Indonesian females' birth interval disparities[*]

Evidence from the Demographic and Health Survey 1997

Yuzhi Pi, Chengle Yang, Yiran Mei

March 30th, 2022

Contents

---

[*]Code and data are available in this GitHub repository: chle1999/Indonesia-Demographic-and-Health-Survey-1997.

Abstract

An interval between consecutive pregnancies is referred to as a birth interval, which may be used to reduce maternal health complications and mortality. In order to examine disparities in the birth interval among different groups by background characteristics, we examined the results of the Indonesia Demographic and Health Survey (1997) by downloading the published data from the DHS site. As a result of the study, females at younger ages, with dead prior births, and with a higher educational level are more likely to give birth after a short interval, whereas work status and gender of prior birth have no significant effects on birth intervals. The results will provide some useful insights for the government and policymakers regarding future strategic planning.

Introduction

Birth interval is a renowned clinical term in perinatology research, referring to the time span between two pregnancies. It is also an inadequately used and poorly understood health intervention (M 2005). In general, a child conceived after an 18 to 23-month birth interval experiences the least adverse perinatal outcomes (B. P. Zhu et al. 1999).

A number of studies have concluded that short birth intervals can increase low birth weight (Adams et al. 1997), cerebral palsy (Pinto-Martin, Cnaan, and Zhao 1998), and morbidity. Pregnancy spacing of fewer than six months is more likely to expose the mother to a higher risk of maternal mortality, premature rupture of membranes, (A. Conde-Agudelo and Belizan 2000) puerperal endometritis, and other maternal health complications(Agustin Conde-Agudelo, Rosas-Bermúdez, and Kafury-Goeta 2007). It should be noted, however, that longer inter-pregnancy intervals may also cause problems, although the symptoms may be caused by other factors. Women who become pregnant after an extended period of time may be aged or infertile, or the pregnancy may be unintended. Consequently, adverse outcomes are more likely to occur(LV, SP, and RL 1998), as well as pre-eclampsia and eclampsia(Pinto-Martin, Cnaan, and Zhao 1998). In general, nutritional depletion caused by insufficient time for replenishment of vital resources may result in maternal death. According to some studies of developing countries, where women are more likely to suffer from undernutrition(B. P. Zhu et al. 1999), support this claim. In developing countries, maternal mortality is the primary cause of premature death and disabilities among women of reproductive age. The fact that most maternal deaths are preventable makes maternal mortality an issue of both human rights and women's rights.

The unique characteristics of Indonesia as one of the few democracies with a majority Muslim population and its long-standing commitment to women's rights make it an attractive candidate among developing

countries(Rinaldo 2019). The population of Indonesia was approximately 202 million(Bank 2022) in 1997, making it ranked first in Muslim majority countries and fourth in the world's population. As indicated in the Indonesia Report on Human Rights Practices for 1997, women in Indonesia are disproportionately illiterate, health-impaired, and malnourished. The President stated that more efforts must be put forth to reduce the maternal mortality rate. The rate was estimated to be 0.425% on every live birth according to official data. In some estimations from other sources, this number may reach as high as 0.625%(Bureau of Democracy and Labor 1997). A "mother-friendly movement" was launched by the government in December 1996 to combat maternal mortality, and it is constantly expanding its scope to include more regions. Taking into account the relevance between the birth interval and maternal mortality, it is pertinent to examine the birth interval in Indonesia in 1997, in order to understand the preliminary outcome of women's rights and maternity protection practices following the implementation of the program.

In this report, we analyze the birth interval results by background characteristics derived from the Indonesia Demographic and Health Survey 1997 (IDHS) conducted by the Central Bureau of Statistics and DHS (Demographic and Health Survey). We will investigate the disparities in the birth interval by age group, gender, and the survival of a prior birth, as well as by region/residence, education level and work status. The results show that shorter intervals of birth are more likely to be found in young females, those with deceased prior children, and those with higher levels of education, whereas work status and gender of prior children are not significant factors. The findings of this study are expected to assist the Indonesian government in assessing the preliminary success of the "mother-friendly movement" program and setting strategic goals for its future.

Specifically, the paper is composed of three parts: a data section that provides extensive information about the source data and methodology of the survey; a result section that presents the visualization of the data extracted from the report and analyzes the results; and a discussion section that discusses the limitations and implications.

For the purpose of analyzing the data in this report, the R statistical programming language (R Core Team 2021) will be used. Furthermore, for cleaning and analysis procedures, the following packages must be installed in R: janitor()(Firke 2021), pdftools() (Ooms 2022), purrr()(Henry and Wickham 2020), tidyverse() (Wickham et al. 2019), stringi()(Gagolewski 2021), and kableExtra() (H. Zhu 2021).

Data

Data Source and Characteristics

According to the Central Bureau of Statistics (CBS), the Indonesia Demographic and Health Survey (IDHS) was conducted at the request of the State Ministry of Population/National Family Planning Coordinating Board (NFPCB). The objectives of the survey include providing data on fertility, family planning, maternal and child health, maternal mortality, and household expenditure data. The IDHS has been taken place in 1991, 1994 and 1997. The information gathered through the questionnaire was strictly confidential. DHS's website provides access to these reports in PDF format. We downloaded the report from the website, imported it into R studio and extracted the data. There were no other reports available that described the background characteristics of the data relating to maternal and child health like the IDHS report.

Our particular interest is in the table of birth intervals provided on page 42 of the report as part of the Fertility section. Below is the original table, including several sections of background characteristics rows, such as the age of the mother, the birth order, the child's gender, and whether or not the child survived his or her previous birth, residence and region, education, and employment status. The columns were the number of months since the previous birth (birth interval), the number of births, and the median number of months since the previous birth. For the column section regarding the number of months since previous birth, the values are expressed as percentages, with each row comprising a group of ages with a total percentage of 100%. Data in the last two columns represent the amount. The number of birth interval recorded for women who had multiple children would reflect the interval between the one prior to the last one and the last one.

As a part of this report, we will carefully analyze the birth interval variable of each categorical characteristic and the median for each group to determine the distribution of birth intervals within each category.

```
## Warning in include_graphics(paste(here(), "/outputs/image/pic.png", sep = "")):
## It is highly recommended to use relative paths for images. You had absolute
## paths: "/Users/chengleyang/Desktop/temp/Indonesia-Demographic-and-Health-
## Survey-1997/outputs/image/pic.png"
```

Table 3.6.1 Birth intervals: background characteristics

Percent distribution of non-first births in the five years preceding the survey by number of months since previous birth, according to demographic and socioeconomic characteristics, Indonesia 1997

