Chapter 7

Analysis of Correlation Coefficient

In view of the large number of sample (199) of teeth that has been collected from the areas of Klang Valley and Malacca, a detailed study has been carried out to examine the relationship between the concentration of each element and the donors' age and also the relationship between the elements themselves. If there is a strong association between two variables, then knowing one will help a lot in predicting the other. But when there is a weak association, information about one variable does not help much in determining the other. Correlation is always between –1 and 1, but can take any value in between. A positive correlation means that the cloud slopes up: as one variable increases, so does the other. A negative correlation means that the cloud slopes down: as one variable increase, the other decreases (D. Freedman, et.al., 1978)

7.1 Relationship between the Elements' Concentration and the Age of Donors

Figures 7.1 to 7.7 are the correlation coefficient graphs plotted to show the relationship between the elements' concentration and the donors' age for the areas of Klang Valley and Malacca. Figure 7.1 reflects the poor relationship between the concentration and the donors' age for each element from the total samples of 199.

The correlation value for each element is as low as 0, except for element Ca and Zn, which represented the value at 0.1 and 0.2 respectively. Therefore, neither will the concentration of each element will increase as the age of sample increases, nor will the concentration of each element decreases as the age increases.

This study is extended to focus on separate areas of Klang Valley and Malacca respectively; the results show that the correlation between the elements and the age gives a slightly higher value as compared with the results mentioned above. The correlation coefficient for the samples of Klang Valley and Malacca are shown in figure 7.2 and 7.3 respectively. From figure 7.2, the positive but weak relationship between the concentration of elements and donors' age has been noted, while the Malaccan sample also shows the same trend of relationship except for the element Cd, which gives a correlation of – 0.1. An obvious difference is found in figure 7.2, where the correlation of 0.1 for the element of Cd if compared with 0 and –1 in figure 7.1 and 7.2 respectively. Therefore, to determine the significance of a difference between the correlation coefficients of the samples from Klang Valley and Malacca, t-test statistic is used as follow (J.R. Green and D. Margerisson, 1977):

$$\mathbf{Z} = 1.1513 \log_{10} \frac{(1+r)}{(1-r)}$$

$$z = \frac{Z_1 - Z_2 - \mu_{Z_1 - Z_2}}{\sigma_{Z_1 - Z_2}}$$

Where $\mu_{Z_1-Z_2} = \mu_{Z_1} - \mu_{Z_2}$ and $\sigma_{Z_1-Z_2} = \sqrt{1/(N_1-3) + 1/(N_2-3)}$ is *normally distributed. (*One Assumption has been made that the samples of Klang Valley and Malacca are distributed normally). Base on the above equations, we can determine that there is a significant difference between the two coefficients at 0.05 level (95% confidence level).

Therefore, we have to decide between the hypotheses H_o: $\mu_{z1} = \mu_{z2}$ and H_1: $\mu_{z1} \neq \mu_{z2}$. Under the hypothesis H_o: $\mu_{z1} = \mu_{z2}$, z is calculated for each element respectively as shown in table 7.1.

Table 7.1: Calculation of z for the Elements of Klang Valley and Malacca Samples

	Klang Valley (r_1)	Malacca (r_2)	z
Ca	0.022	-0.017	0.262
Cd	0.107	-0.076	1.100
Cu	0.026	0.075	-0.305
Pb	0.008	0.107	-0.582
Zn	0.212	0.016	1.243

By using a two-tailed test of the normal distribution, H_o will be rejected only if $z > 1.96$ or $z < -1.96$. Thus, H_o will not be rejected since the z value for each element of the two areas is within the range. So, from the correlation coefficients at 0.05 level for Klang Valley and Malacca it can be concluded that the samples are not significantly different.

Further analysis is carried out on the correlation coefficient between the elements and the age of the female and male donors in Klang Valley (figures 7.4 and 7.5) and Malacca (figures 7.6 and 7.7) respectively. The coefficients of Ca, Pb and Cd reflect the divergence between the samples of female and male for the area of Klang Valley as shown in figures 7.4 and 7.5. The female samples show a high positive coefficient if compared with the male samples. Further, the coefficient of Pb for male samples is about – 0.2, while + 0.4 for female samples. This coefficient of

+ 0.4 implies that the concentration of Pb increases as the age of female increases. To determine whether the increase in coefficient of Pb is significant, two hypotheses had been set up; H_o: $\rho = 0$ (null hypothesis) and H_1: $\rho \neq 0$, using a two-sided T-test, the distribution of

$$T = r\left(\frac{n-2}{1-r^2}\right)^{1/2}$$

Therefore, the realization of test statistic is thus, T = 3.103, which is more than $t_{50}(0.995) = +2.68$. Hence, r is significantly large positive at the 1% level, so H_o is rejected and H_1 is accepted and this signifies some correlation between the Pb concentration and the age of female samples. This had been noted by Steenhout and Purtois (1981) that the level of Pb concentration focused on the teeth of a person increases with age.

Base on the above equations, we can determine if there is a significant difference between the two coefficients of the samples of female and male at a 0.05 level. Therefore, the hypotheses H_o: $\mu_{z1} = \mu_{z2}$ and H_1: $\mu_{z1} \neq \mu_{z2}$ has been made. Under the hypothesis H_o: $\mu_{z1} = \mu_{z2}$, z is calculated for each element respectively as shown in table 7.2.

Table 7.2: The Calculation of Significant Difference (z) for Coefficients of Female and Male for the Klang Valley Samples

	Female (r_1)	Male (r_2)	z	Note
Zn	0.372	0.085	1.458	No significant
Ca	0.182	-0.056	1.234	No significant
Pb	0.402	-0.239	3.205	Significant difference
Cd	0.188	0.053	0.603	No significant
Mg	0.361	0.171	1.034	No significant
Cu	0.299	0.281	0.092	No significant

As application above, a two-detailed test of the normal distribution has been used, H_o will be rejected only if $z > 1.96$ or $z < -1.96$. Thus, H_o will not be rejected since the z value for each element of the two areas is within the range. So, the conclusion can be made that the correlation coefficients for female and male are not significantly different at a 0.05 level except for element Pb.

However, the coefficients of the female samples from Malacca for Cd, Cu and Zn, have simply diverged from the male samples as shown in figures 7.6 and 7.7. To check the significant level of these coefficients of female and male samples, z value is calculated as above.

Table 7.3: The Calculation of Significant Difference (z) for Coefficients of Female and Male for the Malacca Samples

	Female (r_1)	Male (r_2)	Z	Note
Zn	0.333	-0.301	2.454	Significant difference
Ca	-0.024	-0.003	-0.093	No significant
Pb	0.138	0.214	-0.263	No significant
Cd	0.149	-0.337	1.932	No significant
Cu	-0.015	0.255	-1.057	No significant

H_o will be rejected only if $z > 1.96$ or $z < -1.96$. Therefore, no significant difference is found for the elements of Ca, Pb, Cd and Cu, except for Zn between the female and male samples of Malacca. Since the coefficients of Zn for female and male are registered at 0.333 and –0.301 respectively, thus to test the significant of correlation, the distribution of T-test formula has to be applied:

$$T = r\left(\frac{n-2}{1-r^2}\right)^{1/2}$$

Where the n for female and male are 45 and 24 respectively. Two hypotheses has been set up; H_o: $\rho = 0$ and H_1: $\rho \neq 0$. Thus,

	Female	**Male**
T	2.312	-1.478
$t_{50}(0.995) = +2.68, -2.68$ (2 tailed test)	Accepted H_0	Accepted H_0

the T value for female is less than $t_{50}(0.995) = +2.68$, while for the male is more than -2.68. Hence, H_o is accepted and H_1 is rejected. There is no correlation between the Zn concentrations and the age of female and male samples.

Table 7.4: The Calculation of *z* for Female Samples from Klang Valley and Malacca

	Female (Klang Valley) (Correlation Coefficient)	Female (Malacca) (Correlation Coefficient)	*z*	**Note**
Zn	0.3715	0.3325	0.213	No significant difference
Ca	0.1824	-0.0244	1.065	No significant difference
Pb	0.4018	0.1377	1.275	No significant difference
Cd	0.1881	0.1490	0.177	No significant difference
Cu	0.2985	-0.0145	1.471	No significant difference

Table 7.5: The Calculation of *z* for Male Samples from Klang Valley and Malacca

	Male (Klang Valley) (Correlation Coefficient)	Male (Malacca) (Correlation Coefficient)	*z*	**Note**
Zn	0.0849	-0.3006	1.479	No significant difference
Ca	-0.0564	-0.0025	-0.230	No significant difference
Pb	-0.2391	0.2137	-1.609	No significant difference
Cd	0.0529	-0.3374	1.558	No significant difference
Cu	0.2809	0.2554	0.109	No significant difference

The hypothesis of H_o would be rejected only if $z > 1.96$ or $z < -1.96$. Therefore, no significant difference is found for the elements of Ca, Pb, Cd, Cu and Zn among the female and male samples of Klang Valley and Malacca as shown in table 7.4 and 7.5 respectively.

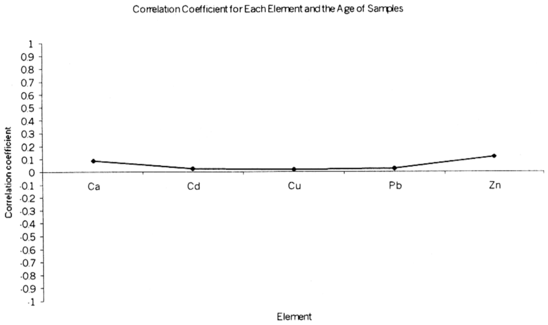

Figure 7.1: The Correlation Coefficient for Each Element versus the Donors' Age of the Total Samples

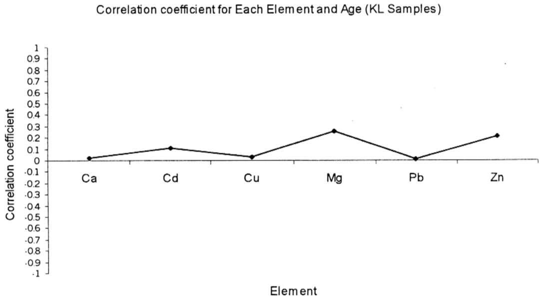

Figure 7.2: The Coefficients between the Elements' Concentrations and The Donors' Age for the Klang Valley Samples

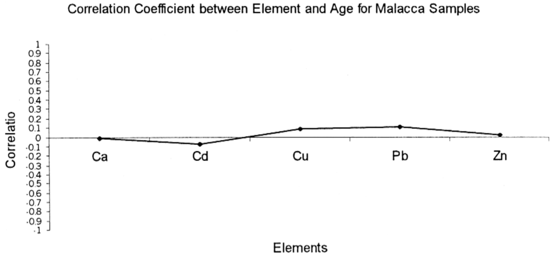

Figure 7.3: The Coefficients between Concentrations of Elements and Donors' Age for Malacca Samples

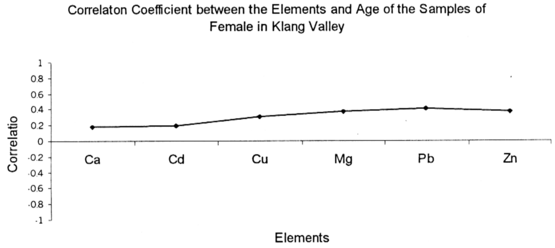

Figure 7.4: The Coefficients between the Elements and Females Donors' Age for the Klang Valley Samples

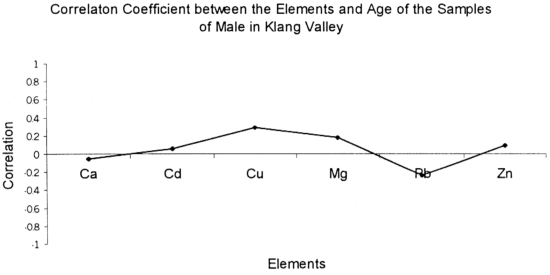

Figure 7.5: The Coefficients between the Elements and Males Donors' Age for the Klang Valley Samples

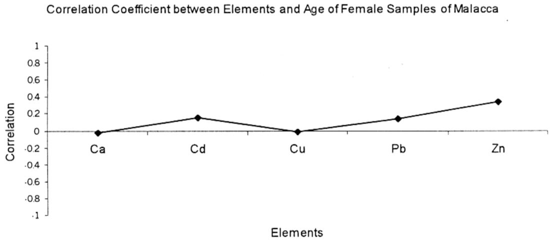

Figure 7.6: The Coefficients between the Elements and Females Donors' Age for the Malacca Samples

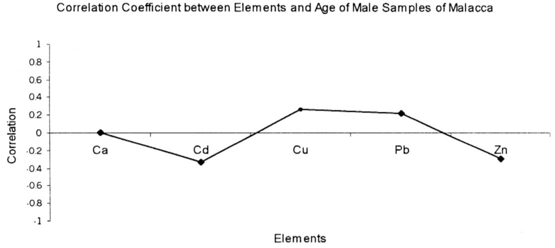

Figure 7.7: The Coefficients between the Elements and Males Donors' Age for the Malacca Samples

7.2 Relationships between Each Element

The analysis then can further be carried out to study the relationship among the elements in these samples. As shown in the table of 7.6, the coefficient of correlation between the concentration of elements has been calculated, with the objective of finding out the relationship of each element.

Table 7.6: Correlation Coefficient among the Relationships of Each Element for Klang Valley and Malacca Areas

		Correlation Coefficient			Correlation Coefficient
Zn	Cu	0.064	Cu	Zn	0.064
	Ca	-0.093		Ca	0.046
	Mg	0.034		Mg	0.020
	Pb	-0.112		Pb	-0.099
	Cd	0.416		Cd	0.010
Ca	Cu	0.046	Mg	Cu	0.020
	Zn	-0.093		Ca	0.251
	Mg	0.251		Zn	0.034
	Pb	0.111		Pb	0.049

	Cd	-0.116		Cd	0.046
Pb	Cu	-0.099	Cd	Cu	0.010
	Ca	0.111		Ca	-0.116
	Mg	0.049		Mg	0.046
	Zn	-0.112		Pb	-0.054
	Cd	-0.054		Zn	0.416

Base on the table 7.6, the correlation coefficient between the elements are near to 0 except for the relationship between Zn and Cd, and Ca and Mg, which the coefficients record at 0.416 and 0.251 respectively. Therefore, by using a two-tailed test, and the set up two hypotheses: H_o: $\rho = 0$ and H_1: $\rho \neq 0$, the test statistic of T applied to determine the significant of the coefficients of 0.416 and 0.251 at the 1% level is carried out.

	Coefficient of Zn and Cd ($r_1 = 0.4155$)	**Coefficient of Ca and Mg** ($r_2 = 0.2508$)
T	6.411	3.636

Thus, the calculated T for both coefficients are more than T_{198} (0.995) = +2.58. Hence r_1 and r_2 are significantly large positively at the 1% level, so, H_0 is rejected, while H_1 is accepted as that there is some correlation between Zn and Cd, and Ca and Mg.

Magnesium is an important element to regulate the Calcium level in the body and 60% of the concentration is found within the bone and teeth and is incorporated as a minor element as apatite crystals and thus, contributes to bone and teeth structure. Hence, the concentration of Mg and Ca has close relationship as calculated as above.

Zinc, which is found in high concentration in teeth and bone, involves in many cellular processes, especially in the proper calcification of bone and teeth. However, Cadmium is a major toxic pollutant in the environment and can cause hazard to the human health. Cd deposits into the human body through food consumption and cigarette smoke inhalation. Thus, the above statistic test shows the significant correlation between Zn and Cd, can be explained from the food intake containing Zn, which is contaminated with Cd. Otherwise, this significant correlation does not necessary imply that one is the cause of the other—there might be one or more other causal factors affecting them. Further, this correlation between Zn and Cd is just based on the 199 samples scatteredly taken from around the 2 locations namely the Klang Valley and Malacca. Hence, this correlation may not reflect the real picture for the concentrations for Zn and Cd in human teeth.

7.3 Conclusion

The correlation coefficients had been calculated to investigate the two relationships, namely (1) the relationship between the element concentration and the age of donor, (2) the relationship between each element.

The findings showed that the concentration of Pb increases as the age of female increases. This result agrees with the research done by Steenhout and Purtois (1981). However, this relationship was not noticeable in male samples.

The results showed that there were some noticeable correlation between the elements of Mg and Ca, and Zn and Cd. The correlation between Mg and Ca could be understood since Mg is an important element to regulate the Ca level in human body. However, further research for confirmation and possible explanation would be needed to resolve the correlation between Zn and Cd.

References:

D. Freedman, R. Pisani and R. Purves, 1978, *Statistics,* Norton & Company, New York, pg. 109-121.

J.R. Green and D. Margerisson, 1977, *Statistical Treatment of Experimental Data,* Elsevier North-Holland Inc., pg.189-197.

Steenhout, A., Pourtois, M., 1981. *Lead accumulation in teeth as a function of age with difference exposure*, Br. J. Ind. Med., 38, 297.