

Aim: Conducting Chi-square tests using chisq.test() (R)

```

1 print("Name: S126 hariprasad")
2
3 # 1. use file.choose() to manually select your 'Fastest Fifties - 2022' csv file
4 data <- read.csv(file.choose())
5
6 # 2. create a contingency table for the two categorical variables
7 # we compare the opponent team (Against) with the stadium (Venue)
8 table_data <- table(data$Against, data$Venue)
9
10 # 3. perform the chi-square test
11 chisq_result <- chisq.test(table_data)
12
13 # 4. View the results (chi-square value, degrees of freedom, and p-value)
14 chisq_result
15
16 # 5. Preview the first few rows of the data
17 head(data)
18

```

Warning message:
In chisq.test(table_data) : chi-squared approximation may be incorrect

```

> # 4. View the results (chi-square value, degrees of freedom, and p-value)
> chisq_result

Pearson's Chi-squared test

data: table_data
X-squared = 38.695, df = 45, p-value = 0.7347

> # 5. Preview the first few rows of the data
> head(data)
  POS   Player Runs BF   SR  Xs  Against Venue Match.Date
1  1 Pat Cummins  56 14 373.33  4  6  MI Maharashtra Cricket Association Stadium 06 April 2022
2  2 Moeen Ali  93 18 516.11  13  3  BR Brabourne - CCI 20 May 2022
3  3 Liam Livingstone  64 21 237.03  7  4  GT Brabourne - CCI 08 April 2022
4  4 Rahul Tripathi  72 21 390.00  4  6  KKR Brabourne - CCI 15 April 2022
5  5 Jonny Bairstow  66 21 227.58  4  7  RCB Brabourne - CCI 13 May 2022
6  6 Evlin Lewis  55 23 239.13  6  3  CSK Brabourne - CCI 11 March 2022

```

Environment: Global Environment
Data: chisq_result List of 9
\$ statistic: Named num 38.7
\$ attr(*, "names")= chr "X-squared"
\$ parameter: Named int 45
\$ attr(*, "names")= chr "df"
\$ p-value: num 0.735
\$ method: chr "Pearson's chi-squared test"
\$ data.name: chr "table_data"
\$ observed: 'table' int [1:10, 1:6] 4 3 3 5 1 3 3 2 4 3 ...
\$ attr(*, "dimnames")=list of 2
.. \$: chr [1:10] "CSK" "DC" "GT" "KKR" ...
.. \$: chr [1:6] "Brabourne - CCI" "DV Patil Stadium" "Eden Gardens" "Maharashtra Cricket Associa..."
\$ expected: num [1:10, 1:6] 2.453 3.68 3.68 3.62 1.84 ...
\$ attr(*, "dimnames")=list of 2
.. \$: chr [1:10] "CSK" "DC" "GT" "KKR" ...
.. \$: chr [1:6] "Brabourne - CCI" "DV Patil Stadium" "Eden Gardens" "Maharashtra Cricket Associa..."
\$ residuals: 'table' num [1:10, 1:6] 0.847 -0.354 -0.354 0.858 -0.619 ...
\$ attr(*, "dimnames")=list of 2
.. \$: chr [1:10] "CSK" "DC" "GT" "KKR" ...
.. \$: chr [1:6] "Brabourne - CCI" "DV Patil Stadium" "Eden Gardens" "Maharashtra Cricket Associa..."
\$ stdres: 'table' num [1:10, 1:6] 1.031 -0.439 -0.439 1.059 -0.743 ...
\$ attr(*, "dimnames")=list of 2
.. \$: chr [1:10] "CSK" "DC" "GT" "KKR" ...
.. \$: chr [1:6] "Brabourne - CCI" "DV Patil Stadium" "Eden Gardens" "Maharashtra Cricket Associa..."
\$ attr(*, "class")= chr "htest"

data: 118 obs., of 10 variables
\$ POS : int 1 2 3 4 5 6 7 8 9 10 ...
\$ Player : chr "Pat Cummins" "Moeen Ali" "Liam Livingstone" "Rahul Tripathi" ...
\$ Runs : int 56 93 64 71 66 55 106 54 58 50 ...
\$ BF : int 14 19 21 21 21 23 23 23 25 25 ...
\$ SR : num 373 163 237 192 228 ...
\$ Xs : int 4 13 7 4 4 6 10 8 9 8 ...
\$ Against : int 6 3 4 6 7 6 9 2 1 ...
\$ Venue : chr "MI" "RR" "GT" "KKR" ...
\$ Match.Date: chr "06 April 2022" "20 May 2022" "08 April 2022" "15 April 2022" ...

table_data 'table' int [1:10, 1:6] 4 3 3 5 1 3 3 2 4 3 ...