In [1]:

```python
import pandas as pd
titanic_data=pd.read_csv('./Downloads/titanic_data.csv')
titanic_data.head()
```

Out[1]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	

In [2]:

```python
titanic_pclass1=(titanic_data.Pclass == 1)
titanic_pclass1
```

Out[2]:

```
0      False
1       True
2      False
3       True
4      False
       ...
886    False
887     True
888    False
889     True
890    False
Name: Pclass, Length: 891, dtype: bool
```

In [3]:

```python
titanic_pclass1=(titanic_data.Pclass == 1)
titanic_pclass1_data = titanic_data[titanic_pclass1]
titanic_pclass1_data.head()
```

Out[3]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000

In [4]:

```
titanic_pclass_data = titanic_data[titanic_data.Pclass == 1]
titanic_pclass_data.head()
```

Out[4]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C
6	7	0	1	McCarthy, Mr. Timothy J	male	54.0	0	0	17463	51.8625	
11	12	1	1	Bonnell, Miss. Elizabeth	female	58.0	0	0	113783	26.5500	C
23	24	1	1	Sloper, Mr. William Thompson	male	28.0	0	0	113788	35.5000	

In [5]:

```
ages=[20,21,22]
age_dataset =titanic_data[titanic_data["Age"].isin(ages)]
age_dataset.head()
```

Out[5]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Ca
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.25	N
12	13	0	3	Saundercock, Mr. William Henry	male	20.0	0	0	A/5. 2151	8.05	N
37	38	0	3	Cann, Mr. Ernest Charles	male	21.0	0	0	A./5. 2152	8.05	N
51	52	0	3	Nosworthy, Mr. Richard Cater	male	21.0	0	0	A/4. 39886	7.80	N
56	57	1	2	Rugg, Miss. Emily	female	21.0	0	0	C.A. 31026	10.50	N

In [6]:

```
ages = [20,21,22]
ageclass_dataset = titanic_data[titanic_data["Age"].isin(ages) & (titanic_data["Pclass"] ==
ageclass_dataset.head
```

Out[6]:

```
<bound method NDFrame.head of      PassengerId  Survived  Pclass
Name   \
102          103         0       1           White, Mr. Richard Frasar
151          152         1       1        Pears, Mrs. Thomas (Edith Wearne)
356          357         1       1          Bowerman, Miss. Elsie Edith
373          374         0       1                    Ringhini, Mr. Sante
539          540         1       1       Frolicher, Miss. Hedwig Margaritha
627          628         1       1         Longley, Miss. Gretchen Fiske
708          709         1       1                    Cleaver, Miss. Alice
742          743         1       1  Ryerson, Miss. Susan Parker "Suzette"

         Sex   Age  SibSp  Parch    Ticket      Fare          Cabin Embarke
d
102     male  21.0      0      1     35281   77.2875            D26
S
151   female  22.0      1      0    113776   66.6000             C2
S
356   female  22.0      0      1    113505   55.0000            E33
S
373     male  22.0      0      0  PC 17760  135.6333            NaN
C
539   female  22.0      0      2     13568   49.5000            B39
C
627   female  21.0      0      0     13502   77.9583             D9
S
708   female  22.0      0      0    113781  151.5500            NaN
S
742   female  21.0      2      2  PC 17608  262.3750  B57 B59 B63 B66
C  >
```

In [7]:

```
titanic_data_filter = titanic_data.filter(["Name","Sex","Age"])
titanic_data_filter.head()
```

Out[7]:

	Name	Sex	Age
0	Braund, Mr. Owen Harris	male	22.0
1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0
2	Heikkinen, Miss. Laina	female	26.0
3	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0
4	Allen, Mr. William Henry	male	35.0

In [8]:

```python
titanic_data_filter = titanic_data.drop(["Name","Sex","Age"], axis = 1)
titanic_data_filter.head()
```

Out[8]:

	PassengerId	Survived	Pclass	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	1	0	PC 17599	71.2833	C85	C
2	3	1	3	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	1	0	113803	53.1000	C123	S
4	5	0	3	0	0	373450	8.0500	NaN	S

In [9]:

```python
titanic_pclass1_data = titanic_data[titanic_data.Pclass ==1]
print(titanic_pclass1_data.shape)

titanic_pclass2_data = titanic_data[titanic_data.Pclass ==2]
print(titanic_pclass2_data.shape)
```

```
(216, 12)
(184, 12)
```

In [10]:

```python
final_data = titanic_pclass1_data.append(titanic_pclass2_data,ignore_index=True)
print(final_data.shape)
```

```
(400, 12)
```

