##### import data code #####

import pandas as pd

df2 = pd.read\_csv("Dataset Generation (Fardina) (Responses) - Form Responses 1.csv")

##### data clean up code #####

#check for duplicates

print(len(df2))

print(len(df2.drop\_duplicates()))

# gets rid of people who we can't identify what their background is, because without any

# background, any analysis is not as important compared to those with

df2.dropna(subset=['What year are you?', 'How old are you?', 'You could describe the adults you

grew up with as...', 'You could describe yourself as...', 'How would you rate your

religiousness / spirituality?', 'What bests represents your gender?'], how='all', inplace = True)

print(len(df2))

def find\_and\_fill(row, parent\_df):

if pd.isna(row['How old are you?']):

year = row['What year are you?']

med\_val = parent\_df.loc[parent\_df['What year are you?'] == year]['How old are you?'].median()

row['How old are you?'] = med\_val

print(row['How old are you?'])

return row

# finds any missing years, fills in with median of what year they are in

df2 = df2.apply(find\_and\_fill, args=(df2,), axis = 1)

print(len(df2))

def fix\_female(value):

if value == 'Famale':

value = 'Female'

return value

# fix famale to female

df2['What bests represents your gender?'] = df2['What bests represents your

gender?'].apply(fix\_female)

import statistics as stat

# fills in blank survey responses with mode value of that col

for column in df2.iloc[:, 8:]:

col = df2[column]

mode\_val = stat.mode(col)

df2[column] = df2[column].fillna(value = mode\_val)

##### chi square test #####

from scipy.stats import chi2\_contingency

alpha = .05

primer = df2[ col name ]

results = {}

for col in df2.iloc[:, 8:]:

contingency\_table = pd.crosstab(primer, df2[col])

chi2, p, dof, ex = chi2\_contingency(contingency\_table)

results[col] = p

for col in results:

value = results[col]

if value <= alpha:

print(value)

print(f'Statistically Significant with alpha value = {alpha}')

print(f'{col}\n')

else:

print(value)

print(f'Not Statistically Significant with alpha value = {alpha}\n')