



**Πανεπιστήμιο Πειραιώς**

**Τμήμα Ψηφιακών Συστημάτων**

**Π.Μ.Σ. “Πληροφορικά Συστήματα και Υπηρεσίες”**

**ΜΕΤΑΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ**  
**ΑΞΙΟΛΟΓΗΣΗ ΠΙΣΤΟΛΗΠΤΙΚΗΣ ΙΚΑΝΟΤΗΤΑΣ ΔΑΝΕΙΟΛΗΠΤΩΝ**  
**ΜΕΣΩ ΣΤΑΤΙΣΤΙΚΩΝ ΤΕΧΝΙΚΩΝ**

**Τμήμα : Ψηφιακών συστημάτων**

**Πρόγραμμα : Προηγμένα Πληροφορικά Συστήματα**

**Ακαδημαϊκό έτος 2025-2026**

**Σπουδαστής: Βιτζηλαίος Γεώργιος**

**Επιβλέπων καθηγητής: Μιχαήλ Φιλιππάκης**

**Μάιος 2026**

**Πανεπιστήμιο Πειραιώς. Κάτοχος όλων των δικαιωμάτων University of Piraeus.**

**All rights reserved.**

**Συγγραφέας: Γεώργιος Βιτζηλαίος**

## ΣΕΛΙΔΑ ΕΓΚΥΡΟΤΗΤΑΣ

Όνοματεπώνυμο Φοιτητή: Γεώργιος Βιτζηλαίος

Τίτλος Μεταπτυχιακής Διπλωματικής Εργασίας: «Αξιολόγηση πιστοληπτικής ικανότητας δανειοληπτών μέσω στατιστικών τεχνικών»

Η παρούσα Μεταπτυχιακή Διπλωματική Εργασία υποβάλλεται ως μερική εκπλήρωση των απαιτήσεων του Προγράμματος Μεταπτυχιακών Σπουδών “Πληροφοριακά Συστήματα & Υπηρεσίες” του Τμήματος Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς και εγκρίθηκε στις 26/05/2026 από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή

Επιβλέπων (Τμήμα Ψηφιακών Συστημάτων, Πανεπιστήμιο Πειραιώς)

Μιχαήλ Φιλιππάκης Διευθυντής του ΠΜΣ Πληροφοριακά συστήματα και υπηρεσίες

Μέλος Εξεταστικής Επιτροπής: Μαρία Χαλκίδη καθηγήτρια ΠΜΣ

Μέλος Εξεταστικής Επιτροπής: Δημοσθένης Κυριαζής καθηγητής ΠΜΣ

## ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ ΑΥΘΕΝΤΙΚΟΤΗΤΑΣ

Ο Γεώργιος Βιτζηλαίος γνωρίζοντας τις συνέπειες της λογοκλοπής, δηλώνω υπεύθυνα ότι η παρούσα εργασία με τίτλο «Αξιολόγηση πιστοληπτικής ικανότητας δανειοληπτών μέσω στατιστικών τεχνικών» αποτελεί προϊόν αυστηρά προσωπικής εργασίας και όλες οι πηγές που έχω χρησιμοποιήσει, έχουν δηλωθεί κατάλληλα στις βιβλιογραφικές παραπομπές και αναφορές. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο ή/και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. Επιπλέον δηλώνω υπεύθυνα ότι η συγκεκριμένη Μεταπτυχιακή Διπλωματική Εργασία έχει συγγραφεί από εμένα προσωπικά και δεν έχει υποβληθεί ούτε έχει αξιολογηθεί στο πλαίσιο κάποιου άλλου μεταπτυχιακού ή προπτυχιακού τίτλου σπουδών, στην Ελλάδα ή στο εξωτερικό. Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου. Σε κάθε περίπτωση, αναληθούς ή ανακριβούς δηλώσεως, υπόκειμαι στις συνέπειες που προβλέπονται τις διατάξεις που προβλέπει η Ελληνική και Κοινοτική Νομοθεσία περί πνευματικής ιδιοκτησίας.

Ο ΔΗΛΩΝ

Όνοματεπώνυμο: Γεώργιος Βιτζηλαίος

Αριθμός Μητρώου: 228213055005

Υπογραφή: Γεώργιος Βιτζηλαίος

## **ΕΥΧΑΡΙΣΤΙΕΣ**

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον επιβλέποντα καθηγητή μου για την καθοδήγηση και την πολύτιμη υποστήριξή του καθ' όλη τη διάρκεια εκπόνησης της παρούσας πτυχιακής εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω την οικογένειά μου για την υπομονή και τη στήριξή της, καθώς και όλους όσους συνέβαλαν με οποιονδήποτε τρόπο στην ολοκλήρωση της εργασίας αυτής.

Περιεχόμενα	
<i>Περίληψη</i> .....	6
<i>Abstract</i> .....	7
1. <i>Εισαγωγή</i> .....	8
2: Πιστωτικός κίνδυνος και καθυστέρηση δανείων .....	12
2.1 Έννοια Πιστωτικού Κινδύνου .....	12
2.2 Καθυστέρηση δανείων (Loan Delinquency).....	13
2.3 Παραδοσιακά μοντέλα Credit Scoring .....	15
2.4 Διαστρωμάτωση κινδύνου (Risk Segmentation).....	16
2.5 Η διαχρονική εξέλιξη του πιστωτικού κινδύνου στο εξωτερικό και στην Ελλάδα .....	18
2.6 Κανονιστικό και νομικό πλαίσιο στην Ελλάδα.....	21
3. <i>Μεθοδολογία έρευνας</i> .....	23
3.1 Σκοπός της εμπειρικής ανάλυσης .....	23
3.2 Περιγραφή του συνόλου των δεδομένων .....	23
3.3 Προεπεξεργασία δεδομένων .....	24
3.4 Στατιστικές τεχνικές ανάλυσης.....	25
3.5 Εργαλεία ανάλυσης δεδομένων .....	26
3.6 Exploratory Data Analysis (EDA).....	27
3.7 Ανάλυση συσχέτισης μεταβλητών .....	27
3.8 Ανάλυση πιθανοτήτων και πίνακες συχνοτήτων .....	28
3.9 Κατηγοριοποίηση δανειοληπτών (Risk Segmentation).....	28
3.10 Διαδικασία Υλοποίησης της Ανάλυσης.....	29
4.1 <i>Εμπειρική εφαρμογή μοντέλου αξιολόγησης πιστωτικού κινδύνου</i> .....	30
4.2 Υλοποίηση της Ανάλυσης σε Περιβάλλον Python.....	30
4.3 Φόρτωση και Έλεγχος Δεδομένων .....	31
4.4 Περιγραφικά ευρήματα της ανάλυσης .....	31
4.5 Οπτικοποίηση και ερμηνεία δεδομένων .....	32
4.6 Υπολογισμός πιθανότητας αθέτησης (Probability of Default).....	33
4.7 Ανάπτυξη μοντέλου κατηγοριοποίησης κινδύνου .....	33
4.8 Επικύρωση και αξιολόγηση του μοντέλου.....	34
4.9 Περιορισμοί της ανάλυσης .....	34

5.1 Αποτελέσματα .....	36
5.2 Κατανομή Πιστωτικού Κινδύνου.....	36
5.3 Ανάλυση συσχέτισης μεταβλητών .....	37
5.4 Κατανομή ηλικίας δανειοληπτών .....	38
5.5 Κατανομή ποσού δανείου .....	39
5.6 Κατανομή διάρκειας δανείου.....	40
5.7 Σχέση ηλικίας και πιστωτικού κινδύνου .....	41
5.9 Σχέση διάρκειας δανείου και πιστωτικού κινδύνου .....	43
5.10 Συνολική ανάλυση αποτελεσμάτων .....	43
5.11 Ανάλυση Odds Ratio .....	45
5.12 Παράδειγμα εφαρμογής.....	46
5.13 Σύγκριση αποτελεσμάτων με θεωρητικό πλαίσιο .....	47
5.14 Πρακτικές επιπτώσεις για τις τράπεζες .....	47
5.15 Περιορισμοί της ανάλυσης .....	48
5.16 Σύγκριση με μεθόδους μηχανικής μάθησης .....	48
6. Συμπεράσματα .....	50
Βιβλιογραφία .....	54
Παράρτημα .....	57

## Περίληψη

Η παρούσα πτυχιακή εργασία έχει ως αντικείμενο την ανάλυση και πρόβλεψη της καθυστέρησης αποπληρωμής δανείων μέσω στατιστικών μεθόδων, χωρίς τη χρήση τεχνικών μηχανικής μάθησης. Στόχος της εργασίας είναι η διερεύνηση των παραγόντων που επηρεάζουν τον πιστωτικό κίνδυνο και η ανάπτυξη ενός απλού μοντέλου αξιολόγησης δανειοληπτών. Για την ανάλυση χρησιμοποιήθηκε το σύνολο δεδομένων German Credit Dataset, το οποίο περιλαμβάνει πληροφορίες για 1000 δανειολήπτες και διάφορα χαρακτηριστικά που σχετίζονται με την οικονομική και κοινωνική τους κατάσταση, καθώς και με τα χαρακτηριστικά των δανείων τους. Η μεθοδολογία βασίστηκε στην περιγραφική στατιστική, την ανάλυση συσχέτισης και την ανάλυση πιθανοτήτων αθέτησης (Probability of Default). Τα αποτελέσματα έδειξαν ότι η διάρκεια και το ποσό του δανείου αποτελούν βασικούς παράγοντες που επηρεάζουν την πιθανότητα καθυστέρησης αποπληρωμής, ενώ η ηλικία εμφανίζει μικρότερη επίδραση. Επιπλέον, αναπτύχθηκε ένα σύστημα κατηγοριοποίησης κινδύνου, το οποίο διαχωρίζει τους δανειολήπτες σε χαμηλού, μεσαίου και υψηλού κινδύνου, παρουσιάζοντας σημαντικές διαφορές στα ποσοστά αθέτησης μεταξύ των ομάδων. Η εργασία καταδεικνύει ότι η στατιστική ανάλυση μπορεί να αποτελέσει ένα αποτελεσματικό και διαφανές εργαλείο αξιολόγησης πιστωτικού κινδύνου, παρέχοντας χρήσιμα συμπεράσματα για τη λήψη αποφάσεων στον τραπεζικό τομέα.

*Λέξεις-κλειδιά:* «πιστωτικός κίνδυνος», «καθυστέρηση δανείων», “probability of default”, «στατιστική ανάλυση», “credit scoring”.

## *Abstract*

This thesis focuses on the analysis and prediction of loan repayment delays using statistical methods, without the use of machine learning techniques. The main objective is to investigate the factors affecting credit risk and to develop a simple borrower risk assessment model. The analysis is based on the German Credit Dataset, which includes data for 1,000 borrowers along with various attributes related to their financial and demographic characteristics, as well as loan features. The methodology relies on descriptive statistics, correlation analysis, and probability of default (PD) estimation. The results indicate that loan duration and loan amount are key factors influencing the probability of default, while age has a less significant impact. Furthermore, a risk segmentation model was developed, classifying borrowers into low, medium, and high-risk categories, revealing substantial differences in default rates across these groups. The findings demonstrate that statistical analysis can serve as an effective and transparent tool for credit risk evaluation, providing valuable insights for decision-making in the banking sector.

*Keywords:* "credit risk", "loan delinquency", "probability of default", "statistical analysis", "credit scoring".

## 1. Εισαγωγή

Η αξιολόγηση του πιστοληπτικού κινδύνου αποτελεί έναν από τους σημαντικότερους και πιο κρίσιμους τομείς της σύγχρονης τραπεζικής λειτουργίας, καθώς συνδέεται άμεσα με τη σταθερότητα, τη βιωσιμότητα και την κερδοφορία των χρηματοπιστωτικών ιδρυμάτων. Φαίνεται μάλιστα ότι η δυνατότητα μιας τράπεζας να εκτιμά με ακρίβεια την πιθανότητα αδυναμίας αποπληρωμής των υποχρεώσεων από τους δανειολήπτες της επηρεάζει ουσιαστικά την ποιότητα του δανειακού της χαρτοφυλακίου, τη διαχείριση της ρευστότητας και τη συνολική της ανθεκτικότητα απέναντι σε οικονομικές διαταραχές και περιόδους κρίσης. Στο πλαίσιο αυτό, η αποτελεσματική διαχείριση του πιστωτικού κινδύνου δεν αποτελεί μόνο εργαλείο προστασίας των τραπεζών από ενδεχόμενες ζημίες, αλλά και βασικό μηχανισμό διασφάλισης της χρηματοπιστωτικής σταθερότητας σε μακροοικονομικό επίπεδο.

Παράλληλα, η συνεχώς αυξανόμενη πολυπλοκότητα των χρηματοπιστωτικών αγορών, σε συνδυασμό με τις αυστηρότερες εποπτικές απαιτήσεις που επιβλήθηκαν μετά τη διεθνή χρηματοπιστωτική κρίση του 2008, ανέδειξαν ακόμη περισσότερο τη σημασία της αξιόπιστης αξιολόγησης των δανειοληπτών. Πλέον, οι τράπεζες καλούνται να εφαρμόζουν πιο σύνθετες και τεκμηριωμένες διαδικασίες ελέγχου της πιστοληπτικής ικανότητας των πελατών τους, προκειμένου να περιορίσουν τον κίνδυνο δημιουργίας μη εξυπηρετούμενων δανείων και να συμμορφώνονται με τα κανονιστικά πλαίσια που έχουν θεσπιστεί από τις εποπτικές αρχές. Μάλιστα, η ανάγκη αυτή κατέστη ιδιαίτερα έντονη στις οικονομίες που επλήγησαν από υψηλά ποσοστά ιδιωτικού χρέους και τραπεζικής αστάθειας, όπου η αύξηση των επισφαλειών επηρέασε σημαντικά τη δυνατότητα χρηματοδότησης της πραγματικής οικονομίας.

Στο σύγχρονο τραπεζικό περιβάλλον, η χρήση ποσοτικών και στατιστικών μεθόδων για την εκτίμηση της πιθανότητας αθέτησης υποχρεώσεων αποτελεί αναπόσπαστο μέρος της διαδικασίας λήψης πιστοδοτικών αποφάσεων. Φαίνεται ότι μέσω της ανάλυσης οικονομικών και δημογραφικών χαρακτηριστικών των δανειοληπτών, οι τράπεζες επιδιώκουν να δημιουργήσουν αξιόπιστα μοντέλα αξιολόγησης, τα οποία επιτρέπουν την κατηγοριοποίηση των πελατών ανάλογα με το επίπεδο κινδύνου που παρουσιάζουν. Τα μοντέλα αυτά εν τέλει συμβάλλουν τόσο

στη μείωση της αβεβαιότητας κατά τη διαδικασία χορήγησης δανείων, όσο και στη βελτιστοποίηση της κατανομής των διαθέσιμων κεφαλαίων.

Σκοπός της παρούσας μεταπτυχιακής εργασίας είναι η διερεύνηση και η πρόβλεψη της καθυστέρησης αποπληρωμής δανείων μέσω της εφαρμογής στατιστικών τεχνικών ανάλυσης δεδομένων, χωρίς τη χρήση προηγμένων αλγορίθμων μηχανικής μάθησης. Η έρευνα επικεντρώνεται στην κατανόηση των βασικών παραγόντων που επηρεάζουν τον πιστωτικό κίνδυνο και στη δημιουργία ενός απλού, ερμηνεύσιμου και λειτουργικού υποδείγματος αξιολόγησης δανειοληπτών. Επιπλέον, η επιλογή της συγκεκριμένης προσέγγισης βασίζεται στην ανάγκη ανάπτυξης μοντέλων που να χαρακτηρίζονται από διαφάνεια και δυνατότητα εύκολης ερμηνείας των αποτελεσμάτων τους, γεγονός που είναι ιδιαίτερα σημαντικό τόσο για τα τραπεζικά στελέχη όσο και για τις εποπτικές αρχές.

Η εργασία επιχειρεί να απαντήσει σε βασικά ερευνητικά ερωτήματα που αφορούν τη σχέση μεταξύ των χαρακτηριστικών των δανειοληπτών και της πιθανότητας εμφάνισης καθυστερήσεων στις αποπληρωμές των δανειακών τους υποχρεώσεων. Παράλληλα, εξετάζεται κατά πόσο συγκεκριμένοι κοινωνικοοικονομικοί και χρηματοοικονομικοί παράγοντες μπορούν να θεωρηθούν αξιόπιστοι δείκτες πρόβλεψης πιστωτικού κινδύνου. Τέλος, μέσα από την ανάλυση πραγματικών δεδομένων επιδιώκεται η εξαγωγή χρήσιμων συμπερασμάτων σχετικά με τη συμπεριφορά των δανειοληπτών και τη δυνατότητα δημιουργίας αποτελεσματικών συστημάτων credit scoring με χρήση κλασικών στατιστικών μεθόδων.

Η σημασία της παρούσας μελέτης έγκειται τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Πιο συγκεκριμένα, από θεωρητικής πλευράς, συμβάλλει στην κατανόηση των βασικών εννοιών και μεθοδολογιών που σχετίζονται με τη διαχείριση πιστωτικού κινδύνου και την αξιολόγηση δανειακών χαρτοφυλακίων. Από πρακτικής πλευράς, η ανάπτυξη ενός κατανοητού και εφαρμόσιμου μοντέλου αξιολόγησης μπορεί να αποτελέσει χρήσιμο εργαλείο υποστήριξης αποφάσεων για τα χρηματοπιστωτικά ιδρύματα, ενισχύοντας τη δυνατότητά τους να εντοπίζουν έγκαιρα πελάτες υψηλού κινδύνου και να περιορίζουν το επίπεδο των επισφαλειών τους. Επιπλέον, η χρήση ερμηνεύσιμων στατιστικών μοντέλων προάγει τη διαφάνεια

στη διαδικασία λήψης αποφάσεων και συμβάλλει στη δικαιότερη και αντικειμενικότερη αξιολόγηση των δανειοληπτών.

Η μεθοδολογία που ακολουθείται στην παρούσα εργασία βασίζεται στη συλλογή και ανάλυση δεδομένων δανειοληπτών, καθώς και στην εφαρμογή στατιστικών τεχνικών για την εκτίμηση της πιθανότητας αθέτησης υποχρεώσεων. Η επεξεργασία των δεδομένων πραγματοποιείται με χρήση κατάλληλων εργαλείων ανάλυσης και προγραμματιστικών τεχνικών σε περιβάλλον Python, με στόχο τη δημιουργία ενός λειτουργικού υποδείγματος πιστοληπτικής αξιολόγησης. Επιπλέον, η ανάλυση περιλαμβάνει στάδια προεπεξεργασίας των δεδομένων, διερεύνησης των μεταβλητών και εφαρμογής στατιστικών μεθόδων πρόβλεψης καθώς και κατηγοριοποίησης.

