

# Financial Analysis: A Case Study of Quarterly Balance sheet Data

BY SAIKARTHIK MANTRI

**Project Objective:** The objective of this project is to leverage Business Intelligence (BI) tools and techniques to analyze quarterly balance sheet data of various companies. The goal is to extract meaningful insights from the data that can aid in financial decision-making and strategic planning.

**Business Case Definition:** The increasing complexity of financial markets and the exponential growth of data volume necessitate the use of advanced Business Intelligence tools for effective financial analysis. This project aims to leverage BI tools to analyze quarterly balance sheet data of various companies, providing valuable insights that can aid in strategic decision-making. The insights derived from this project could potentially lead to more informed and data-driven decisions, thereby improving the financial performance and strategic direction of these companies.

## Scope

1. **Data Storage and Management:** PostgreSQL will be used as the primary database for storing and managing the quarterly balance sheet data. It provides robust data integrity features and supports complex queries.
2. **Data Processing and Transformation:** Alteryx will be used for data processing and transformation. It can connect to PostgreSQL to extract data, perform complex data transformations, and prepare the data for analysis.
3. **Data Visualization and Reporting:** PowerBI will be used for data visualization and reporting. It can connect to Alteryx to import the transformed data and create interactive dashboards and reports for the balance sheet data.
4. **Data Analysis:** The project will include analyzing trends, patterns, and insights from the balance sheet data over different quarters.
5. **Collaboration and Sharing:** PowerBI allows for secure sharing and collaboration of reports and dashboards, enabling stakeholders to make data-driven decisions.

## Limitations

1. **Data Security:** While PostgreSQL provides robust security features, sensitive financial data must be properly protected. This includes implementing proper access controls and encryption.
2. **Data Quality:** The quality of the output is dependent on the quality of the input data. Any inaccuracies in the balance sheet data can lead to incorrect analysis and insights.

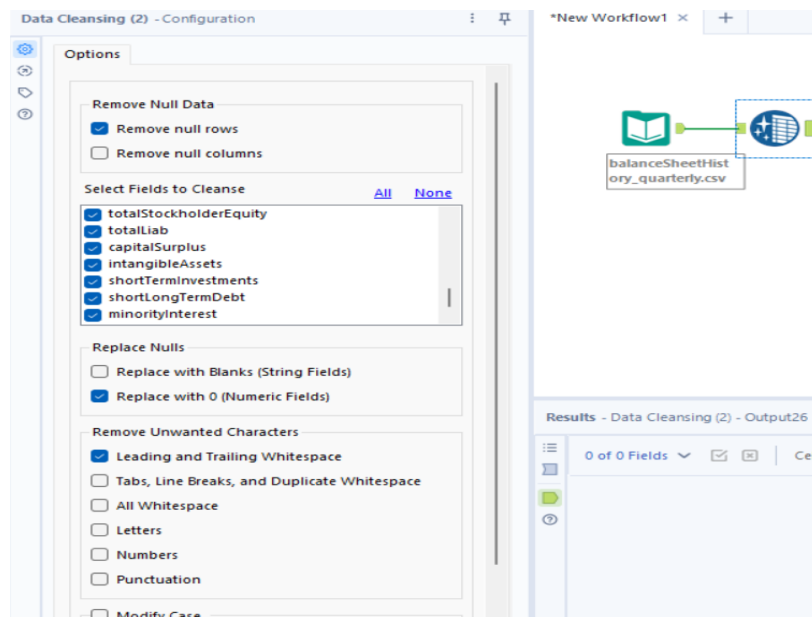
3. **Technical Expertise:** The successful implementation and use of these tools require a certain level of technical expertise. Users need to be familiar with SQL for PostgreSQL, data transformation techniques for Alteryx, and data visualization principles for PowerBI.
4. **Performance:** Large volumes of data can impact the performance of these tools. This can be mitigated by optimizing the data processing and transformation steps in Alteryx and by using efficient data modeling techniques in PowerBI.
5. **Cost:** There are costs associated with the licensing and maintenance of these tools. These costs should be considered in the project budget.

## Source

Source of the data can be found at Kaggle with the following link [Financial data of 4400+ public companies \(kaggle.com\)](https://www.kaggle.com/datasets/rajat-arora/financial-data-of-4400-public-companies).

## ETL PROCESS

1. Extraction, Transformation and Loading are done in the Alteryx software. Basic data preparation, transformation and loading into PostgreSQL are performed in Alteryx.
2. Data has been downloaded from the source into the local computer, using Alteryx input function data has been extracted, data contains multiple null values, hence data cleaning preprocessing is done to the data, rows containing null values have been completely deleted to increase the data accuracy and reduce outliers. Date also contains a date format column, but the Alteryx considers as V\_string hence one more preprocessing step has been added to change the format of a variable known as “Select” in Alteryx



**Input Data (1) - Configuration**

Connect a File or Database  
 C:\Users\mantr\Downloads\archive (3)\balanceSheetHistory\_quarterly.csv  
 Set Up a Connection  
☐ Use Data Connection Manager (DCM)

Options

	Name	Value
1	Record Limit	
2	File Format	Comma Se
3	Search SubDirs	<input type="checkbox"/>

Preview (first 100 records) Refresh

	stock	endDate	accountsPayable	inventory	longTermDebt
1	IVC	2020-03-31	90614000.0	119285000.0	165
2	IVC	2019-12-31	88003000.0	121184000.0	219
3	IVC	2019-09-30	89736000.0	123069000.0	222
4	IVC	2019-06-30	86051000.0	124686000.0	233
5	CSLT	2020-03-31	13299000.0	[Null]	930
6	CSLT	2019-12-31	19596000.0	[Null]	139
7	CSLT	2019-09-30	11989000.0	[Null]	186
8	CSLT	2019-06-30	11355000.0	[Null]	232
9	LPX	2020-03-31	208000000.0	284000000.0	697
10	LPX	2019-12-31	112000000.0	265000000.0	347
11	LPX	2019-09-30	202000000.0	260000000.0	347
12	LPX	2019-06-30	213000000.0	293000000.0	347
13	GDYN	2020-03-31	[Null]	[Null]	[Nu
14	GDYN	2019-12-31	[Null]	[Null]	[Nu
15	GDYN	2019-03-31	[Null]	[Null]	[Nu
16	GDYN	2018-12-31	[Null]	[Null]	[Nu
17	ADA	2020-03-31	5725000.0	7574000.0	572

**Results - Input Data (1) - Output**

0 of 0 Fields

**Select (3) - Configuration**

Options

	Field	Type	Size
<input checked="" type="checkbox"/>	stock	V_String	254
<input checked="" type="checkbox"/>	endDate	Date	10
<input checked="" type="checkbox"/>	accountsPayable	V_String	254
<input checked="" type="checkbox"/>	inventory	V_String	254
<input checked="" type="checkbox"/>	longTermDebt	V_String	254
<input checked="" type="checkbox"/>	netReceivables	V_String	254
<input checked="" type="checkbox"/>	netTangibleAssets	V_String	254
<input checked="" type="checkbox"/>	longTermInvestments	V_String	254
<input checked="" type="checkbox"/>	totalCurrentAssets	V_String	254
<input checked="" type="checkbox"/>	propertyPlantEquipment	V_String	254
<input checked="" type="checkbox"/>	otherStockholderEquity	V_String	254
<input checked="" type="checkbox"/>	deferredLongTermAssetCharges	V_String	254
<input checked="" type="checkbox"/>	totalCurrentLiabilities	V_String	254
<input checked="" type="checkbox"/>	cash	V_String	254
<input checked="" type="checkbox"/>	otherAssets	V_String	254
<input checked="" type="checkbox"/>	treasuryStock	V_String	254
<input checked="" type="checkbox"/>	goodWill	V_String	254
<input checked="" type="checkbox"/>	otherLiab	V_String	254
<input checked="" type="checkbox"/>	retainedEarnings	V_String	254
<input checked="" type="checkbox"/>	otherCurrentAssets	V_String	254
<input checked="" type="checkbox"/>	commonStock	V_String	254
<input checked="" type="checkbox"/>	totalAssets	V_String	254
<input checked="" type="checkbox"/>	otherCurrentLiab	V_String	254

**Results - Select (3) - Output**

0 of 0 Fields

- The transformation process consists of a simple transformation which is summation of columns such as total assets value, total liability, investments, cash, with respect to stock or company and end date of the stock.



- The transformed data is loaded into PostgreSQL server for data storage from the Alteryx function “output”.



Query

Query History

1

2

SELECT \* From public."Fproject"  
limit 100;

Data Output

Messages

Notifications

	stock character	endDate date	Sum_accountsPayable double precision	Sum_inventory double precision
1	MSGN	2019-09-30	412000	[null]
2	ESSA	2020-03-31	1327613000	408000
3	CVU	2019-06-30	11540234	11956006
4	WINT	2019-06-30	724000	[null]
5	LTRPA	2019-09-30	212000000	[null]
6	TW	2020-03-31	98591000	[null]
7	CLGN	2019-12-31	833000	888000
8	WHG	2020-03-31	1648000	[null]

## TOOLS AND TECHNOLOGIES USED IN ETL

Alteryx, PostgreSQL

## DATA STORAGE

### Selection of Data Storage Solution

I have selected **PostgreSQL** as our data storage solution. The reasons for this selection are as follows:

1. **Open-Source:** PostgreSQL is an open-source database system, which means it is free to use and has a large community of developers constantly improving its features and performance.
2. **ACID Compliance:** PostgreSQL is fully ACID compliant (Atomicity, Consistency, Isolation, Durability), ensuring that all transactions are processed reliably.
3. **Extensibility:** PostgreSQL is highly extensible. It supports a number of advanced data types not available in other databases, such as arrays and hstore (for key-value pairs).
4. **Performance:** PostgreSQL has good performance characteristics thanks to its advanced query optimizer and indexing mechanisms. This is particularly important for a BI project where complex analytical queries will be common.
5. **Security:** PostgreSQL has robust security features including strong access controls, views, and granular permissions that help protect sensitive financial data.
6. **Compatibility:** PostgreSQL is compatible with a wide range of BI tools, including Alteryx for data processing and PowerBI for data visualization, which we are using in this project.

In conclusion, PostgreSQL provides a reliable, secure, and high-performance data storage solution that meets the needs of our BI project. Its compatibility with other chosen data processing and visualization tools further makes it an ideal choice for the project.

### Data Structure

All the columns, except stock and end\_date are of string and date format, rest of them all are double format.

Fproject

×

















GeneralColumnsAdvancedConstraintsParametersSecuritySQL



Inherited from table(s)

Select to inherit from...

Columns

+

	Name	Data type	Length/Precision	Scale	Not NULL?	Primary key?	Default
 	stock	character   v	254		<input type="checkbox"/>	<input type="checkbox"/>	
 	endDate	date   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	accountsPayable	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	inventory	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	longTermDebt	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	netReceivables	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	netTangibleAssets	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	
 	longTermInvestme	double precision   v			<input type="checkbox"/>	<input type="checkbox"/>	



×

Close

↺

Reset

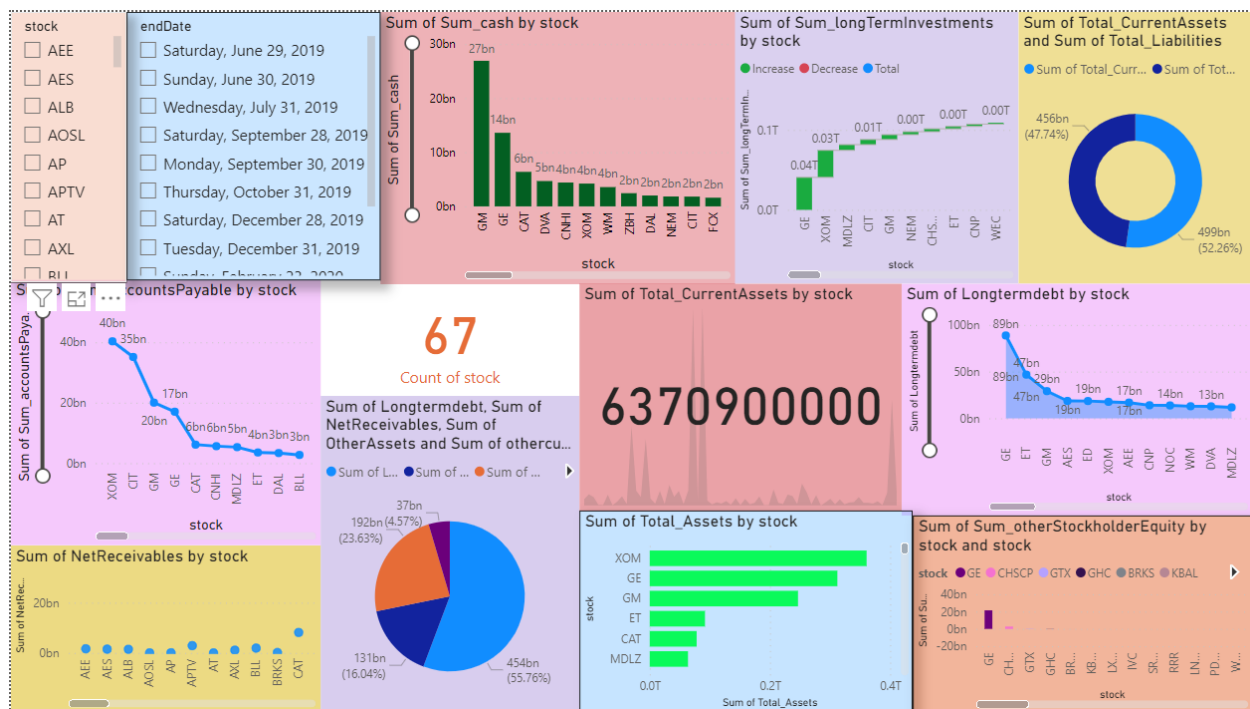
💾

Save

## Data Visualization

I have used Powerbi visualization tool in this project.

Following is the dashboard developed



## Conclusion

Based on the image you've shared, which appears to be a financial dashboard displaying various metrics for different stocks, here's a possible conclusion:

## 6. Conclusion and Summary

The financial dashboard provides a comprehensive view of the performance of various stocks over time. Key metrics such as cash by stock, total current assets, net receivables, total assets, long-term debt, and other stockholder equity are visualized effectively, allowing for easy comparison and analysis.

From the data visualized, it can be inferred that there are significant variations in the financial performance and position of the different stocks. Some stocks have shown a consistent increase in total assets and cash, indicating strong financial health. However, others have shown an increase in long-term debt, which could be a potential area of concern.

The large number "6370900000" displayed in the center-right part of the dashboard represents a sum value of a total current asset's metric across all stocks. This could be a key figure in understanding the overall financial landscape of the stocks in question.

In summary, this dashboard serves as a powerful tool for tracking and analyzing the financial performance of various stocks. It provides valuable insights that can aid in decision-making for investors and stakeholders.

## Data Modeling Diagram

Sure, here's an example of what you might write for the "Transformation Details" section:

### 7.2 Transformation Details

The transformation process is a crucial part of our ETL pipeline, and it is handled using Alteryx. The following steps outline the transformation process:

1. **Data Cleaning:** The first step in the transformation process is data cleaning. This involves removing any irrelevant data, handling missing values, and correcting inconsistent or inaccurate data. For example, we might replace missing values with the mean or median of the column, or drop rows that contain missing values.
2. **Data Normalization:** The next step is data normalization. This involves standardizing the values in numeric columns so that they fall within a similar range. This is particularly important for our balance sheet data, where different financial metrics can have vastly different scales.
3. **Creating Derived Attributes:** We also create new attributes based on existing ones to capture more complex relationships in the data. For example, we might create a new attribute that represents the ratio of current assets to current liabilities, which is a key indicator of a company's short-term liquidity.
4. **Data Aggregation:** The data is then aggregated at the appropriate level for analysis. For our project, this might involve aggregating balance sheet data at the quarterly level for each stock.
5. **Data Formatting:** Finally, the data is formatted to ensure compatibility with our data visualization tool, PowerBI. This might involve converting data types, renaming columns, or restructuring the data.

Throughout the transformation process, we use Alteryx's drag-and-drop interface to create workflows that automate these steps. This not only speeds up the process but also reduces the risk of manual errors. The workflows are also reusable and can be easily adjusted to accommodate changes in the source data or business requirements.

Remember, the goal of the transformation process is to prepare the data for analysis. It's important to ensure that the transformed data accurately represents the original information and is suitable for the intended analysis. This requires a good understanding of both the data and the business context.

Please note that the specific transformation steps and techniques used may vary depending on the nature and quality of your source data, as well as the specific requirements of your project. It's always a good idea to document your transformation process in detail to ensure transparency and reproducibility.



Fact Table: BalanceSheet

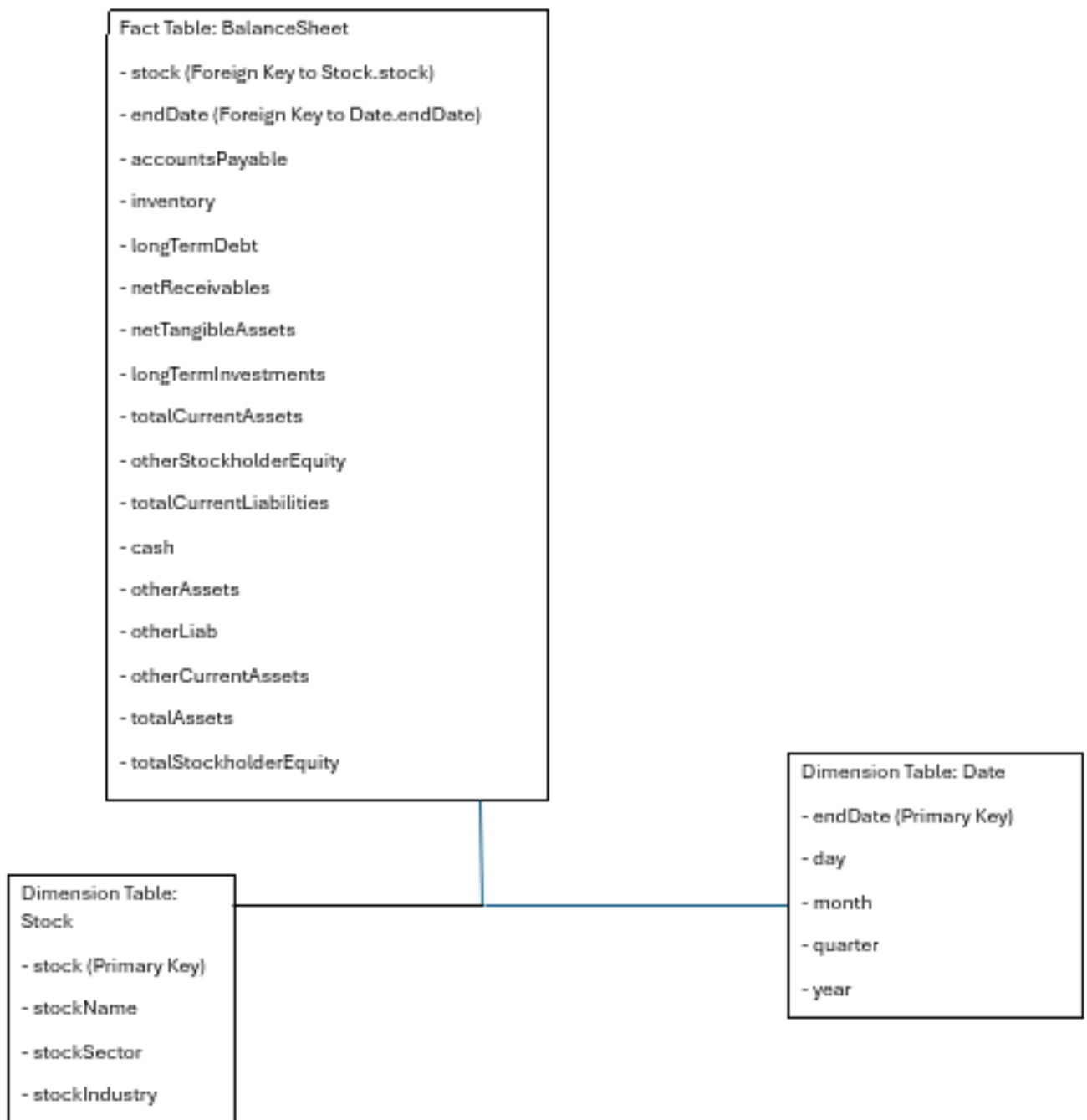
- stock (Foreign Key to Stock.stock)
- endDate (Foreign Key to Date.endDate)
- accountsPayable
- inventory
- longTermDebt
- netReceivables
- netTangibleAssets
- longTermInvestments
- totalCurrentAssets
- otherStockholderEquity
- totalCurrentLiabilities
- cash
- otherAssets
- otherLiab
- otherCurrentAssets
- totalAssets
- totalStockholderEquity

Dimension Table:  
Stock

- stock (Primary Key)
- stockName
- stockSector
- stockIndustry

Dimension Table: Date

- endDate (Primary Key)
- day
- month
- quarter
- year



## Project Implementation Challenges and Solutions

During the implementation of this project, the challenges I faced were:

- **Data Quality:** The quality of the source data was inconsistent, which affected the accuracy of our analysis. We addressed this by implementing rigorous data cleaning procedures in our ETL process.
- **Performance:** Handling large volumes of data affected the performance of our tools. We optimized our ETL processes and used efficient data modeling techniques to address this.

## Future Enhancements and Recommendations

For future enhancements, It is recommended:

- **Incorporating More Data Sources:** To provide a more comprehensive view of the financial landscape.
- **Implementing Advanced Analytics:** Such as predictive modeling to forecast future trends.
- **Improving Data Governance:** To ensure the quality and security of the data.

## References

1. Power BI Documentation. Retrieved from <https://docs.microsoft.com/power-bi/>
2. Practical Analytics by Nithin Kale and Nancy Jones

## Appendices

### Glossary of Terms

- **ETL:** Extract, Transform, Load - a process in data warehousing responsible for pulling data out of source systems and placing it into a data warehouse.
- **Powerbi:** Data visualization tool
- **PostgreSQL:** SQL based data storage software