**Statistical Analysis Report**

**Introduction:**

In this study, data gathered from the "Pulse Survey 2022" data source are statistically analyzed. The analysis focuses on a number of variables that shed light on the activities, educational attainment, demographics, and health-related characteristics of respondents. The questionnaire, data source, and a brief explanation of the variables are all listed below.

**Data Source:**

This analysis's data came from the "Pulse Survey 2022." The Pulse Survey is a thorough dataset that gathers responses from a wide range of people, providing insightful data on numerous facets of their lives.

**Questionnaire and Variables:**

The Pulse Survey 2022's questionnaire covers a wide range of issues, such as socioeconomic variables, health, and daily activities. The following variables were chosen for this analysis:

* ***ACTIVITY2****:* A categorical variable describing respondents' activities.
* ***EEDUC****:* A categorical variable representing respondents' educational levels.
* ***EGENID\_BIRTH****:* A categorical variable indicating the generation or age group of the respondents.
* ***MOBILITY:*** A categorical variable assessing respondents' mobility.
* ***MS***: A categorical variable describing marital status.
* ***REGION:*** A categorical variable identifying the region of the respondents.
* ***RHISPANIC:*** A categorical variable indicating Hispanic ethnicity.
* ***RRACE:*** A categorical variable representing respondents' race.
* ***TBIRTH\_YEAR:*** A categorical variable indicating the year of birth of the respondents.
* ***THHLD\_NUMPER:*** A categorical variable describing the number of people in the household.

**Data Preparation:**

Before conducting the analysis, some data preparation steps were taken. These included handling missing values by removing rows with missing data to ensure the quality of the analysis.

# Handling Missing Values

hps\_data <- na.omit(hps\_data)

I use these lines in my code to remove the missing values.

**Analysis and Results**

ACTIVITY2

The results of the summary of the "activity2" variable are shown below:

Counts.Var1: The dataset contains 2,580 instances of the value -99, or around 3.44% of all observations.

Counts.Freq: The dataset contains the value -88 4,075 times, or around 5.43% of all observations.

Percentages.Freq.Var1: The "activity2" variable's most common value, 1, appears 26,663 times and accounts for about 35.55% of the dataset.

Percentages.Freq.Frequency: The "activity2" variable's value 2, which appears 41,677 times and accounts for roughly 55.57% of the dataset, is the second most common value.

The distribution and occurrence of values inside the "activity2" variable are discussed in this summary. Understanding the make-up of your data is crucial, and it can be helpful.

EEDUC

Summary for EEDUC

Value 1: Appearances 520 times, or approximately 0.69%.

Value 2: Appearances 1,146 times, or roughly 1.53%.

Value 3: Shows up 8,396 times, or around 11.20%.

Value 4: Shows up 15,225 times total, or around 20.30%.

Value 5: Shows up 7,482 times and makes up around 9.98%.

Value 6: Shows up 22,086 times, or around 29.45%.

Value 7: Shows up 20,140 times, or roughly 26.85% of the time.

The "EEDUC" variable seems to indicate various categories or levels of education that are distributed over various levels.

EGENID\_BIRTH

Value 1: Recurs 30,672 times, or around 40.90% of the total.

Value 2: Recurs 44,323 times, or about 59.10% of all occurrences.

It appears that the "EGENID\_BIRTH" variable has two different values, which could signify binary categories.

MOBILITY

Value -99: 344 instances, or approximately 0.46%, are found.

Value -88: Occurs 7,842 times, or 10.46% of the time.

Value 1: Shows up 52,350 times, or around 69.80%.

Value 2: Appearances 11,508 times, making up around 15.35%.

Value 3: Shows up 2,661 times, or about 3.55% of the time.

Value 4: Occurs 290 times and makes up roughly 0.39%.

The "MOBILITY" variable seems to indicate categories of mobility, with values denoting various degrees of mobility or circumstances.

MS

The MS variable describes marital status. The analysis demonstrated:

Value -99: 691 instances, or 0.92%, are represented by this value.

Value 1: Recurs 40,638 times, accounting for about 54.19%.

Value 2: Shows up 3,908 times, or about 5.21% of the time.

