Impact of Women on Country's Development

Team E

- Aishwarya Vijay Saswade
- Anusha Kokkinti
- Neha Agarwal
- Tejaswini Mummadi

Introduction:

United Nations Secretory-General Antanio Guterres stressed on Women's empowerment are "essential to global progress".

The empowerment and women autonomy of women and the improvement of their political, social, economic and health status is highly important for growth of the country.

Data Directory:

Country's Economic Status (Developed/ Developing)

Education
1. Primary Education
2. Secondary School Education
3. Tertiary School Enrollment

Employment
1. Employment in Industry
2. Employment in Agriculture
3. Self-Employment
4. Women in Parliament

Health
1. Infant Mortality

Development Indicators
1. Gender Development Index (GDI)
2. Human Development Index (HDI)
3. Urbanization
4. Income group
5. GDP per capita

EDUCATION

Research Question

- **How the Women Education Impacting the "Development" Status of the Country?**

Model used: Logistic regression

Dependent Variable: Status [Developed, Developing]

Independent Variables: wschoolenrolprimary, wschoolenrolsecondary, wschoolenrolltertiary

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	195.138	112.243
SC	198.156	124.312
-2 Log L	193.138	104.243

R-Square	0.4450	Max-rescaled R-Square	0.6165

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	88.8954	3	<.0001
Score	71.3566	3	<.0001
Wald	38.8688	3	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.6525	0.4430	13.9148	0.0002
wSchool_Enr_Tertiary	1	0.0466	0.0115	16.4555	<.0001
wSchool_Enr_Secondar	1	0.0363	0.0126	8.2593	0.0041
wSchool_Enr_Primary	1	-0.0454	0.0125	13.0967	0.0003

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
wSchool_Enr_Tertiary	1.048	1.024	1.072
wSchool_Enr_Secondar	1.037	1.012	1.063
wSchool_Enr_Primary	0.956	0.932	0.979

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	89.1	Somers' D	0.809
Percent Discordant	8.2	Gamma	0.832
Percent Tied	2.7	Tau-a	0.365
Pairs	5100	c	0.905

Interpretation

- From the results, -2 log L with the parameters (104.243) is significantly less than the model with only intercepts (193.138).
- The Wald's Chi-Square test static's p-value is less than 0.05. Hence the model is significant .
- The log odds of country to gain a Developed status is decreasing by 44% for a percent increase in the women's enrollment in the primary level schools. (Hence If women is just having primary education it doesn't make any difference, to dig deeper we ran separately (with just primary enrollment parameter) and found it's not a significant variable.
- The log odds of country to gain a Developed status is Increasing by 37% for a percent increase in the women's enrollment in the secondary level schools.
- The log odds of country to gain a Developed status is Increasing by 48% for a percent increase in the women's enrollment in the tertiary level schools.
- C = 0.905, means around 90% rows correctly predicted the Status of the Countries.

Inference

Country Status whether being a Developed / developing is significantly impacted by % of women enrollment in schools for higher level of education. Encouraging women to attain progress towards higher levels of education will have positive results on country’s growth.

Research Question

- **Is Infant mortality rate different for different regions?**

Hypothesis:

H0: mu1 = mu2 = mu3 = mu4...

H1: mu1 != mu2 or mu2 !=mu3 or mu3!=mu4 ….

Model Used: ANOVA

Dependent Variable: Female_InfanT_mortality_2018 Female_InfanT_mortality_2018

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	33740.60428	3748.95603	36.66	<.0001
Error	138	14113.12247	102.26900		
Corrected Total	147	47853.72676			

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

Region	Female_InfanT_mortality_2018 LSMEAN	LSMEAN Number
0	2.6600000	1
1	3.9687500	2
2	13.3750000	3
3	17.1666667	4
4	29.8833333	5
5	6.5666667	6
6	14.0272727	7
7	4.0250000	8
8	11.8611111	9
9	42.4236842	10

