A deep Analysis of Female Labour Force Participation in India and Germany (1990–2022)

Abhishek Negi(100004670)
Aakash Vashist(100004661)
Garima (100004523)

Msc. Computer Science : Big Data and AI
SRH University

May 7, 2025

Contents

1 Introduction

The gender gap in labor market participation remains one of the most persistent global inequalities. Female Labor Force Participation Rate (FLFPR) is a key socioeconomic indicator that reflects not only economic engagement but also the effectiveness of gender-inclusive policies, access to education, healthcare, and societal attitudes toward gender roles.

A lower FLFPR often signals broader systemic issues—lack of education access, childcare support, safety, and cultural barriers—while a high FLFPR correlates with inclusive economic growth and institutional support. By comparing India and Germany, we aim to explore how contrasting national policies, cultures, and economic structures influence women's participation in the labor force.

In this analysis, we seek to:

- Statistically compare FLFPR trends between India and Germany from 1990 to 2022,
- Understand distributional patterns and validate the Central Limit Theorem,
- Test whether the differences are statistically significant,
- Reflect on what these results reveal about labor equity and development in both countries.

This study not only provides a statistical comparison but also offers insight into broader development and equity challenges faced by emerging and developed economies.

2 Data: Source and Explanation

The dataset was obtained from the World Bank's open data repository. We extracted yearly FLFPR values for India and Germany for the period 1990–2022. The data were cleaned by:

- Removing null or missing entries,
- Trimming to the target year range,
- Filtering required data columns,
- Ensuring proper labelling,
- Verifying consistent units (percentage).

After cleaning and processing the dataset it provides us with only the required fields, i.e. the FLFPR% rate for the countries - India and Germany.

3 Methods

3.1 Data Cleaning and Preparation

We retained only year and FLFPR columns. The data was filtered to 1990–2022. Missing values were inspected, and the final dataset was stored in a clean format ready for analysis.

3.2 Analysing the trend

We begin with analysing the FLFPR trend in the range 1990-2022. We are able to visualize the general trend by a simple line graph. It shows us the points where there was a major shift in FLFPR over the years.

