```python
In [25]:   import pandas as pd
           import numpy as np
           import matplotlib.pyplot as plt
```

```python
In [29]:   df = pd.read_csv("C:/Users/prajw/Downloads/SPPU-TE-COMP-SEM-2-2019-PATTERN-DSBDA-main/SPPU-TE-COMP-SEM-2-2019-PATTERN-DSBDA-mai
           df
```

Out[29]:

	Rollno	Name	Gender	Branch	Attendence	Phy_marks	Che_marks	EM1_marks	PPS_marks	SME_marks	Total Marks	Percentage
0	1	Mohammed	M	Comp	72.0	62.0	98.0	63.0	89.0	36.0	368	73.6
1	2	Reyansh	M	IT	58.0	62.0	83.0	83.0	88.0	34.0	350	70.0
2	3	Aarav	M	IT	57.0	-20.0	100.0	NaN	56.0	36.0	192	38.4
3	4	Atharv	M	IT	60.0	89.0	83.0	70.0	33.0	23.0	298	59.6
4	5	Vivaan	M	Comp	85.0	90.0	NaN	78.0	23.0	56.0	247	49.4
5	6	Advik	M	ENTC	94.0	99.0	84.0	100.0	56.0	99.0	438	87.6
6	7	Ansh	M	ENTC	98.0	88.0	95.0	81.0	78.0	78.0	420	84.0
7	8	Ishaan	M	ENTC	75.0	66.0	51.0	83.0	-99.0	76.0	192	38.4
8	9	Dhruv	M	ENTC	63.0	NaN	NaN	97.0	56.0	55.0	208	41.6
9	10	Siddharth	M	ENTC	96.0	67.0	78.0	95.0	NaN	98.0	338	67.6
10	11	Vihaan	M	ENTC	82.0	54.0	70.0	88.0	55.0	56.0	323	64.6
11	12	NaN	M	IT	75.0	64.0	67.0	71.0	66.0	87.0	355	71.0
12	13	Aarush	M	IT	67.0	56.0	81.0	NaN	90.0	55.0	282	56.4
13	14	Leo	M	IT	98.0	-34.0	70.0	94.0	77.0	66.0	273	54.6
14	15	Maryam	F	IT	64.0	87.0	60.0	90.0	65.0	90.0	392	78.4
15	16	Saanvi	F	Comp	66.0	90.0	95.0	67.0	99.0	77.0	428	85.6
16	17	Zaranew	F	Comp	93.0	54.0	NaN	75.0	90.0	65.0	284	56.8
17	18	Inaya	F	Comp	74.0	67.0	93.0	93.0	87.0	99.0	439	87.8
18	19	Aarya	F	Comp	72.0	88.0	84.0	81.0	80.0	45.0	378	75.6
19	20	NaN	F	Comp	53.0	76.0	81.0	93.0	65.0	23.0	338	67.6

```python
In [30]:   df.shape
```

Out[30]:   (20, 12)

```python
In [31]:   df.dtypes
```

```
Out[31]:   Rollno          int64
           Name           object
           Gender         object
           Branch         object
           Attendence    float64
           Phy_marks     float64
           Che_marks     float64
           EM1_marks     float64
           PPS_marks     float64
           SME_marks     float64
           Total Marks     int64
           Percentage    float64
           dtype: object
```

```python
In [32]:   df.isna().sum()
```

```
Out[32]:   Rollno         0
           Name           2
           Gender         0
           Branch         0
           Attendence     0
           Phy_marks      1
           Che_marks      3
           EM1_marks      2
           PPS_marks      1
           SME_marks      0
           Total Marks    0
           Percentage     0
           dtype: int64
```

```python
In [33]:   cols_with_na = []
           for col in df.columns:
               if df[col].isna().any():
```

```
        cols_with_na.append(col)

cols_with_na
```

Out[33]: ['Name', 'Phy_marks', 'Che_marks', 'EM1_marks', 'PPS_marks']
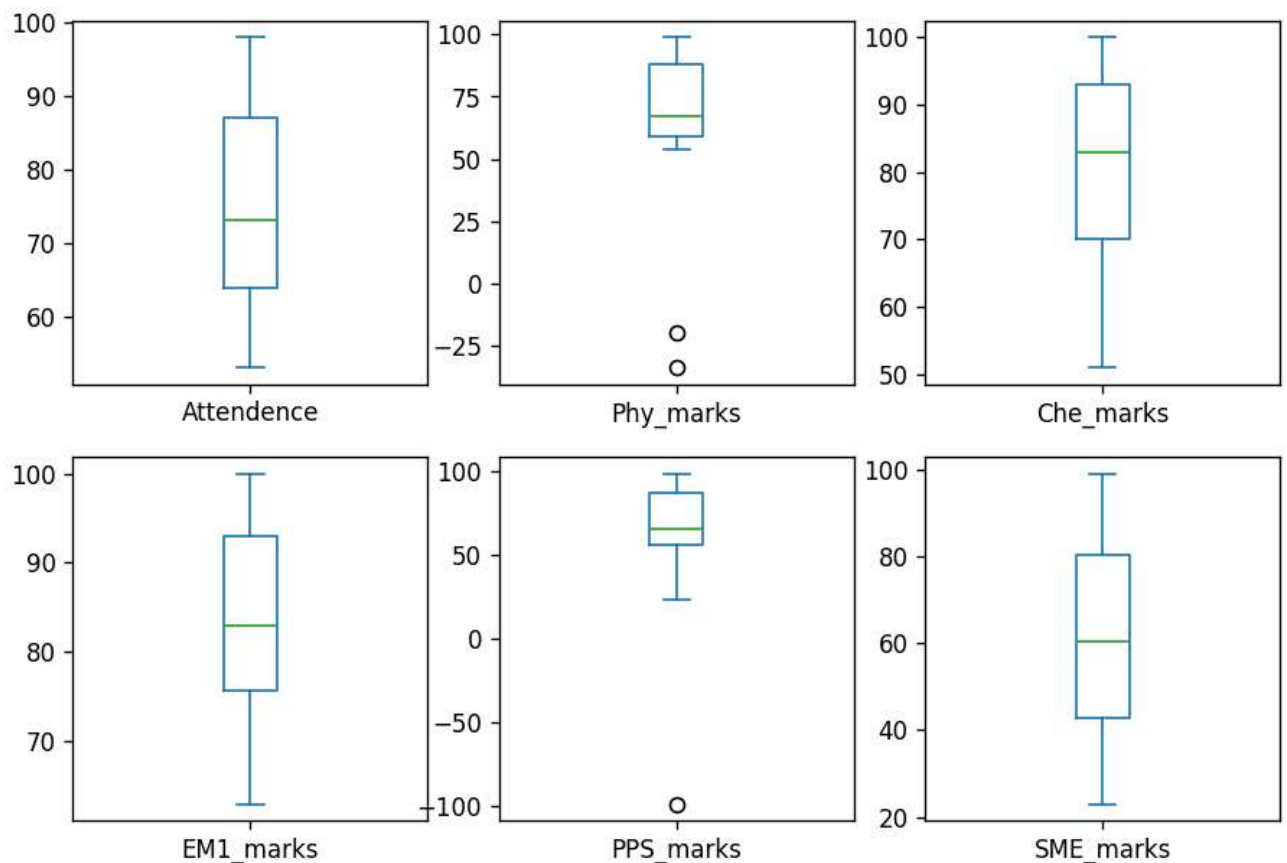
In [37]:
```
df[col] = df[col].bfill()
```

In [38]:
```
df['Total Marks']=df['Phy_marks']+df['Che_marks']+df['EM1_marks']+df['PPS_marks']+df['SME_marks']
df['Percentage']=df['Total Marks']/5

df
```

Out[38]:

	Rollno	Name	Gender	Branch	Attendence	Phy_marks	Che_marks	EM1_marks	PPS_marks	SME_marks	Total Marks	Percentage
0	1	Mohammed	M	Comp	72.0	62.0	98.0	63.0	89.0	36.0	348.0	69.6
1	2	Reyansh	M	IT	58.0	62.0	83.0	83.0	88.0	34.0	350.0	70.0
2	3	Aarav	M	IT	57.0	-20.0	100.0	NaN	56.0	36.0	NaN	NaN
3	4	Atharv	M	IT	60.0	89.0	83.0	70.0	33.0	23.0	298.0	59.6
4	5	Vivaan	M	Comp	85.0	90.0	NaN	78.0	23.0	56.0	NaN	NaN
5	6	Advik	M	ENTC	94.0	99.0	84.0	100.0	56.0	99.0	438.0	87.6
6	7	Ansh	M	ENTC	98.0	88.0	95.0	81.0	78.0	78.0	420.0	84.0
7	8	Ishaan	M	ENTC	75.0	66.0	51.0	83.0	-99.0	76.0	177.0	35.4
8	9	Dhruv	M	ENTC	63.0	NaN	NaN	97.0	56.0	55.0	NaN	NaN
9	10	Siddharth	M	ENTC	96.0	67.0	78.0	95.0	NaN	98.0	NaN	NaN
10	11	Vihaan	M	ENTC	82.0	54.0	70.0	88.0	55.0	56.0	323.0	64.6
11	12	NaN	M	IT	75.0	64.0	67.0	71.0	66.0	87.0	355.0	71.0
12	13	Aarush	M	IT	67.0	56.0	81.0	NaN	90.0	55.0	NaN	NaN
13	14	Leo	M	IT	98.0	-34.0	70.0	94.0	77.0	66.0	273.0	54.6
14	15	Maryam	F	IT	64.0	87.0	60.0	90.0	65.0	90.0	392.0	78.4
15	16	Saanvi	F	Comp	66.0	90.0	95.0	67.0	99.0	77.0	428.0	85.6
16	17	Zaranew	F	Comp	93.0	54.0	NaN	75.0	90.0	65.0	NaN	NaN
17	18	Inaya	F	Comp	74.0	67.0	93.0	93.0	87.0	99.0	439.0	87.8
18	19	Aarya	F	Comp	72.0	88.0	84.0	81.0	80.0	45.0	378.0	75.6
19	20	NaN	F	Comp	53.0	76.0	81.0	93.0	65.0	23.0	338.0	67.6

