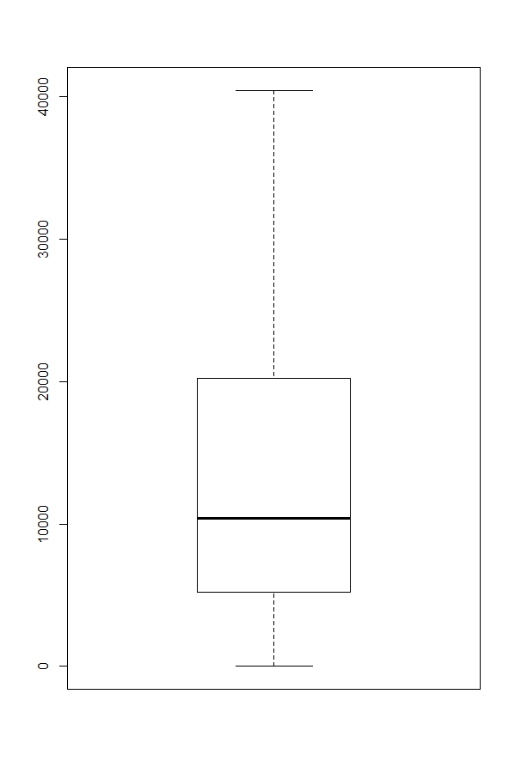
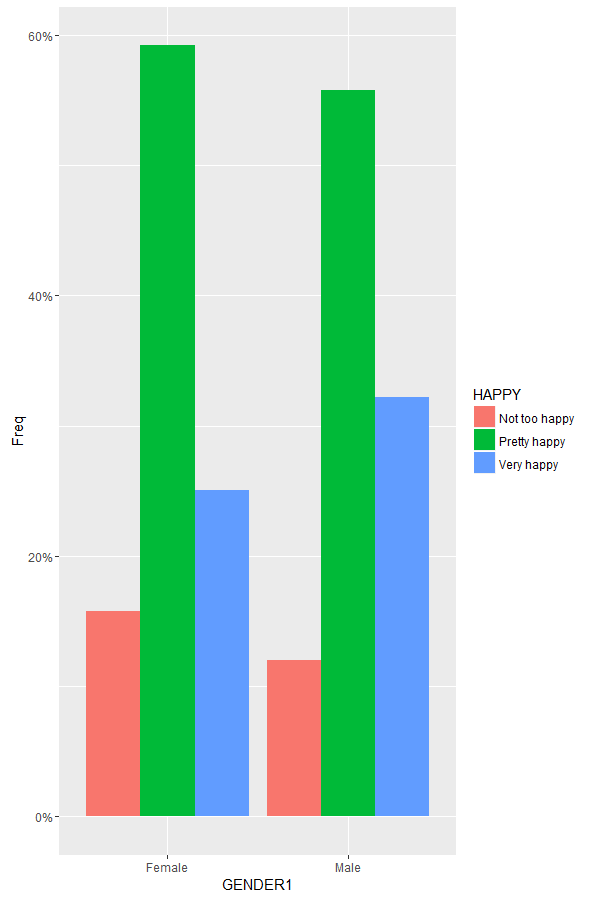
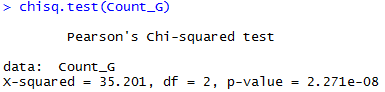
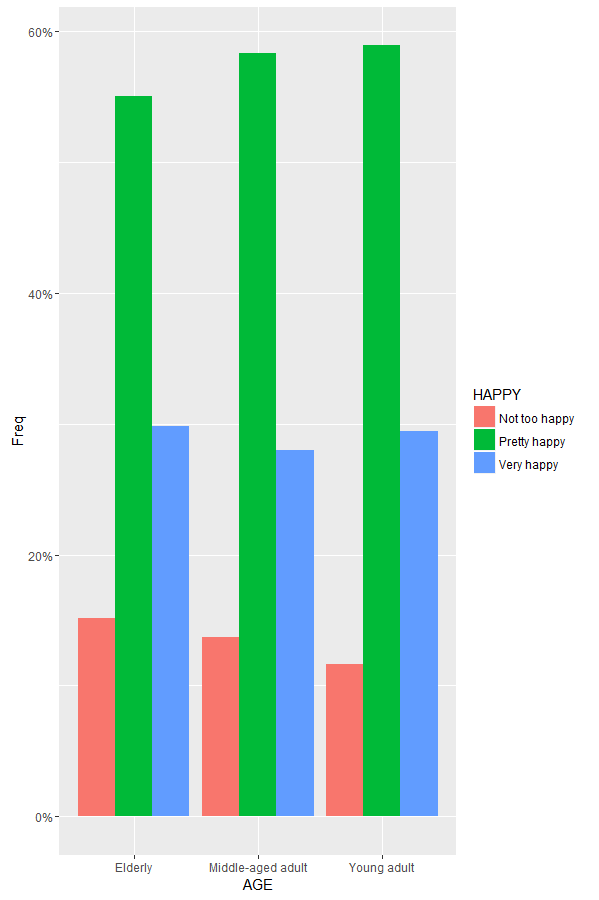
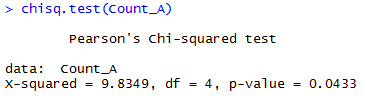
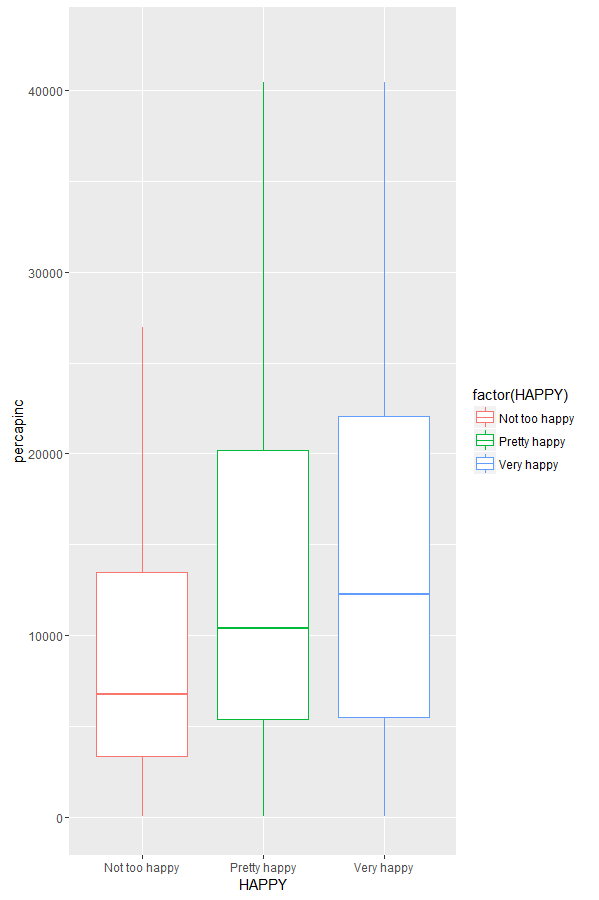
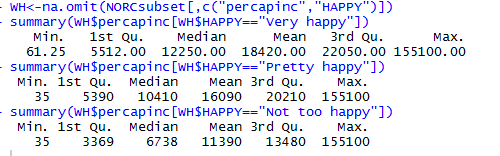
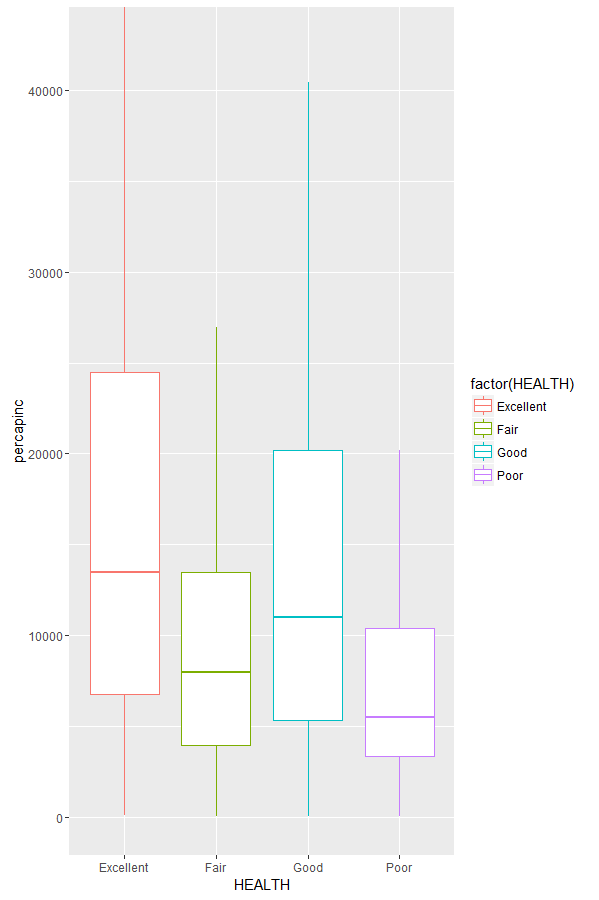
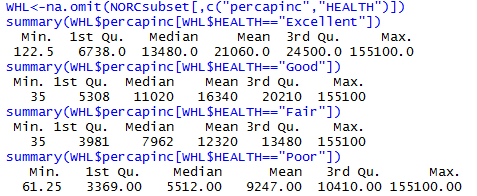
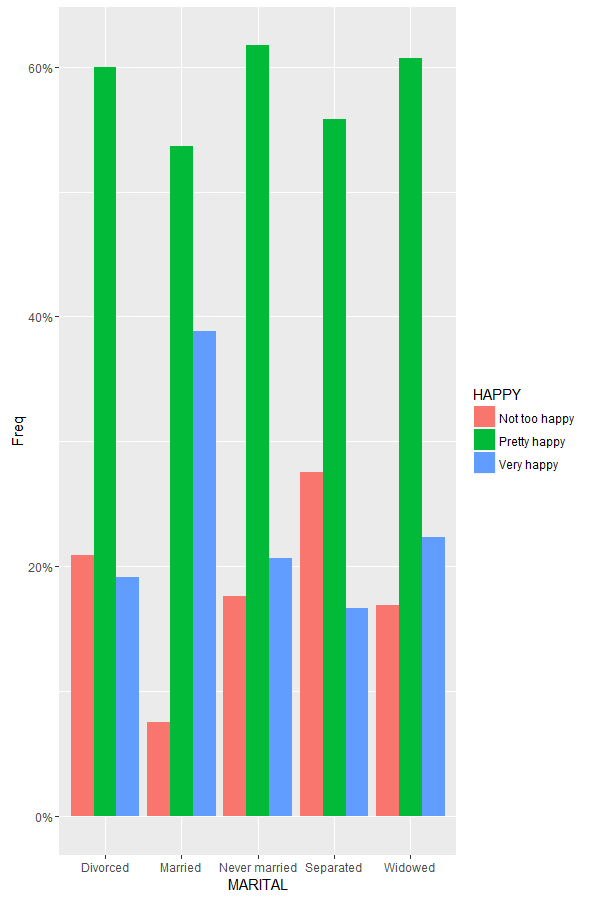
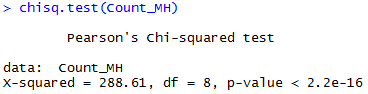
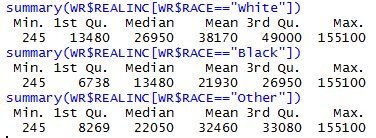
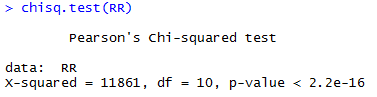
1. There are 2529 entries with missing data.  
   “IncompleteData” contains all the entries with missing data.  
   Please find “ # Q1 “ in the R script for the implementation
2. With the cook book, the missing data are given a specify value, so by finding and replacing them with “ <- NA “ in R can replace the missing value to NA.  
   Please find “ # Q2 & Q3 “ in the R script for the implementation
3. Please find “ # Q2 & Q3 “ in the R script for the implementation
4. Please find “ # Q4 “ in the R script for the implementation
5. There are 446 entries with missing data on either REALINC or HOMPOP which is ~9% of the total amount of entries.  
      
   The Minimum value is 35, the maximum value is 155100.  
   The 1st quartile with 5206, 3rd quartile with 20210.  
   With every quartile, the amount jump  
   by almost 2 times, 3rd quartile is almost  
   half of the 2nd quartile, and 2nd quartile  
   is almost half of 1st   
   The income curve is exponentially   
   increasing.
6.   
      
   By using the chi-square hypothesis testing, it shows that the p-value is less than 0.001. Which indicates that there a highly significant relationship between Gender and Happiness.   
   Though there are higher percentage of female in “Very Happy” category than male, we cannot determine that females are happier, as the overall combine of “Pretty happy” & “Very happy” male have higher percentage. Thus the females are not happier than female based on overall happiness.
7.   
     
   By using the chi-square hypothesis testing, it shows that the p-value is less than 0.05. Which indicates that there a significant relationship between Age and Happiness.  
   Based on the chart, the amount of people who are “Not too happy” increase as the age increase.  
   Thus, we may suggest that younger people are slightly happier, but we shall not imply that age directly cause the increase in happiness, as there may be other factor of each age group affect the happiness level.
8.   
     
   In every quartile, there is a positive correlation between happiness as the income increase.  
   Thus, we may suggest that, there are positive relation between the Income and the Happiness, and so people with higher income will tend to be happier, but this may not mean that the money is the causation of increase in happiness.
9.   
     
   In every quartile, there is a positive correlation between health as the income increase.  
   Thus, we may suggest that, there are positive relation between the Income and the Health, and so people with higher income will tend to be healthier, but the income may not be the causation of the increase in health.
10.   
      
    By using the chi-square hypothesis testing, it shows that the p-value is less than 0.001. Which indicates that there a highly significant correlation between Marital and Happiness.   
    People who are married have lowest percentage in “Not too happy” than the rest, Divorced and Separated have the highest percentage in “Not too happy”. We may suggest that the marital status have positive correlation to happiness, but we may not suggest that marital status is the cause of the happiness, as there might be some other cause that result in the current marital status.
11.   
      
      
    By using the chi-square hypothesis testing, it shows that the p-value is less than 0.001. Which indicates that there a highly significant correlation between Income and Race.  
    Therefore, there are differences in income between races.