



DOCUMENTARY ON NICPS 1987

SMDM PRESENTATION - GROUP 6

OVER POPULATION

WHAT'S THE SOLUTION?



THANOS

SNAP (Kill half the
population)



EARTHLINGS

FAMILY PLANNING
(Control the population)



INDONESIA

National Contraceptive Prevalence Survey 1987

- ❖ Missing Attribute Values : None
- ❖ Number of Instances : 1473
- ❖ Number of Attributes : 10
 - Numerical - 2
 - Categorical – 8

Data Description

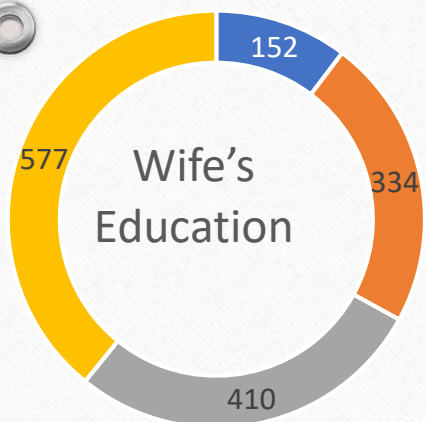
This dataset is a subset of the 1987 National Indonesia Contraceptive Prevalence Survey. The samples are married women who were either not pregnant or do not know if they were at the time of interview.

Problem Statement

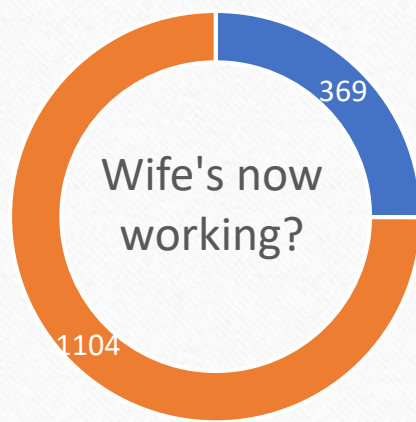
The problem is to predict the current contraceptive method choice (no use, long-term methods, or short-term methods) of a woman based on her demographic and socio-economic characteristics.

Attribute Information

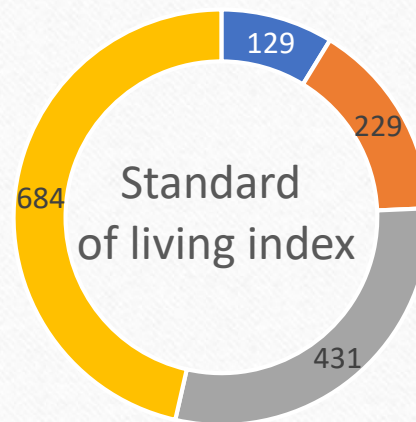
1. Wife's age (numerical)
2. Wife's education (categorical) 1=low, 2, 3, 4=high
3. Husband's education (categorical) 1=low, 2, 3, 4=high
4. Number of children ever born (numerical)
5. Wife's religion (binary) 0=Non-Islam, 1=Islam
6. Wife's now working? (binary) 0=Yes, 1=No
7. Husband's occupation (categorical) 1, 2, 3, 4
8. Standard-of-living index (categorical) 1=low, 2, 3, 4=high
9. Media exposure (binary) 0=Good, 1=Not good
10. Contraceptive method used
 - 1=No-use
 - 2=Long-term
 - 3=Short-term ³



