Question 1:

According to the table below, we can see the top 5 highest coefficient in logistics regression model are `NumCompaniesWorked`, `Education`, `OverTime`, `YearsSinceLastPromotion` and `YearsAtCompany`.

	Variable	Coeff
13	NumCompaniesWorked	0.162387
4	Education	0.082871
14	OverTime	0.081818
24	YearsSinceLastPromotion	0.080499
22	YearsAtCompany	0.076793

Question 2:

1. Decision Variables: defines which employee category should be offered an RCC severance package.

   Represented with binary variables:

$$x_i = 1 \; if \; employee \; category \; is \; offered \; RCC$$
$$x_i = 0 \; if \; employee \; category \; is \; not \; offered \; RCC$$

2. Objective: is to minimize the severance packages cost.

$$Cost = \min\left(\sum_{i=0}^{n}(p_i)\right)$$
$$where \; i \; is \; employees$$
$$where \; p_i \; is \; cost \; per \; employee \; who \; is \; offered \; RCC$$

3. Constraints:

3.1 The proportion of employees per job role should remain the same or within the 10% threshold.

$$(1-T)\{x_0 + (x_1)(p_0)\} \leq (x_1)(p_0) + x_0 \leq (1+T)\{x_0 + (x_1 \times p_0)\}$$

$where\ 1-T\ is\ \%\ of\ employees\ in\ a\ job\ role$

$where\ (x_1)(p_0)\ is\ number\ of\ employees\ in\ a\ job\ role$

$where\ x_0\ is\ employees\ who\ are\ not\ offered\ RCC$

$where\ x_1\ is\ employees\ who\ are\ offered\ RCC$

$where\ p_0\ is\ probability\ of\ staying\ in\ the\ company$

$where\ p_1\ is\ probability\ of\ staying\ in\ the\ company$

$where\ T\ is\ 10\%\ threshold$

3.2 At least forty employees should leave.

$$\sum_{i=0}^{441} (x_1)(p_1) \geq 40$$

3.3 Salary cuts should exceed the expected €3 miilion.

$$\sum_{i=0}^{441} (Salary)(RCC\ offered)(p_i) > €3\ million$$

Question 3:

The reason why I used `monthlyincome` and `overtime` to group employees is because you can see from the result of CAR in notebook file (under title: **CAR**) that those two factors decide if employees leave or doesn't leave. Then, I plotted a histogram graph to observe the distribution of income, and based on the graph, I grouped employees in different income ranges and if they work overtime or not (under title: **Group**). Although I obtained the result from running PCA, it didn't show any useful insights that I could use for group employees (under title: **PCA**).

Question 4:

Answer is in the code file under title: **Question 4**.

Question 5:

The model is shown in the excel file (file name: Starting_model_question4). The model showed group 1, group 2, group 5 and group 6 are given RCC.

Group1: employees whose monthly income is less than 4308 and work overtime.
Group2: employees whose monthly income is less than 4308 and doesn't work overtime.
Group5: employees whose monthly income is larger than 8616 and less than 12924 and work overtime.
Group6: employees whose monthly income is larger than 8616 and less than 12924 and doesn't work overtime.

Question 6:

The model is shown in the excel file (file name: Starting_model_question6). Before adding constraints, the model showed group 1, group 2, group 5 and group 6 are given RCC. After adding constraints, it shows company should give all groups RCC.

Question 7:

Because the result shows in the question 6 model which was already added constrains shows all group should be given RCC, I used the result shows in the question 5 model instead to do the verification process. Firstly, to see if there's gender inequality, I calculated the percentage of males and females who belong to four groups given RCC and the result shows males take up 61% and females 38% respectively. Although it seems there is discrimination, we still have to consider other factors why the percentage of males is higher than that of females. Based on the above approach categorizing employees, we can understand that employees whose monthly income is less than 4308 mostly are men and so is employees whose monthly income is between 8616 and 12924. Plus, the total number of males in the company is larger than that of females. Hence, it is reasonable that the percentage of males is higher than that of females in the groups which are given RCC.

Question 8:

One of the advantages of optimization approach is that it can help solve daily life problems and come up with solutions. For example, in business, it could help business minimize costs or maximise profits. In addition, it can also provide feasible solutions to company. For example, optimization approach accelerates Lyon's making decisions and the result of optimization approach also indicates how resources i.e., the severance packages can be allocated effectively by selecting the appropriate employee groups.

However, one of the disadvantages of optimization approach is that there are limited number of variables can be used in the optimization i.e., decision variables and constraints, which lower the flexibility and efficiency, especially in excel.

The assumption I made was employees who receive relatively low salary and work overtime might have higher chance to take RCC and leave the company. Although the results show in the final model (added constraints) indicates all employees should be given RCC, the results show in the model without constraints fulfill the assumption I made. For example, group 1 are the employees who receive low salary and work overtime which fulfill the assumption I made. Although employees in group 2 do not work overtime, due to the low salary they receive, they still have higher possibility to take RCC and leave the company. Moreover, even though employees in group 5 receive higher salary compare to employees in group 1 and group 2, they work overtime might be the factor affect them to take RCC and leave the company. As for employees in group 5, they do not work overtime but receive very high salary which might be the reason why company wants to offer them RCC to reduce labour costs.