

Korean Demographic Cliff Social Impact**

future issue

Youngho Kim

Invalid Date

Korea is facing an unprecedented demographic cliff. Moreover, this is reported and discussed publically by popular media. In this paper, we will look at the current status of Korea demographic, as well as psychological and economic connection of population cliffs to society.

Table of contents

1	Introduction	2
2	Data	3
2.1	Data Sources	3
2.2	Data Cleaning	4
2.3	Data Visualization	4
3	Model	5
4	Result	6
4.1	Population Trend	6
4.2	Social Survey	6
5	Discussion	6
	Refereces	10

*Code and data are available at: <https://github.com/Alicia-y/Telling-stories-with-data-final-paper>

1 Introduction

Population is the most fundamental element of society. A stable demographic structure is the basis for economic growth and determines the form of social structure. But the population structure doesn't form overnight, rather it's accumulation of birth, natural mortality, and the immigration over decades. And it's influenced heavily from political decisions, social movements, and the trend of the times. Population decline is a common phenomenon in developed countries that are on an economically stable track, and Korea, which has made remarkable economic growth in the 50 years since the end of Korean War, could not avoid it. In fact, as a counter reaction of rapid growth, Korea is suffering more deadly and rapid population decline that was never found in any other country. The current situation in Korea is so serious that the term "Demographic Cliff" created by Harry Dent is not difficult to find in public media.

The "Demographic Cliff" is a theoretical term used to describe a social situation in which a country or region experiences a significant decline in its population due to low birth rates and an aging population. This can lead to a range of economic and social challenges, including a shrinking workforce, increased healthcare costs, and reduced consumer spending. The term was popularized by author and demographer Harry Dent, who has argued that demographic trends are a key driver of economic cycles. However, some economists and demographers have criticized Dent's theories, arguing that other factors such as technological innovation and government policy are also important drivers of economic growth. Also, many modern countries like Germany and Japan, does not followed the Harry Dent's Demographic Cliff model. Therefore it remained as a theoretical model to predict population change in limited circumstances until now Korea's population decline happened.

The population decline in Korea is so rapid that it has become a social phenomenon that anyone can easily observe. From elementary school to university, many schools regardless their the level of education in the metropolitan area are closing due to a lack of students, and national security is threatened by a lack of conscription population. As population issues began to become increasingly visible, they began to negatively affect other social issues, including gender conflicts, generational conflicts, regional disparities, and housing issues. This document will focus on how the popular introduction of rapid population decline influenced people's perception of daily life and social problems.

The data used in this report is obtained from Open API KOSIS("Kosis Korean Statistical Information Service," n.d.), Korean Statistical Information Service. The data is analyzed with R(R Core Team 2022) and Rstudio(RStudio Team 2020) and analyzed with various packages. readxl(Wickham and Bryan 2023) and here(Müller 2020) packages are used to retrieve attained data from KOSIS("Kosis Korean Statistical Information Service," n.d.) to the environment. tidyverse(Wickham et al. 2019), dplyr(Wickham et al. 2023), janitor(Firke 2023), and knitr(Xie 2014) packages are used to manipulate the data in a fit for research purposes. kableExtra(Zhu 2021), patchwork(Pedersen 2022), and scales(Wickham and Seidel 2022) are

used to shape the result into intuitive visual format. For details of each data and the data cleaning process that is tailored to the purpose of the document can be found in Section 2.

Linear regression model is used in this paper to get an insight of public perception and current trend of Korea demographic changes. Section 3 describes how this model works in purpose of this research and the result of the implication of the model on the data. Section 4 discuss the result of data manipulation with organized result of the data. The discussion is carried out in Section 5 about findings from the result regarding social and political impact of current phenomenon, and suggested reaction to it. It also discusses about potential weaknesses and further reaserches of this paper.

2 Data

2.1 Data Sources

The data used in this report is obtained from Open API KOSIS, Korean Statistical Information Service. The data named “Summary of Census Population(by administrative districtsexage) 1960-2010”is demographic data from the census, which was implemented every five years from 1960, five years after the Korean War armistice, to 2010. This data were used to predict the current population status of South Korea and changes in the number of people over the next 30 years, along with the census’s demographics, which were implemented annually from 2015 to 2020 under a changed policy. The second annual census data is named “Population, Households and Housing Units 2015-2020”. The analized result can be found in Section 4.1.

Various survey data were also researched to gain insight into the current social perception of the childbearing age population. The data named “Class_Awareness__19_years_old_and_over” is a survey of where each individual thinks their social position is. It’s important to note that this is not data based on actual capital, occupation, or social influence, but rather a questionnaire about one’s position that an individual feels when considering various factors. The data named “Degree_of_Stress__General_Life__13_years_old_and_over” is a survey of the stress that each individual feels in their overall lives. This data is also a survey of the frequency of stress that individuals feel subjectively, not based on numerical data of work burden, leisure time, or wealth of assets. The data named “Environment_in_the_future__As_Compared_with_5_Years_Ago__13_years_old_and_over” is a prediction of changes in the future environment. This is also a survey of individual guesses, not scientific analysis data through factors such as air quality, degree of industrialization, or future technological development. This data shows how optimistic or pessimistic the Korean people are about future environmental changes. It also indirectly expresses individual thoughts on the current environmental condition of Korea. As environmental climate issues are highlighted in the public, these data are included because the impact of environmental awareness on individual future life design is increasing. The data named “Level_of_Satisfaction_on_Work_Conditions__Efforts_to_prevent_Violence__13_years_old_and_over”

is a satisfaction survey of the workplace environment. Since Korea is mentioned every year as one of the most working hours country within OECD countries, satisfaction in the workplace environment is a very important data to learn about the public's social perception. The data named "Satisfaction_with_Achievement__19_years_old_and_over" is a survey of individuals' achievements in their lives to date that they feel. This is similar to the daily life satisfaction survey, but it does not focus on the present life, but on the satisfaction of the past life so far, and the expectation of the future of how much more you can achieve in the future. The data named "Satisfaction_with_Life__13_years_old_and_over" is currently a survey of individual life satisfaction. This is a direct survey of the level of satisfaction with current life. However, it is not a survey of the current standard of living, but a measure of satisfaction with it, and it is a data to find out how people feel about their current living conditions, not the living conditions themselves. The data named "Selfassessment_of_Health__13_years_old_and_over" is about how individuals feel about their health. This is also subjective data on the health status of individuals, not objective data calculated by examining national health care, access to medical services, or disease incidence. This indirectly shows satisfaction and reliability with the current level of medical services in Korea. The data named "Sentiment_in_the_Present_Surrounding_Environment" is a survey of how individuals feel about the current environmental conditions in Korea. This directly represents the level of people's awareness of the current environment of the Republic of Korea as they focused on the present, not on speculation and expectation of the future. The data named "Trust_in_Korean_Society"

2.2 Data Cleaning

The obtained data from KOSIS([cite-kosis?](#)) is cleaned to meet the purpose of focus of this paper. For most of the dataset, unnecessary categories have been eliminated as the point of this paper is aimed at the social awareness of childbearing age citizens. So that the revised figures can provide the most accurate information according to our analysis and not have any information that is irrelevant to our analysis. In addition, general demographic data has been modified to a suitable form for applying the linear regression model used. The modification includes eliminating district specific data, eliminating undefined data, eliminating data summaries, and merging the different forms of census before and after 2015.

2.3 Data Visualization

The sorted data is summarized in graphs and tables for easy visibility. Census data, a vast quantitative data, are represented at a glance by representing it in linear graphs, and surveys based on individual responses are abbreviated into tables, transforming them into distinctions between important categories. This visual data, which will be discussed in greater detail later in the Section 3 and Section 4 section, has provided a deep understanding of the agenda and opened up the possibility of an insightful discussion on Section 5.

3 Model

A linear regression model was used to estimate the trend of population change in Korea and to predict population change over the next 30 years. The most important thing about predicting population change is to remind you that population change can change as roughly as much as possible and escape speculation with big errors. Birth rates, mortality, immigration, etc. that directly affect the population are easily measurable, but the second level factors that affect these first level factors, such as economic stability felt by young people, social views of underage births, parenting burdens in families and societies, and a country's social and cultural image. Of course, it is also absurd to say that all of this is absolutely impossible to measure, but it is also almost impossible to put all of this into the model and include it in the calculation.

The forecast for population change in this document is limited to 30 years, a relatively short period unlike other population-related studies, because population change itself is not the core content of this document, along with the preceding reasons. The purpose of predicting the trend of population change in this document is not to predict the exact figure, but to understand the overall direction and trend of the current population change in Korea. Through this, the focus was on sharing and understanding the feelings and rough perceptions of the public in Korean society, which is currently facing a population cliff.

Linear regression is a statistical method used to model the relationship between one or more independent variables and the resulting variable. Given sufficient independent variables in measuring the association between the two, you can predict future variations in the variable. In this paper, simple linear regression is selected among linear regression models for estimating population fluctuations because the accuracy of predictions is not of great importance and its purpose is to determine trends. Simple linear regression has only one independent variable, making it easy to apply the model and intuitive to understand. However, since only one independent variable is considered, the results can be significantly fluctuated for other external factors, and it is difficult to speculate on variable data because it assumes a linear relationship between the two variables. The trend of population fluctuations in this study is to see its direction rather than to derive accurate estimates, also since the estimation period is limited simple linear regression analysis is very suitable .