Η παρούσα εργασία οργανώνεται σε έξι βασικά κεφάλαια, τα οποία καλύπτουν σταδιακά το θεωρητικό υπόβαθρο, τη μεθοδολογία, την εμπειρική ανάλυση, τα αποτελέσματα και τα τελικά συμπεράσματα. Αρχικά, το εισαγωγικό μέρος (Κεφάλαια Περίληψη και Εισαγωγή) παρουσιάζει συνοπτικά το αντικείμενο της μελέτης, το ερευνητικό πρόβλημα και τη σημασία της αξιολόγησης του πιστωτικού κινδύνου στο σύγχρονο τραπεζικό περιβάλλον. Εδώ τίθενται οι βασικοί στόχοι της εργασίας και η γενική προσέγγιση που ακολουθείται. Στο δεύτερο κεφάλαιο αναλύεται το θεωρητικό υπόβαθρο του πιστωτικού κινδύνου και της καθυστέρησης δανείων. Παρουσιάζονται βασικές έννοιες όπως το credit scoring, η διαστρωμάτωση κινδύνου και η εξέλιξη των σχετικών μοντέλων, ενώ γίνεται αναφορά τόσο σε διεθνείς πρακτικές όσο και στο κανονιστικό πλαίσιο που ισχύει στην Ελλάδα. Το τρίτο κεφάλαιο αφορά τη μεθοδολογία της έρευνας και περιγράφει αναλυτικά τη διαδικασία ανάλυσης των δεδομένων. Συγκεκριμένα παρουσιάζεται το dataset που χρησιμοποιήθηκε, τα στάδια προεπεξεργασίας, οι στατιστικές τεχνικές που εφαρμόστηκαν, τα εργαλεία ανάλυσης καθώς και η συνολική ερευνητική διαδικασία που ακολουθήθηκε για την αξιολόγηση του πιστωτικού κινδύνου. Το τέταρτο κεφάλαιο επικεντρώνεται στην εμπειρική εφαρμογή της ανάλυσης. Περιγράφεται η υλοποίηση σε περιβάλλον Python, η επεξεργασία των δεδομένων, η παραγωγή στατιστικών ευρημάτων, η οπτικοποίηση, καθώς και ο υπολογισμός της πιθανότητας αθέτησης και η ανάπτυξη ενός απλού μοντέλου κατηγοριοποίησης κινδύνου. Παράλληλα εξετάζεται η αξιολόγηση και οι περιορισμοί της ανάλυσης. Στο πέμπτο

κεφάλαιο παρουσιάζονται και ερμηνεύονται τα αποτελέσματα της έρευνας. Αναλύεται η κατανομή του πιστωτικού κινδύνου και των βασικών μεταβλητών, εξετάζονται οι σχέσεις μεταξύ χαρακτηριστικών και κινδύνου, υπολογίζεται το odds ratio και γίνεται σύγκριση των ευρημάτων με το θεωρητικό πλαίσιο. Επιπλέον, συζητούνται οι πρακτικές επιπτώσεις για τον τραπεζικό τομέα, οι περιορισμοί της ανάλυσης και γίνεται μια σύντομη σύγκριση με προσεγγίσεις μηχανικής μάθησης. Τέλος, το έκτο κεφάλαιο συνοψίζει τα βασικά συμπεράσματα της μελέτης, επισημαίνοντας τη σημασία της στατιστικής ανάλυσης στην αξιολόγηση του πιστωτικού κινδύνου και τη συμβολή της εργασίας στη λήψη τραπεζικών αποφάσεων. Συμπληρωματικά, η εργασία ολοκληρώνεται με τη βιβλιογραφία και το παράρτημα, όπου παρουσιάζονται πρόσθετα στοιχεία και υποστηρικτικό υλικό.

## 2: Πιστωτικός κίνδυνος και καθυστέρηση δανείων

### 2.1 Έννοια Πιστωτικού Κινδύνου

Ο πιστωτικός κίνδυνος (Credit Risk) αποτελεί έναν από τους θεμελιώδεις κινδύνους του χρηματοπιστωτικού συστήματος και συνδέεται άμεσα με τη δυνατότητα ενός δανειολήπτη να τηρήσει τις οικονομικές του υποχρεώσεις απέναντι σε έναν πιστωτικό οργανισμό (Spruchl'áková et al., 2015). Ειδικότερα, ορίζεται ως η πιθανότητα ο οφειλέτης να μην καταβάλει έγκαιρα ή πλήρως τις προβλεπόμενες δόσεις ενός δανείου, οδηγώντας είτε σε καθυστέρηση πληρωμών είτε σε πλήρη αθέτηση της σύμβασης. Μάλιστα, η έννοια αυτή αποκτά ιδιαίτερη σημασία για τις εμπορικές τράπεζες, καθώς η χορήγηση δανείων αποτελεί βασικό πυλώνα της λειτουργίας και της κερδοφορίας τους.

Η ύπαρξη πιστωτικού κινδύνου δεν επηρεάζει μόνο τη χρηματοοικονομική σταθερότητα ενός τραπεζικού ιδρύματος, αλλά δύναται να προκαλέσει ευρύτερες επιπτώσεις στο οικονομικό σύστημα. Ένα αυξημένο ποσοστό μη εξυπηρετούμενων δανείων περιορίζει τη ρευστότητα των τραπεζών, μειώνει την ικανότητά τους να χρηματοδοτούν επιχειρήσεις και νοικοκυριά και ενδέχεται να οδηγήσει σε γενικότερη οικονομική επιβράδυνση. Η διεθνής χρηματοπιστωτική κρίση του 2008 ανέδειξε με ιδιαίτερα έντονο τρόπο τη σημασία της αποτελεσματικής διαχείρισης του πιστωτικού κινδύνου (Financial Stability Board, 2009), καθώς η ανεπαρκής αξιολόγηση της πιστοληπτικής ικανότητας των δανειοληπτών συνέβαλε σημαντικά στην αποσταθεροποίηση του τραπεζικού συστήματος.

Για τον λόγο αυτό, τα χρηματοπιστωτικά ιδρύματα αναπτύσσουν σύνθετα μοντέλα αξιολόγησης και παρακολούθησης του πιστωτικού κινδύνου, αξιοποιώντας τόσο ποσοτικά όσο και ποιοτικά δεδομένα. Η αξιολόγηση αυτή βασίζεται κυρίως σε τρεις θεμελιώδεις παραμέτρους: την Πιθανότητα Αθέτησης (Probability of Default – PD), την Έκθεση κατά την Αθέτηση (Exposure at Default – EAD) και τη Ζημία σε Περίπτωση Αθέτησης (Loss Given Default – LGD) (Li et al., 2021). Πρόκειται για παραμέτρους που αποτελούν τον πυρήνα των σύγχρονων συστημάτων διαχείρισης κινδύνου και χρησιμοποιούνται ευρέως στα διεθνή κανονιστικά πλαίσια, όπως το Basel II και το Basel III, με στόχο την ενίσχυση της κεφαλαιακής επάρκειας και της ανθεκτικότητας των τραπεζών (Dedeloudis et al., 2025).

Παράλληλα, η πιθανότητα Αθέτησης (PD) εκφράζει την πιθανότητα ένας δανειολήπτης να αθετήσει τις υποχρεώσεις του εντός συγκεκριμένου χρονικού ορίζοντα, συνήθως ενός έτους. Η Έκθεση κατά την Αθέτηση (EAD) αναφέρεται στο συνολικό ποσό στο οποίο εκτίθεται η τράπεζα τη στιγμή της αθέτησης, ενώ η Ζημία σε Περίπτωση Αθέτησης (LGD) αποτυπώνει το ποσοστό της ζημίας που τελικά θα υποστεί το πιστωτικό ίδρυμα μετά την αφαίρεση πιθανών ανακτήσεων (Dedeloudis et al., 2025), όπως εμπράγματα εξασφαλίσεις ή ρευστοποίηση περιουσιακών στοιχείων.

Στην παρούσα εργασία δίνεται ιδιαίτερη έμφαση στην Πιθανότητα Αθέτησης (PD), καθώς αποτελεί βασικό δείκτη εκτίμησης της πιστοληπτικής συμπεριφοράς των δανειοληπτών και σημαντικό εργαλείο λήψης αποφάσεων για τα τραπεζικά ιδρύματα. Εξάλλου, η ακριβής πρόβλεψη της πιθανότητας αθέτησης συμβάλλει στη μείωση των επισφαλειών, στη βελτίωση της ποιότητας του χαρτοφυλακίου δανείων και στη διατήρηση της χρηματοπιστωτικής σταθερότητας (Arhinful et al., 2025), ενώ επιτρέπει στις τράπεζες να εφαρμόζουν πιο αποτελεσματικές πολιτικές τιμολόγησης, χορήγησης πιστώσεων και διαχείρισης κεφαλαίων, ενισχύοντας τόσο τη βιωσιμότητα των ίδιων όσο και την εμπιστοσύνη των επενδυτών και των καταθετών.

## 2.2 Καθυστέρηση δανείων (Loan Delinquency)

Η καθυστέρηση δανείων αποτελεί μία από τις βασικότερες ενδείξεις επιδείνωσης της πιστοληπτικής ικανότητας ενός δανειολήπτη και συνδέεται άμεσα με την αύξηση του πιστωτικού κινδύνου για τα χρηματοπιστωτικά ιδρύματα. Ως καθυστέρηση ορίζεται η αδυναμία του δανειολήπτη να καταβάλει έγκαιρα τις προγραμματισμένες δόσεις ή άλλες συμβατικές υποχρεώσεις εντός του χρονικού διαστήματος που έχει συμφωνηθεί στη δανειακή σύμβαση (Chong, 2021). Η εμφάνιση καθυστερήσεων μπορεί να οφείλεται σε πολλούς παράγοντες, όπως η μείωση του εισοδήματος, η ανεργία, η υπερχρέωση, η αύξηση των επιτοκίων ή ακόμη και γενικότερες οικονομικές κρίσεις που επηρεάζουν την οικονομική δυνατότητα νοικοκυριών και επιχειρήσεων.

Στην τραπεζική πρακτική, οι καθυστερήσεις ταξινομούνται συνήθως με βάση τη χρονική διάρκειά τους, καθώς η διάρκεια της αθέτησης αποτελεί σημαντικό δείκτη της πιθανότητας οριστικής μη αποπληρωμής. Συγκεκριμένα, τα δάνεια διακρίνονται

σε κατηγορίες καθυστέρησης 30 ημερών, 60 ημερών και άνω των 90 ημερών. Οι καθυστερήσεις άνω των 90 ημερών θεωρούνται ιδιαίτερα κρίσιμες, καθώς στις περισσότερες περιπτώσεις συνδέονται με τον χαρακτηρισμό του δανείου ως μη εξυπηρετούμενου ανοίγματος (Non-Performing Exposure – NPE) ή υψηλού πιστωτικού κινδύνου. Σύμφωνα με το πλαίσιο της Basel Committee on Banking Supervision (Dedeloudis et al., 2025), τα δάνεια αυτής της κατηγορίας απαιτούν αυξημένη εποπτική παρακολούθηση και μεγαλύτερες προβλέψεις από τις τράπεζες, προκειμένου να αντιμετωπιστούν πιθανές ζημίες και να διασφαλιστεί η σταθερότητα του τραπεζικού συστήματος.

Η έγκαιρη αναγνώριση πιθανής καθυστέρησης έχει ιδιαίτερη σημασία για τα πιστωτικά ιδρύματα, καθώς επιτρέπει την άμεση λήψη προληπτικών μέτρων πριν η οικονομική κατάσταση του δανειολήπτη επιδεινωθεί περαιτέρω. Μέσω της έγκαιρης ανίχνευσης κινδύνου, οι τράπεζες μπορούν να περιορίσουν τις επισφάλειες, να αναθεωρήσουν τα κριτήρια χορήγησης πιστώσεων και να εφαρμόσουν πιο αποτελεσματικές στρατηγικές διαχείρισης του δανειακού τους χαρτοφυλακίου (Rahman, & Awwal, 2024). Παράλληλα, η συστηματική παρακολούθηση των καθυστερήσεων συμβάλλει στη διατήρηση της κεφαλαιακής επάρκειας και στη βελτίωση της συνολικής χρηματοοικονομικής ανθεκτικότητας ενός οργανισμού.

Επιπλέον, η ανάπτυξη μοντέλων πρόβλεψης της πιστοληπτικής συμπεριφοράς των δανειοληπτών έχει αποκτήσει ιδιαίτερη βαρύτητα στη σύγχρονη τραπεζική επιστήμη. Μέσα από τη χρήση στατιστικών και οικονομετρικών τεχνικών, οι τράπεζες επιχειρούν να εντοπίσουν εγκαίρως πρότυπα συμπεριφοράς που συνδέονται με αυξημένη πιθανότητα αθέτησης. Μάλιστα, φαίνεται ότι η πρόβλεψη της πιστοληπτικής συμπεριφοράς αποτελεί βασικό εργαλείο για τη μείωση του πιστωτικού κινδύνου, καθώς επιτρέπει στα χρηματοπιστωτικά ιδρύματα να λαμβάνουν πιο ορθολογικές αποφάσεις σχετικά με την έγκριση δανείων (Rahman, & Awwal, 2024), την τιμολόγηση του κινδύνου και τη διαχείριση των προβληματικών απαιτήσεων. Με τον τρόπο αυτό, ενισχύεται όχι μόνο η κερδοφορία των τραπεζών, αλλά και η συνολική σταθερότητα του χρηματοπιστωτικού συστήματος.

### 2.3 Παραδοσιακά μοντέλα Credit Scoring

Το credit scoring, μία από τις σημαντικότερες και πλέον διαδεδομένες μεθόδους αξιολόγησης της πιστοληπτικής ικανότητας των δανειοληπτών, στηρίζεται στη χρήση στατιστικών και ποσοτικών τεχνικών, μέσω των οποίων αποδίδεται σε κάθε δανειολήπτη μία αριθμητική βαθμολογία (credit score), η οποία αποτυπώνει την πιθανότητα εμφάνισης μελλοντικής αθέτησης ή καθυστέρησης πληρωμών (Markov et al., 2022). Με άλλα λόγια, το credit scoring επιτρέπει στα πιστωτικά ιδρύματα να εκτιμούν το επίπεδο κινδύνου που ενσωματώνει κάθε αίτηση χρηματοδότησης, συμβάλλοντας στη λήψη πιο ορθολογικών και τεκμηριωμένων αποφάσεων.

Η σημασία του credit scoring έχει ενισχυθεί ιδιαίτερα τις τελευταίες δεκαετίες, λόγω της αυξημένης ανάγκης των τραπεζών για ταχύτητα, ακρίβεια και αντικειμενικότητα κατά τη διαδικασία χορήγησης δανείων. Σε ένα περιβάλλον όπου οι πιστωτικοί οργανισμοί διαχειρίζονται τεράστιο όγκο αιτήσεων από ιδιώτες και επιχειρήσεις, η χρήση αυτοματοποιημένων συστημάτων αξιολόγησης μειώνει σημαντικά το λειτουργικό κόστος και περιορίζει τον κίνδυνο ανθρώπινων σφαλμάτων ή υποκειμενικών κρίσεων (Markov et al., 2022). Παράλληλα, συμβάλλει στη διασφάλιση της ίσης μεταχείρισης των δανειοληπτών, καθώς οι αποφάσεις βασίζονται σε προκαθορισμένα και μετρήσιμα κριτήρια.

Τα παραδοσιακά μοντέλα credit scoring βασίζονται κυρίως σε στατιστικές τεχνικές, όπως η ανάλυση συσχετίσεων, η ανάλυση πιθανοτήτων, η διακριτική ανάλυση και ιδιαίτερα η λογιστική παλινδρόμηση (logistic regression), η οποία χρησιμοποιείται ευρέως για την εκτίμηση της πιθανότητας αθέτησης (Haixiang et al., 2017). Μέσα από αυτές τις μεθόδους αναλύονται χαρακτηριστικά του δανειολήπτη, όπως το εισόδημα, η εργασιακή κατάσταση, η ηλικία, το ιστορικό πληρωμών, το ύψος των υφιστάμενων υποχρεώσεων και η συνολική χρηματοοικονομική συμπεριφορά.

Ιδιαίτερα σημαντικό ρόλο στα παραδοσιακά συστήματα αξιολόγησης διαδραματίζουν τα scorecards, τα οποία αποτελούν δομημένα μοντέλα βαθμολόγησης. Σε αυτά κάθε χαρακτηριστικό του δανειολήπτη λαμβάνει συγκεκριμένη βαθμολογία ή «πόντους», ανάλογα με το επίπεδο κινδύνου που αντιπροσωπεύει (Markov et al., 2022). Για παράδειγμα, ένα σταθερό εισόδημα ή ένα

θετικό ιστορικό αποπληρωμών μπορεί να αυξήσει τη συνολική βαθμολογία, ενώ υψηλές οφειλές ή προηγούμενες καθυστερήσεις να τη μειώσουν. Έτσι, το τελικό άθροισμα οδηγεί στην ταξινόμηση του δανειολήπτη σε κατηγορία χαμηλού, μεσαίου ή υψηλού πιστωτικού κινδύνου, διευκολύνοντας εν τέλει τις τράπεζες στη διαδικασία λήψης αποφάσεων.

Οι στατιστικές μέθοδοι που χρησιμοποιούνται στο credit scoring θεωρούνται ιδιαίτερα σημαντικές όχι μόνο λόγω της προβλεπτικής τους ικανότητας, αλλά και επειδή παρέχουν αποτελέσματα που μπορούν να ερμηνευθούν και να αιτιολογηθούν με σαφήνεια. Μάλιστα η ερμηνευσιμότητα αυτή έχει καθοριστική σημασία τόσο για τις ίδιες τις τράπεζες όσο και για τις εποπτικές αρχές, καθώς επιτρέπει την τεκμηρίωση των πιστοδοτικών αποφάσεων με αντικειμενικά και διαφανή κριτήρια (Chang et al., 2024). Επιπλέον, η δυνατότητα κατανόησης των παραγόντων που επηρεάζουν τη βαθμολογία ενός πελάτη ενισχύει την εμπιστοσύνη μεταξύ τραπεζών και δανειοληπτών και συμβάλλει στη δημιουργία ενός πιο αξιόπιστου και σταθερού χρηματοπιστωτικού περιβάλλοντος.

#### 2.4 Διαστρωμάτωση κινδύνου (Risk Segmentation)

Η διαστρωμάτωση κινδύνου (risk stratification) αποτελεί μία ιδιαίτερα σημαντική διαδικασία στη διαχείριση πιστωτικού κινδύνου, καθώς επιτρέπει στις τράπεζες και στα χρηματοπιστωτικά ιδρύματα να κατηγοριοποιούν τους δανειολήπτες ανάλογα με το επίπεδο κινδύνου που παρουσιάζουν. Ουσιαστικά, πρόκειται για μια διαδικασία ταξινόμησης κατά την οποία οι πελάτες ομαδοποιούνται σε κατηγορίες όπως χαμηλού, μεσαίου και υψηλού πιστωτικού κινδύνου, με βάση συγκεκριμένα οικονομικά, δημογραφικά και συμπεριφορικά χαρακτηριστικά (Bhatt et al., 2023). Η λογική πίσω από αυτή τη διαδικασία είναι ότι δεν παρουσιάζουν όλοι οι δανειολήπτες την ίδια πιθανότητα αθέτησης. Επομένως, η τράπεζα οφείλει να αντιμετωπίζει κάθε περίπτωση με διαφορετικό τρόπο, ανάλογα με το επίπεδο κινδύνου που αναλαμβάνει.

Αναλυτικότερα, η σημασία της διαστρωμάτωσης κινδύνου γίνεται ιδιαίτερα εμφανής σε περιόδους οικονομικής αβεβαιότητας, όπου η πιθανότητα καθυστερήσεων και μη εξυπηρετούμενων δανείων αυξάνεται. Για παράδειγμα, ένας πελάτης με σταθερό εισόδημα, χαμηλό επίπεδο υφιστάμενου δανεισμού και

συνεπές ιστορικό πληρωμών είναι πιθανό να ενταχθεί σε κατηγορία χαμηλού κινδύνου. Αντίθετα, ένας δανειολήπτης με ασταθή επαγγελματική κατάσταση, συχνές καθυστερήσεις ή υψηλή επιβάρυνση χρέους ενδέχεται να ταξινομηθεί σε κατηγορία υψηλού κινδύνου. Μέσα από αυτή τη διαφοροποίηση, οι τράπεζες μπορούν να κατανοήσουν καλύτερα τη συμπεριφορά των πελατών τους και να περιορίσουν την πιθανότητα μελλοντικών ζημιών (Bhatt et al., 2023).

Η διαδικασία της διαστρωμάτωσης βασίζεται σε στατιστικές και ποσοτικές τεχνικές ανάλυσης δεδομένων. Συχνά χρησιμοποιούνται πίνακες συχνότητας, αναλογίες αθέτησης (default rates), στατιστικοί δείκτες όπως το odds ratio, καθώς και μεταβλητές που εμφανίζουν στατιστική σημαντικότητα ως προς την πρόβλεψη της αθέτησης. Μέσα από την ανάλυση αυτή εντοπίζονται τα χαρακτηριστικά που διαφοροποιούν τους συνεπείς πελάτες από εκείνους που εμφανίζουν αυξημένη πιθανότητα καθυστέρησης (Bhandary & Ghosh, 2025). Για παράδειγμα, μεταβλητές όπως η αναλογία χρέους προς εισόδημα, η διάρκεια εργασιακής απασχόλησης ή το ιστορικό προηγούμενων δανείων μπορούν να αποτελέσουν κρίσιμους δείκτες για την αξιολόγηση του κινδύνου.

