

Executive Summary

This project analyzes how S&P 500 companies financially recovered from the 2020 slowdown into 2021, using public financial performance from Kaggle. The dataset includes company-level metrics such as Total Revenue, Gross Profit, Operating Income, and Net Income for the years of 2020 and 2021.

The objective is to evaluate growth, profitability, and operational resilience, identifying which firms achieved the strongest recovery.

Introduction

For this project, a publicly available financial performance dataset named **Financial Data S&P 500 companies** from **Kaggle** was used, containing data on the financial results of multiple publicly traded companies. The dataset includes metrics such as *Total Revenue*, *Gross Profit*, *Net Income*, and *Operating Margin*, covering the years **2020 and 2021**.

The objective of this project is to analyze how companies recovered financially from the economic slowdown of 2020 into 2021. By comparing key financial indicators across two consecutive years, this analysis seeks to uncover which companies demonstrated resilience, growth, and profitability during the post-crisis recovery period.

The project follows the **CRISP-DM (Cross-Industry Standard Process for Data Mining)** methodology -- starting with data understanding and exploration in **Microsoft SQL Server**, followed by data preparation, modeling, and visualization in **Power BI**. This approach connects technical analysis with financial storytelling, turning company financial statements into clear insights about performance, growth, and efficiency.

The dataset originally contained quarterly results for multiple firms. To make comparison consistent, all financial metrics were aggregated into annual totals for 2020 and 2021 during the data preparation phase.

Key Questions

This analysis seeks to answer the following financial and analytical questions:

1. How did overall company performance evolve between 2020 and 2021?
2. Which companies experience the strongest post-2020 growth?
3. Which firms achieved the highest profitability in 2021?
4. Is there a relationship between company size (Revenue) and profitability (Net Income Margin)?
5. Which companies combine strong growth with strong profitability -- the ideal balance of performance?

By addressing these questions, this project aims to evaluate financial health, highlight top performers, and explore how companies bounced back after a challenging economic year.

Financial Structure & Key terms

Before beginning the data import and modeling process, it is essential to understand how the financial data is structured and how each column connects within a company's income statement. The following section outlines the logical flow from **Total Revenue** down to **Net Income**, illustrating the key steps that represent a company's profitability journey. It also included concise definitions for each financial term in the dataset, helping to clarify the role of every metric used later in the analysis and visualizations.

Total Revenue

– Cost of Revenue

= Gross Profit

– Research & Development (R&D)

– Selling, General & Administrative (SG&A)

= Operating Income (EBIT)

– Interest Expense

± Total Other Income / Net Expense

= Income Before Tax

– Income Tax Expense

= Net Income

↳ Net Income from Continuing Operations

↳ Net Income Applicable to Common Shares

Research Development (R&D) - money spent on developing new products, technologies, or processes. It appears on the income statement as an operating expense. A high R&Ds signals investments in the future, however it reduces short-term profits

Income Before Tax (IBT) - "Pre-Tax Income" or "Earnings Before Tax". Profit before paying income taxes. Great to show profitability before tax obligations, gives a picture of operational efficiency.

Net Income - bottom line profit after all expenses, interest, and taxes. Key indicator of financial health.

Selling General Administrative (SG&A) - Overhead costs to run the business: salaries, marketing, office rent, admin expenses, etc. It is part of the Operating Expenses.

Gross Profit - Profit after deducting the Cost of Revenue (COGS) but before other expenses. It shows how much profit a company makes from its core business operations before administrative or marketing expenses.

EBIT - Earnings Before Interest & Taxes. Profit from operations excluding interest and income tax expenses. Usually equivalent to Operating Income. Measure used to compare profitability across companies with different capital structures.

Operating Income - Profit from core operations only, excluding non-operating income or expenses. Good to show how efficiently a company runs its main business, without considering financing or investments.

Interest Expense - The cost of borrowing money - interest paid on debt or loans. Important for evaluating leverage and financial risk.

Income Tax Expense - Amount owed to the government in taxes for the period. Income Before Tax x Tax Rate.

Total Revenue - Total amount of money earned from sales and other business activities. Income statement's top line. Usually a growth in total revenue indicates expansion in sales or pricing power.

Total Operating Expenses - the sum of all costs associated with running the business. Includes COGS, R&D, SG&A, depreciation....

Cost of Revenue (COGS) - direct costs of producing goods or services sold by the company. Materials, labor, and manufacturing.

Total Other Income Expense Net - non-operating items such as investment income, gain or losses from currency, or one time charges.

Net Income From Continuing Ops - net income generated from ongoing business operations, excluding discontinued segments or one time events.

Net Income Applicable to Common Shares - the portion of net income available to common shareholders after preferred dividends. What actually belongs to stockholders. |

Exploratory Data Analysis and Data Preparation

The exploratory data analysis and data preparation for this project were conducted in **Microsoft SQL Server Management Studio (SSMS)** using SQL queries. The objective of this stage was to understand the dataset, identify potential quality issues, and ensure that only accurate and relevant records would be used in the visualization and analysis phase in **Power BI**.

Data Import and Structure

Before running analyses, several data quality checks were performed to validate the dataset:

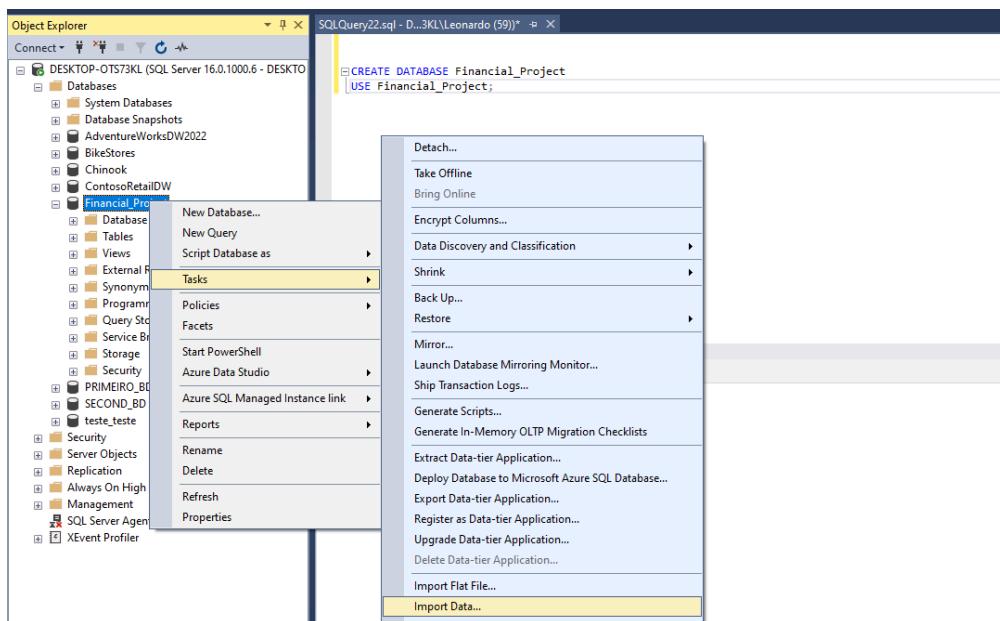
- **Missing values:** Checked for NULLs in revenue, profit, and expense columns.
- **Duplicate records:** Verified no duplicate combinations of (Ticker, Date).
- **Data type consistency:** Ensured all numeric columns were in the same format.
- **Negative values:** Reviewed columns like Total Revenue and Gross Profit for potential errors or reversals.

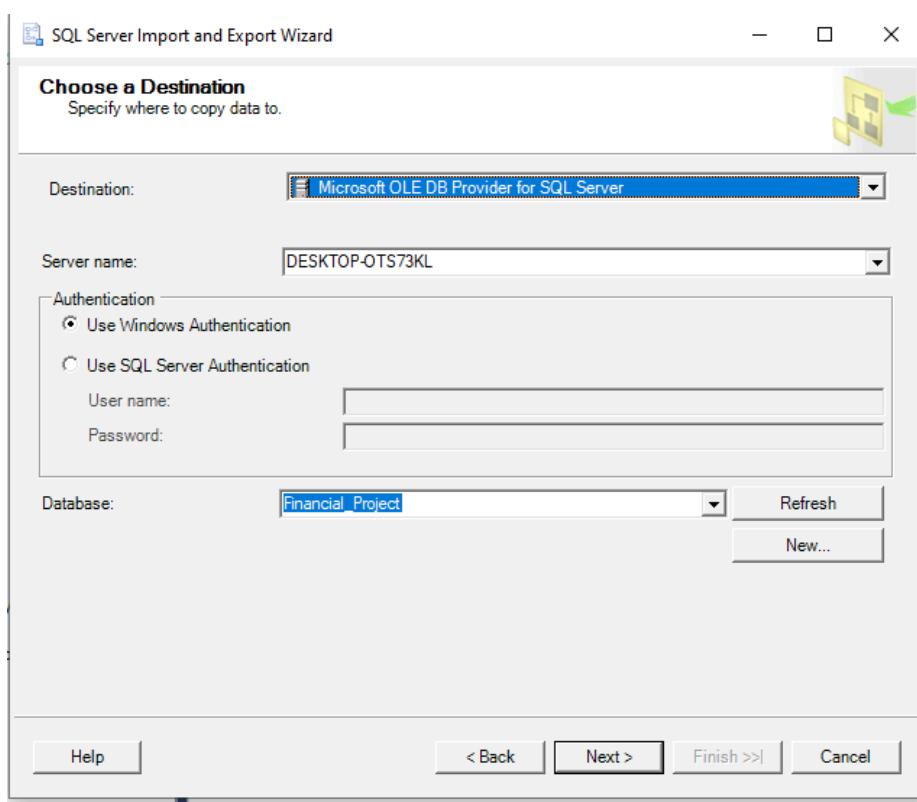
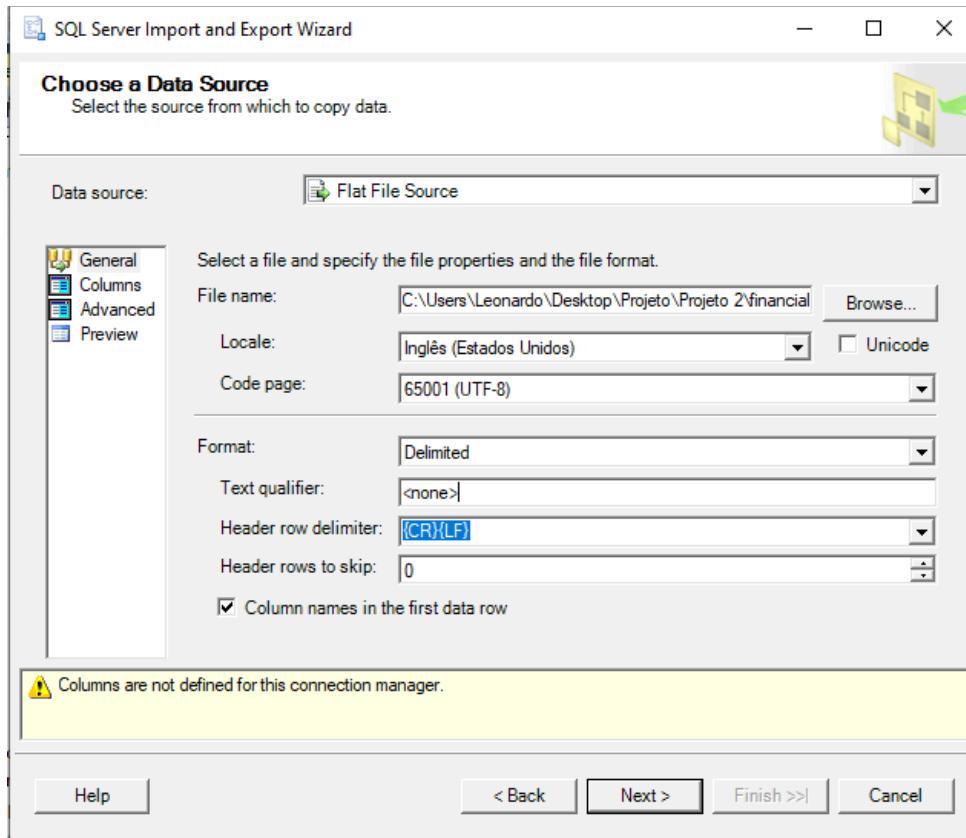
These checks confirmed that the dataset was suitable for financial analysis, with minimal missing or inconsistent records.