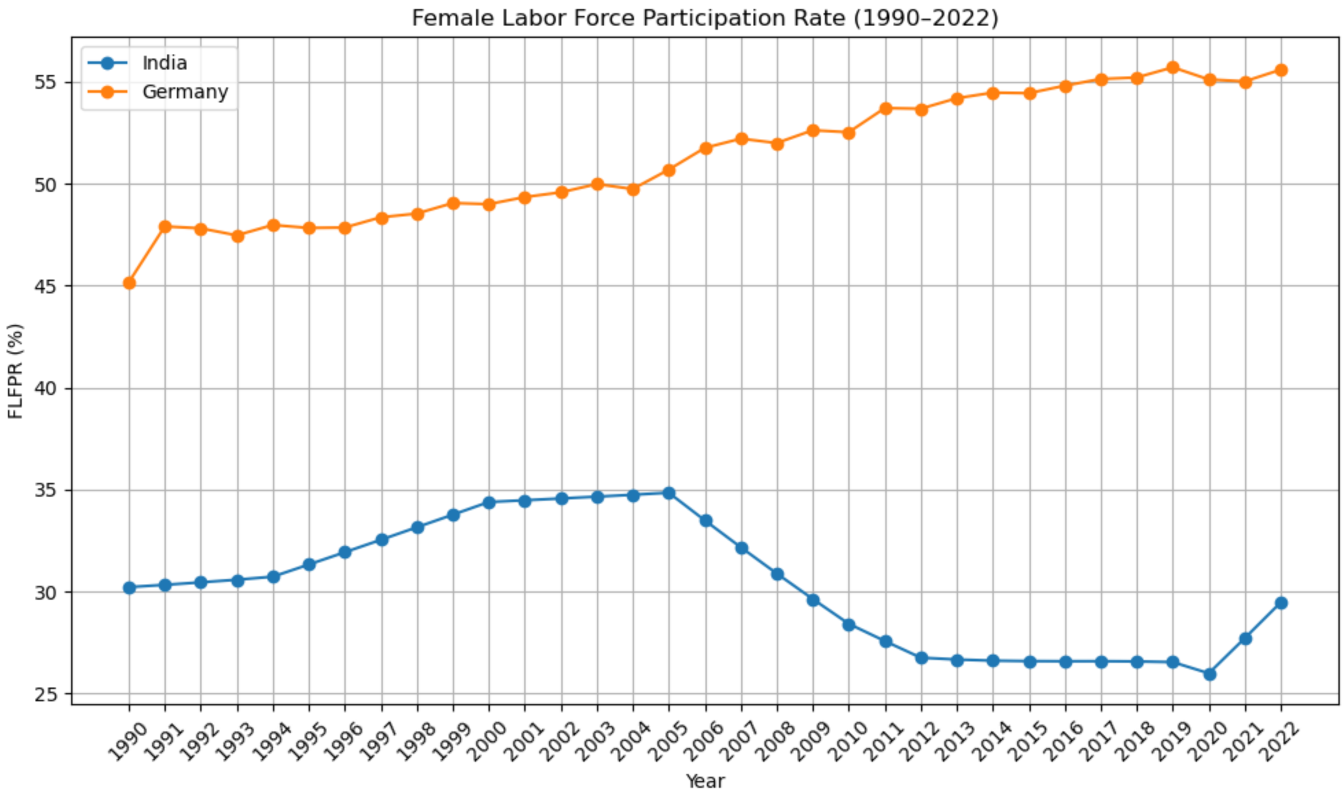

Figure 1: Trend line from 1990-2022

We can see that the Female participation in the workforce in Germany is on a constant rise throughout the years. In the case of India from the years 1990 to 2005 there was a slow but steady increment. However, India witnessed a very steep decline in the next 7 years. The female participation was steady for the next few years and since the year 2020 it's again on a rise.

Basic Statistics:

India

- Mean: 30.33
- Standard Deviation: 3.07
- Skewness: 0.05
- Kurtosis: -1.40

Germany

- Mean: 51.34
- Standard Deviation: 3.07
- Skewness: -0.08
- Kurtosis: -1.33

Even from the basic statistics we can speculate and make some primary assumptions about the reasons behind this trend. Like, since Germany developed earlier than India the females were encouraged to study and have a career. However, since India had started it's developing phase the female participation in the work force was low.

Also one of the major reasons Female workforce was low in India was due to time constraints and them not being comfortable working during odd hours. We can observe this fact while looking at the FLFPR from 2019-2020, since COVID all the companies adopted the Work from Home(WFH) policy, there was a steep rise in the female participation in the workforce.

3.3 Distribution Fitting and Normality Testing

To assess whether the FLFPR data is normally distributed:

- We visualized histograms and KDE plots.
- Applied Q-Q plot and the Shapiro-Wilk normality test:

$$H_0 : \text{Data is normally distributed.}$$

3.3.1 Histogram

A histogram is a visual representation of the distribution of quantitative data. To construct a histogram, the first step is to "bin" (or "bucket") the range of values divide the entire range of values into a series of intervals and then count how many values fall into each interval.

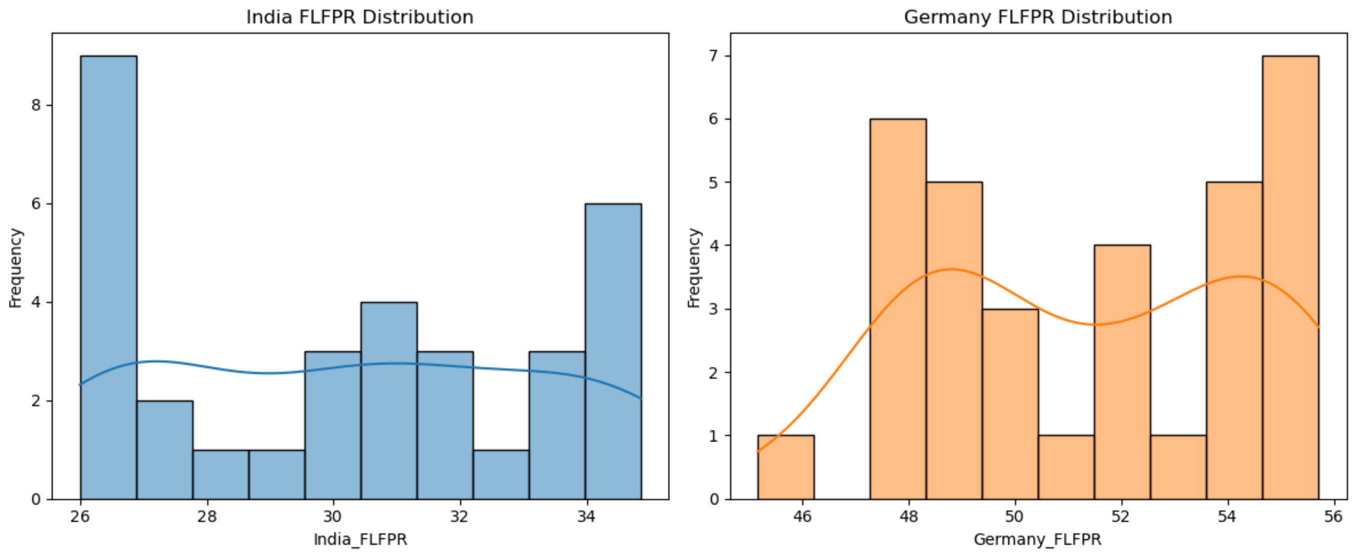

Figure 2: Histogram - Distribution of FLFPR for India and Germany

3.3.2 Kernel Density Estimate

In statistics, kernel density estimation (KDE) is the application of kernel smoothing for probability density estimation, i.e., a non-parametric method to estimate the probability density function of a random variable based on kernels as weights.

$$\hat{f}_h(x) = \frac{1}{nh\sigma} \cdot \frac{1}{\sqrt{2\pi}} \sum_{i=1}^{n} \exp\left(\frac{-(x - x_i)^2}{2h^2\sigma^2}\right)$$

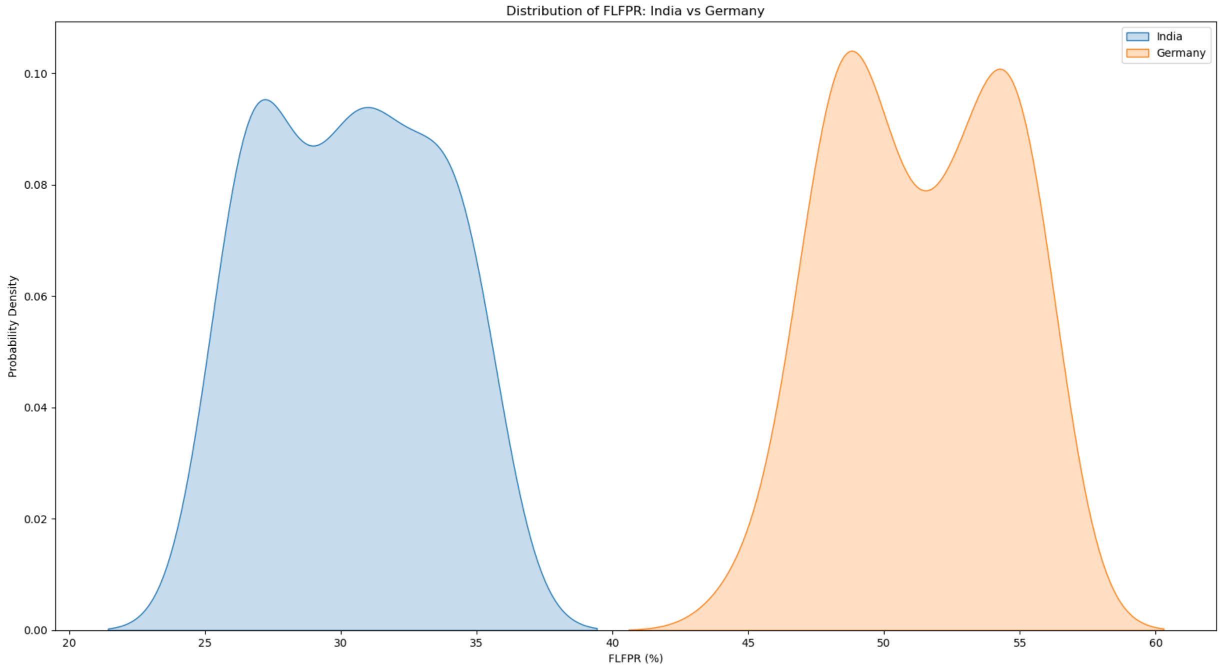

Figure 3: KDE - Distribution of FLFPR for India and Germany

3.3.3 Q-Q plot

In statistics, a Q–Q plot (quantile–quantile plot) is a probability plot, a graphical method for comparing two probability distributions by plotting their quantiles against each other. A point (x, y) on the plot corresponds to one of the quantiles of the second distribution (y-coordinate) plotted against the same quantile of the first distribution (x-coordinate). This defines a parametric curve where the parameter is the index of the quantile interval.

