

# High Dimensional Data Analysis - Group Assignment 18

Group 18

08/09/2021

## Contents

Introduction	2
Acknowledgement	3
Data Description	4

## Introduction

The UCLA-LoPucki Bankruptcy Research Database (BRD) is a UCLA School of Law data gathering, data linking, and data distribution initiative. The goal of the BRD is to encourage bankruptcy research by making bankruptcy data available to academic investigators worldwide. All of the data was gathered when the companies declared bankruptcy. In this study, we will use a variety of high-dimensional analysis approaches like Multidimensional Scaling, Principle Component Analysis and Clustering to extract useful insights from the data.



## Acknowledgement

Our sincere gratitude goes out to Ruben Loaiza-Maya and the our tutor Ari Handayani for their guidance and support. Our heartfelt thanks go to Ruben Loaiza-Maya and our instructor Ari Handayani for their advice and assistance. Their culminating efforts have placed us in a situation where we can produce this report collectively while showcasing our competence to employ diverse solutions that can be used with high dimensional data.

## Data Description

This report is based on data from US businesses that declared bankruptcy between 1980 and 2000. The information is coming from the UCLA-LoPucki Bankruptcy Research Database. Let's take a closer look at these variables and what they mean. Let's go further into the dataset to see if we can find any examples of data cleaning or wrangling.

The dataset has 436 observations and 7 variables with their description explained below.

- Name: Name of the firm
- Assets: Total assets (in millions of dollars)
- CityFiled: City where filing took place
- CPI U.S CPI at the time of filing
- DaysIn: Length of bankruptcy process
- DENYOther: CityFiled, categorized as Wilmington (DE), New York (NY) or all other cities (OT)
- Ebit: Earnings (operating income) at time of filing (in millions of dollars)
- Employees: Number of employees before bankruptcy
- EmplUnion: Number of union employees before bankruptcy
- FilingRate: Total number of other bankruptcy filings in the year of this filing
- FirmEnd: Short description of the event that ended the firm's existence
- GDP: Gross Domestic Product for the Quarter in which the case was filed
- HeadCityPop: The population of the firms headquarters city
- HeadCourtCityToDE: The distance in miles from the firms headquarters city to the city in which the case was filed
- HeadStAtFiling: The state in which firms headquarters is located
- Liab: Total amount of money owed (in millions of dollars)
- MonthFiled: Categorical variable where numbers from 1 to 12 correspond to months from Jan to Dec
- PrimeFiling: Prime rate of interest on the bankruptcy filing date
- Sales: Sales before bankruptcy (in dollars)
- SICMajGroup: Standard industrial classification code
- YearFiled: Year bankruptcy was filed