

2022년 NH투자증권 빅데이터 경진대회 (예선)

팀명	아낌없이 주는 나무				
팀원	성명	생년월일	학교	학과	연락처
	김성연	1998.04.01.	중앙대학교	응용통계학과	010-2398-6822
	김소은	2000.03.07.	중앙대학교	응용통계학과	010-3435-4336
	장효정	2000.09.16.	중앙대학교	응용통계학과	010-6456-2495
주제명	투자 성향에 따른 고객 분류 및 추천				

분석 및 개발 모형

1. 투자 성향을 파악할 수 있는 지표 선정 및 EDA

국내와 해외 주식 거래 내역의 체결 건수를 합한 총거래횟수(고객별 체결 건수의 count)를 선정했다. 주식의 거래횟수가 많다는 것은 그만큼 활발하게 주식 활동을 한다는 의미로 해석했다. 이는 현재와 같은 하락장에서 적극적으로 투자를 진행하는지 혹은 주식을 보유만 하고 있는지 등의 투자 성향을 파악할 수 있다고 예상했다. 총거래횟수에 대한 EDA를 진행한 결과, 이상치로 판단되는 고객 분포가 존재했고 이상치의 특징을 따로 파악하고자 VIP 그룹을 생성했다.

이때, 총거래횟수와 총거래가격을 상관계수는 0.78로 매우 높은 양의 상관관계를 가짐을 확인하였다. 총거래횟수와 총거래가격을 활용하여 상위 그룹을 VIP로 분류한 뒤에 고객 분석을 진행하였다.

2. 고객 분류(VIP/Normal 군집 내 분류)에 사용한 변수의 특징

- 자산의 변동성 : 총자산편차(자산의 표준편차)
- 투자 적극성 : 총거래횟수, 대출총합
- 투자의 신중함 : 주식대기시간(체결시간대 - 주문접수시간대)
- 하락장에서의 주식보유여부 : 체결기준통합잔고수량편차

3. 고객 분류 및 결과

VIP		
공통점	<ul style="list-style-type: none"> - 자산이 많음(3억원 이상까지 고루 분포) - 가족 성숙기_직장인이 대다수 - 대부분이 주식을 보유하고 있으며 주거래 시장은 거래소와 코스닥 	
차이점	Cluster 0	<ul style="list-style-type: none"> - CMA와 퇴직연금의 비중이 존재 - 고객 등급 그린 & 블루 - 보유 주식 월 평균 잔고가 500만원 이상 - 거래유형은 general & swing trader - 주식대기시간이 vip 중에서 긴 편

	Cluster 1	<ul style="list-style-type: none"> - CMA와 신용대출의 비중이 존재 - 고객 등급 로얄 & 그린 / 충성도 최상 - 주로 남성 - 수익금액이 많음 - 총 투자 기간이 매우 긴 편 - 계좌 최초 개설이 1990년부터 꾸준히 있음 - 거의 모든 사람들이 제조 관련 주식 보유 - 보유 주식 월평균 잔고가 3천만원 이상 - 거래 유형은 day & swing trader --> 가장 공격적이고 변동이 큰 투자를 함
--	-----------	--

Normal		
공통점		<ul style="list-style-type: none"> - 자산과 수익이 낮음 - 사회초년생이 대다수 - 대부분이 2020년에 계좌 최초 개설 - 거의 대부분이 비매매
차이점	Cluster 2	<ul style="list-style-type: none"> - 해외 주식을 많이 함 - 투자기간이 6개월 미만
	Cluster 3	<ul style="list-style-type: none"> - 고객 자산이 normal중에 가장 높은 편 - 투자기간은 1-3년 - 주식대기시간이 normal 중에서 가장 긴 편
	Cluster 4	<ul style="list-style-type: none"> - 여성이 많음 - 25세 미만이 대다수로 가장 어린 집단 - 전부 stockholder

사용 외부데이터

1. 코스피_주요주가지수.csv : 2022년 1월에서 6월은 하락장이었기에 평균적으로 코스피보다 자산 변동률이 높은 사람을 확인하고자 사용하였다.

출처 : https://kosis.kr/statHtml/statHtml.do?orgId=343&tblId=DT_343_2010_S0029

2. 섹터(sector) 크롤링 : 섹터는 fnguide의 WICS(WISE Industry Classification Standard) sector Index의 대분류를 사용하였고, 이는 추후에 해외 데이터까지 확장 시 유용한 종목 구분이 될 것이다.

서비스 기획 아이디어

투자 성향 별로 분류한 군집의 특성에 따른 추천 서비스가 가능하다. 기존 고객의 경우 자주 접속하는 시간대에 주 거래 종목(sector)을 추천할 수 있다. 신규 고객의 경우 가입 시 고객 정보 및 투자 성향을 종합적으로 고려하여 비슷한 군집의 고객의 투자 방향성을 제안할 수 있다.