Research Question 4: Ordinal Regression

Objective of Analysis

This analysis aims to examine the relationship between maternal age, work status, and the progression of the pandemic on insomnia severity among mothers. Maternal age is a critical demographic factor, as older mothers may face increased care giving burdens and stress, potentially exacerbating sleep disturbances. Conversely, younger mothers might experience sleep difficulties due to the demands of new motherhood and economic pressures. Work status can influence sleep through structured routines or financial stress, respectively, with research showing significant impacts of unemployment on sleep quality during the pandemic (**?**). The progression of the pandemic introduces cumulative stressors, such as uncertainty, isolation, and disrupted routines, which are known to impair sleep over time (**??**). By analysing these predictors the study seeks to identify key factors influencing maternal insomnia severity and inform targeted interventions.

The data analysed, published in **?**, can be found in the CSV file *'Copy_of_Dataset_ Miranda_et_al.csv'* inside the data sub-directory. The dataset comprises 1000 Argentinian postpartum women, collected through a cross-sectional survey conducted between May and July 2020. This study investigates factors influencing **insomnia severity** during the COVID-19 pandemic, measured using the Insomnia Severity Index (ISI). The ISI is a widely used tool for assessing sleep disturbances, providing a total score ranging from 1 to 4. To facilitate the analysis, the ISI total score was recoded into an ordinal variable with four categories: No insomnia (ISI score < 8), Subthreshold insomnia (ISI score 8-14), Moderate Severity (ISI score 15-21), and Severe insomnia (ISI score > 21), aligning with established guidelines for interpreting ISI scores (**?**). Within this sample, 31.3% of participants reported no significant insomnia ($n = 313$), 43.3% reported subthreshold insomnia ($n = 433$), 22.5% reported moderate severity insomnia ($n = 225$), and 2.9% reported severe insomnia ($n = 29$). The mean score of the sample was 1.97 ($SD = 0.81$). **Maternal age**, a binary variable, was categorised into younger mothers (aged under 35 years) and older mothers (aged 35 years and older). The sample included: 76.2% younger mothers ($n = 762$) and 23.8% older mothers ($n = 238$). The variable **days of pandemic** was calculated as the number of days since March 11, 2020, when the World Health Organisation declared COVID-19 a pandemic, to the date the participant completed the survey. The mean was 409.04 days ($SD = 253.19$), ranging from 52 to 721 days (median = 438). **Work Status** was categorised as employed or unemployed. The sample comprised: 75% employed mothers ($n = 750$) and 25% unemployed mothers ($n = 250$).

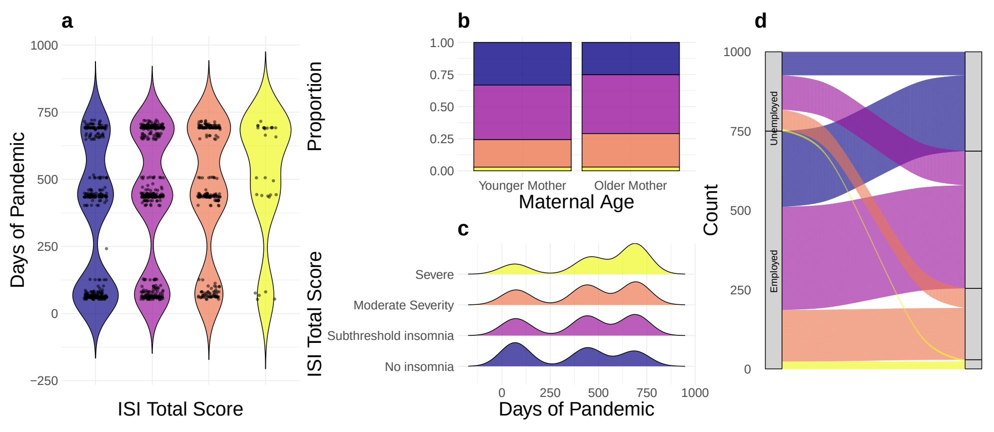

Figure 1: Visualised Distributions of Variables.

Figure **??**: (a) Days of Pandemic by ISI Total Score, (b) Distribution of ISI Categories by Maternal Age, (c) Days of Pandemic by ISI Categories, (d) Sankey Diagram of Work Status and ISI Categories.

Statistical Model

Ordinal logistic regression models, also known as proportional odds models, are a specialised type of generalized linear model used to analyse outcomes that are ordinal. These outcomes have a natural order but do not assume equal spacing between categories, such as levels of insomnia severity (*e.g., no clinically significant insomnia, subthreshold insomnia, clinical insomnia*). This model is particularly useful for analysing ranked data, where the goal is to assess the influence of multiple predictors on the likelihood of observing higher versus lower categories. The ordinal logistic regression model predicts the cumulative probability of an outcome being in or below a specific category. Mathematically, the model is expressed as:

$$\text{logit}(P(Y \leq j)) = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k$$

where $P(Y \leq j)$ is the cumulative probability of being in category j or below, Y is the ordinal outcome (insomnia severity), α_j represents the intercept for category j, and β_k are coefficients associated with the predictors X_k (e.g., days of pandemic, maternal age, and work status). The logit function ensures the linearity of predictors in relation to the log-odds of the cumulative probabilities. The ordinal logistic model assumes proportional odds, meaning the relationship between predictors and log-odds remains constant across categories. This simplifies interpretation and ensures parsimony. If violated models like generalized ordered logistic regression may be more appropriate. Unlike multinomial regression, which treats categories as unordered, ordinal regression respects the ordered nature of the outcome while reducing parameters, enhancing interpretability and generalisability. It extends binary logistic regression by accommodating more than two ordered categories. It can be expressed in R using the MASS package as:

```r
ordinal_model <- polr(isi_total_score ~ maternal_age + days_of_pandemic + work_status,
                      data = mother_df, Hess = TRUE)
```

Model diagnostics show no concerns with multicollinearity (VIF < 5), influential points (Cook's Distance), or residuals, indicating a good fit. The Brant test confirmed that the proportional odds assumption holds, ensuring consistent relationships between predictors and insomnia severity categories, implying a generalized ordered logistic regression was not required.

Model Comparison

```r
glm_linear <- lm(as.numeric(isi_total_score) ~ maternal_age + days_of_pandemic
                 + work_status, data = mother_df)
ordinal_model <- polr(isi_total_score ~ maternal_age + days_of_pandemic + work_status,
                     data = mother_df, Hess = TRUE)
interaction_ordinal <- polr(isi_total_score ~ maternal_age + days_of_pandemic
                           + work_status + work_status * maternal_age, data = mother_df,
                           Hess = TRUE)

anova(ordinal_model, interaction_ordinal)
```

Table 1: Model Comparison Table

Model	AIC	BIC	Residual Deviance	Log Likelihood	McFadden R2	Nagelkerke R2
Linear Model	2392.61	2417.15	634.28	-1191.31	–	–
Ordinal Model	2310.18	2339.62	2298.18	-1149.09	0.01	0.03

Model	AIC	BIC	Residual Deviance	Log Likelihood	McFadden R2	Nagelkerke R2
Interaction Ordinal Model	2309.64	2344.00	2295.64	-1147.82	–	–

The drop test and likelihood ratio tests affirm the critical role of predictors in explaining insomnia severity (Table **??**). The ordinal model significantly improves model fit compared to the linear model, with lower AIC (2310.18 vs. 2392.61) and BIC (2339.62 vs. 2417.15). Adding the interaction term further reduces the AIC (2309.64) but slightly increases the BIC (2344), indicating the improved fit does not offset the added complexity. Including days of pandemic and maternal age significantly improves model fit ($\chi^2(1) = 2.54$, $p = 0.11$), while work status, though less impactful alone, may have theoretical value when interacting with other variables (**?**). Pseudo-R^2 metrics, such as McFadden's ($R^2 = 0.01$) and Nagelkerke's ($R^2 = 0.03$) indicates that the ordinal model explains a modest but meaningful proportion of variance in insomnia severity. These results underscore the importance of retaining key predictors and their interactions to capture nuanced relationships and enhance model explanatory power.

Model Evaluation and Interpretation

Table 2: Ordinal Logistic Regression Coefficients.

Effect	Estimate	Std. Error	Odds Ratio	p-value	95% CI
Older Mother (Maternal Age)	-0.15	0.33	0.86	0.653	[-0.79, 0.5]
Days of Pandemic	0.00	0.00	1.00	$< .001$***	[0, 0]
Employed (Work Status)	-0.31	0.15	0.73	0.039*	[-0.61, -0.02]
Older Mother × Employed	0.58	0.36	1.78	0.112	[-0.13, 1.29]
No Insomnia → Subthreshold Insomnia	-0.48	0.16	0.62	0.002**	[-0.79, -0.18]
Subthreshold Insomnia → Moderate Severity	1.43	0.16	4.19	$< .001$***	[1.11, 1.75]
Moderate Severity → Severe	3.89	0.24	48.83	$< .001$***	[3.42, 4.36]

Note. Significance levels: $*** \ p < .001$, $** \ p < .01$, $* \ p < .05$.

The regression coefficients for the predictors in the model indicate their impact on the log odds of experiencing higher levels of insomnia severity (Table **??**). Older mothers have an estimate of -0.15, suggesting a 13.93% decrease in the likelihood of higher insomnia severity compared to younger mothers (-14%). However, this effect is not statistically significant, as evidenced by the p-value (0.65) and the confidence interval (95% CI: [-0.79; 0.5]) which includes zero. This result suggests that maternal age may not independently influence insomnia severity in this context. Days of pandemic was a highly significant predictor, with an odds ratio of 1 ($p = 0$), indicating a 0% increase in the odds of higher insomnia severity per additional day of the pandemic, highlighting the cumulative mental health burden as the pandemic progresses. Work status showed a significant reduction in the likelihood of higher insomnia severity, with employed mothers showing a 26.66% decrease in odds (0.73, $p = 0.04$, 95% CI: [-0.61; -0.02]), possibly due to financial stability or structured routines. The interaction between older mothers and employment suggested a possible increase in insomnia severity for older employed mothers (1.78, $p = 0.11$), but this effect was not statistically significant.

Thresholds in the model define the cutoffs between categories of insomnia severity and are essential to the ordinal logistic regression model. For example, the threshold separating No Insomnia → Subthreshold Insomnia has an estimate of -0.48 (95% CI: [-0.79; -0.18]) indicating a significant and distinct separation between these levels. Similarly, the threshold for Subthreshold Insomnia → Moderate Severity is 1.43 (95% CI: [1.11; 1.75]), and for Moderate Severity → Severe, the threshold is 3.89 (95% CI: [3.42; 4.36]). These significant thresholds confirm the ordered nature of insomnia severity levels.

The null deviance of the model is 2328.6 on 1 degree of freedom. The model deviance is 2295.6 on 996 degrees of freedom. The reduction in deviance from the null model to the final model shows how much better the model explains the data. A larger reduction in deviance with fewer parameters suggests a significant improvement in fit. Additionally, the small standard errors for significant predictors like days of pandemic and work status indicate precise and reliable estimates. Conversely, larger standard errors for non-significant predictors, such as maternal age, reflect greater uncertainty and wider confidence intervals. These patterns support the model's robustness and the precision of its effect estimates.

Predictions

```r
predict(ordinal_final_model, newdata = new_data, type = "probs")
```

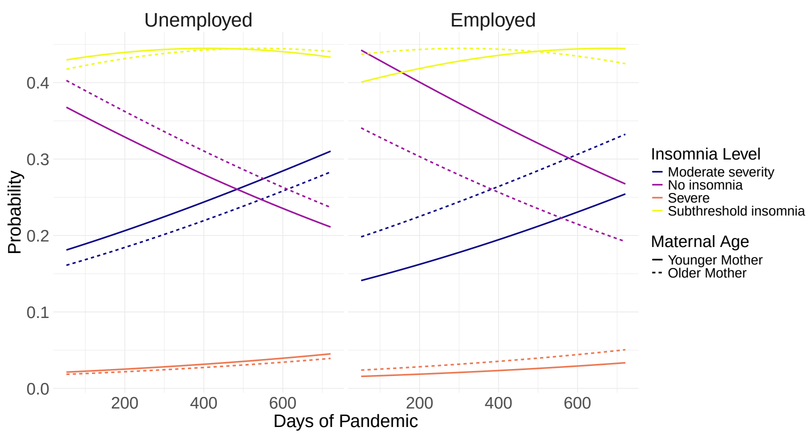

Figure 2: Predicted Probability of ISI Categories.

Figure **??** illustrates the predicted probabilities of various Insomnia Severity Index categories over the course of the pandemic, segmented by maternal age (younger mother vs. older mother) and employment status (employed vs unemployed). It reveals distinct trends, such as the decreasing probability of no insomnia over time, particularly for unemployed individuals, while probabilities for more severe categories, such as moderate severity and severe, gradually increase. The differences between younger mothers generally exhibit higher probabilities for no insomnia and lower probabilities for severe categories compared to older mothers. Employment status also appears to influence these trends, with employed mothers generally showing more stable probabilities across the ISI categories. These patterns suggest that maternal age and employment status interact with stressors from the pandemic to shape insomnia severity.

Conclusion

This analysis highlights how maternal age, employment status, and pandemic duration influence insomnia severity among postpartum women. Prolonged pandemic stress increased the likelihood of severe insomnia, while employment acted as a protective factor, with employed mothers showing lower risks. Younger mothers were more likely to report no insomnia, while older mothers showed slightly higher risks for severe categories, though this effect was minimal. Limitations include self-report biases and limited generalisability beyond the Argentinian context. These findings underscore the need for targeted mental health interventions, particularly for unemployed mothers, alongside policies that promote financial stability and childcare access to reduce insomnia severity during prolonged crises. Future research should explore longitudinal designs to better understand causal relationships expand samples to improve generalisability, and incorporate additional predictors, such as mental health measures to capture a more comprehensive picture of insomnia severity.