In [39]:
```python
import matplotlib.pyplot as plt
plt.rcParams["figure.figsize"] = (9, 6)
df_list = ['Attendence', 'Phy_marks', 'Che_marks', 'EM1_marks', 'PPS_marks', 'SME_marks']
fig, axes = plt.subplots(2, 3)
fig.set_dpi(120)

count=0
for r in range(2):
    for c in range(3):
        _ = df[df_list[count]].plot(kind = 'box', ax=axes[r,c])
        count+=1
```

```
In [40]:  Q1 = df['Che_marks'].quantile(0.25)
          Q3 = df['Che_marks'].quantile(0.75)
          IQR = Q3 - Q1

          Lower_limit = Q1 - 1.5 * IQR
          Upper_limit = Q3 + 1.5 * IQR

          print(f'Q1 = {Q1}, Q3 = {Q3}, IQR = {IQR}, Lower_limit = {Lower_limit}, Upper_limit = {Upper_limit}')
```

          Q1 = 70.0, Q3 = 93.0, IQR = 23.0, Lower_limit = 35.5, Upper_limit = 127.5

```
In [41]:  df[(df['Che_marks'] < Lower_limit) | (df['Che_marks'] > Upper_limit)]
```

Out[41]:

Rollno	Name	Gender	Branch	Attendence	Phy_marks	Che_marks	EM1_marks	PPS_marks	SME_marks	Total Marks	Percentage

```
In [42]:  def BinningFunction(column, cut_points, labels = None) :
              break_points=[column.min()] + cut_points + [column.max( )]
              print('Gradding According to percentage \n>60 = F \n60-70 = B \n70-80 = A\n80-100 = O')
              return pd.cut(column, bins=break_points, labels=labels, include_lowest=True)
```

```
In [43]:  cut_points=[60, 70, 80]
          labels=['F', 'B', 'A', 'O']
          df['Grade']=BinningFunction(df['Percentage'], cut_points, labels)

          df
```

          Gradding According to percentage
          >60 = F
          60-70 = B
          70-80 = A
          80-100 = O

	Rollno	Name	Gender	Branch	Attendence	Phy_marks	Che_marks	EM1_marks	PPS_marks	SME_marks	Total Marks	Percentage	Gra
0	1	Mohammed	M	Comp	72.0	62.0	98.0	63.0	89.0	36.0	348.0	69.6	
1	2	Reyansh	M	IT	58.0	62.0	83.0	83.0	88.0	34.0	350.0	70.0	
2	3	Aarav	M	IT	57.0	-20.0	100.0	NaN	56.0	36.0	NaN	NaN	N
3	4	Atharv	M	IT	60.0	89.0	83.0	70.0	33.0	23.0	298.0	59.6	
4	5	Vivaan	M	Comp	85.0	90.0	NaN	78.0	23.0	56.0	NaN	NaN	N
5	6	Advik	M	ENTC	94.0	99.0	84.0	100.0	56.0	99.0	438.0	87.6	
6	7	Ansh	M	ENTC	98.0	88.0	95.0	81.0	78.0	78.0	420.0	84.0	
7	8	Ishaan	M	ENTC	75.0	66.0	51.0	83.0	-99.0	76.0	177.0	35.4	
8	9	Dhruv	M	ENTC	63.0	NaN	NaN	97.0	56.0	55.0	NaN	NaN	N
9	10	Siddharth	M	ENTC	96.0	67.0	78.0	95.0	NaN	98.0	NaN	NaN	N
10	11	Vihaan	M	ENTC	82.0	54.0	70.0	88.0	55.0	56.0	323.0	64.6	
11	12	NaN	M	IT	75.0	64.0	67.0	71.0	66.0	87.0	355.0	71.0	
12	13	Aarush	M	IT	67.0	56.0	81.0	NaN	90.0	55.0	NaN	NaN	N
13	14	Leo	M	IT	98.0	-34.0	70.0	94.0	77.0	66.0	273.0	54.6	
14	15	Maryam	F	IT	64.0	87.0	60.0	90.0	65.0	90.0	392.0	78.4	
15	16	Saanvi	F	Comp	66.0	90.0	95.0	67.0	99.0	77.0	428.0	85.6	
16	17	Zaranew	F	Comp	93.0	54.0	NaN	75.0	90.0	65.0	NaN	NaN	N
17	18	Inaya	F	Comp	74.0	67.0	93.0	93.0	87.0	99.0	439.0	87.8	
18	19	Aarya	F	Comp	72.0	88.0	84.0	81.0	80.0	45.0	378.0	75.6	
19	20	NaN	F	Comp	53.0	76.0	81.0	93.0	65.0	23.0	338.0	67.6	

In [ ]: