

**Original Article**

# **FAME BY THE STARS: ZODIAC SIGNS MOST LINKED TO CELEBRITY STATUS**

*Ngozi Hannah Okonkwo*

Center for Interdisciplinary Sciences,  
Department of Chemistry and Physics,  
University of Arkansas, United States.  
DOI:<https://doi.org/10.5281/zenodo.15387159>

## **ABSTRACT**

To find any favored zodiac sign for celebrity births, a sample of 100 celebrities were randomly selected from people of different walks of life. The sample contained politicians, natural scientists, social scientists, authors of literary works, social workers, humanitarian workers, business personnel, sports icons, singers, actors, actresses, etc. etc. from history and from the current time. The zodiac signs for the celebrities were found from their known dates of births. In the analysis of data, zodiac signs and the number of celebrities were represented as the independent x and the dependent y variables, respectively. For academic interests for the 9th grade high school juniors (at the time of the project performance), the co-authors of this article, as well as for the potentially illustrative uses in high school mathematics textbooks, bar and scatter plots were made, the line of best-fit and the equation of the line were found, probabilities of occurrences of celebrities for each of the zodiac signs were calculated and the correlation coefficients between the variables were determined for the sample. It was found that the zodiac Aquarius has the largest number of celebrities in the sample and that the two variables are moderately correlated. The sample sizes which were increased to 200 and then 300. By including another 100 more celebrities to find if the trend remains unchanged. In all the three cases, Aquarius turned out to be the zodiac when most of the celebrities are born.

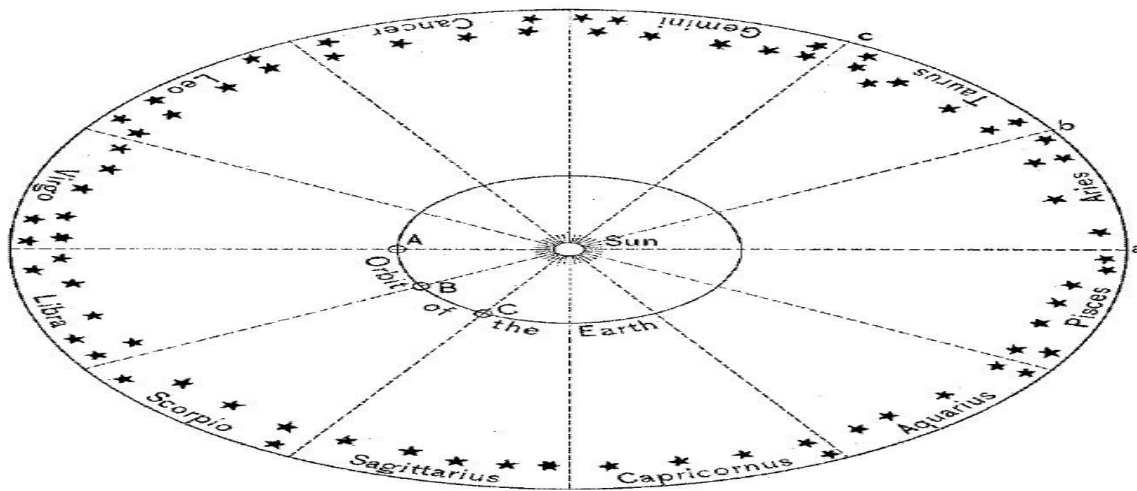
**Keywords:** Bar Plots, Celebrities, Zodiacs, Scatter Plots, Line of Best-Fit, Probability, Correlation Coefficient

## Original Article

### 1. INTRODUCTION

#### 1.1. Apparent Motion of the Zodiacs

Figure 1 shows the sun at the center of the two concentric circles. The inner circle is the orbit of the earth around the sun. The outer circle is drawn through groups of stars called zodiac signs in the sky. In the sky, these groups of stars look about  $8^\circ$  wide in the up-down and  $30^\circ$  long in the left-right directions. During the day time, stars are not visible because of the dazzling light from the sun. The inner circle is marked with three positions A, B and C of the planet earth. Also, the outer circle is marked with the corresponding points a, b and c. When the earth is at position A, the sun would be seen at the position among the stars there if they were visible. When the earth is at position B, the sun would be seen among the group of stars at b if the stars were visible. Similar things would happen for the positions C and c. It is observed that the background stars of the sun are changing which gives the impression that the sun moves around the planet earth. For the planet earth, because of its uniform motion (about 18.5 miles per second) around the sun, it seems stationary. Had there been any change of direction or an increase or a decrease in the motion, its stationary condition would be lost.



**Fig. 1. Illustrates**

The apparent motion of the sun around us (<http://www.astrology.com.au/astrology/12-signs-of-the-zodiac/www.wikipedia.org>)

#### 1.2. Zodiacs and Other Groups of Stars

Zodiac means creatures which refer to the patterns of groups of stars in the clear night sky. There are 12 zodiac signs. The zodiac belt is the great circle (a circle over a sphere having the same center as the sphere) in which the sun apparently moves (<http://www.astrology.com.au/astrology/12-signs-of-thezodiac/www.wikipedia.org>) as illustrated above. Only the groups of stars behind the sun are called the zodiacs or the zodiac signs. There are 88 groups of stars including the zodiacs in the entire sky.

#### 1.3. Solar Crossings of the Zodiacs

The sun enters a zodiac sign about the 21st day of each month. The zodiac constellations are so named because the sun passes through them at a regular time annually (one sign per month). The names of the signs and the months when the sun enters each zodiac along with its brief traits are given in

## Original Article

**Table 1** below

### 1.4. Zodiac Influence Testing for the Celebrities

It is almost a part of universal culture to check one's fate with his/her zodiac sign. People feel quite comfortable with the prediction of his/her zodiac. While there is no scientific proof for the zodiac influences, no proof can be established because of the privacy of personal information, particularly, the date of birth of an individual. The dates of births and so the zodiacs of celebrities are known to the public. Their data could be studied to see if all or most of them are born in a particular zodiac and this is the focus of this article.

### 1.5. Meaning, Duration and Traits of Individual Zodiacs

**Table 1** records the meaning of each of the zodiac signs in column 1, the duration of each of the zodiac signs in column 2 and the brief traits of the individual zodiacs in column 3.

**2. Table 1: Zodiac signs' meanings, durations and traits**

3. Latin name/English name	Time	Brief nature of celebrities
4. Aries/the ram	Mar 21-Apr 20	adventurous, energetic and courageous
5. Taurus/the bull	Apr 21-May 21	patient, reliable and warmhearted
6. Gemini/the twins	May 22-Jun 21	intellectual, youthful and lively
7. Cancer/the crab	Jun 22-Jul 22	emotional, loving and cautious
8. Leo/the lion	Jul 23-Aug 21	generous, creative and faithful
9. Virgo/the virgin	Aug 22-Sep 23	reliable, practical and intelligent
10. Libra/the scales	Sep 24-Oct 23	easygoing, charming and sociable
11. Scorpio/the scorpion	Oct 24-Nov 22	determined, intuitive and passionate
12. Sagittarius/the archer	Nov 23-Dec 22	optimistic, honest and intellectual
13. Capricorn/the goat	Dec 23-Jan 20	practical, ambitious and patient
14. Aquarius/the water carrier	Jan 21-Feb 19	friendly, loyal and original
Pisces/the fishes	Feb 20-Mar 20	compassionate, intuitive and selfless

## MATERIALS AND METHODS

### 2.1. Definition of Celebrities and Their Random Selection

Whoever has got a nationally and/or internationally positive publicity and has earned name and fame in his/her profession in history or in the current days' electronic and/or print media and whose personal information is available to the public was taken to be a celebrity. These celebrities appear in the worlds of politics, natural and social sciences, social and humanitarian services, literature, sports and athletics, recreational productions, arts and sculpture, etc. etc. The randomness and impartiality in the selection of the celebrity samples were maintained by picking the names of the celebrities first and then finding the dates of births and the zodiacs and not the other way around. The picking of the names of the top ones avoided the entire or partial selection in a regular form from a group of celebrities falling in a particular profession.

### 2.2. Celebrity Data Spread

About 40% of the celebrities were from the dead ones and 60% from the living ones. About 70% of them were males and 30% females. Less than 5% were Mexicans and Asians. No grouping was made based on ages because no age records were made as to when an individual became a celebrity.

### 2.3. Data Analysis

First, a random list of 100 celebrities from different walks of life was made. Their dates of births were found from wikipedia.com and found their zodiacs. The number of celebrities falling in each zodiac was counted. For the

## Original Article

academic interests of the 9<sup>th</sup>-grade coauthors when they performed the project, bar plots of the zodiac signs Vs the number of celebrities were made and probabilities for the celebrities falling in a zodiac sign were calculated. Scatter plots were made with the zodiacs as the xvariable and the number of celebrities as the yvariables. The line of best-fit using Microsoft Excel 2010 was found. The equation of the lines of best-fit were determined. The correlation coefficients (<http://www.statisticshowto.com/articles/how-tocompute-pearsons-correlation-coefficients/>) were determined to find how strongly the zodiac signs and the number of celebrities in them are related. The above procedure was repeated for sample sizes of 200 and then for 300 to find any continued feature common in all the sample sizes.

**Table 2** enlists the probabilities for the three sets of samples.

**Table 3 to 5** record the calculations for the correlation coefficients for sample sizes of 100, 200 and 300, respectively.

**Table 6** compares the equations of the line-of-best fit and the correlation coefficients.

### CACULATIONS

#### 3.1. Equation of the Line of Best-Fit

The equation of the line of best-fit was found from its slope and y-intercept. The slope was determined from the coordinates of two points on the line and the intercept, using the coordinates of a point in the equation  $y = mx + b$  and solving for b.

**Table 2.** Probabilities for different sample sizes

Zodiacs	100	200	300
Aries	0.080	0.065	0.067
Taurus	0.070	0.070	0.057
Gemini	0.050	0.075	0.130
Cancer	0.060	0.090	0.070
Leo	0.100	0.095	0.067
Virgo	0.080	0.075	0.057
Libra	0.070	0.075	0.103
Scorpio	0.100	0.085	0.100
Sagittarius	0.070	0.065	0.053
Capricorn	0.100	0.095	0.083
Aquarius	0.140	0.130	0.150
Pisces	0.080	0.080	0.107
	1.000	1.000	1.000

**Table 3.** Correlation coefficient for sample s

ize 100					
Zodiac signs	x	y	xy	$x^2$	$y^2$
Aries	1	8	8	1	64
Taurus		7	14	4	49
Gemini	3	5	15	9	25
Cancer	4	6	24	16	36
Leo	5	10	50	25	100

**Original Article**

Virgo	6	8	42	36	64
Libra	7	7	49	49	49
Scorpio	8	10	80	64	100
Sagittarius	9	7	63	81	49
Capricorn	10	10	100	100	100
Aquarius	11	14	154	121	196
Pisces	12	8	96	144	64
Sum of the cols.	78	100	695	560	896

Correlation Coefficient,  $r = \frac{(n)(\text{Sum of } xy) - (\text{Sum } x)(\text{Sum } y)}{\sqrt{\{(n)(\text{Sum } x^2) - (\text{Sum } x)^2\} \{(n)(\text{Sum } y^2) - (\text{Sum } y)^2\}}} = \frac{540}{\sqrt{(752)(1716)}} = 0.475$

**Table 4.** Correlation coefficient for sample size 200

Zodiac signs	x	y	xy	x <sup>2</sup>	y <sup>2</sup>
Aries	1	13	13	1	169
Taurus	2	14	28	4	196
Gemini	3	15	45	9	225
Cancer	4	18	72	16	324
Leo	5	19	95	25	361
Virgo	6	15	90	36	225
Libra	7	15	105	49	225
Scorpio	8	17	136	64	289
Sagittarius	9	13	117	81	169
Capricorn	10	19	190	100	361
Aquarius	11	26	286	121	626
Pisces	12	16	192	144	256
Sum of the cols.	78	200	1369	650	3478

Correlation Coefficient,  $r = \frac{(n)(\text{Sum of } xy) - (\text{Sum } x)(\text{Sum } y)}{\sqrt{\{(n)(\text{Sum } x^2) - (\text{Sum } x)^2\} \{(n)(\text{Sum } y^2) - (\text{Sum } y)^2\}}} = \frac{828}{\sqrt{(1716)(1736)}} = 0.4797$

**Table 5.** Correlation coefficient for sample size 300

Zodiac signs	x	y	xy	x <sup>2</sup>	y <sup>2</sup>
Aries	1	20	20	1	400
Taurus	2	17	34	4	289
Gemini	3	26	78	9	676
Cancer	4	21	84	16	441
Leo	5	20	100	25	400
Virgo	6	17	102	36	289
Libra	7	31	217	49	961
Scorpio	8	30	240	64	900
Sagittarius	9	16	144	81	256
Capricorn	10	25	250	100	625

## Original Article

Aquarius	11	45	495	121	2025
Pisces	12	32	384	144	1024
Sum of the cols.	78	300	2148	650	8286
Correlation Coefficient, $r = \frac{\sum xy - x(\sum y)/\sqrt{\{(n)(\sum x^2) - x^2\}}}{\sqrt{\{(n)(\sum y^2) - y^2\}}} = \frac{2376/\sqrt{\{(1716)(9432)\}}}{\sqrt{\{(1716)(9432)\}}} = 0.59$					

**Table 6.** Comparison of results

Sample size	Best-line fit	Correlation coefficient	Comments
100	$y = 0.34x + 6.12$	0.48	moderately related
200	$y = 0.48x + 13.56$	0.48	moderately related
300	$y = 1.33x + 16.01$	0.59	moderately related

### 3.1.1. Sample Size of 100 Celebrities

Sample size 100:  $x_1 = 2, y_1 = 6.8, x_2 = 9, y_2 = 9.2$ . Slope of the line,  $m = (y_2 - y_1)/(x_2 - x_1) = 0.34$ . Substitution of this value of the slope in equation  $y = mx + b$  and using  $x = 2$  and  $y = 6.8$ , yields  $b = 6.12$ . So, the equation of the line is:  $y = 0.34x + 6.12$ .

### 3.1.2. Sample Size of 100 Celebrities

Sample size 200:  $x_1 = 3, y_1 = 15, x_2 = 7.2, y_2 = 17$ . Slope of the line,  $m = (y_2 - y_1)/(x_2 - x_1) = 0.48$ . Substitution of this value of the slope in equation  $y = mx + b$  and using  $x = 3$  and  $y = 15$ , yields  $b = 13.56$ . So, the equation of the line is:  $y = 0.48x + 13.56$ .

### 3.1.3. Sample Size of 100 Celebrities

Sample size 300:  $x_1 = 3, y_1 = 20, x_2 = 6, y_2 = 24$ . Slope of the line,  $m = (y_2 - y_1)/(x_2 - x_1) = 1.33$ . Substitution of this value of the slope in equation  $y = mx + b$  and using  $x = 3$  and  $y = 20$ , yields  $b = 16.01$ . So, the equation of the line is:  $y = 1.33x + 16.01$ .

## 3.2. Probabilities

The probabilities of the occurrences of the zodiacs in a sample size were found by dividing the number of celebrities falling in a zodiac by the total number of celebrities in that sample. The probabilities of all samples have been tabulated in **Table 2**.

## 3.3. Correlation Coefficients

The correlation coefficient for each of the sample sizes was determined using the standard statistical procedure. The  $x$ -,  $y$ - and their derived values  $xy$ -,  $x^2$  and  $y^2$ - are tabulated in **Table 3 to 5**. Surprisingly, the two variables- $x$  and  $y$ - show medium correlation.

### 3.3.1. Correlation Coefficient for 100 Sample Size

The quantities for the calculation of correlation coefficient have been tabulated in **Table 3**. The correlation coefficient for this sample size was 0.475 which is an indication of moderate correlation.

### 3.3.2. Correlation Coefficient for 200 Sample Size

The quantities required for the calculation of correlation for this sample size is tabulated in **Table**

4. The correlation coefficient for this sample size was 0.4797 which, too, is an indication of moderate correlation.

## Original Article

### 3.3.3. Correlation Coefficient for 300 Sample Size

The quantities required for the calculation of correlation for this sample size is tabulated in **Table**

5. The correlation coefficient for this sample size was 0.0.59 which is still an indication of moderate correlation.

## 15. RESULTS

### 4.1. Bar Plots

**Figure 2 to 4** shows the bar plots for 100, 200 and 300 sample sizes. It is observed that the number of celebrities in Aquarius outnumbers the same in all other zodiacs which have got fluctuating number of celebrities with the sample sizes.

The scatter plots for 100, 200 and 300 celebrities appear in **Fig. 5 to 7**, respectively. The line of best fit in them has the form  $y = mx + b$  with different values of the slope  $m$  and the y-intercept  $b$ .

The correlation coefficients for the 100, 200 and 300 sample sizes are found to be 0.48, 0.48 and 0.59, respectively.

#### 4.1.1. Bar Plot for Sample Size 100

The bar plot for the sample size of 100 is shown in

**Fig. 2.** The zodiac Aquarius has, by far, the largest number of celebrities. The next ones are the Leo, the Scorpio and the Capricorn.

#### 4.1.2. Bar Plot for Sample Size 200

The bar plot for the sample size of 200 is shown in **Fig. 3.** The zodiac Aquarius has still, by far, the largest number of celebrities. The next ones are the Leo and the Capricorn followed by the Cancer.

#### 4.1.3. Bar Plot for Sample Size 300

The bar plot for the sample size of 300 is shown in **Fig. 4.** Again, the zodiac Aquarius has, by far, the largest number of celebrities. The next one is the Pisces followed by Libra and then by the Scorpio.

### 4.2. Line of Best-Fit

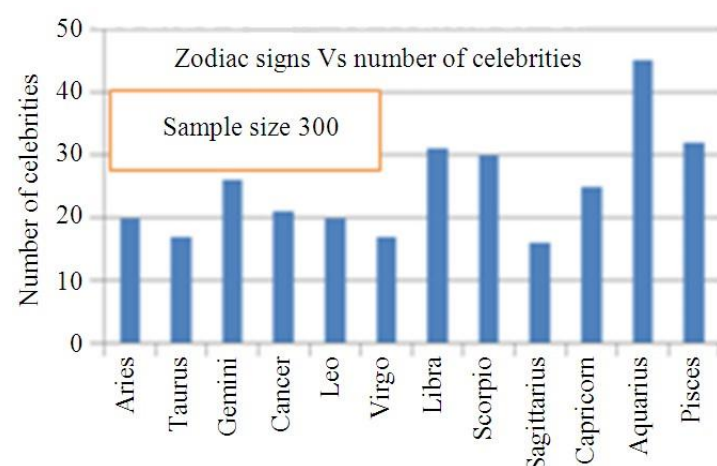
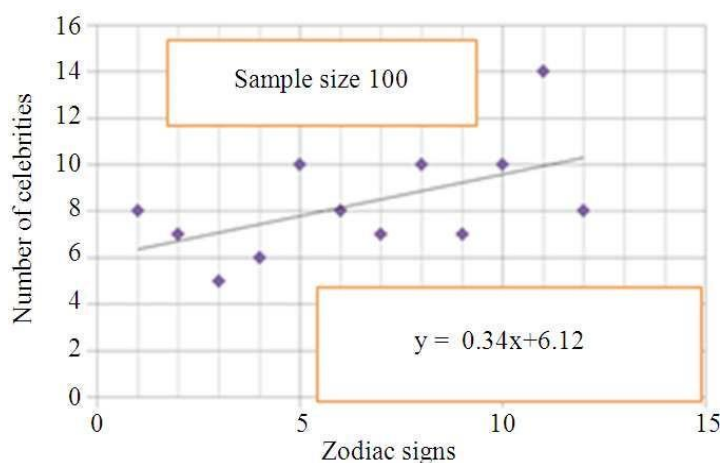
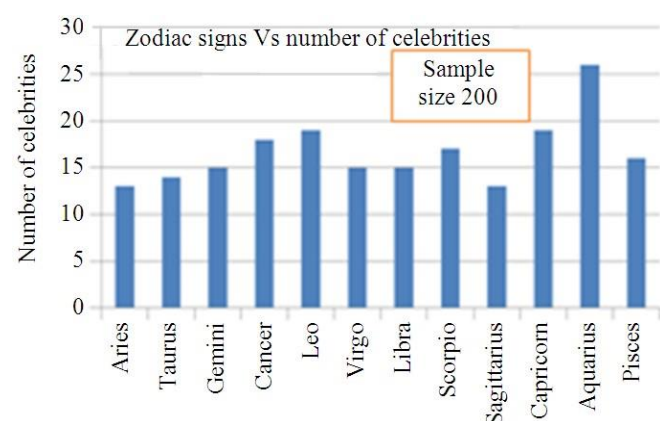
The standard statistical method was applied to find the line of best fit. The lines of best-fit through the scatter plots are shown in **Fig. 5 to 7.**

**Fig. 2.** Bar plot for 100 celebrities

**Fig. 3.** Bar plot for 200 celebrities



## Original Article



### 4.2.1. Line of Best-Fit for Sample Size 100

This is shown in Fig. 5. The equation of the line of best-fit is  $y = 0.34x + 6.12$

### 4.2.2. Line of Best-Fit for Sample Size 200

This is shown in Fig. 6. The equation of the line of best-fit is  $y = 0.48x + 13.56$

### 4.2.3. Line of Best-Fit for Sample Size 300

This is shown in Fig. 7. The equation of the line of best-fit is  $y = 1.33x + 16.01$

## COMPARISON OF RESULTS

The results for the lines of best-fit and the correlation coefficients are shown in Table 6.

## DISCUSSION

### 6.1. Data Spread and Unbiasedness in Data Selection

Samples of individuals of the current time or historical past who excelled in their respective professions nationally or internationally and whose personal information was accessible and whom media give the celebrity status were taken as the subjects of the study. These individuals were among the academicians, discoverers, inventors, statesman, sportsmen, sportswomen, singers, actors, actresses, authors, social workers, etc. etc. recorded in history as well as in the present time. There are individuals in the samples from the time when movie stars and producers were not even borne and sports were not developed. The random selection was not like the generation of random numbers. The names were picked without any choice of the place, time, national origin, sex, of the individuals. The sample selection would be biased if the individuals were selected based on their dates of birth. The dates of birth were not known until the name was picked up and there was no rejection or preference of any individuals



## **Original Article**

for being born under some zodiacs. Setting some conditions or criteria for selection of the sample elements, brings some restrictions to the study. Leaving it open makes the result a more general one out of diversity of age, sex, ethnicity etc.

### **6.2. Limitation of the Study**

It was not possible to study age-wise because individuals might have multiple achievements. It would put some uncertainty as to the age of his/her achieving celebrity status. A sex-wise study would not be informative because of the number of women celebrity would be a few. Even in the sixties, Harvard University was not admitting female students until the early sixties (Dr. Vera Rubin's lecture at the First Vatican Summer School in Observational Astronomy and Astrophysics, 1986, when the principal author was a student). If there is such a blockage as this, the possibility of rising to celebrity status would be stopped. Also, an ethnicity-based study will not be worthwhile because of the majority of the African-Americans are yet to be the first generation college-going ones and a few celebrities are in this ethnic group.

### **6.3. Sample Size Variation Objectives**

The variation of the sample size was made to find if it was a general result independent of the sample of individuals or dependent on any particular sample of individuals. Even though the first selected sample of 100 individuals showed the trend reported in the article, if there was any otherwise effect, it would be reflected with the addition of 100 more and later 200 more individuals in the samples. Further, increasing the sample size in steps should avoid the requirement of a very large number individuals in the sample because of the reflection of the same trends in all the samples.

### **6.4. Academic Values of the Article**

This article has research as well as academic values. The research value lies in its new result. Its academic values relate to the coauthors who are high school juniors and who learn the simple statistics. Mathematics textbooks give examples of scatter plots, least-squares-fit, correlation, etc. Some math textbook author may illustrate these statistical principles with the help of these data points. Three different sample sizes were selected to see if the trend remains the same. This was a part of the procedure in scientific investigations and was followed in the interest of teaching steps in scientific investigation to junior investigators.

### **6.5. Arbitrarily Zodiacs Numbering**

In making the plots, Aries was arbitrarily assigned the number "1". Even if Aries is not marked by "1" by another data analyzer, neither the result nor the academic value changes as long as the consistency is maintained in all cases of assigning the numbers to the zodiacs.

### **6.6. Prohibition to Study for the Public**

The category of subjects could not be extended beyond the celebrities to common people to make a general study because the law makes people's vital records inaccessible. The gives us a good picture of the favored zodiac sign, the Aquarius, for celebrities. The principal author approached the state office of vital records and the CDC, no common person's birth records were made available for this research purpose.

### **6.7. Indeterminate Roles of Zodiacs Traits**

Each of the zodiac signs has several traits. It was not possible to see if all the individuals in a zodiac sign showed all or some of the traits of the sign. Also, it was not possible to see if any particular traits made them celebrities.

### **6.8. Further Academic Value**

The theoretical prediction or the best-fit line cannot be used for any  $x$  beyond  $1 \leq x \leq 12$ . It was in the interest of learning practical lessons of algebra for the juniors. Someday some illustrated algebra textbook can use these

## **Original Article**

relations in function evaluation problems. The probability calculation, determination of the correlation coefficients, scatter plots and the line of best fits were also in the interest of the junior investigators' learning of practical lessons in probability statics and algebra and in the analysis data in research. The results open up a new area of research if common people's vital records were made available for extended research. Astrological beliefs could be tested with practical data.

## **CONCLUSION**

The highest number of celebrities is born in the zodiac sign Aquarius. All other zodiacs have variations in the number of celebrities with the increase of sample sizes. The best-line fit in the scatter plot of the data shows that the number of celebrity births increases linearly as we go down the list of the zodiac signs. Zodiac signs and the number of celebrities are moderately related as shown by the correlation coefficients in the range of 0.50 to 0.60. The result and the procedure used have both the research and academic values.

## **REFERENCES**

- Astrology.com.au. (n.d.). 12 signs of the zodiac. <https://www.astrology.com.au/astrology/12-signs-of-the-zodiac>
- Statisticshowto.com. (n.d.). How to compute Pearson's correlation coefficients.  
<https://www.statisticshowto.com/articles/how-to-compute-pearseons-correlation-coefficients/>
- Wikipedia contributors. (n.d.). Zodiac. Wikipedia. <https://www.wikipedia.org>