Least Squares Means for effect Region
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Female_InfanT_mortality_2018

i/j	1	2	3	4	5	6	7	8	9	10
1		1.0000	0.1151	0.0169	<.0001	0.9998	0.0137	1.0000	0.1467	<.0001
2	1.0000		0.3133	0.0632	<.0001	1.0000	0.0837	1.0000	0.4143	<.0001
3	0.1151	0.3133		0.9976	0.0435	0.9890	1.0000	0.8454	1.0000	<.0001
4	0.0169	0.0632	0.9976		0.3425	0.8591	0.9987	0.4874	0.9554	<.0001
5	<.0001	<.0001	0.0435	0.3425		0.0440	0.0288	0.0045	0.0085	0.1392
6	0.9998	1.0000	0.9890	0.8591	0.0440		0.9715	1.0000	0.9978	<.0001
7	0.0137	0.0837	1.0000	0.9987	0.0288	0.9715		0.7217	0.9996	<.0001
8	1.0000	1.0000	0.8454	0.4874	0.0045	1.0000	0.7217		0.9248	<.0001
9	0.1467	0.4143	1.0000	0.9554	0.0085	0.9978	0.9996	0.9248		<.0001
10	<.0001	<.0001	<.0001	<.0001	0.1392	<.0001	<.0001	<.0001	<.0001	

Interpretation

- P-value < 0.0001 It means the model is significant. And there is at least one pair of region for which the infant mortality rate is significantly different.
- To know which regions have significantly different means we performed the Turkey HSD(Honestly significant different) test. It is clear from the table that there are many pairs of regions for which infant mortality rate is significantly different.
- Example P-value for Region 3 and Region 0 is 0.0169 which is smaller than 0.05 and hence for this pair of regions infant mortality rate is significantly different. Another pair of regions for which infant mortality rate is significantly different includes Region 4 & 0, Region 4 & 1, Region 4 & 2, Region 5 & 4, Region 6 & 0, Region 6 & 4, Region 7 & 4, Region 8 & 4, Region 9 & 0, Region 9 & 1, Region 9 & 2, Region 9 & 3, Region 9 & 5, Region 9 & 6, Region 9 & 7, Region 9 & 8 and Region 9 & 9.

Research Question

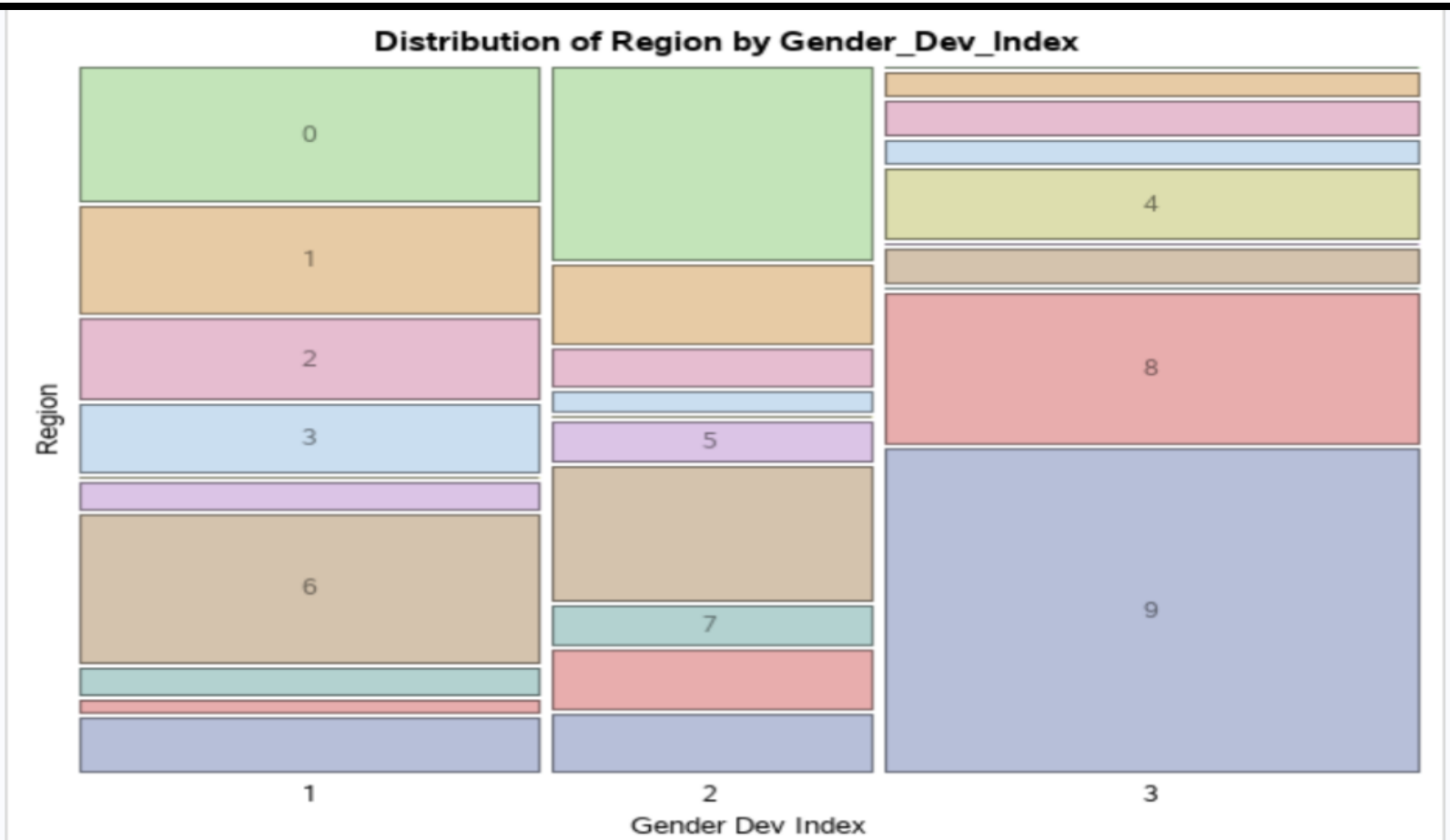
- **Is there any relationship between region and gender development index?**

Model Used: CHI SQUARE

Hypothesis:

H0: Region and GDI are independent

H1: Region and GDI are dependent



Statistics for Table of Region by Gender_Dev_Index

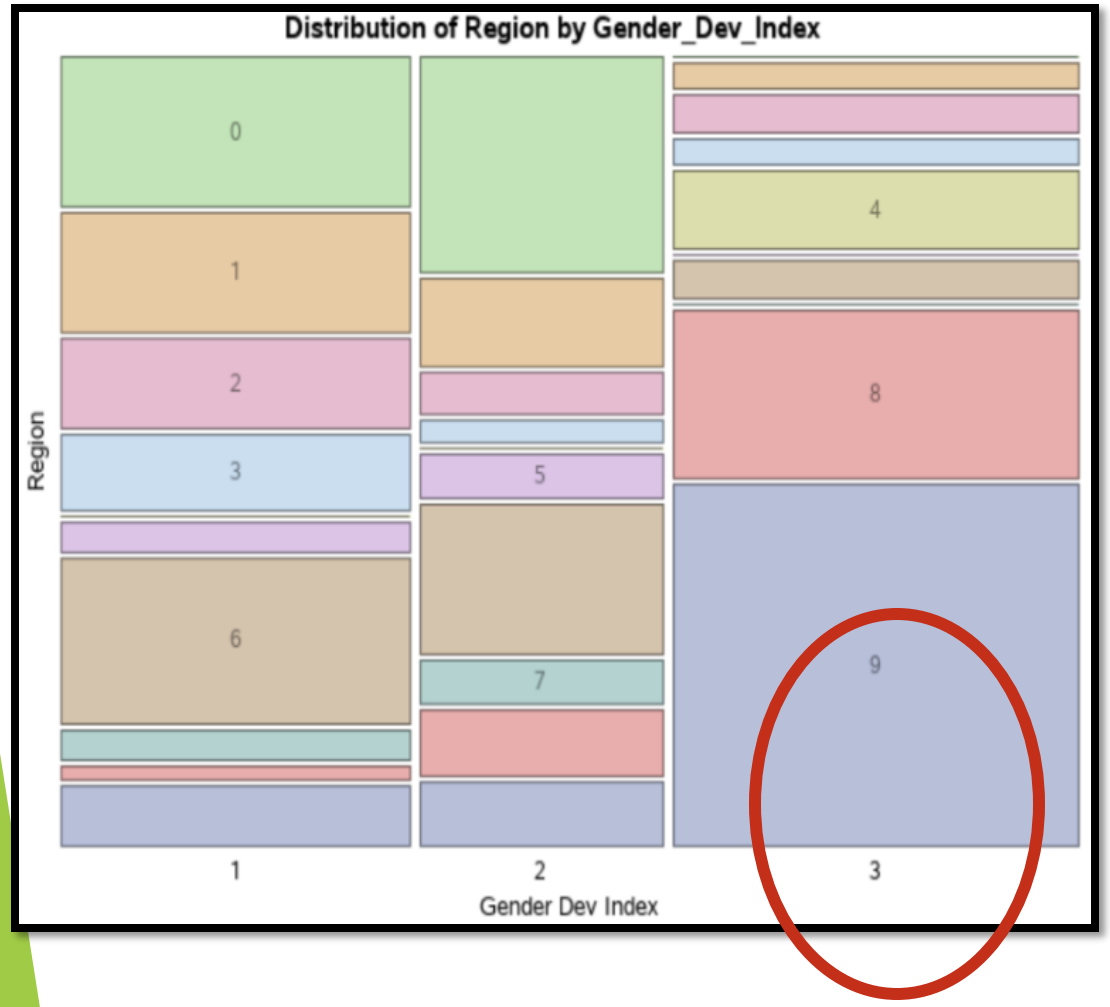
Statistic	DF	Value	Prob
Chi-Square	18	76.8286	<.0001
Likelihood Ratio Chi-Square	18	90.7733	<.0001
Mantel-Haenszel Chi-Square	1	33.5301	<.0001
Phi Coefficient		0.7408	
Contingency Coefficient		0.5953	
Cramer's V		0.5238	

Interpretation

- P-value < 0.0001 -> Reject Null Hypothesis
- It means we have enough evidence to support Alternative Hypothesis.
- Region & GDI are dependent.
- From above “Distribution of Region by GDI” figure it is clear GDI (=1) is high for Region 6, GDI (=2) is high for Region 0 and GDI (=3) is high for Region 9.

Inference

- It is evident that if GDI (=3, gender inequality is more) for certain region is high, the infant mortality rate for that region will also High. **Region 9 is sub Saharan Africa**



Least Squares Means for effect Region
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: Female_InfanT_mortality_2018

i/j	1	2	3	4	5	6	7	8	9	10
1		1.0000	0.1151	0.0169	<.0001	0.9998	0.0137	1.0000	0.1467	<.0001
2	1.0000		0.3133	0.0632	<.0001	1.0000	0.0837	1.0000	0.4143	<.0001
3	0.1151	0.3133		0.9976	0.0435	0.9890	1.0000	0.8454	1.0000	<.0001
4	0.0169	0.0632	0.9976		0.3425	0.8591	0.9987	0.4874	0.9554	<.0001
5	<.0001	<.0001	0.0435	0.3425		0.0440	0.0288	0.0045	0.0085	0.1392
6	0.9998	1.0000	0.9890	0.8591	0.0440		0.9715	1.0000	0.9978	<.0001
7	0.0137	0.0837	1.0000	0.9987	0.0288	0.9715		0.7217	0.9996	<.0001
8	1.0000	1.0000	0.8454	0.4874	0.0045	1.0000	0.7217		0.9248	<.0001
9	0.1467	0.4143	1.0000	0.9554	0.0085	0.9978	0.9996	0.9248		<.0001
10	<.0001	<.0001	<.0001	<.0001	0.1392	<.0001	<.0001	<.0001	<.0001	

EMPLOYMENT

Research Question

- **How is the impact of women's employment in areas like Industry, Agriculture, parliament on country's GDP Per Capita. (Country development indicator)**

- Model Used : Multi Linear Regression
- Dependent variable
  - GDP_Per_Capita
- Independent Variables:
  - % WomenEmploymentIndustry
  - % WomenEmploymentAgriculture
  - % WomenEmploymentParliament

Model: MODEL1
Dependent Variable: GDP_PER_CAPITA_2017 GDP_PER_CAPITA_2017

Number of Observations Read	1048462
Number of Observations Used	143
Number of Observations with Missing Values	1048319

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	27237349060	9079116353	42.24	<.0001
Error	139	29879136737	214957818		
Corrected Total	142	57116485797			

Root MSE	14661	R-Square	0.4769
Dependent Mean	14728	Adj R-Sq	0.4656
Coeff Var	99.55115		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	32496	4547.93005	7.15	<.0001
Employment_in_Industry	Employment in Industry	1	-1013.73815	194.28128	-5.22	<.0001
Employment_in_Agriculture	Employment in Agriculture	1	-492.23416	49.39411	-9.97	<.0001
newwp		1	299.30443	116.27439	2.57	0.0111

Interpretation

- As the f-stat p-value <0.05, model is significant
- A percent increase in Employment of Women in agriculture is Decreasing Country GDP per Capita by 492.2 dollars
- A percent increase in Employment of Women in Industry is Decreasing Country's GDP per Capita by –1013.738 dollars
- A percent increase of women's in parliament is increasing the country's GDP per Capita by 299.30 dollars'

Inference

- Increase in % of Women's in parliament will positively impact country's Economy.

Evaluation: According to International growth center research report, the constituencies that elect women experience significantly higher growth in economic activity through the electoral term than similar constituencies that elect men.

- Increase in % of women's employment in agriculture is not showing a positive impact on country's Economy.

Evaluation: According to FAO report on gendered analysis on agriculture trade, women are generally associated with non-economic and unpaid work. So, their contribution is not registered in the system of national accounts and therefore no market value is given to the labor involved.

Inference

- Increase in % of women's employment in Industry is not showing a positive impact on country's Economy.

Evaluation: Research on women by Catalyst in workforce states that Women's work force participation rate declining globally. Which means that women are more interested into the self-employed leading to negative impact on economy.

1 With Variables:	Wages_or_SalaryEmployed
1 Variables:	GDP_PER_CAPITA_2017

Pearson Correlation Coefficients Number of Observations	
	GDP_PER_CAPITA_2017
Wages_or_SalaryEmployed Wages or SalaryEmployed	0.61941 149

1 With Variables:	Self_Employed
1 Variables:	GDP_PER_CAPITA_2017

Pearson Correlation Coefficients Number of Observations	
	GDP_PER_CAPITA_2017
Self_Employed Self Employed	-0.61941 149

Research Question

- **Testing the moderation effect of self-employment in the urban population on country's economy.**

- Model Used : Moderation Analysis – Linear Regression
- First checking the relationship between Urbanization and GDP Per Capita, Self Employed and GDP Per capita
- Model Used : Linear Regression
- Dependent Variable:
    - GDP_Per_Capita
- Independent Variable:
    - Urban Population, WSelfEmployed

Model: MODEL1
Dependent Variable: GDP_PER_CAPITA_2017 GDP_PER_CAPITA_2017

Number of Observations Read	151
Number of Observations Used	149
Number of Observations with Missing Values	2

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	22000187857	22000187857	88.23	<.0001
Error	147	36654432725	249349882		
Corrected Total	148	58654620582			

Root MSE	15791	R-Square	0.3751
Dependent Mean	14678	Adj R-Sq	0.3708
Coeff Var	107.58116		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	-18320	3743.61295	-4.89	<.0001
urbanpopulation2017	urbanpopulation2017	1	543.54359	57.86631	9.39	<.0001

Model: MODEL1
Dependent Variable: GDP_PER_CAPITA_2017 GDP_PER_CAPITA_2017

Number of Observations Read	151
Number of Observations Used	149
Number of Observations with Missing Values	2

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	22503872157	22503872157	91.51	<.0001
Error	147	36150748425	245923459		
Corrected Total	148	58654620582			

Root MSE	15682	R-Square	0.3837
Dependent Mean	14678	Adj R-Sq	0.3795
Coeff Var	106.83944		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	30619	2104.15016	14.55	<.0001
wSelf_Employed	wSelf Employed	1	-390.17211	40.78755	-9.57	<.0001

Analysis

- Now checking the interaction effect of urbanization with Women being self-Employed as moderator on country's economy.

Least Squares Model (No Selection)

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	31840935628	10613645209	57.40	<.0001
Error	145	26813684955	184921965		
Corrected Total	148	58654620582			

Root MSE	13599
Dependent Mean	14678
R-Square	0.5429
Adj R-Sq	0.5334
AIC	2991.22656
AICC	2991.64614
SBC	2852.24234

Parameter Estimates

Parameter	DF	Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-23252	8025.822334	-2.90	0.0043
wSelf_Employed	1	364.660021	115.930771	3.15	0.0020
urbanpopulation2017	1	738.251853	104.969956	7.03	<.0001
wSelf_Emp*urbanpopul	1	-11.084598	1.925468	-5.76	<.0001

Inference

- The f-stat p value <0.05 , so the model is significant
- The p-value for the interaction effect between self-employment and urban population is < 0.05 and hence significant.
- So, it is evident that, though urbanization have positive impact on country's economy, it will be hampered, if there is increase in more of a self-employed woman.

Research Question

- **Is there an interaction effect between variables Employment-to-Population ration and Vulnerable Employment**

- Model used : Logistic Regression(Interaction Effect)
- Dependent Variable : Status
- Independent Variable : Employment-to-Population Ration, % of Vulnerable Employment women.
- Result : Model is significant

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	194.311	79.803
SC	197.321	88.835
-2 Log L	192.311	73.803

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	118.5078	2	<.0001
Score	81.2320	2	<.0001
Wald	29.6848	2	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.7368	0.6366	1.3395	0.2471
wEmployment_to_Pop_R	1	0.1228	0.0309	15.7816	<.0001
wVulnerable_Employme	1	-0.1355	0.0249	29.6505	<.0001

Odds Ratio Estimates			
Effect	Point Estimate	95% Wald Confidence Limits	
wEmployment_to_Pop_R	1.131	1.064	1.201
wVulnerable_Employme	0.873	0.832	0.917

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	96.0	Somers' D	0.920
Percent Discordant	4.0	Gamma	0.920
Percent Tied	0.0	Tau-a	0.416
Pairs	5049	c	0.960

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	194.311	70.154
SC	197.321	82.196
-2 Log L	192.311	62.154

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	130.1569	3	<.0001
Score	86.2725	3	<.0001
Wald	25.6824	3	<.0001

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-4.5411	1.6111	7.9443	0.0048
wEmployment_to_Pop_R	1	0.2920	0.0779	14.0363	0.0002
wVulnerable_Employme	1	0.0455	0.0536	0.7196	0.3963
wEmployme*wVulnerabl	1	-0.00726	0.00275	6.9754	0.0083

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	97.1	Somers' D	0.942
Percent Discordant	2.9	Gamma	0.942
Percent Tied	0.0	Tau-a	0.426
Pairs	5049	c	0.971

Inference

- The p-value corresponding to Wald's Chi-Square is <0.05
- The p-value corresponding to Interaction element's Estimate is <0.05
- C – value is 0.971 greater than model without interaction effect.
- So, It is evident that the Effect of Employment to population ratio on country's development is dependent on % of Vulnerable Employment(Women)

Research Question:

- **a. Is There any correlation between % of Vulnerable Employment and HDI**

- Model used: Correlation Analysis
- Variables used :
  - % of Vulnerable Employment women.
  - HDI(Human Development Index)
- Result : Higher Negative Correlation exists.

1 With Variables:	Human_Dev_Index
1 Variables:	wVulnerable_Employment

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations	
	wVulnerable_Employment
Human_Dev_Index Human Dev Index	-0.88023 <.0001 150

Research Question:

- **b. Is Vulnerable Employment across countries divided into regions by geographic area.**

- Model used: ANOVA
- Variables used :
  - % of Vulnerable Employment women.
  - regions
- Result : model is significant
- Region 9, 4 are significantly different from most of the regions.

  9 – Sub Saharan Africa

  4- south Asia
- Why?

Dependent Variable: wVulnerable_Employment wVulnerable Employment

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	99557.8092	11061.9788	30.59	<.0001
Error	139	50265.8884	361.6251		
Corrected Total	148	149823.6976			

Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

Region	wVulnerable_Employment LSMEAN	LSMEAN Number
0	8.2264999	1
1	14.5190626	2
2	29.7231664	3
3	49.1377782	4
4	75.6295001	5
5	26.4050002	6
6	37.6735452	7
7	7.9702501	8
8	19.5917777	9
9	75.1710264	10

Least Squares Means for effect Region
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: wVulnerable_Employment

i/j	1	2	3	4	5	6	7	8	9	10
1		0.9926	0.0697	<.0001	<.0001	0.7675	<.0001	1.0000	0.7089	<.0001
2	0.9926		0.5351	0.0010	<.0001	0.9821	0.0110	0.9998	0.9988	<.0001
3	0.0697	0.5351		0.3860	0.0002	1.0000	0.9764	0.6135	0.9159	<.0001
4	<.0001	0.0010	0.3860		0.2080	0.6079	0.8805	0.0154	0.0078	0.0114
5	<.0001	<.0001	0.0002	0.2080		0.0038	0.0011	<.0001	<.0001	1.0000
6	0.7675	0.9821	1.0000	0.6079	0.0038		0.9850	0.9340	0.9997	0.0001
7	<.0001	0.0110	0.9764	0.8805	0.0011	0.9850		0.1232	0.0917	<.0001
8	1.0000	0.9998	0.6135	0.0154	<.0001	0.9340	0.1232		0.9834	<.0001
9	0.7089	0.9988	0.9159	0.0078	<.0001	0.9997	0.0917	0.9834		<.0001
10	<.0001	<.0001	<.0001	0.0114	1.0000	0.0001	<.0001	<.0001	<.0001	

Interpretations

- The Regions 4, 9 are significantly different from other regions but both are not.
- Why?
- Is There any Relation with Gender Inequality?
- Yes.
  - More countries with higher gender inequality.

Inference

- Country's with more suppression towards Women have significant difference in their % of Vulnerable Employment.

Conclusion

- 1.Suppression of women through gender equality is impacting the country's development.
- This is tested through mortality rate and vulnerable employment.
- 2. Progress in women's education definitely has a positive impact on country's development.
- 3. Encouraging women for higher education levels might have positive impact in such cases.
- 4. While urbanization can be positive factor for country's economy, having more self-employed people lead to the negative impact on the country's economy .
- 5. More women in parliament should be encouraged for a country's positive development and gender equality has to be improved for the contribution of women in agriculture and industry to improve economy

THANK YOU!