Αναμφισβήτητα, η εφαρμογή της διαστρωμάτωσης κινδύνου προσφέρει σημαντικά πλεονεκτήματα στις τράπεζες τόσο σε επιχειρησιακό όσο και σε στρατηγικό επίπεδο. Μέσω αυτής, τα πιστωτικά ιδρύματα μπορούν να προσαρμόζουν τις πολιτικές χρηματοδότησης ανάλογα με το προφίλ κάθε πελάτη (Bhatt et al., 2023). Έτσι, ένας δανειολήπτης χαμηλού κινδύνου μπορεί να λάβει ευνοϊκότερο επιτόκιο ή μεγαλύτερο όριο χρηματοδότησης, καθώς θεωρείται περισσότερο αξιόπιστος. Αντίθετα, για πελάτες υψηλού κινδύνου μπορεί να εφαρμοστούν αυστηρότερα πιστοδοτικά κριτήρια, αυξημένες εξασφαλίσεις ή συχνότερη παρακολούθηση της οικονομικής τους συμπεριφοράς.

Παράλληλα, η διαστρωμάτωση κινδύνου συμβάλλει στη συνολική σταθερότητα του τραπεζικού συστήματος, καθώς η σωστή κατανομή και παρακολούθηση του κινδύνου επιτρέπει στις τράπεζες να διατηρούν υγιές δανειακό χαρτοφυλάκιο και να περιορίζουν τη δημιουργία μη εξυπηρετούμενων δανείων. Επιπλέον, διευκολύνει τη συμμόρφωση με τα διεθνή εποπτικά πρότυπα, όπως τα πλαίσια της Basel Committee on Banking Supervision, τα οποία απαιτούν αποτελεσματικές διαδικασίες αναγνώρισης, μέτρησης και διαχείρισης του

πιστωτικού κινδύνου (Dedeloudis et al., 2025). Τέλος, πέρα από τη στατιστική και τεχνική της διάσταση, η διαστρωμάτωση κινδύνου έχει και μια ουσιαστικά ανθρώπινη πλευρά. Δεν αφορά μόνο αριθμούς και πιθανότητες, αλλά και την κατανόηση της οικονομικής συμπεριφοράς των ανθρώπων μέσα σε ένα συνεχώς μεταβαλλόμενο οικονομικό περιβάλλον (Kennedy et al., 2013). Ένας πελάτης μπορεί να θεωρείται υψηλού κινδύνου όχι απαραίτητα λόγω κακής πρόθεσης, αλλά εξαιτίας εξωτερικών παραγόντων, όπως η απώλεια εργασίας, η μείωση εισοδήματος ή μια οικονομική κρίση. Για τον λόγο αυτό, τα σύγχρονα συστήματα αξιολόγησης επιδιώκουν ολοένα και περισσότερο να συνδυάζουν την ποσοτική ανάλυση με μια πιο ισορροπημένη και ρεαλιστική προσέγγιση της οικονομικής συμπεριφοράς των δανειοληπτών.

Λαμβάνοντας υπόψη τα προαναφερθέντα, καθίσταται σαφές ότι η έγκαιρη πρόβλεψη της πιθανότητας καθυστέρησης είναι κρίσιμη τόσο για τις τράπεζες όσο και για τη σταθερότητα του χρηματοπιστωτικού συστήματος. Μέσω αξιόπιστων συστημάτων αξιολόγησης, οι τράπεζες μπορούν να εντοπίζουν τους δανειολήπτες υψηλού κινδύνου και να λαμβάνουν προληπτικά μέτρα, πράγμα που μπορεί να διασώσει μακροπρόθεσμα τόσο τους δανειολήπτες, όσο και τις τράπεζες (Maruf et al., 2024). Εξάλλου, η σωστή αξιολόγηση του πιστωτικού κινδύνου συμβάλλει στη διαφάνεια, στη δίκαιη μεταχείριση των πελατών και στη βελτίωση της διαδικασίας λήψης αποφάσεων.

## 2.5 Η διαχρονική εξέλιξη του πιστωτικού κινδύνου στο εξωτερικό και στην Ελλάδα

Η έννοια του πιστωτικού κινδύνου έχει εξελιχθεί σημαντικά τις τελευταίες δεκαετίες, ακολουθώντας τις μεταβολές της παγκόσμιας οικονομίας, την ανάπτυξη των χρηματοπιστωτικών αγορών και τις συνεχείς αλλαγές στο τραπεζικό περιβάλλον. Φαίνεται ότι ο τρόπος με τον οποίο οι τράπεζες αντιλαμβάνονται και διαχειρίζονται τον κίνδυνο δεν παραμένει σταθερός, αλλά προσαρμόζεται στις οικονομικές συνθήκες, στις τεχνολογικές εξελίξεις και στις εμπειρίες που αφήνουν οι χρηματοπιστωτικές κρίσεις. Όσο αυξάνεται η πολυπλοκότητα των τραπεζικών προϊόντων και η διασύνδεση των διεθνών αγορών, τόσο μεγαλύτερη γίνεται η ανάγκη για αποτελεσματικότερη αναγνώριση και πρόβλεψη του πιστωτικού κινδύνου.

Αναλυτικότερα, κατά τις δεκαετίες του 1980 και 1990, οι τράπεζες βασιζόνταν κυρίως σε πιο παραδοσιακές μορφές αξιολόγησης των δανειοληπτών (Firdavs & Wang, 2026). Η προσωπική εμπειρία και η κρίση των τραπεζικών στελεχών διαδραμάτιζαν καθοριστικό ρόλο στη λήψη αποφάσεων για την έγκριση ή απόρριψη ενός δανείου. Σε πολλές περιπτώσεις, η αξιολόγηση στηριζόταν σε προσωπικές συνεντεύξεις, στη γενική εικόνα του πελάτη και σε βασικά οικονομικά στοιχεία, χωρίς τη χρήση εξελιγμένων συστημάτων ανάλυσης δεδομένων (Bednarik & Marshall, 2024). Παρότι οι μέθοδοι αυτές επέτρεπαν μια πιο ανθρώπινη και εξατομικευμένη προσέγγιση, παρουσίαζαν σημαντικούς περιορισμούς, καθώς συχνά επηρεάζονταν από υποκειμενικές κρίσεις ή αδυνατούσαν να προβλέψουν με ακρίβεια τον πραγματικό βαθμό κινδύνου.

Με την αύξηση του όγκου των χορηγήσεων, την ταχύτερη κυκλοφορία κεφαλαίων και την παγκοσμιοποίηση των αγορών, οι παραδοσιακές πρακτικές αποδείχθηκαν ανεπαρκείς. Εξάλλου, οι τράπεζες άρχισαν να διαχειρίζονται εκατομμύρια πελάτες και δανειακές συμβάσεις, γεγονός που κατέστησε αναγκαία τη δημιουργία πιο συστηματικών και αυτοματοποιημένων μεθόδων αξιολόγησης. Σε αυτό το πλαίσιο αναπτύχθηκαν τα πρώτα οργανωμένα μοντέλα credit scoring (Markov et al., 2022), τα οποία βασιζόνταν σε στατιστικές τεχνικές και επιχειρούσαν να προβλέψουν την πιθανότητα αθέτησης μέσω αντικειμενικών μεταβλητών και έτσι στις αρχές της δεκαετίας του 2000 η χρήση αυτών των μοντέλων και συστημάτων διαχείρισης κινδύνου έγινε πλέον βασικό στοιχείο της τραπεζικής λειτουργίας (Bednarik & Marshall, 2024). Παράλληλα, τα διεθνή κανονιστικά πλαίσια της Basel Committee on Banking Supervision, και ιδιαίτερα το Basel II, καθιέρωσαν την ανάγκη ποσοτικοποίησης του πιστωτικού κινδύνου μέσω συγκεκριμένων δεικτών (Dedeloudis et al., 2025). Η φιλοσοφία αυτών των πλαισίων βασίστηκε στην άποψη ότι οι τράπεζες δεν έπρεπε να βασιζονται αποκλειστικά στην εμπειρική κρίση, αλλά σε μετρήσιμα και επιστημονικά τεκμηριωμένα δεδομένα.

Από την άλλη, η παγκόσμια χρηματοπιστωτική κρίση του 2008 αποτέλεσε σημείο καμπής για την αντίληψη του πιστωτικού κινδύνου, καθώς η κατάρρευση μεγάλων χρηματοπιστωτικών οργανισμών και η μαζική αύξηση των μη εξυπηρετούμενων δανείων αποκάλυψαν σοβαρές αδυναμίες στα συστήματα αξιολόγησης και εποπτείας των τραπεζών (Blake, 2022). Πολλά ιδρύματα είχαν

υποτιμήσει τον πραγματικό κίνδυνο που συνδεόταν με στεγαστικά και καταναλωτικά δάνεια υψηλού ρίσκου, θεωρώντας λανθασμένα ότι η συνεχής ανάπτυξη των αγορών θα διατηρούσε χαμηλά τα επίπεδα αθέτησης. Η κρίση έδειξε ότι ακόμη και πολύπλοκα μοντέλα μπορούν να αποτύχουν όταν αγνοούνται σημαντικοί παράγοντες, όπως η ψυχολογία των αγορών ή οι γενικότερες μακροοικονομικές συνθήκες. Σήμερα, η αξιολόγηση του πιστωτικού κινδύνου βασίζεται σε έναν συνδυασμό στατιστικών μοντέλων, τεχνολογιών ανάλυσης δεδομένων, εποπτικών απαιτήσεων και ολοκληρωμένων συστημάτων risk management. Οι τράπεζες αξιοποιούν πλέον τεχνικές machine learning, ανάλυση μεγάλων δεδομένων (big data), αυτοματοποιημένα πληροφοριακά συστήματα, αλλά και τεχνητή νοημοσύνη προκειμένου να προβλέψουν με μεγαλύτερη ακρίβεια τη συμπεριφορά των δανειοληπτών (Engan et al., 2026). Παράλληλα, η ανθρώπινη διάσταση εξακολουθεί να παραμένει σημαντική, καθώς πίσω από κάθε αριθμό και κάθε στατιστικό μοντέλο βρίσκεται μια πραγματική οικονομική και κοινωνική κατάσταση.

Η εξέλιξη του πιστωτικού κινδύνου στην Ελλάδα παρουσιάζει ιδιαίτερο ενδιαφέρον, καθώς επηρεάστηκε καθοριστικά από την οικονομική κρίση της περιόδου 2009–2018. Κατά τα χρόνια που προηγήθηκαν της κρίσης, το ελληνικό τραπεζικό σύστημα γνώρισε ταχεία ανάπτυξη και σημαντική αύξηση των χορηγήσεων δανείων προς νοικοκυριά και επιχειρήσεις (Στεργίου, 2008). Αναμφισβήτητα, η ένταξη της Ελλάδας στη ζώνη του ευρώ, τα χαμηλότερα επιτόκια και η αυξημένη ρευστότητα διευκόλυναν σημαντικά την πρόσβαση στον τραπεζικό δανεισμό. Επιπλέον, πολλοί πολίτες απέκτησαν στεγαστικά και καταναλωτικά δάνεια, ενώ αρκετές επιχειρήσεις βασίστηκαν στον τραπεζικό δανεισμό για την επέκτασή τους.

Ωστόσο, σε αρκετές περιπτώσεις η πιστοδοτική πολιτική υπήρξε πιο χαλαρή από όσο επέτρεπαν οι πραγματικές δυνατότητες αποπληρωμής των δανειοληπτών. Η αισιοδοξία της εποχής και η πεποίθηση ότι η οικονομική ανάπτυξη θα συνεχιζόταν δημιούργησαν ένα περιβάλλον αυξημένου δανεισμού και υποτίμησης του κινδύνου. Μετά το 2009, η βαθιά ύφεση, η εκτόξευση της ανεργίας, οι μειώσεις μισθών και η συρρίκνωση των εισοδημάτων οδήγησαν μεγάλο μέρος νοικοκυριών και επιχειρήσεων σε αδυναμία αποπληρωμής των υποχρεώσεών τους (Ζωγραφάκης & Σπαθής, 2011).

Κατά την περίοδο 2010–2016, τα μη εξυπηρετούμενα δάνεια στην Ελλάδα αυξήθηκαν δραματικά και αποτέλεσαν μία από τις σοβαρότερες προκλήσεις για το ελληνικό τραπεζικό σύστημα. Χιλιάδες πολίτες δυσκολεύονταν να ανταποκριθούν ακόμη και σε βασικές δανειακές υποχρεώσεις, ενώ πολλές επιχειρήσεις οδηγήθηκαν σε οικονομική ασφυξία ή και σε κλείσιμο. Η κατάσταση αυτή δεν είχε μόνο οικονομικές αλλά και κοινωνικές επιπτώσεις, καθώς συνδέθηκε με αύξηση της ανασφάλειας, κοινωνική πίεση και μείωση της εμπιστοσύνης προς το τραπεζικό σύστημα. Μετά το 2016, οι ελληνικές τράπεζες και οι εποπτικές αρχές προχώρησαν σε σειρά παρεμβάσεων για τη μείωση των μη εξυπηρετούμενων δανείων. Ακολούθως, εφαρμόστηκαν ρυθμίσεις δανείων, αναδιαρθρώσεις οφειλών, πωλήσεις προβληματικών χαρτοφυλακίων και δημιουργήθηκαν εταιρείες διαχείρισης απαιτήσεων. Παράλληλα, υιοθετήθηκαν αυστηρότερα κριτήρια αξιολόγησης νέων δανείων και μεγαλύτερη έμφαση στη διαχείριση κινδύνου (Bank of Greece, 2016). Εν τέλει, η εμπειρία της Ελλάδας ανέδειξε με έντονο τρόπο ότι η έγκαιρη πρόβλεψη του πιστωτικού κινδύνου και η χρήση αξιόπιστων μοντέλων αξιολόγησης δεν αποτελούν μόνο τεχνική ανάγκη των τραπεζών, αλλά βασική προϋπόθεση για τη σταθερότητα της οικονομίας και την προστασία της κοινωνίας από μελλοντικές κρίσεις.

## 2.6 Κανονιστικό και νομικό πλαίσιο στην Ελλάδα

Στην Ελλάδα, η εποπτεία και η σταθερότητα του τραπεζικού συστήματος αποτελούν βασικούς πυλώνες της οικονομικής λειτουργίας και ασκούνται κυρίως από την Bank of Greece, σε συνεργασία με τους ευρωπαϊκούς εποπτικούς μηχανισμούς και ιδιαίτερα με την European Central Bank στο πλαίσιο του Ενιαίου Εποπτικού Μηχανισμού (Single Supervisory Mechanism – SSM) (Bank of Greece, 2024). Η ανάγκη ύπαρξης ενός αυστηρού και αποτελεσματικού κανονιστικού πλαισίου έγινε ακόμη πιο έντονη μετά τη χρηματοπιστωτική κρίση του 2008 και ιδιαίτερα μετά την ελληνική οικονομική κρίση, όταν τα μη εξυπηρετούμενα δάνεια αυξήθηκαν σε εξαιρετικά υψηλά επίπεδα και απείλησαν τη σταθερότητα του τραπεζικού συστήματος (Avgouleas & Goodhart, 2015).

Το κανονιστικό πλαίσιο που εφαρμόζεται σήμερα στην Ελλάδα περιλαμβάνει ένα σύνολο ευρωπαϊκών οδηγιών, κανονισμών, εποπτικών κατευθύνσεων και εθνικών νομοθετικών παρεμβάσεων, οι οποίες στοχεύουν στη σωστή αναγνώριση,

παρακολούθηση και διαχείριση του πιστωτικού κινδύνου. Έτσι οι τράπεζες υποχρεούνται να εφαρμόζουν διαδικασίες αξιολόγησης της πιστοληπτικής ικανότητας των δανειοληπτών, να σχηματίζουν προβλέψεις για πιθανές ζημιές και να διατηρούν επαρκή κεφαλαιακή βάση, ώστε να μπορούν να απορροφούν ενδεχόμενες οικονομικές απώλειες. Όμως ιδιαίτερα σημαντική είναι η έννοια των μη εξυπηρετούμενων ανοιγμάτων (Non-Performing Exposures – NPEs), η οποία χρησιμοποιείται πλέον σε ευρωπαϊκό επίπεδο ως βασικός δείκτης αξιολόγησης της ποιότητας των τραπεζικών χαρτοφυλακίων. Σύμφωνα με τις κατευθυντήριες γραμμές της European Banking Authority (2019), ένα δάνειο δεν χαρακτηρίζεται μη εξυπηρετούμενο μόνο όταν εμφανίζει καθυστέρηση άνω των 90 ημερών, αλλά και όταν υπάρχουν ενδείξεις ότι ο δανειολήπτης είναι απίθανο να αποπληρώσει πλήρως τις υποχρεώσεις του, ακόμη και αν η καθυστέρηση είναι μικρότερη (European Central Bank, 2021). Το στοιχείο αυτό δείχνει ότι η σύγχρονη αξιολόγηση πιστωτικού κινδύνου δεν περιορίζεται αποκλειστικά σε ποσοτικά κριτήρια, όπως ο αριθμός ημερών καθυστέρησης, αλλά λαμβάνει υπόψη και ποιοτικούς παράγοντες που σχετίζονται με τη συνολική οικονομική κατάσταση του δανειολήπτη.

Εμβαθύνοντας και αποσαφηνίζοντας το πλαίσιο, μια επιχείρηση μπορεί να εξυπηρετεί προσωρινά τις δανειακές της υποχρεώσεις, αλλά να εμφανίζει σημαντική μείωση κύκλου εργασιών, υψηλό λειτουργικό κόστος ή αδυναμία κάλυψης βασικών οικονομικών υποχρεώσεων. Αντίστοιχα, ένα νοικοκυριό μπορεί να μην έχει ακόμη καθυστερήσει σημαντικά τις πληρωμές του, αλλά η απώλεια εργασίας, η μείωση εισοδήματος ή η υπερχρέωση να υποδηλώνουν αυξημένη πιθανότητα μελλοντικής αθέτησης (Louzis et al., 2012). Οι περιπτώσεις αυτές καταδεικνύουν ότι η αξιολόγηση του πιστωτικού κινδύνου απαιτεί συνδυασμό οικονομικής ανάλυσης, στατιστικών δεδομένων και ουσιαστικής κατανόησης της πραγματικής οικονομικής κατάστασης του δανειολήπτη.

### 3. Μεθοδολογία έρευνας

#### 3.1 Σκοπός της εμπειρικής ανάλυσης

Βασικός σκοπός της παρούσας εργασίας είναι η διερεύνηση των παραγόντων που επηρεάζουν την πιθανότητα καθυστέρησης αποπληρωμής δανείων από τους δανειολήπτες, μέσω της ανάπτυξης ενός πλαισίου αξιολόγησης πιστωτικού κινδύνου βασισμένου σε πραγματικά δεδομένα. Η ανάλυση επιδιώκει να αναδείξει τη σχέση μεταξύ των δημογραφικών, οικονομικών και πιστωτικών χαρακτηριστικών των πελατών και της πιστοληπτικής τους συμπεριφοράς.

Η σημασία της αξιολόγησης πιστωτικού κινδύνου είναι ιδιαίτερα αυξημένη στο σύγχρονο τραπεζικό περιβάλλον, καθώς η αποτελεσματική πρόβλεψη της πιθανότητας αθέτησης συμβάλλει τόσο στη μείωση των επισφαλειών όσο και στη διαμόρφωση ορθολογικής πιστοδοτικής πολιτικής (Rahman, & Awwal, 2024). Για τον λόγο αυτό η έγκαιρη αναγνώριση πελατών με αυξημένη πιθανότητα καθυστέρησης επιτρέπει στα χρηματοπιστωτικά ιδρύματα να λαμβάνουν καλύτερες αποφάσεις σχετικά με την έγκριση δανείων, τον καθορισμό επιτοκίων και τη διαχείριση του συνολικού χαρτοφυλακίου τους.

Η εμπειρική ανάλυση βασίζεται σε στατιστικές τεχνικές, μεθόδους περιγραφικής στατιστικής και διερευνητικής ανάλυσης δεδομένων (Exploratory Data Analysis – EDA), με στόχο την κατανόηση της δομής των δεδομένων και των σχέσεων μεταξύ των μεταβλητών. Μέσω των διαδικασιών αυτών εξετάζονται οι κατανομές των μεταβλητών, οι συσχετίσεις μεταξύ χαρακτηριστικών και η επίδρασή τους στην πιθανότητα εμφάνισης πιστωτικού κινδύνου. Επιπλέον, η μεθοδολογία στοχεύει στον εντοπισμό των σημαντικότερων μεταβλητών που σχετίζονται με τον πιστωτικό κίνδυνο και στην κατηγοριοποίηση των δανειοληπτών σε διαφορετικά επίπεδα κινδύνου. Με τον τρόπο αυτό δημιουργείται ένα υποστηρικτικό εργαλείο λήψης αποφάσεων, το οποίο μπορεί να αξιοποιηθεί στο πλαίσιο διαδικασιών credit scoring και risk management.

#### 3.2 Περιγραφή του συνόλου των δεδομένων

Για τις ανάγκες της ανάλυσης χρησιμοποιήθηκε το German Credit Dataset, το οποίο αποτελεί ένα από τα πιο διαδεδομένα και ευρέως χρησιμοποιούμενα σύνολα δεδομένων στη διεθνή βιβλιογραφία σχετικά με την αξιολόγηση πιστωτικού

κινδύνου και τα συστήματα credit scoring. Η συχνή χρήση του σε ακαδημαϊκές και ερευνητικές εφαρμογές το καθιστά κατάλληλο για συγκριτικές αναλύσεις και ανάπτυξη μοντέλων πρόβλεψης.

Το συγκεκριμένο dataset περιλαμβάνει δεδομένα για 1.000 δανειολήπτες και αποτελείται από μεταβλητές που σχετίζονται με οικονομικά, δημογραφικά και πιστωτικά χαρακτηριστικά. Οι μεταβλητές αυτές αποτυπώνουν στοιχεία της οικονομικής κατάστασης των πελατών, της πιστωτικής τους συμπεριφοράς καθώς και των χαρακτηριστικών των δανείων που έχουν λάβει.

Ενδεικτικά, το dataset περιλαμβάνει μεταβλητές όπως:

- η διάρκεια του δανείου,
- το ποσό του δανείου,
- το πιστωτικό ιστορικό του δανειολήπτη,
- η επαγγελματική κατάσταση,
- η ηλικία,
- ο αριθμός υφιστάμενων δανείων,
- η ύπαρξη περιουσιακών στοιχείων,
- η κατάσταση τραπεζικού λογαριασμού,
- καθώς και στοιχεία σχετικά με την αποπληρωμή προηγούμενων υποχρεώσεων.

Με βάση τα χαρακτηριστικά αυτά, οι δανειολήπτες διαχωρίζονται σε δύο βασικές κατηγορίες:

1. Δανειολήπτες με καλή πιστοληπτική συμπεριφορά
2. Δανειολήπτες με αυξημένο πιστωτικό κίνδυνο

Ο διαχωρισμός αυτός επιτρέπει την ανάλυση των διαφορών μεταξύ των δύο ομάδων και την αναγνώριση των παραγόντων που σχετίζονται περισσότερο με την πιθανότητα καθυστέρησης αποπληρωμής. Παράλληλα, παρέχει τη δυνατότητα ανάπτυξης μεθόδων ταξινόμησης και αξιολόγησης του κινδύνου με βάση τα χαρακτηριστικά των πελατών.

### 3.3 Προεπεξεργασία δεδομένων

Πριν από την εφαρμογή των στατιστικών τεχνικών πραγματοποιήθηκε διαδικασία προεπεξεργασίας των δεδομένων, με στόχο τη βελτίωση της ποιότητας

και της αξιοπιστίας των αποτελεσμάτων της ανάλυσης. Εξάλλου πρόκειται για κρίσιμο στάδιο σε κάθε διαδικασία ανάλυσης δεδομένων, καθώς πιθανά σφάλματα, ελλιπείς τιμές ή ακραίες παρατηρήσεις μπορούν να επηρεάσουν σημαντικά την ακρίβεια των συμπερασμάτων.

Η διαδικασία περιλάμβανε αρχικά έλεγχο για ελλιπείς ή μη έγκυρες τιμές. Όπου κρίθηκε απαραίτητο, πραγματοποιήθηκε αντικατάσταση των ελλειπουσών τιμών με κατάλληλες στατιστικές μεθόδους, όπως η χρήση μέσων τιμών ή επικρατέστερων κατηγοριών. Σε περιπτώσεις όπου οι εγγραφές παρουσίαζαν σοβαρές ασυνέπειες ή αδυναμία αποκατάστασης, αυτές απομακρύνθηκαν από το dataset ώστε να διασφαλιστεί η αξιοπιστία της ανάλυσης. Στη συνέχεια, οι κατηγορικές μεταβλητές μετατράπηκαν σε αριθμητική μορφή μέσω διαδικασιών κωδικοποίησης (encoding), προκειμένου να μπορούν να χρησιμοποιηθούν από τα στατιστικά εργαλεία και τις διαδικασίες ανάλυσης. Η μετατροπή αυτή είναι απαραίτητη, καθώς οι περισσότερες στατιστικές τεχνικές και τα μοντέλα πρόβλεψης απαιτούν αριθμητικά δεδομένα.

Επιπρόσθετα, πραγματοποιήθηκε έλεγχος για ακραίες τιμές (outliers), οι οποίες ενδέχεται να επηρεάσουν δυσανάλογα τα αποτελέσματα των αναλύσεων. Ο εντοπισμός τους έγινε μέσω στατιστικών και γραφικών μεθόδων, όπως boxplots και μέτρα διασποράς. Σε περιπτώσεις όπου οι ακραίες τιμές θεωρήθηκαν μη ρεαλιστικές ή λανθασμένες, αντιμετωπίστηκαν κατάλληλα ώστε να περιοριστεί η επίδρασή τους στα αποτελέσματα. Τέλος, εξετάστηκε η συνοχή και η ομοιογένεια των δεδομένων, ώστε να διασφαλιστεί ότι όλες οι μεταβλητές βρίσκονται σε κατάλληλη μορφή και κλίμακα για τη συνέχεια της ανάλυσης.

### 3.4 Στατιστικές τεχνικές ανάλυσης

Η ανάλυση των δεδομένων βασίστηκε σε στατιστικές τεχνικές που χρησιμοποιούνται ευρέως στη μελέτη του πιστωτικού κινδύνου και στην ανάπτυξη συστημάτων αξιολόγησης πιστοληπτικής ικανότητας. Οι τεχνικές αυτές επιλέχθηκαν με στόχο την κατανόηση των χαρακτηριστικών των δανειοληπτών, την αναγνώριση προτύπων συμπεριφοράς και την εκτίμηση της πιθανότητας αθέτησης. Αρχικά εφαρμόστηκε περιγραφική στατιστική ανάλυση, μέσω της οποίας υπολογίστηκαν βασικά στατιστικά μέτρα, όπως ο μέσος όρος, η διάμεσος, η τυπική απόκλιση και οι

συχνότητες εμφάνισης των μεταβλητών (Μυλωνάς, 2018). Η διαδικασία αυτή παρείχε μία πρώτη συνολική εικόνα για τη δομή και τη συμπεριφορά των δεδομένων. Στη συνέχεια πραγματοποιήθηκε ανάλυση συσχέτισης μεταξύ των μεταβλητών, με στόχο την αναγνώριση πιθανών σχέσεων μεταξύ των χαρακτηριστικών των δανειοληπτών και της πιθανότητας καθυστέρησης αποπληρωμής. Η ανάλυση αυτή συμβάλλει στον εντοπισμό των σημαντικότερων μεταβλητών που επηρεάζουν τον πιστωτικό κίνδυνο.

Επιπλέον, χρησιμοποιήθηκαν πίνακες συχνοτήτων και αναλογίες αθέτησης, προκειμένου να εξεταστεί η κατανομή των δανειοληπτών ανά κατηγορία κινδύνου και να υπολογιστούν ποσοστά καθυστέρησης για διαφορετικές ομάδες πελατών. Μέσω αυτής της διαδικασίας καθίσταται δυνατή η εξαγωγή συμπερασμάτων σχετικά με τα χαρακτηριστικά που συνδέονται με υψηλότερη πιθανότητα αθέτησης. Τέλος, εφαρμόστηκε διαδικασία κατηγοριοποίησης των δανειοληπτών σε επίπεδα κινδύνου, με σκοπό τη δημιουργία ενός απλού αλλά αποτελεσματικού πλαισίου αξιολόγησης πιστωτικού κινδύνου. Η κατηγοριοποίηση αυτή επιτρέπει τη διάκριση των πελατών σε ομάδες χαμηλού, μεσαίου και υψηλού κινδύνου, διευκολύνοντας τη διαδικασία λήψης πιστοδοτικών αποφάσεων.

### 3.5 Εργαλεία ανάλυσης δεδομένων

Για την υλοποίηση της ανάλυσης χρησιμοποιήθηκε η γλώσσα προγραμματισμού Python, η οποία αποτελεί ένα από τα σημαντικότερα εργαλεία στον τομέα της επιστήμης δεδομένων, της στατιστικής ανάλυσης και της χρηματοοικονομικής μοντελοποίησης. Η ευρεία χρήση της στην ακαδημαϊκή και επαγγελματική κοινότητα οφείλεται στην ευελιξία, την αποδοτικότητα και τη δυνατότητα διαχείρισης μεγάλου όγκου δεδομένων.

Στο πλαίσιο της παρούσας έρευνας αξιοποιήθηκαν εξειδικευμένες βιβλιοθήκες της Python, οι οποίες υποστηρίζουν διαφορετικά στάδια της ανάλυσης (Kabir et al., 2024):

- pandas για τη διαχείριση, οργάνωση και επεξεργασία των δεδομένων,
- numpy για αριθμητικούς και μαθηματικούς υπολογισμούς,
- matplotlib και seaborn για τη δημιουργία γραφημάτων και οπτικοποιήσεων,
- scipy για την εφαρμογή στατιστικών μεθόδων και ελέγχων.

Η χρήση των παραπάνω εργαλείων συνέβαλε στην αποτελεσματική επεξεργασία των δεδομένων, στην οπτικοποίηση των αποτελεσμάτων και στη διεξαγωγή αξιόπιστων στατιστικών αναλύσεων.

### 3.6 Exploratory Data Analysis (EDA)

Πριν από την εφαρμογή των στατιστικών μοντέλων πραγματοποιήθηκε διερευνητική ανάλυση δεδομένων (Exploratory Data Analysis – EDA), με στόχο την καλύτερη κατανόηση της δομής και της ποιότητας του dataset. Η διαδικασία της EDA αποτελεί σημαντικό στάδιο σε κάθε ερευνητική ανάλυση, καθώς επιτρέπει τον εντοπισμό προτύπων, ανωμαλιών και πιθανών προβλημάτων στα δεδομένα πριν από την εφαρμογή πιο σύνθετων τεχνικών. Μέσω της διερευνητικής ανάλυσης εξετάστηκαν:

- οι κατανομές των μεταβλητών,
- η ύπαρξη ακραίων τιμών,
- οι σχέσεις μεταξύ αριθμητικών και κατηγορικών μεταβλητών,
- καθώς και πιθανές ανισορροπίες μεταξύ των κατηγοριών των δανειοληπτών.

Για την καλύτερη κατανόηση των δεδομένων χρησιμοποιήθηκαν διαγράμματα κατανομής, ιστογράμματα, boxplots και heatmaps συσχέτισης. Η οπτικοποίηση αυτή συνέβαλε στην εξαγωγή αρχικών συμπερασμάτων σχετικά με τη συμπεριφορά των μεταβλητών και τη σχέση τους με τον πιστωτικό κίνδυνο.

### 3.7 Ανάλυση συσχέτισης μεταβλητών

Η ανάλυση συσχέτισης αποτελεί βασικό μέρος της μεθοδολογίας, καθώς επιτρέπει την αναγνώριση των μεταβλητών που παρουσιάζουν ισχυρότερη σχέση με την πιθανότητα καθυστέρησης αποπληρωμής. Για τη μέτρηση της σχέσης μεταξύ αριθμητικών μεταβλητών χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson, ο οποίος λαμβάνει τιμές από -1 έως +1. Τιμές κοντά στο +1 υποδηλώνουν ισχυρή θετική συσχέτιση, ενώ τιμές κοντά στο -1 υποδηλώνουν ισχυρή αρνητική συσχέτιση (Μυλωνάς, 2018). Αντίθετα, τιμές κοντά στο μηδέν δείχνουν ασθενή ή μη γραμμική σχέση μεταξύ των μεταβλητών. Η ανάλυση αυτή συμβάλλει στον εντοπισμό σημαντικών χαρακτηριστικών που μπορούν να αξιοποιηθούν για την ανάπτυξη

μοντέλων πρόβλεψης πιστωτικού κινδύνου, ενώ παράλληλα βοηθά στον εντοπισμό πιθανών προβλημάτων πολυσυγγραμμικότητας μεταξύ των μεταβλητών.

### 3.8 Ανάλυση πιθανοτήτων και πίνακες συχνοτήτων

Για την εκτίμηση της πιθανότητας καθυστέρησης αποπληρωμής χρησιμοποιήθηκαν πίνακες συχνοτήτων και αναλογίες αθέτησης (default rates). Μέσω αυτών παρουσιάζεται η κατανομή των δανειοληπτών ανά κατηγορία μεταβλητών και υπολογίζεται το ποσοστό των πελατών που εμφανίζουν προβλήματα αποπληρωμής. Η ανάλυση αυτή επιτρέπει τον υπολογισμό της πιθανότητας αθέτησης (Probability of Default – PD) για κάθε κατηγορία δανειοληπτών. Η Probability of Default αποτελεί βασικό δείκτη στη διαχείριση πιστωτικού κινδύνου, καθώς χρησιμοποιείται ευρέως από τα χρηματοπιστωτικά ιδρύματα για την αξιολόγηση πελατών και την εκτίμηση πιθανών ζημιών. Μέσω των αποτελεσμάτων της ανάλυσης μπορούν να αναγνωριστούν συγκεκριμένα χαρακτηριστικά που συνδέονται με αυξημένο κίνδυνο, όπως υψηλά ποσά δανείων, αρνητικό πιστωτικό ιστορικό ή μεγάλος αριθμός υφιστάμενων υποχρεώσεων.

### 3.9 Κατηγοριοποίηση δανειοληπτών (Risk Segmentation)

Στο πλαίσιο της μεθοδολογίας εφαρμόστηκε διαδικασία κατηγοριοποίησης των δανειοληπτών σε επίπεδα κινδύνου (risk segmentation). Η διαδικασία αυτή βασίστηκε στα χαρακτηριστικά των πελατών και στην εκτίμηση της πιθανότητας αθέτησης που προέκυψε από τις προηγούμενες αναλύσεις (Bhatt et al., 2023).

Οι δανειολήπτες διαχωρίστηκαν σε τρεις βασικές κατηγορίες:

- χαμηλού κινδύνου,
- μεσαίου κινδύνου,
- υψηλού κινδύνου.

Η κατηγοριοποίηση αυτή αποτελεί σημαντικό εργαλείο στη διαχείριση πιστωτικού κινδύνου, καθώς επιτρέπει στις τράπεζες και στους χρηματοπιστωτικούς οργανισμούς να προσαρμόζουν την πολιτική χορήγησης δανείων ανάλογα με το επίπεδο κινδύνου κάθε πελάτη. Επιπλέον, διευκολύνει τη διαδικασία παρακολούθησης του χαρτοφυλακίου δανείων και συμβάλλει στη μείωση πιθανών επισφαλειών.

### 3.10 Διαδικασία Υλοποίησης της Ανάλυσης

Η υλοποίηση της ανάλυσης πραγματοποιήθηκε σε περιβάλλον Python μέσω του λογισμικού Visual Studio Code. Η επιλογή του συγκεκριμένου περιβάλλοντος ανάπτυξης έγινε λόγω της ευκολίας διαχείρισης κώδικα, της υποστήριξης βιβλιοθηκών ανάλυσης δεδομένων και της δυνατότητας οργάνωσης της ερευνητικής διαδικασίας.

Η διαδικασία της ανάλυσης ακολούθησε διαδοχικά στάδια:

1. φόρτωση και ανάγνωση του dataset,
2. έλεγχος και καθαρισμός δεδομένων,
3. προεπεξεργασία και κωδικοποίηση μεταβλητών,
4. διερευνητική και περιγραφική στατιστική ανάλυση,
5. δημιουργία γραφημάτων και οπτικοποιήσεων,
6. ανάλυση συσχετίσεων και πιθανοτήτων αθέτησης,
7. κατηγοριοποίηση δανειοληπτών σε επίπεδα κινδύνου,
8. εξαγωγή και ερμηνεία αποτελεσμάτων.

Η οργανωμένη εφαρμογή των παραπάνω σταδίων συνέβαλε στη δημιουργία μιας ολοκληρωμένης διαδικασίας ανάλυσης πιστωτικού κινδύνου, η οποία συνδυάζει στατιστικές τεχνικές, ανάλυση δεδομένων και υποστήριξη λήψης αποφάσεων. Τα αποτελέσματα της ανάλυσης παρουσιάζονται και ερμηνεύονται στο επόμενο κεφάλαιο της εργασίας.

#### 4.1 Εμπειρική εφαρμογή μοντέλου αξιολόγησης πιστωτικού κινδύνου

Το παρόν κεφάλαιο πραγματεύεται την πρακτική υλοποίηση της εμπειρικής ανάλυσης που πραγματοποιήθηκε για τη μελέτη του πιστωτικού κινδύνου και της πιθανότητας αθέτησης δανειακών υποχρεώσεων. Σε αντίθεση με το προηγούμενο κεφάλαιο, όπου παρουσιάστηκε η μεθοδολογική προσέγγιση της έρευνας, εδώ δίνεται έμφαση στην εφαρμογή των τεχνικών ανάλυσης στα πραγματικά δεδομένα του German Credit Dataset και στην ερμηνεία των αποτελεσμάτων που προέκυψαν. Η ανάλυση υλοποιήθηκε σε περιβάλλον Python και βασίστηκε σε διαδικασίες επεξεργασίας δεδομένων, στατιστικής ανάλυσης και κατηγοριοποίησης πιστωτικού κινδύνου. Μέσα από την εφαρμογή των τεχνικών αυτών επιδιώκεται η ανάδειξη των χαρακτηριστικών που συνδέονται περισσότερο με αυξημένη πιθανότητα καθυστέρησης αποπληρωμής, καθώς και η ανάπτυξη ενός απλού πλαισίου αξιολόγησης δανειοληπτών.

Παράλληλα, το κεφάλαιο αυτό επιδιώκει να συνδέσει τα στατιστικά ευρήματα με την πραγματική τραπεζική πρακτική. Η αξιολόγηση πιστωτικού κινδύνου δεν αποτελεί μόνο μία τεχνική διαδικασία ανάλυσης δεδομένων, αλλά ένα κρίσιμο εργαλείο λήψης αποφάσεων που επηρεάζει τη βιωσιμότητα των τραπεζών, την ποιότητα του δανειακού χαρτοφυλακίου και την οικονομική σταθερότητα.

#### 4.2 Υλοποίηση της Ανάλυσης σε Περιβάλλον Python

Η υλοποίηση της ανάλυσης πραγματοποιήθηκε με τη χρήση της γλώσσας προγραμματισμού Python σε περιβάλλον Visual Studio Code (VS Code). Η επιλογή της Python βασίστηκε στην ευρεία χρήση της στον χώρο της επιστήμης δεδομένων, της τραπεζικής ανάλυσης και των εφαρμογών machine learning, καθώς παρέχει υψηλή ευελιξία στη διαχείριση και ανάλυση οικονομικών δεδομένων.

Κατά την ανάπτυξη της εφαρμογής δημιουργήθηκε οργανωμένη δομή αρχείων, ώστε κάθε στάδιο της ανάλυσης να υλοποιείται ξεχωριστά. Η προσέγγιση αυτή διευκολύνει τόσο την αναγνωσιμότητα του κώδικα όσο και τη δυνατότητα μελλοντικής επέκτασης της εφαρμογής.

Ειδικότερα, αναπτύχθηκαν επιμέρους αρχεία για:

- τη φόρτωση και προεπεξεργασία των δεδομένων,

- την περιγραφική στατιστική ανάλυση,
- την οπτικοποίηση των μεταβλητών,
- τον υπολογισμό της πιθανότητας αθέτησης,
- καθώς και την κατηγοριοποίηση των δανειοληπτών σε επίπεδα κινδύνου.

