Places of child delivery and health: Trends from Turkey in the 1998 Demographic and Health Survey*

Health Trends in 1998 from Turkish Demographic and Health Surveys

Meha Grewal   Ayoon Kim   Shoubhik Bhattacharya

10 April 2022

Abstract

A family's future health and wellness largely depends on the well-being of the mother and the child at birth. Therefore, choice of place of delivery play an important role in determining the wellness of the family. Delivery in a healthcare facility generally brings better outcome for both the mother and the child, so our primary goal is to analyze the factors influencing the choice of child delivery. In order to do that, we obtain a dataset from the Turkish Demographic and Health Surveys from 1998. Then we convert the dataset into a usable dataset to analyze the factors affecting choices of places of delivery. In doing so, we conclude that place of delivery can be indicative of maternal and child health, which can also provide a broader context of the overall health and well-being of the people in Turkey.

Contents

1 Introduction

Place of delivery and fertility is a central issue regarding women's health. To promote safe and informed decisions on health policy, it is essential to understand the place of delivery is for children and mothers. The data originates from the Demographic and Health Surveys (DHS) where their purpose is to assist policy makers in governing a nation. The objective for the DHS is to expand health information and populations in survey countries. This paper includes data from the Turkish Demographic and Health Survey from 1998. The original report includes information regarding fertility, child mortality, indicators of health, and family

---

*Code and data are available at: https://github.com/ShoubhikBhattacharya/STA304_Paper4a

planning. Although, we focus on analyzing places of delivery during childbirth and what factors impact that decision.

Place of delivery can be indicative of maternal and child health, and can provide broader context for overall health in Turkey. An analysis of various variables that can impact places of delivery, can provide knowledge on how those variables can potentially affect other aspects of global health. To create a usable data set we input all the data from the table in DHS into Excel, which became the raw csv file for our analysis. We examined the distribution of each variable in the data set in relation to place of delivery to identify the most influential elements. Furthermore, we created a simulation to understand the most dominant factors and predict the place of delivery for populations in Turkey. Our findings indicate that the location that women utilize for their place of delivery can be influenced by a variety of factors. Our plots showcase how a more highly educated woman, or individual from more urban areas prefer health facilities. These factors can also intersect in the decision of place of delivery and other health factors. Although, this can be representative of a lack of access for women with less privileged socioeconomic backgrounds.

The well-being of a mother and child at birth determines the health of the next generation by helping picture future public health challenges for families, communities, and the health care system. Therefore, by analyzing the factors affecting choice of places of delivery, it can allow us to identify existing health risks in women and hence prevent potential health risks for women and their children. The paper begins by explaining the motivation behind the data and the analysis. Followed by the composition of the data obtained from the 1998 Turkey DHS. Afterwards, we go in depth about the methodology. Our data section contains plots that summarize the variables and an accompanying discussion. Followed by graphs generated from simulated data and its implications. The paper concludes with an explanation of possible weaknesses and limitations.

2 Data

The data is obtained from the Demographic and Health Surveys (DHS) hosted on the DHS website in PDF formats. We obtained data from the Turkey DHS Final Report from 1998 Hacettepe University (1998). The report consists of a wide range of data regarding Children health and its indicators such as mortality, nutrition, water supply and education. Further, other supporting indicators data is provided such as mother's health and nutrition and infant vaccination rates. We obtained table 9.4 from the Turkey 1998 DHS Report for our analysis.This table consists of important statistics regarding birth statistics such as mother's age at birth, child birth order, residence of birth, region of birth, mother's educational attainment levels and total number of antenatal care visits during birth. These variables are further classified by the place of delivery of the child - at health facility or at home. In terms of confidentiality, only aggregate levels were used to construct the table and no micro data containing individual patient information has been obtained or utilized. The method of categorization of this data set (ordering by place of delivery) limits us to compare data to other similar reports.

For conducting this analysis we used the R programming language R Core Team (2020), and libraries such as Wickham et al. (2019), Xie (2022) and Wickham (2016).

As stated previously, we look at Table 9.4 from the Turkey 1998 DHS Final Report on pages 109-110. We consider all sections of this table - Mother's age at birth, birth order, residence, region, education and Antenatal Care (AC) visits. We look at all the variables and match them with their respective columns regarding the place of delivery. We do so to gain a clear understanding on how hygiene patterns during delivery and proper medical attention affect birth outcomes in children. Table 9.4 is shown as Figure 1 below -

As stated previously, we look at Table 9.4 from the Turkey 1998 DHS Final Report on pages 109-110. We consider all sections of this table - Mother's age at birth, birth order, residence, region, education and Antenatal Care (AC) visits. We look at all the variables and match them with their respective columns regarding the place of delivery. We do so to gain a clear understanding on how hygiene patterns during delivery and proper medical attention affect birth outcomes in children. Table 9.4 is shown as Figure 1 below -

Table 9.4 Place of delivery

Percent distribution of births in the five years preceding the survey by place of delivery, according to selected background characteristics, Turkey 1998

Background characteristic	Place of delivery			Total	Number of births
	At a health facility	At home	Don't know/ Missing		
Mother's age at birth					
< 20	73.6	25.8	0.6	100.0	501
20-34	73.4	25.8	0.8	100.0	2,706
35+	61.3	37.6	1.2	100.0	252
Birth order					
1	86.9	12.7	0.4	100.0	1,192
2-3	73.5	25.8	0.7	100.0	1,495
4-5	57.6	40.7	1.6	100.0	435
6+	36.8	61.7	1.4	100.0	337
Residence					
Urban	80.2	19.1	0.6	100.0	2,162
Rural	59.7	39.2	1.1	100.0	1,297
Region					
West	86.6	12.7	0.7	100.0	1,031
South	69.2	29.8	1.0	100.0	490
Central	83.3	16.3	0.3	100.0	795
North	83.7	15.5	0.8	100.0	271
East	44.4	54.4	1.2	100.0	871
Education					
No educ./Pri. incomp.	44.5	54.4	1.1	100.0	936
Pri. comp./Sec. incomp.	80.4	18.8	0.8	100.0	2,118
Sec. comp./+	96.2	3.8	0.0	100.0	405
Antenatal care visits					
None	45.4	54.1	0.5	100.0	1,090
1-3 visits	74.8	24.8	0.4	100.0	859
4+ visits	91.9	7.9	0.2	100.0	1,451
DK/missing	64.4	7.6	28.0	100.0	59
Total	72.5	26.7	0.8	100.0	3,459

