section A

1. **Narrower prediction intervals are more informative and should always be preferred.**
   - I disagree with this statement. While narrower prediction intervals may provide more precise estimates, they may also be overly optimistic and fail to capture the uncertainty inherent in the forecasting process. Wider prediction intervals, on the other hand, acknowledge the possibility of variation and are more conservative in their estimates. Therefore, the preference for narrower or wider prediction intervals depends on the specific context and the desired level of risk tolerance.
2. **The AICC should always be used to select models for forecasting.**
   - I partially agree with this statement. The Akaike Information Criterion corrected for finite sample sizes (AICC) is a useful tool for model selection as it penalizes model complexity to avoid overfitting. However, it should not be the sole criterion for model selection. Other factors such as theoretical considerations, diagnostic tests, and the interpretability of the model should also be taken into account to ensure the chosen model is appropriate for the data and the forecasting task at hand.
3. **An ETS model for Holt's linear trend method is a generalization of an ETS model for simple exponential smoothing. It should therefore always be preferred as it will produce better forecasts.**
   - I disagree with this statement. While Holt's linear trend method, which is an Exponential Smoothing State Space Model (ETS), extends simple exponential smoothing by incorporating a trend component, it may not always be the best choice for forecasting. The appropriateness of Holt's linear trend method depends on the presence and nature of a linear trend in the data. In cases where the data exhibit nonlinear trends or seasonality, alternative methods such as seasonal exponential smoothing or ARIMA models may produce better forecasts. Therefore, the selection of the forecasting method should be based on the characteristics of the data and the forecasting objectives.
4. **The trouble with forecasting is that it assumes the patterns in the past will continue in the future.**
   - I partially agree with this statement. While forecasting often relies on historical patterns to make predictions about the future, it is essential to recognize that past patterns may not always persist due to changing circumstances, unexpected events, or structural shifts in the data-generating process. Therefore, forecasting should not solely rely on historical data but should also incorporate judgment, expertise, and external information to account for potential deviations from historical patterns. While past behavior can provide valuable insights, it is not a guarantee of future outcomes, and forecasting should be approached with caution and flexibility.

Section B

1

The time series plots depicting birth data from 1980 to 2019 reveal intriguing insights into the dynamics of birth rates over the past four decades. The general plot spanning the entire period highlights discernible trends and seasonal patterns, offering a broad perspective on birth rate fluctuations over time. Meanwhile, the yearly plot with distinct colors for each year provides a detailed view of year-to-year variations and potential cohort effects, enabling a closer examination of changes in fertility patterns over different generations. Additionally, the monthly plots for each year offer a granular analysis of intra-year variability and seasonal effects, shedding light on the nuanced dynamics of birth rates within individual years. Together, these plots offer a comprehensive understanding of the complex interplay of factors influencing birth rates, from long-term trends to short-term fluctuations, providing valuable insights for further research and policy considerations.

2

Yes I am happy with it.

these models decomposisont show and explain alot in terms of explainaing how the data behaves Decomposition plots illustrate birth data's components: trend (long-term direction), seasonality (recurring patterns), and remainder (unexplained fluctuations). By analyzing these components, we uncover underlying trends, seasonal effects, and random variation, offering insights into birth rate dynamics over the past four decades.

3

(a) Suitable. The seasonal naïve method is appropriate for forecasting data with strong seasonal patterns, such as birth data, as it simply repeats the previous year's seasonal pattern.

(b) Suitable. STL decomposition helps in capturing seasonality and trend, making it suitable for birth data. The drift method adds a linear trend, enhancing its forecasting capability.

(c) Suitable. STL decomposition on log-transformed data helps stabilize variance, suitable for data with changing variance like birth data. ETS further models the seasonally adjusted component effectively.

(d) Suitable. Holt-Winters method with damped trend and additive seasonality can capture both trend and seasonality, making it suitable for birth data with potentially changing trends.

(e) Suitable. ETS(A,N,A) is appropriate for data with no trend or seasonality, ensuring robust forecasts for birth data with stable patterns.

(f) Suitable. ETS(A,AD,M) can handle data with additive trend and multiplicative seasonality, which might be present in birth data, making it suitable for forecasting.

(g) Not suitable. ARIMA(1,1,4) may not be appropriate as it introduces differencing, which may not be necessary for birth data and can result in overly complex models.

(h) Not suitable. ARIMA(3,0,2)(1,1,1)4 is likely too complex for birth data forecasting, potentially overfitting the data and making interpretation challenging.

(i) Suitable. ARIMA(1,0,2)(2,1,0)12 incorporates seasonal differencing and autoregressive terms, suitable for capturing seasonality and autocorrelation in birth data.

(j) Suitable. Regression with time and Fourier terms can capture both linear trends and seasonal patterns, making it suitable for birth data forecasting.

SECTION C

1

The `fit_ETS` object represents fitted exponential smoothing state space models (ETS) for forecasting monthly birth data in Victoria. The accompanying tibble summarizes key model estimates, including smoothing parameters, seasonal components, and trend characteristics, providing insights into how each model captures and forecasts the data's underlying patterns.

2

the plot explore the models seasonality trend and all compnents of ts such that we can see the data decompose.

3

3. Based on the provided output, the models' fit can be assessed using metrics such as AIC, AICC, BIC, and MSE. Lower values indicate better fit. Among the models, ETS(A) and ETS(Ad) have the lowest AIC, AICC, and BIC values, indicating better fit. However, their MSE values are slightly higher than ETS(N), suggesting that ETS(N) may provide slightly more accurate forecasts. Therefore, I would choose ETS(N) for forecasting the number of births over the next two years due to its lower MSE and comparable AIC, AICC, and BIC values.

4

The estimated ETS(N) model can be written in full as follows:

$$\hat{y}_{t+1} = \ell_t + b_t + s_{t+1-m}$$

where: - $\hat{y}_{t+1}$ is the forecast for the next period $t+1$. - ℓ_t represents the level component at time t. - b_t represents the trend component at time t. - s_{t+1-m} represents the seasonal component for the next seasonal cycle $t+1-m$.

This model assumes no trend ($b_t = 0$) and only considers the level (ℓ_t) and seasonal (s_{t+1-m}) components for forecasting.

5

- For the level component:
  - February 2019: $\ell_{Feb} = 5.90$
- For the seasonal component:
  - February 2019: $s_{Feb} = 0.318$

Now, let's calculate the forecasts:

- For $h = 1$ step ahead (March 2019): $\hat{y}_{Mar} = \ell_{Feb} + s_{Mar} = 5.90 + s_{Feb} = 5.90 + 0.318 = 6.218$
- For $h = 4$ steps ahead (June 2019): $\hat{y}_{Jun} = \ell_{Feb} + s_{Jun} = 5.90 + s_{Feb} = 5.90 + 0.318 = 6.218$
- For $h = 12$ steps ahead (February 2020): $\hat{y}_{Feb} = \ell_{Feb} + s_{Feb} = 5.90 + 0.318 = 6.218$
- For $h = 13$ steps ahead (March 2020): $\hat{y}_{Mar} = \ell_{Feb} + s_{Mar} = 5.90 + s_{Feb} = 5.90 + 0.318 = 6.218$

6

- Forecast for March 2019 ($\hat{y}_{Mar}$): 6.218

We also need the estimated standard deviation of the forecast errors (σ). This information is typically available from the model output but wasn't provided explicitly. We can estimate it using the mean squared error (MSE) from the model output:

- MSE: 0.0173 (from the output)

The standard deviation of the forecast errors (σ) is the square root of the MSE:

$$\sigma = \sqrt{MSE} = \sqrt{0.0173}$$

$$\sigma \approx 0.1316$$

Now, to calculate the 80% confidence interval for the 1-step ahead forecast:

- Margin of error (ME) at 80% confidence level:

  $$ME = Z \times \frac{\sigma}{\sqrt{n}}$$

  where Z is the Z-score corresponding to the desired confidence level, and n is the sample size (which is 1 in this case).

For an 80% confidence level, the Z-score is approximately 1.282.

$$ME = 1.282 \times \frac{0.1316}{\sqrt{1}}$$

$$ME \approx 0.169$$

7 The forecasts from the ETS(N) model, as presented in Figure 7, likely differ from those shown in Figure 5 due to the different characteristics and assumptions of the ETS models used. ETS(N) assumes no trend and only considers the level and seasonal components for forecasting. This may lead to smoother forecasts, especially in the absence of any apparent trend or seasonality changes. In contrast, other ETS models may incorporate trends or additional seasonal patterns, leading to more nuanced and possibly more accurate forecasts, especially over longer horizons.

In longer-term forecasts, differences between the three estimated ETS models are expected to become more pronounced. ETS(N) forecasts may continue to be relatively smooth, capturing primarily the underlying level and seasonal patterns. ETS(A) forecasts may exhibit some trend behavior but with less variability compared to ETS(Ad), which includes a damped trend component. ETS(Ad) forecasts may show more pronounced trend behavior initially, but the damping parameter will likely lead to convergence towards a stable level over longer horizons. Overall, the choice of ETS model for longer-term forecasts depends on the data characteristics and the desired balance between capturing short-term fluctuations and long-term trends.

secion D

1

Differencing with a lag of 12 months is applied to the monthly birth data in Victoria to remove seasonality. The resulting series appears stationary, indicated by a non-significant KPSS p-value of 0.264, suggesting it lacks significant trends or seasonality, making it suitable for further analysis.

3

1. The ARIMA model (1,0,1)(2,1,0)[12] is selected based on its lowest AIC value among candidate models. This model includes one autoregressive term, one moving average term, seasonal differencing, and two seasonal moving average terms.

2. The residuals in Figure 9 appear to have no discernible pattern, indicating that the model adequately captures the underlying structure of the data. The residuals seem to fluctuate around zero without any systematic trend or seasonal pattern.

3. While the model selection process and the appearance of residuals suggest a good fit, the downward trend in the forecasts plotted in Figure 10 raises concerns. This trend indicates that the model may not adequately capture longer-term dynamics or structural changes in the data. Therefore, further investigation and potentially a more complex model may be necessary to improve longer-term forecasting accuracy.

4 The downward trend in the forecasts from the ARIMA model could be due to several factors. Firstly, the model may not adequately capture all the underlying dynamics of the data, such as trend or seasonality. Additionally, if the historical data exhibit a downward trend, the model may extrapolate this trend into the future.

For longer-term forecasts, we would expect the downward trend to persist if the underlying data trend continues unchanged. However, if there are seasonal or cyclical patterns in the data, these may eventually counteract the downward trend, leading to stabilization or even an upward trend in the forecasts over the long term. Therefore, it's important to monitor the forecasts and reassess the model periodically to ensure its accuracy and relevance to the changing data patterns.