```
import pandas as pd
from statsmodels.stats.proportion import proportions_ztest
data=pd.read_csv("/content/train.csv");
meanage=data['age'].mean();
data['age'].fillna(value=meanage,inplace=True)
# 1)
maledata=data[data['sex']=="male"]
femaledata=data[data['sex']=="female"]
surviveddata=data[data["survived"]==1]
nom=len(maledata)
nof=len(femaledata)
nos=len(surviveddata)
nol=len(data)
mu=28;
#h0 : mu<=28
#h1 : mu>28
a=0.05
x=surviveddata['age']
X=x.mean()
n=nol
h0=28
# print(X,n)
stat,p_value=proportions_ztest(count=X,nobs=n,value=h0,alternative='la
rger')
# print(p_value,stat)
if(stat<(a*100)):
    print("Reject null hypothesis .The average age of passengers in
ship who survived is greater than 28 .")
else:
    print("Failed to reject null hypothesis .The average age of
passengers in ship who survived may or may not be than 28 .")
```

```
Reject null hypothesis .The average age of passengers in ship who
survived is greater than 28 .
```

```
# 2)
maledata=data[data['sex']=="male"]
femaledata=data[data['sex']=="female"]
surviveddata=data[data["survived"]==1]
nom=len(maledata)
nof=len(femaledata)
nos=len(surviveddata)
nol=len(data)
fsdata=surviveddata[surviveddata["sex"]=="female"]
msdata=surviveddata[surviveddata["sex"]=="male"]
meanms=msdata['age'].mean()
meanfs=fsdata['age'].mean()
#h0 : diff=0
#h1 : diff!=0
X=meanfs-meanms
a=0.05
```

```
h0=0;
n=nos
# print(X,n)
stat,p_value=proportions_ztest(count=X,nobs=n,value=h0,alternative='two-sided')
# print(p_value,stat)
if(stat<(a*100)):
    print("Reject null hypothesis .There is a difference in average age between the two genders who survived .")
else:
    print("Failed to reject null hypothesis .There might or might not be a difference in average age between the two genders who survived .")
```

```
Reject null hypothesis .There is a difference in average age between the two genders who survived .
```

```
# 3)
reqcount=0
for i in surviveddata['age']:
  if((i>=20)and(i<=40)):
    reqcount+=1;
X=reqcount/nos
# print(surviveddata['age'].std(),X,n,reqcount)
# print(reqcount)
# print(nos)
n=nos
a=0.05
h0=0.5
#h0 : <=0.50
#h1 : >0.50
# print(X,n)
stat,p_value=proportions_ztest(count=X,nobs=n,value=h0,alternative='larger')
# print(p_value,stat)
if(stat<(a*100)):
    print("Reject null hypothesis .Greater than 50% of passengers who survived in ship are in the age group of 20–40 .")
else:
    print("Failed to reject null hypothesis .Greater than 50% of passengers who survived in ship may or may not be in the age group of 20–40 .")
```

```
Reject null hypothesis .Greater than 50% of passengers who survived in ship are in the age group of 20–40 .
```