Project Report

Investment R_{mini}

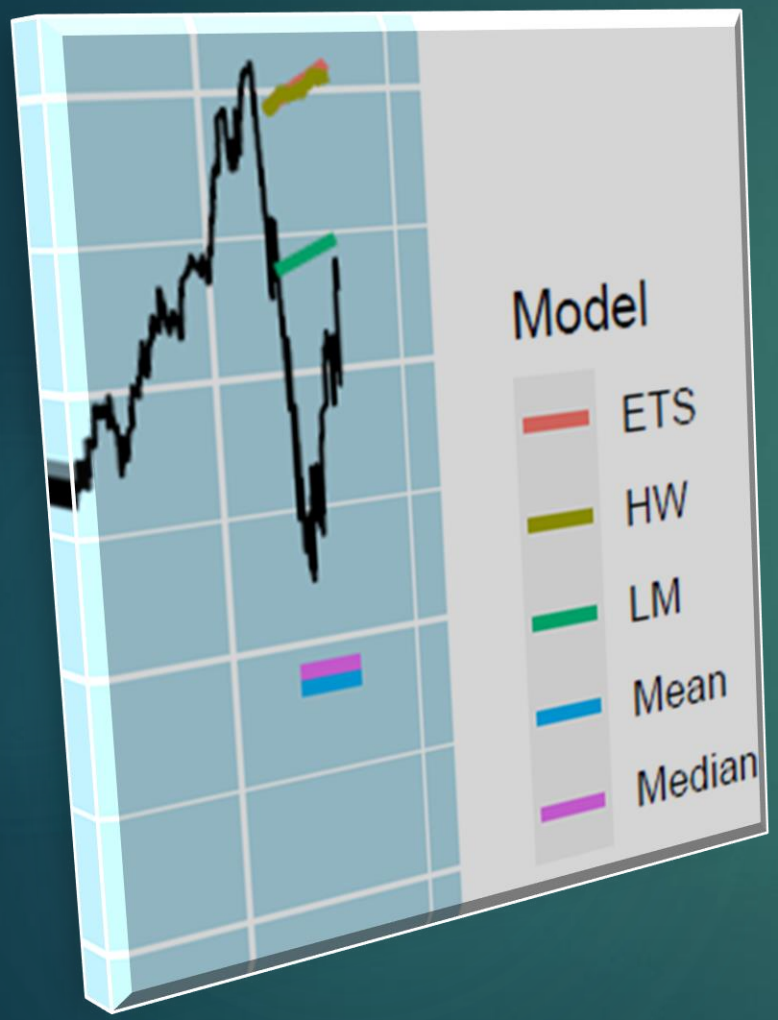
Stock Forecast & Portfolio Recommender

Group 3

Chng Yan Hao

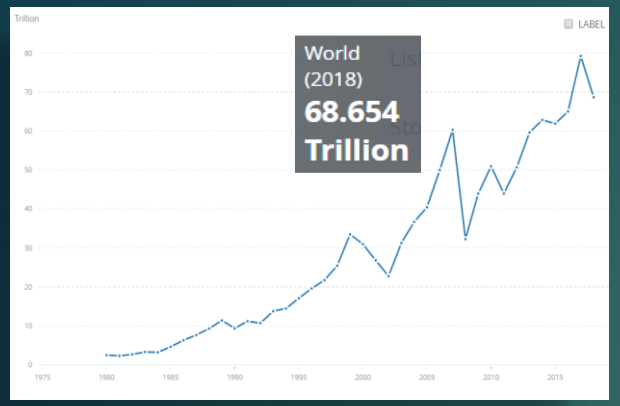
A0024023A

Contents

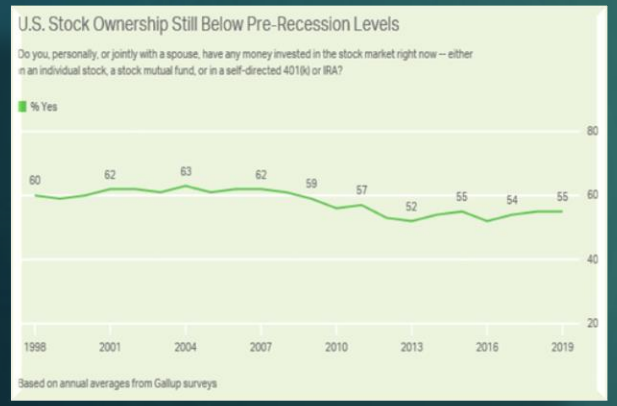


- ✓ Executive Summary
- ✓ Business Problem Background
- ✓ Project Objectives
- ✓ Project Solution
- ✓ Solution Design
- ✓ Project Implementation
- ✓ Project Performance & Validation
- ✓ Project Conclusions: Findings & Recommendation

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 beginning, just the number of stock markets can overwhelm anyone (NYSE, Nasdaq, JPX, LSE, SSE, SEHK, … SGX). The stock products and terms are no easier (bonds, options, shares, long, short, put, call, dividend, margin…). Also, forex exchanges further complicates the whole picture. Potential investors faced high entry barrier.

The primary aim of the project is to help its users to explore more intensively at virtually zero cost. It also offers a second opinion to those whom are already invested. By taking mini step, Investment R_{mini} may just become a 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 give a forecast and a portfolio recommendation.

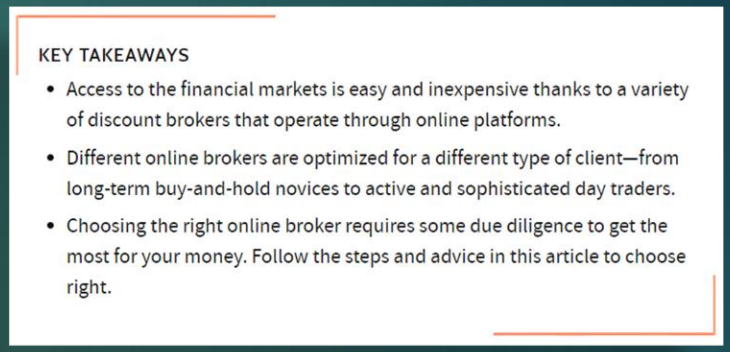
Business Problem Background

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 feeds on any stock pre-selection suitable for DIY users.

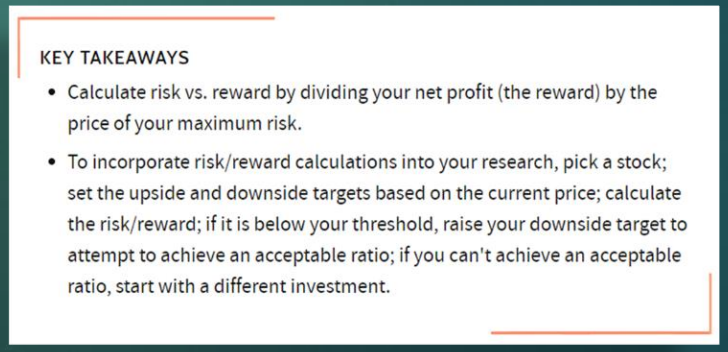
In general, investors have to know your needs and narrow your field.[4]



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

The project has few 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 compare and contrast. i.e. against investment platforms, friends, brokers and even robo-advisors.

4) Demonstrate data discovery using open source tools such as R.
5) Demonstrate a day-on-day tool can do wonders with an add-in activated (Microsoft Excel which is used by millions in the world): Solver -- evolutionary algorithm.
6) Simple soft and hard constraint settings.

Project Solution (Part 1)

Feature 1: Forecast

The most common method for forecasting in stock market is moving average. The main solution chosen in this project is to use Exponential Smoothing[6]. Moving average has constant weights whereas exponential smoothing has exponentially larger weight when nearer.[6]

While some dataset may have obvious seasonality, others may only have obvious trends. Different models attempt to capture 'error', 'trend' & 'seasonality' of a given dataset. ETS() is an in-built exponential smoothing 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.

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

As events and circumstances do not occur as expected, the forecast functionality is mainly to serve more of a second opinion and not 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 a state of art. Here, we are using a simple and well expressed term known as Sharpe Ratio[7]. It can be view loosely 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. The chosen algorithm will be Evolutionary[8] option in the solver. It was chosen for the following:

(i) Problem resembles a combinatorial problem
(ii) No easy formula to calculate in short time
(iii) No one really knows the solution for the dataset

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

where:

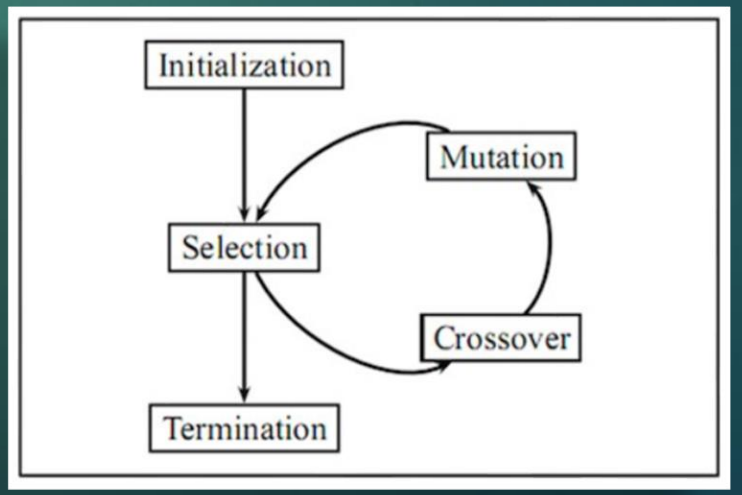
x = The investment

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)

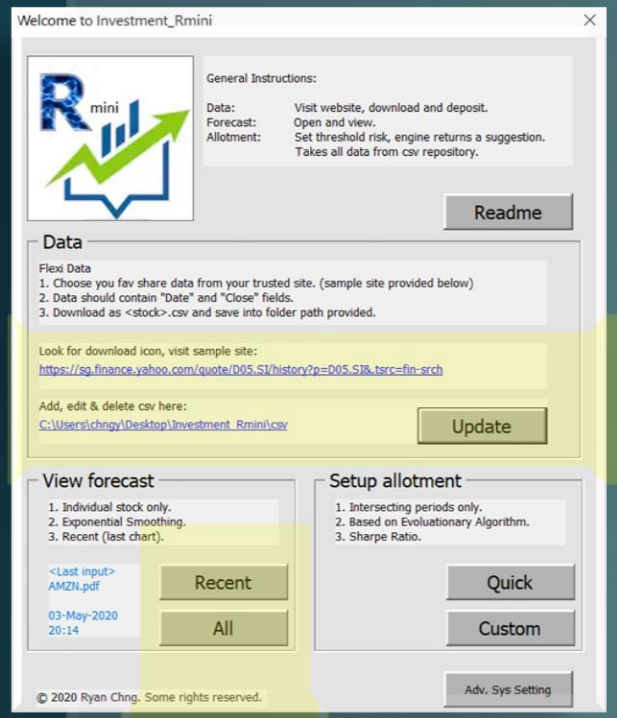
$StdDev(x)$ = The standard deviation of r_x

[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 for small errors in the current dataset, future events are not part of the analysis.

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:

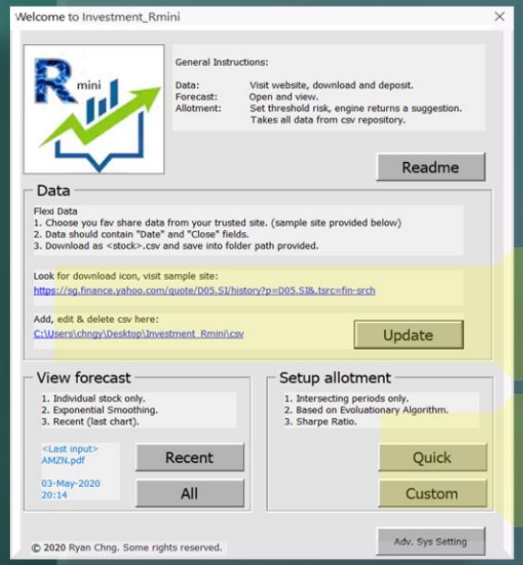
Target = (Est.Total Return – T) / Est. Portfolio Risk + SUM(ss1 + ss2 + ss3)/ Large no.

Sharpe Ratio

~=0 if no violation

Feature 2: Allot

Based on Sharpe Ratio, the solver (Evolutionary Algorithm, EA) look for the maximum. At the end of the search, user can based on the optimal allocation to allot their portfolio.

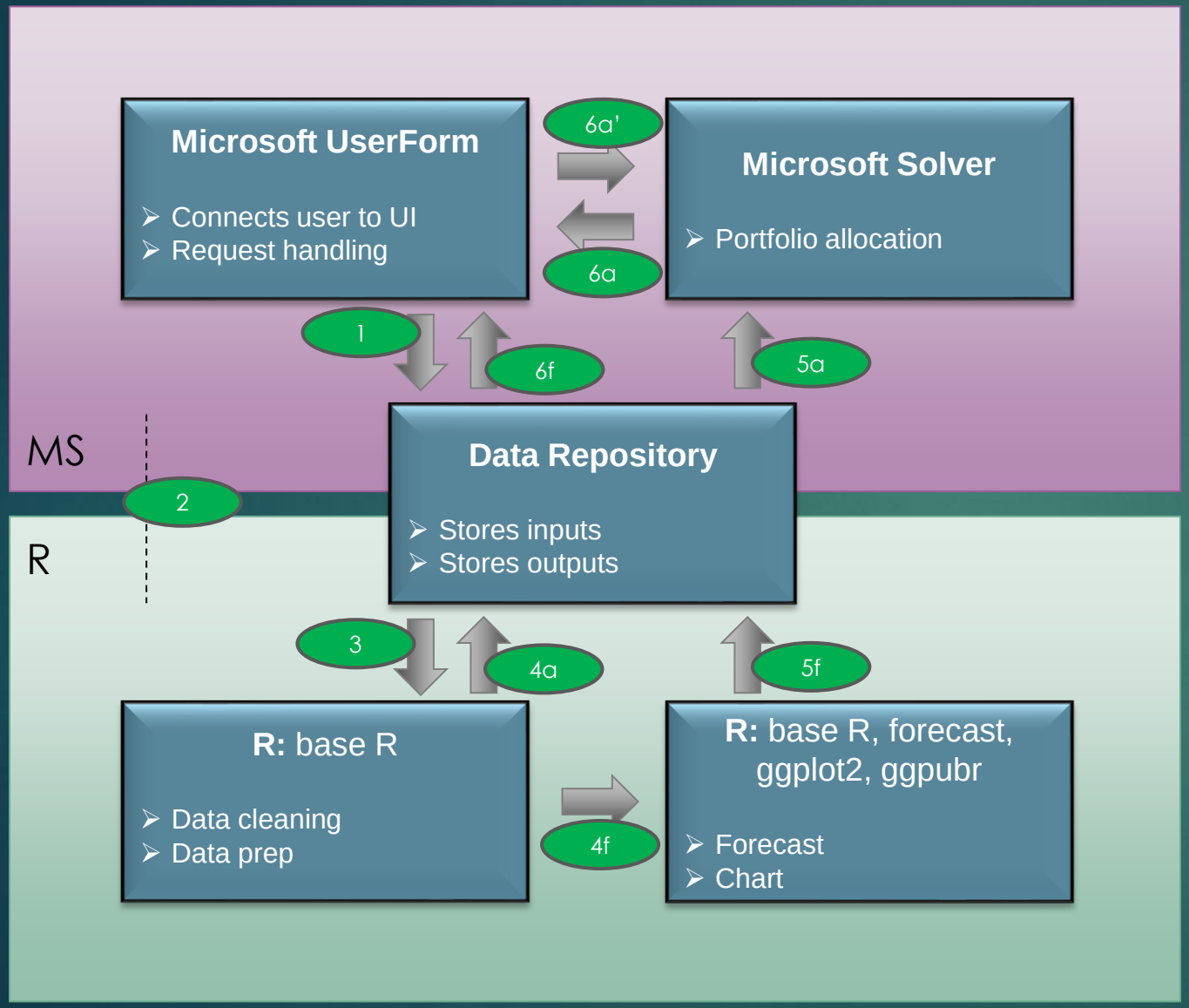


1. Download
2. Update
3. Set
4. View

With 4 Steps:
Download / Update / Set / View

Quick: no threshold set to risk
Custom: User sets a threshold

Project Implementation (Part 1)



Input

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

Forecast

4f: Cleaned data passed on
R requires package forecast()
ETS(best chosen), HW & others applied
R requires package ggplot2(); ggpubr
Forecast charts created
5f: R created charts and push into repository
6f: User calls preferred charts

Allot

4a: Cleaned data passed on
5a: Solver draws required clean sets
6a': User can set preferred parameters
6a: Evolutionary algorithm applied
Optimal allotments returned

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() 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(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.
Feature 2: Allot	Evolutionary Algorithm (from solver)	Pros: Search for solutions for non-linear problem. Cons: Converges to local optimal, different runs may yield slightly different 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 complicated problems. Each sub-section of the problem is usually tackled with a different strategy or method.

In this case, the project here employed a data discovery method to assist it's user to forecast the stocks and make use of search algorithm to tackle the non-linear portfolio optimization. Both methods save the user time and narrow the work to be done to create his portfolio.

In both cases, the methods are not flawless. In forecast feature, the real outcome may deviate from the forecast trend. As the forecast only takes it closing value, parameters beyond the scope plays no impact to the forecast. A even more extensive built can be created, i.e. adding daily volumes or even news which the elements to drive the price can be extracted from some form of sentiment. Similarly to the evolutionary search, the stopping criteria can be set more relax to run more iterations for a better solution however this compromises on the time taken. These are not covered in the scope of this project.

It has demonstrated from the use of just tables of data, insights and knowledge can be simplified and made easily understood. It has reduce great amount of time in many computations and even assist the user to search for an optimal stock portfolio. The project has met its primary objectives and is open to enhancements that relates to the topic.

-- End of Report --