**HW5 (Due date : July.16 PM 19:59, upload in web)**

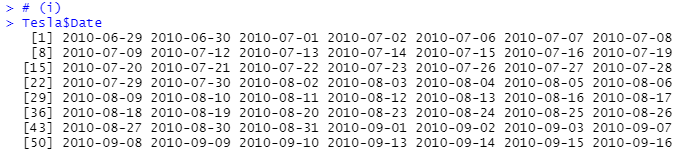
**Submit the homework named as HW5\_이름.doc, HW5\_이름.R (코드도 첨부)**

**1. Public data in kaggle**

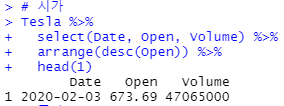
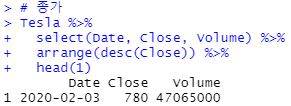
1.1. Search the data related 'tesla' in kaggle, and download a data 'Tesla stock data from 2010 to 2020'

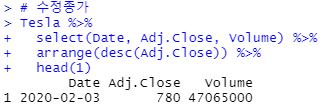
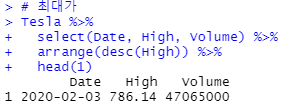
1.2. Answer the questions

(i) Is monthly data or daily data? Daily data

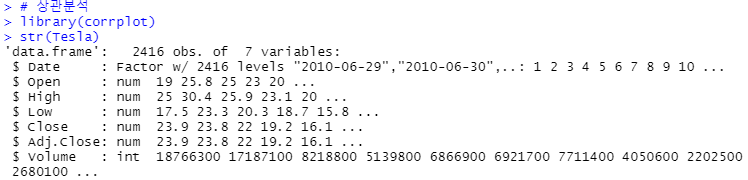


(ii) what is the date with the highest price? The highest price is “786.14” on 2020-02-03

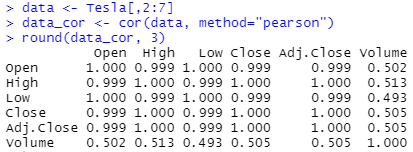
 

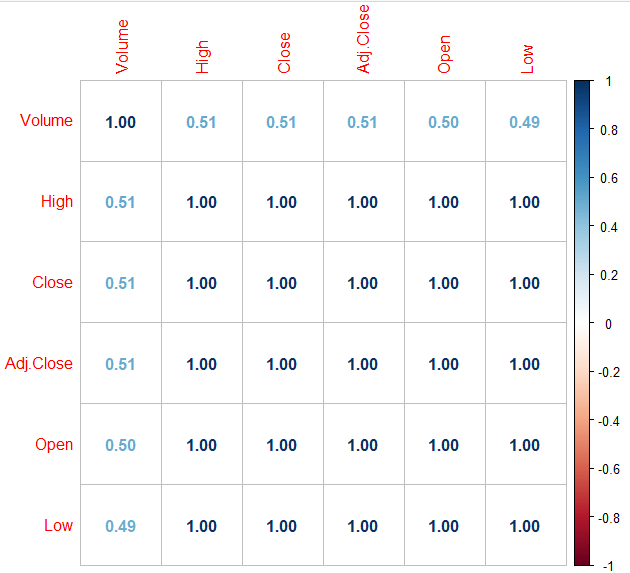
(iii) Any trend you find out from the given information in kaggle?



Factor 타입인 Date변수를 제외하고 상관계수 계산

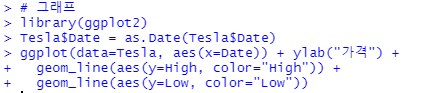


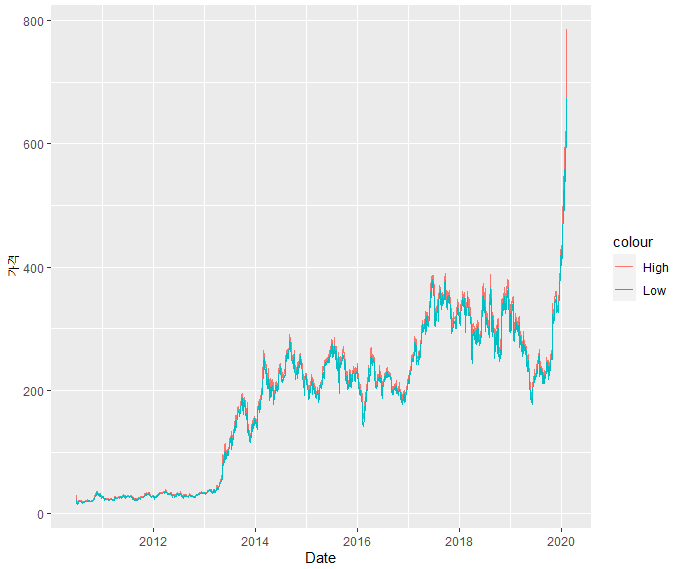




결과로 보았을 때, Volume(거래량)을 제외한 나머지 요인들은 각각 서로의 상관계수가 1또는 거의 1에 가까운 값을 가진다. 그래서 Volume(거래량)을 제외한 시가, 종가, 수정종가, 최고가 최저가의 경우 서로가 강한 상관성이 있다. 역으로 이러한 price들과 거래량은 서로 상관관계가 크게 없다고 볼 수 있다.

시간별 전체 주가 변동에 대한 그래프를 그려보면

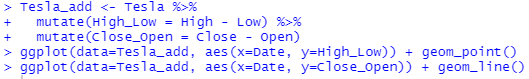


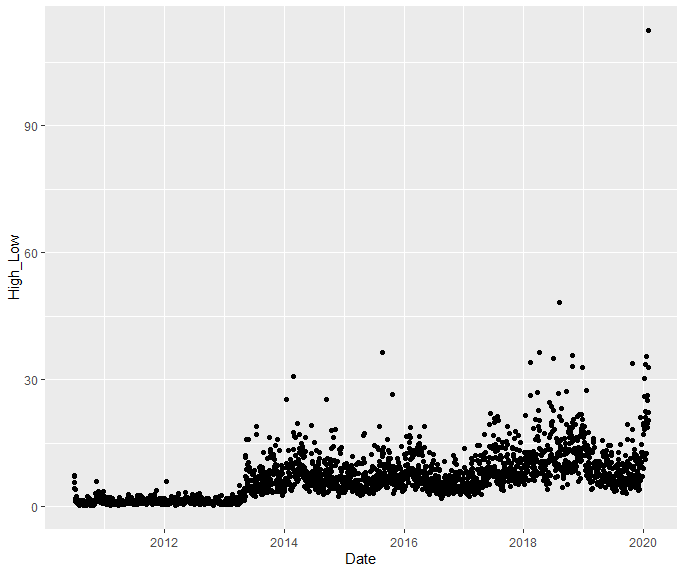
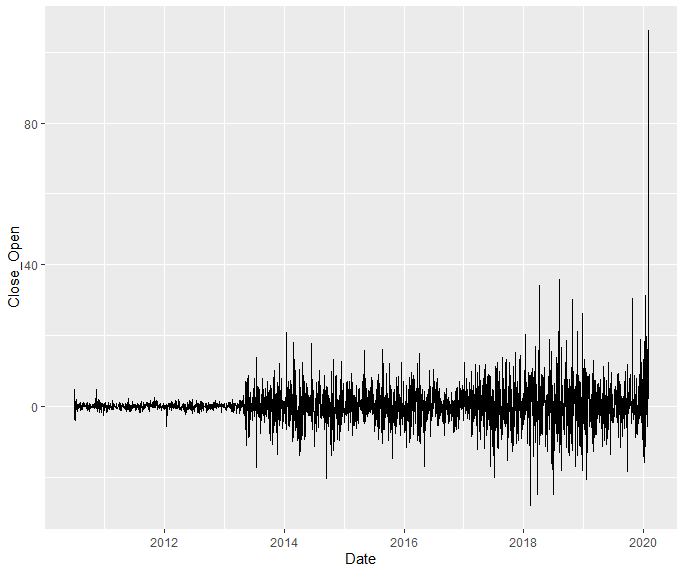


일별 최고가와 최저가를 기준으로 그렸을 때, 위와 같다.

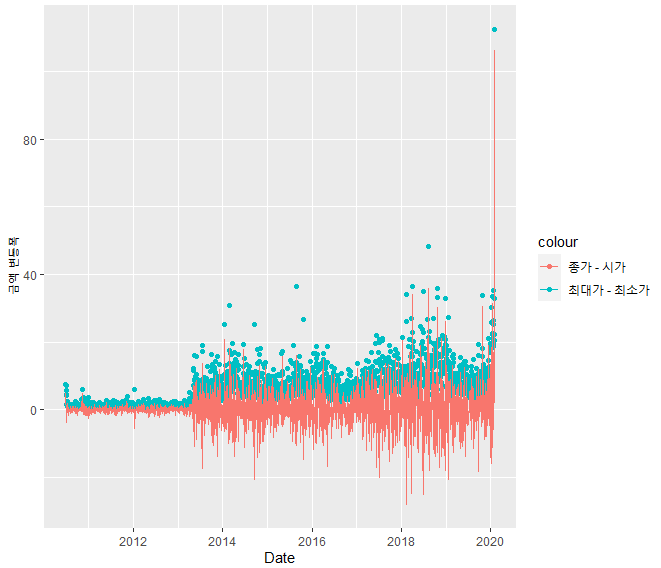
2013년부터 급격한 성장세를 보이다가 2017년에 다시 한번 성장세를 보이고, 2019년 초부터 중순까지 하락세에 있다가 갑자기 주가가 폭등하는 것을 볼 수 있다.

그래서 변동폭은 얼마나 되는지 궁금해서 column을 추가하여 그래프를 그려보았다.



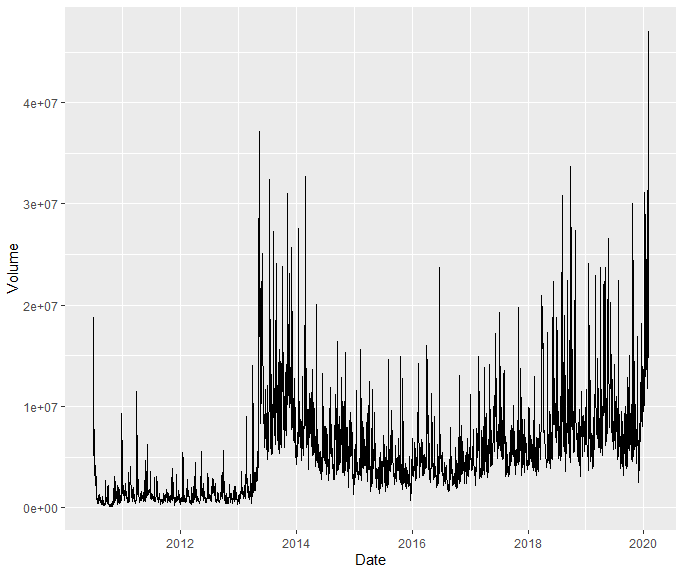
 

좌측은 당일 ‘최고가-최저가’를 일자별로 나타낸 것이고, 그날 얼마나 사람들의 마음을 뒤흔들었는지를 대략적으로 알 수 있다. 우측은 당일 ‘종가-시가’로 시가에 산 사람들이 장이 끝나고 나서의 기분이 어땠을지를 대략적으로 체감할 수 있다.(0보다 작으면 우울했을 것이고 0보다 클수록 기뻤을 것으로 예상)



해당 그래프로 보았을 때, 데이터의 마지막 날인 2020-02-03에 많은 사람들이 폭등하는 가격에 희비가 엇갈렸을 것으로 보인다.





주가가 폭등할 시기에 거래량이 많을까 궁금해서 거래량을 시기별로 시각화 해본 결과, 주가가 폭등 또는 폭락할 때, 거래량이 어느정도 많은 것은 맞지만 그게 금액 변동 폭만큼 차이가 나는 것처럼 보이지는 않는다.