

ECONOMETRICS-I

Course Project

**DOES RAINFALL AND PERCENTAGE
OF TOILETS HAVE AN EFFECT
ON PERCENTAGE OF CHILDREN
WITH MEASLES IN THE
AGE GROUP OF 0 TO 5 YEARS?**

GROUP NUMBER - 3

Presented by:

Aanya Trehan (2020419)

Amolika Bansal (2020424)

Diya Ahuja (2020431)

Pulkit Bazard (2020455)

Shagun Mohta (2020468)

Background

Measles, caused by Morbillivirus, is the 5th most lethal disease globally among the children aged between 0 to 5. In India, measles immunisation coverage is around 48~58.8%. Measles outbreaks generally occur in a pattern of every 3-4 years.

Highest percentage of children infected by Measles in India were found in the state of Uttar Pradesh.

PROCEDURE

- 1** Model Creation : We first theoretically researched and decided the independent variables in our model
- 2** Conducted Model Analysis
- 3** Researched intensively to select creative variables
- 4** Calculated Correlations
- 5** Found out the regression and interpreted the results
- 6** Came up with relevant policy suggestions

There have been several studies and researches done in order to factor out what causes Measles in India particularly. Few of them were:

1. Income
2. Health care facilities
3. drinking water
4. Pregnancy healthcare
5. sanitation facilities
6. Vulnerability to other diseases

Our chosen Variables

Upon researching measles, we decided to go ahead with 2 variables for our creative component-

1. State Wise Annual Rainfall
2. Percentage of Rural Households with access to toilets

Rainfall and Measles

It has been discovered that there has been a strong and consistent pattern of measles outbreaks that have been associated with rainfall.

There has been an inverse relation between occurrence of Rainfall and Measles.

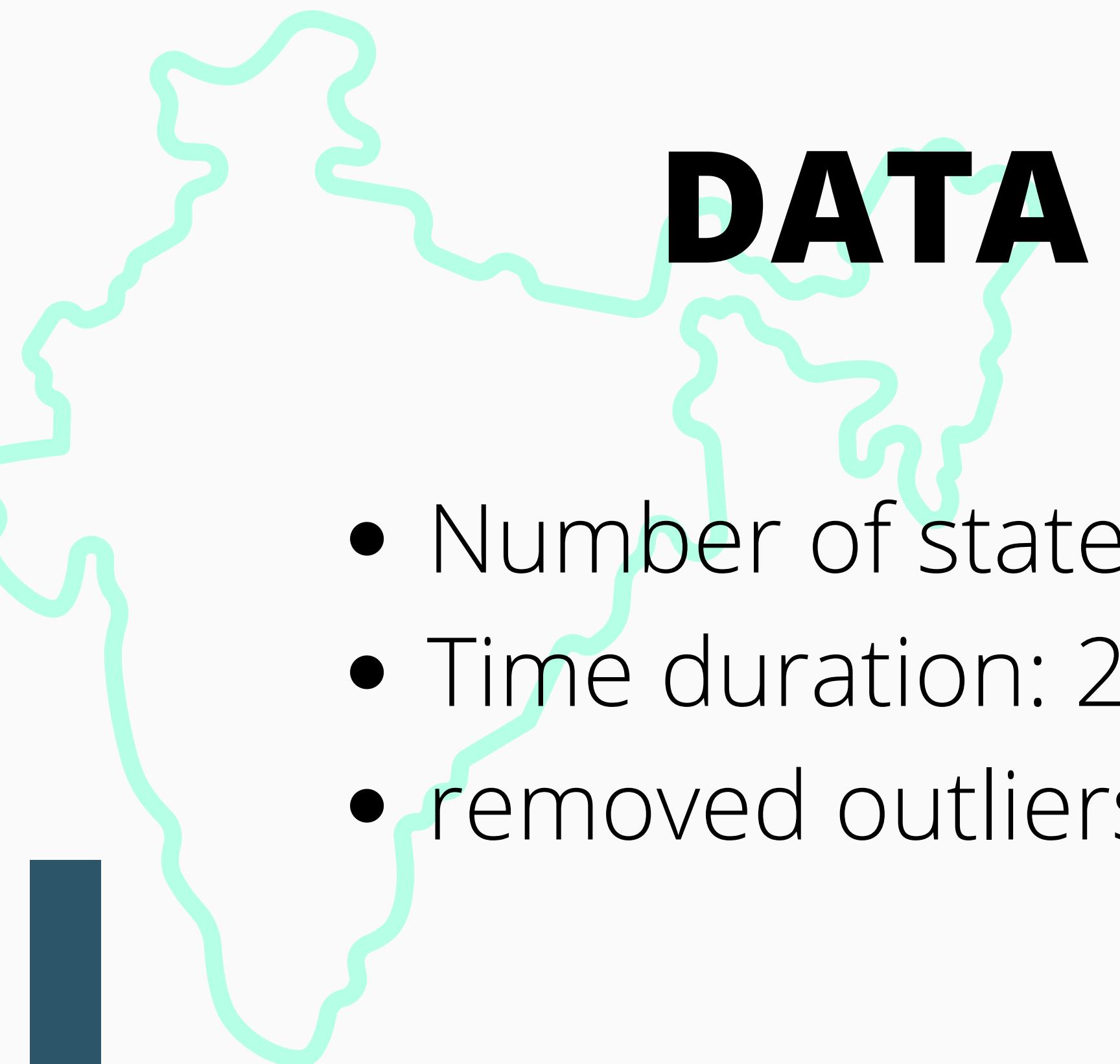
During the rainy season, number of reported Measles cases were less, however they were high in the dry season.

Sanitation Facilities and Measles

Bad sanitisation facilities such as lack of hygiene, immunisation being less, overcrowding etc increases the risk of measles transmission.

There is an inverse relation between Sanitization Facilities and Measles.

When the sanitisation facilities are good, number of reported Measles cases are less, however they are high when the sanitisation facilities are bad.



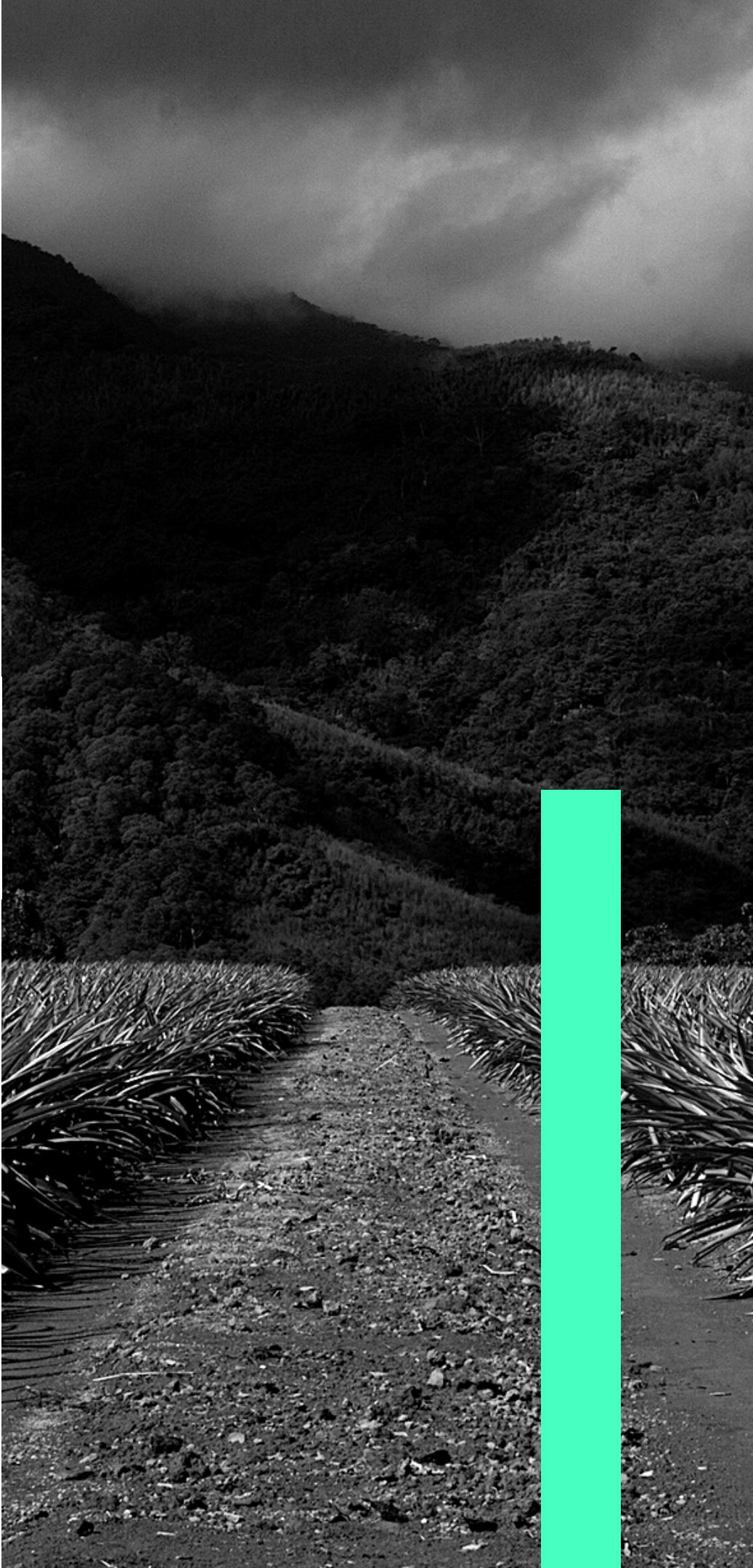
DATA UTILIZED



- Number of states and UTs used :35
- Time duration: 2011 to 2019
- removed outliers wherever necessary

OUR MODEL

$$H_{i,t} = \beta_0 + \beta_1(\log(gdp_{i,t})) + \beta_2(tap_{i,t}) + \beta_3(\log(beds_{i,t})) + \beta_4(\log(rain_{i,t})) + \beta_5(v37_{i,t}) \\ + \beta_6(v16_{i,t}) + \beta_7(v21_{i,t}) + \beta_8(v28_{i,t}) + \beta_9(v46_{i,t}) + \beta_{10}(\log(v34_{i,t})) + \beta_{10}(toilet_{i,t})$$



Hypothesis:

Our health indicator variable (percentage of children infected by measles age 0 to 5) is influenced by rainfall and toilet facilities factors.

Our Variables

VARIABLE	DESCRIPTION	REASON FOR CHOOSING
log(Gdp)	State Wise GDP	If the GDP of the state or district increases then generally it is assumed(because income inequality is not considered in GDP calculation) that the income of the citizens of that area also increases. Consequently increase in income leads to better access to healthcare and basic amenities of life and hence a lower chance of contracting measles.
tap	District Wise Tap Water Access (Percentage of Households) as of 2019	Measles is a highly contagious disease and hence can be easily spread among many people. Still, a huge part of India does not have access to safe and clean drinking water, which can prove to be very dangerous in case of a measles outbreak.

VARIABLE	DESCRIPTION	REASON FOR CHOOSING
log(beds)	State Wise Number of Hospital Beds (as of 2020)	In case of a measles outbreak, the number of beds in hospitals will increase to accommodate the rise in demand for hospital beds, similar to what had happened in covid as both are highly transmissible diseases.
log(v34)	Fully immunized children in the age group of 9 to 11 months	It has been observed that immunization has caused a drastic decline in measles. Fully immunized children in the age group of 9 to 11 months are at a much lower risk of contracting measles and hence affect the total percentage of children with measles in the age group pf 0 to 5 years.

VARIABLE	DESCRIPTION	REASON FOR CHOOSING
v16	Percentage of safe deliveries (to total reported deliveries)	An unsafe delivery typically depicts a weaker immunity in the newborn and a higher chance of contracting infections such as measles. The complications can also be expected to be much worse in the case of a newborn with unsafe delivery.
v37	Percentage of children with Diarrhea and Dehydration in the age group of 0 to 5 years	Diarrhea and dehydration are a few of the symptoms of measles and are generally categorized as complications associated with the disease, specifically in children under the age of 5 years. Thus, the percentage of children with measles is correlated with the percentage of children with diarrhea and dehydration.

VARIABLE	DESCRIPTION	REASON FOR CHOOSING
v28	Percentage of newborns having weight less than 2.5 kg	Children born with low birth weight, that is less than 2.5 kg are at a significant risk of developing diseases like measles in upcoming years
v46	Percentage of infant deaths due to Measles (to total reported infant deaths)	Measles can be fatal in people of all ages. However, complications are more likely in children under the age of five and adults over the age of twenty. Ear infections and diarrhea are common complications. Pneumonia and encephalitis are serious complications.

VARIABLE	DESCRIPTION	REASON FOR CHOOSING
v21	Percentage of women received a postpartum checkup or Post-Natal Care between 48 hours to 14 days of delivery.	After birth, women and infants require assistance and close supervision. Mothers usually take their newborns along with them to their postpartum checkups. According to WHO, the majority of mother and baby fatalities occur within the first six weeks following delivery.

REGRESSION WITH ANNUAL RAINFALL AND TOILET ACCESS

```
lm(formula = v36 ~ lgdp + tap + lbeds + v37 + v16 + v21 + v28 +  
v46 + lv34 + toilet + lr, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-23.031	-1.895	-0.356	1.229	59.402

REGRESSION WITH ANNUAL RAINFALL AND TOILET ACCESS

COEFFICIENTS	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	19.6125396	1.1320118	17.325	< 2e-16*
lgdp	-2.5109970	0.0816572	-30.750	<2e-16*
lbeds	2.6183175	0.0690881	37.898	< 2e-16*
v37	-0.2015847	0.0017321	-116.383	< 2e-16*
v16	0.0239719	0.0024446	9.806	< 2e-16*

REGRESSION WITH ANNUAL RAINFALL AND TOILET ACCESS

COEFFICIENTS	Estimate	Std. Error	t value	Pr(> t)
tap	-0.0071169	0.0012547	-5.672	1.42e-08 *
v21	-0.0044063	0.0009109	-4.837	1.32e-06 *
v28	-0.0623059	0.0030458	-20.457	< 2e-16*
v46	0.5820919	0.0081453	71.464	< 2e-16*
lv34	0.5696952	0.0365827	15.573	< 2e-16 *

REGRESSION WITH ANNUAL RAINFALL AND TOILET ACCESS

COEFFICIENTS	Estimate	Std. Error	t value	Pr(> t)
toilet	0.1158362	0.0052936	21.882	< 2e-16 *
lr	-0.2994553	0.0538399	-5.562	2.68e-08 *

REGRESSION WITH ANNUAL RAINFALL AND TOILET ACCESS

Signif. codes: 0 '*' 0.001 '*' 0.01 " 0.05 '.' 0.1 '' 1

Residual standard error: 5.246 on 41248 degrees of freedom
(29312 observations deleted due to missingness)

Multiple R-squared: 0.3993, Adjusted R-squared: 0.3991
F-statistic: 2492 on 11 and 41248 DF, p-value: < 2.2e-16

REGRESSION WITHOUT ANNUAL RAINFALL AND TOILET ACCESS

```
lm(formula = v36 ~ lgdp + tap + lbeds + v37 + v16 + v21 + v28 +  
v46 + lv34, data = data)
```

Residuals:

Min	1Q	Median	3Q	Max
-22.032	-1.852	-0.338	1.195	60.014

REGRESSION WITHOUT ANNUAL RAINFALL AND TOILET ACCESS

COEFFICIENTS	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	20.2201572	0.4376708	46.199	< 2e-16*
lgdp	-1.4752821	0.0446741	-33.023	<2e-16*
lbeds	2.1007580	0.0456621	46.007	< 2e-16*
v37	-0.1965074	0.0016688	-117.752	< 2e-16*
v16	0.0187042	0.0024073	7.770	8.04e-15*

REGRESSION WITHOUT ANNUAL RAINFALL AND TOILET ACCESS

COEFFICIENTS	Estimate	Std. Error	t value	Pr(> t)
tap	-0.0063062	0.0011298	-5.581	2.40e-08*
v21	-0.0110456	0.0008228	-13.425	< 2e-16*
v28	-0.0686885	0.0029594	-23.211	< 2e-16*
v46	0.5913151	0.0081631	72.438	< 2e-16*
lv34	0.2480585	0.0342240	7.248	4.30e-13*

REGRESSION WITHOUT ANNUAL RAINFALL AND TOILET ACCESS

Signif. codes: 0 '*' 0.001 '*' 0.01 " 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.262 on 42496 degrees of freedom
(28066 observations deleted due to missingness)

Multiple R-squared: 0.3826, Adjusted R-squared: 0.3825
F-statistic: 2926 on 9 and 42496 DF, p-value: < 2.2e-16

DATA SUMMARY

VARIABLES	mean	median	SD	MIN	MAX
gdp	55703048	48530154	42248128	78256	203331431
tap	20.31986	8.3	24.97515	0	100
beds	105389.4	64939	97335.42	1294	281402
v37	87.73024	95.5	18.1507	0	100
v16	91.45482	96.7	11.79509	21.2	100
v21	58.32384	53.6	35.50396	0	416

DATA SUMMARY

VARIABLES	mean	median	SD	MIN	MAX
v28	12.84502	11.1	8.941918	0	106
v46	0.2002345	0	3.134906	0	100
v34	34691.11	27213	28979.83	0	193100
toilet	98.13556	100	5.19903	75.71	100
rain	1299.712	1043.4	906.5147	302.4	4229.4

Null Hypothesis:

Our health indicator variable (percentage of children infected by measles age 0 to 5) is not influenced by the factors of rainfall and toilet facilities.

Findings on the Null Hypothesis->

F test-

ANNUAL RAINFALL

According to the ANOVA table, the value corresponding to the Df and residual Df values given here is 2.7055, less than the F value (=478.43) here. Thus, we can reject the null hypothesis that our health indicator variable(percentage of children infected by measles age 0 to 5) is not influenced by rainfall.

TOILET ACCESS

According to the ANOVA table, the value corresponding to the Df and residual Df values given here is 2.7055, less than the F value (=30.935) here. Thus, we can reject the null hypothesis that our health indicator variable(percentage of children infected by measles age 0 to 5) is not influenced by number of rural households with access to toilets.

Findings on the Null Hypothesis->

T test-

ANNUAL RAINFALL

As can be seen above the p-value (2.68e-08) of rainfall is less than alpha=0.05 hence we can reject the null hypothesis of the test. Therefore, rainfall influences our health indicator variable(percentage of children infected by measles age 0 to 5).

TOILET ACCESS

As can be seen above the p-value (2.68e-08) of rainfall is less than alpha=0.05 hence we can reject the null hypothesis of the test. Therefore, rainfall influences our health indicator variable(percentage of children infected by measles age 0 to 5).

Our null hypothesis is disproved

CONCLUSION

- From our results, we conclude that a significant relationship between rainfall in cm and percentage of children in age 0 to 5 infected with measles was detected.
- A significant relationship between the percentage of rural households with access to toilets in cm and the percentage of children aged 0 to 5 infected with measles was detected.
- The coefficient found for rainfall suggests a negative relationship with our dependent variable.
- The coefficient found for the percentage of access to toilets has a positive relationship with our dependent variable.
- Data inspection shows that areas with higher rainfall have fewer cases of measles
- The coefficient for rainfall acts as suspected while the one for access to toilets came out to be different than what we thought.
- There is a positive relationship between measles and rainfall because of the poor conditions of most toilets in rural households in India. These increase the risk of infections and hence it positively affects the number of people infected with measles in the age group of 0-5 years.

Policy Suggestions

We recommend that the policymakers and the various stakeholders must work upon:

1. Increased immunisation of children in the areas with poor facilities.
2. 2nd dose to be introduced to children between the ages of 5 to 17 years.
3. Better educational programs for social and behavioural change in the worst-hit areas.
4. access to health care facilities needs to be improved in the districts with a lower number of beds.

BIBLIOGRAPHY

- 1 [https://www.spring-nutrition.org/sites/default/files/understanding the linkages between agriculture and health-ifpri 2006.pdf](https://www.spring-nutrition.org/sites/default/files/understanding_the_linkages_between_agriculture_and_health-ifpri_2006.pdf)
- 2 <https://www.statology.org/r-guides/>
- 3 <https://statisticsbyjim.com/regression/interpret-coefficients-p-values-regression/#:~:text=The%20sign%20of%20a%20regression,variable%20also%20tends%20to%20increase.>
- 4 Gupta, S. N., Vidya, R., Gupta, N., & Gupte, M. D. (2011). Factors precipitating outbreaks of measles in district Kangra of North India: A case-control study. International journal of applied & basic medical research, 1(1), 24–30.
<https://doi.org/10.4103/2229-516X.81976>
- 5 <https://www.sciencedaily.com/releases/2020/08/200825110805.htm#:~:text=The%20team%20discovered%20a%20strong,higher%20during%20the%20dry%20season.>



THANK YOU!