

SDG –1: NO POVERTY

AS1117: PROBABILITY AND STATISTICS

FACULTY GUIDE

Dr. Umesh Gupta

Dr. Rabia Kamra

PREPARED BY: Group-1(Section - B)

Kalpana Baghel (2023Btech039)

Sarthak Pandit (2023Btech072)

Siddharth Ranka (2023Btech082)



Institute of Engineering and Technology (IET)

JK LAKSHMIPAT UNIVERSITY, JAIPUR

TABLE OF CONTENTS

S. No.	Content	Page No.
1.	List of Figures	3
2.	List of Tables	4
3.	Abstract	5
4.	Introduction	6
5.	Objectives	11
6.	Methodology	12
7.	Objective 1	14
8.	Objective 2	29
9.	Objective 3	34
10.	Conclusion	38
11.	Appendix (All)	39

LIST OF FIGURES

Figure	Description
Figure 1	Sustainable Development Goals
Figure 2	No Poverty Logo
Figure 3	No poverty Poster
Figure 4	Functional sub centres across different states
Figure 5	correlation between urban and rural health expenditure
Figure 6	Medical and public expenditure on health by rural health service
Figure 7	Total Rural HH vs Tap water Access
Figure 8	Access of tap water in future
Figure 9	Tap water Connections by state
Figure 10	The distribution of House owned by men and women (alone or jointly) in India
Figure 11	The distribution of Land owned by men and women (alone or jointly) in India

LIST OF TABLES

Table	Description
Table 1	Accessible Health Centres
Table 2	Health Expenditure
Table 3	Access of Tap Water in Rural India
Table 4	Percentage of men and women who own a House (alone or jointly) in India
Table 5	Percentage of men and women who own a Land (alone or jointly) in India

ABSTRACT

This report's primary goal is to examine critical socio-economic problems such as lack of access to health services, lack of access to natural resources, and the gap in land ownership by gender. To address these issues three key objectives are set in this study. First, it evaluates the three studies which assesses the implications of unavailability of health services in India leads to poverty through the data for availability of functioning sub-centers, relationship between state spending on health in rural sectors and state trends in rural health expenditure. Second, it assesses the impact of limited availability of natural resources on the rural poor to this end, it explores the relationship between rural population density and access to tap drinking water, uses predictive modelling to project trends in access to tap water in rural areas over the coming decades, and compares the averages of rural households being destitute in the different states. Third, this research examines the gender gap in ownership of property more generally and how that affects women's economic stability, through analysis of the percent of land and houses owned and hypothesis testing of women's equality in ownership of land. Statistical methods are used in these are one-way ANOVA, Z-test, single proportion, correlation, regression, and descriptive statistics to achieve these goals. This meta-analysis provides insight into many factors related to these socio-economic indicators, highlighting the importance of targeted interventions aimed at promoting equity and alleviating poverty.

INTRODUCTION



Fig 1: Sustainable Development Goals

Sustainable Development refers to the development of meeting the basic requirements of people in present to reverse the stock of future generations for satisfying their needs.

The SDGs were planned in 2015 apiece United National General Assembly (UNGA) as one the Post-2015 Development Agenda, that wanted to establish a future worldwide growth foundation to gain the Millennium Development Goals, that done that period. They were in an official manner connected and selected in a UNGA Resolution named the 2030 Agenda, popular colloquially as Agenda 2030.

There are 17 Sustainable Development Goals, and their min motive are: -

- **No Poverty:** Poverty is defined as scarcity of food, clothes, and houses to citizens of the country. So, people were unable to fulfill their basic requirements of life.
- **Zero Hunger:** When a situation of starvation occurs that leads to deficiency of food it could cause health issues to the people.
- **Good Health and Well Being:** There should be proper supply of nutritional food consist of carbohydrates, vitamins, minerals to people with a pollution free environment to decrease deaths rates & health issues of people.
- **Quality Education:** To educate all citizens there should be proper setup of school infrastructure, teachers who deliver lectures on this. Citizens became aware to take education as it is essential part.
- **Clean water and Sanitation:** Water is a basic need of all living beings on earth & it should be available to them. For hygiene there should be proper sewage treatment.

- **Affordable and Clean Energy:** Energy should be consumed by all but in a limited amount as it should be reserved for future generations.
- **Decent Work and Economic Growth:** There should be employment available for all with decent task during working hours to overcome with an obstacle employment
- **Industry, Innovation, and Infrastructure:** More industries should be established to generate more employment. Labors could get wages according to their efforts.
- **Reduced Inequalities:** There should not be any discrimination in policies institutions and society regarding gender caste color and birthplace.
- **Sustainable Cities and communities:** Settlement for humans should be safe with more equal opportunities in cities.
- **Sustainable consumption and Production:** Demand should be met with supply so that there was a proper balance maintained in country regarding production and consumption.
- **Climate Action:** Immediate migration program should be made in advance to control the situations of natural calamity.
- **Life below Water:** Natural water bodies such as ocean and seas water consumed in a limited amount to meet future generation requirement.
- **Life on Land:** Deforestation was the main cause of loss of land due to population growth it could create imbalance.
- **Peace, justice, and strong institutions:** The states should be available to all in accordance with law so punishment should be given so a criminal proper institute such as court should be set up at the specific places for ease of people.

Partnership for the Goals: There should be worldwide relationship among countries for sustainable development

Goal 1: No Poverty



Fig 2: No Poverty Logo

Introduction:

SDG 1: Zero Poverty seeks to eliminate extreme poverty for those who are still living below the poverty line and for the entire world population by the year 2030, and to at least halve the proportion of people living in poverty in all its forms across the globe. The majority of the population in the world lives below the poverty line and this has grossly affected development and caused billions access basic needs and opportunities. For this purpose, a multi-faceted approach is required which includes, among others, strategies aimed at reducing income disparities, increasing resource access, and promoting social inclusion. Equal rights to economic resources and the control of land as well as ownership to property, inheritance, financial opportunity and suitable technology is vital in empowering the poor and the vulnerable population.

Why is it important to work towards the elimination of poverty?

There are several important reasons for the poverty eradication initiative:

- This is a basic and fundamental right: Food, clothing, and education are some necessities and basic needs of every person. Poverty infringes on this provision.
- It enhances economic development: Provision of financial resources and providing opportunities to people would improve productivity at the country level and also reduce income disparities.
- It enhances social stability: When poverty is solved, societies are more strengthened and everyone in the society is able to play their role in the society.
- It makes faster achievement of other SDGs: Poverty is linked to most challenges that relate to health, education, gender and women empowerment, and sustainable cities. If poverty is addressed, then these areas can be addressed as well.

In what ways does poverty influence individuals and communities?

- Absence of basic services and squalor conditions: Most vulnerable communities do not have educational facilities, behavioural healthcare, clean drinking water and sanitation facilities and thus would be hard pressed to thrive.

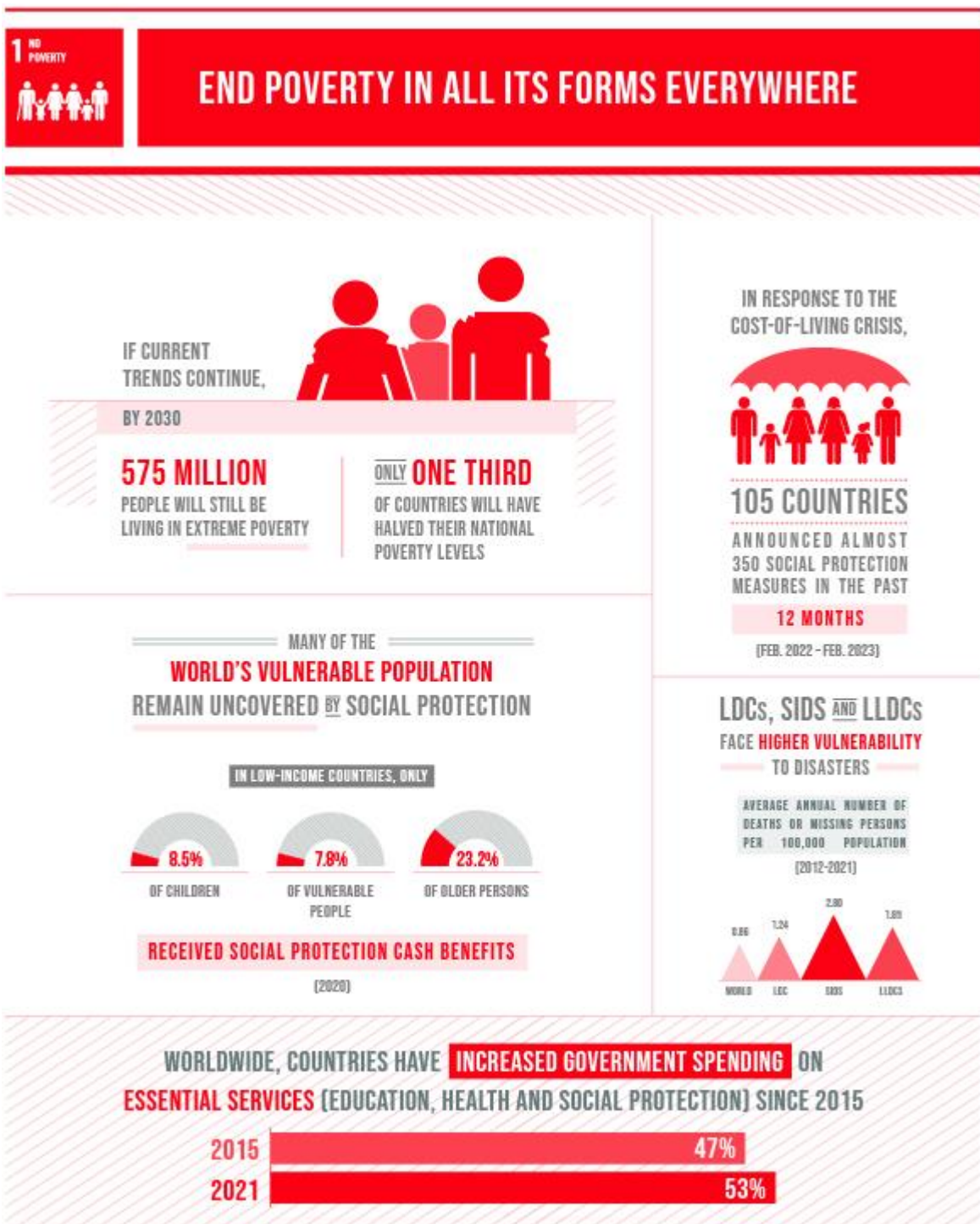
Economic vulnerability: Poverty is a major barrier to availability and access to financial inclusion and even employment opportunities meaning very little or no resources to fetch for risk mitigating measures.

- Intergenerational impact: Children who were brought up in situations of poverty stand reservations of having many opportunities that leave them in poverty situations as a form of cycle.
- Exclusion from decision-making: The poor have insufficient representation in the governance structures and decision-making processes which greatly restricts their possibilities of being able to push for their interests.

Why should people care about basic income and poverty eradication in particular?

- Poverty is an issue of fairness: Providing resources to everyone helps to improve the fairness of the society so that everyone is treated equally.
- It contributes to global security: The reduction of poverty will help to bring about the reduction of social and economic imbalances and resource wars.
- It benefits entire economies: The solutions to the poverty problems empower the masses enabling poverty alleviation to enhance the growth of a nation or even the world as a whole.

Let us effectively cooperate in ensuring that by 2030, all people especially the poor and the oppressed will be able to enjoy the resources and opportunities which are necessary for their proper welfare, dignity and development



THE SUSTAINABLE DEVELOPMENT GOALS REPORT 2023: SPECIAL EDITION- [UNSTATS.UN.ORG/SDGS/REPORT/2023/](https://unstats.un.org/sdgs/report/2023/)

Fig 3: No Poverty Poster

Objective 1: Analysis on how the inaccessibility of health services affect the poverty in India.

- Is there a significant difference in the number of Functional Sub Centres across different states?
- Is there a significant correlation between the medical and public expenditure on health by urban health services (UHS) and rural health services (RHS) in different states?
- Is there a significant trend in the medical and public expenditure on health services in rural areas over the years?

Data Source: <https://ndap.niti.gov.in/dataset/6066>

<https://ndap.niti.gov.in/dataset/6088>

Objective 2: Analysis on How the Inaccessibility of Natural Resources Affects Poverty in Rural and Local Communities.

- Is there a correlation between rural population density and the percentage of rural households with tap water access across states in 2024.
- Can past year's tap water access percentages predict the percentage of rural households with access in 2030?
- Is there a significant difference in the average percentage of rural households with tap water access across different states?

Data Source: <https://ejalshakti.gov.in/jjmreport/JJMIndia.aspx>

Objective 3: Analysis on How the Gender Gap in Property Ownership Affects Women's Economic Stability.

- To access the total percentage of House and Land owned by women and men in each State and Union Territories of India according to data 2019-2021.
- To test the hypothesis whether women own less land than men.
- To test the hypothesis whether there is gender equality in land ownership

Data Source: <https://www.nfhsiips.in/nfhsuser/index.php>

Methodology

Following Methods are used in the report:

Regression Analysis: In linear regression analysis linear regression analysis is done, so as to be able to make a prediction about the value of the dependent variable, having the value of an independent variable. The key objective is to arrive at a linearly related equation, which gives the best solution to the discrepancies between the predicted values and the actual ones.

$$Y = a + bX$$

Here a = y-intercept & b = Slope of Line

$$a = \frac{[(\sum y)(\sum x^2) - (\sum x)(\sum xy)]}{[n(\sum x^2) - (\sum x)^2]}$$
$$b = \frac{[n(\sum xy) - (\sum x)(\sum y)]}{[n(\sum x^2) - (\sum x)^2]}$$

Correlation: Correlation analysis is performed in order to measure the strength as well as the direction of two related variables. It assists in answering the question of whether or not the pairs of variables have any association and if capable, how strong the association is.

$$r = \frac{N(\sum XY) - (\sum X)(\sum Y)}{\sqrt{[N(\sum X^2) - (\sum X)^2][N(\sum Y^2) - (\sum Y)^2]}}$$

ANOVA (Analysis of Variance): ANOVA is applied for the purpose of assessing the degree of variances that exists among randomly selected means of two or more groups. It assists in answering the question as to whether there are any significant statistical differences concerning the means and, if there are, the results could suggest the influence of the independent variables towards the dependent variable.

