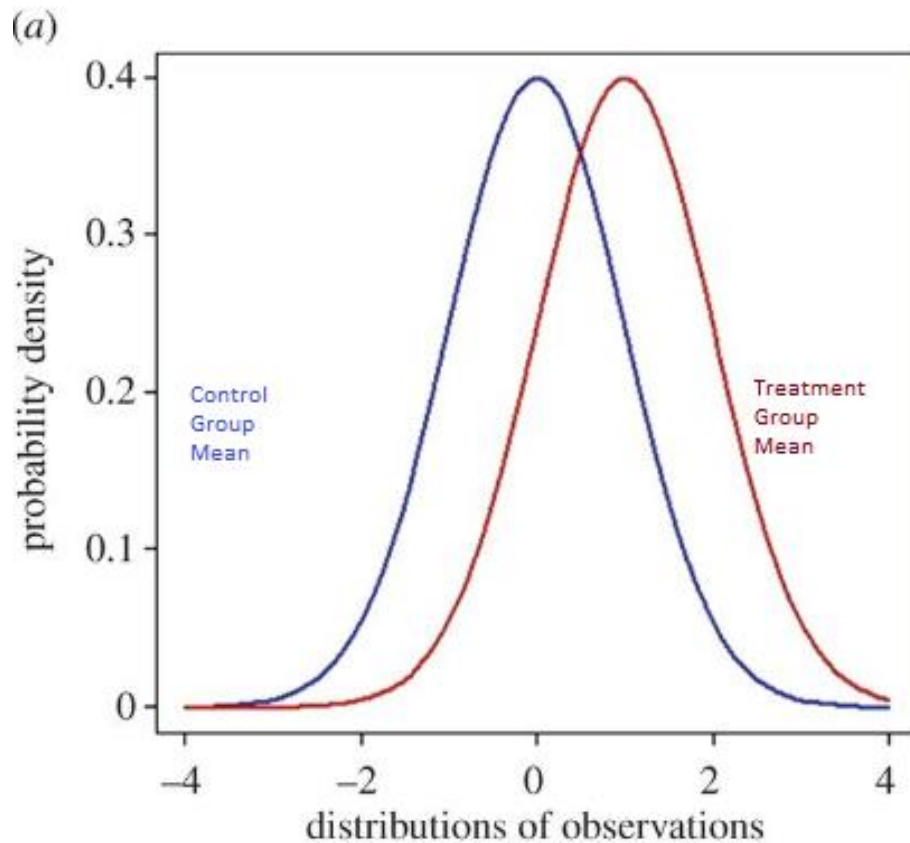
- $SS=$ Sum of Squares. Αυτός ο όρος είναι το άθροισμα των τετραγώνων των αποκλίσεων των παρατηρήσεων από τη μέση τιμή και καλείται σε συντομία άθροισμα τετραγώνων.
- Οι βαθμοί ελευθερίας ( $df$ ).
- $MS=$ Means Square. Αυτός ο όρος είναι το μέσο τετράγωνο της κάθε ομάδας και δίνει μια αμερόληπτη εκτίμηση της αντίστοιχης διακύμανσης  $\sigma^2$ .
- $F$  είναι η τιμή της στατιστικής συνάρτησης.
- $P$ -value όπως εξηγήθηκε προηγουμένως
- Κριτήριο  $F$  ( $F_{crit}$ ) είναι η κρίσιμη τιμή της στατιστικής συνάρτησης με την οποία θα γίνει η σύγκριση.

### Σχολιασμός

Παρατηρείται ότι το  $p$ -value είναι μικρότερο του  $\alpha$  μόνο στις στήλες (κατηγορία μεγέθους). Επίσης το  $F > F_{crit}$  μόνο στις στήλες. Επομένως, συμπεραίνουμε ότι υπάρχει στατιστικά σημαντική διαφορά μεταξύ των συχνοτήτων όσων αφορά το μέγεθος των πλοίων. Δηλαδή τα μικρά πλοία Γ/Φ «έλκουν» τα ατυχήματα και πρέπει να γίνει μια σοβαρή προσπάθεια να επανεξεταστούν κανονισμοί, διαδικασίες συντήρησης, διαδρομές που ακολουθούνται καθώς και το είδος φορτίου που μεταφέρεται από αυτά.

- **4.7.2 t-Test**

Το συγκεκριμένο τεστ εκτιμά κατά πόσο οι μέσες τιμές δύο δειγμάτων είναι στατιστικά σημαντικά διαφορετικές ή όχι.

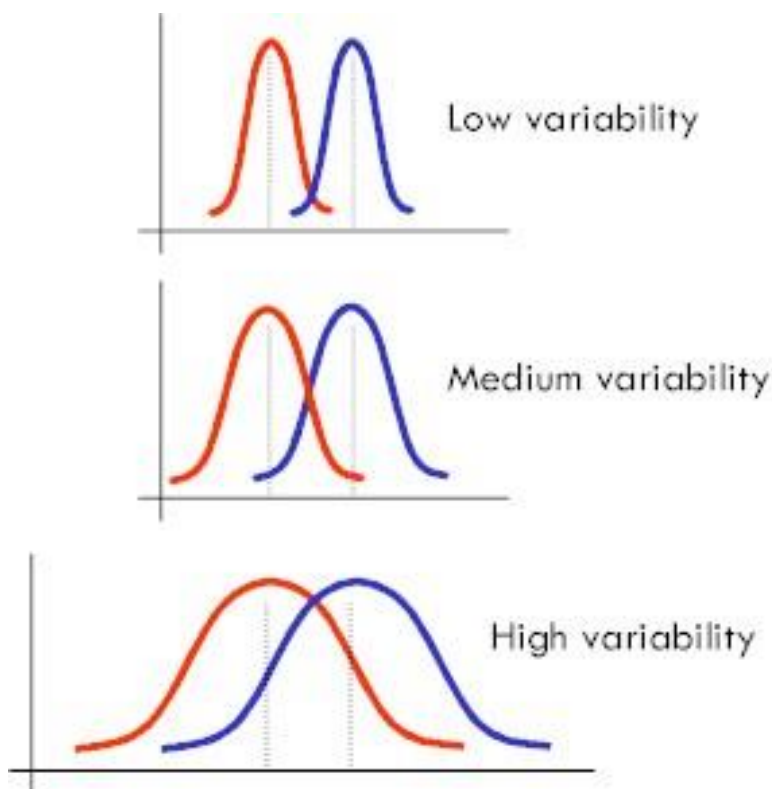


Εικόνα 2 Κατανομή των ομάδων ελέγχου και θεραπείας μιας μελέτης.

Στο παραπάνω σχήμα απεικονίζονται οι κατανομές δύο ομάδων στοιχείων. Βέβαια, οι κατανομές αυτές είναι εξιδανικευμένες. Στην πράξη, τις περισσότερες φορές έχουμε τα ιστογράμματα που προκύπτουν από το κάθε γκρουπ παρατηρήσεων.

Για να ορίσουμε τη σημασία οι μέσες τιμές από δύο γκρουπ στοιχείων-παρατηρήσεων να είναι στατιστικά διαφορετικές πρέπει να εστιάσουμε στο επόμενο σχήμα. Παρατηρούμε ότι και στις τρεις περιπτώσεις η διαφορά των μέσων τιμών είναι η ίδια. Το κάθε γράφημα όμως, παρουσιάζει και μια διαφορετική κατάσταση. Στην πρώτη και στην τρίτη περίπτωση μπορούμε να πούμε ότι οι κατανομές είναι πιο διαφορετικές, διότι η μια επικαλύπτει την άλλη σε πολύ μικρό εύρος. Αντιθέτως, στη δεύτερη περίπτωση, η διαφορά μεταξύ των γκρουπ δείχνει πολύ μικρότερη, καθώς οι δύο κατανομές επικαλύπτονται σχεδόν σε όλο τους το εύρος.

Το παράδειγμα αυτό μας οδηγεί στο εξαιρετικά σημαντικό συμπέρασμα πως όταν συγκρίνουμε δύο κατανομές ή γενικά δύο ομάδες στοιχείων, πρέπει η διαφορά μεταξύ των μέσων τιμών τους να εξετάζεται σε συνδυασμό με το εύρος ή τη μεταβλητότητα των τιμών τους. Αυτό κάνει το t-Test.



Εικόνα 3 Τρία διαφορετικά σενάρια κατανομών με ίση διασπορά μέσων τιμών.

Η συνάρτηση του t-Test είναι στην ουσία ένας λόγος. Ο αριθμητής είναι η διαφορά των μέσων τιμών ή των μέσων όρων των δύο ομάδων. Ο παρονομαστής είναι ένα μέτρο της διασποράς των τιμών. Στο επόμενο σχήμα βλέπουμε πως απεικονίζεται ο κάθε όρος στην κατανομή. Η συγκεκριμένη συνάρτηση χρησιμοποιείται και στην θεωρία Σημάτων.

$$\begin{aligned}
 \frac{\text{signal}}{\text{noise}} &= \frac{\text{difference between group means}}{\text{variability of groups}} \\
 &= \frac{\bar{X}_T - \bar{X}_C}{SE(\bar{X}_T - \bar{X}_C)} \\
 &= \text{t-value}
 \end{aligned}$$

Εικόνα 4 Συνάρτηση στο t-Test.



Ο αριθμητής είναι εύκολα υπολογίσιμος, απλά χρειάζεται να βρούμε τη διαφορά των μέσων τιμών. Ο παρονομαστής καλείται σταθερό σφάλμα της διαφοράς (standard error of difference). Για τον υπολογισμό του απαιτείται η διακύμανση του κάθε γκρουπ την οποία και διαιρούμε με το μέγεθος του δείγματος. Προσθέτουμε τους δύο λόγους και υπολογίζουμε τη ρίζα τους. Δηλαδή ισχύει :

$$SE(\bar{X}_T - \bar{X}_C) = \sqrt{\frac{\text{var}_T}{n_T} + \frac{\text{var}_C}{n_C}}$$

Εικόνα 5 Standard error of difference.

Υπενθυμίζουμε πως η διακύμανση είναι ίση με το τετράγωνο της τυπικής απόκλισης. Συνεπώς η τελική μορφή της συνάρτησης του t-Test είναι η ακόλουθη :

$$t = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{\text{var}_T}{n_T} + \frac{\text{var}_C}{n_C}}}$$

Εικόνα 6 Τελική συνάρτηση του t-Test.

Το t-value θα είναι θετικό ή αρνητικό, ανάλογα με τη σχέση που θα έχει ο αριθμητής με τον παρονομαστή. Για την εφαρμογή του t-Test απαιτείται η επιλογή του επιπέδου εμπιστοσύνης  $\alpha$  (συνήθως  $\alpha=0,05$ ). Οι βαθμοί ελευθερίας είναι ίσοι με τον αριθμό των παρατηρήσεων μείον δύο. Έχοντας το  $\alpha$ , τους βαθμούς ελευθερίας  $df$  και το t-value μπορούμε μέσω πινάκων να δούμε το πόσο στατιστικά σημαντική είναι η διαφορά των μέσων τιμών των δύο ομάδων. Η διαδικασία πλέον γίνεται αυτόματα με τα στατιστικά πακέτα.

Στην περίπτωση μας, έγινε σύγκριση μεταξύ των κατηγοριών  $500 \leq GT \leq 1000$  και  $1000 < GT \leq 20000$  γιατί αυτές συγκεντρώνουν το μεγαλύτερο αριθμό ατυχημάτων με τη δεύτερη φυσικά να κυριαρχεί. Επιλέχθηκε  $\alpha=0,05$  και τα αποτελέσματα που προέκυψαν παρουσιάζονται ακολούθως :

Πίνακας 12 Αποτελέσματα t-Test όπως προέκυψαν από το πρόγραμμα MS EXCEL.

Έλεγχος t δύο δειγμάτων με υποτιθέμενες άνισες διακυμάνσεις		
	500≤GT<1000	1000<GT≤20000
Μέσος	0.000647028	0.022694399
Διακύμανση	2.63953E-07	0.000122255
Μέγεθος δείγματος	22	22
Υποτιθέμενη διαφορά μέσων	0	
βαθμοί ελευθερίας	21	
t	-9.342591701	
P(T<=t) μονόπλευρη	3.14329E-09	
t κρίσιμο, μονόπλευρο	1.720742903	
P(T<=t) δίπλευρη	6.28657E-09	
t κρίσιμο, δίπλευρο	2.079613845	

Παρατηρούμε ότι ισχύει  $t < t_{crit}$  επομένως υπάρχει στατιστική διαφορά μεταξύ των δύο κατηγοριών.

Όταν το t-Test είναι μονόπλευρο, σημαίνει πως γνωρίζουμε πως οι μέσες τιμές των δύο ομάδων είναι διαφορετικές αλλά και ποια ομάδα έχει τη μεγαλύτερη μέση τιμή, ενώ όταν είναι δίπλευρο, αυτό δεν είναι γνωστό κι έτσι πρέπει να πραγματοποιηθεί διπλός έλεγχος.

### • 4.7.3 F-Test

Το F-Test χρησιμοποιείται για να ελέγξει αν οι τυπικές αποκλίσεις δύο πληθυσμών είναι ίσες. Όπως και το t-Test προηγουμένως, μπορεί να είναι μονόπλευρο (one-tailed) ή δίπλευρο (two-tailed). Η επιλογή γίνεται αναλόγως την περίπτωση.

Η υπόθεση F-Test ορίζεται ως εξής:

$H_0$  αν  $\sigma_1 = \sigma_2$

$H_1$  αν  $\sigma_1 > \sigma_2$  για ένα upper one-tailed test

$H_1$  αν  $\sigma_1 < \sigma_2$  για ένα lower one-tailed test

$H_1$  αν  $\sigma_1 \neq \sigma_2$  για ένα two-tailed test

Το test υπολογίζει την παράμετρο  $F = \sigma_1^2 / \sigma_2^2$ , δηλαδή το λόγο των διακυμάνσεων των δύο δειγμάτων. Όσο πιο πολύ αυτός ο λόγος προσεγγίζει το 1, τόσο πιο ισχυρές είναι οι αποδείξεις για ίσες διακυμάνσεις, άρα και ίσες τυπικές αποκλίσεις. Και σε αυτό το test είναι απαραίτητο να οριστεί το επίπεδο εμπιστοσύνης  $\alpha$ .

Η υπόθεση πως δύο τυπικές αποκλίσεις είναι ίσες απορρίπτεται εάν ισχύει

$F > F_{(\alpha, N_1-1, N_2-1)}$ , για ένα upper one-tailed test

$F < F_{(1-\alpha, N_1-1, N_2-1)}$ , για ένα lower one-tailed test

$F < F_{(1-\alpha/2, N1-1, N2-1)}$ , για ένα two-tailed test

Όπου  $F_{(\alpha, k-1, N-k)}$  είναι η κρίσιμη τιμή της κατανομής  $F$  με βαθμούς ελευθερίας  $N1, N2$  αντίστοιχα και για επίπεδο εμπιστοσύνης  $\alpha$ .

Το  $F$ -Test πραγματοποιήθηκε με το MS EXCEL για τις κατηγορίες  $500 \leq GT \leq 1000$  και  $1000 < GT \leq 20000$  και τα αποτελέσματα φαίνονται στον παρακάτω πίνακα. Έγινε η υπόθεση πως οι συχνότητες των ατυχημάτων της δεύτερης κατηγορίας εμφανίζουν μεγαλύτερη τυπική απόκλιση, ενώ επιλέχθηκε  $\alpha=0,05$ .

Πίνακας 13 Αποτελέσματα  $F$ -Test όπως προέκυψαν από το πρόγραμμα MS EXCEL.

Έλεγχος $F$ των διακυμάνσεων δύο δειγμάτων		
	500≤GT<1000	1000<GT≤20000
Μέσος	0.000647028	0.022694399
Διακύμανση	2.63953E-07	0.000122255
Μέγεθος δείγματος	22	22
βαθμοί ελευθερίας	21	21
$F$	0.002159042	
$P(F \leq f)$ μονόπλευρη	0	
$F$ κρίσιμο, μονόπλευρο	0.479803022	

Παρατηρούμε ότι ισχύει  $F < F_{crit}$  και επειδή στην περίπτωση μας ισχύει ότι  $\sigma_1 < \sigma_2$ , άρα έχουμε lower one-tailed test και επομένως απορρίπτεται η μηδενική υπόθεση πως τα δύο δείγματα έχουν ίσες τυπικές αποκλίσεις. Συνεπώς και τα αποτελέσματα αυτού του στατιστικού τεστ αποδεικνύουν τη στατιστικά σημαντική διαφορά που παρουσιάζουν οι δύο κατηγορίες μεγέθους και αναδεικνύει πως δεν είναι τυχαίο ότι τα μικρά σε μέγεθος πλοία τύπου  $G/\Phi$  παρουσιάζουν μια έλξη για ατυχήματα. Για το λόγο αυτό είναι αναγκαίο να εξεταστούν εκ νέου σημαντικοί παράμετροι της λειτουργίας τους.

#### • 4.7.4 Μέθοδος Ελέγχου Ανεξαρτησίας $\chi^2$

Ο έλεγχος  $\chi^2$  αποτελεί επαγωγικό έλεγχο, μέσω του οποίου ελέγχεται η υπόθεση ότι οι δύο μεταβλητές του πίνακα συνάφειας είναι ανεξάρτητες μεταξύ τους (δηλαδή ότι μεταβολές στις τιμές της μίας δεν προκαλούν μεταβολές στις τιμές της άλλης). Όταν δύο μεταβλητές είναι ανεξάρτητες, οι αναμενόμενες συχνότητες για τα κελιά του πίνακα συνάφειάς τους, υπολογίζονται ως το γινόμενο των αντίστοιχων συχνοτήτων κάθε κελιού, διαιρούμενο με το συνολικό πλήθος του πίνακα. Αν δύο μεταβλητές είναι ανεξάρτητες, οι παρατηρούμενες συχνότητες μπορούν να διαφοροποιούνται από τις αναμενόμενες συχνότητες μόνο στο βαθμό που οι διαφορές τους οφείλονται σε τυχαίους παράγοντες και όχι μεγαλύτερης έκτασης που μπορεί να σημαίνει ότι δεν αποδίδονται μόνο στην τύχη. Έτσι η ανεξαρτησία δύο μεταβλητών κατηγοριών, ελέγχεται με τη βοήθεια μιας συνάρτησης που βασίζεται στις διαφορές ανάμεσα

στις παρατηρούμενες και τις αναμενόμενες συχνότητες. Η συνάρτηση αυτή για έναν πίνακα με  $r$  γραμμές και  $c$  στήλες (πίνακας  $r \times c$ ) ορίζεται από την ποσότητα:

$$\chi^2 = \sum_{i=1}^r \pi_{i-1} \left( \frac{(O_i - E_i)^2}{E_i} \right)$$

$O$ =Observed (παρατηρούμενες συχνότητες)

$E$ =Expected (αναμενόμενες συχνότητες)

Οι προς διερεύνηση υποθέσεις είναι οι ακόλουθες:

$H_0$  = Οι μεταβλητές  $X$  και  $Y$  είναι ανεξάρτητες

$H_1$  = Οι μεταβλητές  $X$  και  $Y$  ΔΕΝ είναι ανεξάρτητες

Εφαρμόζεται  $\chi^2$  test σε επίπεδο σημαντικότητας  $\alpha=0,05$ . Αν το  $p$ -value για το  $\chi^2$  test  $>0,05$  δηλαδή το επίπεδο σημαντικότητας  $\alpha$ , τότε απορρίπτεται η  $H_0$  και συνεπώς οι μεταβλητές ΔΕΝ είναι ανεξάρτητες. Αν το  $p$ -value για το  $\chi^2$  test  $\leq 0,05$  τότε δεν απορρίπτεται η  $H_0$  και συνεπώς οι μεταβλητές είναι ανεξάρτητες.

Προϋποθέσεις για να είναι το  $\chi^2$  test αξιόπιστο είναι:

1. Καμία αναμενόμενη συχνότητα δε θα πρέπει να βρεθεί μικρότερη του 1 και
2. Το ποσοστό των αναμενόμενων συχνοτήτων που είναι μικρότερες από το 5, δε θα πρέπει να υπερβαίνει το 20%.

Τα δεδομένα που εξετάστηκαν αν επιδρούν στον Τύπο του Ατυχήματος είναι :

- Age Group
- Flag Color
- GT Category
- Fatalities
- Loss or Disposal
- Pollution
- Severity

Ενδεικτικά παρουσιάζονται στους επόμενους πίνακες τα αποτελέσματα των δύο πρώτων κατηγοριών, ενώ όλες οι μεταβλητές αποδείχτηκαν ανεξάρτητες

Πίνακας 14 Δεδομένα μεταβλητής Age Group με Τύπο Ατυχήματος

Age_Group	Col	Con	Fir	Fou	H-M	W-S	Total
0--4	175	83	44	14	275	162	753
5--9	178	63	50	24	308	184	807
10--14	128	55	30	8	259	115	595
15--19	97	33	29	11	203	110	483
20++	56	19	24	6	68	51	224
Total	634	253	177	63	1113	622	2862

Μέσω της διαδικασίας του στατιστικού ελέγχου που περιεγράφηκε παραπάνω προέκυψε πως η μεταβλητή Age Group είναι ανεξάρτητη του Τύπου Ατυχήματος και  $P\text{-value} = 0.0171 < 0.05$ . Η τιμή αυτή δείχνει ότι η στατιστική σημαντικότητα αυτού του συστήματος δεδομένων είναι μεγάλη.

Πίνακας 15 Δεδομένα μεταβλητής Flag Color με Τύπο Ατυχήματος

Flag_Color	Col	Con	Fir	Fou	H-M	W-S	Total
Black	27	4	12	8	43	27	121
Grey	99	34	34	16	146	103	432
White	508	215	131	39	924	492	2309
Total	634	253	177	63	1113	622	2862

Μέσω της διαδικασίας του στατιστικού ελέγχου που περιεγράφηκε παραπάνω προέκυψε πως η μεταβλητή Flag Color είναι ανεξάρτητη του Τύπου Ατυχήματος και  $P\text{-value} = 0.00018 < 0.05$ . Η τιμή αυτή δείχνει ότι η στατιστική σημαντικότητα αυτού του συστήματος δεδομένων είναι μεγάλη. Η ίδια διαδικασία ακολουθήθηκε και για τις υπόλοιπες μεταβλητές και προέκυψαν ανεξάρτητες του Τύπου Ατυχήματος.

## ΚΕΦΑΛΑΙΟ 5<sup>ο</sup> : ΑΝΑΠΤΥΞΗ ΔΕΝΤΡΩΝ ΚΑΤΗΓΟΡΙΟΠΟΙΗΣΗΣ & ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Στόχος του κεφαλαίου είναι ο προσδιορισμός των αιτιών που οδηγούν στους τύπους ατυχημάτων τους οποίους αναφέραμε και προηγουμένως. Για να γίνει αυτό πραγματικότητα, τα ατυχήματα των πλοίων Γ/Φ που έχουμε στην κατοχή μας μαζί με τις λεπτομερείς τους πληροφορίες, χωρίστηκαν με βάση τον τύπο του ατυχήματος σε έξι κατηγορίες ατυχήματος.

Η μέθοδος που συσχετίζει τις πληροφορίες για το κάθε ατύχημα με βάση μια αρχική μεταβλητή και μας δίνει αποτελέσματα για το τι κοινό είχαν τα ατυχήματα, ονομάζεται παλινδρόμηση (Regression) και κάνοντας χρήση αυτής της μεθόδου προκύπτουν τα αντίστοιχα δέντρα (Regression Trees).

Όσον αφορά την παραπάνω μέθοδο, δε θα είχαμε αποτελέσματα αν δε γινόταν χρήση κάποιου προγράμματος ή κάποιας γλώσσας προγραμματισμού. Επιλέχθηκε να γίνει χρήση της γλώσσας προγραμματισμού R και των βοηθητικών της όπως το R-Studio.

### 5.1 R & R-Studio

Η R αποτελεί μια γλώσσα προγραμματισμού της οποίας το λογισμικό είναι ελεύθερο για το κοινό και χρησιμοποιείται κυρίως για στατιστικούς υπολογισμούς και γραφική απεικόνιση και υποστηρίζεται από τον R Foundation for Statistical Computing. Η γλώσσα R είναι ευρέως χρησιμοποιούμενη από όσους ασχολούνται με τους κλάδους της Στατιστικής (rolls) και την εξόρυξη δεδομένων (data mining). Οι μελέτες των τελευταίων ετών δείχνουν σημαντική αύξηση στη χρήση της συγκεκριμένης γλώσσας στους προαναφερθέντες κλάδους. Συγκεκριμένα τον Αύγουστο του 2018 η R κατέλαβε τη 18<sup>η</sup> θέση στη λίστα TIOBE, η οποία αποτελεί μέθοδο μέτρησης της δημοφιλίας και χρήσης προγραμμάτων στατιστικής παγκοσμίως.



Εικόνα 7 Λογότυπο προγράμματος R.

Η σύνταξη του κώδικα της R βασίστηκε αρχικά στη γλώσσα C και στη FORTRAN. Παρότι η R παρέχει ένα διαδραστικό περιβάλλον εντολών (command line interface), υπάρχουν αρκετά αρχικά κομμάτια προγράμματος σε μια σειριακή ροή διαδικασιών-επεξεργασίας πληροφοριών (graphical front-ends) το κυριότερο από αυτά είναι το R-Studio.

Το πρόγραμμα αυτό είναι εύκολο στην εγκατάσταση και κατεβάζοντας παράλληλα βιβλιοθήκες και διάφορα απαραίτητα εργαλεία, γίνεται εφικτό να αποκτήσει κανείς πολλές

επιλογές για την επεξεργασία δεδομένων και εξόρυξη συμπερασμάτων για τα δεδομένα. Μία πολύ χρησιμοποιούμενη δυνατότητα που δίνεται μέσω του R είναι η κατηγοριοποίηση των μεταβλητών του δείγματος (Classification) και η παλινδρόμηση (Regression) μεταξύ των μεταβλητών για να γίνει αντιληπτό αν αλληλοεπιδρούν, σε ποιο βαθμό και με ποια σειρά. Για να γίνει αυτό εμφανές, το R δίνει τη δυνατότητα αποτύπωσης των αποτελεσμάτων σε δέντρα (Classification trees or Regression trees).

## 5.2 Εισαγωγή Δεδομένων στο πρόγραμμα R

Τα δεδομένα που χρησιμοποιήθηκαν είναι τα ατυχήματα που προέκυψαν για General Cargo Ships με  $GT \geq 500$  την περίοδο 1995-2016 και τα στοιχεία που τα συνόδευαν στην ανάλυση μέσω του προγράμματος R είναι :

1. Age Group
2. Flag Color
3. GT Category
4. Loss or Disposal
5. Pollution
6. Severity
7. Fatalities
8. Casualty Type

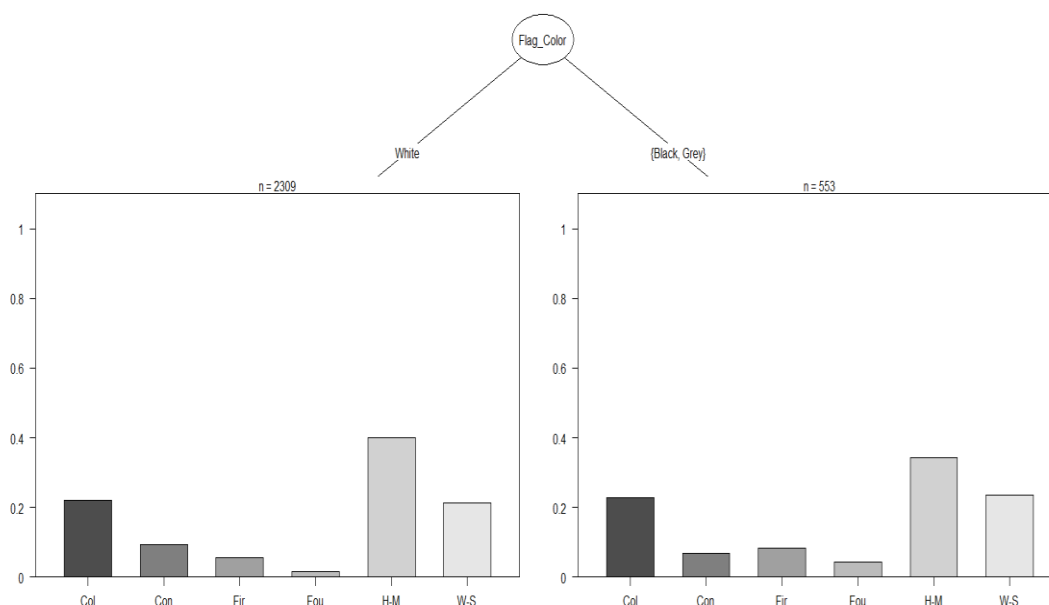
Οι παραπάνω μεταβλητές έχουν κατηγοριοποιημένα τα δεδομένα τους (Categorical values) και για την κάθε μία δίνονται ακολούθως οι πιθανές κατηγορίες που περιέχουν :

1. 0-4 , 5-9 , 10-14 , 15-19 , 20+
2. White , Black , Grey
3.  $GT-1 = 500 \leq GT < 1000$  ,  $GT-2 = 1000 \leq GT \leq 20000$  ,  $GT-3 = GT > 20000$
4. Casualty , Disposal , Total Loss
5. Yes , No , Unknown
6. Serious ,
7. Yes , No
8. Col = Collision, Con = Contact, Fir = Fire/Explosion, Fou = Foundered, H-M = Hull/Mchy Dam. , W-S = Wrecked/Stranded

Τα δεδομένα εισάγονται σε μορφή αρχείου .tab και το πρόγραμμα τα διαβάζει, επιτρέπει στο χρήστη να τα επεξεργαστεί και να βγάλει τα αποτελέσματα που θέλει. Στη δική μας περίπτωση ο αριθμός των ατυχημάτων που έχουμε πλήρη γνώση για τις παραπάνω μεταβλητές ανέρχεται σε 2682 ατυχήματα.

### 5.2.1 Το Συνολικό Δείγμα

Σε αυτή τη διπλωματική σκοπός μέσω των δέντρων παλινδρόμησης (regression trees) να βρεθεί ποιες μεταβλητές αλληλοεπιδρούν και προκύπτουν τα ατυχήματα αυτά στα General Cargo Ships. Για να γίνει αυτό, δίνουμε στο πρόγραμμα μια από τις παραπάνω μεταβλητές ως αρχική και αυτό μας δείχνει ποιες το επηρεάζουν και με ποια σειρά. Στην δική μας περίπτωση όπου έχουμε categorical values έπρεπε να φορτώσουμε στο R μια βιβλιοθήκη κατάλληλη για αυτή την περίπτωση η οποία δόθηκε με την εντολή library (party). Αφού βρέθηκε ο κατάλληλος κώδικας για Regression Trees με Categorical Values, χρησιμοποιήθηκε για τα δεδομένα συνολικά και μας έδωσε τα ακόλουθα αποτελέσματα. Ο κώδικας μαζί με το συνολικό όγκο εισαγωγής δεδομένων και εξαγωγής αποτελεσμάτων παρατίθεται σε παράρτημα.

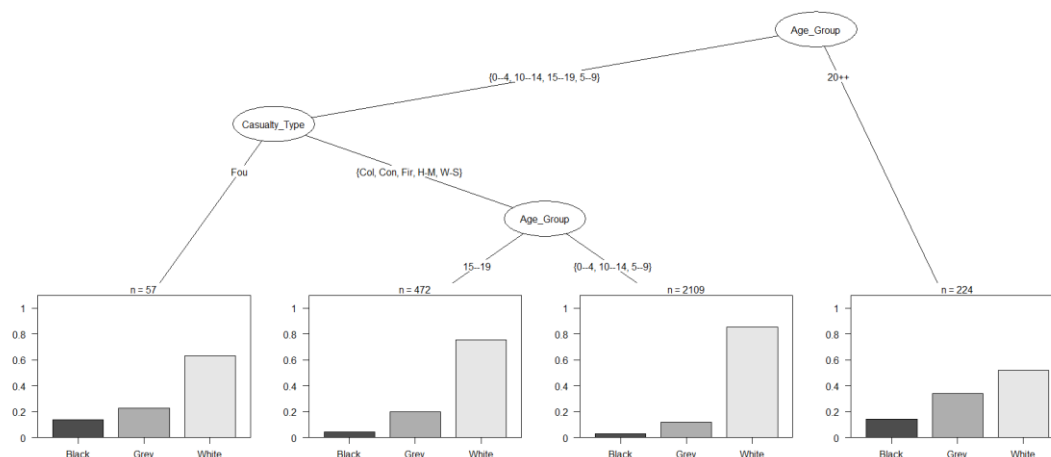


Σχήμα 38 Regression Tree με αρχική μεταβλητή Casualty Type.

Στο παραπάνω παράδειγμα (**Σχήμα 38**), απεικονίζεται το αποτέλεσμα που μας έδωσε το πρόγραμμα για αρχική μεταβλητή τον τύπο του ατυχήματος και φαίνεται πως η μεταβλητή που αλληλοεπιδρά είναι η κατηγορία της σημαίας του πλοίου. Και στις δύο περιπτώσεις το μεγαλύτερο συντελεστή ατυχημάτων έχει ο τύπος Hull/Mchgy Dam, ενώ ακολουθούν Collision και Wrecked/Stranded, αφού έχει γίνει γνωστό από το προηγούμενο κεφάλαιο πως αυτοί οι τύποι ατυχήματος εμφανίζονται πιο συχνά.

Μεγάλο ενδιαφέρον προκαλεί το επόμενο αποτέλεσμα (**Σχήμα 39**), στο οποίο αρχική μεταβλητή είναι το Flag Color, διότι στις 224 περιπτώσεις όπου το πλοίο το οποίο είχε ατύχημα ήταν πάνω των 20 ετών, δε χρειάστηκε άλλη αλληλεπίδραση, ενώ για πλοία μικρότερων ηλικιών δείχνει να παίζει ρόλο ο τύπος του ατυχήματος, με αυτά που ανήκουν στην κατηγορία Foundered (57 περιπτώσεις) να είναι σε μεγάλη πλειοψηφία με σημαία White.



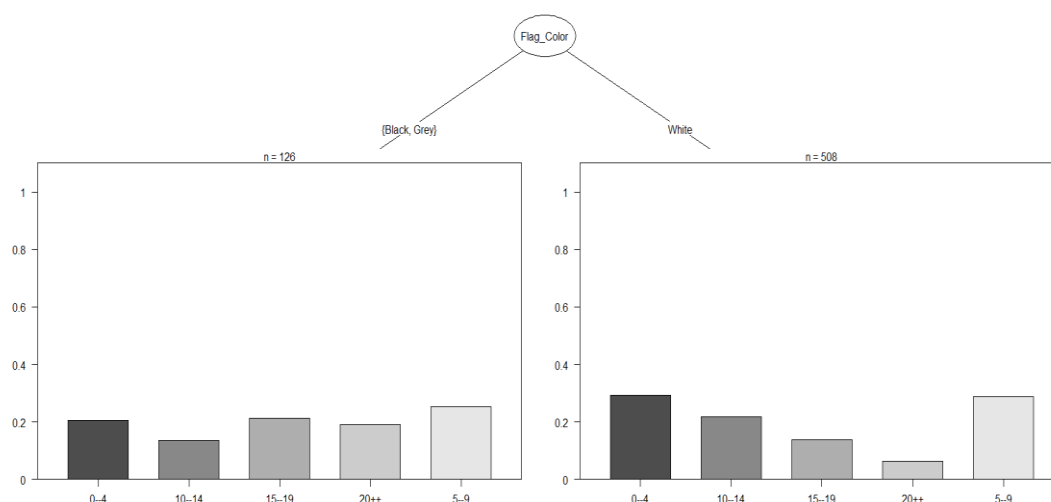


Σχήμα 39 Regression Tree με αρχική μεταβλητή Flag Color.

Επειδή τα δεδομένα μας περιέχουν αρκετές πληροφορίες, κρίθηκε κατάλληλο να διασπαστούν σύμφωνα με τον τύπο του ατυχήματος σε 6 κατηγορίες. Έτσι θα γίνει πιο ξεκάθαρο ποιες μεταβλητές οδηγούν στον εκάστοτε τύπο ατυχήματος, με στόχο τη δραστηκή μείωση των ατυχημάτων για τα General Cargo Ships.

### 5.2.2 Το Δείγμα για ατυχήματα σύγκρουσης (Collision)

Στο δείγμα υπάρχουν 634 περιπτώσεις ατυχημάτων και δεδομένο είναι ο τύπος ατυχήματος πως είναι σύγκρουση. Αν η αρχική μεταβλητή που δίνεται στο πρόγραμμα είναι η ηλικία του πλοίου όταν αυτό είχε ατύχημα, δηλαδή η μεταβλητή Age Group, το αποτέλεσμα που προκύπτει είναι το ακόλουθο :

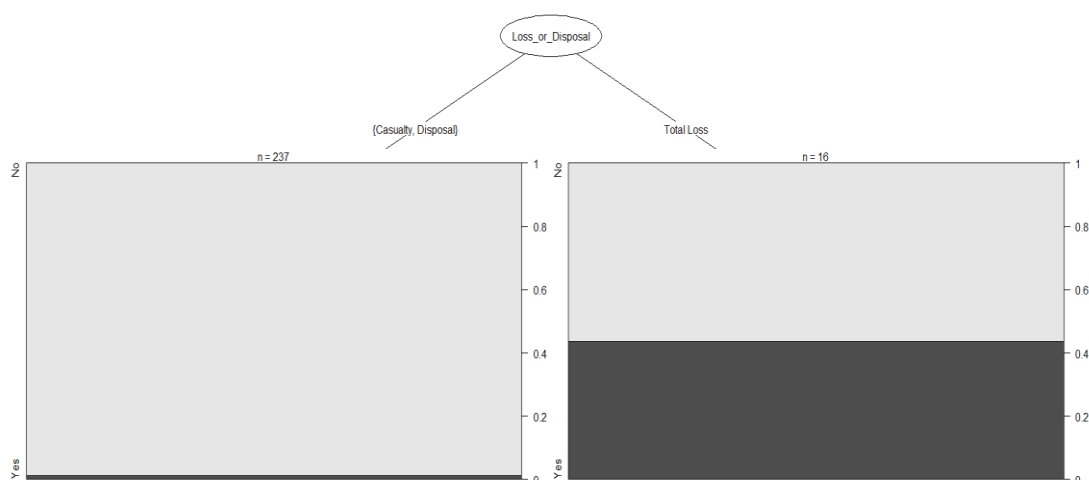


Σχήμα 40 Regression Tree με αρχική μεταβλητή Age Group.

Το ηλικιακό γκρουπ 5-9 καταλαμβάνει το ποσοστό 25% των ατυχημάτων στις περιπτώσεις όπου η σημαία του πλοίου ανήκει στις υποδεέστερες κατηγορίες (Black, Grey) ενώ κατέχει και το 30% στις περιπτώσεις όπου η σημαία είναι White.

### 5.2.3 Το Δείγμα για ατυχήματα επαφής (Contact)

Στο δείγμα υπάρχουν 253 περιπτώσεις ατυχημάτων και δεδομένο είναι ο τύπος ατυχήματος πως είναι επαφή. Αν η αρχική μεταβλητή που δίνεται στο πρόγραμμα είναι αν υπήρχαν νεκροί ή αγνοούμενοι μετά από ατύχημα επαφής, δηλαδή η μεταβλητή Fatalities, το αποτέλεσμα που προκύπτει είναι το ακόλουθο :

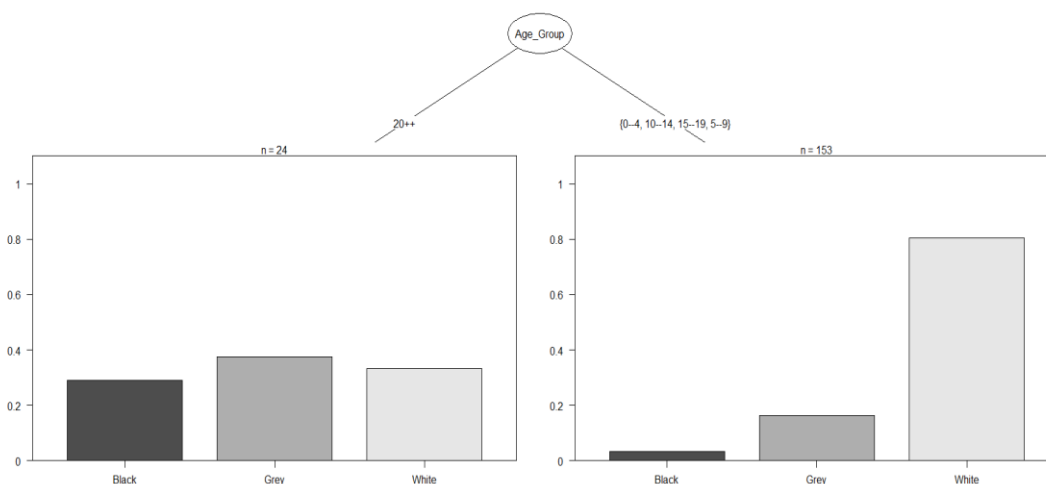


Σχήμα 41 Regression Tree με αρχική μεταβλητή Fatalities.

Φαίνεται πως σε περιπτώσεις ατυχημάτων Contact όπου το πλοίο οδηγείται σε Total Loss το αποτέλεσμα είναι 45% να έχουμε και ανθρώπινες απώλειες.

### 5.2.4 Το Δείγμα για ατυχήματα πυρκαγιάς/έκρηξης (Fire/Explosion)

Στο δείγμα υπάρχουν 177 περιπτώσεις ατυχημάτων και δεδομένο είναι ο τύπος ατυχήματος πως είναι πυρκαγιά ή έκρηξη. Αν η αρχική μεταβλητή που δίνεται στο πρόγραμμα είναι τι κατηγορίας σημαία είχαν τα πλοία που έπαθαν ατύχημα αυτού του τύπου, δηλαδή η μεταβλητή Flag Color, το αποτέλεσμα που προκύπτει είναι το ακόλουθο :



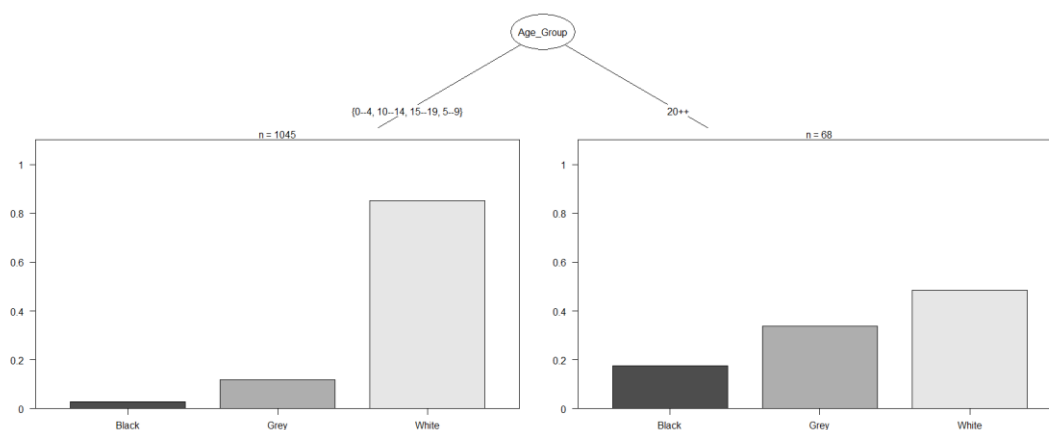
Σχήμα 42 Regression Tree με αρχική μεταβλητή Flag Color.

Η μεγαλύτερη αλληλεπίδραση που βρίσκει το πρόγραμμα σε αυτή την περίπτωση είναι το ηλικιακό γκρουπ στο οποίο άνηκε το πλοίο τη στιγμή του ατυχήματος. Στις περιπτώσεις που είχε λοιπόν ηλικία μεγαλύτερη των 20 ετών, τα πλοία αυτά ήταν από χώρες με κατώτερης κλάσης σημαίες (Black or Grey) σε ποσοστό 70%.

### 5.2.5 Το Δείγμα για ατυχήματα από βλάβη στη γάστρα ή στο μηχανοστάσιο (Hull/Mchy Dam.)

Στην κατηγορία ατυχημάτων Hull/Mchy Dam. ανήκουν 1113 περιστατικά και αποτελεί την πιο συχνά εμφανιζόμενη κατηγορία. Μπορεί πολύ εύκολα ένα περιστατικό αυτής της κατηγορίας να δημιουργήσει σοβαρό ατύχημα λόγω της ζωτικής σημασίας για το πλοίο τόσο της γάστρας όσο και του μηχανοστασίου.

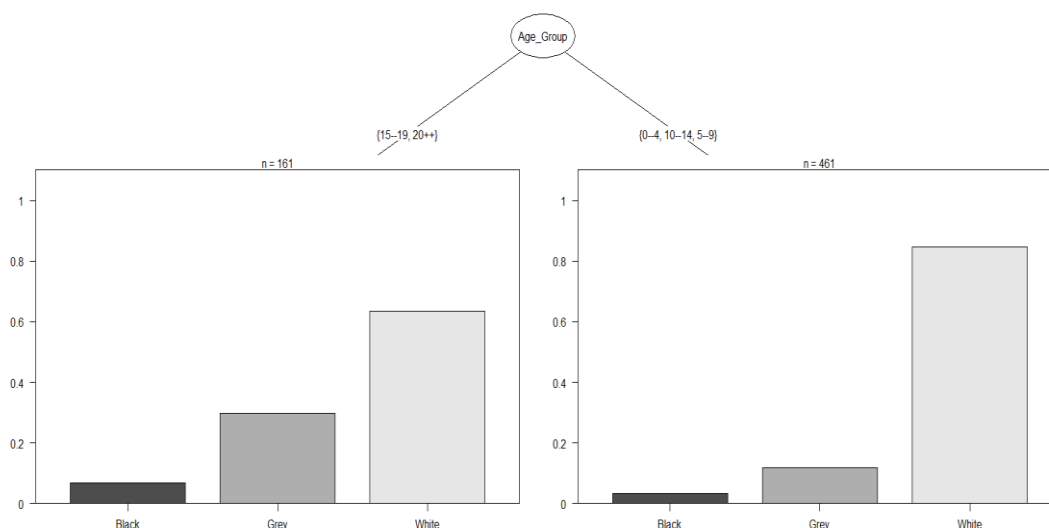
Στο **Σχήμα 43** η αρχική μεταβλητή είναι το χρώμα της σημαίας και φαίνεται πως έχει μεγάλη αλληλεπίδραση η ηλικία του πλοίου. Στις 1045 περιπτώσεις όπου το πλοίο ήταν ηλικίας κάτω των 20 ετών, η πιθανότητα να έχει λευκή σημαία ήταν 80%. Παρόλο που πλοία με λευκή σημαία έχουν αυστηρότερους κανονισμούς και προσέχονται ιδιαίτερα από τις εταιρίες τους, φαίνεται πως τα ατυχήματα στα πλοία αυτά ακμάζουν στην κατηγορία Hull/Mchy Dam.



Σχήμα 43 Regression Tree με αρχική μεταβλητή Flag Color.

### 5.2.6 Το Δείγμα για ατυχήματα πλοίων που καταστράφηκαν ή ακινητοποιήθηκαν (Wrecked/Stranded)

Η κατηγορία ατυχημάτων Wrecked/Stranded κατέχει κι αυτή μεγάλο μερίδιο του συνόλου των ατυχημάτων με 622 περιπτώσεις.



Σχήμα 44 Regression Tree με αρχική μεταβλητή Flag Color.

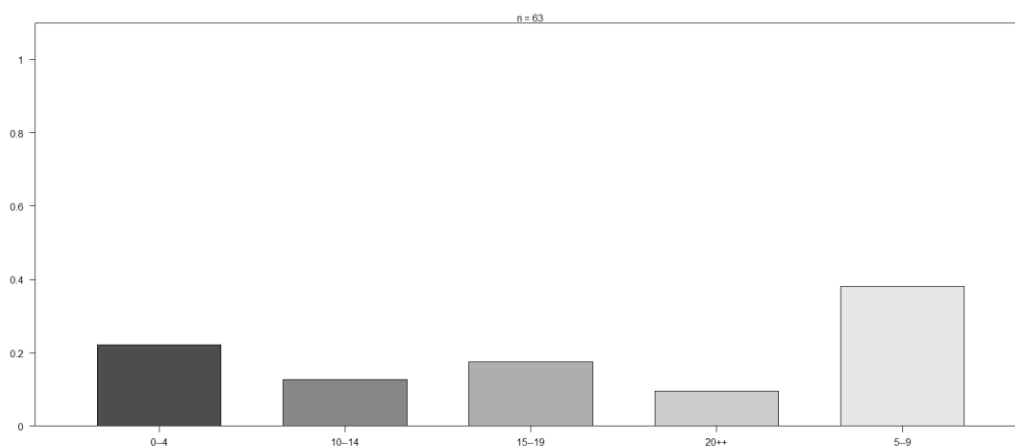
Ορίζοντας στο πρόγραμμα R ως αρχική μεταβλητή τη σημαία των πλοίων που είχαν ατύχημα αυτής της κατηγορίας, μεγάλη αλληλεπίδραση δείχνει να έχει η ηλικία του πλοίου, με τις περισσότερες περιπτώσεις να ανήκουν στα νεότερα σε ηλικία πλοία, και όχι στα πλοία που

ανήκουν στα ηλικιακά γκρουπ 15-19 ή 20 και πάνω ετών. Η σημαία που κατέχουν τα πλοία πάντως σε μεγάλο ποσοστό είναι White.

Ίσως το αποτέλεσμα αυτό να προήλθε από το γεγονός ότι στη Διπλωματική γίνεται μελέτη ατυχημάτων που συνέβησαν την περίοδο 1995-2016 ενώ δεν έχουμε καθόλου πληροφορίες για πλοία που χτίστηκαν πριν το 1990. Αυτό μας φέρνει σε θέση να αποκτάμε πληροφορίες για το ηλικιακό γκρουπ 15-19, μόνο την περίοδο 2005-2016. Όμως ο υψηλός αριθμός ατυχημάτων αυτής της κατηγορίας είναι ανησυχητικός σε κάθε περίπτωση και για κάθε ηλικία πλοίου.

### 5.2.7 Το Δείγμα για ατυχήματα βύθισης (Foundered)

Στα ατυχήματα βύθισης General Cargo Ships υπάρχουν μόνο 63 περιπτώσεις και αυτό καθιστά δύσκολο στο πρόγραμμα να βρει κάποια μεγάλη αλληλεπίδραση με τόσο μικρό δείγμα.



Σχήμα 45 Regression Tree με αρχική μεταβλητή Age Group.

Έτσι λοιπόν παρατίθεται το αποτέλεσμα του προγράμματος χωρίς αλληλεπίδραση αλλά με αρχική μεταβλητή το ηλικιακό γκρουπ. Αυτό που παρατηρείται είναι πως από τα 63 ατυχήματα, η πιθανότητα να ανήκει το πλοίο στα δύο νεότερα σε ηλικία γκρουπ είναι 60%.

## ΚΕΦΑΛΑΙΟ 6<sup>ο</sup> : ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ

Λόγω της ευρείας περιόδου που μελετάται στη Διπλωματική αυτή εργασία (1995-2016) περιλαμβάνει ένα μεγάλο κομμάτι κλιμακωμένης ανάπτυξης του στόλου των πλοίων Γ/Φ. Το συμπέρασμα της μελέτης αυτής της περιόδου 21 ετών δείχνει πως περισσότερο από το στόλο αυξήθηκαν τα ατυχήματα με αποκορύφωμα την τελευταία 10ετία. Αυτό γίνεται εμφανές στα διαγράμματα Συχνότητας Ατυχημάτων του κεφαλαίου 4 με σημαντικότερο παράδειγμα τον τύπο ατυχήματος Hull/Mchy Dam.

Επίσης είναι σημαντικό να αναφερθεί ότι η ολική απώλεια του πλοίου και οι ανθρώπινες απώλειες στα ατυχήματα των πλοίων Γ/Φ δείχνουν να έχουν σταθερές συχνότητες παρά τις προσπάθειες που πάντα γίνονται με σκοπό αυτά τα δύο ατυχήματα να μειωθούν.

Ένα ακόμα σημαντικό συμπέρασμα που βγαίνει από τη μελέτη αυτή είναι πως το έτος κατασκευής αποτελεί σημαντικό παράγοντα στα ατυχήματα πλοίων Γ/Φ διότι τα ατυχήματα πλοίων που χτίστηκαν την τελευταία δεκαετία ήταν ελάχιστα σε σχέση με αυτά της προηγούμενης.

Το πρόβλημα της ηλικίας ήρθε να επιβεβαιωθεί από το κεφάλαιο 5 μέσω του προγράμματος R μαζί με το ζήτημα της σημαίας του πλοίου, όπου παρατηρήθηκε ότι, στους περισσότερους τύπους ατυχημάτων, τα πλοία που είχαν ατύχημα, η πιθανότητα να κατέχουν λευκή σημαία ήταν πάνω από 60%.

Αν και τα αποτελέσματα που έδωσε η ανάλυση των δεδομένων συγκεκριμένης βάσης είναι ικανοποιητικά, μια παρόμοια ανάλυση σε στοιχεία κάποιου άλλου φορέα θα ήταν επιθυμητή.

Άλλοι άξονες που θα μπορούσαν να προστεθούν στην έρευνα αυτή είναι ο καιρός τη στιγμή του ατυχήματος ή το κόστος του ατυχήματος μεταφρασμένο σε χρηματικές μονάδες.

Επίσης σημαντικά θέματα για μελέτη θα αποτελούσαν η σύνθεση του πληρώματος, η εθνικότητα καθώς και η εμπειρία τους, μιας και πολλά ατυχήματα αποδίδονται στον ανθρώπινο παράγοντα.

## BIBΛΙΟΓΡΑΦΙΑ

- Arben Mullaj, Ulf Paulson, «A grounded theory model for Analysis of marine accidents» , Lund University, Science Direct , March 2011
- Eleftheria Eliopoulou & Apostolos Papanikolaou «Statistical analysis of ship accidents and review of safety level», February 2016
- Eleftheria Eliopoulou, Apostolos Papanikolaou, «Casualty Analysis of large Tankers» ,Journal of Marine Science and Technology, June 2007
- EnviroEmerg Consulting Services, «Major Marine Vessel Casualty Risk and Response Preparedness in British Columbia», 2008
- Franz Faul, Edgar Erdfelder, Axel Buchner, Albert-Georg Lang, «Statistical power analyses using G\*Power 3.1 : Tests for correlation and regression analyses», November 2009
- Han J. «Data Mining Concepts and Techniques», 2001
- IACS, «FSA of General Cargo Ships», July 2008
- J.S Paik,D.K Kim, «A New Methology for Assessing the Safety of Damaged Ships» ,International Journal of Maritime Engineering
- Jacquet-Lagrece E., Siskos Y. (1982), «Assessing a set of additive utility functions for multicriteria decision making: the UTA method» , European Journal of Operational Research 10, 151-164
- M. Bogalecka & M. Popek «Analysis of Sea Accidents in 2006»
- Mckay M. ,Beckman R. ,Conover W., «A comparison of three methods for selecting values of Input Variables in the analysis of output from a computer code» , Technometrics,Vol.21,No.2,May 1979
- Nikos Themelis, Kostas Spyrou, «Probabilistic Framework for Safety Onboard Fire Safety» ,contribution of NTUA to task «The Methology for the probabilistic generation of HRR curves» ,July 2010
- Papanikolaou P., Eliopoulou E., Alisafaki A. et al, «Critical review of Aframax tanker incidents», The Third International ENSUS Conference, Newcastle 2005
- Papanikolaou P., Eliopoulou E., Mikelis N.et al, «Casualty analysis of tankers», RINA Learning from Marine Incidents III, London 2006
- Priscilla E. Greenwood, Michael S. Nikulin «A guide to Chi-Squared Testing», April 1996
- Quinlan, J.R. (1986), "Induction of decision trees. Machine Learning"
- Samouilides M.S, Ventikos N.P, Gemelos I.C, «Survey on grounding accidents: statistical analysis and risk assessment», Ships and Offshores Structures,Vol.4,No.1,pp. 55-68,2009
- Schulte J. , «Casualty Investigation Report CA – 63», 2000
- Thomas K. Tiemann, «Introductory Business Statistics with Interactive Spreadsheers – 1<sup>st</sup> Canadian Edition», February 2010

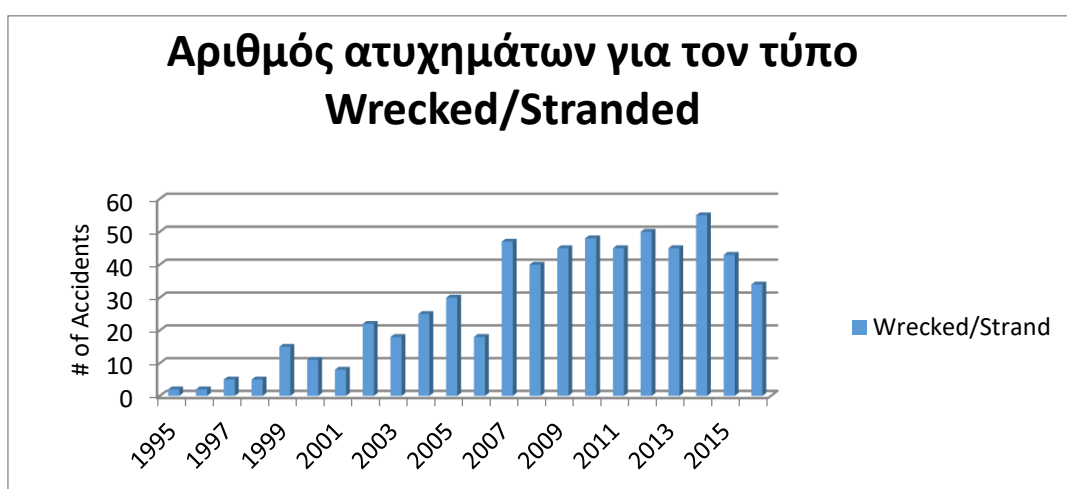
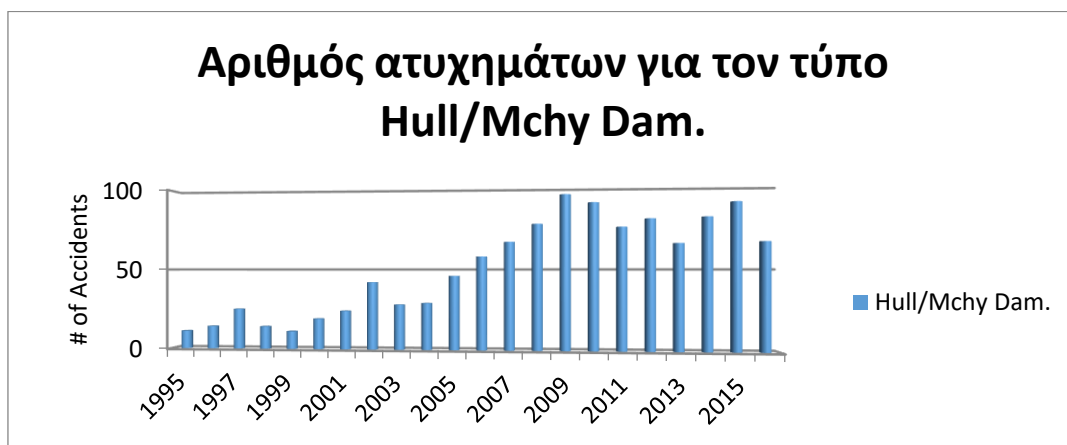
- Tzannatos Ernestos, Kokotos Dimitris, «Analysis of Accidents in Greek shipping during the pre and post-ISM period», July 2009
- Wijnolst N., Wergeland T., «Shipping Innovation», Amsterdam 2009
- Witten I. H. , «Data Mining Practical Machine Learning Tools and Techniques with Java» 2000
- Ηλιοπούλου, Ε. (2006). *Μεθοδολογία ανάπτυξης νέων κανονισμών ευστάθειας πλοίων κατόπιν βλάβης και επιδράσεις στη μελέτη και σχεδίαση*. Αθήνα: ΕΜΠ, ΝΜΜ, Τομέας Μελέτης Πλοίου και Θαλάσσιων Μεταφορών.
- Μπίθα, Κ. (2014). Στατιστική ανάλυση ναυτικών ατυχημάτων κατά την περίοδο 1990-2012. Αθήνα: ΕΜΠ, ΝΜΜ, Τομέας Μελέτης πλοίου και Θαλάσσιων Μεταφορών.
- Σωτήραλης Π. (2011). Πιθανοθεωρητική μελέτη διαρροών πετρελαίου από Δ/Ξ : Ποσότητες και Κόστος. Αθήνα: ΕΜΠ, ΝΜΜ, Τομέας Μελέτης πλοίου και Θαλάσσιων Μεταφορών.