Η οργανωμένη αυτή δομή προσομοιάζει πρακτικές που εφαρμόζονται σε πραγματικά τραπεζικά πληροφοριακά συστήματα, όπου οι διαδικασίες ανάλυσης κινδύνου υλοποιούνται μέσω διακριτών και επαναχρησιμοποιήσιμων υποσυστημάτων.

#### 4.3 Φόρτωση και Έλεγχος Δεδομένων

Το German Credit Dataset ανακτήθηκε μέσω της πλατφόρμας Kaggle και φορτώθηκε σε dataframe της βιβλιοθήκης pandas. Μετά τη φόρτωση πραγματοποιήθηκε αρχικός έλεγχος της δομής των δεδομένων, προκειμένου να επιβεβαιωθεί η ορθότητα των μεταβλητών και η συνοχή των εγγραφών.

Κατά το στάδιο αυτό εξετάστηκαν:

- οι τύποι των μεταβλητών,
- η ύπαρξη ελλিপών τιμών,
- η κατανομή των κατηγοριών,
- καθώς και πιθανές ασυνέπειες στις εγγραφές.

Η διαδικασία αυτή είναι ιδιαίτερα σημαντική, καθώς ακόμη και μικρά προβλήματα ποιότητας δεδομένων μπορούν να επηρεάσουν σημαντικά την αξιοπιστία των αποτελεσμάτων. Στον τραπεζικό τομέα, η λανθασμένη αξιολόγηση ενός δανειολήπτη μπορεί να οδηγήσει είτε σε αύξηση επισφαλειών είτε σε απόρριψη πελατών με καλή πιστοληπτική συμπεριφορά.

#### 4.4 Περιγραφικά ευρήματα της ανάλυσης

Η αρχική στατιστική ανάλυση παρείχε μια συνολική εικόνα των χαρακτηριστικών των δανειοληπτών και των δανείων του dataset. Τα αποτελέσματα έδειξαν σημαντική διαφοροποίηση μεταξύ των πελατών ως προς τη διάρκεια δανείου, το ύψος της χρηματοδότησης και τα δημογραφικά χαρακτηριστικά.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η κατανομή της μεταβλητής πιστωτικού κινδύνου.

Από τις 1.000 συνολικές παρατηρήσεις:

- περίπου το 70% των δανειοληπτών εμφάνιζε καλή πιστοληπτική συμπεριφορά,
- ενώ το 30% παρουσίαζε καθυστέρηση ή αυξημένο κίνδυνο αθέτησης.

Η αναλογία αυτή θεωρείται αρκετά ρεαλιστική για εφαρμογές πιστωτικής αξιολόγησης και αντανακλά τη λειτουργία πραγματικών τραπεζικών χαρτοφυλακίων, όπου η πλειονότητα των πελατών εξυπηρετεί κανονικά τις υποχρεώσεις της, ενώ ένα μικρότερο ποσοστό εμφανίζει προβλήματα αποπληρωμής.

Παράλληλα, παρατηρήθηκε ότι τα δάνεια μεγαλύτερης διάρκειας και υψηλότερου ποσού εμφανίζουν αυξημένη μεταβλητότητα, στοιχείο που συνδέεται με υψηλότερη αβεβαιότητα και αυξημένο πιστωτικό κίνδυνο.

#### 4.5 Οπτικοποίηση και ερμηνεία δεδομένων

Η χρήση γραφημάτων συνέβαλε ουσιαστικά στην κατανόηση της συμπεριφοράς των μεταβλητών και στην αναγνώριση βασικών προτύπων πιστωτικού κινδύνου.

Μέσω histogram και boxplot διαπιστώθηκε ότι:

- οι νεότεροι δανειολήπτες εμφανίζουν μεγαλύτερη συγκέντρωση στις κατηγορίες αυξημένου κινδύνου,
- ενώ τα μεγαλύτερα ποσά δανείων συνδέονται συχνότερα με καθυστερήσεις αποπληρωμής.

Επιπλέον, το heatmap συσχέτισης ανέδειξε θετική σχέση μεταξύ της διάρκειας δανείου και της πιθανότητας αθέτησης, εύρημα το οποίο θεωρείται ιδιαίτερα σημαντικό, καθώς επιβεβαιώνει ότι όσο αυξάνεται ο χρονικός ορίζοντας αποπληρωμής, τόσο αυξάνεται και η αβεβαιότητα σχετικά με τη μελλοντική οικονομική κατάσταση του δανειολήπτη. Φαίνεται μάλιστα ότι η οπτικοποίηση των δεδομένων δεν λειτουργεί μόνο υποστηρικτικά στην ανάλυση, αλλά αποτελεί βασικό εργαλείο λήψης αποφάσεων για τα χρηματοπιστωτικά ιδρύματα. Στην πράξη, οι τράπεζες χρησιμοποιούν αντίστοιχες τεχνικές για την παρακολούθηση χαρτοφυλακίων, τον εντοπισμό υψηλού κινδύνου και τη βελτίωση των πολιτικών χορήγησης δανείων.

#### 4.6 Υπολογισμός πιθανότητας αθέτησης (Probability of Default)

Η εκτίμηση της πιθανότητας αθέτησης αποτέλεσε βασικό στόχο της εμπειρικής ανάλυσης. Με βάση τα δεδομένα του dataset, η συνολική Probability of Default (PD) υπολογίστηκε περίπου στο 30%, γεγονός που σημαίνει ότι σχεδόν ένας στους τρεις δανειολήπτες εμφάνιζε αυξημένο κίνδυνο καθυστέρησης.

Η ανάλυση ανά κατηγορία μεταβλητών έδειξε ότι:

- τα μακροχρόνια δάνεια παρουσιάζουν υψηλότερα ποσοστά αθέτησης,
- τα μεγαλύτερα ποσά χρηματοδότησης συνδέονται με αυξημένο πιστωτικό κίνδυνο,
- ενώ οι νεότερες ηλικιακές ομάδες εμφανίζουν μεγαλύτερη πιθανότητα καθυστέρησης.

Τα αποτελέσματα αυτά μπορούν να ερμηνευθούν και με οικονομικούς όρους. Οι νεότεροι δανειολήπτες διαθέτουν συχνά χαμηλότερη εισοδηματική σταθερότητα και μικρότερο πιστωτικό ιστορικό, γεγονός που αυξάνει την αβεβαιότητα ως προς τη μελλοντική τους δυνατότητα αποπληρωμής. Αντίστοιχα, τα μεγάλα και μακροχρόνια δάνεια συνεπάγονται μεγαλύτερη οικονομική επιβάρυνση και υψηλότερη έκθεση των τραπεζών στον πιστωτικό κίνδυνο.

#### 4.7 Ανάπτυξη μοντέλου κατηγοριοποίησης κινδύνου

Με βάση τα αποτελέσματα της ανάλυσης αναπτύχθηκε ένα απλό μοντέλο risk segmentation, με στόχο την κατηγοριοποίηση των δανειοληπτών σε επίπεδα κινδύνου. Το μοντέλο βασίστηκε κυρίως:

- στη διάρκεια του δανείου,
- στο ποσό χρηματοδότησης,
- και στην ηλικία του δανειολήπτη.

Σε κάθε χαρακτηριστικό αποδόθηκε συγκεκριμένη βαθμολογία κινδύνου και στη συνέχεια υπολογίστηκε ένα συνολικό risk score για κάθε πελάτη. Με βάση το σκορ αυτό οι δανειολήπτες ταξινομήθηκαν σε:

- χαμηλού κινδύνου,
- μεσαίου κινδύνου,
- υψηλού κινδύνου.

Η λογική αυτή προσομοιάζει πρακτικές credit scoring που εφαρμόζονται διεθνώς από τράπεζες και χρηματοπιστωτικούς οργανισμούς. Παρότι το μοντέλο της παρούσας εργασίας είναι απλοποιημένο, καταδεικνύει τον τρόπο με τον οποίο τα στατιστικά δεδομένα μπορούν να μετατραπούν σε εργαλείο υποστήριξης πιστοδοτικών αποφάσεων.

#### 4.8 Επικύρωση και αξιολόγηση του μοντέλου

Για την αξιολόγηση της αποτελεσματικότητας του μοντέλου εξετάστηκε η κατανομή της πιθανότητας αθέτησης μεταξύ των ομάδων κινδύνου που δημιουργήθηκαν.

Τα αποτελέσματα έδειξαν ότι:

- οι δανειολήπτες υψηλού κινδύνου παρουσίαζαν σημαντικά υψηλότερα ποσοστά καθυστέρησης,
- ενώ οι δανειολήπτες χαμηλού κινδύνου εμφάνιζαν αισθητά μικρότερη πιθανότητα αθέτησης.

Η διαφοροποίηση αυτή υποδηλώνει ότι το μοντέλο καταφέρνει σε ικανοποιητικό βαθμό να διαχωρίζει τους πελάτες ανάλογα με το επίπεδο πιστωτικού κινδύνου που παρουσιάζουν. Παρόμοιες διαδικασίες επικύρωσης εφαρμόζονται και στα πραγματικά τραπεζικά συστήματα, καθώς τα μοντέλα πιστωτικού κινδύνου πρέπει να αξιολογούνται συνεχώς ώστε να διασφαλίζεται η αξιοπιστία και η αποτελεσματικότητά τους σε μεταβαλλόμενες οικονομικές συνθήκες.

#### 4.9 Περιορισμοί της ανάλυσης

Παρά τα χρήσιμα ευρήματα της ανάλυσης, η παρούσα εργασία παρουσιάζει ορισμένους περιορισμούς που πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων. Αρχικά, το German Credit Dataset περιλαμβάνει περιορισμένο αριθμό παρατηρήσεων, γεγονός που ενδέχεται να επηρεάζει τη γενίκευση των αποτελεσμάτων σε πραγματικά τραπεζικά περιβάλλοντα μεγάλης κλίμακας.

Επιπλέον, το μοντέλο κατηγοριοποίησης που αναπτύχθηκε βασίζεται σε σχετικά απλές στατιστικές προσεγγίσεις και δεν ενσωματώνει πιο σύνθετες τεχνικές μηχανικής μάθησης ή μακροοικονομικούς δείκτες, όπως:

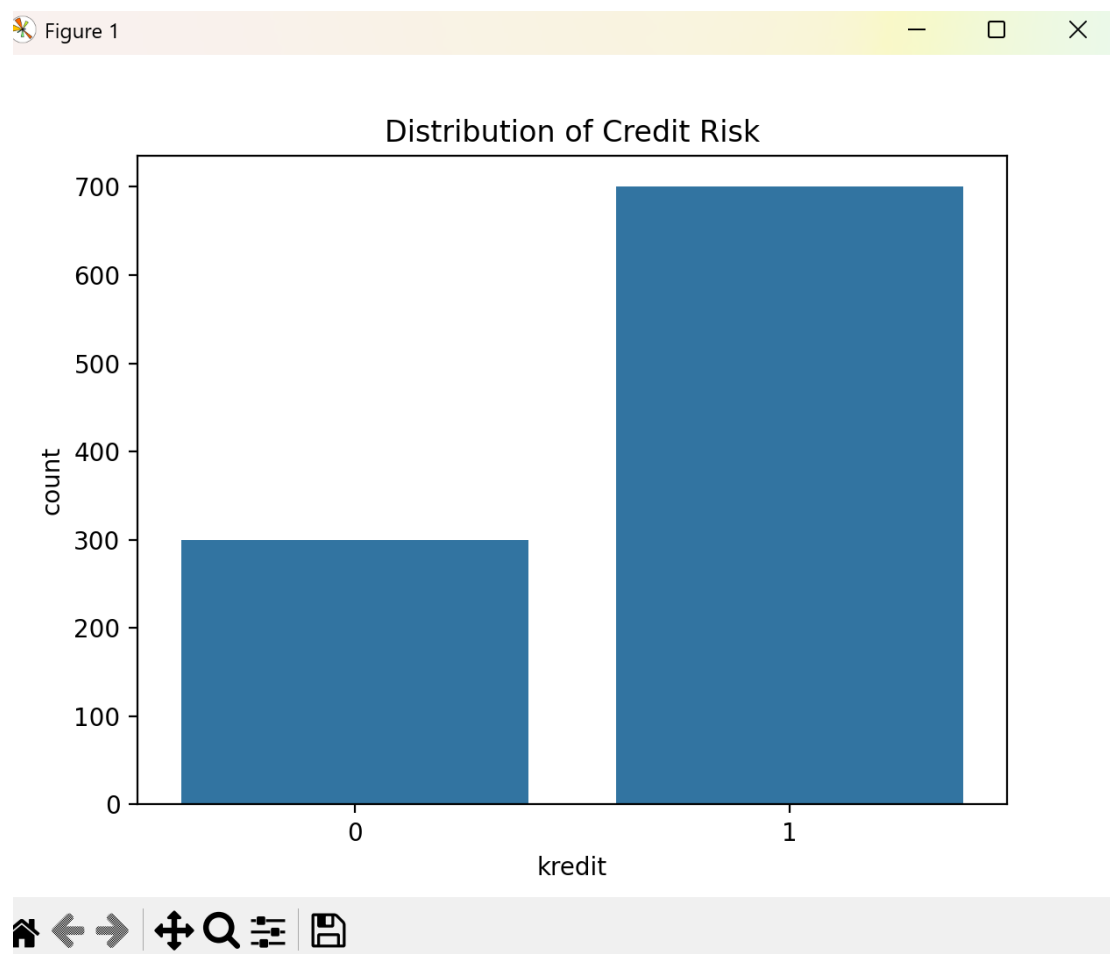
- πληθωρισμό,
- επιτόκια,
- επίπεδα ανεργίας,
- ή μεταβολές εισοδήματος.

Τέλος, η ανθρώπινη οικονομική συμπεριφορά επηρεάζεται συχνά από ψυχολογικούς και κοινωνικούς παράγοντες που είναι δύσκολο να αποτυπωθούν πλήρως μέσα από αριθμητικά δεδομένα. Για τον λόγο αυτό, η αξιολόγηση πιστωτικού κινδύνου δεν αποτελεί αποκλειστικά τεχνική διαδικασία, αλλά απαιτεί και ουσιαστική κατανόηση της οικονομικής πραγματικότητας των δανειοληπτών

### 5.1 Αποτελέσματα

Στο παρόν κεφάλαιο παρουσιάζονται τα αποτελέσματα της ανάλυσης των δεδομένων που πραγματοποιήθηκε με τη χρήση της Python, τα οποία βασίζονται στην οπτικοποίηση και τη στατιστική επεξεργασία του German Credit dataset και αποσκοπούν στην κατανόηση των παραγόντων που επηρεάζουν τον πιστωτικό κίνδυνο.

### 5.2 Κατανομή Πιστωτικού Κινδύνου

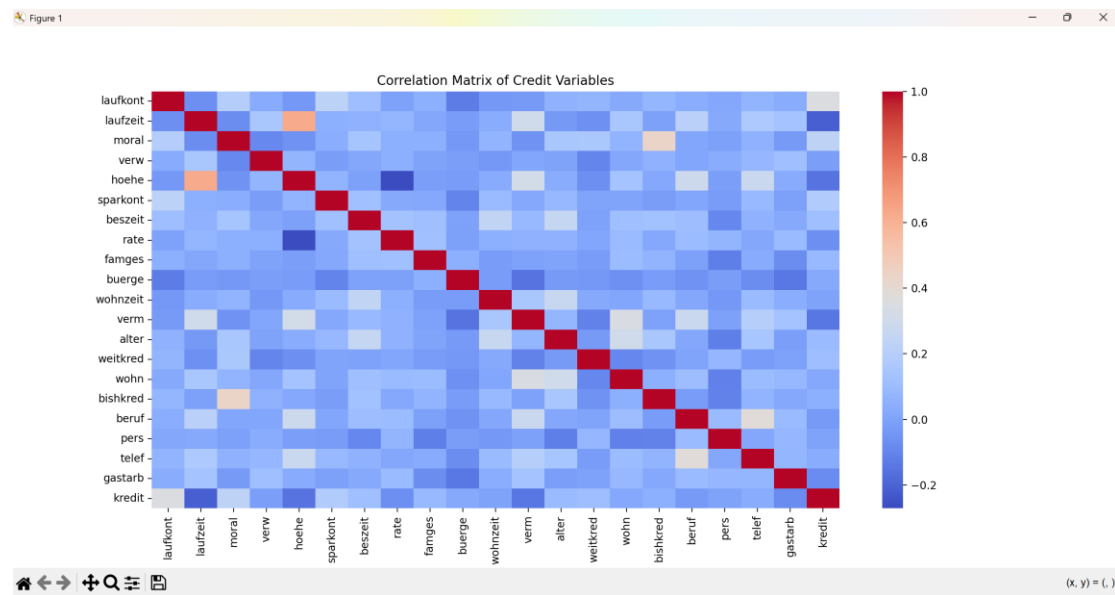


Η κατανομή της μεταβλητής *credit* δείχνει ότι από τους 1000 δανειολήπτες:

3. 700 ανήκουν στην κατηγορία χωρίς καθυστέρηση
4. 300 ανήκουν στην κατηγορία με καθυστέρηση

Το ποσοστό αυτό αντιστοιχεί σε περίπου 30% πιθανότητα αθέτησης, γεγονός που επιβεβαιώνει ότι ένα σημαντικό μέρος των δανειοληπτών παρουσιάζει πιστωτικό κίνδυνο. Η ανισορροπία αυτή είναι χαρακτηριστική σε τραπεζικά δεδομένα και πρέπει να λαμβάνεται υπόψη κατά την ανάλυση.

### 5.3 Ανάλυση συσχέτισης μεταβλητών



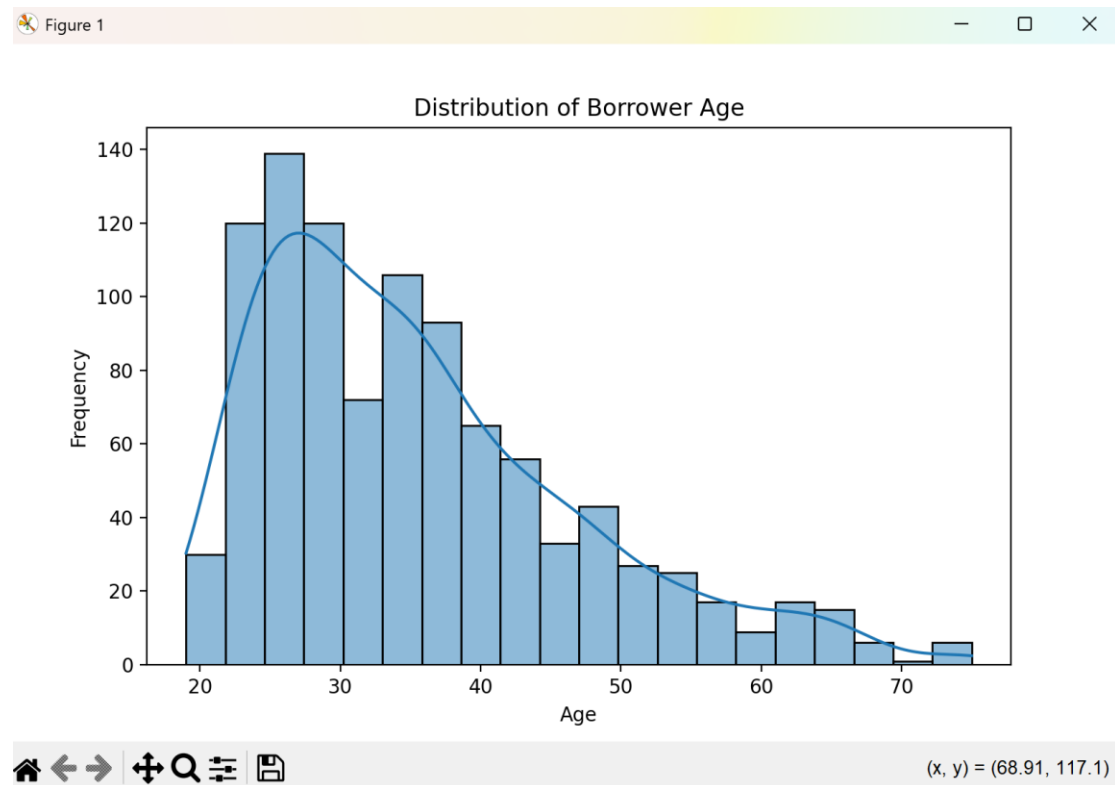
Το διάγραμμα συσχέτισης (correlation matrix) παρουσιάζει τις σχέσεις μεταξύ των μεταβλητών του dataset.

Παρατηρείται ότι:

- οι περισσότερες μεταβλητές έχουν χαμηλή συσχέτιση μεταξύ τους
- δεν υπάρχει μία μεταβλητή που να καθορίζει από μόνη της τον πιστωτικό κίνδυνο
- ο κίνδυνος εξαρτάται από συνδυασμό παραγόντων

Η απουσία ισχυρών συσχετίσεων δείχνει ότι το πρόβλημα είναι πολυπαραγοντικό και απαιτεί συνδυαστική ανάλυση.

## 5.4 Κατανομή ηλικίας δανειοληπτών

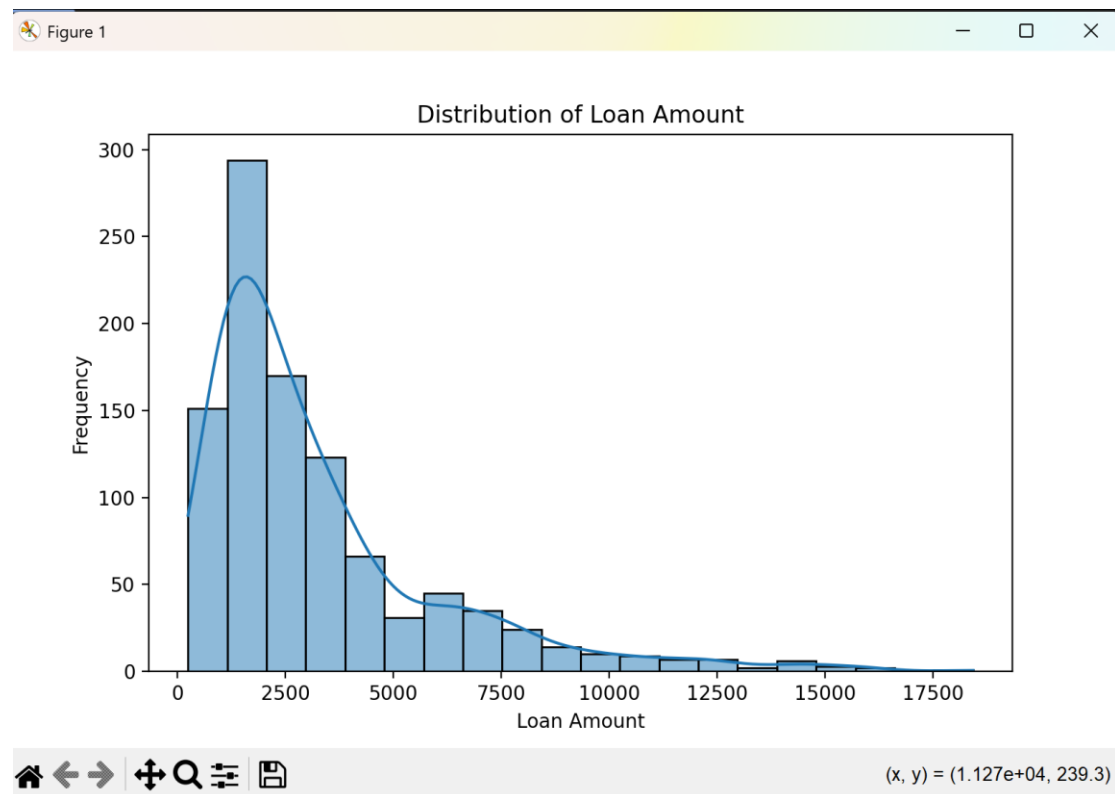


Η κατανομή της ηλικίας δείχνει ότι:

- η πλειοψηφία των δανειοληπτών βρίσκεται μεταξύ 25 και 40 ετών
- υπάρχει μικρότερη συγκέντρωση σε μεγαλύτερες ηλικίες
- η κατανομή παρουσιάζει ελαφριά δεξιά ασυμμετρία

Τα στοιχεία αυτά δείχνουν ότι τα περισσότερα δάνεια δίνονται σε άτομα παραγωγικής ηλικίας.

## 5.5 Κατανομή ποσού δανείου

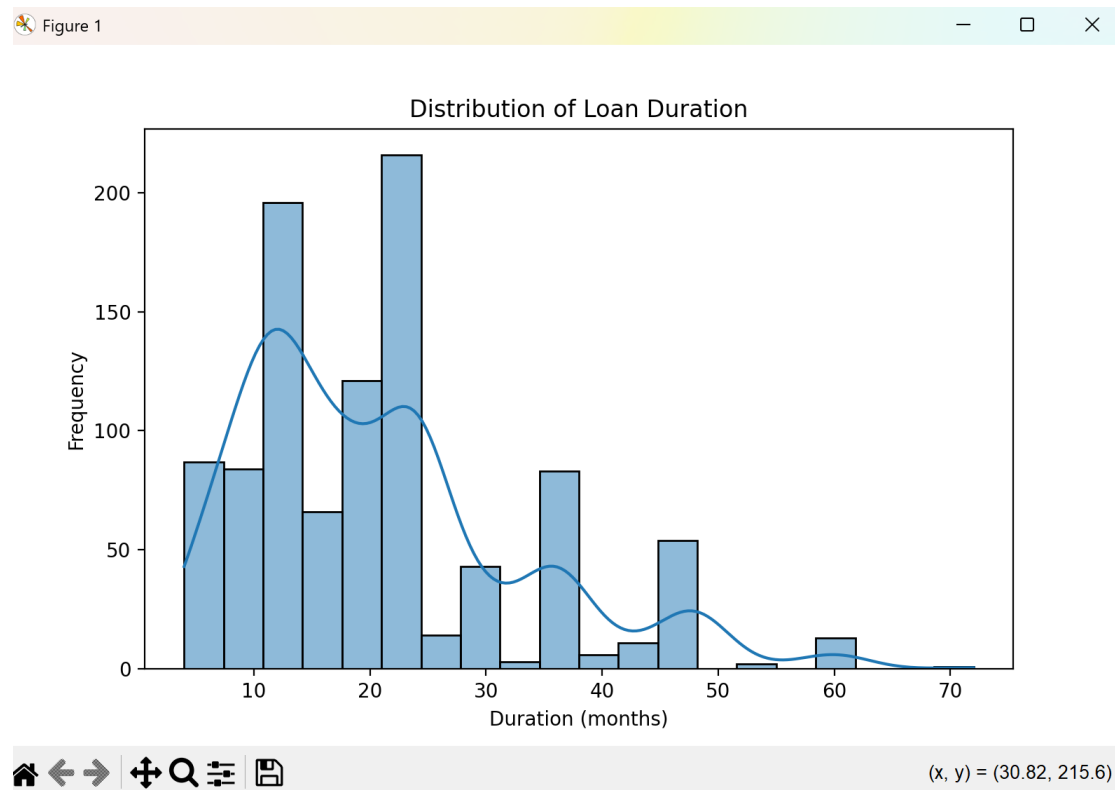


Η κατανομή του ποσού δανείου παρουσιάζει:

- έντονη δεξιά ασυμμετρία (right skewness)
- μεγάλο αριθμό μικρών δανείων
- λίγες περιπτώσεις πολύ μεγάλων ποσών (outliers)

Οι ακραίες τιμές υποδηλώνουν ύπαρξη δανείων υψηλού ρίσκου, τα οποία επηρεάζουν την συνολική εικόνα του πιστωτικού κινδύνου.

## 5.6 Κατανομή διάρκειας δανείου

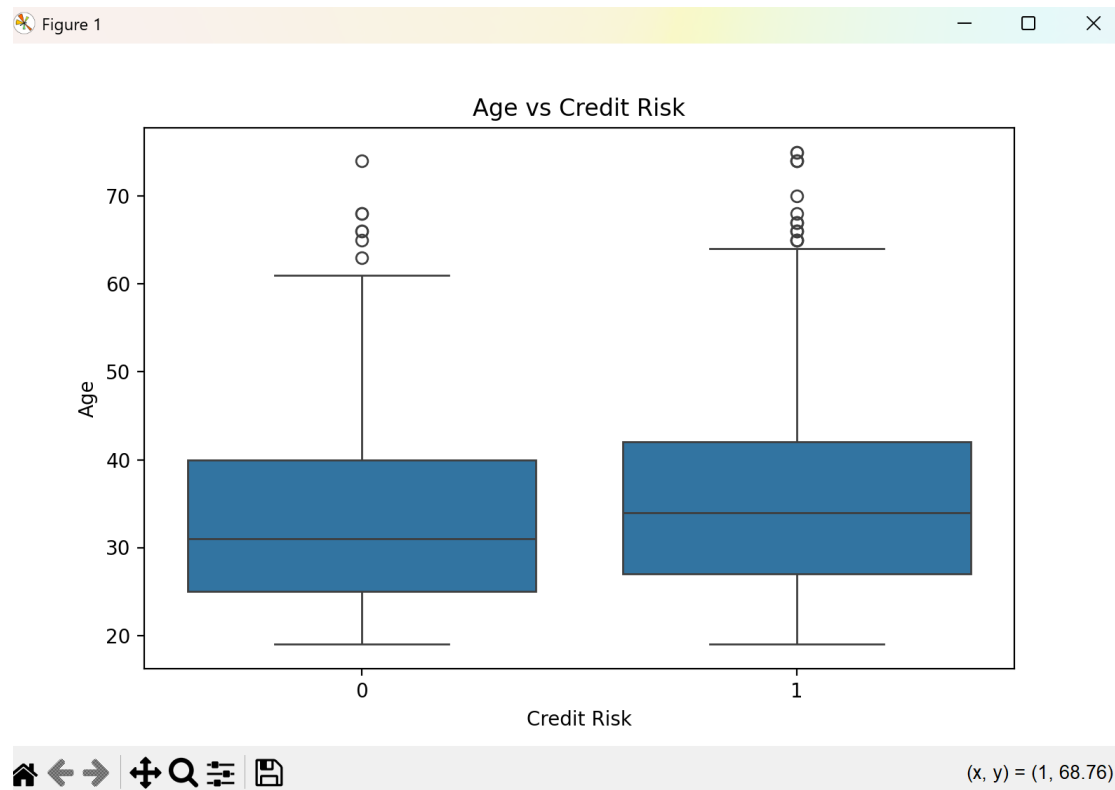


Η διάρκεια των δανείων παρουσιάζει:

- μεγαλύτερη συγκέντρωση στις τιμές 12 έως 24 μήνες
- παρουσία δανείων μεγαλύτερης διάρκειας
- ανομοιόμορφη κατανομή

Η διάρκεια αποτελεί σημαντικό παράγοντα κινδύνου, καθώς μεγαλύτερη διάρκεια συνεπάγεται μεγαλύτερη αβεβαιότητα.

## 5.7 Σχέση ηλικίας και πιστωτικού κινδύνου

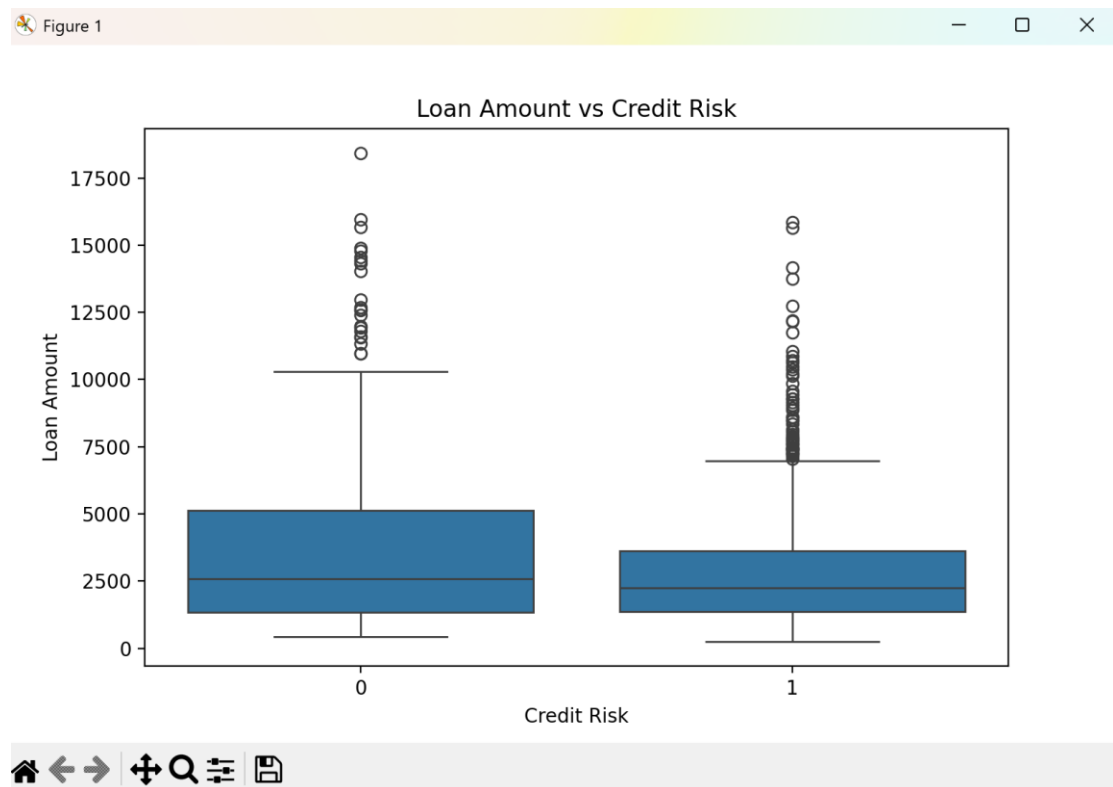


Το boxplot δείχνει ότι:

- οι τιμές ηλικίας στις δύο κατηγορίες είναι παρόμοιες
- υπάρχει ελαφρώς μεγαλύτερη διασπορά στην κατηγορία κινδύνου
- δεν παρατηρείται έντονη διαφοροποίηση

Συνεπώς, η ηλικία δεν αποτελεί τον πιο ισχυρό παράγοντα πρόβλεψης καθυστέρησης.

## 5.8 Σχέση ποσού δανείου και πιστωτικού κινδύνου

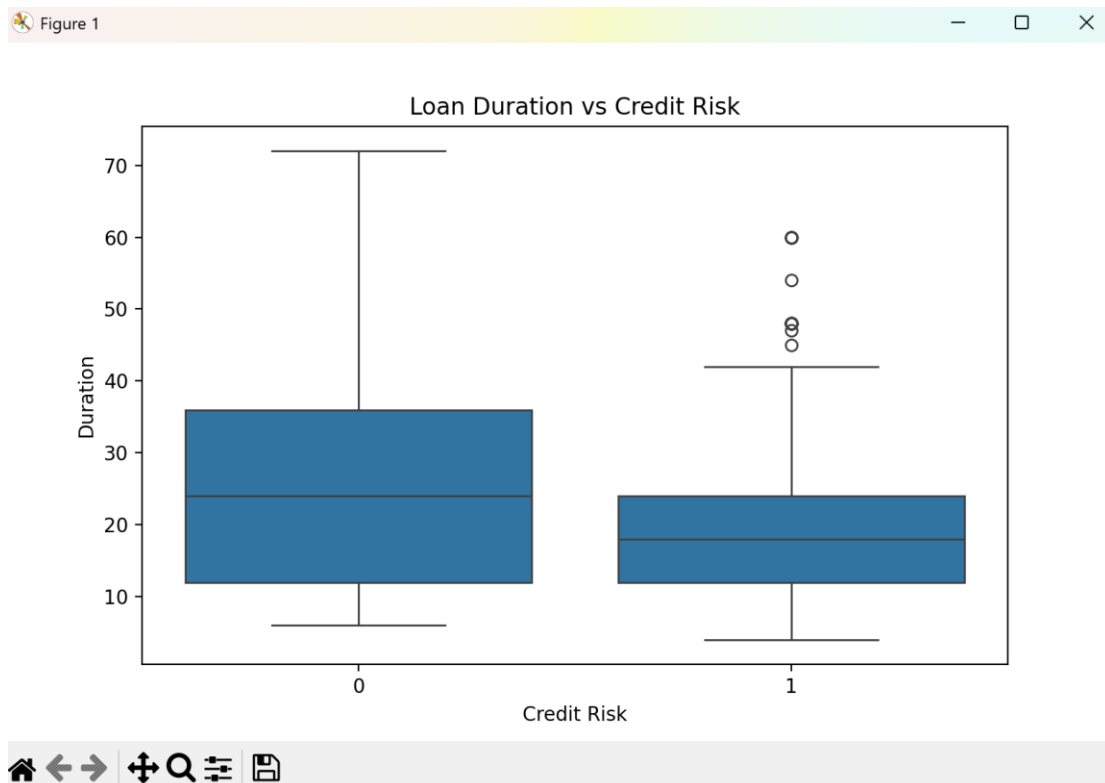


Το διάγραμμα δείχνει ότι:

- τα δάνεια υψηλού ποσού εμφανίζουν μεγαλύτερη διασπορά
- υπάρχουν περισσότερα outliers στην κατηγορία υψηλού κινδύνου
- τα μεγάλα δάνεια συνδέονται με αυξημένη πιθανότητα καθυστέρησης

Αυτό επιβεβαιώνει ότι το ποσό δανείου είναι σημαντικός δείκτης πιστωτικού κινδύνου.

## 5.9 Σχέση διάρκειας δανείου και πιστωτικού κινδύνου



Η ανάλυση δείχνει ότι:

- τα δάνεια μεγαλύτερης διάρκειας εμφανίζουν υψηλότερο κίνδυνο
- υπάρχει μεγαλύτερη διασπορά στην κατηγορία καθυστέρησης
- η διάρκεια σχετίζεται άμεσα με την πιθανότητα αθέτησης

Η διάρκεια αποτελεί έναν από τους πιο σημαντικούς παράγοντες κινδύνου.

## 5.10 Συνολική ανάλυση αποτελεσμάτων

Συνοψίζοντας τα αποτελέσματα:

- η πιθανότητα αθέτησης ανέρχεται περίπου στο 30%
- ο πιστωτικός κίνδυνος εξαρτάται από πολλούς παράγοντες
- οι σημαντικότεροι είναι:
  - διάρκεια δανείου
  - ποσό δανείου
- η ηλικία έχει μικρότερη επίδραση

Η ανάλυση επιβεβαιώνει ότι η χρήση στατιστικών μεθόδων μπορεί να προσφέρει αξιόπιστα αποτελέσματα για την αξιολόγηση του πιστωτικού κινδύνου.

Η ποσοτική ανάλυση της πιθανότητας αθέτησης πραγματοποιείται μέσω πινάκων συχνότητων, οι οποίοι παρουσιάζουν τη σχέση μεταξύ των βασικών χαρακτηριστικών των δανείων και της πιθανότητας καθυστέρησης. Οι πίνακες αυτοί επιτρέπουν την πιο σαφή κατανόηση της επίδρασης κάθε μεταβλητής στον πιστωτικό κίνδυνο.