Descriptive Statistics: Descriptive statistics provide a summary of the features of a sample data set, and describe it in terms of the centre, the spread and the shape of the data. This includes mean, median, mode and standard deviation among others.

Single Proportion Test: The single proportion test helps to identify if the proportion of a particular type feature belongs to a sample as statistically different from a known or hypothesized proportion of a population.

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Z-TEST: The Z-test is used for testing hypotheses regarding population parameters especially in large sample size cases where the population variance is known. This is done to test if the sample mean is actually different from the population mean.

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

OBJECTIVE 1: Analysis on how the inaccessibility of health services affect the poverty in India.

Table 1 Accessible Health Centres

Country	State	District	Functional Sub Centres (UOM:Num)	Functional Primary Health Centres (UOM:Num)	Functionar	Functionar	Functionar	Functionar	Functional District Hospitals (UOM:Num)
India	Goa	South Goa	105	0	2	0	13	2	1
India	Haryana	Sirsa	139	8	8	12	18	1	1
India	Himachal Pradesh	Chamba	177	45	5	0	2	6	1
India	Bihar	Gopalganj	460	53	8	9	7	1	1
India	Assam	Nagaon	200	38	12	34	13,404	0	0.6702
India	Madhya Pradesh	Indore	115	33	6	0	7	3	1
India	Madhya Pradesh	Bhind	210	14	7	0	13	1	1
India	Andhra Pradesh	East Godavari	774	12	26	66	146	3	1
India	Karnataka	Bengaluru Urban	279	96	5	0	0	3	0
India	Haryana	Pahwal	81	6	7	14	11	0	1
India	Gujarat	Porbandar	79	7	4	6	8	0	1
India	Bihar	Vaishali	178	23	2	3	13	2	1
India	Madhya Pradesh	Singrauli	206	13	7	0	2	0	1
India	Chhattisgarh	Bemetara	113	17	4	15	3	1	1
India	Maharashtra	Palghar	264	60	10	50	7	3	0
India	Karnataka	Bidar	275	58	8	109	25	4	0
India	Gujarat	Kachchh	418	32	17	24	39	2	0
India	Chhattisgarh	Rajnandgaon	276	36	10	60	14	1	1
India	Jammu And Kashmir	Baramulla	212	32	6	5	24	0	1
India	Haryana	Fatehabad	131	22	4	0	0	1	1
India	Assam	West Karbi Anglong	46	8	3	11	5,6362	0.3757	0.3757
India	Jharkhand	Goddia	177	9	7	4	3	0	1
India	Gujarat	Morbi	162	14	6	16	20	1	1
India	Bihar	Bhagalpur	354	49	4	8	29	2	1
India	Bihar	Banka	237	30	4	2	12	0	1
India	Karnataka	Haveri	311	69	5	0	0	6	1
India	Chhattisgarh	Jashpur	243	32	8	15	3	2	1
India	Madhya Pradesh	Dhar	475	49	15	0	0	1	1
India	Manipur	Bishnupur	35	8	2	13	1	0	1
India	Jammu And Kashmir	Bandipora	79	13	3	0	0	0	1
India	Chhattisgarh	Sukma	76	9	3	10	3	0	1
India	Madhya Pradesh	Hoshangabad	157	16	6	0	0	2	1
India	Ladakh	Kargil	138	16	4	0	0	0	1
India	Arunachal Pradesh	Anjaw	9	7	5	0	0	0	1
India	Assam	Karimganj	176	20	7	35,288	7,6298	0	0.9537
India	Maharashtra	Amravati	291	20	11	48	48	4	1
India	Maharashtra	Mumbai	0	120	0	0	0	0	0
India	Arunachal Pradesh	Leparada	9	4	2	0	0	0	0.435
India	Maharashtra	Sangli	320	70	14	0	0	2	0
India	Karnataka	Kolar	266	69	2	0	0	4	1
India	Arunachal Pradesh	Upper Subansiri	26	12	4	0	0	0	0.8251
India	Chhattisgarh	Durg	196	24	8	25	8	1	1
India	Madhya Pradesh	Barwani	330	21	8	0	9	2	1
India	Bihar	Lakhisarai	156	8	1	0	9	0	1
India	Karnataka	Chamarajanagara	256	64	3	0	0	3	0
India	Himachal Pradesh	Mandi	325	85	13	0	5	13	1
India	Haryana	Karnal	137	6	6	10	18	1	1
India	Kerala	Alappuzha	366	49	16	0	14	6	2
India	Karnataka	Vijayapura	310	68	9	0	0	4	1
India	Maharashtra	Beed	296	56	10	0	0	3	1
India	Bihar	Sheikhpura	412	58	1	2	3	0	1
India	Gujarat	Kheda	276	25	15	40	39	1	1
India	Chhattisgarh	Raipur	245	40	11	35	9	3	1
India	Chhattisgarh	Raigarh	335	50	9	15	2	3	1
India	Bihar	Jamui	179	28	6	6	8	0	1
India	Maharashtra	Raigad	288	59	9	0	0	5	1
India	Chhattisgarh	Gaurella Pendra Marwahi	44	10	1	2,5721	0.5144	0.1715	0.1715
India	Madhya Pradesh	East Nimar	181	24	7	0	7	1	1
India	Karnataka	Ramanagara	249	63	5	0	0	3	1
India	Bihar	Patna	171	33	4	3	38	4	0
India	Arunachal Pradesh	Dibang Valley	10	1	1	0	0	0	1
India	Andhra Pradesh	Y.S.R.	420	6	12	28	89	1	1
India	Bihar	Aurangabad	254	55	6	0	18	1	1

[illegible]

[illegible]

[illegible]

18

India	Tamil Nadu	Tirunelveli	221.5119759	36.9186264	10.166	20.867	8.5608	8.5608	0.5351				
India	Odisha	Kalahandi	234	20	18	8	26	1	1				
India	Sikkim	South District	39	6	1	14	3	0	1				
India	Uttar Pradesh	Hathras	194	10	7	0	17	0	2				
India	Manipur	Kamjong	18	2	1	0	0	0	0				
India	Punjab	Shahid Bhagat Singh Nag	92	4	3	3	13	1	1				
India	Tamil Nadu	Tiruvannamalai	395	59	18	35	22	9	1				
India	Rajasthan	Bundi	207	27	12	0	7	0	1				
India	Odisha	Kandhamal	170	8	14	2	32	1	1				
India	Rajasthan	Tonk	230	58	9	0	9	0	1				
India	Tamil Nadu	Dindigul	301.0144094	25.00054719	14.001	37.001	34.002	12	1				
India	Odisha	Jharsuguda	62	4	6	4	15	0	1				
India	Uttar Pradesh	Shravasti	32	4	6	33	8	0	1				
India	Uttar Pradesh	Kasganj	169	24	7	1	7	0	1				
India	Meghalaya	Vest Jaintia Hills	46	12	3	0	0	0	1				
India	Odisha	Nuapada	90	5	5	5	12	1	1				
India	Uttarakhand	Uttar Kashi	82	10	4	0	0	0	2				
India	Meghalaya	South Vest Khasi Hills	19	4	2	0	0	0	0				
India	Punjab	Tarn Taran	142	3	11	10	15	2	1				
India	Odisha	Koraput	304	6	16	3	45	0	1				
India	Vest Bengal	Darjeeling	230	28	14	0	0	1	3				
India	Uttar Pradesh	Ghaziabad	130	21	4	16	45	0	3				
India	Telangana	Vikarabad	154	0	4	0	25	0	1				
India	Uttar Pradesh	Mainpuri	206	37	10	0	9	0	2				
India	Manipur	Kakching	18	6	1	0	0	0	0				
India	Uttar Pradesh	Saharanpur	363	43	12	0	17	0	2				
India	Odisha	Cuttack	325	14	22	7	59	2	1				
India	Tamil Nadu	Majiladuthurai	137.4877986	17.61207337	6.2494	19.316	9.0901	6.2494	0.5681				
India	Punjab	Barnala	57	3	4	14	7	1	1				
India	Odisha	Nayagarh	166	16	12	0	22	1	1				
India	Uttar Pradesh	Ambedkar Nagar	272	15	10	0	13	0	1				
India	Telangana	Jangam	102	0	0	10	17	1	0				
India	Odisha	Sonepur	89	1	5	0	19	1	1				
India	Rajasthan	Dungarpur	349	53	15	2	4	1	1				
India	Uttar Pradesh	Lucknow	291	18	17	15	62	0	10				
India	Uttar Pradesh	Mathura	205	20	7	0	17	0	3				
India	Nagaland	Mon	50	15	2	7	0	0	1				
India	Punjab	Pathankot	54	0	4	15	9	0	1				
India	Tamil Nadu	Tenkasi	195.7583535	32.28105213	8.9838	18.595	7.8386	7.539	0.4774				
India	Rajasthan	Alwar	676	123	37	9	10	0	1				
India	Tripura	Sepahijala	146	15	4	0	0	2	0				
India	Vest Bengal	Paschim Bardhaman	153	38	15	14	22	2	1				
India	Uttar Pradesh	Prayagraj	539	58	19	23	26	0	6				
India	Meghalaya	Vest Garo Hills	64	9	5	10	4	0	2				
India	Uttarakhand	Bageshwar	84	13	3	0	0	0	1				
India	Odisha	Puri	237	13	17	4	38	0	1				
India	Tamil Nadu	Theni	171.7099015	2.160143768	7.9757	11.925	30.758	4.9772	0.9937				
India	Tripura	Khowai	104	9	1	0	0	1	1				
India	Uttar Pradesh	Kanpur Dehat	240	15	6	0	20	0	1				
India	Vest Bengal	Dinajpur Uttar	344	24	10	0	0	4	1				
India	Odisha	Sundargarh	386	31	21	4	38	1	2				
India	Uttar Pradesh	Kushi Nagar	352	37	16	16	20	0	1				
India	Manipur	Kangpokpi	18	6	1	4	0	0	0				
India	Telangana	Karimnagar	139	0	3	0	26	0	1				
India	Tamil Nadu	Nagapattinam	104.5122014	13.38792663	4.7506	14.684	6.9099	4.7506	0.4319				
India	Rajasthan	Bhilwara	533	77	25	3	7	0	1				
India	Mizoram	Saiha	24	4	0	0	0	0	1				
India	Telangana	Hyderabad	0	0	10	53	51	3	1				
India	Odisha	Kendujhar	350	30	17	1	37	2	1				
India	Uttar Pradesh	Kaushambi	175	24	5	0	12	0	1				
India	Tamil Nadu	Tiruppur	315.7287719	11.9897002	12.989	23.979	41.964	8.9923	0.9991				
India	Punjab	Fatehgarh Sahib	61	0	5	12	13	1	1				
India	Vest Bengal	Bankura	564	73	22	0	0	5	1				

India	Odisha	Sambalpur	166	2	11	1	34	2	1				
India	Uttar Pradesh	Ballia	367	71	9	0	10	0	2				
India	Puducherry	Mahe	4	1	1	0	0	1	1				
India	Uttar Pradesh	Agra	377	42	16	18	33	0	2				
India	Rajasthan	Sawai Madhopur	266	33	14	1	5	1	1				
India	Mizoram	Mamit	33	6	1	0	0	0	1				
India	Rajasthan	Jhunjhunu	585	104	28	7	11	0	1				
India	The Diadra And Nagar Haveli And Daman An	Daman	6	0	1	14	3	0	1				
India	Tamil Nadu	Cuddalore	336	36	13	19	22	9	1				
India	Tamil Nadu	Ranipet	154.8629605	16.7666591	6.097	17.072	13.413	3.6582	0.3048				
India	Rajasthan	Chittorgarh	381	54	21	3	0	1	1				
India	Odisha	Dhenkanal	165	25	10	2	13	2	1				
India	Odisha	Kendrapara	224	15	8	3	31	1	1				
India	Rajasthan	Nagaur	795	117	31	9	9	3	1				
India	Tamil Nadu	Chengalpattu	222.1541106	15.94670596	7.1485	22.545	19.246	4.949	0.5439				
India	Rajasthan	Kota	185	45	13	4	18	0	0				
India	Uttar Pradesh	Gonda	320	39	16	2	13	0	2				
India	Rajasthan	Dausa	213	48	15	2	0	0	1				
India	Vest Bengal	Howrah	419	36	16	29	38	8	1				
India	Punjab	Amritsar	140	22	6	39	12	2	1				
India	Telangana	Bhadradi Kothagudem	266	0	4	1	34	2	0				
India	Mizoram	Kolasib	26	5	1	0	0	0	1				
India	Mizoram	Champhai	34.59578485	6.450061582	1.1727	0	0	0	0.5884				
India	Rajasthan	Jaisalmer	156	20	8	2	10	0	1				
India	Uttar Pradesh	Firozabad	216	43	9	4	22	0	3				
India	Uttar Pradesh	Baghpat	200	6	7	0	18	0	1				
India	Uttar Pradesh	Sant Kabir Nagar	185	22	4	0	1	0	1				
India	Telangana	Medchal Malkajgiri	103	0	2	24	34	0	0				
India	Uttar Pradesh	Jaunpur	503	73	16	4	10	0	2				
India	Vest Bengal	24 Paraganas South	1068	76	33	0	0	12	2				
India	The Diadra And Nagar Haveli And Daman An	Diu	0	0	1	6	1	0	1				
India	Uttar Pradesh	Budaun	291	40	8	0	17	0	2				
India	Odisha	Khordha	196	20	16	6	54	0	2				
India	Vest Bengal	Coochbehar	406	31	12	0	0	4	1				
India	Telangana	Mahabubabad	171	0	2	2	21	1	0				
India	Telangana	Narayanpet	87	0	2	0	16	1	0				
India	Odisha	Deogarh	38	1	4	4	7	0	1				
India	Rajasthan	Karauli	276	36	11	2	4	1	1				
India	Telangana	Rajanna Sircilla	84	0	0	5	17	1	0				
India	Meghalaya	South Vest Garo Hills	27	9	1	0	0	0	1				
India	Odisha	Malkangiri	158	5	6	0	22	2	1				
India	Punjab	Meer Kotla	47.54341037	18.18600397	2.0784	2.3382	5.4558	1.0392	0.2598				
India	Uttar Pradesh	Deoria	320	66	16	0	11	0	2				
India	Telangana	Jogulamba Gadval	91	0	1	0	13	1	0				
India	Rajasthan	Alwar	370	72	22	2	17	3	1				
India	Tripura	North Tripura	95	14	3	0	0	1	1				
India	Uttarakhand	Almora	206	25	8	0	0	2	2				
India	Telangana	Ranga Reddy	232	0	7	20	56	2	0				
India	Telangana	Jagshankar Bhupalapalli	83	0	2	7	14	0	0				
India	Nagaland	Peren	16	8	1	0	0	0	1				
India	Punjab	Moga	89	10	6	33	11	0	1				
India	Uttar Pradesh	Bijnor	331	48	7	12	16	0	2				
India	Rajasthan	Bikaner	414	45	16	5	26	0	0				
India	Nagaland	Phek	44	23	3	0	0	0	1				
India	Puducherry	Pondicherry	55	10	2	1	17	3	2				
India	Punjab	S.A.S Nagar	51	1	5	25	11	2	1				
India	Vest Bengal	Alipurdwar	209	5	7	27	9	2	1				
India	Odisha	Bargarh	202	19	15	2	30	1	1				
India	Mizoram	Lunglei	57.52004206	9.038863752	0.8217	0	0	0.8217	0.8217				
India	Punjab	Faridkot	54	3	4	8	5	1	1				
India	Telangana	Yadadri Bhuvanagiri	129	0	3	0	21	1	0				
India	Rajasthan	Bharatpur	534	84	17	1	11	0	1				

India	Punjab	Ludhiana	209	5	11	56	26	4	1				
India	Puducherry	Karaikal	17	0	1	0	11	2	1				
India	Telangana	Hanumakonda	81	0	0	0	28	0	0				
India	Nagaland	Tuensang	27.25192781	9.323027935	14343	7.8887	0	0	0.7172				
India	Rajasthan	Pratapgarh	202	28	8	0	4	0	1				
India	West Bengal	Jhargram	219	26	8	0	0	3	1				
India	Tamil Nadu	Virudhunagar	209	3	11	80	44	10	1				
India	Tamil Nadu	The Nilgiris	159.8078471	3.213703797	3.8654	39.924	28.429	4.7981	0.8439				
India	Uttar Pradesh	Gautam Buddha Nagar	110	12	5	18	25	0	1				
India	Telangana	Mahabubnagar	102	0	2	28	17	0	0				
India	Telangana	Nizamabad	225	0	8	0	37	1	0				
India	Tamil Nadu	Thiruvavur	191	28	10	16	12	7	1				
India	Uttarakhand	Pauni Garhwal	228	26	5	10	3	3	2				
India	Manipur	Tamenglong	13	4	1	3	0	0	1				
India	Rajasthan	Ganganagar	414	52	17	2	10	0	1				
India	Uttar Pradesh	Jhansi	311	25	8	27	24	0	2				
India	Uttar Pradesh	Fatehpur	302	35	7	19	13	0	2				
India	Uttar Pradesh	Moradabad	249	31	3	18	29	0	2				
India	Tamil Nadu	Tirupathur	150.6026236	16.30540216	5.9292	16.602	13.044	3.6575	0.2965				
India	Uttar Pradesh	Sonbhadra	173	26	7	0	5	0	1				
India	Telangana	Kumaram Etheem Asifab	108	0	2	0	22	0	0				
India	Tamil Nadu	Thiruvallur	196.135782	23.40975463	8.8577	27.206	10.756	6.327	0.8327				
India	Nagaland	Kiphire	19	2	1	15	2	0	1				
India	Meghalaya	Pili Ehoi	32	8	3	10	2	0	1				
India	Rajasthan	Jodhpur	651	95	25	2	18	0	0				
India	Mizoram	Serchhip	26.68578484	4.941812008	0.9884	0	0	0	0.9884				
India	Rajasthan	Sikar	631	96	30	9	15	2	1				
India	Uttar Pradesh	Maharajan	291	27	12	0	13	0	1				
India	West Bengal	Medinipur West	628	27	22	11	40	5	0				
India	Uttar Pradesh	Gorakhpur	513	57	9	16	34	0	2				
India	Uttar Pradesh	Ghazipur	418	43	6	0	20	0	2				
India	Nagaland	Dimapur	47.04268948	8.014940969	2.0032	6.0107	2.0011	0	1.0011				
India	West Bengal	Murshidabad	832	81	28	0	0	7	1				
India	Tamil Nadu	Pudukkottai	218	3	13	35	60	12	0				
India	Telangana	Mancheri	121	0	3	0	21	1	0				
India	Tripura	Unakoti	70	6	1	0	5	1	1				
India	Mizoram	Aizawl	88.94065284	17.03188884	1.8924	0	3.7847	0.9462	1.8924				
India	Uttar Pradesh	Basti	267	32	11	6	5	0	3				
India	Telangana	Nalgonda	257	0	1	0	39	3	1				
India	Manipur	Ukhul	23	6	1	0	0	0	1				
India	Maharashtra	Buldhana	280	57	13	0	0	1	1				
India	Telangana	Nagarkurnool	178	0	4	0	27	1	0				

Table 2 Health Expenditure

State	Year	Medical And Public Expenditure On Health By Urban	Medical And Public Expenditure On Health By Rural	Health	Medical And Public	Medical And Public	Medical And Public	Medical And Public	Medical And Public	Medical And Public	Medical And Public	Medical And Public	Medical And Public
Andhra Pradesh	2007	888	373	343	243	0	175	48	1131				
Andhra Pradesh	2011	29512718	9568798	4676291	4459904	2661320	0	2163170	863804	1E+07			
Andhra Pradesh	2012	25855013	8586905	5099904	4306682	2661869	0	2258592	1284854	1.4E+07			
Arunachal Pradesh	2011	165	226	5	68	26	2	5	0	19			
Arunachal Pradesh	2012	2160082	3036579	651168	939709	148954	17400	40625	0	77638			
Arunachal Pradesh	2013	1858409	4250714	678015	1223754	1155003	20653	45086	0	120648			
Assam	2011	292	1727	411	180	28	3	170	14	37			
Assam	2012	3319106	15941540	8232115	1770284	124564	30736	1953073	196515	399837			
Assam	2013	3535706	29422825	6468514	1814153	151860	48968	2374524	162788	437621			
Bihar	2011	1917	1737	474	79	0	2	330	4	29			
Bihar	2012	20730954	24365340	4630112	869161	0	31098	3864082	36353	401908			
Bihar	2013	20713631	27447856	7194319	1017346	-81736	36216	4963246	35129	489432			
Chhattisgarh	2011	743	1465	167	146	7	2	162	0	17			
Chhattisgarh	2012	9220709	17877398	2108148	1712961	19089	27508	1780999	0	175865			
Chhattisgarh	2013	12057941	21522930	2292549	1891946	11529	29900	2073687	0	198440			
Delhi	2011	2798	6	156	293	317	4	0	1	59			
Delhi	2012	31106023	402359	1719708	2566430	4003617	69653	1181	0	441110			
Delhi	2013	38278718	95402	2037557	2042345	4341720	61298	0	28173	446880			
Goa	2011	289	103	88	50	31	0	10	0	2			
Goa	2012	3471770	1028505	967164	368072	736357	0	103838	0	22338			
Goa	2013	4616274	1575767	1008382	639634	1226743	0	120471	0	33898			
Gujarat	2011	3034	1213	522	1466	35	108	253	324	170			
Gujarat	2012	32826254	12306523	6328431	14764759	251269	899834	1830620	5446724	2340885			
Gujarat	2013	38251137	10817492	6937883	14850039	167818	2013519	2710944	6495081	2911477			
Haryana	2011	785	862	555	179	2	3	101	7	30			
Haryana	2012	8333562	11528157	7062905	1795210	29776	39984	1113912	237129	303659			
Haryana	2013	11042530	11821712	7274892	1879478	39891	40536	1138760	234209	292737			
Himachal Pradesh	2011	357	476	262	34	0	10	72	0	206			
Himachal Pradesh	2012	4052171	7261502	3070704	373320	0	119716	887921	0	2103130			
Himachal Pradesh	2013	4443337	7427195	4214315	563807	0	96887	884253	0	2428570			
Jammu and Kashmir	2011	309.0888098	898.9006841	541.88355	73.35968586	0	0	16.6282	0	42.0596			
Jammu and Kashmir	2012	7703590.6	4508893.656	5461939.7	650376.6528	0	0	178359	0	61911.7			
Jammu and Kashmir	2013	8319015.764	5182097.887	5769982.8	717618.1409	0	0	192010	0	55460.9			
Jharkhand	2011	1466	329	202	101	0	1	60	3	12			
Jharkhand	2012	15532089	3632608	3745400	989425	0	31354	695601	17323	148574			
Jharkhand	2013	16113661	4013771	6224672	1171500	0	22496	734985	0	147969			
Karnataka	2011	1433	1204	1160	455	941	14	0	64	559			
Karnataka	2012	15666910	14799016	13157981	3985867	1.5E+07	115750	0	776031	5393310			
Karnataka	2013	18473012	20433185	14420939	3427501	1.8E+07	121026	0	1311350	4873018			
Kerala	2011	1700	706	980	728	221	0	372	0	64			
Kerala	2012	20763575	8372351	12322253	9889631	3407704	0	4390738	0	729467			
Kerala	2013	23964333	9967210	12660119	9748828	2725965	0	4909299	0	842206			
Ladakh	2011	6.911190242	20.09931592	12.116454	1.64031414	0	0	0.3718	0	0.94045			
Ladakh	2012	172251.3999	100818.3436	122128.34	14542.34717	0	0	3988.08	0	1384.34			

Ladakh	Calendar Year (Jan - Dec), 201	186012.2358	115871.1128	129016.15	16045.85911	0	0	4293.31	0	1240.1
Madhya Pradesh	Calendar Year (Jan - Dec), 201	3440	871	289	398	74	17	0	0	364
Madhya Pradesh	Calendar Year (Jan - Dec), 201	37050443	11423048	3624441	3199382	377010	159620	0	0	3568708
Madhya Pradesh	Calendar Year (Jan - Dec), 201	44810635	17686387	4088536	3621223	248650	194096	0	0	3837960
Maharashtra	Calendar Year (Jan - Dec), 201	3418	558	813	4567	7	17	472	65	92
Maharashtra	Calendar Year (Jan - Dec), 201	36271242	5603558	9138832	49337851	36123	138044	4938487	673622	1100317
Maharashtra	Calendar Year (Jan - Dec), 201	43399175	6445365	9419183	53957887	299660	128609	5004558	1710945	1380347
Manipur	Calendar Year (Jan - Dec), 201	1151413	1999546	662864	760725	24224	1915	110801	0	80469
Manipur	Calendar Year (Jan - Dec), 201	1078454	3006291	952408	502760	28278	1879	92801	0	139581
Meghalaya	Calendar Year (Jan - Dec), 201	234	137	4	32	155	0	27	8	8
Meghalaya	Calendar Year (Jan - Dec), 201	2511152	1450017	52681	340345	2129125	3550	288816	20193	74778
Meghalaya	Calendar Year (Jan - Dec), 201	5446772	3446268	116655	1208907	2934820	7462	684659	37391	166709
Mizoram	Calendar Year (Jan - Dec), 201	106	98	3	134	6	1	26	0	5
Mizoram	Calendar Year (Jan - Dec), 201	1515059	986057	13333	1135830	38766	6917	237852	4577	51627
Mizoram	Calendar Year (Jan - Dec), 201	1648728	1335649	500826	1486964	227874	8318	274963	5171	50584
Nagaland	Calendar Year (Jan - Dec), 201	202	107	5	127	0	0	0	0	24
Nagaland	Calendar Year (Jan - Dec), 201	2422957	1069033	30464	1175419	0	0	243808	0	14398
Nagaland	Calendar Year (Jan - Dec), 201	2553649	1806445	38402	1407590	0	0	264365	0	11182
Odisha	Calendar Year (Jan - Dec), 201	1388	1489	358	268	10	8	89	0	69
Odisha	Calendar Year (Jan - Dec), 201	12582116	21880292	6994382	2476690	103100	98267	1776786	0	1380857
Odisha	Calendar Year (Jan - Dec), 201	13890281	24282275	4416887	3037524	169280	109840	1922092	0	1442909
Puducherry	Calendar Year (Jan - Dec), 201	282	41	108	25	25	0	4	0	3
Puducherry	Calendar Year (Jan - Dec), 201	2781914	429936	981813	275629	351213	0	50342	0	33000
Puducherry	Calendar Year (Jan - Dec), 201	2936069	463264	863806	304487	383441	0	55563	0	35901
Punjab	Calendar Year (Jan - Dec), 201	1217	514	263	195	222	11	113	0	73
Punjab	Calendar Year (Jan - Dec), 201	16553164	5179139	2728629	1666279	806821	105539	1144804	0	717952
Punjab	Calendar Year (Jan - Dec), 201	15120803	5348080	3113239	1713096	58249	119032	1276072	0	729938
Rajasthan	Calendar Year (Jan - Dec), 201	1829	1892	746	849	0	20	0	0	2422
Rajasthan	Calendar Year (Jan - Dec), 201	20569608	20457303	8951450	9711815	0	204115	0	0	2.3E+07
Rajasthan	Calendar Year (Jan - Dec), 201	21885115	24320500	10153783	17185638	0	229163	0	0	2.6E+07
Sikkim	Calendar Year (Jan - Dec), 201	174	36	1	32	0	1	10	0	7
Sikkim	Calendar Year (Jan - Dec), 201	1861427	421741	10196	349267	-533	5075	108022	0	64206
Sikkim	Calendar Year (Jan - Dec), 201	3692755	405213	35074	321157	0	4790	99206	0	58806
Tamil Nadu	Calendar Year (Jan - Dec), 201	2708	923	863	489	1250	75	347	1179	510
Tamil Nadu	Calendar Year (Jan - Dec), 201	31828082	9486183	9689160	4982816	1.2E+07	753024	3735387	8919434	5458936
Tamil Nadu	Calendar Year (Jan - Dec), 201	36054129	10405610	11669384	5430703	1.9E+07	753137	4135544	1.2E+07	4970327
Telangana	Calendar Year (Jan - Dec), 201	2066	495	242	261	120	0	85	45	445
Telangana	Calendar Year (Jan - Dec), 201	26192202	5087573	2733549	2901662	17696	0	91910	545467	1.2E+07
Telangana	Calendar Year (Jan - Dec), 201	22568912	5126335	2741772	2999512	3267691	0	936090	1976360	1.1E+07
Tripura	Calendar Year (Jan - Dec), 201	266	102	29	45	3	0	0	35	129
Tripura	Calendar Year (Jan - Dec), 201	3208635	1154301	331581	482008	13381	0	0	90560	1621672
Tripura	Calendar Year (Jan - Dec), 201	3752090	1333073	106512	702119	2559	31000	0	324220	1594934
Uttar Pradesh	Calendar Year (Jan - Dec), 201	2686	4378	1486	430	6	49	1039	273	3105
Uttar Pradesh	Calendar Year (Jan - Dec), 201	30303063	53020412	20769220	4103898	8740	512123	1.1E+07	2693287	3.6E+07
Uttar Pradesh	Calendar Year (Jan - Dec), 201	33780651	48283271	26299263	5717406	115268	544019	1.2E+07	4286500	3.8E+07
Uttarakhand	Calendar Year (Jan - Dec), 201	442	632	115	167	0	4	93	0	11
Uttarakhand	Calendar Year (Jan - Dec), 201	4577664	6494412	1473211	1420867	0	40413	79564	634	974713
Uttarakhand	Calendar Year (Jan - Dec), 201	5675785	5848963	1694424	1800517	0	37849	1019182	0	116573
West Bengal	Calendar Year (Jan - Dec), 201	3910	2429	610	288	-2	0	177	0	447
West Bengal	Calendar Year (Jan - Dec), 201	35212844	25014709	8071542	7549897	117851	0	1723279	0	4933225
West Bengal	Calendar Year (Jan - Dec), 201	36418882	30477621	8332096	6390413	-268823	0	2046100	0	5162254

Problem Statement 1.1: Is there a significant difference in the number of Functional Sub Centres across different states?

Hypothesis of Problem statement:

- Null Hypothesis (H0): There is no significant difference in the number of Functional Sub Centres across different states.
 $\mu_1 = \mu_2 = \dots = \mu_k$ (The mean number of Functional Sub Centres is the same across all states).
- Alternative Hypothesis (H1): There is a significant difference in the number of Functional Sub Centres across different states.
 $\mu_1 \neq \mu_2$ for at least one pair.

Descriptive Statistics on Functional Sub Centres

Mean = 204.52

Median = 205

Maximum = 877.26

Minimum = 3

Standard Deviation = 100.38

Some Analysis:

- a) With **204.52 mean**, there is substantial average number of functional. Across the states, indicating that there are many areas that have established local healthcare access through these facilities.
- b) The standard deviation is **100.78** which shows that some districts have numerous centres while others have few.
- c) There are outliers, the maximum value of **877.26** indicate that the state like Tamil Nadu which has maximum no. of subcentres and some states have negligible subcentres of value 3.

ANOVA Methodology

Group Means:

Andaman And Nicobar Islands	32.000000
Andhra Pradesh	526.615385
Arunachal Pradesh	12.200000
Assam	122.124054
Bihar	259.605263
Chandigarh	4.000000

Chhattisgarh	175.663859
Delhi	18.545455
Goa	109.500000
Gujarat	253.121212
Haryana	112.363636
Himachal Pradesh	174.750000
Jammu And Kashmir	131.900000
Jharkhand	151.833333
Karnataka	314.600000
Kerala	384.285714
Ladakh	131.000000
Lakshadweep	14.000000
Madhya Pradesh	196.634615
Maharashtra	277.971429
Manipur	26.812500
Meghalaya	40.454545
Mizoram	33.636364
Nagaland	33.083333
Odisha	219.833333
Puducherry	20.000000
Punjab	109.173913
Rajasthan	406.939394
Sikkim	37.000000
Tamil Nadu	260.815789
Telangana	141.151515
The Dadra & Nagar Haveli And	
Daman And Diu	17.666667
Tripura	120.750000
Uttar Pradesh	267.413333
Uttarakhand	138.769231

Overall Mean: 207.590

SSB: 10029331.062

SSW: 7931223.178

f_Statistics: 25.290

p_value: 6.592683299291285e-101 nearly to 0

df_between: 35

df_within: 700

Here I have made box plot diagram of Functional sub centres in every states. There are outliers also which I have mentioned.

I have taken states on X-axis and Functional sub centre on Y- axis.

RESULT:

The significant level(alpha) is 0.05. And **alpha is > p_value**. It doesn't execute the null hypothesis. Means, there is significant difference in the number of Functional Sub Centres across different states.

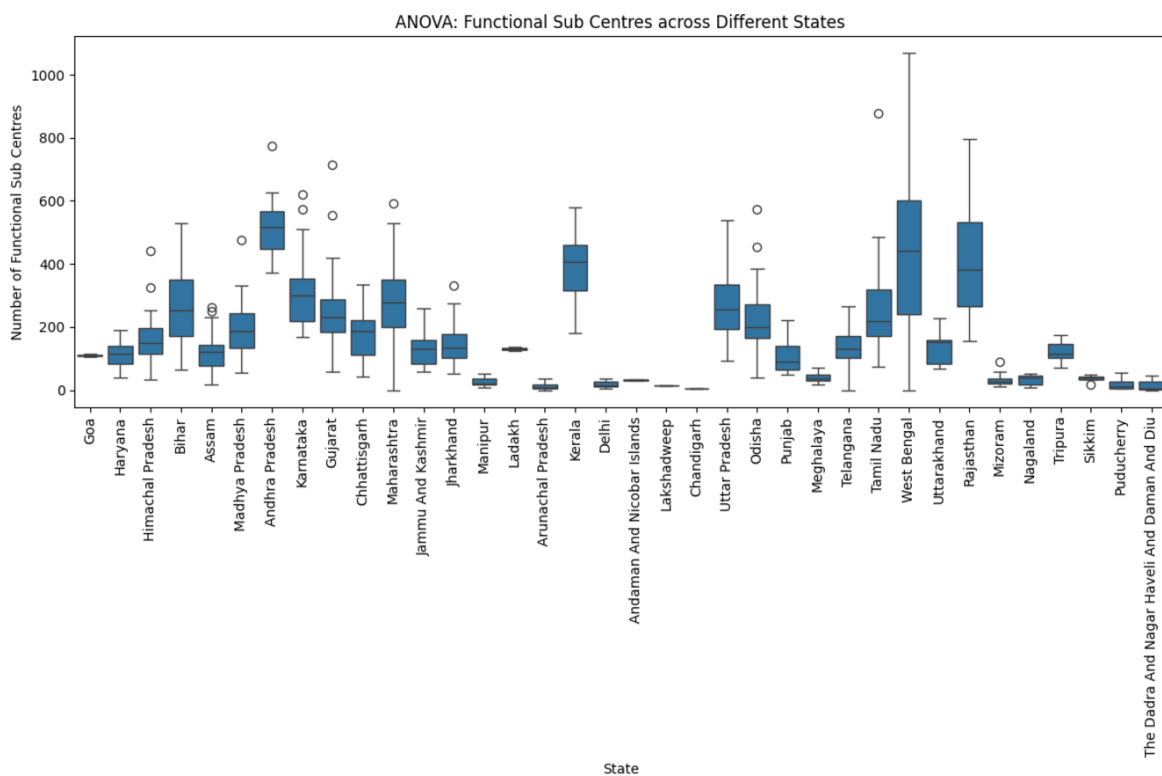


Fig 4: Functional sub centres across different states

Problem Statement 1.2: Is there a significant correlation between the medical and public expenditure on health by urban health services (UHS) and rural health services (RHS) in different states?

Hypothesis of problem statement

- NULL Hypothesis(H0): There is no significant correlation between the medical and public expenditure on health by urban health services (UHS) and rural health services (RHS) in different states. ($r=0$)
- Alternative Hypothesis(H1): There is a significant correlation between the medical and public expenditure on health by urban health services (UHS) and rural health services (RHS) in different states. ($r \neq 0$)

r is the Pearson correlation coefficient(r)

Descriptive Statistics

Stats	Urban Health Services	Rural Health Services
Count	95	95
Mean	10,459,240.52	6,932,797.05
Median	3,471,770	1,806,445
Standard Deviation	13,192,726.80	10,377,660.10
Minimum	6.91	6.00
Maximum	44,810,640	53,020,410.

Some Analysis:

- There is high variability in Urban and Rural health services.
- On average, Urban health services receive higher funding (10.46 million INR) compared to Rural Health Services (6.93 million INR).
- The minimum value is low in both categories, 6.91 rupees for urban and 6 rupees for rural.

CORRELATION:

I have analysed the correlation between Rural Health Services and Urban Health Services.

Pearson correlation coefficient®

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

X_i and Y_i – Data values of Urban Health Services and Rural Health Services respectively.

\bar{X} and \bar{Y} – These are the mean of X and Y .

The numerator represents the covariance between X and Y.

Denominator represents the combine standard deviations of X and Y.

r ranges from -1 to 1:

- **1:** Perfect positive linear relationship.
- **-1:** Perfect negative linear relationship.
- **0:** No linear relationship.

By calculating, we got r is 0.57 and the p value is 1.6×10^{-9} ($p_value < \alpha$).

Conclusion: We reject the null hypothesis. There is a significant correlation between Urban Health Expenditure and Rural Health Expenditure

Scattered plot represents the correlation between Urban Health Services and Urban Health Services

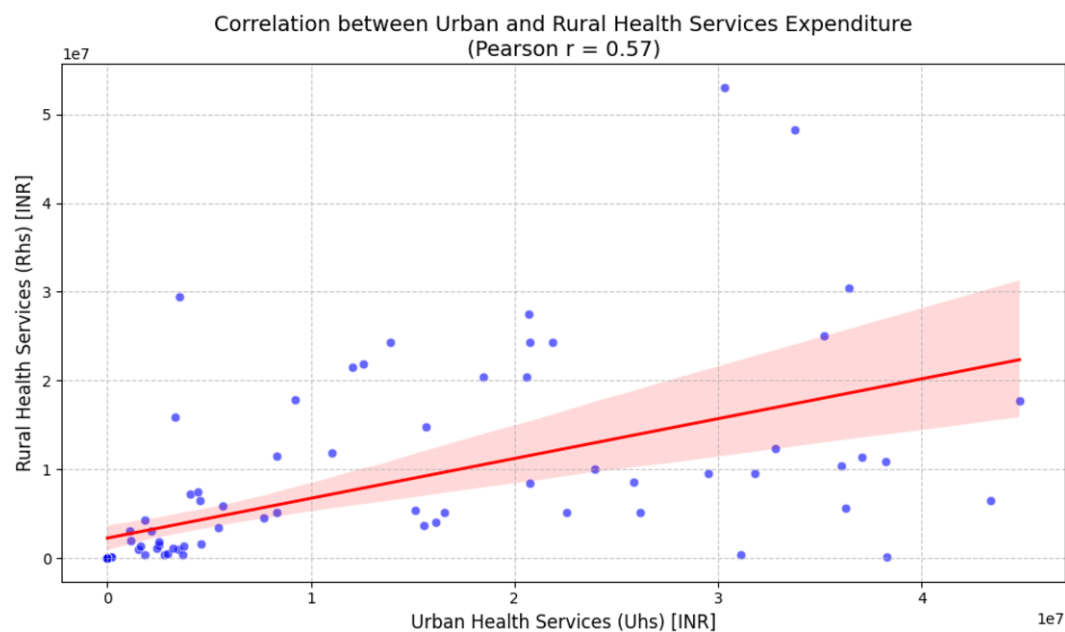


Fig 5: Correlation between urban and rural health expenditure (in rupees)

Analysis of Scatter Plot:

Red line is trend line of expenditure. Blue dots represent a pair of categories.

If the blue dots are in cluster with red line, that represents good relation. And random blue dots are having less correlation.

Problem Statement 1.3: Is there a significant trend in the medical and public expenditure on health services in rural areas over the years?

- Null Hypothesis (H0): There is no significant trend in the medical and public expenditure on health services in rural areas over the years.
($\tau=0$) where τ is Mann-Kendall trend statistic
- Alternative Hypothesis (H1): There is a significant trend in the medical and public expenditure on health services in rural areas over the years.
($\tau \neq 0$) where τ is Mann-Kendall trend statistic

Statistical Data:

Mean = 95

Median=1,806,445

Standard Deviation=6,932,797.05

Minimum=6

Maximum=53,020,410

TIME-SERIES ANALYSIS

Mann-Kendall Trend Test Results:

Trend: Increasing

Mann-Kendall tau: 1.0

P-value: 0.3333333333333333

There is a statistically significant trend in the expenditure over the years.

Here is the graph for the trend in the Rural Health Expenditure over the years.

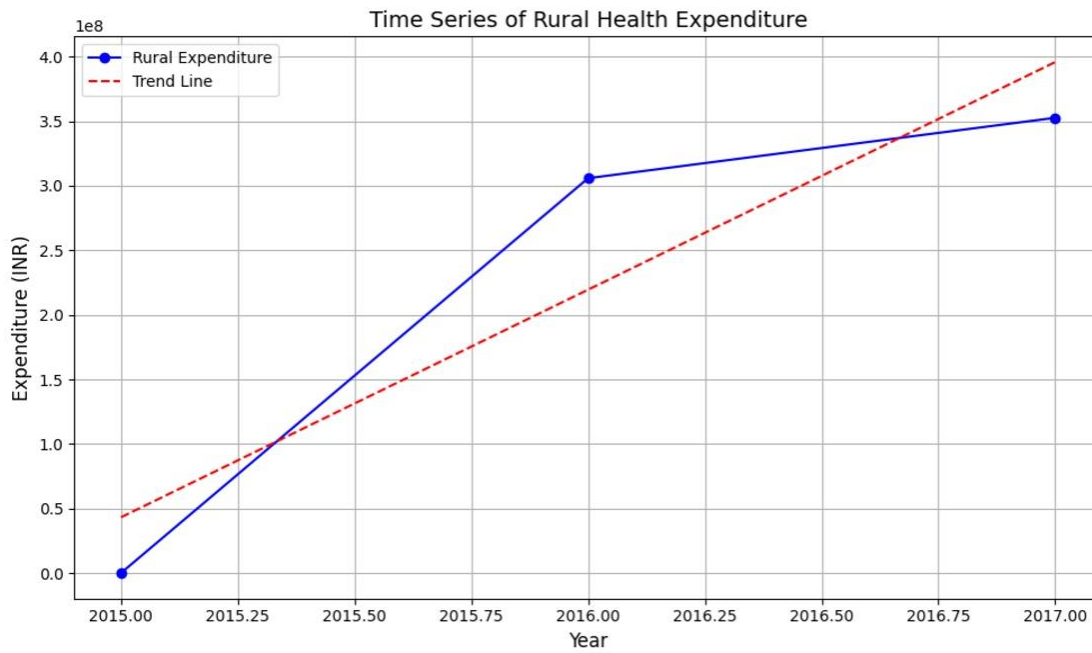


Fig 6: Medical and public expenditure on health by rural health service

Conclusion: ($P_{\text{value}} < \alpha$)

We reject the null hypothesis, there is a statistically significant trend in rural health expenditure over the years.

OBJECTIVE 2: Analysis on How the Inaccessibility of Natural Resources Affects Poverty in Rural and Local Communities

Table 3 Access of Tap Water in Rural India

(Number in lakhs)

S. No.	State/ UT	Total rural HHs as on date	Rural HHs with tap water connection as on 15/08/2019	Rural HHs provided with tap water connections					Total Rural HHs with tap water supply as on 30.01.2024	Balance households as on 30.01.2024
				19-20	20-21	21-22	22-23	23-24*		
1.	A&N Islands	0.62	0.29	-	0.33	-	-	-	0.62	-
2.	Andhra Pr.	95.46	30.74	1.19	12.77	9.59	12.12	3.07	69.43	26.03
3.	Arunachal Pr.	2.29	0.23	0.13	0.65	0.42	0.34	0.52	2.29	0.00
4.	Assam	70.34	1.11	0.49	5.07	16.52	8.75	19.09	51.03	19.31
5.	Bihar	166.30	3.16	28.97	103.46	19.86	3.63	1.26	160.34	5.96
6.	Chhattisgarh	50.00	3.20	0.96	1.51	4.45	10.81	17.38	38.30	11.70
7.	DNH and DD	0.85	0.00	-	0.25	0.61	-	-	0.85	-
8.	Goa	2.63	1.99	0.31	0.33	-	-	-	2.63	-
9.	Gujarat	91.18	65.16	1.06	10.95	8.99	5.03	-	91.18	-
10.	Haryana	30.41	17.66	1.35	7.91	3.49	-	-	30.41	-
11.	Himachal Pr.	17.09	7.63	1.59	3.79	2.87	0.93	0.29	17.09	-
12.	J&K	18.72	5.75	2.07	2.16	0.58	0.40	3.37	14.33	4.38
13.	Jharkhand	62.01	3.45	0.95	3.00	4.19	8.77	10.22	30.58	31.43
14.	Karnataka	101.16	24.51	0.21	3.43	18.70	20.57	6.84	74.27	26.89
15.	Kerala	70.79	16.64	0.85	4.04	6.64	5.29	3.40	36.87	33.92
16.	Ladakh	0.41	0.01	0.01	0.02	0.09	0.18	0.07	0.38	0.03

S. No.	State/ UT	Total rural HHs as on date	Rural HHs with tap water connection as on 15/08/2019	Rural HHs provided with tap water connections					Total Rural HHs with tap water supply as on 30.01.2024	Balance households as on 30.01.2024
				19-20	20-21	21-22	22-23	23-24*		
17.	Lakshadweep	0.13		-	-	-	-	0.11	0.11	0.03
18.	Madhya Pr.	111.86	13.53	4.19	19.87	10.89	9.07	9.79	67.18	44.68
19.	Maharashtra	146.64	48.44	5.45	37.09	10.66	8.22	12.57	122.42	24.22
20.	Manipur	4.52	0.26	0.04	1.96	0.70	0.49	0.07	3.52	1.00
21.	Meghalaya	6.51	0.05	0.02	0.87	1.34	0.80	1.72	4.80	1.71
22.	Mizoram	1.33	0.09	0.16	0.34	0.21	0.30	0.22	1.33	0.00
23.	Nagaland	3.69	0.14	0.02	0.48	0.92	0.87	0.66	3.09	0.60
24.	Odisha	88.63	3.11	4.37	15.47	17.47	12.22	10.01	62.66	25.98
25.	Puducherry	1.15	0.94	0.06	0.08	0.07	0.00	-	1.15	-
26.	Punjab	34.26	16.79	0.76	8.18	8.40	0.13	-	34.26	-
27.	Rajasthan	106.65	11.74	1.02	6.81	5.38	13.88	10.31	49.15	57.50
28.	Sikkim	1.32	0.70	-	0.10	0.08	0.19	0.09	1.17	0.15
29.	Tamil Nadu	125.26	21.76	0.17	16.13	14.90	26.52	20.74	100.22	25.04
30.	Telangana	53.98	15.68	20.18	18.20	-0.19	0.11	-	53.98	-
31.	Tripura	7.45	0.25	0.46	1.42	1.65	0.81	1.08	5.66	1.79
32.	Uttar Pr.	263.63	5.16	4.66	19.16	5.81	59.89	108.07	202.74	60.89
33.	Uttarakhand	14.55	1.30	0.87	4.32	2.75	2.24	1.70	13.10	1.45
34.	West Bengal	174.93	2.15	0.05	12.48	23.31	20.77	15.33	74.08	100.85
Total		19,26.76	3,23.63	82.62	3,22.62	2,01.34	2,33.26	2,57.99	14,21.22	5,05.54

*As on 30.01.2024

HH: Household

Source: JJM-IMIS

Descriptive Stats (Mean, Mode, Median)

Mean: 41.79823529411765

Median: 30.494999999999997

Mode: 0.03

Formula Used:

Mean = Sum of all values/Number of values

Median = Arranging in Ascending Order

If Odd:

Median = Middle value

Else:

Median = Average value of the middle values

Mode = Highest Occurring data

Output explanations:

Mean: 41.79823529411765 – This indicates that, on average, the total number of rural households with tap water is approximately 41.80.

Median: 30.494999999999997 – This is the middle value, indicating that half of the states have tap water access below this value and half are above.

Mode: 0.03 – The most frequently occurring value, indicating that many states have around 0.03.

Results:

Mean: 41.79823529411765, Median: 30.494999999999997, Mode: 0.03
Output explanations:
1. Mean: 41.79823529411765 – This indicates that, on average, the total number of rural households with tap water is approximately 41.80.
2. Median: 30.494999999999997 – This is the middle value, indicating that half of the states have tap water access below this value and half are above.
3. Mode: 0.03 – The most frequently occurring value, indicating that many states have around 0.03.

Problem Statement 2.1: Is there a correlation between rural population density and the percentage of rural households with tap water access across states in 2024.

I have analysed the correlation between Total Rural HH and Tap Water.

Pearson correlation coefficient

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \cdot \sum_{i=1}^n (Y_i - \bar{Y})^2}}$$

X_i and Y_i – Data values of Urban Health Services and Rural Health Services respectively.

\bar{X} and \bar{Y} – These are the mean of X and Y .

The numerator represents the covariance between X and Y.

Denominator represents the combine standard deviations of X and Y.

r ranges from -1 to 1:

- **1:** Perfect positive linear relationship.
- **-1:** Perfect negative linear relationship.
- **0:** No linear relationship.

The correlation coefficient is 0.93. This is a strong positive correlation between the total number of rural households and those with tap water access. A value close to 1 suggests a strong direct relationship; as the number of rural households increases, so does the access to tap water.

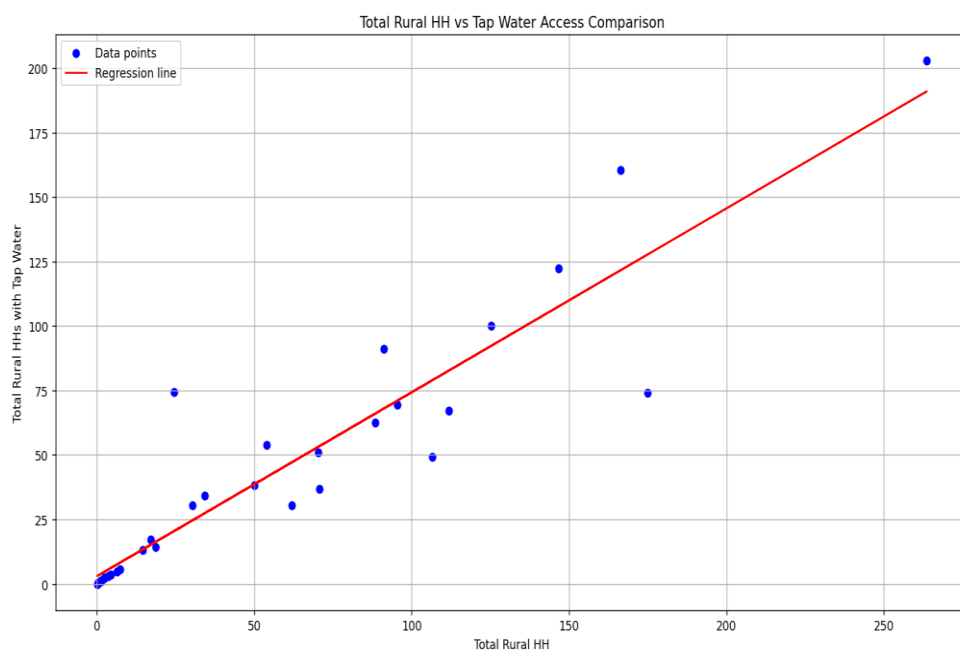


Fig 7: Total Rural HH vs Tap water Access

```
Regression line: Y = 0.712912547014985 * X + 3.0109700928020686
Sxx: 136349.53994411766
Sxy: 97205.29780588235
Syy: 79577.08589411766
SSE: 10278.209451975923
SST: 79577.08589411766
SSR: 69298.87644214171
R^2: 0.8708395848315965
R: 0.9331878614896342
x_bar: 54.40676470588235
y_bar: 41.798235294117646
Slope (m): 0.712912547014985
Intercept (c): 3.0109700928020686
Regression Line: Y = 0.712912547014985 * X + 3.0109700928020686
```

Problem Statement 2.2: Can past year's tap water access percentages predict the percentage of rural households with access in 2030?

