

Hypotheses Testing & More

Student

6/26/2022

Getting the know the data:

Descriptive Statistics:

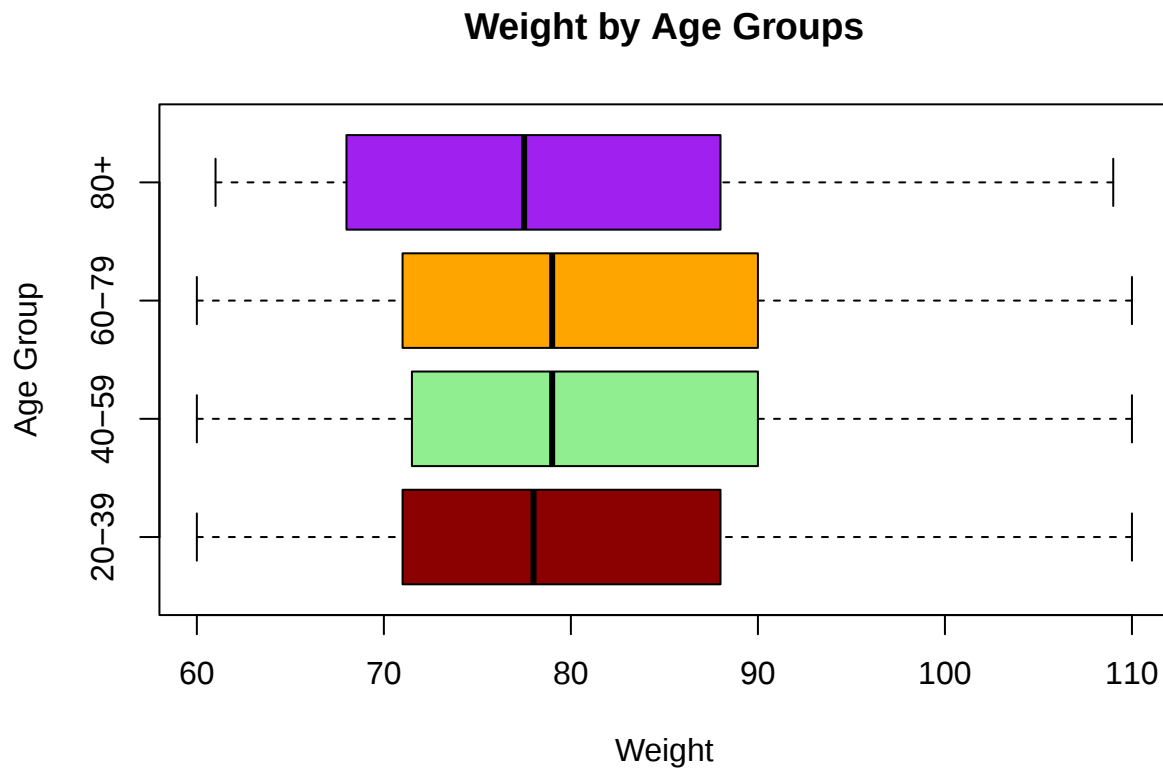
```
##      Success Height Weight Age Age_Group Gender Diabetes
## 1         1    168     69  29    20-39      F          0
## 2         0    162     99  54    40-59      M          1
## 3         1    175     81  46    40-59      M          0
## 4         1    169     75  68    60-79      F          0
## 5         1    181     99  22    20-39      M          1
## 6         1    173     74  20    20-39      M          0
```

```
##
## No. of observations = 3548
##
##   Var. name obs. mean  median  s.d.   min.   max.
## 1 Success    3548 0.49    0      0.5    0      1
## 2 Height     3548 169.53 167    10.87  155    195
## 3 Weight     3548 80.87   79     13.3   60     110
## 4 Age        3548 50.07   51     17.6   20     80
## 5 Age_Group
## 6 Gender
## 7 Diabetes   3548 0.5     0      0.5    0      1
```

```
## [1] Diabetes occurence by Age-Groups Per Gender
```

```
## , , = F
##
##
##      20-39 40-59 60-79 80+
## 0    322   272   301   15
## 1    283   285   311   14
##
## , , = M
##
##
##      20-39 40-59 60-79 80+
## 0    275   301   288   13
## 1    291   281   290    6
```

