1) The mortality rates, as evident from the large experiences, are believed to vary smoothly with age; therefore, the crude estimate of mortality at any age carries information about the mortality rates at adjacent ages.
   By smoothing the experience, we can make use of data at adjacent ages to improve the estimate at each age.
   This reduces the sampling (or random) errors.
   This mortality experience may be used in financial calculations. Irregularities, jumps and anomalies in financial quantities (such as premium rates under life insurance contracts) are hard to justify to customers.
   The main limitation in mortality investigations that graduation will not be able to overcome is to remove any bias in the data arising from faulty data collection or otherwise.
2) 3 main aims of graduation are:
   i) To produce a smooth set of rates that are suitable for a particular purpose.
   ii) To remove random sampling errors.
   iii) To use the information available from adjacent ages.

   The three main features of graduation are:

   i) Smoothness.
   ii) Adherence to data.
   iii) suitability for the purpose in hand.
3)
   i) The crude mortality rates derived from the data will progress more or less roughly. However, it is expected that mortality rates would be smooth functions of age. Hence it follows that a crude estimate at any age 'x' also carries information about the mortality rates at age 'x-1' and 'x+1'.
   Graduating the crude rates will enable this information to be captured, thus resulting in the mortality rates progressing smoothly over age.
   The mortality rates would be used in computing premium rates/annuity rates. If the underlying mortality rates do not progress smoothly, then the derived premium rates etc would also fluctuate randomly, which would be difficult to understand/explain to potential customers.
   The purpose of graduation is:
   - To produce a smooth set of rates that are suitable for a practical purpose.
   - To remove random sampling errors.
   - To use the information available from adjacent ages.
   ii) The methods used for graduation are:
   - Graduation by parametric formula.
   - Graduation by reference to a standard table.
   - Graphical graduation.
   iii) The null hypothesis is that the graduated rates represent the mortality of the members in the pension scheme.

$$\textit{To test the appropriateness of the graduation using chi} - \textit{squared test, we compare} \sum_x z_x^2 \textit{ with } \chi^2 \textit{ with 'm' degrees of freedom.}$$

Age Group	Pensioners	Actual Deaths	Exp Deaths	Stddevn	zx2
50-54	37259	248	227.28	1.3786	1.900553
55-59	28057	392	367.55	1.2839	1.648501

60-64	25654	680	672.13	0.3074	0.094513
65-69	20475	987	997.13	-0.3290	0.108234
70-74	16219	1380	1360.77	0.5445	0.296513
75-79	11843	1625	1584.59	1.0906	1.189511
80-84	7535	1564	1487.57	2.2121	4.893339
85-89	3294	925	891.36	1.3195	1.740956
90+	450	130	155.48	-2.5254	6.377613

The number of age-groups is 9. As the graduation has been done with reference to a standard table, some degrees of freedom are lost.
So, m < 9, say m = 8
Hence, we reject the null hypothesis.

$$The\ values\ of\ \sum_x z_x^2\ is\ 18.2497, which\ is\ higher\ than\ the\ critical\ value\ of\ chi-squared\ distribution\ with\ 8\ degrees\ of\ freedom\ at\ 5\ percent\ level\ (which\ is\ 15.51).$$

iv) Any two from the below:
Signs Test- The signs test checks for any overall bias, whether the graduated rates are too high or too low.
The number of positive signs (or negative) is a Binomial distribution B (9,0.5). There are 7 positive deviations and the probability of obtaining 7 or more positive signs is 0.01953 and since this is lower than 2.5%, we reject the null hypothesis and conclude that the graduated rates do not adhere to the data.
Individual Standardized deviations test- The Individual Standardised Deviations tests looks for individual large deviations at particular ages.
If the graduated rates were the true rates underlying the observed rates, we would expect the individual deviations to be distributed Normal (0,1) and therefore only 1 in 20 zxs should have absolute magnitudes greater than 1.96. Looking at the zxs we see that there are 2 deviations higher than 1.96 and hence we reject the null hypothesis that the graduated rates are the true rates underlying the crude rates.

v) The graduated rates are below the crude mortality rates for most of the ages, the exceptions being 65-69 and 90+ age group.
Since the graduated rates under-estimate, the actual mortality experience, use of the graduated rates may result in over-provisioning of the annuity liabilities under the pension scheme.

