**Q5: Cluster Analysis by Value Retention**

**Introduction**

Though the pandemic fear triggered selling-off of almost all stocks in the first few weeks of the market crash, some industries were hit particularly hard while others stayed relatively strong against the downward trend. Moreover, after traders recovered from panic, they realized some industries might still stay steady during the entire pandemic period. As an analyst from J.P. Morgan noted in a recent article: “The other 78% of market cap [of S&P 500] comes from tech, financials, health care, utilities, communications, consumer staples and internet retailing, all of which should be less impacted by the downturn.” [[1]](#footnote-0) Stocks from these industries quickly bounced back from the recent bottom. Moreover, some companies, such as those from the pharmaceutical industry, benefited from the pandemic. Their stocks rose during the period we are investigating as a result.

On the last day of the dataset we are studying (Mar 23rd), the S&P 500 Index has decreased by one third from its recent highest point on Feb 14th. The market has since bounced back steadily.

**Analysis Journey**

In this part, we have done the clustering analysis according to the % value retention for 505 companies. Value retention is a simplified but effective measure that reflects the strength of a company to retain its value when the storm hits. The formula for calculating the % value retention is:

**Step 1: Drop the NA row**

When calculating the value retention, we fail to find out the prices for United Technologies. In that case, its value retention is NA and we drop the row for further analysis.

**Step 2: k Means Clustering**

We use the kmeans function to do the clustering and set the number of clusters as 3. After getting the result, we define the clusters as HVR, LVR and MVR according to the descending order of their cluster means. Finally, we create a table of cluster names and its size for further comparison.

**Step 3: k Median Clustering**

We use the pam function in the package “cluster” to do the k median clustering and also set the number of clusters as 3. Similarly, we define the clusters as HVR, LVR and MVR according to the descending order of their cluster median value. Finally, we create a table of cluster names and its size for further analysis.

**Step 4: Choose k Median Clustering as our final method**

After comparing the sizes of the smallest group of k means and k median, we find out that the smallest group of k means has more members than that of k median. Therefore, we decide to use k median for the final algorithm since it avoids a cluster with fewer companies in it.

**Step 5: Visualization**

We use plotly in R to draw the clustering visualization. The x-axis is the unique index for each company, the y-axis the % value retention, the colors represent three types of clusters.

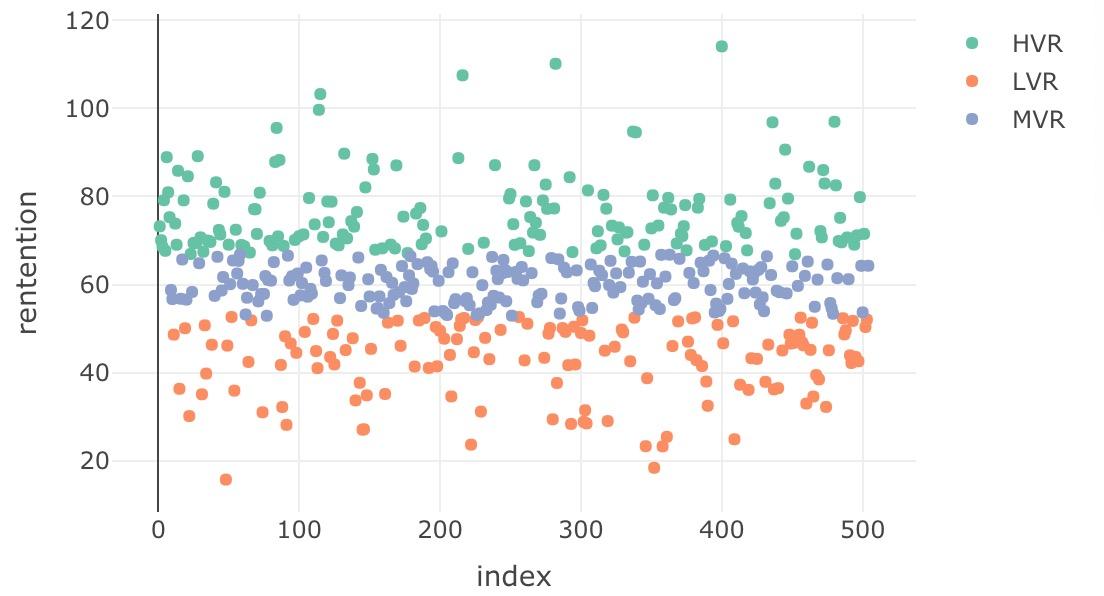


Exhibit 1 - clustering visualization

**Sector Analysis by Clusters**

According to the table below, we could see that consumer discretionary, financials, healthcare, industrials, information technology are the top 5 sectors in terms of composite share in the S&P 500 Index. However, consequences brought by panic are significantly different to all kinds of industry when the US started reporting large coronavirus infection numbers.

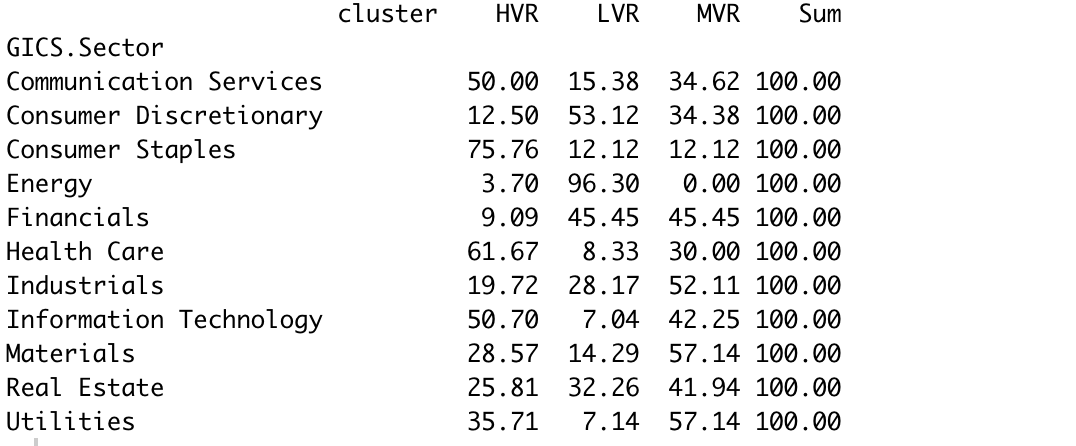
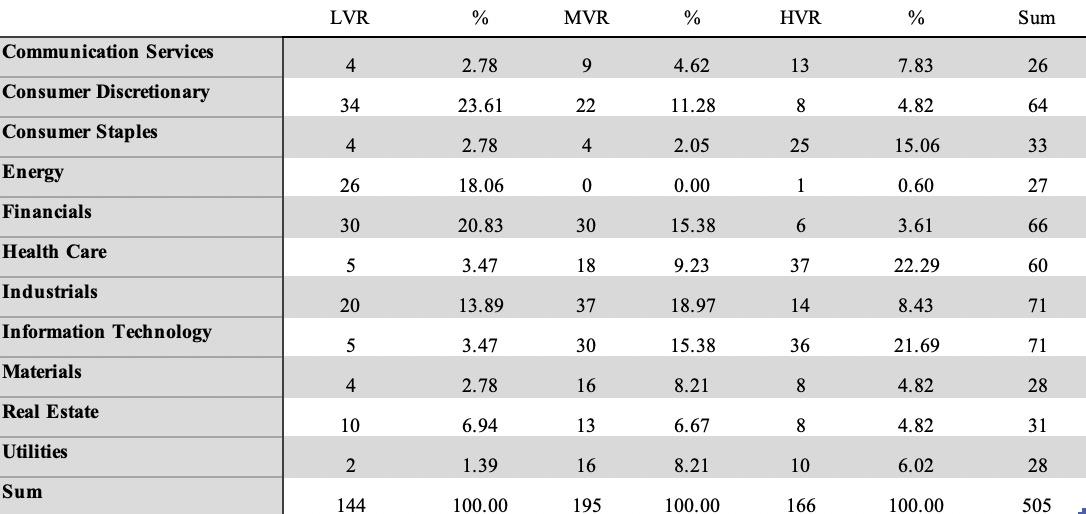


Exhibit 2 - clustering information in sectors

Consumer Discretionary, Energy and Financials ranked top 3 in the Low Value Retention group, each sector accounting for approximately one fifth of the S&P 500 Index in terms of market cap. The Energy industry was hit especially hard in this period, while 96.3% of energetic companies had weak strength to retain its value and survive the storm. Energy industry suffered the greatest losses in this pandemic. The main reason for the crash of energy companies’ stocks is well explained in another sector of the report. The pandemic disrupted the demand for crude oil. For example, China, in which more than a half of industrial products in the world are produced, was hit by the pandemic during the first two months of 2020. It suspended most of its business activities and the demand for crude oil experienced a sudden halt as a result, since China is one of the world’s largest oil importing countries. On the other hand, crude oil prices crashed owing to the failure to reach an agreement between OECD and Russia on a production cut. Saudi Arabia, which led the OECD, announced later that it would recover to a historically high oil production and increase daily oil output to 11 million barrels. As a result, oil prices dropped below 20 USD a barrel. While most American drilling companies target shale oil and are heavy on debt, a price below 20 USD put their business at risk and caused a massive selling-off of their stocks.

At the time I am writing the report, OECD, Russia and the U.S. have reached an agreement on production cut and brought back the oil price to around 30 USD per barrel. However, a large number of American shale oil drilling companies are still in danger of insolvency and possible bankruptcies.[[2]](#footnote-1) Consequently, their stock prices are nowhere near the February peak.

China, the country most heavily affected by the virus initially, is the main global production source of many clean energy technologies, such as solar panels, wind turbines. The disruption of Chinese economy causes potential supply chain bottlenecks for some technologies and components.

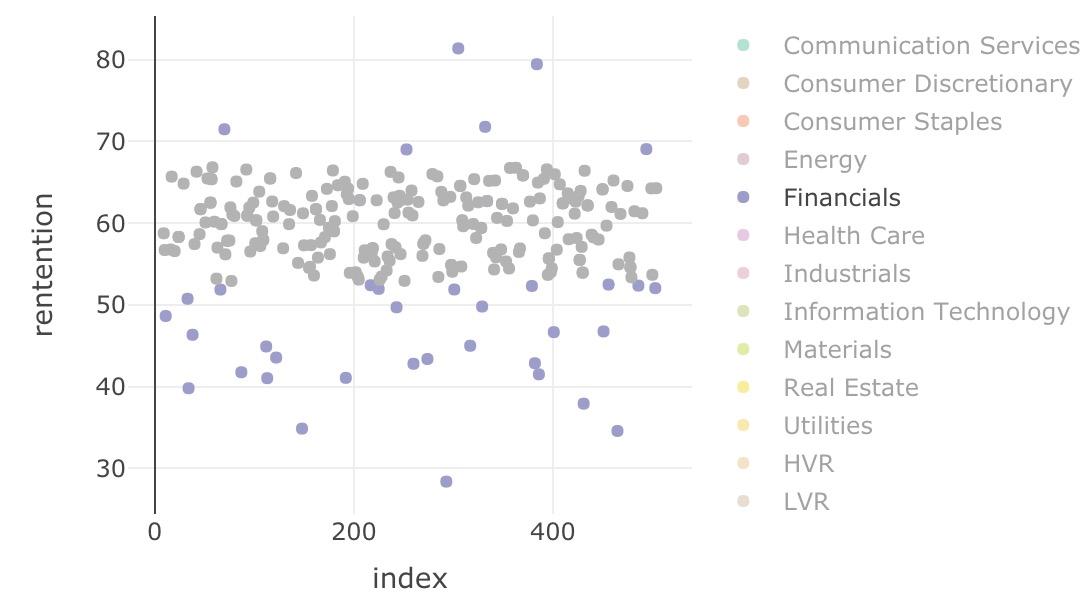
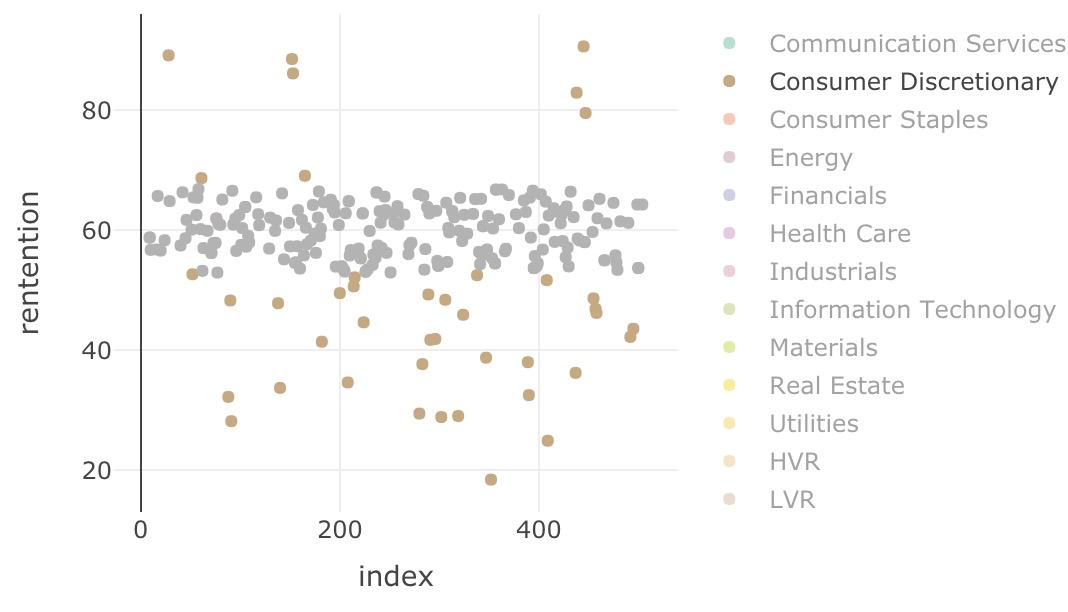


Exhibit 3a - Consumer Discretionary in LVR Exhibit 3b - Financials in LVR

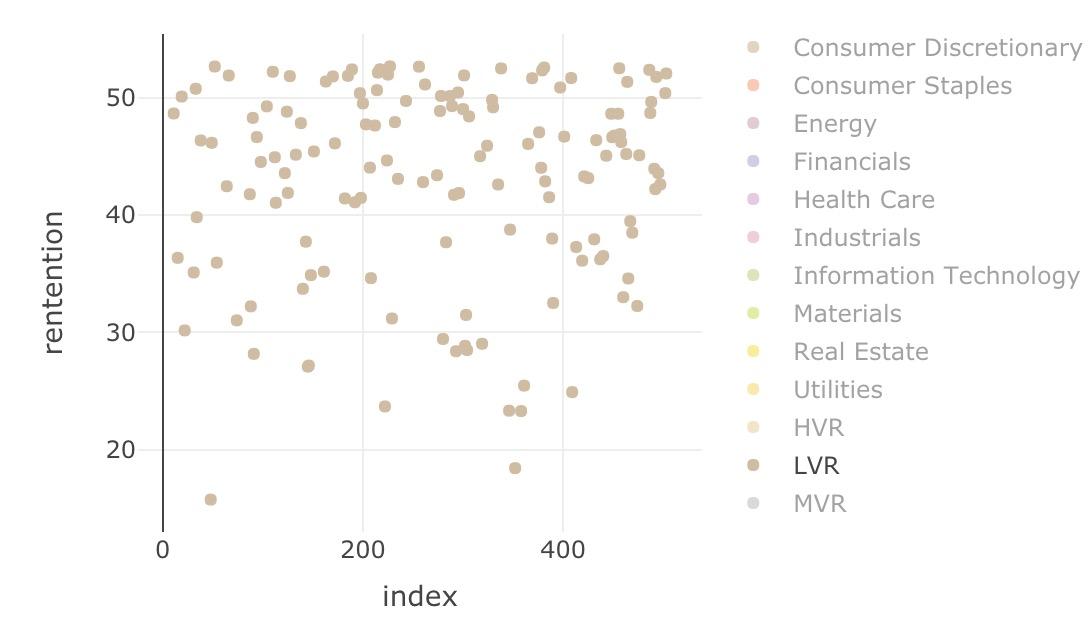
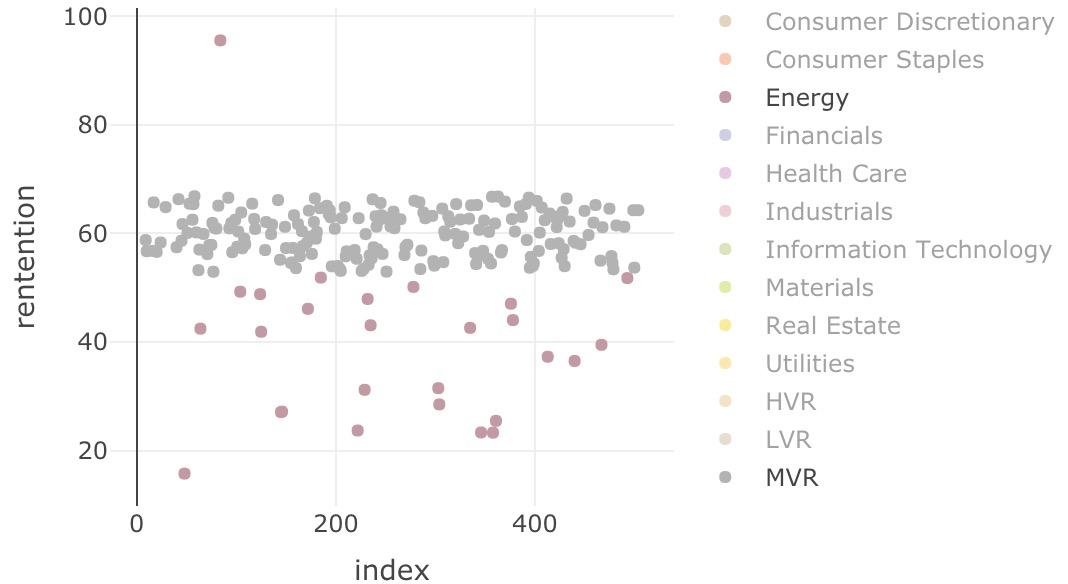


Exhibit 3a - Low Value Retention in LVR Exhibit 3b - Low Value RetentionR

In the Middle Value Retention group, value retention ranged from 53% to 67%. Since the majority of the energy corporation Energy industry suffered heavy losses and clustered in the LVR group, it even disappeared in the MVR group. Information Technology, Industrial, Financials ranked top 3, accounted for 15.4%, 19%,15.4% respectively in MVR group.

Financial institutions have been under rigorous supervision since the financial crisis in 2008, and have been required to perform stress tests each year. Owing to their ability to keep a healthy balance sheet and strong liquidity, most financial institutions have been staying afloat (at least partially) during the market crash.

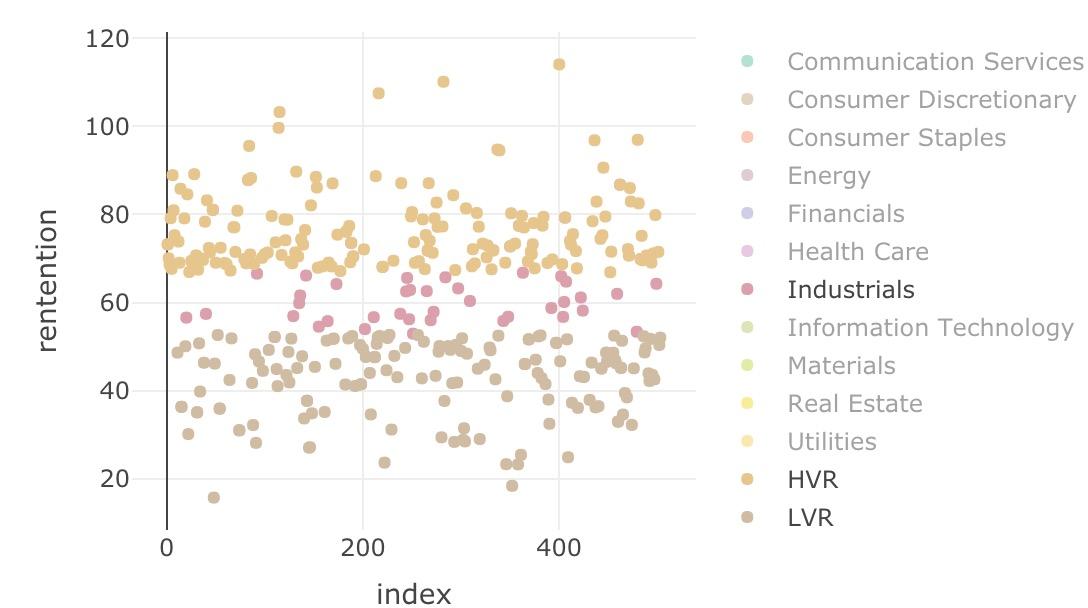
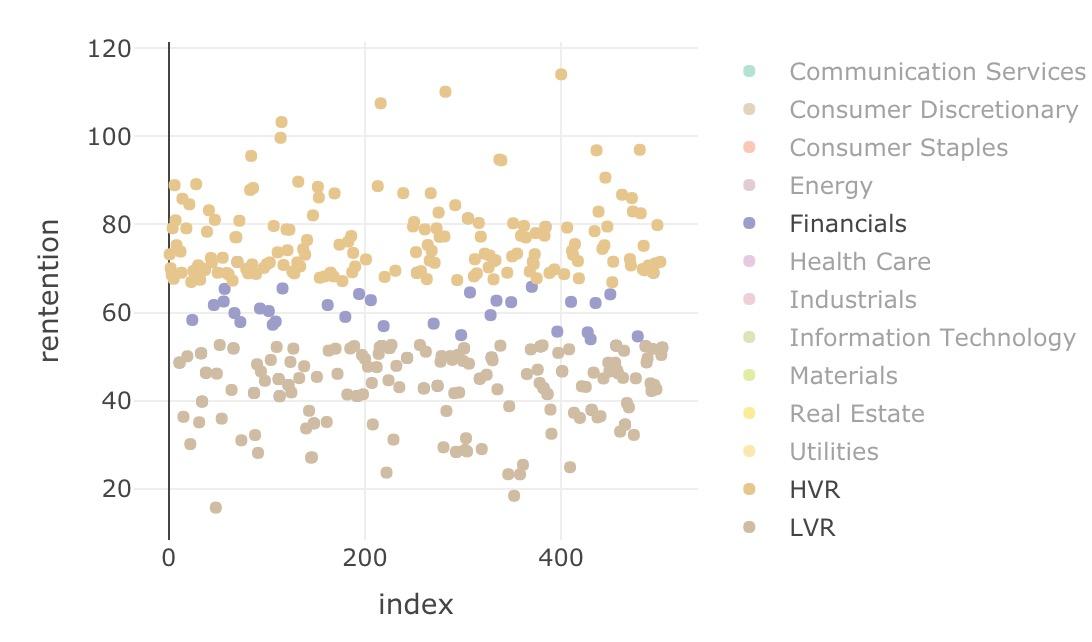


Exhibit 4a - Financials in MVR Exhibit 4b - Industrials in MVR

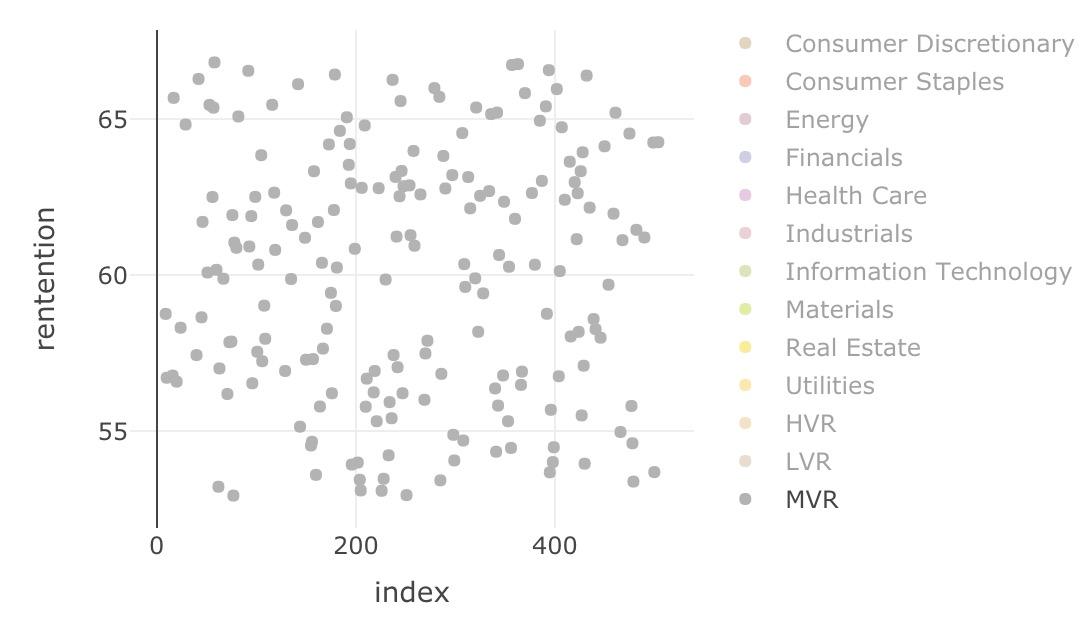
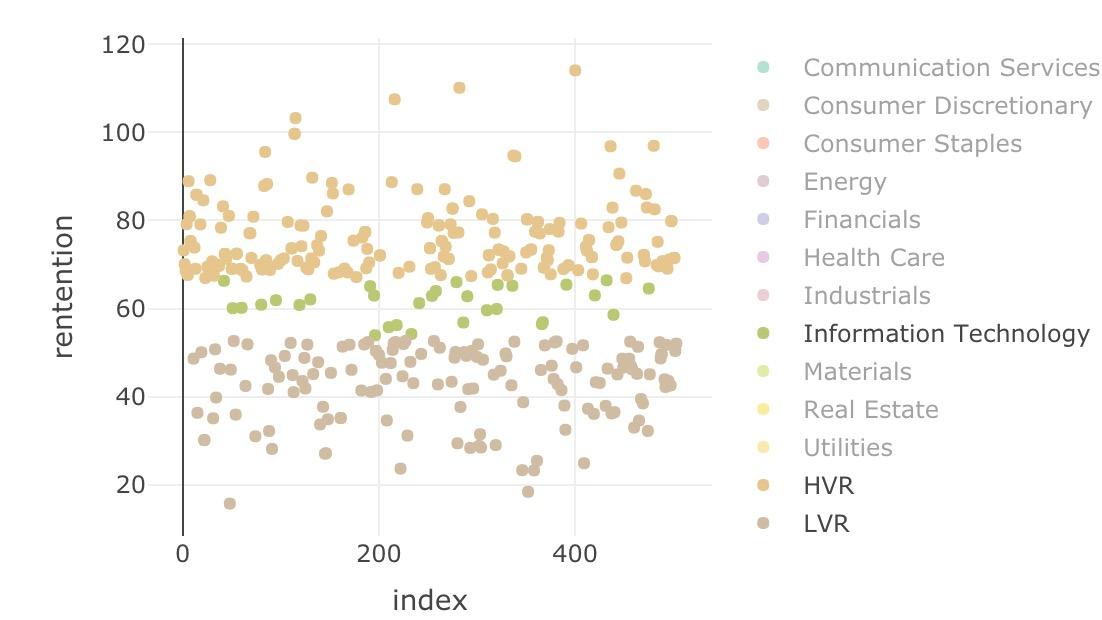


Exhibit 4a - Information Technology in MVR Exhibit 4d - Middle Value Retention

In the high value retention group, Information Technology did a decent job in staying strong against Covid-19, ranked second largest proportion. At the same time, the industrial and financial industry disappeared while the healthcare sector performed quite well in the HRV group, ranked first in the whole market. Besides, the consumer staples sector took up 15% of the HRV group, ranked third with 36 companies.

Most information technology companies have a lot of cash in hand and are light in assets compared to energy companies. A healthy financial status means IT companies can make it through the crisis facing a decline in revenue in the coming months. Moreover, businesses operated by IT companies are less impacted by the pandemic compared to those operated by retail stores. Some might benefit from the pandemic. As a large population are trapped on the sofa during the quarantine, they will turn on Netflix or HBO for entertainment. Subscriptions to streaming services indeed experience an explosive increase in March. Meanwhile, they rely on Zoom or other video conferencing platforms to work. All these companies fall into the category of Information Technology, and as they withstood the crash fairly well, the IT industry as a whole remained intact.

Meanwhile, it is no surprise that the consumer staples industry withstood the crash as well. People still frequented grocery stores for living necessities, e.g. toilet papers. Sometimes, people buy more than regular time. In the case of toilet paper, people pile up toilet papers in storage rooms for a stupid and nonsense fear of global toilet paper shortage.[[3]](#footnote-2) Moreover, restaurants are closed in some states, and as a result, people are forced to cook at home. Those who never cook before now go to grocery stores for food. For example, I myself now spend more than usual time at Costco. As a result, consumer staples companies are expected to deliver better than average earning reports and investors tend to hold their stocks.

Last but not least, the healthcare industry stood strong against the market crash for obvious reasons. The rumors around the soon-available vaccine from a number of companies, such as Moderna, news surrounding curing effects of Remdesivir delivered by Gilead and fast test kits developed by Abbott make some biotech companies the very rare winners in the equity market crash.

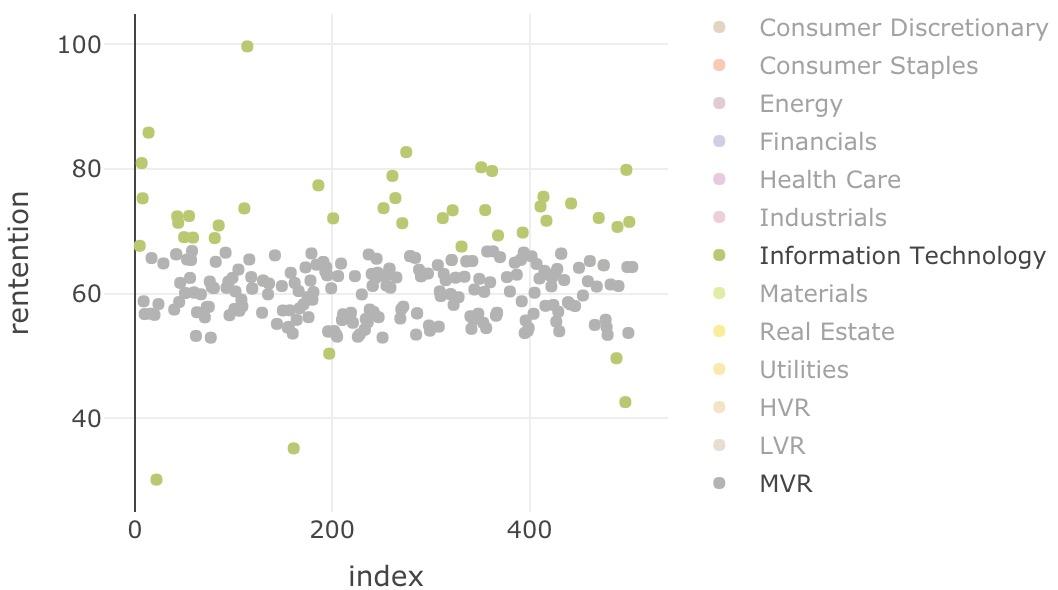
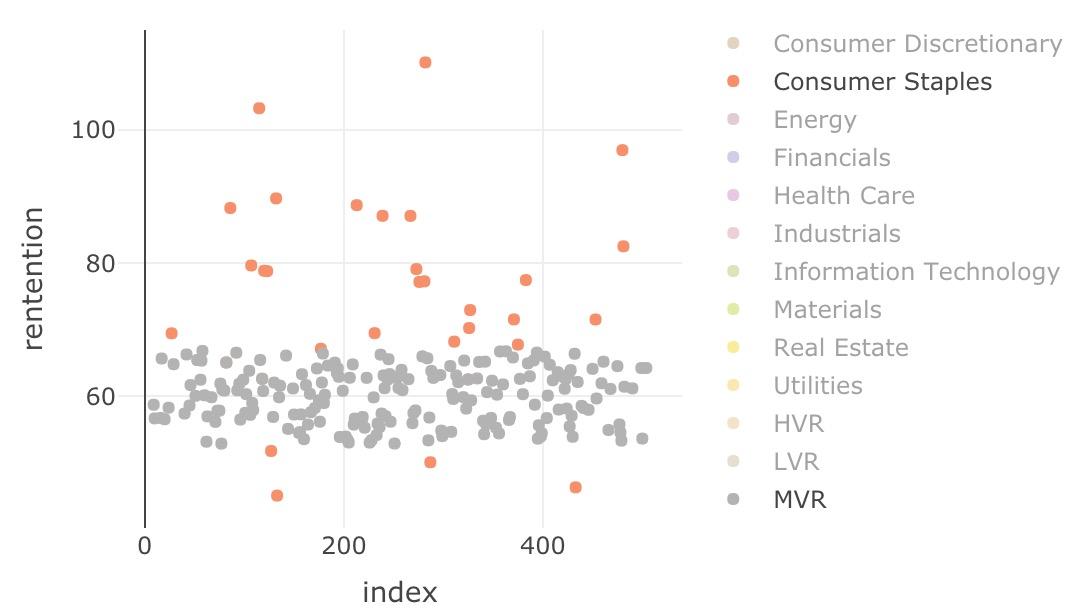
 

Exhibit 5a - Information technology in HVR Exhibit 5b - Consumer Staples in HVR

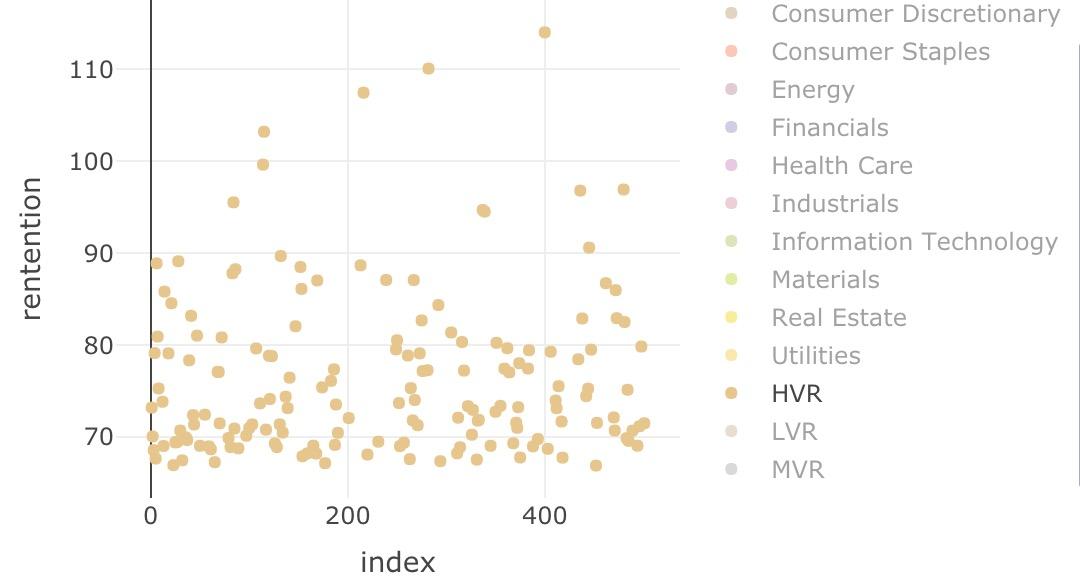
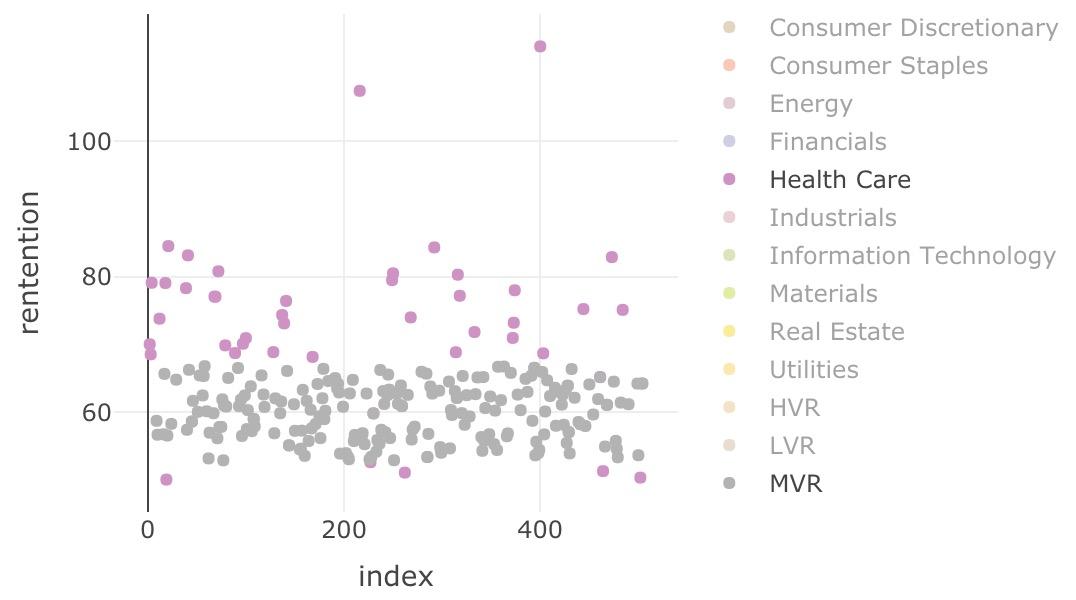


Exhibit 5c - Healthcare in HVR Exhibit 5d - High Value Retention

Energy companies were already struggling with dropping oil prices and uncertain market conditions before covid-19 began. With the virus spreading around the world, energy companies try to understand and contain its effects. It’s no doubt that energy companies suffer greatest losses among industries. In contrast, Industrial and financials stand slightly strongly against the virus, since they are part of people’s needs. On the contrary, healthcare and consumer staples did quite a good job during the pandemic period. Healthcare consists of medical supply companies and services that aim to improve the human body. With increasing numbers of confirmed cases, the demand for Healthcare rapidly grows. Besides, Consumer staples companies provide all necessities of life. During the Self-quarantine period, people need to stay and eat at home. Therefore, the demand for consumer staples dramatically increases, which lead to strong retention ability of consumer staple companies.

1. <https://www.linkedin.com/pulse/why-stock-market-isnt-more-worried-david-kelly/?trackingId=> [↑](#footnote-ref-0)
2. <https://www.reuters.com/article/us-global-oil-layoffs/oilfield-companies-cut-jobs-brace-for-bankruptcies-idUSKBN21P2S5> [↑](#footnote-ref-1)
3. <https://www.washingtonpost.com/national/coronavirus-toilet-paper-shortage-panic/2020/04/07/1fd30e92-75b5-11ea-87da-77a8136c1a6d_story.html> [↑](#footnote-ref-2)