ΠΙΝΑΚΑΣ 5.1 – Συνολική πιθανότητα αθέτησης

<b>Κατηγορία Δανείων</b>	<b>Σύνολο Δανείων</b>	<b>Καθυστερήσεις</b>	<b>Probability of Default</b>
Σύνολο	1000	300	30.00%

Πίνακας 5.2: Πιθανότητα αθέτησης ανά διάρκεια δανείου

<b>Διάρκεια Δανείου</b>	<b>Σύνολο</b>	<b>Probability of Default</b>
Short (0-12)	359	21.17%
Medium (13-24)	411	29.68%
Long (25+)	230	44.35%

Πίνακας 5.3: Πιθανότητα αθέτησης ανά ηλικία

<b>Ηλικιακή Ομάδα</b>	<b>Σύνολο</b>	<b>Probability of Default</b>
Young (<30)	409	36.19%
Middle (30-50)	478	25.31%
Older (50+)	113	27.43%

Πίνακας 5.4: Πιθανότητα αθέτησης ανά ποσό δανείου

<b>Ποσό Δανείου</b>	<b>Σύνολο</b>	<b>Probability of Default</b>
Small (0-2000)	432	28.01%
Medium (2000-5000)	380	26.58%

Ποσό Δανείου	Σύνολο	Probability of Default
Large (5000+)	188	41.49%

Πίνακας 5.5: Κατανομή δανειοληπτών ανά κατηγορία κινδύνου

Κατηγορία Κινδύνου	Αριθμός Δανειοληπτών
Low Risk	438
Medium Risk	377
High Risk	185

Πίνακας 5.6: Ποσοστό αθέτησης ανά κατηγορία κινδύνου

Risk Level	Σύνολο	Default Rate
Low Risk	438	21.46%
Medium Risk	377	32.36%
High Risk	185	45.41%

Από τους παραπάνω πίνακες προκύπτει ότι η πιθανότητα αθέτησης αυξάνεται σημαντικά όσο αυξάνεται η διάρκεια και το ποσό του δανείου. Επιπλέον, η κατηγοριοποίηση κινδύνου παρουσιάζει σαφή διαφοροποίηση μεταξύ των ομάδων, επιβεβαιώνοντας την αποτελεσματικότητα της προσέγγισης risk segmentation.

#### 5.11 Ανάλυση Odds Ratio

Η ανάλυση Odds Ratio αποτελεί μία σημαντική στατιστική τεχνική για την αξιολόγηση της σχέσης μεταξύ των χαρακτηριστικών των δανείων και της πιθανότητας αθέτησης. Το Odds Ratio εκφράζει το πόσο πιο πιθανό είναι να συμβεί ένα γεγονός σε μία κατηγορία σε σχέση με μία άλλη. Στην παρούσα εργασία, το Odds Ratio χρησιμοποιείται για τη σύγκριση της πιθανότητας καθυστέρησης μεταξύ διαφορετικών κατηγοριών δανείων.

Παράδειγμα: Διάρκεια Δανείου

Με βάση τα αποτελέσματα:

- Short loans → 21.17% default
- Long loans → 44.35% default

Υπολογίζεται ότι:

Τα δάνεια μεγάλης διάρκειας έχουν περίπου 2 φορές μεγαλύτερη πιθανότητα καθυστέρησης σε σχέση με τα δάνεια μικρής διάρκειας.

Παράδειγμα: Ποσό Δανείου

- Small loans → 28.01%
- Large loans → 41.49%

Τα μεγάλα δάνεια παρουσιάζουν σημαντικά αυξημένη πιθανότητα αθέτησης, επιβεβαιώνοντας ότι το ύψος του δανείου αποτελεί κρίσιμο παράγοντα κινδύνου.

Η ανάλυση Odds Ratio δείχνει ότι:

- η διάρκεια δανείου έχει ισχυρή επίδραση στον κίνδυνο
- το ποσό δανείου επηρεάζει σημαντικά την πιθανότητα καθυστέρησης
- ο πιστωτικός κίνδυνος αυξάνεται όσο αυξάνεται η «έκθεση» της τράπεζας.

Η χρήση του Odds Ratio ενισχύει την αξιοπιστία της ανάλυσης και επιβεβαιώνει τα αποτελέσματα που προέκυψαν από τα γραφήματα και τους πίνακες.

## 5.12 Παράδειγμα εφαρμογής

Για την καλύτερη κατανόηση της εφαρμογής του μοντέλου αξιολόγησης κινδύνου, παρουσιάζεται ένα ενδεικτικό παράδειγμα δανειολήπτη.

Έστω ένας δανειολήπτης με τα εξής χαρακτηριστικά:

- ηλικία: 25 ετών
- ποσό δανείου: 6000
- διάρκεια δανείου: 36 μήνες

Με βάση τα αποτελέσματα της ανάλυσης:

- ανήκει σε κατηγορία υψηλού ποσού
- ανήκει σε κατηγορία μεγάλης διάρκειας
- ανήκει σε νεαρή ηλικιακή ομάδα

Συνεπώς, ο συγκεκριμένος δανειολήπτης κατατάσσεται στην κατηγορία υψηλού κινδύνου (High Risk), καθώς συνδυάζει πολλούς παράγοντες που αυξάνουν την πιθανότητα καθυστέρησης.

#### 5.13 Σύγκριση αποτελεσμάτων με θεωρητικό πλαίσιο

Τα αποτελέσματα της παρούσας ανάλυσης επιβεβαιώνουν σε μεγάλο βαθμό τα θεωρητικά μοντέλα αξιολόγησης πιστωτικού κινδύνου που παρουσιάστηκαν στα προηγούμενα κεφάλαια. Συγκεκριμένα, διαπιστώθηκε ότι η διάρκεια του δανείου αποτελεί έναν από τους σημαντικότερους παράγοντες κινδύνου, γεγονός που συμφωνεί με τη θεωρία του credit scoring, σύμφωνα με την οποία τα μακροχρόνια δάνεια χαρακτηρίζονται από αυξημένη αβεβαιότητα αποπληρωμής. Παράλληλα, το ποσό του δανείου παρουσιάζει σημαντική επίδραση στην πιθανότητα αθέτησης, καθώς τα μεγαλύτερα δάνεια συνδέονται με αυξημένο πιστωτικό κίνδυνο. Το εύρημα αυτό επιβεβαιώνει τη σημασία της μεταβλητής Exposure at Default (EAD), όπως αναφέρεται στα διεθνή κανονιστικά πλαίσια.

Αντίθετα, η ηλικία του δανειολήπτη εμφανίζει μικρότερη επίδραση, γεγονός που υποδηλώνει ότι τα δημογραφικά χαρακτηριστικά δεν είναι πάντα καθοριστικοί παράγοντες κινδύνου, σε σύγκριση με τα οικονομικά χαρακτηριστικά. Συνολικά, τα αποτελέσματα της ανάλυσης βρίσκονται σε συμφωνία με τη διεθνή βιβλιογραφία και επιβεβαιώνουν τη χρησιμότητα των στατιστικών μεθόδων στην αξιολόγηση πιστωτικού κινδύνου.

#### 5.14 Πρακτικές επιπτώσεις για τις τράπεζες

Τα ευρήματα της παρούσας εργασίας έχουν σημαντικές πρακτικές εφαρμογές στον τραπεζικό τομέα και μπορούν να αξιοποιηθούν στη διαδικασία λήψης αποφάσεων. Αρχικά, η αναγνώριση των βασικών παραγόντων κινδύνου επιτρέπει στις τράπεζες να βελτιώσουν τα συστήματα αξιολόγησης δανειοληπτών, καθώς μέσω της κατηγοριοποίησης των πελατών σε επίπεδα κινδύνου, οι τράπεζες μπορούν να προσαρμόσουν τους όρους δανειοδότησης, όπως:

- επιτόκια
- όρια χρηματοδότησης
- διάρκεια αποπληρωμής

Επιπλέον, η εκτίμηση της πιθανότητας αθέτησης συμβάλλει στη μείωση των μη εξυπηρετούμενων δανείων, καθώς επιτρέπει τον έγκαιρο εντοπισμό δανειοληπτών υψηλού κινδύνου. Έτσι, η εφαρμογή τέτοιων μοντέλων μπορεί επίσης να βελτιώσει τη διαχείριση του χαρτοφυλακίου δανείων, ενισχύοντας τη σταθερότητα του τραπεζικού συστήματος.

#### 5.15 Περιορισμοί της ανάλυσης

Παρά τα σημαντικά ευρήματα της παρούσας μελέτης, υπάρχουν ορισμένοι περιορισμοί που πρέπει να ληφθούν υπόψη. Αρχικά, το dataset που χρησιμοποιήθηκε είναι σχετικά μικρό και δεν αντικατοπτρίζει πλήρως τη σύγχρονη πραγματικότητα του τραπεζικού συστήματος. Επιπλέον, δεν περιλαμβάνει σημαντικές μεταβλητές όπως το εισόδημα, η ανεργία ή άλλοι μακροοικονομικοί δείκτες, οι οποίοι επηρεάζουν σημαντικά τον πιστωτικό κίνδυνο. Επίσης, το μοντέλο που αναπτύχθηκε είναι απλοποιημένο και βασίζεται σε βασικές στατιστικές τεχνικές. Αν και αυτό αποτελεί πλεονέκτημα ως προς την κατανόηση και τη διαφάνεια, ενδέχεται να περιορίζει την ακρίβεια της πρόβλεψης. Τέλος, δεν πραγματοποιήθηκε σύγκριση με πιο σύνθετα μοντέλα, όπως αυτά της μηχανικής μάθησης, γεγονός που θα μπορούσε να προσφέρει πιο ολοκληρωμένη αξιολόγηση της απόδοσης της μεθοδολογίας.

#### 5.16 Σύγκριση με μεθόδους μηχανικής μάθησης

Η παρούσα εργασία βασίστηκε αποκλειστικά σε στατιστικές μεθόδους για την αξιολόγηση του πιστωτικού κινδύνου, χωρίς τη χρήση τεχνικών μηχανικής μάθησης. Ωστόσο, αξίζει να γίνει μία σύντομη σύγκριση μεταξύ των δύο προσεγγίσεων. Αρχικά, οι μέθοδοι μηχανικής μάθησης, όπως τα δέντρα απόφασης, τα νευρωνικά δίκτυα και οι αλγόριθμοι ταξινόμησης, χρησιμοποιούνται ευρέως στη σύγχρονη τραπεζική για την πρόβλεψη πιστωτικού κινδύνου. Οι μέθοδοι αυτές έχουν τη δυνατότητα να εντοπίζουν πολύπλοκες μη γραμμικές σχέσεις μεταξύ των μεταβλητών και συχνά επιτυγχάνουν υψηλότερη ακρίβεια πρόβλεψης. Αντίθετα, οι στατιστικές μέθοδοι που χρησιμοποιήθηκαν στην παρούσα εργασία παρουσιάζουν σημαντικά πλεονεκτήματα ως προς την απλότητα, τη διαφάνεια και την ερμηνευσιμότητα των αποτελεσμάτων. Έτσι, οι αποφάσεις που προκύπτουν είναι εύκολα κατανοητές και μπορούν να αιτιολογηθούν, γεγονός που είναι ιδιαίτερα σημαντικό στον τραπεζικό τομέα. Παρά

το γεγονός ότι οι μέθοδοι μηχανικής μάθησης ενδέχεται να προσφέρουν μεγαλύτερη ακρίβεια, η στατιστική προσέγγιση παραμένει ιδιαίτερα χρήσιμη σε περιπτώσεις όπου απαιτείται σαφής ερμηνεία των αποτελεσμάτων και διαφάνεια στη διαδικασία λήψης αποφάσεων.

## 6. Συμπεράσματα

Η παρούσα πτυχιακή εργασία είχε ως βασικό στόχο τη μελέτη και την πρόβλεψη της καθυστέρησης αποπληρωμής δανείων μέσω στατιστικών μεθόδων ανάλυσης δεδομένων, χωρίς τη χρήση τεχνικών μηχανικής μάθησης. Η επιλογή αυτής της προσέγγισης δεν έγινε τυχαία, αλλά βασίστηκε στην ανάγκη ανάπτυξης ενός μοντέλου αξιολόγησης που να είναι όχι μόνο αποτελεσματικό, αλλά και εύκολα κατανοητό, διαφανές και εφαρμόσιμο σε πραγματικές τραπεζικές συνθήκες. Σε ένα περιβάλλον όπου οι χρηματοπιστωτικοί οργανισμοί καλούνται καθημερινά να λαμβάνουν αποφάσεις υψηλής σημασίας σχετικά με τη χορήγηση δανείων, η δυνατότητα ερμηνείας των αποτελεσμάτων αποκτά ιδιαίτερη αξία. Για τον λόγο αυτό, η εργασία επικεντρώθηκε στη χρήση περιγραφικής στατιστικής, διερευνητικής ανάλυσης δεδομένων και τεχνικών κατηγοριοποίησης κινδύνου, με σκοπό την κατανόηση των βασικών παραγόντων που συνδέονται με την πιθανότητα αθέτησης δανειακών υποχρεώσεων.

Η ανάλυση βασίστηκε στο German Credit Dataset, ένα από τα πλέον γνωστά και αξιόπιστα σύνολα δεδομένων στη διεθνή βιβλιογραφία σχετικά με την αξιολόγηση πιστωτικού κινδύνου. Μέσα από τη μελέτη των δεδομένων προέκυψε μια συνολική εικόνα της συμπεριφοράς των δανειοληπτών, των χαρακτηριστικών των δανείων και των διαφορών μεταξύ πελατών χαμηλού και υψηλού κινδύνου. Τα αποτελέσματα έδειξαν ότι περίπου το 30% των δανειοληπτών εμφάνιζε καθυστέρηση αποπληρωμής, ποσοστό που θεωρείται ιδιαίτερα σημαντικό για ένα τραπεζικό χαρτοφυλάκιο και αναδεικνύει τη σημασία της έγκαιρης πρόβλεψης του πιστωτικού κινδύνου. Το εύρημα αυτό επιβεβαιώνει ότι ακόμη και σε ένα σχετικά ελεγχόμενο δείγμα δεδομένων, ένα σημαντικό ποσοστό πελατών ενδέχεται να αντιμετωπίσει δυσκολίες στην αποπληρωμή των υποχρεώσεών του, γεγονός που μπορεί να επηρεάσει άμεσα τη ρευστότητα, την κερδοφορία και τη σταθερότητα ενός τραπεζικού οργανισμού.

Μέσω της στατιστικής ανάλυσης και της οπτικοποίησης των δεδομένων εξετάστηκαν οι σημαντικότερες μεταβλητές που επηρεάζουν την πιθανότητα αθέτησης. Από την ανάλυση προέκυψε ότι η διάρκεια του δανείου αποτελεί έναν από τους βασικότερους παράγοντες πιστωτικού κινδύνου. Τα δάνεια μεγαλύτερης διάρκειας εμφάνισαν σημαντικά υψηλότερα ποσοστά καθυστέρησης, γεγονός που

μπορεί να εξηγηθεί από την αυξημένη αβεβαιότητα που δημιουργεί ο μεγάλος χρονικός ορίζοντας αποπληρωμής. Όσο αυξάνεται η διάρκεια ενός δανείου, τόσο αυξάνονται και οι πιθανότητες μεταβολής της οικονομικής κατάστασης του δανειολήπτη, όπως απώλεια εισοδήματος, ανεργία, αύξηση οικογενειακών υποχρεώσεων ή γενικότερες οικονομικές κρίσεις. Το εύρημα αυτό συμφωνεί με τη διεθνή τραπεζική πρακτική, σύμφωνα με την οποία τα μακροχρόνια δάνεια θεωρούνται συνήθως πιο επισφαλής σε σχέση με τα βραχυπρόθεσμα.

Παράλληλα, ιδιαίτερα σημαντική αποδείχθηκε και η επίδραση του ποσού του δανείου στον πιστωτικό κίνδυνο. Τα μεγαλύτερα δανειακά ποσά συνδέθηκαν με αυξημένες πιθανότητες καθυστέρησης αποπληρωμής, καθώς συνεπάγονται υψηλότερη οικονομική επιβάρυνση για τον δανειολήπτη. Σε πραγματικές τραπεζικές συνθήκες, ένα υψηλό ποσό δανείου αυξάνει την πιθανότητα οικονομικής πίεσης, ιδιαίτερα όταν συνδυάζεται με ασταθές εισόδημα ή αυξημένες λοιπές υποχρεώσεις. Η παρατήρηση αυτή αναδεικνύει τη σημασία της σωστής αξιολόγησης της αποπληρωτικής ικανότητας πριν από τη χορήγηση χρηματοδότησης και επιβεβαιώνει ότι η σχέση μεταξύ εισοδήματος και δανειακής επιβάρυνσης αποτελεί κρίσιμο στοιχείο της πιστοδοτικής διαδικασίας.

Αντίθετα, η ηλικία των δανειοληπτών δεν φάνηκε να επηρεάζει τόσο έντονα τον πιστωτικό κίνδυνο όσο άλλες μεταβλητές. Παρότι οι νεότεροι δανειολήπτες εμφάνισαν ελαφρώς αυξημένες πιθανότητες καθυστέρησης, η επίδραση της ηλικίας δεν ήταν αρκετά ισχυρή ώστε να θεωρηθεί αυτόνομος παράγοντας πρόβλεψης. Το αποτέλεσμα αυτό δείχνει ότι ο πιστωτικός κίνδυνος δεν μπορεί να ερμηνευτεί μονοδιάστατα, αλλά εξαρτάται από συνδυασμό κοινωνικών, οικονομικών και χρηματοοικονομικών χαρακτηριστικών. Ένας νεότερος δανειολήπτης μπορεί να παρουσιάζει χαμηλό κίνδυνο εφόσον διαθέτει σταθερό εισόδημα και θετικό πιστωτικό ιστορικό, ενώ αντίστοιχα ένας μεγαλύτερος σε ηλικία πελάτης μπορεί να εμφανίζει αυξημένο κίνδυνο λόγω υπερχρέωσης ή μειωμένης οικονομικής σταθερότητας. Η παρατήρηση αυτή επιβεβαιώνει τον πολυπαραγοντικό χαρακτήρα του πιστωτικού κινδύνου και τη σημασία της ολιστικής αξιολόγησης των πελατών.

Σημαντικό μέρος της εργασίας αποτέλεσε επίσης η ανάπτυξη ενός απλού μοντέλου κατηγοριοποίησης κινδύνου (risk segmentation model), μέσω του οποίου οι δανειολήπτες ταξινομήθηκαν σε κατηγορίες χαμηλού, μεσαίου και υψηλού

κινδύνου. Το μοντέλο βασίστηκε σε βασικά χαρακτηριστικά, όπως η διάρκεια και το ποσό του δανείου, και επέτρεψε τη δημιουργία ενός συστήματος βαθμολόγησης που μπορεί να χρησιμοποιηθεί ως υποστηρικτικό εργαλείο στη διαδικασία λήψης αποφάσεων. Η ανάλυση έδειξε ότι οι κατηγορίες υψηλού κινδύνου εμφάνιζαν σημαντικά μεγαλύτερα ποσοστά αθέτησης σε σχέση με τις υπόλοιπες ομάδες, γεγονός που επιβεβαιώνει τη χρησιμότητα της κατηγοριοποίησης στη διαχείριση χαρτοφυλακίων δανείων. Στην πράξη, τέτοιου είδους μοντέλα μπορούν να βοηθήσουν τις τράπεζες να προσαρμόζουν τα επιτόκια, τα πιστωτικά όρια και τις διαδικασίες παρακολούθησης ανάλογα με το επίπεδο κινδύνου κάθε πελάτη.

Ένα ιδιαίτερα σημαντικό συμπέρασμα της εργασίας είναι ότι ακόμη και χωρίς τη χρήση σύνθετων αλγορίθμων τεχνητής νοημοσύνης ή machine learning, η στατιστική ανάλυση μπορεί να προσφέρει αξιόπιστα και ουσιαστικά αποτελέσματα. Σε πολλές περιπτώσεις, τα απλά και ερμηνεύσιμα μοντέλα παρουσιάζουν σημαντικά πλεονεκτήματα, καθώς επιτρέπουν στους αναλυτές και στα τραπεζικά στελέχη να κατανοούν με σαφήνεια τους παράγοντες που επηρεάζουν τον πιστωτικό κίνδυνο. Η διαφάνεια αυτή είναι ιδιαίτερα σημαντική σε ένα αυστηρά εποπτευόμενο τραπεζικό περιβάλλον, όπου οι αποφάσεις πρέπει να μπορούν να αιτιολογηθούν τόσο σε εποπτικές αρχές όσο και στους ίδιους τους πελάτες.

Παρά τη χρησιμότητα των αποτελεσμάτων, η παρούσα μελέτη παρουσιάζει και ορισμένους περιορισμούς. Το μέγεθος του dataset είναι σχετικά μικρό, καθώς περιλαμβάνει 1.000 εγγραφές, γεγονός που περιορίζει τη δυνατότητα γενίκευσης των συμπερασμάτων σε πολύ μεγάλα τραπεζικά χαρτοφυλάκια. Επιπλέον, το dataset δεν περιλαμβάνει σύγχρονες μακροοικονομικές μεταβλητές, όπως επιτόκια, πληθωρισμό ή δείκτες ανεργίας, οι οποίοι επηρεάζουν άμεσα τη δυνατότητα αποπληρωμής των δανειοληπτών. Παράλληλα, το μοντέλο scoring που αναπτύχθηκε είναι σχετικά απλοποιημένο και δεν λαμβάνει υπόψη πιο σύνθετες αλληλεπιδράσεις μεταξύ μεταβλητών. Ωστόσο, οι περιορισμοί αυτοί δεν μειώνουν τη σημασία της έρευνας, αλλά αντίθετα αναδεικνύουν πεδία στα οποία μπορεί να επεκταθεί η μελλοντική μελέτη.

