International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(11s) 2026 pp. 370-378
https://doi.org/10.70917/ijcisim-2026-3760
Article

The Study of density of prime numbers across equal intervals

Kalyani Shrikrishna Namjoshi[1], Prof. Dr. Sachin Kadam[2*]
1IMED Bharati Vidyapeeth (Deemed to be University)
2*Professor of Computer Applications, Head - Ph.D. Section Bharati Vidyapeeth (Deemed to be University)
kalyani.namjoshi@maharshikarvebcapune.org
sachin.a.kadam@bharatividyapeeth.edu

Abstract: Prime numbers are those numbers which are divisible by one and the number itself. Prime numbers are useful in data security and cybersecurity. In this paper all prime numbers between 2 and 100000 are considered. These numbers are divided into five equal intervals. Pairwise proportion tests were conducted over these intervals. The results for the test are as follows: Significant differences in the pairs A-D, B-E, C-E, and D-E are found. The remaining pairs A-B,A-C,B-C,C-D do not have significant differences. This concludes that the distribution of prime numbers is not uniform in the considered intervals. These outcomes confirm that the prime density is not uniform in the intervals, and the last partition has a much higher concentration of primes than the uniform distribution assumption would imply.

Keywords: Prime Numbers, Prime Density, Statistical Analysis, Chi-Square Test, R Programming, Computational Number Theory, Cryptography, Data Visualisation

1. Introduction

A prime number is a positive number which is divisible by 1 and the number itself . Prime numbers are very useful in data security .A prime number is a useful tool in encryption and decryption. It keeps passwords and digital sensitive information safe during the transmission. RSA algorithms in cryptography, which protect the data using private keys, are applications of prime numbers .During online banking transactions, prime number-based public key infrastructure (PKI) is being used to protect the data. Blockchain and Bitcoin use hash functions to secure the data using the prime factorisation method. Generating the post-quantum tool for data security is a need of the future. According to the prime number theorem, the number of primes less than number “N” is N/logN. It can be interpreted as when the value of n increases, the prime numbers' frequency gets decreased.

The mathematical foundation for studying prime density came from the prime number theorem. The theorem states that prime density decreases as numbers become larger, but primes never disappear completely (Apostol, 1976). Researchers have worked on the distribution of primes within finite intervals. Hardy and Wright (2008) demonstrated that although primes become less frequent as numbers increase, their distribution remains sufficiently predictable for practical

cryptographic applications. The decreasing in density implies that larger cryptographic keys require extra candidate numbers to be tested before a prime is found. Cryptographic algorithms require strong skills of prime generation. Menezes et al. (1996) studied that the density of prime number generation is depends upon the statistical density based on the formula log(n)/nlog(n) Koblitz (1994) observed that as the numerical range increases, the density of prime numbers decreases. This increases the computational cost of key generation

To overcome the cost, the Miller–Rabin primality test can be used The security is depends on integer factorisation and discrete logarithms. Rivest et al. (1978) work on RSA algorithm and conclude that factorisation of large prime numbers is key aspect of security. Boneh (1999) contributed on large intervals with more prime numbers reduce the security of key generation . . Bernstein and Lange (2017) conclude that 2048-bit and 4096-bit key sizes, are ideal for cryptographic operations.

The numbers 1 to 1 lakh are divided into five intervals, each of equal size of 20000. The ranges are 1-20000, 20001-40000, 400001-600000, 60001-80000, and 800001-100000. Using R code Find all the prime numbers between corresponding intervals. Find the count of intervalwise prime numbers. For the chi-square test, the null hypothesis (H0) is assumed as “Prime numbers are equally distributed across all ranges." The alternative hypothesis (H_1) is “Prime numbers are not equally distributed." As $p < 0.05$, reject the null hypothesis. . Thus, the prime numbers are not equally distributed .To find which pair has a significant difference, a comparison of all pairs with each other is carried out, and the study identifies which pair has a significant difference .

The workflow of the experiment is shown below

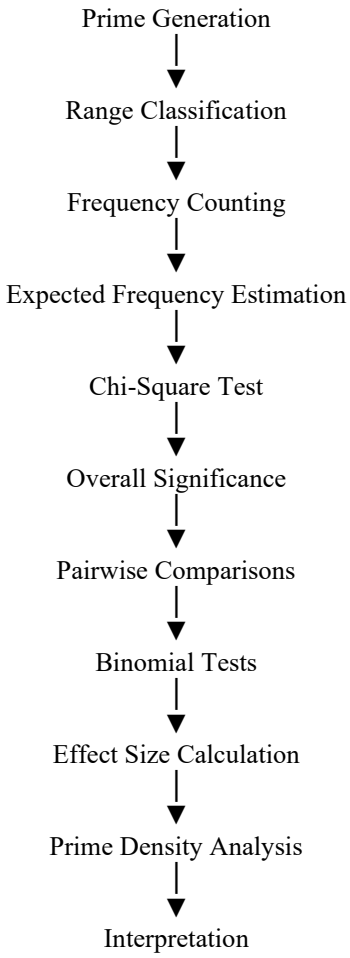

Fig 1: Workflow Diagram

Key Applications Relying on Non-Uniform Primes

- Public-Key Cryptography - In modern systems like RSA, encryption is dependent on the fact that it is difficult for a person to factor very large numbers when those numbers are the product of two large primes. By the irregular way that primes occur and the large distances between them, it is nearly impossible for a system to "guess" the prime factors if the private key is not present.
- Hash Tables & Algorithms - In computer science, it is common for hash functions to use prime numbers for the size of an array or for modulo operations. Because primes do not appear at regular intervals and share no common factors with most smaller integers, they ensure that patterns in data are not predictable. To use the numbers is to ensure that data collisions in a Hash Table are less frequent.
- Error-Correcting Codes - If engineers use the mathematical properties of primes to create finite fields, they can build codes that correct errors in telecommunications, deep space transmission and digital

storage. There is an ability in those codes to detect and automatically fix data that becomes corrupted when a system transfers it.

2. Literature review

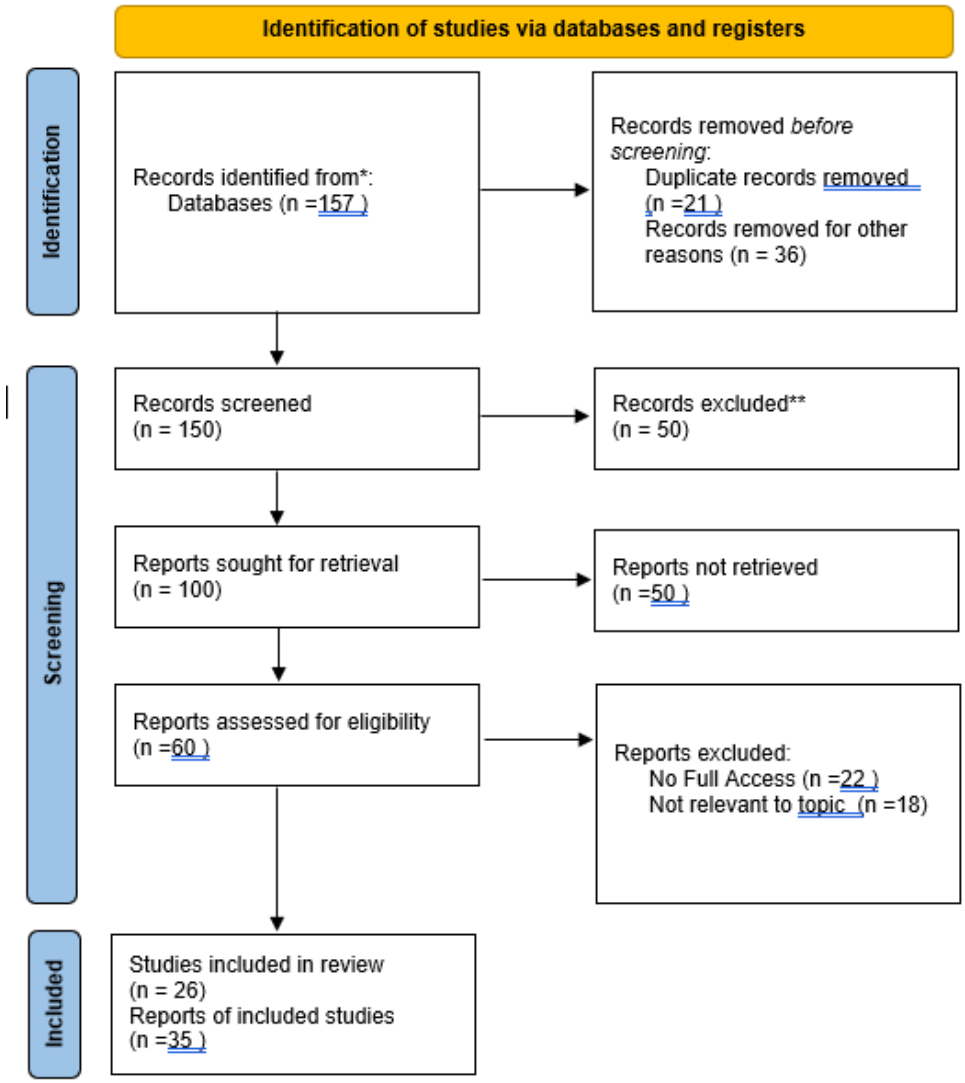

Figure 2: PRISMA Flow Diagram

Public-key cryptography relies on the properties and frequency of prime numbers to stay secure. Algorithms such as RSA and Diffie-Hellman work because it is difficult to factor these numbers (Mamatha & S.S, 2022) To keep systems strong, developers must use large, reliable primes during key generation (Ivanov & Stoianov, 2023). Safe primes are also vital. The process to identify prime is difficult and how often they come helps make these applications more secure (Diaz et al., 1999).New research shows that irregular gaps between primes can help in randomness and hashing .("Primes Don't Play Fair", 2025). Some new tech, like Fowler-Nordheim integrators, might be used to create and study primes in the future (Zhou et al., 2017).

Sr. No.	Author(s)	Year	Study/Contribution	Methodology	Key Findings	Research Gap
1	Hadamard & de la Vallée Poussin	1996	Proof of the Prime Number Theorem	Mathematical analysis	Established that the density of prime numbers decreases approximately as 1/log(n).	Limited computational validation using modern data analytics techniques.
2	Hardy & Wright	1979	Comprehensive study of number theory and prime distribution	Theoretical number theory	Described asymptotic behavior and properties of prime numbers.	Focused on theoretical aspects rather than statistical visualization.
3	Riesel	1994	Computational methods in prime number research	Algorithmic approaches	Presented efficient techniques for prime generation and testing.	Did not analyze density variations using statistical hypothesis testing.
4	Crandall & Pomerance	2005	Prime numbers and computational applications	Computational number theory	Demonstrated applications of prime numbers in cryptography and large-scale computation.	Limited emphasis on interval-based density analysis.
5	Oliveira e Silva et al.	2014	Verification of prime counting functions for large ranges	Large-scale computational experiments	Confirmed theoretical predictions of prime distribution over extensive intervals.	Focused on validation of counting functions rather than density visualization.
6	Granville	2015	Modern perspectives on prime distribution	Analytical number theory	Discussed irregularities and patterns in prime occurrence.	Lacked empirical statistical testing using partitioned intervals.
7	Caldwell & Cheng	2020	Computational analysis of prime records and distributions	Numerical computation	Explored large prime datasets and computational trends.	Did not investigate prime density through goodness-of-fit testing.
8	Present Study	2026	Statistical analysis of prime density in the range 1–100,000	R programming, density plots, Chi-Square test	Demonstrates significant variation in prime density across equal intervals and statistically rejects uniform distribution.	Provides a reproducible framework that combines visualization and hypothesis testing for prime density analysis.

Contributions

This paper looks at how prime numbers are spread out between 1 and 100,000. The prime number theorem already explains this theory. This study adds a practical look at the data using the R programming language. The aim of the study is to divide prime numbers into equal range .and find range wise density of prime numbers. Chi Square test proves that the density of prime numbers is unequal over the intervals The same method could work for bigger number ranges. .Also Applicable to Fermat twin prime numbers The study connects math theory with computer tests. It provides a clear method that others can repeat to study how primes are distributed. The numbers 1 to 1 lakh are divided into five intervals, each of equal size of 20000. The ranges are 1-20000, 20001-40000, 400001-600000, 60001-80000, and 800001-100000. Using R code Find all the prime numbers between corresponding intervals. Find the count of intervalwise prime numbers. For the chi-square test, the null hypothesis (H0) is assumed as “Prime numbers are equally distributed across all ranges." The alternative hypothesis (H_1) is “Prime numbers are not equally distributed." As $p < 0.05$, reject the null hypothesis. . Thus, the prime numbers are not equally distributed .To find which pair has

a significant difference, a comparison of all pairs with each other is carried out, and the study identifies which pair has a significant difference

3. Material and methodology

Data Collection

Prime numbers in the range of 1 to 100,000 were generated using the R programming language. A sequence of integers from 1 to 100,000 was created, and prime numbers were identified using the isprime() function. The resulting dataset consisted of all prime numbers within the specified range.

Data Partitioning

The complete range (1–100,000) was divided into five equal intervals of 20,000 numbers each:

- Range E: 1–20,000
- Range D: 20,001–40,000
- Range C: 40,001–60,000
- Range B: 60,001–80,000
- Range A: 80,001–100,000

Each prime number was assigned to its corresponding interval based on its numerical value.

Data Visualization

The researchers used the ggplot2 package in R to make density plots. These plots showed how prime numbers were spread across the intervals. The facet_wrap() function is used to visualise separate plots for each interval.

Statistical Analysis

A Chi-Square Goodness-of-Fit Test checked if primes were spread evenly across the five intervals. The null hypothesis (H0) was that primes appear with the same frequency in every interval. The alternative hypothesis (H1) was that the frequencies are different.

H0: There is no difference in frequencies of prime number occurrences across all intervals.

H1: Frequencies of prime number occurrences are not equal across all intervals. The team counted the primes in each interval to get the observed frequencies. They then found the expected frequencies by assuming the primes were split evenly. The result shows that the pairs A-D, A-E,B-D ,B-E,C-E and D-E have significant differences .This shows that group E is significantly different from all other groups. This is shown in the table below

Rowwise:

c1 c2 f1 f2 total p1 p2 pair p-value result effect_size

<chr> <chr> <int> <int> <int> <dbl> <dbl> <chr> <dbl> <chr> <dbl>

1 A B 1755 1780 3535 0.496 0.504 A - B 0.686 NoSigDiff 0

2 A C 1755 1854 3609 0.486 0.514 A - C 0.103 NoSigDiff 0

3 A D 1755 1941 3696 0.475 0.525 A - D 0.002 SigDiff 0.05

4 A E 1755 2262 4017 0.437 0.563 A - E 0 SigDiff 0.13

5 B C 1780 1854 3634 0.490 0.510 B - C 0.226 NoSigDiff 0

6 B D 1780 1941 3721 0.478 0.522 B - D 0.009 SigDiff 0.04

7 B E 1780 2262 4042 0.440 0.560 B - E 0 SigDiff 0.12

8 C D 1854 1941 3795 0.489 0.511 C - D 0.163 NoSigDiff 0

9 C E 1854 2262 4116 0.450 0.550 C - E 0 SigDiff 0.1

10 D E 1941 2262 4203 0.462 0.538 D - E 0 SigDiff 0.08

Table 1: pairwise significant difference

Experiment

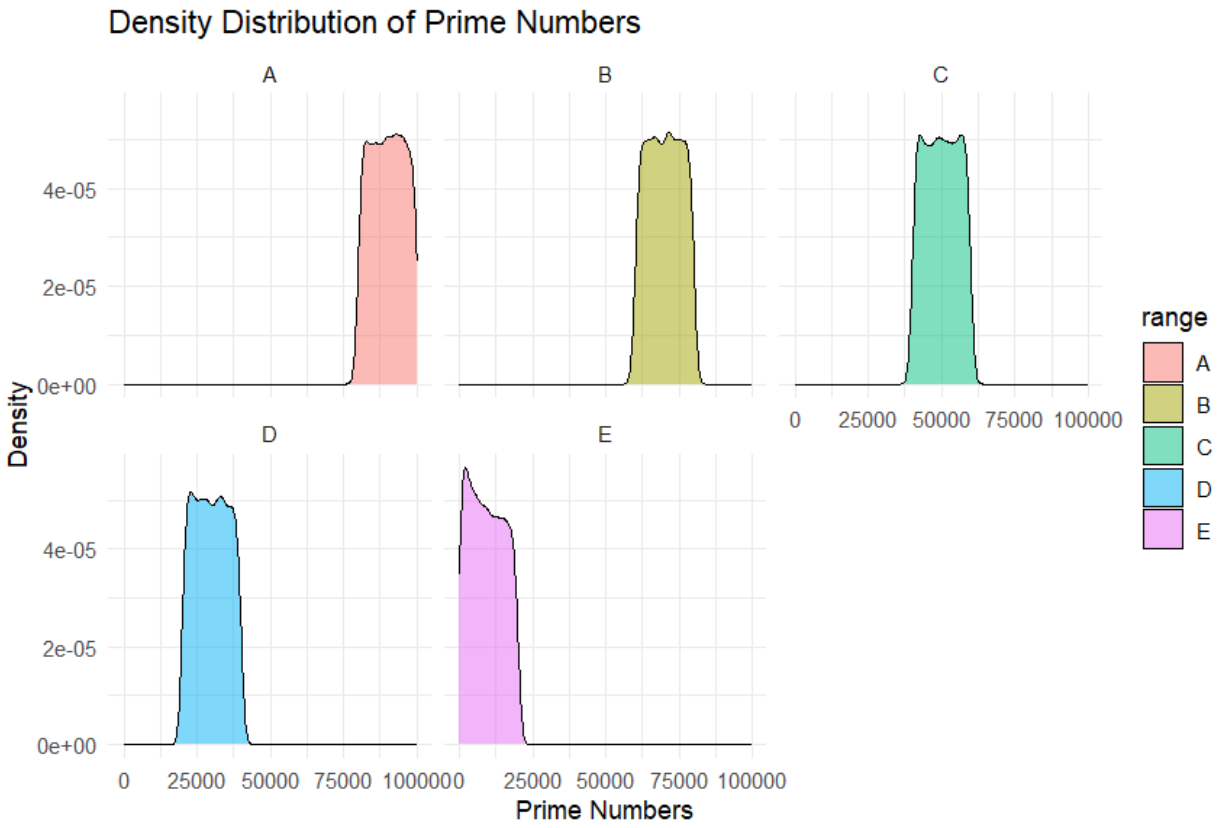

Figure 2: Density Distribution of Prime Numbers

This study looks at prime numbers between 1 and 1,000,000. The total range is split into five equal parts. This allows for a comparison of how primes are spread across different intervals. Each group has the same amount of numbers. The objective of this study is to find groupwise density. using function in R programming language Output drawn as follows 1) count of prime numbers in each interval 2) To find density of prime in each group 3) groupwise comparison to find significant difference between intervals . Result shows that as number increases the density of prime number decsrease

Using R makes the work faster. It handles the counting and sorting without manual errors. This data helps with math research, coding, and security studies.The work uses R to study prime numbers up to one lakh. Dividing the range into parts shows a clear image of how prime density drops over time. And the significant difference is found in pairwise comparison or not .

Validation and verification

Chi-squared test for given probabilities

data: prime_counts

X-squared = 87.872, df = 4, p-value < 2.2e-16

Chi-Square Test Conclusion

Test Result: (from output of R program)

- Chi-square statistic (χ^2) = 87.872
- Degrees of freedom (df) = 4
- p-value < 2.2×10^{-16}

Decision

Since the p-value is much smaller than 0.05, reject the null hypothesis (H_0).

Interpretation

The observed distribution of prime numbers across the five intervals differs significantly from the expected distribution under the assumption of equal proportions. Therefore, prime numbers are not uniformly distributed across ranges A, B, C, D, and E.

Pairwise comparisons show significant differences between the pairs A–D, A–E, B–D, B–E, C–E, and D–E. This indicates that group E differs significantly from all other groups.

Practical Implications of the Research

Studying how prime numbers are distributed benefits mathematics, computer science, and data analysis. Identifying patterns in prime distribution can lead to more efficient computational methods.

- Cryptography: Prime numbers underpin RSA encryption, used for generating public keys and digital signatures.
- Algorithm design: Research on prime density supports the development of faster algorithms for searching and factoring numbers. Understanding whether primes are common or rare in a given range can improve the efficiency of high-performance computing systems.
- Number theory: As a core branch of mathematics, number theory relies on density data to test theoretical models of prime behavior.
- AI and data science: Machine learning tools can analyze prime number data to detect underlying patterns.
- Blockchain and security: Hash functions, which draw on prime number properties, help secure communication channels and login systems in real time.
- Visualization: A visual demonstration of prime density can aid in understanding and communicating these patterns.

Limitations

- If the distribution pattern is fully characterized using the prime number theorem, it could make it easier to break certain security systems that rely on prime unpredictability.
- Analyzing large ranges of prime numbers requires significant time and memory resources.
- Greater randomness in prime distribution increases the difficulty — and thus the computational cost — of breaking associated security systems (note: check this claim's direction, since more randomness usually strengthens rather than weakens security).

4. Conclusion

Studying prime density can help strengthen prime-based security tools. Understanding the randomness and distribution of primes supports the development of more secure cryptographic methods. Since prime density affects encryption strength, this research may contribute to future advances in cryptographic algorithms, including applications in quantum cryptography.

Future Work

- Insights into prime number density can inform the development of more secure cryptographic algorithms.
- This study's methodology could be extended to very large prime numbers.
- Machine learning could be applied to reduce key generation time.
- This work could serve as a foundation for research in big data analysis and parallel computing.

References

1. Ezz-Eldien, A., Ezz, M., Alsirhani, A., Mostafa, A. M., Alomari, A., Alserhani, F., & Alshahrani, M. M. (2024). Computational challenges and solutions: Prime number generation for enhanced data security. PloS one, 19(11), e0311782.
2. Pevnev, V., Yudin, O., Sedlaček, P., & Kuchuk, N. (2024). Method of testing large numbers for primality. Advanced Information Systems, 8(2), 99-106.
3. Assa-Agyei, K., Owa, K., Al-Hadhrami, T., & Asafo, S. K. (2025, February). Advanced Techniques for Prime Number Generation in RSA Encryption: A Path Toward Greater Efficiency. In 2025 International Conference on Computing, Networking and Communications (ICNC) (pp. 479-484). IEEE.
4. Yang, Y., Liu, C., & Chen, S. C. I. (2025, July). Security Analysis of Four-Prime RSA scheme. In 2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET) (pp. 1-4). IEEE.
5. Adusumalli, M. S. (2024). A study of the importance of prime numbers in cryptographic algorithms. Int J Univers Sci Eng (IJUSE), 10, 1-10.
6. MAMATHA, N., & SUNITHA, S. (2022). EXPLORING THE PROPERTIES OF PRIME NUMBERS IN CRYPTOGRAPHY. WORLD, 13(1), 885-894.
7. Stiglic, A. (2025). Prime Number. In Encyclopedia of Cryptography, Security and Privacy (pp. 1869-1876). Cham: Springer Nature Switzerland.
8. Karatsuba, A. A. (1990). The distribution of prime numbers. Russian Mathematical Surveys, 45(5), 99-171.
9. Rives, J. (2025). The Laws of the Minor Prime Factors.
10. Lacey, L. F. Statistical properties of the infinite set of twin prime numbers.
11. Bahrami, D. (2021). An observation on distribution of prime numbers.
12. Maurer, U. M. (1995). Fast generation of prime numbers and secure public-key cryptographic parameters. Journal of Cryptology, 8(3), 123-155.
13. Tenenbaum, G., France, M. M., & Spain, P. G. (2000). The prime numbers and their distribution (Vol. 6). Providence, RI: American Mathematical Society.
14. LeVeque, W. J., Niven, I., & Zuckerman, H. S. (1961). An introduction to the theory of numbers.
15. Singh, D. (2025). Recent Advances and Applications in Modern Number Theory: A Comprehensive Review. Pi International Journal of Mathematical Sciences, 1(1), 33-48.
16. Green, F. (2019). Review of Number Theory: An Introduction via the Density of Primes. ACM SIGACT News, 50(1), 9-13.
17. Apostol, T. M. (1976). Introduction to analytic number theory. Springer.
18. Bernstein, D. J., & Lange, T. (2017). Post-quantum cryptography. Nature, 549(7671), 188–194.
19. Boneh, D. (1999). Twenty years of attacks on the RSA cryptosystem. Notices of the American Mathematical Society, 46(2), 203–213.
20. Hardy, G. H., & Wright, E. M. (2008). An introduction to the theory of numbers (6th ed.). Oxford University Press.
21. Koblitz, N. (1994). A course in number theory and cryptography (2nd ed.). Springer.
22. Menezes, A. J., van Oorschot, P. C., & Vanstone, S. A. (1996). Handbook of applied cryptography. CRC Press.
23. Rivest, R. L., Shamir, A., & Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 21(2), 120–126.
24. Easttom, C. (2025). Quantum Resistant Cryptography and Cyberwarfare. International Conference on Cyber Warfare and Security, (), 71-78.
25. RSA vs ECDSA. https://www.hackingnote.com/en/versus/rsa-vs-ecdsa/
26. Budiman, M. A., Rachmawati, D., & Utami, R. G. (2019). The cryptanalysis of the Rabin public key algorithm using the Fermat factorization method. Journal of Physics. https://doi.org/10.1088/1742-6596/1235/1/012084
27. Find prime groups after dividing the range 1 to N into K equal sections - GeeksforGeeks. https://origin.geeksforgeeks.org/find-prime-groups-after-dividing-the-range-1-to-n-into-k-equal-sections/

28. Anikeeva, O., Brennan, D., & Teusner, D. (2013). Household income modifies the association of insurance and dental visiting. BMC Health Services Research. https://doi.org/10.1186/1472-6963-13-432
29. 150 Fun Math Facts That Will Surprise You - June 2026. https://ponly.com/math-facts/
30. citation()
31. To cite R in publications use:
32. R Core Team (2025). _R: A Language and Environment
33. for Statistical Computing. R Foundation for
34. Statistical Computing, Vienna, Austria.
35. <https://www.R-project.org/>.
36. A BibTeX entry for LaTeX users is
37. @Manual{,
38. title = {R: A Language and Environment for Statistical
39. Computing},
40. author = {{R Core Team}},
41. organization = {R Foundation for Statistical Computing},
42. address = {Vienna, Austria},
43. year = {2025},
44. url = {https://www.R-project.org/},
45. }
46. We have invested a lot of time and effort in creating
47. R, please cite it when using it for data analysis. See
48. also ‘citation("pkgname")’ for citing R packages.