ASSIGNMENTS

1. Find correlation of male_age with female_age.
2. Find correlation of 1st, 2nd and 3rd class male_age and female_age.
3. Find correlation of 1st, 2nd and 3rd class fare.

1. Find correlation of male_age with female_age.

```
In [ ]: import warnings
        warnings.filterwarnings('ignore')
```

```
In [ ]: import pandas as pd
        import seaborn as sns
        import numpy as np
        import matplotlib.pyplot as plt

        df = sns.load_dataset('titanic')
```

```
In [ ]: df1 = df[['sex', 'age']]
        df1.isnull().sum()
        df1 = df1.dropna(subset=['age'], axis=0)
        df1.head()
```

Out[ ]:

	sex	age
0	male	22.0
1	female	38.0
2	female	26.0
3	female	35.0
4	male	35.0

```
In [ ]: df1_male = df1.loc[df1['sex']=='male']
        df1_female = df1.loc[df1['sex']=='female']
        df1_female.head()
```

Out[ ]:

	sex	age
1	female	38.0
2	female	26.0
3	female	35.0
8	female	27.0
9	female	14.0

```
In [ ]: df1_male.rename(columns={'sex':"sex_m", 'age':'age_m'}, inplace=True)
        df1_female.rename(columns={'sex':"sex_f", 'age':'age_f'}, inplace=True)
        df1_female.head()
```

Out[ ]:

	sex_f	age_f
1	female	38.0
2	female	26.0
3	female	35.0
8	female	27.0
9	female	14.0

```
In [ ]: df1_male = df1_male.sample(n=261)
        df1_female = df1_female.sample(n=261)
        df1_female.shape
```

Out[ ]: (261, 2)

In [ ]:
```python
df1_male1=df1_male.loc[::-1].reset_index(drop=True)
df1_female1=df1_female.loc[::-1].reset_index(drop=True)
df1_female1.head()
```

Out[ ]:

	sex_f	age_f
0	female	4.0
1	female	26.0
2	female	4.0
3	female	50.0
4	female	15.0

In [ ]:
```python
df2=pd.concat([df1_male1, df1_female1], axis=1)
df2.head()
```

Out[ ]:

	sex_m	age_m	sex_f	age_f
0	male	40.0	female	4.0
1	male	31.0	female	26.0
2	male	34.0	female	4.0
3	male	31.0	female	50.0
4	male	44.0	female	15.0

In [ ]:
```python
df2_corr = df2.corr(method='pearson')
df2_corr
```

Out[ ]:

	age_m	age_f
age_m	1.000000	-0.117228
age_f	-0.117228	1.000000

In [ ]:
```python
sns.heatmap(df2_corr, annot=True)
```

Out[ ]: <AxesSubplot:>
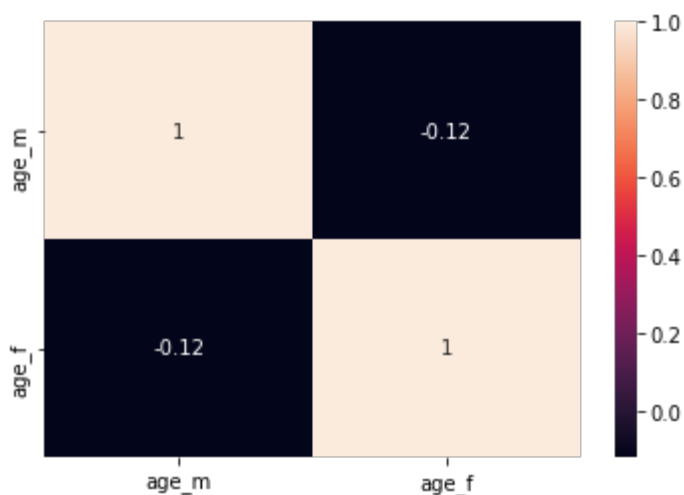
In [ ]:
```python
df1.head()
```

