

A Work Project, presented as part of the requirements for the Award of a Master's degree in  
Finance from the Nova School of Business and Economics.

Investment rationale of Robinhood users in the aviation industry during the COVID-19 pan-  
demic

Christian Lenz (44652)

Work project carried out under the supervision of:

Virginia Gianinazzi

20/05/2022

## **Abstract**

This paper was conducted to analyze the factors that influence retail investors of Robinhood in their decision to invest in airline securities. One can see that retail investors buy airline stocks even though the price is constantly going down. The analysis includes factors such as COVID-19 cases, daily flights, the stock price, Google Search Volume, and mentions in the Reddit forum WallStreetBets to figure out what exactly leads retail investors in investing. We conclude that retail investors seem to be focusing on Google as their primary search tool and less on the Reddit thread.

**Keywords:** Social Media, Fintech, covid-19, Robinhood, Retail Investors, aviation

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

# Table of Contents

<b>1 Introduction .....</b>	<b>1</b>
<b>2 Literature Review .....</b>	<b>3</b>
<b>3 Data and Sample.....</b>	<b>7</b>
3.1 Main Dataset.....	7
3.1.1 Sample 8	
3.2 Further Datasets.....	8
3.3 Limitations.....	9
<b>4 Analysis of influential factors during the pandemic in the aviation industry .....</b>	<b>10</b>
4.1 Robinhood retail investors.....	10
4.2 Stock Prices .....	13
4.3 Volume .....	14
4.4 COVID-19 .....	14
4.5 Daily Flights .....	16
4.6 Google Trends and Reddit WallStreetBets.....	17
<b>5 Empirical Analysis.....</b>	<b>20</b>
5.1 Methodology.....	20
5.2 Results .....	22
5.3 Limitations and considerations.....	23
<b>6 Conclusion.....</b>	<b>24</b>
<b>7 References .....</b>	<b>26</b>
<b>8 Appendix .....</b>	<b>1</b>
8.1 Tables .....	1
8.2 Figures .....	6

Table 1: List of Airlines (by market capitalization) as of 15.05.2020 .....	1
Table 2: Descriptive Statistics AAL of users holding.....	1
Table 3: Descriptive Statistics JETS of users holding .....	2
Table 4: Top 10 Holding of U.S. Global Jets ETF (JETS) as of 31.03.2022.....	2
Table 5: Aircraft types used for dataset OpenSky Network.....	3
Table 6: VIF for AAL .....	3
Table 7: VIF for JETS .....	4
Table 8: OLS Regression Results for AAL.....	4
Table 9: OLS Regression Results for JETS .....	5

Figure 1: Retail investors holding the top 10 airlines by market capitalization.....	6
Figure 2: Retail investors holding the American-based airlines by market capitalization..	6
Figure 3: Retail investors holding American Airlines (AAL) and U.S. Global Jets ETF (JETS).....	7
Figure 4: Stock prices of all airline stocks .....	7
Figure 5: Stock prices of American-based airline stocks.....	8
Figure 6: Stock prices of AAL and JETS.....	8
Figure 7: Stock Volume of all airline stocks.....	9
Figure 8: Stock volume of all American-based airline stocks .....	9
Figure 9: Stock volume of AAL and JETS .....	10
Figure 10: COVID-19 Cases worldwide from 01/2020 until 08/2020.....	10
Figure 11: Comparison daily flights worldwide between 2019 and 2020 .....	11
Figure 12: Daily flights worldwide from 01/2020 until 08/2020.....	12
Figure 13: Daily flights worldwide from 01/2020 until 08/2020 split up by airline.....	12
Figure 14: Google Search Volume for 01/2020 until 08/2020 for AAL.....	13
Figure 15: Google Search Volume for 01/2020 until 08/2020 for JETS .....	13
Figure 16: ADX for AAL.....	14
Figure 17: Reddit WallStreetBets – Mentioning of ticket “AAL” from 01/2020 until 08/2020 .....	14
Figure 18: Reddit WallStreetBets – Mentioning of ticket “JETS” from 01/2020 until 08/2020:.....	15

## 1 Introduction

The COVID-19 virus has brought drastic changes to our daily life since the beginning of December 2019 (Devlin 2021). To contain the spread of the virus, measurements were implemented, such as keeping social distance and staying at home. So, what did individuals do in the meantime to keep themselves busy at home? Likely many found new hobbies such as gardening, but also other large areas of interest emerged during their time at home, namely investing (Bizouati-Kennedy 2021). Many people were suddenly looking for platforms to invest in and found the US-based brokerage platform Robinhood (RH). Robinhood was founded in 2013 and is a financial firm that offers commission-free trade of different asset classes such as stocks, exchange-traded funds (ETF), and cryptocurrencies. Retail investors do not depend on the market hours and a broker from a financial institution anymore. Robinhood's offering is based on a mobile application, and retail investors can trade flexibly. It was a convenient solution during COVID-19 restrictions. Smartphones have been implemented in our daily life for a long time, many Americans use their cellphone daily, even within 30 minutes of waking up and 30 minutes before falling asleep (Mohr 2021). Consequently, a mobile-based application came into play. Investing can now easily be done without any barrier entries as it had before. Every individual with a smartphone can participate in the stock market. Solely in the first quarter of 2020, 3 million people entered the trading platform, increasing its total users to 13 million (Henderson and Kruppa 2020). In 2021, the platform reached about 17 million active investors (Robinhood Markets, Inc. 2022). On some trading days, Robinhood saw trading activity about triple the amount of 2019, even though the pandemic put pressure on the stock market (Rooney 2020). The platform has mostly attracted investors without any deeper knowledge of investments (Adamczyk 2020) and beyond that even those making their first trades on Robinhood (Huang 2015). Therefore, this paper will dive deeper into how retail investors have invested during the

pandemic time. Since there has not been any more detailed industry-specific research done so far, we will focus on the aviation industry. The stock of American Airlines (AAL) has jumped into the top ten traded stocks during the pandemic (Tenebruso 2020), up from attracting only a small base of investors, retail investors were suddenly interested in these so-called old economy stocks. Mainly due to the growing importance of social media within the retail investors' community, decisions are based on the activity within forums such as Reddit. The consequences can be seen in January 2021 when retail investors hyped GameStop and AMC, the so-called “meme stocks,” to jump to extraordinary heights (La Monica 2022), even creating their own language in communicating whether to buy or sell a stock. For our industry, the aviation industry, the expression “buying the dip” is especially of importance. Retail investors bought shares that dropped in price; later, we see this happened to our old economy stocks (S. Pagano, Sedunov and Velthuis 2020). Another example would be the sudden pop-up of new stock industries, such as the cannabis industry (Welch 2020). Hence, it can be essential to compare this particular investor behavior to the aviation stocks as well since they probably show similar behavioral structures such as attention-influenced buying and selling through herd behavior. It can be seen up to a point where stock trading has been more like gambling. Gamblers who typically bet on other platforms were moving to Robinhood as an alternative during the pandemic (Rooney 2020). It is even said that Robinhood amplifies this behavior by implementing features to keep users on the platform as some addictive behavior (Farivar 2021). However, the difference is that the aviation industry has already been trading on the stock market for quite a long time in comparison to cannabis stocks.

I will initiate this paper by introducing a general literature overview of the present results that have been concluded from the same dataset we use. Since the dataset is considerably recent and limited, there are not many academic papers available. Following this, I will continue to introduce the used dataset. I will be extracting the top ten airlines by market capitalization traded on

Robinhood and show how these stocks have been evolving. Since Robinhood mostly includes retail traders, I will further compare them to the whole market by using data from an API from Yahoo Finance, such as the stock price and volume. I will dive deeper into what kind of information retail investors based their decision on in the aviation sector. A crucial factor is the number of COVID-19 cases; hence I will represent how the cases evolved. Since our analysis is focused on the aviation industry, daily flights will also be introduced. Following up we will concentrate more on where retail investors received their information from, using data from Google Trends and Reddit. In our empirical study, we will join all the factors together and figure out what role they played and how significant they were in our observed period and data. Further, we will discuss these findings and give an outlook and possible additional research to be done.

## **2 Literature Review**

In the following, I will give an overview of the papers that have already used our primary dataset, Robintrack, to extract results about Robinhood and its users. In addition, I will include papers, giving insights into the behavior of Robinhood users and likewise papers that can be related to my paper.

In general, the topic of research about retail investors is still relatively new. The first research has generally displayed retail investors as noise traders (Campbell, Y. and Kyle 1993; Barber, Odean and Zhu 2006). Noise traders are investors who decide on their investments without any other information input or causes (Foucault, Sraer and Thesmar 2011). It has already been recognized early that retail investors work with limited information, apply easy decision rules, so-called “rules-of-thumb” and are psychologically biased (Shefrin and Statman 1984). In general, retail investors were seen as suboptimal investors who take on an excessive amount of risk in the past (Mankiw and Zeldes 1991).

Jumping forward to more recent research on retail investors with a focus on our main dataset from Robintrack. Welch (2020) has already concluded this in his paper “The Wisdom of the Robinhood Crowd.” Retail investors have been mainly attracted by stocks with a recent higher trading volume in the last 12 months, thus being drawn to either new stocks such as in the cannabis industry or mature industries such as the one we are focusing on, the aviation industry. As I was assuming, aviation stocks came more into the focus due to the media coverage as a result of the ongoing uncertainty of travel restrictions and hence resulted in higher trading activity. Nevertheless, in the observed period, the Robinhood investors did perform well compared to capital market benchmarking models. Barber et al. (2020) researched the investors' crowd of Robinhood from a behavioral perspective. They concluded that the decisions of retail investors of Robinhood were mostly driven by attention-based buying and herding behavior. Such as more volatile stocks, mostly with higher historically positive returns or trading volume, will be bought; however, experience negative returns afterward. This additionally also holds for more formal information such as earnings announcements as by Friedman and Zeng (2021). The paper shows that Robinhood investors also react stronger to earnings announcements. The before-mentioned behavior has already been concluded by Barber and Terrance in another paper in 2005. Barber and Terrance (2005) demonstrate that attention-based buying might be due to too many alternatives in the investment atmosphere and consequently high search costs. Going with the herd eases decisions and alternatives, such as buying stocks on high-volume trading days or stocks with high media coverage, followed by either extremely positive or negative one-day returns. Institutional investors are not that prone to attention because of having more possibility to restrict their universe of alternatives. Glossner et al. (2021) on the other hand show that retail investors were more drawn to companies with high leverage and low cash levels, so industries were vulnerable during pandemic times. Institutional investors did the contrary and increased their proportion in financially stronger firms with lower leverage and increased

cash holdings. However, not only retail investors are biased by herding behavior, but also institutional investors (Nofsinger and Sias 1999, Sias 2004). Retail and institutional investors seem to act as market contraries in times of falling market prices, with retail investors providing liquidity for more distressed firms, as shown by Barrot, Kaniel, and Sraer (2016) and Hendershott, Menkveld and Praz (2022). The results are similar to Eaton et al. (2021), who also used Robintrack to research that Robinhood investors increase liquidity and lower return volatility. Especially in times of negative market shocks, the role of retail investors in providing liquidity through contrarian buying might seem important. Since we have only been talking about stocks, we will be presenting additional papers on the asset class of Exchange Traded Funds (ETF). ETFs aim to replicate their benchmark indices, by either holding the securities respectively bonds or tracking them through derivatives of the underlying assets. The original idea of ETFs has been to increase the diversification of one's portfolio. During the pandemic, ETFs have also been more within the focus of retail investors. Recently, more ETFs have been set up that focus on a particular industry, hence are more specialized. These new ETFs have received much attention and, similar to the results of Barber et al. (2020) underperformed compared to their respective benchmarks, as concluded in the paper of Ben-David, Byungwook and Moussawi (2021). To further research this result, I will also be including a sector ETF in my analysis.

A measurement to reduce search costs is the usage of Google. Da et. al (2011) show that search frequency can be seen as a measurement of investor attention which becomes an important tool in retail investors' decision-making process. By using Google Search Data, they show a correlation between the investor's attention and the data from Google. However, they conclude that the attention is mostly picked up by uninformed investors. Further, if investors are researching a specific earning announcement, the stock and volume become less volatile after the

announcement since more information seems to be reflected in both measurements (Drake, Roulstone and Thornock 2012). Amihud (2002) found a positive relationship between higher Google search volume and the liquidity of a stock which is in line with the result of Eaton et al. (2021) about Robinhood investors. The same result about the relationship between stock turnover and the volatility of a stock has been concluded by Yuan (2015). Lastly, Preis and Stanley (2013) used Google search queries as a trading strategy to outperform the market based on buying and selling signals of stocks. Throughout time and especially throughout the pandemic, retail investors have been growing stronger and stronger. They have been seen as a non-significant party in the past, however, recent studies by Bohmer et. al (2016), Moss et. al (2020 and Pagano, Sedunov and Velthuis (2020) show the opposite. They concluded that retail investors can be a driving force within unnormal market times, providing liquidity and influencing the variance of securities.

This is mostly due to neobrokers such as Robinhood which simplifies access to individuals to the stock market (Barber and Terrance 2005). In addition, through internet forums such as the Reddit thread WallStreetBets, retail investors were able to collect themselves and act strategically such as on the “meme stocks” (Uusi-Mikkola 2021). It is not new that social media is a valuable source of information for relationships between economic or financial factors. A sentiment analysis executed by Daas and Puts (2014) showed the narrow relationship between social media and the consumer index. In the recent past, there has been researched WallStreetBets, most research focused on sociopolitical aspects such as their contrary position against the establishment or the Wall Street (Di Muzio 2021) or about meme stocks in general including the short squeeze of GameStop (Lyócsa, Baumöhl and Vÿrost 2021; Anand and Pathak 2021; Winkler and Semenova 2021). Since the phenomenon is very recent, not much research of a longer period has been conducted. Buz and de Melo (2021) were one of the first

to conduct research on how retail investors perform against market indices by analyzing their buying and selling behavior, though that building a WallStreetBets portfolio. They showed that the portfolio could have outperformed the S&P 500 for the observed time.

As seen in the literature review, there has only been general research made about retail investors, however, lacking industrial-specific research. Retail investors seem to have a favor for firms that were decreasing in price as in “buying the dip”. Hence, it might seem essential to differ between these industries. Additionally, not much research has been carried out with the inclusion of added information sources such as Reddit. Most research focused on more classic information sources such as Google. We will complement the already existing research in this way.

### **3 Data and Sample**

In this section, I specify what datasets will be used and how the sample was picked. The datasets have been extracted from different sources.

#### **3.1 Main Dataset**

The main dataset is sourced by the website Robintrack.net which used an Application Programming Interface (API) to pull data from Robinhood. The API was made available by Robinhood itself to the public. The data shows how many users of the platform Robinhood hold a specific stock over time. It was updated frequently from about May 2018 until August 2020, as described by Robintrack.net. In August 2020, Robinhood shut down the API out of the concern that third parties might use the data against their clients (Ponczek 2020). However, the dataset can still be downloaded at the website Robintrack.net or via third parties such as Kaggle.com. The stocks are downloaded separately in CSV files, totaling about 8,579 files; hence data has been collected of 8,579 securities.

### **3.1.1 Sample**

In contrast to the above literature, we focus on the aviation industry and hence do not need the complete dataset. We will be focusing on the largest aviation securities sorted by market capitalization. Further details will be delivered in paragraph 4.

### **3.2 Further Datasets**

To analyze the relation of the development of COVID-19 with the plunge in air traffic, a dataset providing information on cases per area between January and August 2020 will be required. By accessing supporting secondary literature and data on the COVID-19 cases and social distancing measurements, we should be able to draw a comparison between the development of COVID-19 cases, restrictions, and air traffic and analyze the relationship between these three factors and understand the strength of the impact of the pandemic on air traffic. Hence, we will be using the coverage of the COVID-19 cases worldwide released by the European Centre for Disease Prevention and Control (ECDC) and travel restrictions by country released by data.humdata.org. The second additional database utilized for the analysis is “Crowdsourced air traffic data from The OpenSky Network 2020”, which aims to illustrate the development of air traffic during the pandemic. On their website, they provide monthly CSV files for the time frame of 1st of January 2019 to the 8<sup>th</sup> of August 2020. I decided to use this dataset because it is built on the comprehensive dataset of the OpenSky Network, which was started by the University of Kaiserslautern, the University of Oxford, and armasuisse. The crowdsourced dataset only provides airline and airport codes. To properly assign the codes to countries and regions, supporting datasets providing the translation of the code would be needed. Therefore, the following datasets will be used: Airline database: assigns airline codes to the respective airline: <https://www.kaggle.com/open-flights/airline-database>.

For financial data, such as stock volume or price, an API by Yahoo Finance implemented into our Python interface will be used. Furthermore, to analyze Google Trend data, an unofficial API will be taken advantage of. The data for the WallStreetBets will be scraped by the API Pushshift made available within a project by Baumgartner et al. (2020).

### **3.3 Limitations**

The first limitation is that Robinhood does not cover all stocks, only the ones traded on Robinhood. The stocks are mostly American-based, creating a geographical bias and not representing all regions correctly. Secondly, some tickets are not displayed or are later displayed due to failures such as bugs that have been fixed later, as mentioned by Welch (2020). Furthermore, there is some missing data because Robinhood had system outages on the 2<sup>nd</sup> and 9<sup>th</sup> of March 2020. The algorithm of Robintrack itself additionally failed to run on two-time periods, namely the 9<sup>th</sup> of August 2018, and between the 24<sup>th</sup> until the 29<sup>th</sup> of January 2019, and the 7<sup>th</sup> until the 15<sup>th</sup> of January 2020. Unfortunately, the data also does not differ between buyers and sellers; hence one can observe the net results out of both. A slight limitation is that Robinhood continuously offered a promotion for new referred accounts who would receive a free stock upon registration. This could slightly falsify our shown figures since users are added to the holding data without buying the security. Another note for our COVID-19 dataset: Analysing the chart also shows that the dataset with COVID-19 cases is not comprehensive as it shows zero new cases for some days in January. Considering that the pandemic outbreak was in December, the dataset did not record every case. Nevertheless, this is presumably because the official recording of COVID-19 started when the magnitude of the pandemic was visible. Of course, minor differences can appear due to different record-taking of the cases throughout countries. In addition, the dataset of OpenSky Network can minimally differ from real data such as limitations of ADS-B data, meaning the receivers sometimes have no coverage or are not allowed to record

flights due to different regulatory environments (Strohmeier et al. 2021). The strongest limitation of Google Search data is that the data is normalized through a formula that is not public. It is questionable consequently how comparable the data from different search queries are. The API used for WallStreetBets, Pushshift, has the usual limitations of APIs such as downtime. Further, there is some delay in storing the data. Since we are not using real-time data, this is not an issue.

#### **4 Analysis of influential factors during the pandemic in the aviation industry**

In the following, I want to analyze further what led retail investors to start focusing more on the aviation stocks, an industry in which stocks are classified as old economy stocks and have not been at the center of attention before the pandemic at all. I will start by providing a glimpse of the main dataset of Robintrack. Following will be factors such as the development of COVID-19 cases and daily flight development during the pandemic. Since retail investors are attention-focused investors (Barber, Huang et al. 2020), we will also include how aviation stocks were displayed in the news by looking at Google Trends. Retail investors have also been heavily influenced by communities such as WallStreetBets on Reddit (Buz and de Melo 2021). That is why we will, finally, include this Reddit thread in our analysis.

##### **4.1 Robinhood retail investors**

First, I want to give an overview of how the aviation stocks listed on Robinhood have been evolving. For better readiness, I picked the top ten airlines by market capitalization out of all the publicly listed airlines in the world as of the 15<sup>th</sup> of May 2022 listed (Companies Market Cap 2022). I produced the following list seen in table 1, with Delta Airlines as the largest airline with a market capitalization of 27.4 billion USD and Copa Holdings, the company behind the Panamanian airline Copa Airlines and its subsidiaries, as the smallest title with a market

capitalization of 3.11 billion USD. I extracted these CSV files of the Robintrack dataset and read them into the Python interface JupyterLab. We plotted these airlines by using the Python library matplotlib. The plot can be seen in figure 1 with the date on the x-axis and the number of users holding the stock on the y-axis. As already written previously, Robintrack had some outages due to updates in January of 2019 and 2020, which can also be seen in our chart. Consequently, for a few days, there is no data. However, it did not have a significant impact since it was before the pandemic started. Interestingly the stocks with the largest market capitalization have not experienced the largest increase of retail investors holding the stock. As already mentioned before, American Airlines has been in the top ten tracking stocks for some time during the first peaks of the pandemic. This is also shown in figure 1 with a huge increase in users holding the security. Although being one of the largest airlines in the world, the non-US airlines do not seem to be playing any role for the retail investors of RH. This might be due to the equity home country bias puzzle when investing in stocks, as identified by Bernanke and Rogoff (2001). Investors typically tend to overweigh stocks from their own country into their portfolio. As Robinhood is an American-based platform and only individuals who are “[...] a US citizen, US permanent resident, or have a valid US visa” (Markets 2022) can sign-up, the major group is also Americans. Thus, it is not surprising that American-based Airlines lead the percentual change of users holding the stock. Consequently, I will be excluding all non-US airlines. To have a better overview of all the six airlines left I attached figure 2. It is seen here that all graphs have a similar course. Therefore, I decided to focus on American Airlines further. Additionally, AAL has been the only aviation stock trading in the top ten of Robinhood for a while (Tenebruso 2020). This reduces the complexity of the overview of all used graphs and statistical measurements used. For verification purposes, I will do all the analyses for the other airlines as well and attach them with the code. Additionally, American Airlines has been the airline that has been in the focus of retail investors, which is surprising since airlines are mostly capital

intensive with respectively low margins (Onvista 2020). However, it fits the results shown in the literature review of (Glossner et al. 2021). A descriptive statistic of the data frame of American Airlines can be found in table 2. The focus lies on American airlines; however, ETFs have also been on the rise during the pandemic, especially since there has been a higher bias from international ETFs to national ETFs (Yousefi and Najand 2022). ETFs are typically an investment vehicle for passive investors as they normally offer low costs and liquidity (Ben-David, Franzoni and Moussawi 2017). The most common ETFs are index funds, meaning they hold the same securities as their respective index. However, in recent times, specialized ETFs have been becoming popular instead of an index following the development of securities of, for example, a certain industry. In the aviation industry, the ETF *U.S. Global Jets Index (JETSX)*, managed by the investment manager US Global Investors, has been in place since the 28<sup>th</sup> of April 2015 and follows major securities of the global airline industry with a clear focus on the American market. As this ETF is also traded on Robinhood and includes most of the airlines we also had in our list as in table 1, I include this sector ETF in my analysis. The US Jets Index includes securities of “the global airline industry, including airline operators and manufacturers from all over the world” “with an emphasis on domestic passenger airlines” (U.S. Global Investors 2022). Looking at table 4, a listing can be seen of all the holdings as of the 19<sup>th</sup> of May 2022. The focus undoubtedly lies on airlines, with United Airlines being the biggest security holding with a total proportion of 12.35%. Currently, according to the factsheet of the industry ETF, airlines have been weighted with 75.99% of the index with a holdings number of 51, the weighting is the market capitalization and fundamental factors based. As shown in figure 3 one can see that development of users holding the securities of American Airlines and JETS have a similar course. This is also confirmed by the correlation index of 0.999262. However, there are some specific peaks, I want to point out before moving on. In May 2020, the stock and the number of users holding stumbled because AAL announced huge layoffs of its staff

(Fonda 2020). Following, in June 2020, the airline industry received an upwind by relaxed pandemic restrictions and rising demand, hence US airlines increased their capacity again giving out a positive signal to analysts (Bushey and Staton 2020).

## **4.2 Stock Prices**

Next, we will further dive deeper into the development of the stock of American Airlines and U.S. Global Jets ETF. Since our dataset from Robintrack does not show up its prices among the users holding the security, we will take access to Yahoo Finance and implement the stock prices of our securities. Again, we will start with the same approach as in the previous paragraph and look at the largest ten airline companies listed on Robinhood by market capitalization. The result can be seen in figure 4. As before, all the airlines have a similar development of stock prices which the correlations also confirm. Since we are focusing just on the US-based airlines, we drop out the non-US securities. Likewise, the stock prices moved very similarly during the pandemic; just at the beginning of 2020, United Airlines had a steeper decrease than its peer group. Continuing our scheme from 4.1, we will be focusing on American Airlines. As we are also interested in how the sector ETF developed, we also add U.S. Global Jets to figure 6. As we expected, both securities crashed after the pandemic rapidly spread around the world and countries began to ban air travel. Hence, investors kept on selling the stock. This behavior is contrary to what Robinhood investors did. More and more users were adding up to the ones holding the securities. This is also shown by the correlation between the chart of stock prices and securities with a correlation of  $-0.685961$  between the Robinhood users holding AAL and the closing price of AAL, respectively  $-0.669534$  for the industry ETF U.S. Global Jets. As mentioned before, analysts received positive signals by growing demand and increasing capacity in June 2020 (Bushey and Staton 2020).

### **4.3 Volume**

Volume is another metric I want to include since it comes the closest to our definition of users holding our main dataset of Robintrack, and we can therefore compare how retail investors traded the stock in comparison to how many people or how many stocks were traded on the overall market. Volume is defined as the following in Yahoo Finance “volume measures the number of a stock’s shares that are traded on a stock exchange in a day or a period of time” (Kristopher 2014). I follow the same approach to pick my sample as in 4.2 and 4.1. The results are shown in Figures 9, showing the volume during the pandemic for American Airlines and for the U.S. Global Jets. Both have a similar course; however, they are less correlated than their closing prices. The correlation of both volumes amounts to 0.877462. With the beginning of the pandemic, the securities were more heavily traded than in the pre-pandemic. Again, a peak can be seen in June 2020 due to positive signals in the airline industry as written previously. However, just by looking at the graph of volume, we can not tell if there were more buyers or sellers. Considering that the price dropped, there were more sellers than buyers. It seems like the overall market behaved contrary to the retail investors of Robinhood, who kept on holding more of the securities as already shown by Pagano, Sedunov and Velthuis (2020) for the airline and cruise line stocks.

### **4.4 COVID-19**

The COVID-19 pandemic that started in December 2019 in Wuhan, China (Devlin 2021) and was declared as a pandemic on the 30<sup>th</sup> of January 2020 by the WHO, rapidly spread across the world within no time (Cucinotta and Vanelli 2020). Air travel was one of the main drivers of rapid diffusion in the world (Ribeiro and Reis 2020). Hence, the first step to slow down the virus was to stop worldwide air travel. As lockdowns started in place at the beginning of March in Europa and the announcement of Donald Trump that the United States bans all air travel

from the Schengen Area on the 11<sup>th</sup> of March 2020 (Ghosh 2021), air travel was one of the most impacted industries by the pandemic with an immense increase in volatility (Baek and Mohanty 2020, Garrow and Lurkin 2021) with the US indices having one of the highest volatilities behind the Brazilian indices (Ganie, Wani and Yadav 2022). No other crisis before, neither 9/11 nor other pandemic situations such as SARS have resulted in such a vast crisis for airline companies (Garrow and Lurkin 2021). The revenues of industry revenue amounted to 328bn USD in 2020, signifying a 60 percent decrease compared to the last year, 2019. The consulting company McKinsey projects in their report that the airline industry will not recover before 2024 (Bouwer, Saxon and Wittkamp 2021). Nevertheless, even if revenue and passenger numbers will return to normal levels, the pandemic will change the airline industry post-pandemic (Faura and Ursula 2021).

Faura and Ursula (2021) assume that restrictions to avoid the spread of the virus are closely linked to the financial and operational impact on an airline. As the United States of America has always had one of the highest case numbers relative to its population shown by ourworldindata.org and hence one of the strictest measurements in place, according to the Oxford Stringency Index, the impacts for US airlines were huge (Financial Times 2022). Since most countries decided how to act on restrictions based on their covid cases, it might be helpful to take a distinctive look at how COVID cases evolved and how the retail investors reacted to these changes. Ganie, Wani, and Yadav (2022) already examined in what range an increase in COVID-19 cases impact the stock prices. They concluded that for US-based airlines, an increase of 1% in COVID cases results in roughly 0.1% price drop with further negative implications in the days after the cases surge. In figure 10, one can see the development of COVID cases worldwide between January and August 2020, with the first cases being officially recognized by the United States health officials on the 20<sup>th</sup> of January 2020 and former President

Donald Trump declaring the pandemic situation on the 31<sup>st</sup> of January 2020 (AJMC 2021). Just 4 days later, on the 24<sup>th</sup> of January 2020, the first cases in Europe were reported in Bordeaux, France (European Centre for Disease Prevention and Control 2020). The World Health Organization (WHO) consequently declares the COVID-19 outbreak a pandemic on the 11<sup>th</sup> of March 2020 (Cucinotta and Vanelli 2020). With the rapid diffusion of the virus worldwide, the cases rapidly rise to above one million at the beginning of April 2020. In figure 10, you can see that from then on, the cases consistently increased up to 30 million cases which have been reached in June 2020. As we have said earlier, with increasing COVID cases, more restrictions were implemented. Therefore, air travel was also stopped, and hence stocks of airlines were heavily impacted (Ganie, Wani and Yadav 2022). Therefore, many investors have sold airline stocks, and the price crashed. However, when comparing retail investors of Robinhood, you can see the number of users holding airline stocks increasing consistently despite price crashes. We will further analyze other factors that made retail investors decide likewise.

#### **4.5 Daily Flights**

For this paragraph, we take advantage of our dataset from the OpenSky Network, which “is a non-profit community-based receiver network which has been continuously collecting air traffic surveillance data” (OpenSky Network 2022). After cleaning the dataset for Not a Numbers (NaN), missing data such as callsigns and smaller aircrafts such as business jets and helicopters, we can evaluate how the daily flights developed overtime during the pandemic. For my analysis, I have only used the main aircraft types used in commercial aviation listed in table 5. We are firstly using the data for 2019, pre-pandemic, and 2020, to compare both years side-by-side seen in figure 11. The impact of COVID-19 on the airline industry can be seen. While the daily flights within 2019 remained relatively stable, with a slight increase throughout the year, the number of flights in 2020 plunged significantly at the beginning of the pandemic. Governments

started to impose restrictions, and airlines were not allowed to operate as before the pandemic. Starting from higher figures than in 2019, from up to 60,000 flights per day, it crashes to its lowest daily number of flights of about 10,000. In the fourth quarter of 2020, however, the industry gained ground again, recovering to up to 30,000 flights per day. The IATA estimated that at the peak of the pandemic, only roughly 5% of the air travel in comparison to one year before was left (IATA 2020). This number could be even lower since airlines were forced to do so-called “ghost flights” to keep take-off and landing slots (Neslen 2022). Consequently, many airlines were forced to ground their fleet reaching a ground level of having 2/3 of the world’s fleet on the ground or in numbers more than 16,000 airplanes being grounded, even being parked on runways such as in Frankfurt International Airport (Garrow and Lurkin 2021) (Kotoky, Stringer and Saxena 2020). Including the securities in our analysis, we can see that as fewer flights are in the air, more retail users are holding the security, resulting in a negative correlation between users holding AAL and the daily number of flights for our period of -0.739673.

#### **4.6 Google Trends and Reddit WallStreetBets**

As Barber and Odean (2007) already concluded, retail investors are influenced by attention-induced trading and due to high search costs, mostly decide on what to buy or sell on the news or their herd group. This rationale is supported by the research findings of Kumari (2022), who conclude that social media plays a bigger and bigger role for retail investors and their decisions. The best examples are the so-called “meme stocks,” which evolved from the Reddit thread WallStreetBets. These are stocks that gain substantial price increases solely through an artificial hype within one more multiple social media platform. One of the first examples is the company GameStop (GME), a US-based store for (used) videogames which have been under pressure for quite a time due to decreasing profitability as e-commerce companies were taking their

customers (Wu 2019). That is why many institutional investors were selling the security short. However, retail investors decided to the contrary in January 2021. They bought the stock until it reached a maximum of roughly \$483 per stock, from a closing price of about \$17 at the beginning of January. Retail investors on Reddit wanted to fight short-sellers, finally ending up in a “short-squeeze” (Hiltzik 2022). A short squeeze is when a stock rapidly increases in price and short sellers must close their short positions rapidly; they get “squeezed” out of the market. Another famous example would be the cinema operator AMC Entertainment or Nokia. Robinhood even restricted trading the GameStop stock for a period, leading to critics from all sides (Adamczyk 2021). In the literature review, we have already shown that herding behavior has been intensively researched for retail investors (Merli and Roger 2013; Hsieh, Chan and Wang 2020; Li, Rhee and Wang 2017; Barber and Terrance 2005). Aloosh, Choi, and Ouzan (2021) further show that herding behavior also exists for meme stocks.

First, we are going to examine the most standard search tool, Google. Google Keywords research has been used in previous studies to find trends in various markets. Bordino et al. (2012) showed a positive correlation between the search interest on Google and the daily trading volumes of the NASDAQ-100. Similar research has been done on the link between Google Trends and abnormal returns of the S&P 500 (Joseph, Larrain and Turner 2017). Research related to retail investors and hence crucial for my paper has been carried out by Da, Engelberg and Pengjie (2011). They found out that Google Trends is a measurement tool for the attention of retail investors and is correlated with different proxies of investors’ attention. Following it makes sense to use Google Trends as a first indication of how they are correlated with our aviation securities. Google is most likely for a major group of retail investors the first source of information since they rarely have access to more professional tools such as Bloomberg. Hence, also to figure out if one can predict stock increases or decreases just by looking at the

development of Google searches of the tickets of our chosen securities. For our analysis, we will continue using American Airlines and the sector ETF U.S. Global Jets. In figure 14, one can see the Google Search Volume (GSV) for the Ticket of American Airlines. We retrieved the data of Google trends by using the API pytrends with which one can download data by keywords. To be notified is that Google uses an index to measure their trends, such as that the number represents the searches relatively to the highest search interest for the chosen time. The highest value is 100. If the value is 0, then there is not enough data available. For example, if the value amounts to 50, the keyword was only half as popular. Both spikes happened in the same period. You can further configure different settings, such as the region; in this case, we will focus on the searches from the US area. The same observation can be made for the ticket of U.S. Global Jets. During the first half of 2020, the search for these tickers increased immensely. To do a second check, we can see how both charts are correlated with each other. We observe the time frame from the beginning of 2020 until Robinhood deactivated the API. As a result, we receive a correlation between the users holding the stock and the audience looking for the specific stock ticket of 0.2762 for AAL and 0.26931 for JETS, both highly significant to the p-value. Hence, as already seen in figures 14 and 15, the interest in searching the stock ticker was extremely high at the beginning of the pandemic but decreased over time with a respective low correlation. The interest in searching for the ticket was still high, but the stock was not traded that much anymore. This might be due to already a large group holding the stock and just observing how the stock is doing. Adding a further indicator that is often used for trend analysis is the Average Directional Index (ADX), developed by Welles Wilder in 1978. ADX is a technical indicator used in trading to get an indication of the strength of a trend. Values of ADX of over 15 imply a trend, and over 25 a strong trend. As shown in figure 16, the peak of the trend was reached in March, indicating ADX values over 45. After that, the trend decreased with a new height in June, as also already mentioned in the previous paragraphs. Since the

trends are very short-term, retail traders likely drove this trend. We could also be looking at other keywords, however, there can be too much cross interest, and just by using the ticket, we can be certain that the searches were related to the stock market. Especially for such a common name, such as American Airlines, which is also used by non-investors to book a flight, it will be different to figure out what group made the Google search, as already shown by Joseph, Wintoki and Zhang (2011). Furthermore, there are always different variations of company names, such as that American Airlines can also be named “AA”. For the ETF US Jets, it is even more critical since an investor does not even know about this specific ETF and will search for more general expressions. Since retail investors are luckily not that wise to directly search for the ticket, a further trend analysis might be useful.

Lastly, we will investigate data from the Reddit thread WallStreetBets. Since the short squeeze of GameStop, the forum has been the focus of retail investors. We collect our data with a script written to scrape comments and submissions by accessing the API Pushshift (Baumgartner et al. 2020). We go through all the archived submissions and comments of /r/wallstreetbets/ and search for the ticket of “AAL” respectively “JETS”.

## **5 Empirical Analysis**

In this part, we will prepare our previous discussions in an equation and test it with panel data regressions if it holds and if the factors from the previous chapters are significant.

### **5.1 Methodology**

To see what factors significantly capture the number of users holding, we use the following baseline model, regressing the amounts of users holding against independent variables:

*Equation 1: OLS Baseline model*

$$RH_{i,d} = \beta_0 + \beta_1 * COVID_d + \beta_2 * Flights_d + \beta_3 * Closing_{i,d} + \beta_4 * Volume_{i,d} + \beta_5 * GSV_{i,d} + \beta_6 * WSB_{i,d} + \varepsilon_{i,d}$$

Whereby  $RH_{i,d}$  is the dependent variable and describes the number of users holding the security  $i$  at the day  $d$  in our observed period, from the 01<sup>st</sup> of January until the 13<sup>th</sup> of August 2020. The first independent variable  $COVID_{i,t}$  defines the COVID cases on the respective day  $d$ .  $Flights_{i,d}$  describes the daily amount of flights on day  $t$ .  $Closing_{i,d}$  and  $Volume_{i,d}$  are the closing price respectively volume for the security  $i$  on the day  $d$ .  $GSV_{i,d}$  and  $WSB_{i,d}$  express how often the stock ticker of stock  $i$  was mentioned on Google or the platform Reddit WallStreetBets per day  $d$ . Since we have many NaNs in our data frame due to weekends and vacations, when the stock market is closed, we removed these days from our sample. Further, we will replace stock prices with stock returns by the following formula:

*Equation 2: Return Formula*

$$R_{i,d} = \frac{P_{t+1}}{P_t} - 1$$

where  $R_{i,d}$  : the return on security  $i$  on day  $d$

$P_{t+1}$  : the price at time  $t + 1$

$P_t$  : the price at time  $t$

Within my methodology, I checked the required assumptions for the multiple regression such as linearity or that the variance of the residuals should be consistent. Most of the assumptions have been met. However, there have been discrepancies with multicollinearity. Hence, we used the variance inflation factor (VIF) which the independent variable is in question. The formula of the VIF is given by:

*Equation 3: Variance inflation factor*

$VIF = \frac{1}{1-R^2}$ , hence the higher  $R^2$ , the higher VIF and the higher the multicollinearity with the variable in question. A typical cut-off value of 5 for moderate or 10 for high multicollinearity is set. For our base model, a high VIF was given for  $Volume_{i,d}$ , as seen in table 6. Hence, to provide non-biased p-values and t-statistics, I decided to remove this variable. Also, in the perspective of retail investors will probably not pay much attention to this measurement. Finally, we are coming up with our regression equation:

*Equation 4: OLS Regression Model*

$$RH_{i,d} = \beta_0 + \beta_1 * COVID_d + \beta_2 * Flights_d + \beta_3 * GSV_{i,d} + \beta_4 * WSB_{i,d} + \beta_5 * Return_{i,d} + \varepsilon_{i,d}$$

## 5.2 Results

The results of the ordinary least squares (OLS) are displayed in table 8. As expected, the daily COVID-19 cases and the daily flights have an inverse relationship with our dependent variable, as global travel restriction regimes were primarily based on the amount of COVID-19 cases. As COVID-19 cases were going up, the number of users holding the security of American Airlines (AAL) increased, the inverse can be seen for the daily flights. The strongest impact on users holding the security can be observed on the independent variable  $GSV_{AAL,d}$  which describes how strong the airline has been searched for on Google. Since the p-value is below the hurdle of 0.05, it is also statistically significant as our previous variables. The variable has a positive relationship with the independent variable which also makes sense as retail investors are especially focused on news (Yuan 2015). Surprisingly, the dependent variable  $WSB_{AAL,d}$  is not statistically significant on a 5% level for the p-value. I would have expected it to be significant with a respectively positive relationship on our dependent variable  $RH_{AAL,d}$ . The same applies to  $Return_{AAL,d}$ , the variable is not statistically significant on a 5% level. The goodness-

of-fit measure  $R^2$  shows that the model fits the given data of Robintrack with an  $R^2$  of 0.895, meaning 89.5% of the variance in the number of users holding the security AAL is explained by the model. The adjusted  $R^2$  with a value of 0.892 is also similarly close and hence we have not been penalized that much for using many independent variables. The model overall is also statistically significant since the probability of rejecting the null hypothesis is approximating zero. Similar results can be reported for JETS as shown in table 9. The only difference is that the independent variable  $WSB_{JETS,d}$  has a positive relationship with the dependent variable. However, the variable is still not statistically significant.

### **5.3 Limitations and considerations**

The first limitation is that the data given by Robintrack does not differ between retail investors buying or selling, we have only been given aggregate holdings. If we had that data available, we could have done a sentiment analysis of the Reddit and Google data to differ between buy and sell signals. This would have given us interesting insights into how both parties differ in their behavioral decision process. Additionally, there has been no relative measurement of the increase of Robinhood users during the pandemic. So, we do not know if the increase was sole because there have been more registered users or if users are holding the stock more relatively than the active users. So, the increase of users holding the stock could have also increased to users. Furthermore, the data is only available until the 13<sup>th</sup> of August 2020 and hence we can only investigate retail investors until that date. It could have been interesting seeing how the behavior changed throughout the whole scope of the pandemic. Multicollinearity could still be a problem. Since we have used a cut-off value for the VIF of 10, which is the upper level, there could still exist some multicollinearity between the independent variable which could bias our results.

In the end, decision-making between retail investors is still truly diverse and not fully observed. Further observable experiments should be made to figure out what leads retail investors to their final buy or sell decision. The crowd of retail investors is diverse and not as homogenous as one the institutional investors who decide on more rational factors such as multiples or other fundamental valuations. This research could be extended by conducting representative surveys throughout the crowd of Robinhood users to include the most important variables in a model. However, this paper can build a foundation for future research since it is one of the first ones to implement social media data. To get certain results, a more intensive implementation of third-party data sets should be done which currently lies out of the scope of this academic paper. In addition, our dataset is only from Robinhood, however, there exist many other platforms similar to RH that could be integrated into future research. Especially, interesting would be how retail investors in different countries decide. One could think that, for example, the risk-awareness of American investors differs from one of the German investors who are opposed to being less risky in their investment decisions.

## **6 Conclusion**

The paper tried to observe what factors a retail investor might consider when investing. We focused on the airline industry since it was one of the main industries in which retail investors bought “the dip”. We have seen from our main dataset that retail investors kept on holding more airline stocks even though the price was constantly going down. Hence, we went along the possible factors for this behavior. Since all Airline securities were behaving similarly, we were focusing on American Airlines and the aviation ETF U.S. Global Jets. First, we started with the main factor such as the COVID cases. COVID cases have an impact the pandemic restrictions such as travel bans with a direct impact on the airline industry. Following we included stock prices and volumes. As said before, stock prices dropped, and the trading volume went up.

Retail investors are confronted with many investment alternatives and hence search costs are high (Barber and Terrance 2005). Thus, retail investors reduce their search costs by using sources such as Google (Da et. al 2011). We, therefore, included Google Trend data in our analysis and analyze how often the respective stock ticker was searched for during the period. As WallStreetBets became a popular place for retail investors to gather, discuss and search for information (Buz and de Melo 2021), I will additionally scrape the data from this platform for my empirical analysis.

I joined the mentioned factors as independent variables in our regression and see how they relate to the dependent variables of the number of users holding the respective security. Our model results are significant and explain most of the variance of our dependent variable. However, surprisingly one of our main observed variables, mentions on WallStreetBets, is not statistically significant. We could interpret that WallStreetBets is just of relevance for a small minority of retail investors on Robinhood. In contrast, Google Trends seems to play a significant role in the decision-making process of retail investors with a strong positive relationship with users holding the security. For both securities, AAL, and JETS, the COVID-19 cases had a positive statistically significant impact on the dependent variable.

My research offers the first attempt of joining multiple factors that could lead to the decision-making process of retail investors, being one of the first papers that take the number of users holding as the dependent variable and try to explain its variance. Since the field of retail investors is so diverse, it has to be further analyzed how retail investors make their decisions and what other factors they might consider. Another research approach is to analyze the data on a more granular level such as on monthly basis for example. This is because as the pandemic progressed, different factors had a different weight in decisions making process. Retail investors adapt consistently in their behavior, consequently, research must also adapt.

## 7 References

- Adamczyk, Alicia. 2021. "Critics view Robinhood restricting GameStop trades as 'an absolute travesty'." *CNBC*. January 28. Accessed May 18, 2022. <https://www.cnbc.com/2021/01/28/critics-call-robinhood-restricting-gamestop-trades-travesty.html>.
- Adamczyk, Alicia. 2020. "Trading apps like Robinhood are having a moment. But users should be careful." *CNBC*. August 21. Accessed April 12, 2022. <https://www.cnbc.com/2020/08/21/robinhood-is-having-a-moment-users-should-be-careful.html>.
- AJMC. 2021. "A Timeline of COVID-19 Developments in 2020." *AJMC*. January 1. Accessed April 26, 2022. <https://www.ajmc.com/view/a-timeline-of-covid19-developments-in-2020>.
- Aloosh, Arash, Hyung-Eun Choi, and Samuel Ouzan. 2021. "Meme Stocks and Herd Behavior."
- Amihud, Yakov. 2002. "Illiquidity and stock returns: cross-section and time-series effects." *Journal of financial markets Vol. 5, No. 1* 31-56.
- Anand, Abhinav, and Jalaj Pathak. IIMB-Working Paper No. 644. "WallStreetBets Against Wall Street The Role of Reddit in the GameStop Short Squeeze." *Working Paper No. 644* 2021.
- Baek, Seungho, and Sunil K.: Glambosky, Mina Mohanty. 2020. "COVID-19 and stock market volatility: An industry level analysis." *Finance Research Letters Volume 37*.
- Barber, Brad M., and Odean Terrance. 2005. "All that Glitters: The Effect of Attention and News on the Buying Behavior of Individual and Institutional Investors." *EFA 2005 Moscow Meetings Paper 51*.
- Barber, Brad M., Terrance Odean, and Ning Zhu. 2006. "Systematic Noise." *AFA 2004 San Diego Meetings, EFA 2005 Moscow Meetings Paper*.
- Barber, Brad M., Xing Huang, Terrance Odean, and Christopher Schwarz. 2020. "Attention Induced Trading and Returns: Evidence from Robinhood Users." *Journal of Finance, Forthcoming 73*.

- Barrot, Jean-Noel, Ron Kaniel, and David Sraer. 2016. "Are retail traders compensated for providing liquidity?" *Journal of Financial Economics, Volume 120, Issue 1* 146-168.
- Baumgartner, Jason, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. "The Pushshift Reddit Dataset." *Fourteenth International AAAI Conference on Web and Social Media*. Pushshift.io.
- Beer, Fiona, Johannes Boße, Dominik Burkert, and Christian Lenz. 2020. "Impact of COVID-19 on the Airline Industry." University Project.
- Ben-David, Itzhak, Francesco Franzoni, and Rabih Moussawi. 2017. "Exchange-Traded Funds." *Annual Review of Financial Economics, Vol. 9* 169-189.
- Bernanke, Ben S., and Kenneth Rogoff. 2001. "The Six Major Puzzles in International Macroeconomics." In *NBER Macroeconomics Annual 2000, Volume 15*, by Ben S. Bernanke and Kenneth Rogoff, 339 - 412. MIT Press.
- Bizouati-Kennedy, Yaël. 2021. "How COVID-19 Turned Day Trading into Gen Investor's Fun – and Lucrative – Hobby." *Go Banking Rates*. May 4. Accessed April 12, 2022. <https://www.gobankingrates.com/investing/strategy/covid-19-turned-day-trading-into-gen-investors-fun-and-lucrative-hobby/>.
- Bordino, Ilaria, Stefano Battiston, Guido Caldarelli, Matthieu Cristelli, Antti Ukkonen, and Ingmar Weber. 2012. "Web Search Queries Can Predict Stock Market Volumes." *PLoS ONE* 7(7): e40014.
- Bouwer, Jaap, Steve Saxon, and Nina Wittkamp. 2021. *Back to the future? Airline sector poised for change post-COVID-19*. McKinsey & Company.
- Bushey, Claire, and Bethan Staton. 2020. "US airline shares surge as American and Virgin add flights." *Financial Times*. June 4. Accessed May 6, 2022. <https://www.ft.com/content/c2697cb3-4ab9-44bf-8f92-82e4d6b1150c>.
- Buz, Tolga, and Gerard de Melo. 2021. "Should You Take Investment Advice From WallStreetBets? A Data-Driven Approach."
- Campbell, John, John Y., and Albert S. Kyle. 1993. "Smart Money, Noise Trading and Stock Price Behaviour." *Review of Economic Studies* 60, no. 1 1-34.

- Companies Market Cap. 2022. *Companies Market Cap*. May 15. Accessed May 15, 2022. <https://companiesmarketcap.com/airlines/largest-airlines-by-market-cap/>.
- Cucinotta, Domenico, and Maurizio Vanelli. 2020. "WHO Declares COVID-19 a Pandemic." *Acta Biomed*.
- Da, Zhi, Joseph Engelberg, and Gao Pengjie. 2011. "In Search of Attention." *The Journal of Finance Volume 66, Issue 5* 1461-1499.
- Daas, Piet J.H., and Marco J.H Puts. 2014. "Social media sentiment and consumer confidence." *ECB Statistics Paper, No. 5*.
- Devlin, Hannah. 2021. "First known Covid case was Wuhan market vendor, says scientist." *The Guardian*. November 19. Accessed April 11, 2022. <https://www.theguardian.com/world/2021/nov/19/first-covid-patient-in-wuhan-was-at-animal-market-study-finds>.
- Di Muzio, Tim. 2021. "GameStop Capitalism. Wall Street vs. The Reddit Rally." *Articles (Working Papers), The Bichler & Nitzan Archives*.
- Dominguez, Hector Rodriguez. 2021. "Using Pushshift API for data analysis on Reddit." *Medium*. September 14. Accessed May 8, 2022. <https://medium.com/mcd-unison/using-pushshift-api-for-data-analysis-on-reddit-b08d339c48b8>.
- Drake, Michael S., Darren T. Roulstone, and Jacob R. Thornock. 2012. "Investor Information Demand: Evidence from Google Searches Around Earnings Announcements." *Journal of Accounting Research Volume 50, Issue* 1001-1040.
- Eaton, Gregory W., T. Clifto Green, Brian S. Roseman, and Yanbin Wu. 2021. "Zero-Commission Individual Investors, High Frequency Traders, and Stock Market Quality."
- European Centre for Disease Prevention and Control. 2020. "Risk assessment: outbreak of acute respiratory syndrome associated with a novel coronavirus, China; First cases imported in the EU/EEA; second update." *European Centre for Disease Prevention and Control*. January 26. Accessed April 26, 2022. <https://www.ecdc.europa.eu/en/publications-data/risk-assessment-outbreak-acute-respiratory-syndrome-associated-novel-0>.

- Farivar. 2021. "Gambling addiction experts see familiar aspects in Robinhood app." *NBC News*. January 30. Accessed April 12, 2022. <https://www.nbcnews.com/business/business-news/gambling-addiction-experts-see-familiar-aspects-robinhood-app-n1256213>.
- Faura, Javier Cifuentes, and Martinez Ursula. 2021. "Situation Of European Airlines Caused By Covid-19: Restrictions, Government Subsidies And Future Prospects." *Aviation* 25(4) 232-240.
- Financial Times. 2022. "Lockdowns compared: tracking governments' coronavirus responses." *Financial Times*. May 11. Accessed May 11, 2022. <https://ig.ft.com/coronavirus-lockdowns/>.
- Fonda, Daren. 2020. "American Airlines Stock Dips After It Announces Deep Staff Cuts." *Barron's*. May 28. Accessed May 06, 2022. <https://www.barrons.com/articles/american-airlines-stock-dips-after-it-announces-deep-staff-cuts-51590678261>.
- Foucault, Thierry, David Sraer, and David J. Thesmar. 2011. "Individual Investors and Volatility." *The Journal of Finance* Vol. 66, No. 4 1369-1406.
- Friedman, Henry L., and Zitong Zeng. 2021. "Retail Investor Trading and Market Reactions to Earnings Announcements." 57.
- Ganie, Irfan Rashid, Tahir Ahmad Wani, and Miklesh Prasad Yadav. 2022. "Impact of COVID-19 Outbreak on the Stock Market: An Evidence from Select Economies." *Business Perspectives and Research*.
- Garrow, Laurie, and Virginie Lurkin. 2021. "How COVID-19 is impacting and reshaping the airline industry." *Journal of Revenue and Pricing Management* 3-9.
- Ghosh, Imam. 2021. "This chart shows how airlines are being grounded by COVID-19." *World Economic Forum*. March 21. Accessed April 25, 2022. <https://www.weforum.org/agenda/2020/03/this-chart-shows-how-airlines-are-being-grounded-by-covid-19/>.
- Glossner, Simon, Pedro Matos, Stefano Ramelli, and Alexander F. Wagner. 2021. "Do Institutional Investors Stabilize Equity Markets in Crisis Periods? Evidence from

COVID-19." *Swiss Finance Institute Research Paper No. 20-56, European Corporate Governance Institute – Finance Working Paper No. 688/2020 70.*

Hendershott, Terrence, Albert J. Menkveld, and Rémy, Seasholes, Mark Praz. 2022. "Asset Price Dynamics with Limited Attention." *The Review of Financial Studies, Volume 35, Issue 2* 962–1008.

Henderson, Richard, and Miles Kruppa. 2020. "Retail trading app Robinhood's value tops \$11bn on new fundraising." *Financial Times*. August 17. Accessed April 12, 2022. <https://www.ft.com/content/b208cbbe-579c-4cbf-9358-01ae02b4381b>.

Hiltzik, Michael. 2022. "Column: Remember the GameStop stock craze? Looks like it's over." *Los Angeles Times*. January 25. Accessed April 29, 2022. <https://www.latimes.com/business/story/2022-01-25/column-gamestop-meme-stocks>.

Hsieh, Shu-Fan, Chia-Ying Chan, and Ming-Chun Wang. 2020. "Retail investor attention and herding behavior." *Journal of Empirical Finance Volume 59* 109-132.

Huang, Daniel. 2015. "Young, Poor and Looking to Invest? Robinhood Is the App for That." *The Wall Street Journal*. January 6. Accessed April 12, 2022. <https://www.wsj.com/articles/BL-MBB-31486>.

IATA. 2020. "Industry Losses to Top \$84 Billion in 2020." *IATA*. June 9. Accessed April 27, 2022. <https://www.iata.org/en/pressroom/pr/2020-06-09-01/>.

Iyyer, Shruti. 2020. "Step by Step Assumptions - Linear Regression." *Kaggle*. May 20. Accessed May 13, 2022. <https://www.kaggle.com/code/shrutimechlearn/step-by-step-assumptions-linear-regression/notebook>.

Joseph, Anthony, Maurice Larrain, and Claude Turner. 2017. "Daily Stock Returns Characteristics and Forecastability." *Procedia Computer Science Volume 114* 481-490.

Joseph, Kissan, Babajide Wintoki, and Zelin Zhang. 2011. "Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search." *International Journal of Forecasting Volume 27, Issue 4* 1116-1127.

Kotoky, Anurag, David Stringer, and Ragini Saxena. 2020. "Two-Thirds of the World's Passengers Jets Are Grounded Amid COVID-19 Pandemic. Here's What That Means."

- TIME*. April 17. Accessed April 28, 2022. <https://time.com/5823395/grounded-planes-coronavirus-storage/>.
- Kristopher, Gordon. 2014. "Interpreting volume in technical analysis." *Yahoo! Finance*. December 4. Accessed April 27, 2022. <https://finance.yahoo.com/news/interpreting-volume-technical-analysis-170127457.html?guccounter=1>.
- Kumari, Sharda. 2022. "Impact of Social Media on biases of individual investor's decision making."
- La Monica, Paul R. 2022. "A year after GameStop mania, meme stocks have tumbled." *CNN Business*. January 26. Accessed April 12, 2022. <https://edition.cnn.com/2022/01/26/investing/meme-stocks-gamestop-amc-reddit/index.html>.
- Li, Wei, Ghon Rhee, and Steven Shuye Wang. 2017. "Differences in herding: Individual vs. institutional investors." *Pacific-Basin Finance Journal Volume 45* 174-185.
- Lyócsa, Štefan, Eduard Baumöhl, and Tomáš Vřrost. 2021. "YOLO trading: Riding with the herd during the GameStop Episode." *ZBW - Leibniz Information Centre for Economics*.
- Mankiw, Gregory N., and Stephen P. Zeldes. 1991. "The consumption of stockholders and non-stockholders." *Journal of Financial Economics* 29 97-112.
- Manu, Jose. 2021. "Financial Data from Yahoo Finance with Python." *Towards Data Science*. January 20. Accessed May 5, 2022. <https://towardsdatascience.com/financial-data-from-yahoo-finance-with-python-b5399743bcc6>.
- Markets, Robinhood. 2022. "What You Need to Get Started." *Robinhood*. January 1. Accessed April 24, 2022. <https://robinhood.com/us/en/support/articles/what-you-need-to-get-started/>.
- Merli, Maxime, and Tristan Roger. 2013. "What Drives the Herding Behavior of Individual Investors?" *Finance*, 34(3).
- Mohr, Justin. 2021. *Pulling the Plug: Retail Traders and Social Media*. SSRN.
- Neslen, Arthur. 2022. "Airlines flying near-empty 'ghost flights' to retain EU airport slots." *The Guardian*. January 26. Accessed April 28, 2022.

- <https://www.theguardian.com/environment/2022/jan/26/airlines-flying-near-empty-ghost-flights-to-retain-eu-airport-slots>.
- Nofsinger, John R., and Richard W. Sias. 1999. "Herding and Feedback Trading by Institutional and Individual Investors." *The Journal of Finance*, Volume 54, Issue 6 2263-2295.
- Onvista. 2020. "Die 10 beliebtesten Aktien auf Robinhood." *Onvista*. December 13. Accessed April 24, 2022. <https://www.onvista.de/news/die-10-beliebtesten-aktien-auf-robinhood-419415027>.
- OpenSky Network. 2022. *OpenSky Network*. <https://opensky-network.org/>.
- Padliya, Palash. 2022. "Analyzing Retail Investors with Robinhood Data." <https://palashp.github.io/>.
- Pagano, Michael S., John Sedunov, and Raisa Velthuis. 2020. "How did retail investors respond to the COVID-19 pandemic? The effect of Robinhood brokerage customers on market quality." *Finance Research Letters*, Forthcoming.
- Ponczek, Sarah. 2020. "Robintrack, Chronicler of Day Trader Stock Demand, To Shut." *Bloomberg*. August 8. Accessed April 14, 2022. <https://www.bloomberg.com/news/articles/2020-08-08/robintrack-chronicler-of-day-trader-stock-demand-to-shut-down>.
- Preis, Tobias, Moat, Susannah Helen, and Eugene H. Stanley. 2013. "Quantifying Trading Behavior in Financial Markets Using Google Trends." *Sci Rep* 3, 1684.
- Ribeiro, Sérgio P., and Alexandre B. Reis. 2020. "Worldwide COVID-19 spreading explained: traveling numbers as a primary driver for the pandemic." *Annals of the Brazilian Academy of Sciences* 92(4).
- Robinhood Markets, Inc. 2022. "Robinhood Reports Fourth Quarter and Full Year 2021 Results." *Robinhood*. January 27. Accessed April 22, 2022. <https://investors.robinhood.com/news/news-details/2022/Robinhood-Reports-Fourth-Quarter-and-Full-Year-2021-Results/default.aspx>.
- Rooney, Kate. 2020. "Fintech app Robinhood is driving a retail trading renaissance during the stock market's wild ride." *CNBC*. June 17. Accessed April 12, 2022.

- <https://www.cnbc.com/2020/06/17/robinhood-drives-retail-trading-renaissance-during-markets-wild-ride.html>.
- . 2020. "Gamblers pivot to stock trading during lockdowns." *CNBC*. May 22. Accessed April 12, 2022. <https://www.cnbc.com/2020/05/22/gamblers-pivot-to-stock-trading-during-lockdowns---barstools-portnoy-revives-old-e-trade-account.html>.
- S. Pagano, Michael S., John Sedunov, and Raisa Velthuis. 2020. "How did Retail Investors Respond to the COVID-19 Pandemic? The Effect of Robinhood Brokerage Customers on Market Quality." *Finance Research Letters*, *Forthcoming* 26.
- Shefrin, Hersh, and Meir Statman. 1984. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." *The Journal of Finance* Vol. 40, No. 3 777-790.
- Sias, Richard W. 2004. "Institutional Herding." *The Review of Financial Studies* Vol. 17, No. 1 165-206.
- Strohmeier, Martin, Xavier Olive, Jannis Lübke, Matthias Schäfer, and Vincent Lenders. 2021. "Crowdsourced Air Traffic Data from the OpenSky Network 2019–20." *Earth System Science Data Volume 13, Issue 2* 357–366.
- Tenebruso, Joe. 2020. "Here Are the Top 10 Most Popular Stocks on Robinhood." *The Motley Fool*. August 3. Accessed April 12, 2022. <https://www.fool.com/investing/2020/08/03/here-are-the-top-10-most-popular-stocks-on-robinho.aspx>.
- U.S. Global Investors. 2022. *Prospectus*. Prospectus, U.S. Global Investors.
- Uusi-Mikkola, Hugo. 2021. "Empirical Analysis of Effects of VIX on Retail Investor Activity during the COVID-19 Pandemic."
- Welch, Ivo. 2020. "The Wisdom of the Robinhood Crowd." *NBER Working Paper No. w27866* 52.
- Winkler, Julian, and Valentina Semenova. 2021. "Reddit's self-organised bull runs: Social contagion and asset prices." *INET Oxford Working Papers 2021-04*.

- Wu, Jasmine. 2019. "GameStop shares tank after earnings miss, cuts sales forecast." *CNBC*. September 10. Accessed April 29, 2022. <https://www.cnbc.com/2019/09/10/game-stop-shares-fall-after-second-quarter-profit-and-revenue-miss.html>.
- Yousefi, Hamed, and Mohammad Najand. 2022. "Geographical Diversification Using ETFs: Multinational Evidence from COVID-19 Pandemic."
- Yuan, Yu. 2015. "Market-Wide Attention, Trading, and Stock Returns." *Journal of Financial Economics (JFE)*, Volume 116, Issue 3 548-564.

## 8 Appendix

### 8.1 Tables

Table 1: List of Airlines (by market capitalization) as of 15.05.2020

Name	Ticket Robinhood	marketcap	Market capitalization (in billion USD)	Price (USD)	Country
Delta Air Lines	DAL	27,395,403,776.00	27.40	42.81	United States
Southwest Airlines	LUV	27,206,268,928.00	27.21	45.93	United States
Ryanair	RYAAY	24,469,673,984.00	24.47	107.95	Ireland
United Airlines Holdings	UAL	15,341,976,576.00	15.34	47.38	United States
American Airlines	AAL	11,959,411,712.00	11.96	17.88	United States
Lufthansa	DLAKY	10,240,087,316.00	10.24	8.57	Germany
Alaska Airlines	ALK	7,391,034,368.00	7.39	58.7	United States
easyJet	ESYJY	7,097,752,039.00	7.10	9.32	United Kingdom
Jetblue Airways	JBLU	4,947,767,296.00	4.95	15.56	United States
Copa Holdings	CPA	3,935,221,248.00	3.94	93.85	Panama

Table 2: Descriptive Statistics AAL of users holding

users_holding	
count	818.00
mean	100225.18
std	205934.97
min	3748.00
25%	6931.28
50%	9191.58
75%	11414.95
max	659935.00

Table 3: Descriptive Statistics JETS of users holding

users_holding	
count	818.00
mean	96227.98
std	187454.88
min	7835.00
25%	9995.81
50%	14311.81
75%	16514.51
max	599367.83

Table 4: Top 10 Holding of U.S. Global Jets ETF (JETS) as of 31.03.2022

HOLDINGS

Data as of 05/19/2022











% Net Assets	Name	CUSIP	Country	Ticker	Shares Held	Market (\$)
12.35%	UNITED AIRLS HLDGS INC	910047109		UAL	8,386,571.00	370,770,303.91
11.32%	DELTA AIR LINES INC DEL	247361702		DAL	8,656,602.00	339,685,062.48
11.15%	AMERICAN AIRLS GROUP INC	02376R102		AAL	19,977,563.00	334,624,180.25
10.01%	SOUTHWEST AIRLS CO	844741108		LUV	6,896,254.00	300,469,786.78
2.97%	SUN CTRY AIRLS HLDGS INC	866683105		SNCY	3,711,183.00	89,031,280.17
2.90%	ALLEGiant TRAVEL CO	01748X102		ALGT	593,902.00	86,941,313.78
2.87%	AIR CANADA	BSDHYK1		AC CN	5,255,687.00	86,111,053.10
2.78%	SKYWEST INC	830879102		SKYW	3,214,411.00	83,510,397.78
2.71%	ALASKA AIR GROUP INC	011659109		ALK	1,742,412.00	81,475,185.12
2.68%	SPIRIT AIRLS INC	848577102		SAVE	4,210,222.00	80,373,137.98

Table 5: Aircraft types used for dataset OpenSky Network

Boeing	McDonnell Douglas	Bombardier	Embraer	Airbus	AT
B721	MD11	CRJ1	E110	A359	AT43
B722	MD52	CRJ-1000	E120	A346	AT46
B732	MD60	CRJ7	E135	A19N	AT72
B733	MD81	CRJ9	E140	A20N	AT73
B734	MD83	CRJ2	E145	A210	AT75
B735	MD87		E170	A318	AT76
B736	MD88		E190	A319	
B737	MD90		E195	A320	
B738	DC93		E290	A321	
B739	DC10		E295	A337	
B742				A338	
B744				A342	
B748				A343	
B74S				A345	
B753				A346	
B752				A388	
B772				A359	
B773				A339	
B783				A330	
B788				A330-200	
B789				A332	
B762				A333	
B763				A359	
B77L					
B77W					
B77F					
B78X					

Table 6: VIF for AAL

	feature	VIF
0	const	21.415336
1	Daily COVID-19 Cases	1.677522
2	Daily Flights 2020	2.670850
3	Daily Volume AAL	5.685743
4	Daily Google Mentions AAL	3.185145
5	Daily WSB mentions	3.352483
6	Return AAL	1.281651

Table 7: VIF for JETS

	feature	VIF
0	const	20.076920
1	Daily COVID-19 Cases	1.877980
2	Daily Flights 2020	2.560718
3	Daily Volume JETS	5.067292
4	Daily Google Mentions JETS	4.324349
5	Daily WSB mentions JETS	3.150007
6	Return JETS	1.082470

Table 8: OLS Regression Results for AAL

OLS Regression Results

<b>Dep. Variable:</b>	RH Users holding AAL	<b>R-squared:</b>	0.895				
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	0.892				
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	256.0				
<b>Date:</b>	Fri, 20 May 2022	<b>Prob (F-statistic):</b>	1.53e-71				
<b>Time:</b>	02:48:43	<b>Log-Likelihood:</b>	-1996.0				
<b>No. Observations:</b>	156	<b>AIC:</b>	4004.				
<b>Df Residuals:</b>	150	<b>BIC:</b>	4022.				
<b>Df Model:</b>	5						
<b>Covariance Type:</b>	nonrobust						
	<b>coef</b>	<b>std err</b>	<b>t</b>	<b>P&gt; t </b>	<b>[0.025</b>	<b>0.975]</b>	
	const	2.446e+05	3.05e+04	8.029	0.000	1.84e+05	3.05e+05
	Daily COVID-19 Cases	2.0976	0.104	20.140	0.000	1.892	2.303
	Daily Flights 2020	-4.3768	0.570	-7.681	0.000	-5.503	-3.251
	Daily Google Mentions AAL	2175.6397	884.410	2.460	0.015	428.129	3923.151
	Daily WSB mentions	-44.1695	63.049	-0.701	0.485	-168.748	80.409
	Return AAL	5419.7296	9.46e+04	0.057	0.954	-1.82e+05	1.92e+05
<b>Omnibus:</b>	5.305	<b>Durbin-Watson:</b>	0.185				
<b>Prob(Omnibus):</b>	0.070	<b>Jarque-Bera (JB):</b>	5.419				
<b>Skew:</b>	-0.435	<b>Prob(JB):</b>	0.0666				
<b>Kurtosis:</b>	2.724	<b>Cond. No.</b>	1.68e+06				

Table 9: OLS Regression Results for JETS

OLS Regression Results

<b>Dep. Variable:</b>	RH Users holding JETS	<b>R-squared:</b>	0.885
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	0.882
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	231.6
<b>Date:</b>	Fri, 20 May 2022	<b>Prob (F-statistic):</b>	1.21e-68
<b>Time:</b>	05:54:11	<b>Log-Likelihood:</b>	-1988.3
<b>No. Observations:</b>	156	<b>AIC:</b>	3989.
<b>Df Residuals:</b>	150	<b>BIC:</b>	4007.
<b>Df Model:</b>	5		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	2.127e+05	3e+04	7.086	0.000	1.53e+05	2.72e+05
Daily COVID-19 Cases	1.8596	0.101	18.492	0.000	1.661	2.058
Daily Flights 2020	-3.7509	0.566	-6.628	0.000	-4.869	-2.633
Daily Google Mentions JETS	1757.0693	622.851	2.821	0.005	526.375	2987.763
Daily WSB mentions JETS	231.5863	597.514	0.388	0.699	-949.044	1412.217
Return JETS	3.463e+04	1.38e+05	0.251	0.802	-2.38e+05	3.07e+05

<b>Omnibus:</b>	6.049	<b>Durbin-Watson:</b>	0.158
<b>Prob(Omnibus):</b>	0.049	<b>Jarque-Bera (JB):</b>	6.132
<b>Skew:</b>	-0.455	<b>Prob(JB):</b>	0.0466
<b>Kurtosis:</b>	2.663	<b>Cond. No.</b>	2.57e+06

8.2 Figures

Figure 1: Retail investors holding the top 10 airlines by market capitalization

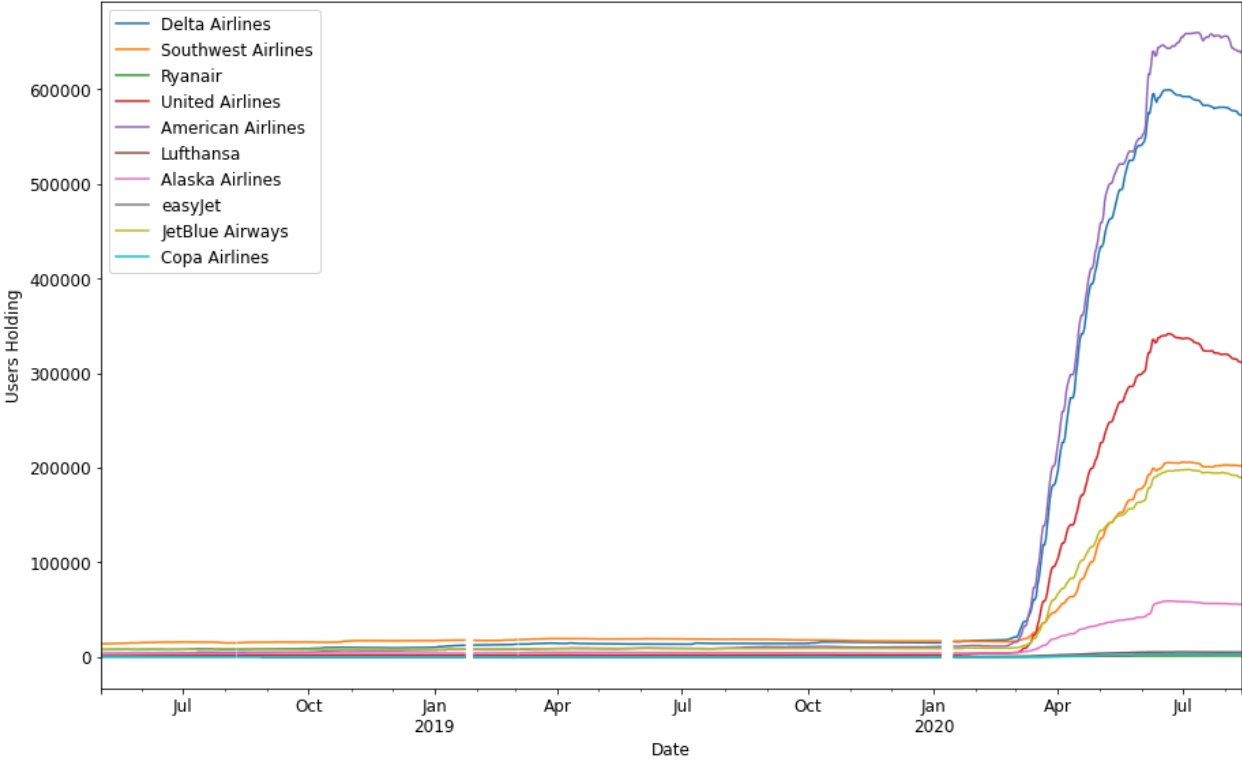


Figure 2: Retail investors holding the American-based airlines by market capitalization

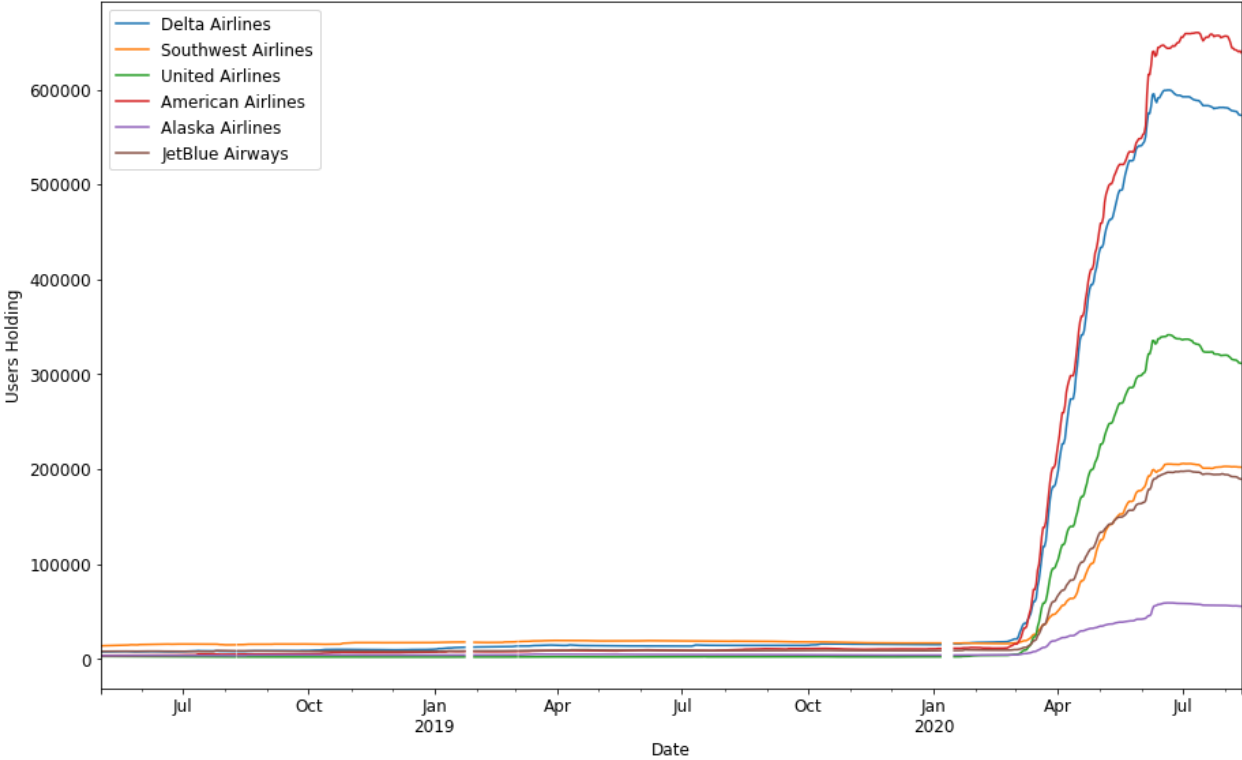


Figure 3: Retail investors holding American Airlines (AAL) and U.S. Global Jets ETF (JETS)

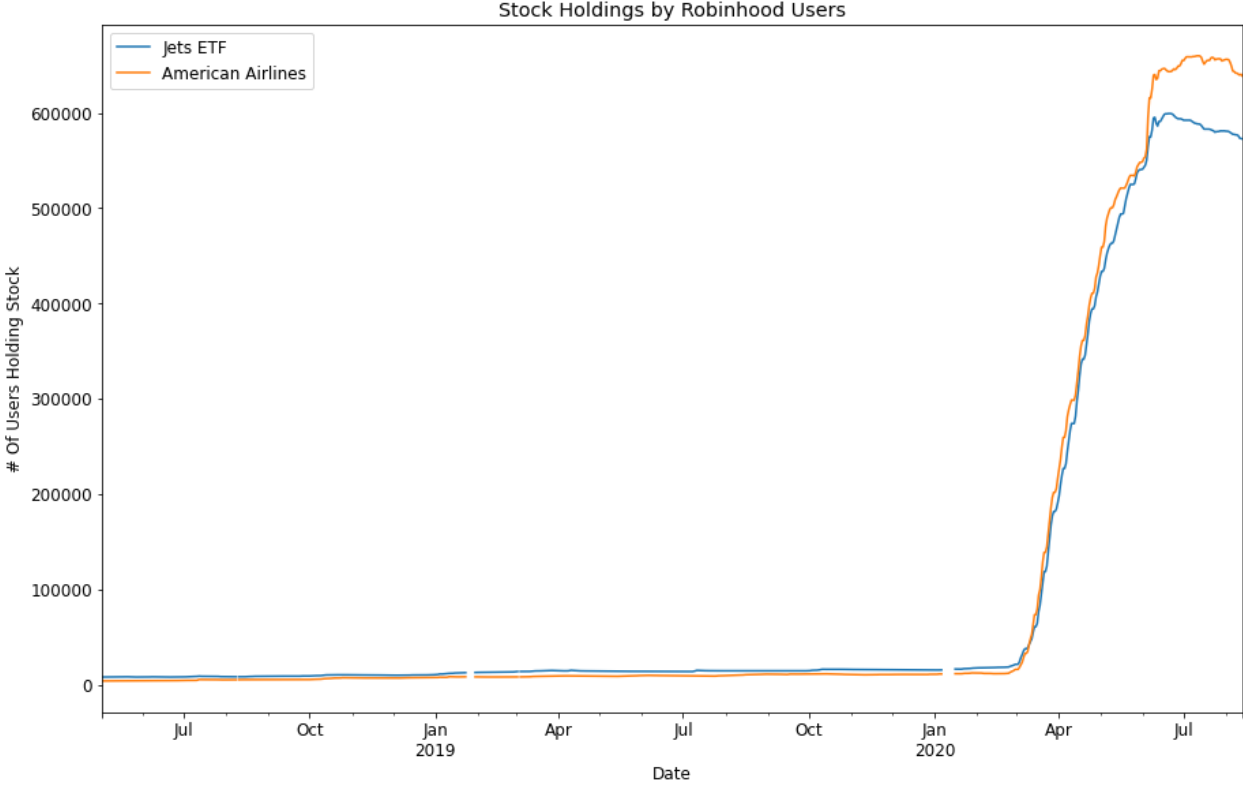


Figure 4: Stock prices of all airline stocks

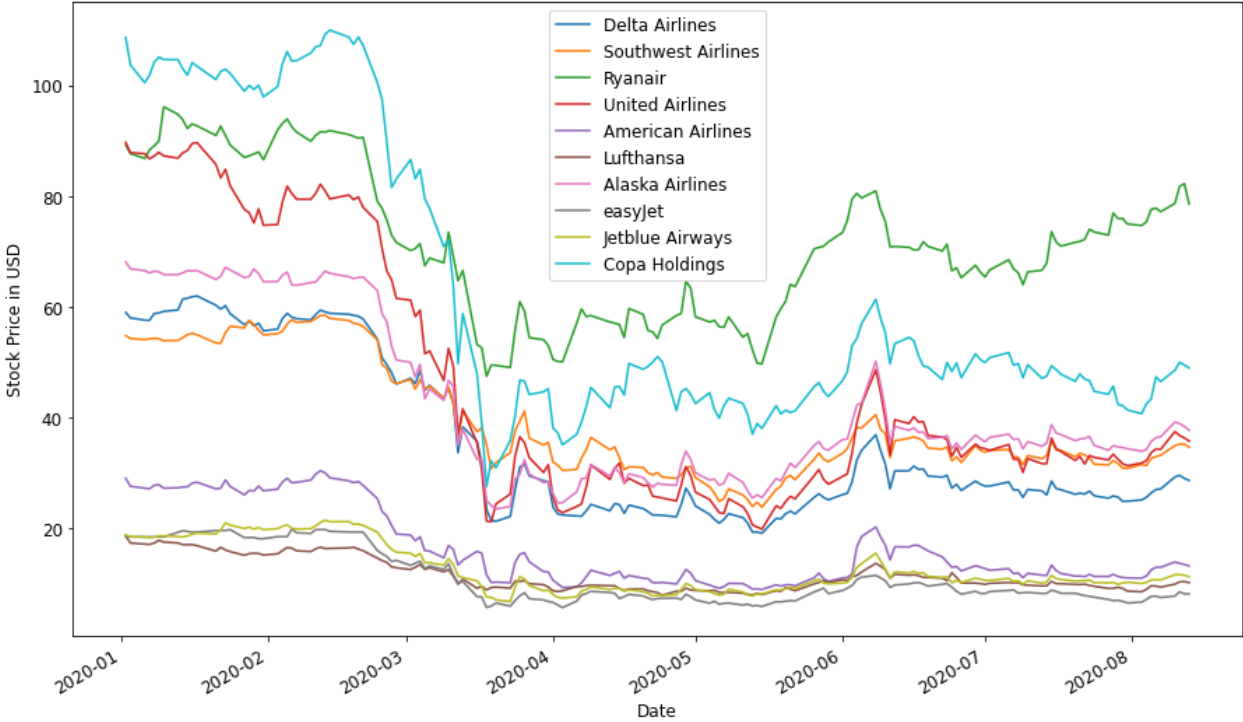


Figure 5: Stock prices of American-based airline stocks

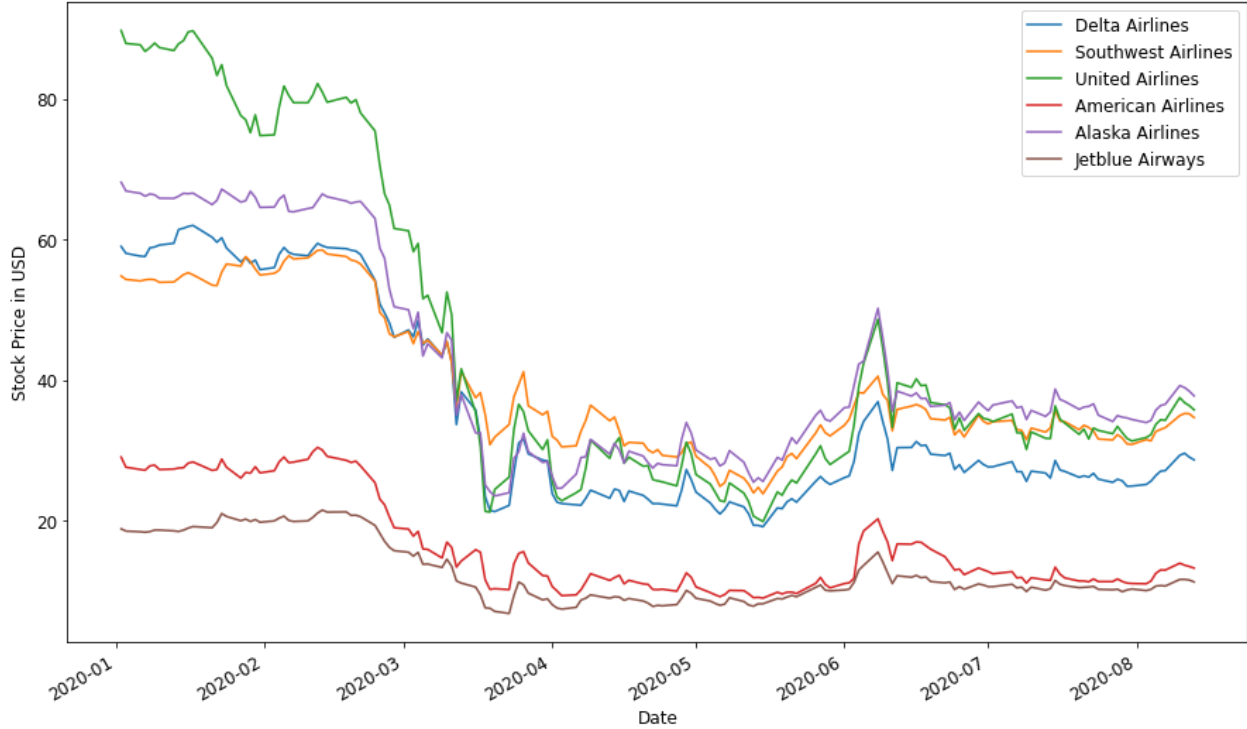


Figure 6: Stock prices of AAL and JETS

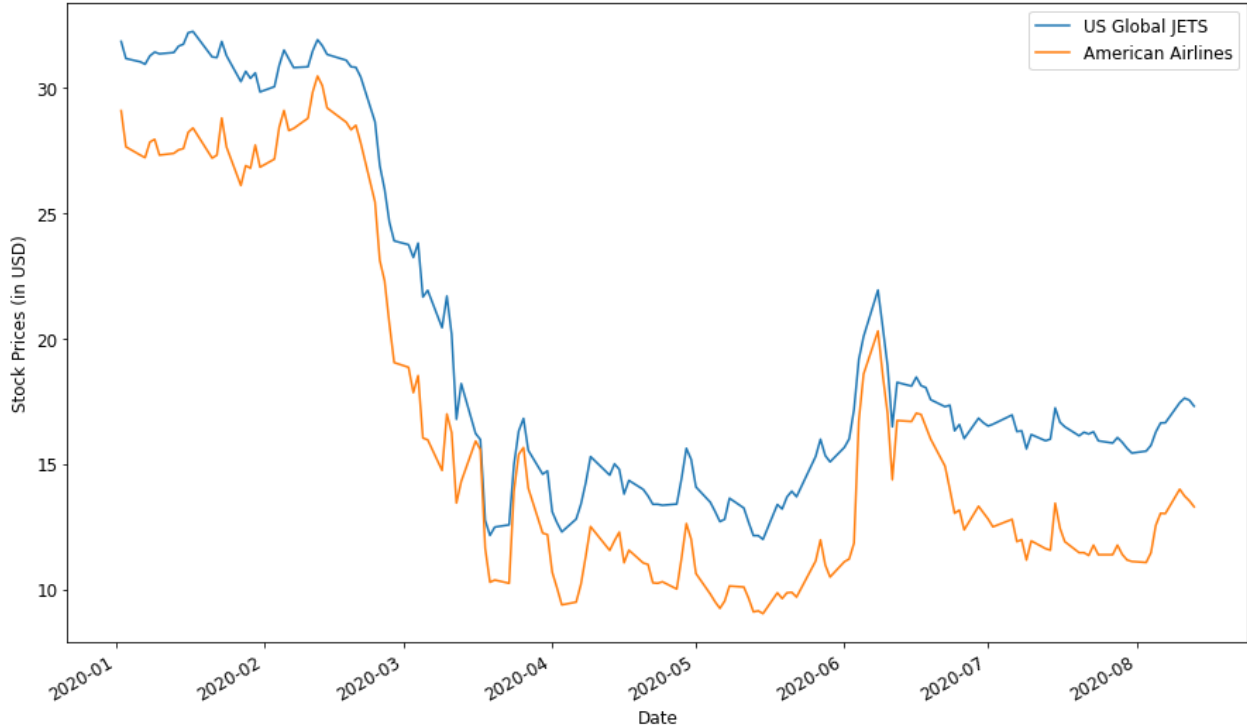


Figure 7: Stock Volume of all airline stocks

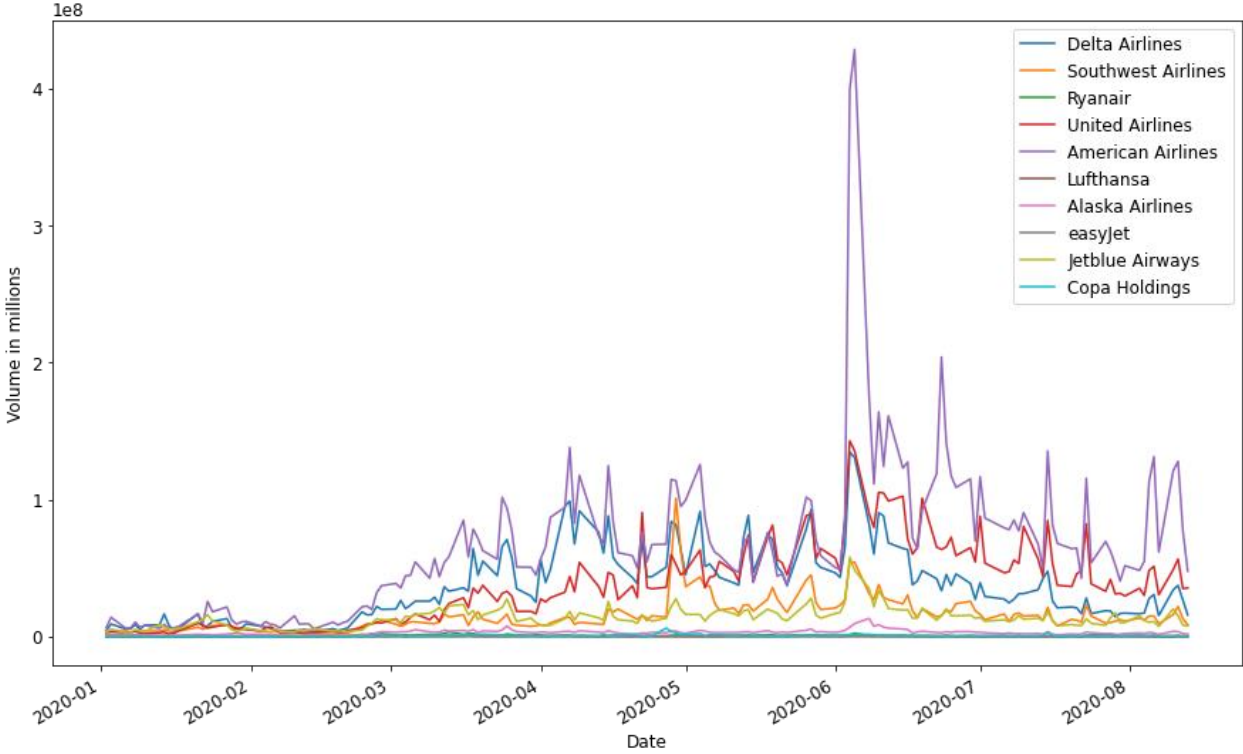


Figure 8: Stock volume of all American-based airline stocks

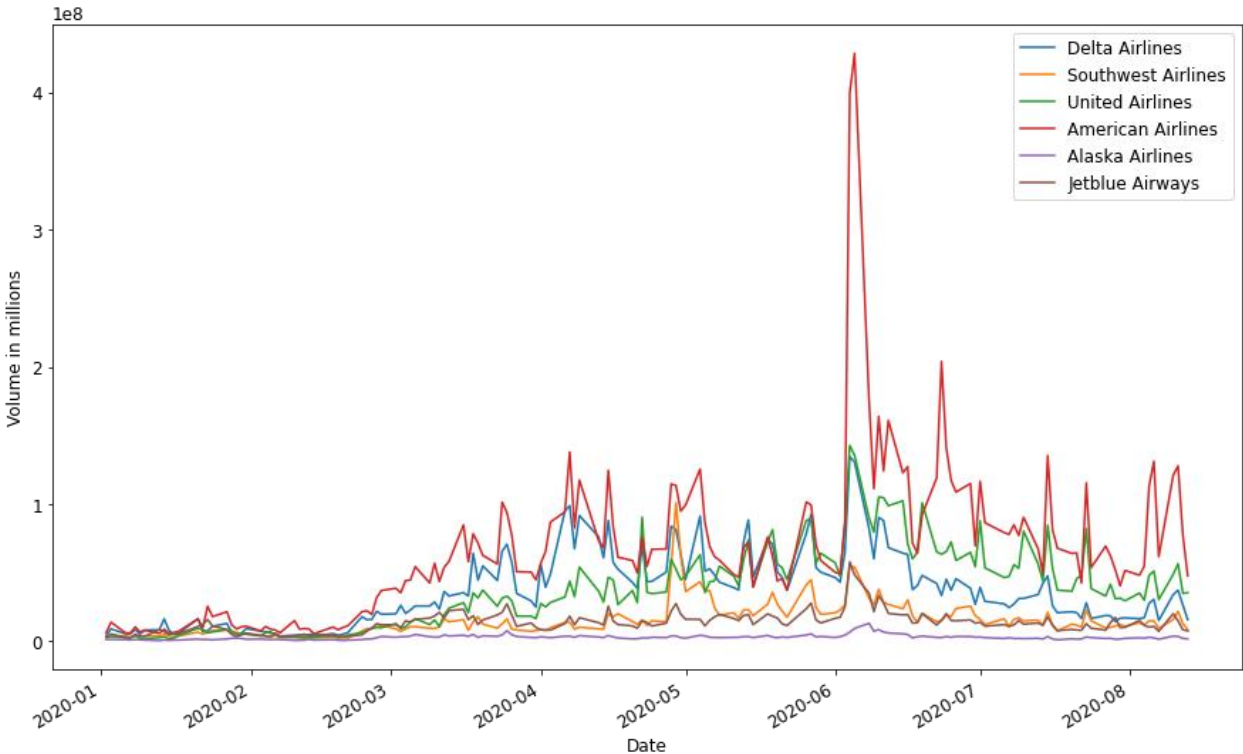


Figure 9: Stock volume of AAL and JETS

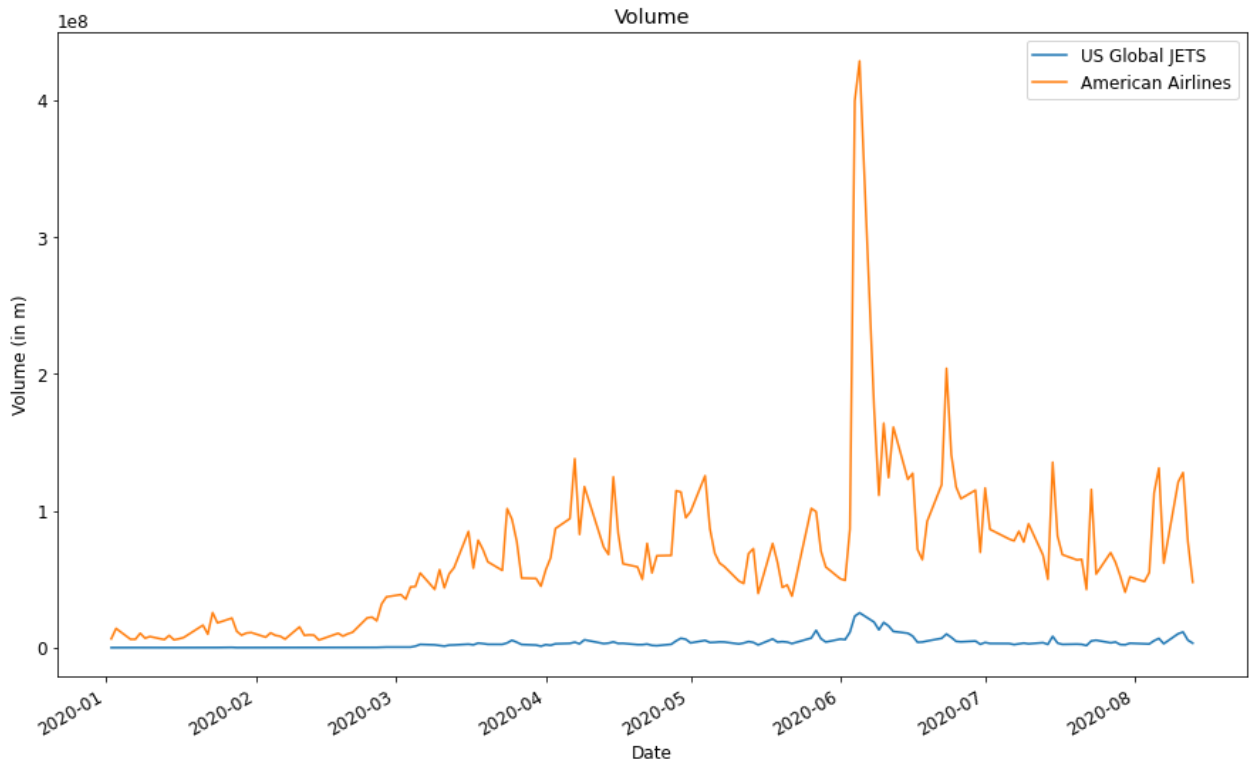


Figure 10: COVID-19 Cases worldwide from 01/2020 until 08/2020

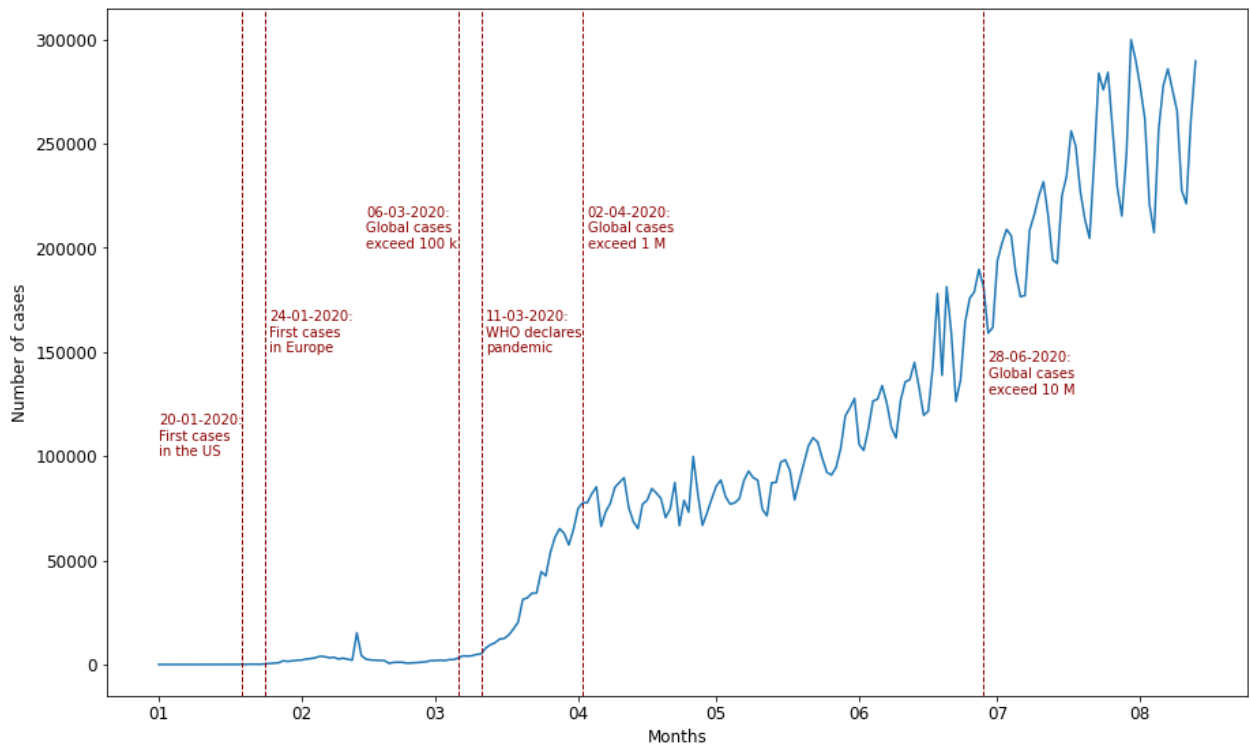


Figure 11: Comparison daily flights worldwide between 2019 and 2020

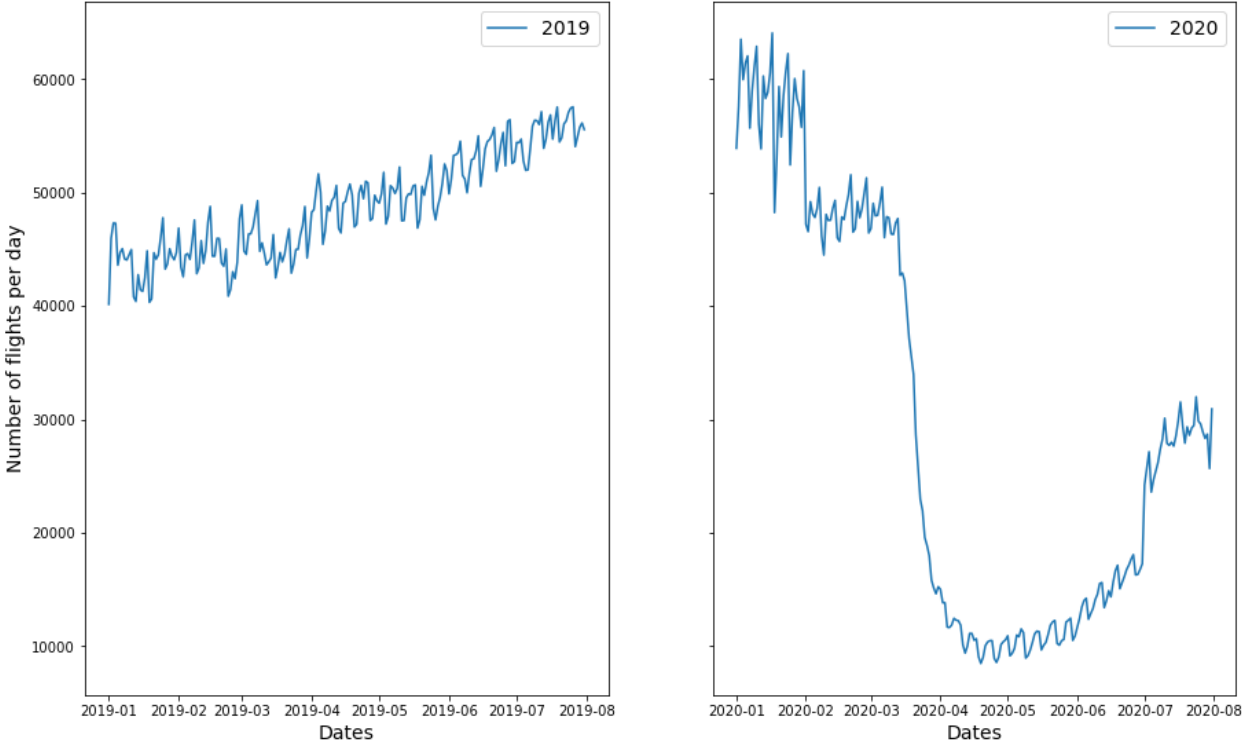


Figure 12: Daily flights worldwide from 01/2020 until 08/2020

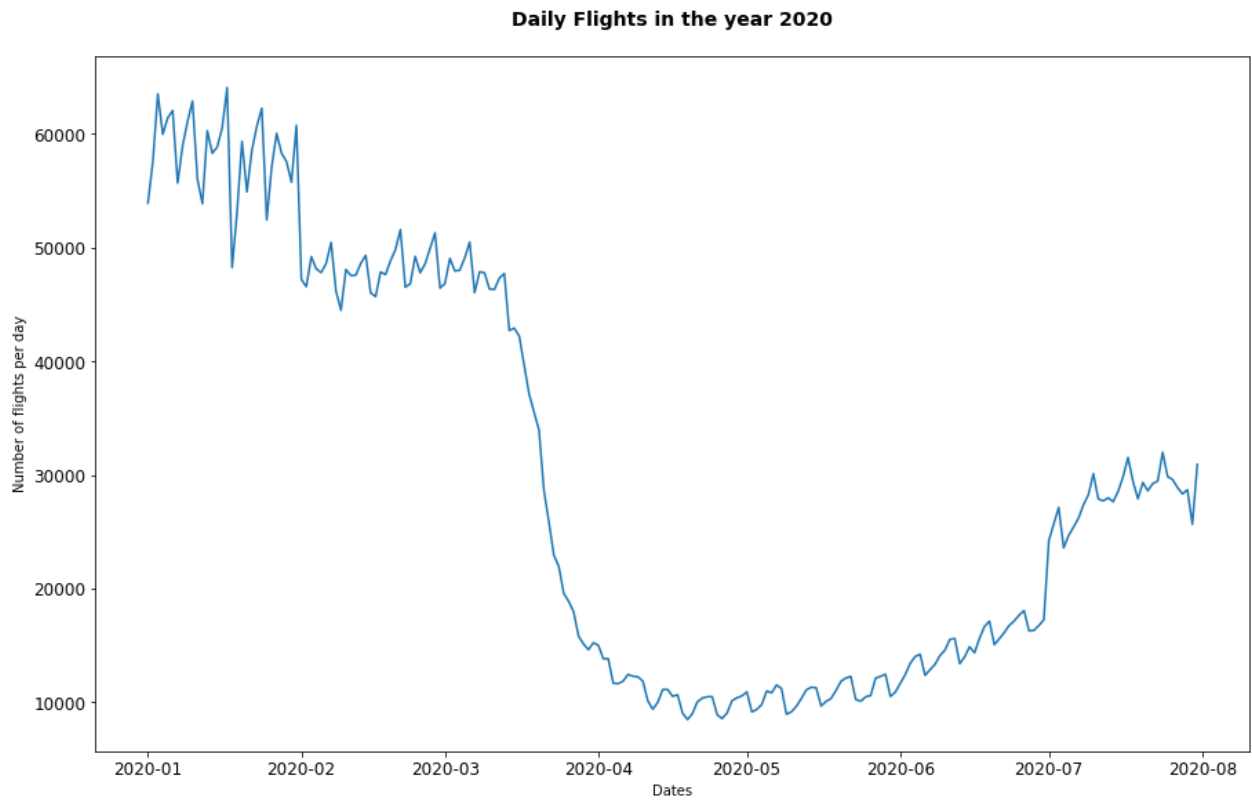


Figure 13: Daily flights worldwide from 01/2020 until 08/2020 split up by airline



Figure 14: Google Search Volume for 01/2020 until 08/2020 for AAL

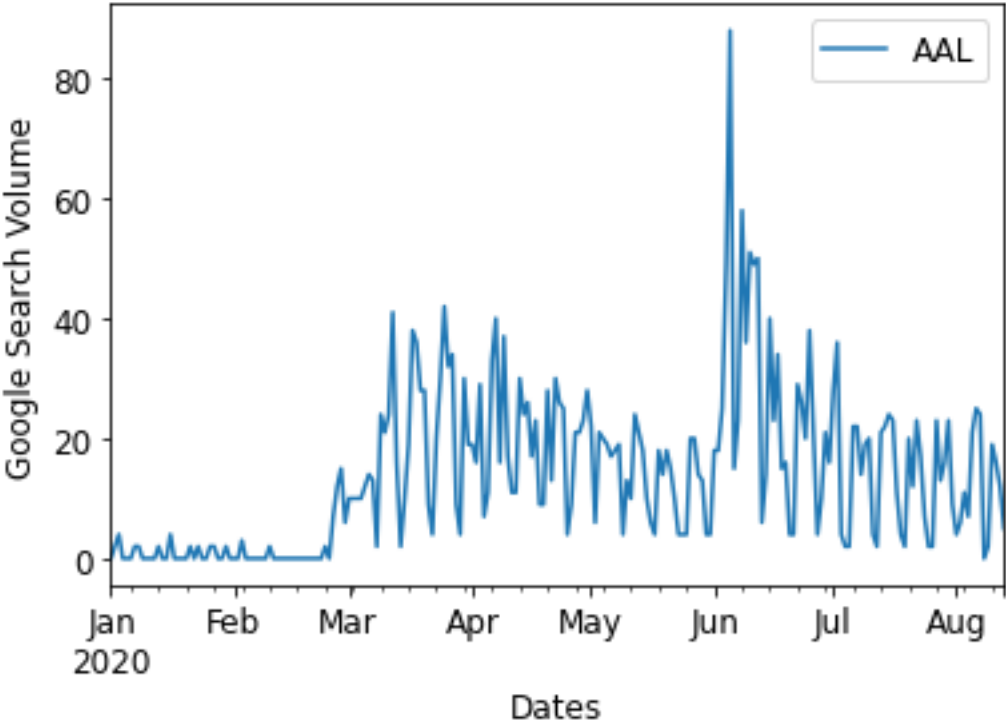


Figure 15: Google Search Volume for 01/2020 until 08/2020 for JETS

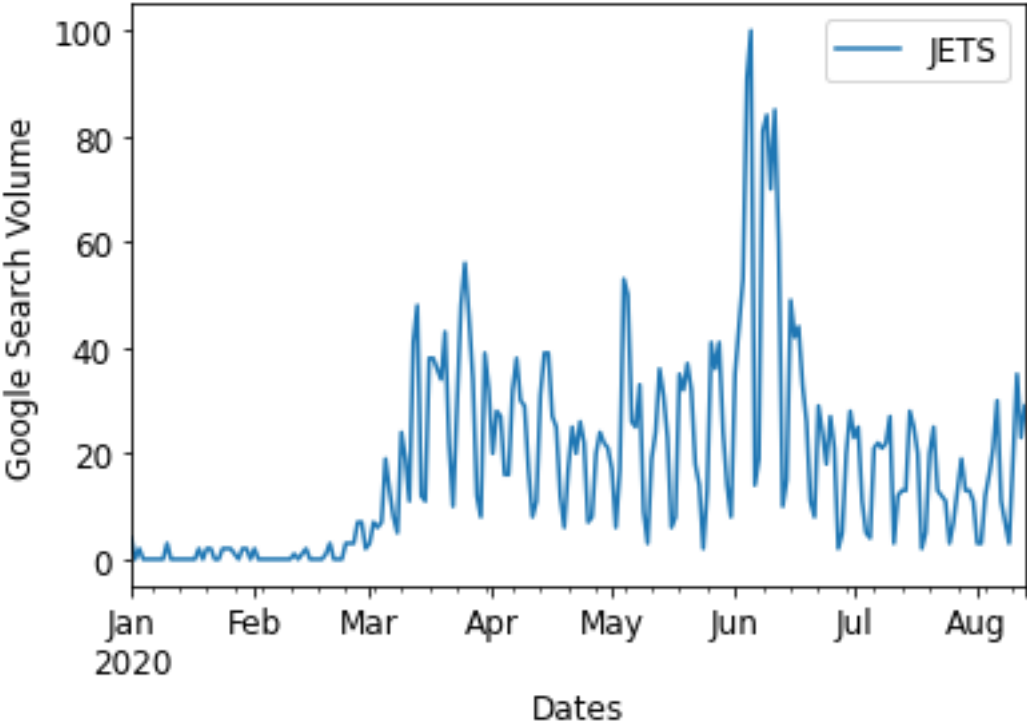


Figure 16: ADX for AAL



Figure 17: Reddit WallStreetBets – Mentioning of ticket “AAL” from 01/2020 until 08/2020

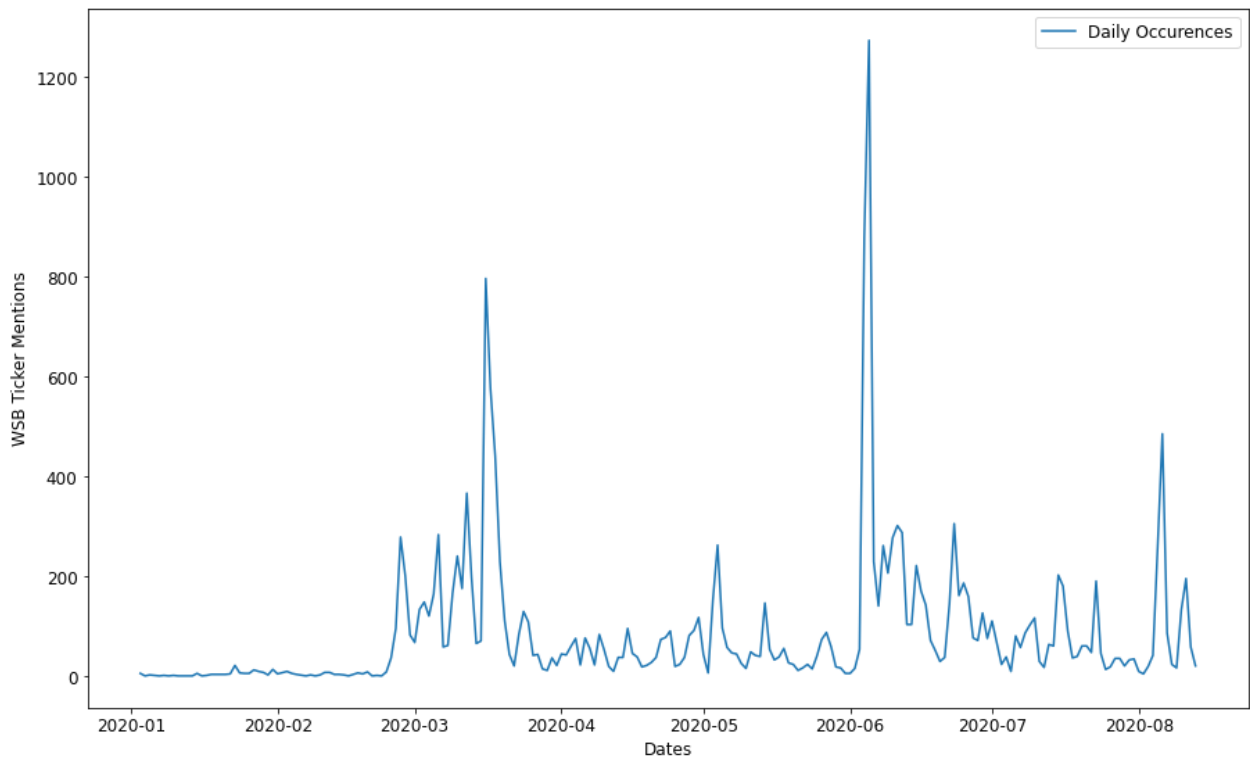


Figure 18: Reddit WallStreetBets – Mentioning of ticket “JETS” from 01/2020 until 08/2020:

