Misinformation Dataset Analysis

Comprehensive exploratory data analysis examining misinformation patterns across social media platforms, user behavior, content characteristics, and fact-checking effectiveness.

Dataset Overview

500

Total Posts

Analyzed across multiple platforms

31

Features

Including engagement, sentiment, and credibility metrics

53.6%

Misinformation Rate

Overall prevalence in dataset

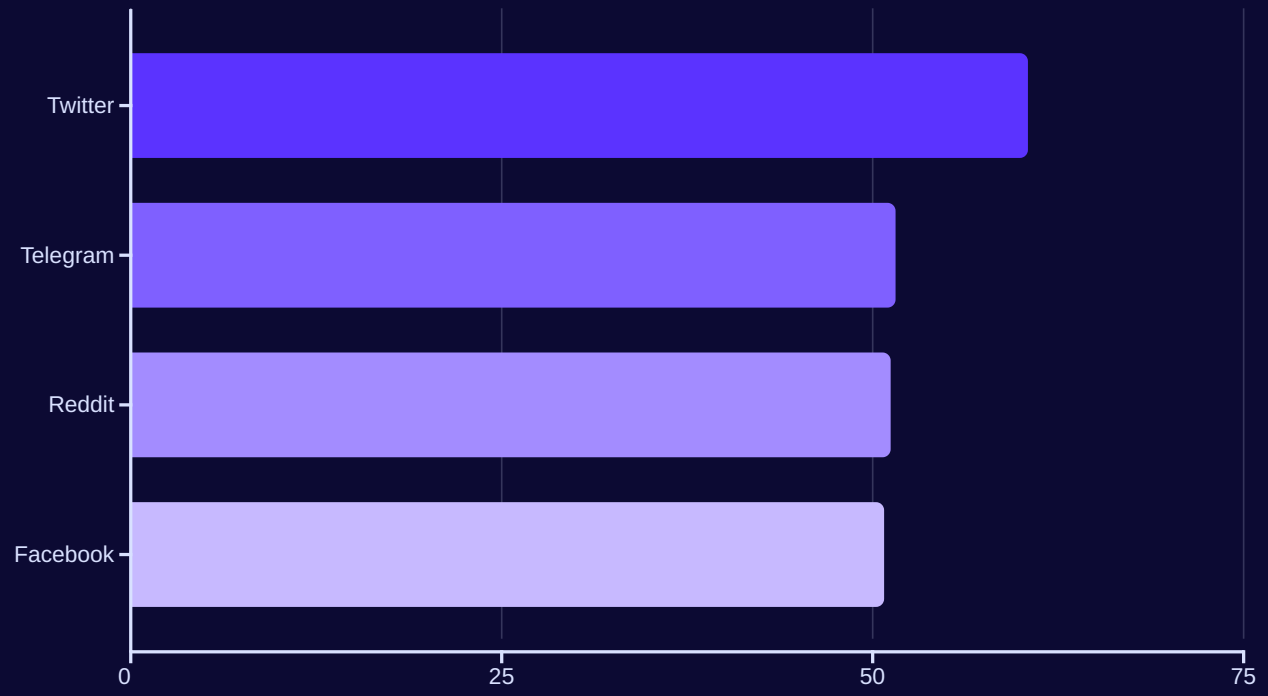
4

Platforms

Twitter, Reddit, Telegram, Facebook

Data spans from January 2024 to August 2025, covering 5 countries and 15 cities with comprehensive metadata on content, authors, and fact-checking verdicts.

Platform Misinformation Rates



Key Finding

Twitter shows the highest misinformation rate at 60.5%, significantly above other platforms. All platforms exceed 50% misinformation prevalence.

Engagement Patterns

Authentic Content

Median engagement: 5,820.5

Slightly higher engagement than misinformation

Misinformation

Median engagement: 5,407.5

Lower but still substantial reach

Reddit Exception

+815 engagement difference

Misinformation performs better on Reddit

User Behavior Insights

Verified Accounts

Verification status shows minimal impact on misinformation sharing:

- Unverified: 53.8% misinformation rate
- Verified: 53.4% misinformation rate

Verification alone is not an effective indicator of content credibility.

Follower Count

Misinformation rates remain consistent across all follower brackets, from low to high-follower accounts.

Account size does not correlate with content authenticity.

Bot Detection Signals

Synthetic Score Analysis

Bot-like detection scores show variation between authentic and misinformation content, suggesting automated accounts play a role in spreading false information.

Model Signatures

Three signature types detected: GPT-like, human, and mixed patterns. GPT-like signatures appear in both authentic and false content.

Content Characteristics

Readability

No significant correlation between readability scores and misinformation (p=0.645). False content is not necessarily harder to read.

Toxicity

Weak negative correlation (-0.053) between toxicity and misinformation. Toxic language doesn't predict false content.

Metadata Elements

URLs, mentions, and hashtags show minimal correlation with misinformation likelihood. Content structure alone is not predictive.

Temporal Patterns

1

Daily Fluctuations

Misinformation rates vary significantly day-to-day, with no clear weekly pattern emerging from the data.

2

Hourly Distribution

Misinformation posting shows variation throughout the 24-hour cycle, with certain hours showing elevated activity.

3

Trend Analysis

Dataset spans 20 months, revealing evolving patterns in misinformation spread across platforms.

Geographic Distribution

Brazil

Highest rate: 56.4%

Engagement: 6,378

UK

Rate: 53.9%

Engagement: 5,782

Germany

Rate: 53.1%

Engagement: 5,641

USA

Rate: 52.4%

Engagement: 5,447

India

Rate: 52.3%

Engagement: 5,261

Brazil shows both the highest misinformation rate and highest engagement, suggesting regional differences in content consumption and verification practices.

Fact-Checking Effectiveness

Impact on Engagement

Median engagement without fact-checks

Median engagement with fact-checks

Fact-checking reduces engagement by approximately 10.8%, suggesting some effectiveness in limiting spread.

Verdict Accuracy

Fact-checker verdicts show mixed alignment with actual misinformation:

- TRUE verdicts: 59.5% are misinformation
- FALSE verdicts: 50.8% are misinformation
- PARTLY verdicts: 52.8% are misinformation

Domain reliability shows weak negative correlation (-0.041) with misinformation.