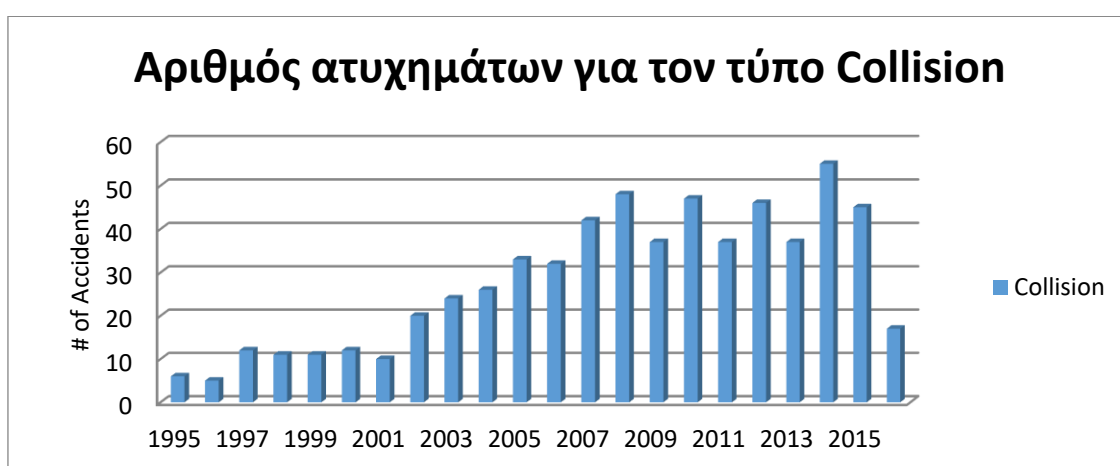
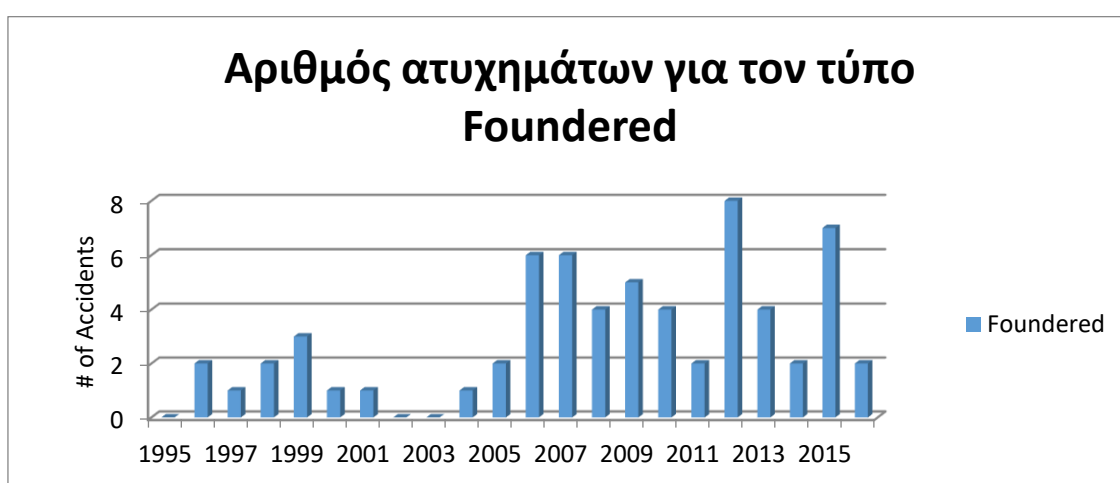
#### Διαδικτυακοί Τόποι

- <https://www.sciencedirect.com/science/article/pii/S0925753516000369>
- <http://dspace.lib.ntua.gr/handle/123456789/5463>
- <http://ftp.auckland.ac.nz/software/CRAN/doc/vignettes/party/party.pdf>
- <https://link.springer.com/article/10.1007%2Fs00038-011-0315-z>
- <https://www.econstor.eu/handle/10419/73524>
- <https://www.sciencedirect.com/science/article/pii/S0308597X09000074>
- <https://seaweb.org>
- <https://www.clarksons.com>
- <http://www.equasis.org>
- <http://www.emsa.europa.eu>



## ΠΑΡΑΡΤΗΜΑ Α : ΣΥΝΟΛΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

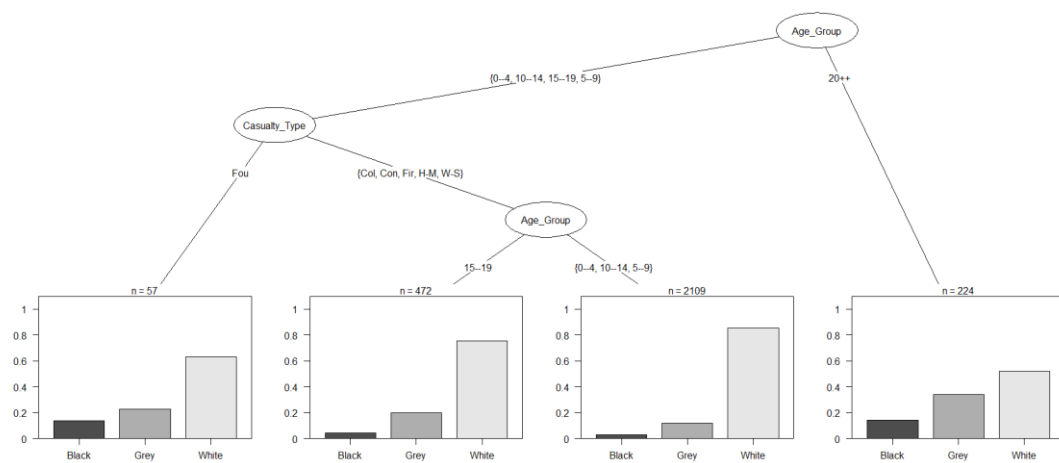
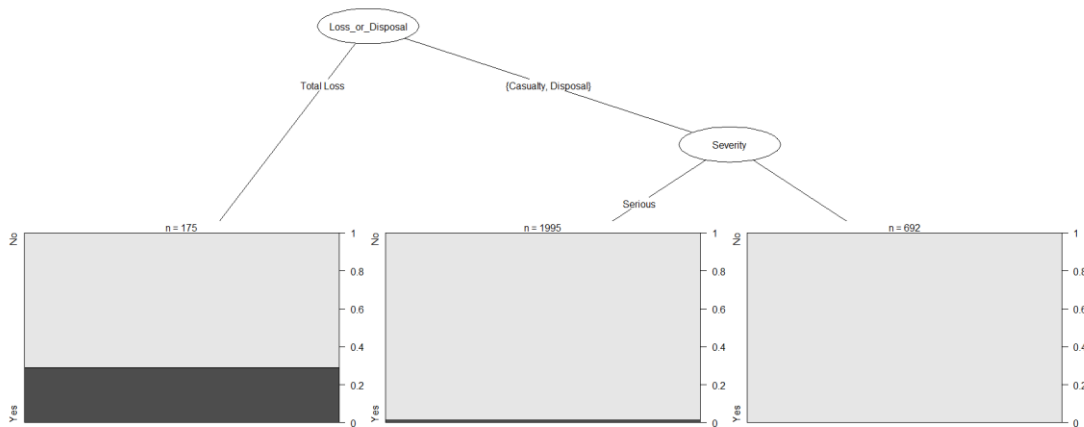
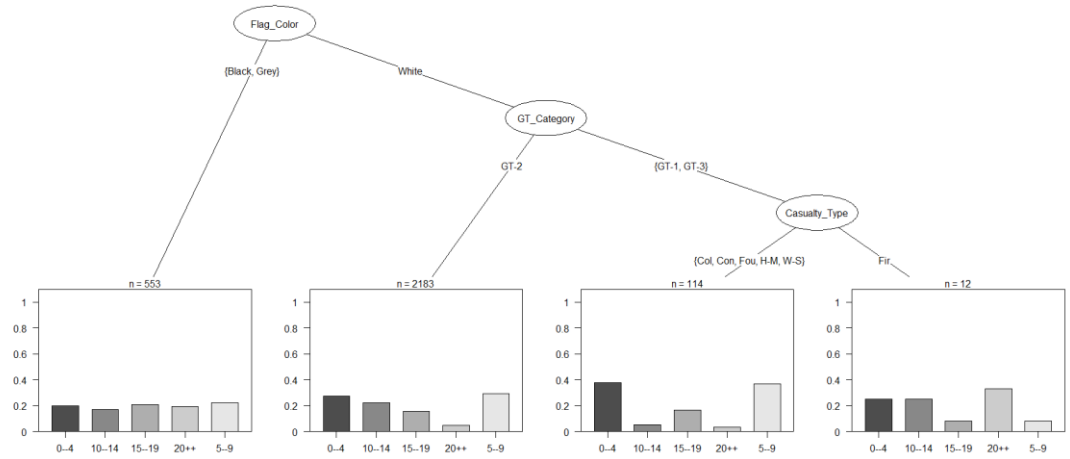


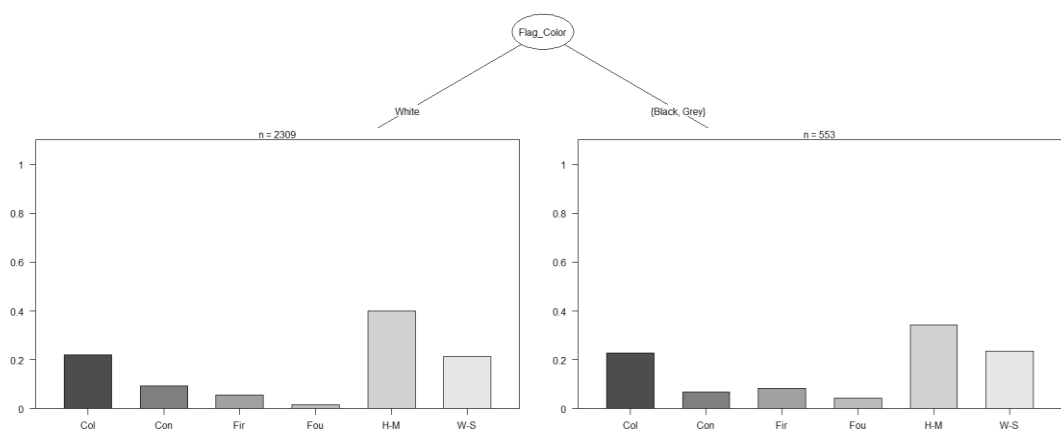
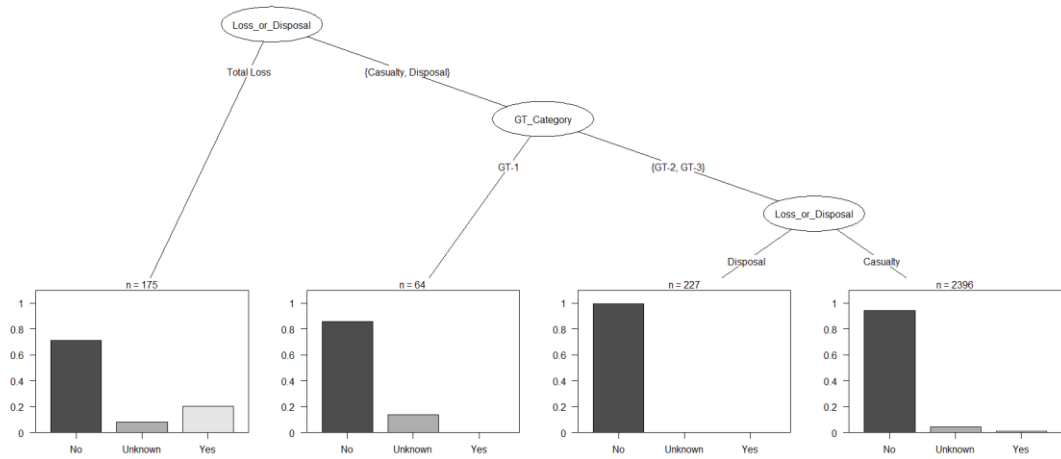
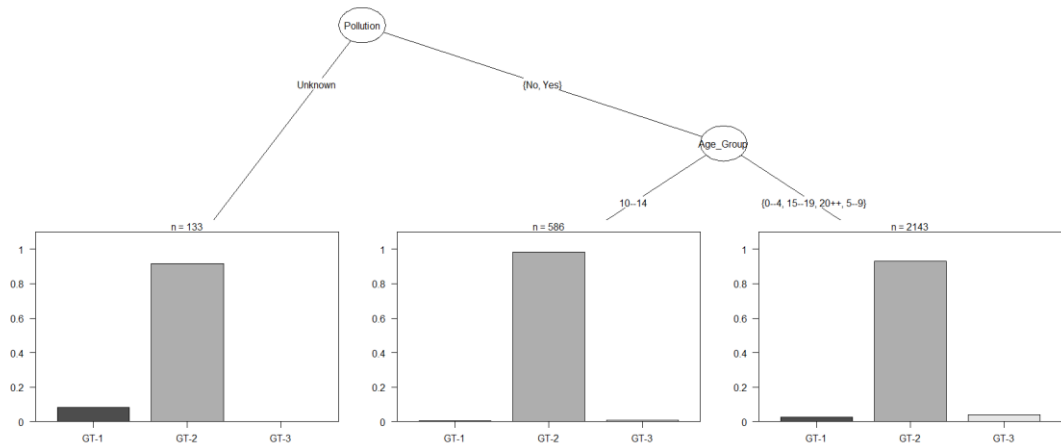




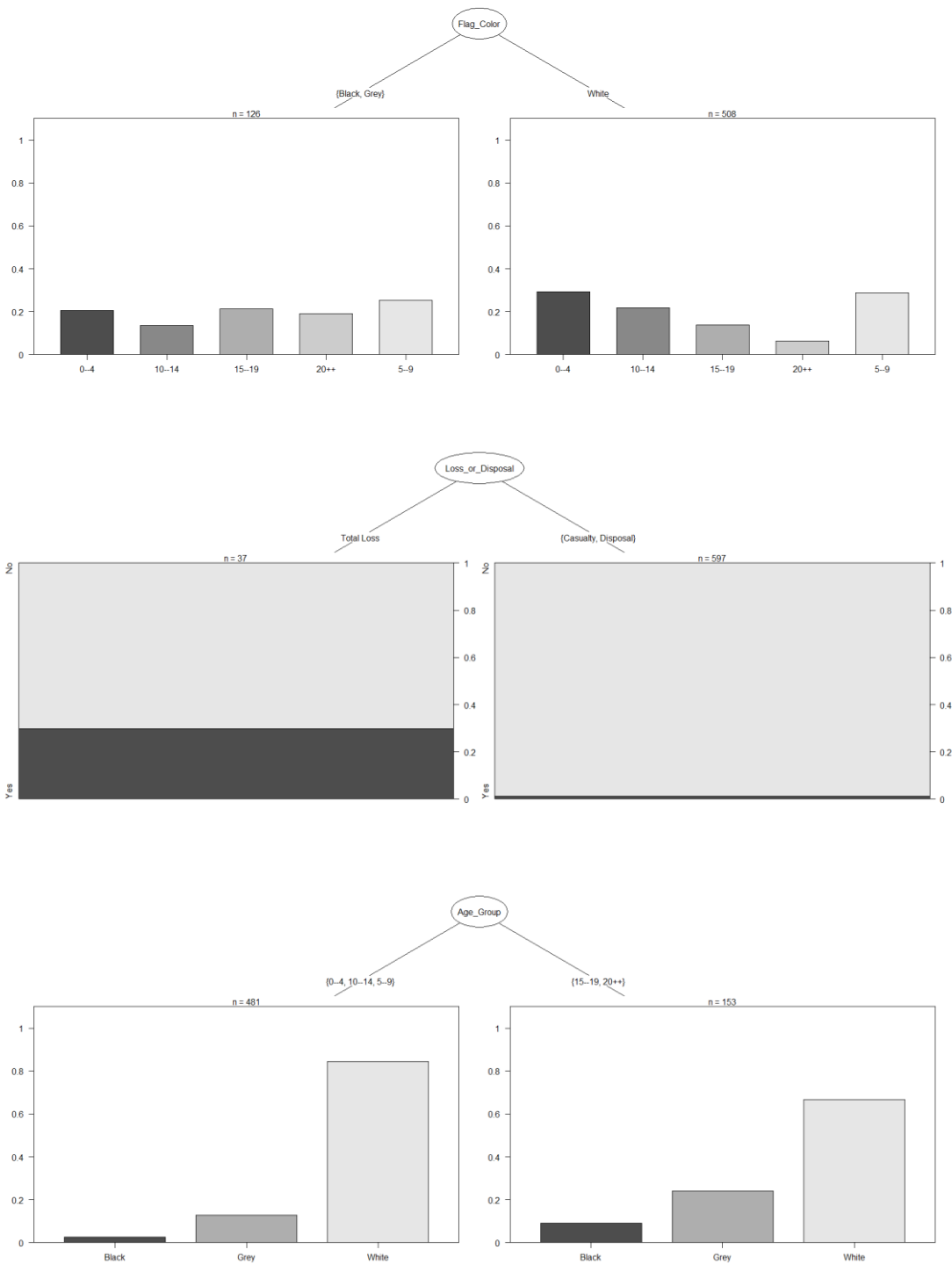
## ΠΑΡΑΡΤΗΜΑ Β : ΣΥΝΟΛΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΔΕΝΤΡΩΝ

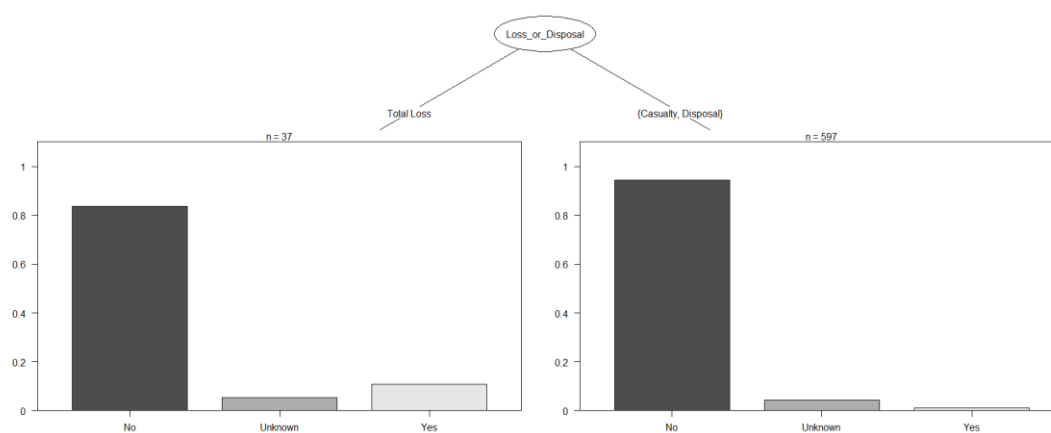
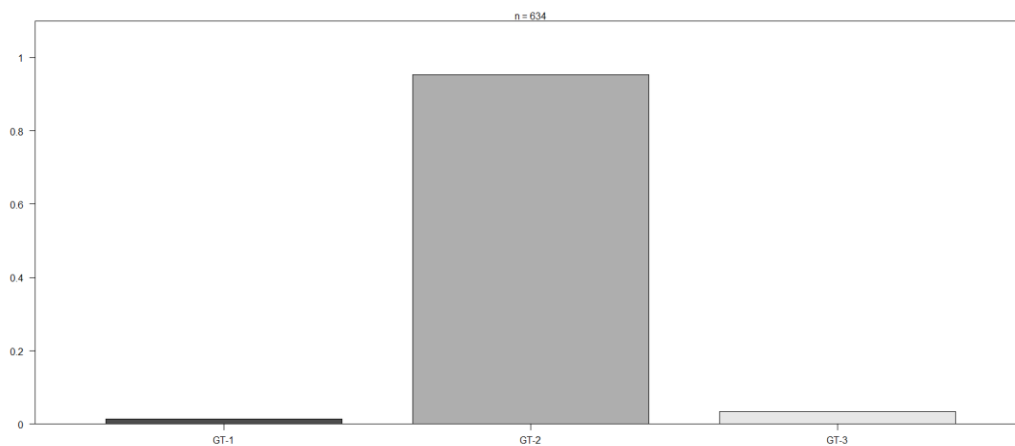
- Δέντρα για το συνολικό Δείγμα



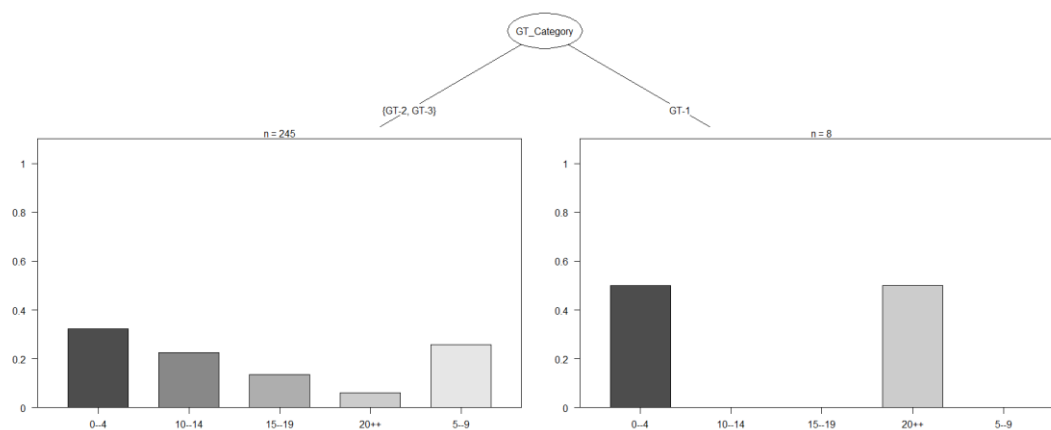


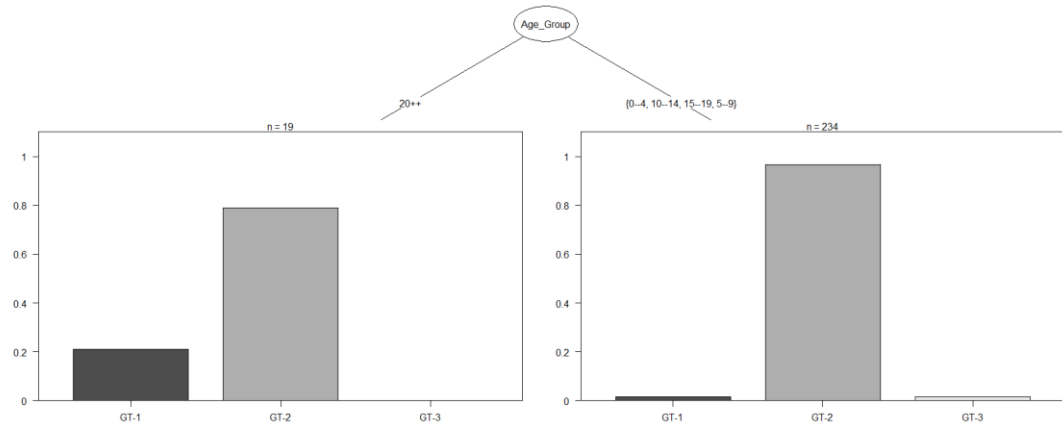
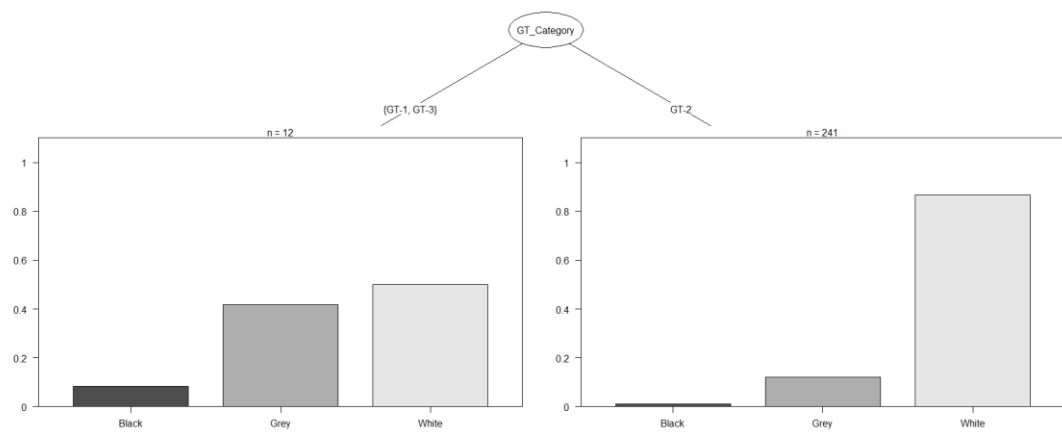
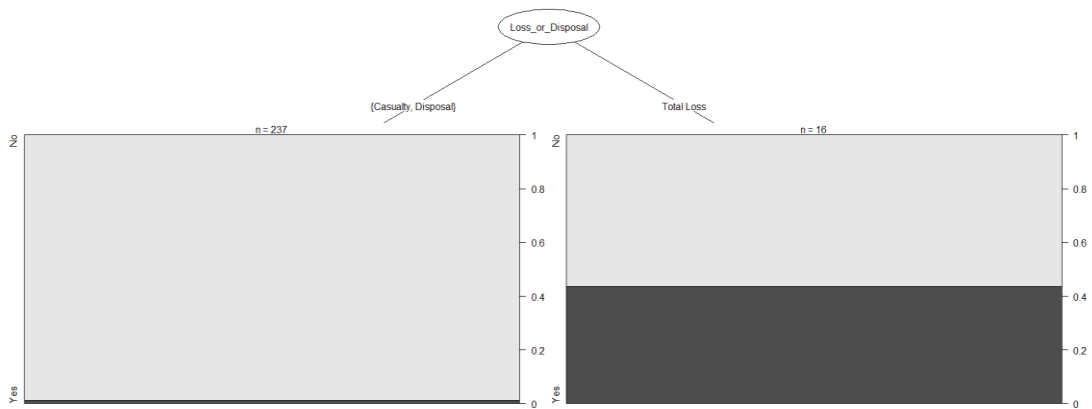
- Δέντρα για το δείγμα των ατυχημάτων Σύγκρουσης



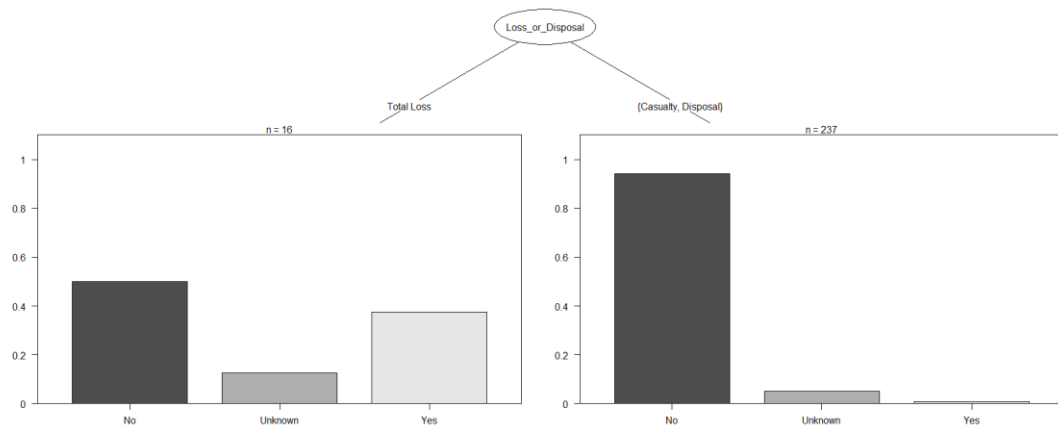


- Δέντρα για το δείγμα των ατυχημάτων Επαφής

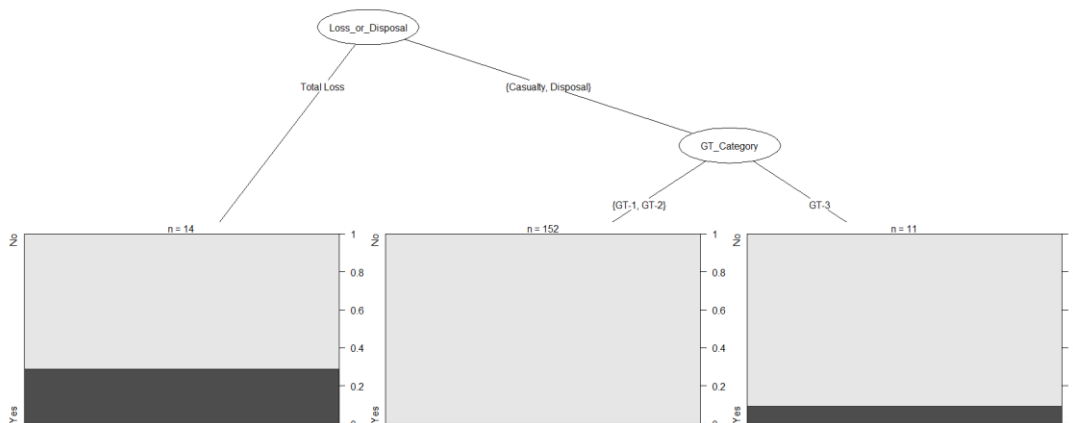
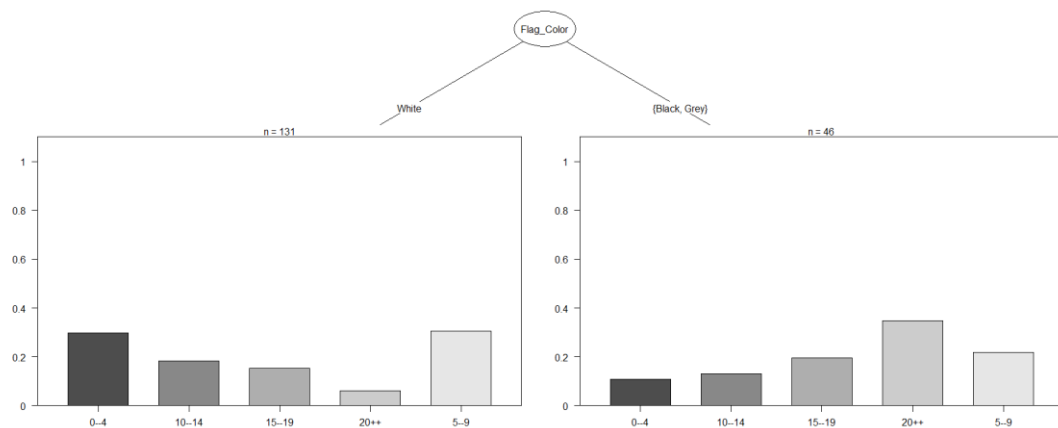


