Harry_Potter_R_Analysis.R

OR

2020-10-18

```
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 3.5.3
```

```
HP1<- read_excel("C:/Users/OR/PycharmProjects/CoffeeAndHappiness/LinesByGenderHP1.xlsx")
```

```
## New names:
## * `` -> ...1
```

```
HP2<- read_excel("C:/Users/OR/PycharmProjects/CoffeeAndHappiness/LinesByGenderHP2.xlsx")
```

```
## New names:
## * `` -> ...1
```

```
HP3<- read_excel("C:/Users/OR/PycharmProjects/CoffeeAndHappiness/LinesByGenderHP3.xlsx")
```

```
## New names:
## * `` -> ...1
```

```
HP4<- read_excel("C:/Users/OR/PycharmProjects/CoffeeAndHappiness/LinesByGenderHP4.xlsx")
```

```
## New names:
## * `` -> ...1
```

```
HP6<- read_excel("C:/Users/OR/PycharmProjects/CoffeeAndHappiness/LinesByGenderHP6.xlsx")
```

```
## New names:
## * `` -> ...1
```

```
############################## Data Extraction ##############################

##############
#### HP1 #####
##############

HP1Female<- HP1[c(1:7),3]
HP1Male<- HP1[c(8:33),3]


##############
#### HP2 #####
##############

HP2Female<- HP2[c(1:8),3]
HP2Male<- HP2[c(9:35),3]


##############
#### HP3 #####
##############

HP3Female<- HP3[c(1:11),3]
HP3Male<- HP3[c(12:35),3]


##############
#### HP4 #####
##############

HP4Female<- HP4[c(1:9),3]
HP4Male<- HP4[c(10:31),3]


##############
#### HP6 #####
##############

HP6Female<- HP6[c(1:9),3]
HP6Male<- HP6[c(10:28),3]


############################## Data Analysis ##############################

print("************************ Data Analysis ************************")
```

```
## [1] "************************ Data Analysis ************************"
```

```
HPFemale<-c(HP1Female, HP2Female, HP3Female, HP4Female, HP6Female)
HPMale<-c(HP1Male, HP2Male, HP3Male, HP4Male, HP6Male)

for(i in 1:(length(HPFemale)-1))
{
  cat("######### Harry Potter", i, "#########")
  cat(sep="\n\n")
  cat(sep="\n\n")
  cat("Variances are Equal in Harry Potter", i)
  cat(sep="\n\n")
  print(var.test(unlist(HPFemale[i]), unlist(HPMale[i]), alternative = "two.sided", sep="\n\n"))
  cat(sep="\n\n")
  cat("Female and Male characters have equal amount of lines in Harry Potter", i)
  cat(sep="\n\n")
  print(t.test(unlist(HPMale[i]), unlist(HPFemale[i]), mu=0 ,paired = FALSE, conf.level = 0.95))
  cat(sep="\n\n")
  cat(sep="\n\n")
}
```

```
## ######### Harry Potter 1 #########
##
## Variances are Equal in Harry Potter 1
```

```
##
##  F test to compare two variances
##
## data:  unlist(HPFemale[i]) and unlist(HPMale[i])
## F = 0.594, num df = 6, denom df = 25, p-value = 0.5357
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2000969 3.0334627
## sample estimates:
## ratio of variances
##          0.5939975
##
##
## Female and Male characters have equal amount of lines in Harry Potter 1
##
##  Welch Two Sample t-test
##
## data:  unlist(HPMale[i]) and unlist(HPFemale[i])
## t = 0.018372, df = 12.076, p-value = 0.9856
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -56.17284  57.12888
## sample estimates:
## mean of x mean of y
##  44.19231  43.71429
##
##
##
## ######### Harry Potter 2 #########
##
## Variances are Equal in Harry Potter 2
##
##  F test to compare two variances
##
## data:  unlist(HPFemale[i]) and unlist(HPMale[i])
## F = 0.50001, num df = 7, denom df = 26, p-value = 0.3483
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1770589 2.1974892
## sample estimates:
## ratio of variances
##          0.5000122
##
##
## Female and Male characters have equal amount of lines in Harry Potter 2
##
##  Welch Two Sample t-test
##
## data:  unlist(HPMale[i]) and unlist(HPFemale[i])
## t = 0.02093, df = 16.221, p-value = 0.9836
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -36.17434  36.89656
## sample estimates:
## mean of x mean of y
##  34.11111  33.75000
##
##
##
## ######### Harry Potter 3 #########
##
## Variances are Equal in Harry Potter 3
##
##  F test to compare two variances
##
## data:  unlist(HPFemale[i]) and unlist(HPMale[i])
## F = 0.56541, num df = 10, denom df = 23, p-value = 0.3505
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2119023 1.9093939
## sample estimates:
## ratio of variances
##           0.565407
```

```
##
##
## Female and Male characters have equal amount of lines in Harry Potter 3
##
##  Welch Two Sample t-test
##
## data:  unlist(HPMale[i]) and unlist(HPFemale[i])
## t = 0.99183, df = 25.499, p-value = 0.3306
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -24.06924  68.87227
## sample estimates:
## mean of x mean of y
##  52.58333  30.18182
##
##
##
## ######### Harry Potter 4 #########
##
## Variances are Equal in Harry Potter 4
##
##  F test to compare two variances
##
## data:  unlist(HPFemale[i]) and unlist(HPMale[i])
## F = 0.28085, num df = 8, denom df = 21, p-value = 0.06983
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.09771946 1.11906469
## sample estimates:
## ratio of variances
##          0.2808457
##
##
## Female and Male characters have equal amount of lines in Harry Potter 4
##
##  Welch Two Sample t-test
##
## data:  unlist(HPMale[i]) and unlist(HPFemale[i])
## t = 0.93145, df = 26.699, p-value = 0.36
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -11.4257  30.4055
## sample estimates:
## mean of x mean of y
##  23.04545  13.55556
```

```
cat("######### Harry Potter 6 #########", sep="\n\n")
```

```
## ######### Harry Potter 6 #########
```

```
cat("Variances Differ in Harry Potter 6")
```

```
## Variances Differ in Harry Potter 6
```

```
var.test(unlist(HP6Female), unlist(HP6Male), alternative = "two.sided")
```

```
##
##  F test to compare two variances
##
## data:  unlist(HP6Female) and unlist(HP6Male)
## F = 0.22068, num df = 8, denom df = 18, p-value = 0.03507
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.07343034 0.89016287
## sample estimates:
## ratio of variances
##          0.2206781
```