Question 1 (20 marks) Is there a statistical significant relationship between attrition and each of the following variables: age and TotalWorkingYear? Perform appropriate hypothesis test.

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.978215	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.082286	Pr > D	<0.0100
Cramer-von Mises	W-Sq	1.327873	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	8.07308	Pr > A-Sq	<0.0050

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.945707	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.119937	Pr > D	<0.0100
Cramer-von Mises	W-Sq	0.690223	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	3.959025	Pr > A-Sq	<0.0050

The above two tables show the Normality of the age variable when the Attrition is No and Yes respectively. Since the p-values of both are less than 0.05, we cannot assume that the distribution of age is a normal distribution. But the Size of both is greater than 30 (in the case of attrition=no, the size of age is 1233, and in the case of attrition=yes, the size of age is 237), therefore, we can consider the sampling distribution to be approximately normal.

Hypothesis test is constructed as below:

1. H0: $\mu1 - \mu2 = 0$, $\mu1$ is mean age of the group of attrition is NO while $\mu2$ is mean age of the group of attrition is Yes
2. H1: $\mu1 - \mu2 \neq 0$
3. α: 0.05

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	1468	6.18	<.0001
Satterthwaite	Unequal	316.93	5.83	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	236	1232	1.19	0.0763

Because p-value = 0.0763 > 0.05, equal variance can be assumed, so the pooled method is used, because the p-value of the pooled method < 0.0001, therefore, at 5% level of significance, H0 can be rejected. So, we can assume that the mean ages of these two groups are different, so age and attrition are statistically related.

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.909822	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.195826	Pr > D	<0.0100
Cramer-von Mises	W-Sq	6.922033	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	37.38619	Pr > A-Sq	<0.0050

Tests for Normality				
Test	Statistic		p Value	
Shapiro-Wilk	W	0.85	Pr < W	<0.0001
Kolmogorov-Smirnov	D	0.171224	Pr > D	<0.0100
Cramer-von Mises	W-Sq	1.373753	Pr > W-Sq	<0.0050
Anderson-Darling	A-Sq	8.4992	Pr > A-Sq	<0.0050

Because p-value = 0.0763 > 0.05, equal variance can be assumed, so the pooled method is used. Because the p-value of the pooled method is < 0.0001, therefore, at 5% level of significance, H0 can be rejected. So, we can assume that the mean ages of these two groups are different, so age and attribute are statistically related.

Hypothesis test is constructed as below:

1. H0: $\mu 1 - \mu 2 = 0$, $\mu 1$ is mean TotalWorkingYears of the group of attrition is NO while $\mu 2$ is mean TotalWorkingYears of the group of attrition is Yes
2. H1: $\mu 1 - \mu 2 \neq 0$
3. α: 0.05

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	1468	6.65	<.0001
Satterthwaite	Unequal	350.88	7.02	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	1232	236	1.17	0.1270

Because p-value = 0.1270 > 0.05, equal variance can be assumed, so the pooled method is used, because the p-value of the pooled method < 0.0001, therefore, at 5% level of significance, H0 can be rejected. So, we can assume that the mean TotalWorkingYears e of these two groups are different, so TotalWorkingYears and attrition are statistically related.

The result of procedure for the raw data contains more than 20% of the cells have expected counts that are less than 5, so a transformation was done for age according to the rule of age < 30 = young, age > 50 = old and middle for others.

Assume, each sample used is independent, and each sample can only fall into just one of a

finite number k of complementary and mutually exclusive outcomes of agegroup.

The FREQ Procedure

Frequency Expected Cell Chi-Square	Table of Attrition by AgeGroup			
	AgeGroup			
Attrition	middl	old	young	Total
No	822 789.29 1.3558	125 119.94 0.213	286 323.77 4.4055	1233
Yes	119 151.71 7.0534	18 23.055 1.1084	100 62.233 22.92	237
Total	941	143	386	1470

Statistics for Table of Attrition by AgeGroup

Statistic	DF	Value	Prob
Chi-Square	2	37.0562	<.0001
Likelihood Ratio Chi-Square	2	34.2815	<.0001
Mantel-Haenszel Chi-Square	1	32.8490	<.0001
Phi Coefficient		0.1588	
Contingency Coefficient		0.1568	
Cramer's V		0.1588	

Sample Size = 1470

Hypothesis test is constructed as below:

4. H0: Attrition and AgeGroup are independent
5. H1: Attrition and AgeGroup are not independent
6. α: 0.05

From the above table, the test statistic is χ 2 = 37.0562 with 2 degrees of freedom. The corresponding P-value < 0.0001. Therefore, At 5% level of significance, H0 can be rejected. There is statistically significant relationship between attrition and AgeGroup. The biggest difference is from “young” and “yes” group.

For the same reason, the result of procedure of total working years does also contains more than 20% counts with expected value less than 5, so a transformation should be done for that variable.

Assume, each sample used is independent, and each sample can only fall into just one of a finite number k of complementary and mutually exclusive outcomes of workingyeargroup.

The FREQ Procedure

Frequency
Expected
Cell Chi-Square

Table of Attrition by WorkingYearsGroup

Attrition	WorkingYearsGroup 0	1	2	3	Total
No	564 604.76 2.7468	450 429.45 0.9831	171 154.33 1.7995	48 44.455 0.2827	1233
Yes	157 116.24 14.29	62 82.547 5.1144	13 29.665 9.3622	5 8.5449 1.4706	237
Total	721	512	184	53	1470

Statistics for Table of Attrition by WorkingYearsGroup

Statistic	DF	Value	Prob
Chi-Square	3	36.0496	<.0001
Likelihood Ratio Chi-Square	3	37.8665	<.0001
Mantel-Haenszel Chi-Square	1	31.0456	<.0001
Phi Coefficient		0.1566	
Contingency Coefficient		0.1547	
Cramer's V		0.1566	

Sample Size = 1470

Hypothesis test is constructed as below:

1. H0: Attrition and WorkingYearGroup are independent
2. H1: Attrition and WorkingYearGroup are not independent
3. α: 0.05

From the above table, the test statistic is χ 2 = 36.0496 with 3 degrees of freedom. The corresponding P-value < 0.0001. Therefore, at 5% level of significance H0 can be rejected. There is statistically significant relationship between attrition and WorkingYearGroup. The biggest difference is from “0” and “y” group.

Attrition	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
No		11.8629	11.4293	12.2965	7.7607	7.4660	8.0798
Yes		8.2447	7.3273	9.1622	7.1692	6.5767	7.8800
Diff (1-2)	Pooled	3.6182	2.5513	4.6851	7.6687	7.4011	7.9565
Diff (1-2)	Satterthwaite	3.6182	2.6044	4.6320			

Based on the table above, we are 95% confident that for employees who did not leave, their average TotalWorkingYears was between 11.4293 and 12.2965. Conversely, we are 95% sure that employees who leave have an average TotalWorkingYears between 7.3273 and 9.1622.

Attrition	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
No		37.5612	37.0646	38.0578	8.8884	8.5509	9.2538
Yes		33.6076	32.3677	34.8475	9.6893	8.8885	10.6500
Diff (1-2)	Pooled	3.9536	2.6985	5.2088	9.0219	8.7071	9.3606
Diff (1-2)	Satterthwaite	3.9536	2.6189	5.2883			

Based on the table above, we are 95% confident that for employees who did not leave, their average age was between 37.0646 and 38.0578. Conversely, we are 95% sure that employees who leave have an average age between 32.3677 and 34.8475.

Question 2 (30 marks) Is there a statistically significant relationship between attrition and each of the following variables: EducationField, BusinessTravel and JobRole? Perform appropriate hypothesis tests and illustrate statistically significant relationships using mosaic plots.

The FREQ Procedure

Frequency
Expected
Cell Chi-Square

Table of Attrition by EducationField

Attrition	EducationField						
	Human Resourc	Life Sciences	Marketing	Medical	Other	Technical Deg	Total
No	20 22.647 0.3094	517 508.3 0.149	124 133.37 0.6577	401 389.19 0.3583	71 68.78 0.0717	100 110.72 1.0376	1233
Yes	7 4.3531 1.6095	89 97.702 0.7751	35 25.635 3.4215	63 74.808 1.8639	11 13.22 0.3729	32 21.282 5.3982	237
Total	27	606	159	464	82	132	1470