One Way ANOVA:



```
##           Df Sum Sq Mean Sq F value Pr(>F)
## Age_Group    3    394   131.2    0.741  0.527
## Residuals 3544 627279   177.0
```

Conclusion:

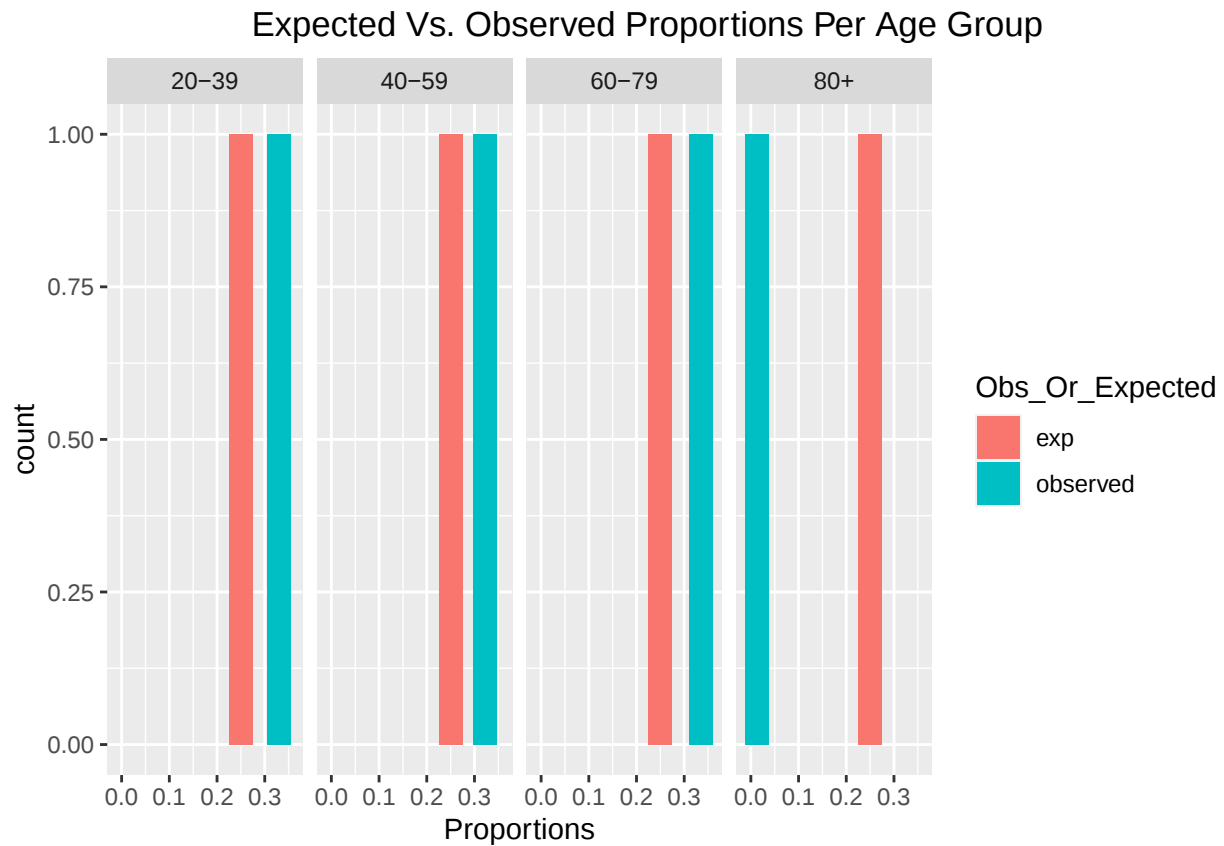
According to the result - there's no strong enough evidence to reject the null hypothesis: there's no significant difference between the different age groups in the weight variable.

Chi-Square Test for Goodness of Fit:

```
##  
## Chi-squared test for given probabilities  
##  
## data: Age_Group  
## X-squared = 1059.6, df = 3, p-value < 2.2e-16
```

Conclusion:

According to the results, there's a strong enough evidence to reject the null hypothesis and support the alternative hypothesis: Age_group distribution, is significantly non-homogeneous.

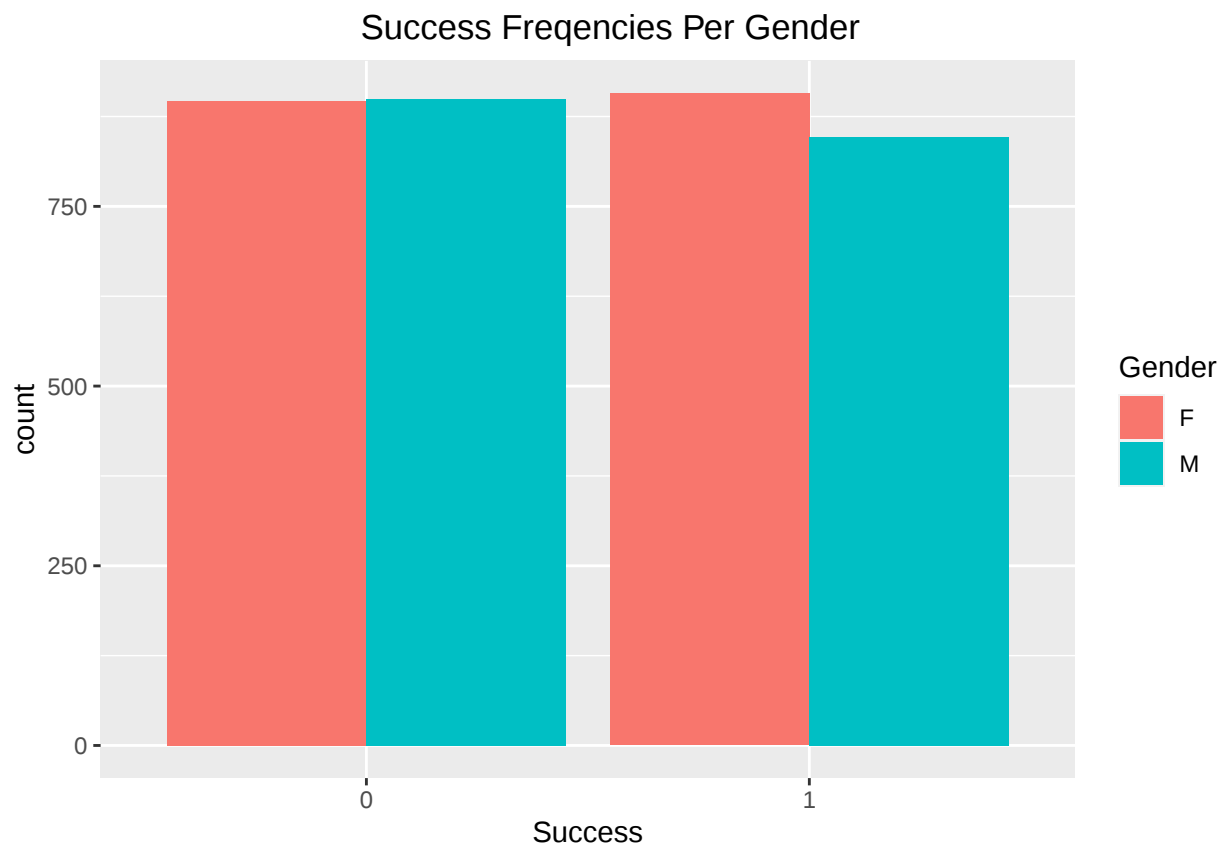


Chi-Square Test for independence:

```
##  
##           None Success  
##   Female  896      907  
##   Male    899      846  
  
##  
## Pearson's Chi-squared test with Yates' continuity correction  
##  
## data:  chisq  
## X-squared = 1.108, df = 1, p-value = 0.2925
```

Conclusion:

According to the results, there's no strong enough evidence to reject the null hypothesis and support the alternative hypothesis. There's no significant dependency between gender and success.



Independent samples T-test for means:

```
##  
## Welch Two Sample t-test  
##  
## data: data$Height by data$Gender  
## t = -54.961, df = 2375.2, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group F and group M is not equal to 0  
## 95 percent confidence interval:  
## -15.35994 -14.30164  
## sample estimates:  
## mean in group F mean in group M  
## 162.2368 177.0676
```

Conclusion:

According to the results, there's a strong enough evidence to reject the null hypothesis and support the alternative hypothesis. There's a significant difference in height between males and females.

