

**WYDZIAŁ ZARZĄDZANIA**

Samodzielna Katedra Zastosowań Matematyki w Ekonomii

Praca licencjacka

*Wykorzystanie analizy dyskryminacyjnej do oceny sytuacji finansowej przedsiębiorstw*

*Use of discriminant analysis to assess the financial situation of enterprises*

Autor: *Maciej Śmiałkowski*

Kierunek studiów: *Informatyka i Ekonometria*

Opiekun pracy: *dr Łukasz Lach*

Kraków, 2020

Spis treści

[1.Wstęp 3](#_Toc44021395)

[1.1 . Aktualność tematu 5](#_Toc44021396)

[1.2 . Cel pracy oraz hipotezy badawcze 6](#_Toc44021397)

[1.3. Przedstawienie metod i rozwiązań najczęściej stosowanych do rozwiązywania tego typu zagadnień 7](#_Toc44021398)

[1.4. Szczegółowy opis przypadków użycia metody 10](#_Toc44021399)

[1.5. Narzędzia wybrane do przeprowadzenia metody 12](#_Toc44021400)

[1.6. Struktura pracy 13](#_Toc44021401)

[1.7. Źródła informacji 14](#_Toc44021402)

[2. Badania Empiryczne 15](#_Toc44021403)

[2.1. Charakterystyka obiektu badań 15](#_Toc44021404)

[Przedsiębiorstwo 15](#_Toc44021405)

[Sektor małych i średnich przedsiębiorstw 17](#_Toc44021406)

[Charakterystyka gospodarki Słowackiej 18](#_Toc44021407)

[Opis danych 21](#_Toc44021408)

[2.2. Analiza literatury 23](#_Toc44021409)

[2.3. Rozdział Teoretyczny 24](#_Toc44021410)

[Geneza 24](#_Toc44021411)

[Istota 25](#_Toc44021412)

[Przykład 29](#_Toc44021413)

[2.4 Rozdział Empiryczny 30](#_Toc44021414)

[Zbiór danych 30](#_Toc44021415)

[Podstawowe statystyki opisowe 30](#_Toc44021416)

[Korelacja 33](#_Toc44021417)

[Implementacja oraz przeprowadzenie metody 35](#_Toc44021418)

[Test Shapiro-Wilka 36](#_Toc44021419)

[Usunięcie obserwacji odstających 39](#_Toc44021420)

[Transformacja danych w celu przybliżenia rozkładów do rozkładu normalnego 44](#_Toc44021421)

[Standaryzacja danych 46](#_Toc44021422)

[Współliniowość zmiennych objaśniających 46](#_Toc44021423)

[2.5 Wyniki eksperymentu badawczego 47](#_Toc44021424)

[2.6 Ocena efektywności 53](#_Toc44021425)

[2.7 Możliwe dalsze kierunki badań 54](#_Toc44021426)

[3. Podsumowanie 54](#_Toc44021427)

[Wnioski końcowe 55](#_Toc44021428)

[Literatura 55](#_Toc44021429)

[Pozycje zwarte 55](#_Toc44021430)

[Strony internetowe 57](#_Toc44021431)

[Akty prawne 57](#_Toc44021432)

[Ilustracje, schematy i wykresy 58](#_Toc44021433)

[Spis ilustracji 58](#_Toc44021434)

[Spis schematów 58](#_Toc44021435)

[Spis wykresów 58](#_Toc44021436)

[Wykorzystanie technik predykcji upadłości do badań prowadzonych na świecie 58](#_Toc44021437)

# 1.Wstęp

Czasy w jakich żyjemy to czasy wzrostu gospodarczego. Wzrost ten jest nieporównywalnie większy niż w poprzednich latach. Wynika to zapewne z faktu, że XXI wiek to okres niebywałego rozwoju technologicznego. To jak zmienia się technologia oraz postrzeganie otaczającego nas świata, sprawia że jednostki organizacyjne nastawione na zysk muszą dostosowywać się do zmian na rynku w zatrważającym tempie. Nie wszystkim z nich to się uda, a te które nie podołają temu zadaniu zbankrutują i znikną z rynku. Gdy jednostka organizacyjna jaką jest przedsiębiorstwo bankrutuje pojawia się wiele problemów. Do problemów tych należą między innymi niewypełnione zobowiązania finansowe. Co więcej wielu ludzi może stracić pracę a inne jednostki organizacyjne mogą stracić odbiorców tudzież dostawców co implikuje kolejne trudności w najbliższym otoczeniu. Nie dziwi zatem fakt, że powstało tak wiele metod i narzędzi prognozujących upadanie spółek. Główną dziedziną nauki wykorzystywaną w predykcji bankructwa jest statystyka. Metody statystyczne sprawdzają się wyśmienicie w tej dziedzinie. Analizując dane i obrabiając je przy pomocy narzędzi statystycznych, jesteśmy w stanie przewidzieć nadchodzące zagrożenia.

Temat tej pracy porusza aspekt jakim jest ocena finansowa przedsiębiorstw. Przedsiębiorstwo jest to wyodrębniona prawnie i ekonomicznie jednostka gospodarcza, której działania ukierunkowane są na świadczenie usług czy też produkcje dóbr[[1]](#footnote-1). W dzisiejszych czasach istnieje ogromna liczba przedsiębiorstw. Konkurują one o klientów, a także o pozycje na rynku, ich głównymi celami oraz zmiennymi świadczącymi o tym jak dobrze prosperują jest utrzymanie płynności finansowej, zysk oraz zwiększenie wartości. Przewidywanie trudnej sytuacji finansowej i bankructwa firmy jest przedmiotem dużego zainteresowania badaczy na całym świecie już od XX wieku. Największą implikacją mającą wpływ na dynamikę tych badań był światowy kryzys przypadający na lata dwudzieste oraz trzydzieste XX wieku. W celu przewidywania bankructwa wykorzystywano przeróżne wskaźniki finansowe i starano się dobrać zmienne tak by zdolności prognostyczne modelu były jak największe.

Badania jakie zostaną przeprowadzone w niniejszej pracy do prognozowania złej kondycji finansowej wykorzystają metodę statystyczną jaką jest analiza dyskryminacyjna. Sama idea metodologiczna jest bardzo stara gdyż liniową funkcję dyskryminacyjną po raz pierwszy wykorzystano do klasyfikacji obiektów już w roku 1936. Natomiast jej pierwsze wykorzystanie do analizy finansowej miało miejsce w roku 1968 a zrobił to amerykański uczony E. Altman.[[2]](#footnote-2) Od tamtej pory metoda jest często wykorzystywana a wcześniej wspomniany rozwój technologiczny ułatwia to w znacznym stopniu. Zastosowanie funkcji dyskryminacyjnej wciąż rośnie. Świadczy o tym choćby fakt, że jest to najpopularniejsza metoda używana do badań nad złą kondycją finansową.

## . Aktualność tematu

To jak dobrze prosperuje dane przedsiębiorstwo jesteśmy w stanie ocenić po poziomie rentowności jakie prezentuje. Istnieje bezpośredni związek między wykorzystaniem zasobów finansowych a generowaniem zysków przez firmy biznesowe.

Analiza finansowa jest wymagana do określenia kondycji finansowej i stabilności firmy. Pomaga w określeniu bieżącej sytuacji finansowej i nadchodzących wymagań finansowych, jest to ważny i niezbędny aspekt przy prowadzeniu działalności gospodarczej. Analiza jaką jest analiza dyskryminacyjna pozwala wykonać taką ocenę, dzieląc grupę firm na dwie klasy, mimo iż podział nie jest szczegółowy, jasno określa które firmy prosperują dobrze a którym grozi upadłość.

Rezultat badań prowadzonych w tej pracy może być przydatny dla takich podmiotów jak pożyczkodawcy, władze lokalne, menadżerowie czy tez inwestorzy. Wynik pomoże sprawdzić czy dana firma będzie w stanie spełnić zobowiązania finansowe albo ocenić czy jej aktualna pozycja na rynku nie wskazuje na bankructwo. Oznacza to, że dzięki wynikom, które otrzymamy, będziemy mogli podjąć decyzje na przykład o sprzedaży akcji zanim jeszcze dana firma upadnie. Władze lokalne będą miały lepszy ogląd na to jakie firmy są odpowiednimi partnerami do współpracy, to samo tyczy się innych przedsiębiorstw ponieważ jeśli firma zostanie sklasyfikowana jako bankrut to nie warto podpisywać z nią długoterminowe umowy, powodem tego jest fakt, iż taka firma będzie miała problem ze spełnieniem swoich zobowiązań. Ważnym aspektem będzie również odpowiedź na pytanie, które zmienne mają największy wpływ na przynależność do poszczególnych grup (bankrutów/nie-bankrutów), to z kolei może pomóc przyszłym przedsiębiorcą lepiej zrozumieć na które wskaźnik powinni w większym stopniu zwrócić uwagę aby ich firma utrzymała się na rynku.

Co więcej w dzisiejszych czasach metody takie jak analiza dyskryminacyjna nie są skomplikowane do przeprowadzenia dla kogoś pałającego się analizą danych. Wiele języków programowania oferuje funkcje które posiadają narzędzia statystyczne pozwalające przeprowadzić takie badania. Do najpopularniejszych języków programowania posiadających odpowiednie narzędzia statystyczne aby przeprowadzi liniową analizę dyskryminacyjna należą takie języki jak Python, R czy też Matlab. Mimo, iż wcześniej wymienione języki z pewnością spełniają najwięcej wymagań stawianych przez analizę danych, to przy pomocy innych języków przeprowadzenie tak prostej metody jaką jest liniowa analiza dyskryminacyjna również byłoby możliwe, choć na pewno kosztowałby to więcej pracy i czasu. Mimo iż sam zamysł klasyfikacji obiektów za pomocą liniowej analizy dyskryminacyjnej pojawił się już w latach 30 dwudziestego wieku[[3]](#footnote-3) sam fakt, iż w dzisiejszych czasach jest tak wiele narzędzi które pozwalają na analizowanie danych za pomocą takich metod świadczy o aktualności tego zagadnienia.

## . Cel pracy oraz hipotezy badawcze

Ocena finansowa jest ważną częścią ekonomii albowiem pozwala sprawdzić w jakiej sytuacji są przedsiębiorstwa oraz jaka jest ich pozycja na rynku. Analiza dyskryminacyjna pozwala taką ocenę wykonać, choć jej wynik nie odpowiada na takie pytania jak to, gdzie znajduje się aktualna pozycja na rynku danego przedsiębiorstwa. Podział otrzymany z badań prowadzonych w tej pracy nie będzie aż tak szczegółowy z uwagi na fakt, że wykorzystanie analizy dyskryminacyjnej do predykcji upadłości przedsiębiorstw opiera się na dzieleniu ich na dwie grupy. W jednej z grup znajdują się przedsiębiorstwa, którym grozi upadek, natomiast w drugiej takie, które nie powinny mieć problemów z utrzymaniem się na rynku.

Głównym zagadnieniem realizowanym w tej pracy będzie przeprowadzenie badań na grupie młodych przedsiębiorstw, by sprawdzić, którym uda się utrzymać na rynku. W szerszej perspektywie wynik otrzymany po przeprowadzeniu metody powinien również odpowiadać na zagadnienia takie jak to które zmienne determinują przynależność do danej klasy. Inaczej mówiąc, jednym z celów tej pracy będzie określenie wskaźników, których konkretny poziom może świadczyć o tym że, firmie grozi upadłość albo że jej pozycja jest stabilna.

Rozważając, co może determinować upadek firmy lub też nastanie czasu w którym firma będzie się znajdować w okresie schyłku należy zastanowić się co może mieć na to największy wpływ. Analizując prace naukowe odnoszące się do podobnego regionu Europy, nasuwają się wnioski, że czynniki najbardziej wpływające na to czy dana firma upadnie to takie wskaźniki jak te związane z aktywami obrotowymi, zobowiązanymi krótko i długo terminowymi czy z tym jakie przychody generowała sprzedaż. Ważnym aspektem w funkcjonowaniu firmy jest płynność finansowa zatem zmienne objaśniające odnoszące się do przepływów pieniężnych również mogą okazać się znaczące.[[4]](#footnote-4),[[5]](#footnote-5)

Mając na uwadze powyższe rozważania, można sformułować następujące hipotezy:

* Na przynależność do danej grupy będą wpływać czynniki związane z przychodami ze sprzedaży.
* Na przynależność do danej grupy będą wpływać czynniki związane z przepływami pieniężnymi.
* Na przynależność do danej grupy będą wpływać czynniki związane z zobowiązanymi krótko- i długoterminowymi.

## 1.3. Przedstawienie metod i rozwiązań najczęściej stosowanych do rozwiązywania omawianych zagadnień

Predykcja upadłości przedsiębiorstw jak już było wcześniej wspomniane jest tematem jak najbardziej aktualnym i takim, który świetnie spełnia wymagania stawiane przez dwudziesty-pierwszy wiek, stąd nie dziwi fakt, że do prognozowania powstało tak wiele różnorakich metod opierających się na zróżnicowanych zasadach matematycznych. Zachodnia literatura dzieli metody badania zagrożeń upadłości w biznesie na trzy grupy metod opierających się na różnych dziedzinach nauki. Pierwsza grupa to metody statystyczne, do grupy tej zaliczamy metody takie jak analiza dyskryminacyjna, której temat jest poruszany tej pracy. To jednak nie jedyne narzędzia statystyczne, które są stanie „przewidzieć” bankructwo, nie sposób nie wspomnieć o modela takich jak model logitowy, probitowy czy też drzewa decyzyjne. Kolejną grupą są metody tak zwanych miękkich technik obliczeniowych. W ich skład wchodzą: sztuczne sieci neuronowe, logika rozmyta, algorytmy genetyczne czy też modele wektorów nośnych. Następna, trzecia i ostatnia grupa to grupa modeli teoretycznych, którą tworzą takie modele jak: modele hazardowe, modele oparte na teorii entropii czy też modele ryzyka kredytowego. Podział jaki został wcześniej opisany przedstawiono również na poniższym schemacie.[[6]](#footnote-6)

Schemat 1.1. Podział metod predykcji upadłości

Źródło: opracowanie własne na podstawie: W. Pierzchalska*, Analiza finansowa oraz modele przewidywania upadłości jako narzędzie oceny kondycji finansowej przedsiębiorstwa*, „Finanse i Prawo Finansowe” , nr ,1 Łódź, 2014.

Wyniki badań przeprowadzone przez Aziza i Dara jasno określają jakie metody i w jakim stopniu są wykorzystywane do predykcji upadłości. Rezultat tychże badań jako najpopularniejszą metodę wskazał właśnie analizę dyskryminacyjną, która była wykorzystywana w ponad 30% przypadków. Kolejnymi najbardziej rozpowszechnionymi metodami były modele logitowe a następnie po nich w rankingu znajdowały się sieci neuronowe, znaczny procent stanowiły też drzewa decyzyjne.[[7]](#footnote-7) Poniżej przedstawiono wykres, który określa jaki procent badań przeprowadzanych na całym świecie jest dokonywany za pomocą danych metod.

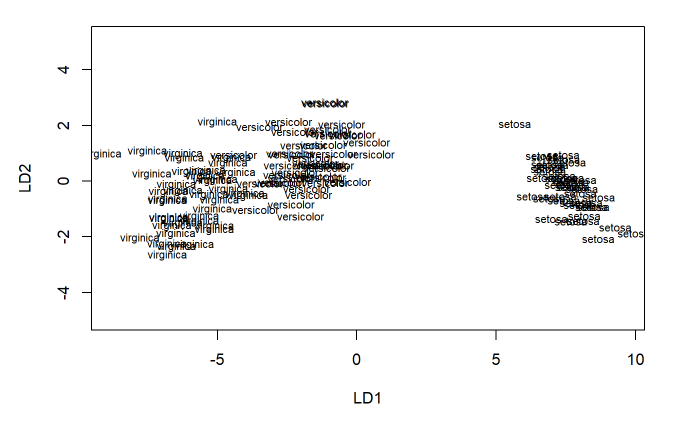
Wykres 1.1. Wykorzystanie technik predykcji upadłości do badań prowadzonych na świecie

Źródło: opracowanie własne na podstawie: W. Pierzchalska*, Analiza finansowa oraz modele przewidywania upadłości jako narzędzie oceny kondycji finansowej przedsiębiorstwa*, „Finanse i Prawo Finansowe” , nr ,1 Łódź, 2014.

Co więcej sama analiza dyskryminacyjna posiada wiele odmian będących w stanie ocenić to czy dana firma upadnie. Najczęściej jednak wykorzystywana jest w tym celu liniowa analiza dyskryminacyjna. Różnią się one przed wszystkim tym w jaki sposób przydzielają obiekty do poszczególnych klas. Możemy wymienić takie rodzaje analizy dyskryminacyjnej jak kwadratowa analiza dyskryminacyjna (QDA), mieszana analiza dyskryminacyjna (MDA), czy tez krokowa analiza dyskryminacyjna.

## 1.4. Szczegółowy opis przypadków użycia metody

Głównym założeniem analizy dyskryminacyjnej jest podział obiektów na klasy. Podział ten nie jest skomplikowany, stąd nie dziwi fakt, że analizę tą można stosować do najróżniejszych badań. Liniowa funkcja dyskryminacyjna, której temat jest podejmowany w tej pracy, z założenia dzieli grupę obiektów na dwie grupy nie jest to jednak jedyny podział jaki możemy uzyskać dzięki funkcji dyskryminacyjnej, gdyż liczba klas, na która dzielimy obiekty, może być większa. Dzięki tej metodzie możemy na przykład wykonać podział kwiatów na konkretne odmiany, jeśli tylko posiadamy odpowiednie dane. W programie RStudio dostępna jest baza danych odnośnie irysów - są to dane na temat konkretnych roślin należących do trzech odmian. Wspomniana baza danych zawiera takie zmienne jak długość czy tez szerokość płatka. Metoda jaką jest analiza dyskryminacyjna jest w stanie podzielić obiekty względem konkretnych odmian czyli w tym przypadku na trzy klasy. Poniżej zaprezentowano działanie LDA dla tej bazy danych.



Wykres 1.1. Przykładowa analiza dyskryminacyjna dla irysów

Źródło: opracowanie własne

Na wykresie możemy zobaczyć, że odmiana setosa znacznie różni się od dwóch pozostałych. Determinują to duże wartości LD1 czyli pierwszej funkcji dyskryminującej.

Spektrum zastosowania analizy dyskryminacyjnej jest jednak dużo szersze, można natknąć się nawet na prace, gdzie przy pomocy funkcji dyskryminacyjnej porównywane są ekstrakty torfowe[[8]](#footnote-8), czyli próbki materii pochodzące z różnego typu torfowisk. Studiując artykuły na temat zastosowania analizy dyskryminacyjnej, nie sposób nie zauważyć jak wiele z nich tyczy się tematów związanych z badaniami nad zdrowiem, na przykład nad badaniem astmy u dzieci[[9]](#footnote-9), diagnozowaniu choroby wieńcowej[[10]](#footnote-10), czy też w kardiochirurgii. [[11]](#footnote-11)

W dziedzinie biznesu można ją wykorzystać do realizacji wielu zadań. Możliwe jest zastosowanie jej do oceny zdolności kredytowej,[[12]](#footnote-12) czy też wspomagania decyzji przetargowych[[13]](#footnote-13) a także do predykcji upadłości spółek, jak to ma miejsce w tych badaniach. Jako pierwszy w taki sposób wykorzystał ją E. Altman kiedy to w Stanach Zjednoczonych pojawiło się zainteresowanie predykcją upadłości. Były to czasy wielkiego światowego kryzysu czyli przełom lat 20-tych i 30-tych XX wieku.[[14]](#footnote-14)

Kolejnym z ważnych badaczy był Beaver wydał on istotną publikacje w roku 1966. Jako pierwszy zastosował dychotomiczny test klasyfikacyjny w jednowymiarowym układzie, a także położył fundamenty pod przyszłe modele predykcyjne. Użył wskaźników finansowych na poziomie 79 upadłych i niewypłacalnych przedsiębiorstw działających w tej samej branży i posiadających taką samą ilość aktywów. Wskaźniki, jakie wytypował w swoim toku badawczym, to Przepływy pieniężne/Ogółem zadłużenia. Określił je jako najlepszy predyktor upadłości.[[15]](#footnote-15)

Wraz z rozwojem technologii oraz programów komputerowych umożliwiających analizowanie danych predykcja upadłości przy pomocy funkcji dyskryminacyjnej czy też innych metod zaczęła się rozwijać bardzo dynamicznie.

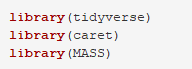
## 1.5. Narzędzia wybrane do przeprowadzenia metody

Do przeprowadzenia w odpowiedni sposób oraz do wyestymowania modelów funkcji dyskryminacyjnej wymagane jest użycie szeregu narzędzi statystycznych. LDA określa średnie dla grupy i oblicza dla każdej obserwacji prawdopodobieństwo przynależności do różnych grup. Jednak zanim to nastąpi konieczne jest zastosowanie różnych narzędzi używanych w statystyce. Instrumenty statyczne pozwolą na wykonanie wykresów oraz przeprowadzenie testów. Wykorzystane zostaną między innymi wykresy pudełkowe, które posłużą do identyfikacji obserwacji odstających, by możliwe było ich usunięcie. Poszczególne zmienne przetestowane zostaną testem Shapiro-Wilka, by zbadać czy ich rozkład jest rozkładem normalnym, gdyż właśnie taki jest pożądany przy prowadzeniu tego typu analiz. Zmienne objaśniające zostaną również poddane standaryzacji.

Głównym programem użytym w tych badaniach będzie Rstudio, jest to program komputerowy umożliwiający wielopoziomową analizę danych. Przy pomocy programowania w języku R pozwala on na używanie narzędzi statystycznych w szybki i intuicyjny sposób. Umożliwia on również korzystanie z gamy bibliotek, które w wielu przypadkach otwierają nowe możliwości lub po prostu ułatwiają pracę. Kodowanie w języku R jest jednym z najczęściej stosowanych narzędzi przy analizie danych. Metoda, jaką jest analiza dyskryminacyjna, zostanie zaimplementowana właśnie w tym języku, który jest zazwyczaj używany właśnie dla badań statystycznych tego typu.

Wykorzystane biblioteki

Poniżej przedstawiono niektóre z bibliotek wykorzystywanych w badaniach prowadzonych w tej pracy.



Ilustracja 1.1. Niektóre z bibliotek wykorzystane do przeprowadzenia badań

Źródło: opracowanie własne

**Tidyverse** – biblioteka dodana dla łatwej manipulacji danymi oraz wizualizacji.

**Caret** – biblioteka odpowiedzialna za kursor ułatwiający proces uczenia maszynowego.

**MASS** – Oblicza przedziały ufności dla jednego lub więcej parametrów w dopasowanym modelu. Pakiet MASS dodaje metody pasowania do glm i nls.

W badaniach zostaną również użyte biblioteki, które pozwolą na określenie jaką transformację należy wykonać na poszczególnych zmiennych, by ich rozkład był zbliżony do rozkładu normalnego. W celu tym zostanie użyta biblioteka „*rcompanion”* która wykorzystując podejście Drabiny Mocy Tukeya. Pozwoli to na znalezienie takich przekształceń, przy pomocy którego będzie możliwa transformacja danych, by ich rozkład był jak najbliższy rozkładowi normalnemu.

## 1.6. Struktura pracy

Struktura pracy będzie miała charakter trzyczęściowy. Pierwsza jej część to wstęp, w którym zostanie przybliżona tematyka analizowanego zagadnienia, środowisko w jakim kształtowała się idea analizy dyskryminacyjnej. Co więcej w rozdziale tym znajdziemy odpowiedź na pytanie: Dlaczego temat pracy jest aktualny? Zostaną podane cel pracy oraz hipotezy badawcze opierające się na pracach o podobnej tematyce i odnoszących się do podobnego regionu. W rozdziale tym zostaną przedstawione także metody wykorzystywane do rozwiązywania tego typu problemów ich popularność oraz w sposób bardziej szczegółowy opisany zostanie sposób w jaki działa docelowa metoda wykorzystywana w tych badaniach. Część ta zawierać będzie także opis środowiska implementacji metody oraz źródła pozyskiwania informacji.

Kolejnym rozdziałem będą badania empiryczne. Rozdział ten w głównej mierze będzie opierać się na statystycznym analizowaniu danych. Jednakże na tym etapie opisane zostaną też dane oraz szczegółowo zostanie zaprezentowany sposób w jaki działa metoda czy też sama geneza jej powstania. Część w której scharakteryzowany zostanie obiekt badań odnosić się będzie również do jednostki organizacyjnej jaką jest przedsiębiorstwo. Opisany będzie cykl jego życia. Podane zostaną również cechy małych i średnich przedsiębiorstw, gdyż właśnie dla takiej grupy będą prowadzone badania oraz krajowa gospodarka Słowacji dla tego sektora, ponieważ właśnie z tego państwa pochodzą dane. Dużo informacji zawartych w tym rozdziale odnosić się będzie do procesu przygotowania danych.

Ostatnią częścią pracy będzie podsumowanie, gdzie zostaną przedstawione konkluzje płynące z przeprowadzonych badań. Co więcej opisane zostaną wnioski wynikające z prowadzonych analiz.

## 1.7. Źródła informacji

W dzisiejszych czasach czyli czasach szybkiego internetu i czasach gdzie źródła takie jak książki odchodzą do lamusa jesteśmy wręcz zasypywani różnego pokroju danymi, nowinkami, czy też informacjami. Wzrost ilości źródeł dostępnych dla każdego spowodował znaczący spadek jakości tej materii. Niejednokrotnie napotykamy na nierzetelne dane czy tez informacje, którym daleko do prawdy. Z uwagi na powyższy fakt źródłami, z których powinniśmy korzystać pisząc pracę analityczną, jaką jest praca licencjacka, nie powinniśmy odnosić się do blogów, czy też subiektywnych opinii, które możemy znaleźć np. na forach. Podstawą rzetelnie napisanej pracy naukowej są inne prace naukowe już zatwierdzone przez uczelnie, artykuły pisane przez ludzi ze świata nauki czy też wyżej wspomniane książki. Mimo że coraz rzadziej korzystamy z klasycznych wydań książek, nie jest to problem, bo ich wersje w postaci najczęściej plików pdf bardzo często można znaleźć w intrenecie.

Chcąc przeprowadzić ocenę finansową, należy wziąć pod uwagę wiele czynników. Każdy z czynników to kolejna zmienna, która w tych badaniach będzie przedstawiana jako szereg danych dla poszczególnych przedsiębiorstw. Informacje używane przy analizie finansowej są najczęściej zawarte w sprawozdaniach finansowych sporządzanych na podstawie ewidencji księgowych oraz kalkulacji. Takie dane są w szczególności dostępne w bilansie, rachunku zysków i strat, rachunku przepływów pieniężnych, informacji dodatkowej a także, księgach i kartotekach inwentarzowych majątku przedsiębiorstwa. Celem tej pracy nie jest jednak zbieranie danych odnośnie przedsiębiorstw zwłaszcza, że nie jest to proste dla grupy jaką są małe i średnie przedsiębiorstwa. Mimo, iż przedsiębiorstwa mają obowiązek publikować informacje finansowe niezbędne do przeprowadzenia tego typu analiz, samo ich szukanie i zestawienie razem wymagało by wiele pracy oraz było by bardzo czasochłonne.

# 2. Badania Empiryczne

Badania empiryczne opierają się na obserwowanych i mierzonych zjawiskach i czerpią wiedzę z rzeczywistych doświadczeń, a nie z teorii lub przekonań[[16]](#footnote-16) stąd głównym zagadnieniem realizowanym w tym rozdziale będzie opis danych oraz charakterystyka środowiska, z jakiego pochodzą. Praca, jaka zostanie wykonana na zestawie danych w tym rozdziale, opierać się będzie o naukowe metody oraz narzędzia statystyczne.

## 2.1. Charakterystyka obiektu badań

### Przedsiębiorstwo

Obiektem tych badań będzie grupa jaką stanowią małe i średnie przedsiębiorstwa należące do czterech różnych gałęzi gospodarki. Podmiot prowadzący działalność gospodarczą a także będący jednostką organizacyjna jest nazywany przedsiębiorstwem. Jednostki organizacyjne tego typu cechują się wyodrębnieniem w zakresie organizacyjnym, prawnym, ekonomicznym a także terytorialnym.

Jednym z głównych celów każdego przedsiębiorstwa jest maksymalizacja zysku czyli prowadzenie działalności w taki sposób aby świadczenie usług czy też sprzedaż produktów była jak najbardziej rentowna. Istnieją jednak inne równie ważne o ile nie ważniejsze cele dla przedsiębiorstwa takie jak wywiązywanie się ze zobowiązań czyli płynność finansowa. Istotna też jest równowaga finansowa czy autonomia.

Działalności jakie prowadzone są przez przedsiębiorstwa to takie działalności jak działalność produkcyjna, usługowa czy też handlowa, co więcej jednostka ta ma zdolności do czynności prawnych a to jaką ma formę i ustrój jest jasno określone w oddzielnym systemie prawnym.

Istnieje wiele rodzajów kryterium klasyfikacji podmiotów gospodarczych. Wynika to z faktu, iż specjalizują się one w prowadzeniu różnego typu działalności oraz że działalność ta może mieć różny zakres czy też formę prawną. Najczęściej jednak przedsiębiorstwa grupowane są ze względu na:

* Formę prawną podmiotu
* Rozmiar podmiotu
* Terytorium, na jakim prowadzona jest działalność
* Formę własności (państwowe/prywatne)
* Typ działalności.[[17]](#footnote-17)

Jednostka organizacyjna, jaką jest przedsiębiorstwo, posiada także określone fazy rozwoju. O ich pochodzeniu dowiadujemy się z tak zwanych teorii fazowych. Ze względu na analogie przedsiębiorstwa do organizmów żywych fazy te nazywane są cyklem życia przedsiębiorstwa. Zazwyczaj cykl w jakim funkcjonuje przedsiębiorstwo posiada cztery fazy i wygląda w sposób następujący[[18]](#footnote-18):

Schemat 2.1 Cykl życia przedsiębiorstwa

Źródło: opracowanie własne

Fazą cyklu życia przedsiębiorstwa, która będzie najważniejsza dla badań prowadzonych w tej pracy, jest schyłek, gdyż przy pomocy narzędzi statystycznych podjęta zostanie próba predykcji upadłości. Inaczej mówiąc, przedsiębiorstwa zostaną pogrupowane na podstawie tego, czy znajdują się w fazie schyłku czy też nie.

### Sektor małych i średnich przedsiębiorstw

Małe i średnie przedsiębiorstwa stanowią grupę, która jest niełatwa do wyodrębnienia, gdyż nie istnieją akceptowalne powszechnie uniwersalne definicje jasno określające przynależność do tejże grupy. Systemy prawne w sposób zróżnicowany klasyfikują ten sektor. Na ogół jednak przedsiębiorstwa dzieli się, biorąc pod uwagę kryteria jakościowe tudzież ilościowe lub oba wcześniej wymienione naraz.[[19]](#footnote-19)

Obiektem badań prowadzonych w tej pracy są przedsiębiorstwa działające w Słowacji. Kraj ten od pierwszego maja 2004 roku podobnie jak Polska jest jednym z krajów członkowskich Unii Europejskiej. Z uwagi na ten fakt to jak klasyfikujemy małe i średnie przedsiębiorstwa w tym państwie opierać się będzie o przepisy ustalone przez UE. To jak owe przepisy grupują przedsiębiorstwa przedstawiono na schemacie 2.2.

Schemat 2.2. Podział przedsiębiorstw ze względu na przepisy unijne

Źródło: opracowanie własne na podstawie: rozporządzenia Komisji Wspólnoty Europejskiej nr 800/2008 z dnia 6 sierpnia 2008r.

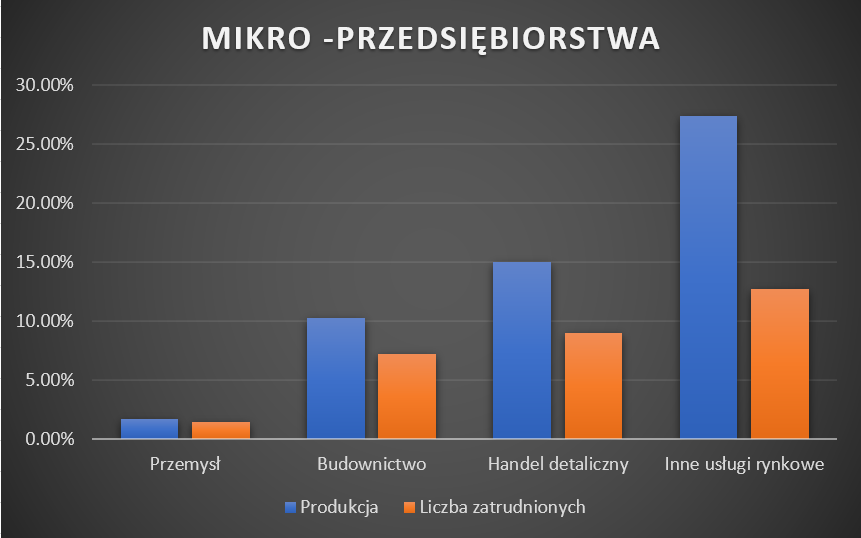
### Charakterystyka gospodarki Słowacji

Słowacja jest państwem, w którym panuje ustrój demokratycznej republiki parlamentarnej. Kraj ten jest zamieszkiwany przez 5 450 421 ludzi a jego powierzchnia wynosi 20 273 tys. , co czyni ją czwartym najmniejszym krajem członkowskim Unii Europejskiej. Mimo, iż państwo to jest niewielkie pod względem czynników takich jak PKB na jednego mieszkańca, w SSN plasuje się lepiej niż Polska i zajmuje 22 pozycję w UE.[[20]](#footnote-20) To jaki udział mają poszczególne sektory gospodarki przedstawiono na wykresie 2.1.

Wykres 2.1. Udział poszczególnych sektorów gospodarki Słowackiej w ogólnej wartości PKB

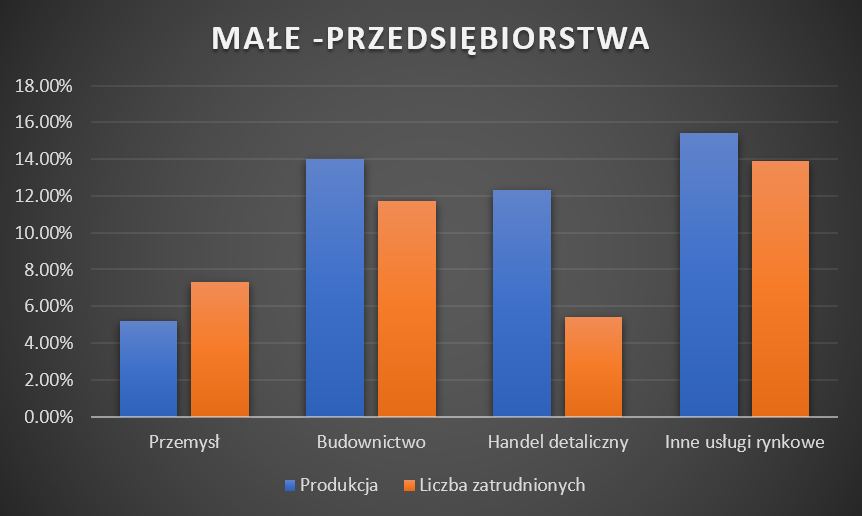
Źródło: opracowanie własne na podstawie: <http://polskoslowackaizba.pl/gospodarka-slowacji-w-2019-roku/> [dostęp:10.06.2020]

Sektor małych i średnich przedsiębiorstw w Słowacji odpowiada za znaczącą cześć PKB. Z 331 021 podmiotów gospodarczych jakie udało się odnotować w tym kraju aż 330 199 to małe i średnie przedsiębiorstwa. Jest to aż 99,75% wszystkich przedsiębiorstw, co prawda jest to stan z roku 1997 i od tamtego czasu sytuacja na rynku z pewnością uległa zmianie lecz daje to ogląd na ogólną sytuację. Poniżej zostaną przedstawione wykresy obrazujące udział sektora MSP w poszczególnych dziedzinach gospodarki w Słowacji.[[21]](#footnote-21)



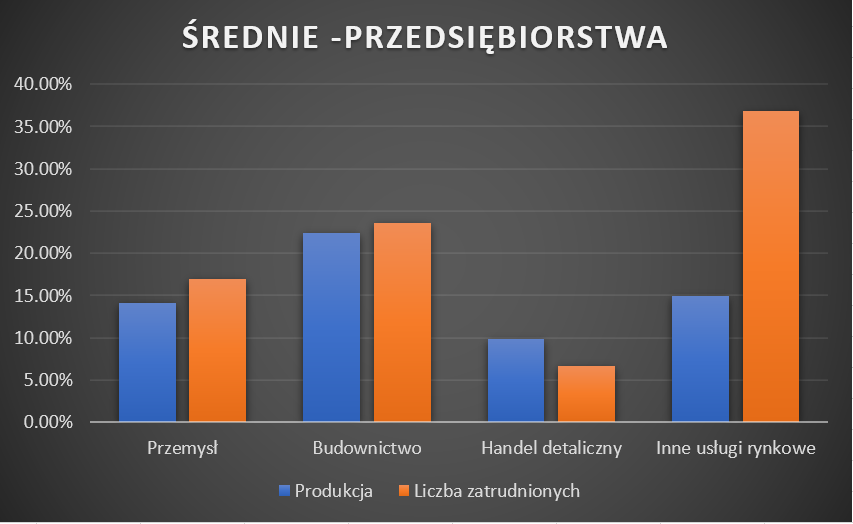
Wykres 2.2. Udział mikro-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji

Źródło: Opracowanie własne na podstawie: B. Korkozowicz,*Małe i średnie przedsiębiorstwa na świecie*, „Polska Fundacja Promocji i Rozwoju Małych i Średnich Przedsiębiorstw”, 2000, s. 39-41.



Wykres 2.3. Udział małych-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji

Źródło: Opracowanie własne na podstawie: B. Korkozowicz,*Małe i średnie przedsiębiorstwa na świecie*, „Polska Fundacja Promocji i Rozwoju Małych i Średnich Przedsiębiorstw”, 2000, s. 39-41.



Wykres 2.4. Udział średnich-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji

Źródło: Opracowanie własne na podstawie: B. Korkozowicz,*Małe i średnie przedsiębiorstwa na świecie*, „Polska Fundacja Promocji i Rozwoju Małych i Średnich Przedsiębiorstw”, 2000, s. 39-41.

### Opis danych

Dane wykorzystywane w badaniach, które zostaną przeprowadzone w tej pracy pochodzą z ogólnodostępnej internetowej bazy danych nazywającej się „Mendeley” i odnoszą się do spółek zarejestrowanych w Słowacji. Baza danych, która została wykorzystana do pozyskania informacji potrzebnych w badaniach prowadzonych w tej pracy nosi tytuł „Dataset of financial ratios of Slovak companies". Spółki te należą do grupy małych i średnich przedsiębiorstw. Branże, na których temat można znaleźć informacje, to sektory takich branż jak przemysłowa, rolnicza, budowlana oraz branża sprzedaży detalicznej. Informacje zawarte w owej bazie danych zostały obliczone na podstawie publicznie dostępnych zapisów. Baza danych zawiera pliki w których osobno umieszczone są spółki, które zbankrutowały i te którym udało utrzymać się na rynku. Każdy z arkuszy zawiera dane finansowe odnoszące się do trzech lat. W zdecydowanej większości przypadków, zestawiając ze sobą pliki z tych samych lat a także te odnoszące się do takich samych gałęzi przemysłu, możemy zauważyć, iż w owych latach firmy na ogół nie bankrutowały. Stosunek liczby firm, które dalej prowadziły działalność w danym okresie, do tych które upadły jest wysoki. Baza danych powstała w maju 2019 roku a jej autorami są P. Drotár, P. Gnip, M. Zoričak, V. Gazda. Jest to baza danych która została dołączona do artykułu naukowego pod tytułem „Small- and medium-enterprises bankruptcy dataset”.

Poszczególne zmienne oznaczają:

V1 - [TAT - Total Asset Turnover]

V2 – 365 \* [ATD- Asset Turnover Days]

V3 – 365 \* [DTR- Days Total Receivables Outstanding]

V4 – 365 \* [ITD- Inventory Turnover Days]

V5 – 365 \* [L1-Cash Ratio]

Ldenominator = Zobowiązania krótkoterminowe + krótkoterminowa pomoc finansowa + bieżące kredyty bankowe + rozliczenia międzyokresowe

V6 - [L2- Quick Ratio]

V7 - [L3- Current Ratio]

V8 - [ROA- Return on Assets]

V9 - [ROE- Return on Equity]

V10 - [ROS- Return on Sales]

V11 - [ROI- Return on Investment]

V12 - [LRR- Labor-to-Revenue Ratio]

V13 - [WAR- Wages to Added Value Ratio]

V14 - [LP- Labor Productivity]

V15 - [DA- Debt-to-Assets Ratio]

V16 - [DE- Debt-to-Equity Ratio]

V17 - [FL- Financial Leverage]

V18 - [DIR- Debt to Income Ratio]

V19 - [DCR- Debt Service Coverage Ratio]

V20 - [ACR- Asset Coverage Ratio]

V21 - [BL- Bank Liabilities to Debt Ratio] [[22]](#footnote-22)

## 2.2. Analiza literatury

Historycznie pierwsze użycie funkcji dyskryminacyjnej, jakie pojawiło się w literaturze statystycznej do klasyfikacji obiektów, to badania Fishera, które miały miejsce w roku 1936. Pracował on na bazie danych dotyczącej irysów, czyli kwiatów popularnych ozdobnych, która zawierała cztery zmienne. Podejście jakie zastosował ten uczony prowadziło do otrzymania reguły dyskryminacyjnej opartej o funkcję liniową.

Kolejnym ważnym krokiem dla kierunku oraz dynamiki rozwoju analizy dyskryminacyjnej było podejście jakie zastosował E. Altman. On jako pierwszy wykorzystał klasyfikacje obiektów jaka można otrzymać przy pomocy dyskryminacji do predykcji upadłości spółek, zrobił to przy pomocy zawansowanego aparatu matematycznego oraz rozległej bazy danych, miało to miejsce w roku 1968. Profesor Altman wyznaczył liniową funkcję dyskryminacyjną na podstawie pięciu wskaźników finansowych. Dane na podstawie których udało mu się to osiągnąć odnosiły się do 66 amerykańskich przedsiębiorstw z których 33 zbankrutowały natomiast pozostałe 33 były w dobrej kondycji finansowej. Liniowa funkcja amerykańskiego uczonego prezentowała się w sposób następujący.

(1)

Liniowa formuła dyskryminacyjna według Altmana

Poszczególne zmienne oznaczały:

Z -

X1 -

X2 -

X3 -

X4 -

X5 -

W dzisiejszych czasach, w literaturze polskiej mamy do czynienia z poglądem, który jako analizę dyskryminacyjną określa zbiór metod mających na celu znalezienie takiej reguły, która pozwoli na klasyfikację obiektów, charakterystyki klas czy też wzoru na funkcję która rozdzieli obiekty na poszczególne klasy[[23]](#footnote-24).

## 2.3. Podstawy Teoretyczne

### Geneza

Analiza dyskryminacyjna to metoda klasyfikacji pierwotnie opracowana w 1936 r. przez R. A. Fishera, który do jej przeprowadzenia używał linowych modeli regresji. Zbiór danych, na którym prowadził badania, był najbardziej znanym zbiorem w wielowymiarowej analizie statystycznej - mowa tu o zestawieniu danych odnoszących się do trzech odmian kwiatu irys - zbiór ten posiada cztery zmienne.

W Polsce jednym z pierwszych uczonych, który zajmował się metodami analizy dyskryminacyjnej był prof. Józef Kolonko, lecz w jego badaniach pojęcie to miało znacznie szersze znaczenie - było to spowodowane tym, iż w naukowej literaturze rosyjskiej dyskryminacje uważa się za podział, który daje najlepszą wartość funkcjonału stanowiącego kryterium (była to tzw. klasyfikacja bez wzorca). Było to inne rozwiązanie niż u Fishera, gdyż jego podział zachodził według wzorca.[[24]](#footnote-25)

### Istota

Możliwości jakie daje wielowymiarowa analiza dyskryminacyjna to przede wszystkim dokonanie klasyfikacji obiektu na podstawie wielu zmiennych objaśniających. Zmienna objaśniana w naszych badaniach będzie się odnosić do tego czy dane przedsiębiorstwo zbankrutowało czy też nie jest to typowa dla tego typu analiz zmienna jakościowa. W wielowymiarowej analizie dyskryminacyjnej klasyfikacja obiektu do jednej grupy dokonuje się na podstawie linowych funkcji klasyfikacji lub linowych funkcji dyskryminacji.[[25]](#footnote-26)

#### Liniowa analiza dyskryminacyjna

Istotą liniowej analizy dyskryminacyjnej jest znalezienie takiej transformacji oryginalnych zmiennych, dzięki którym będzie możliwe jak najlepsze rozdzielenie na dwie klasy. Aby zobrazować jak działa pojęcie rozdzielności, została zdefiniowana następująca funkcja punktacji, jej autorem był Fisher. Zaproponowana funkcja ma na celu znalezienie wektora który ma za zadanie zmaksymalizować funkcje separacji , gdzie:

(2)

(3)

Funkcje punktacji według Fishera

, - średnie zmiennych transformowanych z populacji pierwszej oraz z populacji drugiej

(4)

Wielkość odchylenia standardowego z prób

(5)

Równanie (3) czyli mierzy różnicę pomiędzy przekształconymi średnimi oraz , które są relatywne do wielkości odchylenia standardowego z prób danego równaniem (4). Aby było jak największe, zmienne przekształcone , ,… oraz ,, musiałby by być całkowicie odseparowane. Ilość obserwacji z populacji pierwszej została oznaczona jako n1 natomiast ilość obserwacji w populacji drugiej została oznaczona jako n2.

Wektor odpowiedzialny za maksymalizacje separacji to wektor , który jest dany równaniem:

(6)

Wektor maksymalizujący separacje

Gdzie:

(7)

(8)

(9)

Oraz , są średnimi wektorów 1 i 2

Biorąc za bazę równanie (2) oraz przyjmując, że mamy obserwację , możliwe jest przeniesienie obserwacji do odpowiednich grup populacyjnych na podstawie reguł klasyfikacji przedstawionych poniżej:

* Przenieś obserwacje do populacji pierwszej jeżeli

(10)

Warunek przeniesienia obserwacji do populacji pierwszej

* Przenieś obserwacje do populacji drugiej jeżeli

(11)

Warunek przeniesienia obserwacji do populacji drugiej

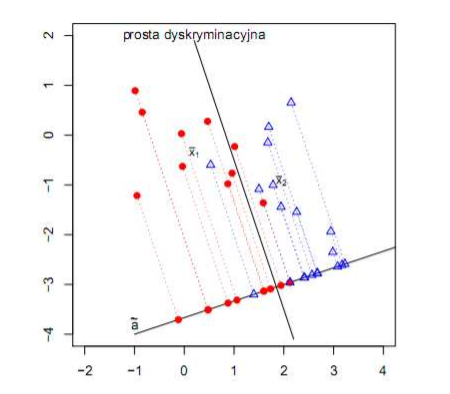
Inaczej rzecz ujmując, jeżeli mamy do czynienia z sytuacją gdzie znajduje się po prawej stronie równania , dyskryminator sklasyfikuje w populacji pierwszej i na odwrót.[[26]](#footnote-27)

#### Ocena skuteczności dyskryminacji

Aby sprawdzić w jakim stopniu dany model sprawdza się w rzeczywistości, najbardziej intuicyjną rzeczą jaką możemy zrobić jest przetestowanie go dla lat wcześniejszych i porównanie wyników ze stanem faktycznym. W przypadku prognozowania upadłości spółek, chcąc sprawdzić czy model działa poprawnie, należałoby ocenić w jakim stopniu taki model poprawnie klasyfikował upadłość spółek w latach ubiegłych, należałoby również wziąć pod uwagę czy w danym czasie nie wystąpił kryzys czy też wahania spowodowane z innych przyczyn. Takie sytuacje mocno zaburzają zdolności prognostyczne modelu. Jedną z praca gdzie temat ten jest podejmowany jest praca Sylwii Godlewskiej pod tytułem „Skuteczność polskich modeli dyskryminacyjnych w ocenie zagrożenia upadłością spółek giełdowych”. Autorka pracy przeprowadza badania na polskich modelach dyskryminacyjnych i dochodzi do wniosków, iż badania takie powinny być dalej prowadzone, gdyż mimo że niektóre z modeli w ogóle nie były w stanie zdiagnozować upadłości spółek, inne poprawnie wykonały to zadanie aż w 83%.[[27]](#footnote-28)

### Przykład

Sposób w jaki działa analiza dyskryminacyjna został pokazany w podręczniku Prof. dr hab. inż. Jacka Koronackiego. Wykres pokazany w podręczniku przedstawiał obserwacje które należały do grupy pierwszej jako trójkąty oraz te należące do grupy drugiej jako kółka. Zaznaczono również dwie zmienne zapisane jako oraz , które mają obrazować średnie grupy pierwszej i drugiej. Jak można zauważyć obserwacje są reprezentowane jako punkty w przestrzeni dwuwymiarowej.



Wykres 2.5. Prosta analiza dyskryminacyjna

Źródło:K. Błonski, O. Brozek, Handwritten digit recognition using discriminant analysis, “Faculty of Applied Mathematics, Sielsian University of Technology Gliwice”, 2019.

Powyższy wykres przedstawia jak obserwacje oraz średnie z próby dla każdej z klas rzutują na prostą oznaczoną literą „a”. Prostą dyskryminacyjną nazywamy prostą, która pada prostopadle do prostej oznaczoną literą „a”.[[28]](#footnote-29)

## 2.4 Badanie empiryczne

### Zbiór danych

Do przeprowadzenia badań posłużono się trzema zestawami danych stworzonymi w oparciu o wyżej wspomnianą bazę danych. Każdy z zestawów odnosi się do innej gałęzi gospodarki. Wzorując się na badaniach jakie prowadził E. Altman, ilość spółek, które zbankrutowały, jest podobna do tych, którym udało utrzymać się na rynku.

Pierwszy zestaw odnosi się do gałęzi gospodarki jaką jest rolnictwo i zawiera dane na temat siedemnastu przedsiębiorstw. Dane odnoszą się do roku 2013.

* 9 z nich nie zbankrutowało (NB)
* 8 z nich zbankrutowało (B)

Drugi zestaw odnosi się do gałęzi gospodarki jaką jest budownictwo i zawiera dane na temat trzydziestu dwóch przedsiębiorstw. Dane odnoszą się do roku 2015.

* 18 z nich nie zbankrutowało (NB)
* 14 z nich zbankrutowało (B)

Trzeci zestaw odnosi się do gałęzi gospodarki jaką jest przemysł i zawiera dane na temat dwudziestu ośmiu przedsiębiorstw. Dane odnoszą się do roku 2015.

* 14 z nich nie zbankrutowało (NB)
* 14 z nich zbankrutowało (B)

### Podstawowe statystyki opisowe

Poniżej zostały przedstawione podstawowe statystyki opisowe dla każdego zbioru danych.

Vars – Poszczególne numery zmiennych

N ­– Liczba obserwacji

Mean – Średnia dla każdej ze zmiennych

Sd – Odchylenie standardowe

Median – Mediana

Mad – Średnie odchylenie bezwzględne

Min – Wartość minimalna

Max – Wartość maksymalna

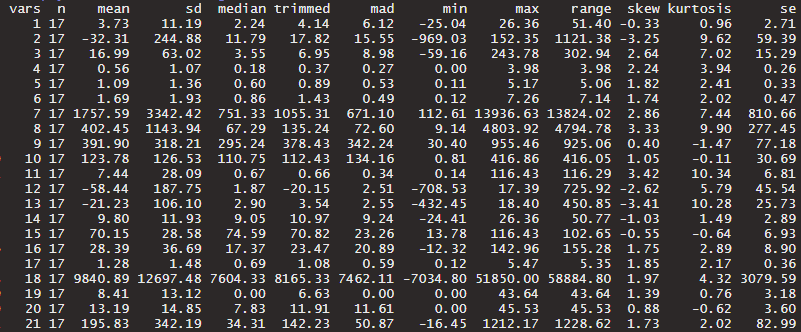
Range – Wielkość przedziału (wartość minimalna + wartość maksymalna)

Skew – Skośność

Kurtosis – Kurtoza

Se – Błąd standardowy

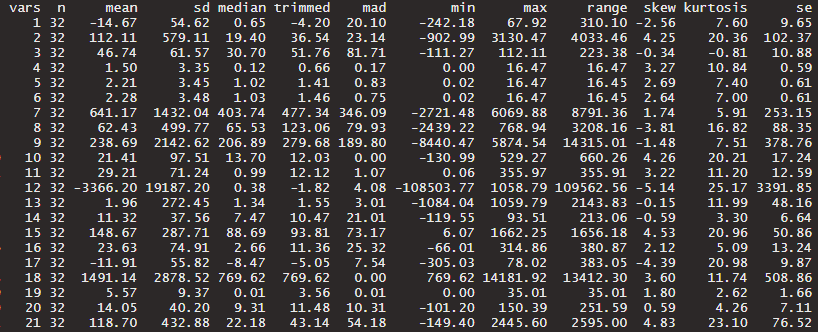
* Dla sektora rolniczego



Ilustracja 2.1. Statystyki opisowe

Źródło: opracowanie własne

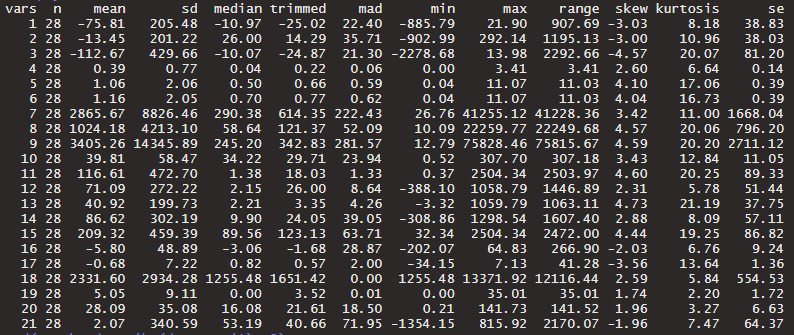
* Dla sektora budowlanego



Ilustracja 2.2. Statystyki opisowe

Źródło: opracowanie własne

* Dla sektora przemysłowego



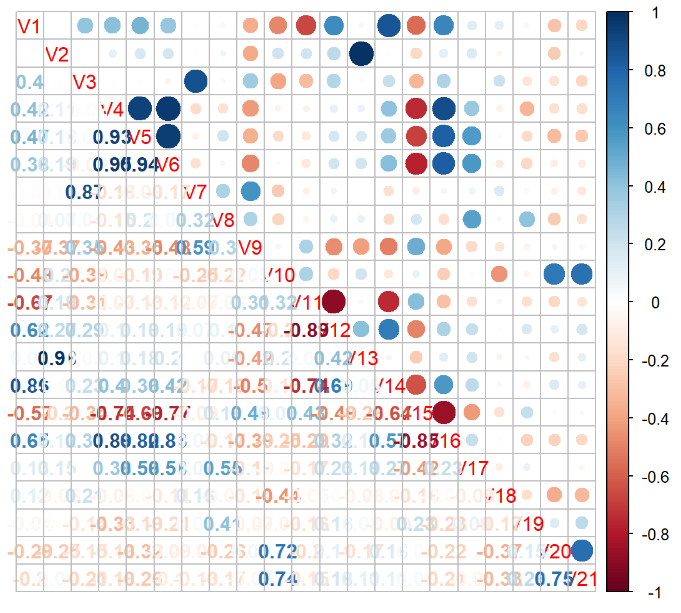
Ilustracja 2.3. Statystyki opisowe

Źródło: opracowanie własne

### Korelacja

Zmienne objaśniające mogą zawierać taką samą informację - aby wykluczyć użycie dwóch takich zmiennych w modelu wykonano analizę współczynnika korelacji.

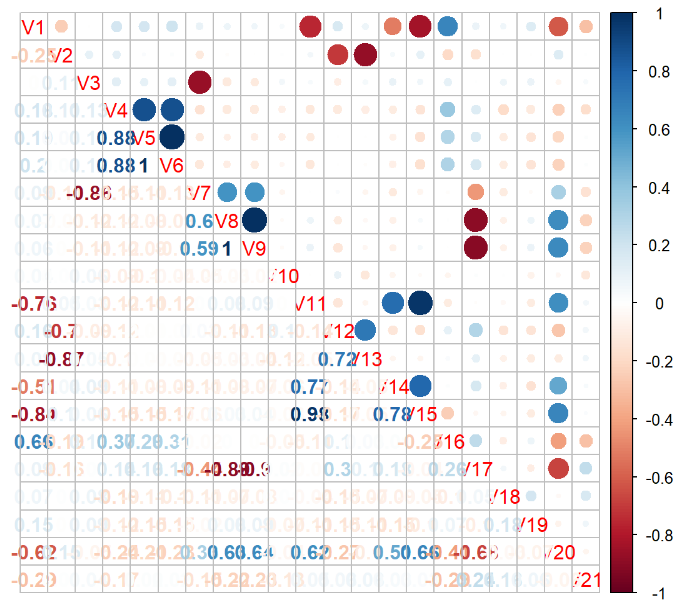
Korelacja Rolnictwo



Wykres 2.6. Korelacja dla poszczególnych zmiennych (rolnictwo)

Źródło: opracowanie własne

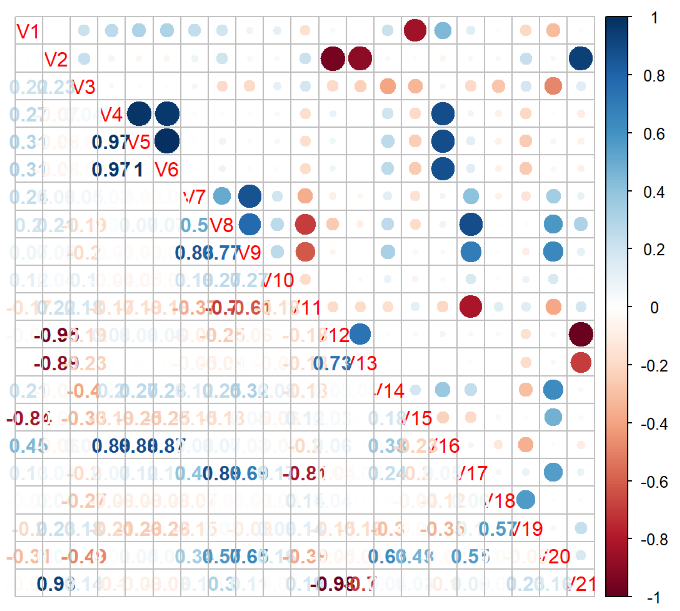
Korelacja Przemysł



Wykres 2.7. Korelacja dla poszczególnych zmiennych (przemysł)

Źródło: opracowanie własne

Korelacja Budownictwo



Wykres 2.8. Korelacja dla poszczególnych zmiennych (budownictwo)

Źródło: opracowanie własne

### Implementacja oraz przeprowadzenie metody

Aby w odpowiedni sposób przeprowadzić analizę dyskryminacyjną, zalecane jest wcześniejsza obróbka danych na którą składają się:

* Transformacja danych w celu przybliżenia rozkładów do rozkładu normalnego
* Usunięcie obserwacji odstających
* Standaryzację danych

Test Shapiro-Wilka

W celu zbadania normalności rozkładów poszczególnych zmiennych podjęto decyzje o przeprowadzeniu testu Shapiro-Wilka. Jest to jeden z najlepszych testów do badania normalności rozkładów, jego główną zaletą jest duża moc.[[29]](#footnote-30). Hipotezy danego testu mają następującą postać:

Jeśli obliczona wartość statystyki jest mniejsza niż wartość krytyczna należy odrzucić hipotezę zerową i przyjąć, że badana cecha ma rozkład inny od rozkładu normalnego. Poniżej przedstawiono wartości krytyczne dla liczby obserwacji jakie występują w tych badaniach. Pierwsza kolumna zawiera liczbę obserwacji. Natomiast następne poziomy istotności to kolejno 0.01, 0.02, 0.05, 0.1.

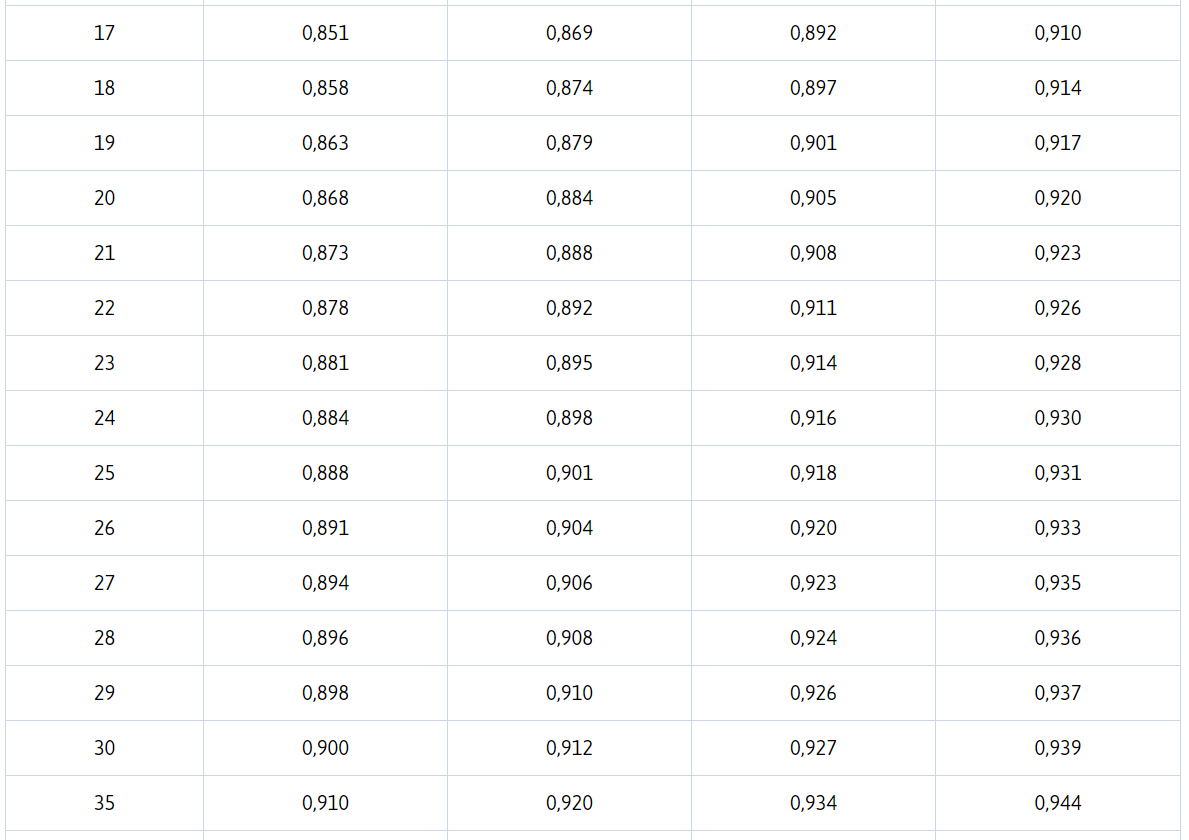
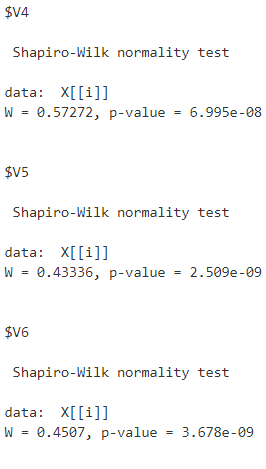


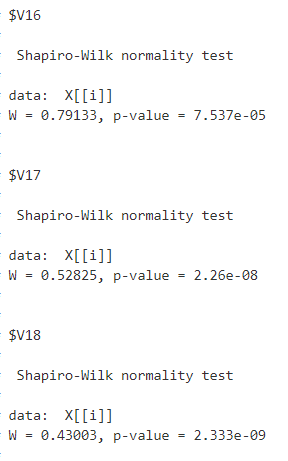
Tabela 2.1. Wartości krytyczne dla testu Shapiro-Wilka

Źródło: <https://www.naukowiec.org/tablice/statystyka/rozklad-w-test-shapiro-wilka-_335.html> [dostęp:20.06.2020]



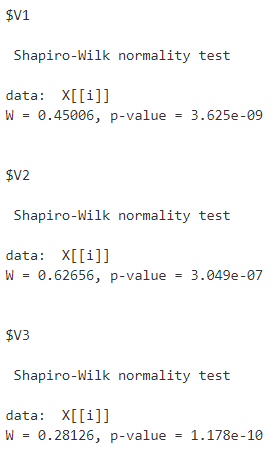
Ilustracja 2.4. Wyniki testu Shapiro-Wilka dla wybranych zmiennych (rolnictwo)

Źródło: opracowanie własne



Ilustracja 2.5. Wyniki testu Shapiro-Wilka dla wybranych zmiennych (przemysł)

Źródło: opracowanie własne



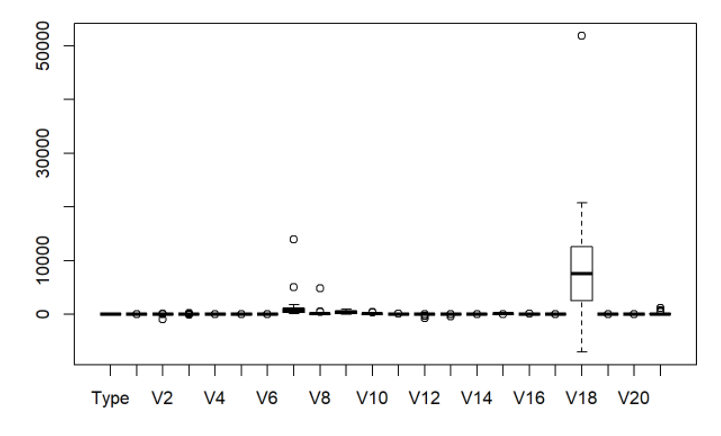
Ilustracja 2.6. Wyniki testu Shapiro-Wilka dla wybranych zmiennych (budownictwo)

Źródło: opracowanie własne

Nie ulega zatem wątpliwości, że na podstawie przeprowadzonych testów należy odrzucić , czyli hipotezę o normalności rozkładów badanych cech. W związku z tym konieczna będzie transformacja danych.

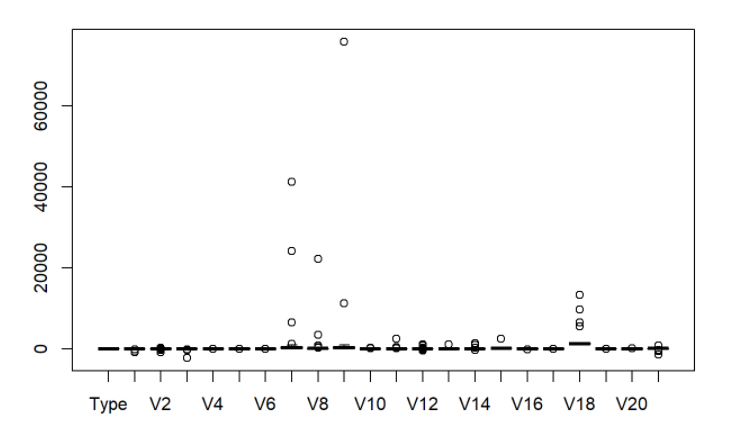
### Usunięcie obserwacji odstających

Obserwacje odstające to takie, które w znacznym stopniu odbiegają od pozostałych. To czy zostaną usunięte czy też nie może mieć znaczący wpływ na wynik testów statystycznych oraz innych przeprowadzonych badań.



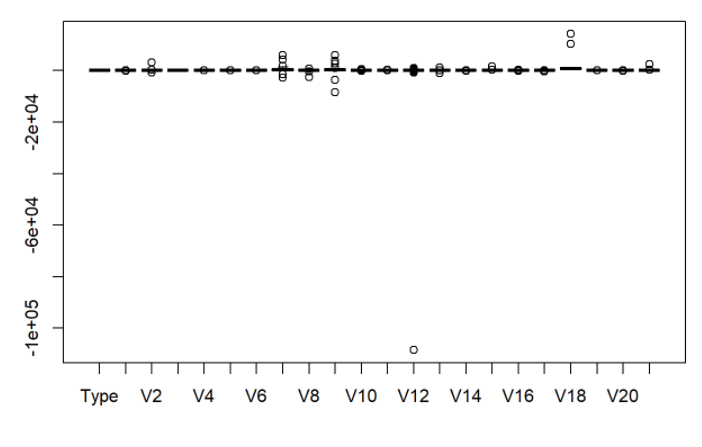
Wykres 2.9. Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo)

Źródło: opracowanie własne



Wykres 2.10. Wykresy pudełkowe dla poszczególnych zmiennych (przemysł)

Źródło: opracowanie własne



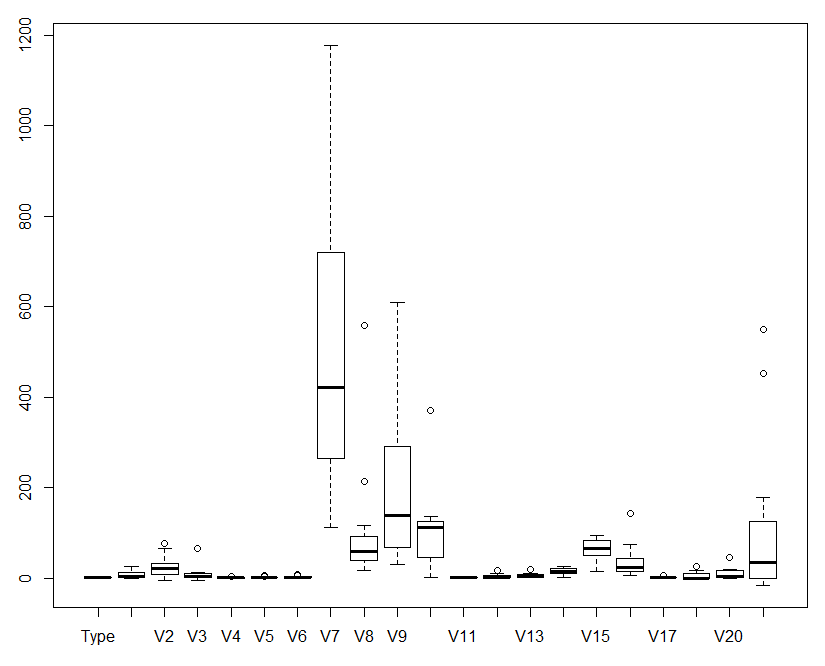
Wykres 2.11. Wykresy pudełkowe dla poszczególnych zmiennych (budownictwo)

Źródło: opracowanie własne

Analizując wykresy pudełkowe dla poszczególnych zmiennych, nie sposób nie zauważyć, iż największe różnicę występują zazwyczaj przy zmiennej V18 wynika to z faktu, iż w przypadku większości obserwacji zmienna ta była nieznana i została uzupełniona na podstawie średniej wyliczonej z dostępnych danych. Zdecydowano zatem o nieuwzględnianiu jej w dalszym toku badań. Ilość obserwacji odstających jakie występowały w zestawie odnoszącym się do przemysłu, czy budownictwa również była bardzo duża, dlatego w dalszych badaniach postanowiono pracować jedynie na zestawie danych odnoszącym się do gałęzi gospodarki jaką jest rolnictwo. Pozostałe zestawy danych posłużą by sprawdzić jaka zmienia się funkcja dyskryminacyjna dla poszczególnych sektorów gospodarki i to w jakim stopniu zmienne objaśniające mają wpływ na przynależność do klas.

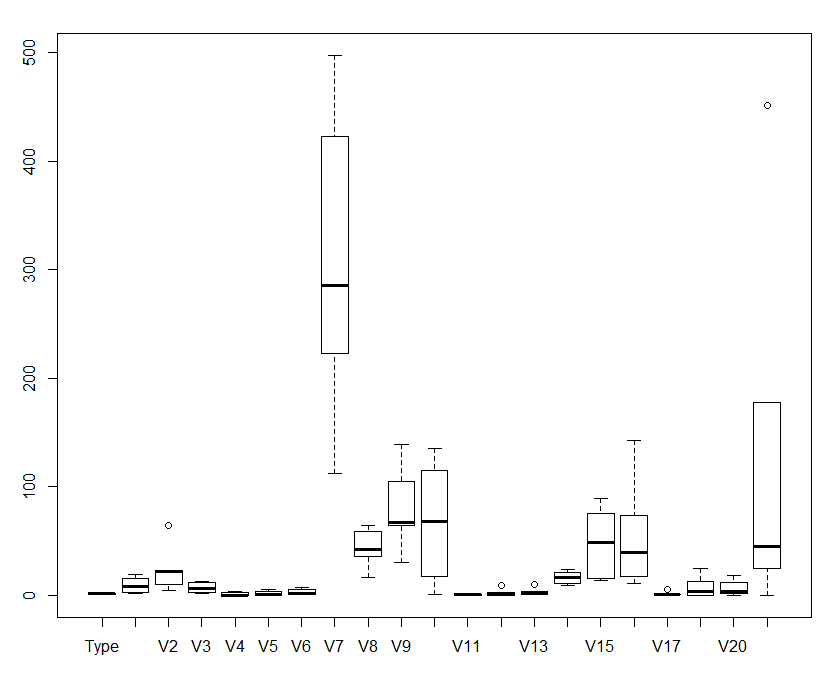
Wykresy pudełkowe - rolnictwo

Po usunięciu zmiennej V18 oraz kilku z najbardziej odstających obserwacji wykresy pudełkowe prezentowały się w sposób następujący (zbiór danych zawierał wtedy już tylko 11 obserwacji oraz 21 zmiennych).



Wykres 2.12. Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo)

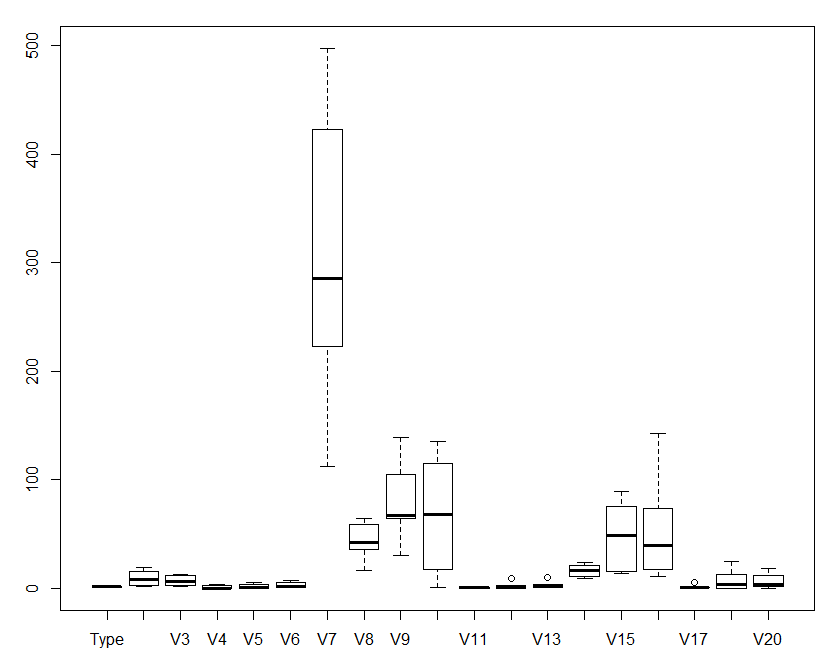
Źródło: opracowanie własne

Okazało się jednak że zestaw danych dalej zawiera obserwacje, które odstają w znacznym stopniu od reszty. Więc postanowiono usunąć kolejne z nich.

Wykres 2.13. Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo)

Źródło: opracowanie własne

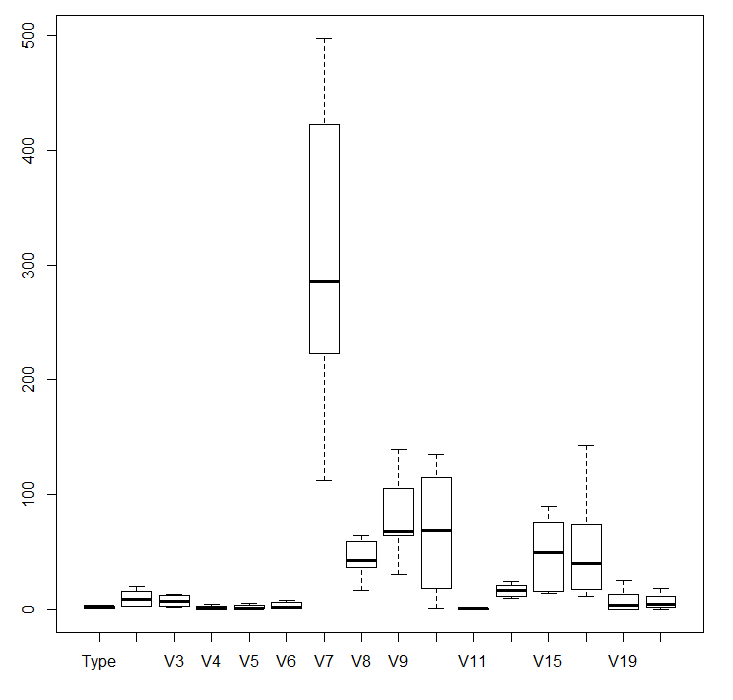
Kolejne zmienne które determinowały powstawanie obserwacji odstających były zmienne V2 oraz V21. Mimo, iż informacja jaką niosły te zmienne była ciekawa z punktu widzenia postawionych hipotez. Patrząc na wykresy korelacji, nie była zupełnie unikatowa, gdyż zarówno jedna jak i druga zmienna posiadały ujemną korelacje wyższą od przynajmniej z jedną zmienną. Zmienna V21 posiadała dwie zmienne, które miały korelacje wyższą od . Ponadto jeśli usunięto by obserwacje które odpowiadały za odstawanie w zbiorze zostały by jedynie obserwacje odpowiadające firmom, które nie zbankrutowały co nie było by korzystne dla tych badań. Poniżej przedstawiono wykres po usunięciu wyżej wymienionych zmiennych i zbiorze danym posiadającym 6 obserwacji 4 odnośnie spółek, które utrzymał się na rynku oraz 2 odnośnie tych którym się to nie udało.



Wykres 2.14. Wykresy pudełkowe dla poszczególnych zmiennych po usunięciu zmiennej V2 oraz V21 (rolnictwo)

Źródło: opracowanie własne

Jak widać na powyższym wykresie z obserwacjami odstającymi mamy do czynienia w przypadku jeszcze trzech zmiennych V12 oraz V13. Obserwacje te tyczą się spółek, które zbankrutowały. Ze względu na fakt, iż w zbiorze danych mamy na chwile obecną jedynie 2 takie spółki, by pozbyć się obserwacji odstających usunięto te zmienne. Obserwacja odstająca w V17 tyczyła się spółki która nie zbankrutowała lecz i w tym przypadku usunięto zmienną. Ostatecznie zbiór nie posiadał żadnych obserwacji odstających.



Wykres 2.15. Wykresy pudełkowe dla poszczególnych zmiennych po usunięciu zmiennych V12, V13 oraz V17 (rolnictwo)

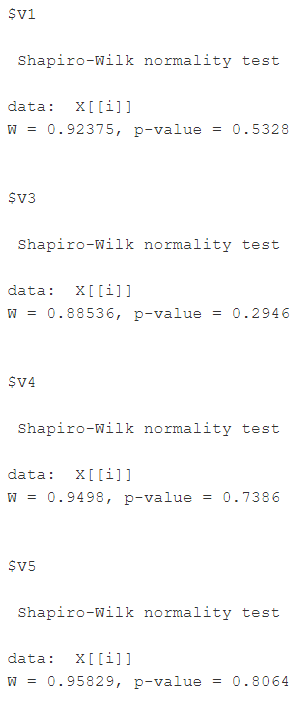
Źródło: opracowanie własne

Po procesie usuwania obserwacji odstających zbiór danych zawierał 6 obserwacji oraz 16 zmiennych, z których jedna mówiła o tym, czy dane przedsiębiorstwo zbankrutowało, czy też nie (Type).

### Transformacja danych w celu przybliżenia rozkładów do rozkładu normalnego

Rolnictwo

Z uwagi na fakt iż, po przeprowadzeniu testu okazało się że praktycznie wszystkie ze zmiennych nie przechodzą testu na normalność, podjęto kroki w celu przybliżenia rozkładów do rozkładu normalnego. W tym celu przetransformowano każdą ze zmiennych. W większości przypadków posłużono się transformacją opartą na potęgowaniu, czyli podnoszeniem poszczególnych zmiennych do potęgi takiej która sprawiała że zmienna objaśniająca miała rozkład zbliżony do rozkładu normalnego.



Ilustracja 2.7. Wyniki testu Shapiro-Wilka dla wybranych zmiennych po transformacji danych (rolnictwo)

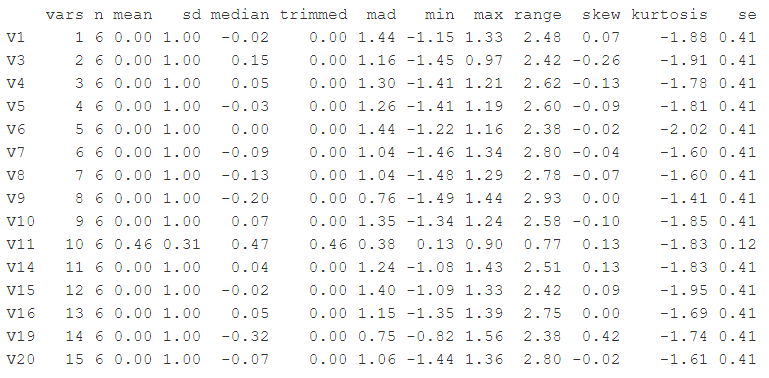
Źródło: opracowanie własne

Jak można zauważyć na powyższej ilustracji statystyka testowa *W* wzrosła w znacznym stopniu w przypadku wszystkich zmiennych. Także p-value znacznie wzrosło. Ukazane wyniki świadczą, iż poszczególne zmienne zbliżyły się do rozkładu normalnego w znacznym stopniu.

### Standaryzacja danych

Rolnictwo

Przed przeprowadzeniem analizy dyskryminacyjnej zaleca się również standaryzację danych . Zmienna uzyskuje wtedy średnią równą zero oraz odchylenie standardowe równe 1.[[30]](#footnote-31) W tym celu wykorzystano funkcję *scale().*

**

Ilustracja 2.8. Podstawowe statystyki opisowe po standaryzacji danych (rolnictwo)

Źródło: opracowanie własne

Jak widać standaryzacja przyniosła zamierzony efekt. Średnie wszystkich zmiennych wynoszą 0 natomiast odchylenie standardowe są równe 1.

### Współliniowość zmiennych objaśniających

Rolnictwo

Zanim ostatecznie wyestymowano model wymagane było jeszcze usunięcie zmiennych współliniowych. Model posiadał aż 16 zmiennych objaśniających jest to bardzo duża liczba w przypadku analizy dyskryminacyjnej. Standardowo modele zawierają dużo mniejszą liczbę zmiennych. Budując model, dobierano zatem zmienne w ten sposób, by współliniowość między nimi była jak najmniejsza, a zatem informacja niesiona ze przez zmienną jak najbardziej unikalna. Ostatecznie model, który zdecydowano się utworzyć, zawierał 4 zmienne. Współczynniki współliniowości wyglądały w sposób następujący.



Ilustracja 2.9. Współczynniki współliniowości pomiędzy zmiennymi użytymi do modelu (rolnictwo)

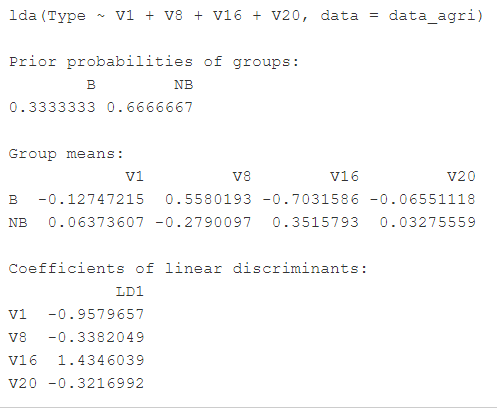
Źródło: opracowanie własne

Nie jest to w prawdzie zupełny brak współliniowości, ale żaden ze współczynników nie przekracza 3.3, co jest wartością jak najbardziej zadowalającą.

## 2.5 Wyniki eksperymentu badawczego

Rolnictwo

Finalny model zawierał cztery zmienne objaśniające V1, V8, V16 oraz V20.



Ilustracja 2.10. LDA (rolnictwo)

Źródło: opracowanie własne

Liniowa funkcja dyskryminacyjna miała następującą postać:

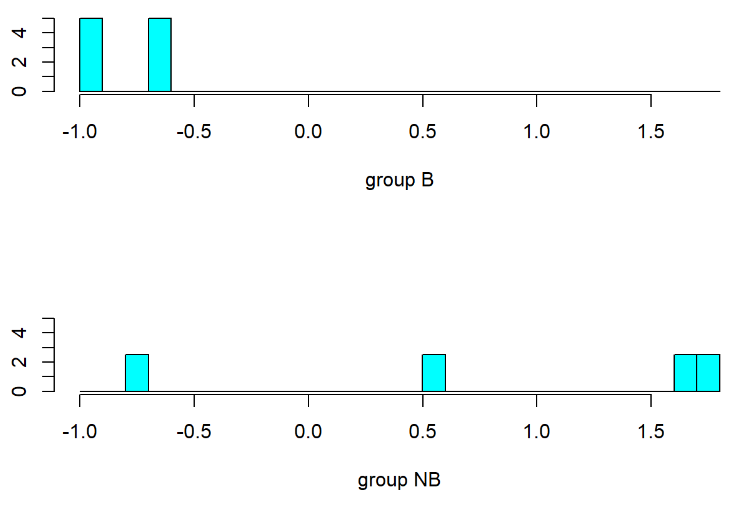
(12)

V1 - [TAT]

V8 - [ROA]

V16 - [DE]

V20 - [ACR]

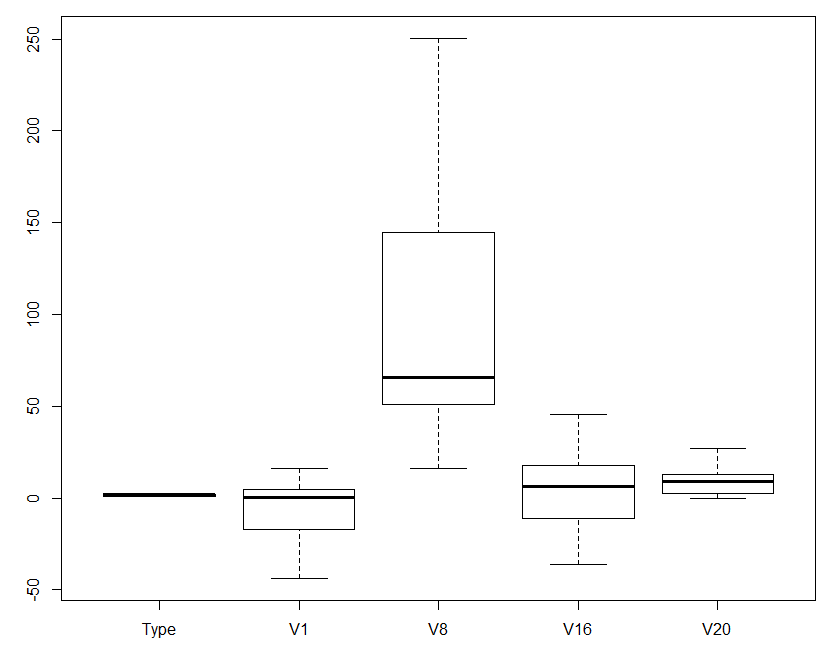


Wykres 2.16. Wykres LDA (rolnictwo)

Źródło: Opracowanie własne

Patrząc na liniową funkcję dyskryminacyjną, można zauważyć, iż zmienną która ma największy wpływ na przynależność do danej grupy była zmienna V16. Współczynnik liniowy przy tej zmiennej wynosi 1,43. Niejako potwierdza to jedną z wcześniej stawianych hipotez, iż najważniejsze dyskryminatory to takie, które wiążą się z zadłużeniem. Kolejną zmienną, która miała bardzo duży wpływ na przynależność do poszczególnych klas, była zmienna V1, w której jedną ze składowych była całkowita sprzedaż (to również znajduje odzwierciedlenie w hipotezach stawianych na początku).

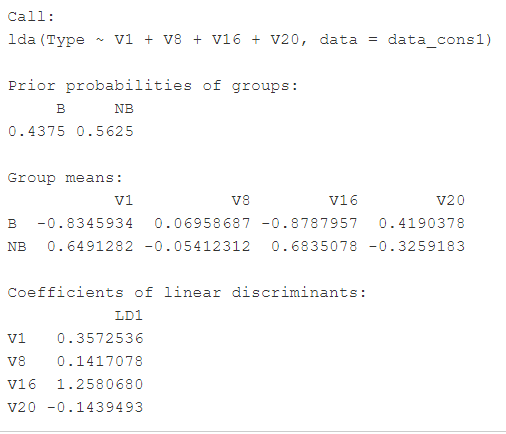
Budownictwo



Wykres 2.17. Wykresy pudełkowe po usunięciu obserwacji odstających (budownictwo)

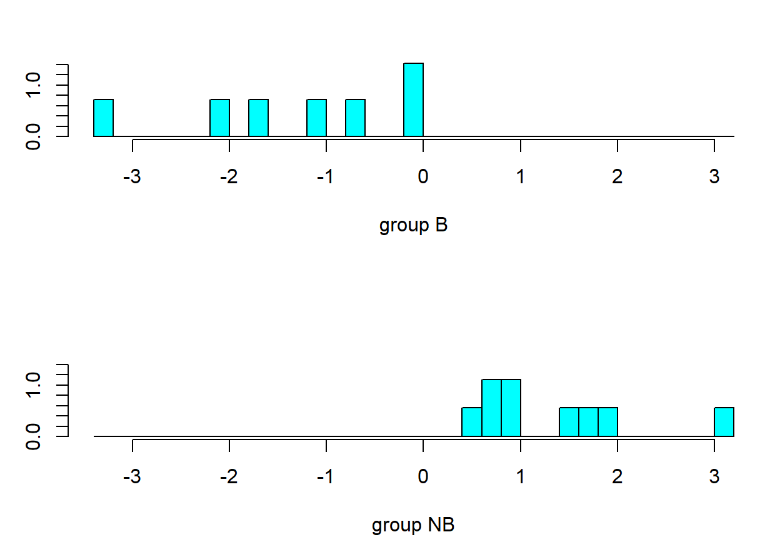
Źródło: Opracowanie własne

Po usunięciu obserwacji odstających zestaw danych zawierał 16 obserwacji z czego 7 odnosiło się do spółek, które zbankrutowały natomiast kolejne 9 to tych, które utrzymały się na rynku. Po sprowadzeniu rozkładów zmiennych do rozkładów normalnych oraz standaryzacji otrzymano następujący model.



Ilustracja 2.11. LDA (budownictwo)

Źródło: opracowanie własne



Wykres 2.18. Wykres LDA (budownictwo)

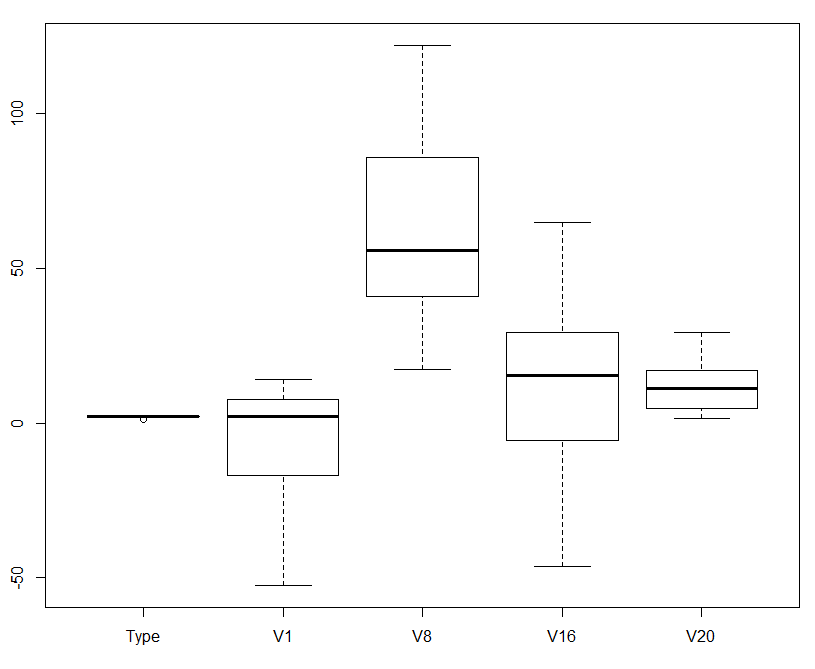
Źródło: Opracowanie własne

Liniowa funkcja dyskryminacyjna miała następującą postać:

(13)

Także i w tym przypadku zmienną objaśniającą mającą największy wpływ na przynależność do grup była zmienna V16. Współczynniki przy pozostałych zmiennych również różniły się znacząco. Ponadto przy niektórych zamiast wartości dodatniej mieliśmy wartość ujemną.

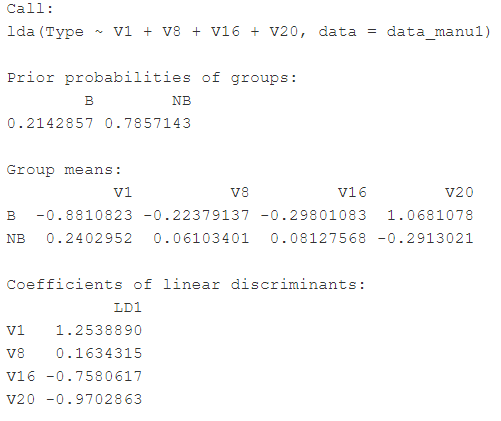
Przemysł



Wykres 2.19. Wykresy pudełkowe po usunięciu obserwacji odstających (przemysł)

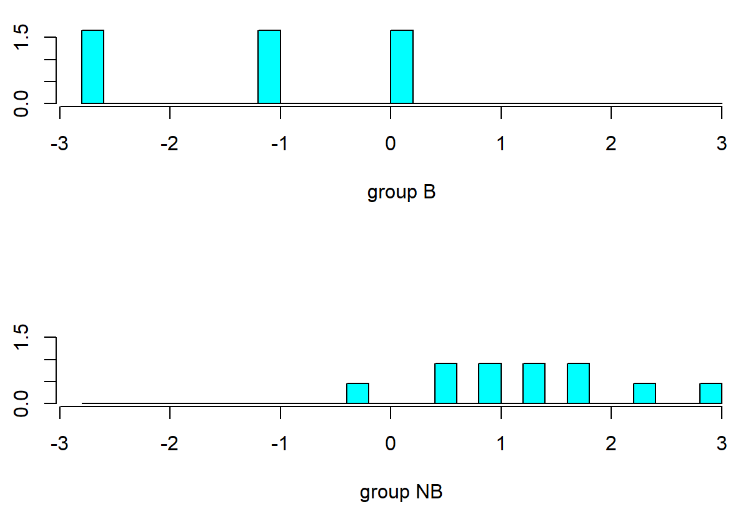
Źródło: Opracowanie własne

Obserwacje odstające w przypadku zmiennej *Type* wynikają z fakt iż w zbiorze danych zostały jedynie 3 przedsiębiorstwa które nie zbankrutowały, natomiast wszystkich obserwacji było 14*.*

**

Ilustracja 2.12. LDA (przemysł)

Źródło: opracowanie własne



Wykres 2.20. Wykres LDA (przemysł)

Źródło: Opracowanie własne

Liniowa funkcja dyskryminacyjna miała w tym przypadku następującą postać:

(14)

Jak widać w tym przypadku najważniejszym dyskryminatorem była zmienna V1.

W pozostałych przypadkach jak wcześniej analizowanych przykładach zmieniły się wartości przy zmiennych objaśniających. Zmienna V1 zawiera wskaźnik jakim jest całkowita sprzedaż. Otrzymane wyniki pokazują że dla gałęzi przemysłu właśnie ten wskaźnik ma największe znaczenie. Znikoma jest rola zmiennej V8 która zawiera wskaźnik związany z zyskiem netto. Nasuwają się wnioski, że dla przedsiębiorstw przemysłowych nie tak ważna jak zysk jest całkowita sprzedaż.

## 2.6 Ocena efektywności

To jak dobrze sprawdza się dany model można ocenić po procencie przedsiębiorstw, które zostały dobrze sklasyfikowane stosując ten model dla innego zbioru danych. Badania wydają się być efektywne jako że ich rezultaty pokazały różnice pomiędzy poszczególnymi gałęziami przemysłu. Inaczej mówiąc, wyniki badań udowodniły, iż w każdym sektorze gospodarki inne zmienne objaśniające wpływają na przynależność do grup. Na podstawie badań można określić nie tylko, które zmienne objaśniające mają największy wpływa na przynależność do grup ale również o ile procent różni się moc dyskryminatorów pomiędzy poszczególnymi gałęziami gospodarki. Przed zaimplementowaniem metody użyto odpowiednich narzędzi statystycznych tak by wyniki analiz były jak najrzetelniejsze. W tym celu zestawy danych poddano wielu testom a obserwacje odstające zostały usunięte. Miało to znaczący wpływ na efektywność.

## 2.7 Możliwe dalsze kierunki badań

Posiadając wiedzę nabytą z badań prowadzonych w tej pracy oraz modele jakie udało się wyestymować. Możliwe było by wykorzystanie ich do predykcja upadłości dla spółek z podobnej gałęzi gospodarki w przyszłych latach. Informacje takie, jeśli model by się sprawdził, były by niezwykle cenne i użyteczne ponieważ pozwalałby na szybkie identyfikowanie nadchodzących zagrożeń finansowych. Mówiąc o predykcji na przyszłe lata należy liczyć się z faktem iż zmiany na rynku jakie przyniosła pandemia koronawirusa w znacznym stopniu ograniczają możliwości prognostyczne. Choć jest też możliwe, że zmiany te nie wpłyną na to jakie wartości wskaźników finansowych wróżą bankructwo.

Co więcej mówiąc o dalszym kierunku badań nie sposób nie wspomnieć o możliwości zastosowania metody dla innych sektorów gospodarki. Mając na uwadze wyniki jakie otrzymano w tych badaniach, można przewidywać, że gdybyśmy dysponowali danymi z gałęzi jaką jest na przykład turystyka moc dyskryminatorów dla zmiennych jakie zostały użyte zmieniła by się.

# 3. Podsumowanie

Analiza dyskryminacyjna jest szerokim pojęciem obejmującym kilka podobnych technik statystycznych wykorzystywanych do testowania hipotezy. W tej pracy skupiono się jednak na liniowej analizie dyskryminacyjnej czyli podejściu jakie zapoczątkowało rozwój technik klasyfikacyjnych. Ta metoda, choć jej idea nie jest niczym nowym w świecie nauki, dalej się sprawdza głównie dzięki rozwojowi technologii, który sprawił że jej użyteczność cały czas rośnie. Problem klasyfikacji obiektów jaki pozwala rozwiązać liniowa funkcja dyskryminacyjna nie tyczy się tylko badań jakie prowadzi się w biznesie jest on obecny nawet w medycynie. To pokazuje jak szerokie spektrum zastosowania ma analiza dyskryminacyjna i jak użyteczna może być dla ludzkości.

# Wnioski końcowe

Przeprowadzone badania miały na celu sprawdzić czy w poszczególnych gałęziach gospodarki inne zmienne determinują przynależność do grupy przedsiębiorstw które nie mają problemów finansowych oraz tych którym grozi bankructwo. Innymi słowy sprawdzono jak wyglądają funkcje dyskryminacyjne dla różnych sektorów. Rezultaty sprawiają że dwie z trzech postawionych hipotezy wydają się być prawdziwe. Dla dwóch gałęzi gospodarki, rolnictwa oraz budownictwa, największą moc miał dyskryminator którego składową było całkowite zadłużenie, co potwierdza hipotezę, że zmienna, która będzie miała duży wpływ na przynależność do grup będzie związana z zobowiązanymi krótko i długo terminowymi. Jeśli spółka ma problem z zobowiązanymi krótko i długoterminowymi siłą rzeczy całkowite zadłużenie - które jest wskaźnikiem wchodzącym w skład zmiennej V16 - rośnie. Inaczej jest w przypadku trzeciego modelu który, odnosi się do przemysłu w tym przypadku zmienna ta ma ujemną wartość współczynnika liniowego. O ile modele odnoszące się do budownictwa i rolnictwa są częściowo podobne to model dla przemysłu jest zupełnie inny wynikać, co może z faktu, że przedsiębiorstwa przemysłowe mają inny charakter. Jest też prawdopodobne to że obserwacje z tej gałęzi są spółkami które zatrudniają większą liczbę pracowników stąd te różnice. Choć badania prowadzone są dla grupy jaką są małe i średnie przedsiębiorstwa to i tej grupie mogą występować znaczące różnice w rozmiarze prowadzonej działalności. Dla sektora przemysłowego najważniejsza jest całkowita sprzedaż co pozwala wnioskować, iż przedsiębiorstwa z tej gałęzi chcąc utrzymać się na rynku powinny maksymalizować wypływ produktów z firmy. Jedyna zmienna która we wszystkich modelach ma współczynnik o wartości ujemnej to zmienna V20 której składowymi są *wartość majątku netto+pasywa /środki trwałe* z uwagi na ten fakt możemy wnioskować że zmienna ta ma podobny wpływ we wszystkich gałęziach gospodarki. Dla przedsiębiorstw przemysłowych wpływ na przynależność do grup jest jednak znacznie większy.

# Literatura

## Pozycje zwarte

N. Grzenkowicz, J. Kowalczyk, A. Kusak, Z. Podgórski, M. Ambroziak, *Podstawy funkcjonowania przedsiębiorstw,* „Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego”, Warszawa, 2008.

S.Juszczyk, R. Balin , *Prognozowanie ryzyka bankructwa przedsiębiorstw z branży transportu drogowego towarów*, „Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego. Ekonomika i Organizacja Gospodarki Żywnościowej” , nr 104, 2013.

E. Gantar, *Analiza dyskryminacyjna - stan aktualny i kierunki rozwoju*, ,,Prace Naukowe Uniwersytetu Ekonomicznego w Katowicach", nr 152, 2013.

E.Grzegorzewska, H.Runowski, *Zdolności prognostyczne polskich modeli dyskryminacyjnych w badaniu kondycji finansowej przedsiębiorstw rolniczych*, „Roczniki Nauk Rolniczych”, nr 95.3/4, 2008.

J. Kisielińska, A. Waszkowski. *Polskie modele do prognozowania bankructwa przedsiębiorstw i ich weryfikacja*, „Zeszyty Naukowe SGGW”, nr 82, 2010.

