

# Προγραμματιστική Εργασία

Χαρά Τσίρκα, Πρόδρομος Αβραμίδης, Γεώργιος Γεροντίδης

May 3, 2024



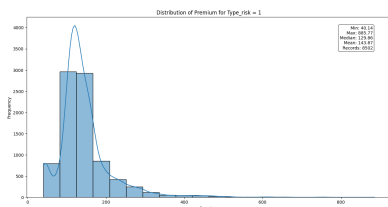
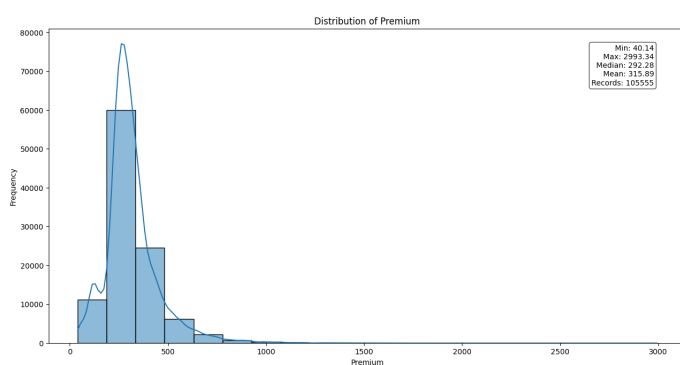
## Εξόρυξη δεδομένων

Εαρινό εξάμηνο 2023-2024

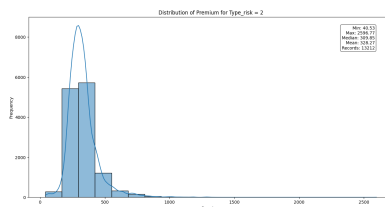
Καθηγητής: Βασιλακόπουλος Μιχαήλ

# Εισαγωγή

Η εργασία μας επικεντρώνεται στην πρόβλεψη του κόστους ασφάλισης μηχανοκίνητων οχημάτων. Η ανάλυση αυτή αποτελεί ένα κρίσιμο ζήτημα στον τομέα της ασφάλισης, καθώς επιτρέπει στους ασφαλιστές να προσδιορίζουν με μεγαλύτερη ακρίβεια τα ασφαλιστικά ασφάλιστρα, λαμβάνοντας υπόψη διάφορους παράγοντες που επηρεάζουν το κόστος. Η διαδικασία της εργασίας ξεκινά με την προ-επεξεργασία των δεδομένων, κατά την οποία πραγματοποιήθηκε εξερευνητική ανάλυση (exploratory analysis) για τον προσδιορισμό των κριτηρίων διαχωρισμού των δεδομένων. Κατά τη διάρκεια αυτής της ανάλυσης, μετρήθηκε ο βαθμός επίδρασης κάθε χαρακτηριστικού (feature) του συνόλου δεδομένων στα αποτελέσματα. Με τη βοήθεια διαγραμμάτων, καταφέραμε να επιλέξουμε τον κατάλληλο διαχωρισμό των δεδομένων για περαιτέρω ανάλυση. Στη συνέχεια, η εργασία θα προχωρήσει στη δημιουργία και αξιολόγηση των μοντέλων πρόβλεψης, λαμβάνοντας υπόψη την εισοδο των χρηστών, ενώ θα ακολουθήσει η οπτικοποίηση και η αξιολόγηση των αποτελεσμάτων.



(a) Motorbikes



(b) Vans

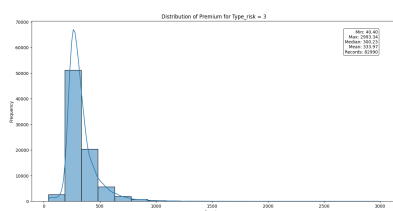


Figure 2: Passenger Cars

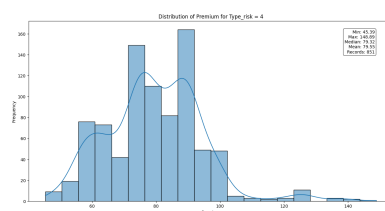


Figure 3: Agricultural Vehicles

Figure 4: Comparison of Different Vehicle Types