

1. Provide a short summary of your project you want to undertake and what sort of analysis you want to perform. This won't lock you in if you realize there's an issue later, but it's good to have a starting plan.

I would like to explore the relationship between menopausal stages and psychological distress among adult women in the United States. The dataset will be from the 2024 National Health Interview Survey (NHIS) Adult, which includes variables related to reproductive health, mental health symptoms, demographic factors, and overall well-being.

The main goal of the project is to determine whether perimenopausal and postmenopausal women report higher levels of distress compared to premenopausal women. I plan to clean and subset the dataset to include women ages 35 – 65, identify their menopausal stage using available reproductive health variables, and compare those groups on their mental health indicators.

The analysis will include a decision tree or linear regression model to predict “high psychological distress” using menopausal stage as one of the key predictors. I also plan to compare this with the reference model scikit learn to verify my results.

2. Provide some background into your area of interest and explain why this problem needs solving.

The effects of perimenopause and menopause is an under researched area of public health. We know that the changes in estrogen and progesterone during this time influence neurotransmitter pathways that regulate mood, including serotonin and dopamine. Because of this, many women experience a range of symptoms, including anxiety, depression, irritability, and emotional instability during this period.

At the same time, women’s mental health during this time is often overlooked or misdiagnosed. Symptoms like low mood or anxiety during midlife are attributed to stress, aging, or chronic illness instead of a hormonal transition. Understanding how psychological distress varies by menopausal stage at population level could help bring more attention to the mental health needs of midlife women.

Mental health during perimenopause and menopause is understudied, even though millions of women experience mood changes during this transition. Using NHIS data, this project explores how reproductive stage relates to mental-health symptoms and medication patterns. By identifying population-level trends, we can bring clarity to an area of women’s health that is too often dismissed or misunderstood.

3. Detail how your approach will contribute to addressing this issue.

By analyzing NHIS data, this project can identify patterns in psychological distress across menopausal stages without relying on clinical samples. sing logistic regression or a

decision-tree model allows me to examine which factors contribute most strongly to distress and whether menopausal stage remains a significant predictor even after accounting for variables such as age, income, education, chronic health conditions, and physical activity.

This approach can help highlight whether perimenopausal women are particularly vulnerable to elevated distress and whether this pattern persists postmenopause. If the analysis shows strong relationships, it supports the idea that menopause-related mental health should be addressed more proactively in public health research and clinical care.

4. Provide a list of citations and LLM prompts used. Please be advised that LLMS should have no part in the writing of this report, just research.

https://www.cdc.gov/nchs/nhis/documentation/2024-nhis.html#cdc_data_surveillance_section_7-sample-adult-interview

<https://pubmed.ncbi.nlm.nih.gov/24951102/>

https://scikit-learn.org/stable/modules/linear_model.html#logistic-regression

LLM Prompts:

>What reference model can I use to verify a decision tree or linear regression analysis?