Background characteristic	Number of months since previous birth						Median number of months since	Number of
	7-17	18-23	24-35	36-47	48+	Total	previous birth	births
Age of mother								
15-19	33.6	37.9	8.9	19.0	0.4	100.0	21.3	63
20-24	11.1	16.3	30.1	22.9	19.7	100.0	32.3	1,062
25-29	7.2	10.0	22.5	18.0	42.2	100.0	42.3	3,185
30-34	4.9	7.8	19.4	15.8	51.9	100.0	49.0	3,205
35-39	4.1	6.9	17.8	14.1	57.1	100.0	54.4	2,244
40-44	3.9	7.5	17.5	15.6	55.4	100.0	54.8	876
45-49	1.7	8.8	19.2	12.9	57.4	100.0	53.0	232
Birth order								
2-3	6.4	9.1	19.2	15.9	49.4	100.0	47.3	6,545
4-6	5.0	7.8	21.9	17.3	47.9	100.0	46.1	3,314
7+	7.2	15.1	28.1	20.1	29.4	100.0	35.8	1,008
Sex of prior birth								
Male	6.3	9.1	21.5	16.4	46.7	100.0	44.9	5,644
Female	5.8	9.5	20.2	17.0	47.5	100.0	45.8	5,222
Survival of prior birth								
Living	4.9	8.8	20.7	16.8	48.9	100.0	46.9	9,946
Dead	19.0	14.6	23.0	15.9	27.6	100.0	31.7	920
Residence								
Urban	6.2	9.9	18.5	14.7	50.8	100.0	48.3	2,777
Rural	6.0	9.1	21.7	17.4	45.8	100.0	44.4	8,090
Region/Residence								
Java-Bali	4.3	7.4	17.9	15.4	54.9	100.0	52.1	5,896
Urban	4.7	8.4	16.8	12.9	57.2	100.0	53.8	1,831
Rural	4.1	7.0	18.4	16.6	53.9	100.0	51.5	4,065
Outer Java-Bali I	8.4	11.5	23.7	18.2	38.2	100.0	39.3	3,405
Urban	9.0	12.4	21.6	18.4	38.8	100.0	40.5	660
Rural	8.2	11.3	24.2	18.2	38.1	100.0	39.0	2,746
Outer Java-Bali II	7.6	11.5	25.9	18.3	36.6	100.0	38.9	1,565
Urban	8.8	13.9	22.5	17.7	37.1	100.0	38.6	286
Rural	7.4	11.0	26.6	18.5	36.5	100.0	38.9	1,279
Education								
No education	5.0	8.6	23.5	17.3	45.6	100.0	44.2	1,283
Some primary	5.2	9.1	19.9	17.0	48.8	100.0	46.7	3,308
Completed primary	5.5	8.3	19.8	15.8	50.6	100.0	48.4	3,273
Some secondary+	8.1	10.8	22.0	17.2	41.9	100.0	41.6	3,002
Work status								
Worked in past year	6.2	9.2	21.2	17.0	46.4	100.0	45.0	4,416
Did not work	6.0	9.3	20.7	16.5	47.5	100.0	45.7	6,451
Total	6.1	9.3	20.9	16.7	47.1	100.0	45.3	10,867

Note: The interval for multiple births is the number of months since the preceding pregnancy that ended in a live birth.

Survey Methodology

A steering committee has been established to coordinate the development of the survey developed by the CBS, the NFPCB, and the Ministry of Health. The 1997 IDHS consisted of three schedules designed to gather information about households, the welfare of families and individuals. The content of the household and individual questionnaire is available in Appendix E of the original report. Survey questions were carefully crafted in order to provide managers, policymakers, and researchers with guidelines for evaluating and improving program effectiveness.

In order to make the survey more accessible, the survey was first designed in English and then translated into Indonesian. As part of the survey, the household schedule was modified from the Susenas survey model, a well-developed national survey conducted by CBS. As an extension of the previous IDHS conducted from 1991 to 1994, the family welfare schedule focuses on assessing the welfare status of family members listed by the households surveyed from the previous schedule throughout the country. The individual questionnaire was distributed only to ever-married women between the reproductive ages of 15-49, in accordance with the survey's objectives to examine fertility and maternal and child health.

A team of 86 interviewers with specific training collected the survey data. It was intended to reflect estimates at both a regional and national level since the sample of the survey was drawn from all 27 provinces. A total of 35,362 households were selected; 34,656 were located and 34,255 were successfully interviewed. According to Table 1.2 on page 7 of the report, it appears that more rural households sampled could not be located. This is perhaps due to the difficulty of finding rural households in developing countries such as Indonesia. Only 28,810 of 29,317 eligible women were interviewed for the individual questionnaire. The response rate was generally high, as the interview team was required to revisit the households as per the rules.

By conducting interviews rather than conducting a thorough investigation, it might impose a self-report bias on the results. As the samples were selected rather than randomly selected, the results may be susceptible to sampling error. Consequently, it was unclear whether the sample could be applied to the population of Indonesia. As well, even though the unfound households and individuals made up a relatively small part of the sample, there may have been some bias involved.

Discussion and Results

Results

All the graphs presented in this report is produced using ggplot() (Wickham 2016).

Maternal Risks and Age

By grouping the respondents by age group, we observed an up-ward sloping trend in median birth intervals. Referring to Figure 1, among mothers aged between 15-19, the median birth interval is 21.3 months. This figure is then increased to 54.8 months among mothers aged between 40-44 years old. However, the median began declined to 53 months as the age continues to increase, for mothers aged between 45 to 49 years old.

Through analyzing this trend, it is evident that more health care resources need to be devoted to younger mothers, as they generally have shorter birth intervals. As discussed previously, children conceived after 18 - 23 months of the previous birth are exposed to smaller risks of adverse perinatal outcomes (B. P. Zhu et al. 1999), while WHO states that a minimum of 33-month interval is recommended for mothers between two consecutive live births (Tesema, Worku, and Teshale 2021). The result of our analysis indicates that in 1997, in Indonesia, more than half of the mothers aged between 15 to 24 have birth intervals that were too short. This could lead to poor maternal and children health, which could further induce at-birth mortality.

On the other hand, although birth intervals are significantly larger as the age increases, mothers are still threatened by the risk for maternal dieases and other complications. Mothers who aged above 35 years

old are considered at an advanced maternal age, together with younger mothers aged below 19, women within these age groups are associated with increased risk of adverse maternal perinatal outcomes. These complications are independent of demographical cofounders such as poverty and weight (Cavazos-Rehg et al. 2015). With this and the previous discussion on the effects of short/pro-longed birth intervals, more support are needed by mothers within these two age groups in order to further reduce maternal risks and infant mortality.
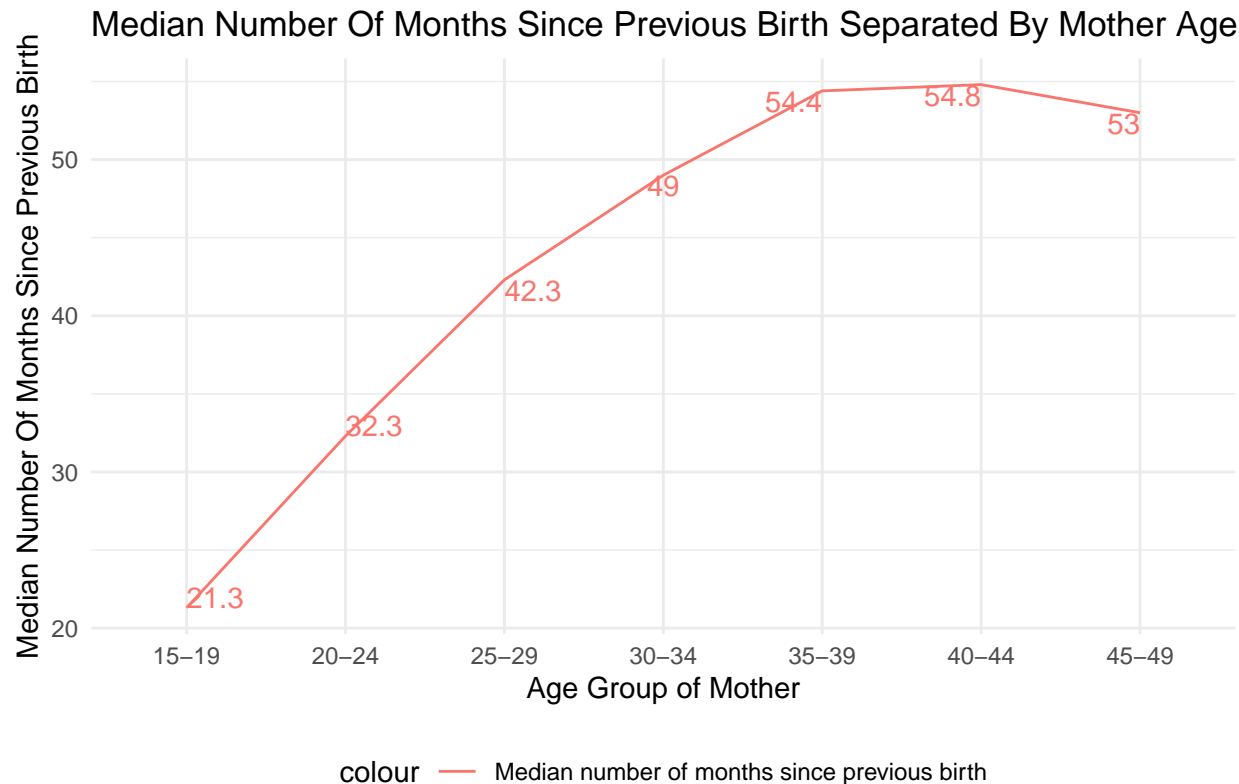
Median Number Of Months Since Previous Birth Separated By Mother Age

