**D207 Performance Assessment**

**Nefisa Hassen**

**A1. Question of the analysis**

Is gender related to hospital readmission ?

H0 = gender is not related to hospital readmission

H1 = gender is related to hospital readmission rate.

**A2. Benefit of the analysis**

An increased hospital readmission leads to penalties for the hospitals. Understanding which patient is at high risk of hospital readmission will benefit hospitals to prepare in advance. In addition, this analysis can be used for hospital based health programs such as ED transition of care programs that are designed to reduce hospital readmission rates for patients (all genders) which will reduce the cost of health care overall.

**A3. Data Identification**

In order to answer the analysis question we will be looking at columns ReAdmist and Gender.

**B1: Code**

The chi-square test will be used.

**Libraries and packages**

import pandas as pd

import numpy as np

import seaborn as sns

import matplotlib.pyplot as plt

%matplotlib inline

import pylab

import statsmodels.api as sm

import statistics

from scipy import stats

from scipy.stats import chisquare

from scipy.stats import chi2\_contingency

importing the data

df = pd.read\_csv("medical\_clean.csv")

**Renaming the 8 survey response columns for univariate and bivariate plots.**

df.rename(columns = {'Item1':'TimelyAdmission',

'Item2':'TimelyTreatment',

'Item3':'TimelyVisits',

'Item4':'Reliability',

'Item5':'Options',

'Item6':'HoursOfTreatment',

'Item7':'CourteousStaff',

'Item8':'EvidenceOfActiveListening'},

inplace=True)

boxplt= df[['TimelyAdmission','TimelyTreatment']].plot(kind = 'box',title = 'TimelyAdmission vs TimelyTreatment')

plt.show()

**Creating data frame**

contingency = pd.crosstab(df['ReAdmis'], df['Gender'])

contingency = pd.crosstab(df['ReAdmis'], df['Gender'])

**B2. Output**

stat, p, dof, expected = chi2\_contingency(contingency )

c, p, dof, expected = chi2\_contingency(contingency)

# interpret p-value

alpha = 0.05

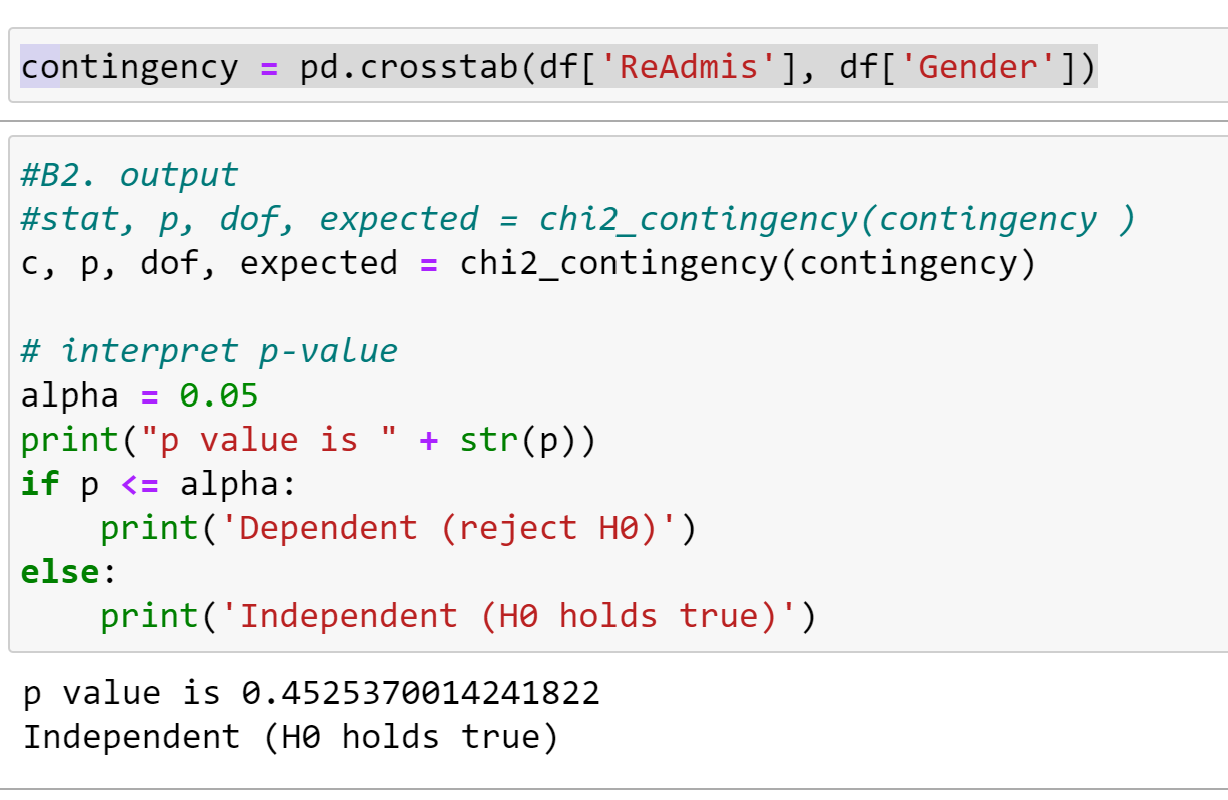
print("p value is " + str(p))

if p <= alpha:

print('Dependent (reject H0)')

else:

print('Independent (H0 holds true)')

****

**B3. Justification**

In this analysis we are trying to determine which patient is more likely to be readmitted to hospital. The question looks at the patient's gender and the ReAdmis columns which are both categorical variables. As a result, the chi square test is the best method.

**C1. Univariate statistics**

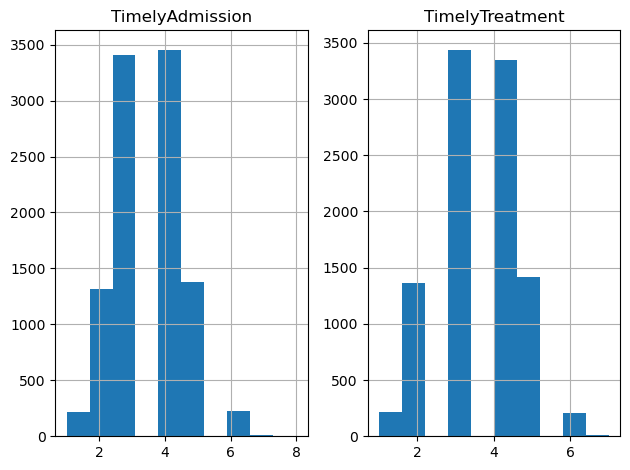
Two categorical variable Readmis, gender and Two continuous variable Timely admission, and Timely treatment

Create histograms of continuous & categorical variables

df[['TimelyAdmission','TimelyTreatment']].hist()

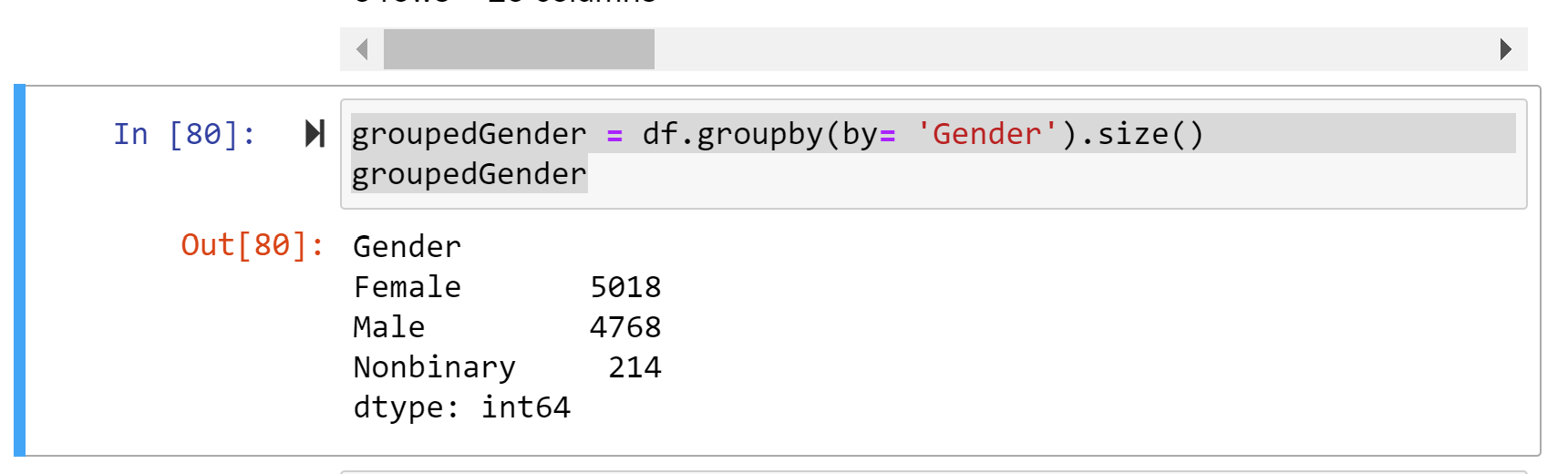
plt.savefig('medical\_clean.jpg')

plt.tight\_layout()



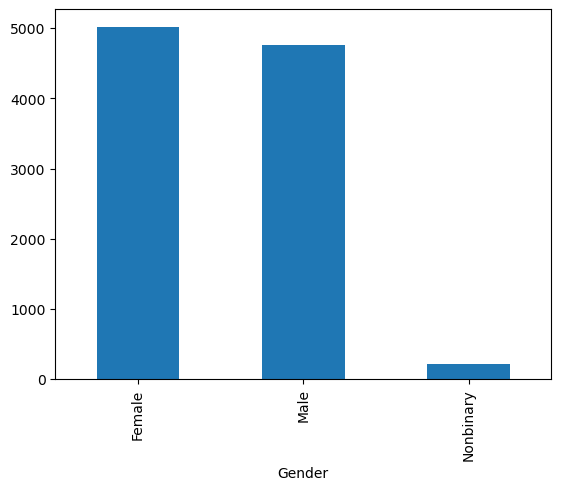
groupedGender = df.groupby(by= 'Gender').size()

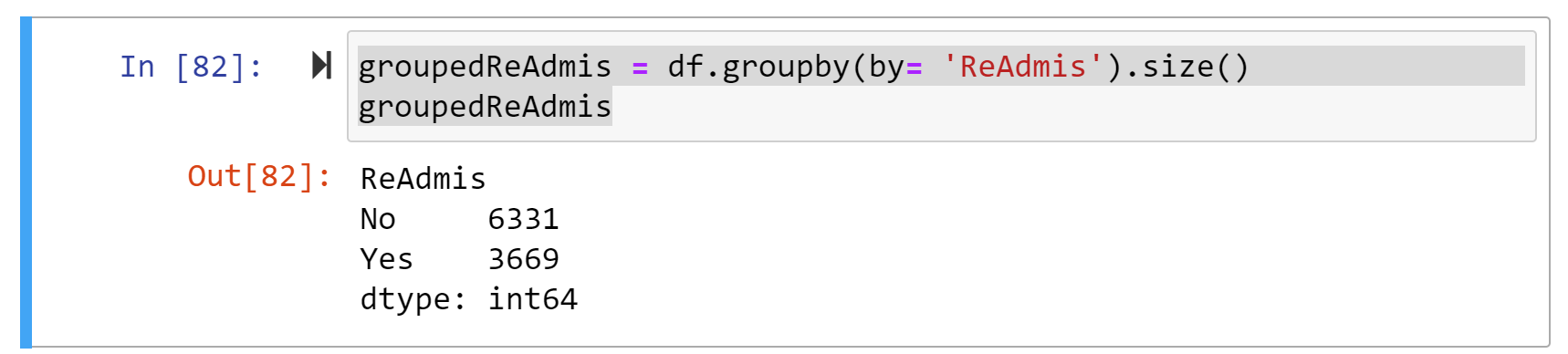
groupedGender



%matplotlib inline

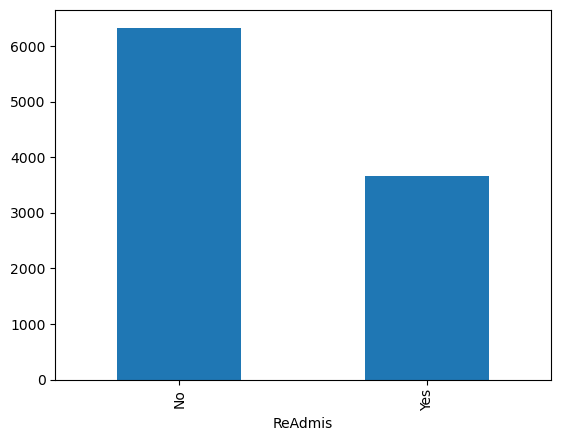
groupedGender.plot.bar()





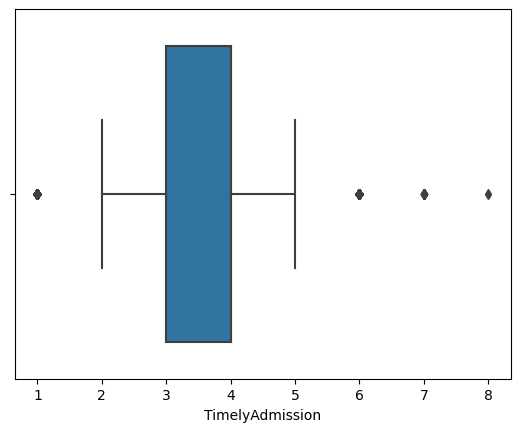
%matplotlib inline

groupedReAdmis.plot.bar()



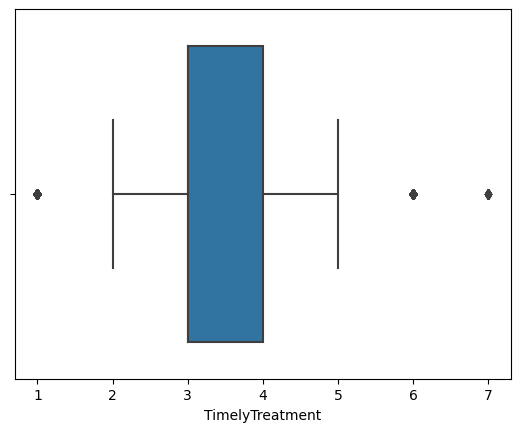
sns.boxplot('TimelyAdmission', data = df )

plt.show()

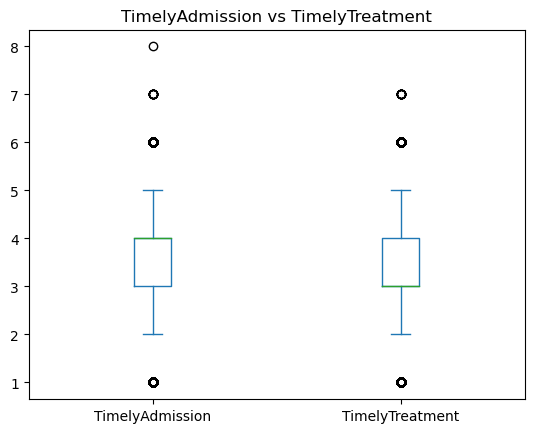


sns.boxplot('TimelyTreatment', data = df)

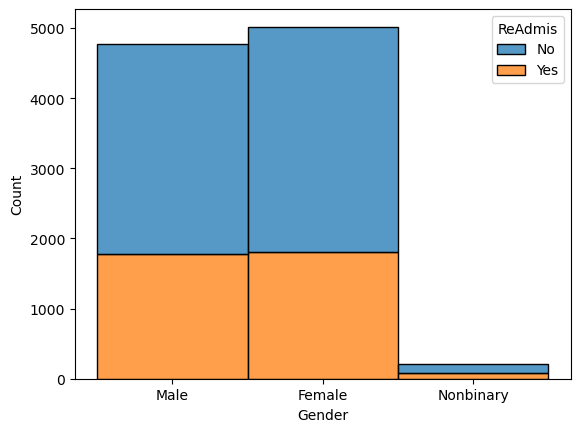
plt.show()



**D1. Bivariate Analysis**



sns.histplot(binwidth=0.5, x= "Gender", hue="ReAdmis", data=df, stat="count", multiple="stack")



**E1. Discuss the results of the hypothesis test.**

The p-value (0.4) is significant and larger than the alpha value of 0.05

Due to the p- value being large, we can not reject the null value which is that gender is not related to hospital readmission.

**E2. Discuss the limitations of your data analysis.**

The chi-square analysis only compares categorical variables. We can not perform a chi-square test between categorical and continuous variables or between two continuous variables.

**E3. Recommend a course of action based on your results.**

Based on the chi-square testing result, the null hypothesis is accepted which means there is no relationship between gender and hospital readmission. This suggests that hospital administrators should provide the same level of care for all genders.

**F. Video**

**G. Third party sources**

De-Yu, Chao. “Chi-Square Test, with Python.” Medium, Towards Data Science, 5 Jan. 2023, https://towardsdatascience.com/chi-square-test-with-python-d8ba98117626.

Gudikandula, P. (2021, March 1). Exploratory data Analysis(Beginner) , Univariate , bivariate and Multivariate - Habberman Dataset. Medium. Retrieved April 16, 2023, from https://purnasaigudikandula.medium.com/exploratory-data-analysis-beginner-univariate-bivariate-and-multivariate-habberman-dataset-2365264b751#:~:text=Bivariate%20analysis%3A,with%20relationship%20between%20two%20values.

“Python - Pearson's Chi-Square Test.” GeeksforGeeks, GeeksforGeeks, 23 June 2020, <https://www.geeksforgeeks.org/python-pearsons-chi-square-test/>.

**H. Source**

**N/A**