1. Create a Contingency table using the inbuilt “HairEyeColor dataset” in R.

Solution steps:

Distribution of hair and eye color and sex in 592 students. Students have been categorised based on their hair colour, eye colour and sex.

```
# Load the HairEyeColor dataset
data("HairEyeColor")

# Convert it to a data frame
hair_eye_df <- as.data.frame(HairEyeColor)
```

#This converts the 3-dimensional table into a data frame. **The resulting data frame will have columns for Hair, Eye, Sex, and Freq (frequency counts).**

```
# Create a contingency table for Hair and Eye color
hair_eye_table <- xtabs(Freq ~ Hair + Eye, data = hair_eye_df)          #You can check the xtabs command in detail by help(xtabs).
print(hair_eye_table)
##################################################################
```

Exercise 1.a

- Manually code that sum of the marginals of the rows = sum of the marginals of the columns =n (i.e. $n_{1+} + n_{2+} = n_{+1} + n_{+2} = n$)
- Also show that $p_{1+} + p_{2+} = p_{+1}+p_{+2} = 1$

You can add the marginals by using “addmargins(hair_eye_table)” .

2. Suppose there are 8000 people in a town. Out of the 8000 people 800 are females. In that town 1600 people are employed. Out of those 1600 employed people, only 120 are female. Create a contingency table of Employee status vs Sex.

Calculate expected frequency of each column by $e_i=n_i/n$ for i=1,2,3,4 . Calculate $\sum_{i=1}^{4}(\mathrm{fi}-\mathrm{ei})\text{^}2/\mathrm{ei}$