**S&P 500 Simulation**

The S&P 500 index is one of the indices that some investment companies offer as a fund, dubbed as "index fund", that you could invest your money in, these funds is bench-marked to the S&P 500 by holding the same share as the index. This index is traded daily thus hold a very large of historical data that goes back to the 1950s.

Now What If? You started to put money on these index funds and continually contribute on a monthly basis. Based on the historical prices of the Index up to the time of upload, we have generated the Viz below to reflect on the potential of profit, as well as show the behaviour of the index for the past 50 plus years. Imagine contributing money since the 80's or 60's or the 50's, how much your wealth would have grown.

<iframe title="S&P500" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiZDljNGM2OGUtMTdjZS00ZmZjLTgyMDctYmQwMDlhNjA2MDE4IiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

We downloaded the SP500 data (^GSPC) from Yahoo Finance from Jan 1950 until Apr 2025. The dataset comprises Date and Close price. The analysis is then generated with Power BI.

Analysis

The analysis helps the investor to simulate the return if they started to contribute fund at a historical date. User can compare the total Return on Investment (ROI) as of the recent date by choosing either one time contribution or monthly occurrence.

Many investment agents like this visual very much as it helps them to demo and compare visually the regular bank saving with unit trust or insurance. We received hot response regularly to update the data and features enhancement request to perfect it as a fund selling tool.

Summary

Looking at SP500 data alone shows you only the up and down movement. You can always make use of the data and relate it to your investment context. If you are the one who plan to invest for the next 30, 40 or 50 years, this simulation paints you a likely picture of the wealth during your retirement.

Top-150 World QS Ranking Universities 2025

Famous universities in the world can assure the quality of education and guarantee the good employment after graduation. The people involve in education line, be it students, employees, colleges, parents follow the report and data to make a wisely decision. QS World Universities Ranking is one of the most sought-after reports. The ranking is updated every year to the public for reference.

<iframe title="QS Ranking2025" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiNWE1MTk3NzktYTFhOS00Yjk0LTg2OWYtOGI1YmM2N2Y4MTMyIiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

Yearly QS Ranking dataset was extracted from internet and combined. The analysis make reference to 3 years data from 2013 to 2015. The data file contains columns such as Institution Name, Ranking, Size, Region, Location (Country), Total Score, Employment Outcomes Score, Academic Reputation Scores and etc.

Analysis

The analysis is mostly the descriptive type. It allows users to view the world class Top-150 university by region or world. One is able to compare the number of university and the university name by country. To understand the underlying scores, especially the popular academic and employment rating in a separate scatter chart. By selecting more than one country in bar chart, you can drill-through to scatter chart and make a comparison on the top-150 universities from the respective countries.

A second visual in tabular format provides the ranking variations of the related university in these 3 years. The list is not limited to the top-150, it is extended to the institutions with ranking beyond 1000.

Summary

Does your target university improve in ranking for the last 3 years?

How are the total scores, employment score or academic reputation compare to the neighbouring country?

By visually, you can get the answer instantly from the analysis.

Data Science Salary (USD$) 2020-2025

Do you know what level of salary for roles In Data Science job market?

This report performs analysis based on job roles mostly from US market from 2020 to 2025. It allows us to understand better from the trending, company size, job experience and employment types.

<iframe title="DS Salary" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiN2U5ODU0ZDUtY2Y5Zi00NTg0LTg4NDctMWE5NjdhZWM1MTUzIiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

Data is downloaded from Kaggle.com. It consists of Year, Job Tile, Salary in USD, experience level (Entry, Mid-level, Senior, Experience), employment type (Full-time, Part-time, Freelance), company size (L, M, S)

Analysis

In overview, we look at the salary trends. It is interesting to see that salary in small company is increasing fast in the past 2 years compared to company with larger size. There is no difference with other profession, the experience and full-time roles earn the higher income.

Highly paid jobs are mostly from senior and experience roles. Whereas, the popular Data Science roles are Data Scientist, Data Engineer and etc, all of them fall into the entry level category.

When we look at the detail jobs breakdown by years, we can see that a lot of new roles emerge in the recent 2 years. The matrix table allows one to drill-down to the experience level. You can also check where is your data science role in the job market.

Summary

Data Science is a very dynamic market, it can tell from the data more new roles are emerged in the recent years. The experience level role can earn more than management role. It is interesting to see that salary in the smaller company keep improving. Does it have any relation with AI revolution? We should continue the analysis with the appropriate data in the coming series…

Covid Cases Estimation

Ministry of Health had launched a new web portal with ton of informative COVID-19 data during MCO period. This information was very details and nicely presented.

However, the information may be too much for a normal audience who is interested to know what is the daily number. Another issue arises when people is flooded with more and more data. When one set of data is shown and supported by another set of detail breakdown, people mind can be unconsciously led to think in the scope of data presented. For example, one will treat the daily reported COVID positive cases is the real situation and the community also uses the same figure in discussion.

The visualization that is presented here focuses on only a single number; the daily positive COVID cases. The intension is to show the more realistic situation so that you can take more precaution.

<iframe title="Covid-MY.adlab" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiM2Q4NGVmYmEtYTg3Mi00NWJlLThjNTYtMTRhNTVhZWQzMGRjIiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

Ministry of Health (MoH) provided a repository of daily COVID data. It was updated daily for the public consumption. Data comprises Date, State, Tested Cases, Positive Cases and etc.

Analysis

Daily COVID positive cases are the statistics based on the carried out COVID tests. However, there is a large number of hidden cases which are not included in the daily reported number. What are the actual positive cases if we increase the testing rate per population on the suspected group?

From the number reported, we know that daily COVID testing rate is only around 0.5% population and still a large number of suspected groups is not tested. What if we increase the population test to a reasonable number like 1% or 1.25% with the same daily positive rate as assumption? The calculated number will tell you what is the possible daily COVID positive number in the community.

Summary

Even though MoH data is no longer updated, this report illustrated the typical Predictive analysis conducted. One can use its own formula or any machine learning algorithm to estimate the outcome using simulation method.

Overview of Malaysia Public Health Sector

Have you ever thought about the distribution of all government clinic and hospital or private clinic and hospital across the states in Malaysia?

Were them distributed proportionately by population?

What are the numbers of government clinic and private clinic in each state?

<div class='tableauPlaceholder' id='viz1746002299168' style='position: relative'><noscript><a href='#'><img alt='Dashboard 1 ' src='https:&#47;&#47;public.tableau.com&#47;static&#47;images&#47;He&#47;HealthCareMalaysiaUpdated&#47;Dashboard1&#47;1\_rss.png' style='border: none' /></a></noscript><object class='tableauViz' style='display:none;'><param name='host\_url' value='https%3A%2F%2Fpublic.tableau.com%2F' /> <param name='embed\_code\_version' value='3' /> <param name='site\_root' value='' /><param name='name' value='HealthCareMalaysiaUpdated&#47;Dashboard1' /><param name='tabs' value='no' /><param name='toolbar' value='no' /><param name='static\_image' value='https:&#47;&#47;public.tableau.com&#47;static&#47;images&#47;He&#47;HealthCareMalaysiaUpdated&#47;Dashboard1&#47;1.png' /> <param name='animate\_transition' value='yes' /><param name='display\_static\_image' value='yes' /><param name='display\_spinner' value='yes' /><param name='display\_overlay' value='yes' /><param name='display\_count' value='yes' /></object></div>

<script type='text/javascript'>

var divElement = document.getElementById('viz1746002299168');

var vizElement = divElement.getElementsByTagName('object')[0];vizElement.style.width='1300px';vizElement.style.height='1500px';

var scriptElement = document.createElement('script');scriptElement.src = 'https://public.tableau.com/javascripts/api/viz\_v1.js';vizElement.parentNode.insertBefore(scriptElement, vizElement);

</script>

Summary of Data

Data is downloaded from internet.

Analysis

This visualization is purely an informative analysis. It described the healthcare landscape and centres distribution across the country.

Purple bubble – government village clinics

Orange bubble – private clinics

The facts discovered from the landscape are:

* The residents in the some less populated states (Perlis, Melaka) or higher density state (Selangor) have much lower village clinics for the size of population.
* They are much higher number of private clinics (Kuala Lumpur, Penang, Selangor) for the population size .

Summary

Health centres distribution may not be solely determined by population, it could be addition factors which needs more data for analysis.

**KLSE Simulation**

This is the part 2 of the previous S&P 500 index simulation. We know that if we invest the fund consistently across 20-30 years, the outcomes can be fruitful. With this, another question comes into my mind, if I invest the same in KLSE index fund, do I get the closer or better return?

For better comparison, I overlap KLSE index with S&P 500 index. The chart will calculate both results within the selected period.

<iframe title="KLSE" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiNWM4YzdhYTMtN2VkNi00ZjYwLWE0NTUtYzFhM2RjOWY5NTlmIiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

KLSE index data (^KLSE) was pulled from Yahoo Finance from Dec 1993 until May 2025. The dataset comprises Date and Close price. Dataset was saved into [Github](https://github.com/jlsiaw/dataset) for anyone who are interested to work on it.

Analysis

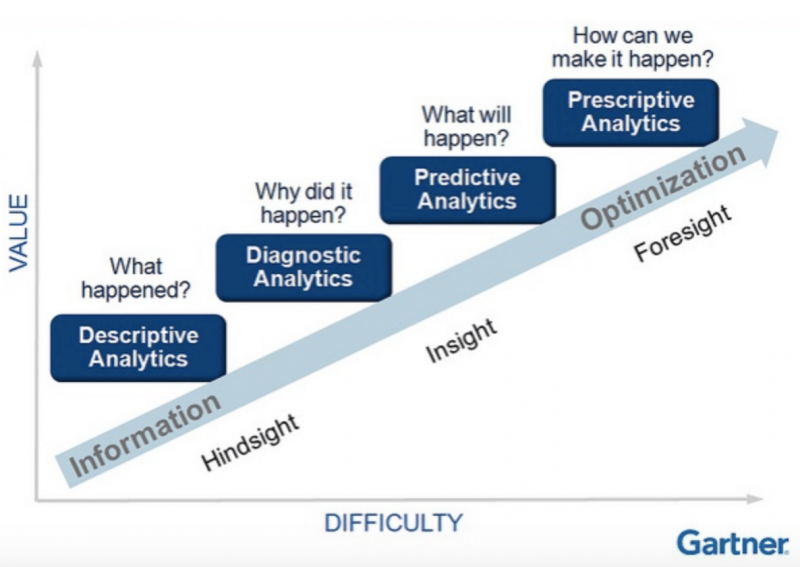
Results show that the gained from SP500 is almost 3 times of KLSE. It was mainly KLSE index did not grow much along these years while SP500 went exponential after 2010.

Summary

It has to note that this analysis is purely based on the historical performance, it should not be a reference for the future growth, a lot of external factors can affect the future movement. Who knows KLSE will follow or even surpass SP500 track due to the extreme volatile world economy situation.

Analytics Types: Descriptive, Diagnostic, Predictive, Prescriptive

In general, there are 4 stages of analysis, results ranging from hindsight to foresight. The easiness level and value created also different in each stage. The simpler analysis takes shorter time but the value can be just nice, however it takes longer time to develop foresight types of analysis, and of course the business impact in term of revenue and changes also very high. Also to note that analytics types are not measured by number of visualizations; descriptive analytics can be many outputs and attractive, whereas prescriptive analytics may have only one page of bar charts to represent result.



Descriptive – periodic report, information tells what happened in the past

Diagnostic – explorative and multiple perspective, answering what caused the trends

Predictive – it can be estimation, machine learning, or simulation, try to get a closer answer of the future outcome

Prescriptive – shed the light for the action to solve the problem. It starts with business motive; discovery drive the direction of action and then it must be supported by the process changes in business. All three aspects have to work hand in hand to see significant business impact.

There are examples in the visualization below, 4 different types of analysis which exhibit the stages as described above.

<iframe title="Analysis Types" width="1024" height="612" src="https://app.powerbi.com/view?r=eyJrIjoiNDI2MjdiZWUtOGZjNy00NTQ4LThiNjYtOGM3MTVhZTM4M2EwIiwidCI6ImFlYzYyYmM2LTMxNDktNGEwZS04MzAyLWJkM2MxYzM2Y2U2NyIsImMiOjEwfQ%3D%3D" frameborder="0" allowFullScreen="true"></iframe>

Summary of Data

A set for processed sales transaction data comprises Date, Category, Sub-Category, Product, Customer Name, Region, Sales, Profit. This dataset has 4 years of sales transactions and data is loaded into Power BI to illustrate analysis types as mentioned above.

Analysis

We want to use the data to tell the story, starting with what had happened in the past, to explore the alerting area, simulate and forecast with the relevant variables and suggest some workable action. In each visual page, mouseover to the information icon for a brief overview.

In Descriptive page, information from charts:

* Trending - year 2021 sales is the highest and 2024 is lower than the previous years
* Profit margin of 2024 is the lowest
* Technology category provides the largest sales value, followed by Furniture and Office Supplies
* Alert for sales revenue and profit margin

In Diagnostic page, the discovery via analysis:

* Issue of low profit products - sub-category 'Table' and 'Bookcases' with high revenue but produced negative margin (at loss), and bring down the overall profit
* Other factors that cause low sales results - 'E.Coast' region generates very low profit margin when comparing to others

In Prediction page, we used simulation and forecasting. Forecasting ignores the latest incomplete month. Mouseover to view forecasting area value.

* We forecast the next 90 days sales with in-built ML algorithm
* Then simulate and forecast the profit margin without involving sub-category Table
* Forecast shows sales does not change much but overall profit margin is improving without selling unprofitable 'Table'

In Prescription page, assuming company allocates large marketing fund to boost the sales after knowing the alerting sales trend and profit margin, we want to identify which action could be more significant? Two options from analysis:

* Target the right products - the sales momentum of Category 'Office Supplies' and 'Technology' are trending down, need more fund for sales campaign compared to ‘Furniture’.
* Customer base - sales was mostly generated from older acquired customers, new customers do not contribute much to sales. Urgently to acquire higher quality customers as action.