3) What sectors and specific stocks have outperformed the S&P 500 index in each year?

SELECT

    sp.Company,

    sp.GICS\_Sector AS Sector,

    sp.GICS\_SubIndustry AS Industry,

    YEAR(CONCAT(sp.Year, '-', sp.Month, '-', sp.Day)) AS Year,

    ROUND((AVG(sp.Close - sp.Open) / AVG(si.Close - si.Open) - 1) \* 100, 2) AS Stock\_Price\_Change\_Percentage,

    ROUND(((si.Close - si.Open) / si.Open) \* 100, 2) AS Index\_Price\_Change\_Percentage

FROM sp500\_prices.sp500\_prices AS sp

INNER JOIN sp500\_index.sp500\_index AS si

    ON sp.Year = si.Year

    AND sp.Month = si.Month

    AND sp.Day = si.Day

WHERE

    sp.Year BETWEEN 2018 AND 2022

GROUP BY

    sp.Company,

    sp.GICS\_Sector,

    sp.GICS\_SubIndustry,

    YEAR(CONCAT(sp.Year, '-', sp.Month, '-', sp.Day))

HAVING

    AVG(sp.Close - sp.Open) / AVG(si.Close - si.Open) > 1

ORDER BY

    YEAR(CONCAT(sp.Year, '-', sp.Month, '-', sp.Day)) ASC,

    sp.GICS\_Sector ASC,

    sp.GICS\_SubIndustry ASC,

    Stock\_Price\_Change\_Percentage DESC;

This SQL code identifies the sectors and specific stocks that have outperformed the S&P 500 index in each year, answering question 3 of the project.

The code utilizes the "sp500\_prices" and "sp500\_index" datasets, comparing the returns of each stock and sector to the returns of the S&P 500 index. The SELECT statement retrieves the necessary columns for analysis, including the stock's name, GICS sector, GICS sub-industry, year, stock price change percentage, and index price change percentage.

Using an INNER JOIN, the code joins the two datasets on the date columns and filters the results to only include data from the years 2018 to 2022. The GROUP BY statement calculates the average stock price change percentage and index price change percentage for each unique combination of company, sector, sub-industry, and year.

The HAVING statement filters the results to only include groups where the average stock price change percentage is greater than 1, indicating outperformance of the index. Finally, the ORDER BY statement sorts the results by year, sector, sub-industry, and stock price change percentage in ascending order by year and sector and in descending order by stock price change percentage.