

In [19]:

```
import os
from IPython.display import display , Markdown
```

한국파생상품학회

In [21]:

```
pdf_file_path = os.listdir('../paper_list/한국파생상품학회/merging_paper/')
len(pdf_file_path) , len(pdf_file_path) == len(set(pdf_file_path))
```

Out[21]:

(247, True)

In [22]:

```
txt_file_path = os.listdir('../paper_list/한국파생상품학회/merging_folder/')
len(txt_file_path) , len(txt_file_path) == len(set(txt_file_path))
```

Out[22]:

(244, True)

In [23]:

```
pdf_file_path[0] , txt_file_path[0]
```

Out[23]:

('09800254.pdf', '09800204.txt')

In [25]:

```
display(Markdown('### **한국파생상품학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.**'))
set([i.split('.')[0].split(' ')[0] for i in pdf_file_path]).difference(\
set([i.split('.')[0].split(' ')[0] for i in txt_file_path]))
```

한국파생상품학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.

Out[25]:

{'09800427', '09800436'}

한국재무관리학회

In [26]:

```
pdf_file_path = os.listdir('../paper_list/한국재무관리학회/merging_paper/')
len(pdf_file_path) , len(pdf_file_path) == len(set(pdf_file_path))
```

Out[26]:

(386, True)

In [27]:

```
txt_file_path = os.listdir('../paper_list/한국재무관리학회/merging_folder/')
len(txt_file_path), len(txt_file_path) == len(set(txt_file_path))
```

Out[27]:

(387, True)

In [28]:

```
pdf_file_path[0] , txt_file_path[0]
```

Out[28]:

('KCI_FI001155882.pdf', 'KCI_FI001233285.txt')

In [30]:

```
display(Markdown('### **한국재무관리학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.**'))
set([i.split('.')[0].split(' ')[0] for i in txt_file_path]).difference(\
set([i.split('.')[0].split(' ')[0] for i in pdf_file_path]))
```

한국재무관리학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.

Out[30]:

```
{'KCI_FI002331159'}
```

한국재무학회

In [32]:

```
pdf_file_path = os.listdir('../paper_list/한국재무학회/merging_paper/')
len(pdf_file_path) , len(pdf_file_path) == len(set(pdf_file_path))
```

Out[32]:

```
(266, True)
```

In [33]:

```
txt_file_path = os.listdir('../paper_list/한국재무학회/merging_folder/')
len(txt_file_path), len(txt_file_path) == len(set(txt_file_path))
```

Out[33]:

```
(258, True)
```

In [34]:

```
pdf_file_path[0] , txt_file_path[0]
```

Out[34]:

```
('vol21_no3_41.pdf', 'vol25_no3_357.txt')
```

In [35]:

```
display(Markdown('### **한국재무학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.**'))
set([i.split('.')[0].split(' ')[0] for i in pdf_file_path]).difference(\
set([i.split('.')[0].split(' ')[0] for i in txt_file_path]))
```

한국재무학회에서 text 파일로 parsing 이 되지 않은 파일이다. 따로 시도해본 결과에도 인코딩 문제로 에러가 발생했다.

Out[35]:

```
{'vol18_no2_139',
 'vol18_no2_185',
 'vol18_no2_209',
 'vol18_no2_239',
 'vol18_no2_263',
 'vol19_no2_105',
 'vol19_no2_155',
 'vol19_no2_189'}
```

한국증권학회지

In [37]:

```
pdf_file_path = os.listdir('../paper_list/한국증권학회지/merging_paper/')
len(pdf_file_path) , len(pdf_file_path) == len(set(pdf_file_path))
```

Out[37]:

```
(486, True)
```

In [38]:

```
txt_file_path = os.listdir('../paper_list/한국증권학회지/merging_folder/')
len(txt_file_path), len(txt_file_path) == len(set(txt_file_path))
```

Out[38]:

(477, True)

In [39]:

```
pdf_file_path[0] , txt_file_path[0]
```

Out[39]:

```
('Conditional_VaR_모형을_사용한_최적자산배분에_관한_연구.pdf',
 '자료시차를_반영한_Continuous_Time_CAPM의_실증에_관한...txt')
```

In [44]:

```
display(Markdown('### **교수님께서 주신 adobe reader 와 웹 상의 converting 서비스를 모두 적용해본 결과,**'))
display(Markdown('### **한국증권학회지에서 글꼴 문제나 권한 문제로 인해, 문제가 가장 많이 발생하였다..**'))
set([i.split('.')[0].split(' ')[0] for i in pdf_file_path]).difference(\
set([i.split('.')[0].split(' ')[0] for i in txt_file_path]))
```

교수님께서 주신 adobe reader 와 웹 상의 converting 서비스를 모두 적용해본 결과,

한국증권학회지에서 글꼴 문제나 권한 문제로 인해, 문제가 가장 많이 발생하였다..

Out[44]:

```
{'Causes_of_Savings_Bank_Insolvency_A_Ko...',
 'ELW_시장에서_요구되는_위험프리미엄이란',
 'ELW_시장에서의_유동성공급자_간_차이',
 'ELW_시장을_통해_살펴_본_시장규제에_대한_재고',
 'ELW_시장의_투자자_인식과_투자행태_분석',
 'ELW시장의_투자자_매매패턴_분석',
 'IPO_락업_제도가_저가발행현상에_미치는_영향',
 'IPO_신디케이트_구성이_인수주선_서비스에_미치는_영향',
 'IPO_주식의_시장조성제도_부활이후_주간사회사의_공모가_산정행태',
 'KOSDAQ_시장에서_히든챔피언_종목은_과연_히든챔피언인가',
 'KOSPI200_옵션시장에서의_변동성지수_산출_및_분석',
 'KOSPI200_옵션의_일중_변동성합수',
 'KOSPI_200_선물_및_옵션의_만기일_효과',
 'KOSPI_200_지수옵션시장에서_모멘텀_기대의_영향',
 'KOSPI_200_지수의_상태_가격_밀도_비교_분석',
 '국내_주식시장에서_발생액_및_투자_이상현상과_수익률_횡단면변동성에_대...',
 '국내_주식시장의_부도위험_이례현상에_관한_연구',
 '국내_증권사에서_예금보호한도와_시장규율에_관한_연구',
 '국면전환_GARCH_모형을_이용한_변동성_구조_분석_및_예측에_관한...',
 '금융위기하에서_포트폴리오_위험척도의_비교',
 '내국인과_외국인의_주식투자_누적손익_요인_분석',
 '단일변량모형과_다변량모형의_포트폴리오_VaR_측정_성과',
 '데이트레이더의_성과지속성과_시장효율성',
 '비자발적_애널리스트_보고서가_지닌_정보로서의_가치',
 '상장지수펀드(ETF)_차익거래전략',
 '신규_분석기업에_대한_재무분석가의_이익예측특성',
 '신용평가등급과_발생액_및_보수주의_이익_인식과의_관련성에_대한_실증적...',
 '월별_포트폴리오_VaR_측정_성과',
 '저변동성_이상현상과_투자전략의_수익성_검증',
 '주식_대량보유의_마찰효과',
 '주식형_펀드의_해지와_부의_이전',
 '한국_상장기업의_자본구조에_관한_재고찰',
 '한국_제조기업의_환노출과_파생상품을_이용한_헤지',
 '한국_주식시장에서_3요인_모형을_이용한_주식수익률의_고유변동성과_기대...',
 '한국_주식시장의_투자주체별_거래행태에_관한_분석',
 '한국_증권시장의_투자자_유형에_따른_성과분석',
 '한국_채권시장의_투명성_제고정책과_유동성에_관하여'}
```