Is Gender a Significant Factor in the Increase of the Number of Heterosexual Partners in the Last Decade?

**The survey**

The surveys chosen for this research question is a survey conducted between 1999 and 2001 and between 2010 and 2012 under the name NSSAL-2 and NSSAL-3, respectively. It stands for the *British National Surveys of Sexual Attitudes and Lifestyles* and it is one of the largest and most detailed studies of sexual attitudes and behaviours in the world. There is also NSSAL-1 that was conducted in 1990, but the question we are going to examine here has been different in that survey so data from 1990 survey was not included in the analysis. The original survey findings were used throughout the 1990s by academics and policymakers to estimate the prevalence and distribution of high-risk behaviours for HIV and sexually transmitted disease in the population and to help plan the provision of sexual health services.

NSSAL-2 sample consisted of 12735 adults aged 16-74. NSSAL-3 sample consists of 15162 adults aged 16-74 from the UK. The surveys consisted of face-to-face interviews, using Computer-Assisted Personal Interviewing (CAPI), and a self-completion part, using Computer-Assisted Self Interviewing (CASI). Face-to-face interviews were covering topics on general health, learning about sex, first sexual experience etc. The self-completion part was mainly used to probe intimate questions regarding different types of sexual practices, sex in the last 4 weeks and condom use, homosexual experience etc. As can be seen, questions in the survey probed various aspects of individual’s sexual behaviour ranging from questions about whether their sexual orientation, number of homosexual and heterosexual partners, age of the first sexual experience and similar.

**Research question**

The research question is: "**Is Gender a Significant Factor in the Increase of the Number of Heterosexual Partners in the Last Decade?**"

Requirements:

* Try to explain why this research questions might be of interest and to whom.

On the side of sociology, gender influences human options, conditions, and experiences. Legal, political, economic, cultural and kinship systems are all profoundly gendered. Deep understanding of gender patterns, dynamics and biases can enhance the accuracy and scope of work in many fields. Gender awareness benefits individuals, communities and organizations.

On the medical side Sexual partner concurrency plays an important role in HIV and STD transmission. Understanding how gender dynamics influence partner choices could inform public health campaigns and initiatives related to sexual education, safe sex practices, and the prevention of sexually transmitted infections.

Understanding how gender influences partner selection and relationship dynamics can contribute to more effective healthcare strategies and policies addressing sexual and reproductive health and also can clarifies the evolving gender dynamics, promoting understanding of modern relationships, and guiding decisions at both personal and societal levels.

* Locate relevant data file and accompanying documentation in the “stata” folder.
* Perform any necessary transformations of data format type.
* Identify variables of interest and specify their type.

**NSSAL\_2000**

* Pos. = 1 Variable = sserial Variable label = full serial number

This variable is numeric, the SPSS measurement level is SCALE

Value label information for sserial

* Pos. = 8 Variable = dateyoi Variable label = year of interview

This variable is numeric, the SPSS measurement level is SCALE

Value label information for dateyoi

* Pos. = 10 Variable = rsex Variable label = respondent's sex

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for rsex

Value = 1.0 Label = male

Value = 2.0 Label = female

* Pos. = 154 Variable = attscale Variable label = sexual attraction

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for attscale

Value = 1.0 Label = opposite sex only

Value = 2.0 Label = more often opposite sex, and at least once same sex

Value = 3.0 Label = about equally often to opposite sex and same sex

Value = 4.0 Label = more often same sex, and at least once opposite sex

Value = 5.0 Label = same sex only

Value = 6.0 Label = never felt sexually attracted to anyone at all

Value = 7.0 Label = refused

Value = 9.0 Label = not answered

* Pos. = 202 Variable = hetlife Variable label = no. of het. sex partners, life

This variable is numeric, the SPSS measurement level is SCALE

Value label information for hetlife

Value = 9995.0 Label = at least one

Value = 9996.0 Label = at least two

Value = 9997.0 Label = at least three

Value = -1.0 Label = not applicable

Value = 9999.0 Label = not answered

* Pos. = 204 Variable = het5yrs Variable label = no. of het. sex partners, last 5 years

This variable is numeric, the SPSS measurement level is SCALE

Value label information for het5yrs

Value = 9995.0 Label = at least one

Value = 9996.0 Label = at least two

Value = 9997.0 Label = at least three

Value = -1.0 Label = not applicable

Value = 9999.0 Label = not answered

* Pos. = 205 Variable = het1yr Variable label = no. of het. sex partners, last year

This variable is numeric, the SPSS measurement level is SCALE

Value label information for het1yr

Value = 995.0 Label = at least one

Value = 996.0 Label = at least two

Value = 997.0 Label = at least three

Value = -1.0 Label = not applicable

Value = 999.0 Label = not answered

**NSSAL\_2010**

* Pos. = 1 Variable = SIN2 Variable label = Scrambled serial number

This variable is numeric, the SPSS measurement level is SCALE

Value label information for SIN2

* Pos. = 2 Variable = wdateyoi Variable label = Year of interview

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for wdateyoi

* Pos. = 5 Variable = wRSex Variable label = Respondent's sex

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for wRSex

Value = 1.0 Label = Male

Value = 2.0 Label = Female

Value = -1.0 Label = Item not applicable

Value = -9.0 Label = Refusal

Value = -8.0 Label = Don't Know

Value = -2.0 Label = Schedule not applicable

* Pos. = 356 Variable = attscale Variable label = Sexual attraction

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for attscale

Value = 1.0 Label = Opposite sex only

Value = 2.0 Label = More often opposite sex, and at least once same sex

Value = 3.0 Label = About equally often to opposite sex and same sex

Value = 4.0 Label = More often same sex, and at least once opposite sex

Value = 5.0 Label = Same sex only

Value = 6.0 Label = Never felt sexually attracted to anyone at all

Value = 7.0 Label = Refused

Value = 9.0 Label = Not answered

* Pos. = 109 Variable = hetlife Variable label = Number of heterosexual sex partners, life

This variable is numeric, the SPSS measurement level is SCALE

Value label information for wHetLife

Value = -1.0 Label = Item not applicable

Value = -9.0 Label = Refusal

* Pos. = 111 Variable = het5yrs Variable label = Number of heterosexual sex partners, last 5 years

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for wHet5Yrs

Value = -1.0 Label = Item not applicable

* Pos. = 112 Variable = het1yr Variable label = Number of heterosexual sex partners, last year

This variable is numeric, the SPSS measurement level is NOMINAL

Value label information for wHet1Yr

Value = -1.0 Label = Item not applicable

* In case a variable is ordinal or nominal, list the possible categories and their meaning. (listed above)
* Define the analysis required to answer the RQ and explain why you have chosen this particular analysis.

Linear Regression

We can assess the relationship between our dependent variable (The Number of Heterosexual Partners in the Last Decade) and the independent variable (Gender). Linear regression provides a quantitative way to measure the strength and direction of the relationship between the Number of Heterosexual Partners and the Gender.

* List all assumptions and how you addressed them.

1. Linearity
2. Homoscedasticity
3. Normality of Residuals
4. Independence of Residuals

* Dependent Variable: Number of Heterosexual Partners in the Last Decade
* Independent Variables: Gender.
* List the null and alternative hypothesis.
* Null Hypothesis (H0): Gender has no significant effect on the increase in the number of heterosexual partners in the last decade.

H0: β1 (coefficient for Gender) = 0

* Alternative Hypothesis (H1): Gender has a significant effect on the increase in the number of heterosexual partners in the last decade.

H1: β1 (coefficient for Gender) ≠ 0

* Provide the results of your analysis and answer the research question.
* Try to explain the impact of the result within the context of the data.

Instructions:

* Submit a PDF or .doc file with your answers and necessary plots.
* Source code [R, Python] or output file [SPSS] should be submitted as separate files.
* List any relevant references you used in a format of your choice.

**You are not allowed to publicly share the dataset(s) provided for this task.**