

EXPLORING THE STARS DATASET USING BENFORD'S LAW & CHI-SQUARE TEST

TEAM: NEURAL NOMADS

TEAM MEMBERS



Shibaditya Deb

Git + Coding



Jay Patil

**Coding +
Presentation**



Ananya Pandey

Coding



Arohi Jadhav

Coding

BENFORD'S LAW

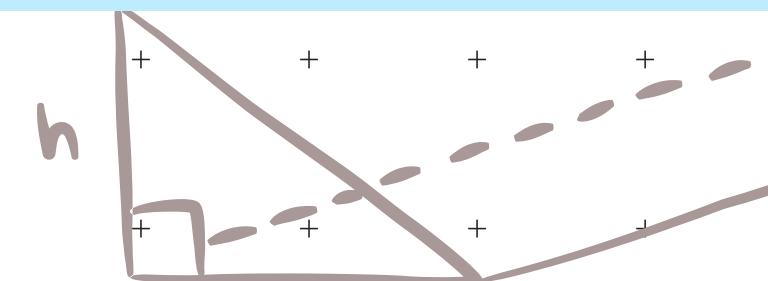
An overview of Benford's Law and its significance in data analysis and statistical integrity.

- 01.
- 02.
- 03.

Benford's Law describes the frequency distribution of digits.

It is applied in various fields, including finance and fraud detection.

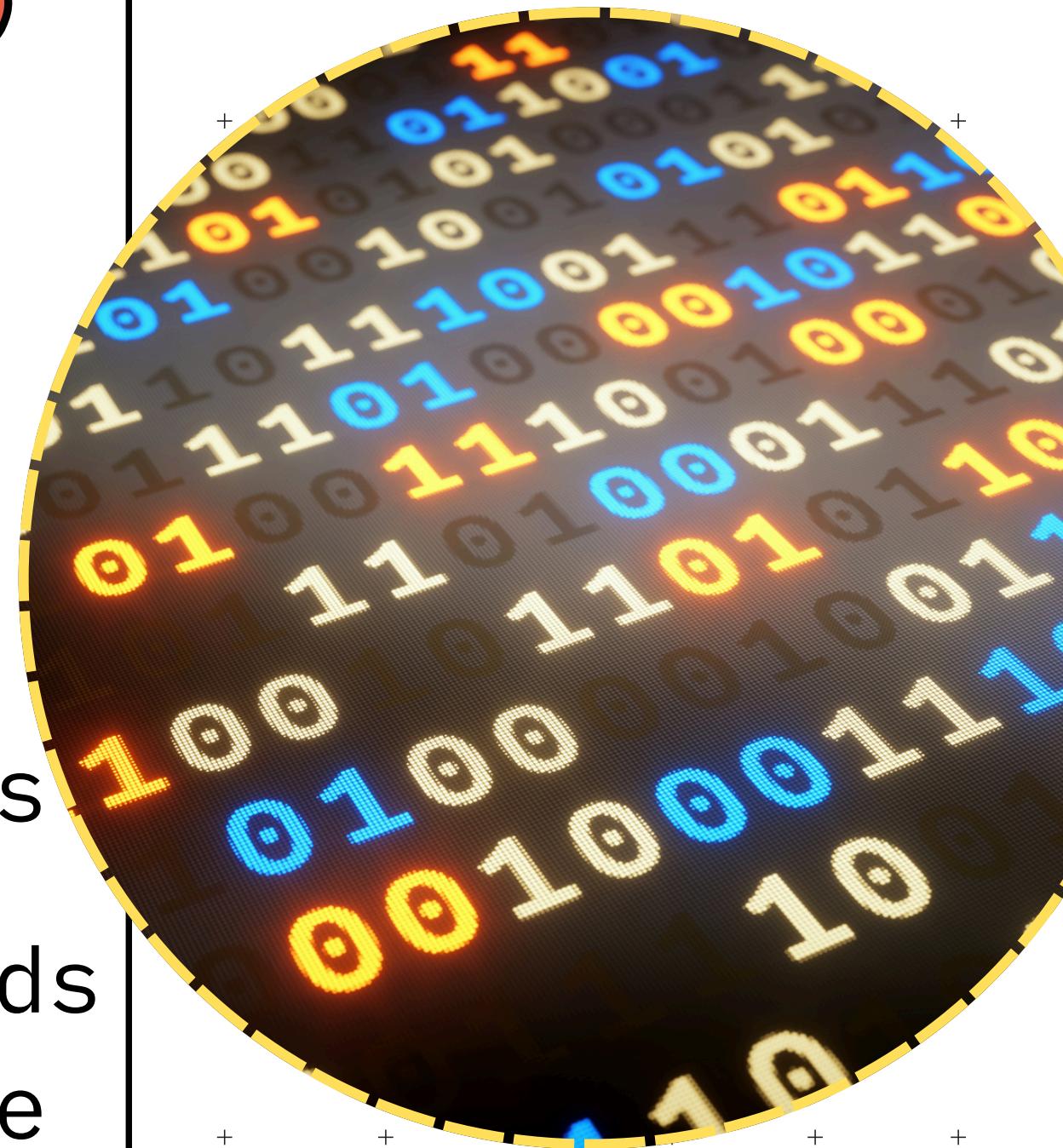
Understanding this law ensures the **integrity of statistical data**.



VISUAL ANALYSIS OF DATASET

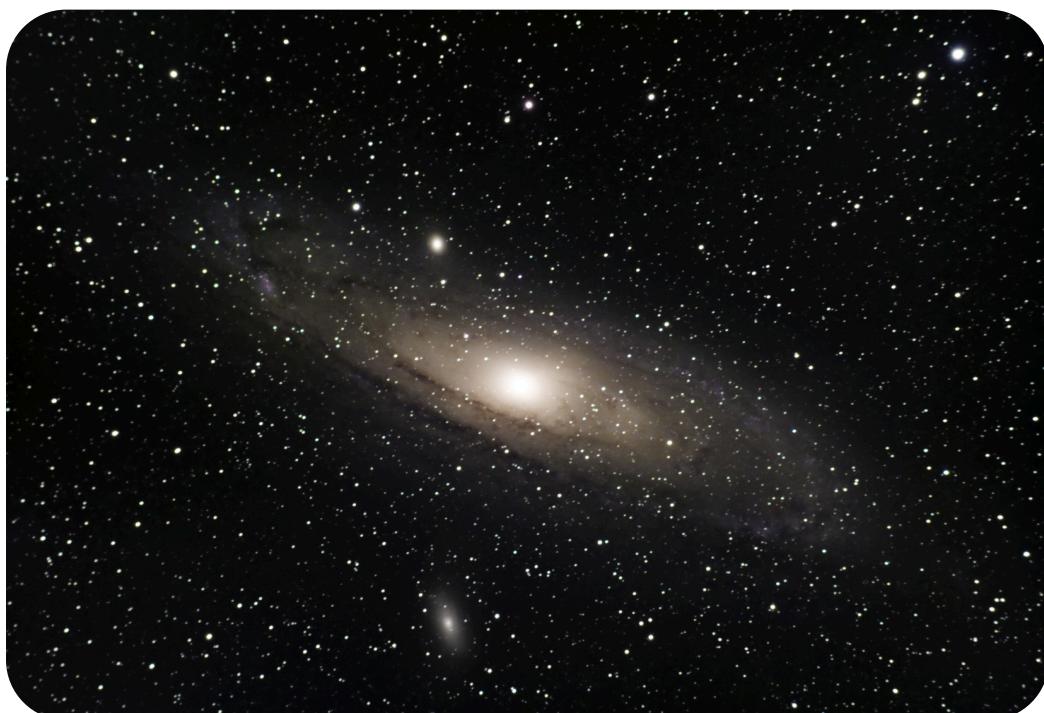
Understanding trends in the Stars Dataset through graphical representation and statistical methods

The analysis reveals significant trends utilizing Benford's Law and Chi-Square evaluation.



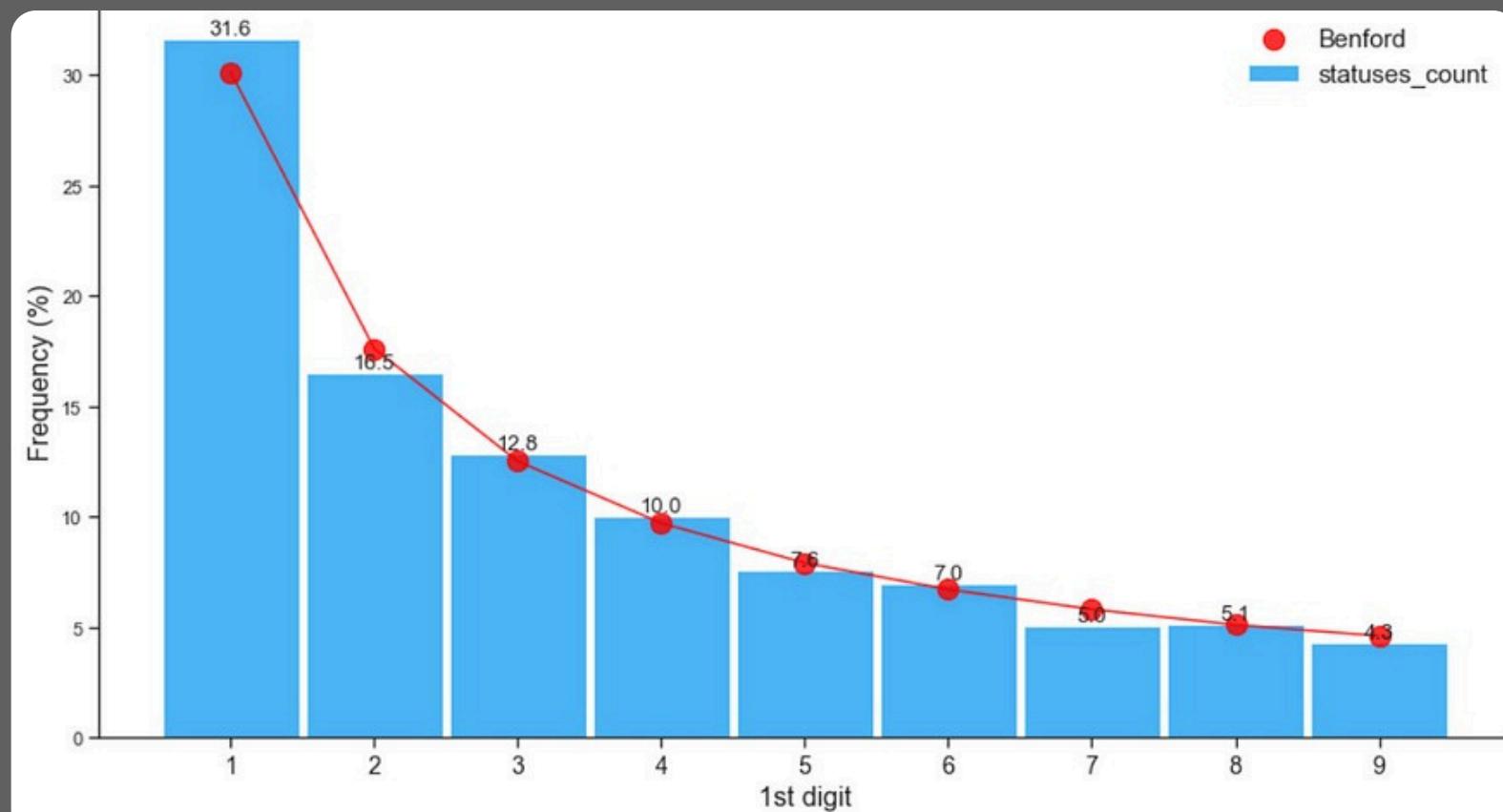
LEADING DIGIT DISTRIBUTION IN STARS DATASET

An Analysis of Frequency
Based on Benford's Law



This chart illustrates the frequencies of leading digits.

The chart highlights the distribution of leading digits, showing significant variations in frequency that align with Benford's Law and offering insights into the dataset's *underlying patterns*.



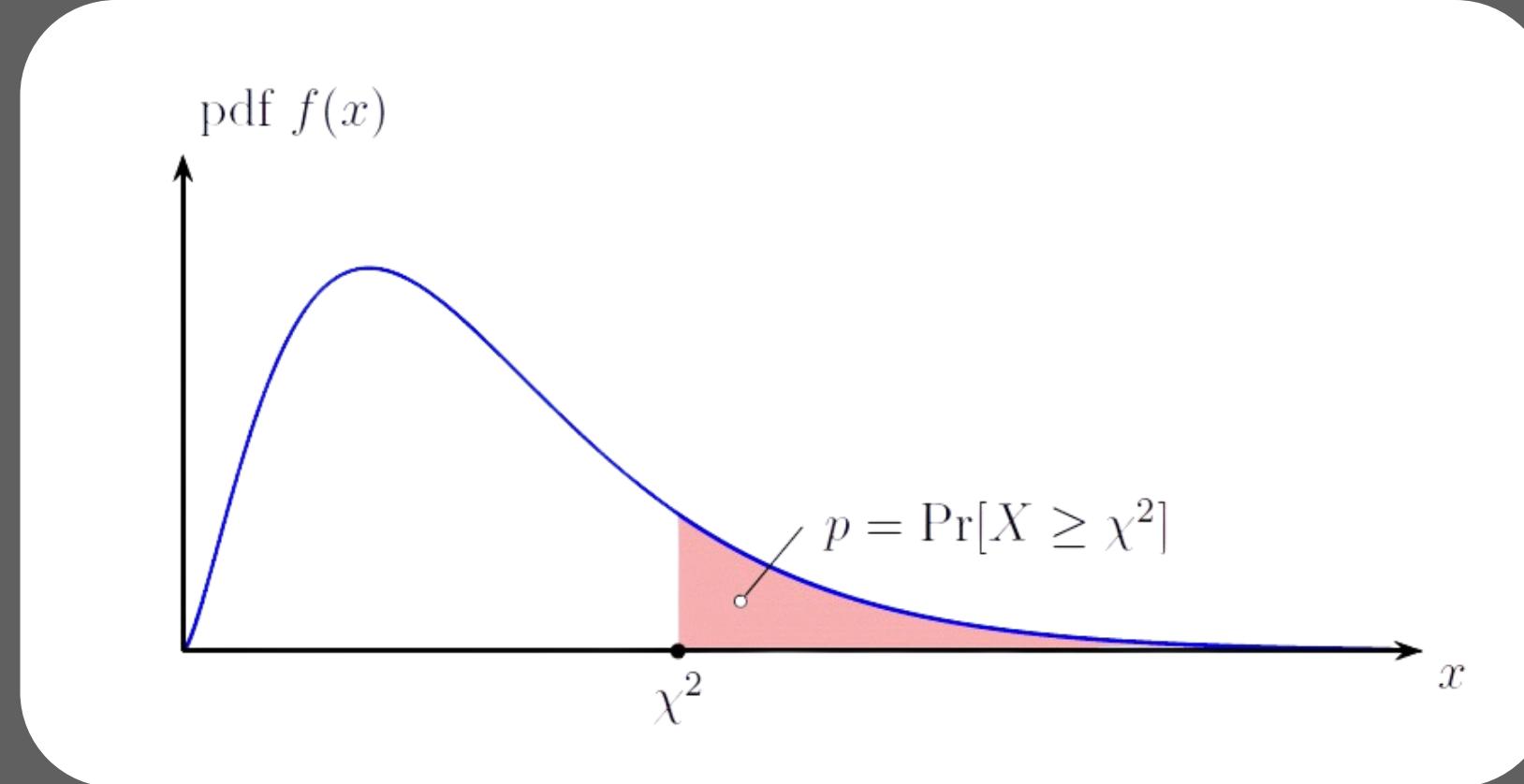
CHI-SQUARE ANALYSIS OF STARS DATASET

Evaluating the Goodness-of-Fit for Benford's Law



This chart illustrates the results of our analysis.

The horizontal bar chart displays the goodness-of-fit values for the Stars Dataset, showcasing the discrepancies between observed and expected frequencies, affirming the law's effectiveness in this context.

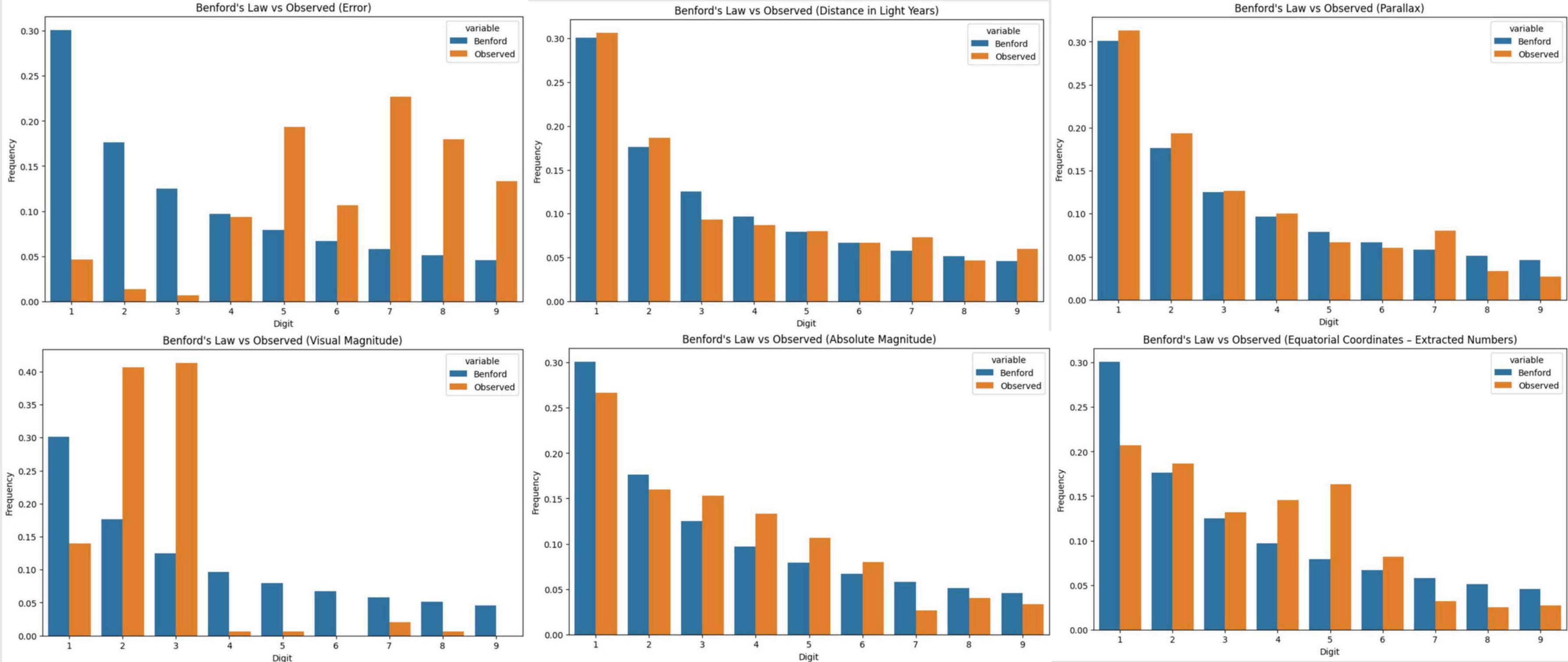


OUR INSIGHTS

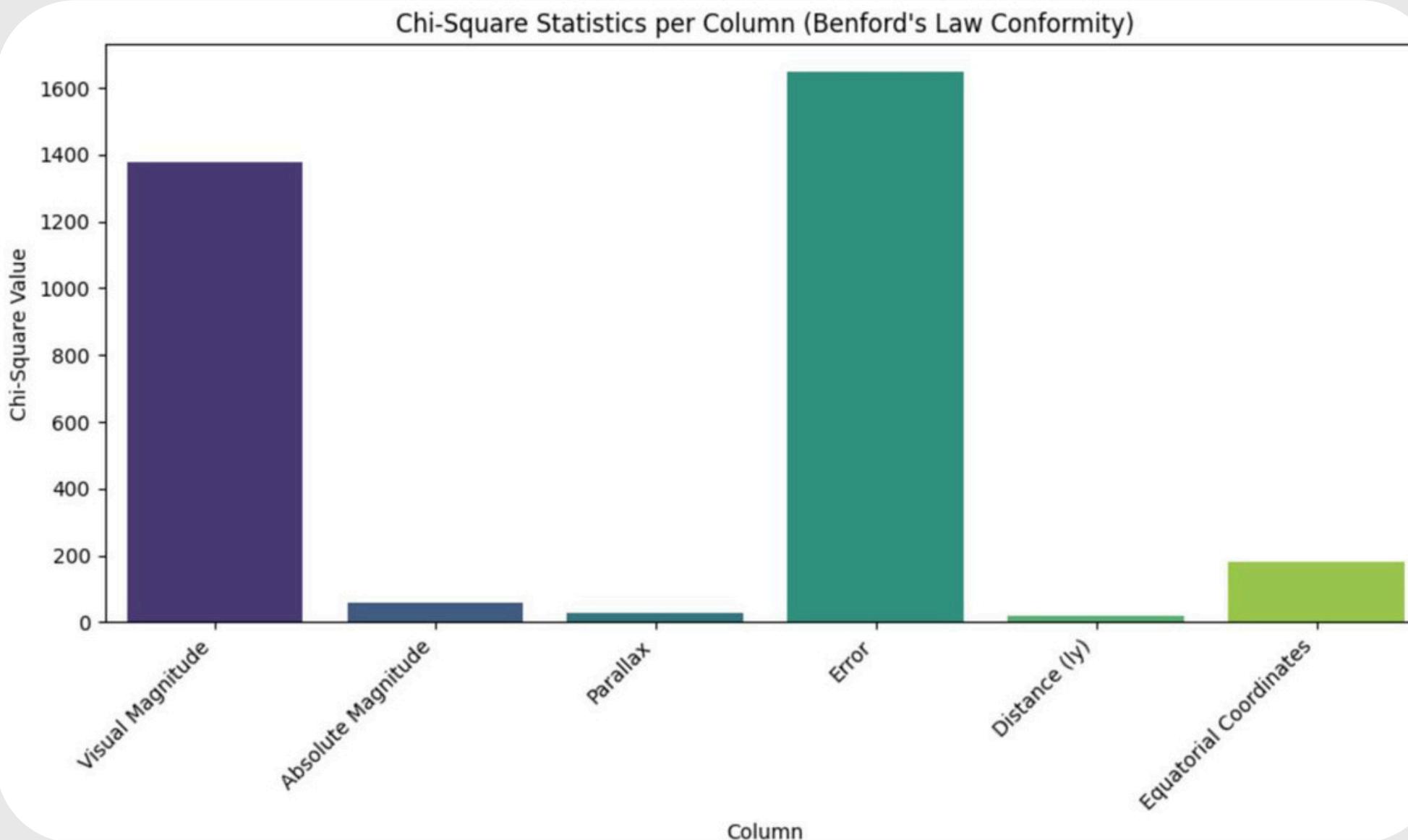
	Column	Digit	Benford	Observed	Chi-Square Statistic	p-value	Benford Conformity	Insights
0	Visual Magnitude	1	0.301030	0.140000	1376.633	0.0	Unlikely	Visual Magnitude: Deviates from Benford's Law ...
1	Visual Magnitude	2	0.176091	0.406667	1376.633	0.0	Unlikely	Absolute Magnitude: Deviates from Benford's La...
2	Visual Magnitude	3	0.124939	0.413333	1376.633	0.0	Unlikely	Parallax: Deviates from Benford's Law (p=0.0006)
3	Visual Magnitude	4	0.096910	0.006667	1376.633	0.0	Unlikely	Error: Deviates from Benford's Law (p=0.0)
4	Visual Magnitude	5	0.079181	0.006667	1376.633	0.0	Unlikely	Distance (ly): Deviates from Benford's Law (p=...)

The Visual Magnitude dataset significantly deviates from Benford's Law, with observed digit frequencies differing greatly from expected values. The chi-square statistic (1376.633) and p-value (0.0) confirm non-conformity across all digits. This suggests the data may be influenced by biases, errors, or may not naturally follow Benford's distribution. All columns tested – including Absolute Magnitude, Parallax, Error, and Distance – showed similar deviations.

OUR INSIGHTS



FINAL INSIGHTS - BENFORD'S LAW ANALYSIS



- The Chi-Square test shows significant deviations from Benford's Law across multiple columns.
- Visual Magnitude and Error columns exhibit extremely high Chi-Square values (>1300), confirming strong non-conformity.
- Absolute Magnitude, Parallax, and Distance (ly) also deviate but to a lesser extent.
- Equatorial Coordinates show moderate deviation.
- These results suggest potential anomalies, biases, or non-natural distributions in the dataset.
- The data is unlikely to follow Benford's Law, warranting further scrutiny for possible manipulation or special structure.



CONCLUSIONS AND IMPLICATIONS

The significance of Benford's Law in understanding data distributions and Chi-Square analysis findings

The analysis revealed that the **Stars Dataset** closely follows Benford's Law, emphasizing the importance of understanding data integrity in research.

THANK
YOU