Project Report

Investment R_{mini}

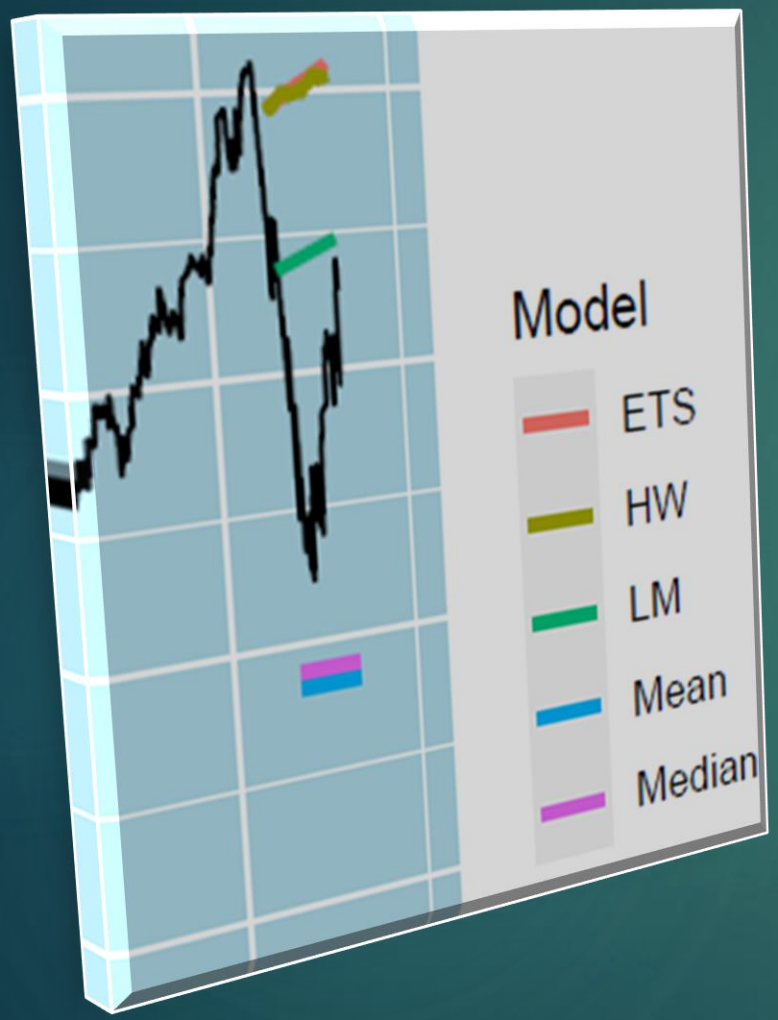
Stock Forecast & Portfolio Recommender

Group 3

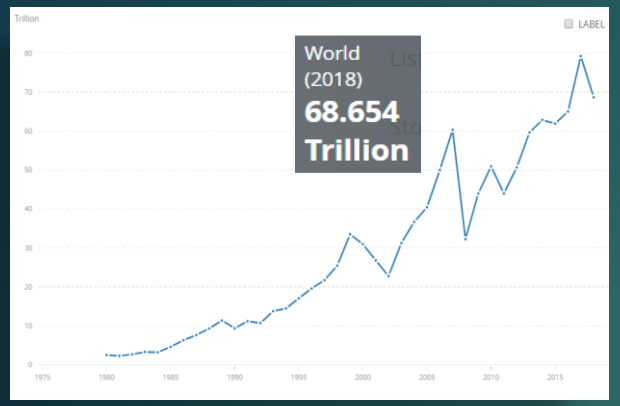
Chng Yan Hao

A0024023A

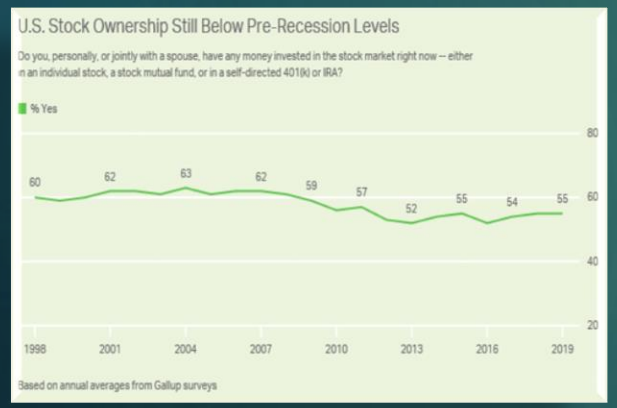
Contents



Executive Summary



[1]Market capitalization of listed domestic companies (current US$), <https://data.worldbank.org/indicator/CM.MKT.LCAP.CD>



[2]Lydia Saad, What Percentage of Americans Owns Stock? <https://news.gallup.com/poll/266807/percentage-americans-owns-stock.aspx>

The world's market capitalization has trillions of dollars (The World Bank 2018, > USD 68T). This amount has grew more than 20 times in the past 40 years.[1] Undoubtably, the stock market is one of the fastest growing where we can find tonnes of money.

It was estimated that only slightly more than half Americans are owning stocks.[2] If Asians are more conservative, this number may well go below half.

For non-investors, just the number of stock markets can be overwhelming (NYSE, Nasdaq, JPX, LSE, SSE, SEHK, … SGX). The stock products and terms are not any easier (bonds, options, shares, long, short, put, call, dividend, margin…). To make things worse, forex exchange may further complex the case. Many potential investors face high entry barrier not only monetarily, also learning fears.

The primary aim of the project is to help those interested to explore more intensively at virtually zero cost. It also offers a non-bias second opinion to those whom are already invested. By taking one mini step, Investment R_{mini} may just become a great stepping stone to begin sailing in the world of stock market. **Contents are designed for users with some statistical, stock and computer knowledge, suitable for DIY users. Application gives a forecast and a portfolio recommendation.**

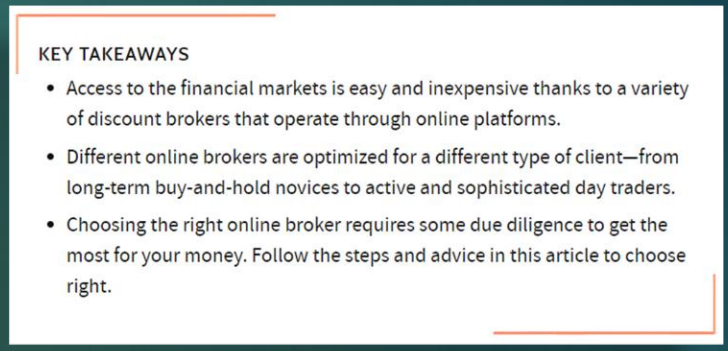
Business Problem Background

With so much money involved, many investment firms offer free trial account with virtual money of no cash value to begin with. Most platforms will take around 15 minutes to sign up. They mainly focus the user for earlier real buy-in, train users on transactions, volumes and historical prices. The range of stock selection is usually limited to what the firm offers, especially custodian accounts.

As an investment is an asset acquired in hope that it grows,[3] the motivation should be coming from how a share may grow and the risks involve. The first questions are likely **"will this stock grow? / how much risk?".**

Investment R_{mini} goes straight to the burning questions. It is built based on statistical methods, models and search algorithms which feed on any selection of stocks suitable, especially suitable for DIY users.

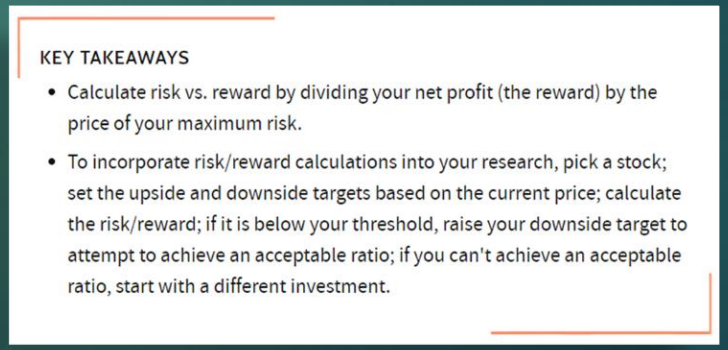
In general, investors have to know their needs and narrow their field.[4] Doing some self-research and gaining some insights is an excellent way to begin.



KEY TAKEAWAYS

- Access to the financial markets is easy and inexpensive thanks to a variety of discount brokers that operate through online platforms.
- Different online brokers are optimized for a different type of client—from long-term buy-and-hold novices to active and sophisticated day traders.
- Choosing the right online broker requires some due diligence to get the most for your money. Follow the steps and advice in this article to choose right.

[4]The Complete Guide to Choosing an Online Stock Broker, <https://www.investopedia.com/investing/complete-guide-choosing-online-stock-broker/>



KEY TAKEAWAYS

- Calculate risk vs. reward by dividing your net profit (the reward) by the price of your maximum risk.
- To incorporate risk/reward calculations into your research, pick a stock; set the upside and downside targets based on the current price; calculate the risk/reward; if it is below your threshold, raise your downside target to attempt to achieve an acceptable ratio; if you can't achieve an acceptable ratio, start with a different investment.

[5]Calculating Risk and Reward, <https://www.investopedia.com/articles/stocks/11/calculating-risk-reward.asp>

[3]Investing Essentials, < https://www.investopedia.com/terms/i/investment.asp>

Project Objectives

Key objectives:

1) **To allow users to explore more intensively with lower entry barriers. Suitable for DIY investors.**
2) Facilitate users with slight statistical, stock and computer knowledge by 'doing the tedious work'.
3) To provide an additional option, opinion and platform for comparison. i.e. against investment platforms, friends, brokers and even robo-advisors.

Sub Objectives:

1) Demonstrate data discovery using open source tools such as R.
2) Demonstrate how an under utilized Microsoft Excel Solver can perform evolutionary algorithm to solve a complex problem.
3) **To have fun exploring shares & stocks in an unconventional manner.**

Project Solution (Part 1)

Feature 1: Forecast

One of the most common methods for forecasting in stock market is to use moving averages. In conjunction with some moving averages, the main solution is to use **Exponential Smoothing, ETS**. Moving averages have constant weights whereas exponential smoothing has exponentially larger weight when closer.[6]

While some datasets may have obvious seasonality, others may only have obvious trends. Different ETS models attempt to capture 'error', 'trend' & 'seasonality' as much as possible. 'ets()' is an in-built ETS function in R that cleverly selects the best ETS model. Holt Winters approach (triple exponential smoothing) is added as reference to capture the trend and seasonality as the best ETS may not show any trend.

Together with the well-known moving averages, quantiles of historical prices and even a linear model, one can have a better stock picture.

As events and circumstances do not occur as expected, the forecast functionality is mainly to serve more of a second opinion. It will not be responsible for damages incurred when forecast deviate from reality.

Simplest form of Exponential Smoothing

$$s_0 = x_0$$
$$s_t = \alpha x_t + (1-\alpha)s_{t-1},\ t > 0$$

where α is the *smoothing factor*, and $0 < \alpha < 1$.

$$\alpha = 1 - e^{\frac{-\Delta T}{\tau}} \qquad \alpha \approx \frac{\Delta T}{\tau}$$

[6]Exponential Smoothing, <https://en.wikipedia.org/wiki/Exponential_smoothing>

Project Solution (Part 2)

Feature 2: Allot

Portfolio allotment and optimization are often done by fund managers. Two of the more computational intensive parts are getting the covariances among shares and iterating the possible combinations to

(i) reduce the risk
(ii) increase return

Optimizing both concurrently to maximize the yield is in a state of art. In Investment R_{mini}, we are using a simple and well expressed term known as **Sharpe Ratio**[7]. It can be view loosely equate as: objective ~ maximize(return / risk)

R has a one line covariance calculation for N items and Microsoft has a great solver add-in to search for an optimal solution. Combining the two will save loads of time and becoming a mini fund manager is no longer a dream. **Evolutionary Algorithm, EA** is the core algorithm to conduct this search. It was chosen as:

(i) Problem resembles a combinatorial problem with no easy mathematical solution.
(ii) Impossible to exhaust all combinations when portfolio is large.
(iii) No one really knows how each solution looks like for different dataset.

$$S(x) = \frac{(r_x - R_f)}{StdDev(r_x)}$$

where:

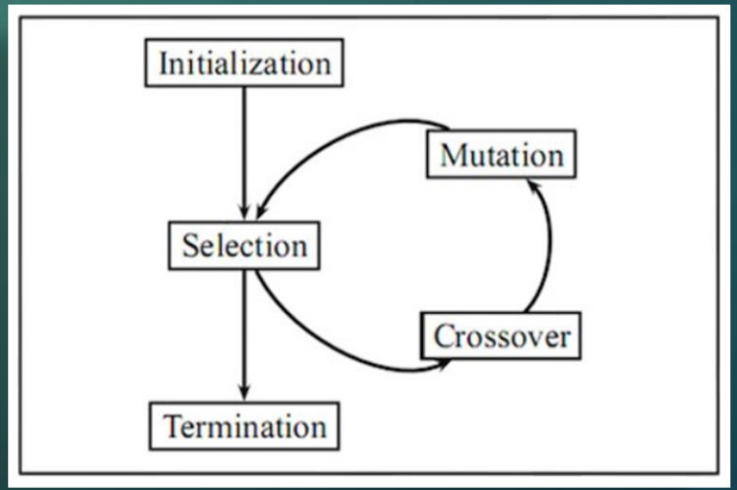
x = The investment

r_x = The average rate of return of x

R_f = The best available rate of return of a risk-free security (i.e. T-bills)

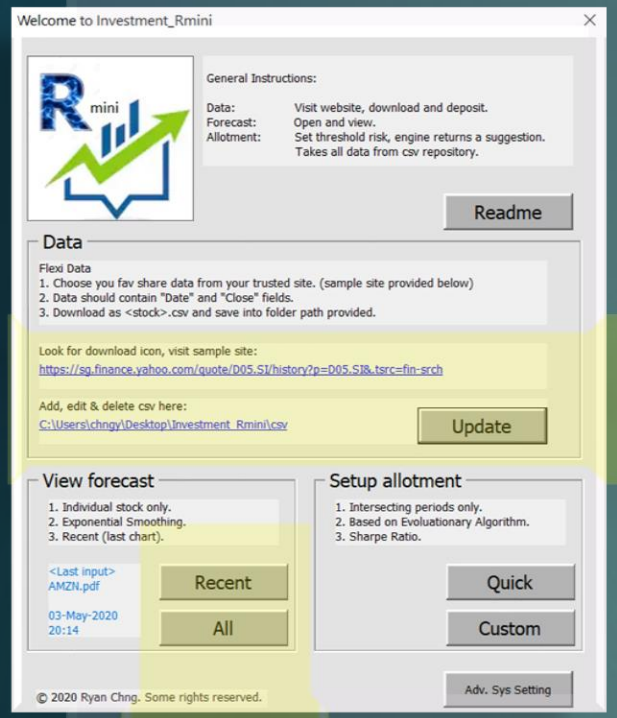
$StdDev(x)$ = The standard deviation of r_x

[7]Sharpe Ratio, <https://www.investopedia.com/articles/07/sharpe_ratio.asp>



[8]Introduction to Evoluntary Alogorithms, <https://towardsdatascience.com/introduction-to-evolutionary-algorithms-a8594b484ac>

Solution Design (Part 1)



1. Download
2. Update

3. View

With 3 Steps:
Download / Update / View

Feature 1: Forecast

As risk and return are hand in hand, different investors have different expectations.

Trend

Past	Med;Ave	Return	Risk
<50d>	1960;2040	60.7%	57.1%
<100d>	1900;1980	77.6%	43.7%
<252d>	1850;1890	20.3%	31.6%
<504d>	1800;1810	24.3%	33.1%

Risk

Test_Set	Pt.Est	MAE	RMSE
ETS		190	221
HW		189	219
LM		206	222
Mean	1200	861	892
Median	1000	1050	1080

Error score

Investor brave: Share will likely go up, I will buy.
Investor timid: Share will likely go up, but too risky.

The forecast can tell one whether the 'share will likely go up or down', but it cannot decide for one whether buying it is worth the risk. Even small errors do not guarantee future stock behaviour will oblige.

Solution Design (Part 2)

Stock	Annual_Return	Risk	Weight
AMZN	38.25%	29.87%	8.40%
BA	6.36%	39.56%	1.49%
BABA	22.90%	32.46%	14.88%
COKE	22.03%	38.35%	19.72%
FB	23.89%	31.49%	10.14%
GOOG	21.55%	27.15%	5.63%
MSFT	29.51%	27.35%	14.74%
NFLX	41.90%	41.86%	0.69%
TSLA	36.26%	52.91%	1.19%
WMT	11.32%	22.45%	23.13%

	Est. Total Return	Est. Portfolio Risk	Total Weight
Overall Portfolio	22.38%	20.00%	100.00%
Threshold	0.00%	20.00%	100.00%
Sub Score	22.38%	20.00%	0.00%
		Score:	1.09

Optimal Allocation

EA will search for the optimal max (Target)

ss1: turns negative if negative return
ss2: turns negative * larger no.
if risk higher than threshold set by user (pseudo soft constraint)
ss3: value = 0 else turns negative * larger no. (pseudo hard constraint)
T: risk free (i.e. 10 year treasury, 0.64%)

Formula:

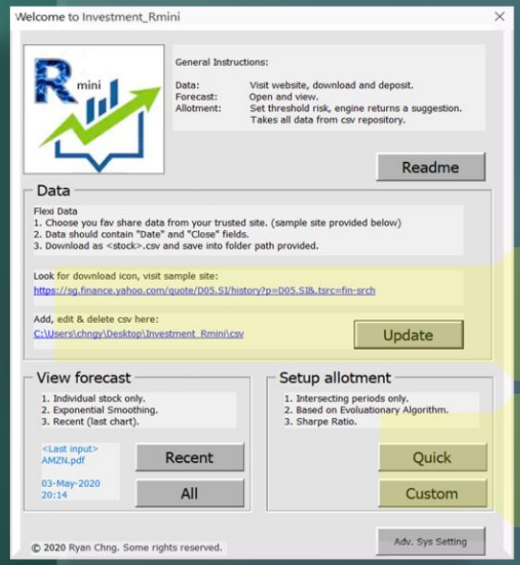
Target = (Est.Total Return – T) / Est. Portfolio Risk + SUM(ss1 + ss2 + ss3)/ Large no.

Sharpe Ratio

~0 if no violation

Feature 2: Allot

Based on Sharpe Ratio, the solver (Evolutionary Algorithm, EA) look for the maximum. This is to get the pair highest return over the lowest risk happen concurrently together.



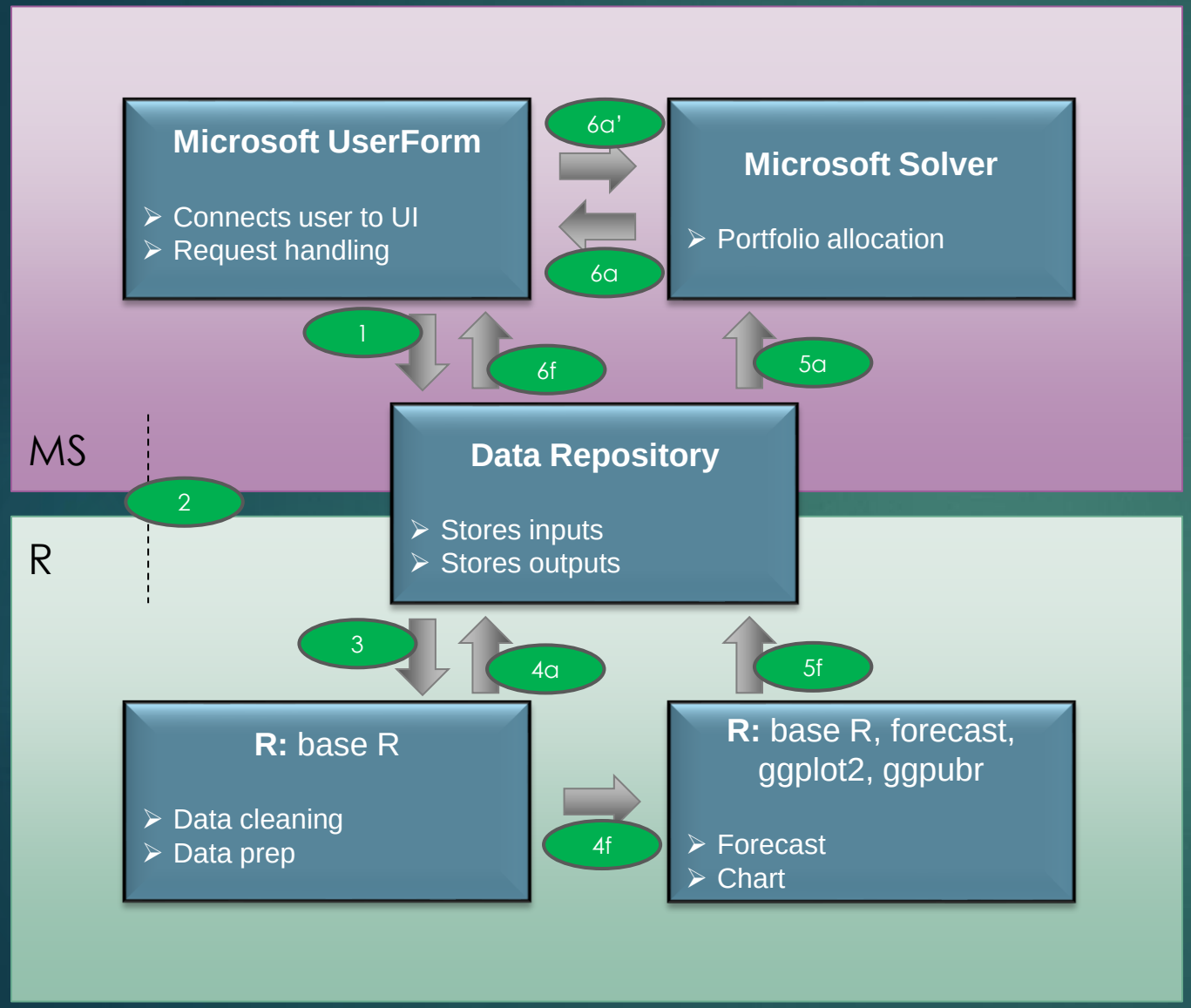
1. Download
2. Update

3. Set
4. View

With 4 Steps:
Download / Update / Set / View

Any violations will bring the ratio negative due to the second piece in place.

Project Implementation (Part 1)



Input

1: Preferred data gets downloaded.
2: User select from UI; R performs data prep.
3: R draws from data repository.

Forecast

4f: Cleaned data passed on.
Training set and test set to be formed.
R requires package forecast.
Best ETS, HW & point estimates calculated.
R requires package ggplot2; ggpubr.
5f: Forecast tables & charts created.
Tables & charts are push into repository.
6f: User calls preferred visualization.

Allot

4a: Cleaned data passed on.
5a: Solver draws required clean sets.
6a': User can set preferred threshold risk.
6a: **EA** applied; optimal allotments return.

Project Implementation (Part 2)

User

Source

FieldName(Exact)	Format
Date	dd/mm/yyyy
Close	double

getData.R

Objects in List	Each object contains
<stock 1>	Date (date format) Close (double)
<stock 2>	Date (date format) Close (double)
…	…

App UI

Call	Type	Content
Link	link	direct to website / folder
Update	button	pulls data in
Recent / All	buttons	open up forecast
Quick / Custom	buttons	activate solver

allot.R

Object	Type	Content
historical	data.frame	all historical values
return	data.frame	daily return
return_risk	data.frame	annual return & risk
covSet	data.frame	covariance table

forecast.R

Object	Type	Content
g1	ggplot	Main line chart
g2	ggplot	boxplot
Quantile	ggtexttable	Quantile values
aScore	ggtexttable	Table of MAE & RMSE
pastTbl	ggtexttable	X-days averages

Project Performance & Validation

Feature	Methods	Comments
Feature 1: Forecast	Exponential Smoothing ('ets()' function from R selects the best from its variants)	Pros: Relative lower MAE, RMSE compared to other methods. Ideal for stock forecast, given only closing price. Cons: Best model may not show trend.
	Subset of ETS, Holt Winters (additive)	Pros: Good to display trend and seasonality. Cons: Slightly larger MAE, RMSE compared to above.
	Linear model	Good bench mark in any case.
	Moving averages, mean, median	Good information and point estimates references
Feature 2: Allot	Evolutionary Algorithm (from solver)	Pros: Search for solutions for non-formulated problem. Cons: Converges to local optimal, different and repeated runs may yield better results.

Project Conclusions: Findings & Recommendation

Techniques, algorithms and models do solve and explain many problems. There is 'no one mould fit all' solution for most of the complicated problems encountered. Each sub-section of the problem is usually tackled with a different strategy or method.

In this case, the project employed provide data discovery to it's user to forecast the stocks; make use of search algorithm to tackle complex portfolio optimization. Both methods save the user time and give the user insights to an optimal portfolio.

In both cases, the methods are not flawless. In forecast feature, the real outcome may deviate from the forecast trend drastically. As the forecast only takes in closing values, parameters beyond the scope plays no impact to the forecast. Many more extensive built can be created, i.e. adding daily volumes or even stock news which the elements that can drive stock price can be siphoned by some form of sentiment analysis. Similarly, the stopping criteria of EA can be set more relax to run more iterations for a better solution or a hybrid solution together with other search techniques can provide more accuracy. These are not covered in the scope of this project as it has already quenched the thirst on the burning questions.

The project has demonstrated that from the use of just tables of data, insights and knowledge can be simplified and made easily understood. Not only to induce knowledge gain, the search for unknown creates some valid optimal answers. **The project has met its primary objectives** and is open to more enhancements that relates to the topic.

-- End of Report --