In [11]:

```python
final_data = pd.concat([titanic_pclass1_data, titanic_pclass2_data])
print(final_data.shape)
```

```
(400, 12)
```

In [12]:

```python
df1 =final_data[:200]
print(df1.shape)

df2 = final_data[200:]
print(df2.shape)
```

```
(200, 12)
(200, 12)
```

In [13]:

```python
df1 =final_data[:200]
print(df1.shape)

df2 = final_data[200:]
print(df2.shape)

final_data2 = pd.concat([df1,df2],axis = 1, ignore_index = True)
print(final_data2.shape)
```

```
(200, 12)
(200, 12)
(400, 24)
```

In [14]:

```python
age_sorted_data = titanic_data.sort_values(by=['Age'])
age_sorted_data.head()
```

Out[14]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
803	804	1	3	Thomas, Master. Assad Alexander	male	0.42	0	1	2625	8.5167
755	756	1	2	Hamalainen, Master. Viljo	male	0.67	1	1	250649	14.5000
644	645	1	3	Baclini, Miss. Eugenie	female	0.75	2	1	2666	19.2583
469	470	1	3	Baclini, Miss. Helene Barbara	female	0.75	2	1	2666	19.2583
78	79	1	2	Caldwell, Master. Alden Gates	male	0.83	0	2	248738	29.0000

In [15]:

```python
age_sorted_data = titanic_data.sort_values(by=['Age','Fare'],ascending =False)
age_sorted_data.head()
```

Out[15]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
630	631	1	1	Barkworth, Mr. Algernon Henry Wilson	male	80.0	0	0	27042	30.0000
851	852	0	3	Svensson, Mr. Johan	male	74.0	0	0	347060	7.7750
493	494	0	1	Artagaveytia, Mr. Ramon	male	71.0	0	0	PC 17609	49.5042
96	97	0	1	Goldschmidt, Mr. George B	male	71.0	0	0	PC 17754	34.6542
116	117	0	3	Connors, Mr. Patrick	male	70.5	0	0	370369	7.7500

In [16]:

```python
updated_class = titanic_data.Pclass.apply(lambda x : x + 2)
updated_class.head()
```

Out[16]:

```
0    5
1    3
2    5
3    3
4    5
Name: Pclass, dtype: int64
```

In [17]:

```python
def mult(x):
    return x * 2

updated_class = titanic_data.Pclass.apply(mult)
updated_class.head()
```

Out[17]:

```
0    6
1    2
2    6
3    2
4    6
Name: Pclass, dtype: int64
```

In [18]:

```python
import matplotlib.pyplot as plt
import seaborn as sns

flights_data = sns.load_dataset('flights')

flights_data.head()
```

Out[18]:

	year	month	passengers
0	1949	January	112
1	1949	February	118
2	1949	March	132
3	1949	April	129
4	1949	May	121

In [ ]:

In [19]:

```python
flights_data_pivot = flights_data.pivot_table(index='month',columns = 'year',values = 'pass
flights_data_pivot.head()
```

Out[19]:

year	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
month												
January	112	115	145	171	196	204	242	284	315	340	360	417
February	118	126	150	180	196	188	233	277	301	318	342	391
March	132	141	178	193	236	235	267	317	356	362	406	419
April	129	135	163	181	235	227	269	313	348	348	396	461
May	121	125	172	183	229	234	270	318	355	363	420	472

In [20]:

```python
import pandas as pd
titanic_data = pd.read_csv('./Downloads/titanic_data.csv')
titanic_data.head()

pd.crosstab(titanic_data.Pclass, titanic_data.Age, margins = True)
```

Out[20]:

Age	0.42	0.67	0.75	0.83	0.92	1.0	2.0	3.0	4.0	5.0	...	63.0	64.0	65.0	66.0	70.0	70
Pclass																	
1	0	0	0	0	1	0	1	0	1	0	...	1	2	2	0	1	
2	0	1	0	2	0	2	2	3	2	1	...	0	0	0	1	1	
3	1	0	2	0	0	5	7	3	7	3	...	1	0	1	0	0	
All	1	1	2	2	1	7	10	6	10	4	...	2	2	3	1	2	

4 rows × 89 columns

In [21]:

```python
import numpy as np
titanic_data.Fare = np.where(titanic_data.Age > 20,titanic_data.Fare +5, titanic_data.Fare)

titanic_data.head()
```

Out[21]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	12.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	76.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	12.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	58.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	13.0500

In [22]:

```python
import numpy as np
titanic_data.Fare = np.where(titanic_data.Age > 20,titanic_data.Fare -10, titanic_data.Fare

titanic_data.head()
```

Out[22]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	2.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	66.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	2.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	48.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	3.0500

In [23]:

```python
updated_class = titanic_data.Fare.apply(lambda x : x -10)
updated_class.head()
```

Out[23]:

```
0    -7.7500
1    56.2833
2    -7.0750
3    38.1000
4    -6.9500
Name: Fare, dtype: float64
```

In [24]:

```python
import seaborn as sns
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import warnings
warnings.filterwarnings('ignore')
```

In [25]:

```
employee_data = pd.read_csv('./Downloads/Absenteeism_at_work.csv')

employee_data.head()
```

Out[25]:

	ID	Reason for absence	Month of absence	Day of the week	Seasons	Transportation expense	Distance from Residence to Work	Service time	Age	Work load Average/da
0	11	26	7	3	1	289	36	13	33	239,55
1	36	0	7	3	1	118	13	18	50	239,55
2	3	23	7	4	1	179	51	18	38	239,55
3	7	7	7	5	1	279	5	14	39	239,55
4	11	23	7	5	1	289	36	13	33	239,55

5 rows × 21 columns

In [26]:

```
ages = [33,50,38]
age_dataset = employee_data[employee_data["Age"].isin(ages) & (employee_data["Age"] ==1)]
age_dataset.head()
```

Out[26]:

	ID	Reason for absence	Month of absence	Day of the week	Seasons	Transportation expense	Distance from Residence to Work	Service time	Age	Work load Average/day

0 rows × 21 columns

In [27]:

```python
ages = [33,50,38]
age_dataset = employee_data[employee_data["Age"].isin(ages) & (employee_data["Age"] >1)]
age_dataset.head()
```

Out[27]:

	ID	Reason for absence	Month of absence	Day of the week	Seasons	Transportation expense	Distance from Residence to Work	Service time	Age	Work Average
0	11	26	7	3	1	289	36	13	33	239
1	36	0	7	3	1	118	13	18	50	239
2	3	23	7	4	1	179	51	18	38	239
4	11	23	7	5	1	289	36	13	33	239
5	3	23	7	6	1	179	51	18	38	239

5 rows × 21 columns

In [28]:

```python
employee_data_filter = employee_data.filter(["Age","Weight","Height"])
employee_data_filter
```

Out[28]:

	Age	Weight	Height
0	33	90	172
1	50	98	178
2	38	89	170
3	39	68	168
4	33	90	172
...	...	...	...
735	33	90	172
736	37	88	172
737	40	98	170
738	39	100	170
739	53	77	175

740 rows × 3 columns

In [29]:

```python
import pandas as pd

titanic_data = pd.read_csv('./Downloads/titanic_data.csv')

titanic_data.head()
```

Out[29]:

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500

In [30]:

```python
import matplotlib.pyplot as plt
titanic_data["Age"].hist()
```

Out[30]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f611c6c088>
```
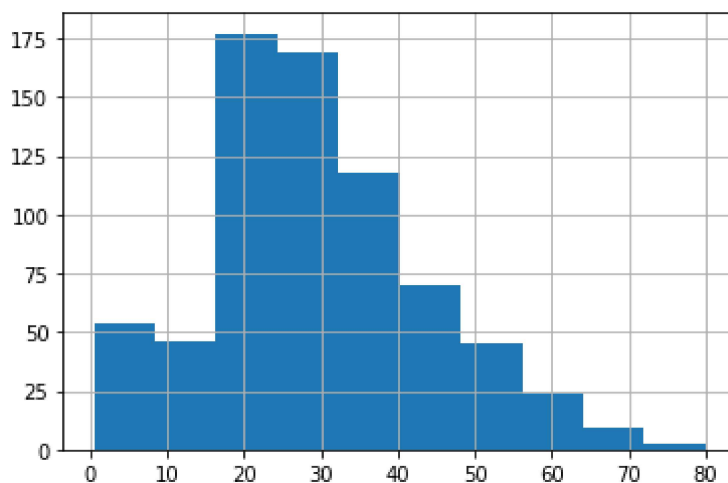
In [31]:

```python
import matplotlib.pyplot as plt
titanic_data['Age'].plot(kind='hist')
```

Out[31]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f613f45048>
```
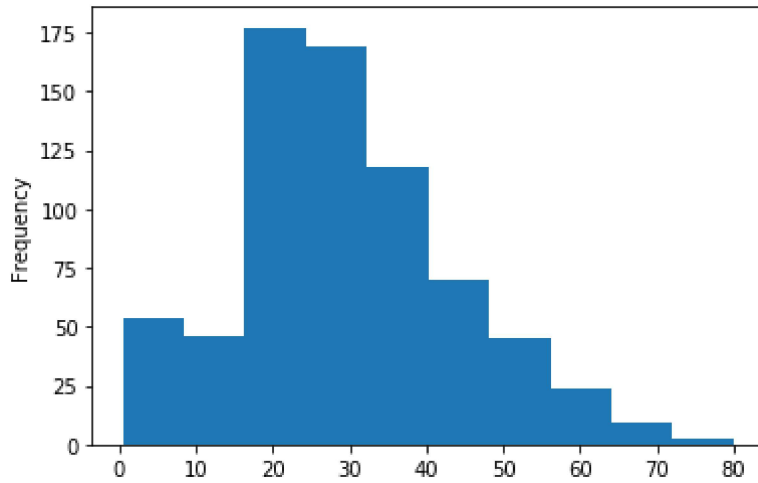

In [32]:

```python
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_style('darkgrid')
titanic_data['Age'].hist(bins =20)
```

Out[32]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f613fdf908>
```
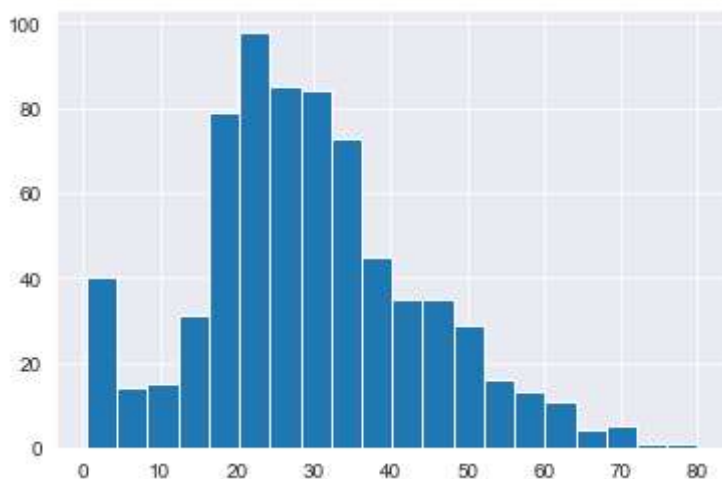
In [33]:

```python
titanic_data['Age'].hist(bins = 20, color = 'orange')
```

Out[33]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f6140679c8>
```
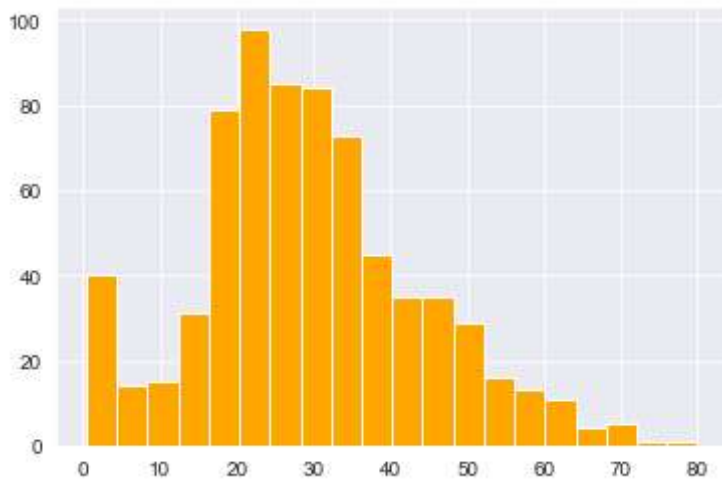

In [34]:

```python
flights_data = sns.load_dataset('flights')

flights_data.head()
```

Out[34]:

	year	month	passengers
0	1949	January	112
1	1949	February	118
2	1949	March	132
3	1949	April	129
4	1949	May	121

