The Relationship Between Petrol Price and Retail Prices Index in the U.K.

Contents

Introduction

I have chosen to study if there is a relationship between the price per liter of petrol and the Retail Prices Index in the U.K. To accomplish this task, I will perform the following steps.

A2

1. Collect data points from a reliable website
2. Organize the information into a table
3. Create a scatter plot of the data and draw the line of the best fit using Microsoft Excel
4. Study the Pearson's Correlation Coefficient using a graphing calculator
5. Place the data into groups depending on the relation with the mean
6. Perform the χ^2 test and analyze the result
7. Conclude my analysis
8. Discuss the validity of conclusions and calculation
9. Discuss the problems with the project design and what can be done to improve it

Data collection

B2

Year	Price per liter(in pence)	Retail Prices Index
1983	36.7	83.1
1984	38.7	87.5
1985	42.8	92.8
1986	38.2	96.7
1987	37.8	100.6
1988	34.7	104.1
1989	38.4	112.3
1990	40.2	121.4
1991	39.5	131.4
1992	40.3	136.7
1993	45.9	139.3
1994	48.9	133.1
1995	50.9	147.5
1996	52.9	151.5
1997	57.9	155.4
1998	60.9	160.8
1999	61.9	164.1

2000	76.9	168.4
2001	77.9	173.1
2002	69.9	174.5
2003	77.9	179.9
2004	77.9	184.6
2005	79.9	190.5
2006	88.9	195.0
2007	87.9	204.4
2008	103.9	212.1

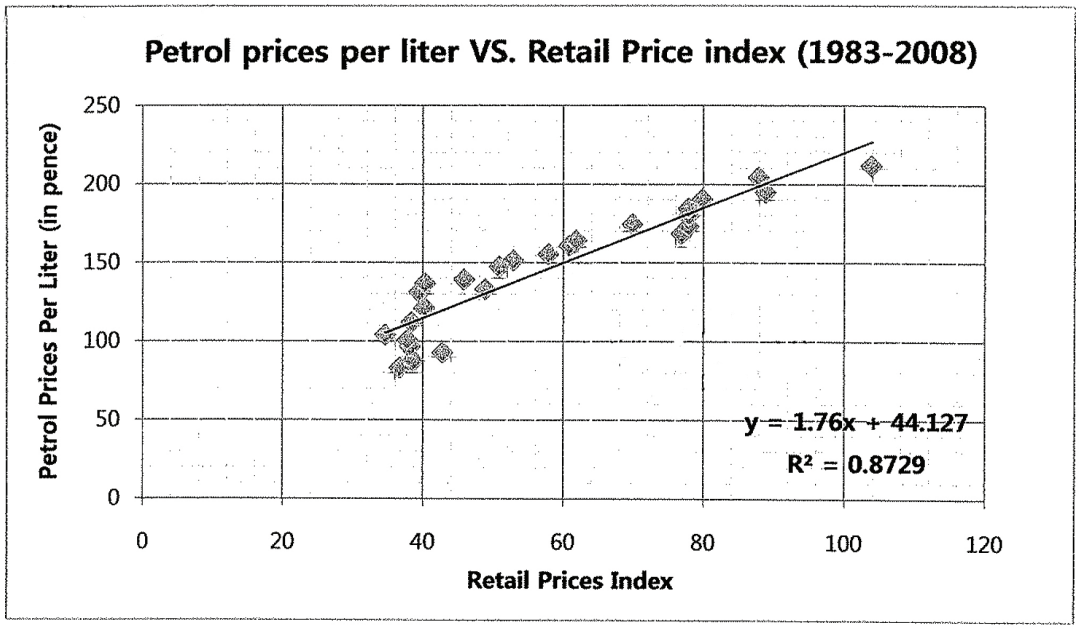
B2

This data shows the mean price per liter of petrol in pence and the Retail Prices Index in the U.K. from the years 1983 until 2008[1]. I obtained these data points from the website speedlimit.org.uk.

Calculations

Using Excel I created a scatter plot of average petrol prices versus Retail Prices Index. I drew a line of the best fit and used Excel to obtain the equation and r^2 value.

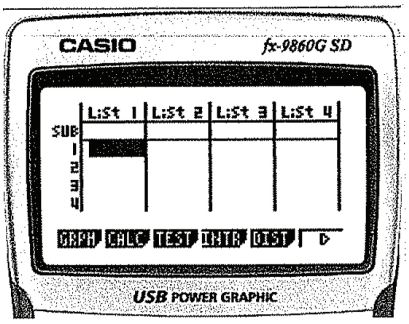
C3



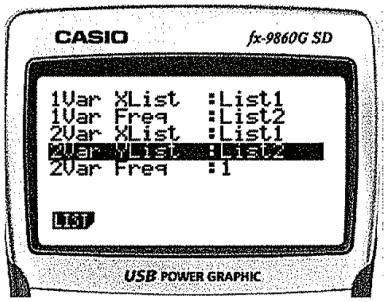
[1] http://www.speedlimit.org.uk/petrolprices.html

We can also find the Pearson's Correlation Coefficient of the two variables and the line of best fit by using a graphing calculator. The calculator will tell us the strength of the correlation.

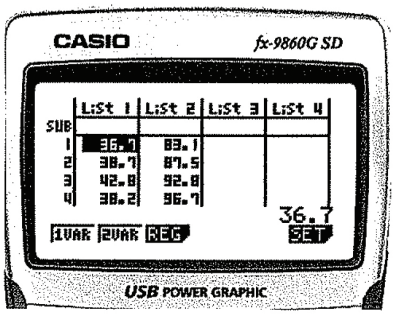
1. Go to menu and press STAT or 2. Then you will see the following.



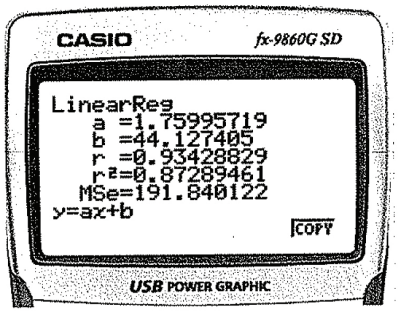
2. Go to CALC(F2), press SET(F6) and then set like this.



3. Press exit and then type price per liter data points on list 1 and retail prices index data on list 2.



4. Then press CALC(F2) -> REG(F3) -> X(F1). Then you will find linear regression and Pearson's Correlation Coefficient as r^2.



Using graphing calculator, it was determined that the equation for the line of best fit is $y = 1.76x + 44.127$ and the Pearson's Correlation Coefficient $r^2 = 0.873$. This agrees with the results from Microsoft Excel. However to further establish that these results are accurate, we can calculate Pearson's Correlation Coefficient manually using the following formulas.

G1

$$r = \frac{s_{xy}}{s_x s_y} \quad \text{where}$$

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}},$$

$$s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n}} \quad \text{and}$$

$$s_{xy} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$$

Using these formulas, I found the following:

C5

Year	Price per liter in pence(x)	Retail Prices Index(y)	$(x-\bar{x})$	$(y-\bar{y})$	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
1983	36.7	83.1	-21.288	-63.084	453.1789	3979.5911	1342.9322
1984	38.7	87.5	-19.288	-58.684	372.0269	3443.8118	1131.8969
1985	42.8	92.8	-15.188	-53.384	230.6753	2849.8515	810.7962
1986	38.2	96.7	-19.788	-49.484	391.5649	2448.6663	979.1894
1987	37.8	100.6	-20.188	-45.584	407.5553	2077.9011	920.2498
1988	34.7	104.1	-23.288	-42.084	542.3309	1771.0631	980.0522
1989	38.4	112.3	-19.588	-33.884	383.6897	1148.1255	663.7198
1990	40.2	121.4	-17.788	-24.784	316.4129	614.2467	440.8578
1991	39.5	131.4	-18.488	-14.784	341.8061	218.5667	273.3266
1992	40.3	136.7	-17.688	-9.484	312.8653	89.9463	167.753
1993	45.9	139.3	-12.088	-6.884	146.1197	47.3895	83.2138
1994	48.9	133.1	-9.088	-13.084	82.5917	171.1911	118.9074
1995	50.9	147.5	-7.088	1.316	50.2397	1.7319	-9.3278
1996	52.9	151.5	-5.088	5.316	25.8877	28.2599	-27.0478
1997	57.9	155.4	-0.088	9.216	0.0077	84.9347	-0.8110
1998	60.9	160.8	2.912	14.616	8.4797	213.6275	42.5628
1999	61.9	164.1	3.912	17.916	15.3037	320.9831	70.0874
2000	76.9	168.4	18.912	22.216	357.6637	493.5507	420.149
2001	77.9	173.1	19.912	26.916	396.4877	724.4711	535.9514
2002	69.9	174.5	11.912	28.316	141.8957	801.7959	337.3002
2003	77.9	179.9	19.912	33.716	396.4877	1136.7687	671.353
2004	77.9	184.6	19.912	38.416	396.4877	1475.7891	764.9394
2005	79.9	190.5	21.912	44.316	480.1357	1963.9079	971.0522
2006	88.9	195.0	30.912	48.816	955.5517	2383.0019	1509.0002
2007	87.9	204.4	29.912	58.216	894.7277	3389.1027	1741.357
2008	103.9	212.1	45.912	65.916	2107.9117	4344.9191	3026.3354

This shows the $\bar{x} = 57.988$ and $\bar{y} = 146.1846$.

From the table we are able to calculate the above formula. $s_x = 19.81461194$ $s_y = 37.32558928$ and $s_{xy} = 690.992307$. Therefore, $r = 0.9342884244$

$r = 0.934$ and $r^2 = 0.873$, which in fact turns out to be same as the Microsoft Excel and the graphing calculator.

G1

The value of $r^2 = 0.873$ indicates a strong correlation and the fact that $r = 0.934$ shows that there is a strong positive correlation between the two variables. As petrol price per liter rise, so

D3

D3

does Retail Prices Index.

To show further statistical evidence that the two variables are dependent we must use a different test as well. We can use the χ^2 test by grouping based on the mean.

	Below $\bar{y}$	Above $\bar{y}$	
Below $\bar{x}$	12	3	15
Above $\bar{x}$	0	11	11
	12	14	26

However, the entry for below $\bar{y}$ and above $\bar{x}$ contains a frequency of 0. Therefore, we can utilize the **Yates' correction for continuity**[2]. This test is useful for a 2 by 2 contingency table and when there is a small amount of data available (frequency of less 5 is present)

The formula for the test is as follows:

$$\chi^2_{calc} = \sum_{i=1}^{n} \frac{\left(|O_i - E_i| - 0.5\right)^2}{E_i}$$

The null hypothesis in this case is that price per liter and the Retail Prices Index are independent of one another. The alternate hypothesis is that the two variables are dependent. I analyzed the data to a 5% significance level.

Using a graphing calculator, I was able to calculate the expected values as follows:

$= \frac{12 \times 15}{26}$

	Below $\bar{y}$	Above $\bar{y}$
Below $\bar{x}$	6.923	8.0769
Above $\bar{x}$	5.0769	5.923

[2] http://en.wikipedia.org/wiki/Yates'_chi-square_test

C5

Therefore, if we put the values into the formula it becomes:

χ^2_{calc} = 3.02599+2.593571+4.126143+3.536878

= 13.283

The formula for Yates' correction for continuity yielded a χ^2_{calc} value of 13.283. Using the formula df=(r-1)(c-1) I found degrees of freedom equal to 1. Using the table of critical values, χ^2_{crit} = 3.84. Because $\chi^2_{calc} > \chi^2_{crit}$ we can reject the null hypothesis. Therefore the results support the notion that price per liter of petrol and the Retail Prices Index are dependent factors in the U.K.

Interpretation

D3

To investigate whether price per liter of petrol and the Retail Prices Index are related factors, I performed the two tests which are Pearson's Correlation Coefficient test and the Yates' correction for continuity (a form of the χ^2 test). Both tests provided support that the two variables are dependent. This is evidence that from the years 1983 until 2008 the price per liter of petrol oil and the retail prices index were related in the U.K.

My calculations show that Pearson's Correlation Coefficient equaled 0.873. This corresponds with strong positive correlation. Moreover, the Yates' correction for continuity test also provided support to a 5% significance level that I must reject the null hypothesis and the two variables are in fact related.

One reason we may see this relationship is that the price of petrol plays a huge role in the economy of a country. Many other sectors of the economy are greatly affected by the price of petrol. Petrol is used for transportation, heating homes, and many other things. It also affects the manufacturing and pricing of consumer goods. Therefore, if the price rises, other prices will also tend to rise, and if the price falls, prices of other goods will also tend to fall. This may explain why the price of oil is so closely tied to the prices of other goods in general.

Validity

I first performed the linear regression to see the relationship between the two variables. The scatter plot clearly showed a strong positive linear correlation between the two variables. A very high r^2 value confirmed that there is a positive correlation between the two variables. To provide even further support, we used the form of χ^2 test known as the Yates' correction for continuity. This statistical test also demonstrated that the two variables are dependent on one another. Taken together, these tests provide strong support that price per petrol oil and the Retail Prices Index are related factors.

I'm confident about my mathematical processes because the calculations were repeated several times to ensure accuracy. For the Pearson's Correlation Coefficient I used three different methods to obtain the r^2 value. The Microsoft Excel, graphing calculator and math by hand all produced the same figures. In addition I performed more than one test to show that the two variables are dependent. Both the Pearson's Correlation Coefficient and the Yates' correction for continuity confirmed that the two variables are related.

G1

G1

E1

Areas of improvement

To know for sure if the two variables are dependent, this study design could be improved. Firstly, the study only involved the years 1983 until 2008. This severely limited the data range and it is possible that it did not give a clear picture of the relationship. In addition the data only pertained two prices in the U.K. To improve the study the data should be extended to include several years prior to 1983 as well. Data could also be taken from across the globe to show whether this relationship is universal or simply just present in a few places. Such changes would drastically improve the study and make the results more valid.

Appendix

Year	Price per liter(in pence)	Retail Prices Index
1983	36.7	83.1
1984	38.7	87.5
1985	42.8	92.8
1986	38.2	96.7
1987	37.8	100.6
1988	34.7	104.1
1989	38.4	112.3
1990	40.2	121.4
1991	39.5	131.4
1992	40.3	136.7
1993	45.9	139.3
1994	48.9	133.1
1995	50.9	147.5
1996	52.9	151.5
1997	57.9	155.4
1998	60.9	160.8
1999	61.9	164.1
2000	76.9	168.4
2001	77.9	173.1
2002	69.9	174.5
2003	77.9	179.9
2004	77.9	184.6
2005	79.9	190.5
2006	88.9	195.0
2007	87.9	204.4
2008	103.9	212.1

Bibliography

Price per liter and the Retail Prices Index in the U.K.

http://www.speedlimit.org.uk/petrolprices.html

retrieved on October 15th, 2008

Yates correction for continuity,

http://en.wikipedia.org/wiki/Yates'_chi-square_test

retrieved on January 4th, 2009

HAESE & HARRIS PUBLICATIONS, Mathematics for the international student Mathematical Studies SL textbook, International Baccalaureate Diploma Programme