[ BTC vs 경제지표 결과 ]

1. 비트코인 급등/급락은 어떤 지표의 영향인지 파악.

* 비트코인 일일수익률 Max 10개, Min 10개 날짜 선택해서 무슨 일이 있었는지 본 결과, (1) 규제했을때 급락하고,

1. 유명인사가 긍정적인 말을 했을 때 급등하며,
2. 도저히 경제지표의 영향이라 볼 수 없을정도로 굉장히 짧은 시간안에(1주일 이내) 급등과 급락을 반복했음..

* 급등과 급락은 경제지표로 설명 불가능. 설명한다면 심리지표로 설명하거나 크롤링으로 ‘비트코인 규제’ 뉴스들 긁어와서 설명할 수 있을 것 같음.

[ 급등, 급락? ]

1. 2017-07-21 : 비트코인이 비트코인 캐시로 하드포크됨.

(하드포크 : 기존의 버전을 개선하면서 과거 버전과 새로운 버전이 완전히 분리가 되는 것을 의미. 하드포크로 새로운 암호화폐가 만들어지면 기존 체인 참여자에게 코인 배당이 돌아간다.)

1. 2017.12.07~08 :

* 근데 하드포크는 호재일수도 악재일 수도 있다고 함.

(<http://it.chosun.com/site/data/html_dir/2018/12/10/2018121000853.html> )

1. 2021-02-09 : 21년 2월 8일, 일론 머스크 테슬라 CEO가 대량으로 비트코인 매입하고 결제수단으로 비트코인 결제를 추가할 것이라고 발표. <https://www.news1.kr/articles/?4214478>
2. 폭락
3. 2013-05-02 : 비트코인 인프라가 충분히 구축되어 있지 않던 당시, 비트코인 거래소인 Mt.Gox와 비트코인 거래를 처리하는 BitInstant가 거래 처리를 멈추었다.

<https://blockinpress.com/archives/2227>

1. 2017-09-15 : 중국의 비트코인 규제. 중국에서 최장기간 운영한 BTCChina 거래소는 정부의 압박에 의해 거래소를 닫았고, 폐쇄 결정. <https://blockinpress.com/archives/2227>
2. 2018-01-17 : 2018-01-11에 박상기 법무부 장관이 가상화폐 거래를 금지하고 거래소 폐쇄를 고려하고 있다고 신년 기자간담회에서 발표. 🡪 가격 폭락.
3. 비트코인이 투자자산으로서 가능성이 있는지.

* (1) 비트코인이 환율, 소비자물가지수 등 실물경제지표와 상관관계가 있는지

(2) 비트코인이 금, 채권, 주식과 대체재/보완재의 관계를 가지는지

를 통해서 투자자산으로서의 가능성을 판단. 만약 ‘관계가 있다’ 라고 나오면 경제지표에 영향을 받고 금융투자자산의 대체재/보완재 역할을 하므로 투자자산으로서의 가능성이 있다고 할 수 있다.라고 결론

* 방법:

1. 시계열 데이터 정상화시킨 후 OLS 적용해서 상관관계 파악하거나
2. VAR모형 사용.

* 정상화 방법: 일단 Y와 각 X들에 대해 정상성 검정을 한 후, 정상성 검정을 만족하지 않는 변수들 각각에 ARIMA 적용. 그렇게 각 변수에 대해서 나온 ARIMA 모델을 회귀모델에 넣은 후 OLS 적용할 수 있을 것 같음. (한가지 문제가 있다면, 정상성 검정할때 type 1 error inflation이 일어날 수 있다는점..? )

1. 비트코인을 화폐로서 사용할 수 있을지

* 급등과 급락이 경제지표로 설명된다고 할 수 없고 변동성이 너무 심하므로 아직까지는 화폐로써의 역할을 잘 할 수 있을지는 미지수이다. 정도의 결론?

ARIMA 복습: <https://predictor-ver1.tistory.com/3>

[ 흐름 ]

1. 경제지표 나열 & 설명:

Y : BTC 가격, X: : BTC 공급량(O), 환율(O), 금가격(O), 채권수익률(O), KOSPI(O), VIX(O), NASDAQ(O), VKOSPI(O), 소비자물가지수(O).

1. 각 변수가 의미하는 것 & 왜 사용하게 되었는지 O
2. 데이터 전처리 과정 설명 O
3. 경제지표와 BTC 가격의 시계열 plot 겹쳐서 상승/하강 추세 보여주기 (O)

* 직관적으로 도저히 말이 안 되는 것들, 예를들어, 비트코인 공급량과 비트코인 가격이 양의 상관관계를 갖고있다, 이런 경우엔 제 3의 변수가 영향을 미치고 있을 것이라고 결론.
* 디자인 : 회귀분석팀 주제분석 1주차 ppt 참고

1. 경제지표와 BTC 가격의 상관관계/상관검정 보여주기. (O)

* 검정을 해본 결과 유의미한 변수들이 많아서 경제지표가 비트코인 가격에 영향을 미침을 알 수 있다. 따라서 앞으로 경제지표들을 활용해서 모델링 진행해볼 예정이라는 흐름.

1. 일일수익률 min, max나온 날짜에서 무슨 일이 있었는지 설명. (O)
2. BTC 가격의 일일수익률 시계열 그래프 그려서 min, max지점 체크한다음에 어떤 일이 있었는지 확인 (O)

* 결과 : 급등, 급락은 경제지표로 설명할 수 없고 국가의 규제나 유명인사의 언급, 또는 심리지표로 설명된다고 볼 수 있다. (O)

중제목 1: 경제지표 설명& 왜 선정했는지

중제목 2: 시각화 (plot 겹친거)

경제지표 설명& 시계열 plot 그리면서 이러한 역사가 있었고 이렇다. 어떤 경제지표는 상승추세다 이런거 보여주고. 비트코인 조사한거를 분석배경으로

시계열 plot 상관관계 이렇고 상관관곅 검정 이렇고. 유의미한 변수들 많아서 비트코인 가격에 영향을미친다. 앞으로 변수들을 ㅘㄹ용해서 모델링 진행해볼 예정이다.

투자자산인지 아닌지 는 킵해둡시다.그건 3주차에 넣으면 좋을 듯.

코드, 데이터프레임을 ppt에 넣지말고 넣을거면 차라리 표로 넣기.

발표 대본써주기 내 파트. 3번 대단원.

일일수익률은 분석배경에 넣기.

저녁 8시까지 -

금3시 스터디.

대본 최대한 빨리 올려주기

앞으로 분석을 어떻게 할것인가?

* 변수선택을 위해서 회귀방법을 통해서 변수선택 하려고 했는데 TSLM 사용가능. 변수선택에

VAR 모형을 쓰면 어떤변수가 얼마만큼 영향을 미치는지를 알 수 있다. 우리가 구한 변수를 벡ㅍㅌ터형태로 회귀에 돌림. X가 y에 얼마만큼여 ㅇ향을 미치는지를 ㅏㅇ룻있음. 우리도 비슷하게 VAR모형을 돌려서 통곚계적으로 유의미한 변수들을 가지고 LSTM, ARIMAX, 랜포나 cotboost같은 머신러닝ㅇ을 돌려보는게 어떨까 생각중. VAR이 통계적인 테스트가 다 된다고 논문에 나와있음.

유의미한 변수들을 찾아낸다까지 1주차에 하고 LSTM 랜포 같은 모델링 그 후에 돌릴 것.

교차검증이 시계열의 경우에는 다르게 함. Rolling window

감성분석. 감성분석이 빨리 되어야 모델을 돌릴 수 있다. 왜냫면 가감ㅅ어분석을 통해서 부정, 긍정 뉴스기사 변수들을 만들어야 모델링 가능… 그래서 VAR이랑 감성분석이랑 같이 진행해서 최종 데이터프레임 완성한다음 모델링 하나씩 맡아서 돌릴 예정.

VAR 돌려서 유의미한 변수 찾고 감성분석 변수 넣어서 final frame 만들어서 그 final frame가지고 모델링 진행하는게 전체적인 분석 flow

2주차부터 감성분석, VAR

토픽모델링 🡪 워드클라우드 같은 방식으로 EDA 할수도 있을 듯

LDA?