In [35]:

```python
flights_data.plot.line(y='passengers', figsize=(8,6))
```

Out[35]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f614125648>
```
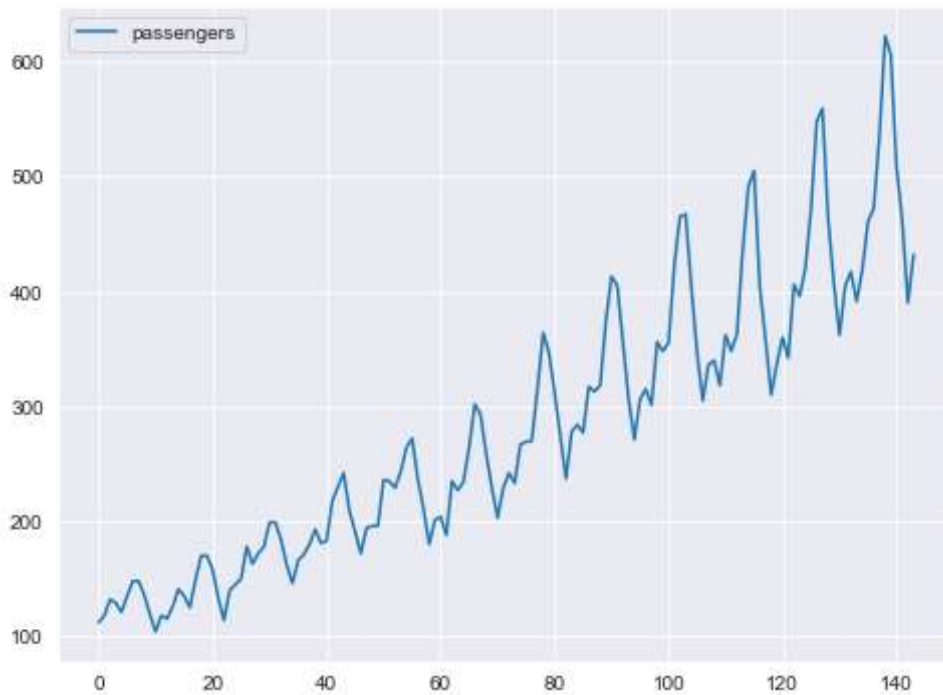
In [36]:

```python
flights_data.plot.line(x = 'year', y='passengers', figsize=(8,6),color ='orange')
```

Out[36]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f61419ab08>
```
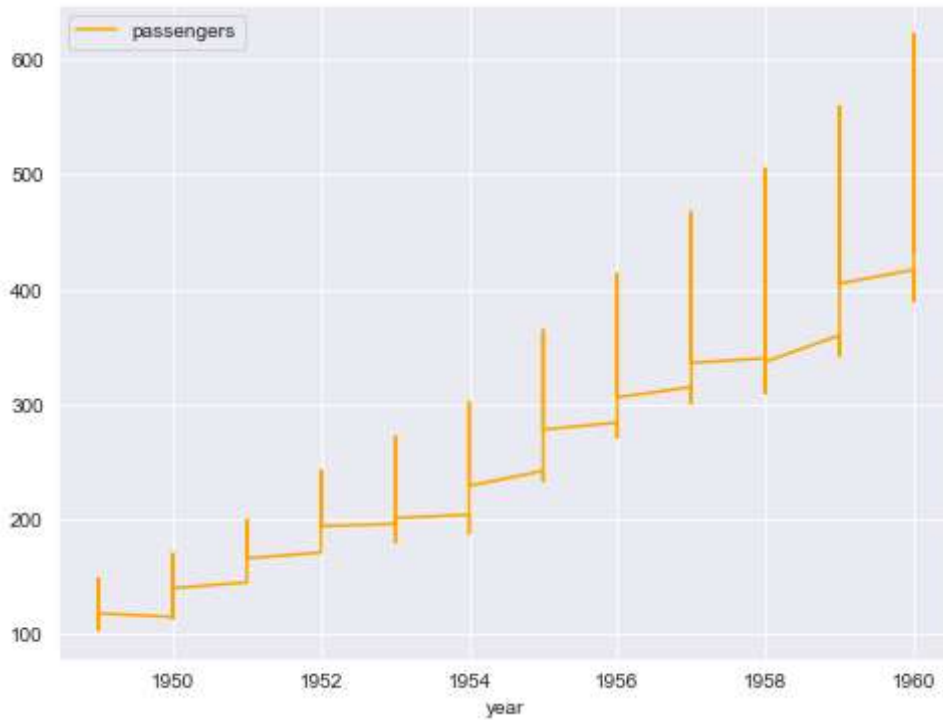
In [37]:

```python
flights_data.plot.scatter(x='year', y ='passengers', figsize =(8,6))
```
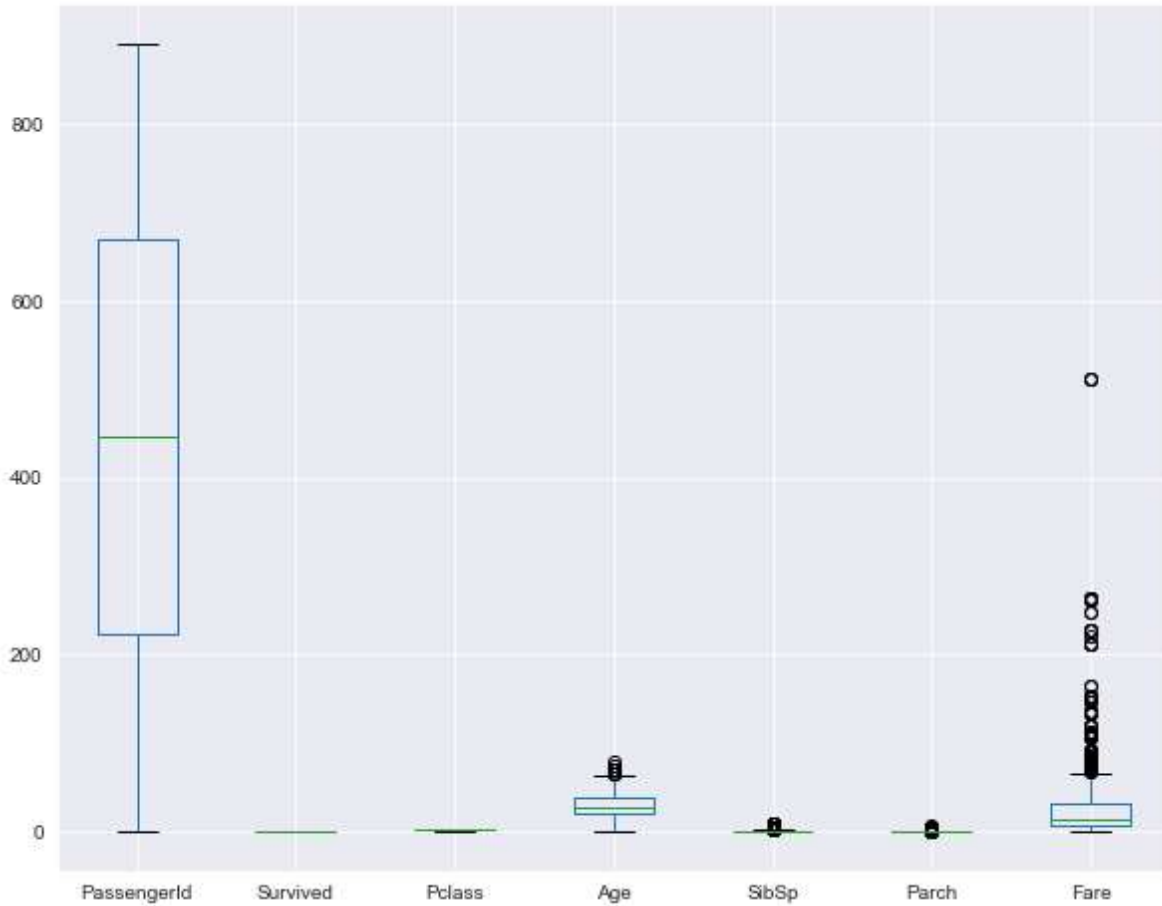
Out[37]:

<matplotlib.axes._subplots.AxesSubplot at 0x1f6142238c8>
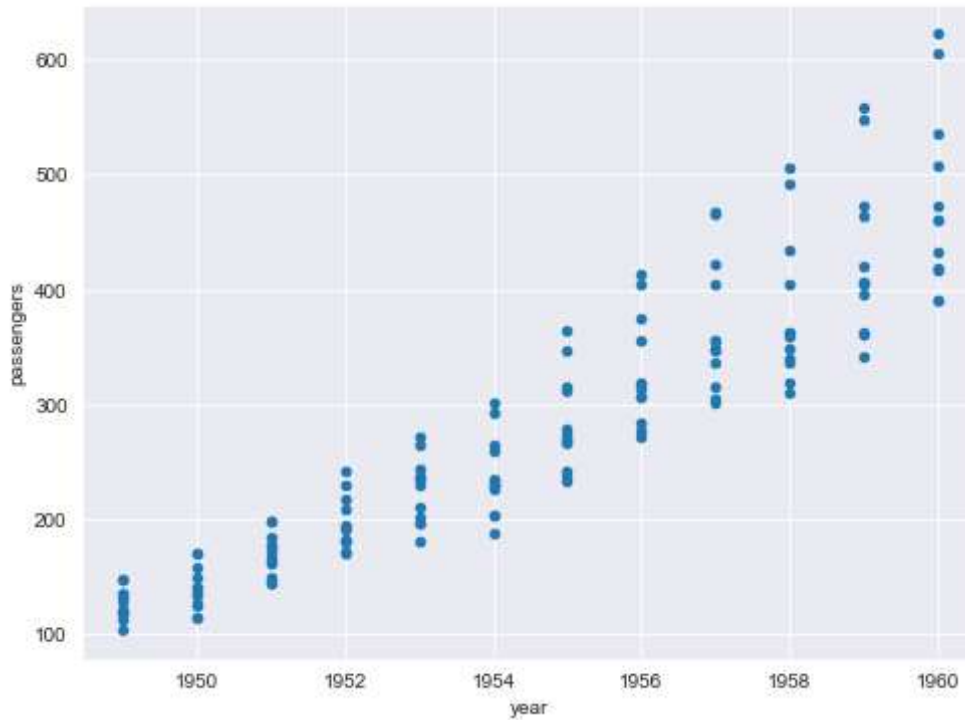
In [40]:

```
titanic_data = pd.read_csv('./Downloads/titanic_data.csv')
titanic_data.plot.box(figsize=(10,8))
```
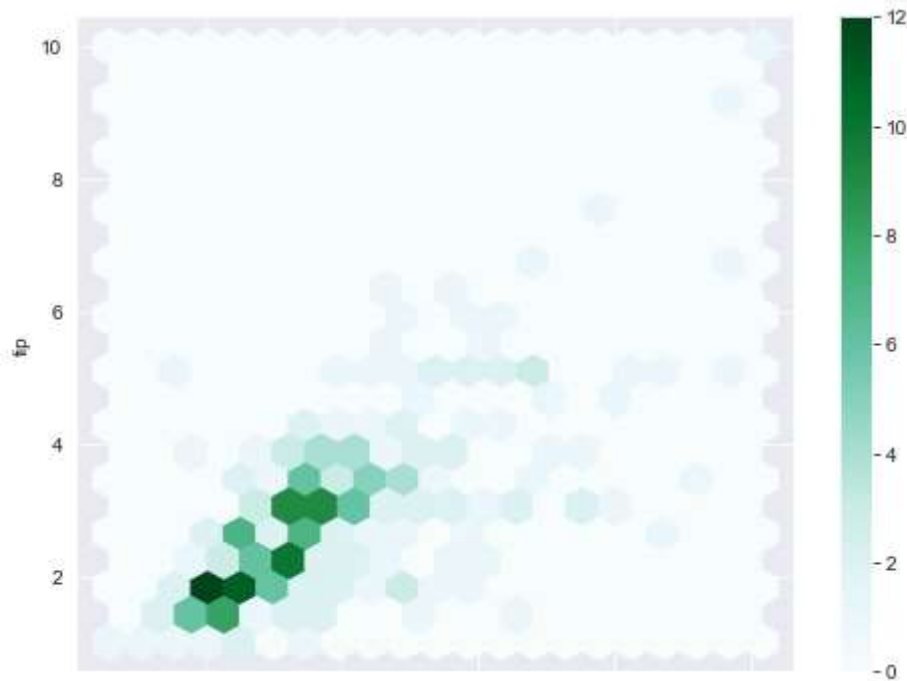
Out[40]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f6142e54c8>
```

In [41]:

```python
tips_data =sns.load_dataset('tips')

tips_data.plot.hexbin(x='total_bill',y='tip',gridsize= 20,figsize=(8,6))
```
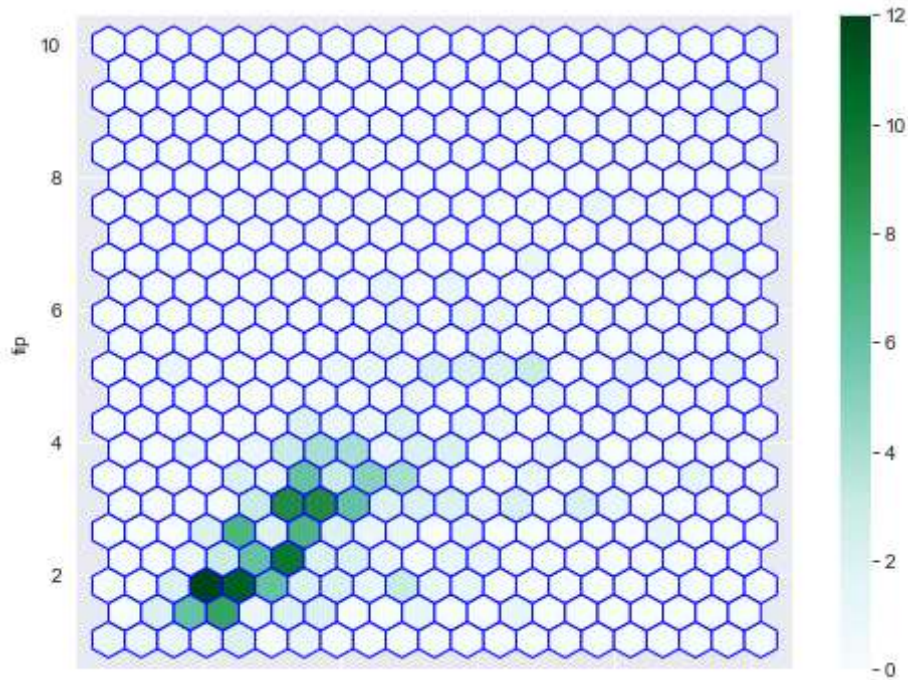
Out[41]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f6145f7108>
```

In [42]:

```python
tips_data.plot.hexbin(x='total_bill', y = 'tip', gridsize=20, figsize=(8,6), color = 'blue'
```
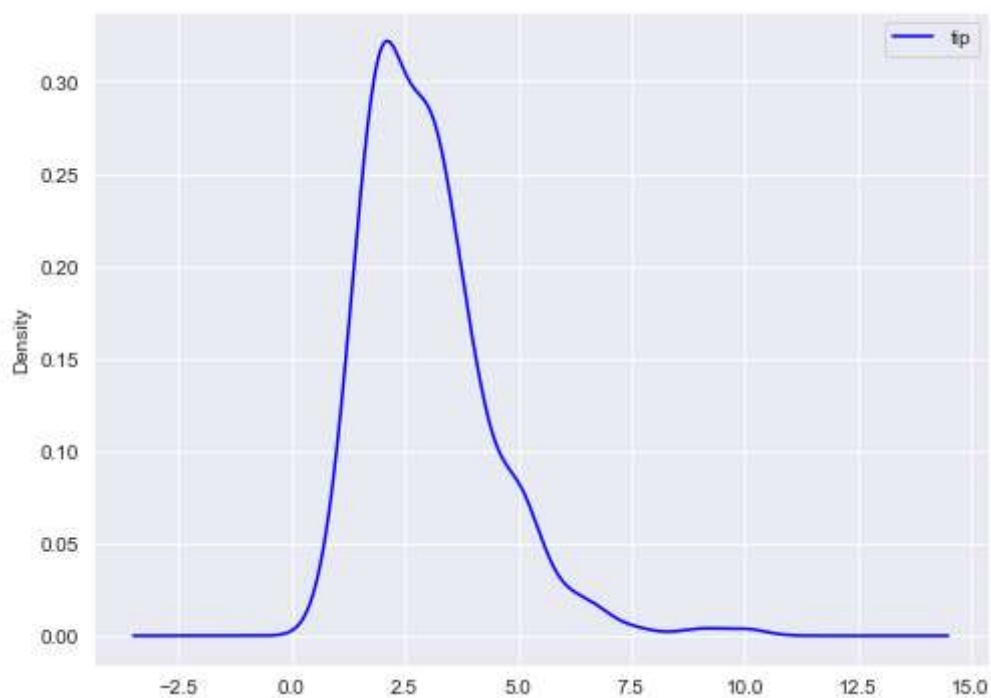
Out[42]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f6144be7c8>
```


In [43]:

```python
tips_data.plot.kde(y='tip', figsize=(8,6), color= 'blue')
```

Out[43]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1f61456ce88>
```

In [44]:

```python
import pandas as pd
```

In [45]:

```python
google_stock = pd.read_csv('./Downloads/google_data.csv')
google_stock.head()
```

Out[45]:

	Date	Open	High	Low	Close	Adj Close	Volume
0	2015-01-06	513.589966	514.761719	499.678131	500.585632	500.585632	2899900
1	2015-01-07	505.611847	505.855164	498.281952	499.727997	499.727997	2065000
2	2015-01-08	496.626526	502.101471	489.655640	501.303680	501.303680	3353500
3	2015-01-09	503.377991	503.537537	493.435272	494.811493	494.811493	2071300
4	2015-01-12	493.584869	494.618011	486.225067	491.201416	491.201416	2326700

In [ ]: