

## Chapter 3

### Preliminary Data Analysis

In this chapter we describe the preliminary analysis of the fertility rates in 2002-2005 in four provinces of southern Thailand. The result of the preliminary analysis has three sections. The first section we show the distribution of variables. Next, we examine the associations between fertility rate and determinants. Finally, we calculate the total fertility rate by year and region. Between January 2002 and December 2005, there were 186,532 births in four provinces, Pattani had 49,860 births, Satun had 18,642 births, Songkhla had 83,552 births and Trang had 34,478 births.

#### 3.1 Variable types

The variables of interest for the first part of study include five determinants and one outcome. Year, month, district and mother's age are the nominal determinants. The fertility rate is the continuous outcome. Their roles and data types are shown in Table 3.1.

Variables	Type	Role
Year	nominal (4)	determinant
Month	nominal (12)	determinant
Province	nominal (4)	determinant
District	nominal (45)	determinant
Mother age	nominal (7)	determinant
Fertility rate	continuous	outcome

*Table 3.1: Variable types and roles*

Table 3.2 shows the number and percentage of births in each year, month, age group, province and district.

Determinant	Number of births	Percent
	186,532	100.0
Year		
2002	46,334	24.8
2003	45,772	24.5
2004	47,053	25.2
2005	47,373	25.4
Month		
January	15,342	8.2
February	13,543	7.3
Mar	14,998	8.0
April	15,154	8.1
May	15,740	8.4
June	15,456	8.3
July	16,530	8.9
August	16,363	8.8
September	16,312	8.7
October	16,133	8.6
November	15,945	8.5
December	15,016	8.1
Age group		
15-19	28,152	15.1
20-24	46,736	25.1
25-29	49,895	26.7
30-34	36,645	19.6
35-39	18,782	10.1
40-44	5,171	2.8
45-49	1,151	0.6
Provinces		
Sogkhla	83,552	44.8
Satun	18,642	10.0
Trang	34,478	18.5
Pattani	49,860	26.7

Table 3.2: Number of births

ฝ่ายหอสมุด  
ศูนย์หญิงหลง อรรถกระวีสุนทร

Determinant	Number of births	Percent
<b>Districts</b>		
<b>Songkhla</b>		
Mueang	16,595	8.9
SathingPhra	811	0.4
Chana	2,845	1.5
NaThawi	3,147	1.7
Thepa	3,329	1.8
SabaYoi	4,115	2.2
Ranot	1,668	0.9
KrasaeSin	333	0.2
Rattaphum	2,783	1.5
Sadao	4,174	2.2
HatYai	41,551	22.3
NaMom	216	0.1
KhuanNiang	807	0.4
BangKlam	226	0.1
Singhanakhon	636	0.3
KhlongHoiKhong	316	0.2
<b>Satun</b>		
Mueang	10,860	5.8
KhuanDon	858	0.5
KhuanKalong	1,913	1.0
ThaPhae	1,159	0.6
La-ngu	2,831	1.5
TungWa	968	0.5
Manang	53	0.03

Table 3.2: Number of births (ctd.)

Determinant	Number of births	Percent
<b>Trang</b>		
Mueang	20,410	1.9
Kantang	2,359	1.3
YanTakhao	2,638	1.4
Palian	1,288	0.7
Sikao	1,393	0.7
HuaiYot	2,788	1.5
WangWiset	1,344	0.7
NaYong	1,415	0.8
Rutsada	792	0.4
HatSamran	51	0.03
<b>Pattani</b>		
Mueang	15,241	8.2
KhohPho	3,417	1.8
NongChik	4,631	2.5
Panare	2,734	1.5
Mayo	3,492	1.9
TungYangDaeng	2,267	1.2
SaiBuri	4,349	2.3
MaiKaen	929	0.5
Yaring	5,403	2.9
Yarang	5,694	3.1
Kapho	780	0.4
MaeLan	923	0.5

*Table 3.2: Number of births (ctd.)*

The birth certificate form contained the age of the mother in years, is classified as 20 or less, 21-25, 26-30, 31-35, 36-40, 41-45, and 46 or more. We estimated the age of the mother in new group as the age of female in census include 7 groups as 15-19, 20-24, 25-29, 30-34, 35-39, 40-44 and 45-49.

Figure 3.1 shows the distribution of the fertility rates.

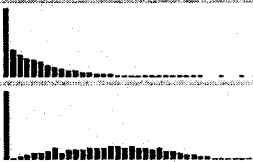

Variable name	Size	Graph	Minimum	Maximum	Mean	StDev	Skewness	Kurtosis
fertility rate	15120		0	32.349	3.779	4.101	1.526	2.711
lnFertility rate	15120		0	3.507	1.203	0.874	0.020	-1.136

Figure 3.1: Histograms and numerical summaries of transformed data

The distribution of fertility rate is positively skewed, with skewness coefficient of 1.5; therefore taking appropriate transformations of the data can reduce the skewness. The data were transformed by adding 1 to account for the zeros, and taking the natural logarithm.

For the second part of the study, we would like to fit a logistic regression model to the fertility. The variables of interest include five determinants and one outcome. Period is a binary determinant and season, region and mother’s age are the nominal determinants. Birth is the status of birth, is the binary outcome (1 is birth, 0 is no birth). Their roles and data types are shown in Table 3.3.

Variables	Type	Role
Period	binary	determinant
Season	nominal (4)	determinant
Region	nominal (8)	determinant
Mother age	nominal (7)	determinant
Birth	binary	outcome

Table 3.3: New group variable types and roles

### 3.2 Associations between the fertility rate and determinants

In this section we examine the associations between fertility rate and determinants.

Figure 3.2 shows the association between the fertility rate and year.

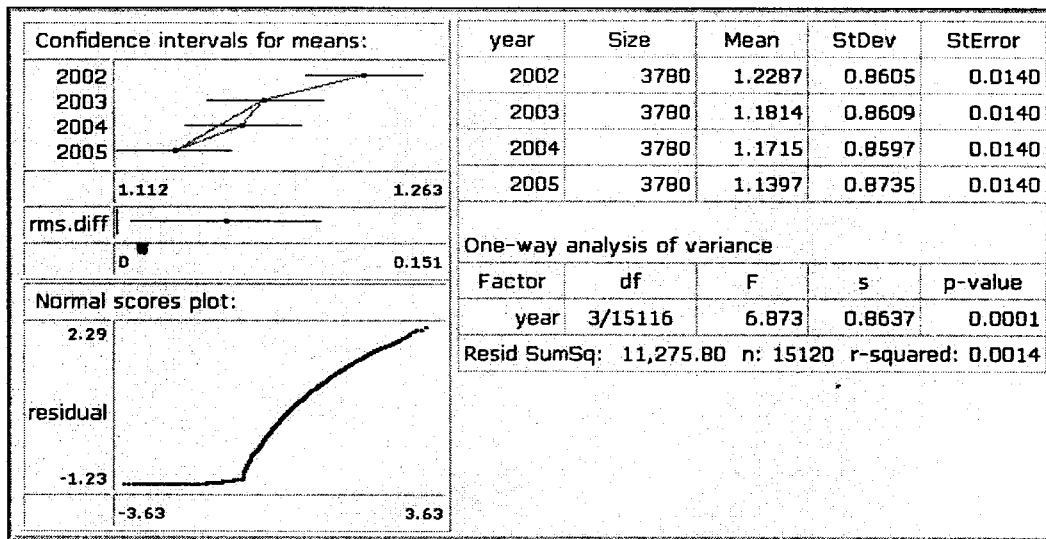


Figure: 3.2 Comparison of fertility rate by year

The association between fertility rate and year is statistically significant (p-value < 0.05). The fertility rates steadily declined from an average of 1.23 in 2002 to 1.14 in 2005.

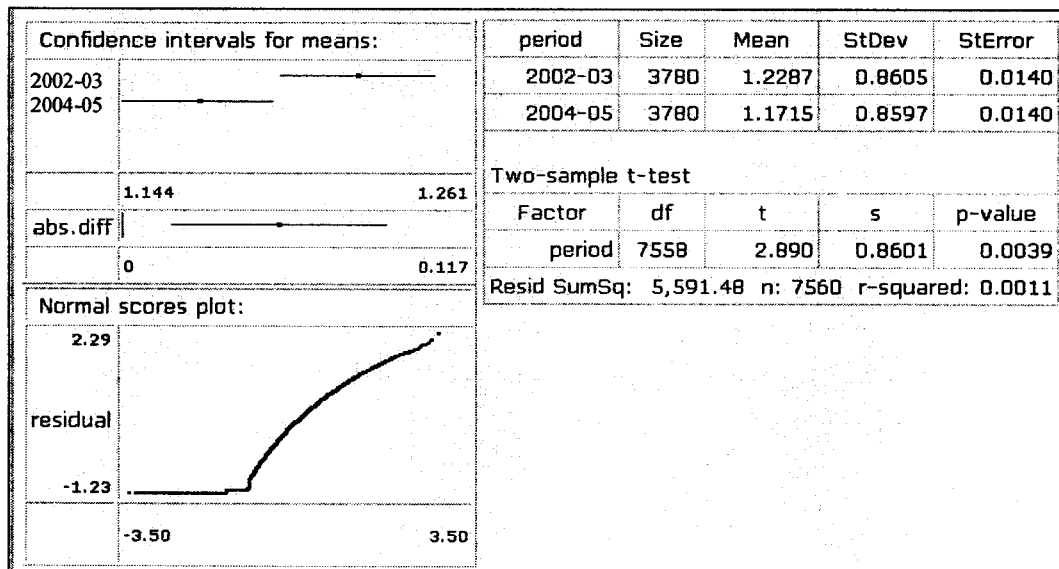


Figure: 3.3 Comparison of fertility rate by period

Figure 3.3 shows the association between the fertility rate and period. The association is statistically significant ( $p\text{-value} < 0.05$ ). The fertility rate in period 1 (2002-2003) was higher than in period 2 (2004-2005).

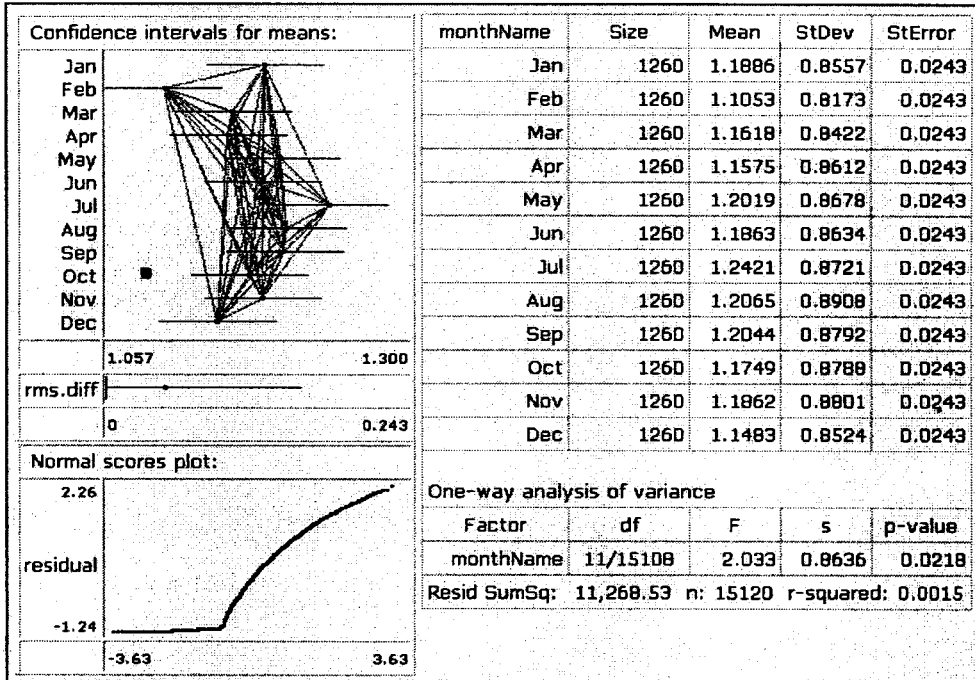


Figure 3.4: Comparison of fertility rate by month

Figure 3.4 shows the association between the fertility rate and month. The association is statistically significant ( $p\text{-value} < 0.05$ ). The fertility rates in July, August and September were higher than in another months.

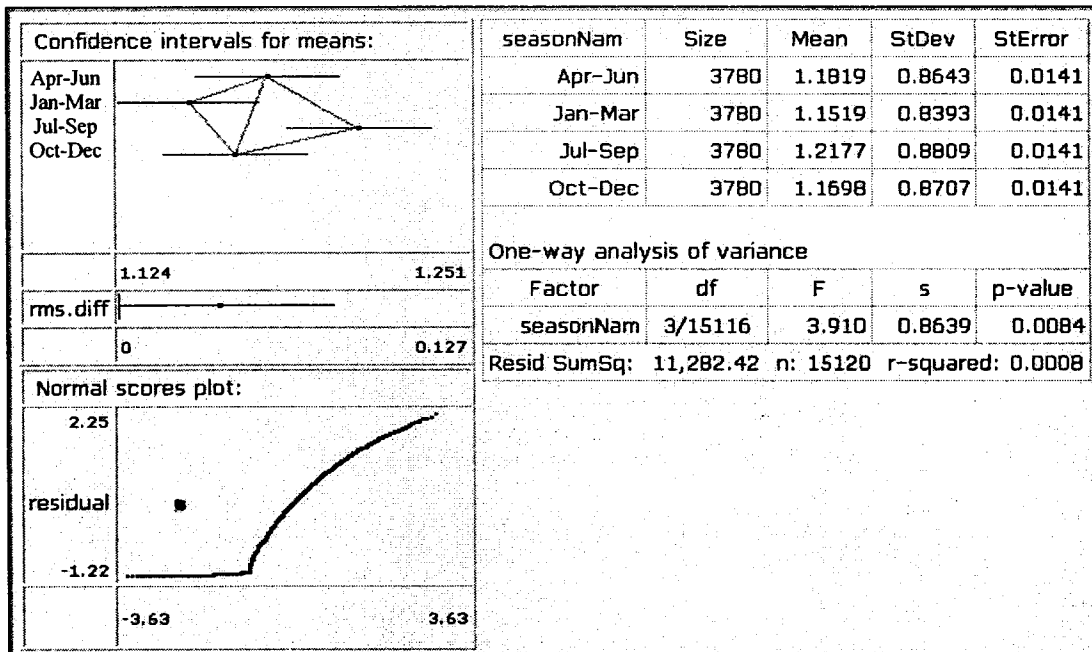


Figure 3.5: Comparison of fertility rate by season

Figure 3.5 shows the association between the fertility rate and season. The association is statistically significant ( $p\text{-value} < 0.05$ ). The fertility rates in the third season (July-September) were higher than other seasons.

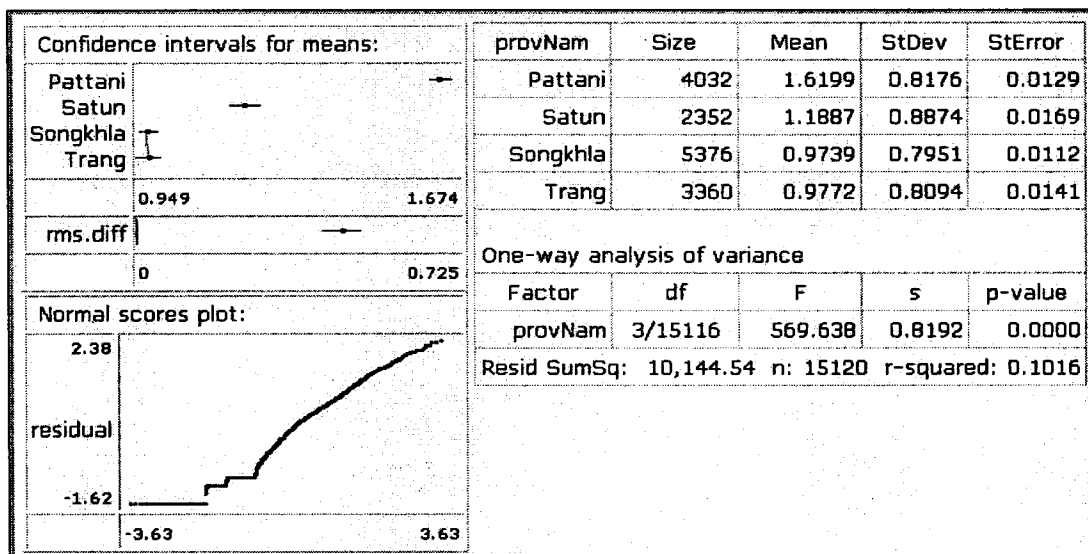


Figure 3.6: Comparison of fertility rate by province



Figure 3.6 shows the association between the fertility rate and province. The association is statistically highly significant ( $p\text{-value} < 0.05$ ). The fertility rates of Pattani were highest and the fertility rates between Songkhla and Trang were not different.

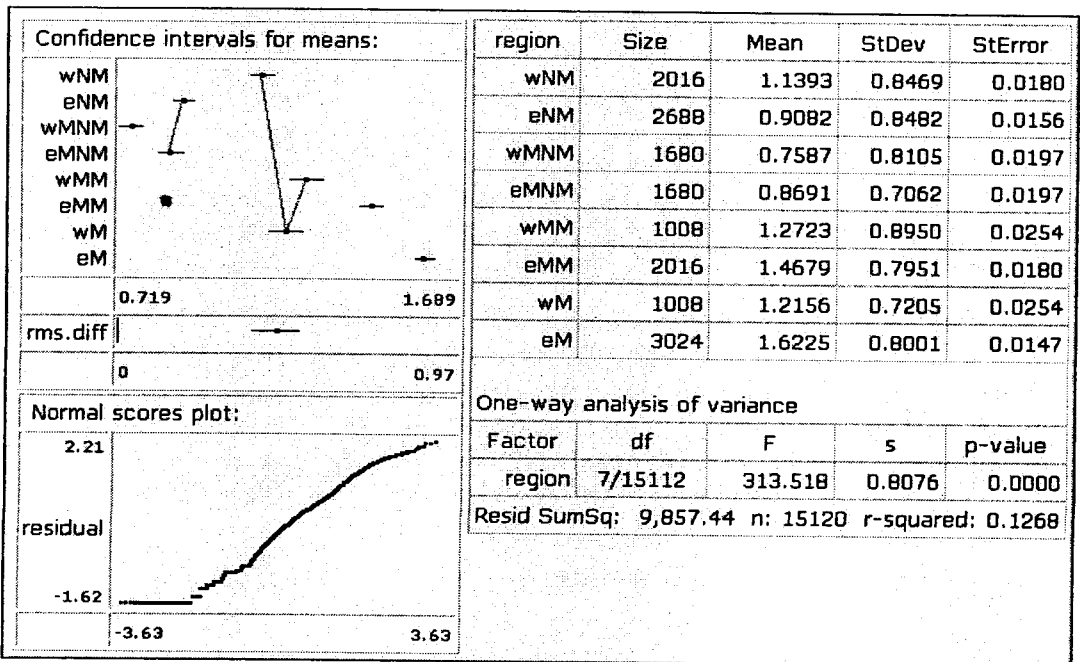
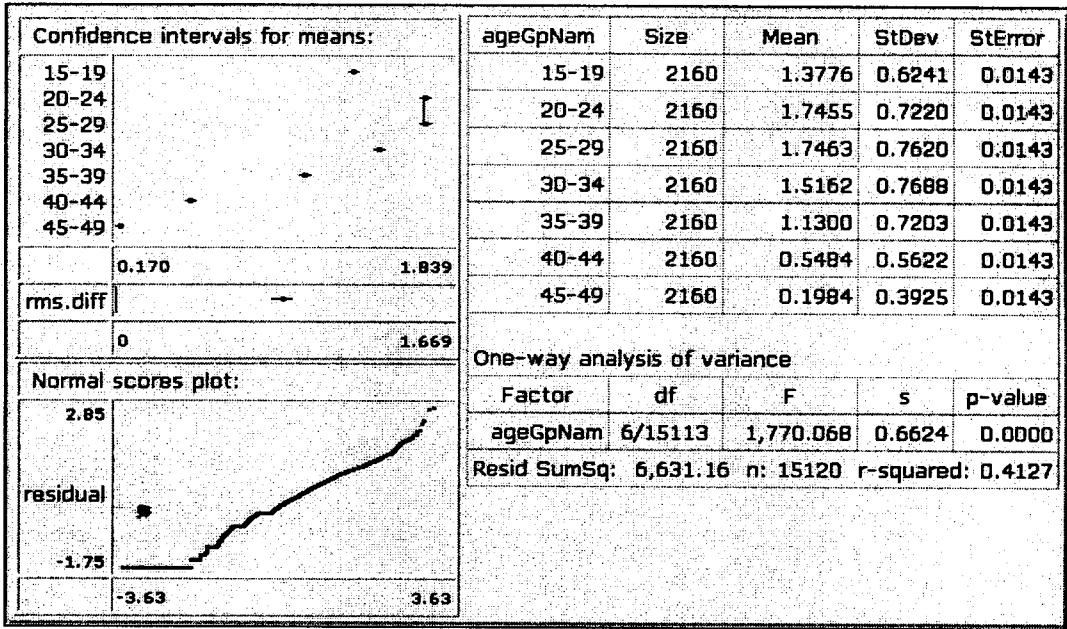


Figure 3.7: Comparison of fertility rate by region

Figure 3.7 shows the association between fertility rate and region. The association is statistically significant ( $p\text{-value} < 0.05$ ). The fertility rates of Muslims in the east coast (eM is Muslims more than 80% in the east region) were higher than Muslims in the west coast (wM is Muslims more than 80% in the west region). The fertility rates between non-Muslims in the east and west coast were not different.

Figure 3.8 shows the association between the fertility rate and mother's age group.



*Figure 3.8: Comparison of fertility rate by mother age group*

This association is statistically highly significant ( $p\text{-value} < 0.05$ ). The mother's age groups with the highest fertility rates were 20-24 and 25-29 years.

### 3.3 Total fertility rate

In this section, we calculate the total fertility rate (TFR). Table 3.4 shows the TFR and its age-specific components based on registered births in all four provinces for years 2002-2005. It indicates that the overall fertility appears to be near the replacement level (replacement level is 2.1)

year	age group							TFR
	15-19	20-24	25-29	30-34	35-39	40-44	45-49	
2002	0.248	0.481	0.523	0.394	0.212	0.066	0.023	1.946
2003	0.262	0.459	0.507	0.383	0.211	0.064	0.021	1.907
2004	0.285	0.459	0.521	0.389	0.206	0.067	0.015	1.943
2005	0.299	0.458	0.526	0.387	0.202	0.056	0.010	1.938

Table 3.4: Total fertility rate (TFR) and age-specific components based on births registration from four provinces

Table 3.5 shows the total fertility rate separated by demographic region. The estimates given in Table 3.5 are biased because many women moved from their district of residence to give birth in a hospital located in another region where the birth was subsequently registered, so calculating the birth rate in a region based on the number of mothers resident in the region (as in Table 3.5) yields rates that are biased upward or downward, depending on this flow of movement.

year	% Muslim	west coast				east coast			
		<20	20-49	50-79	80+	<20	20-49	50-79	80+
2002		2.050	1.126	2.187	1.432	2.218	0.942	2.202	2.424
2003		2.053	0.987	1.996	1.450	2.255	0.880	2.141	2.337
2004		2.136	0.951	2.234	1.336	2.391	0.965	2.150	2.116
2005		2.288	0.966	2.181	1.206	2.423	0.878	2.148	2.016

Table 3.5: Total fertility rate (TFR) by demographic region based on birth registrations

Estimates of the net movement of mothers between regions is shown in Table 3.6, which lists the numbers of registered births in each region on the bottom line and the population in these regions aged 0-4 years in the right-hand column. The numbers in the body of the table are estimates of the numbers of births registered to residents of the regions, obtained as follows. First, we estimate the total number of births to residents of each region in the 5-year period from 2001 to 2005 by reducing the populations in the age group 0-4 in the 2000 Population Census by the constant 3.05%, the factor needed to match the total number of registered births in the four provinces. If this total is less than the number of births registered in the region, we simply transfer it to its corresponding diagonal cell in the table. Next, we fill in the remaining diagonal cells using the corresponding totals in the bottom row. Finally, we fill in as many remaining cells as are necessary to make all the row and column totals match, based on assumptions about where mothers are likely to go if they register their births outside their region of residence.

Region	west	west	west	west	east	east	east	east	estimate	0-4 in
%Muslim	0-19	20-49	50-79	80+	0-19	20-49	50-79	80+	of total	2000
W 0-19	30611								30611	31574
W 20-49	4155	8988			4566				17709	18262
W 50-79			14850						14850	15317
W 80+			1661	5891	1548				9100	9386
E 0-19					51127				51127	52735
E 20-49					13820	13961			27781	28655
E 50-79					156		37072		37228	38399
E 80+					5682			37969	43651	45024
total	34766	8988	16511	5891	76899	13961	37072	37969	232057	239356

*Table 3.6: Estimates of numbers of births registered in eight demographic regions classified by region of residence of mother*

The information in Table 3.6 forms the basis for recalculating total fertility, by replacing the denominators used to obtain the estimates in Table 3.5 by weighted sums using the row proportions given in Table 3.4. Let  $P_{ijt}$  denote the number of women resident in region  $i$  (ordered as in the top row of Table 3.4) and age group  $j$  in month  $t$ . Then the adjusted estimate of the number of mothers in age-group  $j$  in month  $t$  for the east coast region with percent Muslim less than 20% ( $i = 5$ ), for example, is given by the formula  $0.258P_{2jt} + 0.170P_{4jt} + P_{5jt} + 0.497P_{6jt} + 0.004P_{7jt} + 0.130P_{8jt}$ .

Using the method described in Section 2 to estimate the total numbers of women resident in any specified month in 2002-2005 based on the numbers of women recorded in the 2000 population census, we thus obtain the numbers of women (in all seven age groups) resident in the eight regions in January 2002 (say) as  $P_1 = 108,767$ ,  $P_2 = 219,004$ ,  $P_3 = 55,578$ ,  $P_4 = 94,040$ ,  $P_5 = 49,021$ ,  $P_6 = 110,448$ ,  $P_7 = 27,714$  and  $P_8 = 107,117$ , respectively. Substituting these values into the above formula, we estimate the total number of women giving birth in the region to be 190,440. In the same way, estimates may be obtained of the numbers of women from each age group giving birth to babies registered in each region in each month, and these numbers provide appropriate denominators for obtaining relatively unbiased estimates of the total fertility rate.

We use a formula based on the row proportions  $p_{ij}$  in this table to get more accurate values of  $P_{ijt}$  using the formula.

$$P_{ijt}^{adjusted} = \sum_{k=1}^8 \rho_{ik} P_{kjt} \quad (3.1)$$

Table 3.7 gives the corresponding adjusted estimates of total fertility rate using this method. This gives more reasonable results.

% Muslim	west coast				east coast			
	<25	25-49	50-79	80+	<25	25-49	50-79	80+
year 2002	1.825	2.217	1.976	2.213	1.613	1.873	2.211	2.786
2003	1.827	1.943	1.803	2.242	1.634	1.750	2.150	2.686
2004	1.900	1.873	2.017	2.065	1.725	1.918	2.159	2.432
2005	2.034	1.901	1.967	1.865	1.741	1.745	2.157	2.318

*Table 3.7: Total fertility rate by demographic region based on birth registrations and numbers of mothers in denominator adjusted for region of registration*

The total fertility rate by demographic region as shown in Table 3.7 decreased substantially in the Muslim regions on both sides of the peninsula. There is a substantial difference between Muslim and Buddhist fertility on the east, but not on the west coast where Muslim fertility was below the replacement level in 2005.

age group	west coast				east coast			
	<20	20-49	50-79	80+	<20	20-49	50-79	80+
15-19	0.290	0.355	0.282	0.354	0.237	0.355	0.271	0.270
20-24	0.489	0.529	0.468	0.482	0.399	0.470	0.518	0.539
25-29	0.563	0.503	0.540	0.493	0.487	0.433	0.538	0.590
30-34	0.358	0.333	0.362	0.399	0.357	0.301	0.441	0.545
35-39	0.151	0.177	0.200	0.235	0.160	0.175	0.274	0.397
40-44	0.040	0.057	0.065	0.100	0.034	0.063	0.097	0.155
45-49	0.005	0.027	0.022	0.030	0.004	0.023	0.029	0.054
	1.896	1.981	1.940	2.093	1.678	1.821	2.169	2.551

*Table 3.8: Total fertility rate (TFR) by demographic region and mother's age group based on birth registrations*

Table 3.8 shows the total fertility rate by demographic region separated by mother's age group. Mothers aged 30 and over among Muslim in west coast have lower fertility rates than mother in east coast.