Value 3: Shows up 11,962 times and makes up roughly 15.95%.

Value 4: Shows up 1,500 times and accounts for about 2.00%.

Value 5: 16,296 instances, or approximately 21.73%, are present.

The "MS" field seems to indicate marital status-related categories, with values denoting various marital statuses.

RHISPANIC

The RHISPANIC variable indicates Hispanic ethnicity. The results revealed:

17.35% is represented by value 1.

Represents 31.38% as Value 2.

Value 3: Denotes a 20.45% value.

Represents 30.82% as value 4.

With respect to each value's % contribution to the dataset, the "Region" variable appears to reflect several regions or geographical areas.

RRACE

The RRACE variable represents respondents' race. The analysis indicated:

Value 1: Recurs 67,528 times, or about 90.04% of the time.

Value 2: 7,467 occurrences, or 9.96%, of the total.

A binary category pertaining to Hispanic ethnicity appears to be represented by the "RHISPANIC" variable.

THHLD\_YEAR

THHLD\_YEAR indicates the number of people in the household. The analysis revealed:

Of course, based on the provided summary, below is an examination of the "TBIRTH\_YEAR" variable:

TBIRTH\_YEAR Analysis in Summary:

The "TBIRTH\_YEAR" variable contains details on the years that people were born. With counts and percentages for each birth year, it appears to be an extensive breakdown of birth year data.

(1) Historical Distribution The information spans a large range of birth years, from 1934 to 2004.

2. Distribution Patterns

- With 1,467 occurrences, or roughly 1.96% of the total, 1961 ranks as the most prevalent birth year in the dataset.

- The frequency of births between 1954 and 1963 is substantially greater.

- The dataset is more sparsely distributed and has lower frequencies for birth years before 1950 and after 2000.

Peaks of Note: Some years, such 1947 and 1953, show higher frequencies, suggesting possible trends or events in the population during those years.

Historical Context: You can examine and comprehend the age distribution in your dataset with this variable. It might be useful for demographic research, generational trends, or cohort analysis.

Data completeness: The data provides a detailed picture of the ages of the individuals in the sample and covers a wide range of birth years.

Potential Outliers: Birth years with low frequencies, such as 1934 and 1935, could be outliers or represent particular demographic traits. Further research into these cases might yield new information.

Data Quality: To prevent problems with data quality, make sure that the birth years are precise and correctly cleansed.

Any research where age or the year of birth is an important consideration, such as demographic studies, marketing segmentation, or market analysis, can benefit from the examination of the "TBIRTH\_YEAR" variable. Please share more information if you have particular queries or need extra analysis pertaining to this variable. I'll be pleased to help you further.

"THHLD\_NUMPER" variable is presented:

Summary analysis for THHLD\_NUMPERThe number of individuals living in each home appears to be represented by the "THHLD\_NUMPER" field. The summary displays the totals and %s for various household sizes.

1. Distribution of household sizes

- The most prevalent household size is 2, with 28,090 occurrences, or roughly 37.46% of the sample.

- The second most frequent household size is one, with 14,944 instances, or around 19.93%.

- There are 12,525 (16.70%) and 10,966 (14.62%) more instances of household sizes 3 and 4, respectively.

- Households with five to ten members have gradually lower frequencies and make up a diminishing portion of the sample.

2. Predominance of Smaller Households According to the data, families with one or two persons are the most common and account for a sizable share of the dataset. This shows that a large portion of the dataset's population resides in smaller homes.

3. Declining Frequency Beyond two people, the frequency declines substantially, with larger households being less typical.

4. Outliers: Households with a size of 7, 8, 9, or 10 people may exist, but they are comparatively uncommon, making up less than 1% of the dataset for each one. There may be a need for additional research or validation for these bigger family sizes.

5. Repercussions: Understanding the makeup of the homes in the dataset can benefit from knowing the distribution of household sizes. Various demographic and marketing analysis.

6. Data Quality: To prevent problems with data quality, make sure the household size information is correct and well cleansed.

For a variety of research projects, this study offers information about the dataset's distribution of household sizes. Please share more information if you have particular queries or need extra analysis pertaining to this variable. I'll be pleased to help you further.