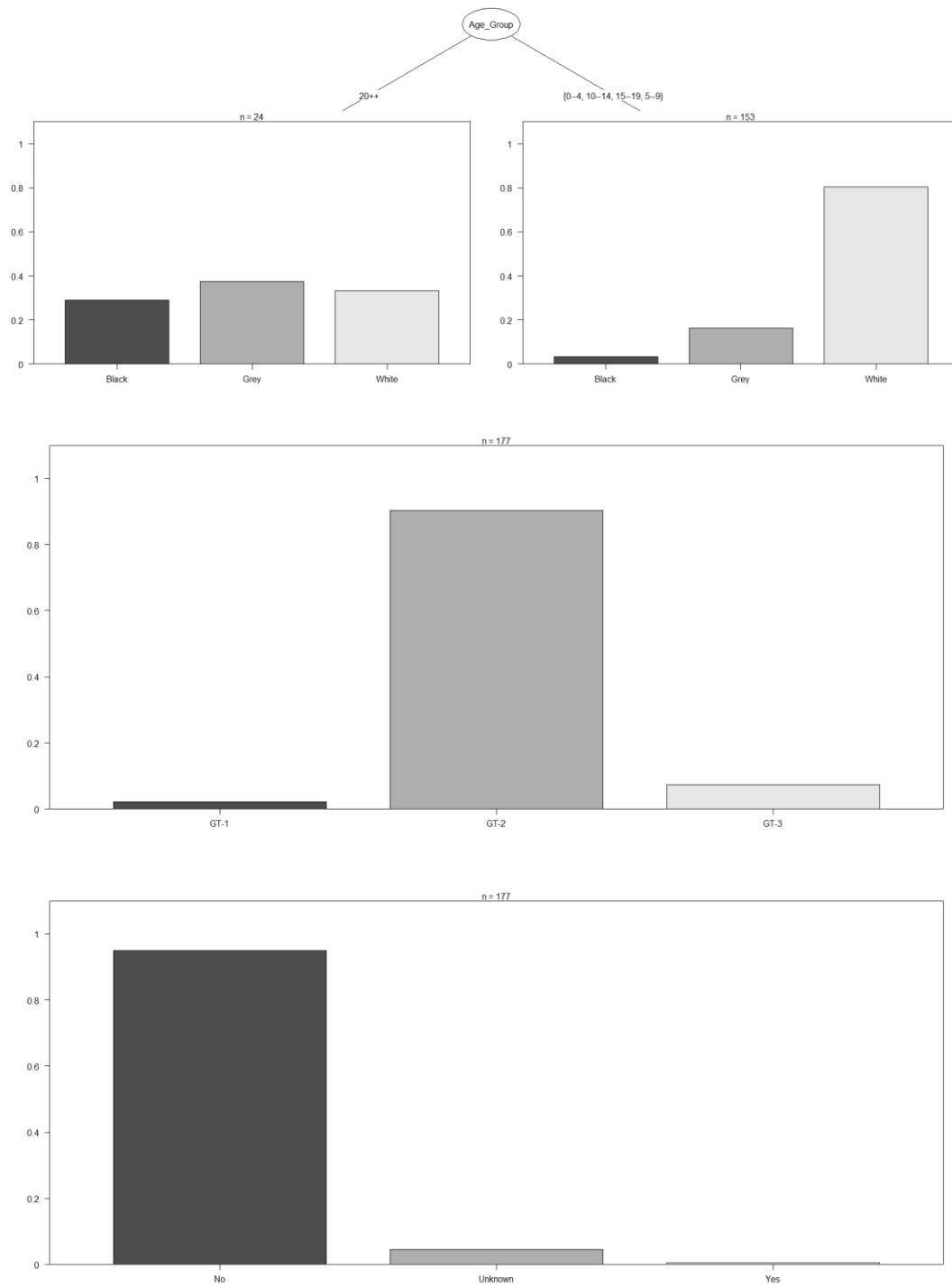




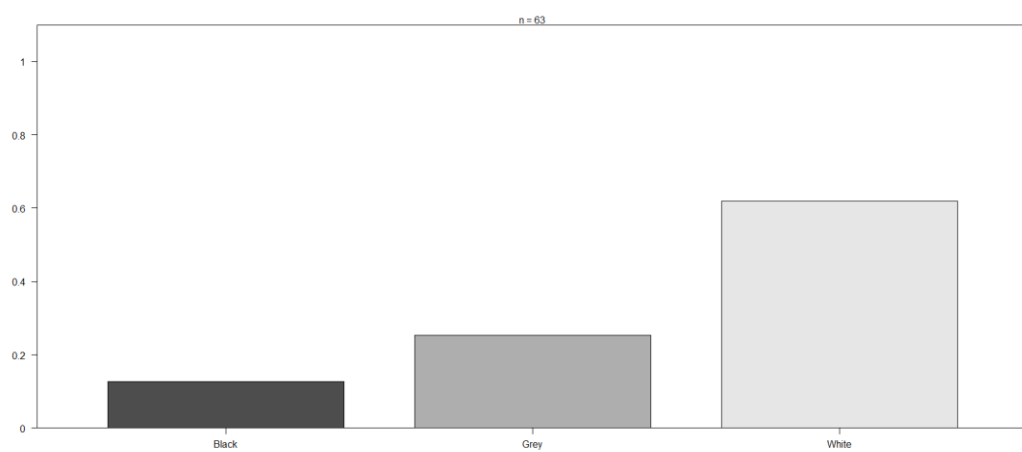
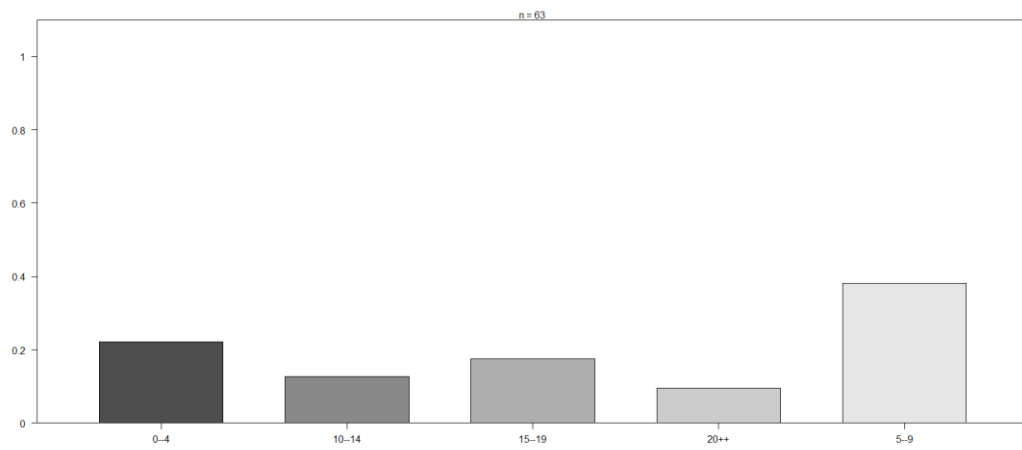


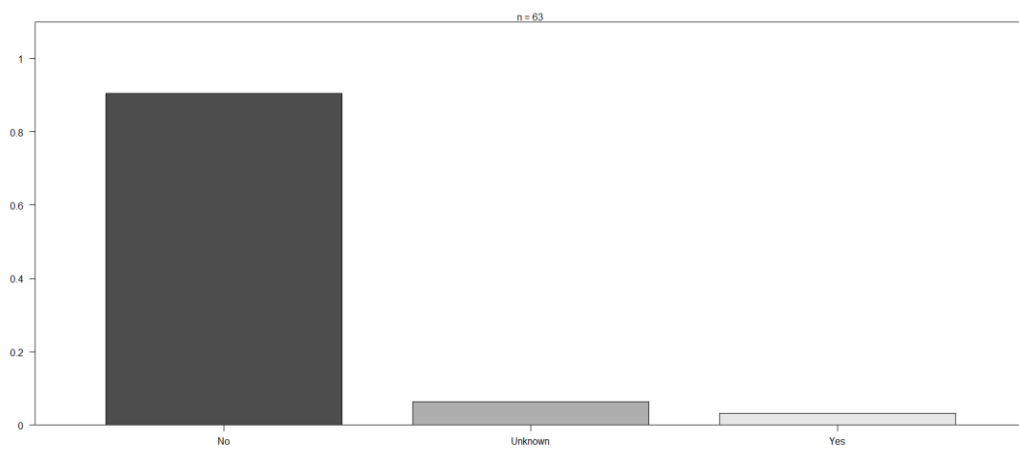
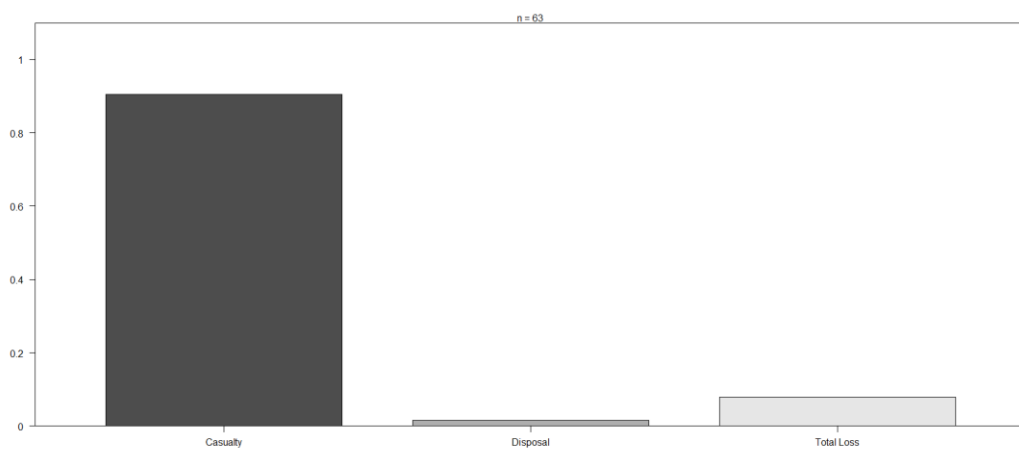
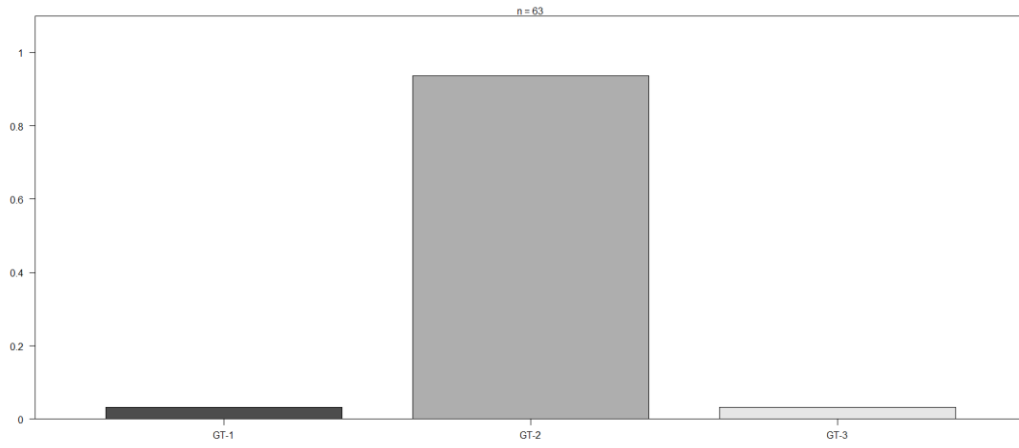
- Δέντρα για το δείγμα των ατυχημάτων Φωτιάς ή Έκρηξης



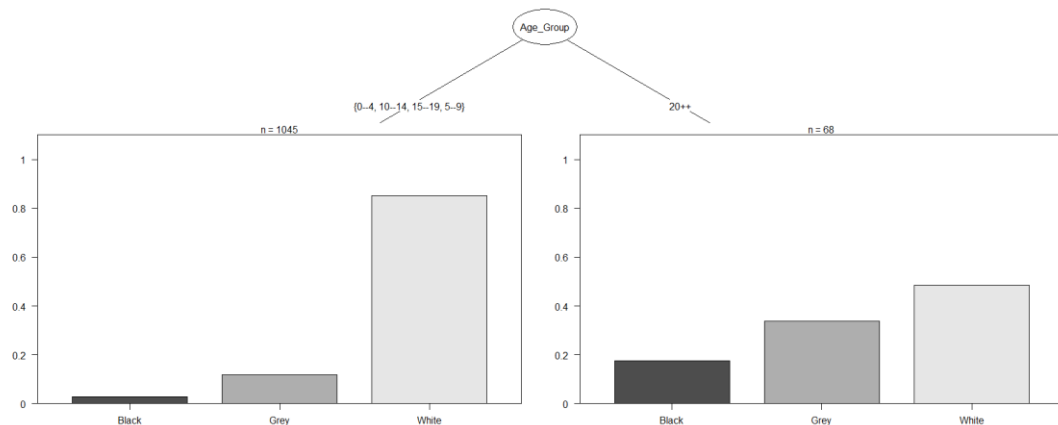
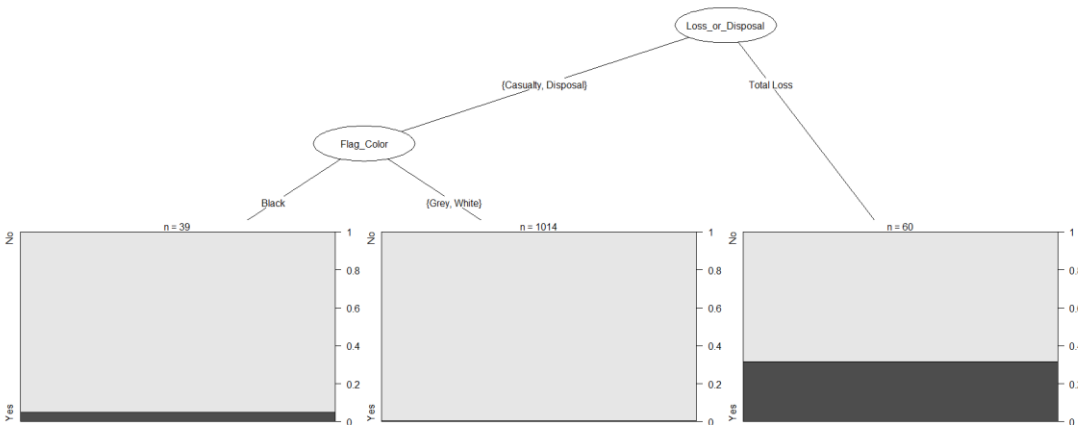
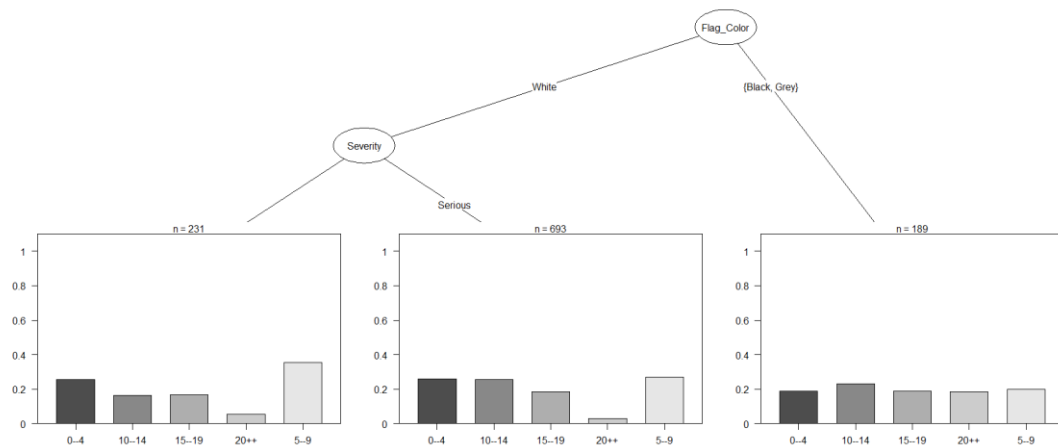


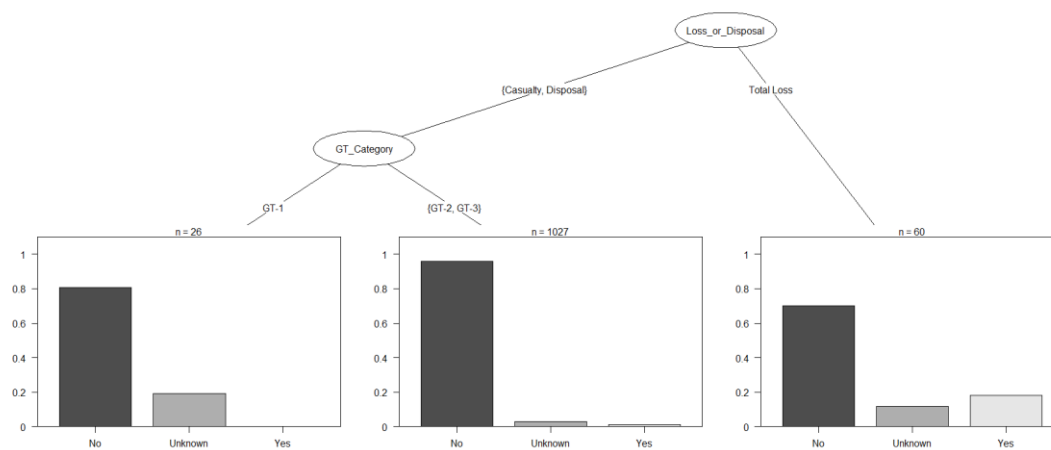
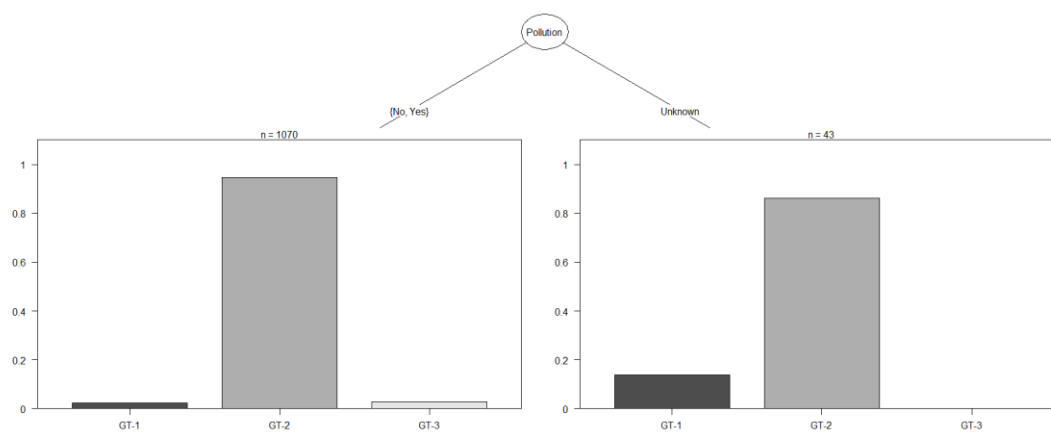
- Δέντρα για το δείγμα των ατυχημάτων Βύθισης