colour — Median number of months since previous birth

Figure1 created by Group 16 in STA304, Winter 2022

Sex of the Previous Birth and its Effect on Birth Interval

The culture of gender preference is deeply rooted in many Asian countries. Due to both cultural and economic factors, sons are much more perceived than daughters. As a result, families whose first child is a daughter are much more likely to have a second child in the hope of its a son (Aly and Shields 1991). Similarly, prior to any research, we expected the trend to continue in Indonesia - reflecting in birth intervals, we expected that those whose previous birth is a daughter should have shorter birth intervals for their next child as they are more incentivized to have a next child hoping its a son.

However, as shown in Figure 2, there is no obvious discrepancy between birth intervals, regardless of the gender of the previous birth. A potential explanation is that most of the ethnic groups in Indonesia, including the Javanese and the Madurese, have a bilateral kinship system. (Guilmoto 2016) Within the bilateral kinship system, children are considered equally belonged to both the father and mother side of the family. The word "belong" encompasses both emotional ties and inheritance rights such as the transfer of properties. Therefore, Indonesia is less prone to the son-preference culture, which is reflected in the result of this research that birth interval is relatively independent of the sex of prior birth.
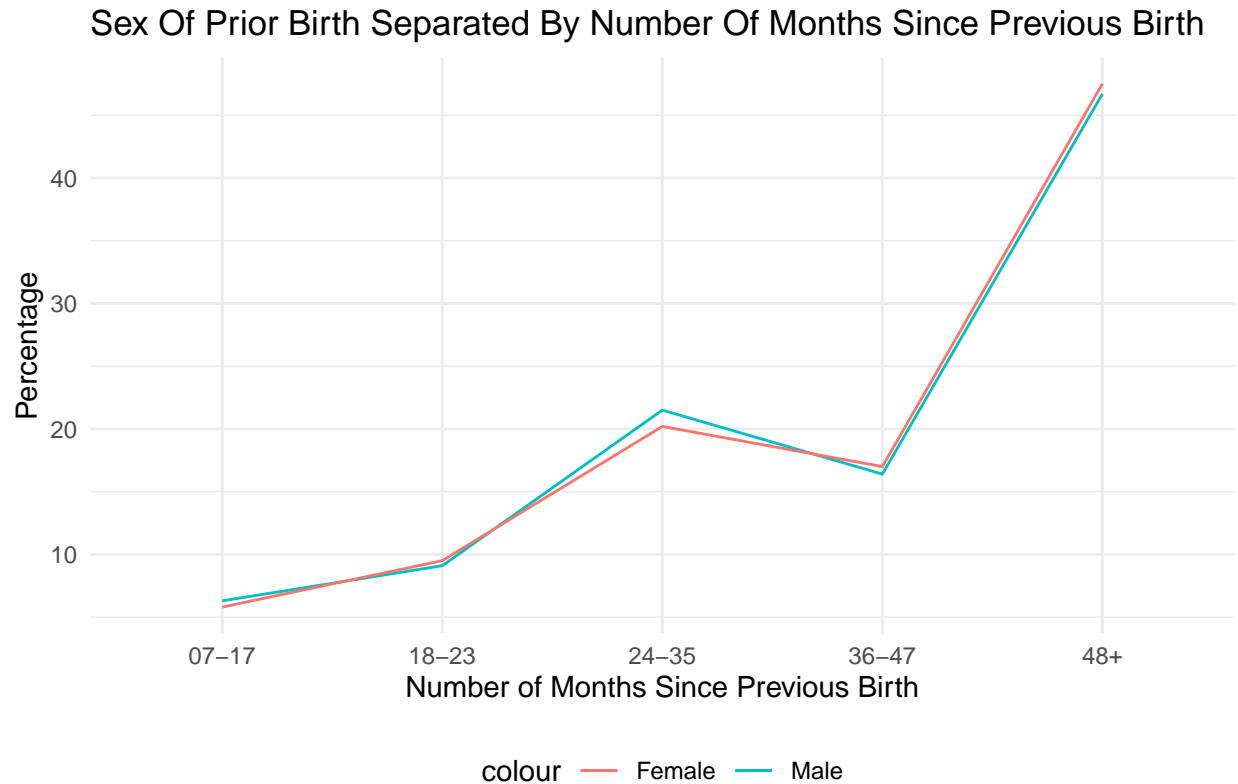
Sex Of Prior Birth Separated By Number Of Months Since Previous Birth

Figure2 created by Group 16 in STA304, Winter 2022

Survival of Prior Birth and its Effect on Birth Interval

Another factor affecting birth interval discussed in the 1997 Demographic and Health Survey is the survival of prior birth. As demonstrated in Figure 3, among mothers whose previous child is still alive, nearly 50% of them (48.9%) waited more than 48 months to have their next child. Only 4.9% chose to have a next child within 7 to 17 months of its previous birth. In contrast, 19% of mothers whose previous child passed away chose to have children within the first 7 to 17 months. In another word, mothers who lost its previous children are more likely to bear the risks associated with extremely short birth intervals.