2. Describe the correlation between each of the following variables and receiving at least one dose of the COVID-19 vaccination using counts and percentages: ACTIVITY2, EEDUC, MOBILITY, MS, REGION, RHISPANIC, RRACE, TBIRTH\_YEAR, and THHLD\_NUMBER.

In this report, we give a brief summary of the statistical analysis of Selected Variables by "Vaccine Receipt." This can be expanded upon and altered to meet my own data and analysis.

cat("Counts of ACTIVITY2 by Vaccine Receipt:\n")

Counts of ACTIVITY2 by Vaccine Receipt:

> print(activity2\_counts)

-99 1 2

-99 227 2141 212

-88 248 2986 841

1 50 25127 1486

2 147 36755 4775

>

> # Print percentages

> cat("\nPercentages of ACTIVITY2 by Vaccine Receipt:\n")

Percentages of ACTIVITY2 by Vaccine Receipt:

> print(activity2\_percentages)

-99 1 2

-99 33.779762 3.195093 2.898551

-88 36.904762 4.456118 11.498496

1 7.440476 37.497948 20.317200

2 21.875000 54.850841 65.285753

Counts of EEDUC by Vaccine Receipt:

> print(eeduc\_counts)

-99 1 2

1 7 386 127

2 7 860 279

3 77 6647 1672

4 160 12898 2167

5 69 6429 984

6 189 20552 1345

7 163 19237 740

>

> # Print percentages

> cat("\nPercentages of EEDUC by Vaccine Receipt:\n")

Percentages of EEDUC by Vaccine Receipt:

> print(eeduc\_percentages)