Figure 1: Place of Delivery by Background Characteristics

3   Results

We apply the legend here to specify the place of delivery for the following -
- H - Birth delivery at Home
- HF - Birth delivery at Health Facility
- DK - Birth delivery information is unknown or missing

Plotted below are the graphs for place of delivery by each background characteristic -

## Figure 2: Place of Delivery by Mother's age at birth
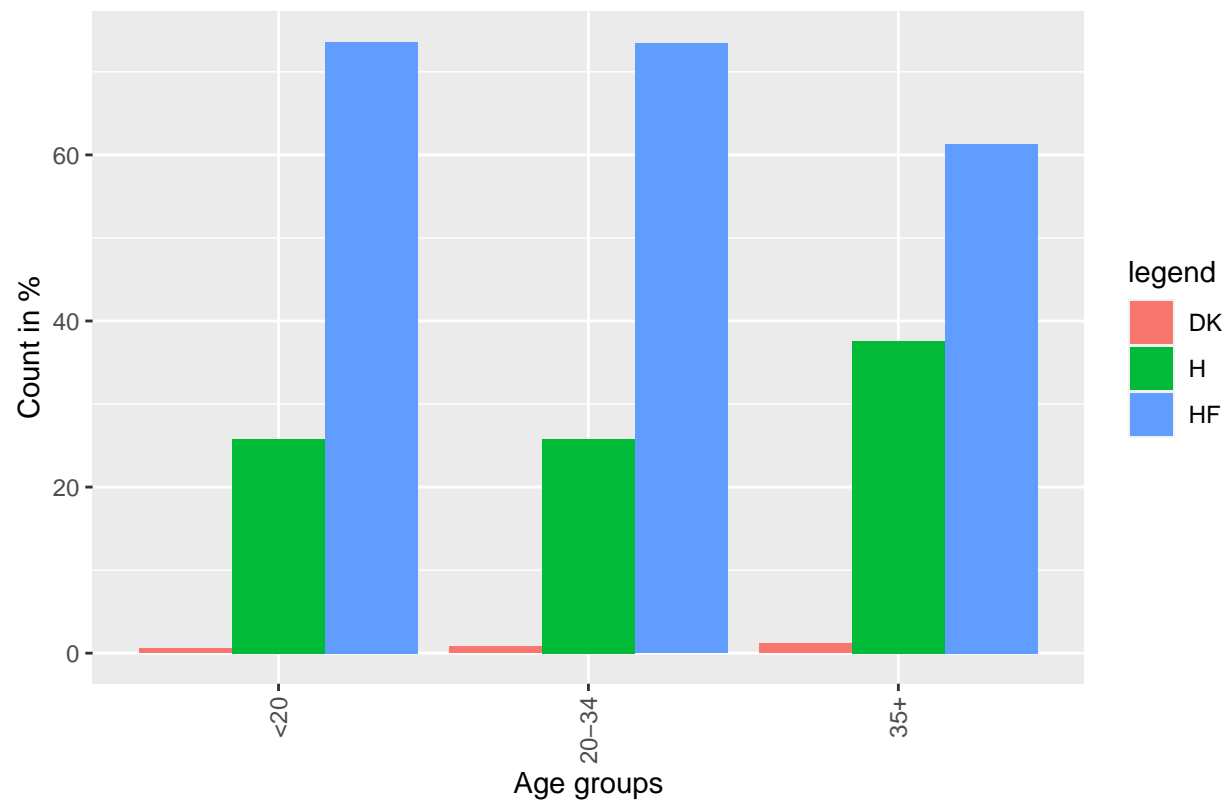
These plots summarize the variables in our data set and how it impacts the place of delivery. To begin with mother's age in Figure 2**??**, the vast majority of mothers give birth at a health facility. However, a larger portion of mothers aged 35 and above give birth at home.

Figure 3: Place of Delivery by Child's birth order
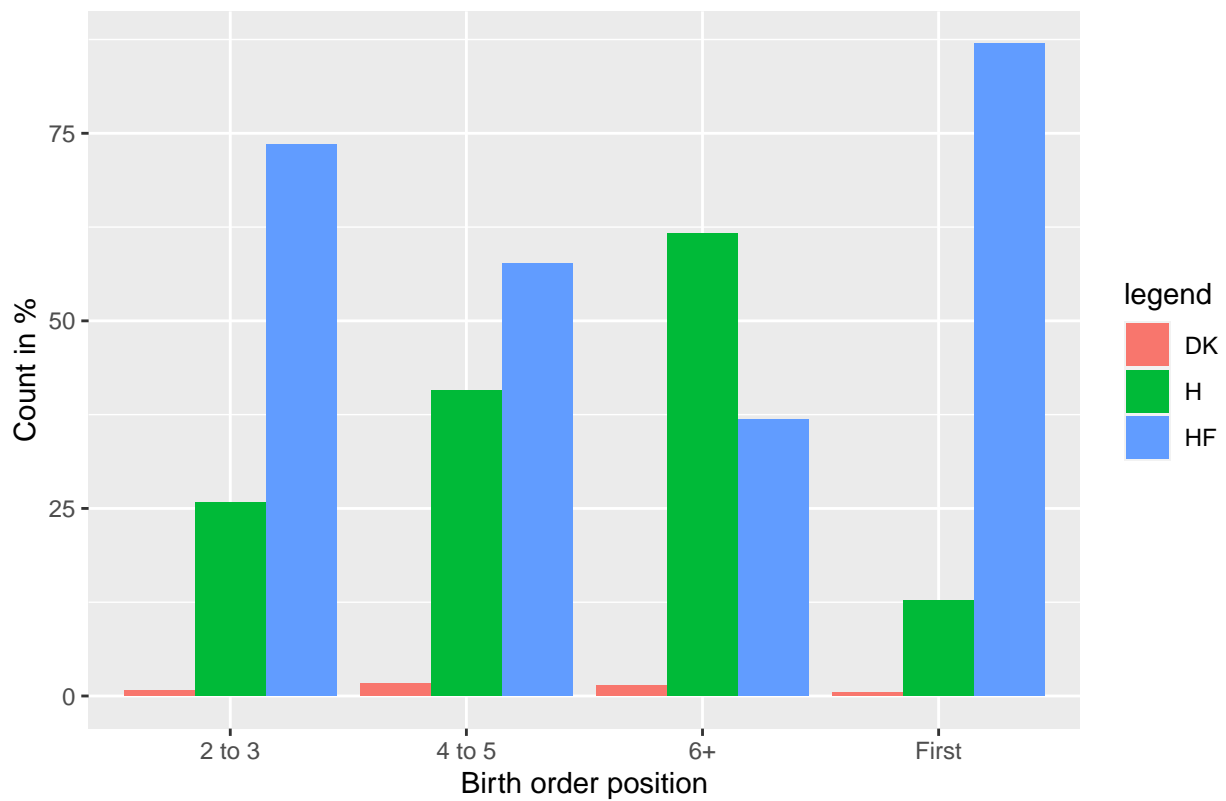
Also, looking at place of delivery based on child birth order in Figure 3 (Figure **??**), a very large part of children born first were at a health facility. That portion goes down as the child birth order increases and at 6 or more kids, majority of children were birthed at home.

Figure 4: Place of Delivery by Residence
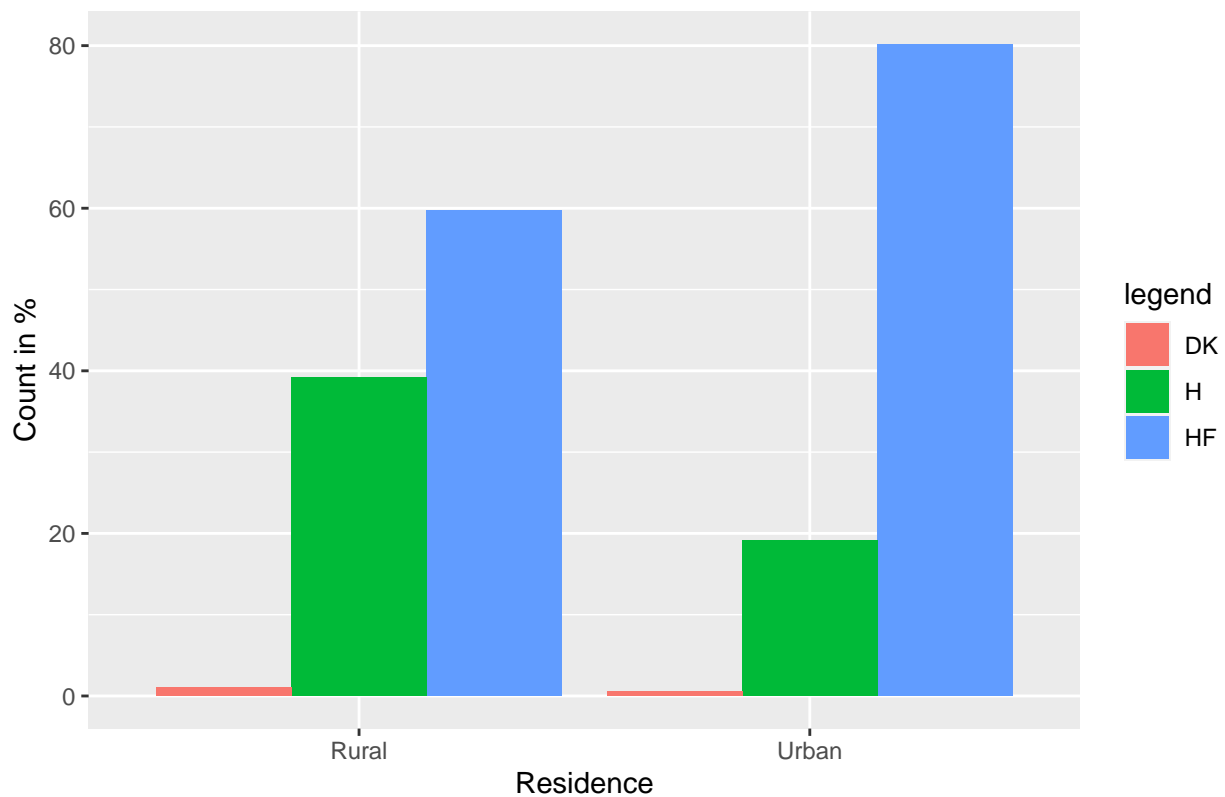
Based on Figure 4 (Figure **??**), births at health facilities are much higher in urban residences and smaller for rural areas. This can possibly be attributed to more accessibility to health facilities in urban regions as there are more resources closer together, while residents in rural areas must travel farther distances to gain access to health facilities.