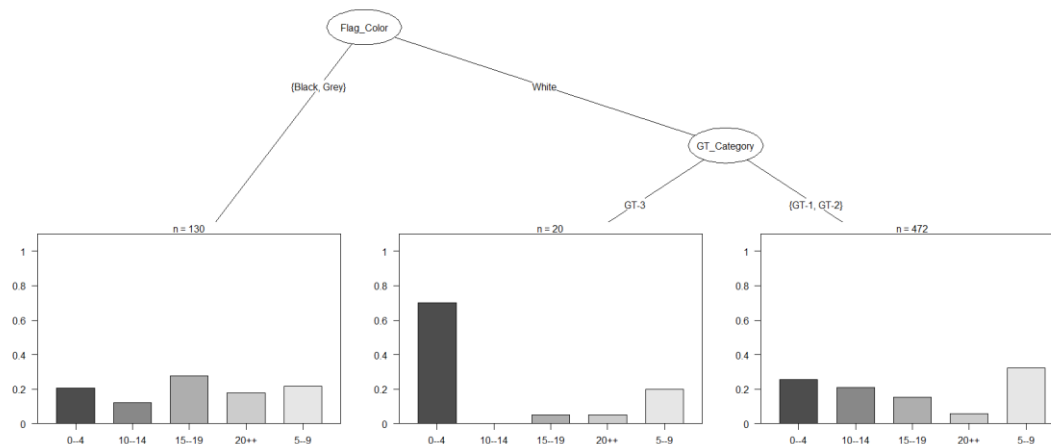


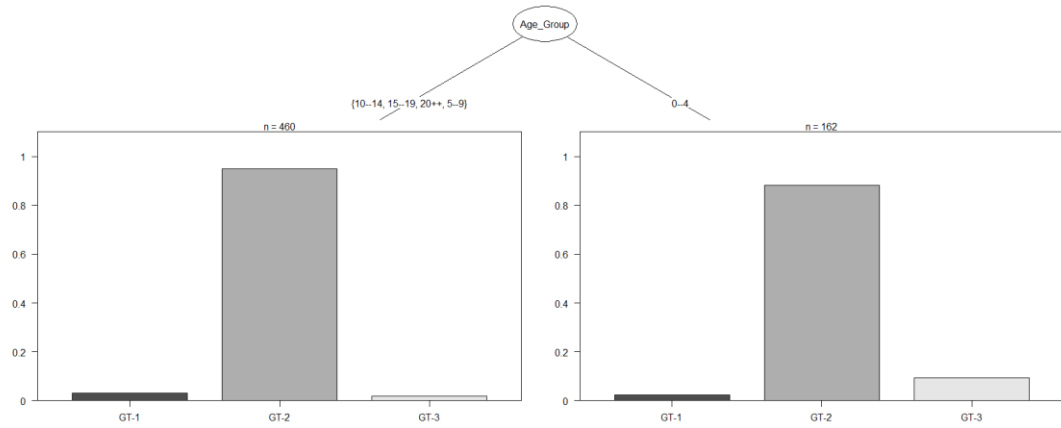
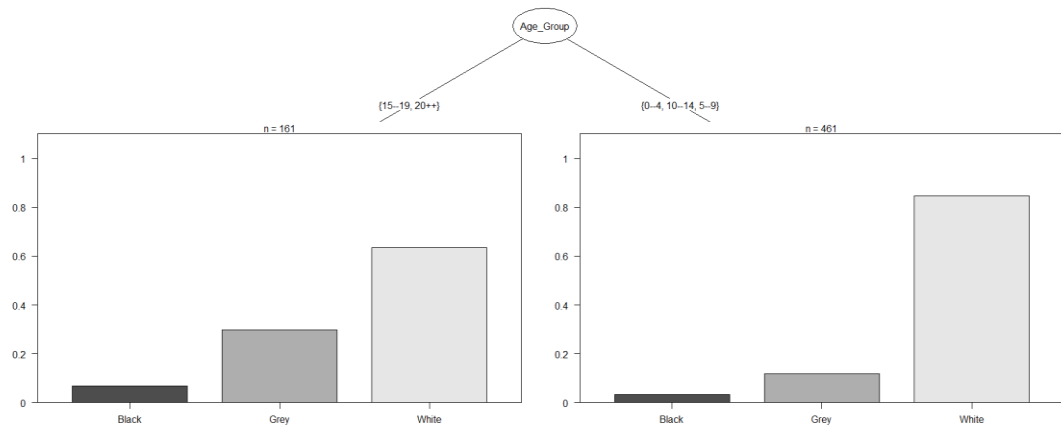
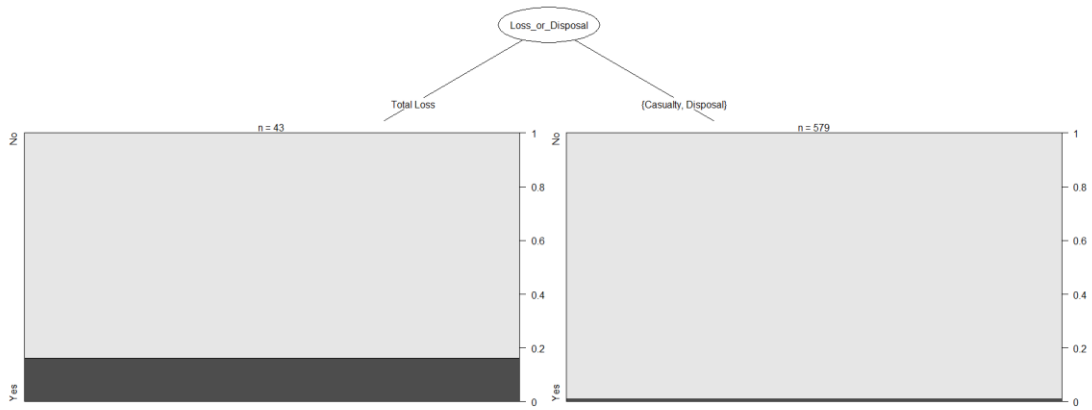
- Δέντρα για το δείγμα των ατυχημάτων με ζημιά στη Γάστρα ή το Μηχανοστάσιο

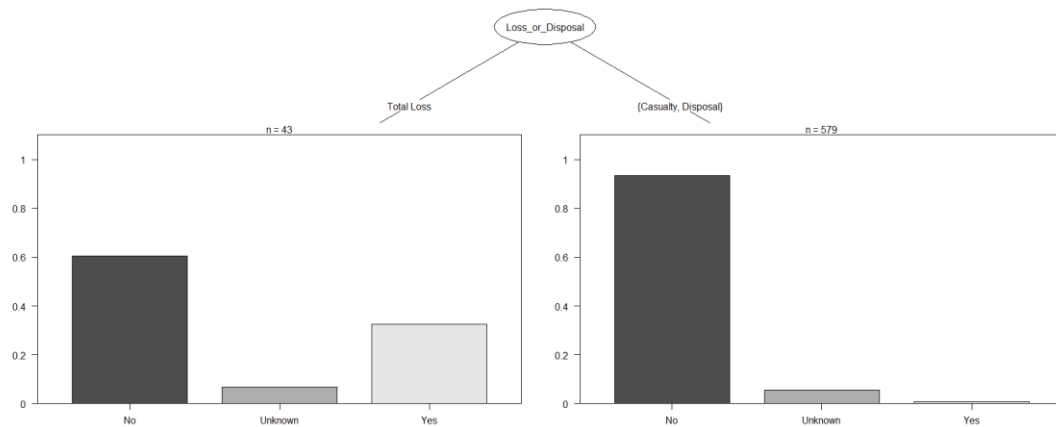




- Δέντρα για το δείγμα των ατυχημάτων που τα πλοία καταστράφηκαν ή ακινητοποιήθηκαν







Για να παραχθούν τα παραπάνω αποτελέσματα στο πρόγραμμα R, χρησιμοποιήθηκε ο παρακάτω κώδικας για το συνολικό δείγμα αλλά και για τους 6 τύπους ατυχημάτων.

```
> library(party)
```

```
> data11 <- read.delim(file.choose(),header=T)
```

```
> str(data11)
```

```
'data.frame': 2862 obs. of 8 variables:
 $ Age_Group      : Factor w/ 5 levels "0--4","10--14",...: 1 1 1 1 1 1
1 1 1 1 ...
 $ Flag_Color     : Factor w/ 3 levels "Black","Grey",...: 3 3 3 3 3 3
3 3 3 3 ...
 $ GT_Category    : Factor w/ 3 levels "GT-1","GT-2",...: 2 2 2 2 2 2
2 2 2 ...
 $ Loss_or_Disposal: Factor w/ 3 levels "Casualty","Disposal",...: 1 1 1
1 1 1 1 1 1 3 ...
 $ Pollution      : Factor w/ 3 levels "No","Unknown",...: 1 1 1 1 1 1
1 1 1 1 ...
 $ Severity       : Factor w/ 2 levels "", "serious": 2 2 2 2 2 2
2 2 2 ...
 $ Fatalities     : Factor w/ 2 levels "No","Yes": 1 1 1 1 1 1
1 1 1 1 2
...
 $ Casualty_Type  : Factor w/ 6 levels "Col","Con","Fir",...: 1 1 1 1 1
1 1 1 1 1 ...
> m11 <- ctree(Casualty_Type~.,data=data11)
> m11
```

Conditional inference tree with 2 terminal nodes

```
Response: Casualty_Type
Inputs: Age_Group, Flag_Color, GT_Category, Loss_or_Disposal,
Pollution, Severity, Fatalities
Number of observations: 2862
```

```
1) Flag_Color == {white}; criterion = 0.999, statistic = 35.588
```

```
2)* weights = 2309
```

```
1) Flag_Color == {Black, Grey}
```

```
3)* weights = 553
```

```
> plot(m11, inner_panel=node_inner(m11, pval=FALSE),
id=FALSE),terminal_panel=node_barplot(m11, id=FALSE))
```



## ΠΑΡΑΡΤΗΜΑ Γ : ΣΗΜΑΙΕΣ ΠΛΟΙΩΝ

Ακολουθεί η λίστα από την Paris MOU που αφορά το χρώμα της σημαίας του κάθε πλοίου ανάλογα με τη χώρα στην οποία ανήκει.

Effective from 1 July 2018

### White List

RANK	FLAG	INSPECTIONS 2015-2017	DETENTIONS 2015-2017	BLACK TO GREY LIMIT	GREY TO WHITE LIMIT	EXCESS FACTOR
<b>WHITE LIST</b>						
1	France	258	0	25	11	-1.89
2	Cayman Islands, UK	438	2	40	21	-1.84
3	Netherlands	3,028	35	236	188	-1.83
4	Denmark	1,189	12	98	68	-1.79
5	United Kingdom	1,194	13	99	69	-1.76
6	Norway	1,429	17	116	84	-1.75
7	Italy	1,087	12	90	62	-1.74
8	Bahamas	2,239	30	177	136	-1.73
9	Isle of Man, UK	694	7	60	37	-1.71
10	Singapore	1,887	27	151	113	-1.68
11	Sweden	316	2	30	14	-1.66
12	Marshall Islands	4,022	68	309	254	-1.65
13	Hong Kong, China	1,997	31	159	121	-1.64
14	Belgium	220	1	22	9	-1.59
15	Germany	588	8	52	30	-1.52
16	Ireland	120	0	13	3	-1.41
17	Cyprus	1,921	44	153	116	-1.36
18	Gibraltar, UK	716	16	62	38	-1.32
19	Liberia	4,105	110	315	260	-1.29
20	Finland	408	7	38	20	-1.26
21	Luxembourg	199	2	20	8	-1.24
22	China	198	2	20	7	-1.24
23	Malta	4,681	135	357	298	-1.23
24	Greece	903	22	76	50	-1.18
25	Bermuda, UK	243	4	24	10	-1.04
26	Estonia	90	0	11	2	-1.04
27	Japan	117	1	13	3	-0.88
28	Philippines	150	2	16	5	-0.86
29	Antigua and Barbuda	2,866	110	224	178	-0.84
30	Portugal	762	26	65	41	-0.76
31	Poland	102	1	12	2	-0.67
32	Latvia	100	1	12	2	-0.64
33	Barbados	341	11	32	16	-0.55
34	Saudi Arabia	59	0	8	0	-0.40
35	Panama	6,110	334	461	394	-0.34
36	Korea, Republic of	82	1	10	1	0.28
37	Croatia	105	2	12	3	0.26
38	Faroe Islands, DK	251	9	25	10	-0.24
39	Turkey	1,133	59	94	65	-0.19
40	Spain	156	5	17	5	0.05

Paris MoU  
on Port State Control

Effective from 1 July 2018

## Grey List

RANK	FLAG	INSPECTIONS 2015-2017	DETENTIONS 2015-2017	BLACK TO GREY LIMIT	GREY TO WHITE LIMIT	EXCESS FACTOR
<b>GREY LIST</b>						
41	Kazakhstan	66	1	9	1	0.04
42	Algeria	83	2	10	1	0.06
43	United States of America	203	9	21	8	0.10
44	Russian Federation	1,182	71	98	68	0.11
45	Thailand	53	1	7	0	0.12
46	Lithuania	123	5	14	3	0.15
47	Switzerland	119	6	13	3	0.27
48	Libya	32	1	5	0	0.28
49	Egypt	43	2	6	0	0.34
50	Saint Vincent and the Grenadines	557	36	49	29	0.36
51	Morocco	57	3	8	0	0.37
52	Iran, Islamic Republic of	104	6	12	2	0.37
53	Curacao	119	7	13	3	0.37
54	Azerbaijan	45	3	6	0	0.48
55	Bulgaria	33	3	5	0	0.62
56	Lebanon	74	7	9	1	0.72
57	Tuvalu	34	4	5	0	0.77
58	Albania	66	7	9	1	0.80
59	Tunisia	40	5	6	0	0.85
60	India	68	8	9	1	0.91

Paris MoU  
on Port State Control

Effective from 1 July 2018

## Black List

RANK	FLAG	INSPECTIONS 2015-2017	DETENTIONS 2015-2017	BLACK TO GREY LIMIT	GREY TO WHITE LIMIT	EXCESS FACTOR
<b>BLACK LIST</b>						
61	Belize	399	40	37	Medium Risk	1.24
62	Cook Islands	414	43	38		1.36
63	Vanuatu	251	28	25		1.38
64	Cambodia	161	22	17		1.85
65	Saint Kitts and Nevis	266	35	26	Medium to High Risk	1.98
66	Sierra Leone	292	46	28		2.80
67	Ukraine	94	18	11		2.98
68	Palau	180	36	19	High Risk	3.77
69	Moldova, Republic of	480	89	43		3.90
70	Tanzania United Rep.	259	53	25	Very High Risk	4.16
71	Togo	422	85	39		4.34
72	Comoros	282	60	27		4.46
73	Congo, Republic of the	104	28	12		5.33

Paris MoU  
on Port State Control