-99 1 2

1 1.041667 0.576042 1.736396

2 1.041667 1.283410 3.814602

3 11.458333 9.919563 22.860268

4 23.809524 19.248161 29.628110

5 10.267857 9.594234 13.453651

6 28.125000 30.670507 18.389390

7 24.255952 28.708084 10.117583

>

>

|  |
| --- |
| > cat("Counts of EGENID\_BIRTH by Vaccine Receipt:\n")  Counts of EGENID\_BIRTH by Vaccine Receipt:  > print(egenid\_birth\_counts)    -99 1 2  1 271 27574 2827  2 401 39435 4487  >  > # Print percentages  > cat("\nPercentages of EGENID\_BIRTH by Vaccine Receipt:\n")  Percentages of EGENID\_BIRTH by Vaccine Receipt:  > print(egenid\_birth\_percentages)    -99 1 2  1 40.32738 41.14970 38.65190  2 59.67262 58.85030 61.34810  >  >  >  >  > # Calculate counts and percentages for MS by vaccine receipt  > ms\_counts <- table(selected\_vars$MS, selected\_vars$RECVDVACC)  > ms\_percentages <- prop.table(ms\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of MS by Vaccine Receipt:\n")  Counts of MS by Vaccine Receipt:  > print(ms\_counts)    -99 1 2  -99 323 340 28  1 203 36800 3635  2 23 3577 308  3 69 10655 1238  4 7 1234 259  5 47 14403 1846  >  > # Print percentages  > cat("\nPercentages of MS by Vaccine Receipt:\n")  Percentages of MS by Vaccine Receipt:  > print(ms\_percentages)    -99 1 2  -99 48.0654762 0.5073945 0.3828275  1 30.2083333 54.9179961 49.6992070  2 3.4226190 5.3380889 4.2111020  3 10.2678571 15.9008491 16.9264424  4 1.0416667 1.8415437 3.5411540  5 6.9940476 21.4941277 25.2392672  >  >  >  >  >  > # Calculate counts and percentages for MOBILITY by vaccine receipt  > mobility\_counts <- table(selected\_vars$MOBILITY, selected\_vars$RECVDVACC)  > mobility\_percentages <- prop.table(mobility\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of MOBILITY by Vaccine Receipt:\n")  Counts of MOBILITY by Vaccine Receipt:  > print(mobility\_counts)    -99 1 2  -99 143 162 39  -88 342 6139 1361  1 129 47681 4540  2 46 10447 1015  3 12 2323 326  4 0 257 33  >  > # Print percentages  > cat("\nPercentages of MOBILITY by Vaccine Receipt:\n")  Percentages of MOBILITY by Vaccine Receipt:  > print(mobility\_percentages)    -99 1 2  -99 21.2797619 0.2417586 0.5332240  -88 50.8928571 9.1614559 18.6081488  1 19.1964286 71.1561134 62.0727372  2 6.8452381 15.5904431 13.8774952  3 1.7857143 3.4666985 4.4572054  4 0.0000000 0.3835306 0.4511895  >  >  >  >  >  >  > # Calculate counts and percentages for REGION by vaccine receipt  > region\_counts <- table(selected\_vars$REGION, selected\_vars$RECVDVACC)  > region\_percentages <- prop.table(region\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of REGION by Vaccine Receipt:\n")  Counts of REGION by Vaccine Receipt:  > print(region\_counts)    -99 1 2  1 103 12102 807  2 208 20747 2575  3 173 13294 1871  4 188 20866 2061  >  > # Print percentages  > cat("\nPercentages of REGION by Vaccine Receipt:\n")  Percentages of REGION by Vaccine Receipt:  > print(region\_percentages)    -99 1 2  1 15.32738 18.06026 11.03363  2 30.95238 30.96151 35.20645  3 25.74405 19.83913 25.58108  4 27.97619 31.13910 28.17884  >  >  >  > # Calculate counts and percentages for RHISPANIC by vaccine receipt  > rhisp\_counts <- table(selected\_vars$RHISPANIC, selected\_vars$RECVDVACC)  > rhisp\_percentages <- prop.table(rhisp\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of RHISPANIC by Vaccine Receipt:\n")  Counts of RHISPANIC by Vaccine Receipt:  > print(rhisp\_counts)    -99 1 2  1 588 60449 6491  2 84 6560 823  >  > # Print percentages  > cat("\nPercentages of RHISPANIC by Vaccine Receipt:\n")  Percentages of RHISPANIC by Vaccine Receipt:  > print(rhisp\_percentages)    -99 1 2  1 87.50000 90.21027 88.74761  2 12.50000 9.78973 11.25239  >  >  >  > # Calculate counts and percentages for RRACE by vaccine receipt  > rrace\_counts <- table(selected\_vars$RRACE, selected\_vars$RECVDVACC)  > rrace\_percentages <- prop.table(rrace\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of RRACE by Vaccine Receipt:\n")  Counts of RRACE by Vaccine Receipt:  > print(rrace\_counts)    -99 1 2  1 559 54597 5908  2 57 5322 766  3 27 4130 103  4 29 2960 537  >  > # Print percentages  > cat("\nPercentages of RRACE by Vaccine Receipt:\n")  Percentages of RRACE by Vaccine Receipt:  > print(rrace\_percentages)    -99 1 2  1 83.184524 81.477115 80.776593  2 8.482143 7.942217 10.473065  3 4.017857 6.163351 1.408258  4 4.315476 4.417317 7.342084  >  >  >  >  >  > # Calculate counts and percentages for TBIRTH\_YEAR by vaccine receipt  > tbirth\_year\_counts <- table(selected\_vars$TBIRTH\_YEAR, selected\_vars$RECVDVACC)  > tbirth\_year\_percentages <- prop.table(tbirth\_year\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of TBIRTH\_YEAR by Vaccine Receipt:\n")  Counts of TBIRTH\_YEAR by Vaccine Receipt:  > print(tbirth\_year\_counts)    -99 1 2  1934 17 299 53  1935 2 103 6  1936 1 134 7  1937 2 156 7  1938 1 187 7  1939 3 241 8  1940 2 286 5  1941 2 321 10  1942 9 425 10  1943 5 521 23  1944 1 537 20  1945 9 578 26  1946 4 818 25  1947 11 1004 28  1948 4 953 40  1949 13 1102 38  1950 15 1175 39  1951 12 1246 43  1952 13 1296 57  1953 10 1385 58  1954 14 1422 77  1955 14 1483 81  1956 12 1480 90  1957 10 1409 88  1958 10 1392 103  1959 17 1426 95  1960 23 1419 110  1961 23 1328 116  1962 18 1311 102  1963 18 1368 130  1964 15 1358 134  1965 10 1243 120  1966 12 1209 122  1967 9 1149 125  1968 17 1108 148  1969 18 1273 158  1970 13 1330 146  1971 9 1247 122  1972 8 1154 147  1973 8 1159 165  1974 13 1122 164  1975 10 1131 159  1976 15 1181 169  1977 7 1240 164  1978 14 1223 182  1979 9 1256 177  1980 20 1334 212  1981 15 1287 197  1982 10 1262 203  1983 10 1294 204  1984 18 1253 198  1985 12 1275 195  1986 12 1198 168  1987 10 1149 177  1988 11 1170 175  1989 5 1049 198  1990 7 1078 212  1991 5 1089 168  1992 4 961 150  1993 8 935 134  1994 12 835 144  1995 5 752 123  1996 4 711 117  1997 3 624 88  1998 4 420 94  1999 7 372 71  2000 3 282 67  2001 1 193 39  2002 2 143 26  2003 0 102 32  2004 2 53 18  >  > # Print percentages  > cat("\nPercentages of TBIRTH\_YEAR by Vaccine Receipt:\n")  Percentages of TBIRTH\_YEAR by Vaccine Receipt:  > print(tbirth\_year\_percentages)    -99 1 2  1934 2.52976190 0.44620872 0.72463768  1935 0.29761905 0.15371070 0.08203445  1936 0.14880952 0.19997314 0.09570686  1937 0.29761905 0.23280455 0.09570686  1938 0.14880952 0.27906699 0.09570686  1939 0.44642857 0.35965318 0.10937927  1940 0.29761905 0.42680834 0.06836205  1941 0.29761905 0.47904013 0.13672409  1942 1.33928571 0.63424316 0.13672409  1943 0.74404762 0.77750750 0.31446541  1944 0.14880952 0.80138489 0.27344818  1945 1.33928571 0.86257070 0.35548264  1946 0.59523810 1.22073154 0.34181023  1947 1.63690476 1.49830620 0.38282745  1948 0.59523810 1.42219702 0.54689636  1949 1.93452381 1.64455521 0.51955154  1950 2.23214286 1.75349580 0.53322395  1951 1.78571429 1.85945172 0.58791359  1952 1.93452381 1.93406856 0.77932732  1953 1.48809524 2.06688654 0.79299973  1954 2.08333333 2.12210300 1.05277550  1955 2.08333333 2.21313555 1.10746514  1956 1.78571429 2.20865854 1.23051682  1957 1.48809524 2.10270262 1.20317200  1958 1.48809524 2.07733290 1.40825814  1959 2.52976190 2.12807235 1.29887886  1960 3.42261905 2.11762599 1.50396500  1961 3.42261905 1.98182334 1.58599945  1962 2.67857143 1.95645361 1.39458573  1963 2.67857143 2.04151681 1.77741318  1964 2.23214286 2.02659344 1.83210282  1965 1.48809524 1.85497470 1.64068909  1966 1.78571429 1.80423525 1.66803391  1967 1.33928571 1.71469504 1.70905113  1968 2.52976190 1.65350923 2.02351654  1969 2.67857143 1.89974481 2.16024063  1970 1.93452381 1.98480801 1.99617173  1971 1.33928571 1.86094405 1.66803391  1972 1.19047619 1.72215673 2.00984413  1973 1.19047619 1.72961841 2.25594750  1974 1.93452381 1.67440195 2.24227509  1975 1.48809524 1.68783298 2.17391304  1976 2.23214286 1.76244982 2.31063713  1977 1.04166667 1.85049769 2.24227509  1978 2.08333333 1.82512797 2.48837845  1979 1.33928571 1.87437508 2.42001641  1980 2.97619048 1.99077736 2.89855072  1981 2.23214286 1.92063753 2.69346459  1982 1.48809524 1.88332911 2.77549904  1983 1.48809524 1.93108388 2.78917145  1984 2.67857143 1.86989807 2.70713700  1985 1.78571429 1.90272948 2.66611977  1986 1.78571429 1.78781955 2.29696473  1987 1.48809524 1.71469504 2.42001641  1988 1.63690476 1.74603411 2.39267159  1989 0.74404762 1.56546136 2.70713700  1990 1.04166667 1.60873912 2.89855072  1991 0.74404762 1.62515483 2.29696473  1992 0.59523810 1.43413571 2.05086136  1993 1.19047619 1.39533496 1.83210282  1994 1.78571429 1.24610127 1.96882691  1995 0.74404762 1.12223731 1.68170632  1996 0.59523810 1.06105150 1.59967186  1997 0.44642857 0.93121819 1.20317200  1998 0.59523810 0.62678148 1.28520645  1999 1.04166667 0.55514931 0.97074104  2000 0.44642857 0.42083899 0.91605141  2001 0.14880952 0.28802101 0.53322395  2002 0.29761905 0.21340417 0.35548264  2003 0.00000000 0.15221836 0.43751709  2004 0.29761905 0.07909385 0.24610336  >  >  >  >  > # Calculate counts and percentages for THHLD\_NUMPER by vaccine receipt  > thhld\_numper\_counts <- table(selected\_vars$THHLD\_NUMPER, selected\_vars$RECVDVACC)  > thhld\_numper\_percentages <- prop.table(thhld\_numper\_counts, margin = 2) \* 100  >  > # Print counts  > cat("Counts of THHLD\_NUMPER by Vaccine Receipt:\n")  Counts of THHLD\_NUMPER by Vaccine Receipt:  > print(thhld\_numper\_counts)    -99 1 2  1 167 13696 1081  2 298 25817 1975  3 97 11068 1360  4 61 9563 1342  5 24 4107 788  6 9 1662 368  7 2 557 168  8 2 256 85  9 0 98 26  10 12 185 121  >  > # Print percentages  > cat("\nPercentages of THHLD\_NUMPER by Vaccine Receipt:\n")  Percentages of THHLD\_NUMPER by Vaccine Receipt:  > print(thhld\_numper\_percentages)    -99 1 2  1 24.8511905 20.4390455 14.7798742  2 44.3452381 38.5276605 27.0030079  3 14.4345238 16.5171843 18.5944763  4 9.0773810 14.2712173 18.3483730  5 3.5714286 6.1290274 10.7738584  6 1.3392857 2.4802638 5.0314465  7 0.2976190 0.8312316 2.2969647  8 0.2976190 0.3820382 1.1621548  9 0.0000000 0.1462490 0.3554826  10 1.7857143 0.2760823 1.6543615 |
|  |
| |  | | --- | | > | |

Interpreting counts and percentages in the context of data analysis is essential to draw meaningful conclusions from your results. Let's use your example of counts and percentages of "ACTIVITY2" by "Vaccine Receipt" to provide interpretations:

Counts of SELECTED VARIABLE by Vaccine Receipt

In your analysis, you calculated the counts of different activities ("SELECTED VARIABLE") based on whether individuals received a vaccine ("Vaccine Receipt" or "RECVDVACC"). Here's how you can interpret these counts:

Count Values: The count values represent the number of individuals in your dataset who reported engaging in each specific activity ("SELECTED VARIABLE"). For example, if the count for "Activity 1" is 150 for those who received the vaccine, it means 150 individuals reported engaging in "Activity 1" and received the vaccine.

Comparison: By comparing the counts for different activities across groups (vaccinated vs. unvaccinated), you can understand the distribution of these activities among the two groups. For instance, if the count for "SELECTED VARIABLE" is higher among the vaccinated group (220) compared to the unvaccinated group (50), it suggests that "SELECTED VARIABLE" is more prevalent among those who received the vaccine.

Percentages of SELECTED VARIABLE by Vaccine Receipt

Next, you calculated the percentages of individuals engaging in different activities by vaccine receipt. Here's how you can interpret these percentages:

Percentage Values: The percentage values represent the proportion of individuals in each group (vaccinated and unvaccinated) who reported engaging in a specific activity. For example, if "Activity 1" is 60% for the vaccinated group, it means that 60% of the vaccinated individuals engage in "Activity 1."

Comparative Insight: By examining the percentages, you gain insights into the relative prevalence of activities within each group. For instance, a higher percentage of vaccinated individuals engaging in "selected variable" (81%) compared to unvaccinated individuals (19%) suggests that "selected variable" is more common among the vaccinated group.

Interpreting counts and percentages is crucial for drawing conclusions about the relationship between variables and identifying patterns or differences among subgroups. It allows you to make inferences about the impact of vaccination on individuals' engagement in various activities in this specific context.

**Conclusion:**

Based on each person's status as having had a vaccination, the study indicates disparities in the distribution of activities among them. As an illustration, more vaccinated people than unvaccinated people reported taking part in Activity 2.

Researchers and public health decision-makers who want to understand how immunization affects daily activities may find this material useful.

**Appendix:**

The selected variables from the "Pulse Survey 2022" data source are summarized in the following tables, together with the relevant counts and percentages. Activities, education, demography, and mobility are just a few of the topics covered in the analysis.

Table : 1 ACTIVITY\_2

| **Counts.var1** | **Counts.frequency** | **Percentages.frequency** | **Percentages.var1** |
| --- | --- | --- | --- |
| -99 | 2,580 | 3.44% | -99% |
| -88 | 4,075 | 5.43% | -88% |
| 1 | 26,663 | 35.55% | 1% |
| 2 | 41,677 | 55.57% | 2% |

Table.2 EEDUC

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1 | 520 | 1 | 0.69% |
| 2 | 1,146 | 2 | 1.53% |
| 3 | 8,396 | 3 | 11.20% |
| 4 | 15,225 | 4 | 20.30% |
| 5 | 7,482 | 5 | 9.98% |
| 6 | 22,086 | 6 | 29.45% |
| 7 | 20,140 | 7 | 26.86% |

Table.3 EGENID\_BIRTH

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1 | 30,672 | 1 | 40.90% |
| 2 | 44,323 | 2 | 59.10% |

Table.4 MOBILITY

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| -99 | 344 | -99 | 0.46% |
| -88 | 7,842 | -88 | 10.46% |
| 1 | 52,350 | 1 | 69.80% |
| 2 | 11,508 | 2 | 15.35% |
| 3 | 2,661 | 3 | 3.55% |
| 4 | 290 | 4 | 0.39% |

Table.5 MS

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| -99 | 691 | -99 | 0.92% |
| 1 | 40,638 | 1 | 54.19% |
| 2 | 3,908 | 2 | 5.21% |
| 3 | 11,962 | 3 | 15.95% |
| 4 | 1,500 | 4 | 2.00% |
| 5 | 16,296 | 5 | 21.73% |

Table.6 REGION

| **Counts.Var1** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- |
| 1 | 1 | 17.35% |
| 2 | 2 | 31.38% |
| 3 | 3 | 20.45% |
| 4 | 4 | 30.82% |

Table.7 RHISPANIC

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1 | 67,528 | 1 | 90.04% |
| 2 | 7,467 | 2 | 9.96% |

Table .8 RRACE

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1 | 61,064 | 1 | 81.42% |
| 2 | 6,145 | 2 | 8.19% |
| 3 | 4,260 | 3 | 5.68% |
| 4 | 3,526 | 4 | 4.70% |

Table 9 TBIRTH\_YEAR

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1934 | 369 | 1934 | 0.49% |
| 1935 | 111 | 1935 | 0.15% |
| 1936 | 142 | 1936 | 0.19% |
| 1937 | 165 | 1937 | 0.22% |
| 1938 | 195 | 1938 | 0.26% |
| 1939 | 252 | 1939 | 0.34% |
| 1940 | 293 | 1940 | 0.39% |
| 1941 | 333 | 1941 | 0.44% |
| 1942 | 444 | 1942 | 0.59% |
| 1943 | 549 | 1943 | 0.73% |
| 1944 | 558 | 1944 | 0.74% |
| 1945 | 613 | 1945 | 0.82% |
| 1946 | 847 | 1946 | 1.13% |
| 1947 | 1,043 | 1947 | 1.39% |
| 1948 | 997 | 1948 | 1.33% |
| 1949 | 1,153 | 1949 | 1.54% |
| 1950 | 1,229 | 1950 | 1.64% |
| 1951 | 1,301 | 1951 | 1.73% |
| 1952 | 1,366 | 1952 | 1.82% |
| 1953 | 1,453 | 1953 | 1.94% |
| 1954 | 1,513 | 1954 | 2.02% |
| 1955 | 1,578 | 1955 | 2.10% |
| 1956 | 1,582 | 1956 | 2.11% |
| 1957 | 1,507 | 1957 | 2.01% |
| 1958 | 1,505 | 1958 | 2.01% |
| 1959 | 1,538 | 1959 | 2.05% |
| 1960 | 1,552 | 1960 | 2.07% |
| 1961 | 1,467 | 1961 | 1.96% |
| 1962 | 1,431 | 1962 | 1.91% |
| 1963 | 1,516 | 1963 | 2.02% |
| 1964 | 1,507 | 1964 | 2.01% |
| 1965 | 1,373 | 1965 | 1.83% |
| 1966 | 1,343 | 1966 | 1.79% |
| 1967 | 1,283 | 1967 | 1.71% |
| 1968 | 1,273 | 1968 | 1.70% |
| 1969 | 1,449 | 1969 | 1.93% |
| 1970 | 1,489 | 1970 | 1.99% |
| 1971 | 1,378 | 1971 | 1.84% |
| 1972 | 1,309 | 1972 | 1.75% |
| 1973 | 1,332 | 1973 | 1.78% |
| 1974 | 1,299 | 1974 | 1.73% |
| 1975 | 1,300 | 1975 | 1.73% |
| 1976 | 1,365 | 1976 | 1.82% |
| 1977 | 1,411 | 1977 | 1.88% |
| 1978 | 1,419 | 1978 | 1.89% |
| 1979 | 1,442 | 1979 | 1.92% |
| 1980 | 1,566 | 1980 | 2.09% |
| 1981 | 1,499 | 1981 | 1.99% |
| 1982 | 1,475 | 1982 | 1.97% |
| 1983 | 1,508 | 1983 | 2.01% |
| 1984 | 1,469 | 1984 | 1.96% |
| 1985 | 1,482 | 1985 | 1.98% |
| 1986 | 1,378 | 1986 | 1.84% |
| 1987 | 1,336 | 1987 | 1.78% |
| 1988 | 1,356 | 1988 | 1.81% |
| 1989 | 1,252 | 1989 | 1.67% |
| 1990 | 1,297 | 1990 | 1.73% |
| 1991 | 1,262 | 1991 | 1.68% |
| 1992 | 1,115 | 1992 | 1.49% |
| 1993 | 1,077 | 1993 | 1.44% |
| 1994 | 991 | 1994 | 1.32% |
| 1995 | 880 | 1995 | 1.17% |
| 1996 | 832 | 1996 | 1.11% |
| 1997 | 715 | 1997 | 0.95% |
| 1998 | 518 | 1998 | 0.69% |
| 1999 | 450 | 1999 | 0.60% |
| 2000 | 352 | 2000 | 0.47% |
| 2001 | 233 | 2001 | 0.31% |
| 2002 | 171 | 2002 | 0.23% |
| 2003 | 134 | 2003 | 0.18% |
| 2004 | 73 | 2004 | 0.10% |

Table 10 THHLD\_NUMPER

| **Counts.Var1** | **Counts.Freq** | **Percentages.Freq.Var1** | **Percentages.Freq.Freq** |
| --- | --- | --- | --- |
| 1 | 14,944 | 1 | 19.93% |
| 2 | 28,090 | 2 | 37.46% |
| 3 | 12,525 | 3 | 16.70% |
| 4 | 10,966 | 4 | 14.62% |
| 5 | 4,919 | 5 | 6.56% |
| 6 | 2,039 | 6 | 2.72% |
| 7 | 727 | 7 | 0.97% |
| 8 | 343 | 8 | 0.46% |
| 9 | 124 | 9 | 0.17% |
| 10 | 318 | 10 | 0.42% |