QBUS6860 GROUP ASSIGNMENT

Group 180

GROUP MEMBER & SID

Hongjin Gou 510332280
Chen Yang 520397585
Xingyu Tao 510360537
Chaoying Shi 520030855

WORD COUNT: 3425

Table of Contents

1. Introduction **2**
2. Key Findings **2**
3. Main analysis of key business questions I **2**
4. Main analysis of key business questions II **8**
5. Summary **11**
6. Limitations **11**
7. References **12**
8. Appendix **13**

1. Introduction

The main objective of this analysis is to investigate the relationship between fuel prices and crude oil prices, focusing specifically on petrol stations in NSW. Additionally, it aims to assess the impact of fluctuations in crude oil prices on petrol prices in the broader society. The analysis involves three datasets, including fuel prices for all suburbs in NSW, crude oil price data from 2019 to 2020, and the Socio-Economic Index for Areas (SEIFA), which provides socio-economic advantage and disadvantage scores for various regions. During the data cleaning process, the suburb names in the fuel dataset will be standardized to follow a consistent title format. This standardization will facilitate further analysis of the data. Furthermore, outliers in the data are also identified and removed (as shown in the appendix 1*).* In addition, dates are filtered in accordance with the Excel short date format and arranged in chronological sequence from earliest to latest. In addition, the method of filling in the missing data is to use machine learning and random forest models to train the data set to get relatively accurate filling data.

The more obvious outliers are Postcode and price. As shown in the box plot, some outliers are greater than 3000, which is unreasonable, because we know that the postcode of NSW is between 2000 and 3000. In addition, most of the fuel prices in NSW are concentrated at 120 to 140, and some outliers are 50 or 60, which is literally out of range.

2. Key Findings:

- On a monthly basis, fluctuations in crude oil prices have an immediate impact on changes in gasoline prices.
- Fuels that contain a higher proportion of unleaded petrol are significantly influenced by changes in crude oil prices.
- The price of the fuel of P95 and P98, which is of higher quality for unleaded oil, tends to fluctuate in a smaller magnitude than P91.
- Fuel prices in NSW are strongly influenced by crude oil prices with a certain time delay, which can be caused by external factors and consumer behavior.
- The fuel price in the regions with higher social-economic scores tends to fluctuate in a larger magnitude.
- Relative socioeconomic advantage and disadvantage scores appear to be the most impactful on fuel prices.
- Economic determinants predominantly influence prices, likely due to the local economic development strategies and market circumstances in NSW.
- The Index of Socio-economic Advantages and Disadvantages has a greater impact on average fuel prices across zip codes than the Index of Economic Resources.

3. Main analysis of key business questions I

According to the Australian Government, the report indicates that global oil prices were influenced by the COVID-19 pandemic, which subsequently affected the pattern of domestic petrol prices (2020). Additionally, the COVID-19 pandemic led to limitations on business operations and global production as a result of traffic restrictions implemented by most countries to prevent the spread of the virus (2021). To gain deeper insights into these dynamics, a comprehensive analysis of the actual trends for both crude oil prices and retail petrol prices is required. Therefore, we will focus on the period spanning from the third quarter of 2019 to the third quarter of 2020, by creating average price plots on a monthly basis, which would possibly visualize and assess the degree of significance of the influence of crude oil on the Australian retail petrol market.

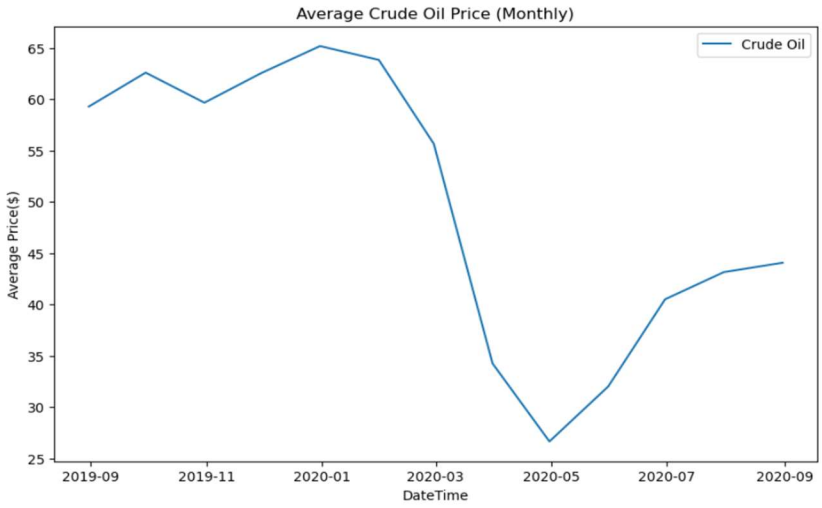


Figure 1. Monthly Crude Oil Price

The provided graph illustrates the mean price of crude oil spanning from August 2019 to August 2020. Initially starting at approximately $60 per barrel, the price reached its peak at $65 in January 2020. However, over the next four months, there was a substantial decline, with the price experiencing a sharp decline to reach a low point of $26. The decrease in crude oil prices can be attributed to the reduced demand resulting from the implementation of restrictions by countries.

The merged graph displays both the previously mentioned graphs for crude oil and petrol prices, allowing for a direct comparison. It is evident that any fluctuations observed in crude oil prices correspondingly influence petrol prices in the same direction. This drastic drop can be attributed to various factors, including global market conditions, reduced demand due to the COVID-19 pandemic, and geopolitical events impacting oil supply and demand dynamics.

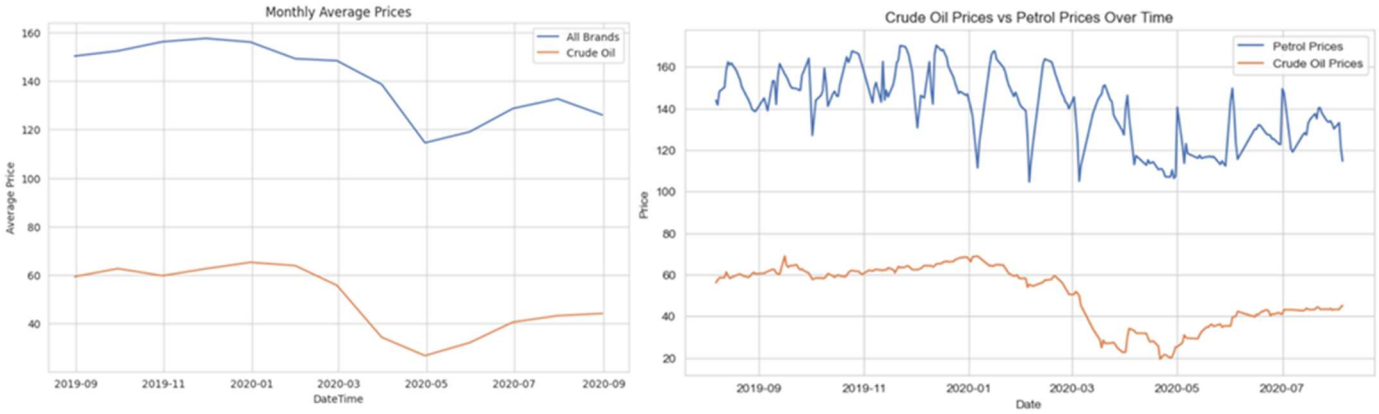


Figure 2. Monthly vs Average Comparison

Monthly price curves were plotted for different types of fuel in NSW and they were compared against the fluctuations in crude oil prices. The analysis revealed that traditional gasoline, including P95, P98, and E10, showed a strong correlation with changes in crude oil prices. The fuels, which heavily relied on unleaded petrol derived from crude oil, were highly sensitive to its fluctuations. On the other hand, fuels like Diesel (DL) and B20 exhibit lower sensitivity to crude oil price changes. E85, which combined ethanol and unleaded petrol, was not significantly affected by crude oil fluctuations. Additionally, electric vehicles (EV) are not impacted by crude oil prices at all.

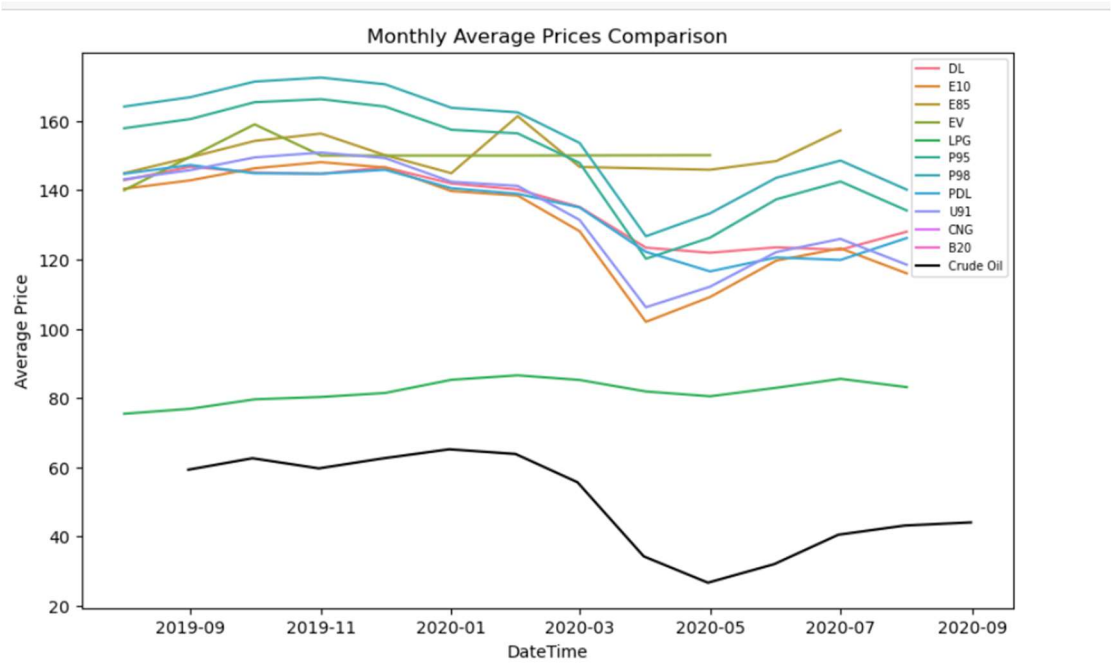


Figure 3. Monthly Average Fueltype vs Crude Oil Prices

In the discussion of how prices of different fuel types are affected in different geographical segments, we will first divide the regions into three segments: The region near the City, the region in the suburbs, and the regions far from the City. The basis of the division will be discussed in Q2.

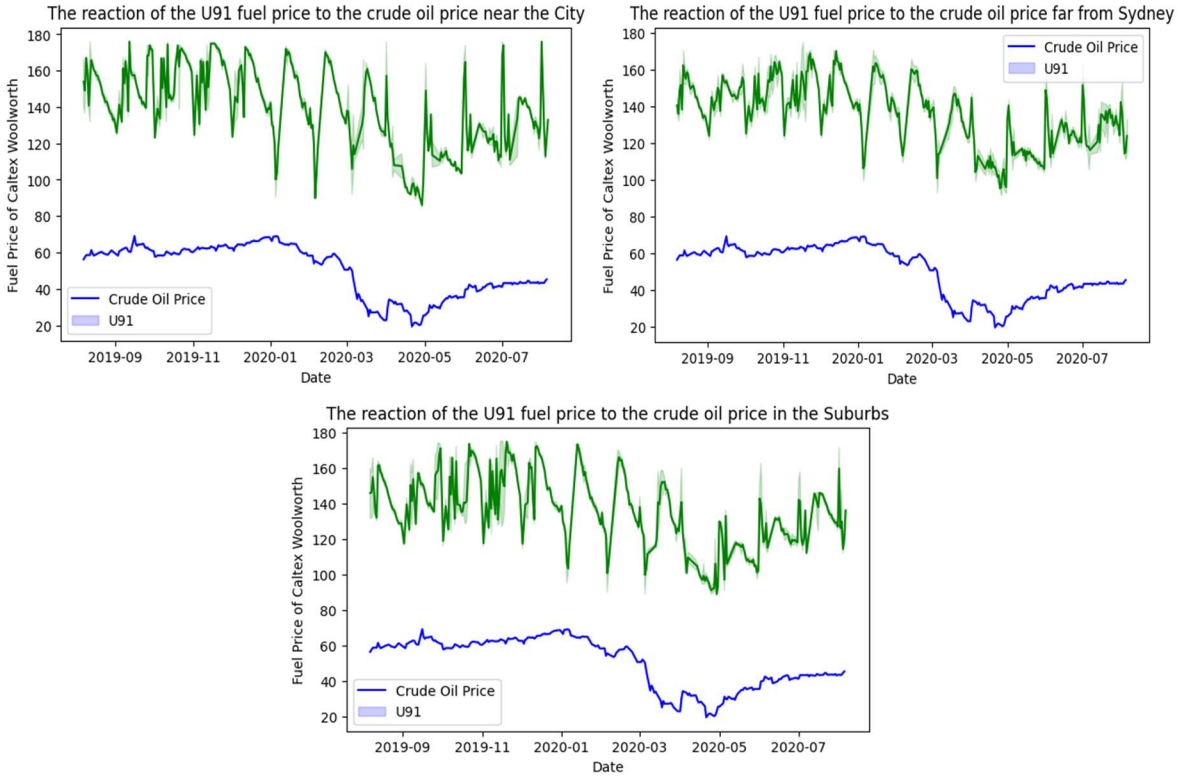


Figure 4. The fluctuation of U91 fuel

From the above three graphs, we could roughly see that the price of fuel type U91 is reacting simultaneously to the price of the crude oil. Further discussion of whether there is actually any time lag in the reaction of the fuel would be dependent on the analysis of residual over time.

We could also imply from the graph that the aptitude of fluctuation of the U91 is larger than that of crude oil price. The order of fluctuation magnitude of U91 in those three regions, from largest to smallest, is U91 near the City, in the suburbs and far from the city.

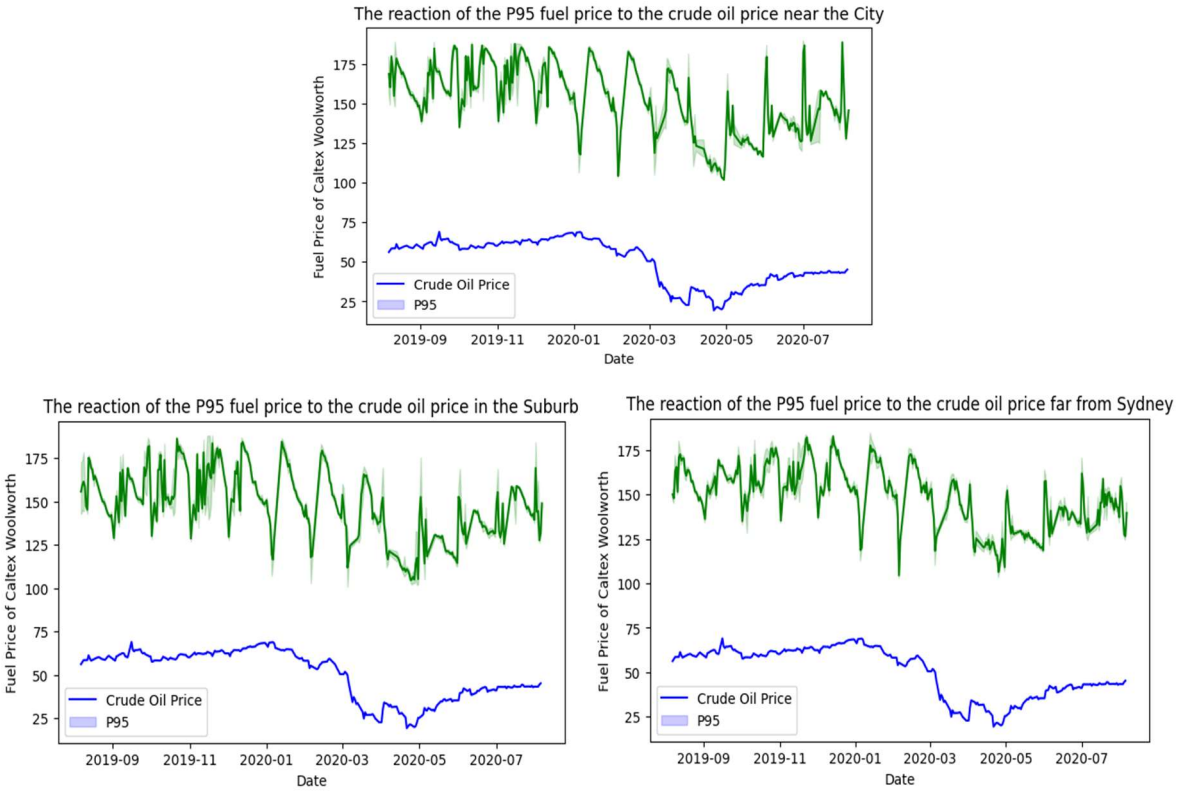
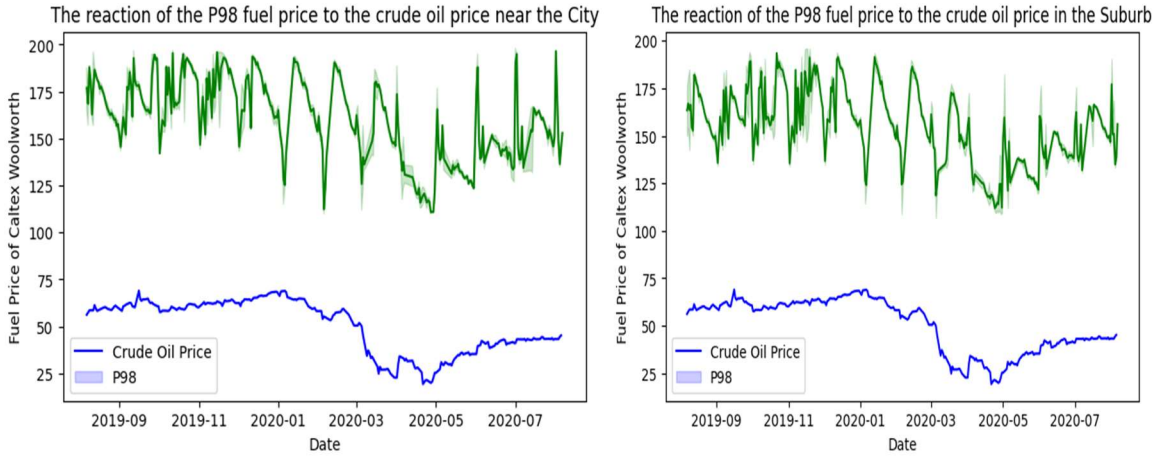


Figure 5. The fluctuation of P95 fuel

According to the graph, we could roughly see that the fluctuation of the price of the fuel P95 shows a tendency of reacting monotonously and simultaneously with the crude oil price, but not as significant as those of U91. The graph roughly shows that aptitude of fluctuation of the fuel is larger in all three regions than the crude oil prices, and there is no significant difference in the fluctuation magnitude in those three regions.



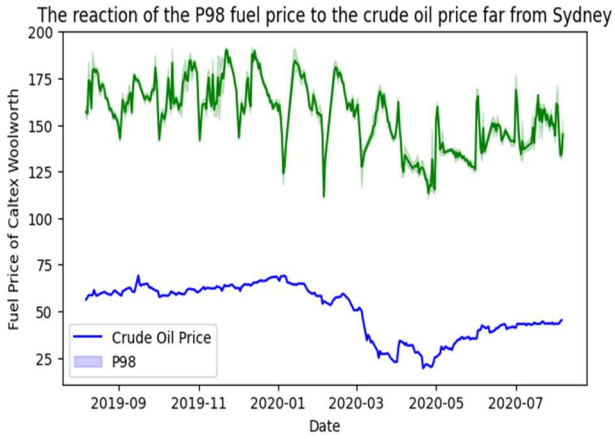


Figure 6. The fluctuation of P98 fuel

The fluctuation pattern and magnitude of the fuel P98 is similar to those of P95, and it is reacting roughly simultaneously to the crude oil price. The pattern and the magnitudes of the fluctuation in the three regions near the city, in the suburbs and far from the city in a similar pattern.

Overall, the price of the three different fuels reacts with the same tendency and simultaneously to the crude oil prices, but with a larger scale of magnitude. Further implication that we could imply from the graphs is that the period of one rise and fall of the fuel is one month, while the price tends to be high in the middle of the month and the price tends to be lower in the beginning and the end of the month.

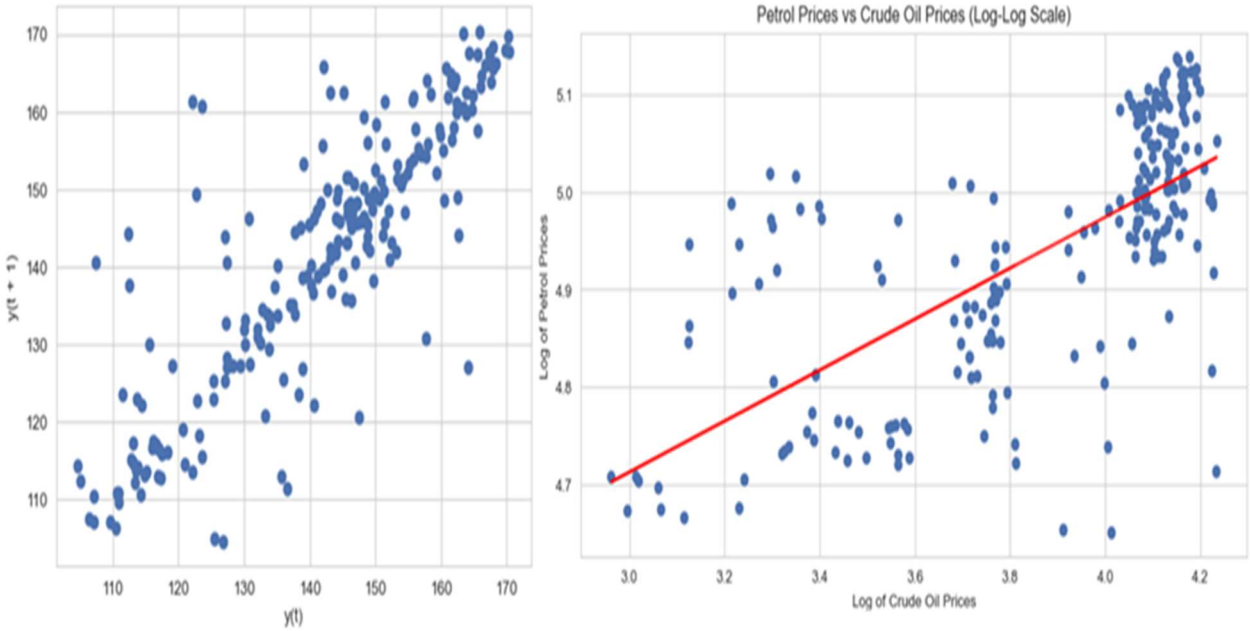


Figure 7. Time series relationship between crude oil price and fuel price

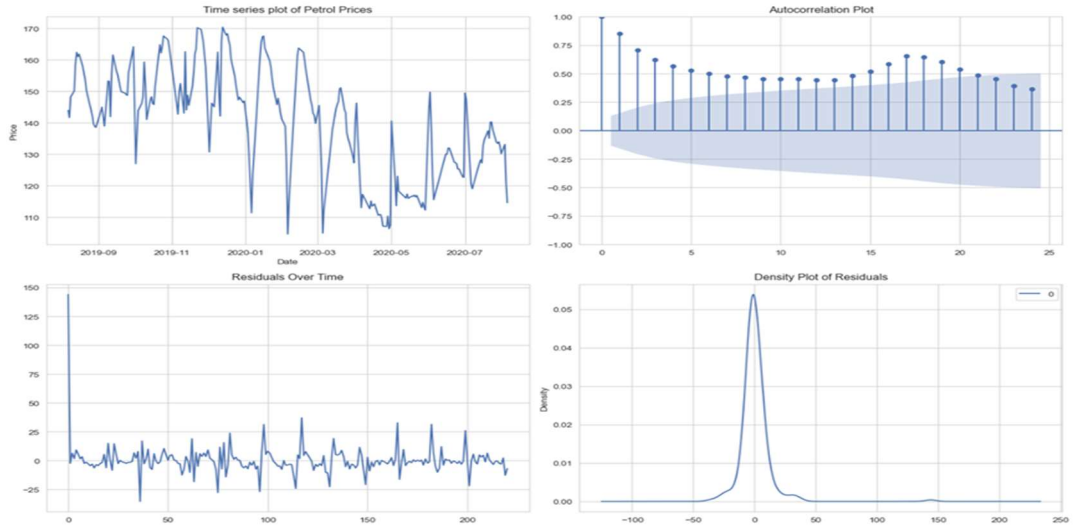


Figure 8. SARIMA visualization of crude oil price and fuel price

According to the above findings, Petrol prices in NSW were significantly influenced by the prices of crude oil. This relationship was observed to be linear in figure 7, which means that a time series influenced crude oil prices directly impacts the petrol prices consumers face (Blair et al., 2017). However, the response was not immediate; there was a clear time lag between changes in crude oil prices and their reflection in petrol prices. For instance, the SARIMA model in appendix 2 shows a noticeable delay between November 2019 and March 2020. During this period, each unit increase in crude oil prices resulted in a delayed 0.2 unit increase in petrol prices. This is illustrated in the autocorrelation plot in figure 8 which cleared that the impact of crude oil price changes on petrol prices increases with time. Furthermore, this delay could be exacerbated by external factors. In March 2019, the crude oil price surged, but the subsequent increase in petrol prices was delayed significantly due to the Covid-19 lockdown, which limited people's mobility for about a month (HT, 2013). As a result, the impact of this crude oil price increase was most pronounced in May 2019.

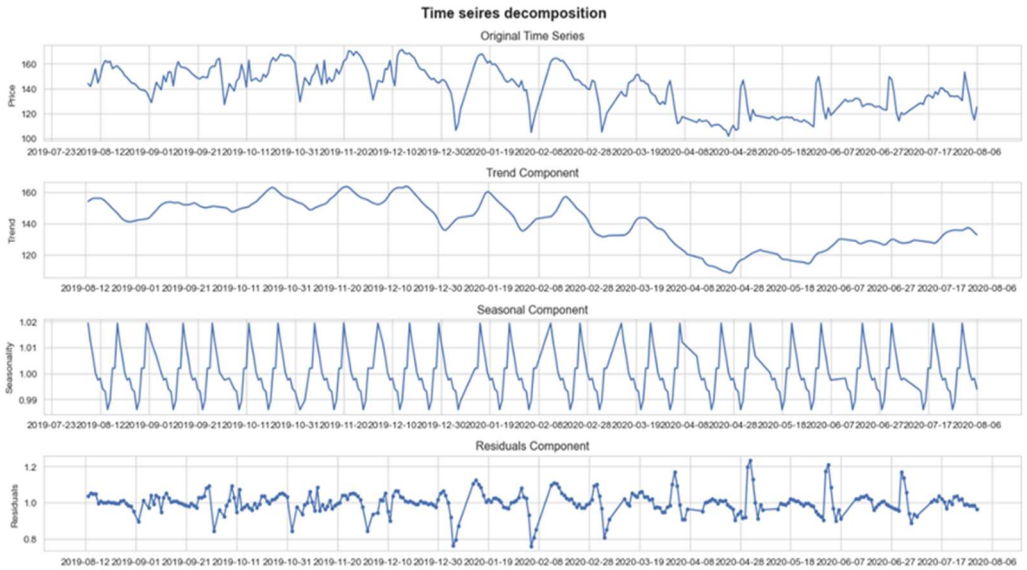


Figure 9. Time series decomposition

In figure 9, the time series decomposition further supports these observations. The trend component indicates that the underlying trends in crude oil and petrol prices tend to stabilize over time, suggesting that the lag in price changes becomes more predictable. Regarding

seasonality, there's a clear pattern where both crude oil and petrol prices are lower at the beginning and end of each month, peaking mid-month. This suggests consumer habits may influence these trends, potentially due to the typical timing of refueling (Blair et al., 2017).

Overall, it's clear that while petrol prices in NSW are strongly influenced by crude oil prices, this relationship manifests with a certain time delay, which can be influenced by external factors and consumer behavior. For recommendations, fuel suppliers in NSW are likely to focus on forecasting and inventory management. Anticipate potential increases or decreases in gasoline prices by monitoring trends in crude oil prices and adjust inventories accordingly (Xiong et al., 2021). For example, given signs of peak demand in the middle of the month, suppliers should build up inventories in response to increased demand (Shi et al., 2013). On the other hand, retailers at gas stations could consider implementing a dynamic pricing strategy that takes into account observed seasonality and delayed responses to crude oil prices. In other words, they may adjust prices in real time based on supply and demand. Uber, for example, offers dynamic pricing through its "dynamic pricing model" (Shi et al., 2013). During periods of high demand, prices rise to balance driver supply and demand (Shi et al., 2013). Gas stations could consider a similar strategy, perhaps offering lower prices at the beginning and end of the month and higher prices mid-month when demand peaks.

4. Main analysis of key business questions II

In dataset Seifa which contains all Australian postcodes, to specifically target New South Wales state, create a new data frame by extracting postcodes from 2000 to 2999, and combining mean fuel prices from dataset 1 and selecting two index scores (IER and IRSAD) from dataset 3 to investigate the relationship among postcode, mean fuel price and index scores.

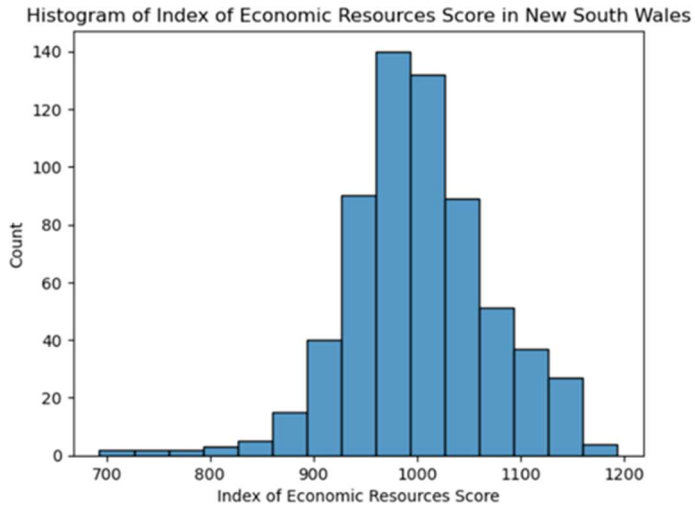


Figure 10. Distribution of ERS in New South Wales

Investigating the relationship among the postcodes, mean fuel price and index scores in the following analysis, we use postcodes to represent regions in New South Wales, it is noticeable that one postcode may stand for more suburbs. And two main indexes are used to examine the socioeconomic impact degree: the Index of Economic Resources Score (IER) and the Index of Relative Socio-economic Advantage and Disadvantage Score (IRSAD). IER was selected to measure a region's economic resources and IRSAD was selected to measure a region's advantage score.

Broadly, plotted a histogram to show the overall distribution of the Index of IER in New Sales Wales. IER stands for financial aspects of incomes; it is measured by areas with higher scores having relatively greater access to economic resources than areas with lower scores. In this

case, the figure 10 shows the IER overall distribution in New South Wales. IER in New South Wales has a positively left-skewed distribution with a mean of 1001, which indicates that most of the area has a higher IER.

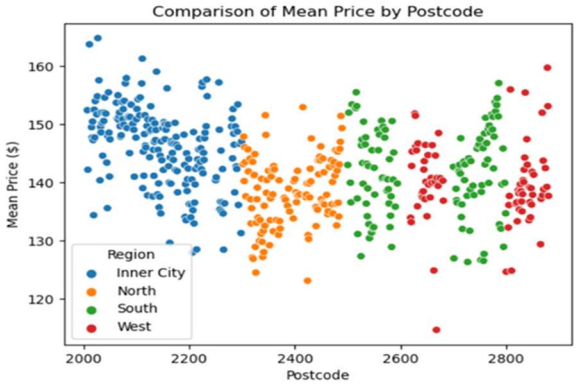


Figure 11. Average fuel price vs Postcode

Geographically, *Postcode* starts from 2600 and 2800 close to the west, 2500 and 2700 close to the south, and 2300 and 2400 close to the north, otherwise considered as the inner city in this context. The above graph shows the comparison of the mean price of regions were divided by postcodes. It shows that the main fuel price in the inner city is higher than the other three suburbs.

Figure 12. Comparison of average IER

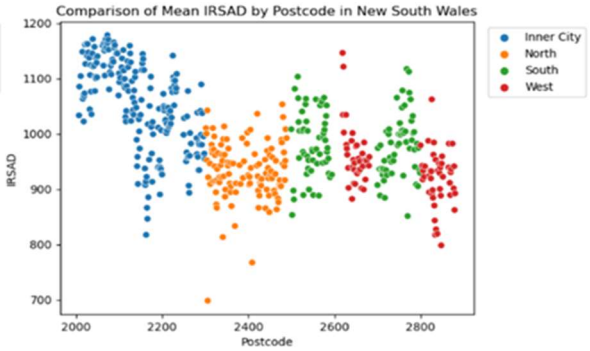
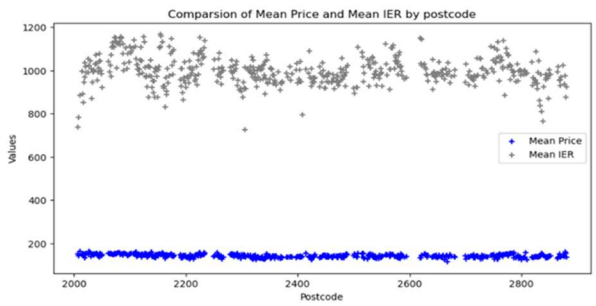


Figure 13. Comparison of average IRSAD

To detail investigate whether the socioeconomic factor impact on average fuel price is higher than other suburbs. Two index scores were selected based on the correlation of the index scores and postcodes (as shown in the Appendix 3). From comparing the two index scores moving trends, figure 12 shows that there is not much difference related to the IER score in different regions. So, the IER score will have less influence on fuel price change. Oppositely, figure 13 shows that the Inner city has the highest IRSAD score, followed by the South, and the North and West have the lowest score. (It is assumed that people who live close to the inner city have a relatively high incidence of advantage of fuel price.)



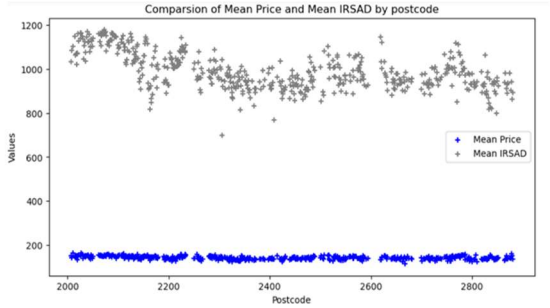


Figure 14. Comparison of two average scores by postcode

In general, from plotting two comparable graphs as figure 14 showed of the relationship between mean price with IER and IRSAD scores, the single index score has less impact on the fuel price. So, there is not much lag reaction related to the influence of crude oil on the petrol price. And the fluctuation of the mean price of different regions indicates that the influence effect will affect New South Wales as a whole.

We could divide the postcodes into the other three geographical regions based on the proximity to the city of Sydney. For instance, postcodes with numbers 2000 to 2100 are the closest to the city, with numbers 2100 to 2200 are in the suburbs near the city, while 2200 to 2300 is far from the city of Sydney. The other basis of separating those postcodes is the socio-economic advantage and disadvantage score. The regions near the city with the postcode from 2000 to 2100 is with the social and economic score of 1137.0, while the regions in the suburbs with the postcode from 2100 to 2200 and the far from the city of Sydney with the postcode from 2200 to 3000 is with the score of 1048.0 and 963.0, respectively. We could infer from the data that the further the suburb is from the city, the lower the social and economic advantage and disadvantage score is, which indicates that in regions far from the city people might be having a lower standard of life than the areas near the city.

As mentioned above, the fluctuation of the fuel price is in larger magnitude in the regions that have higher social-economic scores. This might be caused by the inherently higher price and higher price elasticity of the fuel near the city (since the demand in respective regions is relatively stable in the long run) (Wadud et al., 2009).

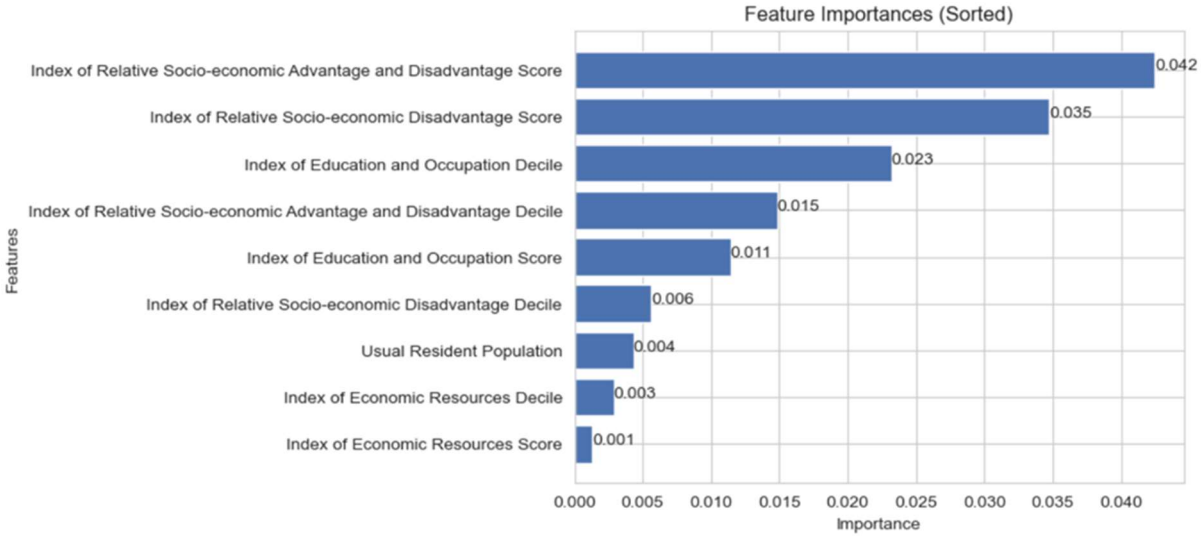


Figure 15. logistic model of fuel prices vs economic indicators

Further exploration of fuel prices and various socio-economic indicators may be useful for improving pricing strategies. As depicted in figure 15, the logistic model shows the influence of each index on fuel prices. The Relative Socio-economic Advantage and Disadvantage Score appears to be the most impactful, constituting 4.2% of the influence. This indicates a robust positive correlation between oil prices and relative economic value.

However, factors such as education, occupation score, and resident population account for less than 1% each, demonstrating a minor role of population-related factors on the pricing of fuel. This suggests that economic determinants predominantly influence prices, likely due to the local economic development strategies and market circumstances in NSW.

In terms of recommendations, it would be beneficial for the government to devise some strategies like fuel taxation, aimed at matching fuel supply with the demand of local

communities, thus fostering sustainable development (Xiong et al., 2021).

For further suggestions in the future, since trade of fuel has been negatively impacted by the Covid-19 pandemic (AbdelMaksoud et al., 2022), it is predictable that the demand of the fuel could be higher, which draws up the price and the magnitude of fluctuation of price of the fuel near the city of Sydney. Thus, it is recommended that people traveling long distances in NSW could refill the fuel in areas far from Sydney rather than those near the city. Furthermore, although the premium fuel tends to be price higher, people who are risk averse could choose it instead of standard unleaded oil since the price of the premium fuel is relatively stable and of higher quality.

5. Summary

Overall, there is a clear positive correlation between crude oil prices and fuel prices. However, there is a delay in the impact of changes in crude oil prices on fuel prices, with significant delays observed in specific periods, such as from November 2019 to March 2020. In addition, for CBDs in more economically developed coastal areas, the sales volume is higher, which also leads to higher prices than other areas. On the other hand, in NSW, fuel prices are lower in suburbs to the west and north, possibly due to relatively low economic indicators. The higher prices of suburbs in the east and south may be due to the more developed economy, education, occupation and population. In addition, relative socioeconomic advantage and disadvantage scores have a strong positive correlation with fuel prices, indicating that economic factors have the most significant impact on fuel prices. However, education level, occupational scores, resident population and other factors have little influence, each accounting for less than 1%. Based on these conclusions, some suggestions can be made to make fuel pricing closer to social needs and market reality. Forecasting and inventory management, dynamic pricing strategies, and government intervention can all play a key role in achieving this alignment (Xiong et al., 2021).

6. Limitations

However, it's important to note the shortcomings of the analysis. Firstly, the data set used for this analysis is drawn exclusively from sales records of New South Wales (NSW) for the period from the last half of 2019 to the first half of 2020. To further improve the accuracy and generalizability of these findings, it would be beneficial to incorporate data sets from a wider range of regions and time periods, such as national sales data across Australia for the entirety of 2020. Secondly, the method used to address missing data, namely, filling in gaps with a random forest algorithm, although useful, isn't perfect and might introduce some bias into the analysis. Future studies could benefit from using multiple imputations or other advanced statistical techniques for handling missing data. Lastly, this study did not include any text analysis that could enable the matching of more detailed addresses to their respective suburbs. This potentially reduces the accuracy of analyses regarding the impact of local community economic indices on fuel prices.

Therefore, while this analysis provides useful insights and recommendations, these limitations should be addressed in future research to provide a more comprehensive and accurate understanding of the factors influencing fuel prices.

7. Reference list

AbdelMaksoud, K. M., Hathout, H. M. ., & Albagoury, S. H. (2022). The socio-economic impact of COVID-19 on the petroleum sector in Egypt: a descriptive analysis. *International Journal of Social Economics*, *49*(1), 21–33. https://doi.org/10.1108/IJSE-10-2020-0688

Australia Government (2020). Gov.au. Retrieved May 23, 2023, from

https://www.bitre.gov.au/sites/default/files/documents/bitre-is111-petrol_prices.pd

Blair, B. F., Campbell, R. C., & Mixon, P. A. (2017). Price pass-through in US gasoline markets. *Energy Economics*, 65, 42–49.https://doi.org/10.1016/j.eneco.2017.04.011

PMC(2021). Nih.gov. Retrieved May 23, 2023, from

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9756000/

Shi, W., Yang, Z., & Li, K. X. (2013). The impact of crude oil price on the tanker market. *Maritime Policy and Management*, 40(4), 309–322https://doi.org/10.1080/03088839.2013.777981

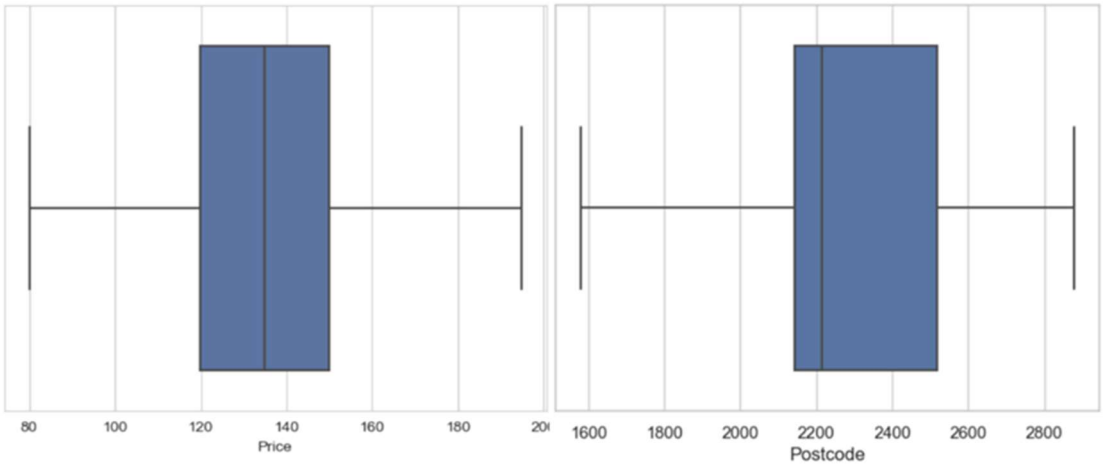
Wadud, Z., Graham, D. J., & Noland, R. B. (2009). Modelling fuel demand for different socio-economic groups. *Applied Energy*, *86*(12), 2740–2749. https://doi.org/10.1016/j.apenergy.2009.04.011

Will Covid-19 Impact On Crude Oil Tanker Market? - *Identify Which Types Of Companies Could Potentially Benefit*. (2020). HT Digital Streams Limited.

Xiong, F., Xu, S., & Zheng, D. (2021). An investigation of the uber driver reward system in China - an application of a dynamic pricing model. *Technology Analysis & Strategic Management*, 33(1), 44–57.

https://doi.org/10.1080/09537325.2020.1787975f

8. Appendix



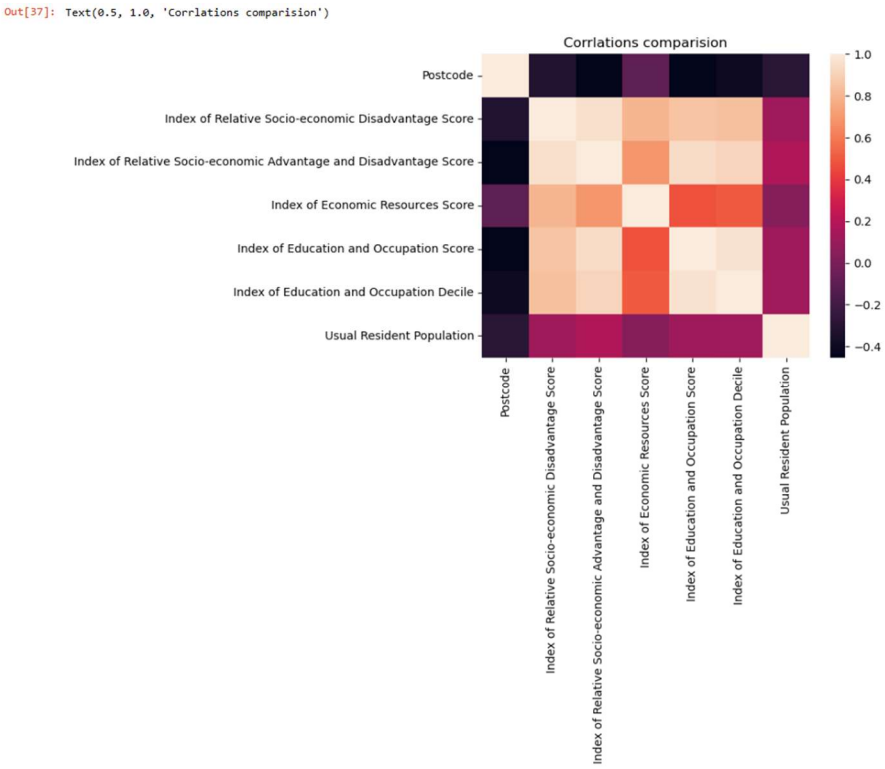
Appendix 1. Data cleaning of the fuel dataset

SARIMAX Results

Dep. Variable:		Price	No. Observations:			221
Model:		ARIMA(5, 1, 0)	Log Likelihood			-795.827
Date:		Tue, 23 May 2023	AIC			1603.654
Time:		14:46:48	BIC			1624.016
Sample:		0	HQIC			1611.877
		- 221				
Covariance Type:		opg				
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.0645	0.062	-1.041	0.298	-0.186	0.057
ar.L2	-0.2079	0.062	-3.359	0.001	-0.329	-0.087
ar.L3	-0.1418	0.066	-2.135	0.033	-0.272	-0.012
ar.L4	-0.1232	0.086	-1.435	0.151	-0.291	0.045
ar.L5	-0.0652	0.102	-0.642	0.521	-0.264	0.134
sigma2	81.1300	4.935	16.440	0.000	71.458	90.802

Ljung-Box (L1) (Q):	0.01	Jarque-Bera (JB):	176.60
Prob(Q):	0.92	Prob(JB):	0.00
Heteroskedasticity (H):	1.17	Skew:	0.44
Prob(H) (two-sided):	0.51	Kurtosis:	7.30

Appendix 2. Results of SARIMAX model



Appendix 3. Correlations of SEIFA

Median of Social-economic Indexes Near the City

```
Index of Relative Socio-economic Advantage and Disadvantage Score    963.0
Index of Relative Socio-economic Disadvantage Score                  979.0
Index of Economic Resources Score                                    991.0
Index of Education and Occupation Score                              963.5
dtype: float64
```

Median of Social-economic Indexes In the Suburb

```
Index of Relative Socio-economic Advantage and Disadvantage Score    1048.5
Index of Relative Socio-economic Disadvantage Score                  1025.5
Index of Economic Resources Score                                    1009.0
Index of Education and Occupation Score                              1057.5
dtype: float64
```

Median of Social-economic Indexes Far From Sydney

```
Index of Relative Socio-economic Advantage and Disadvantage Score    1137.0
Index of Relative Socio-economic Disadvantage Score                  1097.0
Index of Economic Resources Score                                    1031.5
Index of Education and Occupation Score                              1157.0
dtype: float64
```

Appendix 3. Median of Social-economic Indexes in Various Regions