(Figure 5 2) showcases how education level impacts place of delivery. Mothers with no education largely delivered at home. The higher the mother's educational level, the larger the likelihood of them delivering their child at a health facility.
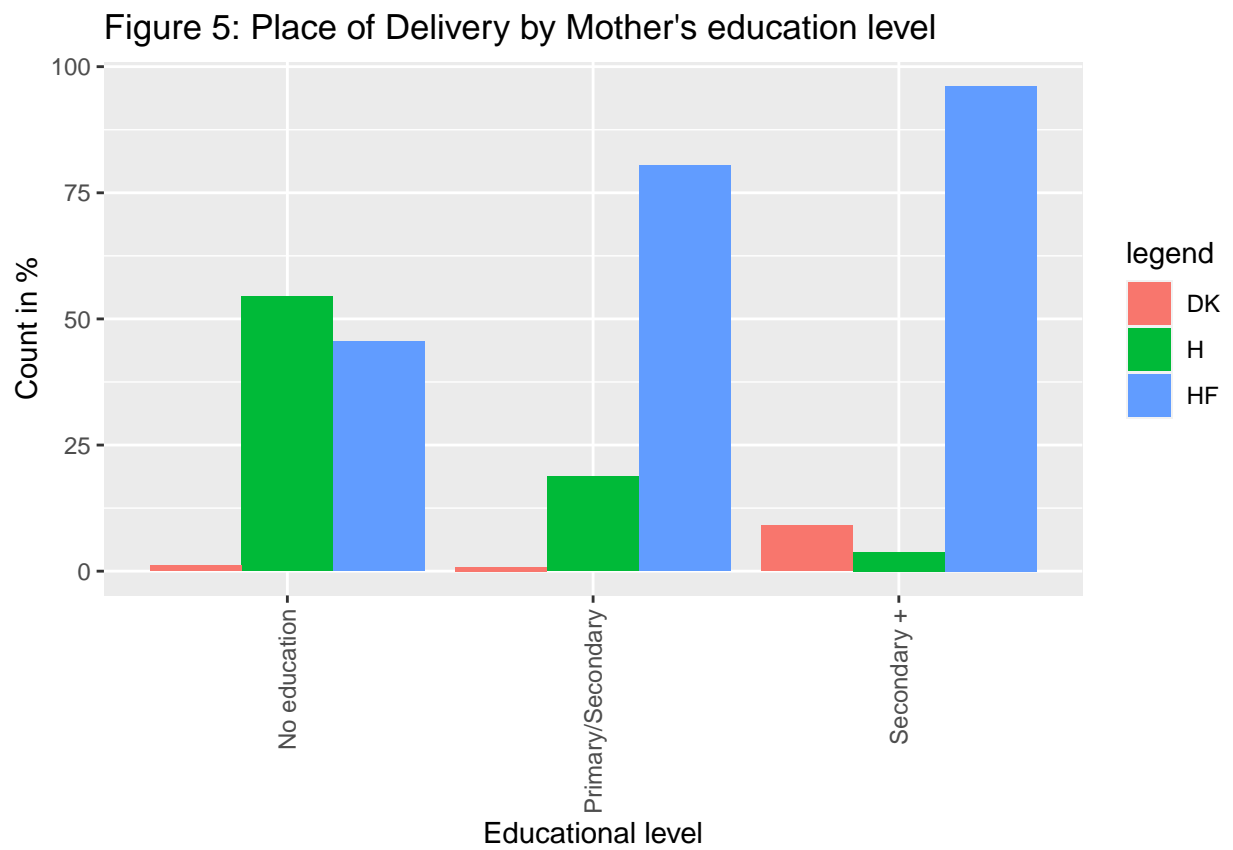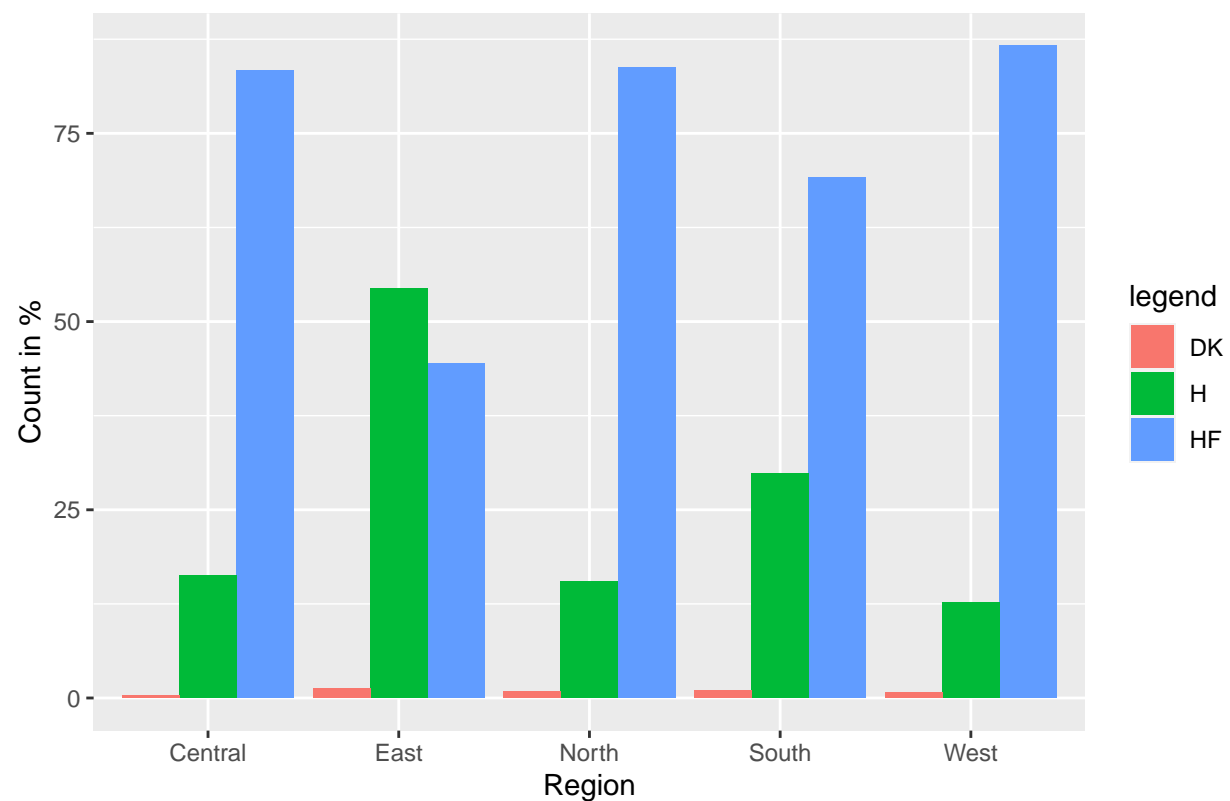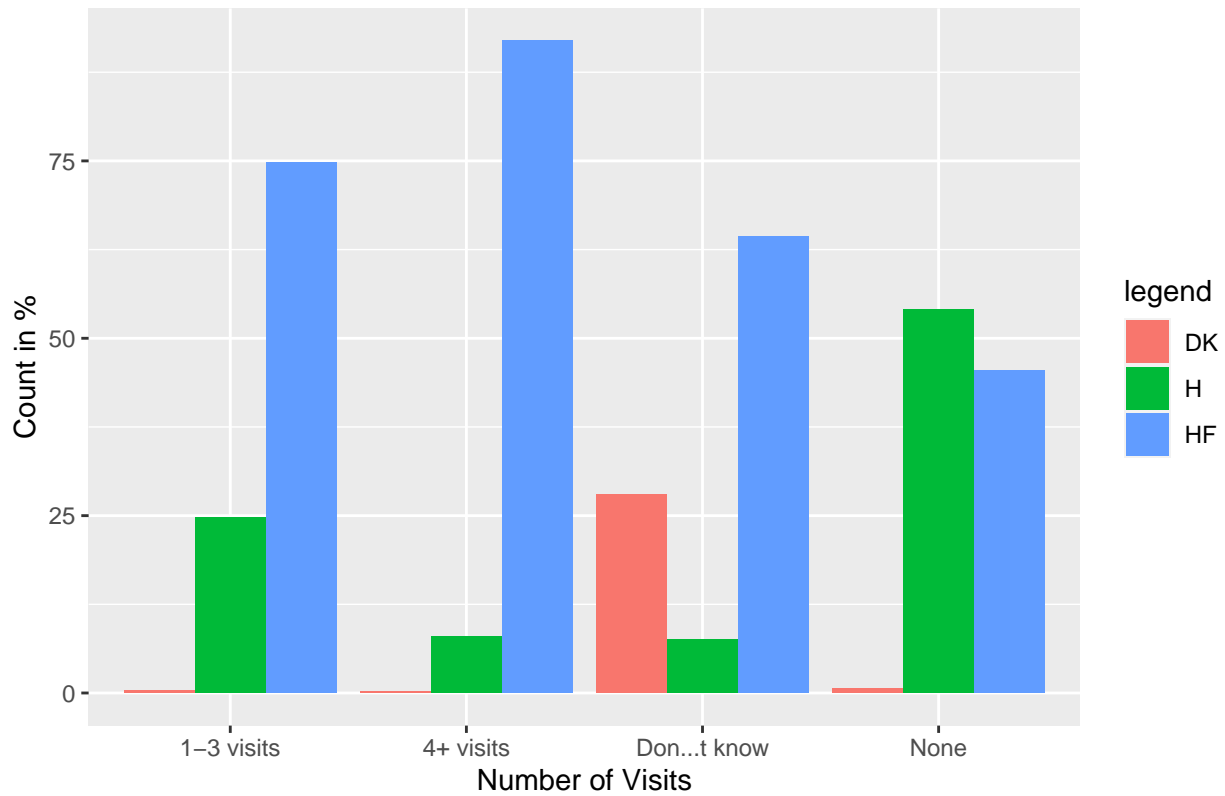
Figure 2: Place of Delivery by Mother's Educational level

Figure 6: Place of Delivery by Region

In figure 6 (Figure **??**) we visualize place of delivery by region in Turkey, every region except East has greater number of births delivered at health facilities.

Figure 7: Place of Delivery by AC visits

Lastly, in figure 7 (Figure **??**)we can conclude that patients with more antenatal care visits used health facilities, but patients with no visits had mostly at home deliveries. This may indicate a lack of access for those mothers. Overall, we can conclude that mothers with less education, living in rural conditions, and with more children tend to deliver their children at home, compared to mothers in urban areas with higher education.
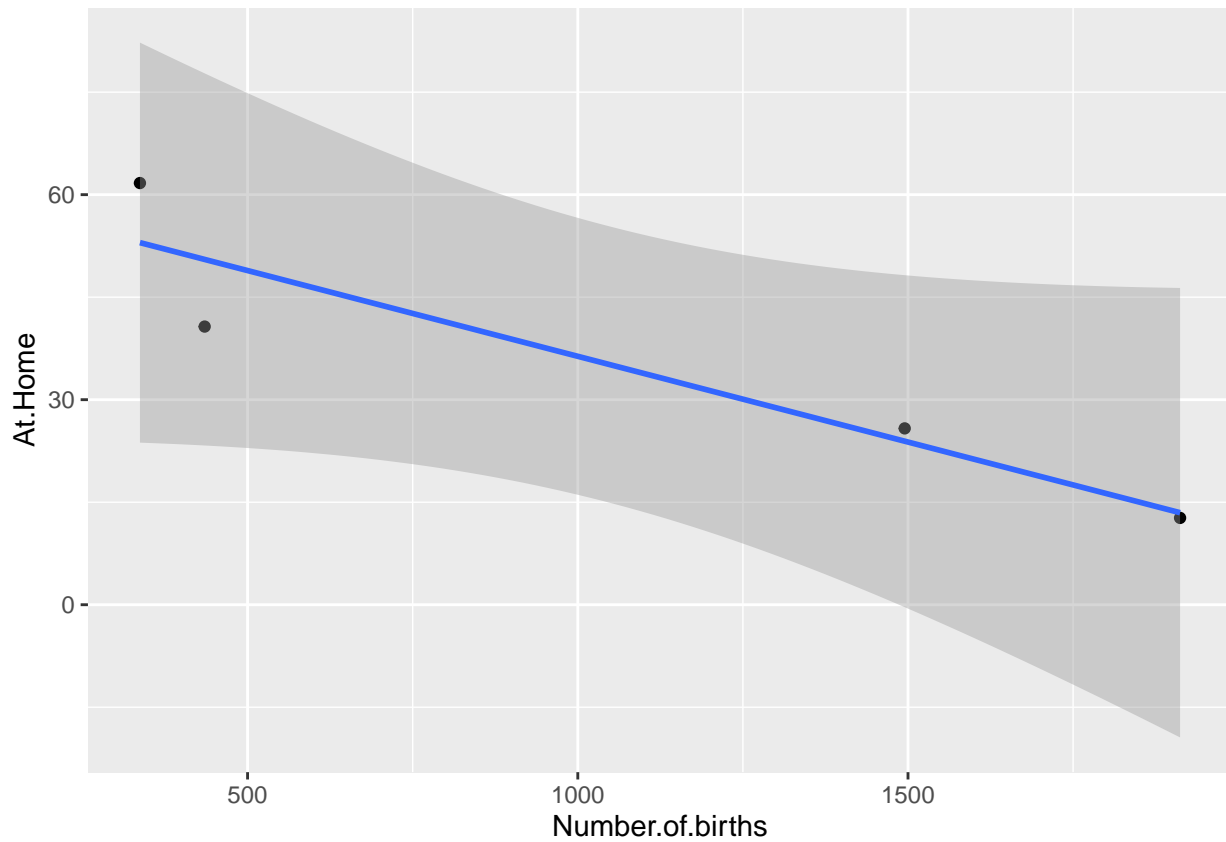
4 Discussion

4.1   First discussion point
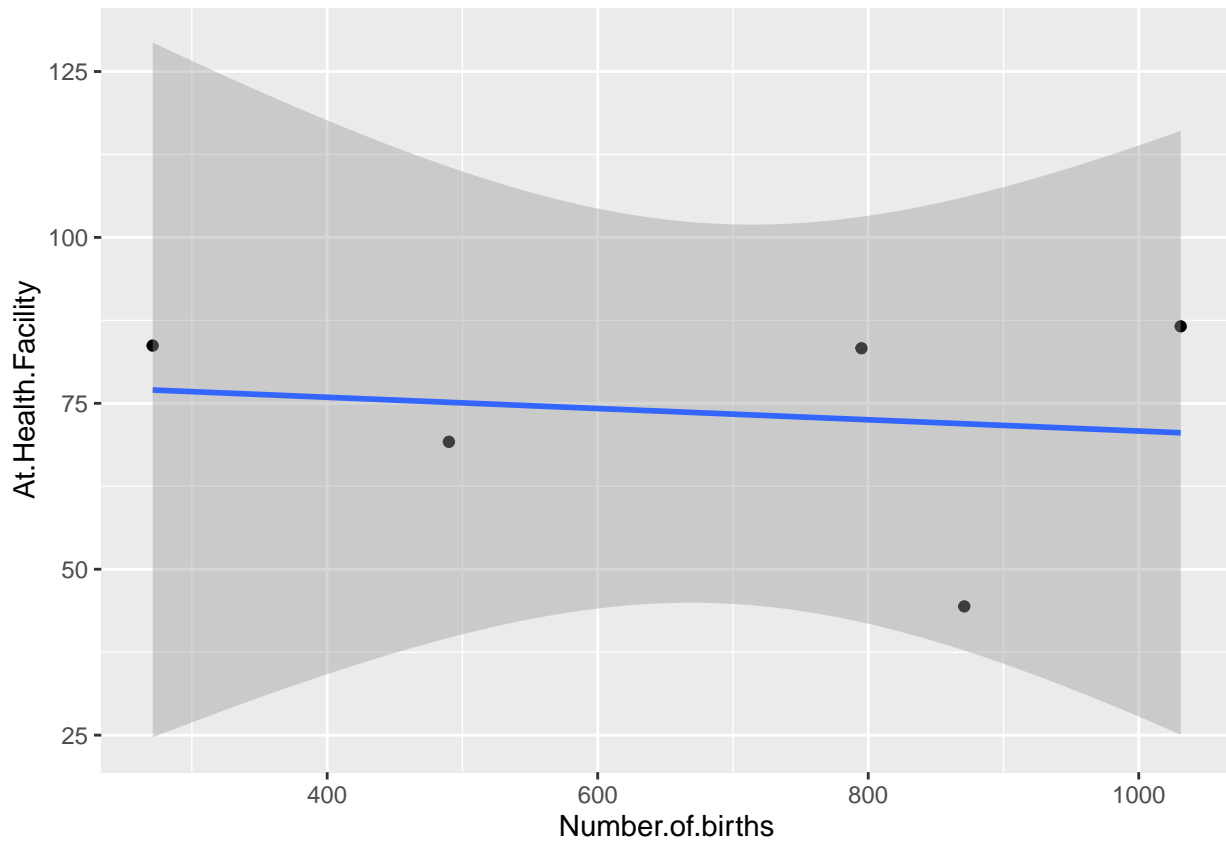
This plot depicts the number of births at a health facility using simulated data of regions in Turkey from 1998. The line of best fit is horizontal with a very slight negative relationship. Essentially, there is no clear pattern and thus region is not a significant indicator of place of delivery.

4.2 Second discussion point

This second plot shows the number of births and amount of births at home, based on birth order. This illustrates the trend from a simulation that only includes information on the birth order. There is a clear negative correlation, indicating that the number of at-home births decrease as the number of births grows. Meaning that the birth order with higher number of births, hence first births and the 2-3 birth, have lower amounts of home as the place of delivery.

