Princess Sumaya University for Technology
جامعة الأميرة سميّة للتكنولوجيا

31 May 2023

FinTech

Business Intelligence Project

Done By:

Yasmeen Sabbara 20200899

Alanoud Hijazi 20160304

Mays Abdallah 20190838

Talal Alkilani 20200782

Supervised By:

Dr. Dalia Al-Eisawi

Table of Contents

Introduction to FinTech

For every good or service an individual needs or wants throughout their day, a payment should be made first in order to acquire it, which entices us to make many payments on a daily basis for different entities. For instance, consider the ways you can pay your internet bill, Of course, the most obvious way is to drive yourself there and pay using cash, credit or debit cards which primarily does the job perfectly fine. Until recently, a different perspective has been emphasized since the COVID-19 pandemic due to the need for digital connectivity to replace physical interactions, Fintech has come to light where you can pay your bills using your bank's mobile application on your phone or using online payment processing applications while simply lying comfortably on your couch without having to be exposed to contact of any kind.

Financial technology or Fintech is defined as "advances in technology that have the potential to transform the provision of financial services spurring the development of new business models, applications, processes, and products" (IMF and World Bank 2019). Fintech has reshaped the financial industry and changed the way people access, manage and interact with their finances by removing the geographical barriers and high entry costs barriers through offering digital platforms and mobile applications which are accessible to users anytime and anywhere. Fintech transformed traditional financial operations into practical and effective digital experiences whilst having user-friendly interfaces that provide better user experience. It has also proven to be a hub for innovation and development of new ideas which are a great threat to traditional financial institutions, in addition to the emerging technologies like artificial intelligence and cloud computing which are used to launce new business models that cater to changing consumer needs. In terms of enterprises, customers and financial institutions, Fintech has aided in cost savings and improved operational efficiency and allowed for many tasks to be

done more quickly, affordably, and with less paper work. The importance of Fintech also emerges from making use of data analytics and artificial intelligence to create tailored financial solutions, risk assessments and fraud detection systems which aid in decision-making. Fintech is positioned to influence the future of finance as technology further develops therefore empowering people and organizations.

Business intelligence in the FinTech industry

In today's rapidly changing and constantly expanding environment, thanks to the technological advancements and innovations, client expectations have also become skyrocketing. As a result, Fintech organizations must turn to business intelligence solutions and integrate them with their business in order to gain competitive advantage and deliver better customer experience. In Fintech, data is abundant therefore BI solutions are critical because they assist businesses in transforming raw data into actionable insights thus leading to well-informed decisions, improved operations, improved customer experience, and allows the organization to shave high competitive advantage. To begin with, the employment of BI solutions allows Fintech companies to manage and analyze large volumes of complex data and to obtain a holistic view of their operations by integrating data from transactions, customer profiles, market trends and many more. They also allow for the discovery of patterns, anomalies and trends resulting in more knowledge of customer behavior, risk management, fraud detection and market prospects. Monitoring and reporting in real-time are especially important when it comes to financial technology, BI systems enable them to do so by monitoring key performance indicators like transaction volumes, revenue, growth, client acquisition and retention rates in addition to creating thorough reports which ensures regulatory compliance and transparency for stakeholders. Furthermore, access to reliable data enables fast decision-making, identifying latest

trends and responding quickly to market developments. Fintech organizations require an effective risk management and compliance in order to ensure a regulated business which can be achieved through BI solutions that provide full risk assessment capabilities by evaluating historical data, detecting abnormalities and identifying fraudulent activities. In finance, businesses may frequently have to acquire insights to individual customer requirements and preferences with the help of BI solutions by examining customer data such as demographics, transaction history, etc. this data is then used to customize financial goods and services thus increasing client's satisfaction, loyalty and as a result profitability. BI technologies also help Fintech organizations to offer proactive customer service by detecting problems before they have an impact on the consumer. Implementing BI solutions provides a competitive advantage to Fintech organizations as it allows them to have a deeper insight into market dynamics, consumer behavior, and trends in the business environment. This information enables the identification of opportunities and allows them to be ahead of rivals in a market that is constantly evolving by having quick, flexible, and responsive services. In conclusion, Fintech companies that embrace business intelligence have the capacity to make educated decisions, deliver more convenient services and prosper in a highly competitive field.

Before business intelligence VS after business intelligence

As one can imagine, the Fintech industry before implementing business intelligence faced many challenges in effectively analyzing their data. For instance, organizations found it harder to get a complete picture of their operations and performance because of the data being lost in different departments as compared to after the implementation of BI solutions, data is consolidated from various sources in a centralized data repository thus making it much easier during analysis and reporting. Secondly, data was processed by hand before BI which increased

the cost and time on the organizations and leading to more mistakes, unlike with BI which allows automated data processing, cleaning and transformation therefore reducing manual errors and increasing data accuracy. Without business intelligence, Fintech organizations would have trouble extracting useful insights due to the absence of effective data analysis leading to uninformed decision making, but rather, decisions that might be biased or based on gut feelings. As mentioned previously, BI solutions enabled monitoring and reporting using KPIs and real-time reports, however, if BI is absent, the creation of reports and keeping track of KPIs would require a create deal of time and effort for the organization.

To sum up, it is safe to say that the adoption of business intelligence in the Fintech industry has transformed their data management and analysis. It also enhanced the way decisions are made leading to improved operational efficiency, better insights and higher competitive position in the Fintech industry.

Business Intelligence

Business intelligence is an umbrella that covers the processes and methods of collecting, storing, and analyzing data from business operations or activities to optimize performance (*Business Intelligence: A Complete Overview*, n.d.). It includes a range of tools and techniques like data integration tools, data marts, data warehouses, and analytics platforms that allow businesses to transform raw data into information that is helpful to the organization. Organizations also use BI processes which involve the extraction, transformation, modeling, and analysis of data. For instance, when a FinTech organization gathers data from transactions in their core banking systems including information on customer transactions or account balances and more, in addition to customer profiles, spreadsheets, and external systems, they are performing data extraction. Then, the data they collected is cleansed for errors, inconsistencies, or duplicates so

that they can work with accurate and reliable data, hence it is integrated to provide a comprehensive view and to check for the correctness and consistency of this data, data transformation is used which comprises cleaning and formatting the data in Fintech companies. Data modelling entails organizing and structuring the data to make analysis and reporting easier using techniques like star schema or snowflake schema. Finally, in order to find patterns, trends, and correlations in the data, data analysis employs statistical and analytical methods. Storing data in the context of business intelligence refers to the technologies used to manage huge volumes of data which include data lakes that store data in its raw format and NoSQL databases that offer adaptable and scalable storage solutions. Not everyone across the enterprise would be professional enough to make use of the information obtained from the BI processes, therefore, performance visualization is conducted for the presentation of data and insights in a visual manner to aid comprehension and decision-making in business intelligence. Charts, graphs, dashboards, and reports are examples of representations that help users easily understand complex information and see trends, in addition to study the data and drill down to specifics using interactive dashboards. Performance visualization is very important in a Fintech organization as it enhances data comprehension thus enabling informed decision-making and facilitating effective insight communication within all business functions in the organization. Overall, these components of business intelligence work together to enable firms to take advantage of the power of data and make informed decisions that will lead to their business' success.

Data Warehouse architecture:

While researching a suitable architecture for FinTech, we found that this was a general term to use as we got multiple results and none of them were too compatible to what we were looking for. So, we narrowed our research scope down to PayPal.

To recommend a suitable data warehouse for PayPal, we have to consider the concepts available. These concepts are one-tier, two-tier, or three-tier. These typically refer to the number of tiers involved in the system. Each tier represents a unique component responsible for conducting specific functions or tasks. In the case of PayPal, we would recommend the use of a three-tier architecture as it is commonly used and benefits the company in many areas which include better maintainability, flexibility, and scalability. The following describes each tier in detail:

1. Presentation tier:

   This tier is the first tier in this architecture, it focuses on the presentation, which contains the user interfaces, reporting tools, and dashboards. This tier enables the user to have access to the query data, data warehouse, and extract and generate reports. Commonly used tools in this tier include Tableau, power BI, and looker.

2. Application tier:

   This is the middle tier, or as it is more commonly known as the logic tier, it manages the technical part of handling the processing and transformation of data**.** It contains the components that perform data integration, ETL processes, and data modeling. This tier focuses on the use of technologies and software like Apache spark, Apache Hadoop or data integration tools like Apache airflow or AWS glue.

3. Data tier:

   This is the bottom or last tier. It represents the data management and storage layer. This is where the data warehouse is found, where structured and processed data is stored for analysis. Commonly used software technologies in this tier include Amazon Redshift, Google BigQuery, or snowflake.

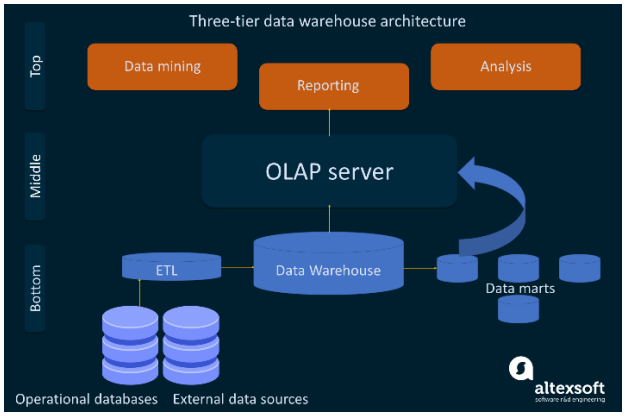


Figure 1: Three-Tier Visual

Data integration

To gain a comprehensive view of the business, customers, transactions, and many other operational aspects integrating different data sources is important. The following describes why integration is important:

1. Full view of customers:

   Integrating data from various sources allows PayPal to consolidate customer information, these include demographics, transaction history, activity across different channels, and preferences. This view helps in understanding the customer behavior and patterns and provides a personalized experience for the customer.

2. Enhanced fraud detection:

This point goes into the security part of the integration. By using the data such as transaction logs, user activity, and external fraud databases, PayPal can develop fraud detection models. These models can reveal suspicious patterns, anomalous behavior, and potential fraud attempts.

3. Accurate financial reporting:

   Integrating financial data from multiple sources, such as transactional systems, payment logs, and accounting platforms, ensures accurate and timely financial reporting. This consolidation of data helps in the effectiveness in identifying trends and generating accurate and comprehensive financial reports.

How PayPal Should Integrate its Data

1. Data integration technologies:

   Data integration technologies like ETL or ELT are used to extract data from various data sources, transform it into generic format, and load it into the data repository. Tools used can include Apache Kafka and Apache NiFi.

2. Data virtualization:

   In this stage virtualization techniques and technologies are used where a logical layer is created to integrate data from various sources. This method presents a unified view of the data, eliminating data duplication and optimizing the integration process.

3. Data warehouse or data lake:

   It is crucial to create a central data repository like a data warehouse or data lake to store the integrated data. And it is important that this repository is accurate because it acts as the main source of data for analytics, reporting, and data-driven decision making. In this

case we would recommend PayPal to use as a data warehouse due to the large amount of data.

PayPal Data Set

This data set includes the price of the PayPal stock each day from 2015 until 2022. It includes the price of the stock when it first opens, its high, its low, its close, its volume of shares by day, and the target volume of shares.

Descriptive Statistics

We will include descriptive statistics for each column of the data set.

Open Price

Mean	110.7273866
Standard Error	1.801002521
Median	88
Mode	39.400002
Standard Deviation	74.97450414
Sample Variance	5621.176271
Kurtosis	-0.057281722
Skewness	1.037418989
Range	278.840004
Minimum	30.82
Maximum	309.660004
Sum	191890.5609
Count	1733

Figure 2: Descriptive Statistics Open Price

This data set is not normally distributed, asymmetrical, and has outliers.

Its histogram should be right-skewed (Positive Skewness)

We have a right tail.

The mode < median < mean

As seen the Kurtosis is not exceedingly high which indicates outliers but not very much.

High

	High
Mean	112.1225069
Standard Error	1.822762311
Median	88.987499
Mode	36.5
Standard Deviation	75.90223893
Sample Variance	5761.149875
Kurtosis	-0.076990502
Skewness	1.030537559
Range	278.880003
Minimum	31.280001
Maximum	310.160004
Sum	194420.427
Count	1734

Figure 3: Descriptive Statistics High Price

This data set is not normally distributed, asymmetrical, and has outliers.

Its histogram should be right-skewed (Positive Skewness)

We have a right tail.

The mode < median < mean

As seen the Kurtosis is not exceedingly high which indicates outliers but not very much.

Low

Mean	108.9945093
Standard Error	1.769886652
Median	86.7350005
Mode	38.099998
Standard Deviation	73.70042638
Sample Variance	5431.752849
Kurtosis	-0.032260061
Skewness	1.042789507
Range	275.230011
Minimum	30
Maximum	305.230011
Sum	188996.4791
Count	1734

Figure 4: Descriptive Statistics Low Price

This data set is not normally distributed, asymmetrical, and has outliers.

Its histogram should be right-skewed (Positive Skewness)

We have a right tail.

The mode < median < mean

As seen the Kurtosis is not exceedingly high which indicates outliers but not very much.

Close

Mean	110.5985929
Standard Error	1.797206966
Median	87.884998
Mode	33.77
Standard Deviation	74.83808051
Sample Variance	5600.738294
Kurtosis	-0.052377765
Skewness	1.03714972
Range	277.9
Minimum	30.629999
Maximum	308.529999
Sum	191777.9602
Count	1734

Figure 5: Descriptive Statistics Close Price

This data set is not normally distributed, asymmetrical, and has outliers.

Its histogram should be right-skewed (Positive Skewness)

We have a right tail.

The mode < median < mean

As seen the Kurtosis is not exceedingly high which indicates outliers but not very much.

Volume of Shares

Mean	9277086.563
Standard Error	156159.617
Median	7561150
Mode	4653300
Standard Deviation	6502693.462
Sample Variance	4.2285E+13
Kurtosis	91.59179668
Skewness	6.340265156
Range	134584000
Minimum	1680000
Maximum	136264000
Sum	16086468100
Count	1734

Figure 6: Descriptive Statistics Volume of Shares

This data set is not normally distributed, asymmetrical, and has outliers.

Its histogram should be right-skewed (Positive Skewness)

We have a right tail.

The mode < median < mean

As seen, the Kurtosis is exceedingly high which indicates so many outliers.

Quartiles Analysis

Open Price:

Q1 = 42.89 Q2 = 87.97 Q3 = 143.44

Minimum Price 30.82 Maximum price 305.50

Outliers= 294.90, 300.14

High Price:

Q1 = 43.2 Q2 = 88.9 Q3 = 146.3

Minimum Price 31.2 Maximum price 309.1

Outliers= 301.1

Low Price:

Q1 = 42.6 Q2 = 86.7 Q3 = 141.7

Minimum Price 30 Maximum price 305.2

Outliers= 290.55, 298.17

Close Price:

Q1 = 43.0 Q2 = 87.8 Q3 = 145.4

Minimum Price 30.6 Maximum price 306.7

Outliers= 300.2

Volume Price:

Q1 = 5,794,929 Q2 = 7,561,150 Q3 = 10,603,125

Minimum Price 1680000 Maximum price 136264000

Outliers= 63995100, 52101200, etc.…

Box Plot

Figure 7: Open Price Box Plot

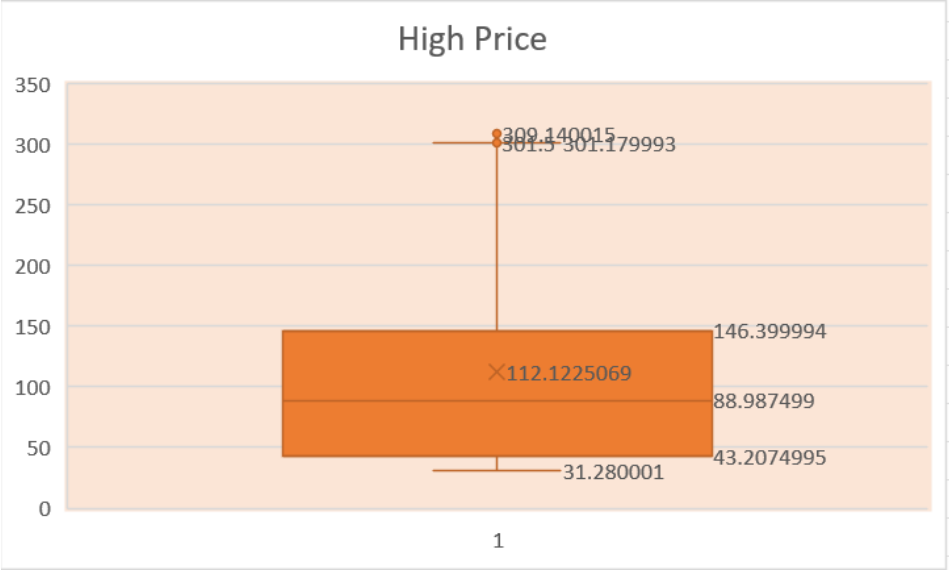
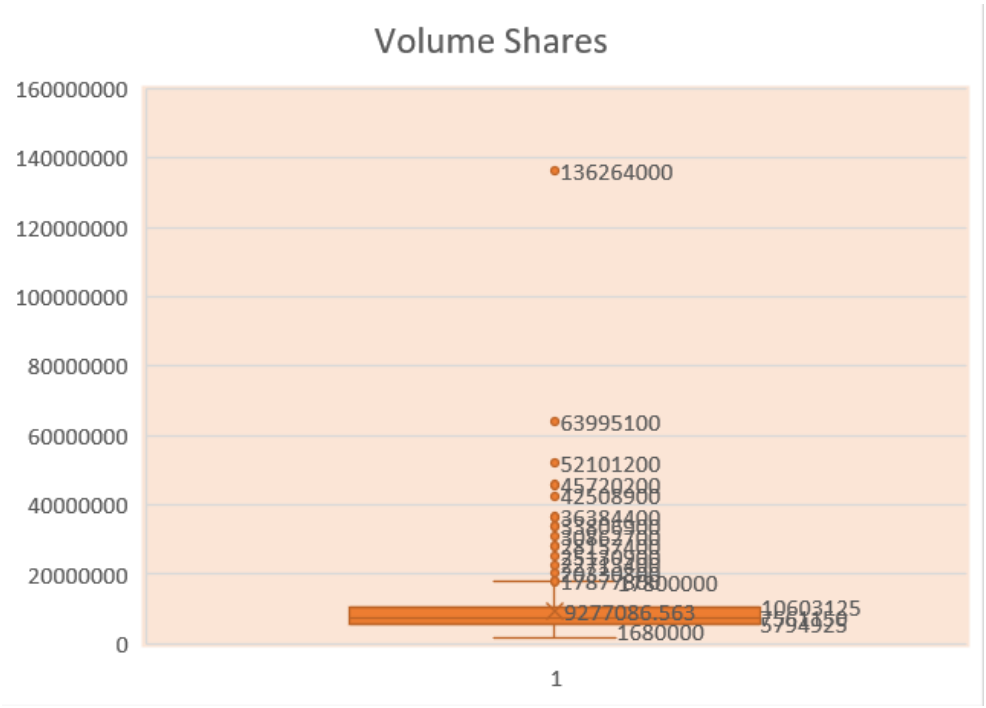


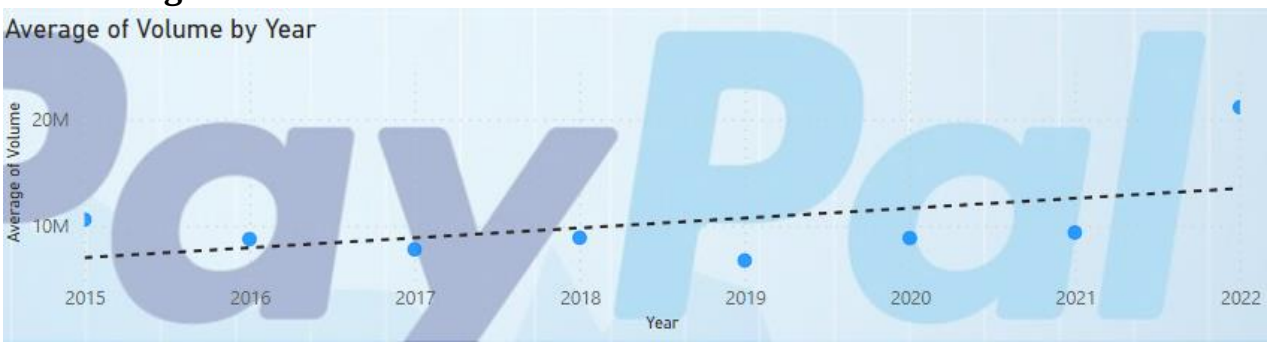
Figure 8: High Price Box Plot

Figure 9: Low Price Box Plot

Figure 10: Close Price Box Plot



Linear Regression



As seen in the scatter plot above, there is an incremental relationship between the years(independent variable) and the average of volume shares(dependent variable). As noted above it increases with the years, especially between 2020 till 2022 after Covid-19 because it is a growing market and expected to grow further in the future with all the technological advancements happening. We predict that in 2023 the average volume of shares will approximately be 10.5M.

KPI Analysis

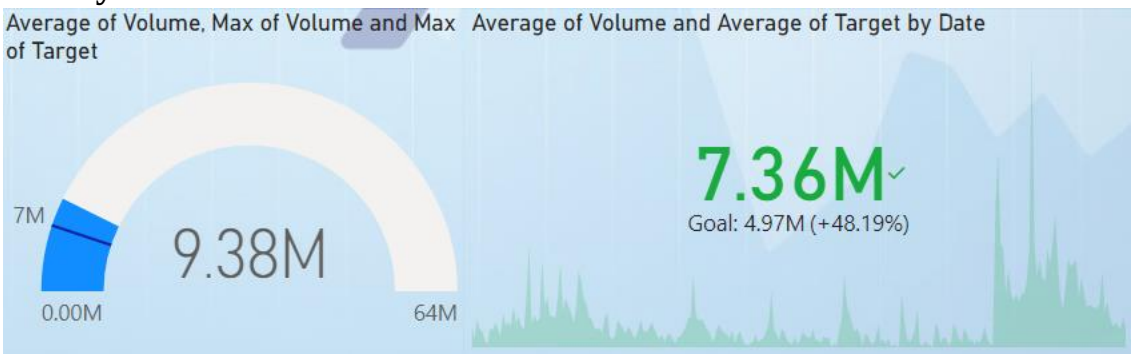


Figure 13: KPI Analysis

In this KPI analysis we measured the volume of shares with the target volume of shares. In this example we see the target reached in 2021 since it was the year that people knew and started to get more comfortable with buying stocks and trusting PayPal as a safe way to make Financial transactions online.

Brief Reporting and Examples

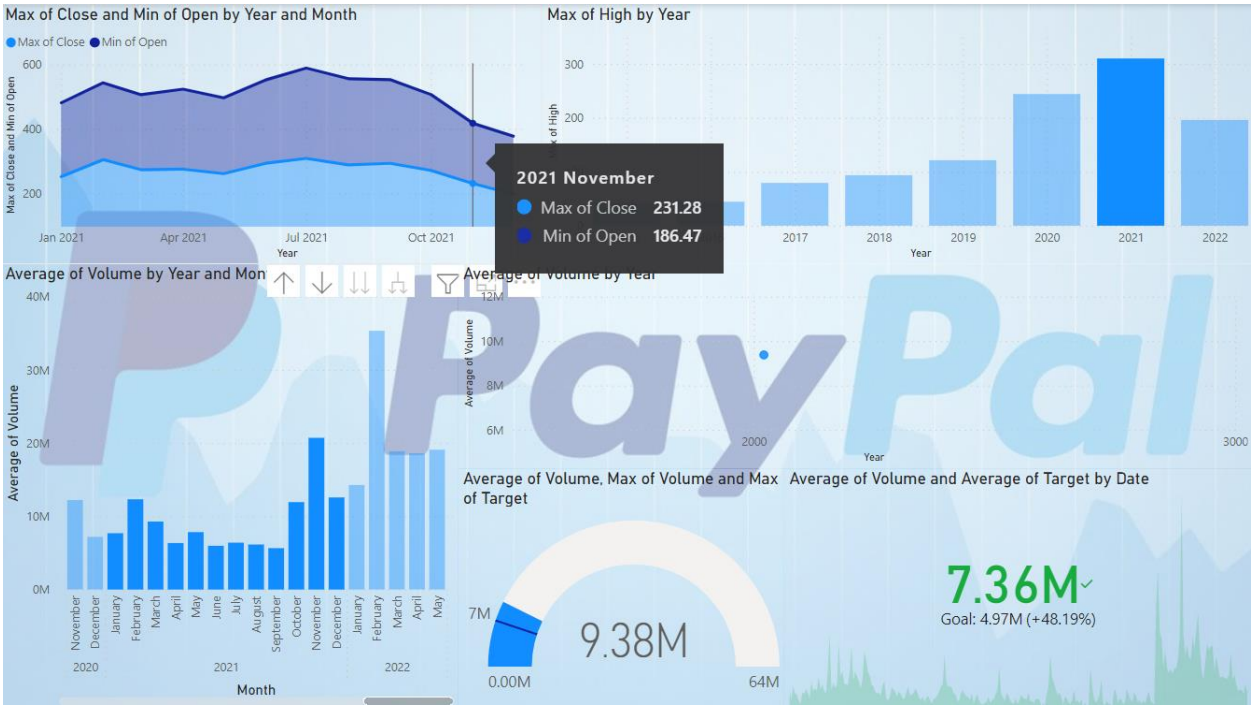


Figure 14: 2021 Example

As we can see in this example the average volume of shares increased drastically in November 2021. As a result of the decrease in price suddenly after a stabilization in the price of approximately 250 as the lowest price from January till November.

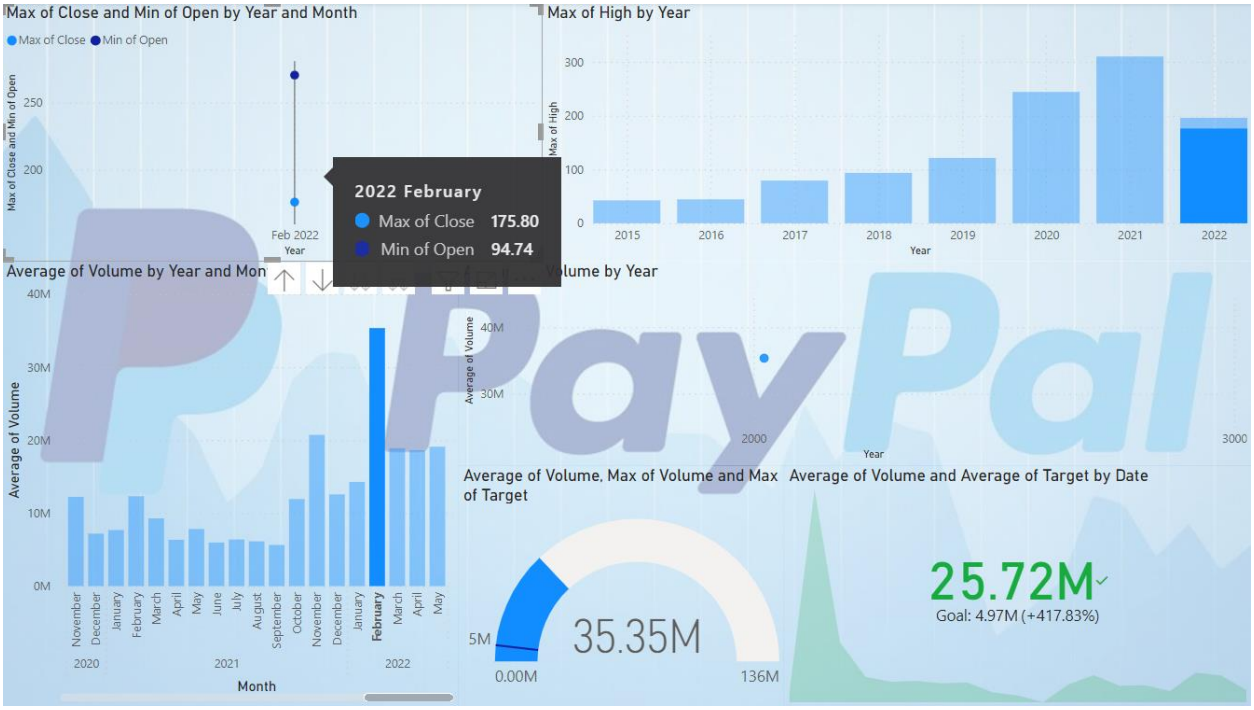


Figure 15: February 2022 Example

We noted in February 2022 we had the highest average of shares of all years and months. It was a result of the price drastically decreasing to a price of 94.74 dollars. As a result, we reached our target and exceeded it. So, there is a direct relationship between the price of the stock and the volume of shares.

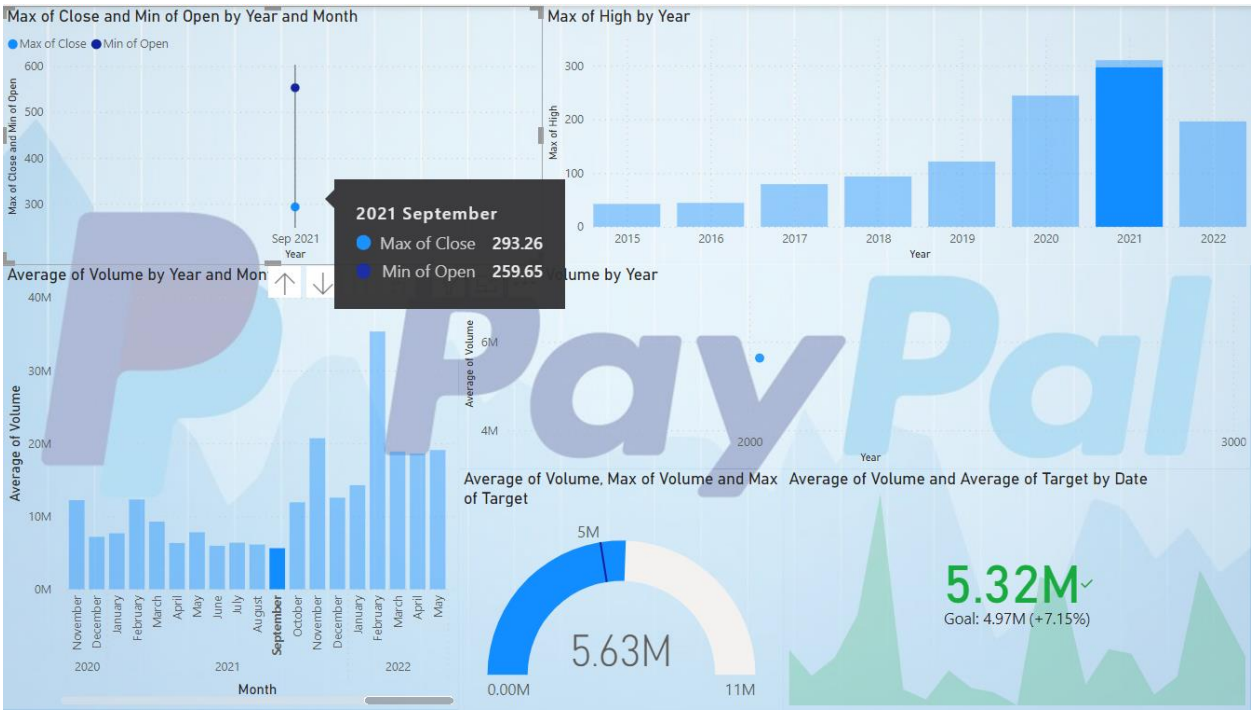


Figure 16: September 2021 Example

In contrast, here is an example of when the price of stock was very high. We noted in September 2021 we had the lowest average of shares of all years and months. It was a result of the price being very high with a price of 293.26 dollars. As a result, we barely reached our target.

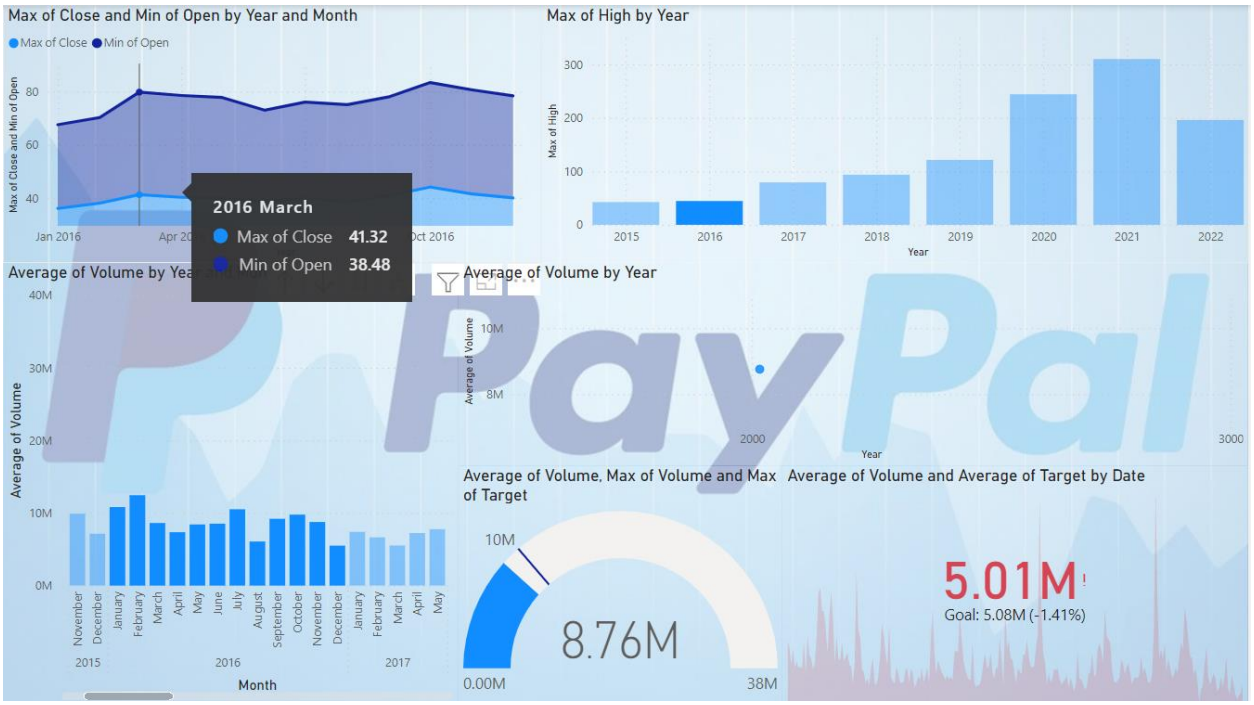


Figure 17: 2016 Example

In this example, we note in 2016 we did not reach our target. Even though the price was very low throughout the year we did not have the expected volume of shares. After doing our research, we saw that in 2016 people did not trust the concept of FinTech in general resulting in that low amount of shares. More specifically, people did not use PayPal as much as they do from 2020 onwards.

Dashboard Design

The intended use of this dashboard is to help strategic people make the desired decision on whether to invest in PayPal stocks and to visualize its growth as a market.

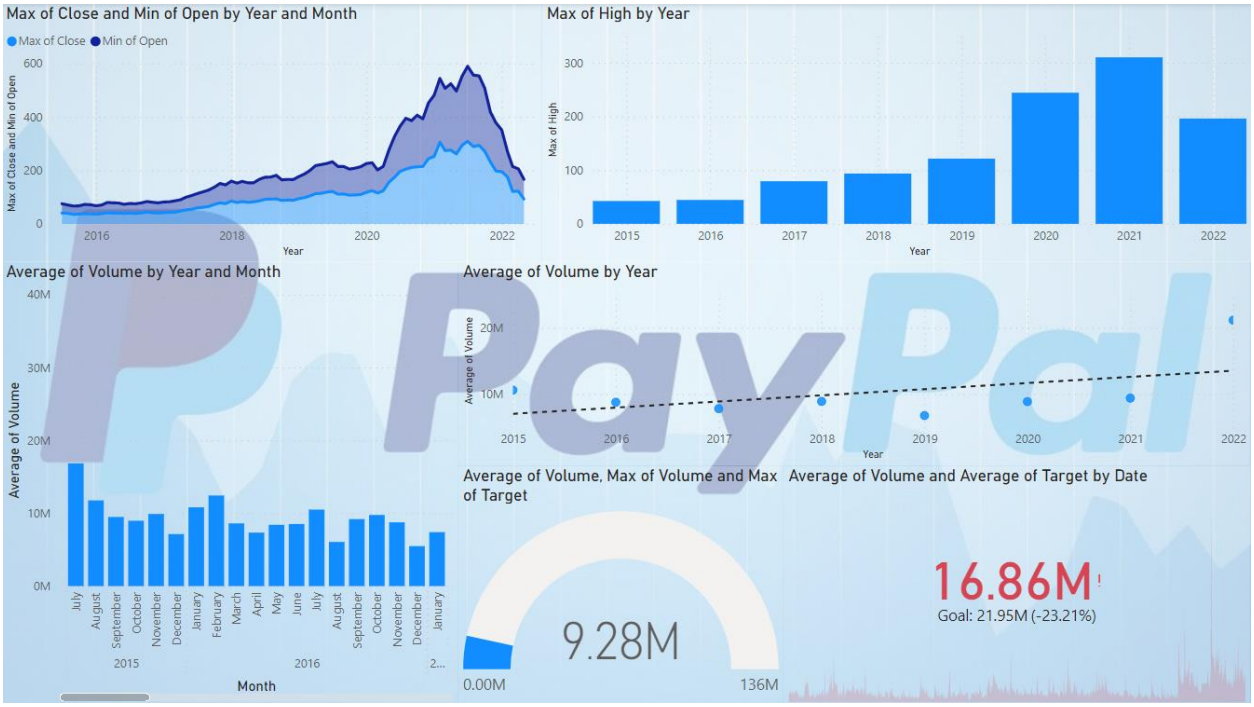


Figure 18: Dashboard Design

References

- *Kagan, J. (2023). Financial Technology (Fintech): Its Uses and Impact on Our Lives. Investopedia. https://www.investopedia.com/terms/f/fintech.asp*
- *PayPal. (n.d.). How does online payment processing work? PayPal. https://www.paypal.com/us/brc/article/how-online-payments-processing-works*
- *Business intelligence: A complete overview. (n.d.). Tableau. https://www.tableau.com/learn/articles/business-intelligence*
- *Damini Dutta, (30 April 2022), Fintech With An Approach Of Business Analyst*
- *Ibrahim A. Zeidy ,(n.d.) The Role Of Financial Technology in Changing Financial Industry and Increasing Efficiency in the Economy.*
- *Rajiv Dahiya, Son A.Le, John Kirk Ring, (February 2021), Big Data analytics and competitive advantage: the strategic role of firm-specific knowledge.*
- *David Varga, (November 2017), Fintech, the new era of financial services.*