Out[ ]:

	sex	age
0	male	22.0
1	female	38.0
2	female	26.0
3	female	35.0
4	male	35.0

```
In [ ]:  plt.figure(figsize=(10, 5))
         sns.histplot(data=df1, x='age', hue='sex', multiple='dodge', binwidth=5)
```

Out[ ]: <AxesSubplot:xlabel='age', ylabel='Count'>
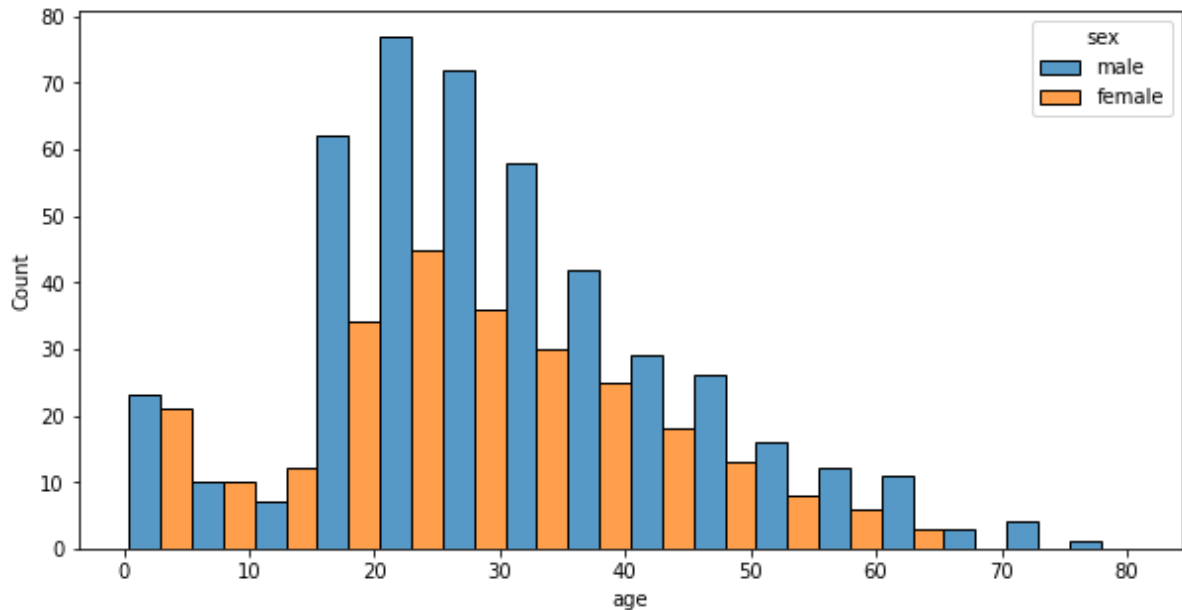
2. Find correlation of 1st, 2nd and 3rd class male_age and female_age.

```
In [ ]:  df.head()
```

Out[ ]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class	who	adult_male	dec
0	0	3	male	22.0	1	0	7.2500	S	Third	man	True	Na
1	1	1	female	38.0	1	0	71.2833	C	First	woman	False	
2	1	3	female	26.0	0	0	7.9250	S	Third	woman	False	Na
3	1	1	female	35.0	1	0	53.1000	S	First	woman	False	
4	0	3	male	35.0	0	0	8.0500	S	Third	man	True	Na

```
In [ ]:  df1 = df[['sex', 'age', 'class']]
         df1.isnull().sum()
         df1 = df1.dropna(subset=['age'], axis=0)
         df1.head()
```

Out[ ]:

	sex	age	class
0	male	22.0	Third
1	female	38.0	First
2	female	26.0	Third
3	female	35.0	First
4	male	35.0	Third

```
In [ ]:  df1.isnull().sum()
```

Out[ ]: sex   0
         age   0
         class   0
         dtype: int64

```
In [ ]:  df1_male = df1.loc[df1['sex']=='male']
         df1_female = df1.loc[df1['sex']=='female']
         df1_female.head()
```

Out[ ]:

	sex	age	class
1	female	38.0	First
2	female	26.0	Third
3	female	35.0	First
8	female	27.0	Third
9	female	14.0	Second