The original dataset was a CSV file named '*financial_data_sp500_companies.csv*', containing 19 columns and approximately 500 unique companies. Each company has four quarterly financial reports for the years **2020 and 2021**, including metrics such as:

- Total Revenue
- Gross Profit
- Net Income
- Operating Income
- Total Operating Expenses

During import, all columns were initially assigned a **VARCHAR(50)** data type. Since this format would limit numeric analysis, the financial columns were later converted to **FLOAT** using the **ALTER TABLE** command. This ensured accurate calculations and aggregations for year-over-year and comparative analyses.





FinancialData
[Column 0]
date
firm
Ticker
[Research Development]
[Income Before Tax]
[Net Income]
[Selling General Administrative]
[Gross Profit]
Ebit
[Operating Income]
[Interest Expense]
[Income Tax Expense]
[Total Revenue]
[Total Operating Expenses]
[Cost Of Revenue]
[Total Other Income Expense Net]
[Net Income From Continuing Ops]
[Net Income Applicable To Common Shares]

Data Exploration

After importing the dataset, an initial exploration was conducted to understand the data and confirm its quality:

1. Previewing Data:

The first few rows of the table were inspected to verify that the import was successful and that all columns were correctly mapped:

```

USE Financial_Project;
SELECT TOP 10
*
FROM
[financial data sp500 companies]

```

Results

	date	firm	Ticker	Research Development	Income Before Tax	Net Income	Selling General Administrative	Gross Profit	Ebit	Operating Income	Interest Expense
2021-09-30	3M	MMM	482000000.0	1761000000.0	1434000000.0	1739000000.0	4089000000.0	1868000000.0	1868000000.0	-117000000.0	
2021-06-30	3M	MMM	514000000.0	1940000000.0	1524000000.0	1666000000.0	4231000000.0	2051000000.0	2051000000.0	-121000000.0	
2021-03-31	3M	MMM	520000000.0	1946000000.0	1624000000.0	1720000000.0	4327000000.0	2087000000.0	2087000000.0	-132000000.0	
2020-12-31	3M	MMM	449000000.0	1706000000.0	1389000000.0	1814000000.0	4246000000.0	1983000000.0	1983000000.0	-141000000.0	
2021-09-30	3M	MMM	482000000.0	1761000000.0	1434000000.0	1739000000.0	4089000000.0	1868000000.0	1868000000.0	-117000000.0	
2021-06-30	3M	MMM	514000000.0	1940000000.0	1524000000.0	1666000000.0	4231000000.0	2051000000.0	2051000000.0	-121000000.0	
2021-03-31	3M	MMM	520000000.0	1946000000.0	1624000000.0	1720000000.0	4327000000.0	2087000000.0	2087000000.0	-132000000.0	
2020-12-31	3M	MMM	449000000.0	1706000000.0	1389000000.0	1814000000.0	4246000000.0	1983000000.0	1983000000.0	-141000000.0	
2021-09-30	A. O. Smith	AOS		1664000000.0	1316000000.0	1743000000.0	3403000000.0	1660000000.0	1660000000.0	-1000000.0	
0	2021-06-30	A. O. Smith	AOS		1513000000.0	1182000000.0	1698000000.0	3214000000.0	1516000000.0	1516000000.0	-900000.0
Income Tax Expense	Total Revenue	Total Operating Expenses	Cost Of Revenue	Total Other Income Expense Net	Net Income From Continuing Ops	Net Income Applicable To Common Shares					
324000000.0	8942000000.0	7074000000.0	4853000000.0	-107000000.0	1437000000.0	1434000000.0					
415000000.0	8950000000.0	6899000000.0	4719000000.0	-111000000.0	1525000000.0	1524000000.0					
319000000.0	8851000000.0	6764000000.0	4524000000.0	-141000000.0	1627000000.0	1624000000.0					
316000000.0	8583000000.0	6600000000.0	4337000000.0	-277000000.0	1390000000.0	1389000000.0					
324000000.0	8942000000.0	7074000000.0	4853000000.0	-107000000.0	1437000000.0	1434000000.0					
415000000.0	8950000000.0	6899000000.0	4719000000.0	-111000000.0	1525000000.0	1524000000.0					
319000000.0	8851000000.0	6764000000.0	4524000000.0	-141000000.0	1627000000.0	1624000000.0					
316000000.0	8583000000.0	6600000000.0	4337000000.0	-277000000.0	1390000000.0	1389000000.0					
34800000.0	914600000.0	748600000.0	574300000.0	400000.0	131600000.0	131600000.0					
33100000.0	859800000.0	708200000.0	538400000.0	-300000.0	118200000.0	118200000.0					

2. Checking Original Data Types:

After import, all columns were assigned **VARCHAR(50)**. This was confirmed by inspecting the table schema.

o.financial data sp500 companies

Columns

- Column 0 (varchar(50), null)
- date (varchar(50), null)
- firm (varchar(50), null)
- Ticker (varchar(50), null)
- Research Development (varchar(50), null)
- Income Before Tax (varchar(50), null)
- Net Income (varchar(50), null)
- Selling General Administrative (varchar(50), null)
- Gross Profit (varchar(50), null)
- Ebit (varchar(50), null)
- Operating Income (varchar(50), null)
- Interest Expense (varchar(50), null)
- Income Tax Expense (varchar(50), null)
- Total Revenue (varchar(50), null)
- Total Operating Expenses (varchar(50), null)
- Cost Of Revenue (varchar(50), null)
- Total Other Income Expense Net (varchar(50), null)
- Net Income From Continuing Ops (varchar(50), null)
- Net Income Applicable To Common Shares (varchar

EXEC sp_help 'financial data sp500 companies'

Results

Name	Owner	Type	Created_datetime
1 financial data sp500 companies	dbo	user table	2025-10-24 14:06:26.280
<hr/>			
Column_name	Type	Computed	Length
1 Column 0	varchar	no	50
2 date	varchar	no	50
3 firm	varchar	no	50
4 Ticker	varchar	no	50
5 Research Development	varchar	no	50
6 Income Before Tax	varchar	no	50
7 Net Income	varchar	no	50
8 Selling General Admini...	varchar	no	50
<hr/>			
Identity	Seed	Increment	Not For Replication

3. Counting Rows and Companies:

The total number of records and unique companies were calculated to ensure dataset completeness:

```

[=] SELECT
    COUNT(*) AS total_rows,
    COUNT(DISTINCT firm) AS total_companies
  FROM
    FinancialData

```

total_rows	total_companies
2012	502

4. Null Value Checks:

All financial columns were checked for missing values to guarantee data completeness:

```

[=] USE Financial_Project;
[=] SELECT
    SUM(CASE WHEN date IS NULL THEN 1 ELSE 0 END) AS null_date,
    SUM(CASE WHEN firm IS NULL THEN 1 ELSE 0 END) AS null_firm,
    SUM(CASE WHEN Ticker IS NULL THEN 1 ELSE 0 END) AS null_ticker,
    SUM(CASE WHEN [Research Development] IS NULL THEN 1 ELSE 0 END) AS null_rd,
    SUM(CASE WHEN [Income Before Tax] IS NULL THEN 1 ELSE 0 END) AS null_ibt,
    SUM(CASE WHEN [Net Income] IS NULL THEN 1 ELSE 0 END) AS null_netincome,
    SUM(CASE WHEN [Selling General Administrative] IS NULL THEN 1 ELSE 0 END) AS null_sga,
    SUM(CASE WHEN [Gross Profit] IS NULL THEN 1 ELSE 0 END) AS null_gp,
    SUM(CASE WHEN Ebit IS NULL THEN 1 ELSE 0 END) AS null_ebit,
    SUM(CASE WHEN [Operating Income] IS NULL THEN 1 ELSE 0 END) AS null_oi,
    SUM(CASE WHEN [Interest Expense] IS NULL THEN 1 ELSE 0 END) AS null_ie,
    SUM(CASE WHEN [Income Tax Expense] IS NULL THEN 1 ELSE 0 END) AS null_ite,
    SUM(CASE WHEN [Total Revenue] IS NULL THEN 1 ELSE 0 END) AS null_totalrev,
    SUM(CASE WHEN [Total Operating Expenses] IS NULL THEN 1 ELSE 0 END) AS null_toe,
    SUM(CASE WHEN [Cost Of Revenue] IS NULL THEN 1 ELSE 0 END) AS null_cor,
    SUM(CASE WHEN [Total Other Income Expense Net] IS NULL THEN 1 ELSE 0 END) AS null_toiet,
    SUM(CASE WHEN [Net Income From Continuing Ops] IS NULL THEN 1 ELSE 0 END) AS null_netops,
    SUM(CASE WHEN [Net Income Applicable To Common Shares] IS NULL THEN 1 ELSE 0 END) AS null_netcommon
  FROM
    [financial data sp500 companies]

```

null_date	null_firm	null_ticker	null_rd	null_ibt	null_netincome	null_sga	null_gp	null_ebit	null_oi	null_ie	null_ite	null_totalrev	null_toe	null_cor	null_toiet	null_netops	null_netcommon
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

No missing values were found.

5. Duplicate Check & Cleaning:

A check for duplicates (same firm, ticker, date) revealed that some companies, such as **3M (MMM)**, and **Advance Auto Parts (AAP)** had duplicate entries for the same quarter. The duplicate rows were removed so that each company has only one record per reporting date.

```

SELECT
    firm,
    Ticker,
    date,
    COUNT(*) AS cnt
FROM
    FinancialData
GROUP BY
    firm,
    Ticker,
    date
HAVING
    COUNT(*) > 1

WITH DUPLICATES
AS
(
    SELECT
        *
    ,   ROW_NUMBER() OVER (PARTITION BY firm, Ticker, date ORDER BY date) AS RRN
    FROM
        FinancialData
)
DELETE FROM DUPLICATES WHERE RRN > 1

```

6. Negative Values Awareness:

Some companies reported negative Net Income or Gross Profit in certain quarters. These values were retained as they reflect real financial losses, which are important for profitability analysis.

7. Quarterly Coverage Check:

Most companies had 4 annual reports with 4 different dates as listed on the Data Card on Kaggle. However, one company was missing a quarter, which was noted for transparency.

```

SELECT
    firm,
    COUNT(*) AS REPORT_CNT
FROM
    FinancialData
GROUP BY
    firm
HAVING
    COUNT(*) <> 4
ORDER BY
    REPORT_CNT DESC

```

A screenshot of a SQL query results window. The window has tabs for 'Results' and 'Messages'. The 'Results' tab is selected and shows a single row of data in a table format. The table has two columns: 'firm' and 'REPORT_CNT'. The 'firm' column contains 'Advance Auto Parts' and the 'REPORT_CNT' column contains '3'.

firm	REPORT_CNT
Advance Auto Parts	3

8. Ticker and Firm Verification:

All tickers were correctly matched to their respective companies:

```
SELECT
    firm,
    COUNT(DISTINCT Ticker) AS tickers
FROM
    FinancialData
GROUP BY
    firm
HAVING
    COUNT(DISTINCT Ticker) > 1
```

A screenshot of a SQL query results window. The window has a header bar with a dropdown menu and a search icon. Below the header are two tabs: "Results" and "Messages", with "Results" being the active tab. The results table has two columns: "firm" and "tickers". A single row is displayed, showing "firm" and "tickers" both as empty values.

firm	tickers

9. Data Type Conversion:

All numeric columns were converted from **VARCHAR** to **FLOAT** to allow for calculations such as sums, averages, and growth percentages:

```
|  
| USE Financial_Project  
|  
+ALTER TABLE FinancialData ALTER COLUMN firm NVARCHAR(100);  
ALTER TABLE FinancialData ALTER COLUMN date DATE;  
ALTER TABLE FinancialData ALTER COLUMN [Research Development] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Income Before Tax] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Net Income] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Selling General Administrative] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Gross Profit] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Ebit] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Operating Income] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Interest Expense] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Income Tax Expense] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Total Revenue] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Total Operating Expenses] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Cost Of Revenue] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Total Other Income Expense Net] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Net Income From Continuing Ops] FLOAT;  
ALTER TABLE FinancialData ALTER COLUMN [Net Income Applicable To Common Shares] FLOAT;
```

0 %

Messages

Commands completed successfully.

Completion time: 2025-10-24T15:22:21.0778315-03:00

```

SELECT
    COLUMN_NAME,
    DATA_TYPE
FROM
    INFORMATION_SCHEMA.COLUMNS
WHERE
    TABLE_NAME = 'FinancialData'

```

% ▾ ◀

Results Messages

	COLUMN_NAME	DATA_TYPE
	date	date
	firm	nvarchar
	Ticker	varchar
	Research Development	float
	Income Before Tax	float
	Net Income	float
	Selling General Administrative	float
	Gross Profit	float
)	Ebit	float
	Operating Income	float
	Interest Expense	float
	Income Tax Expense	float
	Total Revenue	float
	Total Operating Expenses	float
	Cost Of Revenue	float

10. Numeric Ranges:

The range of key numeric columns was checked to identify extreme values:

- Total Revenue: min = -552,600,00; max = 152,079,000,000; avg = 6,547,387,779.
- Extreme negative values exist but are retained as they may reflect real adjustments or losses.

```

SELECT
    MIN([Total Revenue]) AS MIN_REV ,
    MAX([Total Revenue]) AS MAX_REV,
    AVG([Total Revenue]) AS AVG_REV
FROM
    FinancialData

```

MIN_REV	MAX_REV	AVG_REV
-552600000	152079000000	6544777676,88092

SQL Analysis / Calculations

After the dataset was cleaned, validated, and prepared, SQL queries were used to extract key financial insights from the data.

1. Year-over-Year (YoY) Growth

Identify the top 10 companies with the highest year-over-year (YoY) revenue growth from 2020 to 2021.

```

WITH Growth
AS
(
    SELECT
        date,
        firm,
        [Total Revenue],
        LAG([Total Revenue]) OVER (PARTITION BY firm ORDER BY date) AS Last_year,
        ROUND([Total Revenue] - LAG([Total Revenue]) OVER (PARTITION BY firm ORDER BY date), 2) AS YOY_Difference,
        ROUND(
            (([Total Revenue] - LAG([Total Revenue]) OVER (PARTITION BY firm ORDER BY date))
            / NULLIF(LAG([Total Revenue]) OVER (PARTITION BY firm ORDER BY date), 0)
            ) * 100,
            2
        ) AS YOY_Revenue_Growth
    FROM
        FinancialData
)
SELECT TOP 10
    G.date,
    G.firm,
    G.Last_year,
    G.[Total Revenue],
    G.YOY_Difference,
    G.YOY_Revenue_Growth
FROM
    Growth AS G
WHERE
    date = (
        SELECT
            MAX(date)
        FROM
            FinancialData AS F2
        WHERE
            f2.firm = g.firm
    )
ORDER BY
    YOY_Revenue_Growth DESC

```

	date	firm	Last_year	Total Revenue	YOY_Difference	YOY_Revenue_Growth
1	2021-09-30	Royal Caribbean Group	50914000	456958000	406044000	797,51
2	2021-03-31	NRG Energy	202700...	8091000000	6064000000	299,16
3	2021-04-30	Intuit	157600...	4173000000	2597000000	164,78
4	2021-06-30	Modema	193700...	4354000000	2417000000	124,78
5	2021-06-30	Live Nation Entertain...	290609...	575946000	285337000	98,19
6	2021-06-30	Booking Holdings	114100...	2160000000	1019000000	89,31
7	2021-05-31	Carnival Corporation	26000000	49000000	23000000	88,46
8	2021-06-30	American Airlines Gro...	400800...	7478000000	3470000000	86,58
9	2021-06-30	United Airlines	322100...	5471000000	2250000000	69,85
10	2021-03-31	Kinder Morgan	311500...	5211000000	2096000000	67,29

2. Company Ranking by Total Revenue

Rank all companies based on their Total Revenue across both 2020 and 2021 combined.

```
WITH CompanyRevenue
AS
(
    SELECT
        firm,
        SUM([Total Revenue]) AS Total_Revenue_Both_Years
    FROM
        FinancialData
    GROUP BY
        firm
)
SELECT TOP 10
    RANK() OVER(ORDER BY Total_Revenue_Both_Years DESC) AS RNK,
    firm,
    Total_Revenue_Both_Years
FROM
    CompanyRevenue
ORDER BY
    RNK
```

% ▾

Results Messages

RNK	firm	Total_Revenue_Both_Years
1	Walmart	566145000000
2	Amazon	443298000000
3	Apple	347155000000
4	CVS Health	277285000000
5	UnitedHealth Group	272099000000
6	McKesson Corporation	238228000000
7	Alphabet (Class A)	220265000000
7	Alphabet (Class C)	220265000000
9	ExxonMobil	215879000000
10	AmerisourceBergen	204321190000

3. Yearly Financial Summary:

Summarize and compare total financial performance across both years (2020 and 2021). NULL appears for 2020 since no 2019 data exists. Quantified overall financial growth between 2020 and 2021, highlighting total gains in revenue, profitability, and margin improvements across the dataset.

```
WITH YearlyTotal
AS
(
    SELECT
        DATEPART(YEAR, date) AS 'YEAR',
        SUM([Total Revenue]) AS Revenue_Total,
        SUM([Gross Profit]) AS Total_Gross_Profit,
        SUM([Net Income]) AS Total_Net_Income
    FROM
        FinancialData
    GROUP BY
        DATEPART(YEAR, date)
)
SELECT
    YEAR,
    Revenue_Total,
    Total_Gross_Profit,
    Total_Net_Income,
    ROUND(Revenue_Total - LAG(Revenue_Total) OVER(ORDER BY YEAR), 2) AS Revenue_Change,
    ROUND(Total_Gross_Profit - LAG(Total_Gross_Profit) OVER(ORDER BY YEAR), 2) AS Gross_Profit_Change,
    ROUND(Total_Net_Income - LAG(Total_Net_Income) OVER(ORDER BY YEAR), 2) AS Net_Income_Change
FROM
    YearlyTotal
```

YEAR	Revenue_Total	Total_Gross_Profit	Total_Net_Income	Revenue_Change	Gross_Profit_Change	Net_Income_Change
2020	5594368341000	2131387400000	475062314000	NULL	NULL	NULL
2021	7541000456500	3016061020500	960653670000	1946632115500	884673620500	485591356000

Summary

An initial exploration, validation, and preparation phase was completed using the *Financial Data S&P 500 Companies* dataset in **Microsoft SQL Server**. The dataset contains 19 columns and approximately 500 unique companies. Each company includes **two available financial reports per year**, though the specific quarters vary. In 2020, most records correspond to **Q3 and Q4**, while in 2021, the coverage typically ranges from **Q1 to Q3**, with **Q4 data not reported** for several firms.

During the data preparation, the CSV file was imported through the SSMS import wizard and structured into a relational table. All columns were initially set to **VARCHAR(50)** and later converted to proper data types -- financial metrics to **FLOAT**, date to **DATE**, and company identifiers to **VARCHAR/NVARCHAR**.

Data quality checks were performed to confirm: successful import and column mapping, accurate data types for financial calculations, absence of null values, removal of duplicate

records, awareness of negative values for Net Income and Gross Profit, correct quarterly coverage, and validation of unique firm and ticker identifiers.

Numeric ranges were also verified to ensure realistic financial magnitudes --Total Revenue spanned from **-552.6M** to **152.1B**, with an average of approximately **6.5B USD**.

After validation, SQL analysis was conducted to extract key insights, including year-over-year revenue growth (YoY), company ranking by total revenue, and annual financial summaries. The dataset is now fully cleaned, verified, and ready for visualization and trend modeling in Power BI and R.

Power BI Data Preparation, Cleaning, and Modeling

1. Data Import

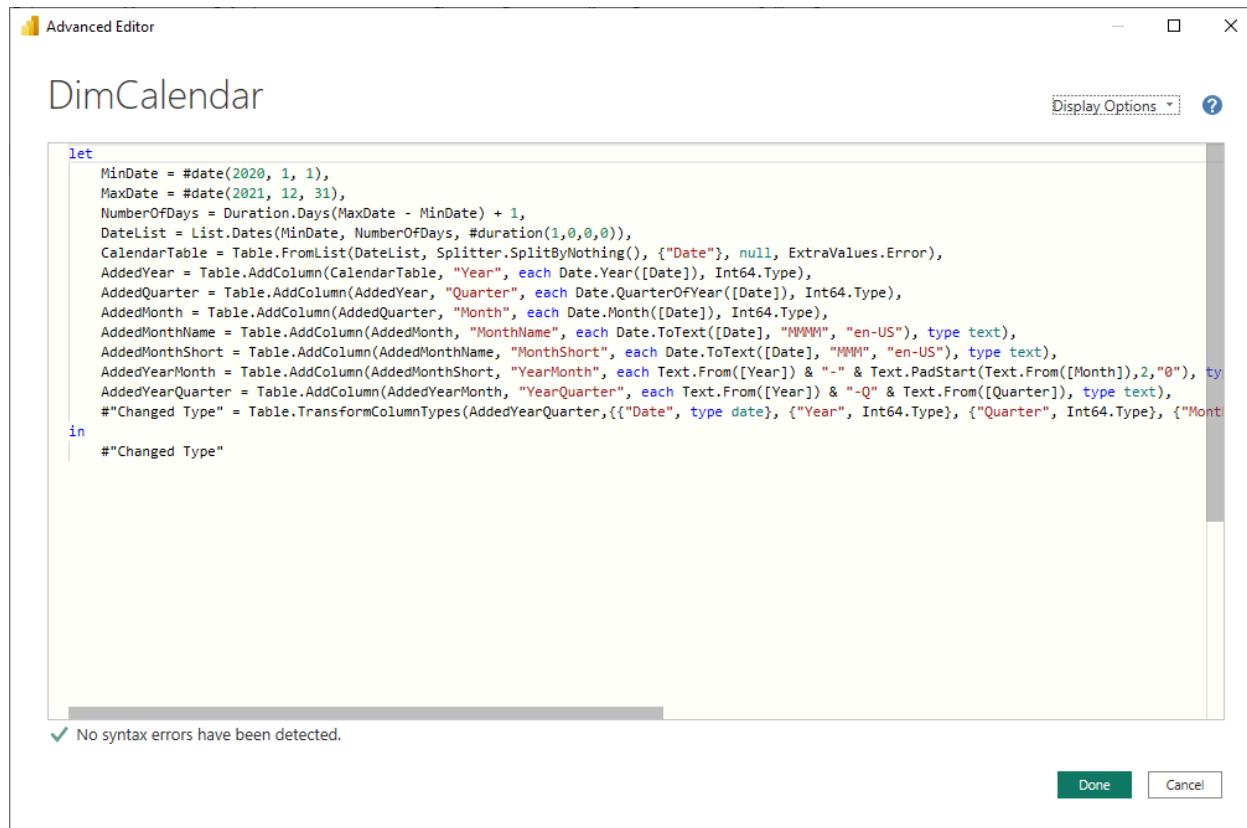
The cleaned financial dataset from SQL Server was imported into Power BI Desktop for visualization and further analysis. All columns were included. Column types were validated: **numeric columns (decimal number)**, **date column (date)**, and **text columns (firm (NVARCHAR, Ticker (VARCHAR))**.

The screenshot shows the Power BI Desktop interface with the 'FinancialData' table selected in the Data view. The table contains 27 rows of financial data with columns: date, firm, Ticker, 1.2 Research Development, 1.2 Income Before Tax, and 1.2 Net Income. The Properties pane on the right shows the table's name is 'FinancialData' and lists an applied step named 'Changed Type' under the Source and Navigation sections.

	date	A ^B _C firm	A ^B _C Ticker	1.2 Research Development	1.2 Income Before Tax	1.2 Net Income
1	30/09/2021	3M	MMM	482000000	1761000000	
2	31/03/2021	3M	MMM	520000000	1946000000	
3	31/12/2020	3M	MMM	449000000	1706000000	
4	30/06/2021	3M	MMM	514000000	1940000000	
5	30/09/2021	A. O. Smith	AOS	0	166400000	
6	30/06/2021	A. O. Smith	AOS	0	151300000	
7	31/03/2021	A. O. Smith	AOS	0	126100000	
8	31/12/2020	A. O. Smith	AOS	0	151900000	
9	30/09/2021	Abbott Laboratories	ABT	672000000	2493000000	
10	30/06/2021	Abbott Laboratories	ABT	654000000	1348000000	
11	31/03/2021	Abbott Laboratories	ABT	654000000	2043000000	
12	31/12/2020	Abbott Laboratories	ABT	643000000	2388000000	
13	30/09/2021	AbbVie	ABBV	1655000000	3688000000	
14	30/06/2021	AbbVie	ABBV	1784000000	1163000000	
15	31/03/2021	AbbVie	ABBV	1731000000	3867000000	
16	31/12/2020	AbbVie	ABBV	1821000000	-1507000000	
17	30/09/2021	Abiomed	ABMD	41041000	67271000	
18	30/06/2021	Abiomed	ABMD	36837000	-9350000	
19	31/03/2021	Abiomed	ABMD	32599000	73494000	
20	31/12/2020	Abiomed	ABMD	32105000	80735000	
21	31/08/2021	Accenture	ACN	0	1917805000	
22	31/05/2021	Accenture	ACN	0	2094001000	
23	28/02/2021	Accenture	ACN	0	1762443000	
24	30/11/2020	Accenture	ACN	0	1986867000	
25	30/09/2021	Activision Blizzard	ATVI	329000000	759000000	
26	30/06/2021	Activision Blizzard	ATVI	335000000	1002000000	
27	31/03/2021	Activision Blizzard	ATVI	353000000	765000000	

2. Calendar Table Creation

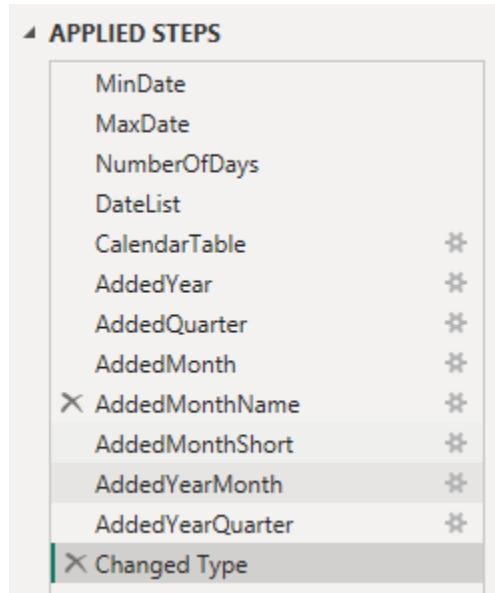
A Calendar Table was created using **Power Query's M language**, covering the full data range of the dataset (January 2020 - December 2021). The table also includes custom columns to support visualizations and time intelligence calculations.



The screenshot shows the Power Query Advanced Editor window. The title bar says "Advanced Editor". The main area contains the following M code:

```
let
    MinDate = #date(2020, 1, 1),
    MaxDate = #date(2021, 12, 31),
    NumberOfDays = Duration.Days(MaxDate - MinDate) + 1,
    DateList = List.Dates(MinDate, NumberOfDays, #duration(1,0,0,0)),
    CalendarTable = Table.FromList(DateList, Splitter.SplitByNothing(), {"Date"}, null, ExtraValues.Error),
    AddedYear = Table.AddColumn(CalendarTable, "Year", each Date.Year([Date]), Int64.Type),
    AddedQuarter = Table.AddColumn(AddedYear, "Quarter", each Date.QuarterOfYear([Date]), Int64.Type),
    AddedMonth = Table.AddColumn(AddedQuarter, "Month", each Date.Month([Date]), Int64.Type),
    AddedMonthName = Table.AddColumn(AddedMonth, "MonthName", each Date.ToText([Date], "MMMM", "en-US"), type text),
    AddedMonthShort = Table.AddColumn(AddedMonthName, "MonthShort", each Date.ToText([Date], "MMM", "en-US"), type text),
    AddedYearMonth = Table.AddColumn(AddedMonthShort, "YearMonth", each Text.From([Year]) & "-" & Text.PadStart(Text.From([Month]),2,"0"), type text),
    AddedYearQuarter = Table.AddColumn(AddedYearMonth, "YearQuarter", each Text.From([Year]) & "-Q" & Text.From([Quarter]), type text),
    #"Changed Type" = Table.TransformColumnTypes(AddedYearQuarter,{{"Date", type date}, {"Year", Int64.Type}, {"Quarter", Int64.Type}, {"Month", Int64.Type}, {"MonthName", type text}, {"MonthShort", type text}, {"YearMonth", type text}, {"YearQuarter", type text}})
in
    #"Changed Type"
```

At the bottom left, there is a green checkmark icon followed by the text "No syntax errors have been detected." On the right side, there are "Done" and "Cancel" buttons.



The screenshot shows the "APPLIED STEPS" pane. It lists the following steps:

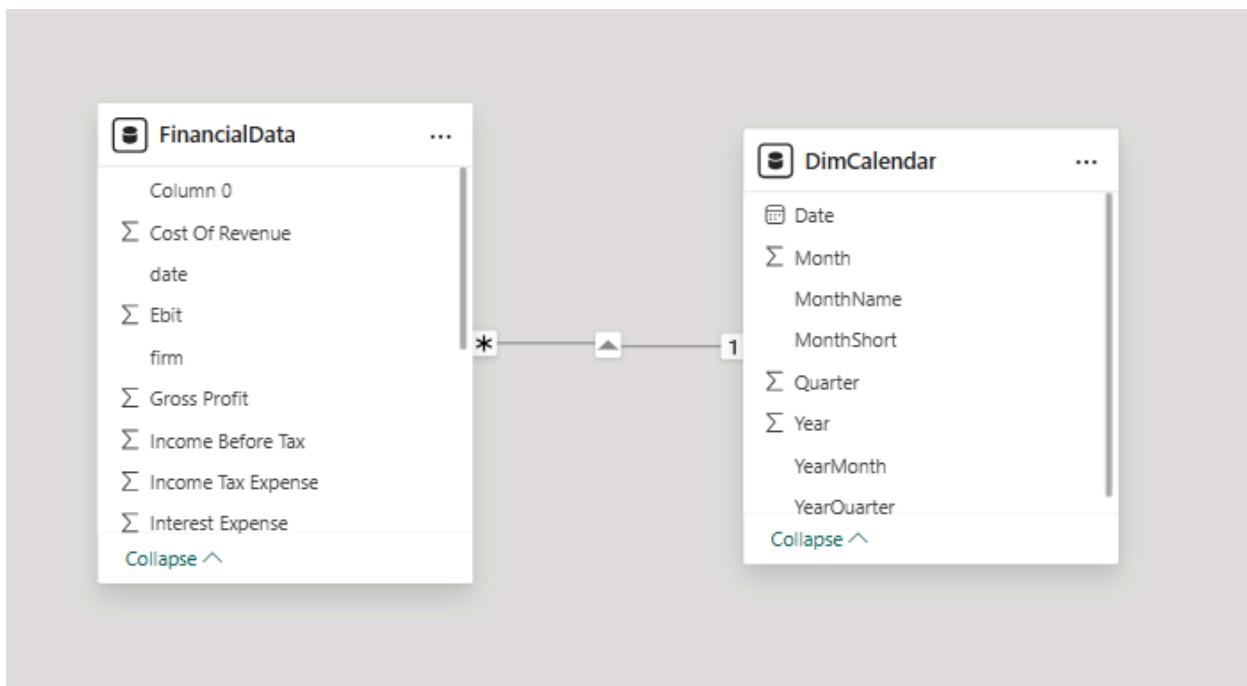
- MinDate
- MaxDate
- NumberOfDays
- DateList
- CalendarTable
- AddedYear
- AddedQuarter
- AddedMonth
- AddedMonthName
- AddedMonthShort
- AddedYearMonth
- AddedYearQuarter
- Changed Type

The "Changed Type" step is highlighted with a green border.

	1 ² ₃ Year	1 ² ₃ Quarter	1 ² ₃ Month	A ^B _C MonthName	A ^B _C MonthShort	A ^B _C YearMonth	A ^B _C YearQuarter
1	2020		1	1 January	Jan	2020-01	2020-Q1
2	2020		1	1 January	Jan	2020-01	2020-Q1
3	2020		1	1 January	Jan	2020-01	2020-Q1
4	2020		1	1 January	Jan	2020-01	2020-Q1
5	2020		1	1 January	Jan	2020-01	2020-Q1
6	2020		1	1 January	Jan	2020-01	2020-Q1
7	2020		1	1 January	Jan	2020-01	2020-Q1
8	2020		1	1 January	Jan	2020-01	2020-Q1
9	2020		1	1 January	Jan	2020-01	2020-Q1
10	2020		1	1 January	Jan	2020-01	2020-Q1
11	2020		1	1 January	Jan	2020-01	2020-Q1
12	2020		1	1 January	Jan	2020-01	2020-Q1
13	2020		1	1 January	Jan	2020-01	2020-Q1
14	2020		1	1 January	Jan	2020-01	2020-Q1
15	2020		1	1 January	Jan	2020-01	2020-Q1
16	2020		1	1 January	Jan	2020-01	2020-Q1
17	2020		1	1 January	Jan	2020-01	2020-Q1
18	2020		1	1 January	Jan	2020-01	2020-Q1
19	2020		1	1 January	Jan	2020-01	2020-Q1
20	2020		1	1 January	Jan	2020-01	2020-Q1
21	2020		1	1 January	Jan	2020-01	2020-Q1
22	2020		1	1 January	Jan	2020-01	2020-Q1
23	2020		1	1 January	Jan	2020-01	2020-Q1
24	2020		1	1 January	Jan	2020-01	2020-Q1
25	2020		1	1 January	Jan	2020-01	2020-Q1
26	2020		1	1 January	Jan	2020-01	2020-Q1
27	2020		1	1 January	Jan	2020-01	2020-Q1

3. Linking Tables

The Calendar Table was linked to the Financial Data table via the **Date column**, relationship **one-to-many (Calendar -> Financial Data)**, and cross-filter single direction.



Dashboard Construction

The first step in creating the dashboard was to set up DAX measures. These measures will help turn the data into clear numbers. Creating these measures will make it possible to analyze the dataset and also answer the main business questions with confidence. Each one has a specific role in the project, helping to understand the data and find useful insights.

DAX Measures:

The screenshot shows a list of DAX measures in a Power BI Data View. The measures are listed under a folder named '#Measures'. Each measure is preceded by a small icon representing its type (e.g., calculator for calculations).

- ✓ #Measures
- Gross Profit Margin %
- Growth-Profit Score
- Net Income
- Net Income Margin %
- Net Income YTD
- Operating Income
- Revenue Rank
- Revenue YoY
- Revenue YoY Delta
- Revenue YoY Delta %
- Revenue YTD
- Total Gross Profit
- Total Revenue

1. Total Revenue: Calculates total revenue.

```
1 Total Revenue = SUM(FinancialData[Total Revenue])
```

2. Total Gross Profit: Calculates total gross profit.

```
Total Gross Profit = SUM(FinancialData[Gross Profit])
```

3. Net Income: Calculates total net income.

```
Net Income = SUM(FinancialData[Net Income])
```

4. Operating Income: Calculates total operating income.

```
Operating Income = SUM(FinancialData[Operating Income])
```

5. Revenue YoY (Previous Year Revenue): Calculates total revenue for the previous year.

```
1 Revenue YoY =
2 VAR YOY =
3     CALCULATE(
4         [Total Revenue],
5         SAMEPERIODLASTYEAR(DimCalendar[Date])
6     )
7 RETURN
8 IF(
9     [Total Revenue],
10    YOY
11 )
```

6. Revenue YoY Delta: Calculates the difference in revenue to the prior year.

```
1 Revenue YoY Delta =
2 VAR YOYDELTA =
3     [Total Revenue] - [Revenue YoY]
4 RETURN
5 IF(
6     [Revenue YoY],
7     YOYDELTA
8 )
```

7. Revenue YoY Delta %: Calculates the percentage revenue growth year-over-year.

```
1 Revenue YoY Delta % =
2 DIVIDE(
3     [Revenue YoY Delta],
4     [Revenue YoY]
5 )
```

8. Net Income Margin %: Calculates the profitability relative to total revenue.

```
1 Net Income Margin % =
2 DIVIDE(
3     [Net Income],
4     [Total Revenue]
5 )
```

9. Gross Profit Margin %: Calculates the efficiency in generating gross profit relative to revenue.

```
1 Gross Profit Margin % =  
2 DIVIDE(  
3     [Total Gross Profit],  
4     [Total Revenue]  
5 )
```

10. Growth-Profit Score: Calculates the best overall performing companies.

```
1 Growth-Profit Score =  
2 [Revenue YoY Delta %] * [Net Income Margin %]
```

11. Revenue YTD: Calculates the cumulative revenue for the year.

```
1 Revenue YTD =  
2 TOTALYTD(  
3     [Total Revenue],  
4     DimCalendar[Date]  
5 )
```

12. Net Income YTD: Calculates the cumulative net income for the year.

```
1 Net Income YTD =  
2 TOTALYTD(  
3     [Net Income],  
4     DimCalendar[Date]  
5 )
```

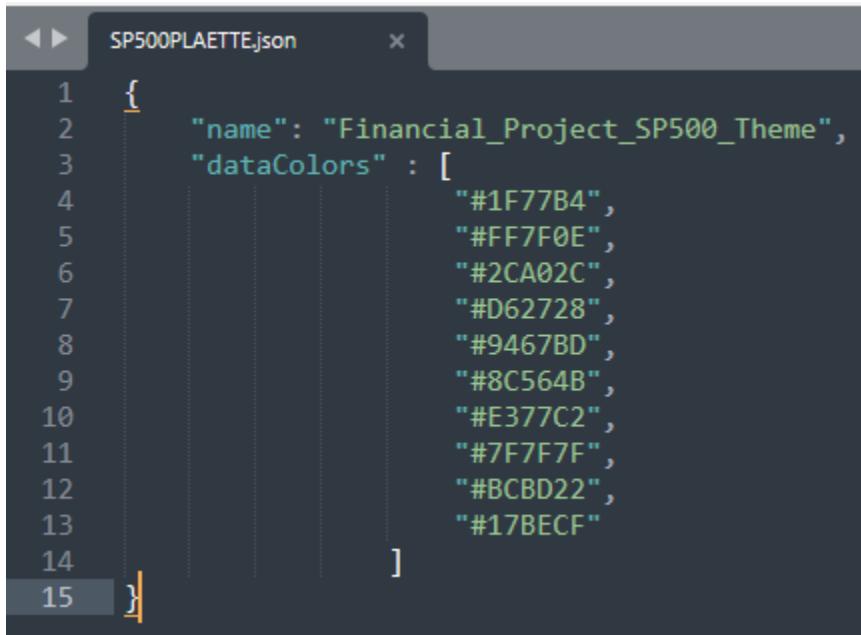
13. Rank by Total Revenue: Ranks the companies based on their total revenue over the selected period.

```
1 Revenue_Rank =
2 RANKX(
3     ALL(FinancialData[firm]),
4     [Total_Revenue],,
5     DESC,
6     Skip
7 )
```

Dashboard Design

After building the data model and importing the cleaned financial dataset into **Power BI**, the next step was to design a clear and professional dashboard layout. A soft off-white canvas was used as the main background to keep the interface clean, while a dark navy-blue sidebar provides contrast and houses the navigation menu.

A custom color theme was imported into **Power BI** via a **JSON file** to ensure visual consistency across all visuals. Accent colors were carefully selected to differentiate metrics. This ensures that charts, KPIs, and tables are easy to interpret at a glance while maintaining a professional, executive style.



```
SP500PLAETTE.json
```

```
1 {
2     "name": "Financial_Project_SP500_Theme",
3     "dataColors" : [
4         "#1F77B4",
5         "#FF7F0E",
6         "#2CA02C",
7         "#D62728",
8         "#9467BD",
9         "#8C564B",
10        "#E377C2",
11        "#7F7F7F",
12        "#BCBD22",
13        "#17BECF"
14    ]
15 }
```



Dashboard Presentation

To answer the project's question, six interactive dashboards were developed. Each dashboard focuses on a different analysis area.

Overview

The **Overview** dashboard delivers a consolidated view of financial performance for selected **S&P 500 companies** between **2020 and 2021**. It highlights four key metrics — **Total Revenue**, **Gross Profit**, **Total Operating Income**, and **Net Income Margin %** — providing a snapshot of profitability and operational efficiency across the market.

The **Income Statement Waterfall** visual illustrates how each component of the income statement contributes to **Net Income**, showing the impact of major expenses such as **Cost of Revenue**, **R&D**, **SG&A**, **Interest Expense**, and **Taxes** on overall earnings. The **Quarterly Revenue vs. Gross Profit** chart tracks revenue and margin performance over time, while the **Top 5 Companies by YoY Revenue Growth** highlights firms such as **Royal Caribbean** and **Moderna**, which recorded the strongest post-pandemic rebounds and expansion rates.

Together, these visuals provide a comprehensive picture of how revenue evolves into profit and reveals which companies are driving the most significant year-over-year growth across the S&P 500.



Income Statement Waterfall

The **Waterfall Chart** illustrates how each financial component impacts overall profitability — from **Total Revenue** down to **Net Income**. It was built step-by-step to replicate a real income statement flow:

1. **Created Measures: Calculated the main financial values using SUM().**
2. **Structured the Order:** Arranged the categories to follow the logical flow of the income statement, showing how each cost reduces profit step-by-step.
3. **Configured the Waterfall Visual:** Used **Category (X-Axis)**, **Values (Y-axis)**.

```
1 Income Statement Steps =
2 DATABASE(
3   "Order", INTEGER,
4   "Step", STRING,
5   "IsSubtotal", BOOLEAN,
6   {
7     {1, "Total Revenue", TRUE},
8     {2, "Cost of Revenue", FALSE},
9     {3, "Gross Profit", TRUE},
10    {4, "Research Development", FALSE},
11    {5, "Selling General Administrative", FALSE},
12    {6, "Total Operating Expenses", FALSE},
13    {7, "Operating Income", TRUE},
14    {8, "Interest Expense", FALSE},
15    {9, "Total Other Income Expense Net", FALSE},
16    {10, "Income Before Tax", TRUE},
17    {11, "Income Tax Expense", FALSE},
18    {12, "Net Income", TRUE}
19  }
20 )
21
```

✓  Income Statement Steps

-  Is Subtotal
- IsSubtotal
- \sum Order
- Step
-  Waterfall Value

```

1 Waterfall Value =
2 VAR s = SELECTEDVALUE('Income Statement Steps'[Step])
3 RETURN
4 SWITCH(
5     s,
6     "Total Revenue",      [Total Revenue],
7     "Cost of Revenue",    -[Cost of Revenue],
8     "Gross Profit",       [Gross Profit],
9     "Research Development", -[Research Development],
10    "Selling General Administrative", -[Selling General Administrative],
11    "Total Operating Expenses", -[Total Operating Expenses],
12    "Operating Income",   [Operating Income],
13    "Interest Expense",   -[Interest Expense],
14    "Total Other Income Expense Net", [Total Other Income Expense Net],
15    "Income Before Tax",   [Income Before Tax],
16    "Income Tax Expense",  -[Income Tax Expense],
17    "Net Income",          [Net Income]
18 )
19

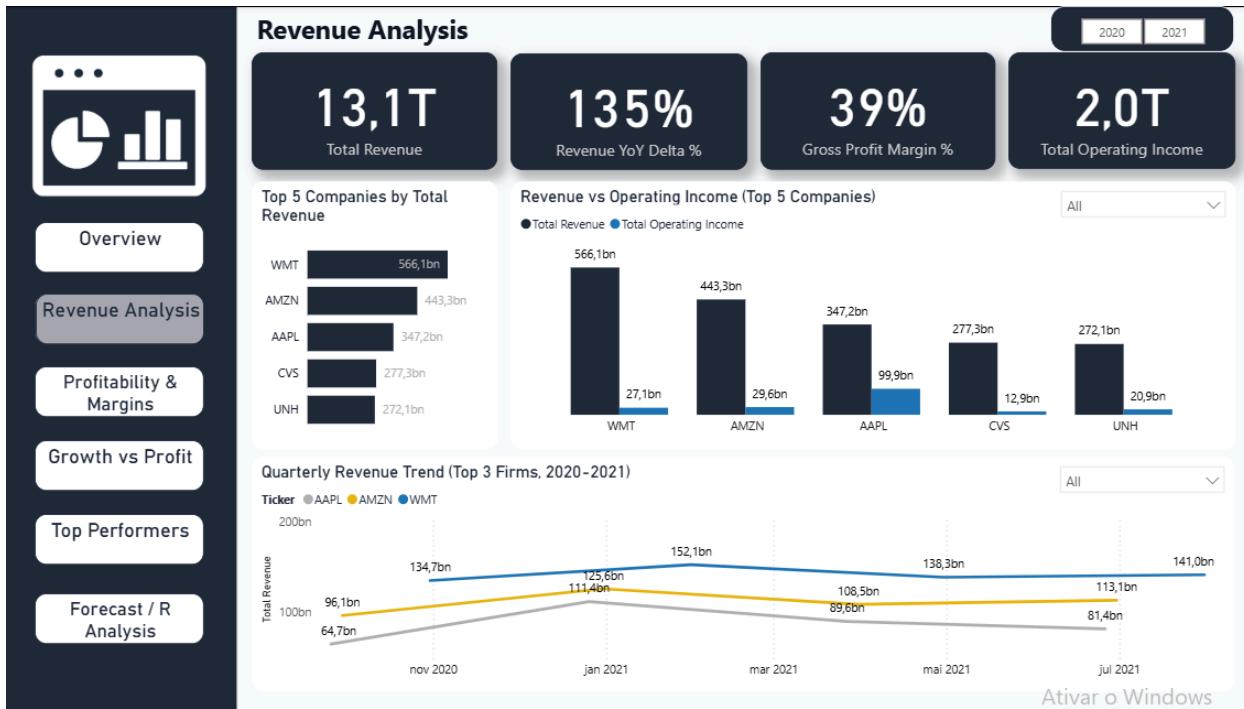
```

Revenue Analysis

This dashboard focuses on identifying which companies generate the highest revenue and how effectively they convert sales into profit. It features four key indicators – **Total Revenue**, **Revenue YoY Delta %**, **Gross Profit Margin %**, and **Total Operating Income** – that summarize key financial performance metrics and highlight growth efficiency.

The **Top 5 Companies by Total Revenue** visual highlights leading firms such as **Walmart**, **Amazon**, and **Apple**, providing a clear comparison of market scale. The **Revenue vs. Operating Income** clustered column chart further compares the same companies, illustrating how revenue translates into operating profit. At the bottom, the **Quarterly Revenue Trend** line chart tracks revenue progression for the top three companies (**Apple**, **Amazon**, and **Walmart**) between 2020 and 2021, offering insight into performance trends and seasonality.

Together, these visuals deliver a comprehensive view of revenue generation and operational efficiency, helping pinpoint which companies are driving growth and maintaining profitability over time.



Profitability & Margins

This dashboard focuses on understanding how efficiently companies convert their revenue into profit, highlighting operational performance and margin strength across leading firms. Four KPI cards — **Total Operating Income**, **Net Income Margin %**, **Gross Profit Margin %**, and **Operating Income YoY %** — summarize key profitability indicators and year-over-year trends.

The **Company-Level Profitability Metrics** table provides a detailed comparison across top firms, showing how **Total Revenue**, **Operating Income**, **Net Income**, and **Net Income Margin %** differ by company. Conditional formatting highlights which firms achieve the strongest margins and operational results. The **Total Operating Income and Net Income by Ticker** chart visualizes profitability distribution among major players such as **Apple**, **Amazon**, and **Walmart**, while the **Net Income Margin Trend** line chart (2020-2021) tracks profitability progression for these top three firms over time.

Together, these visuals offer a comprehensive view of how revenue translates into profit, revealing which companies manage costs and operations most effectively.

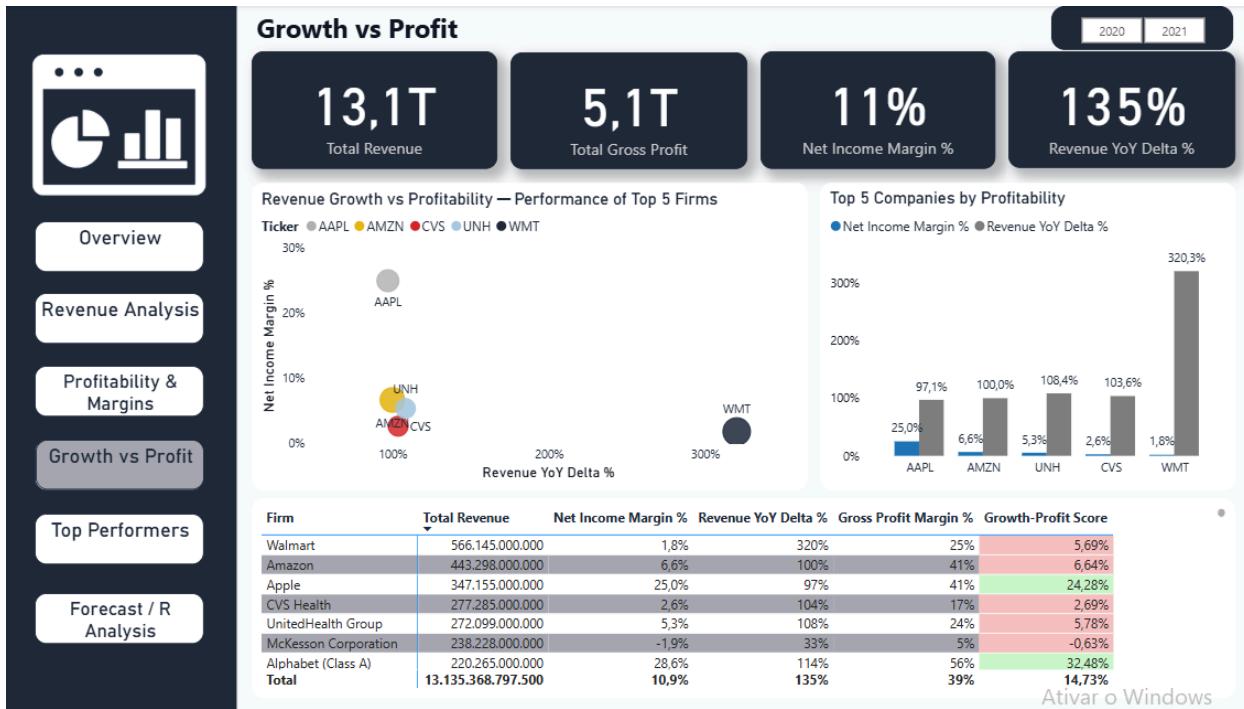


Growth vs Profit

This dashboard analyzes how company growth correlates with profitability, helping identify which firms achieve strong financial performance while maintaining sustainable margins. It features four key indicators — **Total Revenue**, **Total Gross Profit**, **Net Income Margin %**, and **Revenue YoY Delta %** — that summarize overall growth and profitability trends.

The **Revenue Growth vs Profitability** scatterplot compares **Revenue YoY Delta %** and **Net Income Margin %** for the top five firms, highlighting the balance between growth and efficiency. Companies positioned in the upper-right quadrant, such as **Apple** and **Walmart**, demonstrate both strong revenue growth and healthy margins. The **Top 5 Companies by Profitability** bar chart ranks firms based on **Net Income Margin %** and **Revenue YoY Delta %**, revealing which ones successfully combine expansion with profit stability. At the bottom, the **Company-Level Table** presents detailed performance metrics — **Total Revenue**, **Net Income Margin %**, **Revenue Growth %**, **Gross Profit Margin %**, and a calculated **Growth-Profit Score**, which measures how efficiently each company converts revenue gains into profitability.

Together, these visuals provide a clear view of how top firms balance growth and efficiency, offering valuable insight into financial health and long-term performance sustainability.

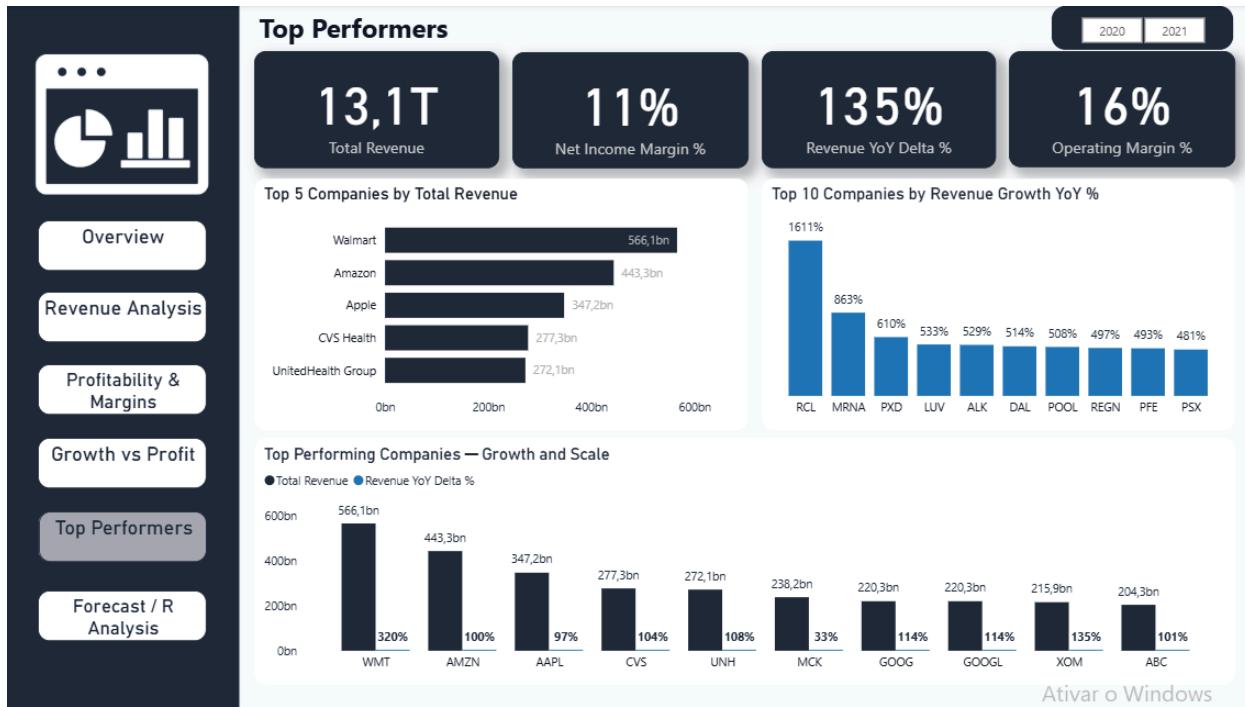


Top Performers

This dashboard highlights the leading companies in both total revenue generation and revenue growth performance, helping identify market leaders and high-growth firms. Four key indicators – **Total Revenue**, **Net Income Margin %**, **Revenue YoY Delta %**, and **Operating Margin %** – summarize overall financial performance and provide context for comparison.

The **Top 5 Companies by Total Revenue** chart showcases dominant players such as **Walmart**, **Amazon**, and **Apple**, emphasizing their scale and market strength. To the right, the **Top 10 Companies by Revenue Growth YoY %** visual highlights firms with the fastest year-over-year expansion, including **Royal Caribbean**, **Moderna**, and **Pioneer Natural Resources**. At the bottom, the **Revenue vs. YoY Growth** clustered column chart combines **Total Revenue** with **Revenue YoY Delta %**, illustrating how the largest companies balance market scale with growth acceleration.

Overall, this page provides a clear snapshot of top-performing firms from both a scale and growth perspective, offering valuable insight into market leadership, expansion potential, and performance momentum.



Forecast / R Analysis

The final dashboard integrates **predictive analytics** into the financial report by applying an **ARIMA (AutoRegressive Integrated Moving Average)** time-series model to forecast quarterly revenue. It combines key financial KPIs with a forward-looking projection to help anticipate near-term revenue trends and support strategic planning. The page includes four consistent KPI cards — **Total Revenue**, **Gross Profit**, **EBIT Margin %**, and **Net Income Margin %** — to maintain alignment across all report sections.

The **Forecasted Quarterly Revenue — ARIMA Model Projection** line chart displays historical quarterly revenue alongside a four-quarter forecast generated in R. The **blue confidence band** represents the expected range of potential outcomes based on historical data, illustrating the uncertainty around the projection. Below, the **Quarterly Actuals — Revenue and Operating Income** chart compares real historical values, providing a baseline for evaluating forecast accuracy and performance alignment. Finally, the text box titled “**About this Forecast**” explains the model’s methodology — emphasizing that the ARIMA model uses historical revenue data to estimate the next four quarters and visualize possible variation in future performance.

Together, these visuals deliver a data-driven forecast that blends historical insights with predictive modeling, enhancing financial decision-making and forward-looking strategy.



Technical Implementation — R Integration in Power BI

- Data Preparation:** Selected quarterly revenue data (YearQuarter and Revenue_Q) from Power BI and removed duplicates or missing values.
- Libraries Loaded:** Imported the forecast, ggplot2, and dplyr packages for modeling and visualization.
- Time-Series Conversion:** Converted revenue data into a quarterly time series using ts() with a frequency of 4.
- Modeling:** Applied auto.arima() to automatically find the best ARIMA parameters for forecasting.
- Forecasting:** Generated predictions for the next four quarters using the forecast function.
- Visualization:** Plotted the forecast with autoplot() from ggplot2, highlighting the blue confidence interval to show expected variation.

R script editor

⚠ Duplicate rows will be removed from the data.

```
1 # The following code to create a dataframe and remove duplicated rows is always executed and acts as a preamble for your script:
2
3 # dataset <- data.frame(YearQuarter, Revenue_Q)
4 # dataset <- unique(dataset)
5
6 # Paste or type your script code here:
7 suppressPackageStartupMessages({
8   library(forecast)
9   library(ggplot2)
10  library(dplyr)
11  library(tidyr)
12 })
13
```

R script editor

⚠ Duplicate rows will be removed from the data.

```
14 #Cleaning and Order
15 df <- dataset %>%
16   tidyrr::drop_na(YearQuarter, Revenue_Q) %>%
17   mutate(
18     Year = as.integer(substr(YearQuarter, 1, 4)),
19     Quarter = as.integer(substr(YearQuarter, 7, 7))
20   ) %>%
21   arrange(Year, Quarter)
22
23 ##Convert Revenue to billions
24 df$Revenue_Q <- df$Revenue_Q / 1e12
25
26 #Creating a quarterly time series
27 y <- ts(df$Revenue_Q, frequency = 4,
```

Ativar o Wi

R script editor

⚠ Duplicate rows will be removed from the data.

```
28 start = c(min(df$Year), min(df$Quarter[df$Year == min(df$Year)])))
29
30 #Fit ARIMA and forecast 4 quarters
31
32 fit <- auto.arima(y)
33 fc <- forecast(fit, h = 4)
34
35 #Creating ggplot
36 autoplot(fc, fill = "skyblue", color = "#003366") +
37   ggtitle("Revenue Forecast - Next 4 Quarters") +
38   xlab("Quarter") +
39   ylab("Total Revenue (Tn USD)") +
40   theme_minimal(base_size = 14) +
41   theme(
```

Ativar o Wi

R script editor

⚠ Duplicate rows will be removed from the data.

```
37   ggtitle("Revenue Forecast - Next 4 Quarters") +
38   xlab("Quarter") +
39   ylab("Total Revenue (Tn USD)") +
40   theme_minimal(base_size = 14) +
41   theme(
42     plot.title = element_text(face = "bold", hjust = 0.5, size = 16, color = "#003366"),
43     axis.text = element_text(color = "gray25"),
44     axis.title = element_text(face = "bold"),
45     panel.grid.minor = element_blank(),
46     panel.grid.major.y = element_line(color = "gray85"),
47     plot.background = element_rect(fill = "white", color = NA)
48   )
49
50
```

Ativar o Wi

Navigation Menu

To improve the dashboard's interactivity and user experience, a **navigation menu** was added to the left side of the report. Each button was configured with the “**Page Navigation**” **action**, allowing users to switch between dashboard pages. To enhance usability, the buttons were formatted with a **press color effect**, changing appearance when selected. This visual feedback helps users identify the active page, and ensures a smooth navigation throughout the report.



Questions Answered

How did overall company performance evolve between 2020 and 2021?

Between 2020 and 2021, overall company performance showed **significant recovery and growth**. **Total Revenue** increased from **5.6T in 2020** to **13.1T in 2021**, representing a substantial expansion in overall sales activity across the S&P firms analyzed. **Total Gross Profit** more than doubled from **2.1T** to **5.1T**, indicating improved cost management and stronger pricing power during the recovery phase. **Total Operating Income** rose sharply from **777.9B** to **2T**, highlighting better operational efficiency and expense control. **Net Income Margin** improved from **8%** to **11%**, showing enhanced profitability despite higher revenue levels.

The **Income Statement Waterfall** further illustrates this improvement: major cost components such as **Cost of Revenue** and **Operating Expenses** grew proportionally less than total revenue, allowing firms to retain a larger share of earnings. The **Quarterly vs Gross Profit** chart also reveals consistent upward trends through 2020-2021, peaking around early 2021 before stabilizing — a typical pattern following recovery from an economic downturn.

In summary, **companies demonstrated strong revenue recovery, improved margins, and higher operating income efficiency** between 2020 and 2021, signaling a successful transition from contraction to sustained growth.

Which companies experience the strongest post-2020 growth?

The post-2020 recovery was led by a select group of firms that achieved extraordinary revenue expansion as global economic activity rebounded. The **Top 10 Companies by Revenue Growth YoY %** highlight several standout names:

- **Royal Caribbean Cruises Ltd. (RCL)** posted a remarkable **1611% YoY growth**, the highest in the dataset, driven by the rebound of the travel and tourism sector after pandemic-related shutdowns.
- **Moderna (MRNA)** followed with **863% growth**, fueled by surging global demand for its COVID-19 vaccine.
- **Pioneer Natural Resources (PXD)** recorded **610% growth**, reflecting the strong recovery in oil and gas markets.
- Airlines such as **Southwest Airlines (LUV)** and **Alaska Air Group (ALK)** also ranked among the leaders, each exceeding **500% annual revenue growth**, underscoring the travel industry's sharp rebound.

While these firms posted the fastest percentage increases, the **largest revenue generators** overall remained **Walmart (566B)**, **Amazon (443B)**, and **Apple (347B)**. Each maintained substantial absolute growth and market dominance, combining scale with resilience.

While smaller firms experienced the most dramatic rebounds, mega-caps like **Walmart** and **Amazon** still delivered consistent, large-scale expansion — balancing stability and recovery momentum.

In summary, **post-2020 growth was driven by pandemic-sensitive industries — travel, biotechnology, and energy — while large-cap leaders sustained dominant revenue positions**, demonstrating both recovery and resilience in the S&P 500 landscape.

Which firms achieved the highest profitability in 2021?

In 2021, profitability improved significantly across several major firms as the economy stabilized and demand normalized after 2020's disruptions. The **Profitability & Margins** dashboard highlights the top-performing companies by key profitability indicators: **Total Operating Income**, **Net Income Margin %**, and **Gross Profit Margin %**.

- **Apple (AAPL)** recorded the **highest overall profitability**, with **\$51.6 billion in operating income**, and a **net income margin of 26.5%**, reflecting exceptional pricing power, strong product margins, and efficient cost management.
- **Microsoft (MSFT)** and **Google (GOOG/GOOGL)** followed closely, each posting **net income margins around 31-36%**, driven by high-margin software, cloud, and digital advertising businesses.
- **JPMorgan Chase (JPM)** also delivered a strong **37% net margin**, reflecting strong financial sector rebound and robust investment returns.

Smaller firms such as **Bio-Rad Laboratories (BIO)** and **eBay (EBAY)** stood out with **extraordinarily high margins (above 100%)**, largely due to one-time gains or accounting adjustments rather than recurring operational profits.

In summary, **Apple, Microsoft, Google, and JPMorgan** emerged as the **most profitable companies in 2021**, combining high revenue scale with outstanding efficiency. Meanwhile, **Bio-Rad** and **eBay** achieved exceptional profit margins, showing how certain industries experienced temporary post-pandemic boosts.

Is there a relationship between company size (Revenue) and profitability (Net Income Margin)?

Based on the **Growth vs Profit dashboard**, there is **no strong direct correlation** between company size (measured by total revenue) and profitability (measured by net income margin %). Large companies such as **Walmart**, **Amazon**, and **CVS Health** generated the **highest total revenues** — ranging from \$277 billion to \$566 billion — but maintained **relatively modest profit margins** between **1.8% and 6.6%**. These firms operate in **high-volume**,

low-margin sectors (retail and healthcare), where scale drives revenue but limits margin expansion.

In contrast, **smaller technology and biotech firms** such as **Apple, Microsoft, and Google** combined **moderate revenue scale** with **exceptionally high margins (25% to 36%)**. Their business models depend on **intellectual property, digital platforms, and recurring service revenue**, which naturally support higher profitability despite lower total revenue relative to retail giants. The **scatter plot** reinforces this pattern: larger firms cluster at the **right side of the chart (higher revenue)** but **lower on the vertical axis (lower margins)**, while smaller innovation- driven companies sit **higher on the margin scale**. The **overall average** net income margin across the dataset was **10.9%**, suggesting that **efficiency and business model — rather than scale alone — drive profitability**.

There is **no linear relationship** between company size and profitability. **High revenue does not guarantee high margins**; instead, **sector dynamics, cost structure, and innovation intensity** determine a firm's ability to convert revenue into profit.

Which companies combine strong growth with strong profitability -- the ideal balance of performance?

The companies that achieved the **best balance between revenue expansion and profitability** are:

- **Moderna** — *Net Income Margin: 63.9%, Revenue YoY Growth: 763%, Growth-Profit Score: 485.5%*. Moderna clearly dominates this category, combining explosive growth during the vaccine rollout with extremely high profitability, representing the strongest recovery momentum.
- **The Cooper Companies** — *Net Income Margin: 131%, Revenue Growth: 217%, Growth-Profit Score: 284.8%*. Demonstrated efficient profit conversion and consistent post-pandemic expansion.
- **Regeneron Pharmaceuticals** — *Net Income Margin: 52.6%, Revenue Growth: 510%, Growth-Profit Score: 208.8%*. Balanced rapid growth in pharmaceutical sales with solid cost efficiency.
- **Public Storage** — *Net Income Margin: 51.9%, Revenue Growth: 236%, Growth-Profit Score: 122.4%*. Reflected strong property demand and margin discipline in the storage sector.
- **Pfizer** — *Net Income Margin: 33.6%, Revenue Growth: 133%, Growth-Profit Score: 126.8%*. Achieved sustained revenue growth while keeping a high profitability level.

Together, these companies represent the **ideal mix of expansion and financial efficiency**, standing out as post-crisis leaders that maximized both **growth potential** and **bottom-line profitability**. Among large, diversified firms, **Apple** and **Alphabet (Google)** also

performed strongly, maintaining **profit margins above 25%** with **steady revenue growth (30-35%)**, showcasing scalability and consistent returns.

Summary

The data tells a powerful story of resilience and transformation. In **2020**, the pandemic sent shockwaves across industries — revenues collapsed, margins tightened, and many companies struggled to stay afloat. Yet by **2021**, a remarkable recovery had begun. Total revenue more than doubled, profitability strengthened, and several industries reinvented themselves through innovation and operation efficiency.

Technology and healthcare led the comeback. Companies like **Apple, Google, and UnitedHealth Group** maintained strong margins despite global uncertainty, while biotech firms such as **Moderna** and **Regeneron** emerged as breakout performers, combining record-high revenue growth with exceptional profitability. In contrast, **travel and energy companies** saw explosive growth as demand rebounded, though margins remained pressured by recovery costs.

By the end of 2021, the landscape had shifted. The firms that thrived were those able to **balance growth with profitability**. **Moderna, The Cooper Companies, and Public Storage** stood out as ideal examples, achieving the rare mix of expansion and efficiency. This post-pandemic period highlighted not just recovery, but a deeper transformation — a move toward agility, innovation, and financial strength that defined the new era of corporate performance.

Conclusion

To wrap up, this project provided a comprehensive understanding of how major S&P 500 companies performed during and after the pandemic period. By analyzing data from 2020 and 2021, I was able to uncover clear trends in how different sectors recovered, highlighting the strong resilience of technology and healthcare companies such as Apple, Google, and Moderna. These firms demonstrated the importance of innovation adaptability, and efficiency in driving both growth and profitability. On the other hand, industries like travel and energy faced greater volatility but showed signs of gradual recovery as market conditions stabilized. Overall, this analysis emphasized the value of using data-driven insights to evaluate company performance, understand market shifts, and identify the factors that contribute to long-term financial resilience.

