King Abdulaziz University (KAU)
Faculty of Engineering – Girls' Campus
Electrical Engineering Department
EE-495 – Special Topics In Computer Engineering

EE-495

Special Topics In Computer Engineering

Case Study No. 1

Hypothesis Testing For Coronavirus Using Python

Haneen Alamoudi - 1708436

Submission Date: 11 March 2021 (WEEK 8)

Contents

Introduction

In this case study, the implementation of a two-sample Z test for coronavirus dataset was done, and the formula is shown below. John Hopkin`s repository coronavirus dataset has been studied, the dataset includes information about state, country, last update, confirmed cases, deaths cases, recovered cases latitude and longitude. The case study concentrates on the feature of temperature and humidity for latitude and longitude. In the dataset, a temperature below 24 considered to be a cold climate and above 24 considered as hot climate

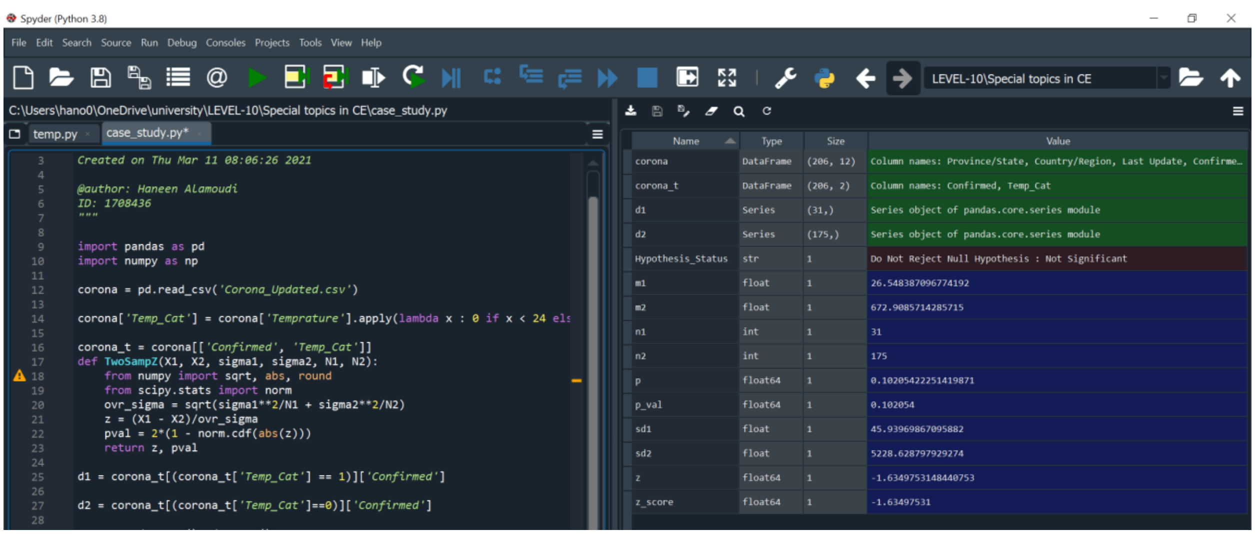
$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$

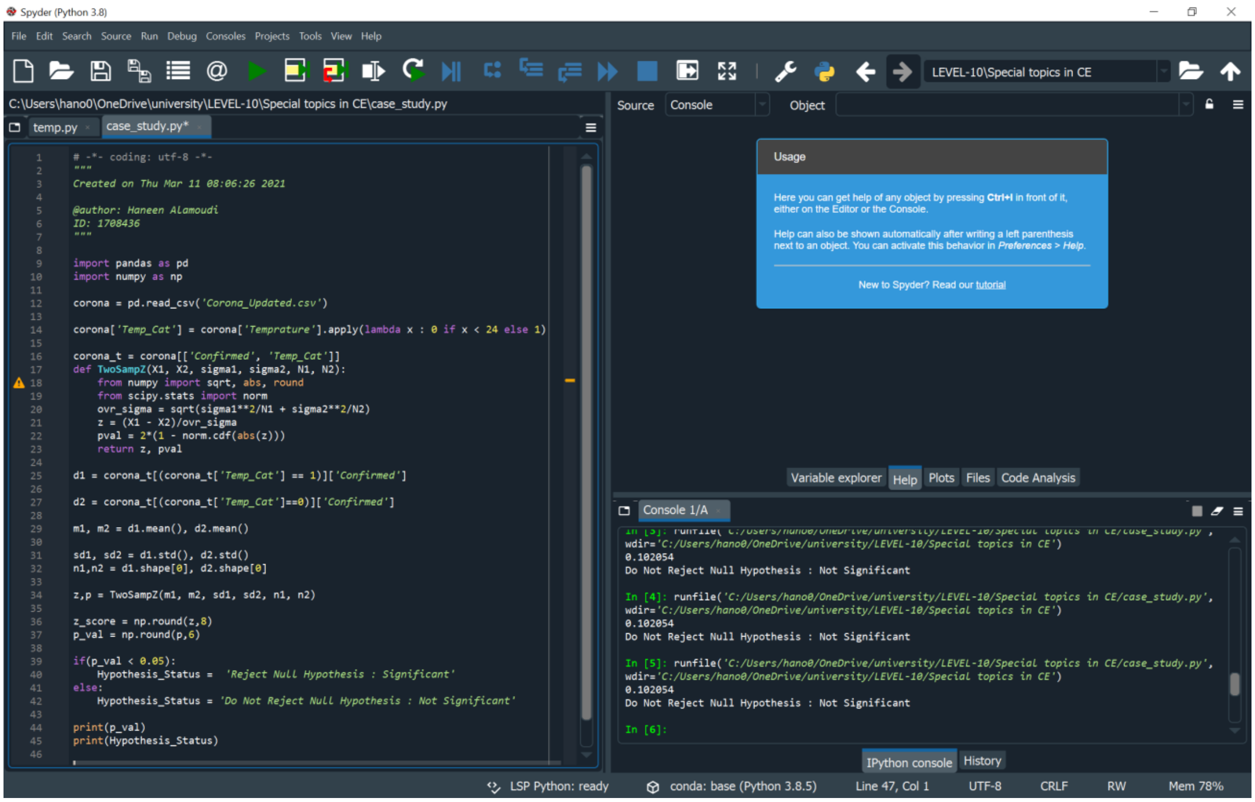
Main Hypothesis for testing and current perception: Warm climate is more resistant to the corona outbreak.

Null Hypothesis: Temperature doesn`t affect COV-19 Outbreak

Alternate Hypothesis: Temperature does affect COV-19 Outbreak

Output





Code

```
# -*- coding: utf-8 -*-
"""
Created on Thu Mar 11 08:06:26 2021

@author: Haneen Alamoudi
ID: 1708436
"""

import pandas as pd
import numpy as np

corona = pd.read_csv('Corona_Updated.csv')

corona['Temp_Cat'] = corona['Temprature'].apply(lambda x : 0 if x < 24 else 1)

corona_t = corona[['Confirmed', 'Temp_Cat']]
def TwoSampZ(X1, X2, sigma1, sigma2, N1, N2):
    from numpy import sqrt, abs, round
    from scipy.stats import norm
    ovr_sigma = sqrt(sigma1**2/N1 + sigma2**2/N2)
    z = (X1 - X2)/ovr_sigma
    pval = 2*(1 - norm.cdf(abs(z)))
    return z, pval

d1 = corona_t[(corona_t['Temp_Cat'] == 1)]['Confirmed']
```

```
d2 = corona_t[(corona_t['Temp_Cat']==0)]['Confirmed']

m1, m2 = d1.mean(), d2.mean()

sd1, sd2 = d1.std(), d2.std()
n1,n2 = d1.shape[0], d2.shape[0]

z,p = TwoSampZ(m1, m2, sd1, sd2, n1, n2)

z_score = np.round(z,8)
p_val = np.round(p,6)

if(p_val < 0.05):
    Hypothesis_Status =  'Reject Null Hypothesis : Significant'
else:
    Hypothesis_Status = 'Do Not Reject Null Hypothesis : Not Significant'

print(p_val)
print(Hypothesis_Status)
```

Conclusion

Based on the calculation done using the Python code, z = -1.634975 and p = 0.102054. since P-value is greater than 0.05, we cannot reject the null hypothesis because it is not significant. And conclude that temperature does not effect COVID-19.