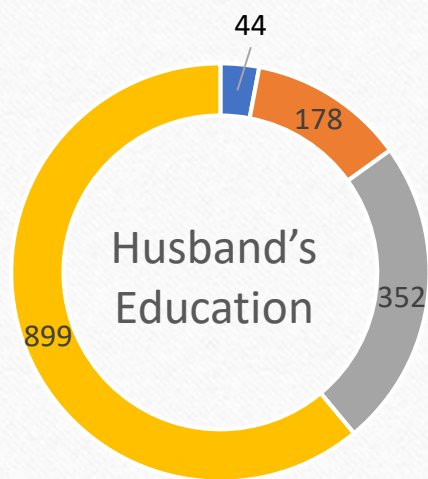
■ 1 ■ 2 ■ 3 ■ 4



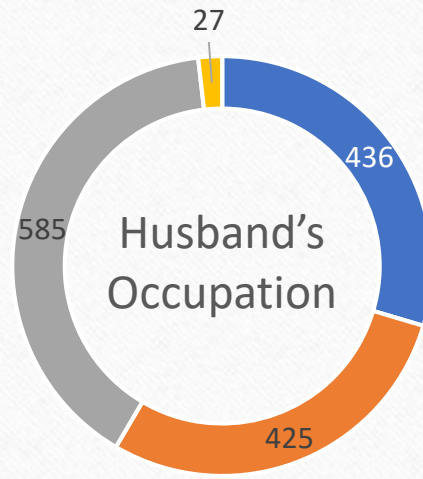
■ 0 ■ 1



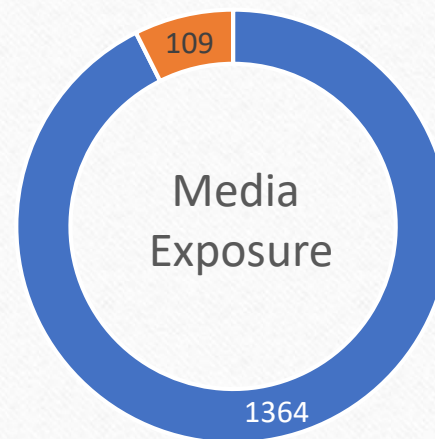
■ 1 ■ 2 ■ 3 ■ 4



■ 1 ■ 2 ■ 3 ■ 4



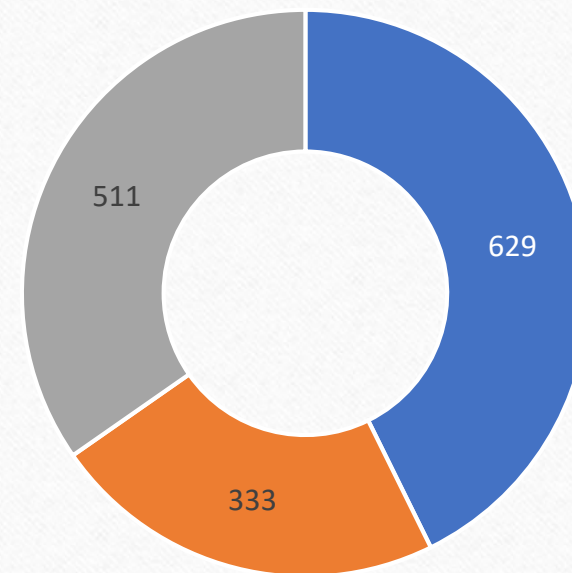
■ 1 ■ 2 ■ 3 ■ 4



■ 0 ■ 1

CATEGORICAL DATA

Contraceptive method



■ 1 ■ 2 ■ 3

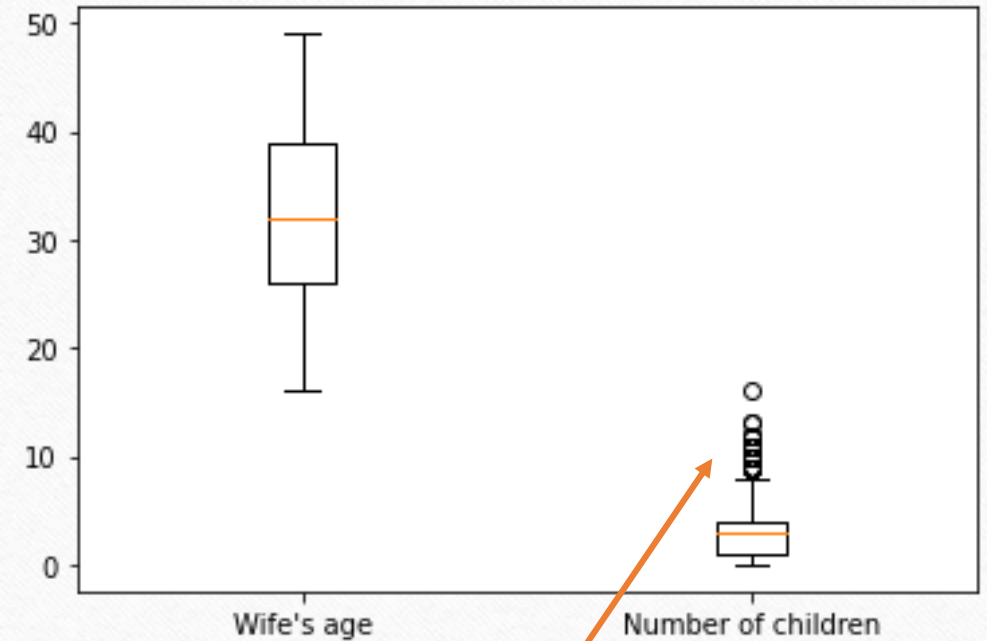
NUMERICAL DATA

Wife's age	Number of children
30	2
23	1
25	3
42	6
29	3
33	2
33	3
39	8
33	4
17	1

Mean Value

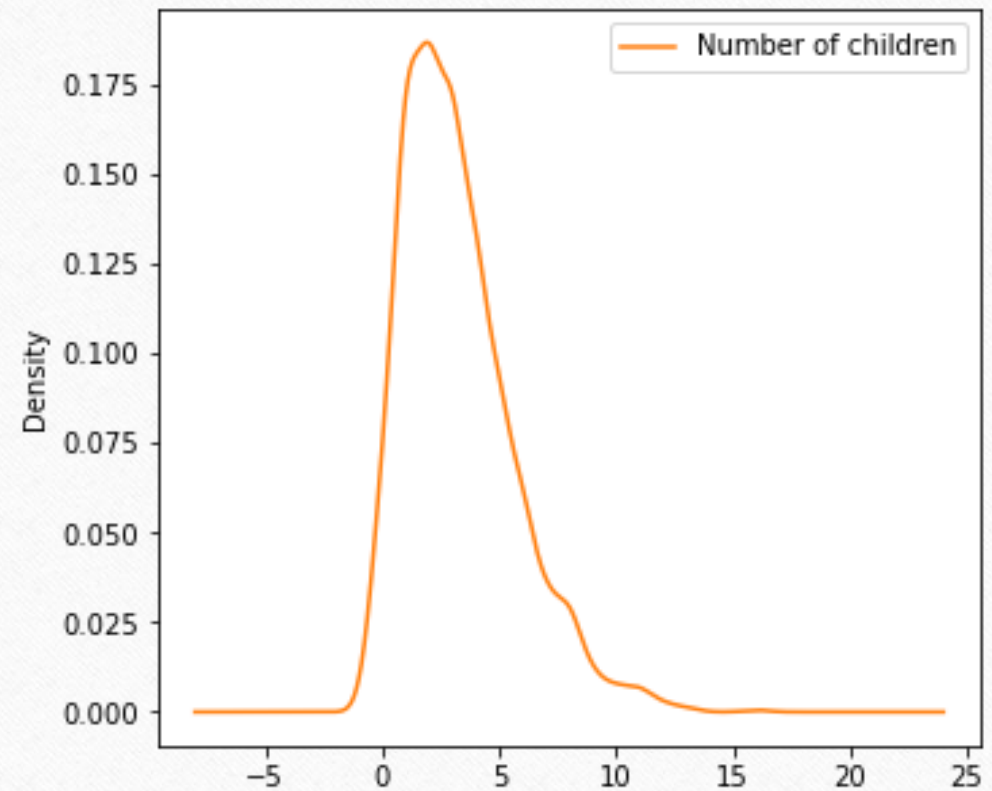
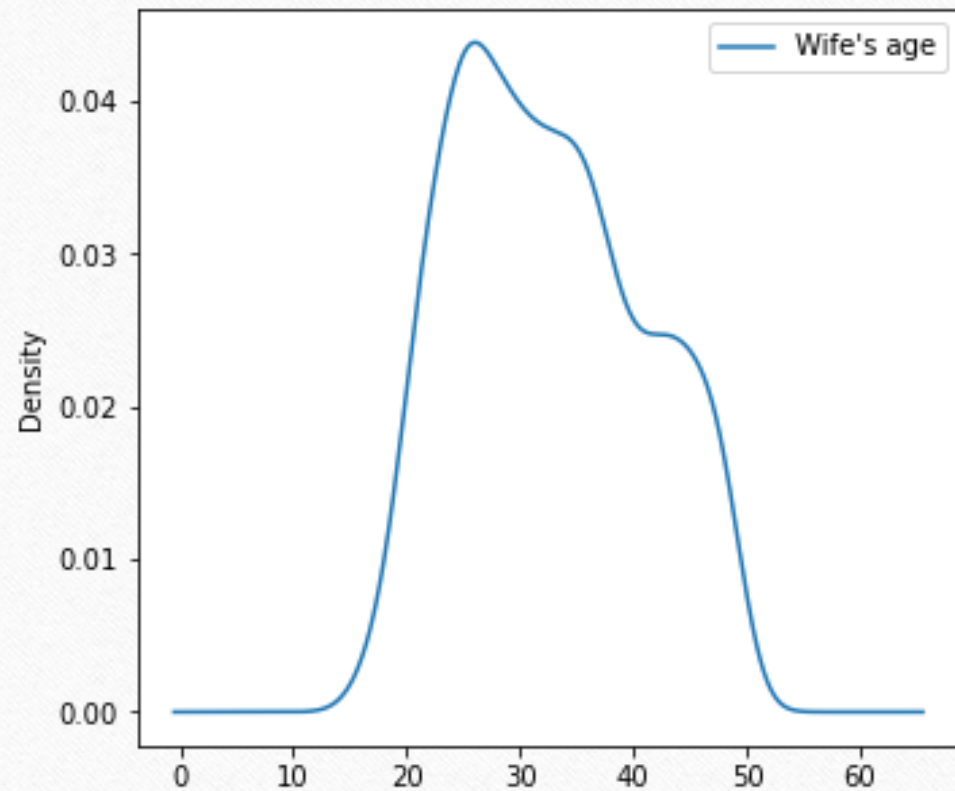
Numerical Summary:

	Wife's age	Number of children
count	1473.000000	1473.000000
mean	32.538357	3.261371
std	8.227245	2.358549
min	16.000000	0.000000
25%	26.000000	1.000000
50%	32.000000	3.000000
75%	39.000000	4.000000
max	49.000000	16.000000



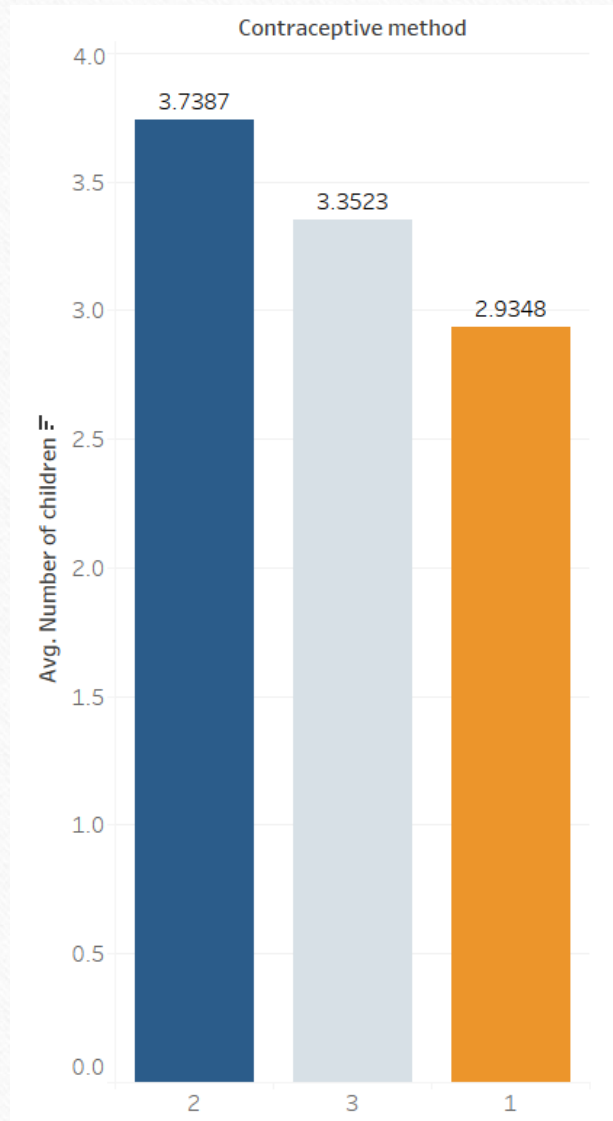
Extreme values present in Number of Children

NUMERICAL DATA DISTRIBUTION



SURVEY CLAIMS

- Average number of children is the same for each type of Contraceptive Method.
- The contraceptive method used by women of different ages is the same.
- 40% of the women do not use any contraceptive method, 35% of the women uses long term contraceptive methods and the rest use short term contraceptive methods
- Women currently working have more number of children than women currently not working.
- Couples with higher standard of living have more children than couples with lower standard of living.
- Contraceptive method used depends on media exposure of the couple.



Claim1:

Average number of children is the same for each type of Contraceptive Method.

Test Used: Kruskal Wallis H

Test Result: P_value for Kruskal : $1.1082473929391914e-13$
Significance level: 0.05

Inference: Here P value is less than alpha thus we reject null hypothesis. And we conclude that average number of children is different for at least one type of Contraceptive Method.

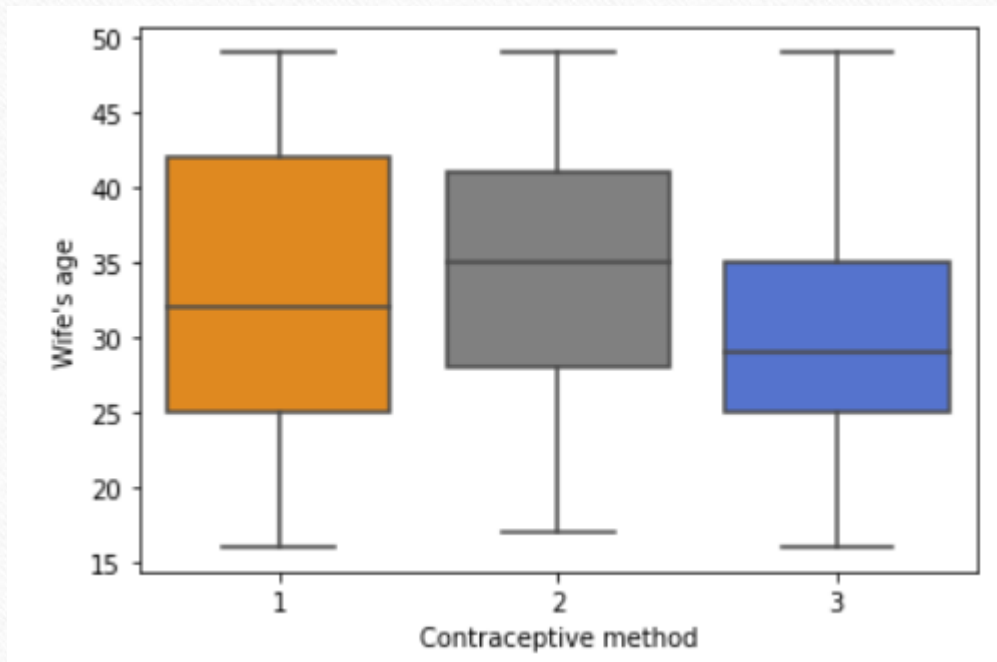
Number of children with different Contraceptive Method:

	1	2	3
1	1.000000e+00	2.205238e-13	7.985241e-08
2	2.205238e-13	1.000000e+00	1.046469e-02
3	7.985241e-08	1.046469e-02	1.000000e+00

The P value for pairs: no use(1) and long term(2) & no use(1) and short term(3) long term(2) and no use(1) & long term(2) and short term(3), short term(3) and no use(1) & short term(3) and long term(2), is less than significance level.₈

Claim2:

The contraceptive method used by women of different ages is the same.



Test Used: Oneway ANOVA

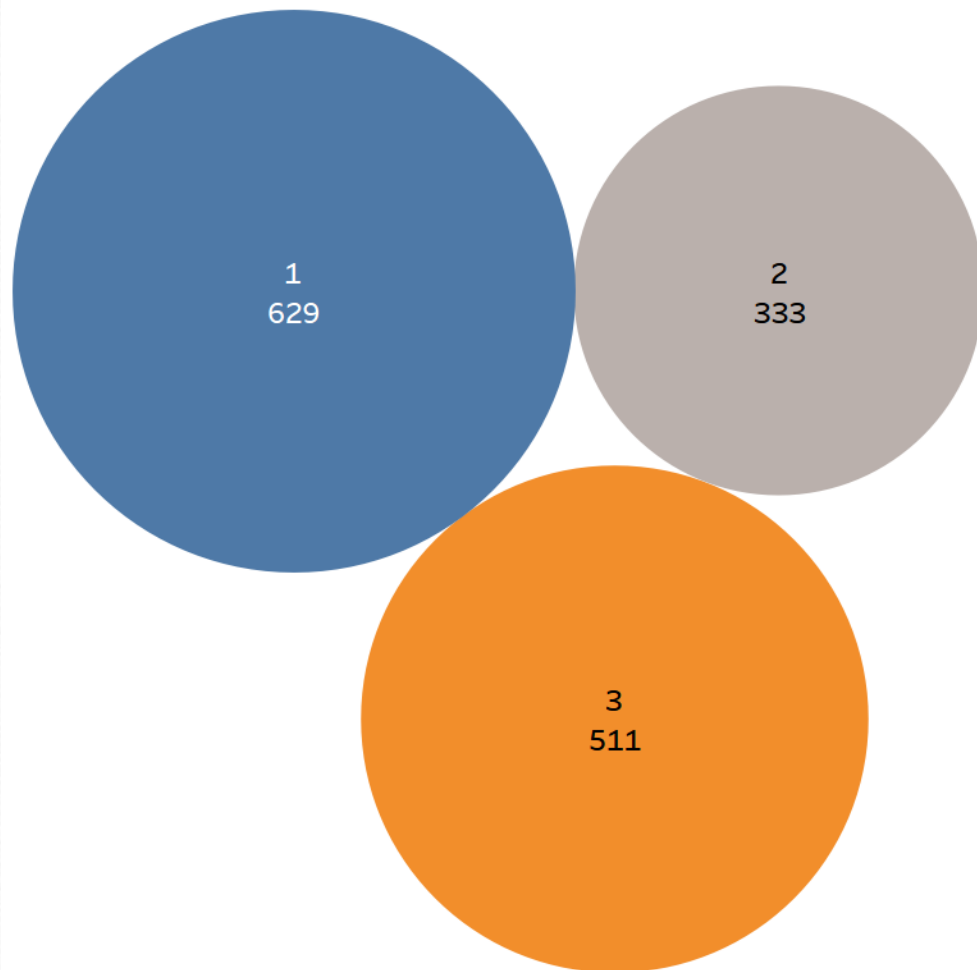
Test Result: p_val

0.010301352958241352

Identify the Number of children with different Contraceptive Method

group1	group2	meandiff	p-adj	lower	upper	reject
1	2	0.9599	0.1841	-0.3204	2.2402	False
1	3	-3.1799	0.001	-4.305	-2.0548	True
2	3	-4.1398	0.001	-5.4703	-2.8093	True

Inference: The reject=False for pair (1,2) denotes that we fail to reject the null hypothesis; and conclude that the mean age of women using type 1 and type 2 of contraceptive method is the same. For the pairs (1,3) and (2,3) the mean age of women is not the same.



Claim3:

40% of the women do not use any contraceptive method, 35% of the women uses long term contraceptive methods and the rest use short term contraceptive methods

Test Used: Chisquare Goodness of fit test

Test Result:

Chi-stat: 6.102851323828913
p-value: 0.04729145468225175

Inference: Here p-value is less than 0.05 thus we can reject the null hypothesis at 5% level of significance and conclude that the ratio of Contraceptive Method used is different from the claimed value.

Claim4:

Women currently working have more number of children than women currently not working

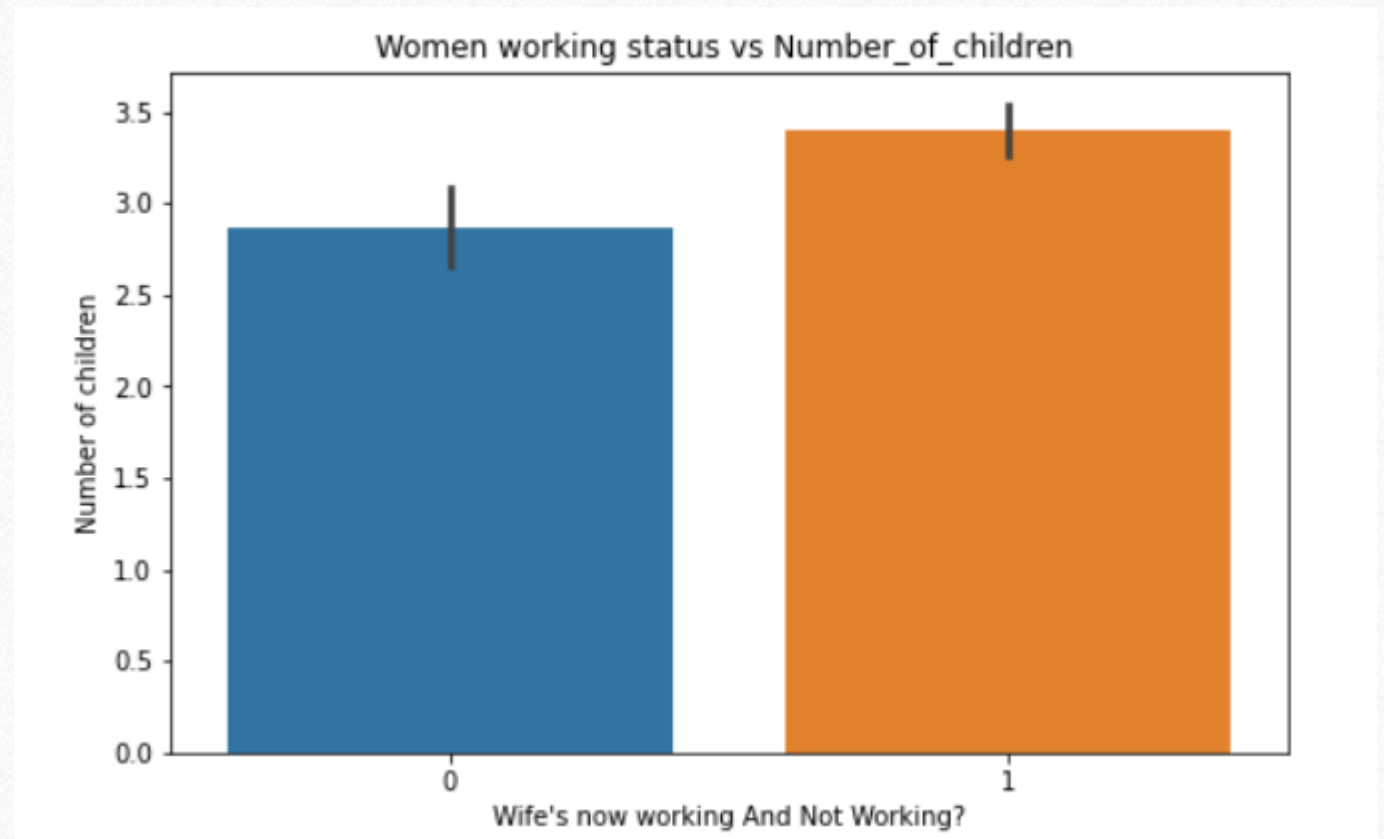
Test Used: MANN-WHITNEY U TEST

Test Result:

Test statistic: 177047.5

p-value: 0.9999302180431572

Inference: Here the p-value is greater than 0.05. Thus we fail to reject (i.e. accept) the null hypothesis and we can conclude that Women currently working have the same or less number of children than women currently not working.



Claim5:

Couples with higher standard of living have more children than couples with lower standard of living.

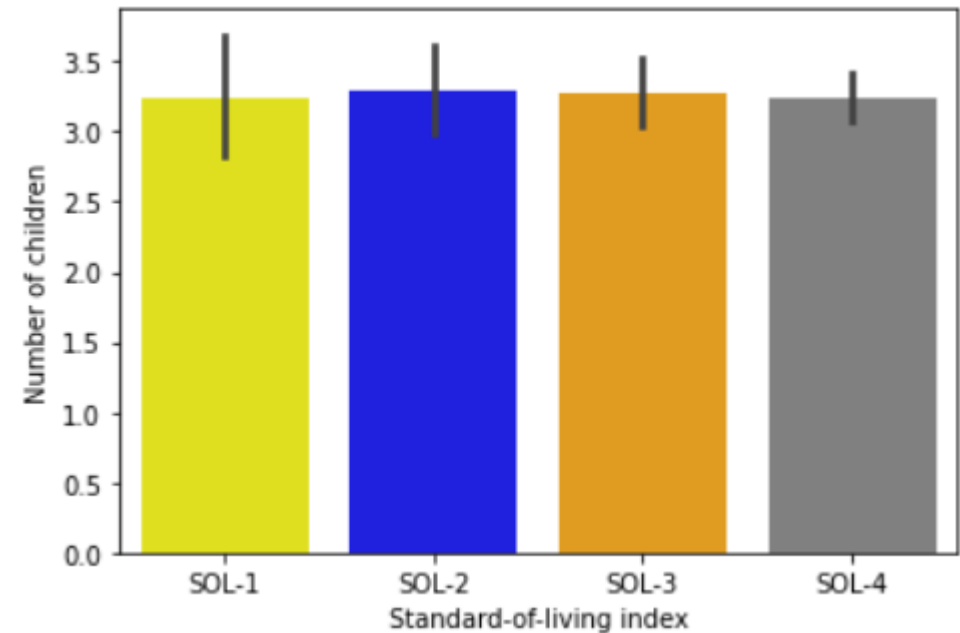
Test Used: Wilcoxon test

Test Result:

Critical value for one-tailed Z-test: 1.64
z-stat: -0.7388117954724889

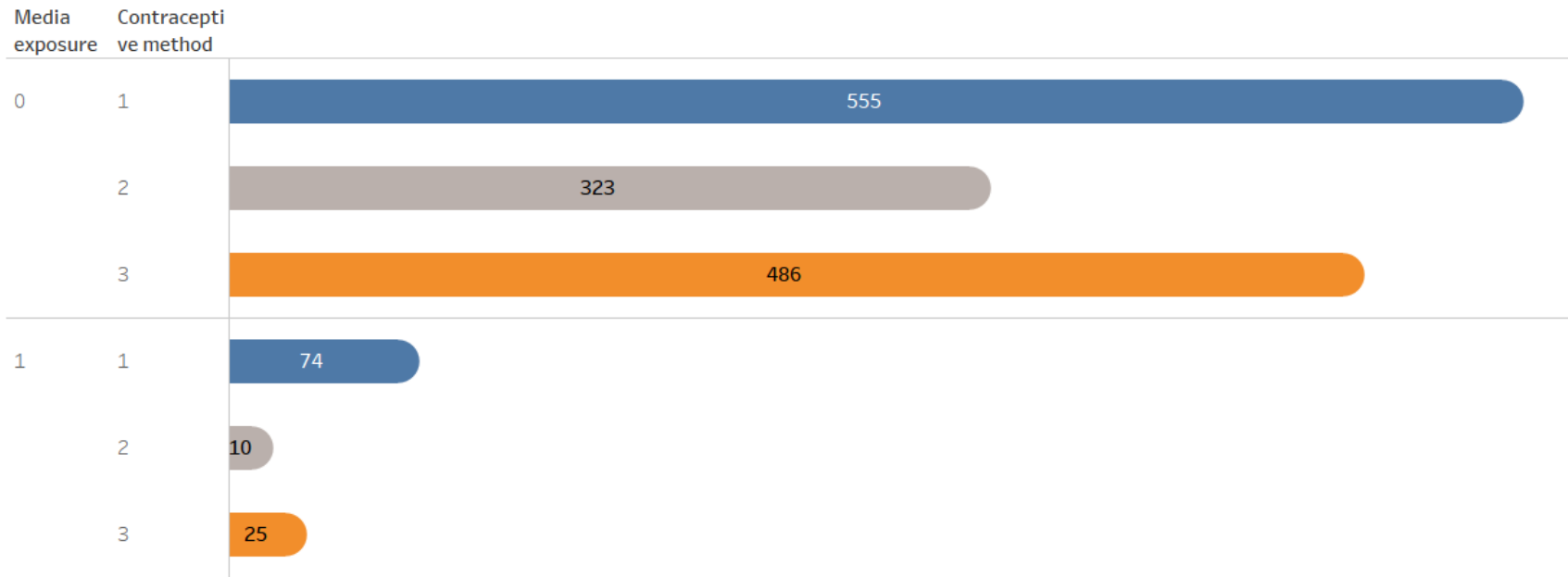
Inference:

Since test statistic is less than critical value with 95% confidence it falls in the acceptance region and so we fail to reject the null hypothesis and conclude that Couples with higher standard of living have the same or less children than couples with lower standard of living.



Claim6:

Contraceptive method used does not depend on media exposure of the couple.



Test Used:
Chi Square Test of Independence

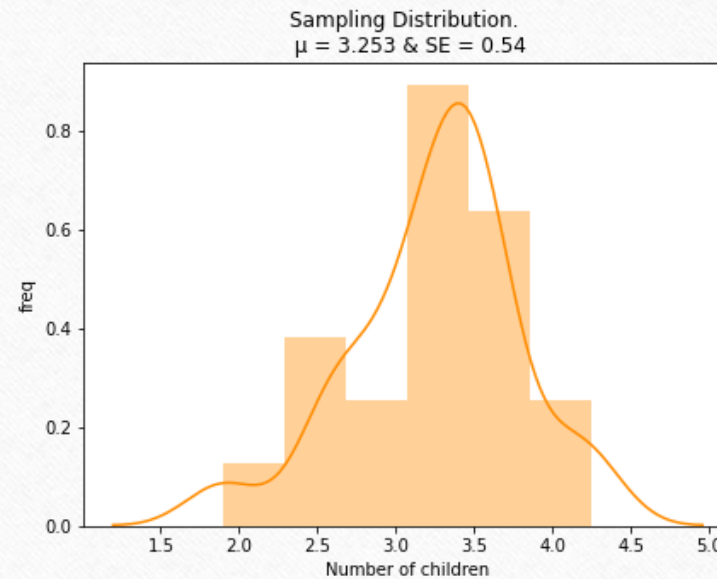
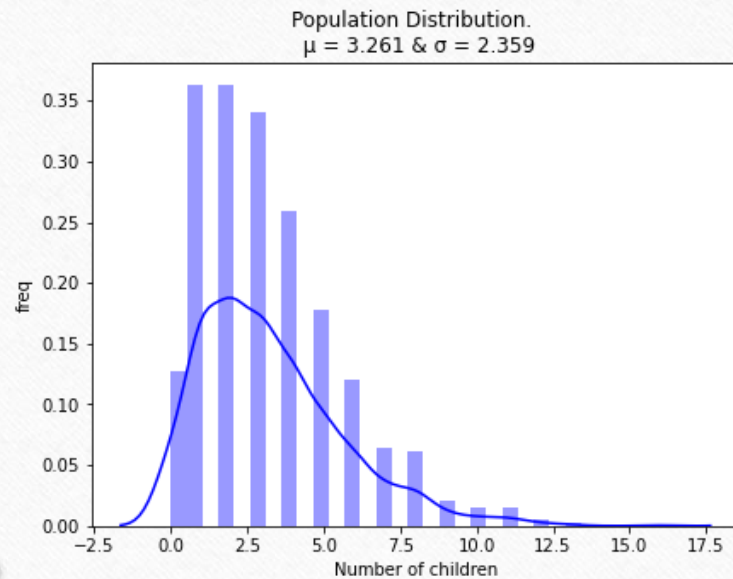
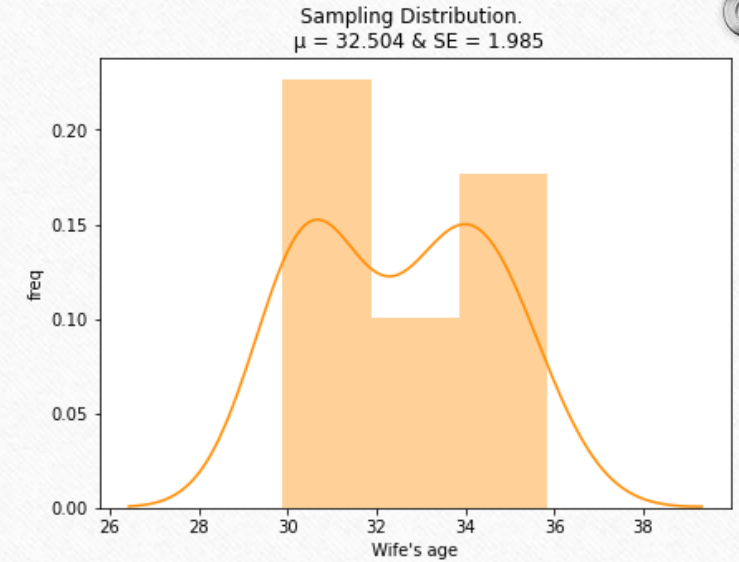
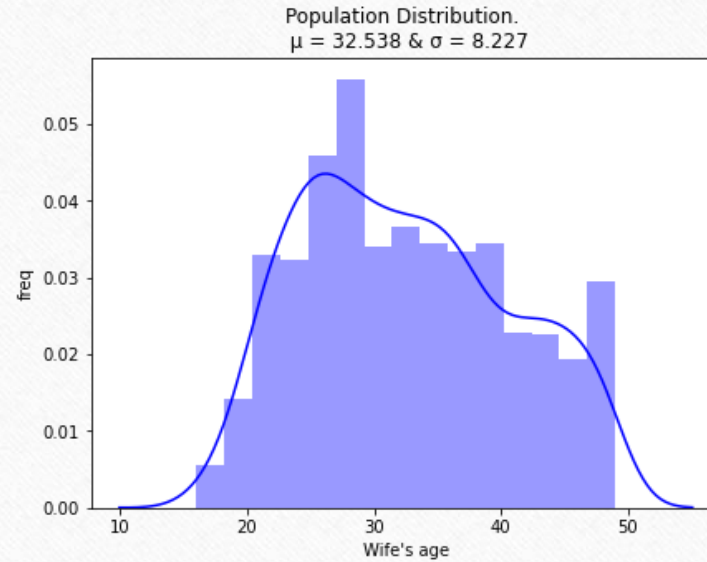
Test Result:

Chi-statistic: 31.572283442858545
P-value: 1.393688639141483e-07
Degrees of freedom: 2

Inference:

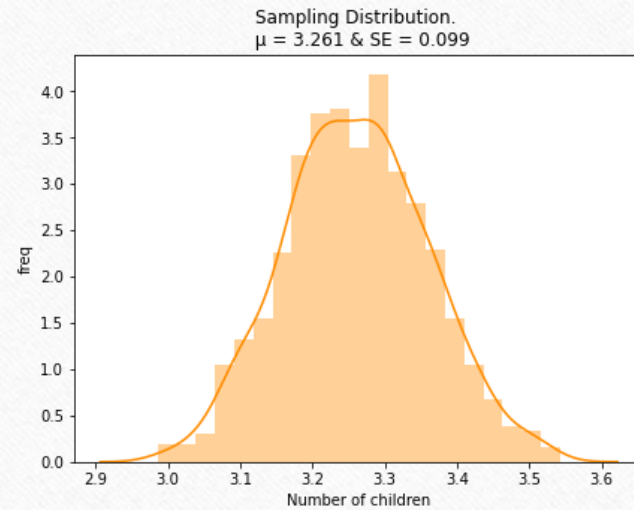
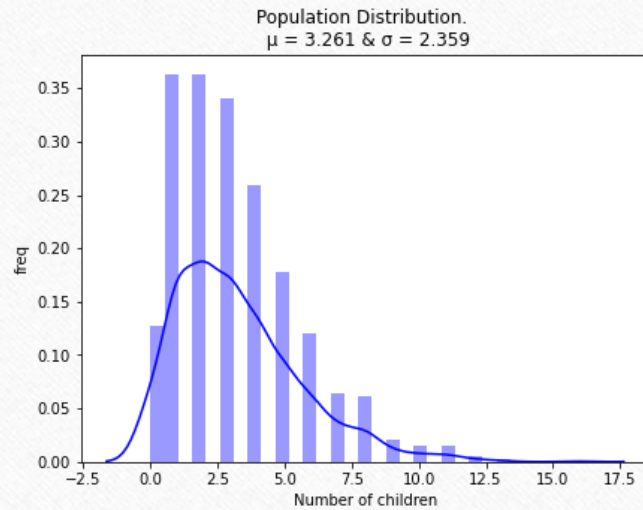
Since the p-value($1.39e^{-07}$) is less than 0.05, we reject the null hypothesis. Therefore we can say with 95% confidence that contraceptive method used depends on media exposure.

Simulation Of CLT



Sample Size = 20
No of samples = 20

Simulation Of CLT



Sample Size = 1000
No of samples = 500

