INTERNATIONAL PAPER COMPETITION 2022
Department of Research and Technology
Ambarrukmo Tourism Institute (STIPRAM)
Yogyakarta, Indonesia
Jl. Ring Road Timur, Bantul, Yogyakarta 55584
Email: ristek@stipram.ac.id
Phone: +6281 229881414

INTERNATIONAL PAPER COMPETITION 2022

ANALYSIS OF MALAYSIA'S FUTURE TALENTS USING TIME SERIES

Data Science

Mohamad Rai'zuddin Farhan Arshad
Nuryasmin Ab Rashid

Nor Azuana Ramli

Centre for Mathematical Sciences, College of Computing & Applied Sciences,
Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.

raiezarshd@gmail.com (corresponding author)
ysminrshid@gmail.com, azuana@ump.edu.my

DEPARTMENT OF RESEARCH AND TECHNOLOGY
AMBARRUKMO TOURISM INSTITUTE (STIPRAM)
YOGYAKARTA, INDONESIA
2022

Analysis of Malaysia's Future Talents using Time Series

Mohamad Rai'zuddin Farhan Arshad*, Nuryasmin Ab Rashid, and Nor Azuana Ramli

Centre for Mathematical Sciences, College of Computing & Applied Sciences,
Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.

raiezarshd@gmail.com, ysminrshid@gmail.com, azuana@ump.edu.my

*Corresponding Author
Received: day month 202x, Revised: day month 202x, Accepted: day month 202x
Published online: day month 202x

Abstract: The unemployment rate has been increased after the pandemic and it also effects the youth unemployment rate. This research aims to investigate the issues relating to youth talents and employment. In achieving this aim, the main objective is set to forecast the future trend of Malaysia's labour force and employed rate in the upcoming 10 years by using time series analysis. Data of labour force are collected through Department of Statistics Malaysia Official Portal. After data cleaning and pre-processing, the data will be analysed by using Double Moving Average and Holt's Method. The model will be evaluated using R-squared, root mean squared error and mean absolute percentage error. Based on the analysis, the Holt's Method is the best for the number of employed persons by age group (25 to 59 years old) as the model achieves better in each evaluation metric. To conclude, the demand for jobs also increases as the population increases. If this is not manageable, it can cause a higher rate of youth unemployment and the Twelfth Malaysia Plan to develop future talents might not be successful.

Keywords: Unemployment; Youth; Time Series; 12th Malaysia Plan; Holt's Method

1. Introduction

The Twelfth Malaysia Plan, 2021-2025 (henceforth referred to as RMK-12) is a medium-term plan to achieve "A Prosperous, Inclusive, Sustainable Malaysia" and marks the beginning of a new phase in Malaysia's development path – the Shared Prosperity Vision 2030 (WKB 2030). To become a high-income nation, RMK-12 offers a new transformative strategy based on three themes, four catalytic policy enablers, and fourteen game-changers. Focus on creating future talent is one of RMK-12's four primary catalytic policy enablers. The growth objectives of RMK-12 will necessitate the cultivation of highly trained individuals to meet the needs of an expanding economy.

However, the government and policy facilitators must eliminate poverty to focus on fostering future potential. According to United Nations, unemployment and underemployment are the defining characteristics of poverty [1]. The unemployment rate in Malaysia is low before the pandemic. The COVID-19 pandemic has caused an increase in the unemployment rate and altered the way individuals work, and businesses function. According to [2], some private-sector workers have lost their jobs, while nearly half of self-employed respondents have lost their jobs and have less than one month's savings. Currently, it is difficult for fresh graduates to find employment, as most employers prefer to hire experienced personnel. In 2019, the youth unemployment rate in Malaysia was 10.5%, which is six times the adult average of 1.7%. The youth unemployment ratio to the national average has increased during the past decade [3]. The rising unemployment rate in Malaysia will have a detrimental effect on the country's economy.

To understand more about the future of Malaysia's talents, three questions will be discussed in this paper. The research questions are as follows:

a. What is the future trend of Malaysia's labour force and employment rate in the upcoming ten years?
b. What approach can be used to analyse which state contributes the most significant labour force?
c. How to evaluate and prepare youth to be future-ready talent based on six types of assessments?

To answer all the questions raised above, this research is set to achieve all the objectives which are:

a. To forecast the future trend of Malaysia's labour force and employed rate in the upcoming ten years by using time series analysis.

b. To analyse which state that contribute the most significant labour force by using geographical information system.
c. To suggest a decision-making process on deciding the evaluation process based on six types of assessments

In order to achieve all the objectives, this research is conducted by phases. All the phases will be further explained in Section 2. Then, the results obtained will be discussed in Section 3 before this technical paper is ended with the conclusion and some recommendations for future study in Section 4.

2. Material & Methodology

2.1. Data

This research used secondary data, where the data were acquired from a variety of sources and agencies. The data were downloaded from the Department Of Statistics Malaysia Official Portal (https://www.dosm.gov.my/) and the data required are labour force by age group, employed persons by age group, labour force by state, and Malaysian map (GADM). All the datasets consist of the records of the labour force in thousand units, the records of employed persons in thousand units, and the records of the labour force in thousand units for every year starting from 1982 until 2020. All the data were separated by age group and clustered by states. This research will only focus on the age group from 25 to 59 years old.

2.2. Research Framework

This research was conducted by following the research flow as illustrated in Figure 1 below:

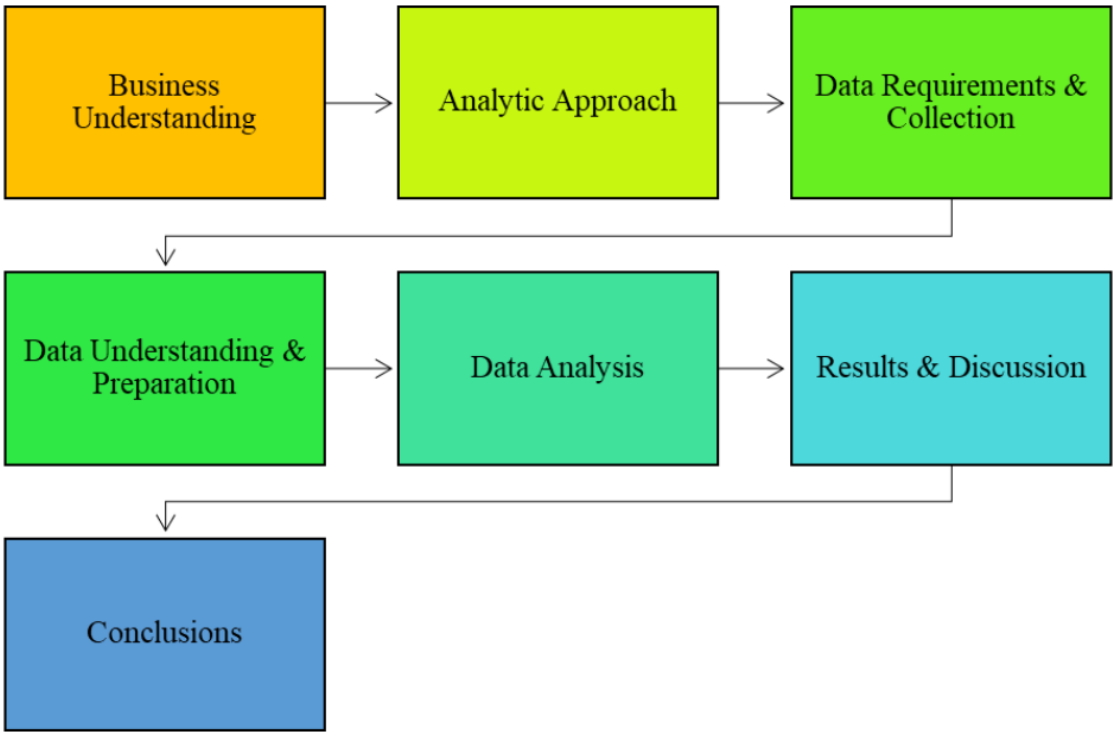


Figure 1. Research framework

a) Business understanding
According to the problem statement of this project, the problem has been defined in terms of Malaysian talents. Thus, each step is generalized in overcoming Malaysian talents problem related.

b) Analytic approach
This research applied three different analytical approach in order to achieve all the objectives.

c) Data requirements and collection
The data required are the labour force by age group, Malaysia/states, 1982 – 2020 and the employed persons by age group, Malaysia/states, 1982 – 2020. The details of the data collection process have been explained in 2.1.

d) Data preparation
The data collected will be prepared and cleansed according to the desires of this research. As the data is secondary, data must be transformed and formatted according to each objective requirement. The details will be explained in Section 3.

e) Data analysis, results, and discussion
The final data will be applied in the detailed step of the analysis, from the method and approach used in this research referring to each objective required. It will also explain the interpretation of each finding appropriate for the research objectives and the data collected. The details will be explained in Section 3.

f) Conclusion
The conclusions are needed to conclude the relations between the data collected, project questions and objectives, and the findings from the analysis. The details will be explained in Section 4.

3. Results and Discussion

3.1. Time Series Analysis

Before any analysis can be done, the line plot for the data is plotted to observe the pattern of the data throughout the time. The plot is shown in Figure 2.

Figure 2. Labour force and employed persons by age group (25-59) line plot from 1982 until 2020

The plot of the data suggests a relatively strong upward trend in the data from the year 1982 to the year 2020 for both data. Both data are considered discrete-time series data where the data are measured at discrete points. The increasing trend is believed to be due to changes in population, technology, demographics, and other preferences. Thus, in order to forecast the value of the future number of the labour force and employed persons, two methods are suggested to be tested and evaluated, which are :

a) Double Moving Average (DMA)
b) Double Exponential Smoothing (Holt's Method)

The Moving Average method is a forecasting method that takes a group of observations and looks for the average value as a forecast for the next period. For the Double Moving Average (DMA) method, the moving average process occurs twice. DMA will be tested in this research, and the steps in forecasting using the DMA method are as follows:

1) Calculate Single Moving Average (*S*')

$$S'_t = \frac{Y_t + Y_{t-1} + \cdots + Y_{t-n+1}}{n}$$

2) Calculate Double Moving Average (*S*")

$$S''_t = \frac{S'_t + S'_{t-1} + \cdots + S'_{t-n+1}}{n}$$

3) Determine the amount of constant value

$$E_t = 2S'_t - S"_t$$

4) Determine the amount of trend

$$Tr_t = \frac{2}{n-1}(S'_t - S"_t)$$

5) Determine the forecast value

$$\hat{Y}_{t+k} = E_t + kTr_t$$

The Exponential Smoothing method is a forecasting method that makes continuous improvements to forecasting the latest observation objects. In other words, the latest observation will be given a higher priority for forecasting than the more extended observation. The DES method is used when data shows a trend. A trend is a smoothed estimate of the average growth at the end of each period. DES has two methods Brown Method and DES Holt. Double Exponential Smoothing (Holt's Method) will be tested for this research.

For the historical data of the labour force by age group (25 to 59 years old), the plot is as shown in Figure 2. The plot suggests a relatively strong upward trend in the data from the year 1982 to the year 2020 for both data. The data also does not vary about a fixed level, exhibits an overall slow upward trend, and the variances increase as the series increases, suggesting that the data are stationary in variance but non-stationary in mean.

After all the data had been plotted, now the analysis can be continued with DMA and Holt's Method. All analyses were done using Microsoft Excel as shown in Appendix. Both models were evaluated by using R-squared, Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), and the results are as tabulated in Table 1.

Table 1. Comparison between DMA and Holt's Method

	DMA	DES	Condition Choosing
R-Squared	0.971	**0.980**	Higher
MSE	37221.934	**28212.909**	Lower
RMSE	192.930	**167.967**	Lower
MAPE	0.365	**0.030**	Lower

Based on the table, the Holt's Method is the best for the number of employed persons by age group (25 to 59 years old) as the model achieves better in each evaluation metric. The higher the R-squared indicates that the model fits the actual data better. The model also has lower value of errors evaluations (MSE, RMSE and MAPE) even though the model is compared using the line chart, both forecasting model shows almost perfect forecasting as the lines almost match and fit with the actual data as shown in Figure 3.

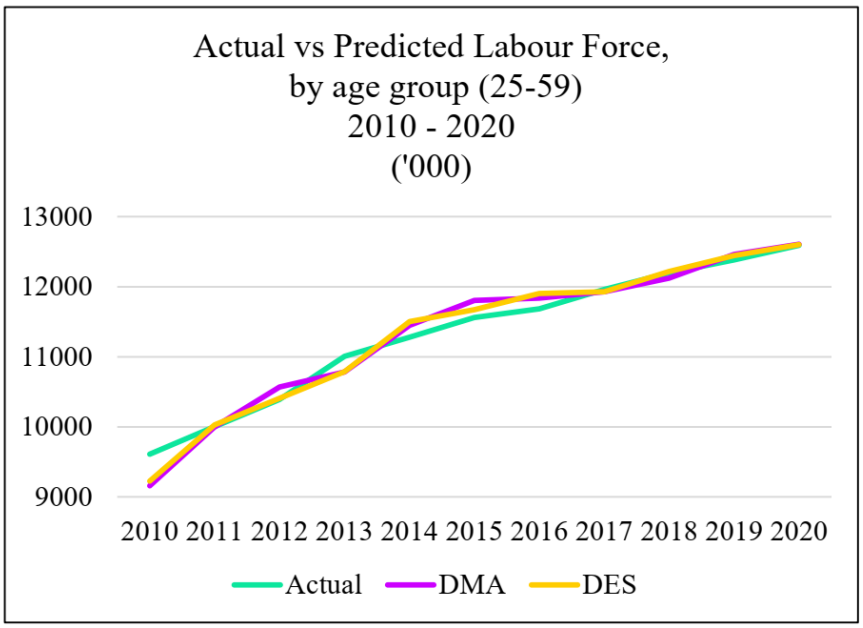


Figure 3. Comparison between actual and predicted labour force by age group (25-59) from 2010-2020

Similar analysis was conducted on the employed persons by age group dataset. The line plot is illustrated in Figure 2. The plot of the historical data of the number of employed persons by age group (25 to 59 years old) suggests a relatively strong upward trend in the data from the year 1982 to the year 2020 for both data. The data also does not vary about a fixed level, exhibits an overall slow upward trend, and the variances increase as the series increases, suggesting that the data are stationary in variance but non-stationary in mean.

Then, the analysis continued with DMA and Holt's Method as shown in Appendix. The performance of the model based on this dataset show that the Holt's Method is the best model as the model achieves better in each evaluation metric similar to previous dataset. The model forecasting for both data will be using Double Exponential Smoothing (Holt's Method) with the optimal value of alpha (α) and beta (β). Both forecasting showed acceptable values as the number of labour force should always be higher than employed person as tabulated in Table 2. The results in a plot graph can be seen as in Figure 4.

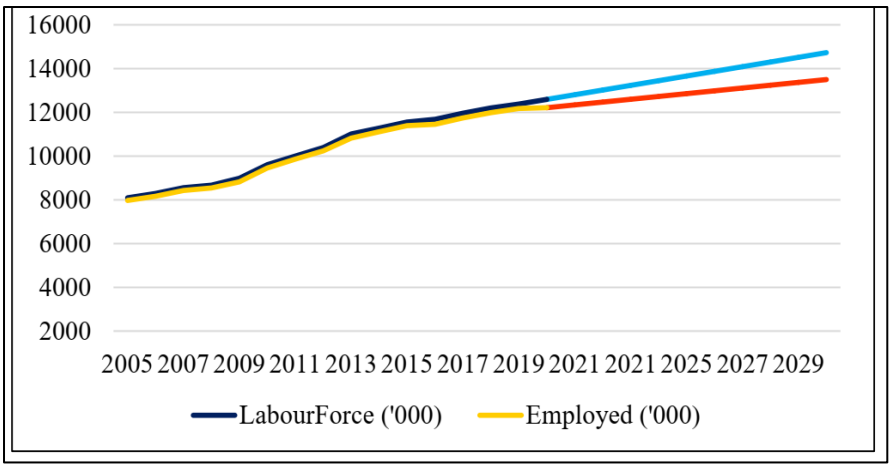


Figure 4. Labour force and employed persons by age group (25-59) forecasting from 2021-2029

Table 2. Optimal value of alpha (α) and beta (β) with results of the forecasting

	alpha (α)	beta (β).
Labour Force	1	0.49
Employed	1	0.47

Year	Labour Force Forecasting	Employed Person Forecasting
2021	12804.5	12339.1
2022	13018.3	12467.8
2021	13232.0	12596.6
2024	13445.7	12725.4
2025	13659.5	12854.1
2026	13873.2	12982.9
2027	14087.0	13111.7
2028	14300.7	13240.4
2029	14514.4	13369.2
2030	14728.2	13497.9

3.2. Geographical Information System

By using ArcMap and Excel, we can see which state has the highest labour force through the heat map. The results from Figure 5 shows that the highest labour force is in Selangor with 35027.20 while the lowest is in Perlis with only 1093.63.

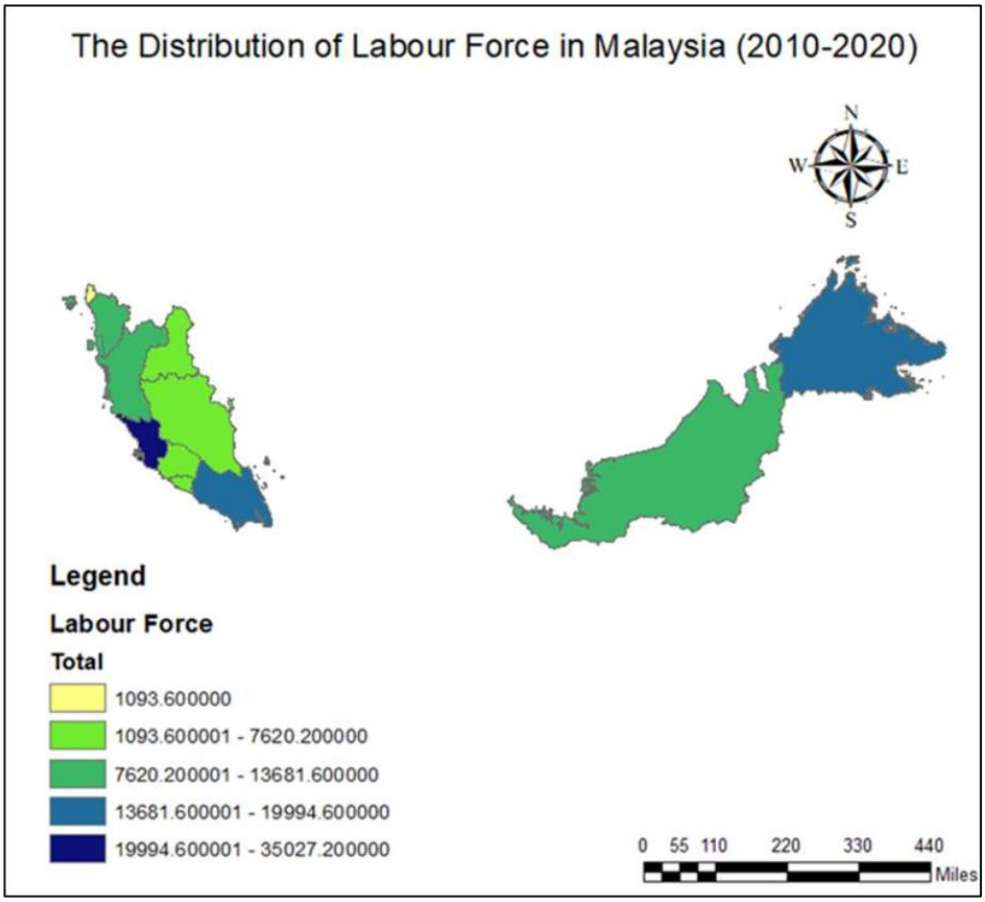


Figure 5. The distribution of labour force in Malaysia from 2010 until 2020 using GIS.

3.3. Determination of Talent Assessment Importance

The last part of this research is focusing on the decision analysis where Analytic Hierarchy Process (AHP) was applied to obtain the results. The AHP speciality is it is one of the Multi-Attribute Decision Making. In this research, we want to identify which graduates are the best to be future-ready talent based on the six criteria which are work values, work interest, personality and motivation, employability, the future of work, and English proficiency. The hierarchy structure model is illustrated as in Figure 6.

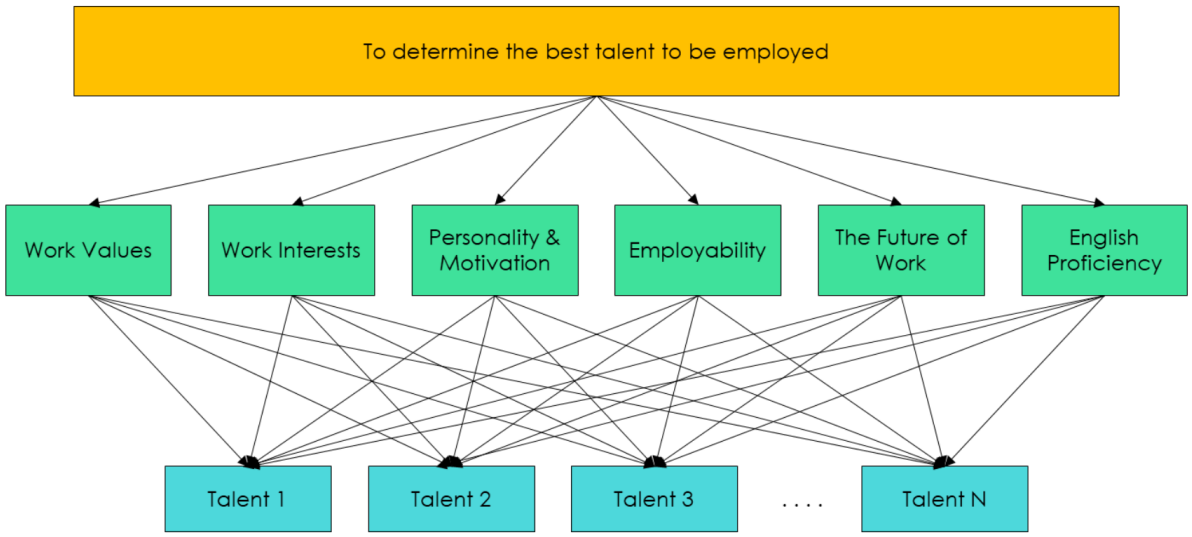


Figure 6. The hierarchy structure model

For weightage calculations, the AHP Excel Template is used to calculate automatically. The criteria are assigned its keyword accordingly. In Excel, the scale is entered based on the Pairwise comparison scale and assumptions on its importance. In the table, the values represent the scale of the importance of each criterion. If the value is a single digit, the criteria are more important; if it is in a fraction, the other is more important. Based on the analysis, it is suggested that "The Future of Work" is the most crucial criterion for future-ready talent, given that it weighs 27%. The value explained that to determine whether the graduate's aspirations for their future employment are realistic. The second essential aspect is "The Work Values." The recruiter examines the graduate's perspective on work-related characteristics and standards. By monitoring specific metrics, recruiters can determine whether a graduate meets a high level. As "The Employability" and "The Work Interest" are virtually equally weighted, recruiters should analyze how graduates' employability level corresponds with their work interests. If they have deficiencies in their desired sector, they can be trained to become more employable and valued. Finally, "The English Proficiency" and "The Personality & Motivation" have not always been met in the opinion of recruiters. All the results are summarized as in Figure 7.

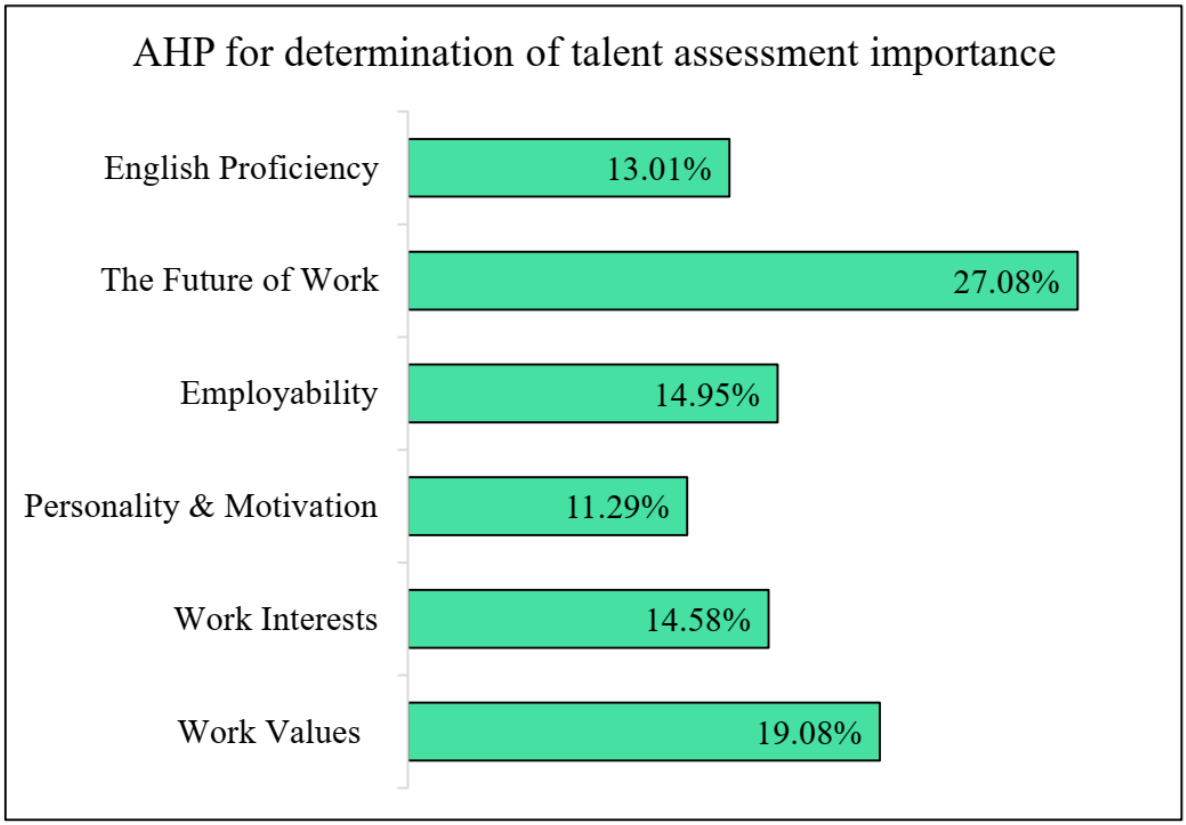


Figure 7. Determination of talent assessment importance by using AHP

4. Conclusion

This research focuses on the strategic priorities that will help transform Malaysia in developing future talents. Following the Twelfth Malaysia Plan (RMK-12), this research provided insightful findings regarding the problems related to the nation's talents. First, this research succeeded in forecasting the future trend of Malaysia's labour force and employed rate in the upcoming ten years by using Double Exponential Smoothing (Holt's Method), which the model evaluated resulted in the lowest MSE, RMSE and MAE. Second, this research determined that Selangor has the highest labour force while Perlis has the lowest in Perlis by using ArcGIS. Lastly, this research ables to develop and propose the decision-making process on deciding the evaluation process based on six types of assessments by using AHP.

For future recommendation, this project can be employed as a dashboard for the talent organisations. The design draft for dashboard demo as in Figure 8 below. The dashboard will have the features such as :

1) Filters for states and year
2) Total for Labour Force and Employed Person
3) User can filter the total, and the total will change according to “Year” and “States”
4) Visualizations for time series data and heat map

In conclusion, the demand for jobs also increases as the population increases. If this is not manageable, it can cause a higher rate of youth unemployment and this RMK-12 to developing future talents might not be successful. The youth feels less motivated and insecure to find any jobs. Thus, these findings can give an idea to the government and policy changes to open more job opportunities, especially in the state with the lowest labour force. As the number projections of the labour force and employed increase yearly, choosing suitable graduates or talents can be challenging. Thus, the decision-making process can help them in providing a systematic way to evaluate and choose the right talents that fit the company's needs.

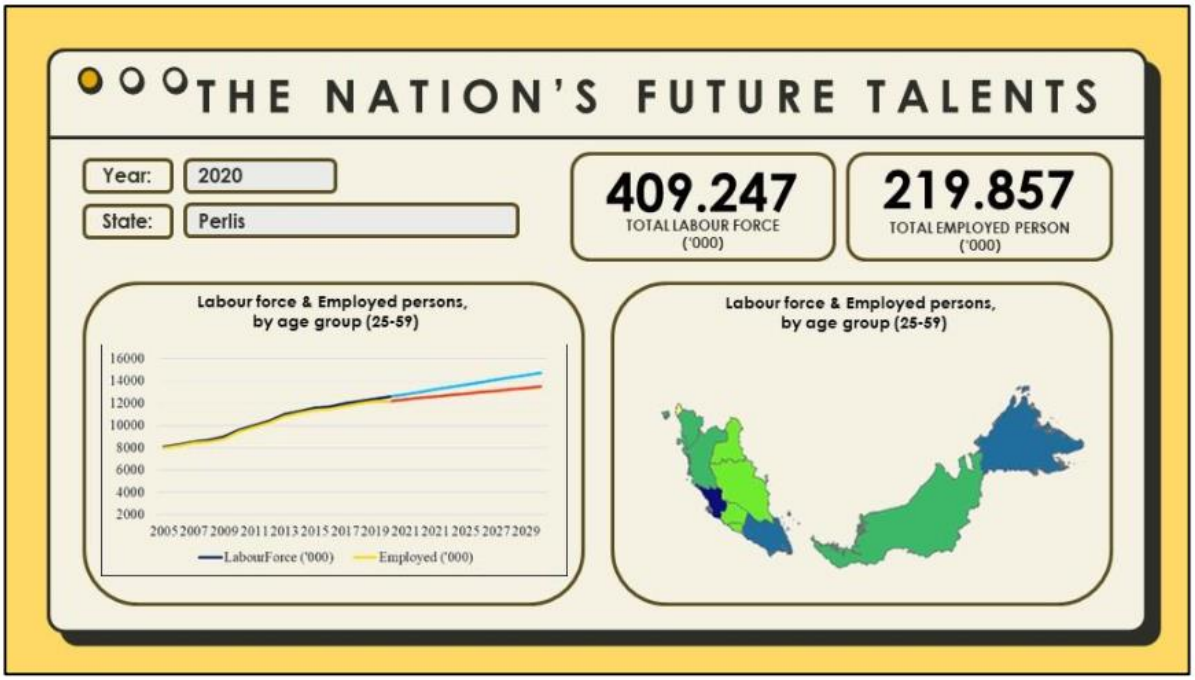


Figure 8. Proposed dashboard Graphical User Interface (GUI)

References

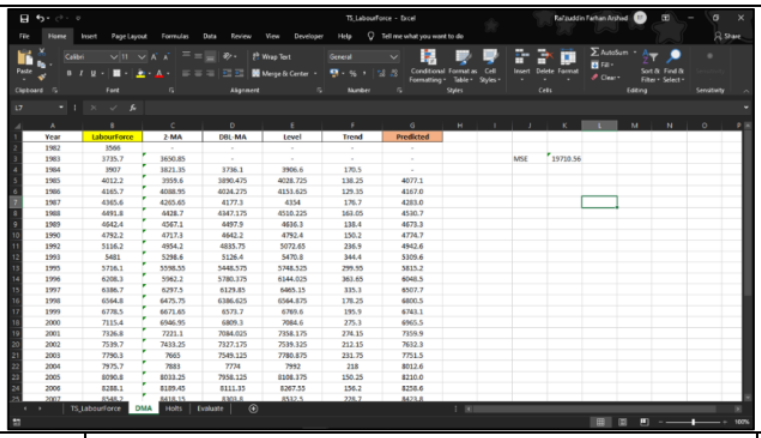
[1] United Nations’ Employment and decent work (poverty eradication), https://www.un.org/development/desa/socialperspectiveondevelopment/issues/employment-and-decent-work.html/ Retrieved June 13, 2022.

[2] Rahman, A. A., & Jasmin, A. F., “The Vulnerability of Jobs to COVID-19: The Case of Malaysia,” *ISEAS Yusof Ishak Institute Economics Working Paper*, No. 2020 – 09, 1-19 (2020). http://hdl.handle.net/11540/12748.

[3] Lee, H. A., “Unemployment among Malaysia’s Youth: Structural Trends and Current Challenges,” *Ideas – Yusof Ishak Institute*, 65, 1–16. (2020).

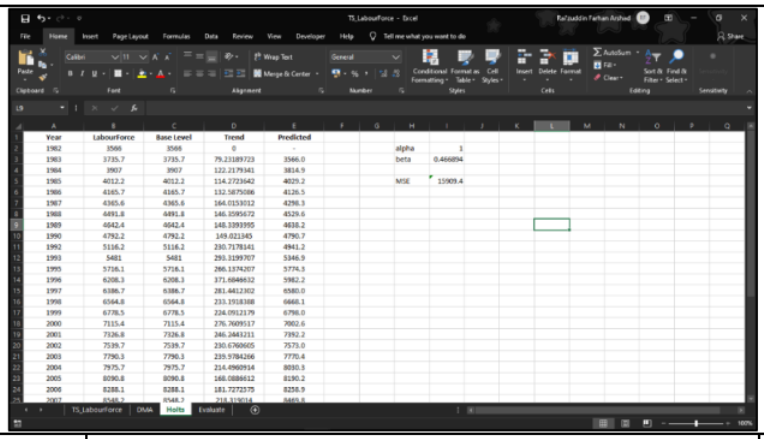
Appendix

1) Double Moving Average for Labour Force



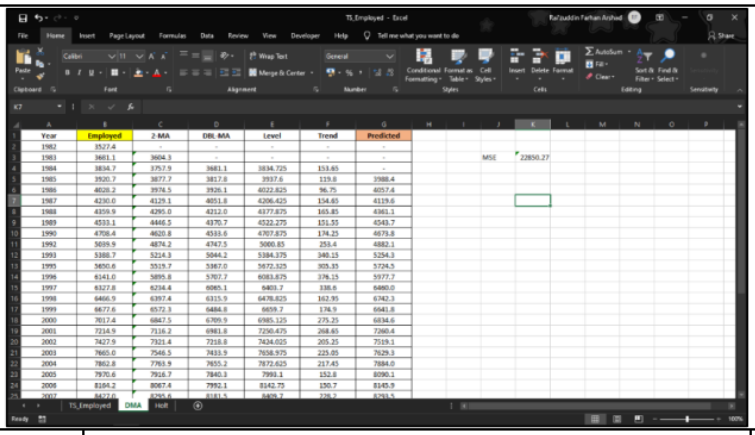
Cell	Name	Formula	Copied to
C3	3-MA	=AVERAGE(B2:B3)	C4:C38
D4	DBL-MA	=AVERAGE(C3:C4)	D4:D38
E4	Level	=2*C4-D4	E5:E38
F4	Trend	=2*(C4-D4)/(2-1)	F5:F38
G5	Predicted	=E4+F4	G5:G38
K5	MSE	=SUMXMY2(B5:B38,G5:G38)/COUNT(G5:G38)	--

2) Double Exponential Smoothing (Holt's Method) for Labour Force



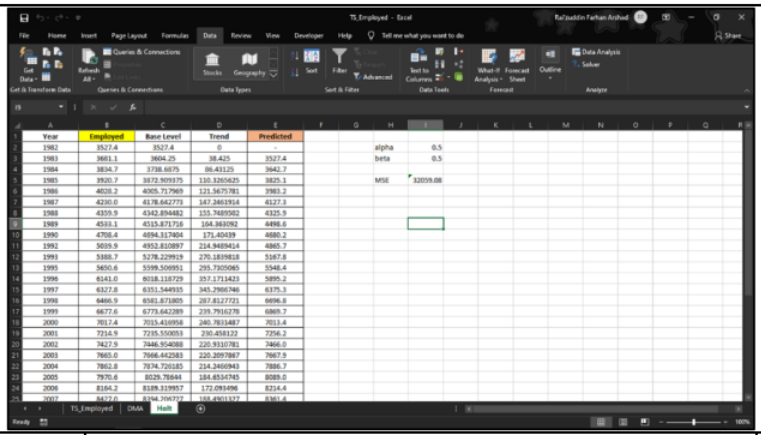
Cell	Name	Formula	COPIED TO
C2	Base Level	=B2	--
C3	Base Level	=$I$2*B3+(1-$I$2)*(C2+D2)	C4:C38
D2	Trend	=0	--
D3	Trend	=$I$3*(C3-C2)+(1-$I$3)*D2	D4:D38
E3	Predicted	=SUM(C2:D2)	E4:E38
I5	MSE	=SUMXMY2(B3:B38,E3:E38)/COUNT(E3:E38)	--

3) Double Moving Average for Employed Persons by age group



Cell	Name	Formula	Copied to
C3	3-MA	=AVERAGE(B2:B3)	C4:C38
D4	DBL-MA	=AVERAGE(C3:C4)	D4:D38
E4	Level	=2*C4-D4	E5:E38
F4	Trend	=2*(C4-D4)/(2-1)	F5:F38
G5	Predicted	=E4+F4	G5:G38
K5	MSE	=SUMXMY2(B5:B38,G5:G38)/COUNT(G5:G38)	--

4) Double Exponential Smoothing (Holt's Method) for Employed Persons by age group



Cell	Name	Formula	COPIED TO
C2	Base Level	=B2	--
C3	Base Level	=$I$2*B3+(1-$I$2)*(C2+D2)	C4:C38
D2	Trend	=0	--
D3	Trend	=$I$3*(C3-C2)+(1-$I$3)*D2	D4:D38
E3	Predicted	=SUM(C2:D2)	E4:E38
I5	MSE	=SUMXMY2(B3:B38,E3:E38)/COUNT(E3:E38)	--