Formula Used:

Regression line: $Y = a + bx$

$$b = S_{xy}/S_{xx}$$

$$a = \bar{y} - b\bar{x}$$

$$R^2 = SST/SSR$$

$$SSR = \sum(\hat{Y}_i - \bar{Y})^2$$

$$SST = \sum(Y_i - \bar{Y})^2$$

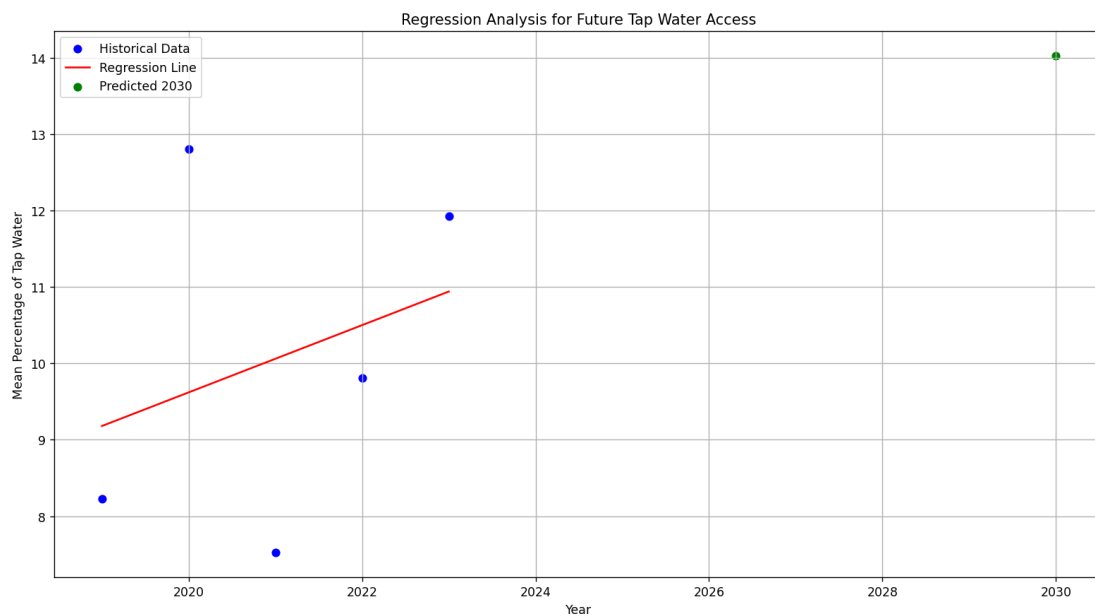


Fig 8: Access of tap water in future

Regression line: $Y = 0.4400952380952381 x + -879.3721904761906$

Predicted percentage for 2030: 14.021142857142763

The value of R^2 is: 0.0930993926668704

Problem Statement 2.3: Is there a significant difference in the average percentage of rural households with tap water access across different states?

To determine whether there is significant difference in the average percentage of rural household with tap water access across different states, ANOVA test is performed using the data for the years 2021 to 2024.

NULL Hypothesis: There is no significant difference in the average percentage of rural household with tap water access across different states from 2021-2024

Alternative Hypothesis: There is significant difference in the average percentage of rural household with tap water access across different states from 2021-2024

After the ANOVA results:

We got F (Ms (TR)/MSE) value as 0.128489 which is smaller than $F_{0.05, 2, 99}$ which equals 3.12.

$F (Ms (TR)/MSE) < F_{0.05, 2, 99}$

So, we fail to ignore the NULL Hypothesis.

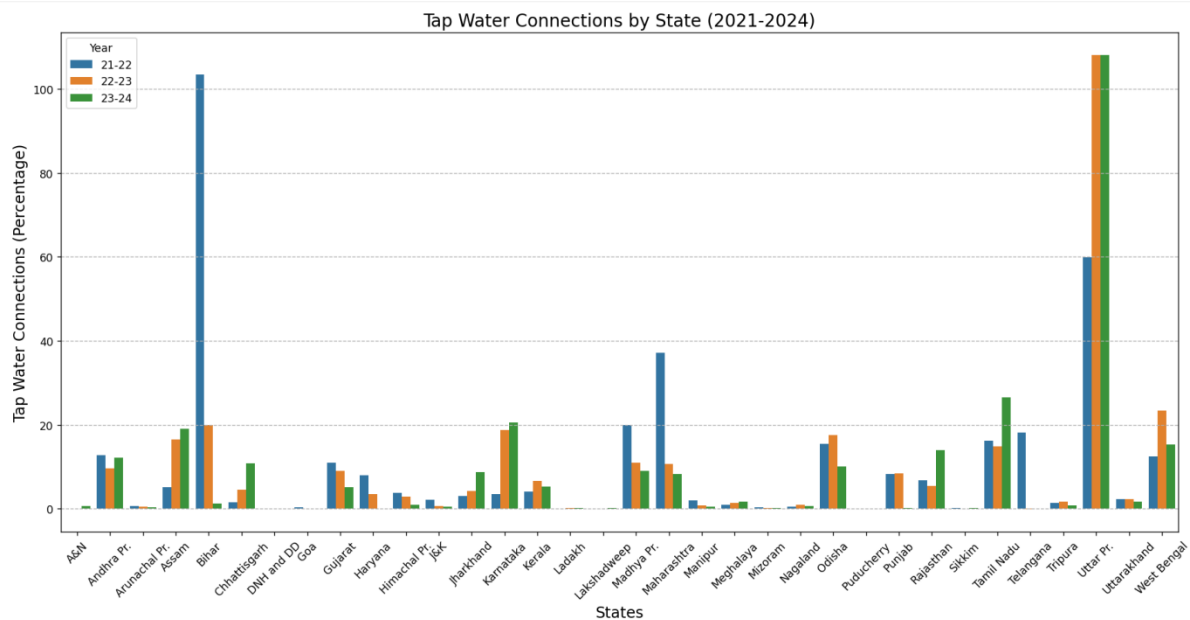


Fig 9: Tap water Connections by state

```

ANOVA Results:
              sum_sq   df      F    PR(>F)  Mean Square
C(Year)      97.013649   2.0  0.128489  0.87957    48.506825
Residual  37374.178297  99.0      NaN     NaN    377.516952

Conclusion:
Fail to reject the null hypothesis: This indicates that the ANOVA test did not find a statistically significant difference in tap water connection means among the years.

```

OBJECTIVE 3: Analysis on How the Gender Gap in Property Ownership Affects Women's Economic Stability.

Table 4 Percentage of men and women who own a House (alone or jointly) in India

State/Union Territory	Women	Men
Chandigarh	30.4	40.2
Delhi	21.9	37.2
Haryana	38.6	50.4
Himachal Pradesh	22.4	52.8
Jammu & Kashmir	56.6	78.5
Ladakh	71.9	74.6
Punjab	63.2	58.2
Rajasthan	26.0	55.9
Uttarakhand	23.8	52.3
Chhattisgarh	45.1	71.7
Madhya Pradesh	38.9	57.3
Uttar Pradesh	51.2	65.5
Bihar	54.4	67.8
Jharkhand	63.6	82.6
Odisha	42.5	73.1
West Bengal	22.0	51.6
Arunachal Pradesh	68.7	76.0
Assam	42.2	74.2
Manipur	57.3	76.8
Meghalaya	64.1	48.0
Mizoram	19.4	50.4
Nagaland	25.5	52.1
Sikkim	52.4	75.6
Tripura	15.8	53.4
Dadra & Nagar Haveli and Daman & Diu	55.8	44.1
Goa	22.8	21.2
Gujarat	42.2	63.7
Maharashtra	21.5	44.0
Andaman & Nicobar Islands	14.5	47.2
Andhra Pradesh	45.6	67.1
Karnataka	66.2	67.9
Kerala	24.5	53.6
Lakshadweep	29.7	29.3
Puducherry	33.6	47.9
Tamil Nadu	47.0	60.4
Telangana	63.6	74.3
INDIA	42.3	60.1

Source: National Family Health Survey (NFHS - 5), 2019-21-INDIA REPORT

Table 5 Percentage of men and women who own a Land (alone or jointly) in India

State/ Union Territory	Women	Men
Chandigarh	9.0	26.5
Delhi	12.7	21.1
Haryana	30.8	36.8
Himachal Pradesh	20.3	48.2
Jammu & Kashmir	51.1	69.8
Ladakh	63.8	70.7
Punjab	27.1	31.5
Rajasthan	20.7	43.3
Uttarakhand	17.5	38.8
Chhattisgarh	38.8	66.3
Madhya Pradesh	32.3	45.4
Uttar Pradesh	42.7	55.8
Bihar	43.8	52.9
Jharkhand	54.6	73.7
Odisha	36.6	58.0
West Bengal	16.7	33.6
Arunachal Pradesh	62.8	71.5
Assam	34.8	63.5
Manipur	24.8	44.0
Meghalaya	44.6	38.3
Mizoram	13.9	40.4
Nagaland	15.8	41.4
Sikkim	39.7	70.1
Tripura	10.7	34.8
Dadra & Nagar Haveli and Daman & Diu	50.1	39.0
Goa	9.2	5.2
Gujarat	35.0	42.4
Maharashtra	14.7	27.0
Andaman & Nicobar Islands	8.9	29.0
Andhra Pradesh	24.5	34.8
Karnataka	53.7	51.1
Kerala	11.5	25.1
Lakshadweep	6.1	12.6
Puducherry	10.0	7.2
Tamil Nadu	21.9	26.3
Telangana	42.6	51.1
INDIA	31.7	42.3

Source: National Family Health Survey (NFHS - 5), 2019-21-INDIA REPORT

Problem Statement 3.1: To access the total percentage of House and Land owned by women and men in each State and Union Territories of India according to data 2019-2021.

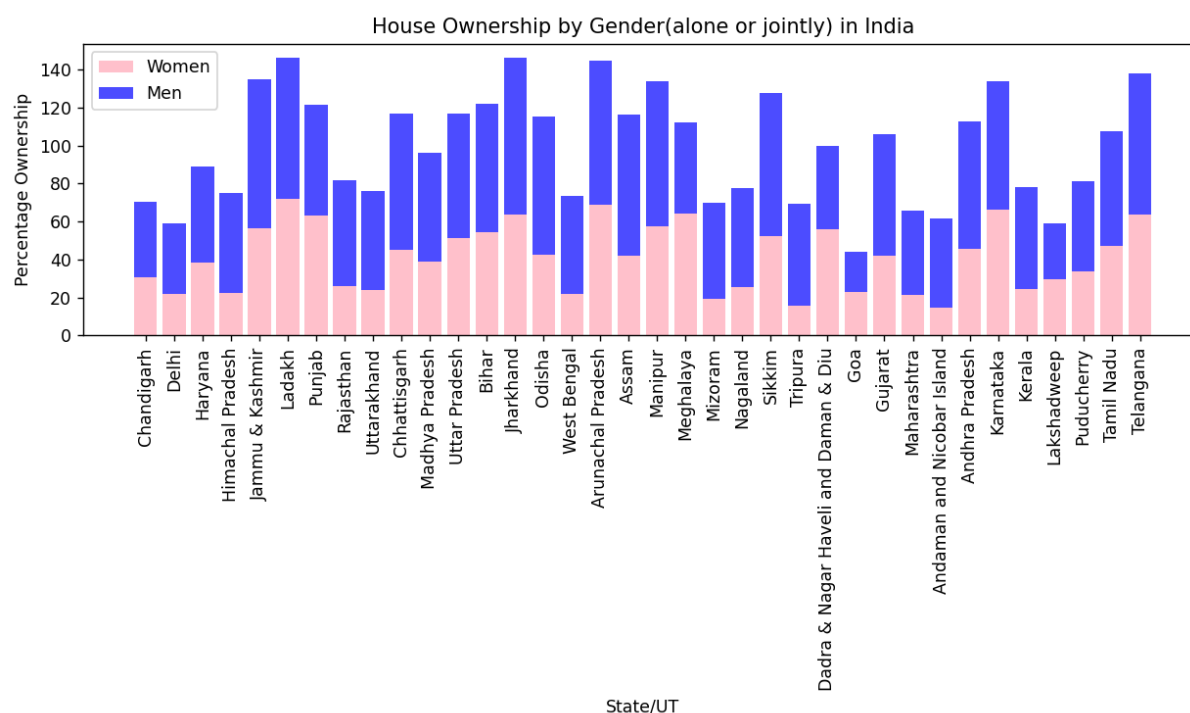


Fig 10: The distribution of House owned by men and women (alone or jointly) in India

From the above graphs we can access the percentage of House owned by women and men in India in 2019-2021.

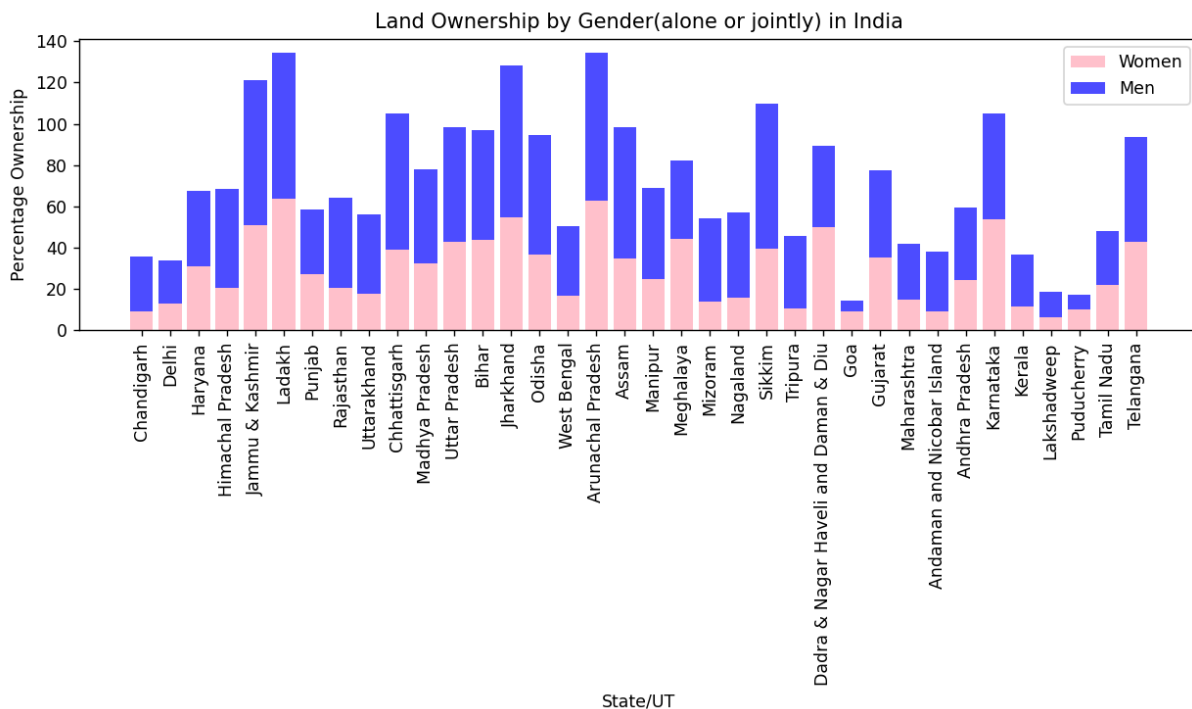


Fig 11: The distribution of Land owned by men and women (alone or jointly) in India

From the above graphs we can access the percentage of Land owned by women and men in India in 2019-2021.

Problem Statement 3.2: To test the hypothesis whether women own less land than men

It is one tailed Hypothesis test and two means.

Null Hypothesis: $\mu_{\text{female}} \geq \mu_{\text{male}}$

Alternative Hypothesis: $\mu_{\text{female}} < \mu_{\text{male}}$

Steps:

$$\mu_{\text{male}} = 42.42$$

$$\mu_{\text{female}} = 29.27$$

$$\sigma^2 = 273.68$$

$$\sigma^2 = 318.73$$

$$N = 35$$

Now we will calculate Z because $N > 30$ and the variance are known.

$$Z = \frac{29.27 - 42.42}{\sqrt{273.68 + 318.73 / 35}}$$

$$Z = -13.15 / 4.12$$

$$Z = -3.191$$

Now we will take the level of significance as $\alpha = 0.05$.

It is one tailed test. So, $Z < -Z_{\alpha}$

$$Z = -1.645$$

So, we can see our calculated Z-score is $-3.91 < -1.645$. So, we reject the Null hypothesis.

We conclude that women own less land than men.

Problem Statement 3.3: To test the hypothesis whether there is gender equality in land ownership.

Let's apply the **single proportion test** using 50% as the hypothesized proportion.

Null Hypothesis: The proportion of land owned by women is equal to 50% i.e. $p = 0.50$.

Alternative Hypothesis: The proportion of land owned by women is not equal to 50 % i.e. $p \neq 0.50$.

Steps:

$$\mu_{\text{female}} = 29.27$$

$$p_{\text{obs}} = 29.27/100 = 0.2927$$

$$\text{Hypothesized proportion } (p) = 0.50$$

$$n = 35$$

$$Z = \frac{p_{\text{obs}} - p}{\sqrt{\frac{p(1-p)}{n}}}$$

$$Z = 0.2927 - 0.50 / \sqrt{0.50(1-0.50)/35}$$

$$Z = -0.2073/0.0845$$

$$Z = -2.45$$

$$\alpha = 0.05$$

It is two tailed tests. $|Z| < Z_{\alpha}$

$$Z = -1.96 \text{ and } +1.96$$

So, we can see -2.45 lies outside -1.96 and $+1.96$. So, we reject the null hypothesis.

The proportion of land owned by women is different from 50 %

CONCLUSION

From the first objective we can decide that there are significant variations in number of functional health sub-centres across the states. This indicates unequal healthcare availability, with some states being better served while others lack basic health facilities. This highlights the urgent need for interventions to ensure equal healthcare access across the country.

From the second objective, we can conclude that limited access to natural resources, particularly water in rural areas. The study shows that rural with higher population density struggle more to get enough water. This highlights that there is need to improve water access to help improving living conditions in rural areas.

From the third objective, we can conclude that women own less land than men, which makes it harder for them to be financially independent. This inequality in property ownership shows the need for policies that give women the same rights as men when it comes to owning land and property.

APPENDIX

KALPANA BAGHEL

ANOVA CODE

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats

df = pd.read_csv('/content/drive/MyDrive/Accessible_health_Centres.csv')

data = df['State', 'Functional Sub Centres (UOM:Number)']
group_means = data.groupby('State')['Functional Sub Centres (UOM:Number)'].mean()
overall_mean = data['Functional Sub Centres (UOM:Number)'].mean()
SSB = sum(data.groupby('State').size() * (group_means - overall_mean)**2)
SSW = sum((data['Functional Sub Centres (UOM:Number)'] - data.groupby('State')['Functional Sub Centres (UOM:Number)'].transform('mean'))**2)

df_between = len(group_means) - 1
df_within = len(data) - len(group_means)
MSB = SSB / df_between
MSW = SSW / df_within

F_statistic = MSB / MSW
```

```
p_value = stats.f.sf(F_statistic, df_between, df_within)
print(f"ANOVA result: F-statistic = {F_statistic}, p-value = {p_value}")
```

```
print("Group Means:\n", group_means)
print("\nOverall Mean:", overall_mean)
print("\nSSB:", SSB)
print("\nSSW:", SSW)
print("\nF_stats: ", F_statistic)
print("\np_value: ", p_value)
print("\ndf_between:", df_between)
print("\ndf_within:", df_within)
```

Plot the ANOVA graph

```
plt.figure(figsize=(12, 8))
sns.boxplot(x='State', y='Functional Sub Centres (UOM:Number)', data=data)
plt.xticks(rotation=90)
plt.title('ANOVA: Functional Sub Centres across Different States')
plt.xlabel('State')
plt.ylabel('Number of Functional Sub Centres')
plt.tight_layout()
plt.show()
```

CORRELATION:

```
import pandas as pd
```

```
file_path = '/content/drive/MyDrive/health_expenditure.csv'
data = pd.read_csv(file_path)
```

```
columns_of_interest = [
    'Medical And Public Expenditure On Health By Urban Health Services ( Uhs ) (UOM:INR(IndianRupees))',
    'Medical And Public Expenditure On Health By Rural Health Services ( Rhs ) (UOM:INR(IndianRupees))'
```



```
]
```

```
data = data.rename(columns={
    columns_of_interest[0]: "Urban Health Services (Uhs)",
    columns_of_interest[1]: "Rural Health Services (Rhs)"
})
```

```
descriptive_stats = data[["Urban Health Services (Uhs)", "Rural Health Services (Rhs)"]].describe()
print(descriptive_stats)
```

Scatter plot

```
import matplotlib.pyplot as plt
import seaborn as sns
```

```
# Calculate the Pearson correlation coefficient
```

```
correlation_coefficient = data["Urban Health Services (Uhs)"].corr(data["Rural Health Services (Rhs)"])
```

```
plt.figure(figsize=(10, 6))
```

```
sns.scatterplot(
    x=data["Urban Health Services (Uhs)"],
    y=data["Rural Health Services (Rhs)"],
    color="blue", alpha=0.6
)
```

```
sns.regplot(
    x=data["Urban Health Services (Uhs)"],
    y=data["Rural Health Services (Rhs)"],
    scatter=False, color="red", line_kws={"linewidth": 2}
)
```

```
plt.title(f"Correlation between Urban and Rural Health Services Expenditure\n(Pearson r = {correlation_coefficient:.2f})", fontsize=14)
```

```
plt.xlabel("Urban Health Services (Uhs) [INR]", fontsize=12)
```

```
plt.ylabel("Rural Health Services (Rhs) [INR]", fontsize=12)
plt.grid(True, linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

TIME ANALYSIS

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import kendalltau

file_path = '/content/drive/MyDrive/health_expenditure.csv'
df = pd.read_csv(file_path)

cleaned_df = df[['Year',
    'Medical And Public Expenditure On Health By Rural Health Services ( Rhs )
    (UOM:INR(IndianRupees))']].dropna()

cleaned_df.columns = ['Year', 'Rural_Expenditure']
cleaned_df['Year'] = cleaned_df['Year'].str.extract(r'(\d{4})').astype(int)
aggregated_data = cleaned_df.groupby('Year')['Rural_Expenditure'].sum().reset_index()

plt.figure(figsize=(10, 6))
plt.plot(aggregated_data['Year'], aggregated_data['Rural_Expenditure'], marker='o', linestyle='-',
color='blue', label='Rural Expenditure')
plt.title('Time Series of Rural Health Expenditure', fontsize=14)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Expenditure (INR)', fontsize=12)
plt.grid(True)
plt.legend()

z = np.polyfit(aggregated_data['Year'], aggregated_data['Rural_Expenditure'], 1)
p = np.poly1d(z)
plt.plot(aggregated_data['Year'], p(aggregated_data['Year']), "r--", label='Trend Line')
```

```

plt.legend()
plt.tight_layout()
plt.show()

def mann_kendall_test(years, values):
    tau, p_value = kendalltau(years, values)
    return tau, p_value

tau, p_value = mann_kendall_test(aggregated_data['Year'], aggregated_data['Rural_Expenditure'])

print(f"Mann-Kendall tau: {tau}")
print(f"P-value: {p_value}")

alpha = 0.05
if p_value < alpha:
    print("Reject the null hypothesis (H0). There is a significant trend in the expenditure.")
else:
    print("Fail to reject the null hypothesis (H0). There is no significant trend in the expenditure.")

```

SARTHAK PANDIT

DESCRIPTIVE

```

import pandas as pd

# Read data from CSV
df_rural = pd.read_csv("water resources data .csv")

# Mean, Median, Mode
mean_rural = df_rural["Total Rural HHs with tap water supply as on 30.01.2024"].mean()
median_rural = df_rural["Total Rural HHs with tap water supply as on 30.01.2024"].median()
mode_rural = df_rural["Total Rural HHs with tap water supply as on 30.01.2024"].mode()[0]

```

```
# Output results
```

```
print(f'\nMean: {mean_rural}, Median: {median_rural}, Mode: {mode_rural}')
```

```
print("Output explanations:")
```

```
print(f'1. Mean: {mean_rural} - This indicates that, on average, the total number of rural households with tap water is approximately {mean_rural:.2f}.')
```

```
print(f'2. Median: {median_rural} - This is the middle value, indicating that half of the states have tap water access below this value and half are above.')
```

```
print(f'3. Mode: {mode_rural} - The most frequently occurring value, indicating that many states have around {mode_rural}.')
```

CORRELATION

```
import numpy as np
```

```
import matplotlib.pyplot as plt
```

```
import pandas as pd
```

```
df = pd.read_csv("water resources data .csv")
```

```
df["Total rural HH as on date"] = pd.to_numeric(df["Total rural HH as on date"], errors='coerce')
```

```
df["Total Rural HHs with tap water supply as on 30.01.2024"] = pd.to_numeric(df["Total Rural HHs with tap water supply as on 30.01.2024"], errors='coerce')
```

```
df.dropna(subset=["Total rural HH as on date", "Total Rural HHs with tap water supply as on 30.01.2024"], inplace=True)
```

```
xbar = df["Total rural HH as on date"].mean()
```

```
ybar = df["Total Rural HHs with tap water supply as on 30.01.2024"].mean()
```

```
X = df["Total rural HH as on date"].values.flatten()
```

```
Y = df["Total Rural HHs with tap water supply as on 30.01.2024"].values.flatten()
```

```
sxx = np.sum((X - xbar) ** 2)
```

```
sxy = np.sum((X - xbar) * (Y - ybar))
```

```
syy = np.sum((Y - ybar) ** 2)
```

```
m = sxy / sxx
```

```
c = ybar - m * xbar
```

```
print('Regression line: Y =', m, '* X +', c)
```

```
y_hat = m * X + c
```

```
sst = syy
```

```
ssr = np.sum((y_hat - ybar) ** 2)
```

```
sse = np.sum((Y - y_hat) ** 2)
```

```
R_sq = ssr / sst
```

```

R = sxy / np.sqrt(sxx * syy)

print("Sxx:", sxx)
print("Sxy:", sxy)
print("Syy:", syy)
print("SSE:", sse)
print("SST:", sst)
print("SSR:", ssr)
print("R^2:", R_sq)
print("R:", R)

print("x_bar:", xbar)
print("y_bar:", ybar)
print("Slope (m):", m)
print("Intercept (c):", c)
print("Regression Line: Y =", m, "* X +", c)

plt.scatter(X, Y, color='blue', label='Data points')
plt.plot(X, y_hat, color='red', label='Regression line')
plt.title('Total Rural HH vs Tap Water Access Comparison')
plt.xlabel("Total Rural HH")
plt.ylabel("Total Rural HHs with Tap Water")
plt.legend()
plt.grid(True)
plt.show()

```

REGRESSION

```

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

data = {
    "State": [
        "A&N", "Andhra Pr.", "Arunachal Pr.", "Assam", "Bihar", "Chhattisgarh",
        "Goa", "Gujarat", "Haryana", "Himachal Pr.", "J&K", "Jharkhand",
        "Karnataka", "Kerala", "Madhya Pr.", "Maharashtra", "Manipur", "Meghalaya",
        "Nagaland", "Odisha", "Puducherry", "Punjab", "Rajasthan", "Sikkim",
        "Tamil Nadu", "Telangana", "Tripura", "Uttar Pr.", "Uttarakhand", "West Bengal"
    ],
    "2019": [0.29, 30.74, 0.23, 1.11, 3.16, 3.20, 1.99, 65.16, 17.66, 7.63, 5.75, 3.45, 0.21, 16.64, 13.53,
48.44, 0.26, 0.05, 0.14, 3.11, 0.94, 16.79, 11.74, 0.70,
21.76, 15.68, 0.25, 5.16, 1.30, 2.15],
    "2020": [0.33, 12.77, 0.65, 5.07, 103.46, 1.51, 0.33, 10.95, 7.91, 3.79, 2.16, 3, 3.43, 4.04, 19.87,
37.09, 1.96, 0.87, 0.48, 15.47, 0.08, 8.18, 6.81, 0.1,
16.13, 18.2, 1.42, 19.16, 4.32, 12.48],

```

```

"2021": [None, 9.59, 0.42, 16.52, 19.86, 4.45, None, 8.99, 3.49, 2.87, 0.58, 4.19, None, 6.64, 10.89,
10.66, 0.7, 1.34, 0.92, 17.47, 0.07, 8.4, 5.38, 0.08,
14.9, -0.19, 1.65, 5.81, 2.75, 23.31],
"2022": [None, 12.12, 0.34, 8.75, 3.63, 10.81, None, 5.03, None, 0.93, 0.4, 8.77, 20.57, 5.29, 9.07,
8.22, 0.49, 0.8, 0.87, 12.22, None, 0.13, 13.88, 0.19,
26.52, 0.11, 0.81, 59.89, 2.24, 20.77],
"2023": [None, 3.07, 0.52, 19.09, 1.26, 17.38, None, None, None, None, 3.37, 10.22, 6.84, 3.4,
9.79, 12.57, 0.07, 1.72, 0.66, 10.01, None, None, 10.31, 0.09,
20.74, None, 1.08, 108.07, 1.7, 15.33],
"2024": [0.62, 69.43, 2.29, 51.03, 160.34, 38.3, 2.63, 91.18, 30.41, 17.09, 14.33, 30.58, 74.27,
36.87, 67.18, 122.42, 3.52, 4.8, 3.09, 62.66, 1.15, 34.26,
49.15, 1.17, 100.22, 53.98, 5.66, 202.74, 13.1, 74.08]
}

```

```

df = pd.DataFrame(data)
df.replace({None: np.nan}, inplace=True)
df = df.dropna(subset=["2019", "2020", "2021", "2022", "2023"])
years = np.array([2019, 2020, 2021, 2022, 2023]).reshape(-1, 1)
mean_values = df[["2019", "2020", "2021", "2022", "2023"]].mean(axis=0).values.reshape(-1, 1)

```

```

model = LinearRegression()
model.fit(years, mean_values)
y_hat = model.predict(years)
R_sq = model.score(years, mean_values)

```

```

predicted_2030 = model.predict(np.array([[2030]]))

```

```

print("Regression line: Y =", model.coef_[0][0], "x +", model.intercept_[0])
print("Predicted percentage for 2030:", predicted_2030[0][0])
print("The value of R^2 is:", R_sq)

```

```

plt.scatter(years, mean_values, color='blue', label='Historical Data')
plt.plot(years, y_hat, color='red', label='Regression Line')
plt.scatter(2030, predicted_2030, color='green', label='Predicted 2030')
plt.xlabel('Year')
plt.ylabel('Mean Percentage of Tap Water')
plt.title('Regression Analysis for Future Tap Water Access')
plt.legend()
plt.grid(True)
plt.show()

```

ANOVA

```

import pandas as pd
import statsmodels.api as sm
from statsmodels.formula.api import ols
import matplotlib.pyplot as plt

```

```

import seaborn as sns

# Provided data in a structured format
data = {
    'State': [
        'A&N', 'Andhra Pr.', 'Arunachal Pr.', 'Assam', 'Bihar', 'Chhattisgarh',
        'DNH and DD', 'Goa', 'Gujarat', 'Haryana', 'Himachal Pr.', 'J&K',
        'Jharkhand', 'Karnataka', 'Kerala', 'Ladakh', 'Lakshadweep', 'Madhya Pr.',
        'Maharashtra', 'Manipur', 'Meghalaya', 'Mizoram', 'Nagaland', 'Odisha',
        'Puducherry', 'Punjab', 'Rajasthan', 'Sikkim', 'Tamil Nadu', 'Telangana',
        'Tripura', 'Uttar Pr.', 'Uttarakhand', 'West Bengal'
    ],
    '21-22': [
        None, 12.77, 0.65, 5.07, 103.46, 1.51, None, 0.33, 10.95, 7.91,
        3.79, 2.16, 3.00, 3.43, 4.04, 0.02, None, 19.87, 37.09, 1.96,
        0.87, 0.34, 0.48, 15.47, 0.08, 8.18, 6.81, 0.10, 16.13, 18.20,
        1.42, 59.89, 2.24, 12.48
    ],
    '22-23': [
        None, 9.59, 0.42, 16.52, 19.86, 4.45, None, None, 8.99, 3.49,
        2.87, 0.58, 4.19, 18.70, 6.64, 0.09, None, 10.89, 10.66, 0.70,
        1.34, 0.21, 0.92, 17.47, 0.07, 8.40, 5.38, 0.08, 14.90, -0.19,
        1.65, 108.07, 2.24, 23.31
    ],
    '23-24': [
        0.62, 12.12, 0.34, 19.09, 1.26, 10.81, None, None, 5.03, None,
        0.93, 0.40, 8.77, 20.57, 5.29, 0.18, 0.11, 9.07, 8.22, 0.49,
        1.72, 0.22, 0.66, 10.01, None, 0.13, 13.88, 0.19, 26.52, None,
        0.81, 108.07, 1.70, 15.33
    ]
}

# Create a DataFrame
df = pd.DataFrame(data)

# Fill NaN values with 0 for better visualization (or choose to drop them)
df.fillna(0, inplace=True)

# Prepare the data for visualization
df_long = pd.melt(df, id_vars=['State'], value_vars=['21-22', '22-23', '23-24'],
                  var_name='Year', value_name='Tap_Water_Connections')

# Ensure Tap_Water_Connections is a numeric type
df_long['Tap_Water_Connections'] = pd.to_numeric(df_long['Tap_Water_Connections'],
errors='coerce')

# Drop any rows that still have NaN values after conversion
df_long = df_long.dropna()

```

```

# Conduct ANOVA on selected years
model = ols('Tap_Water_Connections ~ C(Year)', data=df_long).fit()
anova_result = sm.stats.anova_lm(model, typ=2)

# Calculate Mean Square for each source
anova_result['Mean Square'] = anova_result['sum_sq'] / anova_result['df']

# Print ANOVA results including Mean Squares
print("\nANOVA Results:\n")
print(anova_result)

# Determine whether to reject or fail to reject the null hypothesis
alpha = 0.05 # significance level
p_value = anova_result["PR(>F)"].iloc[0] # getting p-value for the main effect

# Conclusion
if p_value < alpha:
    conclusion = "Reject the null hypothesis: There are significant differences among the year means."
else:
    conclusion = "Fail to reject the null hypothesis: This indicates that the ANOVA test did not find a statistically significant difference in tap water connection means among the years."

# Print the conclusion
print("\nConclusion:")
print(conclusion)

# Visualization: Bar plot of Tap Water Connections by State for selected years
plt.figure(figsize=(15, 8))
sns.barplot(data=df_long, x='State', y='Tap_Water_Connections', hue='Year', errorbar=None) # Use errorbar instead of ci
plt.title('Tap Water Connections by State (2021-2024)', fontsize=16)
plt.xlabel('States', fontsize=14)
plt.ylabel('Tap Water Connections (Percentage)', fontsize=14)
plt.xticks(rotation=45) # Rotate x-axis labels for clarity
plt.legend(title='Year')
plt.grid(axis='y', linestyle='--')
plt.tight_layout() # Adjust layout
plt.show()

```

SIDDHARTH RANKA

DESCRIPTIVE

```
import pandas as pd
```



```

import matplotlib.pyplot as plt

data = {
    'State/UT': [
        "Chandigarh", "Delhi", "Haryana", "Himachal Pradesh", "Jammu & Kashmir", "Ladakh",
        "Punjab",
        "Rajasthan", "Uttarakhand", "Chhattisgarh", "Madhya Pradesh", "Uttar Pradesh", "Bihar",
        "Jharkhand",
        "Odisha", "West Bengal", "Arunachal Pradesh", "Assam", "Manipur", "Meghalaya", "Mizoram",
        "Nagaland",
        "Sikkim", "Tripura", "Dadra & Nagar Haveli and Daman & Diu", "Goa", "Gujarat",
        "Maharashtra",
        "Andaman and Nicobar Island", "Andhra Pradesh", "Karnataka", "Kerala", "Lakshadweep",
        "Puducherry",
        "Tamil Nadu", "Telangana"
    ],
    'Female_House': [
        30.4, 21.9, 38.6, 22.4, 56.6, 71.9, 63.2, 26, 23.8, 45.1, 38.9, 51.2,
        54.4, 63.6, 42.5, 22, 68.7, 42.2, 57.3, 64.1, 19.4, 25.5, 52.4, 15.8,
        55.8, 22.8, 42.2, 21.5, 14.5, 45.6, 66.2, 24.5, 29.7, 33.6, 47, 63.6
    ],
    'Male_House': [
        40.2, 37.2, 50.4, 52.8, 78.5, 74.6, 58.2, 55.9, 52.3, 71.7, 57.3, 65.5,
        67.8, 82.6, 73.1, 51.6, 76, 74.2, 76.8, 48, 50.4, 52.1, 75.6, 53.4, 44.1,
        21.2, 63.7, 44, 47.2, 67.1, 67.9, 53.6, 29.3, 47.9, 60.4, 74.3
    ]
}

df = pd.DataFrame(data)
fig, ax = plt.subplots(figsize=(10, 6))
plt.bar(df['State/UT'], df['Female_House'], label='Women', color='pink')
plt.bar(df['State/UT'], df['Male_House'], bottom=df['Female_House'], label='Men', color='blue',
alpha=0.7)
plt.title('House Ownership by Gender(alone or jointly) in India')
plt.xlabel('State/UT')
plt.ylabel('Percentage Ownership')
plt.xticks(rotation=90, fontsize=10)

```

```

plt.legend()
plt.tight_layout()
plt.show()

import pandas as pd
import matplotlib.pyplot as plt

data = {
    'State/UT': [
        "Chandigarh", "Delhi", "Haryana", "Himachal Pradesh", "Jammu & Kashmir", "Ladakh",
        "Punjab",
        "Rajasthan", "Uttarakhand", "Chhattisgarh", "Madhya Pradesh", "Uttar Pradesh", "Bihar",
        "Jharkhand",
        "Odisha", "West Bengal", "Arunachal Pradesh", "Assam", "Manipur", "Meghalaya", "Mizoram",
        "Nagaland",
        "Sikkim", "Tripura", "Dadra & Nagar Haveli and Daman & Diu", "Goa", "Gujarat",
        "Maharashtra",
        "Andaman and Nicobar Island", "Andhra Pradesh", "Karnataka", "Kerala", "Lakshadweep",
        "Puducherry",
        "Tamil Nadu", "Telangana"
    ],
    'Female_Land': [
        9, 12.7, 30.8, 20.3, 51.1, 63.8, 27.1, 20.7, 17.5, 38.8, 32.3, 42.7, 43.8, 54.6, 36.6, 16.7, 62.8, 34.8,
        24.8, 44,
        13.9, 15.8, 39.7, 10.7, 50.1, 9.2, 35, 14.7, 8.9, 24.5, 53.7, 11.5, 6.1, 10, 21.9, 42.6
    ],
    'Male_Land': [
        26.5, 21.1, 36.8, 48.2, 69.8, 70.7, 31.5, 43.3, 38.8, 66.3, 45.4, 55.8, 52.9, 73.7, 58, 33.6, 71.5,
        63.5, 44, 38.3,
        40.4, 41.4, 70.1, 34.8, 39, 5.2, 42.4, 27, 29, 34.8, 51.1, 25.1, 12.6, 7.2, 26.3, 51.1
    ]
}

df = pd.DataFrame(data)

fig, ax = plt.subplots(figsize=(10, 6))

plt.bar(df['State/UT'], df['Female_Land'], label='Women', color='pink')

plt.bar(df['State/UT'], df['Male_Land'], bottom=df['Female_Land'], label='Men', color='blue',
alpha=0.7)

```

```

plt.title('Land Ownership by Gender(alone or jointly) in India')
plt.xlabel('State/UT')
plt.ylabel('Percentage Ownership')
plt.xticks(rotation=90,fontsize=10)
plt.legend()
plt.tight_layout()
plt.show()

```

Z-TEST

```

import math

mu_male=42.42
mu_female=29.27
variance_male=273.68
variance_female=318.73
n=35
s_e=math.sqrt((variance_male+variance_female)/n)
z_score=(mu_female - mu_male)/s_e
z_alpha=-1.645
print("Mean (Men):",mu_male)
print("Mean (Women):",mu_female)
print("Sample Size (n):",n)
print("Standard Error:",round(s_e,4))
print("Z-Score:",round(z_score,3))
if z_score < z_alpha:
    print("Reject the null hypothesis.")
else:
    print("Fail to reject the null hypothesis.")

```

SINGLE PROPORTION TEST

```

import math

p_obs=29.27 / 100
p=0.50

```

```
n=35
s_e=math.sqrt((p * (1 - p)) / n)
z_score=(p_obs - p)/s_e
z_alpha=1.96
print("Observed Proportion (p_obs):", p_obs)
print("Hypothesized Proportion (p):", p)
print("Sample Size (n):", n)
print("Standard Error:", round(s_e, 4))
print("Z-Score:", round(z_score, 2))
if abs(z_score) > z_alpha:
    print("Conclusion: Reject the null hypothesis.")
else:
    print("Conclusion: Fail to reject the null hypothesis.")
```