4.3 Third discussion point
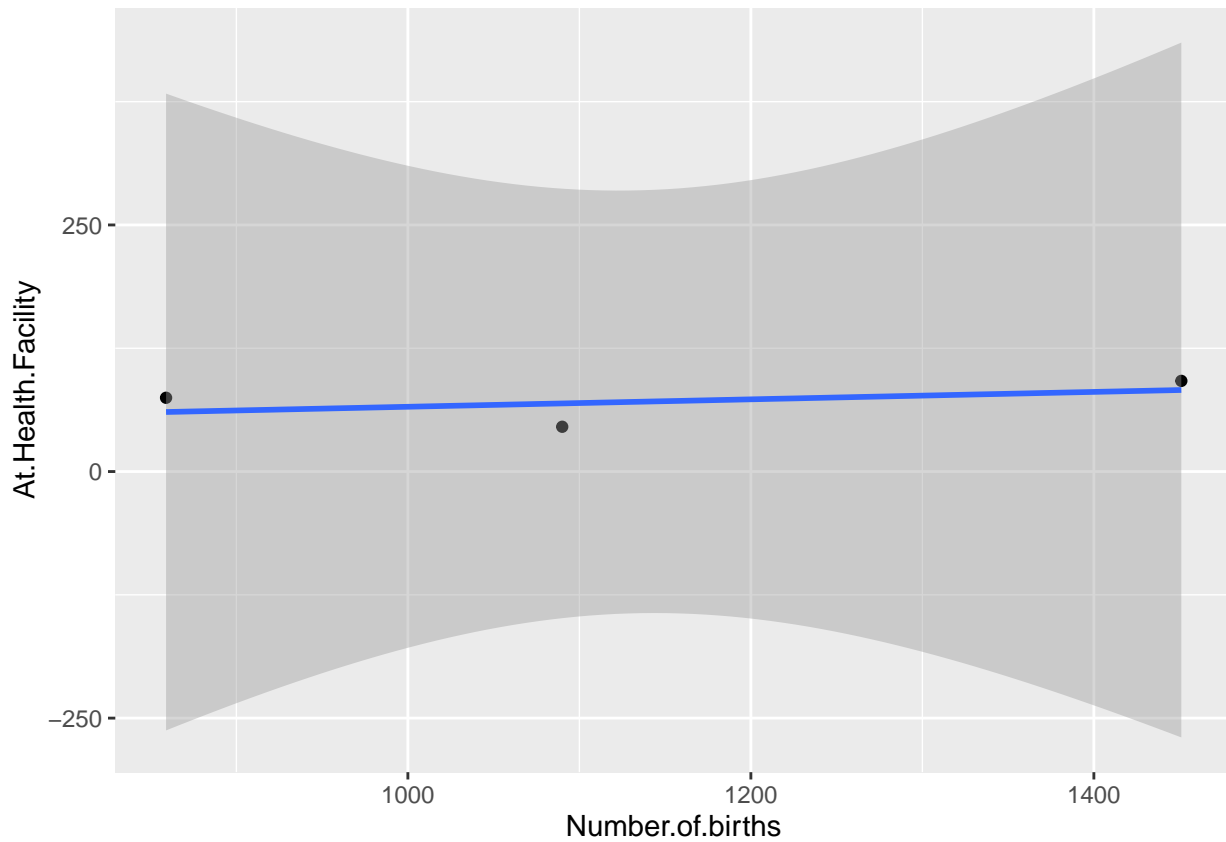
This plot showcases number of births at health facilities using simulations from antenatal care visits. The line of best fit is completely horizontal, meaning there is no trend between the variables. Antenatal visits has no correlation with number of births, it fluctuates regardless of amount. However, a higher amount of antenatal care visits means a larger portion of births at health care facilities.

4.4 Weaknesses and next steps

Some possible weaknesses in regards to this analysis include the lack of information on the socioeconomic conditions of the respondents of the survey. This could play a substantial role in where women deliver their children. Additionally, the survey does not provide any information about the conditions of at home births. Those births could include health professionals but still be categorized as an at home birth rather than a health facility.

Appendix

Motivation

1. *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.*
    - The dataset was created to evaluate hygiene conditions during delivery of babies in Turkey 1998. Information about background characteristics of the mother and infant provides useful insights into the development of children and future outcomes.
2. *Who created the dataset (for example, which team, research group) and on behalf of which entity (for example, company, institution, organization)?*
    - The dataset was created by research groups in Hacettepe University, Institute of population studies in collaboration with Macro International Inc.
3. *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.*
    - Technical and Financial support were provided both by the United Nations Population Fund (UNFPA) and primarily sponsored by the United States Agency for International Development (USAID)

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, photos, people, countries)? Are there multiple types of instances (for example, movies, users, and ratings; people and interactions between them; nodes and edges)? Please provide a description.*
    - The instances represent the background characteristics of children and mothers in Turkey in 1998. The types are Mother's age at birth, Birth order, Residence, Region, Education and Antenatal Care Visits.
2. *How many instances are there in total (of each type, if appropriate)?*
    - There are 21 instances.
3. *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 (for example, 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 (for example, to cover a more diverse range of instances, because instances were withheld or unavailable).*
    - Yes. The data combines the total percentage of responses and includes values for "don't know and missing values."
4. *What data does each instance consist of? "Raw" data (for example, unprocessed text or images) or features? In either case, please provide a description.*
    - The instances are set up based on the group of background characteristics defined.
        - The first background characteristic consists of mother's ages during giving birth split by ages below 20, 20-34, and 35 and above.
        - The second background characteristic consists of the order of birth of the child - first, second to third, fourth to fifth, sixth and above
        - The third background characteristic consists of the residential setting during birth - urban or rural
        - The fourth background characteristic consists of the region in Turkey during birth - West, South, Central, North and East
        - The fifth background characteristic consists of the educational level attained by the mother during time of birth - no education, primary education, secondary education and above
        - The final background characteristic consists of the number of Antenatal Care (AC) visits - 0, 1-3, 4+, don't know or missing
    - The columns are split based on the place of delivery - at a health facility, at home or don't know/missing. Additionally, the table accumulates the total number and proportion of births.
5. *Is there a label or target associated with each instance? If so, please provide a description.*
    - The first column of each instance corresponds to the background characteristic being measures, labelled accordingly.

6. *Is any information missing from individual instances? If so, please provide a description, explaining why this information is missing (for example, because it was unavailable). This does not include intentionally removed information, but might include, for example, redacted text.*
    - Data in the table is aggregated and hence, have proportions of missing values of place of delivery for each background characteristic. Values were missing because they were not answered for in the survey.
7. *Are relationships between individual instances made explicit (for example, users' movie ratings, social network links)? If so, please describe how these relationships are made explicit.*
    - There are no relationships between individual instances
8. *Are there recommended data splits (for example, training, development/validation, testing)? If so, please provide a description of these splits, explaining the rationale behind them.*
    - There are no recommended data splits
9. *Are there any errors, sources of noise, or redundancies in the dataset? If so, please provide a description.*
    - There are no errors, sources of noise or redundancies in the dataset.
10. *Is the dataset self-contained, or does it link to or otherwise rely on external resources (for example, 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 (that is, including the external resources as they existed at the time the dataset was created); c) are there any restrictions (for example, 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.*
    - The dataset is self-contained
11. *Does the dataset contain data that might be considered confidential (for example, 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.*
    - The data is aggregated and hence, does not affect confidentiality of individual responses.
12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting, threatening, or might otherwise cause anxiety? If so, please describe why.*
    - The dataset does not consist of data that might be offensive, insulting, threatening or axniety-inducing.
13. *Does the dataset identify any sub-populations (for example, by age, gender)? If so, please describe how these subpopulations are identified and provide a description of their respective distributions within the dataset.*
    - the dataset does not consist of any sub-populations.
14. *Is it possible to identify individuals (that is, one or more natural persons), either directly or indirectly (that is, in combination with other data) from the dataset? If so, please describe how.*
    - It is not possible to identify individuals directly or indirectly as the raw data is provided in aggregated format.
15. *Does the dataset contain data that might be considered sensitive in any way (for example, 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.*
    - The dataset doesn't consist of data that might be considered sensitive.
16. *Any other comments?*
    - None

Collection process

1. *How was the data associated with each instance acquired? Was the data directly observable (for example, raw text, movie ratings), reported by subjects (for example, survey responses), or indirectly inferred/derived from other data (for example, 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.*
    - The data associated was acquired through the survey conducted by DHS.
2. *What mechanisms or procedures were used to collect the data (for example, hardware apparatuses or*

sensors, manual human curation, software programs, software APIs)? How were these mechanisms or procedures validated?

- Manual human curation was used to collect the data

3. *If the dataset is a sample from a larger set, what was the sampling strategy (for example, deterministic, probabilistic with specific sampling probabilities)?*

-

4. *Who was involved in the data collection process (for example, students, crowdworkers, contractors) and how were they compensated (for example, how much were crowdworkers paid)?*

- University students across all universities in Ankara, Turkey was primarily used to collect data. No information was reported about compensation.

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

- The data was collected from the first week of August 1998 to the end of November 1998.

6. *Were any ethical review processes conducted (for example, 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.*

- Ethical review processes were not conducted

7. *Did you collect the data from the individuals in question directly, or obtain it via third parties or other sources (for example, websites)?*

- We obtained the data via the Demographic and Health Surveys website: dhsprogram.com

8. *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.*

- The interviews with data collectors were conducted on a voluntary basis. No notice is available.

9. *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.*

- The individuals consented to the collection and use of their data. No information of consent agreement language is provided.

10. *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 mechanism to revoke consesnt was not provided

11. *Has an analysis of the potential impact of the dataset and its use on data subjects (for example, 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.*

- No analysis of the potential impact of the dataset and use on data subjects was conducted.

12. *Any other comments?*

- None

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, 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.*

- The data was obtained originally in the PDF format. The table was extracted manually to a usable dataset in Excel by entering data into a .csv format. This usable dataframe was applied in RStudio for conducting analysis.

2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for example, to support unanticipated future uses)? If so, please provide a link or other access point to the "raw" data.*

- The raw data obtaines was saved in inputs/data/Turkey1998.csv

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

- R Software - https://www.R-project.org/

4. *Any other comments?*
   - None

Uses

1. *Has the dataset been used for any tasks already? If so, please provide a description.*
   - No
2. *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.*
   -
3. *What (other) tasks could the dataset be used for?*
   - The datset could be used to examine other factors regarding birthplace delivery characteristics
4. *Is there anything about the composition of the dataset or the way it was collected and prepro-cessed/cleaned/labeled that might impact 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 (for example, stereotyping, quality of service issues) or other risks or harms (for example, legal risks, financial harms)? If so, please provide a description. Is there anything a dataset consumer could do to mitigate these risks or harms?*
   - None
5. *Are there tasks for which the dataset should not be used? If so, please provide a description.*
   - The dataset should not be used to examine factors other than place of delivery at birth in Turkey in 1998.
6. *Any other comments?*
   - None

Distribution

1. *Will the dataset be distributed to third parties outside of the entity (for example, company, institution, organization) on behalf of which the dataset was created? If so, please provide a description.*
   - No, the dataset is only avaiable for personal uses only.
2. *How will the dataset be distributed (for example, tarball on website, API, GitHub)? Does the dataset have a digital object identifier (DOI)?*
   - The dataset will be distributed using Github.
3. *When will the dataset be distributed?*
   - The dataset will be distributed in April 2022.
4. *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.*
   - The dataset will be released under the MIT license
5. *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.*
   - No
6. *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.*
   - No
7. *Any other comments?*
   - None

Maintenance

1. *Who will be supporting/hosting/maintaining the dataset?*
   - Shoubhik Bhattacharya, Ayoon Kim and Meha Grewal will be responsible for supporting, hosting and maintaining this dataset.
2. *How can the owner/curator/manager of the dataset be contacted (for example, email address)?*

- Can be contacted via Github

3. *Is there an erratum? If so, please provide a link or other access point.*
   - No

4. *Will the dataset be updated (for example, to correct labeling errors, add new instances, delete instances)? If so, please describe how often, by whom, and how updates will be communicated to dataset consumers (for example, mailing list, GitHub)?*
   - Not at the moment.

5. *If the dataset relates to people, are there applicable limits on the retention of the data associated with the instances (for example, 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.*
   - Since we use aggregate data, the data does not relate to a specific individual.

6. *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.*
   - No

7. *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.*
   - No mechanisms are in place as of now

8. *Any other comments?*
   - None

A   Additional details

References

Hacettepe University, Macro International, Institute of Population Studies/Turkey. 1998. "Turkey DHS - Final Report." *Demographic and Health Surveys.*

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis.*

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.

Xie, Yihui. 2022. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.*