W. Pierzchalska, *Analiza finansowa oraz modele przewidywania upadłości jako narzędzie oceny kondycji finansowej przedsiębiorstwa*, „Finanse i Prawo Finansowe”, nr 1 Łódź, 2014.

R. Bejger, D. Gołębiowska, P. Nicia, „*Wykorzystanie analizy dyskryminacyjnej do porównania ekstraktów torfowych*” , „Roczniki gleboznawcze”, nr 2, tom LX Warszawa 2009.

D. Raedler,. *Identyfikacja nowych immunofenotypów w przypadku alergicznej i niealergicznej astmy u dziec*, „Alergologia Polska-Polish Journal of Allergology”, nr 3.2 , 2016.

E.Mazur-Stążka, *Komputerowe wspomaganie procesu diagnozowania choroby wieńcowej*, „Cardiovascular Forum. Vol. 5. No. 1. Via Medica Medical Publishers”, 2000.

J. Moczko *Niestandardowe techniki wspomagania diagnostyki kardiologicznej* , 2008

M. Markowska, *Zastosowanie analizy dyskryminacyjnej do oceny zdolności kredytowej–case study*, „zeszyty naukowe wwszip", nr 4, 2004, s. 86.

A. Leśniak, *Wspomaganie decyzji przetargowych wykonawców z zastosowaniem metod matematycznych* 2018

E. Mączyńska, M. Zawadzki *Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw*, „Ekonomista ” nr 2, 2006.

W. H. Beaver, *Financial ratios as predictors of failure*, “Journal of accounting research”, nr 2, 1966.

R. Krupski,. *Kontekst turbulencji otoczenia w badaniach empirycznych nad strategiami przedsiębiorstw*, „Przegląd Organizacji”, nr 9, 2011.

R. Wolański, *Wpływ otoczenia finansowego na konkurencyjność małych i średnich przedsiębiorstw*, „Wolters Kluwer”, 2013.

B. Korkozowicz, *Małe i średnie przedsiębiorstwa na świecie*, „Polska Fundacja Promocji i Rozwoju Małych i Średnich Przedsiębiorstw”, 2000.

M. Kasjaniuk, *Zastosowanie analizy dyskryminacyjnej do modelowania i prognozowania kondycji przedsiębiorstw*, „Barometr Regionalny”, nr 2.6, 2006.

E. Gantar, *Analiza dyskryminacyjna - stan aktualny i kierunki rozwoju*, ,,Prace Naukowe Uniwersytetu Ekonomicznego w Katowicach", nr 152, 2013.

J. Pociecha *Problemy prognozowania bankructwa firmy metodą analizy dyskryminacyjnej* „[Acta Universitatis Lodziensis. Folia Oeconomica”, nr 205, 2006.](http://dspace.uni.lodz.pl/xmlui/handle/11089/5938)

Błonski, K., & Brozek, O. *Handwritten digit recognition using discriminant analysis*, “Faculty of Applied Mathematics, Sielsian University of Technology Gliwice”, 2019.

S.Godlewska, *Skuteczność polskich modeli dyskryminacyjnych w ocenie zagrożenia upadłością spółek giełdowych*,  „Annales Universitatis Mariae Curie-Skłodowska. Sectio H. Oeconomia”, nr  44.2, 2010.

## Strony internetowe

<https://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tec00114>

<http://polskoslowackaizba.pl/gospodarka-slowacji-w-2019-roku/>

<https://www.sciencedirect.com/science/article/pii/S2352340919307140#bbib2>

<https://www.statystyka.az.pl/centrum-statystyki/test-normalnosci-shapiro-wilka.php>

<https://www.naukowiec.org/tablice/statystyka/rozklad-w-test-shapiro-wilka-_335.html>

<http://manuals.pqstat.pl/statpqpl:usepl:arkpl:normstandpl>

## Akty prawne

Komisji Wspólnoty Europejskiej nr 800/2008 z dnia 6 sierpnia 2008r.

# Ilustracje, schematy i wykresy

## Spis ilustracji

|  |  |
| --- | --- |
| Ilustracja 1.1 | Niektóre z bibliotek wykorzystane do przeprowadzenia badań |
| Ilustracja 2.1 | Statystyki opisowe |
| Ilustracja 2.2 | Statystyki opisowe |
| Ilustracja 2.3 | Statystyki opisowe |
| Ilustracja 2.4 | Wyniki testu Shapiro-Wilka dla wybranych zmiennych (rolnictwo) |
| Ilustracja 2.5 | Wyniki testu Shapiro-Wilka dla wybranych zmiennych (przemysł) |
| Ilustracja 2.6 | Wyniki testu Shapiro-Wilka dla wybranych zmiennych (budownictwo) |
| Ilustracja 2.7 | Wyniki testu Shapiro-Wilka dla wybranych zmiennych po transformacji danych (rolnictwo) |
| Ilustracja 2.8 | Podstawowe statystyki opisowe po standaryzacji danych (rolnictwo) |
| Ilustracja 2.9 | Współczynniki współliniowości pomiędzy zmiennymi użytymi do modelu (rolnictwo) |
| Ilustracja 2.10 | LDA (rolnictwo) |
| Ilustracja 2.11 | LDA (budownictwo) |
| Ilustracja 2.12 | LDA (przemysł) |

## Spis schematów

|  |  |
| --- | --- |
| Schemat 1.1 | Podział metod predykcji upadłości |
| Schemat 2.1 | Cykl życia przedsiębiorstwa |
| Schemat 2.2 | Podział przedsiębiorstw ze względu na przepisy unijne |

## Spis wykresów

|  |  |
| --- | --- |
| Wykres 1.1 | Wykorzystanie technik predykcji upadłości do badań prowadzonych na świecie |
| Wykres 1.2 | Przykładowa analiza dyskryminacyjna dla irysów |
| Wykres 2.1 | Udział poszczególnych sektorów gospodarki Słowackiej w ogólnej wartości PKB |
| Wykres 2.2 | Udział mikro-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji |
| Wykres 2.3 | Udział małych-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji |
| Wykres 2.4 | . Udział średnich-przedsiębiorstw w poszczególnych dziedzinach gospodarki w Słowacji |
| Wykres 2.5 | Prosta analiza dyskryminacyjna |
| Wykres 2.6 | Korelacja dla poszczególnych zmiennych (rolnictwo) |
| Wykres 2.7 | Korelacja dla poszczególnych zmiennych (przemysł) |
| Wykres 2.8 | Korelacja dla poszczególnych zmiennych (budownictwo) |
| Wykres 2.9 | Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo) |
| Wykres 2.10 | Wykresy pudełkowe dla poszczególnych zmiennych (przemysł) |
| Wykres 2.11 | Wykresy pudełkowe dla poszczególnych zmiennych (budownictwo) |
| Wykres 2.12 | Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo) |
| Wykres 2.13 | Wykresy pudełkowe dla poszczególnych zmiennych (rolnictwo) |
| Wykres 2.14 | Wykresy pudełkowe dla poszczególnych zmiennych po usunięciu zmiennej V2 oraz V21 (rolnictwo) |
| Wykres 2.15 | Wykresy pudełkowe dla poszczególnych zmiennych po usunięciu zmiennych V12, V13 oraz V17 (rolnictwo) |
| Wykres 2.16 | Wykres LDA (rolnictwo) |
| Wykres 2.17 | Wykresy pudełkowe po usunięciu obserwacji odstających (budownictwo) |
| Wykres 2.18 | Wykres LDA (budownictwo) |
| Wykres 2.19 | Wykresy pudełkowe po usunięciu obserwacji odstających (przemysł) |
| Wykres 2.20 | Wykres LDA (przemysł) |

## Spis tabel

|  |  |
| --- | --- |
| Tabela 2.1 | Wartości krytyczne dla testu Shapiro-Wilka |

1. N. Grzenkowicz, J. Kowalczyk, A. Kusak, Z. Podgórski, M. Ambroziak, Podstawy funkcjonowania przedsiębiorstw, „Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego”, Warszawa, 2008. [↑](#footnote-ref-1)
2. 2 S.Juszczyk, R. Balin , Prognozowanie ryzyka bankructwa przedsiębiorstw z branży transportu drogowego towarów, „Zeszyty Naukowe Szkoły Głównej Gospodarstwa Wiejskiego. Ekonomika i Organizacja Gospodarki Żywnościowej” , nr 104, 2013. [↑](#footnote-ref-2)
3. E. Gantar, *Analiza dyskryminacyjna - stan aktualny i kierunki rozwoju*, ,,Prace Naukowe Uniwersytetu Ekonomicznego w Katowicach", nr 152, 2013. [↑](#footnote-ref-3)
4. E.Grzegorzewska, H.Runowski, *Zdolności prognostyczne polskich modeli dyskryminacyjnych w badaniu kondycji finansowej przedsiębiorstw rolniczych*, „Roczniki Nauk Rolniczych” ,nr 95.3/4, 2008. [↑](#footnote-ref-4)
5. J. Kisielińska, A. Waszkowski, *Polskie modele do prognozowania bankructwa przedsiębiorstw i ich weryfikacja*, „Zeszyty Naukowe SGGW”, nr 82, 2010. [↑](#footnote-ref-5)
6. W. Pierzchalska*, Analiza finansowa oraz modele przewidywania upadłości jako narzędzie oceny kondycji finansowej przedsiębiorstwa*, „Finanse i Prawo Finansowe” , nr ,1 Łódź, 2014. [↑](#footnote-ref-6)
7. W. Pierzchalska*, Analiza finansowa oraz modele przewidywania upadłości jako narzędzie oceny kondycji finansowej przedsiębiorstwa*, „Finanse i Prawo Finansowe” , nr ,1 Łódź, 2014. [↑](#footnote-ref-7)
8. R. Bejger, D. Gołębiowska, P. Nicia, *Wykorzystanie analizy dyskryminacyjnej do porównania ekstraktów torfowych* , „Roczniki gleboznawcze”, nr 2, tom LX, Warszawa, 2009. [↑](#footnote-ref-8)
9. D. Raedler,. *Identyfikacja nowych immunofenotypów w przypadku alergicznej i niealergicznej astmy u dziec*, „Alergologia Polska-Polish Journal of Allergology”, nr 3.2 , 2016. [↑](#footnote-ref-9)
10. E.Mazur-Stążka, *Komputerowe wspomaganie procesu diagnozowania choroby wieńcowe*j, „Cardiovascular Forum. Vol. 5. No. 1. Via Medica Medical Publishers”, 2000. [↑](#footnote-ref-10)
11. J. Moczko A. *Niestandardowe techniki wspomagania diagnostyki kardiologicznej*, 2008. [↑](#footnote-ref-11)
12. M. Markowska, *Zastosowanie analizy dyskryminacyjnej do oceny zdolności kredytowej–case study*, „zeszyty naukowe wwszip", nr 4, 2004, s. 86. [↑](#footnote-ref-12)
13. A. Leśniak, *Wspomaganie decyzji przetargowych wykonawców z zastosowaniem metod matematycznych* 2018 [↑](#footnote-ref-13)
14. E. Mączyńska, M. Zawadzki *Dyskryminacyjne modele predykcji bankructwa przedsiębiorstw*, „Ekonomista ” nr 2, 2006. [↑](#footnote-ref-14)
15. W. H. Beaver*, Financial ratios as predictors of failure*, “Journal of accounting research”, nr 2, 1966. [↑](#footnote-ref-15)
16. R. Krupski,. *Kontekst turbulencji otoczenia w badaniach empirycznych nad strategiami przedsiębiorstw*, „Przegląd Organizacji”, nr 9, 2011. [↑](#footnote-ref-16)
17. N. Grzenkowicz, J. Kowalczyk, A. Kusak, Z. Podgórski, M. Ambroziak, *Podstawy funkcjonowania przedsiębiorstw,* „Wydawnictwo Naukowe Wydziału Zarządzania Uniwersytetu Warszawskiego*”* Warszawa, 2008. [↑](#footnote-ref-17)
18. H. Nowak*, Determinanty przetrwania nowo tworzonych przedsiębiorstw w województwie wielkopolskim – podejście instytucjonalne, „*Uniwersytet Ekonomiczny”Poznań, 2013. [↑](#footnote-ref-18)
19. R. Wolański, *Wpływ otoczenia finansowego na konkurencyjność małych i średnich przedsiębiorstw,* „Wolters Kluwer”, 2013, s. 19-20. [↑](#footnote-ref-19)
20. <https://ec.europa.eu/eurostat/tgm/table.do?tab=table&plugin=1&language=en&pcode=tec00114> [dostęp:10.06.2020] [↑](#footnote-ref-20)
21. B. Korkozowicz,*Małe i średnie przedsiębiorstwa na świecie*, „Polska Fundacja Promocji i Rozwoju Małych i Średnich Przedsiębiorstw”, 2000, s. 39-41. [↑](#footnote-ref-21)
22. <https://www.sciencedirect.com/science/article/pii/S2352340919307140#bbib2>, [dostęp: 04.06.2020] [↑](#footnote-ref-22)
23. E. Gantar*, Analiza dyskryminacyjna - stan aktualny i kierunki rozwoju*, ,,Prace Naukowe Uniwersytetu Ekonomicznego w Katowicach", nr 152, 2013. [↑](#footnote-ref-24)
24. E. Gantar*, Analiza dyskryminacyjna - stan aktualny i kierunki rozwoju*, ,,Prace Naukowe Uniwersytetu Ekonomicznego w Katowicach", nr 152, 2013. [↑](#footnote-ref-25)
25. ### M. Kasjaniuk, *Zastosowanie analizy dyskryminacyjnej do modelowania i prognozowania kondycji przedsiębiorstw*, „[Barometr Regionalny](http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ekon-element-issn-1644-9398)” 2006, [nr 2(6)](http://yadda.icm.edu.pl/yadda/element/bwmeta1.element.ekon-element-191520f9-c8e0-3c8a-b67a-71c730fed1f0) s, 95-100

    [↑](#footnote-ref-26)
26. J. Pociecha *Problemy prognozowania bankructwa firmy metodą analizy dyskryminacyjnej* „[Acta Universitatis Lodziensis. Folia Oeconomica”, nr 205, 2006.](http://dspace.uni.lodz.pl/xmlui/handle/11089/5938)  [↑](#footnote-ref-27)
27. S.Godlewska, *Skuteczność polskich modeli dyskryminacyjnych w ocenie zagrożenia upadłością spółek giełdowych*,  „Annales Universitatis Mariae Curie-Skłodowska. Sectio H. Oeconomia”, nr  44.2, 2010. [↑](#footnote-ref-28)
28. K. Błonski, O. Brozek, *Handwritten digit recognition using discriminant analysis*, “Faculty of Applied Mathematics, Sielsian University of Technology Gliwice”, 2019. [↑](#footnote-ref-29)
29. <https://www.statystyka.az.pl/centrum-statystyki/test-normalnosci-shapiro-wilka.php>,

    [dostęp: 20.06.2020] [↑](#footnote-ref-30)
30. <http://manuals.pqstat.pl/statpqpl:usepl:arkpl:normstandpl> [24.06.2020] [↑](#footnote-ref-31)