If the two distributions being compared are similar, the points in the Q–Q plot will approximately lie on the identity line $y = x$.

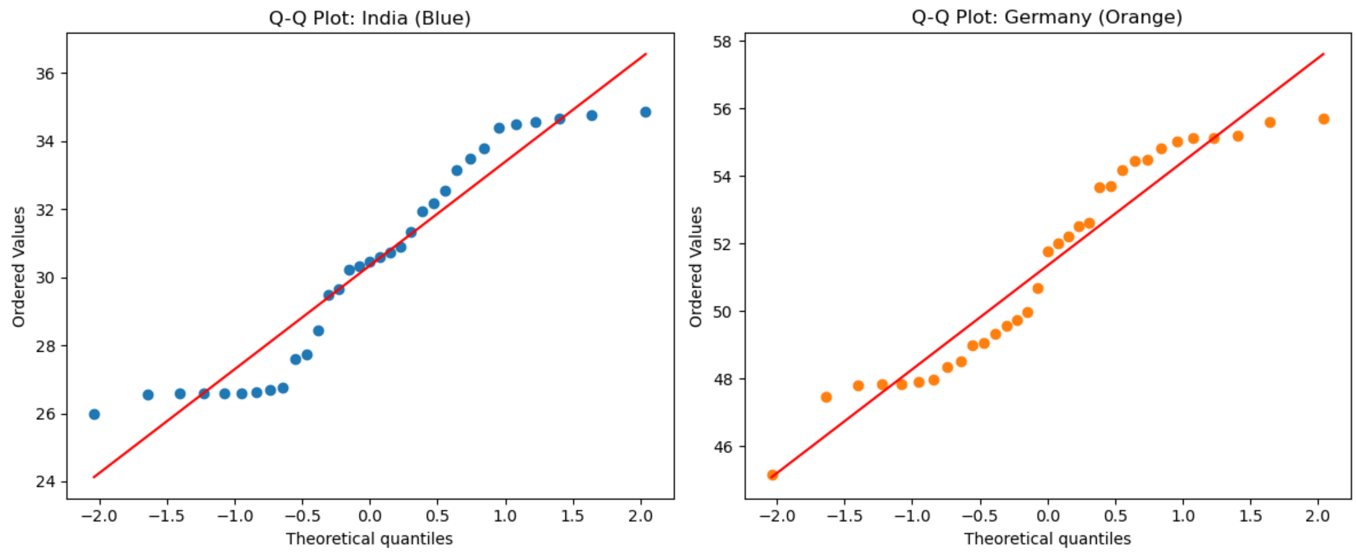

Figure 4: KDE - Distribution of FLFPR for India and Germany

As per observation, we can see that the plot doesn't lie on the $y = x$ line, hence the data is not normally distributed. To confirm our initial observation we will apply the Shapiro-Wilk test.

3.3.4 Shapiro-Wilk Normality Test

The Shapiro–Wilk test is a test of normality. The null-hypothesis of this test is that the population is normally distributed. If the p value is less than the chosen alpha level, then the null hypothesis is rejected and there is evidence that the data tested are not normally distributed.

$$W = \frac{\left(\sum_{i=1}^{n} a_i x_{(i)}\right)^2}{\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

where $x_{(i)}$ is the ith order statistic, i.e., the ith-smallest number in the sample (not to be confused with x_i).

$$\overline{x} = \frac{x_1 + \cdots + x_n}{n}$$

The coefficients a_i are given by:

$$(a_1, \ldots, a_n) = \frac{m^\mathsf{T} V^{-1}}{C}$$

where C is a vector norm:

$$C = \left\|V^{-1}m\right\| = \left(m^\mathsf{T} V^{-1} V^{-1} m\right)^{1/2}$$

and the vector m is:

$$m = (m_1, \dots, m_n)^\mathsf{T}$$

Result

Generally $\alpha = 0.05$,

- **India Shapiro Test p-value: 0.0051**
- **Germany Shapiro Test p-value: 0.0157**

As p-value $< \alpha (= 0.05)$. We reject the null hypothesis. Hence we can conclude the data for both India and Germany is not normally distributed.

3.4 Central Limit Theorem (CLT) Assessment

In probability theory, the central limit theorem (CLT) states that, under appropriate conditions, the distribution of a normalized version of the sample mean converges to a standard normal distribution. This holds even if the original variables themselves are not normally distributed.

To demonstrate the Central Limit Theorem:

- We drew random samples of sizes 5, 10, and 30.
- Repeated sampling 100 times for each size and plotted the sample means.
- As sample size increased, the sampling distribution approached normality.

Below given plots show us the probability distribution or our data after applying CLT. We can observe as the sample size (n) increases the sampling distribution approaches a normal distribution. This step is necessary as almost all the hypothesis test require the data to be normally distributed.

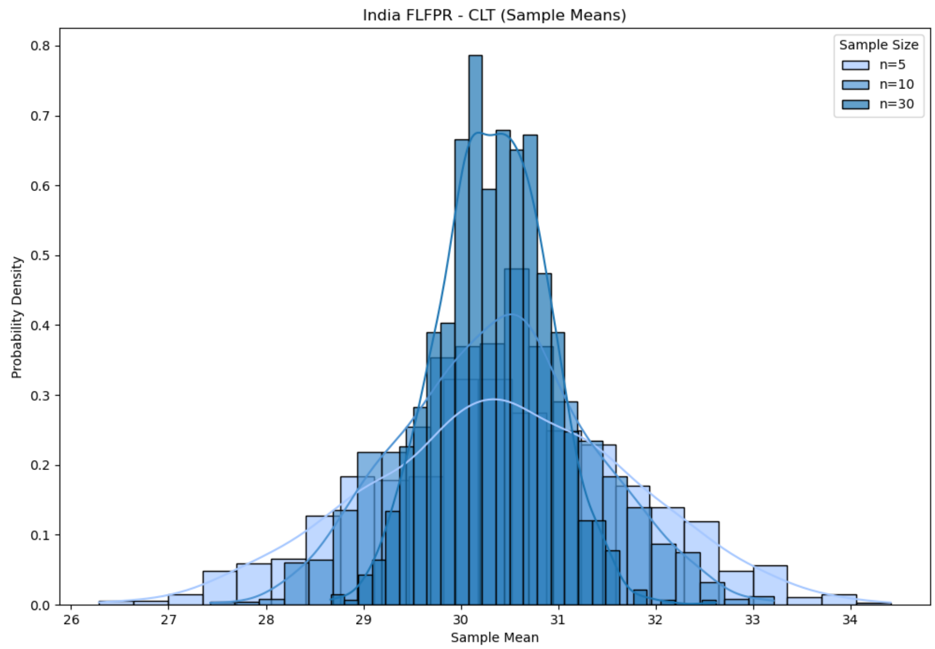

Figure 5: CLT demonstration with various sample sizes – India

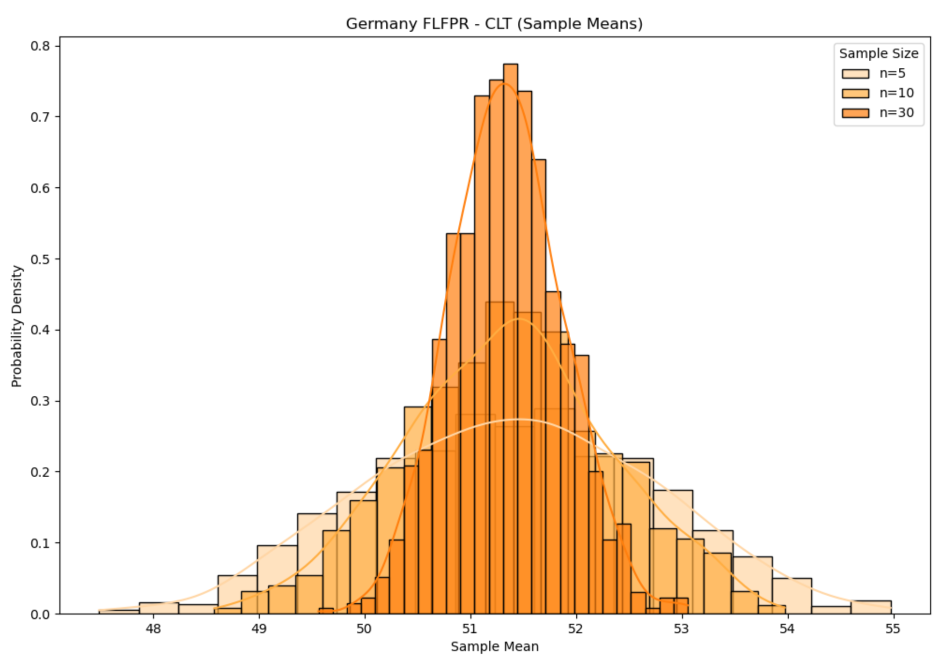

Figure 6: CLT demonstration with various sample sizes – Germany

From these graphs, we can see that CLT holds true.

We apply Shapiro-Wilk normality test again to confirm that our sampling distribution now follows normal distribution.

India FLFPR Sample Means

- India (Sample size $n = 5$) → Shapiro p-value: 0.0040 (Not Normal)
- India (Sample size $n = 10$) → Shapiro p-value: 0.7459 (Normal)
- India (Sample size $n = 30$) → Shapiro p-value: 0.1827 (Normal)

Germany FLFPR Sample Means

- Germany (Sample size $n = 5$) → Shapiro p-value: 0.4943 (Not Normal)
- Germany (Sample size $n = 10$) → Shapiro p-value: 0.7947 (Normal)
- Germany (Sample size $n = 30$) → Shapiro p-value: 0.0936 (Normal)

Hence we can safely conclude that our sampling distribution is now normally distributed.

3.5 Hypothesis Testing

Now that we have a normal distribution, we move to test our hypothesis. To evaluate whether there is a statistically significant difference in Female Labor Force Participation Rate (FLFPR) between India and Germany from 1990 to 2022, we selected appropriate hypothesis testing methods based on data properties.

Levene's Test

In statistics, Levene's test is an inferential statistic used to assess the equality of variances for a variable calculated for two or more groups. This test is used because some common statistical procedures assume that variances of the populations from which different samples are drawn are equal. Levene's test assesses this assumption. It tests the null hypothesis that the population variances are equal (called homogeneity of variance or homoscedasticity). If the resulting p-value of Levene's test is less than some significance level (typically 0.05), the obtained differences in sample variances are unlikely to have occurred based on random sampling from a population with equal variances. Thus, the null hypothesis of equal variances is rejected and it is concluded that there is a difference between the variances in the population.

$$W = \frac{(N-k)}{(k-1)} \cdot \frac{\sum_{i=1}^{k} N_i (Z_{i\cdot} - Z_{\cdot\cdot})^2}{\sum_{i=1}^{k} \sum_{j=1}^{N_i} (Z_{ij} - Z_{i\cdot})^2}$$

where:

- k is the number of different groups to which the sampled cases belong,
- N_i is the number of cases in the ith group,
- N is the total number of cases in all groups,
- Y_{ij} is the value of the measured variable for the jth case from the ith group,
- Z_{ij} is defined as:

$$Z_{ij} = \begin{cases} |Y_{ij} - \bar{Y}_{i\cdot}|, & \text{if } \bar{Y}_{i\cdot} \text{ is the mean of the } i\text{th group} \\ |Y_{ij} - \tilde{Y}_{i\cdot}|, & \text{if } \tilde{Y}_{i\cdot} \text{ is the median of the } i\text{th group} \end{cases}$$

According to our data, when calculating p-value using Levene's test we got

- **p-value** $= 0.73(< \alpha = 0.05)$

Hence, we can conclude the data has equal variance.

Despite Levene's test showing equal variances (p = 0.73), the non-normality and moderate sample size (n = 33) necessitate a more robust approach.

We therefore adopted the following hypothesis tests:

- **Welch's t-test:**

  In statistics, Welch's t-test, or unequal variances t-test, is a two-sample location test which is used to test the (null) hypothesis that two populations have equal means. It is named for its creator, Bernard Lewis Welch, and is an adaptation of Student's t-test, and is more reliable when the two samples have unequal variances and possibly unequal sample sizes. These tests are often referred to as "unpaired" or "independent samples" t-tests, as they are typically applied when the statistical units underlying the two samples being compared are non-overlapping.

  Welch's t-test defines the statistic t as:

$$t = \frac{\Delta\overline{X}}{s_{\Delta\overline{X}}} = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_{\overline{X}_1}^2 + s_{\overline{X}_2}^2}}$$

  where the standard error of the sample mean is:

$$s_{\overline{X}_i} = \frac{s_i}{\sqrt{N_i}}$$

  Here,

  - $\overline{X}_i$ is the mean of the i-th sample,
  - s_i is the corrected sample standard deviation,

- N_i is the sample size.

Unlike Student's *t*-test, the denominator is not based on a pooled variance estimate.

The degrees of freedom ν are approximated using the Welch–Satterthwaite equation:

$$\nu \approx \frac{\left(\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}\right)^2}{\frac{s_1^4}{N_1^2 \nu_1} + \frac{s_2^4}{N_2^2 \nu_2}}$$

This can be simplified when $N_1 = N_2$:

$$\nu \approx \frac{s_{\Delta \overline{X}}^4}{\nu_1^{-1} s_{\overline{X}_1}^4 + \nu_2^{-1} s_{\overline{X}_2}^4}$$

where $\nu_i = N_i - 1$ is the degrees of freedom associated with the i-th sample.

It was selected as the primary test because it does not assume equal variances and is more reliable under non-normal data conditions. It directly compares the means of the two groups and is suitable given our moderate sample size and violation of normality.

- **Mann-Whitney U test:**

The Mann–Whitney U test (also called the Mann–Whitney–Wilcoxon (MWW/MWU), Wilcoxon rank-sum test, or Wilcoxon–Mann–Whitney test) is a nonparametric statistical test of the null hypothesis that randomly selected values X and Y from two populations have the same distribution.

A very general formulation of the hypothesis test is as follows:

- The variable under investigation is at least **ordinal**.
- The null hypothesis is:

$$H_0 : P(X > Y) = P(X < Y)$$

or equivalently,

$$H_0 : P(X < Y) + 0.5\, P(X = Y) = 0.5$$

- The alternative hypothesis is:

$$H_1 : P(X > Y) \neq P(X < Y)$$

or equivalently,

$$H_1 : P(X < Y) + 0.5\, P(X = Y) \neq 0.5$$

- If the null hypothesis is true, then the two distributions are equal:

$$F_X = F_Y$$

Under these assumptions, the test is accurate and consistent.

The Mann–Whitney U statistic is given by:

$$U = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

or equivalently,

$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

where:

- n_1 and n_2 are the sample sizes of the two groups,
- R_1 is the sum of the ranks for the first group,
- R_2 is the sum of the ranks for the second group.

The smaller of U_1 and U_2 is typically used to determine significance.

It is used as a non-parametric alternative to validate our results. It tests for differences in the central tendency (typically medians or rank distributions) without assuming normality. While less powerful for mean comparisons, it offers a robust secondary check.

These tests provide a comprehensive and statistically sound foundation for comparing FLFPR between the two countries.

We tested whether there is a significant difference in mean FLFPR between India and Germany. The summary of results is given below:

For all the hypothesis test we took the same null hypothesis.

H_0 = **India and Germany have equal mean FLFPR**

- **Levene's Test** (equal variances): p = 0.7300
- **Welch's t-test** (unequal variances): p = 0.0000
- **Mann-Whitney U test** (non-parametric): p = 0.0000

All tests indicate a statistically significant difference in mean FLFPR.

Note: p-value of 0.0000 just means that the p-value is so small that it gets round off to 0. This indicate that there is a big difference between the means.

Below mentioned graph compares the mean FLFPR of both the countries.

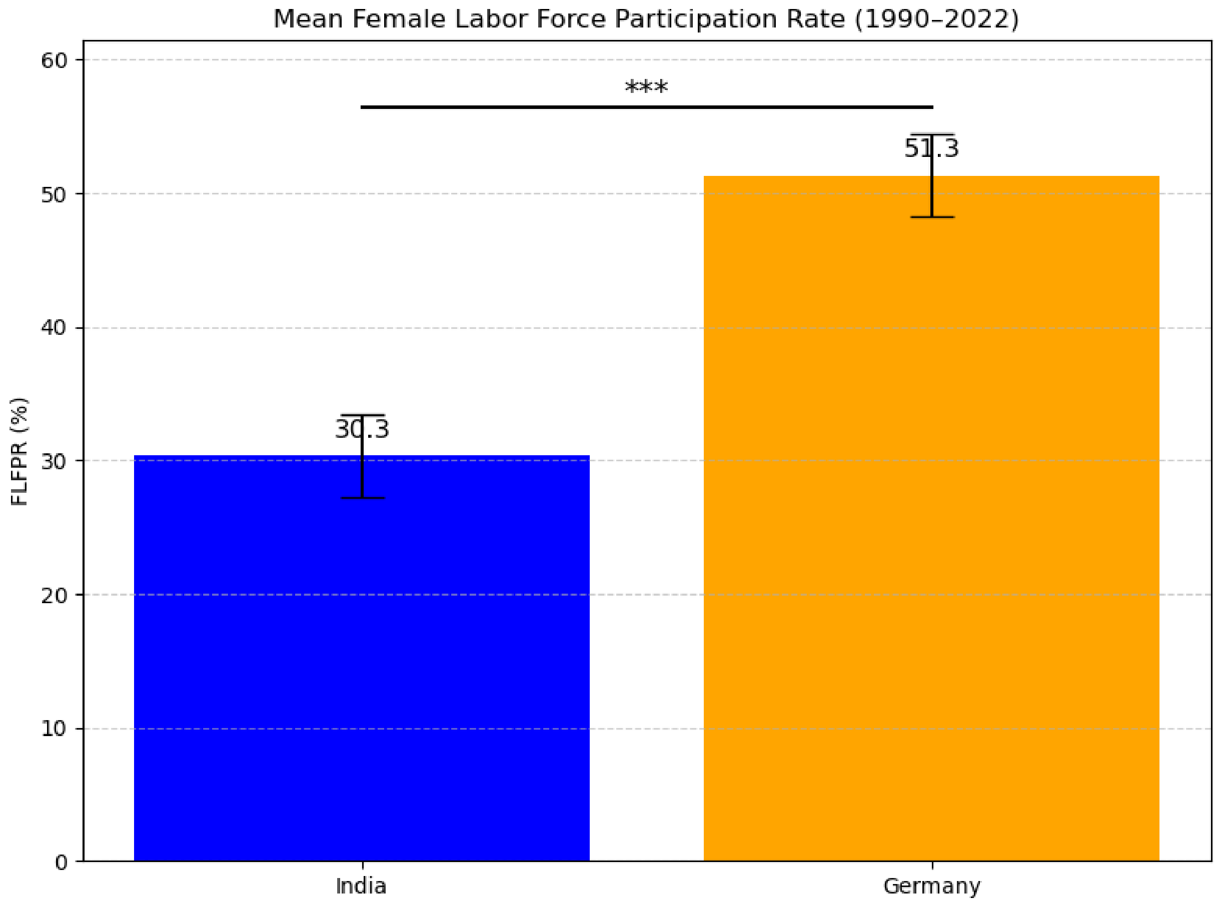

Figure 7: Mean FLFPR with hypothesis test indication

4 Discussion

Germany consistently outperforms India in terms of female labor force participation. Factors contributing to this difference could include:

- Social and economic policies supporting women in Germany,
- Better access to childcare and education,
- Stronger enforcement of labor laws and anti-discrimination policies.

India, on the other hand, faces cultural, infrastructural, and economic barriers that limit female participation.

5 Conclusion

Our analysis confirms that Germany has a significantly higher FLFPR than India across 1990–2022. The findings emphasize the importance of structural and policy interventions to improve labor participation among Indian women.

6 References

- World Bank Open Data: https://data.worldbank.org/
- Shapiro, S. S., and Wilk, M. B. (1965). An Analysis of Variance Test for Normality.
- Central Limit Theorem. (Various sources in theoretical statistics).
- Wikipedia contributors. (n.d.). Female labor force participation. Wikipedia. Retrieved May 7, 2025, from https://en.wikipedia.org/wiki/Female_labor_force_participation

A Appendix: Figures

- All the plots and graphs have been generated using python. You can find the entire code in the final project file at https://github.com/53-97/Tools_-_Methods_Data_Analysis.git