Statistics for Table of Attrition by EducationField

Statistic	DF	Value	Prob
Chi-Square	5	16.0247	0.0068
Likelihood Ratio Chi-Square	5	14.9001	0.0108
Mantel-Haenszel Chi-Square	1	1.0587	0.3035
Phi Coefficient		0.1044	
Contingency Coefficient		0.1038	
Cramer's V		0.1044	

Sample Size = 1470

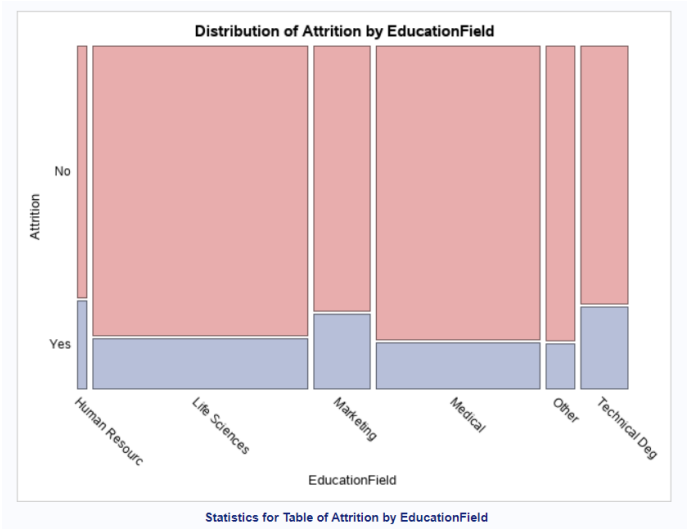
According to the figure above, there is only one count with expected value less than 5, which is only 10% of the total counts, so we can still perform chi-square test directly.

Assume, each sample used is independent, and each sample can only fall into just one of a finite number k of complementary and mutually exclusive outcomes of EducationField.

Hypothesis test is constructed as below:

1. H0: Attrition and EducationField are independent
2. H1: Attrition and EducationField are not independent
3. α: 0.05

From the above table, the test statistic is χ 2 = 16.0247 with 5 degrees of freedom. The corresponding P-value = 0.0068 which is smaller than 0.05. Therefore, at 5% level of significance H0 can be rejected. There is statistically significant relationship between attrition and EducationField. The biggest difference is from “Technical Deg” and “yes” group.



That is the MOSAIC diagram, from the diagram, statistically speaking, we can find that not all EducationFields have the same Attrition for the same type of proportion. It is not difficult to see from the above figure that when EducationField is Life Sciences, Medical and Other, the proportion of the value of attrition is no is higher than the other cases.

The FREQ Procedure

Frequency
Expected
Cell Chi-Square

Table of Attrition by BusinessTravel

Attrition	BusinessTravel			
	Non-Travel	Travel_Frequently	Travel_Rarely	Total
No	138 125.82 1.1798	208 232.34 2.55	887 874.84 0.1689	1233
Yes	12 24.184 6.1381	69 44.659 13.267	156 168.16 0.8789	237
Total	150	277	1043	1470

According to the data in the table above, it can be found that there is no case where the expected value is less than 5, so we can still perform chi-square test directly.

Assume, each sample used is independent, and each sample can only fall into just one of a finite number k of complementary and mutually exclusive outcomes of BusinessTravel.

Hypothesis test is constructed as below:

1. H0: Attrition and BusinessTravel are independent
2. H1: Attrition and BusinessTravel are not independent
3. α: 0.05

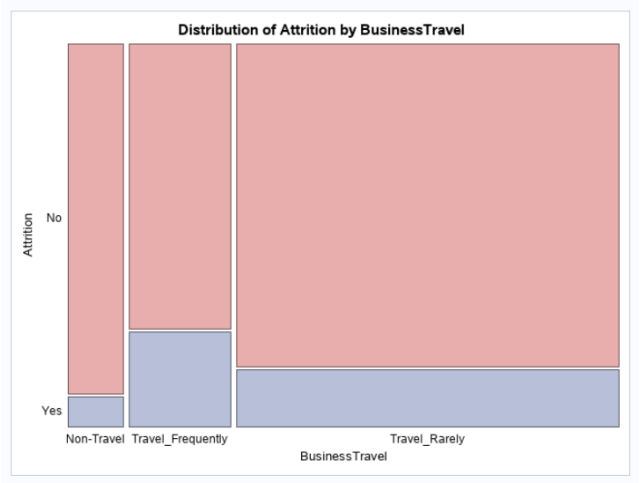
From the following table, the test statistic is χ 2 = 24.1824 with 2 degrees of freedom. The corresponding P-value is smaller than 0.001. Therefore, at 5% level of significance H0 can be rejected. There is statistically significant relationship between attrition and BusinessTravel. The biggest difference is from “Travel_Frequently” and “yes” group.

Statistics for Table of Attrition by BusinessTravel

Statistic	DF	Value	Prob
Chi-Square	2	24.1824	<.0001
Likelihood Ratio Chi-Square	2	23.7602	<.0001
Mantel-Haenszel Chi-Square	1	0.0000	0.9977
Phi Coefficient		0.1283	
Contingency Coefficient		0.1272	
Cramer's V		0.1283	

Sample Size = 1470

The result of MASIC diagram is shown below:



That is the MOSAIC diagram, from the diagram above, statistically speaking, we can find that not all BusinessTravel have the same Attrition for the same type of proportion. It is not difficult to see from the above figure that when BusinessTravel is Non-Travel, the proportion of the attrition value of No is much higher than the other two, and the lowest proportion of the Attrition value of No is when BusinessTravel is Travel_Frequently.

The FREQ Procedure

Frequency
Expected
Cell Chi-Square

Table of Attrition by JobRole

Attrition	JobRole									
	Healthcare Representative	Human Resources	Laboratory Technician	Manager	Manufacturing Director	Research Director	Research Scientist	Sales Executive	Sales Representative	Total
No	122 109.88 1.337	40 43.616 0.2998	197 217.24 1.8862	97 85.555 1.531	135 121.62 1.4714	78 67.102 1.7699	245 244.92 2.46E-7	269 273.44 0.0721	50 69.618 5.5284	1233
Yes	9 21.12 6.9556	12 8.3837 1.5599	62 41.757 9.8132	5 16.445 7.9651	10 23.378 7.6552	2 12.898 9.2081	47 47.078 0.0001	57 52.559 0.3752	33 13.382 28.762	237
Total	131	52	259	102	145	80	292	326	83	1470

According to the data in the table above, it can be found that there is no case where the expected value is less than 5, so we can still perform chi-square test directly.

Assume, each sample used is independent, and each sample can only fall into just one of a finite number k of complementary and mutually exclusive outcomes of JobRole.

Statistics for Table of Attrition by JobRole

Statistic	DF	Value	Prob
Chi-Square	8	86.1903	<.0001
Likelihood Ratio Chi-Square	8	88.9087	<.0001
Mantel-Haenszel Chi-Square	1	6.6242	0.0101
Phi Coefficient		0.2421	
Contingency Coefficient		0.2353	
Cramer's V		0.2421	

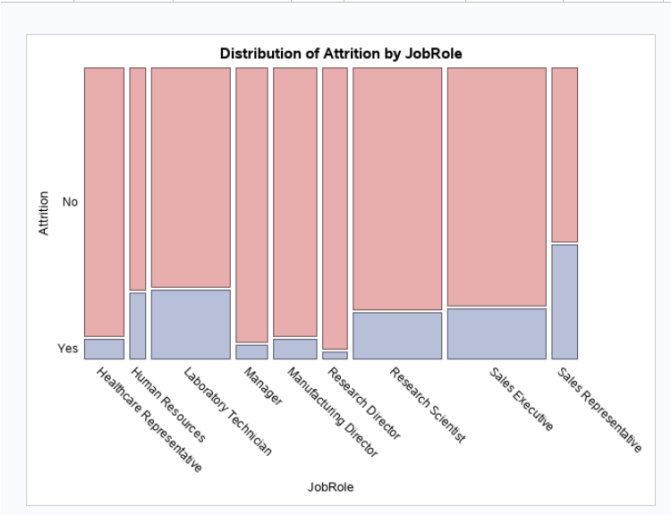
Sample Size = 1470

Assume, each sample used is independent, and each sample can only fall into just one of a finite number k of complementary and mutually exclusive outcomes of JobRole.

Hypothesis test is constructed as below:

4. H0: Attrition and JobRole are independent
5. H1: Attrition and JobRole are not independent
6. α: 0.05

From the above table, the test statistic is χ 2 = 86.1903 with 8 degrees of freedom. The corresponding P-value is smaller than 0.0001 which is also smaller than 0.05. Therefore, at 5% level of significance H0 can be rejected. There is statistically significant relationship between attrition and JobRole. The biggest difference is from “Sales Representative” and “yes” group.



That is the MOSAIC diagram, from the diagram above, statistically speaking, we can find that not all JobRole have the same Attrition for the same type of proportion. It is not difficult to see from the above figure that when JobRole is Healthcare Representative, Manager, Manufacturing Director, Research Director, the proportion of the attrition value of No is much higher than other cases. When JobRole is Sales Representative, the proportion of Attrition with NO value is much smaller than others.

Build a logistic model to predict attrition events. Investigate the factors that affect attrition. You can use backward selection to help you find the simplest model with the highest predictive power. Report and interpret in detail only your final model but do indicate how it was obtained and why it was considered the best. Consider as many numerical and categorical variables as possible. You may also wish to fit a model with interactions, but this is not strictly required. Interpret the odds ratios produced by your model.

First try to use all the variables to fit the logistic model, the results are shown below.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	81.3	Somers' D	0.625
Percent Discordant	18.7	Gamma	0.625
Percent Tied	0.0	Tau-a	0.169
Pairs	292221	c	0.813

It can be found that the value of the c-value of the model is 0.813.
Then remove the variable with the highest p-value among the variables:

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Age	1	7.0495	0.0079
EducationField	5	11.5712	0.0412
Department	2	0.0976	0.9524
BusinessTravel	2	19.6715	<.0001
Education	1	0.2008	0.6541
Gender	1	2.9847	0.0841
JobRole	8	29.5805	0.0003
MaritalStatus	2	35.7539	<.0001
MonthlyIncome	1	0.5740	0.4487
NumCompaniesWorked	1	20.7546	<.0001
OverTime	1	93.8501	<.0001
TotalWorkingYears	1	2.3324	0.1267
YearsAtCompany	1	0.0063	0.9369

Removing the department variable, and perform the fitting again:

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	81.2	Somers' D	0.623
Percent Discordant	18.8	Gamma	0.623
Percent Tied	0.0	Tau-a	0.169
Pairs	292221	c	0.812

Summary of Backward Elimination					
Step	Effect Removed	DF	Number In	Wald Chi-Square	Pr > ChiSq
1	Department	2	12	0.0976	0.9524
2	YearsAtCompany	1	11	0.0058	0.9394
3	Education	1	10	0.1942	0.6595
4	MonthlyIncome	1	9	0.5735	0.4489
5	TotalWorkingYears	1	8	2.6563	0.1031
6	Gender	1	7	3.1086	0.0779

It can be found that the c-value has not changed significantly. Based on the same logic, the value with the largest p-value can be continuously deleted until the variable with the p-value less than 0.05 is retained at the end.

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	1300.583	1064.612
SC	1305.876	1149.301
-2 Log L	1298.583	1032.612

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	265.9704	15	<.0001
Score	255.8522	15	<.0001
Wald	190.6648	15	<.0001

Type 3 Analysis of Effects			
Effect	DF	Wald Chi-Square	Pr > ChiSq
Age	1	17.2032	<.0001
BusinessTravel	2	19.0049	<.0001
JobRole	8	55.4842	<.0001
MaritalStatus	2	34.0640	<.0001
NumCompaniesWorked	1	21.4799	<.0001
OverTime	1	92.4475	<.0001

Logistic model was fit to predict the probability of Attrition, i.e. p = P (Attrition = Yes). Dummy variables for categorical predictors BusinessTravel, JobRole, MaritalStatus, OverTime were defined using reference coding. Odds ratios for Attrition will therefore be estimated Non-Travel, Sales Representative, Single and Yes for OverTime.

Model fit statistics indicate the model with intercept only to be inferior to the model that includes categorical predictors BusinessTravel, JobRole, MaritalStatus, OverTime. All three tests for the global hypothesis of zero beta indicate a highly statistically significant model, P-value < 0.0001. Type 3 analysis of effects shows that all six predictors are statistically significant (P-value < 0.0001).

From the parameters estimates section in Table:

Analysis of Maximum Likelihood Estimates						
Parameter		DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	1.4170	0.5247	7.2927	0.0069
Age		1	-0.0436	0.0105	17.2032	<.0001
BusinessTravel	Travel_Frequently	1	1.3300	0.3578	13.8168	0.0002
BusinessTravel	Travel_Rarely	1	0.6694	0.3333	4.0339	0.0446
JobRole	Healthcare Representative	1	-2.0221	0.4512	20.0816	<.0001
JobRole	Human Resources	1	-0.4940	0.4394	1.2643	0.2608
JobRole	Laboratory Technician	1	-0.5797	0.3009	3.7110	0.0541
JobRole	Manager	1	-1.9681	0.5537	12.6340	0.0004
JobRole	Manufacturing Director	1	-2.1538	0.4389	24.0804	<.0001
JobRole	Research Director	1	-3.0683	0.7905	15.0674	0.0001
JobRole	Research Scientist	1	-1.3890	0.3119	19.8317	<.0001
JobRole	Sales Executive	1	-1.0075	0.3052	10.8981	0.0010
MaritalStatus	Divorced	1	-1.1529	0.2355	23.9738	<.0001
MaritalStatus	Married	1	-0.8363	0.1756	22.6869	<.0001
NumCompaniesWorked		1	0.1490	0.0321	21.4799	<.0001
OverTime	No	1	-1.5941	0.1658	92.4475	<.0001

The estimated model for log odds of attrition is:

log (p/(1-p)) = 1.4170 – 0.0436xAge + 1.3300xBusinessTravel_TravelFrequently + 0.6694xBusinessTravel_TravelRarely - 2.0221xJobRole_HealthcareRepresentative - 0.4940xJobRole_HumanResources - 0.5797xJobRole_LaboratoryTechnician - 1.9681xJobRole_Manager - 2.1538xJobRole_ManufacturingDirector - 3.0683xJobRole_ResearchDirector - 1.3890xJobRole_ResearchScientist - 1.0075xJobRole_SalesExecutive - 1.1529xMaritalStatus_Divorced - 0.8363xMaritalStatus_Married + 0.1490x NumCompaniesWorked - 1.5941xOverTime_No

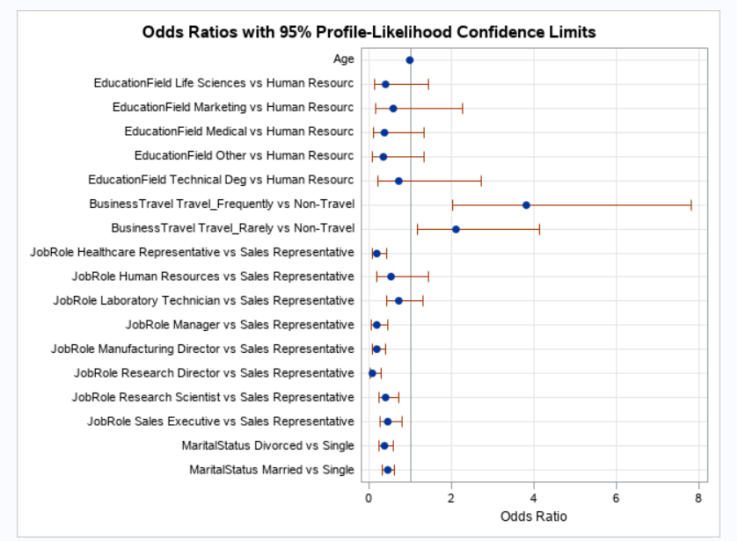
This equation confirms that young employees who do not travel, working as Sales Representative, single, working in more companies and work overtime have a higher chance to leave company.

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	80.3	Somers' D	0.605
Percent Discordant	19.7	Gamma	0.605
Percent Tied	0.0	Tau-a	0.164
Pairs	292221	c	0.803

Model performance statistics in Table are based on analysis of all possible pairs of employees in which one leave the company and the other did not. These statistics show 80.30% concordant pairs, 19.70% discordant pairs and nearly no ties. Based on the c statistic, the probability is 80.30% that an employee who leaves the company has higher predicted probability than does an employee who did not leave the company. Therefore, the model works quite well overall.

Odds Ratio Estimates and Profile-Likelihood Confidence Intervals				
Effect	Unit	Estimate	95% Confidence Limits	
Age	1.0000	0.975	0.957	0.993
EducationField Life Sciences vs Human Resourc	1.0000	0.400	0.114	1.441
EducationField Marketing vs Human Resourc	1.0000	0.578	0.152	2.255
EducationField Medical vs Human Resourc	1.0000	0.366	0.103	1.327
EducationField Other vs Human Resourc	1.0000	0.323	0.080	1.313
EducationField Technical Deg vs Human Resourc	1.0000	0.720	0.196	2.709
BusinessTravel Travel_Frequently vs Non-Travel	1.0000	3.812	2.004	7.806
BusinessTravel Travel_Rarely vs Non-Travel	1.0000	2.094	1.155	4.135
JobRole Healthcare Representative vs Sales Representative	1.0000	0.186	0.075	0.425
JobRole Human Resources vs Sales Representative	1.0000	0.521	0.173	1.421
JobRole Laboratory Technician vs Sales Representative	1.0000	0.720	0.405	1.289
JobRole Manager vs Sales Representative	1.0000	0.167	0.052	0.454
JobRole Manufacturing Director vs Sales Representative	1.0000	0.178	0.075	0.395
JobRole Research Director vs Sales Representative	1.0000	0.078	0.012	0.289
JobRole Research Scientist vs Sales Representative	1.0000	0.398	0.221	0.717
JobRole Sales Executive vs Sales Representative	1.0000	0.452	0.257	0.799
MaritalStatus Divorced vs Single	1.0000	0.368	0.235	0.563
MaritalStatus Married vs Single	1.0000	0.442	0.317	0.613

Odds ratios together with 95% confidence intervals are listed in Table and illustrated in Figure. there were significant differences in the chances of leaving the company in Travel_Frequently vs Non-Travel and Travel_Rarely vs Non-Travel, all JobRole except Human Resources vs Sales Representative and Laboratory vs Sales Representative, all MaritalStatus vs single and age.



The estimated of all significant result are as shown below:

1. Age: For every 1 year decrease in the age of employees, the increase in the odds ratio is 0.975 with 95% CI (0.957, 0.993).
2. Compared to Non-Travel, employees who traveled frequently and traveled rarely had 3.812 and 2.094 times higher odds of leaving the company.
3. Compared with Sales Representative, employees who works as Healthcare Representative had 0.186 times lower leaving the company, same, Manager had 0.167 times lower, Manufacturing Director had 0.178 times lower, Research Director had 0.078 times lower, Research Scientist had 0.398 times lower, Sales Executive had 0.452 times lower.
4. Compared with employee who is single, MaritalStatus Divorced had 0.368 times lower and MaritalStatus Married had 0.442 lower odds of leaving the company.

Question 4 (20 marks) Write a summary of your findings from Questions 1– 3. In particular, indicate the important factors that affect attrition, and what policies company can implement to address it. Keep the technical details of the analysis that led you to these conclusions to the absolute minimum. Rather, focus on practical significance and present your findings in non-specialist terms. One to two paragraphs (up to a page) will be sufficient.

An employee who is young and has short TotalWorkingYears more likely to leave the company. Employees who have an education background is Human Resource, Marketing and Technical Deg is more likely to leave the company. At the same time, if an employee travels rarely or frequently, that employee would have a higher chance to leave the company. For employee who works as sales representative will have a higher chance to leave the company. Single employee will have a higher chance to leave the company.

For young employees, it is necessary to provide some benefits to enhance the attractiveness of the company to them, including but not limited to organizing regular activities, allowing old employees to provide work plans and suggestions in the company, etc., to strengthen their stickiness to the company. At the same time, young employees can be promised that the longer they work in the company, the more benefits they can get.

For employees with Human resource, Marketing and Tech Deg backgrounds, it is necessary to pay attention to their career planning and link their salary with their length of service in the company.

In addition, it is necessary to reduce the number of times employee's travel. If travel is unavoidable, then try to let the same employees travel as much as possible. In this way, the number of people who leave because of travel can be reduced.

For those single employees, the policy can provide opportunities for marriage incentives. For example, if employees get married, they can get longer marriage holidays compared to other companies, or paid marriage holidays. At the same time, the company can also organize blind dates between single employees in different departments.