PBL Exploratory Data Analysis

Import Necessary Libraries

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
```

Import the data

In [2]:
```
# df=pd.read_csv('Preprocessed Data.csv')
df_PH1=pd.read_csv('PH1.csv')
df_PH2=pd.read_csv('PH2.csv')
df_PH3=pd.read_csv('PH3.csv')
df_PH4=pd.read_csv('PH4.csv')
df_PH5=pd.read_csv('PH5.csv')
df_PH6=pd.read_csv('PH6.csv')
df_PH7=pd.read_csv('PH7.csv')
```

Understanding the data

In [4]:
```
print('The shape of phase 1 data is: {}'.format(df_PH1.shape))
print('The shape of phase 2 data is: {}'.format(df_PH2.shape))
print('The shape of phase 3 data is: {}'.format(df_PH3.shape))
print('The shape of phase 4 data is: {}'.format(df_PH4.shape))
print('The shape of phase 5 data is: {}'.format(df_PH5.shape))
print('The shape of phase 6 data is: {}'.format(df_PH6.shape))
print('The shape of phase 7 data is: {}'.format(df_PH7.shape))
```

```
The shape of phase 1 data is: (52395, 64)
The shape of phase 2 data is: (1186587, 64)
The shape of phase 3 data is: (109020, 64)
The shape of phase 4 data is: (1167633, 64)
The shape of phase 5 data is: (1962381, 64)
The shape of phase 6 data is: (1196532, 64)
The shape of phase 7 data is: (22190, 64)
```

Understanding each attribute

1. **ABRK** - AIRBRAKE POSITION
2. **ELEV_1** - ELEVATOR POSITION LEFT
3. **ELEV_2** - ELEVATOR POSITION RIGHT
4. **FLAP** - T.E. FLAP POSITION
5. **N1CO** - N1 COMPENSATION
6. **PACK** - PACK AIR CONDITIONING ALL
7. **PH** - FLIGHT PHASE FROM ACMS
8. **SAT** - STATIC AIR TEMPERATURE
9. **SPLG** - SPOILER DEPLOY GREEN
10. **SPLY** - SPOILER DEPLOY YELLOW

11. **TAI** - TAIL ANTICE ON
12. **TAT** - TOTAL AIR TEMPERATURE
13. **TMODE** - THRUST MODE
14. **WAI_1** - INNER WING DEICE
15. **WAI_2** - OUTER WING ANTICE
16. **CCPC_Mean** - CONTROL COLUMN POSITION CAPT
17. **CCPF_Mean** - CONTROL COLUMN POSITION F/O
18. **CWPC_Mean** - CONTROL WHEEL POSITION CAPT
19. **CWPF_Mean** - CONTROL WHEEL POSITION F/O
20. **MSQT_1_Mean** - SQUAT SWITCH LEFT MAIN GEAR
21. **PI_Mean** - IMPACT PRESSURE LSP
22. **PS_Mean** - STATIC PRESSURE LSP
23. **PSA_Mean** - AVERAGE STATIC PRESSURE LSP
24. **PT_Mean** - TOTAL PRESSURE LSP
25. **RUDD_Mean** - RUDDER POSITION
26. **RUDP_Mean** - RUDDER PEDAL POSITION
27. **ALT_Mean** - PRESSURE ALTITUDE LSP
28. **ALTR_Mean** - ALTITUDE RATE
29. **AOAC_Mean** - CORRECTED ANGLE OF ATTACK
30. **BAL1_Mean** - BARO CORRECT ALTITUDE LSP
31. **BAL2_Mean** - BARO CORRECT ALTITUDE LSP
32. **CAS_Mean** - COMPUTED AIRSPEED LSP
33. **DA_Mean** - DRIFT ANGLE
34. **GS_Mean** - GROUND SPEED LSP
35. **LATG_Mean** - LATERAL ACCELERATION
36. **LONG_Mean** - LONGITUDINAL ACCELERATION
37. **MACH_Mean** - MACH LSP
38. **MH_Mean** - MAGNETIC HEADING LSP
39. **N1T_Mean** - N1 TARGET LSP
40. **NSQT_Mean** - SQUAT SWITCH NOSE MAIN GEAR
41. **TAS_Mean** - TRUE AIRSPEED LSP
42. **TRKM_Mean** - TRACK ANGLE MAG LSP
43. **VIB_1_Mean**- ENGINE VIBRATION 1
44. **WD_Mean** - WIND DIRECTION TRUE
45. **WS_Mean** - WIND SPEED AVERAGE
46. **WS_Min** - WIND SPEED MINIMUM
47. **ROLL_Mean** - ROLL ANGLE LSP
48. **VRTG_Mean** - VERTICAL ACCELERATION
49. **CTAC_Mean** - CROSS TRACK ACCELERATION
50. **FPAC_Mean** - FLIGHT PATH ACCELERATION
51. **IVV_Mean** - INERTIAL VERTICAL SPEED LSP
52. **AIL_1** - AILERON POSITION LH
53. **AIL_2** - AILERON POSITION RH
54. **A_T** - THRUST AUTOMATIC ON
55. **BLV** - BLEED AIR ALL VALVES
56. **EAI** - ENGINE ANTICE ALL POSITIONS
57. **FF** - FUEL FLOW RATE - **TARGET VARIABLE**
58. **Total Fuel**- TOTAL FUEL

59. **OIT_1** - OIL TEMPERATURE 1
60. **OIT_2** - OIL TEMPERATURE 2
61. **OIT_3** - OIL TEMPERATURE 3
62. **OIT_4** - OIL TEMPERATURE 4
63. **datetime** - DATE AND TIME OF FLIGHT

In [43]:
```
df_PH1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 52395 entries, 0 to 52394
Data columns (total 64 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   ACID         52395 non-null  int64  
 1   ABRK         52395 non-null  float64
 2   ELEV_1       52395 non-null  float64
 3   ELEV_2       52395 non-null  float64
 4   FLAP         52395 non-null  int64  
 5   N1CO         52395 non-null  int64  
 6   OIT_1        52395 non-null  float64
 7   OIT_2        52395 non-null  float64
 8   OIT_3        52395 non-null  float64
 9   OIT_4        52395 non-null  float64
 10  PACK         52395 non-null  int64  
 11  PH           52395 non-null  int64  
 12  SAT          52395 non-null  float64
 13  SPLG         52395 non-null  int64  
 14  SPLY         52395 non-null  int64  
 15  TAI          52395 non-null  int64  
 16  TAT          52395 non-null  float64
 17  TMODE        52395 non-null  int64  
 18  WAI_1        52395 non-null  int64  
 19  WAI_2        52395 non-null  int64  
 20  CCPC_Mean    52395 non-null  float64
 21  CCPF_Mean    52395 non-null  float64
 22  CWPC_Mean    52395 non-null  float64
 23  CWPF_Mean    52395 non-null  float64
 24  MSQT_1_Mean  52395 non-null  float64
 25  PI_Mean      52395 non-null  float64
 26  PS_Mean      52395 non-null  float64
 27  PSA_Mean     52395 non-null  float64
 28  PT_Mean      52395 non-null  float64
 29  RUDD_Mean    52395 non-null  float64
 30  RUDP_Mean    52395 non-null  float64
 31  ALT_Mean     52395 non-null  float64
 32  ALTR_Mean    52395 non-null  int64  
 33  AOAC_Mean    52395 non-null  float64
 34  BAL1_Mean    52395 non-null  float64
 35  BAL2_Mean    52395 non-null  float64
 36  CAS_Mean     52395 non-null  float64
 37  DA_Mean      52395 non-null  float64
 38  GS_Mean      52395 non-null  float64
 39  LATG_Mean    52395 non-null  float64
 40  LONG_Mean    52395 non-null  float64
 41  MACH_Mean    52395 non-null  float64
 42  MH_Mean      52395 non-null  float64
 43  N1T_Mean     52395 non-null  float64
 44  NSQT_Mean    52395 non-null  float64
 45  TAS_Mean     52395 non-null  float64
 46  TRKM_Mean    52395 non-null  float64
 47  VIB_1_Mean   52395 non-null  float64
 48  WD_Mean      52395 non-null  float64
 49  WS_Mean      52395 non-null  float64
```

