

Chi Square Test for Multiple columns comparisons

```
# Chi2 sqaure test for the relationship to be established
import math

# Cramer's Value to measure the chi square to find the association.

def cramers_V(statistical_value, contingency_table, total_observation,
row_count=0, column_count=0):

    # Finding the columns length and rows length
    for rows in contingency_table:
        row_count += 1

    for columns in contingency_table.T:
        column_count += 1

    # Degree Of Freedom
    k = min(row_count, column_count)
    denominator = total_observation * k # Total Observations as well as the
mininum Degree Of Freedom
    numerator = statistical_value # Chi2 Statistics

    # Cramer's V Formula
    return math.sqrt(numerator / denominator)

comparee = ['Biking Habit', 'Jogging Habit', 'Walking Habit'] # Add the
columns you want to use against
comparer = ['Smoking Habit', 'Drinking Habit'] # Add the columns you want
to use in the first place

def chi_square(healthy_columns, unhealthy_columns):

    try:
        for first_columns in range(len(unhealthy_columns)): # For Unhealthy
Columns
            print()

            print(f'
-----{unhealthy_columns[fir
st_columns]}-----
-----')

            print()
            for second_columns in range(len(healthy_columns)): # For
Healthy Columns

                print(f'{unhealthy_columns[first_columns]} to
{healthy_columns[second_columns]}-----')
                table = pandas.crosstab(data['Smoking Habit'],
data[healthy_columns[second_columns]])

                chi2_statistics, p_value, degree_of_freedom, expected =
chi2_contingency(table)

                print(
                    f"Cramer's V Score for {
                        unhealthy_columns[first_columns]} and
{healthy_columns[second_columns]} -",
                    cramers_V(statistical_value=chi2_statistics,
contingency_table=table,
total_observation=sum(
                        sum(
                            table.values
                        )
                    )
                ) # Initiating the Cramer's V to check the accuracy for
the association of Chi2 Test

                if p_value < 0.05:
                    print('There is a strong evidence for the association
to be identified for this relationship')
                else:
                    print('There is not a strong evidence for the
association to be identified for this relationship')
                    print()
            except (IndexError, ValueError, SyntaxError):
                if IndexError:
                    print('I want you to perform correct indexing')
                elif ValueError:
                    print('I cannot work witht the values you provided, change it!'
)
                elif SyntaxError:
                    print('Could you please write clear for me to understand it and
then get you the output')
                else:
                    print('Its an error! Please change your code once')

# Final Chi2 Test For Multiple Columns Comparisions
chi_square(comparee, comparer)
```