#Hypothesis testing

import numpy as np

from scipy import stats

import pandas as pd

data=pd.read\_csv("/content/yulu\_bike\_sharing\_dataset.csv")

import matplotlib.pyplot as plt

data['temp'].value\_counts().plot.hist(bins=40)

data['temp'].value\_counts().mean()

X=data['temp'].value\_counts()

def t\_value(X, h\_0):

    se = np.sqrt(np.var(X) / len(X))

    return (np.mean(X) - h\_0) / se

def p\_value(t):

    # Two-sided p-value, so we multiply by 2.

    return stats.norm.sf(abs(t))\*2

t = t\_value(X, 2.75)

p = p\_value(t)

t,p

import scipy.stats as stats

stats.ttest\_1samp(a=X, popmean=2.75)

if p < 0.05:

     print("Null hypothesis is reject")

else:

    print("fail to reject the null hypothesis")

if t < 0.05:

     print("Null hypothesis is reject")

else:

    print("fail to reject the null hypothesis")