4)

a) Cumulative deviations test
H0 = the standard table rates are the true underlying mortality rates for the term assurance policyholders
We first calculate the individual deviations using the formula:

$$z \approx \frac{\theta_x - E_x q_x^s}{\sqrt{E_x q_x^s (1 - q_x^s)}} \approx \frac{\theta_x - E_x q_x^s}{\sqrt{E_x q_x^s}}$$

$$Since\ q_x^s\ \approx 0, we\ can\ use\ the\ approximation\ that\ 1 - q_x^s \approx 1$$

x	Ex	dx	qx (s)	Ex . qx (s)	dx - Ex.qx
40	50,000	87	0.2053%	102.65	-15.65
41	48,560	84	0.2247%	109.11	-25.11
42	47,190	101	0.2418%	114.11	-13.11
43	44,100	112	0.2602%	114.75	-2.75
44	43,600	123	0.2832%	123.48	-0.48
45	40,400	110	0.3110%	125.64	-15.64
46	37,280	108	0.3438%	128.17	-20.17
47	35,370	122	0.3816%	134.97	-12.97
48	32,100	150	0.4243%	136.20	13.80
49	29,000	139	0.4719%	136.85	2.15
50	26,200	151	0.5244%	137.39	13.61
				1,363.32	**-76.32**

$Test\ statistic = \frac{\sum \theta_x - E_x q_x^s}{\sqrt{\sum E_x q_x^s}}$=-2.07

Under the null hypothesis, this has a normal distribution. Since this is a two-tailed test, we compare the test statistic, at 95% confidence level: |-2.07| > 1.96. Therefore, we reject the null hypothesis and conclude that the standard table rates do not adequately reflect the true mortality rates for the term assurance policyholders.

5)

i)

a) Maximum smoothness would be achieved by ignoring the plotted estimated rates and drawing a graduation curve which is very smooth, e.g., straight line. The deviations between the rates read from such a curve and the observed rates are likely to be very large. The graduation curve will be very smooth but have poor adherence to data.
On the other hand, joining up a plot of the estimated values will give perfect adherence to data but is likely to produce a curve with rapidly changing curvature which would not satisfy the smoothness criteria.

b) Graduation aims to resolve these conflicts by striking a fair balance between the adherence to data and smooth progression of rates.
Graduated rates can be obtained by many methods, some ensure smoothness, e.g., graduation by a mathematical form (the chosen functional form will ensure smoothness), reference to a standard table (a simple relationship with an already smooth set of standard table rates will ensure smoothness). In this case the graduated rates just need to satisfy tests of adherence to data.

Graphical graduation does not ensure smoothness, so graduated rates must be checked for smoothness and adherence to data. The graduation process must be repeated until both criteria are satisfied.

ii) By rearranging we get

loge(qx/px) = a+b*x

(qx/px) = exp (a+b*x)

qx/(1-qx) = exp (a+b*x)

qx = 1/(1+exp(-(a+bx)))

Age	a+bx	qx	E	A	A-E	(A-E)^2/E
30	-7.0140	0.000898	224.60	245.00	20.40	1.85
31	-6.8978	0.001009	252.25	276.00	23.75	2.24
32	-6.7816	0.001133	283.29	313.00	29.71	3.12
33	-6.6654	0.001273	318.16	323.00	4.84	0.07
34	-6.5492	0.001429	357.30	339.00	-18.30	0.94
Total			1,435.60	1,496.00	60.40	8.22

Chi-square value: 8.22

However, we have used two estimated parameters. So, the number of degrees of freedom to use is 5-2 = 3. The upper 5% point of chi square (3) distribution is 7.815. So, on the basis of chi square test, our hypothesis of good fit is rejected.

6)

a) The low volume of data is the principal problem in the given situation, hence;

- The crude rates may not be suitable for the purpose.
- There are random sampling errors.
- The rate at a particular age cannot be linked to the rates at adjacent ages.
- Rates will not progress smoothly from age to age which allows a practical smooth set of premium rates to be produced.

b)

Age	z_x	z_x^2	$\mathring{\mu}_{x+1/2}$	$\Delta \mathring{\mu}_{x+1/2}$	$\Delta^2 \mathring{\mu}_{x+1/2}$	$\Delta^3 \mathring{\mu}_{x+1/2}$
18-22	0.5249	0.27552	0.0061	0.0070	0.0061	0.0033
23-27	0.3615	0.13068	0.0131	0.0131	0.0094	0.0033
28-32	1.0266	1.05391	0.0262	0.0225	0.0127	0.0020
33-37	-1.0912	1.19072	0.0487	0.0352	0.0147	-0.0009
38-42	1.2394	1.53611	0.0839	0.0499	0.0138	-0.0044
43-47	0.5949	0.35391	0.1338	0.0637	0.0094	-0.0076
48-52	0.6346	0.40271	0.1975	0.0731	0.0018	
53-57	0.1787	0.03193	0.2706	0.0749		
58+	-0.0367	0.001349	0.3455			
		4.976839				

$$Test\ Statistics: X = \sum z_i^2$$

Null Hypothesis: H0: X has a Chi Square distribution

The observed value of X is 4.98

The degrees of freedom are equal to the number of age groups less some allowances for the method of graduation. So, degree of freedom is 9 (Max).

This is a one-sided test. Upper 5% point of Chi Square distribution exceeds the observed value 4.98 for all degrees of freedom (<= 9) except one. There is insufficient evidence to rejectH0.

$The\ third\ difference\ of\ \mu^{o}_{x+\frac{1}{2}}\ is\ not\ very\ insignificant\ and$

$hence\ the\ graduates\ rates\ are\ not\ very\ smooth.$

7)

8)

9)

i) H0: Standard mortality table represents the true underlying mortality rates of the experience.

Age: x	*Initial exposed to risk:* Ex	*Standard Mortality rates* qx	*Expected Deaths* E	*Actual deaths* A	*A-E*	*Chi-squared*
50	2305	0.0064	14.75	15	0.25	0.00417
51	2475	0.0069	17.08	16	-1.08	0.06798
52	2705	0.0075	20.29	22	1.71	0.14455
53	2900	0.0081	23.49	23	-0.49	0.01022
54	3170	0.0087	27.58	27	-0.58	0.01216
55	6730	0.0094	63.26	66	2.74	0.11850
56	6875	0.0101	69.44	67	-2.44	0.08556
57	8190	0.0109	89.27	88	-1.27	0.01810
58	8200	0.0117	95.94	102	6.06	0.38278
59	7680	0.0119	91.39	80	-11.39	1.42001
60	7160	0.0121	86.64	85	-1.64	0.0309
Total	58390		599.12	591.00	-8.12	2.295

a) If H0 is true, the Chi-squared statistic will follow the χ^2 sampling distribution.
Since we are an experience with a standard table, the degrees of freedom = 11
Observed values of χ^2 = 2.295
Tabular value of χ^2 for upper 95% point = 19.68
The observed value of test statistic is less than 5% significance point. Hence, there is no evidence to reject the null hypothesis.

b) If H0 is true test statistic Z(c) ~ N (0, 1)
Here, observed value of Z(c) = -8.12/√599.12 = -0.332
This is a two-tailed test. We compare the value of statistic with the upper and lower 2.5% points of N (0,1). As -1.96 < -0.332 < 1.96, there is insufficient evidence to reject the null hypothesis.

c) If H0 is true:

Age	*zx*	*zx^2*
50	0.0646	0.00417
51	-0.2607	0.06798
52	0.3802	0.14455
53	-0.1011	0.01022
54	-0.1103	0.01216
55	0.3442	0.11850
56	-0.2925	0.08556
57	-0.1345	0.01810
58	0.6187	0.38278
59	-1.1916	1.42001
60	-0.1758	0.03089

Number of groups of positive signs = 4
Number of positive signs = 4
Number of negative signs = 7
From the table, we can find that the value for k for which

$$\sum_{t=1}^{k} \frac{\binom{n1-1}{t-1}\binom{n2+1}{t}}{\binom{m}{n1}} < 0.05$$

Where n1=4; n2=7 & m=4+7=11
The tabular value of k is 1 which is less than number of groups of signs and hence there is no evidence to reject the null hypothesis.

10)

i) The hypothesis requires the expected claims to be calculated based on the observed mortality experience in the last year.
E(A_LY) = observed mort rate (LY) * Exp_risk_CY = Deaths_LY/Exp_risk_LY * Exp_risk_CY
This requires the exposed to risk at each age. As the exposed to risk is not directly available the expected deaths based on actual experience in last year can be derived as
E(A_LY) = Deaths_LY/(Exp_risk_LY * std_mort_rate) *(Exp_risk_CY * std_mor_rate) = Deaths_LY/Exp_deaths_LY * Exp_deaths_CY

Age	Deaths_LY	Exp_deaths_LY	Exp_deaths_CY	E (A_LY) E	Act_claims A	X^2 = (A-E)^2/E
25	808	810	724	722	719	0.01
30	1,851	1,708	1,433	1,553	1,322	34.46
35	1,400	1,084	444	573	500	9.28
40	1,562	1,705	1,397	1,280	1,207	4.15
45	1,366	1,572	1,465	1,273	1,177	7.24
50	1,296	1,643	1,209	954	905	2.48
55	2,200	2,911	2,436	1,841	1,798	1.00

There are 7 ages. Hence the degrees of freedom is 7.
The upper 95% for X2 7for 7 degrees of freedom is 15.5. The observed value is 58.6 which is much higher than this, hence the hypothesis is rejected.

11) The null hypothesis is that the graduated rates are the same as the true underlying rates for the block of business.

$$\chi_n^2 = \sum_{x}^{all\ ages} z_x^2$$

where n is the degrees of freedom.

Age	Exposed to risk	Observed Death	Graduated Rates (qx)	Expected Death	Zx	Zx^2
20	1060	12	0.0125	13.2500	-0.3434	0.1179
21	1250	14	0.0118	14.7500	-0.1953	0.0381
22	1210	16	0.0134	16.2140	-0.0531	0.0028
23	700	9	0.0102	7.1400	0.6961	0.4845
24	875	11	0.0111	9.7125	0.4131	0.1707
25	950	15	0.0137	13.0150	0.5502	0.3027
26	805	10	0.0126	10.1430	-0.0449	0.0020
27	1390	16	0.0127	17.6530	-0.3934	0.1548
28	1080	17	0.0122	13.1760	1.0535	1.1098
29	1310	14	0.0131	17.1610	-0.7631	0.5822
			Total		0.9197	2.9657

Observed Test Statistic is 2.9657

The number of age groups (n) is 10, but we lose at least one number of degrees since we are comparing our data with a set of graduated data derived from our data, perhaps 2. So, the number of degree of freedom is 8.

The critical value of the chi-squared distribution with 8 degrees of freedom at the 5% level is 15.51.

Since 2.9657 < 15.51

We do not reject the null hypothesis.

12)

i) Role of Graduation in producing life table:

- Produce smooth set of results that are suitable for a particular purpose.
- Remove the random sampling errors.
- Use the information available from adjacent ages to improve the reliability of estimates. This is particularly important at the older ages where exposure numbers are small and data are sparse.

Procedure for graduating the rates:

- Choose a method of graduation. The preferable method is the graduation by parametric formula as the availability of data is large.
- Then select a statistical model to provide the graduated rates. This step involves selecting a graduation formula and determine the parametric values and then calculate the graduated rates.
- Test the graduated rates – the graduated rates must be compared with original data to see if they are acceptably close according to some graduation tests. Some general checks like whether mortality of male at a particular age is higher than the mortality of female at that age can also be performed in this stage.
- The process will be repeated with different models and/or with different number of parameters till we reach a suitable level of graduated rates. The final choice will be influenced by the level of goodness of fit (adherence to

data) required against the smoothness required in the rates produced. As the purpose is to produce a national life table for general use, the balance will be tilted towards goodness of fit.

ii) Shape of the mortality curve over the range of ages and the level of mortality rates.

iii) This is a test for over graduation. The above test detects grouping of signs of deviations. It does this by analysing the relationship between the deviations at nearby ages taking into account the magnitude of the values. This test will address the problem of the inability of the chi square test to detect excessive clumping of deviations of the same sign.

iv)

Age	E	A	z_x	z_{x+1}	$A = z_x - \bar{z}$	$B = z_{x+1} - \bar{z}$	A2	AB
60	38.3	36	-0.37165	-1.01415	-0.40343	-1.04593	0.162758	0.421964
61	40.45	34	-1.01415	-0.69317	-1.04593	-0.72496	1.093978	0.758262
62	42.52	38	-0.69317	-0.68311	-0.72496	-0.7149	0.525569	0.518275
63	44.56	40	-0.68311	0.6332	-0.7149	0.601412	0.511083	-0.42995
64	47.63	52	0.6332	-0.45398	0.601412	-0.48577	0.361697	-0.29215
65	51.25	48	-0.45398	0.221781	-0.48577	0.189993	0.23597	-0.09229
66	55.35	57	0.221781	0.585212	0.189993	0.553424	0.036098	0.105147
67	60.45	65	0.585212	0.726943	0.553424	0.695155	0.306278	0.384716
68	63.22	69	0.726943	0.755509	0.695155	0.723721	0.483241	0.503099
69	67.78	74	0.755509	0.643078	0.723721	0.61129	0.523772	0.442403
70	71.56	77	0.643078		0.61129		0.373676	0

z_x is calculated using the formula $z_x = (A-E) / \text{SQRT}(E)$ derived from the Poisson Distribution assumption.

Now $r(j) = \sum_{i=1}^{m-j} (z_x - \bar{z}) * (z_{x+1} - \bar{z}) / \{ (m-j)/m * \sum_{i=1}^{m} (z_x - \bar{z})^2 \}$

$\Rightarrow r_1 = 0.55296$

And the T ratio is $r_1 * \text{Sqrt}(11) = 1.83$

The T Ratio is positive which suggest that the rates are over graduated and value 1.83 < 1.96 and hence we do not reject the Null Hypothesis that the graduated rates are the true rates underlying the observed data.

13)

a) Signs Test- The number of positive signs is 8 and negative is 12. The number of signs is Binomial (20,1/2). The probability of obtaining 8 positive signs in 20 observations is 2 * 0.2517, which is higher than 5%. Hence there is no evidence to reject the graduated values.

b) Grouping of signs test: The number of positive deviations is 8 and negative is 12. From the tables, the critical value when n1 is 8 and n2 is 12 is 2. Since we have observed 2 groups of positive deviations, we reject the graduated rates and conclude there is evidence of grouping of deviations of the same sign.

14)

i) “Under graduation” occurs when too much emphasis is given to goodness of fit. Under graduated rates adhere closely to the crude rates, but the resulting rates do not show a smooth progression from age to age.

“Over graduation” occurs when too much emphasis is given to smoothness. Over graduated rates show a smooth progression from age to age, but the resulting rates do not adhere closely to the crude rates.

ii) The chi-squared test is for the overall fit of the graduated rates to the data. The test statistics is ∑ zx^2, where

$$z_x = \frac{(\theta_x - E_x \mathring{q}_x)}{\sqrt{E_x \mathring{q}_x (1 - \mathring{q}_x)}}$$

However, considering the fact that qx is very small, we redefined the zx as below

$$z_x \approx \frac{(\theta_x - E_x \mathring{q}_x)}{\sqrt{E_x \mathring{q}_x}}$$

The calculations are given in the table below:

Average Age	**Θx**	**qx**	**qx**	**Expected Deaths**	**Zx**	**zx^2**
23	2	0.00889	0.00220	1.98	0.01421	0.00020
28	4	0.00833	0.00240	2.88	0.65997	0.43556
33	5	0.00923	0.00260	3.38	0.88116	0.77645
38	7	0.01000	0.00360	5.40	0.68853	0.47407
43	8	0.01091	0.00540	5.94	0.84523	0.71441
48	9	0.01500	0.00900	7.20	0.67082	0.45000
53	9	0.01385	0.01500	9.75	-0.24019	0.05769
58	5	0.01429	0.02300	8.05	-1.07498	1.15559

$$\sum z_x^2 = 4.06397$$

The test statistic has a chi-squared distribution with degrees of freedom given by number of age groups less 1 for the parametric function and further reduction for using the standard table.

The critical value of chi-squared distribution with 6-degree freedom at 5% level is 12.59.

Since 4.06397 < 12.59, there is no evidence to reject the null hypothesis that the graduated rates are the true rates underlying the crude rates.

iii) Signs test

a) The Signs test looks for overall bias.

b) If the null hypothesis is true, the number of positive signs is distributed Binomial (8, 0.5). From the table, we observe that there are 6 positive signs. Prob (observed number of positive signs <=6) = 1- Prob(positive signs >6).

$$= 1 - \left\{\binom{8}{7} + \binom{8}{8}\right\} * (0.5)^8 = 1 - 0.035156 = 0.965$$

This is greater than 0.025 (two tailed test).

c) We cannot reject the null hypothesis and conclude that the graduated rates are not systematically higher or lower than the crude rates.

Grouping of Signs test

a) The grouping of signs test looks for run or clumping of deviations with the same sign for over graduation.

b) We have total 8 age groups with 6 positive signs and 2 negative signs. There is only one run in this analysis.

Pr (one positive run) =

$$\frac{\left\{\binom{5}{0}\binom{3}{1}\right\}}{\binom{8}{6}} = 3/28 = 0.10714$$

This is greater than 0.05 (using one-tailed test).

c) We accept the null hypothesis that graduated rates are true underlying the crude rates.

15)

a) Null Hypothesis Ho: The true underlying rates of withdrawal are the graduated rates.

Chi Square test

t	Z_t^2
0	11.77
1	2.69
2	5.01
3	0.45
4	1.07
5	1.62
6	2.01
7	10.11
8	0.47
9	1.88
	37.07

Degrees of freedom = 10 – 3(no. of parameters) = 7

Now Pr (X7^2 >= 20.3) = 0.005

The tabulated value with 7 degrees of freedom at upper 5% point is 14.07.

37.07 > 14.07

Hence Ho is strongly rejected.

b) Standardised deviation test

According to the null hypothesis, each standardized deviation should be a random event from a unit normal distribution.

t	Standardised deviation Zt
0	3.43
1	-1.64
2	-2.24
3	0.67
4	1.03
5	-1.27
6	1.42
7	3.18
8	-0.69
9	-1.37

SD	Exp No	Observed No
-∞ → -3	0	0
-3 → -2	0.2	1
-2 → -1	1.4	3
-1 → 0	3.4	1
0 → 1	3.4	1
1 → 2	1.4	2
2 → 3	0.2	0
3 → ∞	0	2

The above is far removed from a unit normal distribution. Hence reject Ho. Graduated rates clearly not a good representation of the crude rates.

c) Sign Test

According to Ho, we would expect that the number of positive deviations to have a binomial distribution with B (10,0.5). Expected no equal to 5.

Above we observe 5 positive and 5 negative deviations. This is acceptable. Hence not enough evidence to reject Ho.

Summary Conclusion:

- Chi squared test and individual standardized deviation tests indicate that, adherence to data at many of individual durations is not acceptable. This is particularly the case at duration 0,2 and 7.
- Satisfactory sign test indicates that the graduation runs centrally through the data.
- Grouping of signs test if carried out would indicate if there are clumps of deviation over wide duration ranges.
- Summing up, it appears that data cannot be adequately represented by a quadratic function. There are severe adherence problems at duration 0-2 and 6-7. Smoothness however is expected to be satisfactory as quadratic function has been fitted.