



Statistical Analysis of Per-Capita Alcohol Consumption in Iran, India, and Germany (2000–2020)

Achal Ajith (100004415)
Mohammadreza Hendiani (100001915)
Negin Ghanei (100003179)

May 2025

Abstract

This report presents a comprehensive statistical analysis of annual per-capita alcohol consumption (liters of pure alcohol, age 15+) in Iran, India, and Germany over the period 2000–2020. Employing descriptive statistics, distributional comparisons, parametric fitting, normality and trend tests, and bootstrap resampling, we characterize consumption patterns, assess theoretical assumptions, and quantify uncertainties across contrasting socio-cultural contexts. All analyses are implemented in a single reproducible Jupyter notebook environment.

1 Introduction

Alcohol consumption constitutes a major global health concern, with misuse impairing immune defenses and influencing disease risk and progression. The relationship between consumption and health outcomes varies across socio-economic and cultural environments. This study leverages World Bank data on per-capita alcohol consumption from 2000 to 2020 for three exemplar countries: Iran (strict regulation), India (moderate consumption), and Germany (high consumption). Our aims are to:

1. Characterize temporal trends and summary statistics per country.
2. Compare empirical distributions across countries via visual (histograms, ECDFs) and quantitative (Kolmogorov–Smirnov) methods.
3. Assess parametric assumptions through distribution fitting (normal and log-normal) and normality tests (Shapiro–Wilk, Anderson–Darling).
4. Detect long-term trends using non-parametric methods (Mann–Kendall).
5. Evaluate the sampling distribution of the mean through bootstrap resampling, illustrating the Central Limit Theorem in small samples.

2 Methodology

2.1 Data Collection

Data were sourced from the World Bank’s World Development Indicators, specifically annual total per-capita alcohol consumption (liters of pure alcohol, age 15+) for Iran, India, and Germany from 2000 to 2020.

2.2 Data Preparation

Raw data were filtered to remove duplicates and missing entries, retain only the three target countries, and ensure age 15+ consistency. Statistical features (mean, median, mode, IQR, min, max, standard deviation, variance) were computed. Visualizations include histograms with KDE overlays, empirical cumulative distributions, and Q–Q plots.

2.3 Analytical Methods

- Descriptive statistics (central tendency, dispersion).
- Distributional comparison: kernel density estimates, Kolmogorov–Smirnov tests.
- Parametric fitting: estimation of normal and log-normal parameters, Q–Q plots, Shapiro–Wilk tests.
- Trend detection: Mann–Kendall non-parametric test.
- Bootstrap resampling: 1,000 replicates of sample means per country, 95% confidence intervals, normality tests on bootstrap means.

3 Results and Discussion

3.1 Temporal Trends

The temporal evolution of consumption is shown in Figure 1. Germany starts at approximately 13.8 L in 2000, declining to 11.9 L by 2020. India rises from 2.0 L to a peak of 5.1 L in 2013, then falls to 4.1 L. Iran remains below 1.5 L throughout, peaking in 2011 and decreasing thereafter.

3.2 Distributional Comparisons

Table 1 summarizes empirical statistics. Iran exhibits strong right skew (skewness = +1.8), India left skew (skewness = -0.7), and Germany slight right skew (skewness = +0.5). Kolmogorov–Smirnov tests confirm significant differences between each pair (all $p < 0.01$).

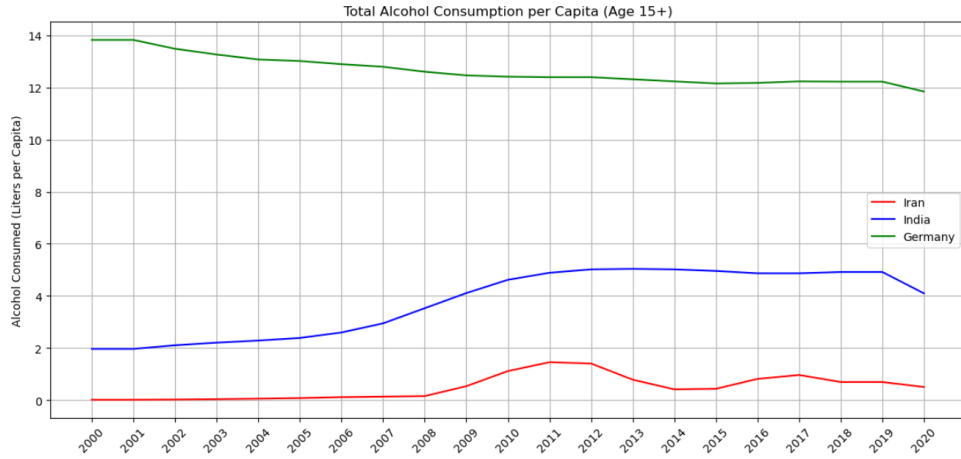


Figure 1: Annual per-capita alcohol consumption (2000–2020).

Table 1: Descriptive statistics of per-capita alcohol consumption (2000–2020).

Country	Mean	Median	Mode	IQR	Min	Max	Skewness
Iran	0.50	0.44	0.02	0.70	0.02	1.46	+1.8
India	3.78	4.11	1.97	2.53	1.97	5.04	−0.7
Germany	12.66	12.41	12.22	0.78	11.84	13.82	+0.5

3.3 Visual Distributional Comparisons

3.4 Parametric Fitting and Normality

Shapiro–Wilk tests (Table 2) reject normality for Iran and India ($p < 0.01$) and marginally for Germany ($p = 0.04$).

Table 2: Shapiro–Wilk normality test results.

Country	p -value	Conclusion
Iran	< 0.01	Reject normality
India	< 0.01	Reject normality
Germany	0.04	Marginally non-normal

3.5 Trend Tests

Mann–Kendall tests (Figure 6) indicate significant increasing trends in Iran ($p = 0.00016$) and India ($p = 0.00003$), and a significant decreasing trend in Germany ($p < 10^{-8}$).

3.6 Bootstrap and Central Limit Theorem

Bootstrap resampling yields distributions approximating normality for India and Germany but remaining non-normal for Iran.

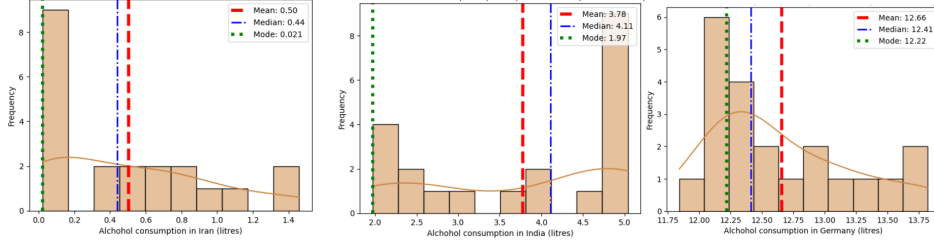


Figure 2: Histograms and KDE plots for alcohol consumption by country.

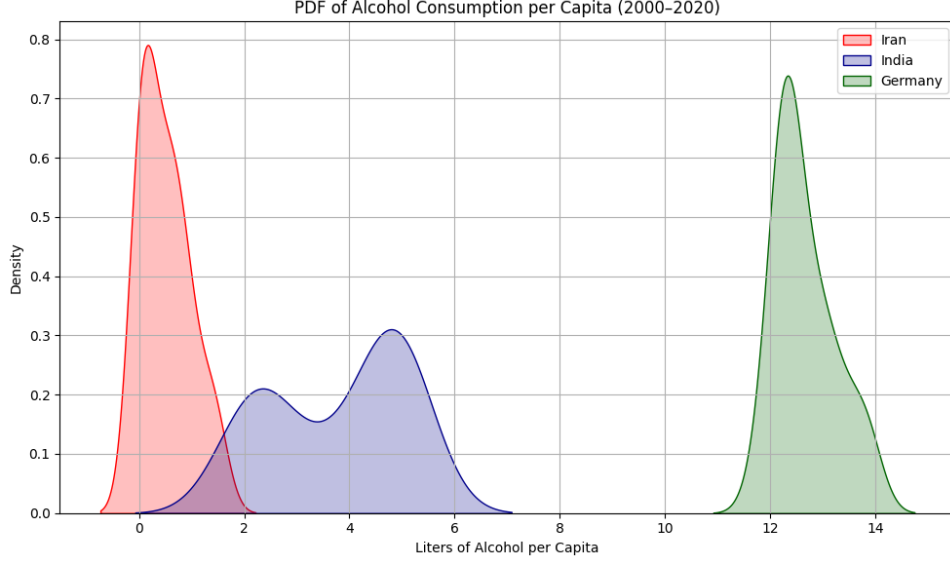


Figure 3: Empirical cumulative distribution functions (ECDFs).

4 Conclusion

This study elucidates distinct consumption patterns in three countries, with Germany exhibiting high but declining use, India rising then moderately declining, and Iran consistently low. Distributional analyses and normality tests highlight non-normality in original series. Bootstrap results affirm the Central Limit Theorem for moderate-to-large variability series (India, Germany) but caution in small, skewed samples (Iran). These findings demonstrate the utility of combined descriptive, inferential, and resampling methods in public-health data analysis.

References

1. World Bank (2025). World Development Indicators: Alcohol consumption, total (liters of pure alcohol, age 15+). Retrieved May 1, 2025, from <https://databank.worldbank.org/>
2. Wikipedia (2024). Central Limit Theorem. Retrieved April 29, 2024, from https://en.wikipedia.org/wiki/Central_limit_theorem
3. OpenAI (2025). ChatGPT response to "refer ChatGPT in Chicago format". Retrieved May 1, 2025, from <https://chat.openai.com/>

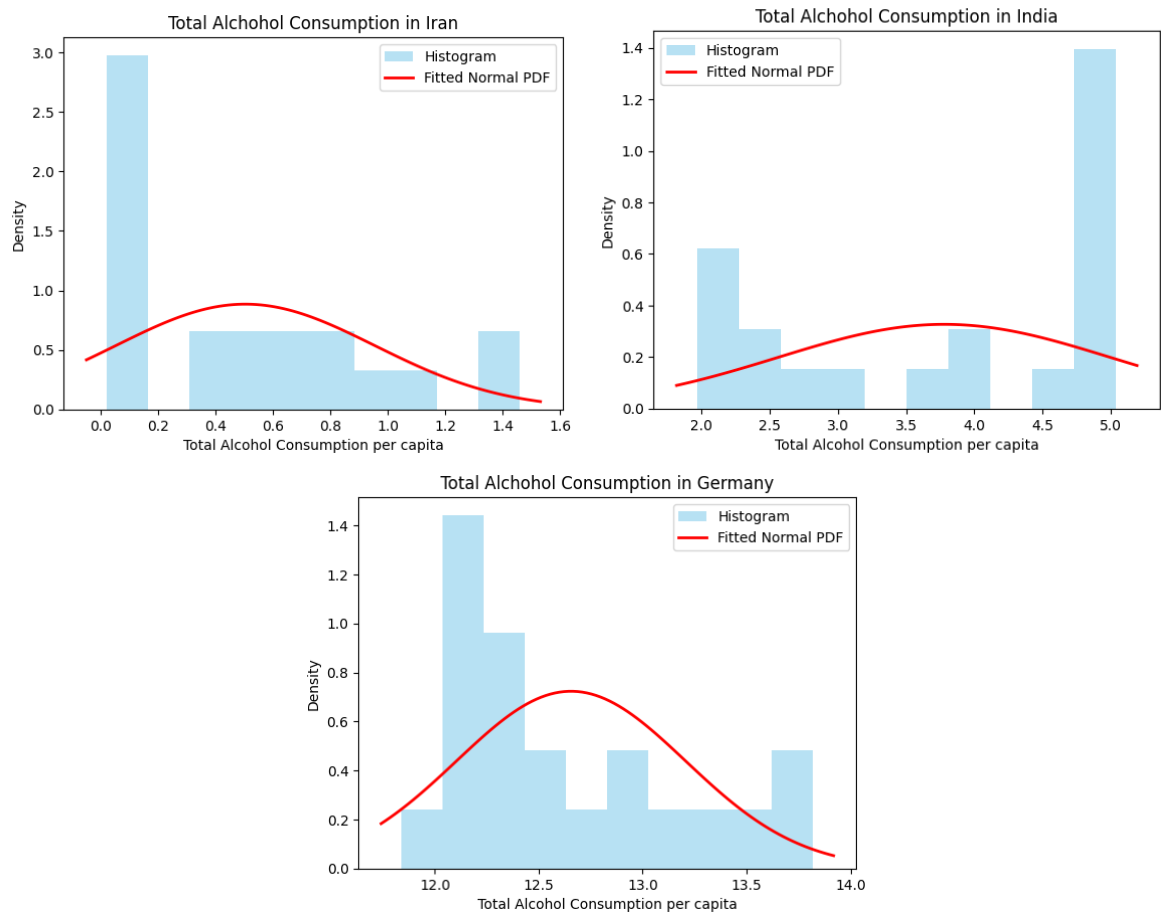


Figure 4: Q-Q plots for Iran (a), India (b), and Germany (c).

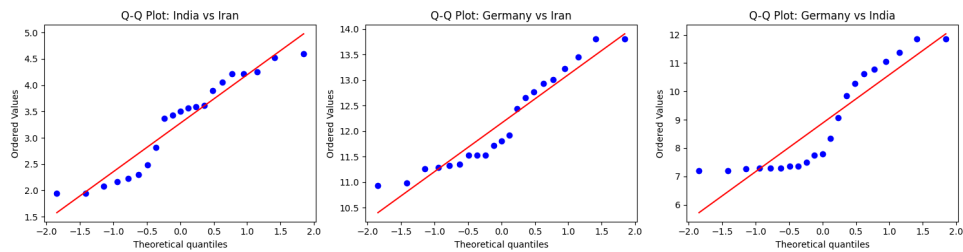


Figure 5: Shapiro-Wilk test statistics for each country.

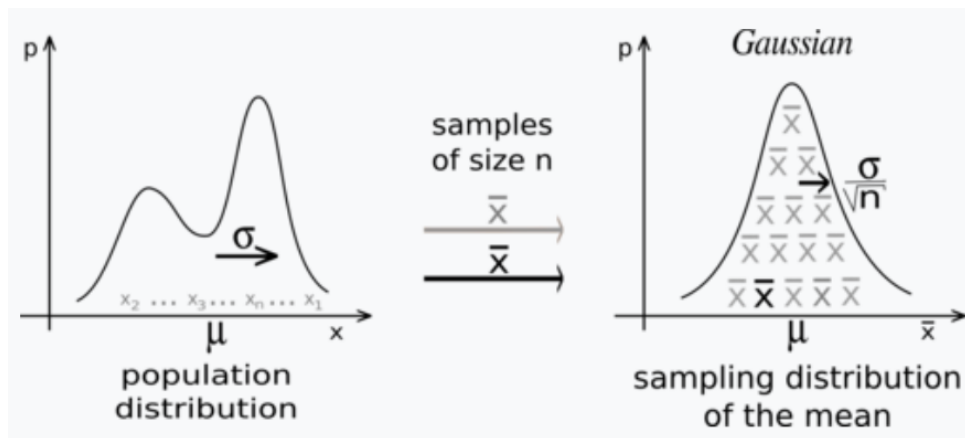


Figure 6: Mann-Kendall trend test results.

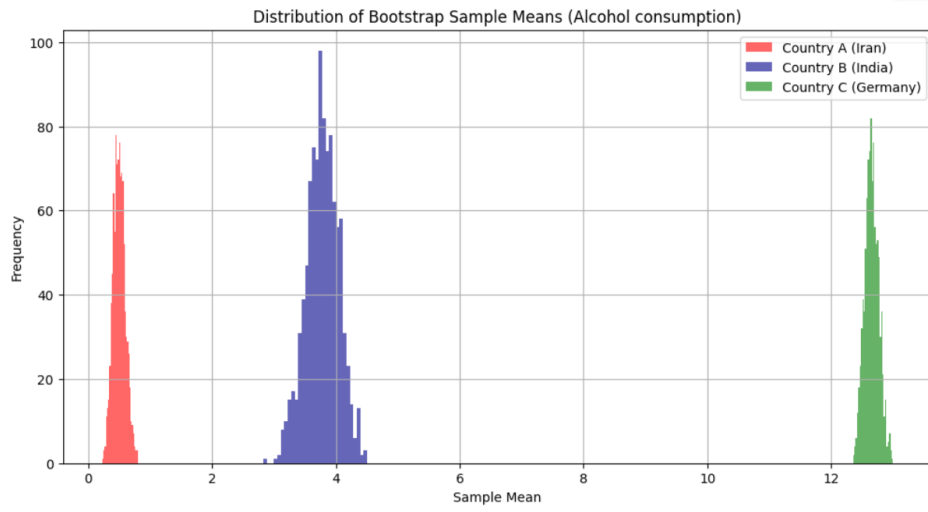


Figure 7: Bootstrap distributions of sample means.

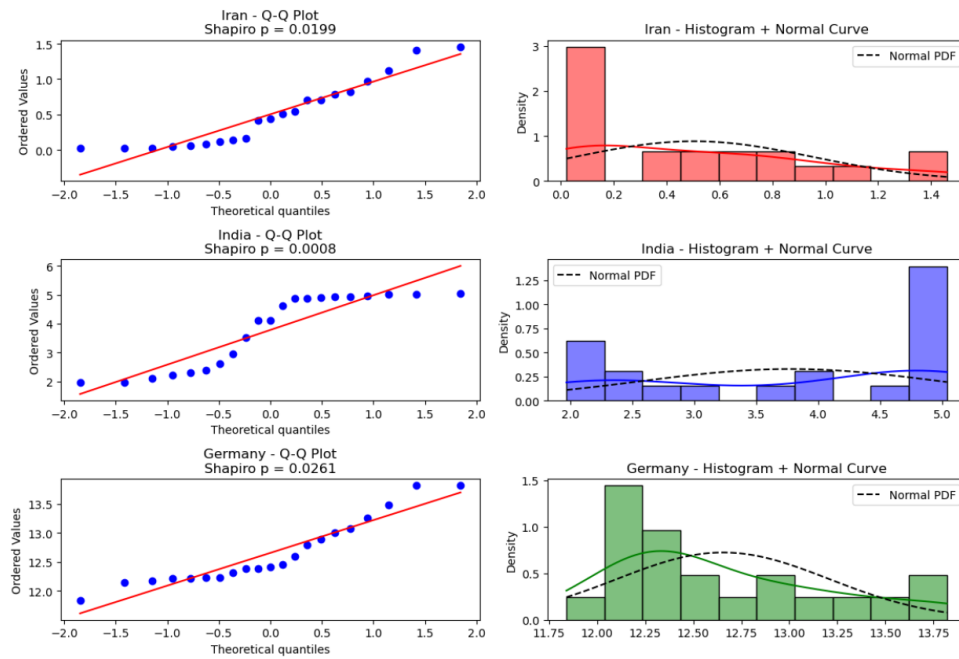


Figure 8: Normality test results on bootstrap means.