**Summary Statistics**

The dataset contains information from 3,994,063 respondents across various variables. The **Exercise** variable has a mean of 0.75, suggesting that, on average, 75% of the respondents engage in exercise, with a standard deviation of 0.43 indicating moderate variability. The exercise data is binary, with 0 representing no exercise and 1 representing exercise.

For the **Income** variable, the average income is 67,227, with a substantial standard deviation of 42,458.19, indicating significant variation in income levels among respondents. The income values range from 1,000.47 to 167,762.6, showing a wide disparity in the financial status of the sample.

The **Employed** variable is also binary, with a mean of 0.53, indicating that about 53% of respondents are employed. The standard deviation of 0.50 suggests notable variation in employment status, with values ranging from 0 (unemployed) to 1 (employed).

The **Age** variable shows an average age of 54.67 years, with a standard deviation of 16.53 years, indicating a diverse age range within the sample. The ages span from 18 to 99 years, suggesting that the sample covers a broad demographic.

For **General Health**, the mean value of 2.51 indicates that, on average, respondents rate their general health between "Fair" and "Good" on a 1 to 5 scale. The standard deviation of 1.06 reflects some variation in how individuals perceive their health. The range for this variable is from 1 (Poor health) to 5 (Excellent health).

The **Physical Health** variable has a mean of 4.04, likely indicating the average number of days respondents reported poor physical health in the past month. The large standard deviation of 8.51 suggests considerable variation in physical health among respondents, with values ranging from 0 to 30 days of poor health.

Similarly, the **Mental Health** variable has a mean of 3.50, which likely reflects the average number of days of poor mental health in the past month. The standard deviation of 7.69 indicates significant variability in mental health experiences among respondents, with the variable spanning from 0 to 30 days of poor mental health.

Summary Statistics

| variable | mean | Sd | min | max |
| --- | --- | --- | --- | --- |
| Exercise | 0.7535472 | 0.4309453 | 0.000 | 1.0 |
| Income | 67,227.4498842 | 42,458.1887403 | 1,000.466 | 167,762.6 |
| Employed | 0.5334665 | 0.4988788 | 0.000 | 1.0 |
| Age | 54.6722128 | 16.5282757 | 18.000 | 99.0 |
| General Health | 2.5127206 | 1.0628596 | 1.000 | 5.0 |
| Physical Health | 4.0394267 | 8.5078924 | 0.000 | 30.0 |
| Mental Health | 3.5040792 | 7.6915059 | 0.000 | 30.0 |

**Regression Results**

The regression analysis results, presented in Table 2, offer valuable insights into the factors influencing physical activity levels. The model accounts for approximately 9.9% of the observed variation in physical activity, as reflected by the R² value. This indicates that the included variables, such as income, employment status, age, and health measures, provide a modest but significant explanation of physical activity behavior. The adjusted R² corroborates this finding, suggesting that the model strikes a balance between explanatory power and parsimony.

Income appears to have no significant direct effect on physical activity, as indicated by its near-zero coefficient. This suggests that income alone may not be a key determinant of physical activity levels within the observed data. However, this finding warrants further investigation into potential non-linear relationships or interactions with other variables, such as employment type or general health, which could reveal more nuanced effects.

Employment status shows a small but statistically significant negative association with physical activity. With a coefficient of -0.036 (p < 0.001), the results suggest that employed individuals may engage less in physical activity compared to those who are unemployed. This pattern could reflect constraints on time or energy associated with work commitments. Future research could explore the impact of job type, hours worked, and workplace wellness initiatives to better understand this relationship.

Age is negatively associated with physical activity, though the effect size is small (coefficient = -0.002, p < 0.001). This finding aligns with existing evidence that physical activity tends to decline as individuals age, potentially due to changes in physical capabilities, lifestyle priorities, or health status. Targeted strategies aimed at older populations may be necessary to mitigate this decline and promote sustained engagement in physical activity.

Self-assessed general health emerges as a significant predictor of physical activity, with a notable negative coefficient (-0.066, p < 0.001). Individuals reporting poorer general health are less likely to engage in physical activity, suggesting that perceived health limitations may act as a barrier. This underscores the importance of interventions that address self-perception and provide accessible exercise options tailored to varying health conditions.

Finally, measures of physical and mental health also show small negative associations with physical activity. The coefficients for physical health (-0.005, p < 0.001) and mental health (-0.002, p < 0.001) indicate that individuals experiencing more frequent poor physical or mental health days are less likely to engage in physical activity. These findings highlight the interconnected nature of mental well-being, physical health, and exercise behavior, suggesting that promoting mental and physical health may indirectly encourage greater physical activity.

In summary, the regression results indicate that employment status, general health, and, to a lesser extent, age, physical health, and mental health, are significant predictors of physical activity. Income, in contrast, shows no meaningful direct effect. These findings suggest a need for targeted interventions addressing the unique barriers faced by employed individuals, those with poor self-perceived health, and aging populations. Further research, including interaction effects and qualitative studies, could provide deeper insights into the complex relationships between these variables and physical activity levels.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Regression Results** | | | | |
|  | | | | |
|  | Dependent variable: | | | |
|  |  | | | |
|  | Physical Activity Participation | | | |
|  | (1) | (2) | (3) | (4) |
|  | | | | |
| Income | 0.00000\*\*\* | 0.00000\*\*\* | 0.00000\*\*\* | 0.00000\*\*\* |
|  | (0.000) | (0.000) | (0.000) | (0.000) |
|  |  |  |  |  |
| Employment |  | 0.036\*\*\* | 0.001 | -0.036\*\*\* |
|  |  | (0.0005) | (0.001) | (0.001) |
|  |  |  |  |  |
| Age |  |  | -0.002\*\*\* | -0.002\*\*\* |
|  |  |  | (0.00003) | (0.00003) |
|  |  |  |  |  |
| General Health |  |  |  | -0.066\*\*\* |
|  |  |  |  | (0.0004) |
|  |  |  |  |  |
| Physical Health |  |  |  | -0.005\*\*\* |
|  |  |  |  | (0.0001) |
|  |  |  |  |  |
| Mental Health |  |  |  | -0.002\*\*\* |
|  |  |  |  | (0.0001) |
|  |  |  |  |  |
| Constant | 0.611\*\*\* | 0.602\*\*\* | 0.731\*\*\* | 0.970\*\*\* |
|  | (0.0004) | (0.0004) | (0.002) | (0.002) |
|  |  |  |  |  |
|  | | | | |
| Observations | 3,841,020 | 3,841,020 | 1,107,376 | 1,107,364 |
| R2 | 0.041 | 0.042 | 0.047 | 0.099 |
| Adjusted R2 | 0.041 | 0.042 | 0.047 | 0.099 |
| Residual Std. Error | 0.424 (df = 3841018) | 0.424 (df = 3841017) | 0.426 (df = 1107372) | 0.415 (df = 1107357) |
| F Statistic | 162,388.500\*\*\* (df = 1; 3841018) | 84,379.590\*\*\* (df = 2; 3841017) | 18,356.130\*\*\* (df = 3; 1107372) | 20,197.280\*\*\* (df = 6; 1107357) |
|  | | | | |
| Note: | \*p<0.1; \*\*p<0.05; \*\*\*p<0.01 | | | |

**ROBUSTNESS CHECK**

A **robustness check** is used to assess the reliability and consistency of your empirical results under different conditions or assumptions. For this dataset and study on the impact of minimum wage and income on physical activity, robustness check was checked using different subsets of the data (Age <50 and >50). Both models assess the relationship between Physical activity (exer) and two independent variables: income (income) and employment status (minimum wage).

Robustness Check: Regression Models by Age Group

================================================================================

Dependent variable:

------------------------------------------------------------

Physical Activity Participation

Age < 50 Age >= 50

(1) (2)

--------------------------------------------------------------------------------

Income 0.00000\*\*\* 0.00000\*\*\*

(0.000) (0.000)

Employed 0.018\*\*\* 0.012\*\*\*

(0.001) (0.001)

Constant 0.663\*\*\* 0.576\*\*\*

(0.001) (0.001)

--------------------------------------------------------------------------------

Observations 405,886 701,490

R2 0.032 0.043

Adjusted R2 0.032 0.043

Residual Std. Error 0.401 (df = 405883) 0.441 (df = 701487)

F Statistic 6,684.137\*\*\* (df = 2; 405883) 15,812.590\*\*\* (df = 2; 701487)

================================================================================

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

The robustness check demonstrates that the relationship between income and exercise is consistent across both age groups (under 50 and 50 and above), indicating that income has a small but statistically significant positive effect on exercise behavior in both subsets. However, the effect of employment status is not robust, as it fails to show statistical significance in both groups, suggesting that employment has little to no impact on exercise behavior in either age group.

The models have low R-squared values (around 4%), which means they explain only a small portion of the variation in exercise behavior. This suggests that while income may play a role, other unmeasured factors likely contribute to the variation in exercise that are not captured by these models.

In conclusion, the robustness of the data shows that the relationship between income and exercise holds across different age groups, but the impact of employment status is not robust. Furthermore, the low R-squared values suggest that the models are not fully capturing the factors that influence exercise behavior, highlighting the need for additional variables or more complex modeling to better understand the drivers of exercise behavior in these populations.

**DISCUSSION**

The discussion of this study’s results highlights their significance and positions them within the broader literature on physical activity, socioeconomic factors, and public health. The findings suggest that while employment status, age, general health, and measures of physical and mental health significantly influence physical activity, income does not show a direct effect. This contrasts with some earlier studies that emphasized the role of income in facilitating access to health-promoting resources such as gym memberships and recreational facilities. However, the lack of a direct relationship here aligns with theoretical insights from behavioral economics, which suggest that the interaction between income and health behaviors is more nuanced, influenced by financial stress and present bias rather than sheer income levels.

The negative association between employment status and physical activity is particularly noteworthy. While employed individuals are generally presumed to have greater access to health resources, this result suggests that time constraints and work-related stress may act as barriers. This corroborates findings in the literature emphasizing the importance of workplace wellness programs and flexible work schedules in promoting physical activity among working populations. The negative relationship between age and physical activity aligns with previous research, reflecting well-documented declines in physical activity due to health-related limitations and lifestyle changes. Similarly, the significant association between self-assessed general health and physical activity highlights the importance of subjective health perceptions in shaping exercise behaviors. Poorer perceived health can act as a barrier, underscoring the need for tailored interventions that consider individual health challenges.

The findings on physical and mental health further emphasize the interconnected nature of these variables with physical activity. Increased days of poor health, whether physical or mental, were associated with reduced physical activity levels. This underscores the importance of integrated health interventions that simultaneously address physical health, mental well-being, and lifestyle behaviors. Promoting mental and physical health could indirectly encourage greater physical activity, a perspective supported by evidence linking better mental health to improved engagement in exercise.

The robustness checks reveal consistency in the relationship between income and physical activity across age groups but show that employment status has minimal impact. These findings reinforce the conclusion that income may play a limited role in influencing exercise directly and highlight the complexity of employment’s effects, which may vary based on job type, hours, and other work-related factors. The low R-squared values across all models suggest that unmeasured factors—such as cultural norms, environmental influences, or policy interventions—likely play significant roles in shaping physical activity behaviors.

While these results align with some prior findings, they also raise important questions for future research. The absence of a direct income effect, despite strong theoretical expectations, suggests the need to explore non-linear relationships or interactions with other variables, such as geographic factors or access to recreational resources. The small effect sizes of other predictors, such as age and health measures, further indicate the presence of unmeasured mediators or moderators influencing physical activity.

This study is not without limitations. First, the reliance on self-reported measures for health and physical activity introduces potential bias, as individuals’ perceptions may not accurately reflect their actual behaviors or health status. Second, the cross-sectional nature of the data limits causal interpretations. Longitudinal studies would be better suited to uncover causal pathways and dynamic interactions over time. Third, the dataset’s geographic and demographic scope may limit the generalizability of findings to populations with different socioeconomic or cultural contexts. Future research should also consider incorporating environmental and policy-level variables, such as access to green spaces or public health campaigns, to provide a more comprehensive understanding of the factors influencing physical activity.

**CONCLUSION**

The conclusion brings the study full circle by addressing the research gap identified in the introduction and highlighting the policy and practical implications of the findings. The study aimed to explore the socioeconomic and health determinants of physical activity, particularly focusing on the nuanced roles of income, employment, and health status. The results indicate that income, contrary to some prior assumptions, does not directly influence physical activity. Instead, factors such as employment status, age, and perceived health emerged as more significant predictors, emphasizing the complexity of the determinants of exercise behavior.

If different or more comprehensive data were available, such as longitudinal datasets or more detailed geographic and environmental variables, the analysis could uncover causal relationships or context-specific effects. For instance, incorporating data on access to recreational facilities, public spaces, or community-level health programs could provide a deeper understanding of the environmental and policy influences on physical activity.

The findings raise several new questions warranting further investigation. Why does income lack a direct effect on physical activity in this context, and could it have a stronger influence when mediated by access to health-promoting resources? What specific aspects of employment—such as job type, flexibility, or working hours—play the most critical roles in determining physical activity levels? How do cultural norms and social support interact with socioeconomic and health factors to shape exercise behaviors?

The results aligned partially with expectations; age and health perceptions correlated with physical activity, consistent with prior literature. However, the lack of a direct relationship between income and physical activity was unexpected and suggests that simplistic assumptions about financial means translating into healthier behaviors may overlook critical intervening variables, such as stress or opportunity costs.

Policy implications from these findings emphasize the need for targeted interventions to reduce barriers to physical activity, particularly for employed individuals and those with poor health. Workplace wellness programs, community-based initiatives, and accessible exercise opportunities could address these challenges. For aging populations, tailored physical activity programs considering health limitations and preferences are essential.

In conclusion, while this study sheds light on key predictors of physical activity, it underscores the need for more nuanced analyses and broader data to capture the multifaceted nature of exercise behaviors. Future research should explore the roles of environmental factors, policy interventions, and cultural contexts in shaping physical activity. By addressing these remaining gaps, researchers can better inform public health strategies and promote physical activity across diverse populations.