Research Question 3: Multilevel Regression

Objective of Analysis

This analysis investigates how workplace and demographic factors, such as job category, gender, and age, influence task demand and depression, while accounting for group-level variability like job roles and pandemic experiences. The hierarchical nature of the data, where individuals are nested within job categories, necessitates the use of multilevel regression models. These models decompose variance into individual- and group-level components. They address the non-independence of observations within groups, improving parameter estimates. By modelling group-level effects, the analysis captures shared experiences within job categories that shape perceptions of workload and mental health (**??**).

The data analysed are published in **?** and can be found in the CSV file *'workplace_stress.csv'* inside the data sub-directory. The data comprised 1575 full-time Italian office workers recruited online. The primary outcome variable, **Task Demand** (TD), is a composite score calculated as the mean of four items (e.g., "doing multiple things at once"). Higher scores indicate greater task demand. In this sample, the mean score is 4.05 ($SD = 1.32$), with scores ranging from 1 to 7 (*median* = 4). **Job Category**, a categorical variable, serves as the grouping variable and represents different job roles in the sample, with 18 unique categories. **Gender** is coded as a binary variable, 50.86% of participants identified as male ($n = 801$) and 49.14% of participants identified as female ($n = 774$). **Age** is a continuous variable, with participants' ages ranging from 22 to 63 (*mean* = 34.9, $SD = 9.5$, *median* = 31) years. Figure **??** shows (a) Task Demand by Job Category and (b) Distribution of Job Categories.
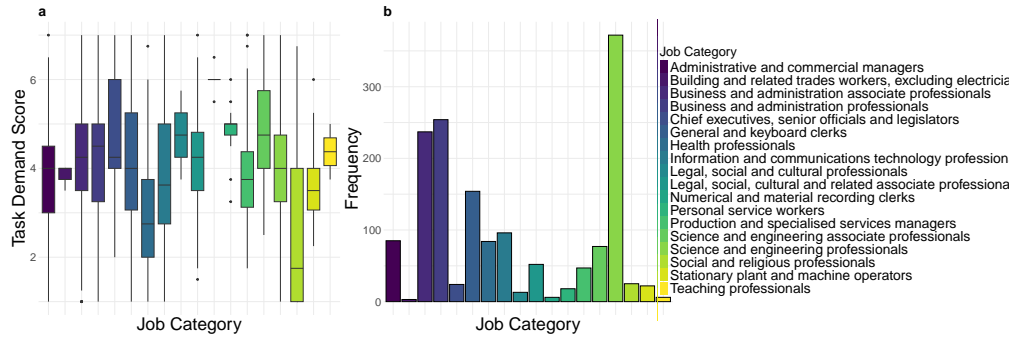

Figure 1: Distribution of Job Categories.

Statistical model

A **linear mixed-effects model** (also known as a multilevel regression model) was used to examine the relationship between predictors and task demand. This approach extends traditional regression techniques by incorporating random effects to account for hierarchical data structures. Since individuals are nested within job categories, task demand scores are likely correlated within groups. Multilevel models address this non-independence of observations by modelling both the group-specific variability (random effects) and common relationships across groups (fixed effects). Unlike general regression models, which assume independent observations, multilevel models account for nested data. While generalized linear models can handle non-Gaussian distributions, they still assume independent observations. The multilevel model is expressed mathematically as:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + u_j + \epsilon_{ij}$$

where Y_{ij} is the outcome variable for individual i in group j (job category), β_0 is the fixed intercept, representing the average task demand across all groups, $\beta_1 X_{ij}$ represents the fixed effect of predictor X_{ij} (e.g., gender or age), u_j is the random effect for group j, capturing deviations in baseline task demand scores

across job categories, ϵ_{ij} is the residual error term, accounting for variability within groups. Including u_j, capturing individual-level variability within groups.

This model incorporates a random intercept for job categories, allowing the baseline task demand scores to vary across groups. By capturing group-level variability, the model improves the accuracy and robustness of fixed effect estimates, providing insights into both individual-level and group-level factors influencing task demand. Multilevel models rely on residuals at both the within-group levels and between-group levels are assumed to follow a normal distribution. As well as, random effects are assumed to be independent of fixed effects and residual errors and the model assumes homogeneity of variance within groups, meaning variability does not differ substantially across clusters. The linear mixed-effects model was implemented in R using the lme4 package:

```r
mixed_effects_model <- lmer(task_demand ~ gender + age + (1 | job), data = work_df)
```

This model decomposes variance into between-group variability (job categories) and within-group variability (individual-level differences), improving the robustness and accuracy of estimates. Residual diagnostics and Variance Inflation Factor analysis revealed no major issues, confirming the robustness of the results.

Comparison of Models

```r
null_mixed_model <- lmer(task_demand ~ 1 + (1 | job), data = work_df)
fixed_effects_model <- lm(task_demand ~ gender + age, data = work_df)
polynomial_model <- lmer(task_demand ~ gender + age + I(age^2) + (1 | job),
                        data = work_df)
random_slope_model <- lmer(task_demand ~ gender + age + I(age^2) + (age | job),
                          data = work_df)

# Compare Models
anova(null_mixed_model, fixed_effects_model, polynomial_model, random_slope_model)
```

Table 1: Model Comparison: AIC, BIC, and R-squared.

Model	AIC	BIC	Marginal R2	Conditional R2
Null Model	5257.07	5273.16	–	–
Fixed Effects Model	5347.94	5369.39	–	–
Polynomial Model	5248.90	5281.07	0.02	0.24
Random Slope Model	5226.43	5269.33	0.01	0.26

A series of models were compared to assess the influence of age, gender, and job categories on task demand (Table **??**). The random slope model, which allowed the slope of age to vary across job categories, provided the best fit with the lowest AIC (5226.43) and BIC (5269.33), outperforming the polynomial (AIC = 5248.9, BIC = 5281.07) and fixed-effects models. The null model showed the poorest fit. The conditional R^2 for the random slope model indicated that 26.23% of the variance in task demand was explained by both fixed and random effects, while the marginal R^2 ($R^2_{\text{marginal}} = 0.01$) reflected the small but significant contribution of fixed effects alone. The random slope model highlights the variability in how task demand changes with age across job categories. Steeper slopes suggest task demand increases more with age, requiring targeted interventions like mentoring or workload adjustments. Flatter slopes indicate a more stable relationship between age and task demand. The likelihood test with the polynomial model confirmed the random slope model's superiority ($\chi^2(2) = 24.15$, $p < 0.001$). Despite additional parameters, its lower AIC and BIC indicate that the improvement in fit outweighed the added complexity. Simpler models, such as the polynomial and fixed-effects models failed to capture variability across job categories effectively. The quadratic term for age proved

critical, capturing nonlinear relationships that a linear term alone could not address. Removing it resulted in a significant reduction in explained variance (Sum Sq = 18.8, $p < 0.001$). Residual deviance decreased from 5251 (null deviance) to 5210, demonstrating that the predictors significantly improved explanatory power.

Model Evaluation and Interpretation

Table 2: Fixed Effects Coefficients with 95% Confidence Intervals.

	Estimate (Beta)	Std. Error	t-value	p-value	95% CI (Beta)
(Intercept)	1.94	0.79	2.45	0.014*	[0.39, 3.48]
Gender: Male	0.01	0.07	0.12	0.902	[-0.13, 0.15]
Age	0.12	0.04	3.30	$< .001$***	[0.05, 0.19]
Age (Quadratic)	0.00	0.00	-3.53	$< .001$***	[0, 0]

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

Table **??** provides insights into the relationship between task demand and key predictors, accounting for variability across job categories through random effects. The baseline task demand, represented by the intercept ($\beta = 1.94$, $p = 0.01$; 95% *CI*: [0.39, 3.48]), is statistically significant, indicating that the average task demand across all groups is non-zero when all predictors are held constant. The narrow confidence interval suggests precision in estimating the baseline task demand. The coefficient for gender ($\beta = 0.01$, $p = 0.9$; 95% *CI*: [-0.13, 0.15]) is not statistically significant. This result, along with a confidence interval that includes zero, suggests no measurable impact of being male on task demand. The lack of significance may reflect balanced task assignments across genders within job categories. The linear effect of age ($\beta = 0.12$, $p < 0.001$; 95% *CI*: [0.05, 0.19]) is positive and statistically significant, indicating that for every one-year increase in age, task demand increases by approximately 12.14%. The quadratic term for age ($\beta = -0.002$, $p < 0.001$; 95% *CI*: [0, 0]) is negative and statistically significant, indicating a diminishing return in the effect of age on task demand. The small standard errors for significant predictors, such as age ($SE = 0.04$) indicate precise estimates, enhancing the reliability of the findings. These results suggest that while task demand increases with age initially, the effect plateaus or declines among older employees, emphasizing the need for age-sensitive workload management.

The intraclass correlation coefficient (ICC) evaluates the proportion of variance in task demand attributable to differences between job categories. Using the random intercept variance alone, the ICC was estimated at 0.66, indicating that approximately 66.2% of the total variance in task demand is attributable to job category variability. After accounting for random slopes and covariance between random intercepts and slopes, the ICC remained at 0.66, suggesting that the inclusion of random slopes did not substantially alter the proportion of variance explained by job categories. The variance of the random intercept for job categories was $\sigma^2 = 2.95$, while the random slope variance for age was 0.002. This small random slope variance indicates minimal variability in the age-task demand relationship across job categories. The covariance (Cov = -0.075) suggests that job categories with higher baseline task demand tend to have a weaker relationship between age and task demand, while those with lower baseline task demand exhibit a stronger age effect.

Predictions

```
predict(random_slope_model, newdata = mixed_prediction_data)
```

Figure **??** illustrates the impact of random effects on task demand across job categories, capturing both varying intercepts (a) and varying slopes (b), as well as the trends of predicted task demand by job categories.
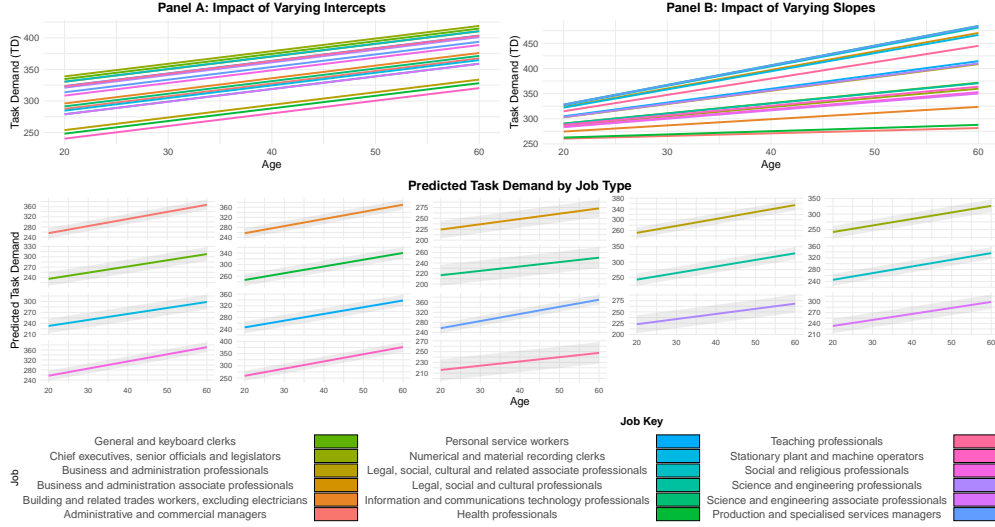
Figure 2: Predicted Impact of Random Effects on Task Demand across Job Categories.

In Panel a, the vertical separation of lines reflects differences in baseline task demand across job categories, with higher intercepts indicating roles such as Chief executives, senior officials and legislators that consistently experience higher task demands, and lower intercepts representing roles such as Numerical and material recording clerks with consistently lower demands. Panel b shows how the relationship between age and task demand varies by job category, as evidenced by differences in the slopes. Steeper slopes, seen in roles such as Teaching professionals, suggest a stronger age effect where task demand increases more substantially with age. Conversely, flatter slopes, as observed for General and keyboard clerks, indicate weaker age effects, where task demand is less influenced by age. The variability in task demands across job categories aligns with psychological stress theories, which suggest that unbalanced demands, especially in high-pressure roles, can exacerbate stress and contribute to burnout if not managed effectively. Structured routines and supportive workplace policies may act as coping mechanisms, reducing stress and improving task performance in such environments. The faceted plots (c) illustrate job-specific task-demand trends, with shaded ribbons representing the confidence intervals. These intervals highlight the precision and variability of predictions, aiding in understanding the degree of certainty associated with each job category.

Conclusion

This analysis underscores the critical role of job categories, age, and gender in shaping task demand, with job categories accounting for approximately 66.2% of the variability. Variability in baseline task demand and age-task demand relationships highlight the importance of tailored workplace policies, such as balancing workload for roles with high baseline demands, implementing mentoring programs to manage increasing age-related demands, and providing flexible work arrangements for older employees to mitigate burnout. However, the study's findings are limited by its reliance on industry-specific sampling reduces the generalisability of findings to other occupations or work environments. Potential over-simplification of clustering may arise from excluding predictors like education or organisational structure. Additionally, the cross-sectional design restricts insights into dynamic changes in task demand over time. Future research should expand to diverse industries, explore additional predictors, and use longitudinal data to capture task demand over time. These insights provide a foundation for targeted interventions and equitable workplace policies to enhance employee well-being and organisational productivity.