

Dataset Info.

****티커(Ticker)****는 주식 시장에서 특정 회사나 금융 상품을 식별하기 위해 사용하는 **고유한 문자 코드**입니다. 각 기업이나 금융 자산은 거래소에서 유일한 티커코드를 가진다.

1. **문자 조합**: 티커는 보통 **1~5개의 대문자**로 구성됩니다. 예를 들어, "AAPL"은 애플(Apple Inc.)을, "TSLA"는 테슬라(Tesla Inc.)를 나타냅니다.
2. **거래소별 다름**: 동일한 기업이라도 거래소에 따라 다른 티커코드를 가질 수 있습니다. 예를 들어, **Altria Group**는 뉴욕 증권거래소(NYSE)에서는 "MO"로, NASDAQ에서는 "MO"로 표기됩니다.
3. **유용성**: 티커코드는 투자자들이 빠르게 정보를 찾고 거래할 수 있도록 도와줍니다. 각 기업의 주가, 거래량, 재무정보 등을 쉽게 조회할 수 있습니다.
4. **다양한 자산군**: 주식뿐만 아니라, ETF(상장지수펀드), 채권, 원자재 등 다양한 금융 상품에도 티커가 사용됩니다.

[ETF란 무엇인가요? ETF 주식 차이점, ETF 상품 고르는 법 \(tistory.com\)](#)

NH_CONTEST_NW_FC_STK_IEM_IFO.csv

- 설명: 8월 27일 기준 해외종목 정보.
- TCK_IEM_CD: 티커종목코드
- FC_SEC_KRL_NM: 외화증권한글명
- FC_SEC_ENG_NM: 외화증권영문명
- STK ETF_DIT_CD: 주식/ETF구분코드
- LTG_TOT_STK_QTY: 상장주식총수량
- FC_MKT_DIT_CD: 외화시장구분코드
- CO_ADR: 회사주소
- WEB_ADR: 웹주소
- BTP_CFC_NM: 업종분류명
- CEO_NM: CEO명
- ENG_UTK_OTL_CTS: 영문사업개요내용
- SER_CFC_NM: 섹터분류명
- IDS_NM: 산업명
- MKT_PR_TOT_AMT: 시가총액

NH_CONTEST_STK_DT_QUT.csv

- 설명: 2024년 05월 28일부터 2024년 08월 27일까지
NH_CONTEST_NW_FC_STK_IEM_IFO.csv에 있는 종목 일자별 시세 정보
- BSE_DT: 거래일자
- TCK_IEM_CD: 티커종목코드
- IEM_ONG_PR: 종목시가
- IEM_HI_PR: 종목고가
- IEM_LOW_PR: 종목저가
- IEM_END_PR: 종목종가
- BF_DD_CMP_IND_PR: 전일대비증감가격
- BF_DD_CMP_IND_RT: 전일대비증감율
- ACL_TRD_QTY: 누적거래수량
- TRD_CST: 거래대금 = 거래된 주식 수 x 주식 가격
- SLL_CNS_SUM_QTY: 매도체결합계수량 = 특정 기간 동안 매도 주문이 실제로 체결된 총 주식 수량
- BYN_CNS_SUM_QTY: 매수체결합계수량 = 특정 기간 동안 매수 주문이 실제로 체결된 총 주식 수량
- SBY_BSE_XCG_RT: 환율

NH_CONTEST_NHDATA_STK_DD_IFO.csv

- 설명: 2024년 05월 28일부터 2024년 08월 27일까지 NH데이터 기반 주식 일별 정보.
- BSE_DT: BSE_DT
- TCK_IEM_CD: 티커종목코드
- TOT_HLD_ACT_CNT: 총보유계좌수 : 주식을 보유하고 있는 계좌 수. 한 투자자가 여러 계좌에서 같은 주식을 보유하고 있어도 계좌 수는 1로 계산
- TOT_HLD_QTY: 총보유수량 : 주식의 총 보유 수량. 한 계좌에서 100주, 다른 계좌에서 200주를 보유하고 있다면 총보유수량은 300주
- TCO_AVG_HLD_QTY: 당사평균보유수량 : 총보유수량 / 총보유계좌수
-> 투자자들이 평균적으로 보유하고 있는 주식의 수량
- TCO_AVG_HLD_WHT_RT: 당사평균보유비중비율 : 투자자들이 평균적으로 보유하고 있는 주식의 비율. (당사평균보유수량 / 총주식수) x 100
- TCO_AVG_EAL_PLS: 당사평균평가손익 : 투자자가 보유하고 있는 자산의 현재 가치와 해당 자산의 매입가 간의 차이
- TCO_AVG_PHS UIT_PR: 당사평균매입단가 : 총 매입금액 / 총 매입수량
- TCO_AVG_PFT_RT: 당사평균수익율 : (총 수익 / 총 투자금) x 100
- TCO_AVG_HLD_TE_DD_CNT: 당사평균보유기간일수
- DIST_HNK_PCT10_NMV: 분포상위10퍼센트수치
- DIST_HNK_PCT30_NMV: 분포상위30퍼센트수치
- DIST_HNK_PCT50_NMV: 분포상위50퍼센트수치
- DIST_HNK_PCT70_NMV: 분포상위70퍼센트수치
- DIST_HNK_PCT90_NMV: 분포상위90퍼센트수치
- BSE_END_PR: 기준종가 : 특정 날짜의 종가를 기준으로 하여 이후의 가격 변동을 비교하는 데 사용되는 지표. 기준종가가 10,000원, 현재 가격이 12,000원이라면, 해당 주식은 20% 상승한 것으로 평가함.
- LSS_IVO_RT: 손실투자자비율
- PFT_IVO_RT: 수익투자자비율
- IFW_ACT_CNT: 신규매수계좌수
- OFW_ACT_CNT: 전량매도계좌수 : 보유하고 있는 모든 주식을 매도한 계좌의 총 수
- VW_TGT_CNT: 종목조회건수

- RGS_TGT_CNT: 관심종목등록건수

NH_CONTEST_NHDATA_IFW_OFW_IFO.csv

- 설명: 2024년 05월 28일부터 2024년 08월 27일까지 유입/유출 종목 데이터
- 해외종목의 유입/유출으로 한정, 최대 TOP5까지 제공
- BSE_DT: 기준일자
- TCK_IEM_CD: 티커종목코드
- IFW_OFW_DIT_CD: 유입/유출구분코드
- IFW_OFW_TCK_CD: 유입/유출티커코드
- IFW_OFW_AMT_WHT_RT: 유입/유출금액비중 : (유입, 유출금액 / 총 거래 금액) X 100
- IFW_OFW_RNK: 유입/유출랭크

NH_CONTEST_NHDATA_CUS_TP_IFO.csv

- 설명: 2024년 05월 28일부터 2024년 08월 27일까지 NH데이터 기반 고객 보유 정보.
- BSE_DT: 기준일자
- TCK_IEM_CD: 티커종목코드
- CUS_CGR_LLF_CD: 고객구성대분류코드
- CUS_CGR_MLF_CD: 고객구성중분류코드
- CUS_CGR_ACT_CNT_RT: 고객구성계좌수비율 : (특정 고객의 보유 계좌 수 / 전체 고객 수) x 100
- CUS_CGR_IVS_RT: 고객구성투자비율 : (고객의 총 투자 금액 / 전체 고객 자산) x 100

고객구성계좌수비율 (CUS_CGR_ACT_CNT_RT)

이 비율은 고객이 평균적으로 얼마나 많은 계좌를 보유하고 있는지를 나타냅니다. 예를 들어, 이 비율이 20%라면, 전체 고객 수 대비 20%의 고객이 AAL에 대한 계좌를 보유하고 있다는 의미입니다. 이는 해당 종목에 대한 고객의 관심이나 이용도를 나타내며, 고객이 다양한 금융 상품에 참여하고 있는지를 평가할 수 있는 중요한 지표입니다.

고객구성투자비율 (CUS_CGR_IVS_RT)

이 비율은 고객의 총 투자 금액이 전체 고객 자산에서 차지하는 비율을 나타냅니다. 만약 이 비율이 15%라면, AAL에 대한 고객의 투자 금액이 전체 고객 자산의 15%를 차지한다는 의미입니다. 이는 고객들이 AAL에 대해 얼마나 적극적으로 투자하고 있는지를 보여주며, 해당 종목에 대한 시장의 신뢰도와 투자 성향을 평가하는 데 유용합니다.

NH_CONTEST_DATA ETF_HOLDINGS.csv

- 설명: 2024년 8월 27일 기준 ETF의 구성 종목.
- ETF_TCK_CD: 대상 ETF 티커
- TCK_IEM_CD: ETF 개별 구성 종목 티커
- MKT_VLU: 보유 종목의 가치 (USD)
- FC_SEC_ENG_NM: 보유 종목의 영문명
- FC_SEC_KRL_NM: 보유 종목의 한글명
- STK_QTY: 보유 종목의 주수 (주)
- WHT_PCT: 보유 종목의 비중 (%)
- SEC_TP: 보유 종목의 타입 (ST: 주식, EF: ETF, EN: ETN, SSEF: Single-Stock ETF)

NH_CONTEST_DATA HISTORICAL_DIVIDEND.csv

- 설명: 2022년 8월 27일부터 2024년 8월 27일까지 ETF 배당 내역.
- ETF_TCK_CD: 대상 ETF 티커
- EDIV_DT: 배당락일 (YYYYMMDD) : 배당금을 받을 권리가 사라지는 날을 의미합니다. 즉, 배당락일 전날까지 주식을 보유해야 배당금을 받을 자격이 주어집니다. 배당락일 당일에 주식을 매수한 사람은 배당금을 받을 수 없습니다.
- DDN_AMT: 배당금
- AED_STKP_DD_N_AMT: 수정 배당금
- DDN_BSE_DT: 배당 기준일 (YYYYMMDD) : 기업이 배당금을 지급할 주주를 확정하는 날짜입니다. 즉, 이 날짜에 주주 명부에 이름이 올라가 있는 주주가 배당금을 받을 권리를 가집니다.
- DDN_PYM_DT: 지급일 (YYYYMMDD)
- PBA_DT: 공시일 (YYYYMMDD) : 공시일은 기업이 법적 또는 규제상의 이유로 **주요 경영 정보나 재무 사항**을 공개해야 할 때, 그 정보를 공시하는 날짜
- DDN_PYM_FCY_CD: 배당 주기 (Quarterly: 분기배당, Weekly: 주배당, Monthly: 월배당, SemiAnnual: 반기배당, Annual: 연배당, Other: 알 수 없음)

NH_CONTEST ETF_SOR_IFO.csv

- 설명: 2024년 05월 28일부터 2024년 08월 27일까지 ETF 점수 정보.

$$Z\text{점수? } z = \frac{x - \mu}{\sigma}$$

- BSE_DT: 거래일자
- ETF_TCK_CD: 대상 ETF 티커
- MM1_TOT_PFT_RT: 1개월총수익율
- MM3_TOT_PFT_RT: 3개월총수익율
- YR1_TOT_PFT_RT: 1년총수익율
- ETF_SOR: ETF점수
- ETF_Z_SOR: ETFZ점수
- Z_SOR_RNK: Z점수순위
- ACL_PFT_RT_Z_SOR: 누적수익율Z점수
누적수익률 = $\{(\text{최종 투자금액} / \text{초기 투자금액}) - 1\} \times 100$

- IFO_RT_Z_SOR: 정보비율Z점수

정보 비율 : 포트폴리오의 초과 수익률(벤치마크 대비 성과)과 그 성과의 변동성을 비교하여, 포트폴리오 관리자의 능력을 평가하는 데 사용되는 지표입니다. 주로 투자 성과 분석에서 활용되며, 포트폴리오가 벤치마크보다 얼마나 더 나은 성과를 냈는지, 그리고 그 성과가 얼마나 일관된지를 보여줍니다. 값이 높을수록 안정적인 수익을 달성한다.

양수? 포트폴리오의 성과 > 벤치마크
-> 초과수익률 / 초과수익률의 표준편차

벤치마크 대비 성과 : 투자한 포트폴리오나 자산의 수익률을 특정 기준과 비교하여 평가하는 것을 의미합니다. 여기서 **벤치마크**(Benchmark)는 투자 성과를 비교하기 위해 설정한 기준으로, 주로 시장의 평균적인 수익률이나 특정 지수(예: S&P 500, KOSPI, MSCI World Index 등)가 사용됩니다.

- SHPR_Z_SOR: 샤프지수Z점수

샤프지수(Sharpe Ratio)는 투자 성과를 평가할 때 사용하는 지표로, **투자가 얼마나 효율적으로 위험을 감수하면서 초과 수익을 냈는지**를 나타냅니다.

-> (포트폴리오 수익률 - 무위험 수익률) / 포트폴리오 수익률의 표준편차

높은 샤프지수: 높은 위험 대비 높은 초과 수익을 얻었다는 뜻. 즉, 위험을 감수한 만큼 성과가 좋았다는 의미.

낮은 샤프지수: 위험을 감수했음에도 불구하고, 그에 상응하는 충분한 초과 수익을 얻지 못했다는 의미.

음수 샤프지수: 포트폴리오가 무위험 자산보다도 낮은 수익률을 냈다는 것을 의미.

무위험 자산은 거의 손실이 발생하지 않는 매우 안정적인 자산입니다. 예를 들어, 정부가 발행한 국채는 정부의 신용을 기반으로 하므로 파산 위험이 거의 없어 보통 **무위험 수익률**로 간주됩니다.

투자자가 위험을 감수하고 주식이나 다양한 금융 자산에 투자하는 이유는, 더 높은 수익을 기대하기 때문입니다. 그런데 포트폴리오의 수익률이 무위험 자산보다 낮다

면, 투자자는 더 큰 위험을 감수하고도 그에 상응하는 보상을 받지 못한 상황인 것이죠. 이는 매우 부정적인 투자 결과로 해석될 수 있습니다.

- CRR_Z_SOR: 상관관계Z점수

- TRK_ERR_Z_SOR: 트래킹에러Z점수

투자 포트폴리오의 수익률이 **벤치마크 수익률과 얼마나 차이가 나는지**를 측정하는 지표입니다. 이를 통해 포트폴리오가 벤치마크를 얼마나 잘 따르고 있는지, 혹은 얼마나 벗어났는지를 확인할 수 있습니다. 주로 **액티브 펀드**와 같이 시장 수익률을 초과하려는 투자 전략을 평가할 때 자주 사용됩니다. **계산은 초과수익률의 표준편차로 구한다.**

낮은 트래킹 에러: 포트폴리오가 벤치마크를 **매우 잘 따르고 있다**는 뜻입니다. 즉, 벤치마크와의 수익률 차이가 크지 않다는 의미입니다. 일반적으로 패시브 펀드(지수를 그대로 추종하는 펀드)는 트래킹 에러가 낮습니다.

높은 트래킹 에러: 포트폴리오가 벤치마크와 수익률 차이가 **크다**는 뜻입니다. 이는 포트폴리오가 더 적극적으로 벤치마크와 다른 전략을 취하고 있으며, 액티브 펀드가 더 높은 초과 수익을 추구하려 할 때 나타납니다.

- MXDD_Z_SOR: 최대낙폭Z점수

최대낙폭은 투자 기간 동안 자산의 가치가 얼마나 크게 하락할 수 있는지를 보여주기 때문에, **리스크**를 측정하는 중요한 지표로 사용됩니다.

최대낙폭이 크면, 해당 자산이나 포트폴리오의 **변동성**이 크고, 위험이 높다고 평가할 수 있습니다. 반대로, 최대낙폭이 작으면 상대적으로 **안정적인 자산**이라고 판단할 수 있습니다.

계산은 $\{(\text{최고점} - \text{최저점}) / \text{최고점}\} \times 100$

- VTY_Z_SOR: 변동성Z점수

변동성 = 표준편차