Στο πλαίσιο αυτό, η εργασία μπορεί να αποτελέσει βάση για περαιτέρω ερευνητική ανάπτυξη στον τομέα της αξιολόγησης πιστωτικού κινδύνου. Μελλοντικές έρευνες θα μπορούσαν να αξιοποιήσουν μεγαλύτερα και πιο σύγχρονα

datasets, να ενσωματώσουν μακροοικονομικούς παράγοντες και να συγκρίνουν τις παραδοσιακές στατιστικές μεθόδους με τεχνικές μηχανικής μάθησης, όπως logistic regression, random forests ή neural networks. Μια τέτοια σύγκριση θα μπορούσε να προσφέρει πιο ολοκληρωμένη εικόνα σχετικά με την ακρίβεια, τη διαφάνεια και την πρακτική εφαρμοσιμότητα των διαφορετικών προσεγγίσεων. Επιπλέον, θα μπορούσαν να αναπτυχθούν ολοκληρωμένα συστήματα υποστήριξης αποφάσεων που να συνδυάζουν στατιστική ανάλυση, risk scoring και αυτοματοποιημένες διαδικασίες αξιολόγησης.

Συνολικά, η εργασία ανέδειξε τη μεγάλη σημασία της ανάλυσης δεδομένων στη σύγχρονη τραπεζική πρακτική και επιβεβαίωσε ότι η σωστή αξιολόγηση του πιστωτικού κινδύνου αποτελεί βασική προϋπόθεση για τη σταθερότητα του χρηματοπιστωτικού συστήματος. Η δυνατότητα έγκαιρης αναγνώρισης δανειοληπτών με αυξημένο κίνδυνο επιτρέπει στις τράπεζες να περιορίζουν τις επισφάλειες, να διαχειρίζονται αποτελεσματικότερα τα χαρτοφυλάκιά τους και να λαμβάνουν πιο ορθολογικές αποφάσεις χρηματοδότησης. Παράλληλα, η χρήση απλών, κατανοητών και διαφανών μεθόδων καθιστά τη στατιστική ανάλυση ιδιαίτερα χρήσιμη όχι μόνο σε ακαδημαϊκό επίπεδο αλλά και σε πραγματικές εφαρμογές του τραπεζικού τομέα. Συνεπώς, η παρούσα εργασία επιβεβαιώνει ότι ακόμη και βασικές τεχνικές ανάλυσης δεδομένων μπορούν να συμβάλουν ουσιαστικά στην κατανόηση και στη διαχείριση του πιστωτικού κινδύνου, προσφέροντας πρακτικά και αξιόπιστα εργαλεία για τη σύγχρονη τραπεζική λειτουργία.

## Βιβλιογραφία

- Arhinful, R., Mensah, L., Gyamfi, B. A., & Obeng, H. A. (2025). The Impact of Non-Performing Loans on Bank Growth: The Moderating Roles of Bank Size and Capital Adequacy Ratio—Evidence from U.S. Banks. *International Journal of Financial Studies*, 13(3), 165. <https://doi.org/10.3390/ijfs13030165>
- Avgouleas, E., & Goodhart, C. (2015). Critical reflections on bank bail-ins. *Journal of Financial Regulation*, 1(1), 3–29. <https://doi.org/10.1093/jfr/fju009>
- Bednarik, Z., & Marshall, M.I. (2024). Personal relationships of rural small businesses with community banks in times of crisis. *Journal of Small Business and Enterprise Development*, 31, 5, 881-904, <https://doi.org/10.1108/JSBED-04-2023-0196>.
- Bhandary, R. & Ghosh, B. (2025). Credit Card Default Prediction: An Empirical Analysis on Predictive Performance Using Statistical and Machine Learning Methods. *Journal of Risk and Financial Management*, 18. [10.3390/jrfm18010023](https://doi.org/10.3390/jrfm18010023).
- Bhatt, T. K., Ahmed, N., Iqbal, M. B., & Ullah, M. (2023). Examining the Determinants of Credit Risk Management and Their Relationship with the Performance of Commercial Banks in Nepal. *Journal of Risk and Financial Management*, 16(4), 235. <https://doi.org/10.3390/jrfm16040235>
- Blake, D. (2022). The Great Game Will Never End: Why the Global Financial Crisis Is Bound to Be Repeated. *Journal of Risk and Financial Management*, 15(6), 245. <https://doi.org/10.3390/jrfm15060245>
- Chang, V., Sivakulasingam, S., Wang, H., Wong, S. T., Ganatra, M. A., & Luo, J. (2024). Credit Risk Prediction Using Machine Learning and Deep Learning: A Study on Credit Card Customers. *Risks*, 12(11), 174. <https://doi.org/10.3390/risks12110174>
- Chong, F. (2021). Loan Delinquency: Some Determining Factors. *Journal of Risk and Financial Management*, 14, 320. [10.3390/jrfm14070320](https://doi.org/10.3390/jrfm14070320).
- Dedeloudis, G., Lois, P., & Repousis, S. (2025). Banking Supervision and Risk Management in Times of Crisis: Evidence from Greece's Systemic Banks (2015–2024). *Journal of Risk and Financial Management*, 18(7), 386. <https://doi.org/10.3390/jrfm18070386>

- Engan, A., Hjelkrem, L. O., & Ristad, M. (2026). A review of credit risk assessment in banking — From deterministic models to probabilistic AI. *Finance Research Open*, 2(2), 100125. <https://doi.org/10.1016/j.finr.2026.100125>
- Financial Stability Board. (2009). Principles for sound stress testing practices and supervision. [https://www.fsb.org/uploads/r\\_0910a.pdf](https://www.fsb.org/uploads/r_0910a.pdf)
- Firdavs, U., & Wang, W. (2026). Credit Risk Modeling in Banking: A Comparative Analysis of Logistic Regression and Machine Learning Approaches. *Journal of Computer and Communications*, 14, 155-180. doi: 10.4236/jcc.2026.144008.
- Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from Class-Imbalanced Data: Review of Methods and Applications. *Elsevier*, 73, 220-239. 10.1016/j.eswa.2016.12.035
- Kabir, M. A., Ahmed, F., Islam, Md. M. & Ahmed, Md. R. (2024). Python For Data Analytics: A Systematic Literature Review Of Tools, Techniques, And Applications. *Academic Journal of Science, Technology, Engineering & Mathematics Education*, 4. 134-154. 10.69593/ajsteme.v4i04.146.
- Kennedy, K., Mac Namee, B., Delany, S.J., O’Sullivan, M., & Watson, N. (2013). A window of opportunity: Assessing behavioural scoring. *Expert Systems with Applications*, 40, 4, 1372-1380, <https://doi.org/10.1016/j.eswa.2012.08.052>.
- Li, K., Zhou, F., Li, Z., Yao, X., & Zhang, Y. (2021). Predicting loss given default using post-default information. *Knowledge-Based Systems*, 224, 107068. <https://doi.org/10.1016/j.knosys.2021.107068>.
- Louzis, D. P., Vouldis, A. T., & Metaxas, V. L. (2012). Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios. *Journal of Banking & Finance*, 36(4), 1012–1027. <https://doi.org/10.1016/j.jbankfin.2011.10.012>
- Markov, A., Seleznyova, Z., Victor Lapshin, V. (2022). Credit scoring methods: Latest trends and points to consider. *The Journal of Finance and Data Science*, 8, 180-201. <https://doi.org/10.1016/j.jfds.2022.07.002>.
- Maruf, A. A., Kowsar, M.M, Mohiuddin, M., & Mohna, H. (2024). Behavioral factors in loan default prediction: a literature review on psychological and socioeconomic risk indicators. *American Journal of Advanced Technology and Engineering Solutions*, 04, 43-70. 10.63125/0jwbtbn29.

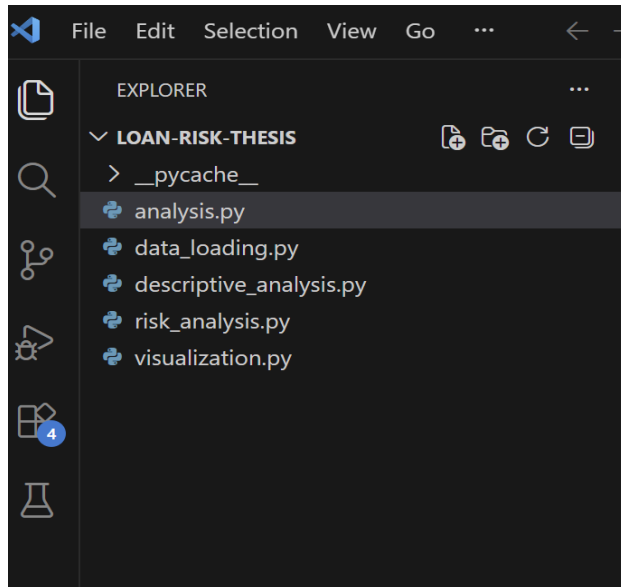
- Rahman, Md & Sarker, Awwal. (2024). The Impact of Credit Risk and Bank-Specific Variables on Financial Performance of the Listed Commercial Banks in Bangladesh. *International Journal of Research and Innovation in Social Science (IJRISS)*, vol. 8, 3826-3839.
- Spuchl'áková, E., Valašková, K., Adamko, P. (2015). The Credit Risk and its Measurement. Hedging and Monitoring, *Procedia Economics and Finance*, 24, 675-681, [https://doi.org/10.1016/S2212-5671\(15\)00671-1](https://doi.org/10.1016/S2212-5671(15)00671-1).
- Ζωγραφάκης, Σ. & Σπαθής, Π. (2011). «Οικονομική κρίση και ελληνική οικονομία: το τέλος των αποκλίσεων και επιπτώσεις, Στο *Τράπεζα της Ελλάδας, Η ελληνική αγορά εργασίας: χαρακτηριστικά, εξελίξεις και προκλήσεις*. Ομιλίες Ημερίδας 22 Μαρτίου 2010.
- Μυλωνάς, Κ. (2018). *Στατιστική ανάλυση επαγωγών στις κοινωνικές επιστήμες*. Πεδίο.
- Στεργίου, Λ. (2008). Πώς ξεκίνησε η Οικονομική Κρίση του 2008. *Οικονομικό ένθετο της Εφημερίδας Καθημερινή* (26-10-2008). [http://news.kathimerini.gr/4dcgi/\\_w\\_articles\\_economyagor\\_1\\_26/10/2008\\_289798](http://news.kathimerini.gr/4dcgi/_w_articles_economyagor_1_26/10/2008_289798)

#### *Ιστότοποι*

- Bank of Greece (2016). *Η ελληνική οικονομία και το τραπεζικό σύστημα στην περίοδο της κρίσης*. <https://www.bankofgreece.gr/enimerosi/grafeio-typoy/anazhthsh-enhmerwsewn/enhmerwseis?announcement=88fb89ab-4d90-4d4b-8170-2c9dfd81f626>
- Bank of Greece (2024). *Financial stability report 2024*. <https://www.bankofgreece.gr>
- European Banking Authority (2019). *Guidelines on management of non-performing and forborne exposures*. <https://www.eba.europa.eu>
- European Central Bank (2021). *Guidance to banks on non-performing loans*. <https://www.ecb.europa.eu>

## Παράρτημα

### Παράρτημα A.1 – Δομή φακέλου εργασίας



- analysis.py
- data\_loading.py
- descriptive\_analysis.py
- visualization.py
- risk\_analysis.py

### Παράρτημα A.2 – Κώδικας φόρτωσης δεδομένων

```
import kagglehub
from kagglehub import KaggleDatasetAdapter

def load_data():
    file_path = "german_credit_data.csv"

    df = kagglehub.load_dataset(
        KaggleDatasetAdapter.PANDAS,
        "varunchawla30/german-credit-data",
        file_path
    )

    return df
```

### Παράρτημα Α.3 – Κώδικας περιγραφικής ανάλυσης

```
def dataset_info(df):
    print("\nDataset Info:")
    print(df.info())

def statistical_summary(df):
    print("\nStatistical Summary:")
    print(df.describe())

def risk_distribution(df):
    print("\nCredit Risk Distribution:")
    print(df["kredit"].value_counts())
```

### Παράρτημα Α.4 – Κώδικας οπτικοποίησης δεδομένων

```
import matplotlib.pyplot as plt
import seaborn as sns

def plot_risk_distribution(df):
    plt.figure(figsize=(8, 5))
    sns.countplot(x="kredit", data=df)
    plt.title("Distribution of Credit Risk")
    plt.xlabel("Credit Risk")
    plt.ylabel("Count")
    plt.show()

def correlation_heatmap(df):
    plt.figure(figsize=(12, 8))
    sns.heatmap(df.corr(), annot=False, cmap="coolwarm")
    plt.title("Correlation Matrix of Credit Variables")
    plt.show()

def age_distribution(df):
    plt.figure(figsize=(8, 5))
    sns.histplot(df["alter"], bins=20, kde=True)
    plt.title("Distribution of Borrower Age")
    plt.xlabel("Age")
    plt.ylabel("Frequency")
    plt.show()
```

```

def loan_amount_distribution(df):
    plt.figure(figsize=(8, 5))
    sns.histplot(df["hoehe"], bins=20, kde=True)
    plt.title("Distribution of Loan Amount")
    plt.xlabel("Loan Amount")
    plt.ylabel("Frequency")
    plt.show()

def loan_duration_distribution(df):
    plt.figure(figsize=(8, 5))
    sns.histplot(df["laufzeit"], bins=20, kde=True)
    plt.title("Distribution of Loan Duration")
    plt.xlabel("Duration (months)")
    plt.ylabel("Frequency")
    plt.show()

def age_vs_risk(df):
    plt.figure(figsize=(8, 5))
    sns.boxplot(x="kredit", y="alter", data=df)
    plt.title("Age vs Credit Risk")
    plt.xlabel("Credit Risk")
    plt.ylabel("Age")
    plt.show()

def loan_amount_vs_risk(df):
    plt.figure(figsize=(8, 5))
    sns.boxplot(x="kredit", y="hoehe", data=df)
    plt.title("Loan Amount vs Credit Risk")
    plt.xlabel("Credit Risk")
    plt.ylabel("Loan Amount")
    plt.show()

def loan_duration_vs_risk(df):
    plt.figure(figsize=(8, 5))
    sns.boxplot(x="kredit", y="laufzeit", data=df)
    plt.title("Loan Duration vs Credit Risk")
    plt.xlabel("Credit Risk")
    plt.ylabel("Duration")
    plt.show()

```

## Παράρτημα Α.5 – Κώδικας ανάλυσης πιστωτικού κινδύνου

```
import pandas as pd

def probability_of_default(df):

    print("\nOverall Probability of Default:")

    total = len(df)
    defaults = len(df[df["kredit"] == 0])

    pd_rate = defaults / total

    print(f"Total loans: {total}")
    print(f"Defaults: {defaults}")
    print(f"Probability of Default: {pd_rate:.2%}")

def pd_by_loan_duration(df):

    print("\nProbability of Default by Loan Duration Group:")

    df = df.copy()

    df["duration_group"] = pd.cut(
        df["laufzeit"],
        bins=[0, 12, 24, 72],
        labels=["Short (0-12)", "Medium (13-24)", "Long (25+)"]
    )

    pd_table = df.groupby("duration_group")["kredit"].apply(
        lambda x: (x == 0).mean()
    )

    counts = df["duration_group"].value_counts().sort_index()

    result = pd.DataFrame({
        "Total Loans": counts,
        "Probability of Default": pd_table
    })

    result["Probability of Default"] = result["Probability of Default"].apply(
        lambda x: f"{x:.2%}"
    )

    print(result)
```

```

def pd_by_age_group(df):

    print("\nProbability of Default by Age Group:")

    df = df.copy()

    df["age_group"] = pd.cut(
        df["alter"],
        bins=[18, 30, 50, 100],
        labels=["Young (<30)", "Middle (30-50)", "Older (50+)"]
    )

    pd_table = df.groupby("age_group")["kredit"].apply(
        lambda x: (x == 0).mean()
    )

    counts = df["age_group"].value_counts().sort_index()

    result = pd.DataFrame({
        "Total Loans": counts,
        "Probability of Default": pd_table
    })

    result["Probability of Default"] = result["Probability of Default"].apply(
        lambda x: f"{x:.2%}"
    )

    print(result)

def pd_by_loan_amount(df):

    print("\nProbability of Default by Loan Amount:")

    df = df.copy()

    df["loan_amount_group"] = pd.cut(
        df["hoehe"],
        bins=[0, 2000, 5000, 20000],
        labels=["Small (0-2000)", "Medium (2000-5000)", "Large (5000+)"]
    )

    pd_table = df.groupby("loan_amount_group")["kredit"].apply(
        lambda x: (x == 0).mean()
    )

    counts = df["loan_amount_group"].value_counts().sort_index()

    result = pd.DataFrame({

```

```

    "Total Loans": counts,
    "Probability of Default": pd_table
})

result["Probability of Default"] = result["Probability of Default"].apply(
    lambda x: f"{x:.2%}"
)

print(result)

def risk_segmentation(df):

    print("\nRisk Segmentation Results:")

    df = df.copy()

    df["risk_score"] = 0

    df.loc[df["laufzeit"] > 24, "risk_score"] += 1
    df.loc[df["alter"] < 30, "risk_score"] += 1
    df.loc[df["hoehe"] > 5000, "risk_score"] += 1

    df["risk_level"] = pd.cut(
        df["risk_score"],
        bins=[-1, 0, 1, 3],
        labels=["Low Risk", "Medium Risk", "High Risk"]
    )

    result = df["risk_level"].value_counts()

    print(result)

def default_rate_by_risk(df):

    print("\nDefault Rate by Risk Level:")

    df = df.copy()

    df["risk_score"] = 0

    df.loc[df["laufzeit"] > 24, "risk_score"] += 1
    df.loc[df["alter"] < 30, "risk_score"] += 1
    df.loc[df["hoehe"] > 5000, "risk_score"] += 1

    df["risk_level"] = pd.cut(
        df["risk_score"],
        bins=[-1, 0, 1, 3],

```

```

    labels=["Low Risk", "Medium Risk", "High Risk"]
)

default_rate = df.groupby("risk_level")["kredit"].apply(
    lambda x: (x == 0).mean()
)

counts = df["risk_level"].value_counts().sort_index()

result = pd.DataFrame({
    "Total Borrowers": counts,
    "Default Rate": default_rate
})

result["Default Rate"] = result["Default Rate"].apply(
    lambda x: f"{x:.2%}"
)

print(result)

```

#### Παράρτημα Α.6 – Κεντρικό αρχείο εκτέλεσης

```

from data_loading import load_data
from descriptive_analysis import dataset_info, statistical_summary, risk_distribution

from visualization import (
    plot_risk_distribution,
    correlation_heatmap,
    age_distribution,
    loan_amount_distribution,
    loan_duration_distribution,
    age_vs_risk,
    loan_amount_vs_risk,
    loan_duration_vs_risk
)

from risk_analysis import (
    probability_of_default,
    pd_by_loan_duration,
    pd_by_age_group,
    pd_by_loan_amount,
    risk_segmentation,
    default_rate_by_risk
)

# Load dataset
df = load_data()

```

```
# Descriptive statistics
dataset_info(df)
statistical_summary(df)
risk_distribution(df)

# Visualizations
plot_risk_distribution(df)
correlation_heatmap(df)
age_distribution(df)
loan_amount_distribution(df)
loan_duration_distribution(df)
age_vs_risk(df)
loan_amount_vs_risk(df)
loan_duration_vs_risk(df)

# Credit Risk Analysis
probability_of_default(df)
pd_by_loan_duration(df)
pd_by_age_group(df)
pd_by_loan_amount(df)

# Risk Segmentation
risk_segmentation(df)

# Model Validation
default_rate_by_risk(df)
```