In [ ]:
```python
df1_male.head()
```

Out[ ]:

	sex	age	class
0	male	22.0	Third
4	male	35.0	Third
6	male	54.0	First
7	male	2.0	Third
12	male	20.0	Third

In [ ]:
```python
df1_male_first = df1_male.loc[df1_male['class']=='First']
df1_male_second = df1_male.loc[df1_male['class']=='Second']
df1_male_third = df1_male.loc[df1_male['class']=='Third']
```

In [ ]:
```python
df1_male_third.head()
```

Out[ ]:

	sex	age	class
0	male	22.0	Third
4	male	35.0	Third
7	male	2.0	Third
12	male	20.0	Third
13	male	39.0	Third

In [ ]:
```python
df1_male_first = df1_male_first.sample(n=74)
df1_male_second = df1_male_second.sample(n=74)
df1_male_third = df1_male_third.sample(n=74)
```

In [ ]:
```python
df1_male_third.head()
```

Out[ ]:

	sex	age	class
532	male	17.0	Third
296	male	23.5	Third
146	male	27.0	Third
743	male	24.0	Third
225	male	22.0	Third

In [ ]:
```python
df1_male_first.rename(columns={'sex':"1st class male", 'age':'1st class male age','
df1_male_second.rename(columns={'sex':"2nd class male", 'age':'2nd class male age',
df1_male_third.rename(columns={'sex':"3rd class male", 'age':'3rd class male age','
```

In [ ]:
```python
df1_male_third.head()
```

Out[ ]:

	3rd class male	3rd class male age	3rd class male class
532	male	17.0	Third
296	male	23.5	Third
146	male	27.0	Third
743	male	24.0	Third
225	male	22.0	Third

In [ ]:
```python
df1_male_first=df1_male_first.loc[::-1].reset_index(drop=True)
df1_male_second=df1_male_second.loc[::-1].reset_index(drop=True)
df1_male_third=df1_male_third.loc[::-1].reset_index(drop=True)
```

In [ ]:
```python
df1_male_third.head()
```

Out[ ]:

	3rd class male	3rd class male age	3rd class male class
0	male	10.0	Third
1	male	37.0	Third
2	male	32.0	Third
3	male	9.0	Third
4	male	19.0	Third

In [ ]:
```python
df1_male_third.shape
```

Out[ ]: (74, 3)

In [ ]:
```python
df1_female_first = df1_female.loc[df1_female['class']=='First']
df1_female_second = df1_female.loc[df1_female['class']=='Second']
df1_female_third = df1_female.loc[df1_female['class']=='Third']
```

In [ ]:
```python
df1_female_first.head()
```

Out[ ]:

	sex	age	class
1	female	38.0	First
3	female	35.0	First
11	female	58.0	First
52	female	49.0	First
61	female	38.0	First

In [ ]:
```python
df1_female_first = df1_female_first.sample(n=74)
df1_female_second = df1_female_second.sample(n=74)
df1_female_third = df1_female_third.sample(n=74)
```

In [ ]:
```python
df1_female_third.shape
```

Out[ ]: (74, 3)

In [ ]:
```python
df1_female_first.rename(columns={'sex':"1st class female", 'age':'1st class female
df1_female_second.rename(columns={'sex':"2nd class female", 'age':'2nd class female
df1_female_third.rename(columns={'sex':"3rd class female", 'age':'3rd class female
```

In [ ]:
```python
df1_female_second.head()
```

Out[ ]:

	2nd class female	2nd class female age	2st class female class
312	female	26.0	Second
98	female	34.0	Second
416	female	34.0	Second
211	female	35.0	Second
426	female	28.0	Second

In [ ]:
```python
df1_female_first=df1_female_first.loc[::-1].reset_index(drop=True)
df1_female_second=df1_female_second.loc[::-1].reset_index(drop=True)
df1_female_third=df1_female_third.loc[::-1].reset_index(drop=True)
```

In [ ]:
```python
df1_female_third.head()
```

Out[ ]:

	3rd class female	3rd class female age	3rd class female class
0	female	6.0	Third
1	female	8.0	Third
2	female	9.0	Third
3	female	37.0	Third
4	female	18.0	Third

In [ ]:
```python
df2 = pd.concat([df1_male_first,df1_male_second,df1_male_third,
                 df1_female_first,df1_female_second,df1_female_third], axis=1)
```

In [ ]:
```python
df2.head()
```

Out[ ]:

	1st class male	1st class male age	1st class male class	2nd class male	2nd class male age	2st class male class	3rd class male	3rd class male age	3rd class male class	1st class female	1st class female age	1st class female class	2nd class female	f
0	male	36.0	First	male	34.0	Second	male	10.0	Third	female	18.0	First	female	
1	male	4.0	First	male	42.0	Second	male	37.0	Third	female	39.0	First	female	
2	male	17.0	First	male	34.0	Second	male	32.0	Third	female	39.0	First	female	
3	male	25.0	First	male	1.0	Second	male	9.0	Third	female	54.0	First	female	
4	male	19.0	First	male	24.0	Second	male	19.0	Third	female	56.0	First	female	

In [ ]:
```python
df2.shape
```

Out[ ]: (74, 18)

In [ ]:
```python
df2_corr = df2.corr(method='pearson')
df2_corr
```
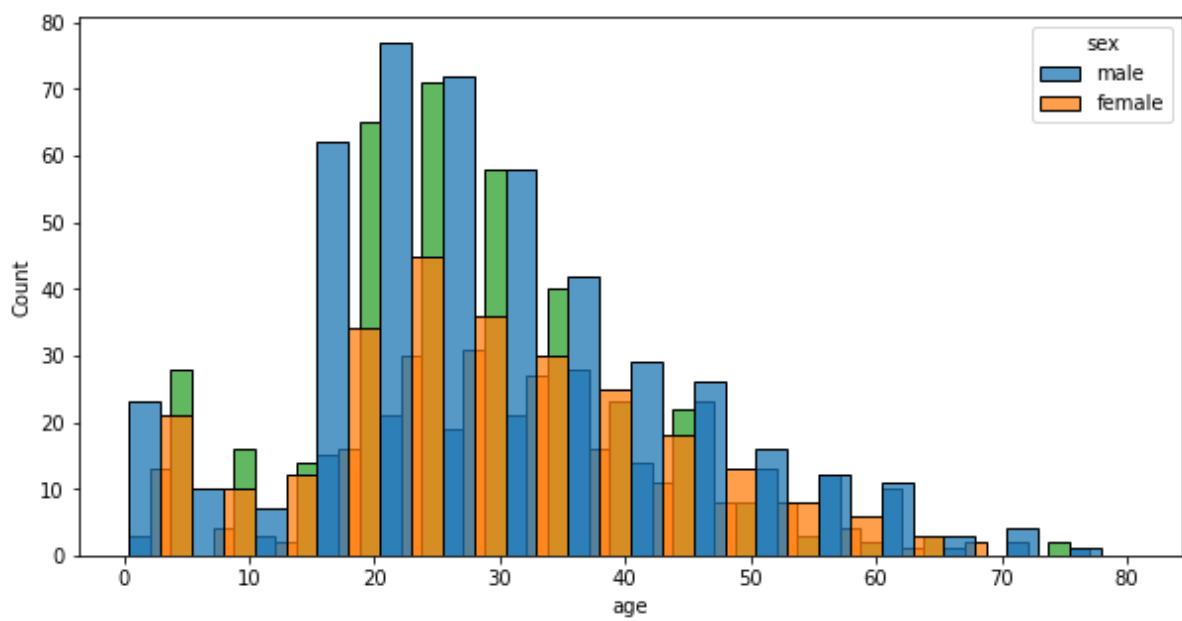
Out[ ]:

	1st class male age	2nd class male age	3rd class male age	1st class female age	2nd class female age	3rd class female age
1st class male age	1.000000	-0.015920	-0.027114	-0.253996	0.245717	0.016768
2nd class male age	-0.015920	1.000000	0.152031	-0.187000	0.078146	0.043788
3rd class male age	-0.027114	0.152031	1.000000	-0.136611	-0.111154	0.046124
1st class female age	-0.253996	-0.187000	-0.136611	1.000000	0.080752	0.028905
2nd class female age	0.245717	0.078146	-0.111154	0.080752	1.000000	0.066025
3rd class female age	0.016768	0.043788	0.046124	0.028905	0.066025	1.000000

In [ ]:
```python
plt.figure(figsize=(15,10))
sns.heatmap(df2_corr, annot=True)
```

Out[ ]: <AxesSubplot:>
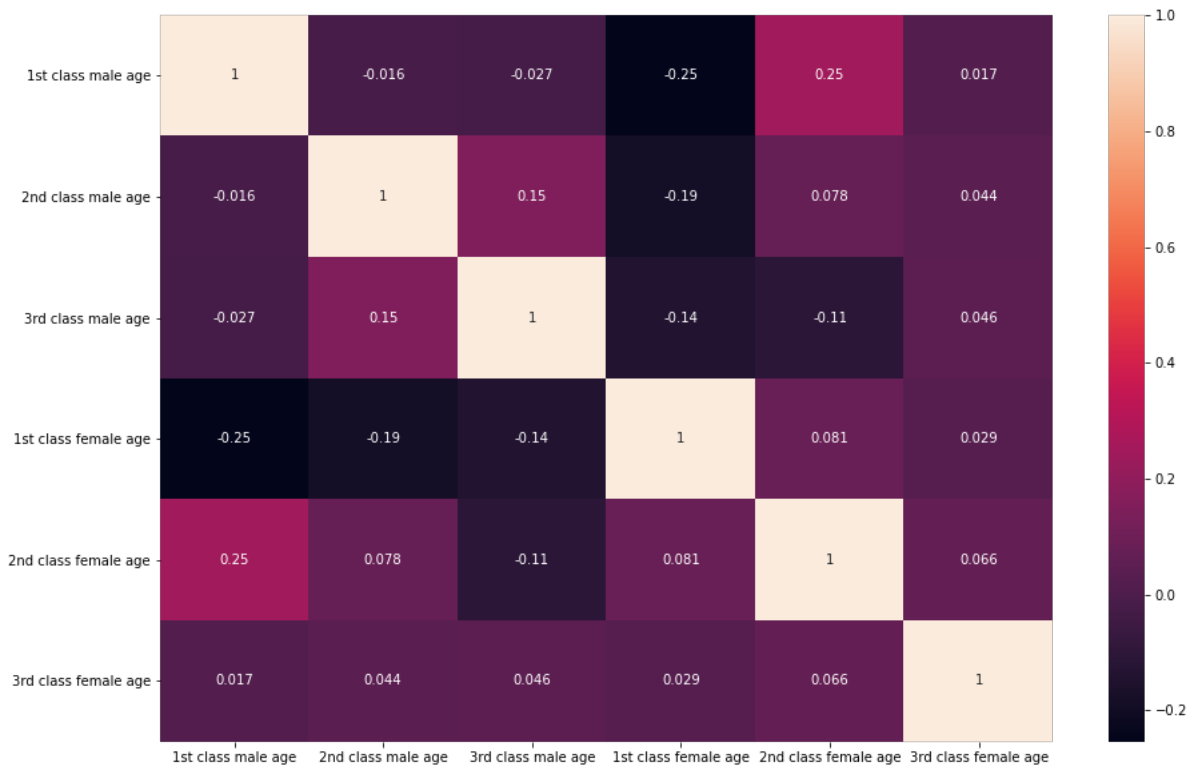
In [ ]:
```python
plt.figure(figsize=(10, 5))
sns.histplot(data=df, x='age', hue='class', multiple='dodge', binwidth=5,)
sns.histplot(data=df, x='age', hue='sex', multiple='dodge', binwidth=5)
plt.show()
```

In [ ]: