



# FELLOWSHIP PROGRAM ASSIGNMENT COVER SHEET

**THIS FORM MUST BE AT THE FRONT OF EACH ASSIGNMENT  
CANDIDATES MUST KEEP A COPY OF THEIR ASSIGNMENT**

Candidate to complete the following section (and update details in header and footer):

<p><b>SUBJECT:</b></p> <p><b>Data Analytics Applications</b></p>
<p><b>DATE DUE:</b> Monday 27 September 2021 at 9.00am (AEST)</p>

- Please ensure that your candidate number and subject name is located on the header and footer of each page of the assignment.
- By completing and submitting this cover sheet you are confirming that this assignment is your own work, and all material that is used is correctly referenced and cited.

# Data Analytics Applications Assignment

## Question 1: Domain knowledge

Word count excluding titles, charts, tables, references: 998

We present research on the characteristics of this industry, including its competitive dynamics, future prospects, ownership structure, key people and regulations.

### Market dynamics

**Commercial free-to-air (FTA) television in Australia is dominated by Seven West Media (SWM), Nine Entertainment (NEC) and the ViacomCBS owned Network 10 (TEN).** These companies comprise the majority of the market share in this industry. This is supplemented by the Australian Broadcasting Corporation (ABC) and the Special Broadcasting Service (SBS), both of which are government broadcasters. There also exist regional networks including NBN Television and WIN, some of which are owned by SWM or NEC.

Company	Market Share
SWM	33%
NEC	28%
TEN	16%
ABC	16%
SBS	8%
Total	100%

Source: [Compare TV – Australia TV ratings](#)

The oligopolistic nature of the FTA market is reflected in the market capitalisation and share prices of these companies. For example, NEC has observed a stable share price over the longer-term. This is expected for a company that has retained high market share in an industry with stagnant growth, like the television industry.

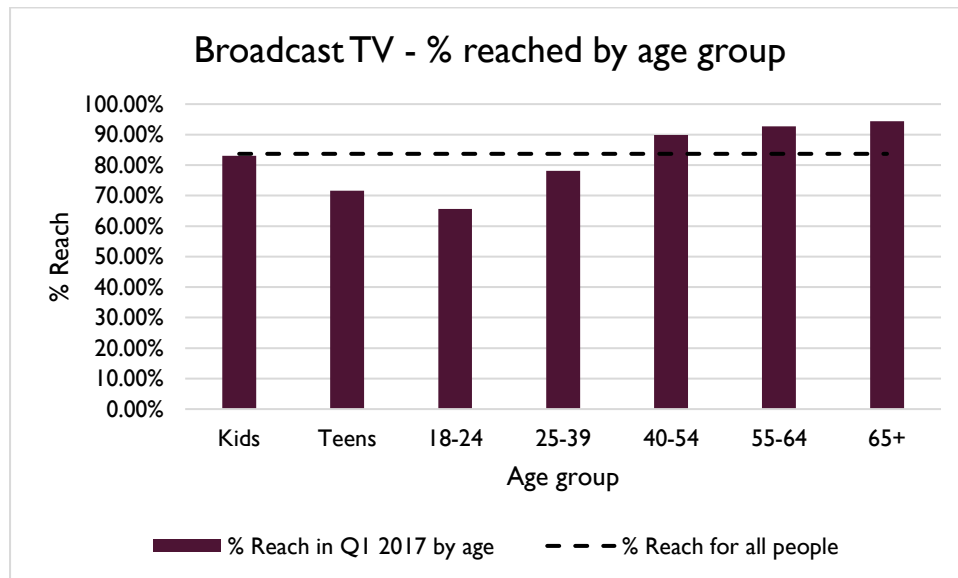


Source: [ASX](#)

Analysing the impact of tweets on the TEN share price is not possible since the takeover by ViacomCBS. ViacomCBS is a multinational conglomerate operating many networks across several continents, and it would thus be impossible to isolate how Viacom's stock moves in relation to TEN.

## Industry projections

The viewership of commercial FTA television has been in decline in recent years. Moreover, it is unpopular amongst younger people, who watch fewer hours of television per day. This is due primarily to the rise of streaming services like Netflix, which afford viewers more convenience by allowing them to consume a broader range of content, whenever it suits them.



Source: [OzTAM Q1 2017 Australian Video Viewing Report](#)

Due to low viewership, we would expect declining company revenues and profits, and therefore share prices. This may be somewhat offset by each network introducing its own streaming service to mitigate some of the lost viewership.

## Ownership of Seven and Nine Network

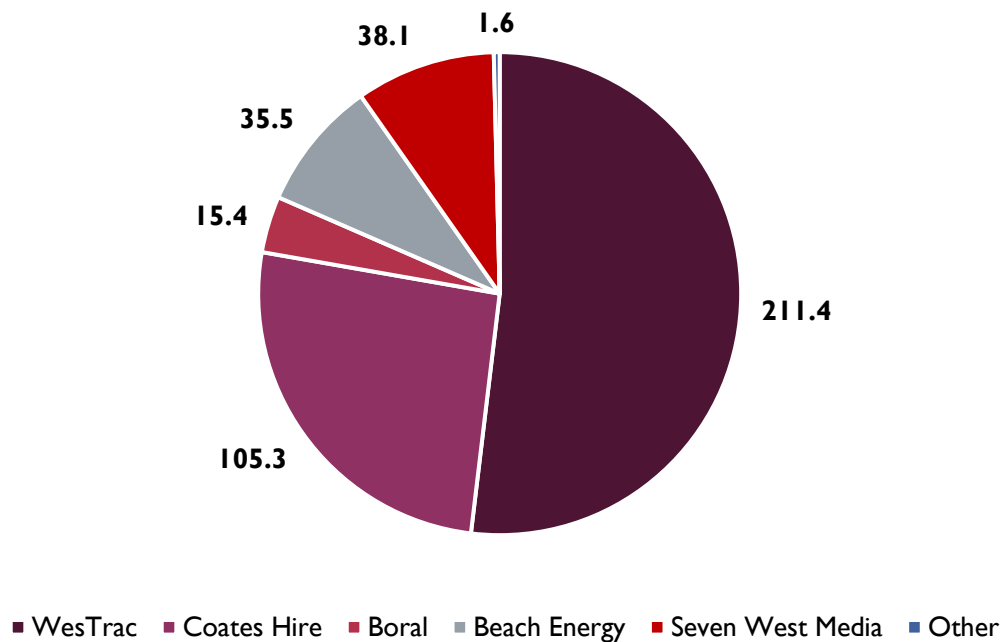
All Seven Network channels are wholly owned by SWM. This company is ASX-listed and trades under the ticker ASX:SWM; the largest shareholder of which is Seven Group Holdings (40% ownership). The latter is the parent company, trading under the ticker ASX:SVW.

This has severe impacts on how we structure the analysis. The investment manager has requested an analysis of the relationship between television-related tweets and SVW. However, SVW is a conglomerate owning businesses in numerous industries. Therefore, the analysis is only valid if a large proportion of SVW's business relates to media.

The chart below depicts the breakdown of SVW's profit (EBIT) by segment. Channel Seven only contributes AUD 38.1M (9%) to SVW's profit. SVW's other businesses relate to equipment hire and dealing for the construction and infrastructure industries (WesTrac and Coates Hire), construction materials (Boral) and oil/gas (Beach Energy). These industries obviously bear no relation to media, suggesting that tweets relating to Channel Seven should have immaterial impact on the SVW share price. On the other hand, we show that most of

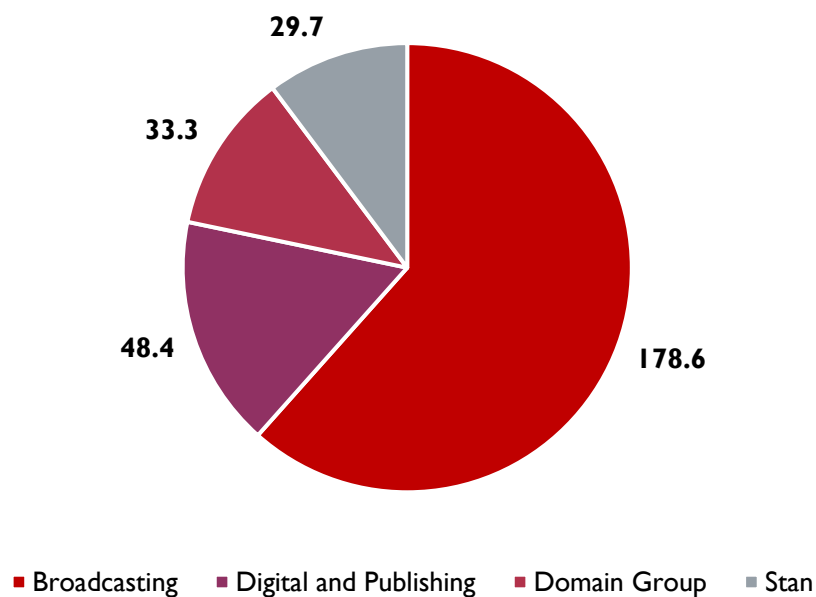
NEC's business relates to broadcasting and streaming; therefore, it only makes sense to analyse how tweets could impact NEC's share price.

Seven Group Holdings (SVW) - Earnings Before Interest and Tax  
breakdown by business segment



Source: [Seven Group Holdings Limited – Results for the half-year ended 31 December 2020](#)

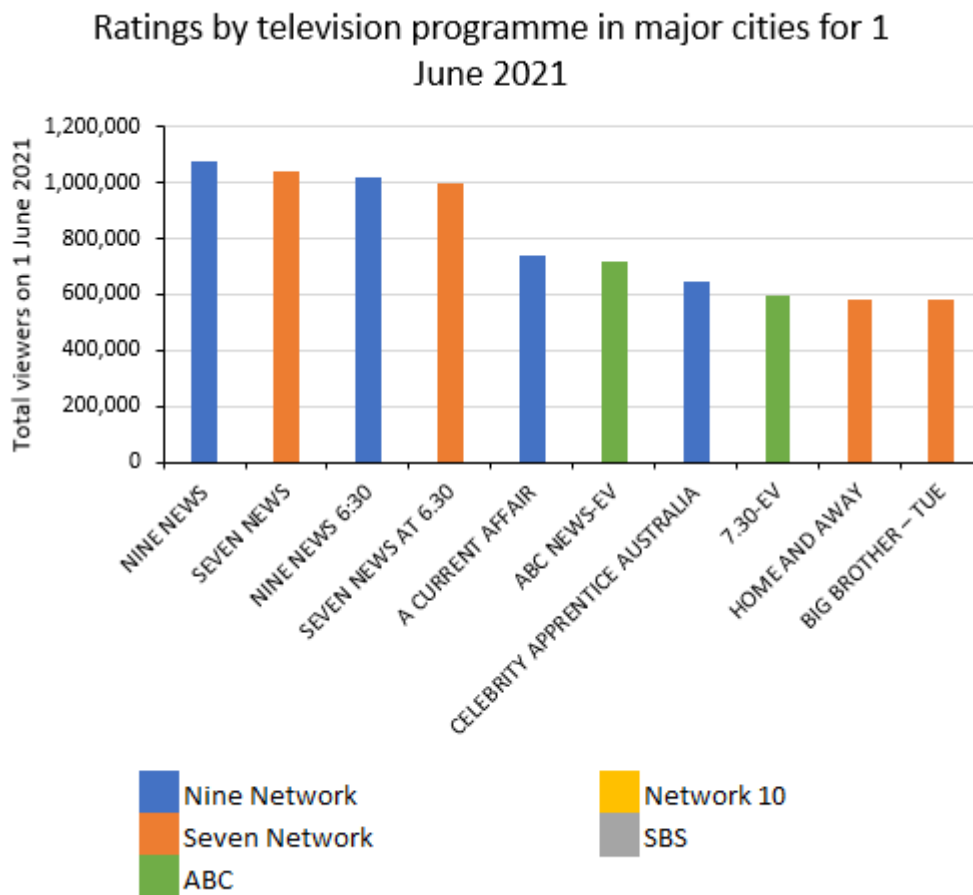
Nine Entertainment (NEC) - Earnings Before Interest and Tax  
breakdown by business segment



Source: [Nine Entertainment Co. Holdings Limited – Results for the half-year ended 31 December 2020](#)

## Influential people in the Australian television industry

In Australia, a few television shows typically dominate the ratings. The evening news programmes ran by the commercial FTA channels dominate ratings; with Nine News, Seven News, A Current Affair and ABC News often topping ratings (in the absence of major sporting broadcasts like State of Origin).



Source: [TVTonight ratings for Tuesday 1 June 2021, from OzTAM](#)

Highly followed television presenters featured on these news programs include Mark Ferguson, Peter Overton, Sandra Sully, David Koch, Waleed Aly, Tracy Grimshaw and Leigh Sales.

As a result of their dominance, a few well-known television presenters may have enough of a following to be able to move stock prices based on their tweets. Tweets referencing or tagging these presenters may receive greater interaction than others, thus potentially move prices.

## Local content quotas on Australian television

The Australian Communications and Media Authority imposes quotas on commercial FTA television. Specifically, these channels are required to “broadcast at least 55% Australian content between 6am and midnight on primary channels, and 1460 hours of Australian

content (per year)”<sup>1</sup>. The purpose of these quotas is to promote television programs reflective of Australian culture.

As a result, local television networks like Seven and Nine are required to invest in developing locally produced television shows. In recent years, FTA networks have argued that such investment is a “waste of money”<sup>2</sup> due to low viewership. This is especially true when producing local children’s content and may be reflective of children’s greater engagement with online platforms like Youtube Kids. This is relevant as such investments likely result in depressed profits for the FTA networks, thereby contributing to the downward trend in share prices.

### **Release of programming and other sensitive information**

It is necessary to understand the ASX’s Continuous Disclosure laws before designing such an analysis. These laws apply to all ASX-listed companies, including SVW and NEC.

Many of the tweets covered in this study are just marketing for that network’s current or upcoming shows; for example, many tweets promote the upcoming series of Big Brother on Seven Network. It’s worth noting that none of these tweets should be “new” information. Under the ASX’s Continuous Disclosure rules, publicly listed entities like SVW and NEC are required to disclose any sensitive information that could potentially impact share prices. Such information must be disclosed directly via an announcement on the ASX, rather than by Twitter or other social media.

***Once an entity is or becomes aware of any information concerning it that a reasonable person would expect to have a material effect on the price or value of the entity’s securities, the entity must immediately tell ASX that information.***

---

Source: [ASX Continuous Disclosure: An Abridged Guide](#)

While Twitter is frequently used as a source of up-to-date and new information, it cannot be used to trade on new price-sensitive information unless the company violates its Continuous Disclosure obligations.

For example, while the @Channel7 Twitter account may announce they’ll broadcast the Tokyo 2020 Olympics (a major contract worth millions in revenue, and thus price-sensitive), this must be announced directly to the ASX (and also on their Investor Relations page, as done [here](#)) and **cannot** be announced first via Twitter. Continuous disclosure therefore decreases the potential profit that could be generated by automating trades using tweets,

---

<sup>1</sup> [Australian Communications and Media Authority – Content rules for commercial TV](#)

<sup>2</sup> [Sydney Morning Herald – Australian commercial television networks want to scrap children’s content quota \(July 2017\)](#)

because such information is already public and under the (semi-strong) Efficient Market Hypothesis should already be incorporated into the stock price.

## Question 2: Compliance with Developer Policy

Word count excluding titles, charts, tables, references: 709

### Off-Twitter Matching

This analysis will make use of a large volume of publicly available data from Twitter. In addition, it is commonplace for data scientists to harness a wide range of data sources and information to improve the predictive power of their models. This is achieved through extensive data linkage techniques.

However, Twitter limits the extent to which this can be performed to protect the safety and privacy of its users. This is defined as part of its rules on **Off-Twitter Matching**. Twitter disallows the linking of Twitter content like handles, tweets, and IDs to external information such as their real identity or household location, **unless** we obtain express, opt-in consent, or the information is otherwise publicly available.

In the context of this analysis, this could take many forms. For example, an employee of Nine Entertainment may pseudonymously post a rumour on Twitter that Nine has secured the rights to the AFL Grand Final, a highly profitable contract. This information would likely have a material impact on the share price. A trading algorithm may be able to scrape the Internet to connect this user's handle to their real life identity, and determine they are a Nine employee, assigning greater validity to the rumour.

Such a practice would be in violation of the developer policy. Our analysis will comply with the Off-Twitter Matching policy by exclusively using data from contained within the tweets, and stock price data. An internal "model audit" role could also be established to identify if the model violates this policy and report to the appropriate management. This will help ensure that no external data is matched to and used as part of the analysis.

### Pay to Engage

In 2021, it is commonplace for algorithms to make trades based on tweets, and tweets from certain individuals can have material impacts on the prices of securities like shares and cryptocurrency. While this is most well-known regarding Elon Musk's tweets about Dogecoin, it is also prevalent in smaller markets like the ASX. For example, the chart below shows how the share price of Archer Materials (ASX:AXE) fell by over 40% immediately after a negative tweet was made by popular trader Nick Fabrio.





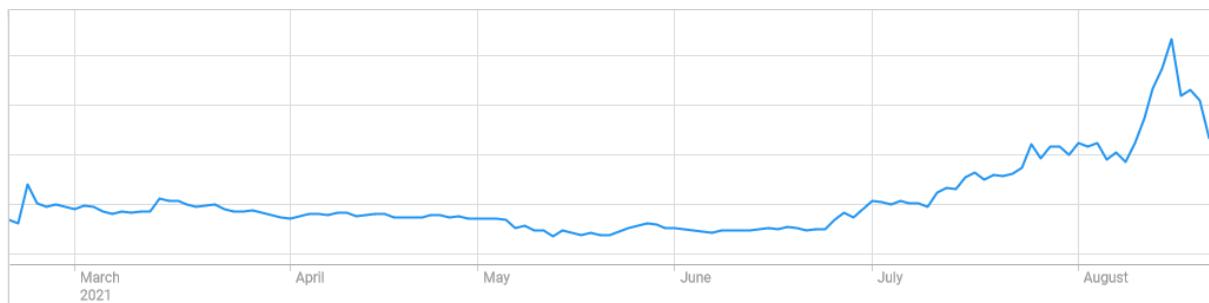
Nick Fabio  
@longhorncapital

...

Sooo anyone care to tell me what kind of hijinks this company deals with? So far all i see is a few patents... Also, is there such thing as a patent in China? lol \$AXE



1:03 PM · Aug 16, 2021 · Twitter Web App



Source: [ASX](#), [Twitter](#) (Nick Fabio)

An investment company seeking to make profits on these trades may find it more consistently profitable to pay well-known or highly followed individuals to make tweets about stocks (or other traded securities) and front-run these trades accordingly.

Twitter's Pay to Engage policy disallows services that compensate individuals for actions such as posting tweets, retweeting, commenting, etc. Moreover, such actions are considered market manipulation by the Australian Securities and Investments Commission (ASIC) and are therefore deemed illegal<sup>3</sup>.

<sup>3</sup> [Australian Securities and Investments Commission – “Market Integrity Update – February 2021”](#)

To avoid violating Twitter's policies and Australian financial services laws, the trading algorithm designed **must not** have the ability to interact directly with tweets (apart from viewing them). Furthermore, the users of any algorithms written from this analysis must agree to not use Twitter to make stock-related posts. If no tweets are written as part of this analysis (either by the algorithm or by a human using it), then the analysis will be compliant with Twitter's Pay to Engage policy.

### **Content compliance**

Twitter's content compliance policy requires any Twitter content stored offline to be kept up-to-date with its current form on Twitter. This is relevant because large volumes of Twitter data (including the actual content of tweets) will be stored offline to build a training dataset for the model. This could cause problems if, for example, a tweet accidentally made by a celebrity containing sensitive or private information is saved offline. Under Twitter's content compliance policy, the investment fund must destroy this data "as soon as reasonably possible", or within 24 hours of Twitter or the poster requesting so.

Numerous controls can be used to ensure the analysis is compliant:

- Establishing a compliance team or manager to respond to deletion/modification requests from Twitter or posters.
- Running a script that checks if the Tweet IDs of any saved tweets are no longer active, and deleting those tweets accordingly.
- Destroying all Twitter data one year after its collection or once no longer needed, to ensure that any tweets the compliance team may have missed are eventually deleted.

These controls will help ensure that the company knows when a tweet it has saved is deleted or modified, and react quickly.

## Question 5a: Key risks

Word count excluding titles, charts, tables, references: 514

### Operational risks

If the fund uses these models to influence and automate trading, there is significant risk associated with coding errors. For example, a single line of incorrect logic could reverse the bids and offers, causing the wrong trades being performed. An extra digit could cause bids to be off by an order of magnitude.

This risk may be considered acceptable if a human manually vets all trades before executing them. However, this would waste time on every trade, and often a profitable opportunity could disappear by the time a trade a human executes the trade. Funds will often automate thousands of trades per minute, resulting in extreme losses by the time a problem is identified.

This risk should be considered low frequency and extreme severity. This risk materialised for Knight Capital Group in 2012, where an errant trading software caused Knight to purchase over \$7 billion in stocks within one hour, resulting in \$460 million of losses and requiring an emergency capital raising to avoid bankruptcy<sup>4</sup>.

A related risk is that the model relies on many dependencies. Updates to Python or critical packages like Scikit-Learn could introduce bugs into the model, as was the case when Python 3 was introduced, deliberately breaking backwards compatibility<sup>5</sup>.

### Regulatory risks

ASIC may take issue with the use of tweets for trading because it may result in individuals having the ability to materially influence share prices on their own. If ASIC bans the practice, this could eliminate a (potentially) profitable part of the business.

We consider this risk as moderate severity and high likelihood. This is because ASIC has already expressed concern in relation to the use of social media<sup>6</sup> in trading stocks, due to the extreme market volatility it has caused throughout 2020-21. It could also cause significant negative impact on the company's profits by entirely preventing the use of this strategy.

### Model risks

The model may generate trades that underperform the market average. This model in particular is at risk due to its structure. This is because the neural network and tree-based models were designed to optimise the success rate in identifying a stock price increase or decrease, with no allowance for the size of a price movement. Yale University research suggests share price decreases are larger and more violent than increases<sup>7</sup>, as the (red) 20%

---

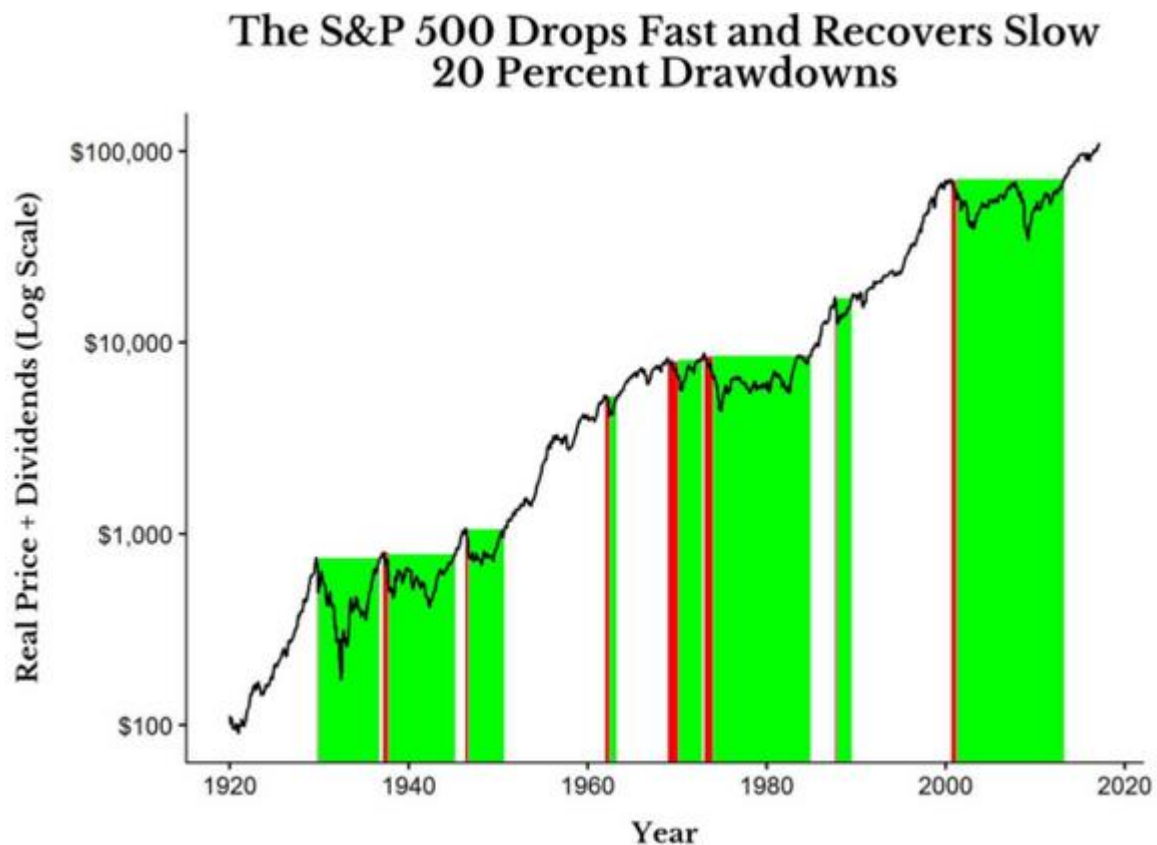
<sup>4</sup> [Securities and Exchange Commission – Administrative Proceeding in the Matter of Knight Capital Americas LLC – October 2013](#)

<sup>5</sup> [Ben James – “Stop Using Python 2” – August 2018](#)

<sup>6</sup> [Australian Financial Review – “Social trading fuels market speculation: ASIC” – August 2020](#)

<sup>7</sup> [Data sourced from Robert Shiller – Yale University](#)

drops occur much more quickly than the 20% gains (green) in the chart below. By not accommodating for the size of movements, the investment fund **risks underperforming the market even if the model classifies the direction of movement correctly**. Benchmarking to an index could resolve this, but significant market risk remains – as evidenced by the Nine Entertainment share price falling over 50% in one month<sup>8</sup> due to the COVID-19 pandemic. Therefore, there is significant risk of underperformance and loss due to market and model-related risks.



<sup>8</sup> [ASX – Nine Entertainment Co. Holdings Limited](#)

## Question 5b: Implementation considerations

Word count excluding titles, charts, tables, references: 485

### Model deployment

When implementing this model, we must consider the best method of deployment. These include:

- **An insight tool**, where the model could compile summary statistics such as what keywords are likely to move stock prices and in which direction, which Twitter accounts have the biggest influence on stock prices, etc. This could assist a trader in deciding whether to buy or sell shares.
- **A decision support tool**, where the model could estimate the probability of a stock moving up or down and the size of the movement based on keywords in the tweet. A human trader can then decide whether to trade on this information.
- **A decision engine**, where the model automatically trades the stock if the predicted probability of a stock movement is high enough. No human trader is involved here.

The deployment form will depend on the model's results and goal. For example, if the model shows that profitable trading windows are quickly arbitrated out, then a decision engine is needed to automate these trades. However, executives and other stakeholders may not trust that the model is reliable. If the goal is to influence trades, it would be easier to gain buy-in for a decision support tool as a human can assess the model results before trading.

### Testing the model

Producing a trading model that outperforms the market is notoriously difficult. Testing is necessary to convince executives to deploy the model. Faults can also be identified, reducing the risk of large trading losses. An experimental setup to test it could include:

- Measuring the annualised return, volatility and **alpha** produced by the model. Investors often select active funds based on these metrics.
- Running the model on other media stocks, such as Seven West Media or ViacomCBS.
- Using the ASX200 or MSCI Media/Entertainment Index as a control group to assess whether the model produces alpha, or above-market returns.
- Monitoring the win rate, drawdown rates, Sharpe ratio and tracking error of the model. For example, executives may pull the model if drawdown rates are too high, as the fund would be exposed to significant losses.

If the model produces above-market returns without significant volatility, it will be easier to convince executives to deploy the model, and investors to become clients.

### Deployment architecture

It is important to consider the best architecture pattern for the use case. For example, if analysis suggests stocks must be bought quickly after a tweet to profit, then realtime processing of tweets is necessary to capitalise on this. If the Twitter API has a rate limit

capping how quickly data requests can be made, then realtime processing is infeasible. Excessive queries can also flood the API and cause outages, potentially causing model failure.

On the other hand, a batch process that collects tweets every hour and predicts the resulting stock movement may miss profitable trading opportunities. An on-demand architecture may suffer from similar problems, as a human trader may not always be present to request tweets and make trades.

## **Question 6: Video executive summary**

<https://youtu.be/XFvVcHuw9Mw>