An Analysis of Suicide in Western and Northern Europe and the United States.

Suicide rates have been dropping and perhaps religion is playing a role.

By Christian Miyares, Qing He

Table of Contents

Introduction.

Body

- <u>Data</u>
- <u>Methods</u>
- Analysis
- Results and Inferences

Conclusion

<u>Appendix</u>

- <u>Sources</u>
- Results
- Multiple Regression

Introduction

Suicide is the leading cause of death in many countries. It sits among other disease such as heart disease or stroke. However, the Center for Disease and Control Prevention (CDC) lists suicide or intentional self-harm as being in the leading causes of death for the United States. With the CDC listing self-harm so high on the list I wonder if suicide has increased in the US. The data I use provides data from 1985-2016. However, I want to explore the relationship between suicide and religion. The United States is the most religious western country, with many people practicing or believing in the Christianity. Although the US is heavy on religion, more people are identifying with irreligion in the United States. I want to compare countries in Western and Northern Europe that tend to have a greater population of irreligious people to the suicide rates over time as well as the United States.

Body

A. Data

The data I got about the suicide as from Kaggle and uploaded by a random user. However, the user pulled the data from the United Nations Development Program, World Bank, the World Health Organization and another Kaggle data set. All in all, the data seems to come from good sources and shows no signs of anomaly. The dataset is vast and provides most countries in the world. It also provides amount of suicide per country and per one hundred thousand people in that country. The "per one hundred thousand" variable is very useful in comparing countries with different populations. Some problems I had with the dataset were that although the data set claims to provide equal data from countries from the years 1985-2016 some countries did not have this data. This is most likely that the governments of those countries were not tracking suicide numbers during those years. However, due to that I had choose to not include certain countries. As an additional resource to the data set, I used data from Pew Research Center on importance of religion. I was not able to access the raw data from this but will be making inferences based on my conclusions with the suicide data. For the full data set and other sources please see the appendix.

B. Methods

My goal for my analysis was to first see how suicide rates trended whether they trended up or down over the years. Therefore, I decided to use Linear and Multiple Regression. The data set provided age ranges from 5-75+ however the variability and development of these age made it too great to include all of them in my analysis. Therefore, I decided to explore the group aged 35-54. This is one of the bigger groups at about 20 years. I believe at this age most adults have settled down and potentially have children or are married. I wanted to study this group because they seem the most interesting of all the ages. I began with splitting the groups into male and female as I believe different factors can affect them. I made female and male groups for each country. Those countries include the United States. Then I wanted to study other countries comparable to the United States in terms of wealth and culture, therefore I chose other western countries in Western Europe and Northern Europe. More specifically, in Western Europe, Austria, Belgium, Czech Republic, France, Germany, Ireland, Luxembourg, Netherlands, and Switzerland. In Northern Europe, Denmark, Finland, Iceland, Norway, and Sweden. I then

referenced the pew research center to make inferences on the religiosity of the countries mentioned above and determine if religion is influential in suicide. Finally, my hypothesis for the analysis is that I believe that countries with lower religiosity will tend to commit suicide less often especially as time goes on from 1985-2016.

C. Analysis

As mentioned in the methods I used the Pearson Correlation test in order to determine the strength of each individual male and female variable by country. Beginning with the United States males I got a 0.99 r or the linear correlation coefficient. "R measures the strength and the direction of a linear relationship between two variables". Therefore, as time goes on males aged 35-54 in the United States have been committing more and more suicide. In fact, it rose from 22 to over 24 cases per 100k in 30 years. The R-Square or the goodness-of-fit measure was 0.72 when performing simple linear regression. Females were similar with R of .98 and was lower at 0.49. However, a key difference is that although both are rising in suicide cases, the female group rose from 6.5 to 7.5 per 100k in 30 years. Next, I moved onto the Western European countries and performed that same two tests as performed on the US on both separate male and female groups. Beginning with Austrian females, we immediately get opposite results from those of the US. R is -0.96 and R-squared 0.92 making the model a strong negative linear correlation. Also, the Suicides dropped from about 17.5 to 6 per 100k people. Similarly, with Austrian males r=-0.95 and r-squared=0.8941. In order to avoid redundancy, I will now summarize the findings for the rest of the countries and will make note on anomalous ones. The rest of the simple linear models will be provided in the appendix. For the rest of the females in Western and Northern Europe they all trended relatively strong negatively (r < -0.5). There were a few anomalies. Irish females were trending negative, but the Pearson r coefficient was too small and close to zero to make any inferences. Iceland females also got a 0 r, and Norway and United Kingdom (UK) got negative but greater than -0.5. The male results were much more all over the place. Most of them trended negatively with some (Iceland, and Norway), trending negatively but with low r coefficient. Anomalous countries were Belgium Ireland Netherland and UK with all of them trending positively and a high r coefficient. Later, we will explore why these countries have high suicide rates for men but not women. Next, I performed multiple regression on all the countries. The variables I wished to compare were country and sex and their effects on suicides per 100k people. Then I made separate models using only males and only females. Starting with the first model mentioned. I used the stepwise function in order to find significant variables in my data. After removing insignificant variables, I ended up with the countries, Austria, Belgium, Czech Republic, Finland and France. All the countries trended down from 1985 to 2015. Similarly, for other two models I performed the same procedure using the stepwise setting. The female only model was better than the male only and the female only model also trended negatively. Finally, in the male it trended negatively but the starting suicides per 100k was much higher than the female and the combined male and female group.

D. Results/Inferences

From the analysis I learned a lot about different trends with suicide in terms of time. In it should be noted once again that this analysis refers to the group of people aged 34-54. In the United States suicide is on the rise for both males and females. This was confirmed by the CDC already, but we further confirmed it with this age group. Males are obviously more susceptible to suicide in terms of numbers. Although the female rate is rising in US it has risen a small amount compared to that of males. Moving onto Western Europe we see a completely opposite trend. Suicide rates are dropping in drastic amounts across the board. To use Austrian females for example in 1985 there were 17.5/100k suicides while in 2015 there were about 7.5k/100k. The model we made for Austrian females also shows a tight confidence interval which means the variables are closely related. As time passes suicide falls. Why is this happening? My hypothesis is that less religion importance in a country brings about less suicide. To begin this inference, the countries I am picking are generally wealthier countries. A Gallup poll from 2009 concludes that religiosity is highest in countries that are poorer. There is an exception to that however and that is the United States. More specifically I will list all the countries we tested by religion being unimportant. Sweden(83%), Denmark(81%), Czech Republic(75%), Norway(78%), UK(73%), Finland(70%), France(69%), The Netherland(67%), Belgium(58%), Luxembourg(59%), Germany(59%), Switerland(57%), Ireland(46%), US(31%). Unfortunately, there is not data I could find on Iceland. Starting with Sweden females and males. In both groups suicide rates have dropped since 1985 more for males than females. In general, all the rest of the female groups except for Ireland and US tend trend negatively. This is interesting as less than half of the total population both of those countries believe that religion is unimportant. For the males all the countries trend negative except for Belgium, Ireland, Netherlands, UK and US. Belgium is a strange statistic because it has an r=0.09 therefore cannot make any determinations about it. Ireland trends positive r=0.7 which is rather high while US trends positive as well with an even higher r=0.98. The UK, and Netherlands trend positive but less than r=0.5. My results for men are rather inconclusive. However, the countries with higher coefficients of correlation (Ireland and US) also tend to be more religious, whereas the ones with lower r coefficients although positive are less conclusive. Although we cannot say that religion is the main driver in suicide because of lack of data. We

Conclusion

countries tend to be more religious. This makes sense, since in a country that is more poor people will have less education and may more easily fall into a religious group whether it be for comfort

can infer that there is some relationship. Perhaps my theory is that the countries have higher standards of living where people tend to be happier and do not rely on a deity or group in order to find such happiness. As mentioned before, the Gallup poll on religion states that poorer

or for inclusion. The United States and Ireland although they are more wealthy countries, religion still plays a major influence. Although these a rich country briefly perhaps the people

within are not as wealthy as it may seem.

In general, discover many connections between suicide over time but some stuff remains in question. For example, men suicide rates did drop in a lot of countries but in others it is still trending upwards. We do not have data to know why such a thing is happening. Perhaps there are

factors that cannot even be measured and may be a host of things. Regardless the good news is that female suicide rates have fallen in almost every single country sampled. Some limitations on the analysis is the religious data. Although I have the percentage of people that say that religion is unimportant, we cannot know if there was equal number of males and females sampled here. Therefore, all in all I conclude that as a whole suicide has been falling in Western and Northern Europe for both sexes but rising in the United States. Also, I can conclude to an extent that religion maybe playing a role. Perhaps the United States and Ireland are going through a transformation where less and less people become religious and coping with such a change my lead to more suicide. Therefore, to explore such questions time will tell and we hope the rates continue to trend down not just in the west but in the whole world.

APPENDIX

A. Sources

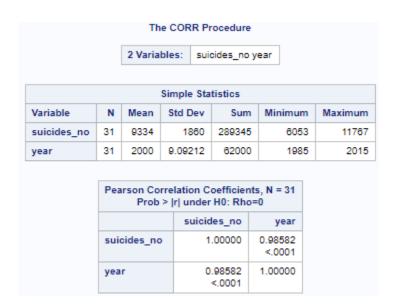
Crabtree, Steve. "Religiosity Highest in World's Poorest Nations." *Gallup.com*, Gallup, 18 Nov. 2020, news.gallup.com/poll/142727/religiosity-highest-world-poorest-nations.aspx.

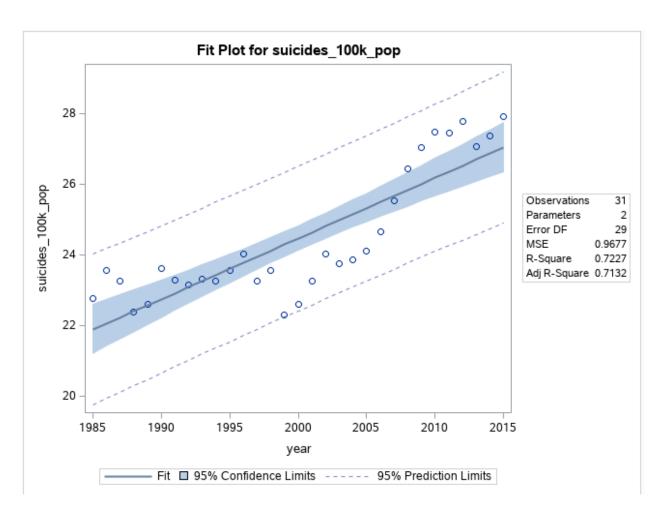
- "Importance of Religion by Country." *Wikipedia*, Wikimedia Foundation, 9 Nov. 2020, en.wikipedia.org/wiki/Importance_of_religion_by_country.
- Poushter, Jacob, and Janell Fetterolf. "How People around the World View Religion's Role in Their Countries." *Pew Research Center's Global Attitudes Project*, Pew Research Center, 30 Dec. 2019, www.pewresearch.org/global/2019/04/22/how-people-around-the-world-view-religions-role-in-their-countries/.
- GDP Ranked by Country 2020, worldpopulationreview.com/countries/countries-by-gdp.
- "FastStats Leading Causes of Death." *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 30 Oct. 2020, www.cdc.gov/nchs/fastats/leading-causes-of-death.htm.
- "In U.S., Decline of Christianity Continues at Rapid Pace." *Pew Research Center's Religion & Public Life Project*, 9 June 2020, www.pewforum.org/2019/10/17/in-u-s-decline-of-christianity-continues-at-rapid-pace/.

B. Graphs/Plots/Results

Correlation and Simple Linear Regression by Country and Sex

Male US:





Female US

The CORR Procedure 2 Variables: suicides_no year Simple Statistics Variable N Std Dev Sum Minimum Mean Maximum 633.46170 91572 suicides_no 31 2954 2105 31 2000 9.09212 62000 1985 2015 year Pearson Correlation Coefficients, N = 31 Prob > |r| under H0: Rho=0

suicides_no

1.00000

0.98064

<.0001

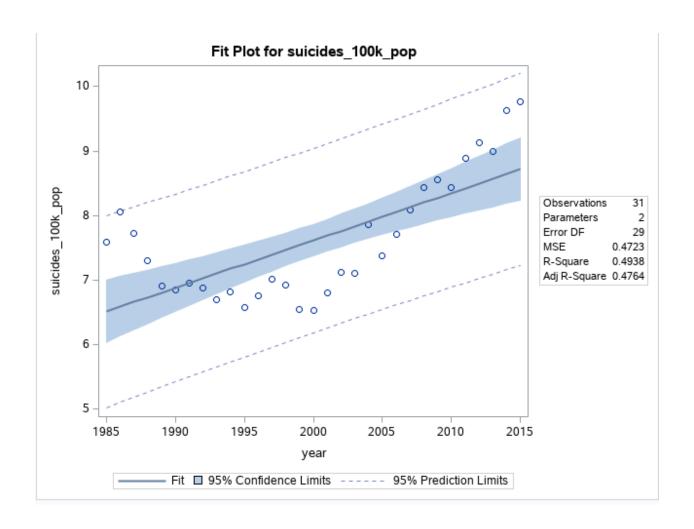
suicides_no

year

year

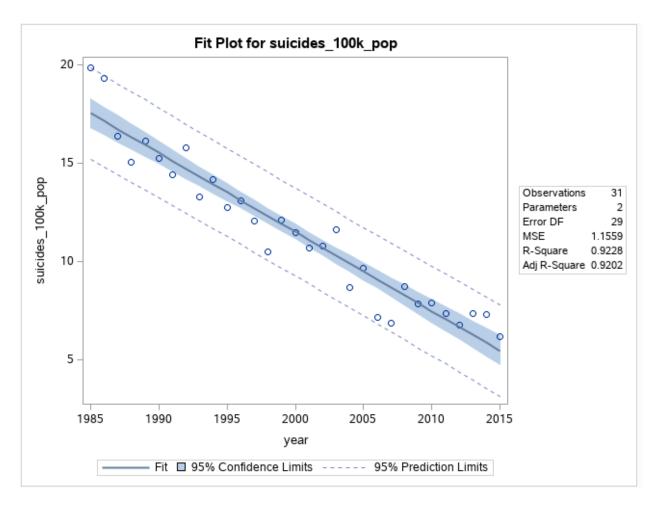
0.98064

1.00000

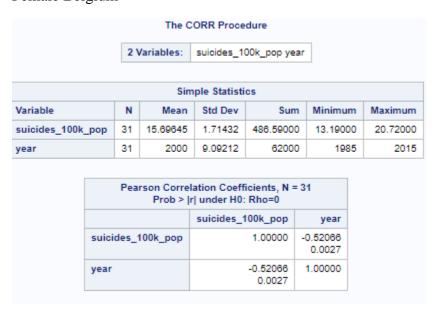


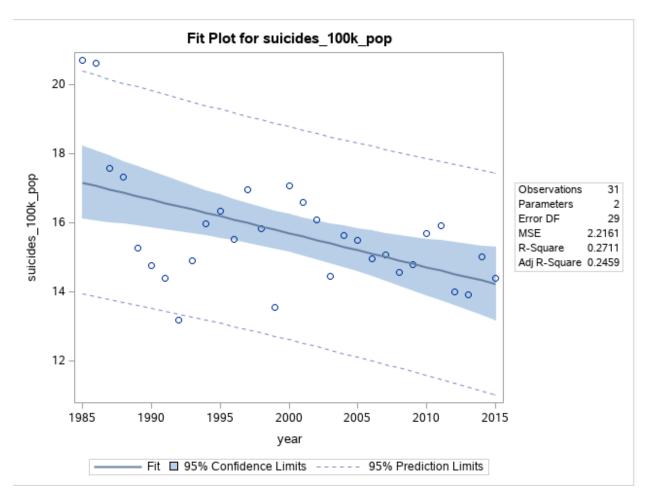
Female Austria

			The C	CORR Proce	dure		
		2١	/ariables:	suicides_1	00k_pop ye	ar	
			Sir	nple Statisti	ics		
Variable		N	Mean	Std Dev	Sun	n Minimu	m Maximum
suicides_100k_	рор	31	11.50387	3.80548	356.6200	0 6.2100	00 19.85000
year		31	2000	9.09212	6200	0 198	85 2015
		Pea		lation Coeff r under H0		= 31	
-		Pear			Rho=0	= 31 year	
-	suicio			r under H0	Rho=0		

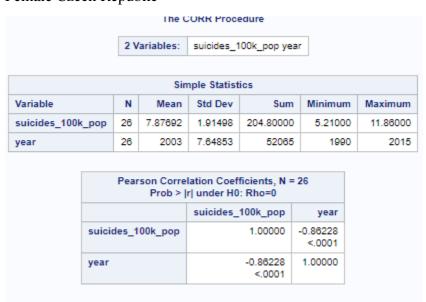


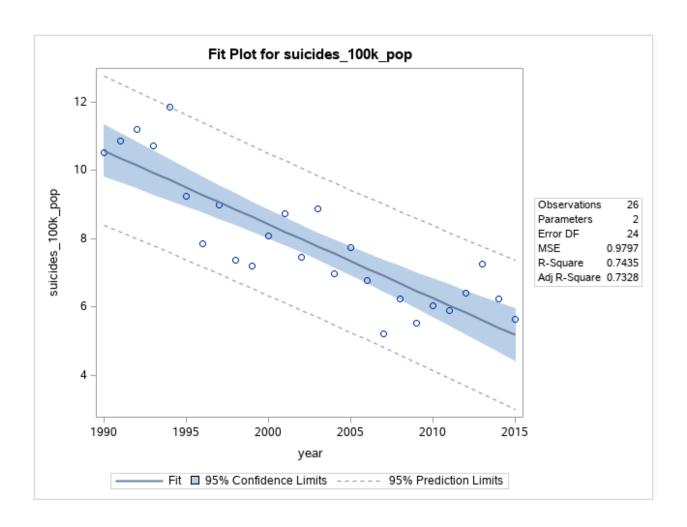
Female Belgium



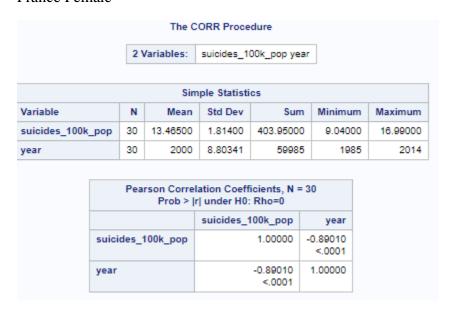


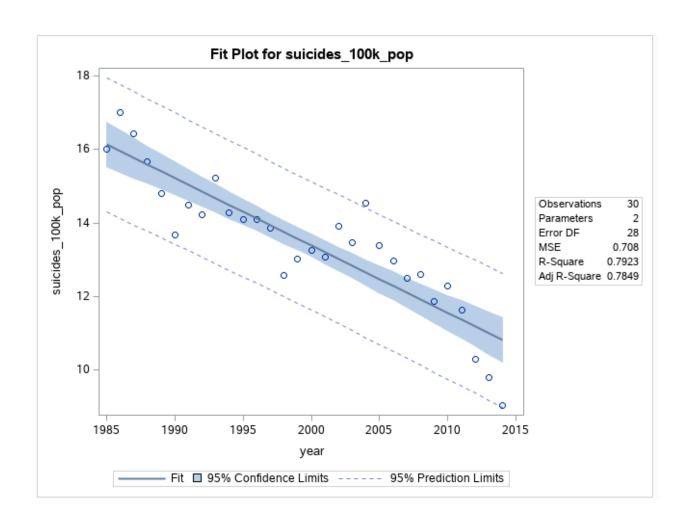
Female Czech Republic





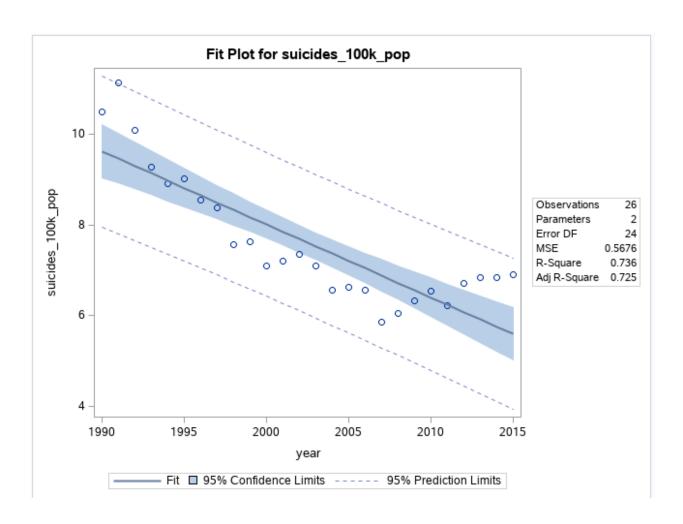
France Female



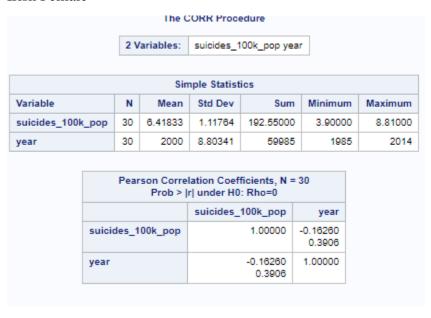


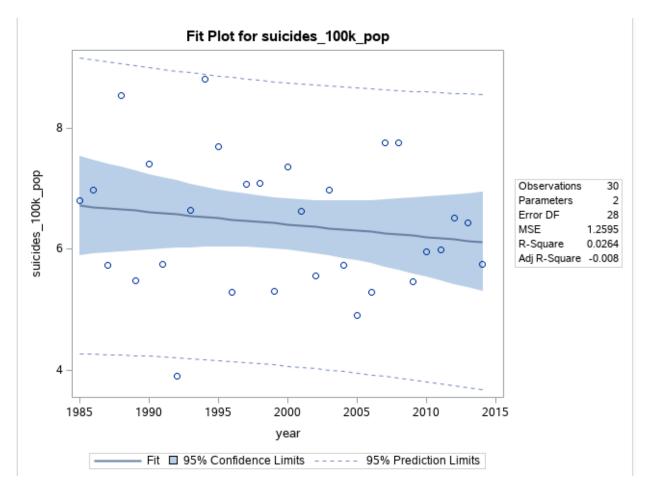
German Female

			The C	ORR Proce	edure			
		2 V	ariables:	suicides_1	00k_pop ye	ar		
			Sin	nple Statist	ics			
Variable		N	Mean	Std Dev	Sun	Minimu	m	Maximum
suicides_100k	_рор	26	7.60423	1.43681 197.71000 5.8600		5.8600	00	11.13000
year		26	2003	7.64853	52068	5 199	0	2015
		Pear		lation Coef r under H0		= 26		
				suicides_	100k_pop	year		
	suicio	des_1	00k_pop		1.00000	-0.85793 <.0001		
	year				-0.85793 <.0001	1.00000		

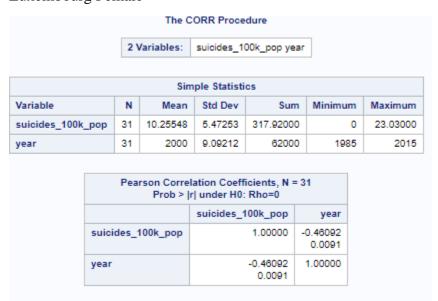


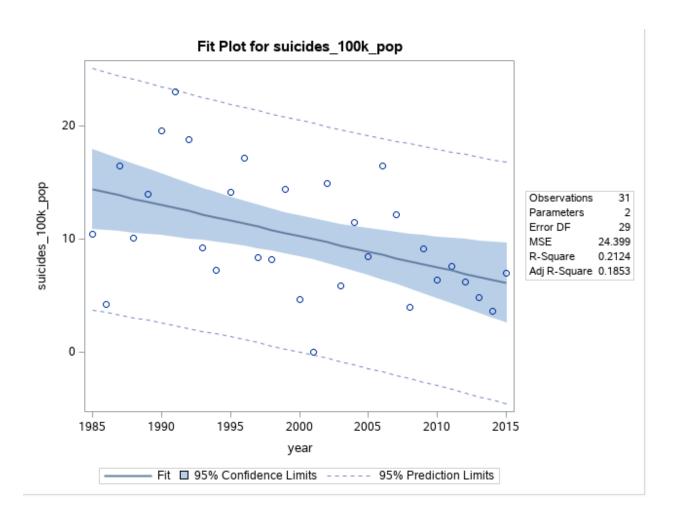
Irish Female





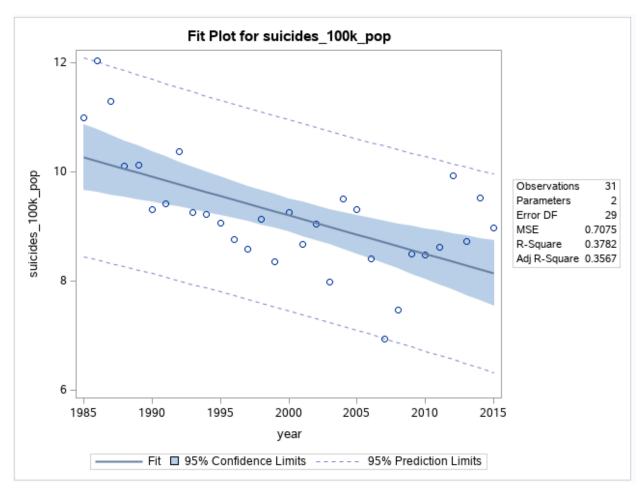
Luxembourg Female



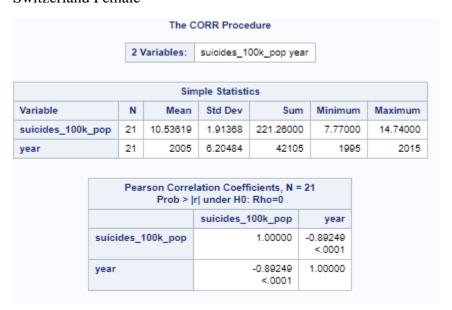


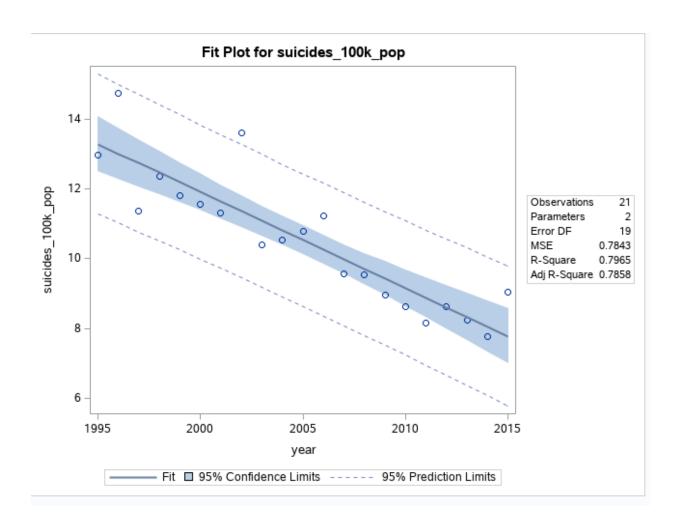
Netherlands Female

			rne C	The CORR Procedure								
		2 V	ariables:	suicides_1	00k_pop ye	ar						
			Sin	nple Statist	ics							
Variable		N	Mean	Std Dev	Sum	Minimum	Maximum					
suicides_100k	_рор	31	9.20968	1.04877	877 285.50000 6.9400		12.05000					
year		31	2000	9.09212	62000	1985	2015					
		Pear		lation Coef r under H0	ficients, N = : Rho=0	: 31						
				suicides_	100k_pop	year						
	suicid	les_1	00k_pop		1.00000	-0.61496 0.0002						
	year				-0.61496 0.0002	1.00000						



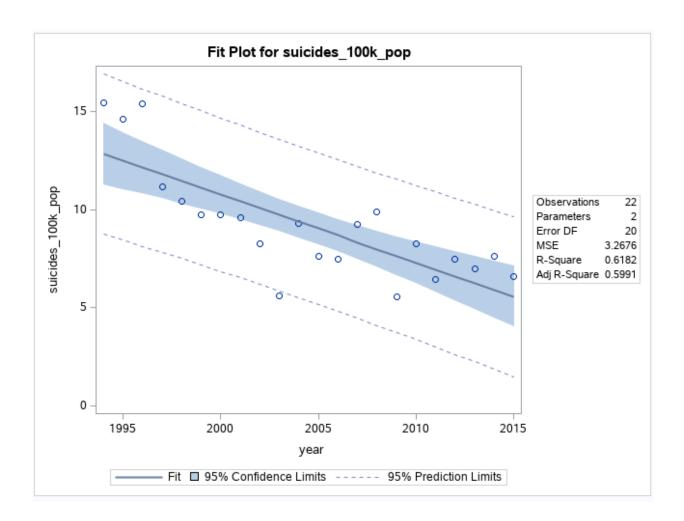
Switzerland Female





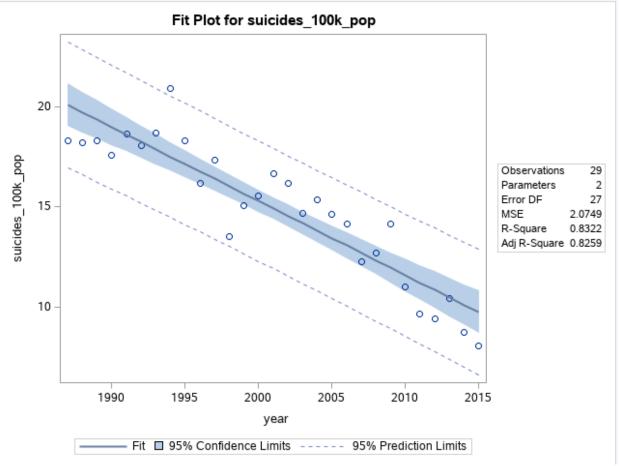
Denmark Female

		The C	ORR Proce	edure		
	2 V	/ariables:	suicides_1	100k_pop ye	ar	
		Sin	nple Statist	ics		
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
suicides_100k_po	22	9.19682	2.85503	202.33000	5.56000	15.44000
year	22	2005	6.49359	44099	1994	2015
		0	I-4: C4			
	Pear		r under H0	ficients, N = : Rho=0	: 22	
	Pear		r under H0		year	
sui			r under H0	: Rho=0		

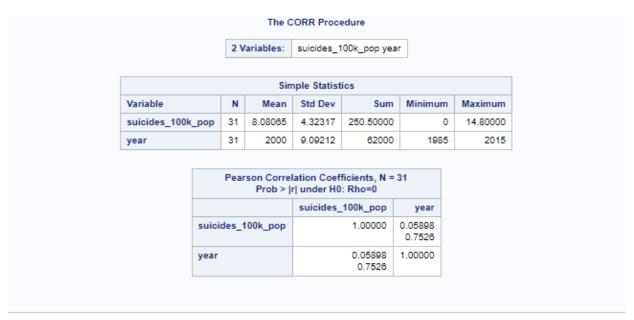


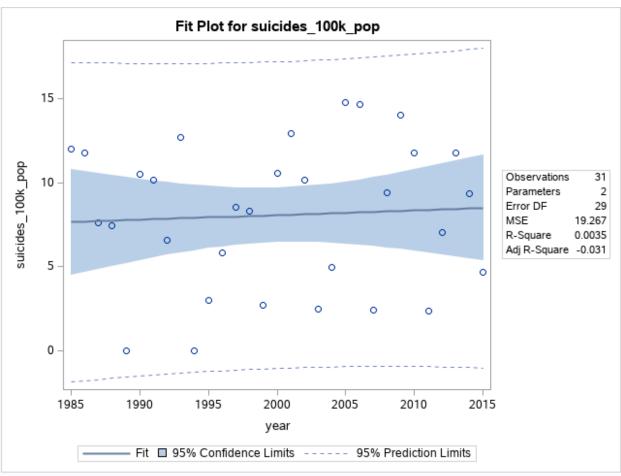
Finland Female



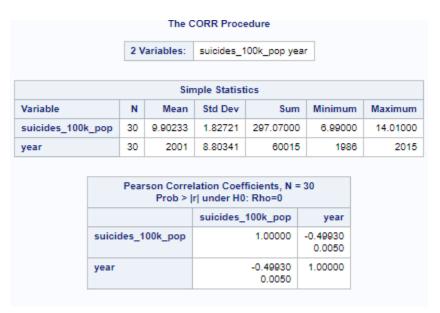


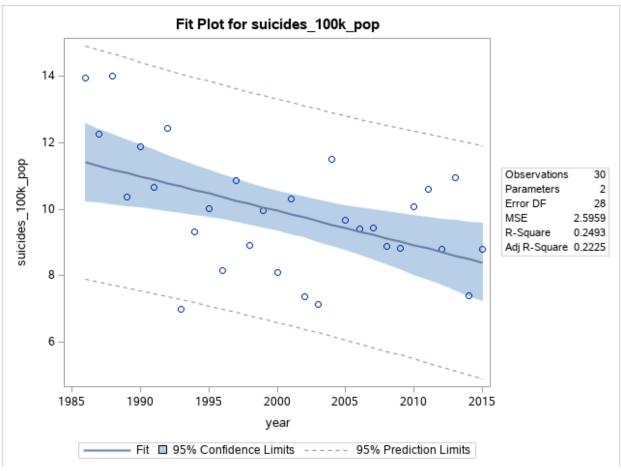
Iceland Female



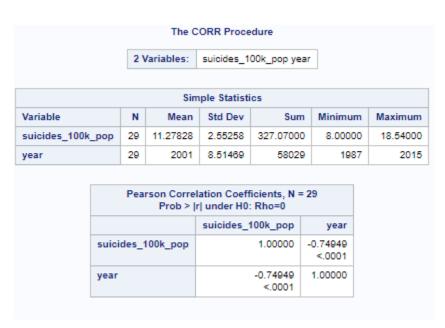


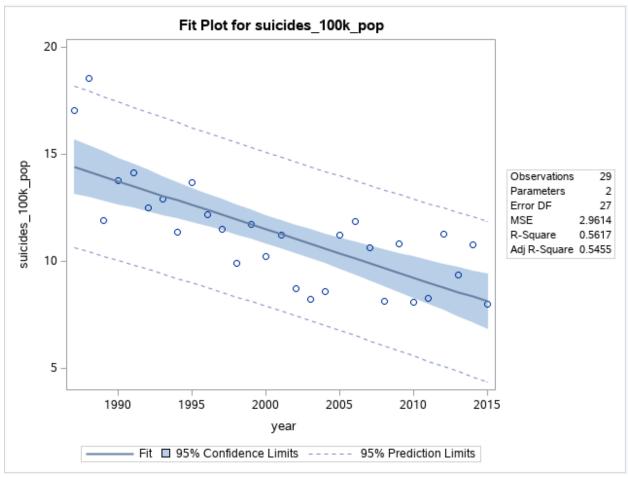
Norway Female



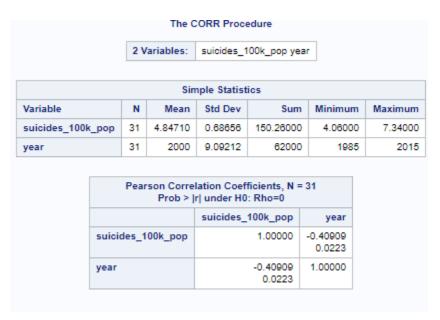


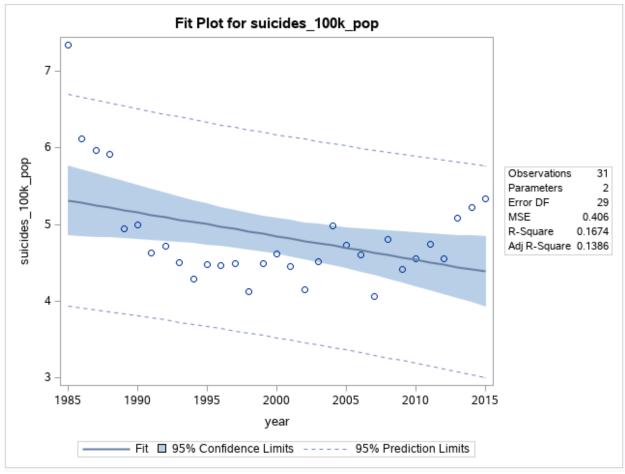
Sweden Female





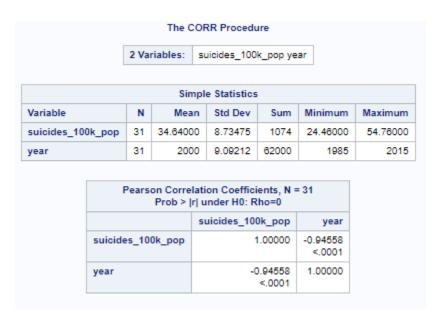
UK Female

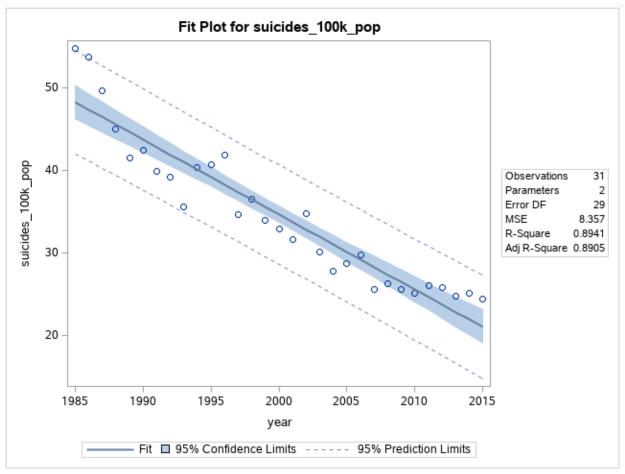




Austria male

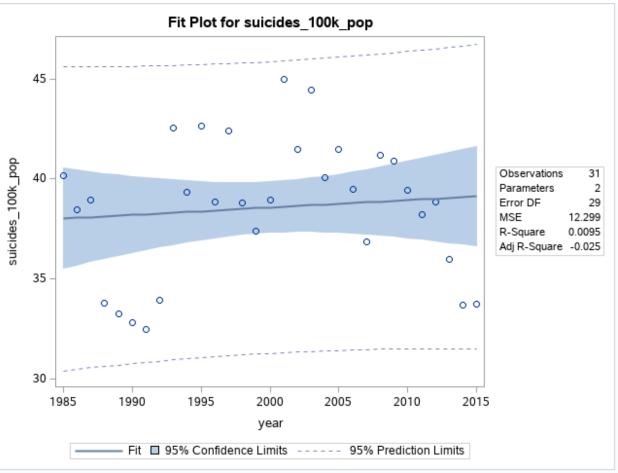
25





Belgium male

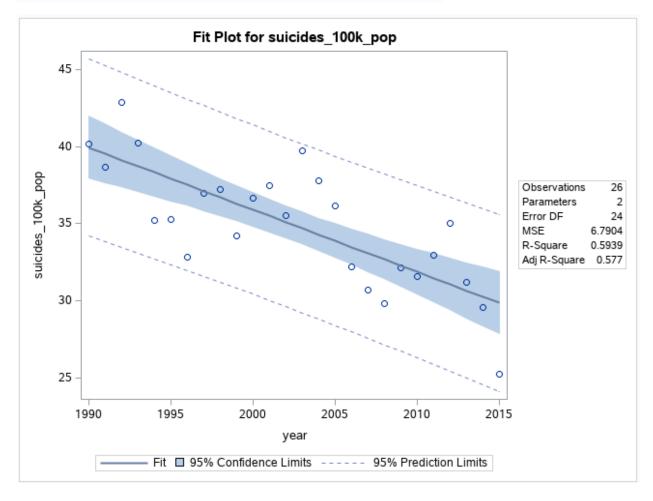




Czech Male

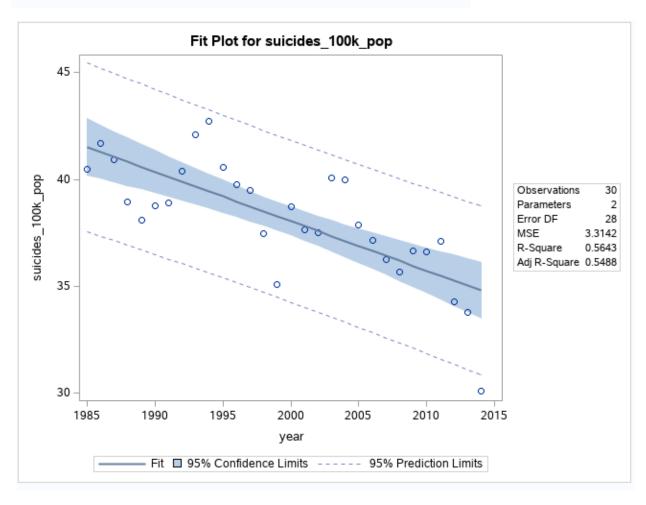
27

			The C	ORR Proce	dure		
		2 \	Variables:	suicides_1	00k_pop ye	ar	
			Sin	nple Statisti	CS		
Variable		N	Mean	Std Dev	Sur	n Minimum	Maximum
suicides_100k	рор	26	34.88462	4.00670	907.0000	0 25.21000	42.84000
year		26	2003	7.64853	5208	5 1990	2015
		Pea		lation Coeff r under H0:		= 26	
				suicides_	100k_pop	year	
	suici	des_	100k_pop		1.00000	-0.77067 <.0001	
					-0.77067	1.00000	

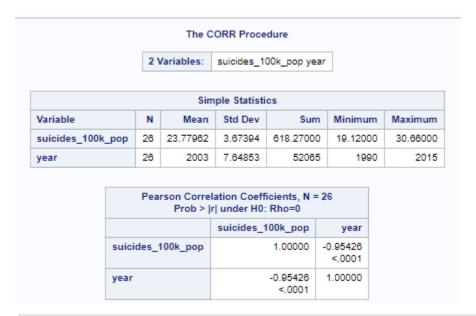


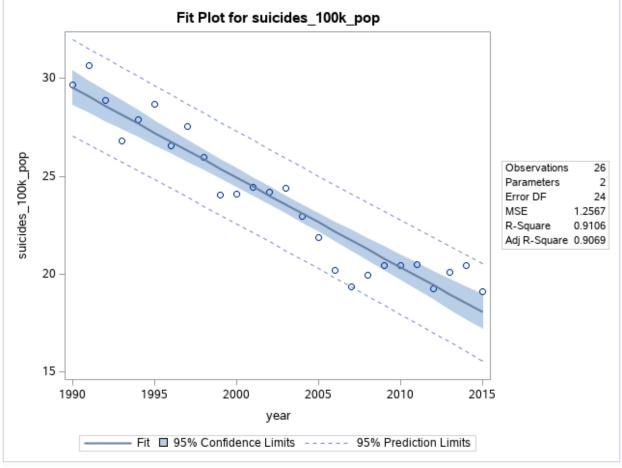
France male

			The C	COF	RR Procedu	ire		
		2 Va	riables:	SI	uicides_100	k_pop ye	ear	
			Sir	npl	e Statistics			
Variable		N	Mea	an	Std Dev	Sum	Minimum	Maximum
suicides_10	0k_pop	30	38.1596	37	2.71013	1145	30.10000	42.73000
year		30	200	00	8.80341	59985	1985	2014
		Pears		r u	on Coeffici Inder H0: R	ho=0	= 30	
				SI	uicides_10	0k_pop	year	
	suicide	es_10	0k_pop		1	1.00000	-0.75122 <.0001	
	year				-(0.75122 <.0001	1.00000	



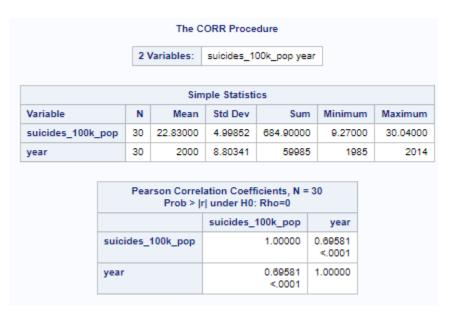
German male

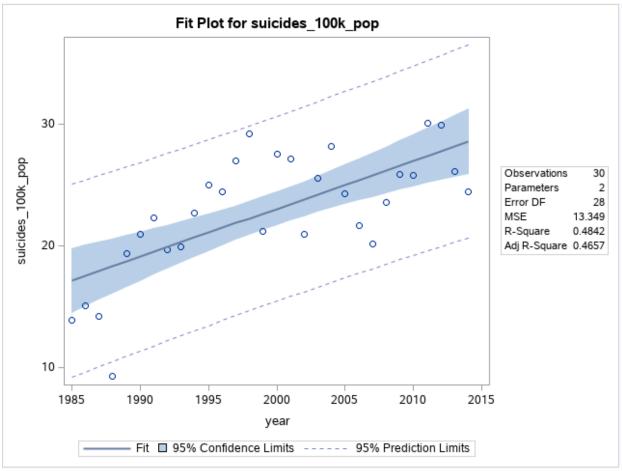




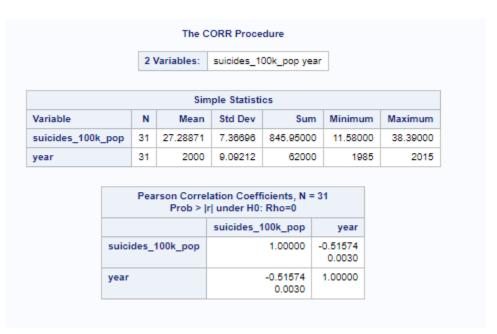
29

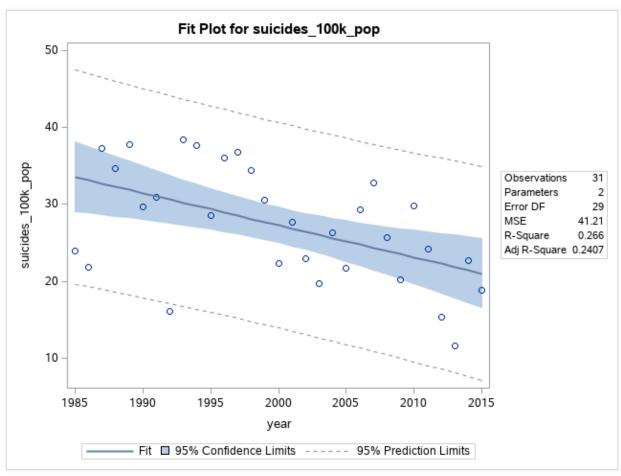
Irish male



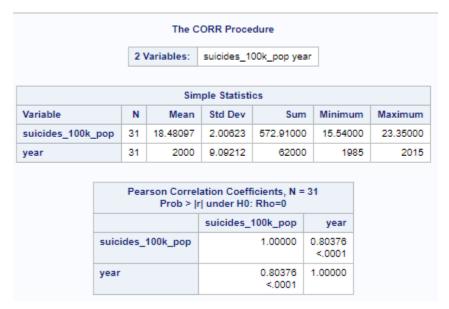


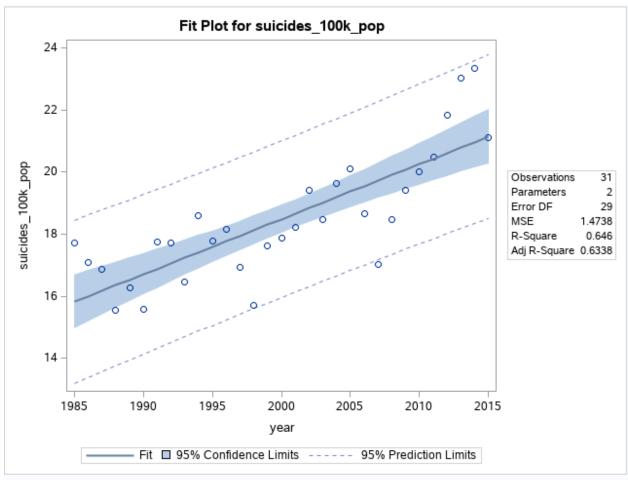
Luxembourg male





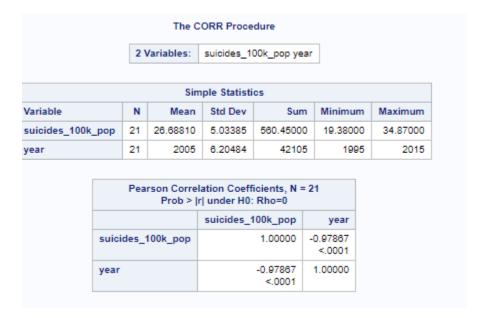
Netherlands male

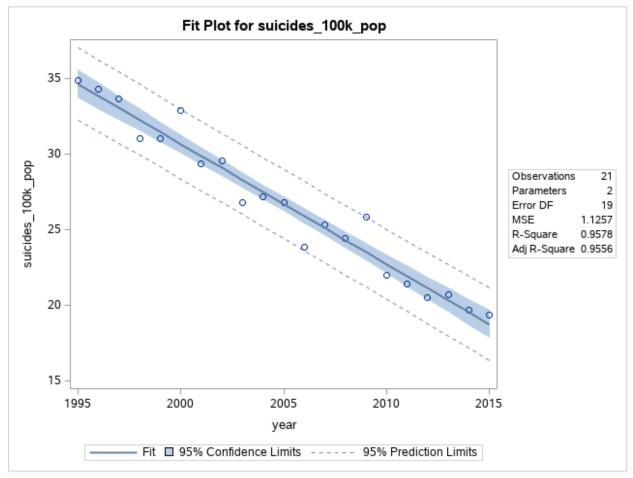




33

Switzerland male

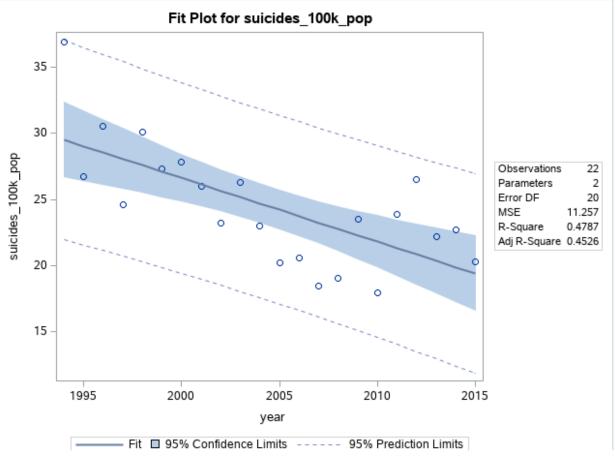




Denmark male

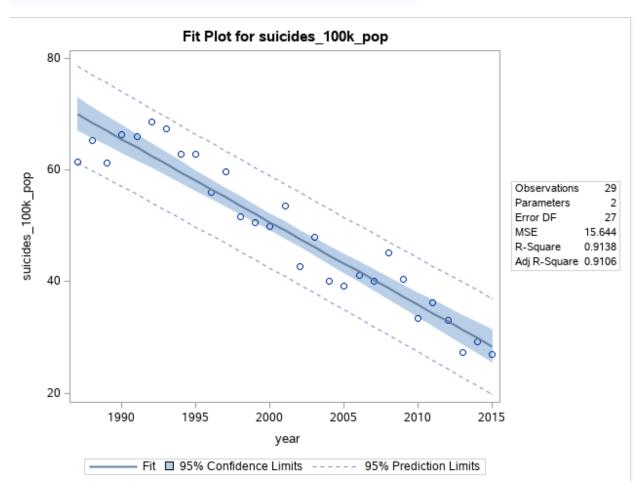
34



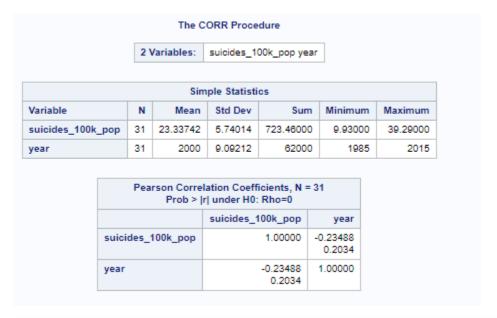


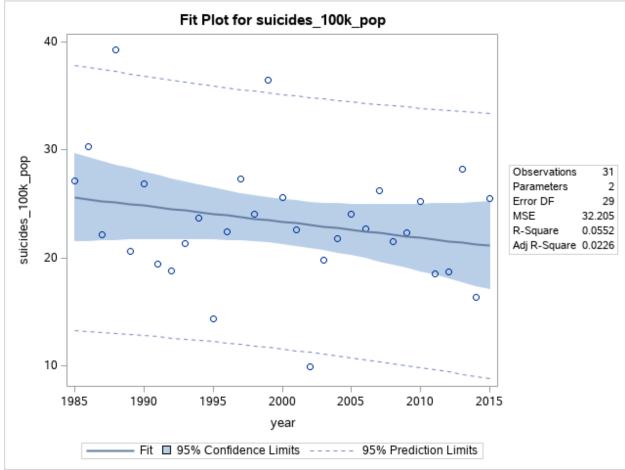
Finland male

			The C	ORR Procedu	ire		
		2 Va	riables:	suicides_100	k_pop ye	ar	
			Sim	ple Statistics			
Variable		N	Mear	n Std Dev	Sum	Minimum	Maximum
suicides_100	es_100k_pop		49.17724	4 13.22907	1426	26.98000	68.56000
year		29	200	1 8.51469	58029	1987	2015
year			on Correl	ation Coeffici	ents, N = ho=0	: 29	2015
year		Pears	on Correl	ation Coeffici under H0: R suicides_100	ents, N = ho=0		2015



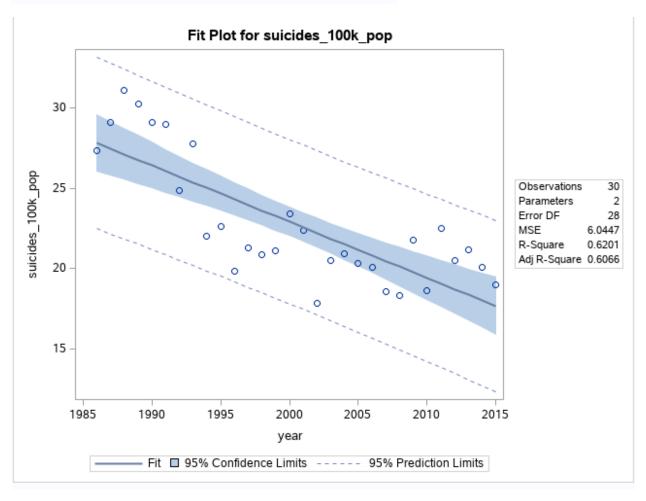
Iceland male





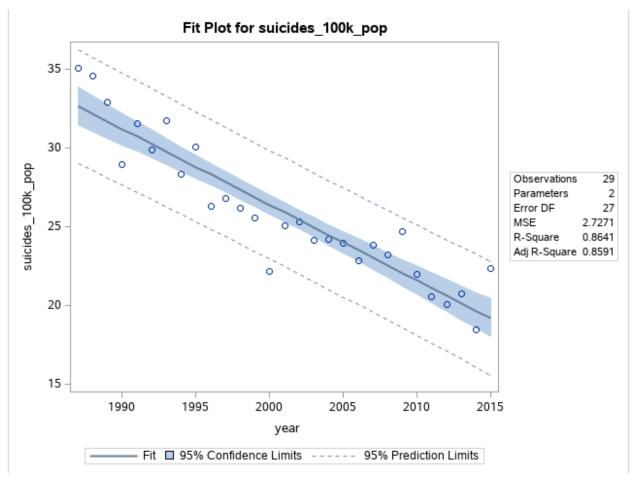
Norway male

		The	CORR Proc	edure		
		2 Variable	es: suicide	es_no ye	ar	
		Si	mple Statis	tics		
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
suicides_no	30	142.30000	13.08105	4269	116.00000	163.00000
year	30	2001	8.80341	60015	1986	2015
	F	earson Corre Prob >	elation Coef			
	5	uicides_no	suicides	_no	year 0.20287	



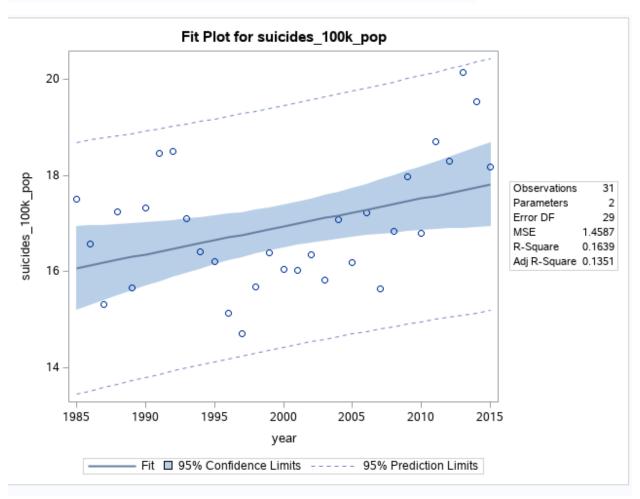
Sweden male

			The C	ORR Proce	dure		
		2 \	/ariables:	suicides_1	00k_pop ye	ar	
			Sin	nple Statisti	CS		
Variable		N	Mean	Std Dev	Sun	n Minimun	Maximum
suicides_100k	es_100k_pop		25.91034	4.39919	751.4000	0 18.46000	35.05000
year		29	2001	8.51469	5802	9 1987	7 2015
		Pea	rson Corre	lation Coeff	inionte N.	- 20	
						- 25	
				r under H0:	Rho=0	- 23	
					Rho=0	year	
	suici			r under H0:	Rho=0		



UK male

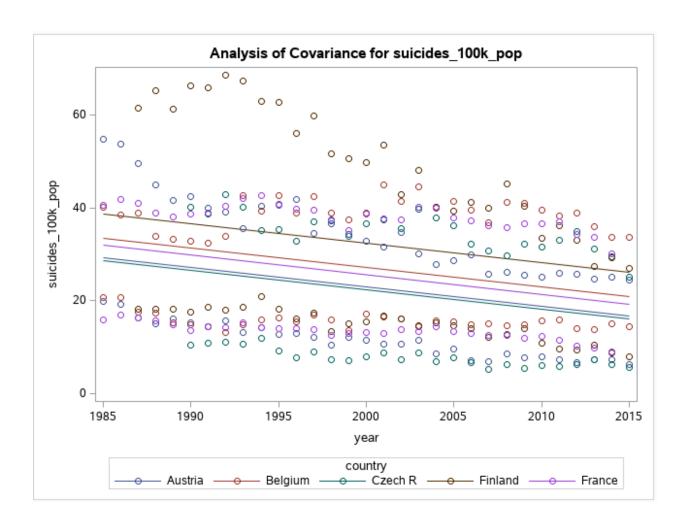
		The C	ORR Proce	dure		
	2 \	/ariables:	suicides_1	00k_pop yea	ar	
		Sin	nple Statisti	cs		
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
suicides_100k_pop	31	16.93645	1.29867	525.03000	14.72000	20.13000
year	31	2000	9.09212	62000	1985	2015
	Pea		ation Coeff under H0:	Rho=0		
		Prob > ı		Rho=0 100k_pop	year	
sui			under H0:	Rho=0		



Multiple Regression

Combination of male and females:

				TI	he GLM	M Proc	edure					
				Cla	ss Lev	el Info	rmation					
	Class	\top	Lev	els \	Values							
	countr	у		5 4	Austria	Belgiu	m Czech F	R Finlar	nd Fr	ance		
	Number of Observations Read 294											
	Number of Observations Used 294											
				T	he GLM	M Proc	edure					
		п	lono	ndent	Variab	ام: جيين	cides_100	k non				
			repe	lidelit	variab	ie. suit	Jues_100	к_рор				
Source	•	I	DF	Sum	of Squ	iares	Mean So	uare	F۱	/alue	Pr >	F
Model			5		7630.8	1434	1528.1	6287		7.65	<.000	01
Error		2	288	5	7486.8	0641	199.6	0697				
Correc	ted Total	2	293	6	5117.6	2075						
	R-Squar	e (Coef	ff Var	Root	MSE	suicides	5_100k	c_po	p Mear	n	
	0.11718	5	54.4	12458	14.1	2823			25	5.95929	9	
5	Source	DF	:	Туре	el SS	Mean	Square	F Va	lue	Pr >	F	
c	ountry	4	3	841.16	5275	960	0.291319	4	.81	0.000	09	
y	/ear	1	3	789.64	9062	3789	9.649062	18	.99	<.000	01	
5	Source	DF	:	Type	III SS	Mean	Square	F Va	lue	Pr >	F	
C	country	4	1 3	695.84	2534	923	3.960634	4	.63	0.001	12	
	/ear	1		789.64			9.649062		.99	<.000		



Female Only

The GLM Procedure

Class Level Information								
Class	Levels	Values						
country	9	Austria Belgium Czech R Finland France Luxembo Netherl Sweden Switzer						

Number of Observations Read	259
Number of Observations Used	259

The GLM Procedure

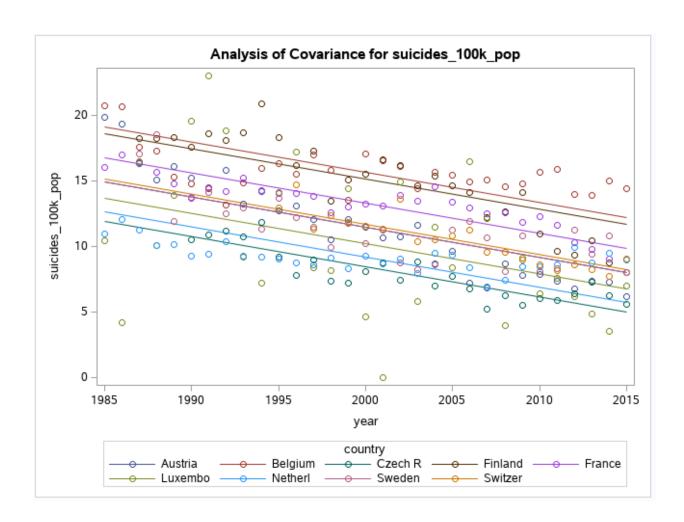
Dependent Variable: suicides_100k_pop

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	2538.030746	282.003416	56.11	<.0001
Error	249	1251.500664	5.026107		
Corrected Total	258	3789.531410			

R-Square	Coeff Var	Root MSE	suicides_100k_pop Mean
0.669748	19.12504	2.241898	11.72232

Source	DF	Type I SS	Mean Square	F Value	Pr > F
country	8	1558.884501	194.860563	38.77	<.0001
year	1	979.146245	979.146245	194.81	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
country	8	1439.778673	179.972334	35.81	<.0001
year	1	979.146245	979.146245	194.81	<.0001



Male Only

The GLM Procedure

Class Level Information						
Class	Levels	Values				
country	4	Austria Belgium Czech R France				

Number of Observations Read	118
Number of Observations Used	118

The GLM Procedure

Dependent Variable: suicides_100k_pop

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4	1623.937586	405.984396	22.61	<.0001
Error	113	2029.215258	17.957657		
Corrected Total	117	3653.152844			

R-Square	Coeff Var	Root MSE	suicides_100k_pop Mean
0.444530	11.57088	4.237648	36.62339

Source	DF	Type I SS	Mean Square	F Value	Pr > F
country	3	389.842482	129.947494	7.24	0.0002
year	1	1234.095104	1234.095104	68.72	<.0001

Source	DF	Type III SS	Mean Square	F Value	Pr > F
country	3	307.631404	102.543801	5.71	0.0011
year	1	1234.095104	1234.095104	68.72	<.0001