From a psychological stand of point, this is reasonable as parents of a healthy newborn are more occupied, having a new child in such short period of time will increase their financial and psychological burden. For those faced infant mortality or the loss of loved ones, although they suffer from grief, ambivalence, and insecurities, research shown that the majority of women chose to become pregnant within a year after a perinatal loss. (Lamb 2022)

Nevertheless, becoming pregnant after perinatal loss is a lot more challenging due to the coping mechanism during the new pregnancy. Women are more prone to anxiety and depression, which are both causes to other maternal complications, during the perinatal period (Lamb 2022). Combining with the negative effects of short birth intervals, Indonesian women who has lost its previous child are exposed to greater pregnancy risks.
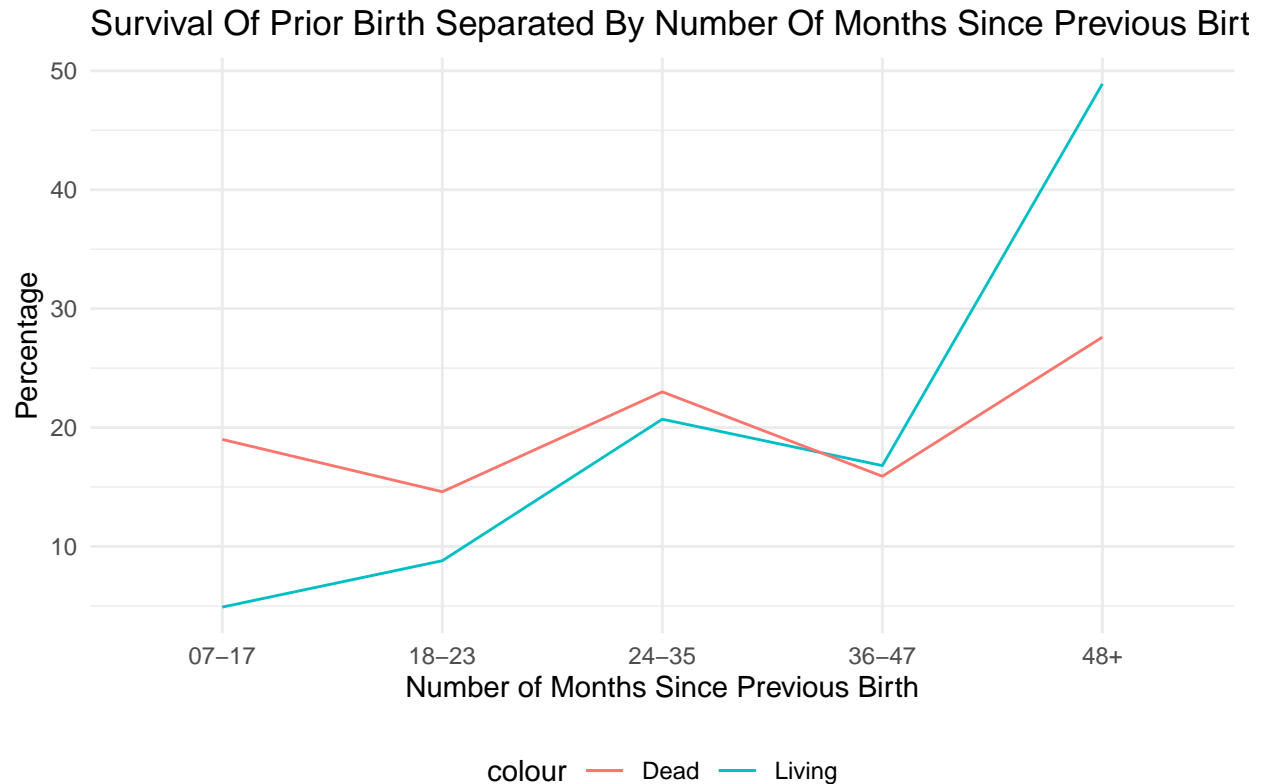
Survival Of Prior Birth Separated By Number Of Months Since Previous Birt

Figure3 created by Group 16 in STA304, Winter 2022

Demographical Factors and Their Effect on Birth Interval

This report also looks at other demographical factors of the mothers, including geographical location, work status, and education level, and how they affect birth intervals.

Figure 5 groups participants by their residential location and the area they live in. The IDHS categorized all participants into three different regions: Java-Bali; Outer Java-Bali I; and Outer Java-Bali II.Then under each region, the participants were further grouped by the areas they live in (rural or urban). The result of the analysis shows that the birth interval varies widely among different regions. The overall median birth interval in Java-Bali is 12.8 months longer than it in Outer Java-Bali I and 13.2 months longer in Outer Java-Bali II.

This phenomenon could be explained by the types of occupations women chose in these regions. According to IDHS, in 1997, only 36.1% of women in Java-Bali worked in land, while over 20% of women had industrial jobs. Meanwhile, in other regions, over 50% women worked in land. Agricultural jobs are significantly more flexible compared to industrial jobs, which provided more incentives for women to bear children. At the same time, job flexibility provides more time for mothers to recover from previous birth so that they are able to have birth within a shorter interval.
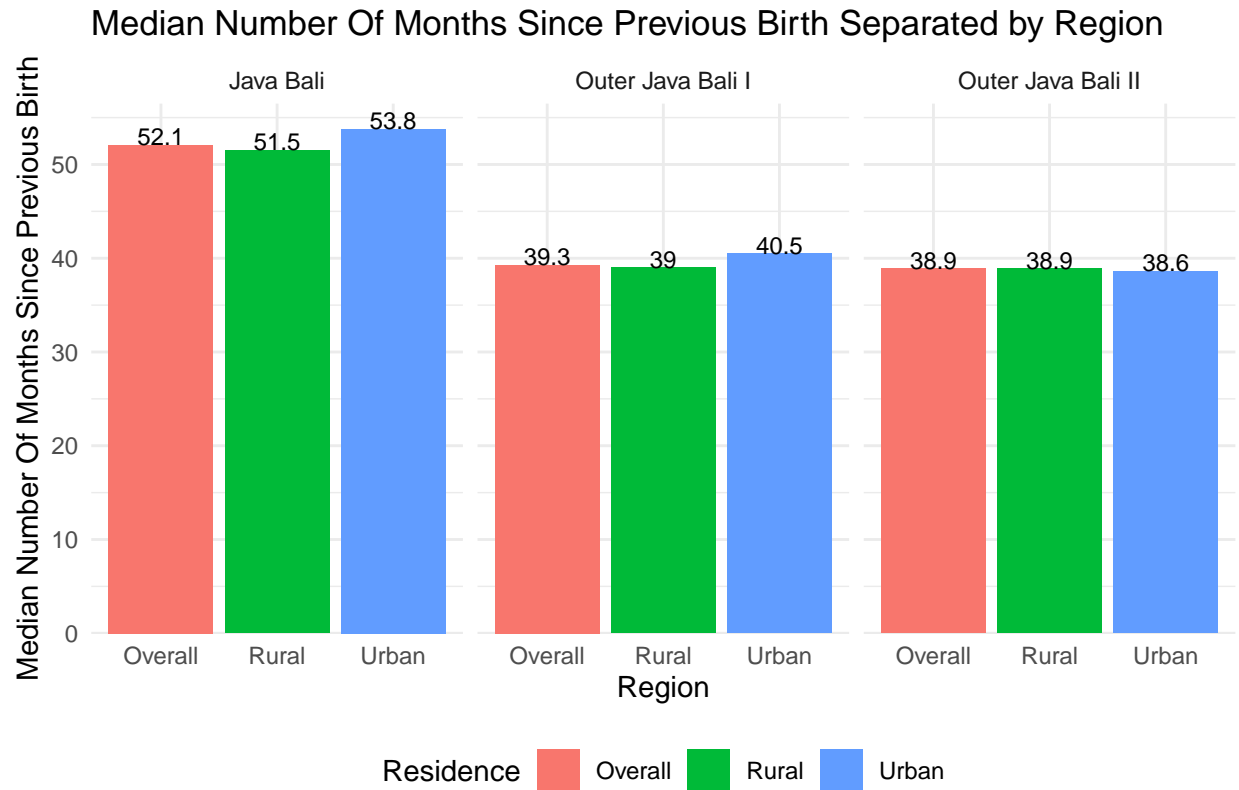
Median Number Of Months Since Previous Birth Separated by Region

Similarly, referring to Figure 6, where mothers were categorized based on their levels of education, there is no obvious discrepancy among those who had no education, some primary education, and completed primary education. Interestingly, it is notable that among those who had some secondary education and above, 8.1% had extremely short birth intervals (7 to 17 months), while the number was around 5% among other groups. According to the research conducted by Jungho Kim, who conducted in-depth research on the positive correlation between Indonesian women's education level and birth intervals, this phenomenon is extremely common in developing countries including Indonesia, as in the first stages of development, coupled with higher income have more incentive to bear children to pass down their properties and wealth (Kim 2010).

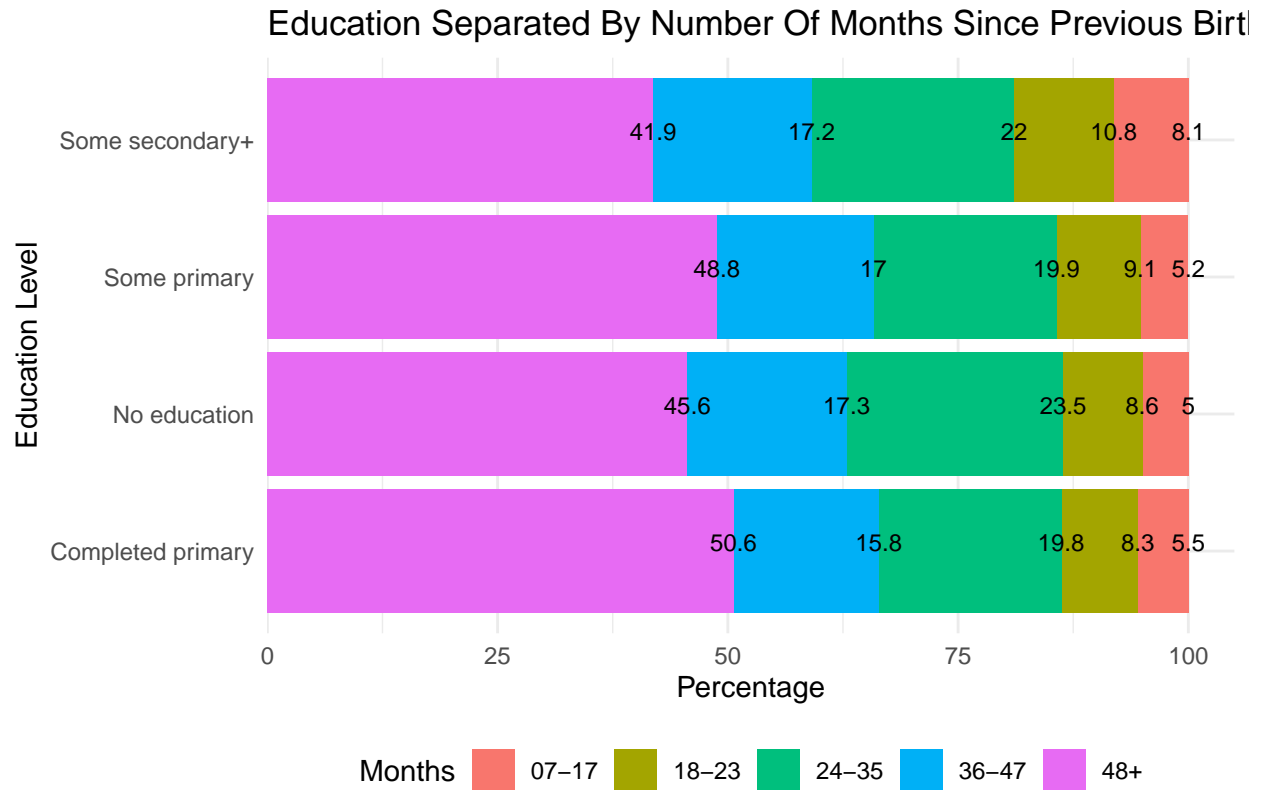Education Separated By Number Of Months Since Previous Birth

Figure5 created by Group 16 in STA304, Winter 2022

In contrast, as shown in Figure 6, there are no obvious disparities in birth intervals among women with different work status. That is, work status seems to have no impact on the length between giving births.
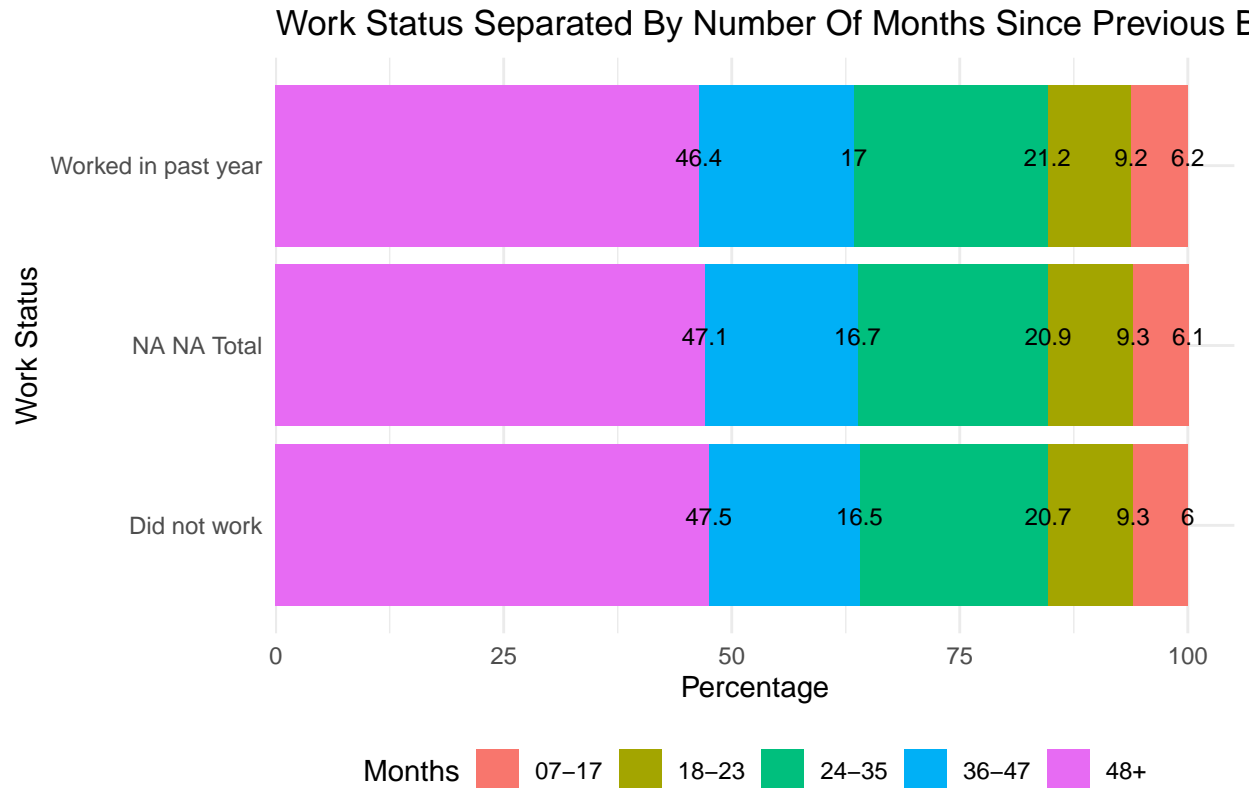
Work Status Separated By Number Of Months Since Previous E

Figure6 created by Group 16 in STA304, Winter 2022

Discussion

General Findings:

Overall, through the analysis conducted in this report, there are five general findings:

1. There is a positive relationship between mothers' age and their birth intervals - younger mothers tend to give birth to newborns within a short period from the birth of their previous child. (implication)

2. No obvious discrepancy is found in birth intervals among mothers with different work status in the past year and among those whose previous child is of a different gender. That is, it seems birth interval is independent of mothers' previous-year-work-status and previous child's gender. This finding is particularly interesting as it reflects Indonesia's unique bilateral kinship system and the government's ongoing effort in promoting women's rights.

3. Survival of the previous child has a significant impact on birth intervals - 19% of mothers who had lost their previous child got pregnant within 7 to 17 months of the previous birth while for those whose previous child is alive, only 4.9% gave birth within this short interval. 50% chose to wait more than 48 months instead.

4. Looking at the survey participants' demographic characteristics, birth intervals vary substantially between mothers living in different regions. Women who lived in the Java-Bali Region of Indonesia

had a median birth interval that was 35% longer than women who lived in the other two regions, but no obvious difference was observed among women who lived in rural or urban areas within each region. However, this brings no urgent implication as all median numbers are well above the birth interval threshold (33 months) recommended by WHO (Tesema, Worku, and Teshale 2021)

5. Another interesting finding is that unlike results indicated from other research, which suggests that there exists a negative relationship between a woman's education level and fertility (Kim 2010), there seems to be a negative relationship between Indonesian women's education level and their birth intervals. According to research, this is a common phenomenon in developing countries, and this provides another perspective for considering factors affecting women's fertility in the future

Implications

More Maternal Care Needs to be Provided for Mothers aged below 24 and above 35

As discussed previously in the result section, more than half of the mothers aged below 24 are having birth intervals that are way too short (below WHO recommended standard), while mothers aged above 35 are considered at an advanced maternal age. Giving birth at an advanced maternal age increases the risks of maternal complications. Together with the fact that having longer birth intervals increases the risk of pre-eclampsia and eclampsia, women aged above 35 also require additional care from the healthcare system.

The "mother-friendly movement" should devote more resources to pregnant women of these age groups in order to achieve its goal of reducing maternal mortality. From the result of the 1997 IDHS, it seems like the Indonesian government still had plenty of space for improvements in this area.

Indonesia Stands out In Terms of Gender Equality at Birth

The culture of son preference drives fertility and encourages families with daughters to have an additional child in order to achieve the expected number of sons (Aly and Shields 1991). This culture is especially common in Asian countries, even in certain European countries, and it could lead to great harm to the family and the mothers. Mothers who are "unable" to "deliver" sons to the family are often emotionally abused. Forced abortion, physical abuse, and other violent activities could also arise under this toxic culture (Hesketh, Lu, and Xing 2011).

From finding (2), it seems like Indonesian women were not pressured to immediately have another child after giving birth to a daughter. Other than the effect of Indonesia's unique bilateral kinship system, the Indonesian government's commitment in women's rights also contributed to gender equality at birth.

Mothers Whose Previous Child Passed Away Also Require More Maternal Care

Other than age, as discussed in finding (3), birth intervals were noticeably short for women who lost their previous child. Other than the risks associated with shorter birth intervals, these women are also exposed to more emotional burdens such as maternal depression (Lamb 2022). With this, the government need to devote more resource to reduce the risks of these complications, ultimately lower maternal mortality. This aligns with the goal of the "mother-friendly movement".

Education Level Could Impact Birth Interval Differently in the Future

Although in 1997, education level seems to have a negative relationship with birth interval, this relationship could change in the future. As discussed previously, this phenomenon is often prevalent in developing countries: as the economy is in the early stage of development, the benefit of having children to act as an

additional income source for the family outweighs the opportunity cost of being pregnant. However, in the long-run, education level seems to have a negative impact on fertility (Kim 2010). This means that the government might shift the resource allocation in the future when mothers with lower education levels in general have shorter birth intervals.

Ethics and biases

One bias of this research is that only ever-married women were chosen to participate in this survey. Women who gave birth without marrying were excluded from this dataset. This creates an ethical issue as according to research, "Births outside of marriage are often associated with disadvantage for both children and their parents. Women and men who have children outside of marriage are younger on average, have less education, and have lower income than married parents" (Shattuck and Kreider 2013). That is, as these women often have lower income and less access to resources, they are the group that the government should particularly focus on when it comes to reducing maternal mortality. Excluding them from this survey decreases the representativeness of the result and increases bias if the result is used for policy-making.

Another bias is that although there is no significant difference between birth intervals among mothers whose previous child is of different gender, others' research suggests that son preference still exists in Indonesia, and it should be prominent if the data is grouped by provinces. As some provinces, for example, Bali, have a patrilineal kinship system, the lack of a son does have an impact on the timing of the next child: "among parents without sons, the third birth is sooner and almost twice as frequent as among parents who already had a son (59% vs. 31%)" (Guilmoto 2016). Therefore, this survey failed to prevail the actual relationship between gender of the previous child and birth interval, which is a limitation of this report.

Research Outlook

For future research, an attempt to establish a model to better visualize the relationship between birth intervals and the factors discussed should be made. However, additional data points need to be collected for model building purposes. Individual-level data are much more preferred than estimations as it allows us to better train the model. Similarly, responses should be better curated to have quantitative meaning, in order to run regressions and other analyses. The sample size could be revised to incorporate premarital birth women, as they are subjected to more maternal risks which could amplify the impact of short birth intervals. Finally, geographical data should be collected in response to the limitation mentioned above to provide more insight into a previous child's gender's impact on birth intervals.

Appendix

A. Datasheet

Extract of the questions from Gebru et al. (2021)

A.1 Motivation

Q: For what purpose was the dataset created? Was there a specific task in mind? Was there a specific gap that needed to be filled? Please provide a description.

A: In the same manner as previous surveys, it was designed to provide policymakers and managers of population and health programs with comprehensive information regarding fertility, infant and child mortality, family planning, and maternal and child health.

Q: Who created the dataset (e.g., which team, research group) and on behalf of which entity (e.g., company, institution, organization)?

A: On behalf of the State Ministry for Population and the National Family Planning Coordinating Board (NFPCB), IDHS surveys were conducted by the Central Bureau of Statistics.

Q: Who funded the creation of the dataset? If there is an associated grant, please provide the name of the grantor and the grant name and number.

A: The Government of Indonesia funded all local costs of the survey through the development budget of the NFPCB.

Q: Any other comment?

A: The purpose of the 1997 IDHS was to extend the contents of previous IDHS by collecting data that could be used as indicators of family well-being.

A.2 Composition

Q: What do the instances that comprise the dataset represent (e.g., documents, photos, people, countries)? Are there multiple types of instances (e.g., movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.

A: The observations are constituted by selected people in all states over Indonesia. People are classified by different characteristics including ages, regions, work status, and other features of households and individuals.

Q: How many instances are there in total (of each type, if appropriate)?

A: The IDHS 1997 interviewed 34,255 households and 28,810 ever-married women aged 15 to 49.

Q: Does the dataset contain all possible instances or is it a sample (not necessarily random) of instances from a larger set? If the dataset is a sample, then what is the larger set? Is the sample representative of the larger set (e.g., geographic coverage)? If so, please describe how this representativeness was validated/verified. If it is not representative of the larger set, please describe why not (e.g., to cover a more diverse range of instances, because instances were withheld or unavailable).

A: The dataset was a sample selected from the whole population, the same as the 1994 IDHS sample which was a subsample of the 1994 National Socio-Economic Survey. The sample methodologies were It was developed in respect of providing reliable estimates of major variables in each province, as well as for urban and rural areas.

Q: What data does each instance consist of? "Raw" data (e.g., unprocessed text or images) or features? In either case, please provide a description.

A: The data analyzed in this report represents the background features of age, birth order, sex of prior birth, survival of prior birth, residence/region, education and work status.

Q: Is there a label or target associated with each instance? If so, please provide a description.

A: Yes, the first column of the dataset clearly labelled and classified each instance into groups of distinct background characteristics.

Q: Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (e.g., because it was unavailable). This does not include intentionally removed information, but might include, e.g., redacted text.

A: No.

Q: Are relationships between individual instances made explicit (e.g., users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.

A: There exists no explicit relationship between individual instances.

Q: Are there recommended data splits (e.g., training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.

A: There are no recommended data splits.

Q: Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.

A: There are no errors, sources of noise, or redundancies in the dataset presented.

Q: Is the dataset self-contained, or does it link to or otherwise rely on external resources (e.g., websites, tweets, other datasets)? If it links to or relies on external resources, a) are there guarantees that they will exist, and remain constant, over time; b) are there official archival versions of the complete dataset (i.e., including the external resources as they existed at the time the dataset was created); c) are there any restrictions (e.g., licenses, fees) associated with any of the external resources that might apply to a dataset consumer? Please provide descriptions of all external resources and any restrictions associated with them, as well as links or other access points, as appropriate.

A: The dataset is self-contained without the link to external sources.

Q: Does the dataset contain data that might be considered confidential (e.g., data that is protected by legal privilege or by doctor-patient confidentiality, data that includes the content of individuals' non-public communications)? If so, please provide a description.

A: Yes, the data in the dataset of the individual level are all confidential and strictly applied to the privacy restrictions.

Q: Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.

A: No.

Q: Does the dataset identify any subpopulations (e.g., by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset. Is it possible to identify individuals (i.e., one or more natural per- sons), either directly or indirectly (i.e., in combination with other data) from the dataset? If so, please describe how.

A: The dataset contains subpopulations of women aged 15-49 which is the observations analyzed in this report. It is unable to identify individuals.

Q: Does the dataset contain data that might be considered sensitive in any way (e.g., data that reveals race or ethnic origins, sexual orientations, religious beliefs, political opinions or union memberships, or locations; financial or health data; biometric or genetic data; forms of government identification, such as social security numbers; criminal history)? If so, please provide a description.

A: Yes, the survey that produces the data records ethnic origins, sexual activity, information about AIDS, religious beliefs, financial and health data, and biometric data.

Q: Any other comments?

A: N/A.

A.3 Collection Process

Q: How was the data associated with each instance acquired? Was the data directly observable (e.g., raw text, movie ratings), reported by subjects (e.g., survey responses), or indirectly inferred/derived from other data (e.g., part-of-speech tags, model-based guesses for age or language)? If the data was reported by subjects or indirectly inferred/derived from other data, was the data validated/verified? If so, please describe how.

A: The data was collected from survey responses from a total of 34,255 households and 28,810 ever-married women aged between 15 and 49. Therefore it is reported by subjects (self-reported data).

Q: What mechanisms or procedures were used to collect the data (e.g., hardware apparatuses or sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?

A: A team of 86 interviewers with specific training collected the survey data. Interviews were conducted in all 34,255 households, and the sample was designed to produce estimate at the provincial level.

Q: If the dataset is a sample from a larger set, what was the sampling strategy (e.g., deterministic, probabilistic with specific sampling probabilities)?

A: According to the IDHS report, sample households were selected from a provincial level - they did not constitute a proportional representation at the country level. The results are "based on data that were weighted to take account of differential sampling probabilities and non response at both the household and individual levels" ().

Q: Who was involved in the data collection process (e.g., students, crowdworkers, contractors) and how were they compensated (e.g., how much were crowdworkers paid)?

A: Data were collected by 284 interviewers, including 86 filed editors and 86 supervisors. The survey fieldworkers consist mostly of CBS staff therefore they were compensated that way. No informationon how much exactly they were paid were provided in IDHS report.

Q: Over what timeframe was the data collected? Does this timeframe match the creation timeframe of the data associated with the instances (e.g., recent crawl of old news articles)? If not, please describe the timeframe in which the data associated with the instances was created.

A: The survey fieldwork began on September 1, 1997 and was completed on December 31, 1997.

Q: Were any ethical review processes conducted (e.g., by an institutional review board)? If so, please provide a description of these review processes, including the outcomes, as well as a link or other access point to any supporting documentation.

A: N/A

Q: Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (e.g., websites)?

A: The data was accessed from the IDHS report in the form of a table and then transformed into R dataset.

Q: Were the individuals in question notified about the data collection? If so, please describe (or show with screenshots or other information) how notice was provided, and provide a link or other access point to, or otherwise reproduce, the exact language of the notification itself.

A: The individuals were notified regarding the collection, but it was unclear on how they were notified.

Q: Did the individuals in question consent to the collection and use of their data? If so, please describe (or show with screenshots or other information) how consent was requested and provided, and provide a link or other access point to, or otherwise reproduce, the exact language to which the individuals consented.

A: No information was provided in the IDHS report regarding consent.

Q: If consent was obtained, were the consenting individuals provided with a mechanism to revoke their consent in the future or for certain uses? If so, please provide a description, as well as a link or other access point to the mechanism (if appropriate).

A: N/A

Q: Has an analysis of the potential impact of the dataset and its use on data subjects (e.g., a data protection impact analysis) been conducted? If so, please provide a description of this analysis, including the outcomes, as well as a link or other access point to any supporting documentation.

A: N/A

Q: Any other comments?

A: N/A

A.4 Preprocessing/cleaning/labelling

Q: Was any preprocessing/cleaning/labelling of the data done (e.g., discretization or bucketing, tokenization, part-of-speech tagging, SIFT feature extraction, removal of instances, processing of missing values)? If so, please provide a description. If not, you may skip the remaining questions in this section.

A: yes. we use ocr to extract the table, tokenization and replace the wrong instance

Q: Was the "raw" data saved in addition to the preprocessed/cleaned/labelled data (e.g., to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.

A: https://github.com/chle1999/Indonesia-Demographic-and-Health-Survey-1997/tree/main/inputs/FR95.pdf

Q: Is the software that was used to preprocess/clean/label the data available? If so, please provide a link or other access point.

A: https://www.rstudio.com

Q: Any other comments?

A: N/A

A.5 Uses

Q: Has the dataset been used for any tasks already? If so, please provide a description.

A: The dataset was used for analysis regarding fertility in the 1997 IDHS report. The IDHS report included a preliminary research on the factors impacting women's birth interval.

Q: Is there a repository that links to any or all papers or systems that use the dataset? If so, please provide a link or other access point.

A: N/A

Q: What (other) tasks could the dataset be used for?

A: The dataset could potentially be used for comparisons - comparing changes in birth interval among IDHS report results from multiple years.

Q: Is there anything about the composition of the dataset or the way it was collected and preprocessed/cleaned/labelled that might im- pact future uses? For example, is there anything that a dataset consumer might need to know to avoid uses that could result in unfair treatment of individuals or groups (e.g., stereotyping, quality of service issues) or other risks or harms (e.g., legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?

A: The users need to be aware of how the participants were grouped in the dataset - there are two types of grouping: by age group and by regions. When conducting analysis, be aware of which grouping method was selected and its impact on the analysis.

Q: Are there tasks for which the dataset should not be used? If so, please provide a description.

A: N/A

Q: Any other comments?

A: N/A

A.6 Distribution

Q: Will the dataset be distributed to third parties outside of the entity (e.g., company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.

A: N/A

Q: How will the dataset will be distributed (e.g., tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?

A: it does not have DOI

Q: When will the dataset be distributed?

A: 1998

Q: Will the dataset be distributed under a copyright or other intellectual property (IP) license, and/or under applicable terms of use (ToU)? If so, please describe this license and/or ToU, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms or ToU, as well as any fees associated with these restrictions.

A: N/A

Q: Have any third parties imposed IP-based or other restrictions on the data associated with the instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any relevant licensing terms, as well as any fees associated with these restrictions.

A: N/A

Q: Do any export controls or other regulatory restrictions apply to the dataset or to individual instances? If so, please describe these restrictions, and provide a link or other access point to, or otherwise reproduce, any supporting documentation.

A: N/A

Q: Any other comments?

A: N/A

A.7 Maintenance

Q: Who will be supporting/hosting/maintaining the dataset?

A: Central Bureau of Statistics Jakarta, Indonesia

Q: How can the owner/curator/manager of the dataset be contacted(e.g., email address)?

A: Telephone 345-6285; Fax 384-1545

Q: Is there an erratum? If so, please provide a link or other access point.

A: N/A

Q: Will the dataset be updated (e.g., to correct labelling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (e.g., mailing list, GitHub)?

A: N/A

Q: If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (e.g., were the individuals in question told that their data would be retained for a fixed period of time and then deleted)? If so, please describe these limits and explain how they will be enforced.

A: N/A

Q: Will older versions of the dataset continue to be supported/hosted/maintained? If so, please describe how. If not, please describe how its obsolescence will be communicated to dataset consumers. N/A

Q: If others want to extend/augment/build on/contribute to the dataset, is there a mechanism for them to do so? If so, please provide a description. Will these contributions be validated/verified? If so, please describe how. If not, why not? Is there a process for communicating/distributing these contributions to dataset consumers? If so, please provide a description.

A: N/A

Q: Any other comments?

A: N/A

Reference

Adams, M. M., K. M. Delaney, P. W. Stupp, B. J. McCarthy, and J. S. Rawlings. 1997. "The Relationship of Interpregnancy Interval to Infant Birthweight and Length of Gestation Among Low Risk Women, Georgia." *Paediatric and Perinatal Epidemiology*, 48–62. https://doi.org/10.1046/j.1365-3016.11.s1.8.x.

Aly, Hassan Y., and Michael P. Shields. 1991. "Son Preference in Egypt." *Economic Development and Cultural Change.* https://doi.org/https://doi.org/10.1086/451874.

Bank, The World. 2022. "Population, Total - Indonesia." https://data.worldbank.org/indicator/SP.POP.TOTL?locations=ID.

Bureau of Democracy, Human Rights, and Labor. 1997. "Indonesia Report on Human Rights Practices for 1997." https://irp.fas.org/world/indonesia/indonesia-1997.htm.

Cavazos-Rehg, Patricia A., Melissa J. Krauss, Edward L. Spitznagel, Kerry Bommarito, Tessa Madden, Margaret A. Olsen, Harini Subramaniam, Jeffrey F. Peipert, and Laura Jean Bierut. 2015. "Maternal Age and Risk of Labor and Delivery Complications." *Maternal and Child Health Journal.* https://doi.org/10.1007/s10995-014-1624-7.

Conde-Agudelo, A., and J. M. Belizan. 2000. "Maternal Morbidity and Mortality Associated with Interpregnancy Interval: Cross Sectional Study." *BMJ*, 1255–59. https://doi.org/10.1136/bmj.321.7271.1255.

Conde-Agudelo, Agustin, Rosas-Bermúdez Anyeli, and Ana C Kafury-Goeta. 2007. "Effects of Birth Spacing on Maternal Health: A Systematic Review." *American Journal of Obstetrics and Gynecology*, 297–308. https://doi.org/10.1016/j.ajog.2006.05.055.

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Gagolewski, Marek. 2021. *Stringi: Fast and Portable Character String Processing in r.* https://stringi.gagolewski.com/.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

Guilmoto, Christophe Z. 2016. "Is There a Gender Bias in Births and Child Mortality in Indonesiai?" *NIUSS.* https://www.niussp.org/health-and-mortality/gender-bias-births-child-mortality-indonesiapeut-parler-de-discrimination-sexuelle-en-indonesie/.

Henry, Lionel, and Hadley Wickham. 2020. *Purrr: Functional Programming Tools.* https://CRAN.R-project.org/package=purrr.

Hesketh, Therese, Li Lu, and Zhu Wei Xing. 2011. "TThe Consequences of Son Preference and Sex-Selective Abortion in China and Other Asian Countries." *Canadian Medical Association Journal.* https://doi.org/10.1503/cmaj.101368.

Kim, Jungho. 2010. "Women's Education and Fertility: An Analysis of the Relationship Between Education and Birth Spacing in Indonesia." *Economic Development and Cultural Change.* https://doi.org/10.1086/649638.

Lamb, Elizabeth H. 2022. "The Impact of Previous Perinatal Loss on Subsequent Pregnancy and Parenting." *J Perinat Educ.*, 33–40. https://doi.org/10.1624/105812402X88696.

LV, Klerman, Cliver SP, and Goldenberg RL. 1998. "The Impact of Short Interpregnancy Intervals on Pregnancy Outcomes in a Low-Income Population. Am j Public Health." *AM J Public Health*, 1182–85. https://doi.org/10.2105/ajph.88.8.1182.

M, Norton. 2005. "New Evidence on Birth Spacing: Promising Findings for Improving Newborn, Infant, Child, and Maternal Health." *International Journal of Gynecology and Obstetrics*, S1–6. https://doi.org/10.1016/j.ijgo.2004.12.012.

Ooms, Jeroen. 2022. *Pdftools: Text Extraction, Rendering and Converting of PDF Documents*. https://CRAN.R-project.org/package=pdftools.

Pinto-Martin, J. A., A. Cnaan, and H. Zhao. 1998. "Short Interpregnancy Interval and the Risk of Disabling Cerbral Palsy in a Low Birthweight Population." *Pediatric Physical Therapy*, 53–54. https://doi.org/10.1097/00001577-200214010-00010.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rinaldo, Rachel. 2019. "The Women's Movement and Indonesia's Transition to Democracy." *SocArXiv*. https://doi.org/10.31235/osf.io/cfyva.

Shattuck, Rachel M., and Rose M. Kreider. 2013. "Social and Economic Characteristics of Currently Unmarried Women with a Recent Birth: 2011."

Tesema, Getayeneh Antehunegn, Misganaw Gebrie Worku, and Achamyeleh Birhanu Teshale. 2021. "Duration of Birth Interval and Its Predictors Among Reproductive-Age Women in Ethiopia: Gompertz Gamma Shared Frailty Modeling." *PLoS ONE*. https://doi.org/10.1371/journal.pone.0247091.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Zhu, B. P., R. T. Rolfs, B. E. Nangle, and J. M. Horan. 1999. "Effect of the Interval Between Pregnancies on Perinatal Outcomes." *The New England Journal of Medicine*, 589–94. https://doi.org/10.1056/NEJM199902253400801.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.