**Impact of demographic factors on naming trends**

Abbas Khanzada

# **Abstract**

This study investigates the nature of the association between demographic factors (gender and ethnicity) and baby naming trends alongside the impact that the year of birth has on individuals’ names. The results from the tests showed that baby naming trends vary across demographic groups, and that the year of birth impacts individuals first names.

# **Introduction**

This study aimed to determine the relationship between demographic factors (gender and ethnicity) and frequency of baby names and whether an individuals’ first name depended on the year in which they were born. The findings from this project are valuable insights into cultural trends, social norms, and naming patterns among others. The implication of this study is for parents, policymakers, and marketers, among others. This study used a data visualization tool (grouped bar chart) and inferential statistics (ANOVA, t-test, and Chi-square test). After conducting the inferential tests, we found out that there was a significant association between variables. The results from the t-test showed a significant difference in the number of babies with the same name between males and females. An ANOVA test showed a significant association between ethnicity and the number of babies with the same name. This highlights the influence of ethnic groups’ culture on naming. The Chi-square test showed that an individuals’ first name depended on the year of birth.

In the other sections below (from background to conclusion), this document has provided a detailed explanation of why each method was used, and the study’s importance to the interpretation of the results. The sections further highlight what the stakeholders stand to gain from the findings of these studies.

# **Background**

The aim of this study is to investigate the existence of a potential relationship between the social-demographic factors (gender and ethnicity) and the frequency of a baby’s name. This study also investigated whether the first name of the participant was associated with the year that they were born. The study of the relationship between demographic factors and the frequency of baby names will provide valuable insights into social norms, cultural trends, and baby-naming biases. The information and insight derived from this study will be pertinent to a wide range of stakeholders, who include parents, sociologists, legislators, and marketers. The parents may use this information to make informed decisions concerning the naming of their children considering factors like popularity and cultural significance. The results from this study could also be used by sociologists to understand the patterns of cultural assimilation and to identify the evolution of naming practices among the various ethnic groups. Marketers can use the findings from this study to tailor their products and services to specific customer demographic categories. Policymakers may use the information from this study to consider the implications of naming trends on equality, social cohesion, and discrimination.

Prior studies have shown that there is a significant association between cultural, historical, and economic factors and naming customs. There are certain names that might be prevalent in particular ethnic groups due to cultural beliefs. The attitude towards the allocation of names to female and male children might be affected by the views of society and culture regarding gender. This study has incorporated theoretical principles like symbolic interactions, socialization, and identity construction among others.

# **Data and methods**

In this study, we used the popular data names dataset. Data was collected through civil birth registration. Each record represents the ranking of a baby name in the order of frequency. This data was intended to be used to represent the popularity of a name and factors associated with them. The data has got 6 variables and 57582 observations. Out of the 6 variables, this study only focused on five variables: Sex, ethnicity, year of birth, first name and number of babies with that name. The description of the variables is as shown in the table below:

|  |  |
| --- | --- |
| **Column Name** | **Column Description** |
| Year of Birth | Year the baby was born |
| Sex | Sex of the baby |
| Ethnicity | Mother's Ethnicity |
| Baby's First Name | Baby's first names |
| Count | Number of babies with this name |

This study intends to use statistical techniques to determine whether there is a significant association between number of babies with similar name and sex and ethnicity. This study used a two-sample independent t-test to determine whether there was a significant association between the number of babies with same name and sex. This was the most appropriate test since the dependent variable “number of babies” was measured using a numeric scale while the independent variable “sex” was measured using a nominal scale and it had only two categories. This study used ANOVA test to determine whether there was a significant association between the number of babies with similar names and ethnicity. This was the most appropriate test since the dependent variable “number of babies” was measured using a numeric scale while the independent variable “ethnicity” was measured using a nominal scale and it had more than two groups.

This study used a Chi-square test to determine whether there was a significant association between the year of birth and the first name of the participant. This test was appropriate since both variables were non-numeric.

# **Analytical results**

Descriptive statistics

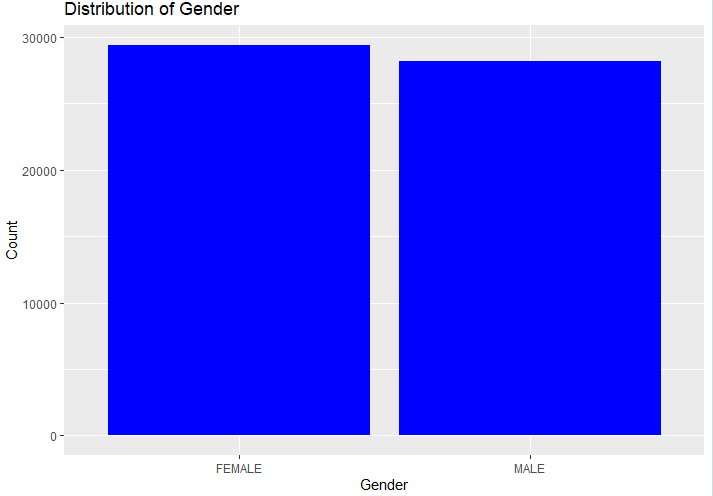


Figure 1

Based on the results from figure 1 above, there were 51.06% (n=29402) female and 48.94% (n=28180) male names that were involved in the study.

Figure showing distribution of names by ethnicity

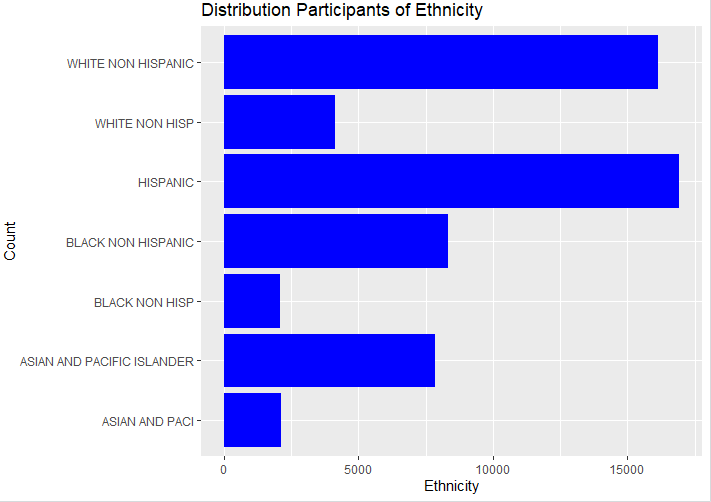


Figure 2

Based on the results from figure 2 above, most of the participants were either Hispanic or white non-Hispanic. There were very few Asian, White Non-Hispanic and Black Non-Hispanic participants.

**Ethnicity and Number of babies with that name**

Table showing the ANOVA tables

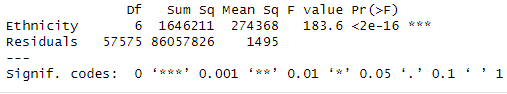


Table 1

ANOVA test was conducted to determine whether there was a significant association between ethnicity and number of babies with the same name. The results from the ANOVA table show that there was a significant association between the number of babies with same name and ethnicity, F (6, 57575) = 183.6, p < 0.05.

**Gender and Number of babies with that name**

Table showing the t-test table

Welch Two Sample t-test

data: Count by Gender

t = -30.102, df = 49972, p-value < 2.2e-16

alternative hypothesis: true difference in means between group FEMALE and group MALE is not equal to 0

95 percent confidence interval:

-10.426453 -9.151672

sample estimates:

mean in group FEMALE mean in group MALE

29.13894 38.92800

We conducted a two-sample t-test to determine whether there was a significant difference in the average number of babies with the same name between male and female participants. The results of the t-test indicated a significant difference between male (*M* = 38.92) and female (*M* = 29.14) in the average number of babies with the same names; t (49,972) = -30.10, *p* < 0.05.

**Year of birth and child first name**

Table showing the Chi square test results

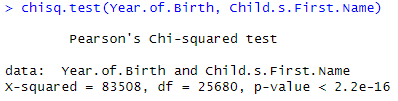


Table 3

A chi-square test was used to check whether there was a significant association between year of birth and child’s first name. Both the variable year and child’s first name had more than 2 groups. The contingency table analysis showed a significant association between the between year of birth and child’s first name, χ² (25680, N=57582) = 83508, *p <* 0.05. The association indicates that the child’s first name was dependent on the year of birth.

# **Conclusion**

The main aim of this study was to determine the relationship between demographic factors and number of babies with same name and whether the name of the participant was dependent on the year of birth. The results from the t-test showed that there was a meaningful association between gender groups and the similarity in the names. This indicates that the gender of a person was a critical determinant of the name that they will be given.

The results from the ANOVA test showed that there was a meaningful relationship between the ethnic group of a person and their name. These results are consistent with the findings by Ogihara (2021) which showed that culture which is ethnic based has a significant impact on how the kids are named. From this study, we can conclude that ethnicity plays an important role in a child’s naming.

The results from the Chi-square test of independence showed that the first name of the individuals was dependent on the year that that person was born. This finding can be explained by the results from a study conducted by Zelinsky (2010) which showed that certain names experience cyclical patterns with some names experiencing peaks and troughs in their popularity over time. From this study, there is enough evidence to conclude that the first name of an individual is dependent on the year that they were born.

# **Reference**

Zelinsky, W. (2010). Cultural variation in personal name patterns in the Eastern United States. *Annals of the Association of American Geographers*, *60*(4), 743-769.

Ogihara, Y. (2021). Direct evidence of the increase in unique names in Japan: The rise of individualism. *Current Research in Behavioral Sciences*, *2*, 100056.