Part 3. Answer the following questions with support of analytic evidence. Please provide

- o The R code (how do you realize it?)
- o The R output (Copy and paste <u>only</u> those relevant)
- o Your observations (What do you observe from the output?)
- o Your conclusion/decision/action

Q1. Again, please work on the cleaned FAA data set you prepared by carrying out Steps 1-9 in Part 1 of the project. Create a multinomial variable and attach it to your data set.

Y = 1 if distance < 1000

Y = 2 if 1000 < = distance < 2500

Y = 3 otherwise

Discard the continuous data for "distance", and assume we are given this multinomial response only without knowing its order.

In your meeting with an FAA agent who wants to know "what are risk factors in the landing process and how do they influence its occurrence", you are allowed to present:

- One model
- One table
- No more than five figures
- No more than five bullet statements. Please use statements that she can understand.

What model/table/figures/statements would you include in your presentation? Be selective!

Q2. The number of passengers is often of interest of airlines. What distribution would you use to model this variable? Do we have any variables (in the FAA data set) that are useful for predicting the number of passengers on board?