

## 가. 프로젝트 제목

- 재무제표 시각화 라이브러리 financialPlot의 개발과 이를 활용한 기업의 도산가능성 예측

## 나. 프로젝트 목적

재무제표 분석이란, 이해관계에 있는 기업의 재무적 안정성과 수익성을 분석하여 도산가능성이 있는지 판단하는 것이다. (상장된 기업의 경우 미래에 관리종목으로 지정될 가능성이 있는지 판단한다.) 그러나 DART에서 제공하는 재무제표는 테이블과 텍스트로만 이루어져 있고, 주요 평가 지표가 계산 되어있지 않다. 따라서 분석가는 주요 지표를 일일이 계산하고 이를 다시 시각화 하여 분석을 시작해야 하는 단점이 있다.

이러한 문제점을 해결하기 위하여 재무제표 시각화 라이브러리 'financialPlot'을 개발, 배포하였다. 'financialPlot'은 재무제표의 안정성, 수익성, 부실가능성 분석에 자주 사용되는 15개의 차트를 제공한다. 재무정보 데이터를 입력으로 받고, 비율과 같은 지표를 자동으로 계산하여 시각화 한다. 분석가는 시각화 라이브러리에 대한 특별한 공부 없이 사용할 수 있어 재무제표 분석의 효율성을 제고할 수 있을 것으로 기대한다.

또한 'financialPlot'과 분류 모델을 사용하여 실제 기업 데이터를 분석하고 도산가능성이 높은 기업을 예측하였다. 일반적으로 재무제표 데이터는 사람이 분석하는데, 도산 기업의 예측 과정에서 기업에게 가지고 있는 이미지와 편견이 반영된다. 해당 문제를 해결하기 위해, 데이터와 분류모델을 이용하여 분석결과를 도출하였다.

## 다. 주요 내용

: 프로젝트 구현을 위한 기반 지식

### 1) Plotly

Plotly는 2013년 「Plotly」사가 배포한 시각화 라이브러리이다. 기본적인 차트에 Colormap, Annotation, Hover, Spike lines과 같은 옵션을 사용하여 특정 정보를 강조하거나 숨길 수 있다. 또한 Button, Click 기능을 추가하여 다른 라이브러리와 차별화된 인터랙티브 그래프를 구현한다. 구현에 이용한 주요 옵션은 다음과 같다.

함수	옵션	기능
fig.add_trace()	customdata	외부 변수를 끌어와 hovertemplate의 요소로 사용
	hovertemplate	사용자가 마우스를 올렸을 때 표시할 정보 설정
	hoverlabel	Hover의 font family, color, size 설정
fig.update_layout()	showspikes	spikes line 표시 여부 설정
	spikedash	dot, solid 등 spike line style 설정
fig.add_annotation()	text	annotation으로 표시할 text 입력
	x, y	annotation의 위치 설정
go.bar()	animation_frame	애니메이션 효과의 프레임 기준 설정
fig.update_layout()	updatemenus = dict(type = "button")	그래프에 버튼 탑재, 버튼을 클릭했을 때 동작 정의

## 2) Pypi 모듈화

financialPlot을 모듈화하여 Python Package Index에 등록하였다.

[Pypi financialPlot Page] <https://pypi.org/project/financialPlot/>

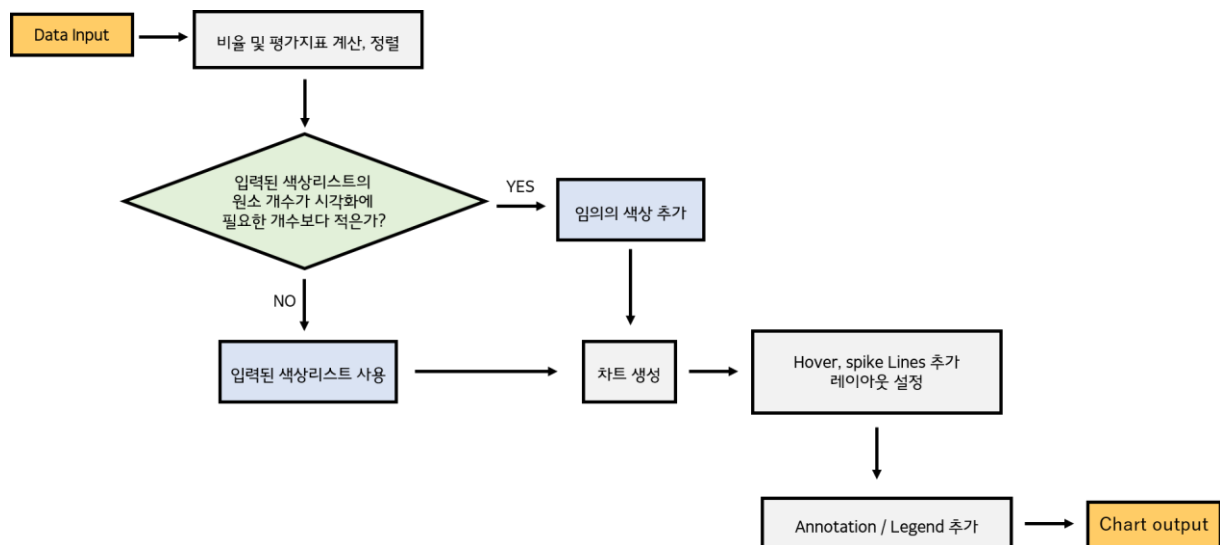
## 라. 개발 환경

- 개발 환경 및 언어: Jupyter Notebook, Python
- 사용 오픈 소스: Pandas, Plotly

## 2 프로젝트 구성도 및 세부기능

### 가. financialPlot의 구성도 및 세부기능

- Workflow



· 세부기능(모듈)

1) 도넛(최상위계정명, 세부계정리스트, 금액리스트, 색상맵)

: 최상위계정에 속하는 각 세부계정의 비율을 자동으로 계산하고 비율이 높은 순으로 정렬하여 도넛차트로 시각화

2) 번버스트\_BS(자산세부계정리스트, 부채세부계정리스트, 자본세부계정리스트, 자산금액리스트, 부채금액리스트, 자본금액리스트, 색상맵)

: "자산, 부채, 자본의 세부계정 비율을 자동으로 계산하고 시각화

3) 번버스트\_IS(매출원가\_판매비와관리비\_영업이익금액리스트, 색상맵)

: 매출액 대비 매출원가, 매출총이익, 판매비와관리비, 영업이익의 비율의 자동으로 계산하여 시각화

4) 비율비교(제목, 기업리스트, 세부계정리스트, 색상맵, 높이, \*\*기업명및금액리스트)

: 기업별 세부계정의 비율을 자동으로 계산하여 비교할 수 있도록 시각화

5) 버블(제목, 기업리스트, x축이름, x축리스트, y축이름, y축리스트, 크기라벨명, 크기리스트, 색상라벨명, 색상리스트, 크기 최대값, x축범위, y축범위, 색상맵)

: 버블 차트 시각화

6) 도넛\_02(비율명, 분모계정명, 분자계정명, 분모계정금액, 분자계정금액, 색상)

: 분자계정금액 / 분모계정금액 비율을 자동으로 계산하여 해당 비율을 시각화

7) 방사형(제목, 비율이름리스트, 색상맵, \*\*기업및비율리스트);

: 기업별 비율 분석 결과를 방사형 그래프로 시각화

8) 막대(제목, 기업리스트, y축이름, 수치리스트, 색상맵, 범례명리스트, 범례색상리스트, 수치선색상)

: x축은 기업, y축은 데이터를 입력하여 기업별 데이터 수치를 막대차트로 시각화, 분위수와 평균을 자동으로 계산하여 옵션으로 제공

9) 막대\_년도별(제목, 항목이름리스트, 년도리스트, 수치단위, 강조색상, 일반색상, 추세선색상, \*\*수치리스트)

: 각 항목의 수치를 년도별로 비교할 수 있는 막대차트를 시각화, 각 년도 중에서 가장 큰 값을 가지는 막대는 강조 색상으로, 나머지는 일반 색상으로 표시

10) 추세선(제목, 항목이름리스트, 년도리스트, y축이름, 색상맵, \*\*항목별수치리스트)

: 년도별, 항목별 수치 추세선을 시각화

11) 증감률(제목, 항목이름리스트, 년도리스트, y축이름, y축단위, 색상맵, \*\*항목별수치리스트)

: 각 항목의 년도별 증감액을 자동으로 계산하여 전년도 대비 증감액을 시각화

12) 스택(제목, 매출원가리스트, 판매비와관리비리스트, 영업이익리스트, 년도리스트, 색상맵)

: 년도별 매출액 대비 매출원가, 판매비와관리비, 영업이익의 구성비율을 그래프 면적을 통해 시각화

13) 막대\_애니메이션(제목, 년도리스트, x축이름, x축리스트, y축이름, y축리스트, 색상맵)

: x축과 y축 데이터를 년도별로 프레임화 하여 애니메이션 형태로 시각화

15) 기업부실예측(제목, 기업리스트, 유동자산리스트, 비유동자산리스트, 자산총계리스트, 유동부채리스트, 비유동부채리스트, 부채총계리스트, 자본금리스트, 이익잉여금리스트, 자본총계리스트, 영업이익리스트, 매출액리스트, 부실가능성낮음\_색상, 판단보류\_색상, 부실가능성높음\_색상)

: 기업의 계정별 데이터를 입력받아 자동으로 Altman Score를 계산하고 부실 기업 예측 결과를 시각화

## · Installation

```
Pip install financialPlot
```

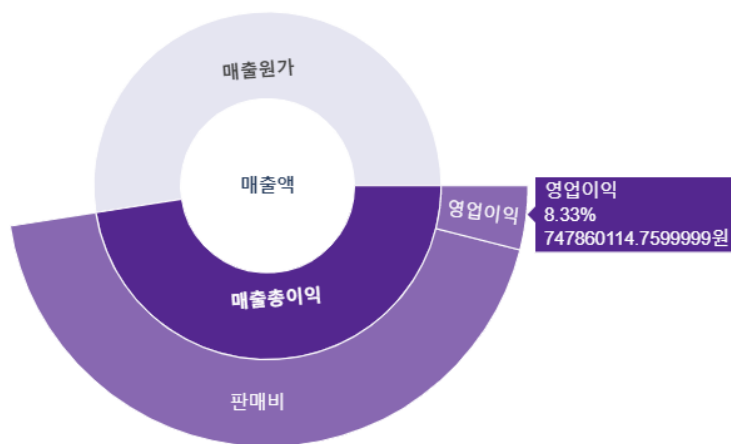
```
from financialPlot import financialPlot
```

## · Using Examples

Input

```
financialPlot.썬버스트_JS(profit_amt, colors)
```

Output



## 나. 데이터 수집 및 전처리

### · 데이터 수집 범위

- 국내 기업의 2015 ~ 2019년 재무제표 데이터
- DART OPEN API를 통해 수집

### · 데이터 전처리

- JSON 형태의 데이터를 DataFrame으로 변환
- 결측치는 DART 사이트를 참고하여 실제 값으로 대체
- KRW으로 화폐 단위 통일

· 데이터 실루엣

기업명	관리종목여부	년도	...	현금및현금성자산의순증가	기초현금및현금성자산	기말현금및현금성자산
케이씨씨	N	2015	...	-153422000000	462184000000	313072000000
케이씨씨	N	2016	...	182507000000	313072000000	507291000000
케이씨씨	N	2017	...	-148018000000	507291000000	347248000000
케이씨씨	N	2018		-95125769085	347248000000	247935000000
케이씨씨	N	2019		127381000000	247935000000	123099000000

## 다. financialPlot을 활용한 탐색적 데이터 분석

### 1) 구성비율 분석

공통재무상태표분석: 기업의 자산, 부채, 자본을 구성하는 세부계정의 비율 분석

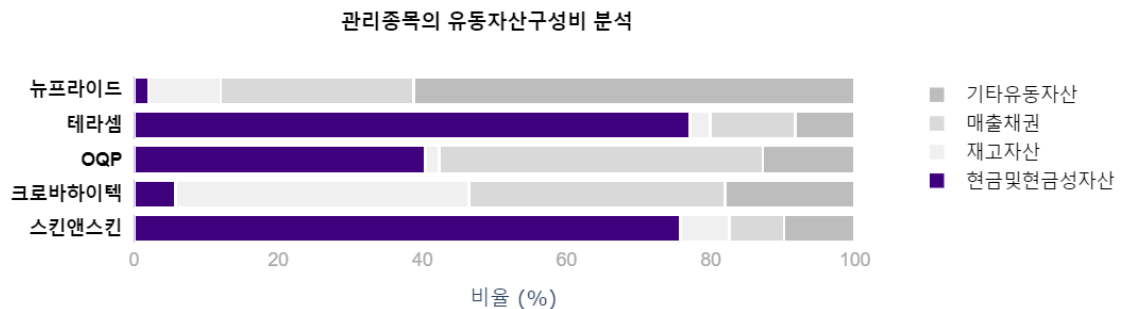
공통형포괄손익계산서분석: 손익계산서 항목 중 매출액을 구성하는 매출원가, 판매비, 영업이익의 비율 분석

financialPlot.비율비교(

"관리종목의 유동자산구성비 분석", man\_corplist, asset\_details, colorlist,

400, 스킨앤스킨 = man\_amtlist[0], 크로바하이텍 = man\_amtlist[1],

OQP = man\_amtlist[2], 테라셀 = man\_amtlist[3], 뉴프라이드 = man\_amtlist[4])

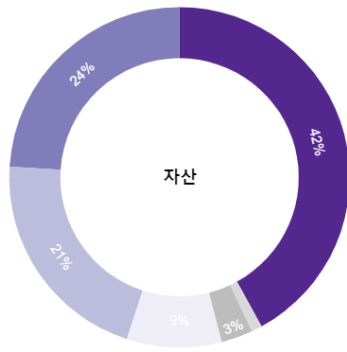


관리 종목의 중 현금및현금성자산의 비율이 지나치게 높거나 지나치게 낮은 사례가 빈번하게 발견되었다. 현금및 현금성자산의 비율이 지나치게 높다는 것은 기업의 영업활동과 투자활동이 제대로 이루어지지 않는다는 것을 뜻한다. 지나치게 현금의 비율이 낮은 경우 단기의 차입금이나 매입채무를 제 때 상환하지 못할 가능성이 있다. 따라서 현금의 자산에 대한 비율이 지나치게 높거나 낮은 재무상태는 주의를 기울여야 한다.

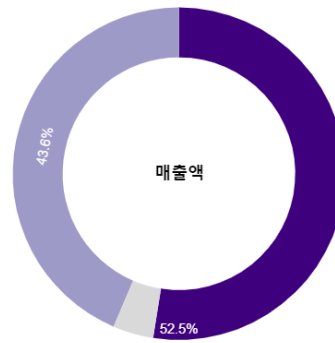
관리종목 중 스킨앤스킨의 자산 구성비와 매출액 구성비를 자세히 살펴보기 위해 도넛 차트를 이용하였다.

financialPlot.도넛("자산", asset\_details, asset\_amt, px.colors.sequential.Purples\_r[1::2])

financialPlot.도넛("매출액", profit\_details, profit\_amt, px.colors.sequential.Purples\_r[::4])



■ 기타비유동자산  
 ■ 유형자산  
 ■ 기타유동자산  
 ■ 매출채권  
 ■ 재고자산  
 ■ 현금및현금성자산  
 ■ 무형자산  
 ■ 투자부동산



■ 매출원가  
 ■ 판매비와관리비  
 ■ 영업이익(순실)

스킨앤스킨의 자산 구성비의 경우 12개월 내에 현금으로 전환하지 못하는 기타비유동자산과 유형자산의 비율이 66%를 차지하였다. 전체 매출액에서 매출원가와 판매비를 제외한 순이익, 영업이익의 비율은 3.9%로 현저히 낮았다.

## 2) 안정성 비율 분석

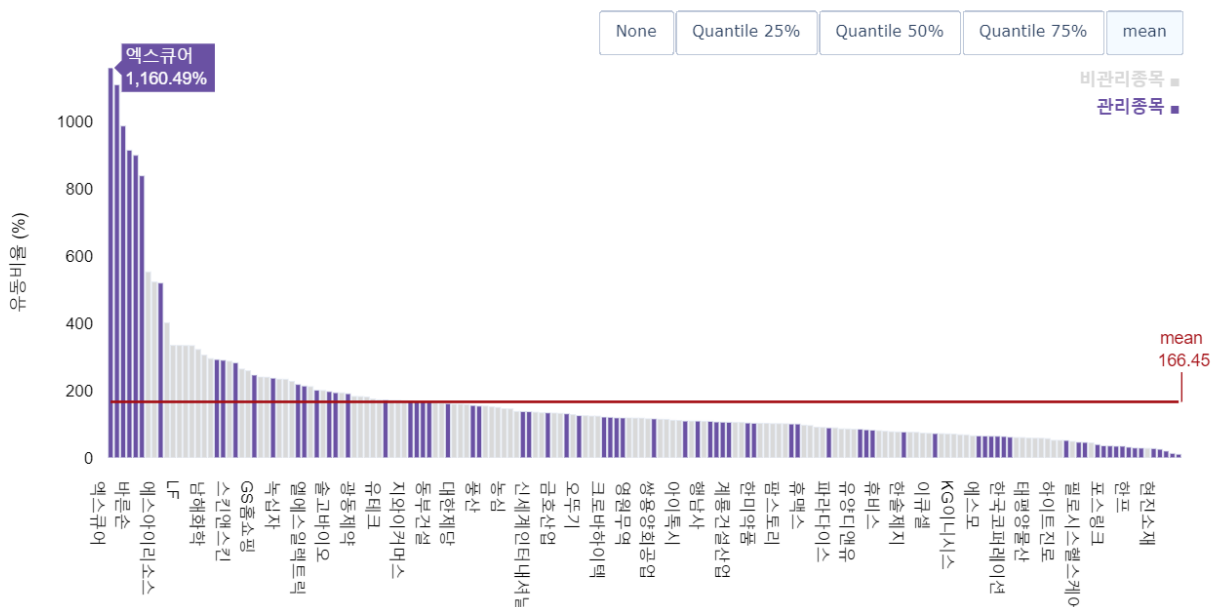
유동비율(유동자산 / 유동부채 \* 100), 당좌비율(유동자산 - 재고자산) / 유동부채 \* 100)

부채비율(부채총계 / 자본총계 \* 100), 유동부채비율(유동부채 / 자기자본 \* 100)

financialPlot.막대(

"유동비율 분석", corp\_list, "유동비율 (%)", ratio\_list, color\_list, legend, legend\_colors, px.colors.sequential.Redr[1])

### 유동비율 분석

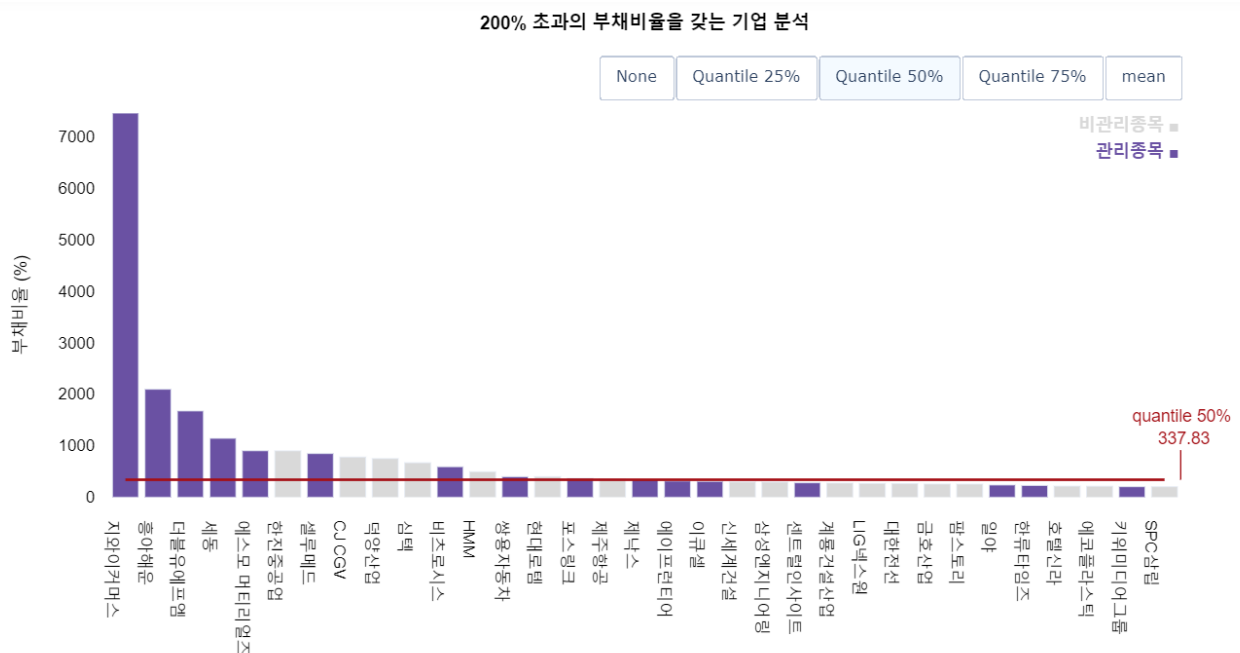


유동비율은 유동자산을 유동부채로 나누어 산출한 값으로, 기업의 단기 채무능력을 보여준다. 일반적으로 유동비율이 높은 기업일수록 상환능력이 좋다는 것을 의미한다. 따라서 겉으로 보기에는 괜찮은 기업일지라도, 유동비율이 동종업계의 평균에 비해 너무 낮다면 재무적 안정성을 의심해볼 필요가 있다. 위의 그래프에서 일부 관리종목

의 유동비율이 지나치게 높은 이유는 유동자산을 영업에 사용하지 않고 묻어두었기 때문이다. 결론적으로 유동비율은 높을 수록 좋지만, 과도한 유동비율은 매우 낮은 수익성을 의미한다.

financialPlot.막대(

"200% 초과 부채비율을 갖는 기업 분석", new\_corplist, "부채비율 (%)", ratio\_list,  
color\_list, legend, legend\_colors, px.colors.sequential.Red1)

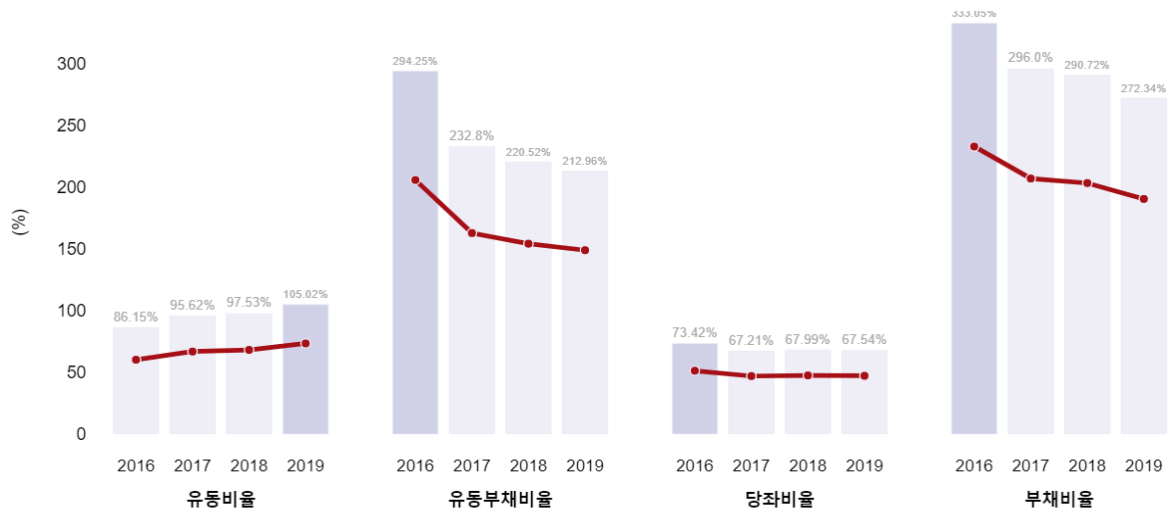


부채비율은 기업의 운영에서 자기자본에 비해 타인자본이 얼마나 개입되었는가를 보여주는 지표로, 200% 수준을 적정 수준으로 본다. 부채비율이 200%가 넘는 기업만을 따로 추출하여 시각화 하였을 때, 관리종목의 기업들이 높은 부채비율을 보인다. 그러나 관리종목으로 지정되지는 않았지만 높은 부채비율을 보이는 기업에 대해 면밀히 살펴볼 필요가 있다. 그 중 계룡건설산업의 부채비율 및 재무적 안정성을 면밀히 살펴보기 위하여 년도 별 안정성 비율을 시각화 하였다.

financialPlot.막대\_년도별(

"계룡건설산업의 년도별 안정성 분석", ["유동비율", "유동부채비율", "당좌비율", "부채비율"], [2016, 2017, 2018, 2019], "(%)",  
"rgba(188,189,220,0.7)", px.colors.sequential.Purples[1], px.colors.sequential.Red1,  
유동비율 = 유동비율리스트, 유동부채비율 = 유동부채비율리스트, 당좌비율 = 당좌비율리스트, 부채비율 = 부채비율리스트)

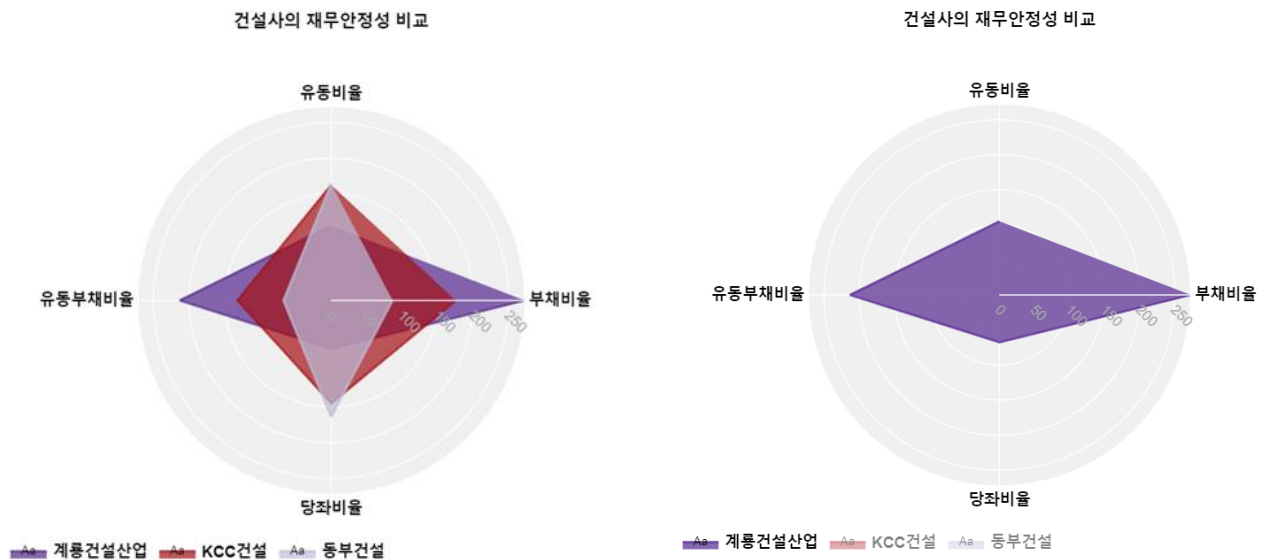
### 계룡건설산업의 연도별 안정성 분석



높은 부채비율로 인하여 재무적 안정성이 우려되나, 시간이 지남에 따라 부채비율이 감소하고 있는 추세를 확인할 수 있다. 그러나 유동비율과 부채비율 모두 적정수준에서 벗어나 있으므로 재무공시가 업데이트 될 때마다 추세를 확인하는 것이 중요할 것으로 보인다.

financialPlot에는 여러 기업의 여러 평가지표를 한번에 비교할 수 있도록 방사형 차트를 제공하고 있다. 동종 건설 업계에 속하는 타 건설사들의 재무적 안정성을 방사형 차트를 통해 시각화 하였다.

```
financialPlot.방사형(
    "건설사의 재무안정성 비교", ["유동비율", "유동부채비율", "당좌비율", "부채비율"], colorlist,
    계룡건설산업 = ratio_list[0], KCC건설 = ratio_list[2], 동부건설 = ratio_list[4])
```





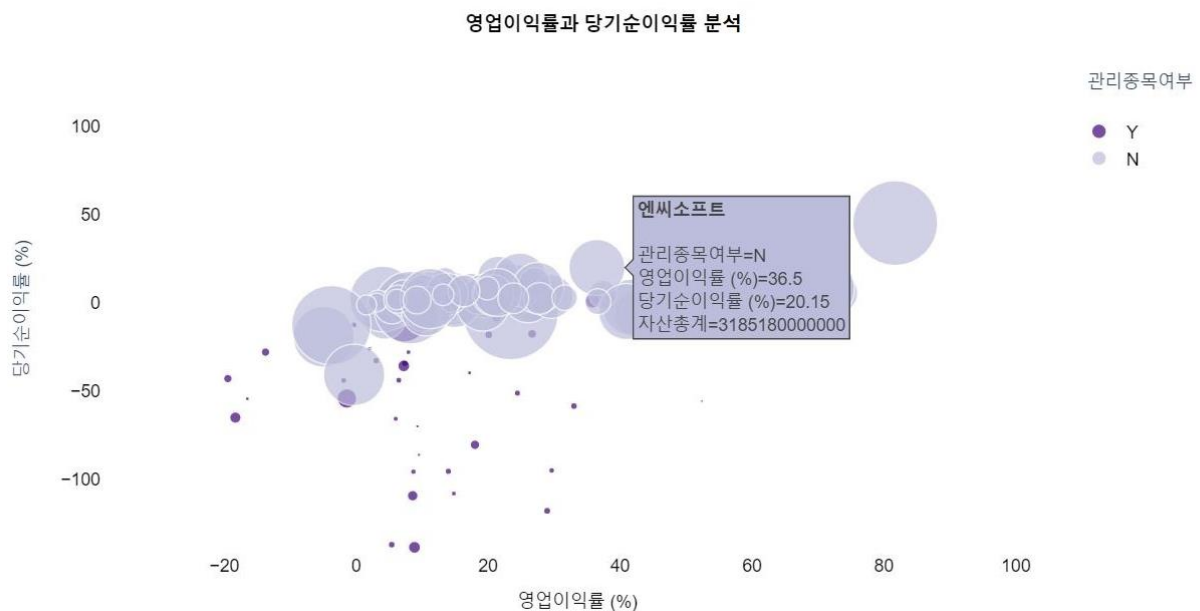
계룡건설산업, KCC건설, 동부건설의 2019년 안정성비율을 종합적으로 분석하였을 때, 동부건설 > KCC 건설 > 계룡건설산업 순으로 재무적 안정성이 높다. 동부건설은 당좌비율과 유동비율은 높고 부채비율과 유동부채비율은 낮는데 반해, 계룡건설산업은 부채비율이 높고 유동비율이 낮다. 만약 재무적 안정성이 높은 건설사 파트너를 검토하는 경우 동부건설이 가장 유력할 것으로 보인다.

### 3) 수익성 비율 분석

매출액 대비 매출총이익, 영업이익, 당기순이익이 차지하는 비율을 분석한다. 더 세분화하여 분석하는 경우, 매출원가, 판매비와관리비, 영업이익이 매출액 대비 차지하는 비율을 분석한다.

매출총이익률(매출총이익/매출액 \* 100), 영업이익률(영업이익/매출액 \* 100), 순이익률(당기순이익/매출액 \* 100)

```
financialPlot.버블(
    "영업이익률과 당기순이익률 분석",
    corp_list,
    "영업이익률 (%)", operating_margin_ratio_list, "당기순이익률 (%)", net_income_ratio_list,
    "자산총계", asset_list, "관리종목여부", man_list,
    50, [min(operating_margin_ratio_list), max(operating_margin_ratio_list)],
    [min(gross_margin_ratio_list), max(gross_margin_ratio_list)],
    [px.colors.sequential.Purples_r[0], px.colors.sequential.Purples[3]]
)
```



전체 데이터의 영업이익률과 순이익률을 시각화 하였을 때, 당기순이익률이 높을수록 영업이익률 또한 높아지는 패턴을 보인다. 다만 관리종목의 경우 원의 크기인 자산의 규모가 비교적 작고, 당기순이익률과 영업이익률 또한 낮은 값을 갖는 것을 확인할 수 있다. 상위의 이익률을 가지는 기업은 신세계, 강원랜드, 엔씨소프트이며, 농심, 오뚜기와 같은 제조업의 경우 중위권에 머물렀다.

도넛\_02 차트를 통하여 상위권의 이익률을 가지는 엔씨소프트와 하위권의 이익률을 가지는 한진중공업의 매출총 이익률을 비교하였다.

```
financialPlot.도넛_02(
  "엔씨소프트_매출총이익률", "매출액", "매출총이익",
  엔씨소프트_매출액, 엔씨소프트_매출액 - 엔씨소프트_매출원가, px.colors.sequential.Purples_r[0])
```

```
financialPlot.도넛_02(
  "한진중공업_매출총이익률", "매출액", "매출총이익",
  한진중공업_매출액, 한진중공업_매출액 - 한진중공업_매출원가, px.colors.sequential.Purples_r[0])
```



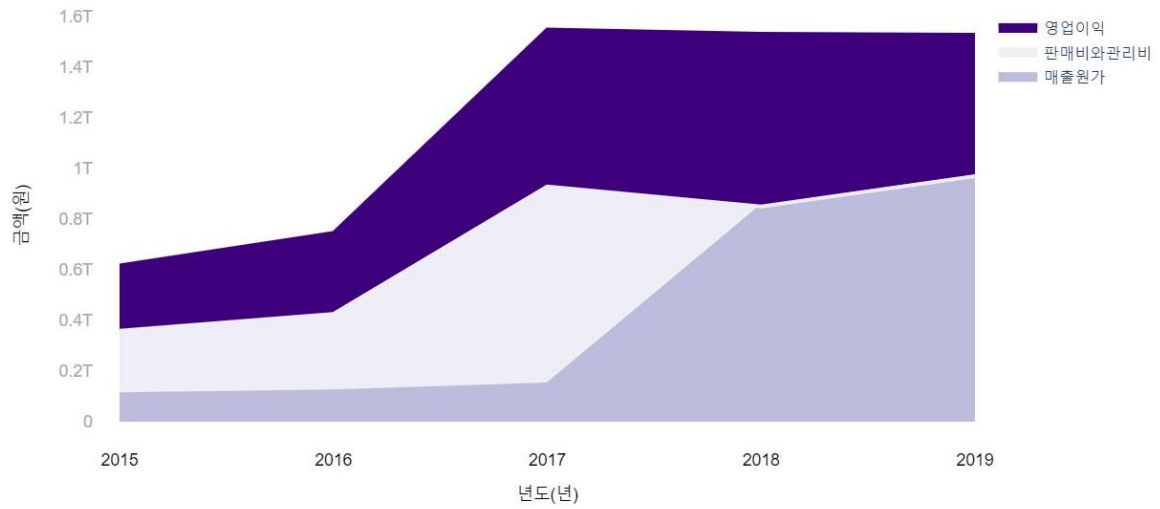
엔씨소프트의 2019년 매출액은 1조 5천만원, 한진중공업의 매출액은 1조 6천만원으로 비슷한 수준이었지만, 매출 원가를 제외한 매출총이익의 비율은 엔씨소프트가 3배 이상 높은 것으로 나타났다.

financialPlot은 버블차트와 도넛\_02차트 이외에도 수익성 분석에 용이한 스택 차트를 제공한다. 스택 차트를 통해 엔씨소프트와 한진중공업의 매출액과 매출액 구성비의 추이를 분석하였다.

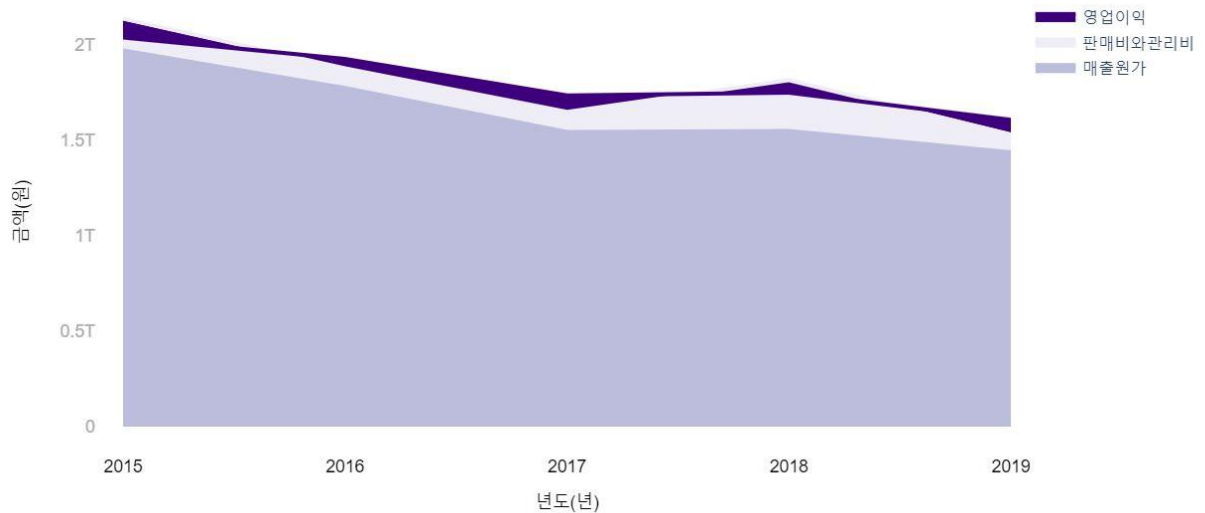
```
financialPlot.스택(
  "엔씨소프트의 매출액 추이와 구성비 분석",
  cost_list, selling_cost_list, operating_margin_list, years, colorlist)
```

```
financialPlot.스택(
  "한진중공업의 매출액 추이와 구성비 분석",
  cost_list, selling_cost_list, operating_margin_list, years, colorlist)
```

엔씨소프트의 매출액 추이와 구성비 분석



한진중공업의 매출액 추이와 구성비 분석



엔씨소프트의 경우 2016년에서 2017년 사이에 매출액이 급격히 증가하였으며, 이에 따라 영업이익의 금액도 크게 늘어난 것으로 보인다. 해마다 엔씨소프트의 매출액의 절반 정도가 영업이익으로 남는 것을 확인할 수 있다. 2018년에 판매비와 관리비가 0이 된 것은 기존에 판매비와 관리비로 계상되던 금액이 매출원가로 편입된 것으로 보인다. 한진중공업의 경우 매출액의 금액은 엔씨소프트와 비슷하지만 꾸준히 감소하고 있다. 또한 매출액의 대부분이 매출원가로 이루어져 실질적 영업이익은 매우 적은 것으로 나타난다.

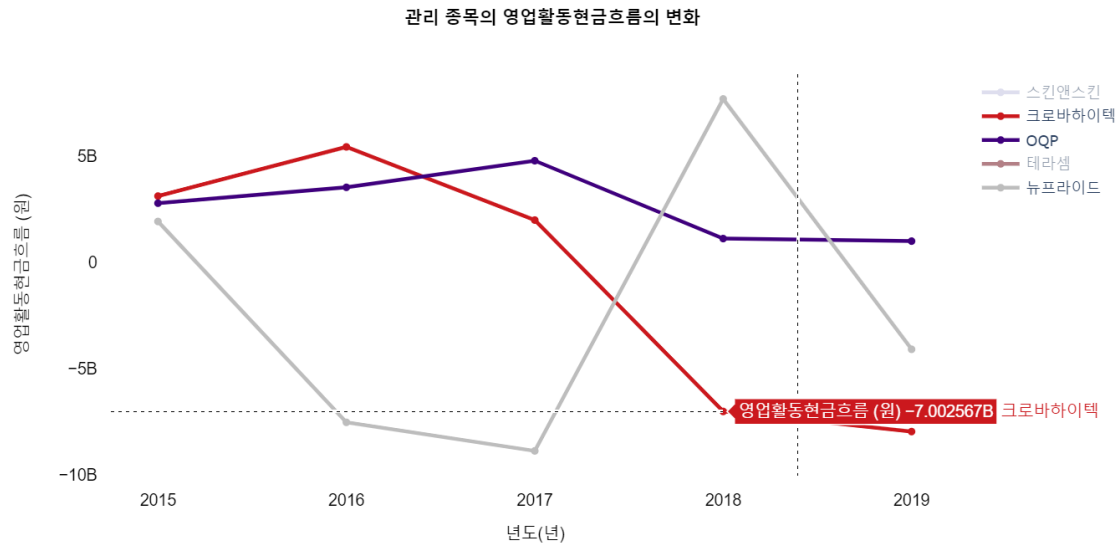
#### 4) 성장성 분석

총 자산, 유동자산, 매출액, 영업이익, 현금흐름 등의 계정이 전기에 비해 얼마나 증감하였는지 분석

관리종목에 속하는 임의의 기업에 대하여 영업현금흐름의 성장성을 분석하였다.

### financialPlot.추세선(

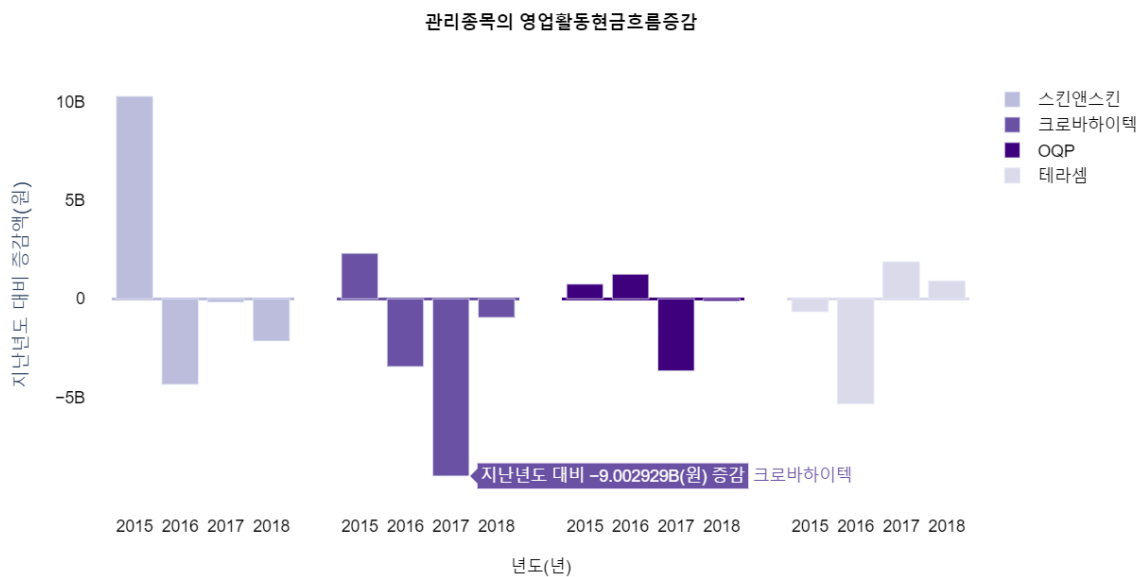
"관리 종목의 영업활동현금흐름의 변화", corp\_list, [2015, 2016, 2017, 2018, 2019], "영업활동현금흐름 (원)", colorlist, 스킨앤스킨 = investment\_list[0], 크로바하이텍 = investment\_list[1], OQP = investment\_list[2], 테라셀 = investment\_list[3], 뉴프라이드 = investment\_list[4])



추세선 그래프를 활용하면 각 년도에 따라 각 기업 혹은 항목이 증가하는지, 감소하는지 쉽게 파악할 수 있다. 그러나 정확히 얼마정도 증가하는가를 한눈에 파악하기 어렵다. 이러한 점을 보완하기 위해 financialPlot의 증감률 차트를 이용하여 지난해 대비 증감 금액을 시각화 하였다.

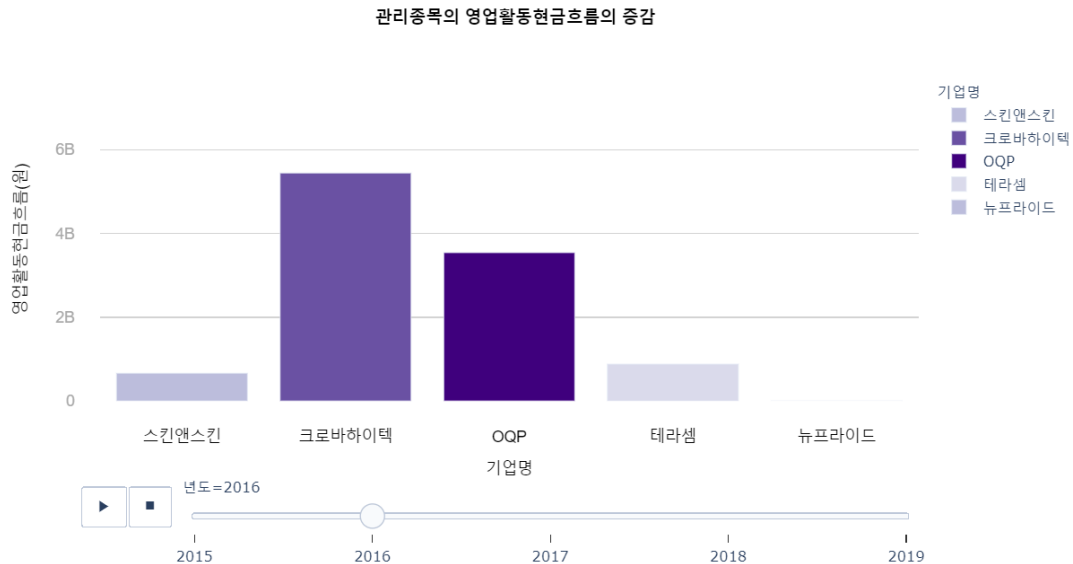
### financialPlot.증감률(

"관리종목의 영업활동현금흐름증감", corp\_list[:-1], list(range(2015, 2020)), "지난년도 대비 증감액", "(원)", colorlist, 스킨앤스킨 = investment\_list[0], 크로바하이텍 = investment\_list[1], OQP = investment\_list[2], 테라셀 = investment\_list[3])



성장성의 높낮이 변화를 파악하기 위해 막대\_애니메이션 차트를 사용하여 시각화 할 수 있다.

```
financialPlot.막대_애니메이션(  
    "관리종목의 영업활동현금흐름의 증감", year_data,  
    "기업명", corp_data, "영업활동현금흐름(원)", value_data, colorlist)
```



플레이 버튼을 누르면, 년도에 따라 프레임이 이동한다. 플레이버튼을 누르지 않고 아래의 년도 바를 움직이면서 각 년도의 영업활동현금흐름 금액을 파악할 수 있다.

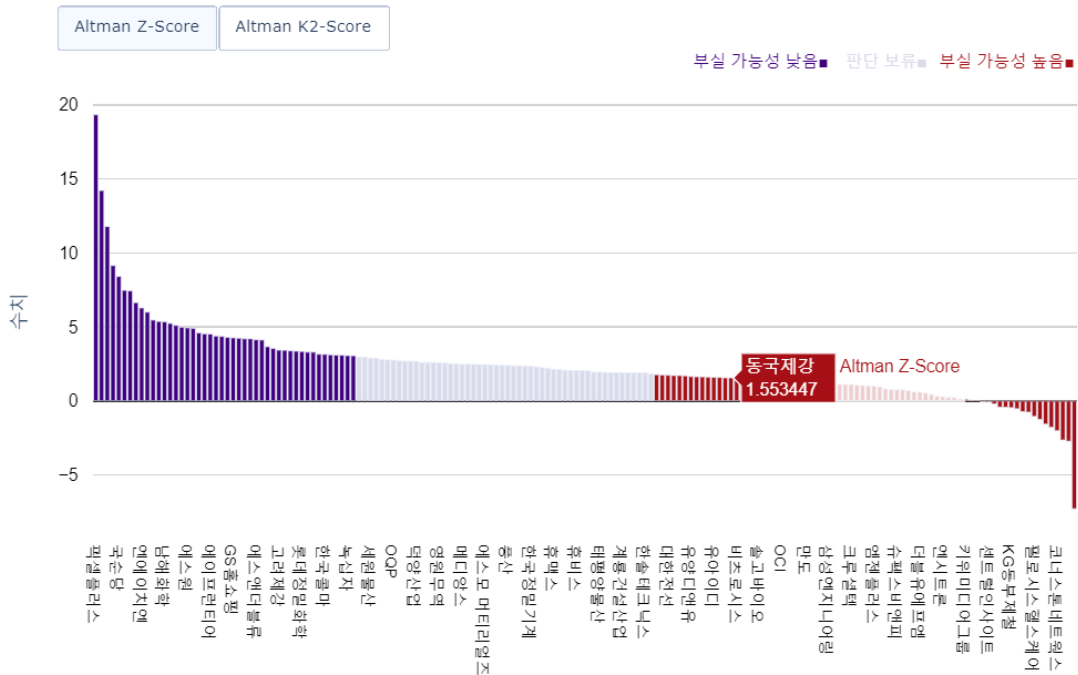
## 5) 기업부실예측

Altman Z-Score, Altman K2-Score 값을 통해 기업의 도산가능성(부실가능성)을 예측

재무적 안정성과 수익성에 따른 Altman Z-Score, Altman K2-Score를 분석하여 부실가능성이 높은 기업을 확인하였다.

```
financialPlot.기업부실예측(  
    "기업 부실 예측",  
    corp_list,  
    유동자산리스트, 비유동자산리스트, 자산총계리스트,  
    유동부채리스트, 비유동부채리스트, 부채총계리스트,  
    자본금리스트, 이익잉여금리스트, 자본총계리스트,  
    영업이익리스트,  
    매출액리스트,  
    px.colors.sequential.Purples_r[0], px.colors.sequential.Purples[2], px.colors.sequential.Reds_r[1])
```

## 기업 부실 예측



보라색으로 분류된 기업은 높은 Z-score을 갖는 기업으로 부실 가능성이 낮은 기업이다. 반면, 빨간색으로 분류된 기업은 낮은 Z-score의 값을 갖는 기업으로, 부실 가능성이 높은 기업이다. 관리종목의 대부분이 '부실가능성 높음'으로 분류된다. 다만, 비관리종목임에도 불구하고 부실가능성이 높은 집단으로 분류된 기업들이 있다. **동국제강**의 경우 대외적으로 언급되는 경영적 위기는 없다. 그러나 Altman Z-score 값이 낮은 점을 고려하여 회계적 오점이 있는지 면밀히 살펴볼 필요가 있다.

### 3 관리종목 분류 모델링

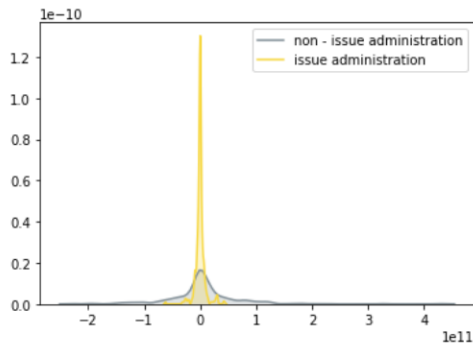
#### 가. 문제 정의

각 기업의 주식 종목을 재무 데이터를 통하여 관리 종목과 비관리종목으로 분류

- Features: 각 기업의 유동자산, 비유동자산, 영업현금흐름, 매출액과 같은 재무제표 당기액
- Target: 각 기업의 관리종목여부

#### 나. 데이터 적합성 확인

Target 변수인 관리종목여부의 각 데이터 개수가 균형을 이루는지 확인하였다. 비관리종목인 "N" 값의 데이터 개수가 130개 더 많아 해당 변수를 다운사이징 해주었다. 아래의 그래프는 '현금및현금성자산'을 비관리종목집단과 관리종목 집단으로 나누어 정규성을 확인하고 Student's t-test를 적용한 예시이다.



현금및현금성자산의순증가  
비관리종목: Normality test: p-value = 0.000  
관리종목: Normality test: p-value = 0.000  
비관리종목: Mean(SD) = 5006879800.09(62144821984.52)  
관리종목: Mean(SD) = 120450303.84(8906920975.31)  
Student's t-test: t stats = -1.497, p-value = 0.135

#### 다. 분류 결과

데이터를 features\_train, features\_test, targets\_train, targets\_test 로 구분하고 교차검증을 거쳐 SVM 모델을 통해 예측한 결과, Bac, Sens, Spec의 정확도는 97% 수준이었다.

CV results  
Bac: Mean(SD) = 0.974(0.025)  
Sens: Mean(SD) = 0.975(0.032)  
Spec: Mean(SD) = 0.973(0.041)

Permutation = 20으로 설정하여 Post-hoc 분석을 수행한 결과, Bac, Sens, Spec의 p-value 값은 0.48로 0.05 이내였다, Feature variable에 따른 계수와 p-value 값은 다음과 같다. 유의미한 수준의 p-value를 가지는 칼럼은 매출채권, 기타비유동자산, 유형자산, 영업이익(손실), 무형자산 및 자본금이었다.

	coefficients	p value							
매출채권	0.78	0.05	재고자산	0.16	0.43	기타유동자산	0.08	0.86	
기타비유동자산	0.74	0.05	자본총계	0.15	0.57	금융수익	0.07	0.81	
유형자산	0.63	0.05	법인세차감전순이익(손실)	0.14	0.76	현금및현금성자산의순증가	0.07	0.90	
영업이익(손실)	0.51	0.05	매출총이익(손실)	0.14	0.81	비유동부채	0.07	0.95	
매출원가	0.49	0.24	유동부채	0.14	0.19	영업활동현금흐름	0.07	0.95	
무형자산	0.45	0.05	자산총계	0.12	0.33	충포괄손익	0.06	0.95	
매출액	0.43	0.33	부채와자본총계	0.12	0.33	부채총계	0.06	0.67	
유동자산	0.32	0.10	판매비와관리비	0.11	0.90	기말현금및현금성자산	0.05	1.00	
자본금	0.32	0.05	당기순이익(손실)	0.08	1.00	투자부동산	0.05	1.00	
이익잉여금	0.27	0.24	기초현금및현금성자산	0.08	1.00	투자활동현금흐름	0.04	0.95	
매입채무	0.25	0.24	기타유동부채	0.08	0.67	현금및현금성자산	0.04	1.00	
법인세비용	0.25	0.24	기타자본구성요소	0.08	0.67	금융비용	0.03	1.00	

#### 라. 대안의 제시

분류 모델 개발의 목적은 미래가 불분명한 기업에 대하여 관리종목지정 가능성을 정확하게 검토하는 것이다. 따라서 분류 모델이 테스트 데이터의 관리종목지정여부를 정확하게 분류하는 것은 중요하다. 그러나, 분류 모델이 정확히 예측하지 못한 테스트 데이터를 자세히 살펴볼 필요성도 있다. 실제로는 관리종목이 아니나 분류 모델이 관리종목으로 예측한 경우, 이는 기업의 재무 데이터가 관리종목으로 예측될 만큼 위험하다는 것을 나타낼 수 있다. 따라서 해당 데이터 분석의 결과로, **사조대림**의 재무적 수익성과 안정성을 자세히 검토할 필요성을 제시한다.

	기업명	년도	관리종목여부	predicted
241	사조대림	2016	N	Y

해당 프로젝트를 통하여 시각화 라이브러리의 레퍼런스를 전체적으로 구현해보는 경험을 얻었다. 또한 분류 모델링에 대한 새로운 지식을 습득하였다. 데이터 분석이 사이클에 따라 어떻게 이루어지고 있는지에 대해서는 익히 알고 있었다. 그러나 confounding variable, Post-hoc analysis와 같은 개념은 새로운 것이었다. 이를 추가하여 보다 완성도 있는 모델링 과정에 대해 학습하였다. 마지막으로, 전공 분야와 관련 있는 회계 지식과 시각화 기술을 이용하여 여러 기업의 재무적 데이터에 대한 가정을 설립하고 확인해 볼 수 있었다.

- Plotly 공식 레퍼런스: <https://plotly.com/python/> 의 모든 세부자료
- 경희대학교 소프트웨어융합학과 데이터시각화(SWCON425) 이원희 교수님 수업자료: Practical Machine Learning for Data Scientists