The goal of linear regression is to find the best-fitting line that minimizes the sum of squared errors between the predicted values and the actual values of the dependent variable. This line is represented by a linear equation in the form of following

$$y = b_0 + b_1x + e$$

where Y represents the dependent variable, that is, the expected number of people, and X is the independent variable, in this case, the time in years. b_0 and b_1 are the coefficients or parameters of the model, and e is the error term.

The coefficients in the model represent the slope of the line or hyperplane and the intercept using the y-axis. This is the best value calculated from the previous data. For 1, which represents the slope, it represents the degree to which the dependent variable changes when the independent variable changes on a unit basis. For 0, which represents an intercept, it represents the value of the dependent variable when the independent variable is 0.

The results of the model application can be found on Figure 2 in the Section 4.1.

4 Result

4.1 Population Trend

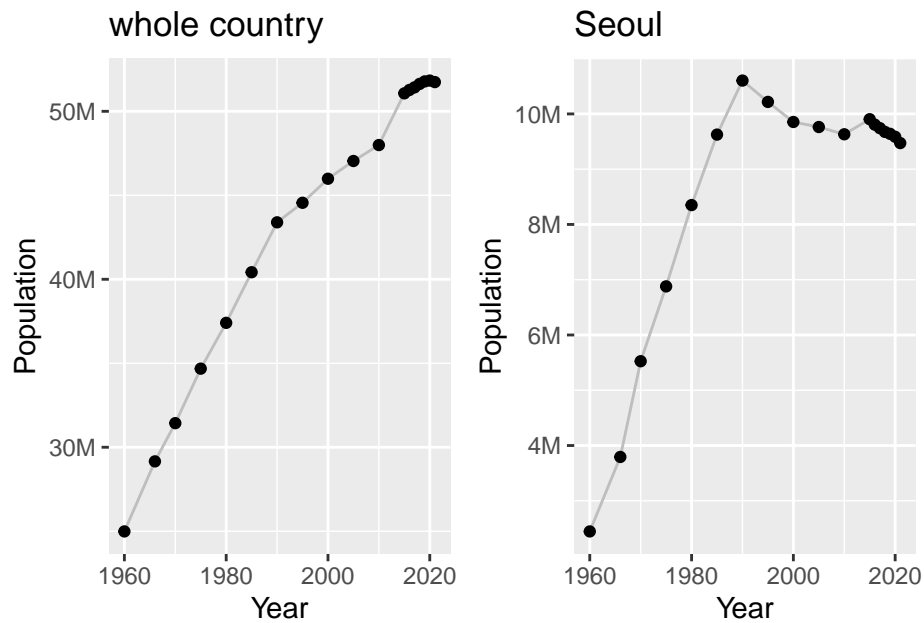


Figure 1: Population changes from 1960 to 2020

4.2 Social Survey

5 Discussion

asasgasg

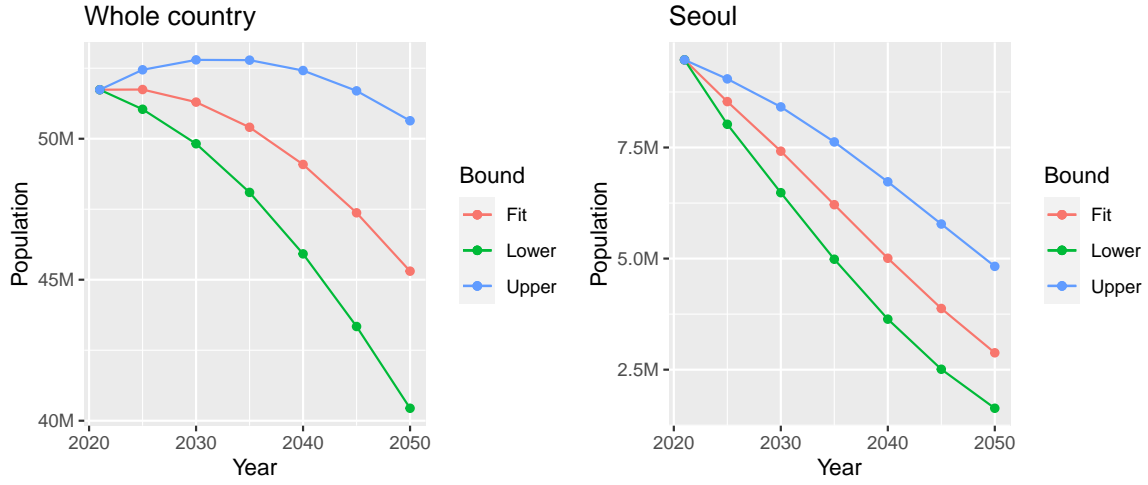


Figure 2: Population Trend Prediction from 2020 to 2050

Work Satisfac- tion	Total	Very satisfied	Moderately satisfied	Neither satisfied nor dis- satisfied	Moderately dissatis- fied	Very dis- satisfied	Dont Know
Total	100	22.0	28.0	38.0	5.0	2.3	4.7
20~29 years old	100	24.3	26.2	36.0	5.0	2.7	5.8
30~39 years old	100	23.9	27.6	36.5	5.6	2.5	3.9

Figure 3: Social Perception Survey

Enviroment in Future	Total	Much better	Slightly better	No changes	Slightly worse	Much worse
Total	100	5.9	33.9	47.1	11.3	1.7
20~29 years old	100	7.5	33.4	47.8	9.6	1.7
30~39 years old	100	7.7	37.3	44.3	8.9	1.8

Figure 4: Social Perception Survey

Sentiment in Sur- rounding	Total	Very good	Somewhat good	Average	Somewhat bad	Very bad
Total	100	12.3	30.0	40.1	14.4	3.2
20~29 years old	100	11.2	26.9	42.9	15.2	3.9
30~39 years old	100	9.3	26.8	40.9	17.1	5.8

Figure 5: Social Perception Survey

Class Awareness	Total	Upper	Middle	Lower
Total	100	2.7	58.8	38.5
19~29 years old	100	2.6	59.1	38.3
30~39 years old	100	2.7	64.8	32.4

Figure 6: Social Perception Survey

Trust in Society	Total	Very Trustworthy	Moderately Trustworthy	Little Trustworthy	Never Trustworthy
Total	100	4.5	51.1	39.1	5.3
20~29 years old	100	3.9	44.0	43.7	8.3
30~39 years old	100	4.2	48.0	40.7	7.0

Figure 7: Social Perception Survey

Life Satis- faction	Total	Very Satisfied	Moderately Satisfied	Neither satisfied nor dissat- isfied	Moderately Dissatis- fied	Very Dis- satisfied
Total	100	12.1	31.2	42.6	11.6	2.5
20~29 years old	100	15.5	32.2	40.1	10.1	2.1
30~39 years old	100	13.5	32.3	40.9	10.6	2.8

Figure 8: Social Perception Survey

Achievement Satisfac- tion	Total	Very Satisfied	Moderately Satisfied	Neither satisfied nor dissat- isfied	Moderately Dissatis- fied	Very Dis- satisfied
Total	100	7.5	27.5	46.1	15.1	3.8
19~29 years old	100	8.5	24.8	47.1	14.9	4.6
30~39 years old	100	8.9	28.7	43.0	14.6	4.8

Figure 9: Social Perception Survey

Degree of Stress	Total	Severe	Moderate	Weak	None
Total	100	4.1	40.8	44.2	10.8
20~29 years old	100	3.5	36.1	46.1	14.3
30~39 years old	100	4.8	46.7	40.8	7.7

Figure 10: Social Perception Survey

Self Health As- sessment	Total	Very good	Good	Average	Bad	Very bad
Total	100	11.5	41.6	34.9	10.8	1.2
20~29 years old	100	23.2	49.1	23.9	3.7	0.1
30~39 years old	100	15.2	50.2	29.7	4.6	0.2

Figure 11: Social Perception Survey

Refereces

- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://CRAN.R-project.org/package=janitor>.
- “Kosis Korean Statistical Information Service.” n.d. *Open API KOSIS*. <https://kosis.kr/eng/>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- Pedersen, Thomas Lin. 2022. *Patchwork: The Composer of Plots*. <https://CRAN.R-project.org/package=patchwork>.
- R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- RStudio Team. 2020. *RStudio: Integrated Development Environment for r*. Boston, MA: RStudio, PBC. <http://www.rstudio.com/>.
- Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Golemund, et al. 2019. “Welcome to the tidyverse.” *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>.
- Wickham, Hadley, and Jennifer Bryan. 2023. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.
- Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.
- Wickham, Hadley, and Dana Seidel. 2022. *Scales: Scale Functions for Visualization*. <https://CRAN.R-project.org/package=scales>.
- Xie, Yihui. 2014. “Knitr: A Comprehensive Tool for Reproducible Research in R.” In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.