HOMEWORK 4 – Austin Rippee

QUESTION #1

Text

Description automatically generated

QUESTION #2A screenshot of a computer

Description automatically generated with medium confidence

QUESTION #3A screenshot of a computer

Description automatically generated with medium confidence

import pandas as pd

df = pd.read\_csv('ratings.csv', header=None)

movieID = df[1].value\_counts().index[0]

movieCount = df[1].value\_counts().values[0]

print('Movie ID:', movieID, 'is rated most often,', movieCount, 'times.')

QUESTION #4

Text

Description automatically generated

This is all the code I tried doing for problem 4. For some reason, I could not get myself to find the answer. I may have missed something trivial, but it was not to be in the time allotted.

import pandas as pd

df = pd.read\_csv('ratings.csv', header=None)

print('Movie ID \t Rating \t # Rating')

#dfs = df[df[2] > 0].groupby(df.index,[1])

#dfs = df.groupby(1)[2].mean().sort\_values(ascending=False).head(5)

# movieIDRating = df.groupby([1])[2].mean().sort\_values(ascending=False).head(5)

#ratingNum1 = df[1].value\_counts().sort\_values(ascending=False).head(5)

#ratingCount1 = df[1].value\_counts().sort\_values(ascending=False).head(5)

#q3 = df.sort\_values([2], ascending=False).groupby([1]).head(5)

# q3 = (df[1].to\_string(index=False)).head(5)

# q1 = df[1].value\_counts().head(5)

#dfSize = df.groupby([1]).size().sort\_values(ascending=False)

#df1 = df.groupby([1])[2].mean()

#df1[df[] >= 10]

#df1.sort\_values(by=[2L], ascending=False)

#df.groupby([1]).filter(.size() >= 10).groupby([1]).agg({}).sort\_values([2],ascending=False)

df1 = df.groupby([1]).filter(.size() >= 10).groupby([1]).agg({}).sort\_values([2].mean(), ascending = False)

print(df1)