|  |
| --- |
| 미국 10년 만기 채권 수익률과  주요 증시와의 상관관계분석 |
|  |

|  |  |
| --- | --- |
|  |  |
| **작성자** | 김광수  빅데이터 UI 전문가 양성(파이썬,R)  2021-04-08 목요일 |
|  |  |

목 차

|  |  |
| --- | --- |
| 들어가기 앞서 | **3** |
|  |  |
| 1 서론 | **4** |
|  |  |
| 2 연구방법 | **6** |
|  | |
| 3 분석 준비 | **7** |
|  | |
| 4 분석 | |
| 1. 나스닥과 미국 10년만기 국채 수익률 | **9** |
| 2. S&P500과 미국 10년만기 국채 수익률 | **12** |
| 3. KOSPI 200과 미국 10년만기 국채 수익률 | **15** |
|  | |
| 5 분석 외 시각화 자료 | **18** |
|  | |
| 6 분석 결론 | **22** |
|  |  |
| 7 분석 한계 | **23** |
|  |  |
| 8 참조 | **24** |

|  |
| --- |
|  |
|  |
|  |

들어가기 앞서

2020년 코로나로 인해 많은 것 들이 바뀌게 된 것 같습니다. 본인 또한 코로나로 피해도 봤지만, 그 덕분에 주식이라는 세계에 처음 발을 담가보게 되었습니다. 2020년 폭풍과도 같은 주식시장을 겪고 주식과 경제에 새로운 재미와 느낀 바도 크고, 전국적인 대유행을 했다고 생각합니다. 그러다 보니 프로젝트 주제 선정에 큰 영향을 주었습니다. 경제와 주식은 긴밀한 연관이 있고, 그 안에 세밀한 내용들 중 최근 문제가 되었던 미국 10년만기 채권 금리 와 주요증시와의 관계에 대해 분석하기로 하였습니다. 수많은 경제지표들 중 한가지 만으로 증시를 정확히 분석을 한다는 것은 불가능하다고 생각합니다. 다만, 미 10년만기 국채가 증시에 미쳐왔던 영향을 확인하는 것은 가능하다고 생각합니다. 모쪼록 부족한 내용이지만, 잘 봐주시면 감사하겠습니다.

1 서론

앞서 말했던 것 과 같이 미 국채 10년 만기 금리와 주요 증시(Nasdaq, S&P500, KOSPI)와 의 상관관계에 대해 분석해보고 21년 이후 투자포지션의 선택에 도움을 줌에 목적이 있다. 추가로 원,달러환율과 KOSPI의 상관관계 분석을 하였다. 주요 증시에 있어 Nasdaq을 채택한 이유는 세계적으로 큰 미국증시에서도 금리의 영향을 가장 크게 받는 종목들이 많이 포함되어 있기 때문에 채택하게 되었다. S&P 500의 채택 이유는 뉴욕증권거래소에 상장된 기업 중 시가총액 상위 500위의 기업들 전체를 아우르는 종합지수이기 때문에 채택하였다. KOSPI의 채택이유는 미 국채 10년 금리의 변화가 한국증시에 어떤 영향을 미치는지 확인하고자 채택하게되었다.

본론에 들어가기에 앞서 기본배경이 되는 내용을 숙지해야 어느정도 데이터간 연관성에 대한 이해가 가능함에 따라 기본개념을 먼저 간단하게 서술하겠다. 먼저 주식과 국채의 차이점을 설명하겠다. 주식은 기본적으로 개인의 의사에 따라 자유롭게 거래가 가능하다. 그렇다 보니 매번 주식이 오른다는 보장은 없고 떨어질 수 있다는 큰 리스크가 있다. 반면 국채의 경우 발행자가 국가인 채권이다. 한 국가에서 제일 안전한 재산 취급을 받다 보니 국채수익률을 무위험이자율 이라고 부르기도 한다. 다만, 국가가 발행인이라고 모두 무위험인 것은 아니다. 그러나 미국에서 발행한 국채라면 어떻게 될까? 세계의 경제 대국에서 발행하는 채권은 말 그대로 무위험에 가까운 자산이 된다. 그리고 국채의 만기전까지 거래가 가능하다. 다만, 채권수익률은 주식투자 수익률에 비해 낮은 수준에 속한다는 단점 아닌 단점이 있다. 그래도 굉장히 안전한 수익률이 보장되며 어느정도 예측가능 하여 리스크 대비가 가능하기 때문에 거대자본을 가진 주체들이 선호하는 측면이 있다. 그런 이유로 세계경제가 위기에 처하면 큰 자본을 가진 투자자들이 리스크를 최대한 줄이기 위해 여러 군대로 분산한 자금을 미국 국채 시장으로 옮기게 된다. 이 과정에서 외화들이 세계 주요 증시에서 빠져나가기 때문에 증시에 악영향을 미친다. 이런 상황이 오면 채권거래가 활발해진다. 이 말은 채권의 가격은 올라간다는 이야기이다. 그래서 세계경제가 좋지 않을 때 미 국채 수익률은 내려가게 된다. 반대로 세계 경제가 회복기나 호황기에 접어들어가면 미 국채시장에 몰려 있던 자본들이 세계 각국의 투자처로 빠져나가게 된다. 이는 낮은 수익률보단 경제의 상승에 힘입어 국채보다 리스크는 크지만 투자 대비 수익률을 극대화할 수 있는 방향으로 주식시장으로 이동한다는 것을 알 수 있다. 이렇게 되면 채권의 가격이 내려가 경제가 회복되는 상황이 오거나 예상이 되면 미 국채 수익률은 상승하게 된다. 이것이 단순한 미국 국채금리 하락, 상승 요인이다. 내용만 보면 주식시장과, 국채수익률이 서로 양의 상관관계에 있다고 생각할 수 있는데, 자산시장이 그렇게 단순하지 않다. 이유는 투자자들이 항상 주식시장 지수 배당수익률과 국채수익률을 비교하고 있기 때문이다. 따라서 투자자들이 미 국채수익률의 매력이 주식시장의 리스크보다 좋아 보이게 된다면 아무리 경제가 좋다 하더라도 무조건 적으로 자본이 주식시장으로 가는 것만은 아니라는 것이다. 경제의 회복은 투자의 과열로 향하기 마련이다. 그렇다 보니 저렴해진 채권이 더 매력적으로 느껴지면 주식시장에서 돈이 빠져 나간다. 그래서 국채수익률 상승은 주식시장에서 악재라고 생각되는게 일반적 이였다. 주식과 채권은 역사적으로 이러한 관계에 있었는데 이번 분석의 주된 목적은 ‘코로나 이후 증시에서의 금리상승이 증시에 미치는 영향은 어떠한가’ 이다. 왜 이시기가 중요한가 하면 일단 20년도의 증시 상태를 어느정도 파악해야 한다..

일반적으로 경기의 하락의 원인은 새로운 성장의 모멘텀이 없다 보니 장기적으로 내려가는 추세를 잡거나 특정 사건이 발생하여 영향을 미치는 경우가 있다. 전자의 경우 전세계적으로 4차산업 5차산업 이런 방식으로

경제 성장을 이뤄나가는 것이 해결책인 반면, 후자는 갑작스럽게 외압으로 경제가 무너지게 되는 측면이 있어서 대응방법이 달라지게 된다. 이번 코로나의 경우 급격한 속도로 전염되고 치사율도 높았기 때문에 세계경제

에 큰 타격을 주었다. 그렇게 됨 에 따라 각 나라들은 당장의 나라의 빚보다는 경제를 살려야 했기 때문에 내수경제 진작과 수출의 증가를 위해 경기부양정책을 실행하게 된다. 이 때 미국은 채권을 발행하여 급한 돈을 끌어들여 내수에 공급하게 된다 그와 더불어 기준금리를 거의 0에 가까운 수준으로 낮춰버리게 되었다. 그러하다 보니 20년도 미채권금리는 거의 제로 금리에 가까울 정도로 낮아 졌고, 기업들 또한 제로 금리로 자금 조달 비용이 줄어들고 20년도 3월 이후 엄청난 현금부양정책을 지속적으로 추진함으로 시장에 현금 유동성이 크게 된 바, 그리하여 주식시장은 코로나 이전보다 더 큰 성장을 하게 된다. 문제는 일반적으로 나라에서 돈을 푸는 이유가 시민들이 그 돈을 쓰고 그에 따라 실물경제가 살아나서 그 실물경제를 기반으로 경제가 회복세로 돌아서게 만드는 것에 있는데, 이번은 시장에 풀린 돈이 넘쳐흐르고 코로나 라는 특수성에 의해서 돈을 소비할 곳이 줄어들다 보니 노는 돈들이 주식시장으로 흘러가면서 주식시장이 성장하게 되었다는 것이다. 이로 인해 기업의 성장이 주식시장의 성장을 이끈게 아닌, 단지 현금의 유동성에 의하여 주식시장이 올랐기 때문에 현금의 유동성이 현재 주식시장에 가장 중요한 핵심이 되었다. 그 후 21년도로 넘어와 미국은 더 많은 경기부양책을 사용해 더 강하게 경기부양을 하겠다는 정책을 펼치는데 이번에도 채권은 발행될 것이다. 기존 채권의 수익률이 굉장히 낮은 상황에서 시중금리와 경기상승에 맞게 발행될 채권의 수익률이 더 높아질 가능성이 있고 그와 달리 기준금리는 올리지 않은 현재 상황에서 주식보단 채권이 더 매력적으로 부각 될 가능성이 높아지고 있는데 더불어 국채금리도 상승추세라 증시의 하락이 예견되고 있다. 투자자들은 FED 가 다시 채권을 매입해줌으로 채권수익률의 상승을 막아 줄 것인가에 관해 주시하는데 달러를 발행하는 FED입장에서 추가적인 부양정책을 하게 된다면 달러 인덱스의 인하 그리고 실물경제와 물가상승의 괴리에서 인플레이션의 발생을 대비해야하는 측면이 부각되고 있어 금리상승을 막지 못할 것 같다는 불안감이 커졌다. 앞서 말했는데 경기가 회복되면 국채금리가 올라가기 마련인데, 현재는 경제가 회복되는 것은 확실하나 아직 가시적으로 경기회복이된게 별로 없고, 코로나 또한 종식되지 못했다는 것이 문제다. 거기에다 말도 안되는 단위의 금액을 발행하다 보니 인플레이션 발생에 대한 우려가 심각해진 점. 그렇다 보니 주식 같은 위험자산보다 국채 같은 안전자산이 매력이 더 높아지는 것도 증시엔 악재로 받아들여진다. 이는 증시의 하락을 야기할 것이라는 게 현재 화두로 떠오르고 있다. 결국엔 현재 경기회복에 대한 불확실성으로 인해 국채금리는 상승되고 증시는 하락하는 상황이 발생되고있다. 또한 국채금리 상승은 기준금리 상승에도 영향을 주기 때문에 기업들의 투자이익에도 손실을 줄 수 밖에 없는 구조로 주식시장은 굉장히 흔들리고 있다.

이러한 배경을 바탕으로 국채금리 중에서도 기준금리에 영향을 미치는 10년만기 미국 국채금리와 주요증시간의 상관관계를 분석하고자 한다.

2 분석방법

분석에는 프로그래밍 언어 Python을 사용하였다.

1. 그래프시각화

Python의 API pandas를 통해 분석 데이터를 DataFrame화 하였고 Matplotlib과 seaborn을 사용하여

시각화 하였다. 시각화는 기본적으로 선 그래프를 사용하였고, 각 데이터 들 간 시계열순으로 추세를 파악함에 사용하였다.

2. 상관분석

상관분석은 확률론과 통계학에서 두 변수간 에 어떤 선형적 관계를 갖고 있는지를 분석하는 방법이다. 두 변수는 서로 독립적인 관계로부터 서로 상관된 관계일 수 있으며 이때 두 변수간의 관계의 강도를 상관관계라고 한다. 상관관계의 정도를 파악하는 상관계수는 두 변수간의 연관된 정도를 나타낼 뿐 인관관계를 설명하지 않는다. 상관 계수의 값의 범위는 -1부터 +1까지이며 계수의 절대값이 클수록 변수 사이에 강한 관계가 있다. 이번 분석에선 Pearson 상관계수를 사용하였다. Pearson 상관에서 절대값 1은 완전 선형관계를 나타내고, 0에 가까운 상관 값은 변수 사이에 선형관계가 없음을 나타낸다. pandas 패키지 내에 corr()함수를 사용하였다.

3. 회귀분석

통계학에서, 선형회귀는 종속변수 y 와 한 개 이상의 독립 변수 X와의 선형 상관관계를 모델링하는 회귀분석 기법이다. 한 개의 독립변수에 기반한 경우에는 단순 선형회귀, 둘 이상의 독립변수에 기반한 경우에는 다중 선형회귀라고 한다. 그리고 데이터들의 경향을 알고 싶을 때 최적의 추세선을 그리고 싶을 때 사용하는 방법이 최소자승법이라고 한다. 최소자승법은 잔차의 제곱의 합을 최소로 하는 방법이다. Statsmodels 패키지내에 포함된 OLS를 사용하였다. 최소자승법을 활용한 회귀선 Fitting 으로 선형 회귀식은 y = α + βX (y=종속 변수, α=y절편(intercept), β=회귀계수) 이고 , OLS는 최소자승법으로 회귀 모델을 구하는 메소드이다.

3 분석준비

역사적으로 지켜본 봐 수치상 미국 10년 만기 국채수익률 상승 시 주식시장이 떨어진 적이 더 많았던 부분이 많았다.

그 기조에는 변화가 없을 것으로 생각한다. 이를 확인해 보고자 분석에 사용한 데이터는 다음과 같다

분석에 사용된 데이터

|  |  |  |
| --- | --- | --- |
| 변수명 | 데이터 내용 | 개수 |
| Nasdaq | 2002.01.02~2021.04.01 Nasdaq 일별 웹 크롤링 데이터 | 4846개 |
| Snp500 | 2002.01.02~2021.04.01 S&P 500 일별 웹 크롤링 데이터 | 4849개 |
| Ndq\_70 | 1970.01.02~2001.12.31 Nasdaq 일별 CSV 데이터 | 8081개 |
| Sxp\_70 | 1970.01.02~2001.12.31 S&P500 일별 CSV 데이터 | 8082개 |
| kospi | 1980.01.04 ~ 2021.04.01 KOSPI 200 일별 데이터 | 11027개 |
| Ten\_usy | 1970.01.02 ~ 2021.03.26 US 10 Year Treasury Note 일별 데이터 | 12720개 |
| Usd\_krw | 1981.04.01 ~ 2021.03.26 원 달러 환율 일별 데이터 | 10052개 |

Nasdaq Composite과 S&P500 데이터의 경우 2002.01.02 ~ 2021.04.01의 데이터는 네이버 웹사이트

증권 페이지에서 selenium을 사용하여 웹 크롤링하여 list로 저장한 후 , pandas를 통해 dataframe으로 변환 하였다. 네이버 증권에 없던 이전 데이터들 (1970.01.02~2001.12.31) 의 데이터 들은 CSV파일로 pandas의 read\_csv() 함수로 dataframe으로 불러들였다. 이후 두 데이터를 Concat하여 하나의 데이터로 정제하였다.

그 외 데이터는 모두 CSV파일을 pandas의 read\_csv()함수를 통해 datafame으로 불러들였다. 해당 데이터들의 Columns는 모두 동일했다.

정제 전 데이터 Columns

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Columns | Date | Open | High | Low | Close |
| Dtype | Datetime64 | Float64 | Float64 | Float64 | Float64 |

위의 순서 그대로 Columns였고, Date는 측정한 날짜를 의미한다. Open은 장 개장 시점의 지수, 금리 수치를나타낸다. High는 해달 날짜에 고점인 지수, 금리 수치를 의미한다.Low는 해당 날짜에 저점인 지수, 금리 수치를 의미한다. Close는 해당 날짜의 장 마감 시점의 지수, 금리 수치를 의미한다.

Date Columns 같은 경우 CSV파일은 불러올 당시object Dtype을 Datetime64 Dtype으로 변환하였고, 웹 크롤링한 개채는 Dataframe으로 변환 후 replace와 pandas의 to\_datetime() 함수를 사용하여 object Dtype을 datetime64 Dtype으로 변환하였다.

분석을 위해 필요한 Columns 는 Date 와 Close 였기 때문에 나머지 Columns는 분석 시작 전 제거하였다.

이후 지수 변동률 비교를 위해 per Columns를 Nasdaq Composite , S&P500 , KOSPI 200 , US 10 Year Treasury Note 에 추가하였다. 계산식은 US 10 Year Treasury Note 제외, 해당 날짜의 Close 에 전날의 Close를 차감하여 해당 날짜의 Close를 나눈 후 100을 곱하여 계산하였다. US 10 Year Treasury Note 의

Columns에 있는 데이터들이 이미 수익률을 나타냄으로 해당 날짜의 Close에서 전날의 Close를 차감하여 계산하였다. 원, 달러 환율의 경우 per Columns를 추가하지 않았다.

정제 후 데이터 Columns

|  |  |  |  |
| --- | --- | --- | --- |
| Columns | Date | Close | Per |
| Dtype | Datetime64 | Float64 | Float64 |

기본적으로 Matplotlib이나 Seaborn을 사용한 시각화에 이용된 데이터는 위의 데이터와 같다.

lineplot만을 시각화 하여 시계열에 따른 추세에 관해 분석함에 있어 부족하지 않았다.

이후 상관관계 분석과 OLS회귀분석에 필요한 데이터를 마련하기 위해 데이터의 재편성이 필요하였다.

그에 따라 데이터들을 Pandas 의 Merge함수를 사용하여 Date를 기준으로 inner merge 하였다.

상관관계분석에 사용한 데이터

|  |  |  |  |
| --- | --- | --- | --- |
| 변수명 | Merge한 데이터 | | 개수 |
| df | Ndq | Ten\_usy | 12718개 |
| df2 | Spx | 12719개 |
| df3 | kospi | 9761개 |

df의 Columns

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 변수명 | Columns | | | | |
| df | Date | Ndq\_C | Ndc\_p | Us10y\_c | Us10y\_p |
| df2 | Spx\_C | Spx\_p |
| df3 | Kospi\_c | Kospi\_p |

Ndq\_c 는 ndq의 Close 값이며, Ndq\_p는 ndq의 per값이다. 이하 다른 데이터도 같은 방법으로 변수명을 지정해 주었다.

df, df2, df3 를 베이스로 df\_corr, df2\_corr, df3\_corr 로 재편하였는데, 재편 방식은 일별로 구성된 데이터를

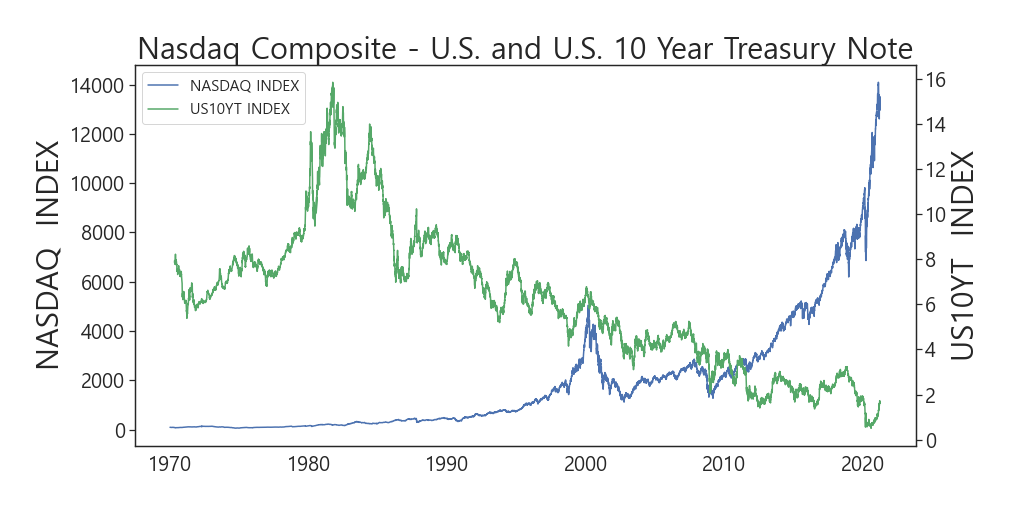
1차적으로는 연도별로 2차적으로는 월별로 재구성하여 correlation을 계산한 것을 추가하였다.

df\_corr의 Columns

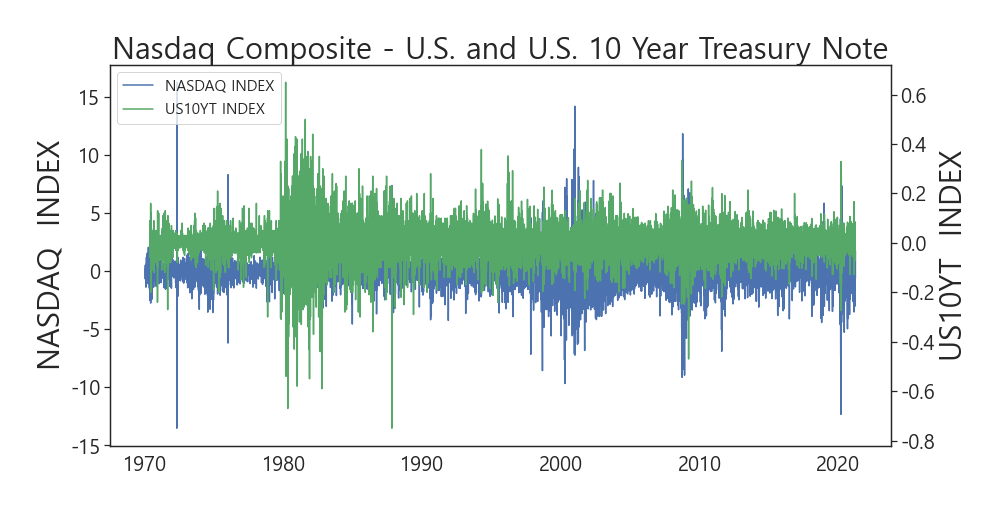
|  |  |  |
| --- | --- | --- |
| Date | Month | Corr |

4 분석 내용

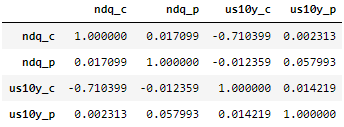
4.1 나스닥과 미국 10년 만기 국채 수익률



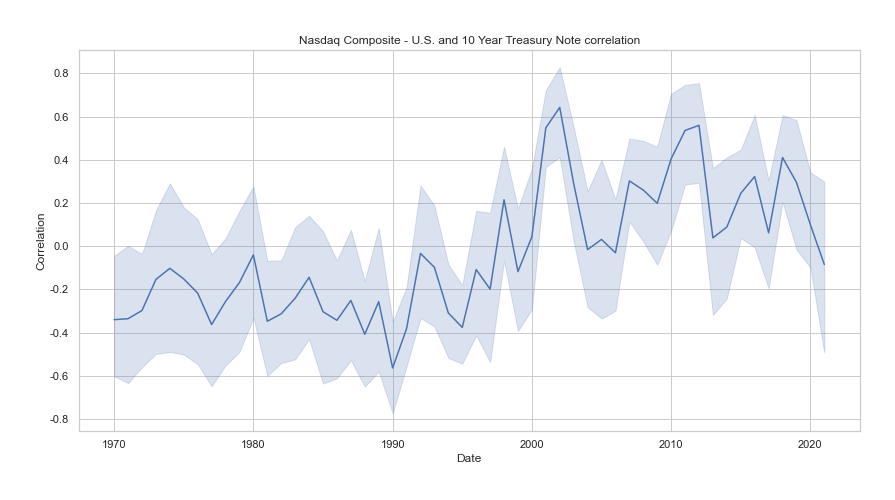
시각적인 정보로 확인할 수 있는 것들은 과거엔 미국 10년만기 국채 수익률이 굉장히 높았다는 사실이고 시간이 지남에 따라 현대에는 점점 더 하락하고 있는 모습을 확인할 수 있다. 나스닥 지수의 경우 우상향 하는 모습을 보이고 있다. 21년 초 금리상승으로 하락반전하여 고전하고 있다.



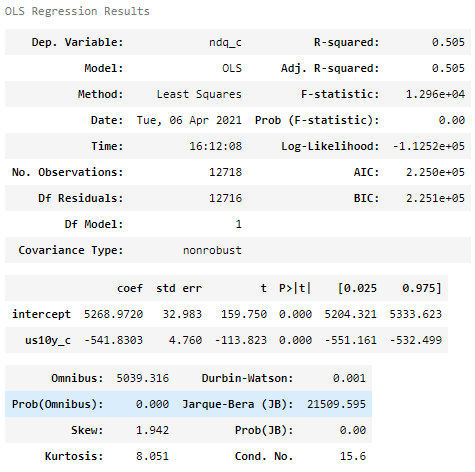
Nasdaq 지수선에선 1995~2000년 닷컴버블로 상승하다가 추락하는 모습을 볼 수 있다. 2000년대 초기에는 지수와 금리가 같이 하락하는 모습을 볼 수 있다. 이 같은 이유에는 9,11테러 , 아프간/이라크 전쟁등에 원인이 있다. 미국은 초 저금리 대책으로 경기부양을 노렸고 금리가 인하되자 주택융자 금리도 인하되었고 그러자 부동산 가격이 상승하였다. 이것은 이후 서브프라임 모기지사태의 시발점이 된다.



피어슨 상관계수를 확인한 결과 나스닥 종합 지수와 미국채 10년물 수익률과의 상관관계는 강한 음의 상관관계인 -0.710399를 나타내고 있다. 이외 나머지 변수들은 유의미한 상관계수를 발견하기 어려웠다. 예상했던 수치를 확인할 수 있었다.



위 그래프는 df 데이터를 베이스로 연도별 월별로 correlation을 계산한 것을 시간 순으로 lineplot을 사용하여 시각화한 그래프이다. 앞선 df데이터를 사용하여 확인한 나스닥과 미국 10년만기 국채 상승률 간 상관계수는 강한 음의 상관관계로 나타내고 있었으나. 이를 연도별, 월별로 나누어 그린 그래프를 확인하면 2000년도 이전까지는 0이하인 상관계수 밖에 없어 음의 상관관계 의 기조가 유지되었다면 2000년도 이후의 상관계수에서도 하락의 모습이 보이고 있지만, 2000년도 이전에 비하여상대적으로 양적 상관관계로 가까워지고 있음을 알 수 있다. 20년이후나 하락은 나스닥에 첨단기술관련회사가 많기 때문에 나타나는 현상으로, 첨단기술산업은 돈을 빌려 투자를 해야하는데 국채금리가 상승하니 이자비용이 증가할 수 있다. 이는 매출액의 감소로 이어질 수 있기 때문에 투자자들이 대거 주식을 팔아치움으로써 나타난 결과로 볼 수 있다.



해당 표는 나스닥과 미국 10년 만기 국채 수익률 간 OLS회귀분석을 한 결과표로 df 데이터에 intercept(절편) columns만 추가하여 진행하였다.

선형 회귀식은 y = α + βX (y=종속 변수, α=y절편(intercept), β=회귀계수)이다.

분석 결론

intercept의 coef는 5268.9720 ,us10y\_c 의 coef 는 -541.8303

이는 선형회귀식에서 α=5268.9720,β=-541.8303을 말한다.

따라서 선형회귀식은 Y = 5269 - 542X (소수점 반올림 ) 이다.

β가 1증가할때마다 α는 541 감소한다.

intercept의 p-value는 0, us10y\_c의 p-value도 0 에 수렴한다.

us10y\_c에 대한 정보는 통계적으로 유의하고, intercept에 대한 정보도 통계적으로 유의했다.

R-squared는 결정계수로 결정 계수(coefficient of determination)는 추정한 선형 모형이 주어진 자료에 적합한 정도를 재는 척도이다. 0과 1사이에 값이 있고 종속 변인과 독립변인 사이에 상관관계가 높을수록 1에 가까워진다. 결정계수가 0.505 이므로, us10y\_c가 ndq\_c에 대해 50.5%만큼 설명한다고 볼 수 있다.

결과 보고

모형 적합도

ndq\_c에 대하여 us10y\_c로 예측하는 회귀분석을 실시한 결과 이 회귀모형은 통계적으로 유의미하였다. (F(1,12716)=1.296e+04,p<0.05)

회귀계수

us10y\_c의 회귀계수는 -514.8303으로, ndq\_c에 대하여 유의미한 예측변인인 것으로 나타났다.

(t(12716)=-113.823,p<0.05)

결정계수

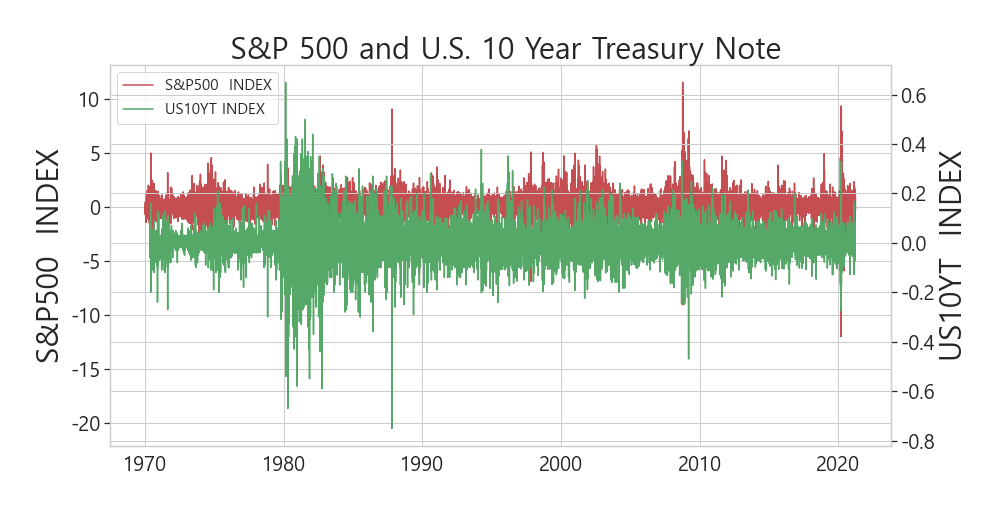
us10y\_c가 ndq\_c를 50.5프로만큼 설명한다.

4.2 S&P500 과 미국 10년 만기 국채 수익률

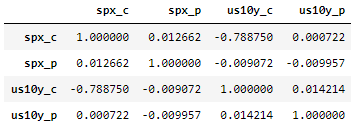
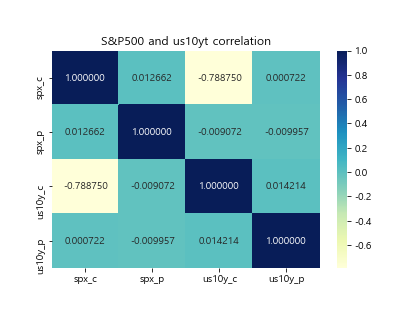


전체적으로 그래프의 추이가 나스닥과 큰 차이가 없음을 알 수 있다. 다만 지수에 편입되어 있는 종목들의

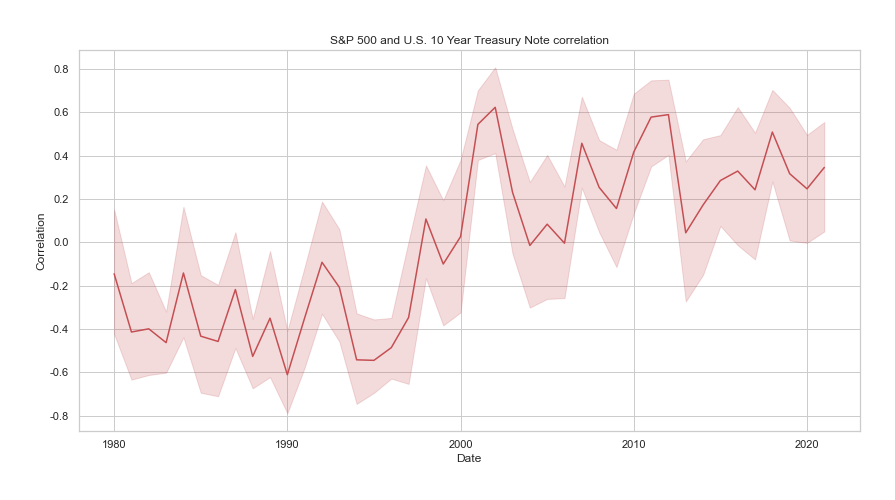
주 업종이 다르다 보니 닷컴버블때는 나스닥보다 낙폭이 크지 않았지만, 서브프라임모기지사태 때의 하락은 나스닥과 비교했을 때 지수가 더 많이 빠지는 것을 볼 수 있다. 이후 상승과 하락은 나스닥과 같은 추세를 보이며 다른 점이라면, 나스닥은 21년도들어오며 살짝 빠지는 모습을 보이는데, S&P500은 살짝 눌림을 주고 꾸준히 올라가고 있음을 볼 수 있다.



여담으로 1987년쯤 보면 20프로 이상 지수가 빠진 날이 있다. 미국은 80년도 레이거노믹스 아래서 경제호황을 누렸지만, 시간이 지날수록 재정적자와 경상수지 적자가 악화되어 금융시장은 과열조짐을 보이기 시작했다. 이에 대한 투자자들의 우려가 반영된 것이 검은 월요일이라고 하며 1987.10.19에 일어난 지수 폭락 사건을 말하게 된다. 이후 FED에서 금리를 낮추고 통화량을 증가시키면서 경제위기를 모면했다.

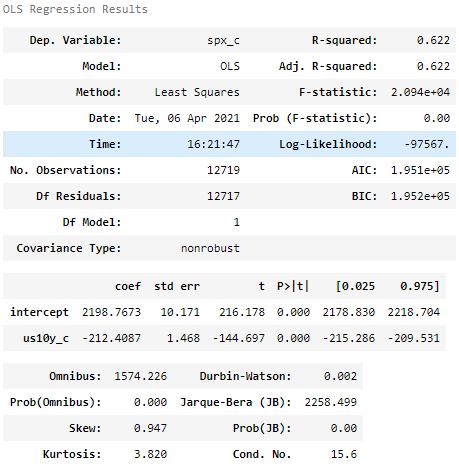


피어슨 상관계수를 확인한 결과 S&P500지수와 미국 국채 10년물 수익률과의 상관관계는 강한 음의 상관관계인 -0.788750을 나타내고 있다. 이외 나머지 변수들은 유의미한 상관계수를 발견하기 어려웠다. 예상했던 수치를 확인할 수 있었다.



위 그래프는 df2 데이터를 베이스로 연도별 월별로 correlation을 계산한 것을 시간 순으로 lineplot을 사용하여 시각화한 그래프이다. 앞선 df2데이터를 사용하여 확인한 S&P500과 미국 10년만기 국채 상승률 간 상관계수는 강한 음의 상관관계로 나타내고 있었으나. 이를 연도별, 월별로 나누어 그린 그래프를 확인하면 2000년도 이전까지는 0이하인 상관계수 밖에 없어 음의 상관관계 의 기조가 유지되었다면 2000년도 이후의 상관계수에서도 하락의 모습이 보이고 있지만, 2000년도 이전에 비하여상대적으로 양적 상관관계로 가까워지고 있음을 알 수 있다. 앞선 나스닥 그래프와 굉장히 유사한 모습을 하고 있는 것을 볼 수 있다. 특이점이라면

20년도 이후의 그래프가 나스닥은 하락한 반면 S&P500은 상승했다는 것인데, 이는 금리가 올라가면서 금융주가 많이 포함된 S&P500지수에는 오히려 도움이 되는 방향으로 작용한것으로 볼 수 있다.



해당 표는 S&P500과 미국 10년 만기 국채 수익률 간 OLS회귀분석을 한 결과표로 df2 데이터에 intercept(절편) columns만 추가하여 진행하였다.

선형 회귀식은 y = α + βX (y=종속 변수, α=y절편(intercept), β=회귀계수)이다.

분석 결론

intercept의 coef는 2198.7673 , us10y\_c 의 coef 는 -212.4087

이는 선형회귀식에서 α=2198.7673 , β=-212.4087 을 말한다.

따라서 선형회귀식은 Y = 2199 - 212X (소수점 반올림 ) 이다.

β가 1증가할때마다 α는 212 감소한다.

intercept의 p-value는 0, us10y\_c의 p-value도 0 에 수렴한다.

us10y\_c에 대한 정보는 통계적으로 유의하고, intercept에 대한 정보도 통계적으로 유의했다.

R-squared는 결정계수로 결정 계수(coefficient of determination)는 추정한 선형 모형이 주어진 자료에 적합한 정도를 재는 척도이다. 0과 1사이에 값이 있고 종속 변인과 독립변인 사이에 상관관계가 높을수록 1에 가까워진다. 결정계수가 0.622 이므로, us10y\_c가 spx\_c에 대해 62.2%만큼 설명한다고 볼 수 있다.

결과 보고

모형 적합도

spx\_c에 대하여 us10y\_c로 예측하는 회귀분석을 실시한 결과 이 회귀모형은 통계적으로 유의미하였다. (F(1,12717)=2.094e+04,p<0.05)

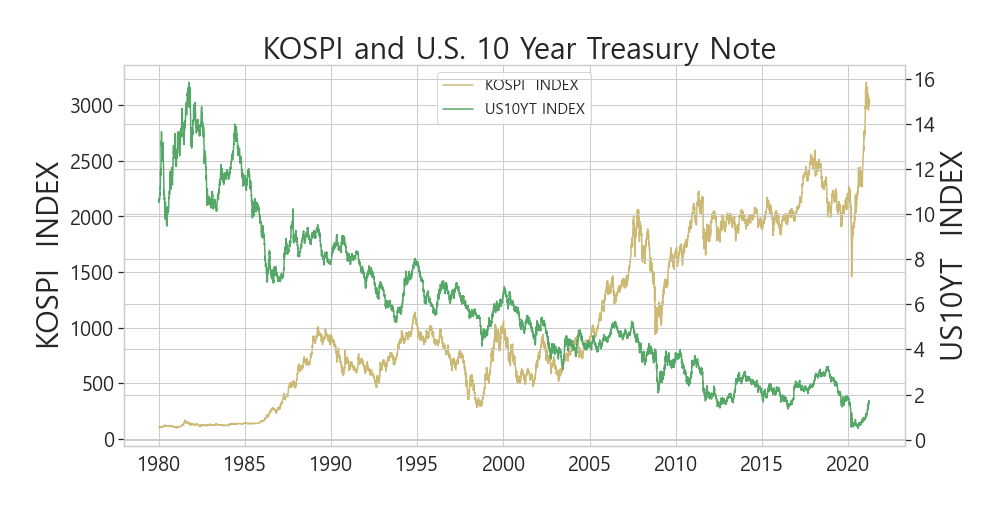
회귀계수

us10y\_c의 회귀계수는 -212.4087으로, spx\_c에 대하여 유의미한 예측변인인 것으로 나타났다. (t(12717)=-144.697,p<0.05)

결정계수

us10y\_c가 spx\_c를 62.2프로만큼 설명한다.

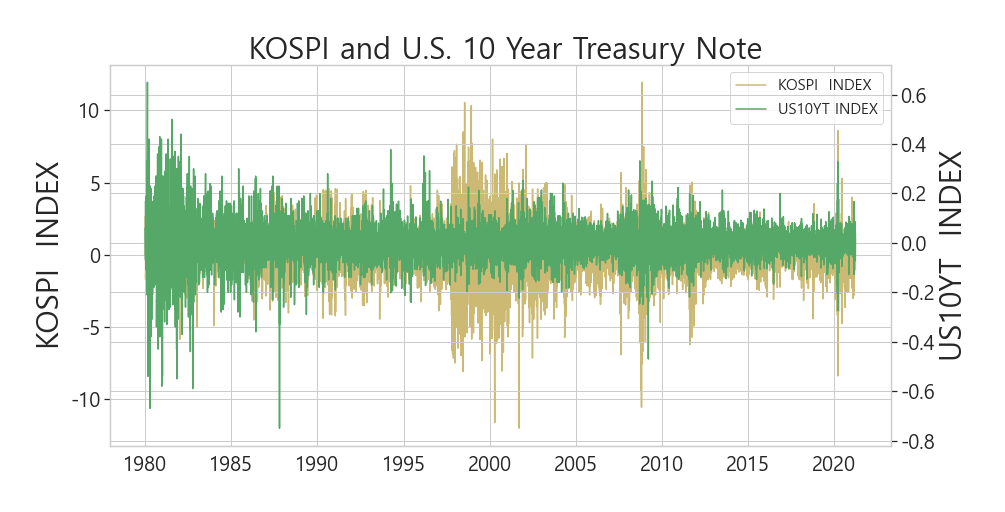
4.3 코스피200과 미국 10년 만기 국채 수익률



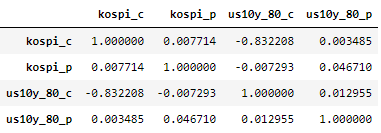
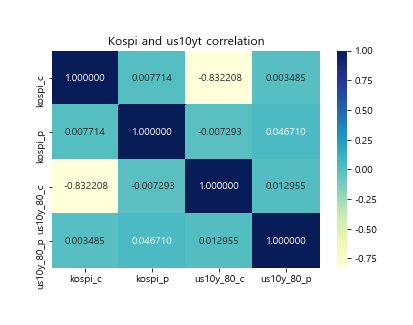
kospi 데이터 정보가 1980.01.04부터 시작하기 때문에 us10y또한 1980.01.04부터 시작하는 데이터로 재편

변수명은 tens\_usy80으로 하였다.

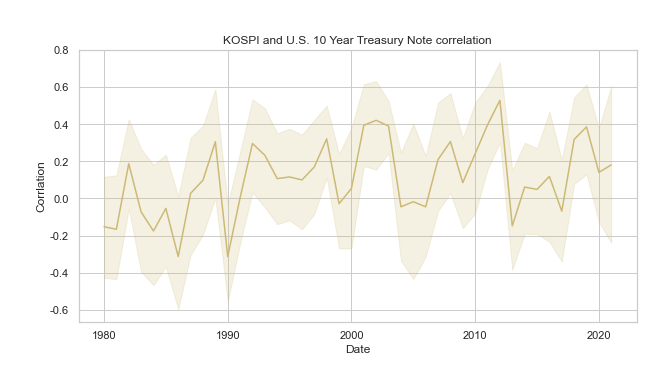
미국채금리와 전혀 관계없어 보이지만 시각적으로 봤을 때, 한국증시 자체가 미국증시와 동기화 되어 영향을 받는 일이 많음으로 코스피 지수 또한 우상향 하는 모습을 보이고 있다. 최근기조로는 미국증시보다 중국증시에 영향을 더 크게 받는 모습을 보인다.



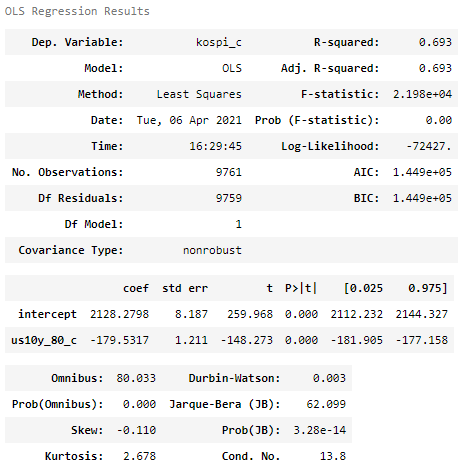
해당 그래프로는 변동률에 따른 유의미한 관계를 찾기 어렵다.



피어슨 상관계수를 확인한 결과 코스피 200지수와 미국 국채 10년물 수익률과의 상관관계는 강한 음의 상관관계인 -0.832208을 나타내고 있다. 이외 나머지 변수들은 유의미한 상관계수를 발견하기 어려웠다. 예상했던 수치를 확인할 수 있었다. 놀라운 점은 해당 국가에 있는 증시시장보다 다른 나라의 증시가 오히려 더 음의 상관계수 수치가 높았다는 점이다. 그만큼 미국이 대한민국에 미치는 경제적 여파가 큼을 알 수 있다.



위 그래프는 df3 데이터를 베이스로 연도별 월별로 correlation을 계산한 것을 시간 순으로 lineplot을 사용하여 시각화한 그래프이다. 앞선 df3 데이터를 사용하여 확인한 코스피 200과 미국 10년만기 국채 상승률 간 상관계수는 강한 음의 상관관계로 나타내고 있었으나. 의외로 크게 의미 없는 수치수준일지라도 90년도 이후부터 음의상관관계를 벗어나 양의 상관관계 0. ~ 0.4 사이를 오고 가는 것을 볼 수 있다. 이 또한 미국만큼은 아니더라도 미국채금리 상승 시 코스피200의 지수 또한 상승 했었음 을 의미할 수 있다.



해당 표는 코스피200 과 미국 10년 만기 국채 수익률 간 OLS회귀분석을 한 결과표로 df3 데이터

에 intercept(절편) columns만 추가하여 진행하였다.

선형 회귀식은 y = α + βX (y=종속 변수, α=y절편(intercept), β=회귀계수)이다.

분석 결론

intercept의 coef는 2128.2798 , us10y\_80\_c 의 coef 는 -179.5317

이는 선형회귀식에서 α=2128.2798 , β=-179.5317 을 말한다.

따라서 선형회귀식은 Y = 2128 - 179X (소수점 반올림 ) 이다.

β가 1증가할때마다 α는 179 감소한다.

intercept의 p-value는 0, us10y\_80\_c의 p-value도 0 에 수렴한다.

us10y\_80\_c에 대한 정보는 통계적으로 유의하고, intercept에 대한 정보도 통계적으로 유의했다.

R-squared는 결정계수로 결정 계수(coefficient of determination)는 추정한 선형 모형이 주어진 자료에 적합한 정도를 재는 척도이다. 0과 1사이에 값이 있고 종속 변인과 독립변인 사이에 상관관계가 높을수록 1에 가까워진다. 결정계수가 0.693 이므로, us10y\_80\_c가 kospi\_c에 대해 69.3%만큼 설명한다고 볼 수 있다.

결과 보고

모형 적합도

kospi\_c에 대하여 us10y\_80\_c로 예측하는 회귀분석을 실시한 결과 이 회귀모형은 통계적으로 유의미하였다. (F(1,9759)=2.198e+04,p<0.05)

회귀계수

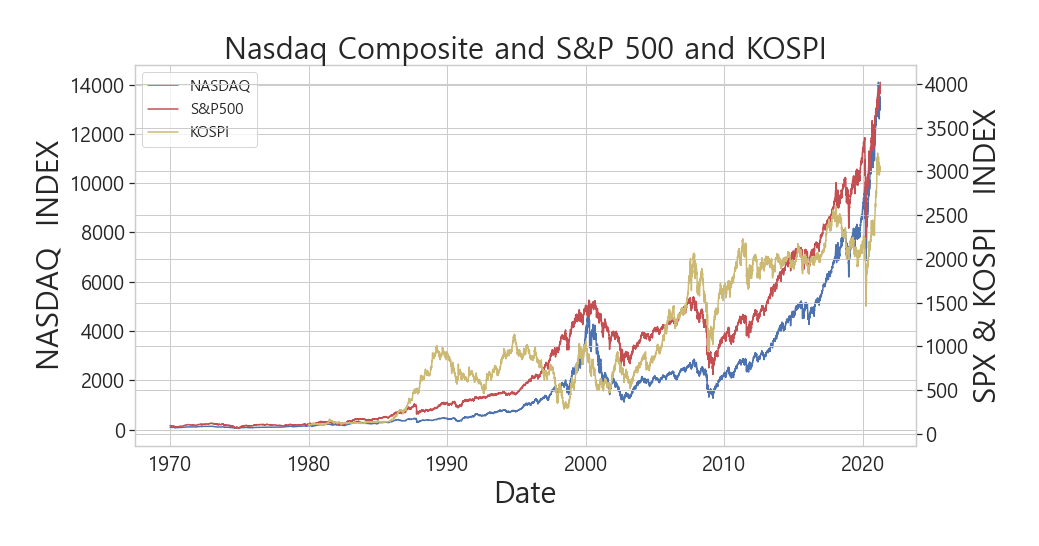
us10y\_80\_c의 회귀계수는 -179.5317으로, kospi\_c에 대하여 유의미한 예측변인 인 것으로 나타났다. (t(9759)=-148.273,p<0.05)

결정계수

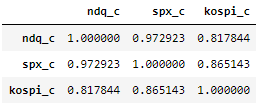
us10y\_80\_c가 kospi\_c를 69.3프로만큼 설명한다.

5 분석 외 시각화 자료

5.1 Nasdaq Composite - U.S , S&P 500, KOSPI 간 선그래프

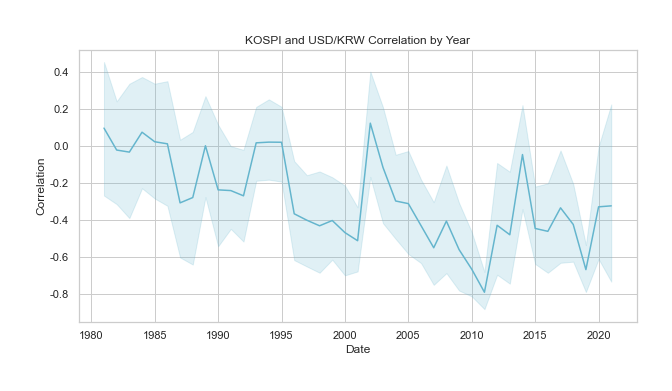
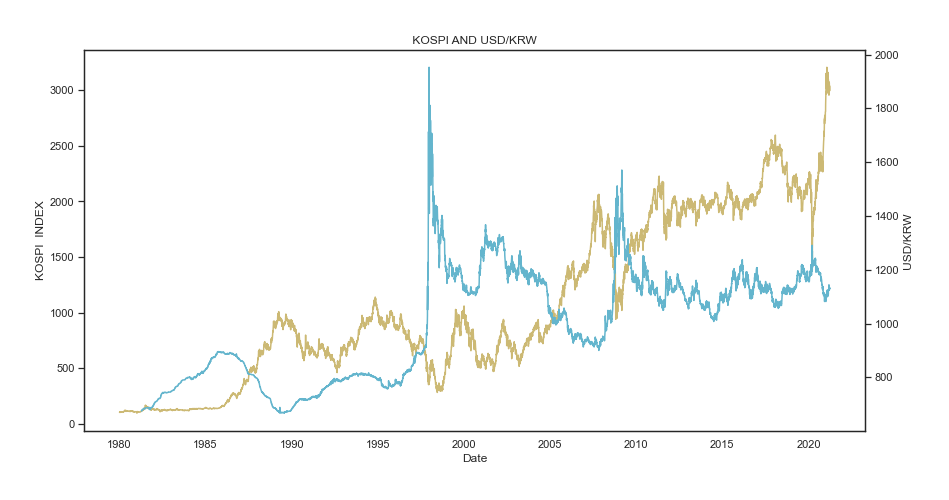


세 지수를 연도별로 선 그래프 시각화 하였을 때 확인되는 바, 개별적으로 한국의 IMF같은 특정 사건이 아니라면 큰 흐름으로 같은 방향으로 움직이고 있음을 알 수 있다. 2000년도 닷컴버블로 인해 거의 전 세계가 증시하락을 경험했고, 2008년 전세계적으로 금융위기가 닥쳐오고 리먼사태를 겪으며 동반하락하기도 한다.이후 경제회복은 더디나 증권 지수는 꾸준히 상승하였고 2020년 코로나를 겪고 또 한번 폭락했으나 거의 모든 증시가 1년사이 엄청난 속도로 회복하게 된다. 결론적으로 우상향 하고있는 그래프를 보여준다. 이는 거의 전세계적으로 나타나는 흐름이라고 보면 된다. 사실상 1970년대를 넘어가면서부터 세계경제는 어떻게 보면 서로 동기화 되어있다고 생각된다. 그 중에 미국은 세계경제의 한 축이고 우리나라는 그러한 미국과 중국에 영향을 크게 받기 때문에 증시그래프가 비슷하게 나타나는게 아닌가 생각된다.



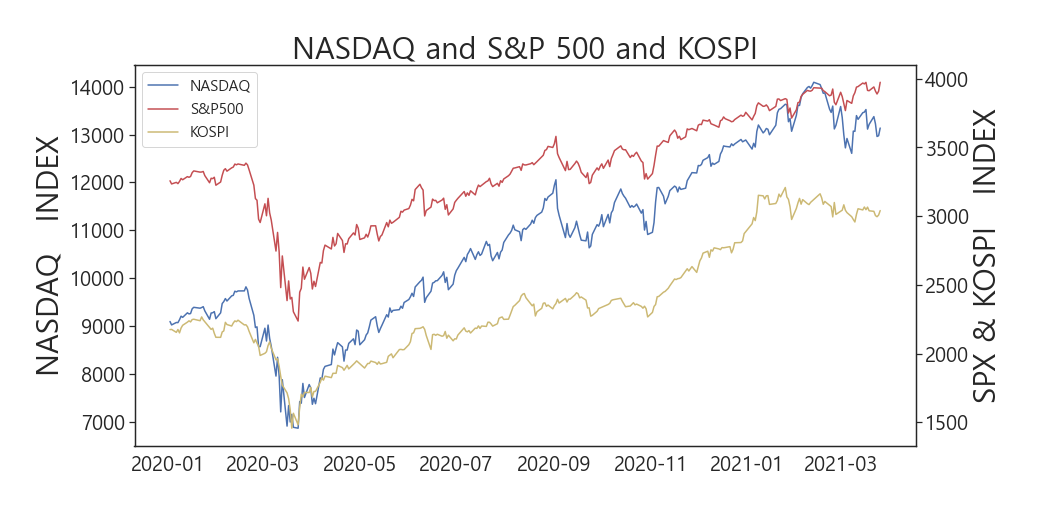
Pearson 상관계수 수치를 보면 코스피와 다른 두개의 미국증시의 상관계수가 0.8이상으로 강한 양의상관관계에 있음을 알 수 있다. 이렇다 보니 잠을 자고 일어나 아침에 꼭 확인하게 되는게 미국 증시가 된다. 오늘 새벽에 미증시가 폭락하면, 오늘의 코스피 또한 폭락을 피하기가 어렵기 때문이다. 이러한 기조는 국내 기업들이 첨단산업, 신기술 계발이나 IT쪽 기업들이 다수 포진 되어있고 성장잠재력이 높은 기업들이 많다 보니 어쩔 수 없는 상황이다. 좋은 측면으로 보면 미국 나스닥 지수가 오르면 국내 지수도 오를 가능성이 높다 정도이다.

5.2 KOSPI와 원달러환율과의 lineplot사실상 증시라는게 금리 하나만으로 오르고 내리는 곳이 아니라 다양한 경제지표들이 그날그날 발표되면 그 데이터들을 보고 상승과 하락의 방향을 정하게 된다. 그 다양한 경제지표중 하나인 원, 달러 환율을 한번 확인해보고자 간단하게 그래프로 시각화 해보았다.

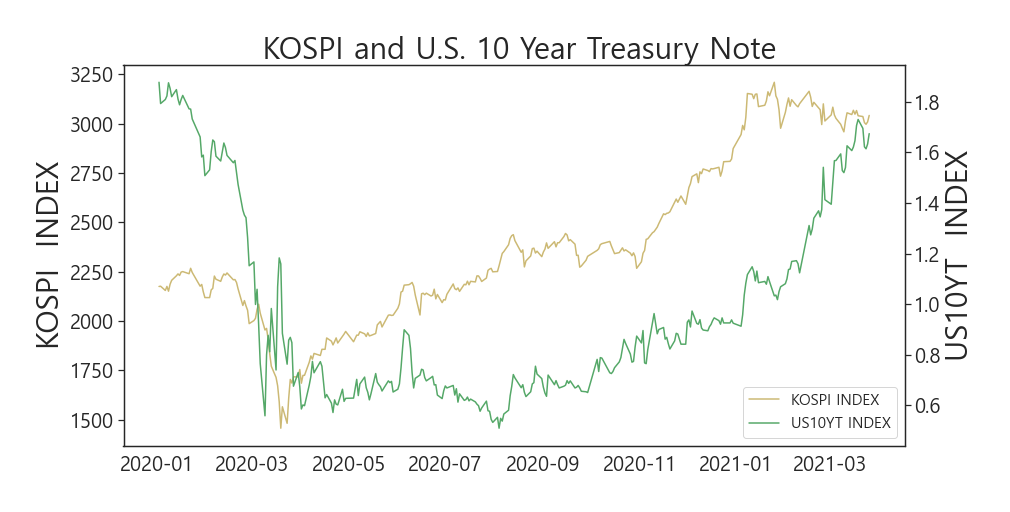


연도별 월별로 나눈 correlation 선 그래프 이다. 전체기간으로 상관관계수치가 0을 넘은적이 거의 없다. 코스피 지수와 원달러 환율은 음의 상관관계에 놓여있다. 이유인 즉 환차익,환손실의 연유가 크다. 환율이 상승 한다는 것은 원화가치 하락을 말하고 이렇게 되면 코스피에서 원화를 달러화로 환전할 때 손실이 발생한다. 따라서 환율이 오를 것 같으면 외국인들은 환손실을 우려해 원화를 달러로 바꿔서 kospi시장에서 나가게됨으로 지수가 하락한다. 반대로 환율이 떨어지는 추세라면 원화가치 상승을 말하기 때문에 환차익을 노리기위해서 외국자본이 kospi시장으로 들어와 지수가 올라간다.

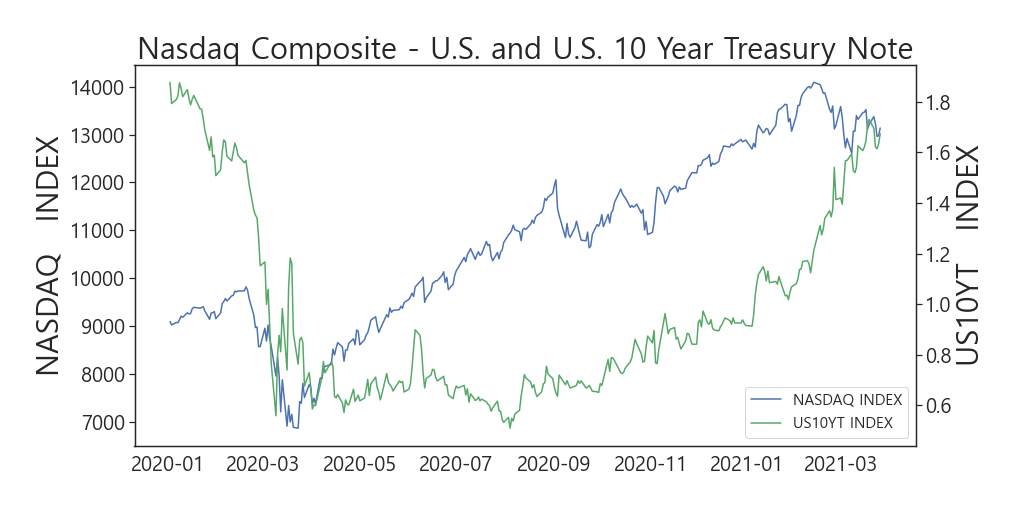
5.3 20~21년도 지수 변화 및 미국10년만기 국채 linplot

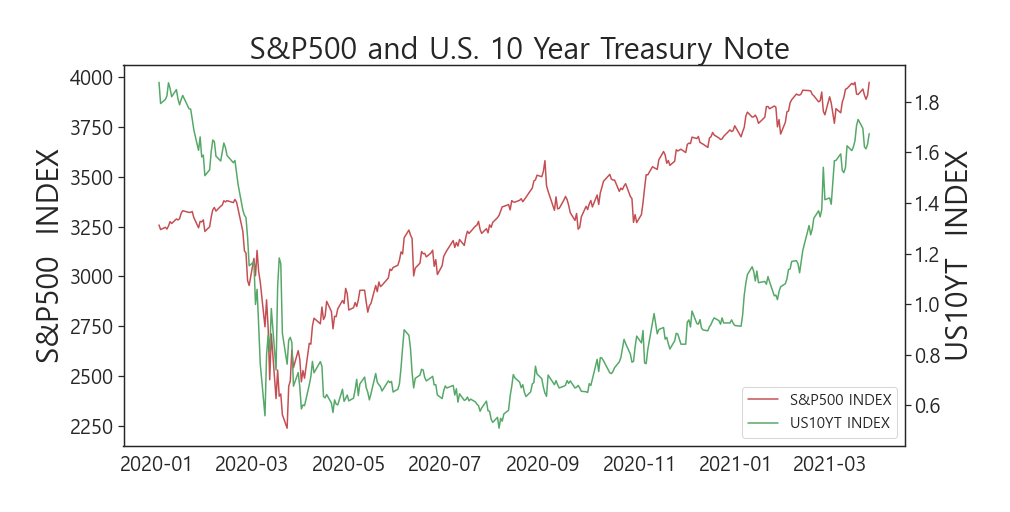


본 분석과 전혀 연관이 없는 그래프는 아니나 특별히 분석할 내용은 없기에 분석 외 시각화 자료에 삽입하였다. 보는 바와 같이 20년도 코로나가 폭발적으로 전염되는 시기에 주가가 함께 폭락하며, 다시금 올라가는 그래프로 보여진다. 그리고 21년을 지난 그래프를 보면 S&P500은 점진적으로 상승하는 반면에 나스닥은 고꾸라지는 모습을 볼 수 있고, 코스피 같은 경우 1월이후로 옆으로 기는 모습을 확인할 수 있다. 이 시기가 코로나 이후 제로금리에 가깝게 내려간 미국 국채 10년만기 채권 금리가 폭등하는 시기로 지금까지 동기화되어 가던 세 증시의 모습에서 차별하 된 모양을 보이게 된다. 앞서 설명한 것 과 같이 S&P500 같은 경우는 금융주가 포함되어 있기 때문에 금리가 오른다 한들 지수가 밀리지 않게 되었고, 나스닥은 성장주와 첨단 산업 기술주들이 많이 포진해 있기 때문에 금리인상에 민감한 반응을 보인다. 코스피는 국내 시가총액 200위 기업주들이 들어가 있음으로 금융주 또한 많기 때문에 대한민국 자체가 성장주,기술주를 기반으로 하는 시장이라 하락은 면하지 못했지만 폭락하는 모습은 보이지 않는다.



금리가 상승하며 증시의 상승이 둔화되는게 보인다.





같은 미국 내에 있는 증시 지만, 미국 10년만기 국채 수익률이 상승할 때 보이는 추세는 다르게 나타난다.

6 분석 결론

위 분석해서 나온 데이터들을 종합해 본 결과, 미국 10년 만기 국채 금리 상승률이 올라갈수록 증시가 하락할 확률이 높다는 것이 밝혀 졌다. 그 이유의 첫째로 그래프로 보이는 바 또한 그러한 추세를 보여주며, 두번째로 상관계수 수치를 확인할 때 모두 강한 음의 상관관계를 나타내고 있었다. 세번째로 OLS회귀분석을 실시하였을 때 모형적합도상 비교데이터간 유의성을 확인했고 회귀계수가 전부다 음수로 나타난 점, 낮은 수치이지만 결정 계수가 모두 50프로이상은 되었다는 점이 있다. 이처럼 결구 과거에서부터 연구된 내용과 큰 차이는 없는 결과가 나왔다.

그렇다면 이 분석 결론을 바탕으로 21년 이후의 투자포지션을 어떻게 잡아야 하는게 맞을까?

미국 국채 금리가 오를 때 주가도 같이 상승할 때는 경기가 전반적으로 좋아질 것이란 기대감이 크기 때문이다. 그러나 최근에는 금리가 올라가는 속도가 너무나 가팔랐고 그러다 보니 금리상승의 요인이 실물경기의 활성화가 아니라 통화팽창으로 인한 인플레이션 때문이라고 생각했기 때문에 증시가 하락 한 것으로 생각된다. 현재의 상황은 그 두가지가 모두 엉켜있는 상황으로 투자자들의 판단이 혼란스런 시기가 되었다. 과거 시장에선 대체로 미국 국채와 투자자산들과의 상관관계에 대한 통설이 역의 관계에 있단 거였고, 방금 확인하기도 하였다. 그러나 최근 경제전문가들은 양적완화 제로금리 재난지원금등 지난 10년간 예측 불가능한 수준으로 돈을 풀어, 처음맞이하는 국면이라 시장이 아주 복잡하고 엉망이라 기존의 이론이 깨지고 있음을 말하고 있다. 이렇게 주식시장에 불확실성이 커지다 보니 투자심리가 위축될 수 밖에 없는 노릇이다.

그러나, 선 그래프에서 확인할 수 있는 것 같이, 미국주식시장과 코스피는 어쨋든 우상향 그래프를 보여주고 있다.미 국채 10년물 금리상승과 비교하였을 때도 2000년도 이전엔 음의 상관관계로 금리가 상승하면 주식시장엔 좋지 않았으나 2000년도 이후는 같은 방향으로 올라가는 모습이 보인다. 이는 경제회복과 더불어 주식시장도 같이 활성화되고 있다는 증거로 생각된다. 현재는 미래 경제에 대한 불확실성으로 변동성이 심하긴 하나 미 국채금리의 단기적인 상승에 투자자들이 발작하는 것 이라 보고 fed 또한 경제가 완전히 회복되기전까지는 테이퍼링을 하지 않겠다고 했다. 하지만 분기마다 새로운 보고를 할 것이라고 말했기 때문에 언제나 관심을 가지고 FED의 발언에 유의해야한다. 결론적으로 이미 풀린 돈과 계속 나오는 부양정책으로 시간이 지날수록 경제의 회복은 가시화되고 실업률이 회복되고, 코로나백신의 보급이 활발해지면서 미국채금리도 상승되겠지만 경제회복이 뒷받침되면서 인플레이션우려가 줄어들면 주식시장 지수도 상승할 수 있을 것이라고 생각된다. 시간이 더 지나 결국엔 시장에 풀린 돈을 인플레이션의 여파가 오기 전에 회수하는 시점은 올 것이라고 생각되지만 지금 당장은 아니라고 생각한다. 그건 FED에서도 말하고 있고, 아직 경제회복은 시작에 불과하다고 생각한다. 결국엔 기존 금리 인상을 하겠지만, 지금 당장의 미국 10년만기 국채 수익률의 상승을 크게 걱정할 필요는 없다고 생각한다. 단기적으로 1년간은 주식시장은 더 상승할 것이라고 예측한다.

6 분석 한계

첫 번째로 증시와 미국 10년 만기 국채 수익률의 관계를 분석하려 했지만, 사실상 증시가 미국 10년만기 국채 수익률 하나만으로 좌지우지되는 영역이 아니라는 것이 가장 큰 한계라고 본다. 사실상 독립변수가 될 수 있는 경제지표들이 너무 많아 어떤 것을 추가해야 하는가 고민하였으나 경제영역에 지식이 많지 않고 국채와 증시에 대한 이해도도 높지 않았기 때문에 데이터 표본의 다양성을 반영하지 못하였다. 추가적 설명에 있어서도 미흡한 부분이 많았고 , 권위있는 논문들의 내용들도 찾는데 실패하였다. 국내 비슷한 논문들에 있는 외국경제학자들의 논문들이 표현된 것을 보긴 했으나 찾지 못하였다. 두 번째로는 일단 분석 영역에도 알고 있는 지식이 부족하다 보니 분석기술의 한계가 있겠다. 그래프 시각화의 경우 plotline 이외에도 좀 더 다양하게 여러 그래프를 시각화 하여 분석할 수 있지 않을까 했지만 적용하기가 힘들었고, pearson 상관계수의 경우도 연관성만 나타낼 뿐 인과성을 나타내지 못하는 것에 한계가 있었다. OSL회귀분석의 경우 수치상으로 유의미 하다고 나오긴 하였지만 회귀분석에 대해 지식이 미흡한 부분이 있어 이게 과연 제대로된 분석결과 인지에 대한 한계가 있다. 추후 경제지표에 관한 추가적인 지식과 상관관계분석 과 회귀분석에 대한 이해도를 높이는 방안으로 나가야 더 좋은 결론을 낼 수 있다고 생각한다.

7 자료 참조

주식시장, 금리에 관한 정보 - https://blog.boolileo.com/investment-analysis/economic-analysis/

장단기 금리역전 , 미국채 금리 상승 해석 - <https://blog.naver.com/skyat23/222210367128>

금리와 주식간 관계 - <http://www.hani.co.kr/arti/economy/finance/835229.html>

회귀분석 방법 - https://mindscale.kr/course/basic-stat-python/11/

OLS 회귀분석 방법 - <https://bkshin.tistory.com/entry/DATA-17-Regression>

회귀분석 방법 - <https://velog.io/@pyose95/Data-Analysis-15>

경제지표 정보 - <https://ecodemy.cafe24.com>

분석데이터 - <https://stooq.com/>

분석데이터 - <https://finance.naver.com>