```
 50  WS_Min        52395 non-null  float64
 51  ROLL_Mean     52395 non-null  float64
 52  VRTG_Mean     52395 non-null  float64
 53  CTAC_Mean     52395 non-null  float64
 54  FPAC_Mean     52395 non-null  float64
 55  IVV_Mean      52395 non-null  float64
 56  AIL_1         52395 non-null  float64
 57  AIL_2         52395 non-null  float64
 58  A_T           52395 non-null  int64
 59  BLV           52395 non-null  int64
 60  EAI           52395 non-null  int64
 61  FF            52395 non-null  int64
 62  datetime      52395 non-null  object
 63  Total Fuel    52395 non-null  int64
dtypes: float64(46), int64(17), object(1)
memory usage: 25.6+ MB
```

In [ ]:

```
# Explain Each column
```

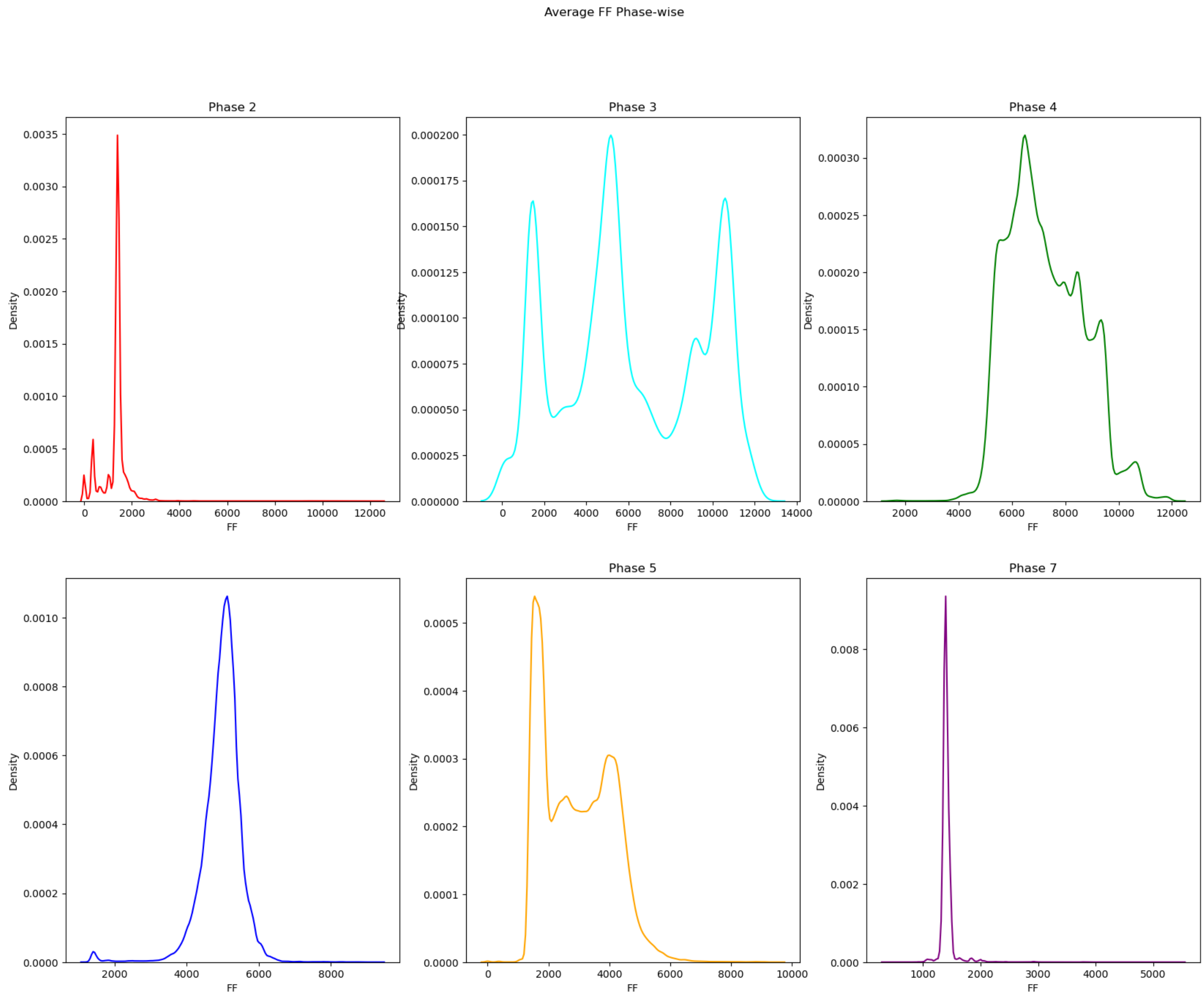
Phase wise FF rate

In [41]:

```
fig, axes = plt.subplots(2, 3, figsize=(20, 15))
sns.kdeplot(ax=axes[0,0],x="FF",data=df_PH2,color='red')
axes[0,0].set_title('Phase 2')
sns.kdeplot(ax=axes[0,1],x="FF",data=df_PH3,color='cyan')
axes[0,1].set_title('Phase 3')
sns.kdeplot(ax=axes[0,2],x="FF",data=df_PH4,color='green')
axes[0,2].set_title('Phase 4')
sns.kdeplot(ax=axes[1,0],x="FF",data=df_PH5,color='blue')
axes[1,1].set_title('Phase 5')
sns.kdeplot(ax=axes[1,1],x="FF",data=df_PH6,color='orange')
axes[1,2].set_title('Phase 6')
sns.kdeplot(ax=axes[1,2],x="FF",data=df_PH7,color='purple')
axes[1,2].set_title('Phase 7')
fig.suptitle('Average FF Phase-wise')
```

Out[41]:

```
Text(0.5, 0.98, 'Average FF Phase-wise')
```



Phase 1 (Pre-flight):

Fuel flow rate during pre-flight is relatively low, indicating that this phase has a minimal impact on fuel consumption.

Phase 2 (Taxi):

Taxi phase has a significant impact on fuel consumption with a wide range of fuel flow rates observed. Possible optimization strategies for reducing fuel consumption during taxi include minimizing the taxi time or using electric taxiing.

Phase 3 (Takeoff):

The takeoff phase has a relatively high fuel flow rate, indicating the potential for significant fuel savings through optimization. Optimization strategies may include optimizing the takeoff weight or reducing the thrust setting.

Phase 4 (Climb):

Fuel flow rate during climb is relatively stable and lower than the takeoff phase, indicating efficient engine performance during this phase. Possible optimization strategies may include optimizing the thrust setting and the climb rate.

Phase 5 (Cruise):

Cruise phase has a lower fuel flow rate compared to other phases, indicating that efficient engine performance can be maintained during this phase. Optimization strategies may include optimizing the cruising altitude and speed.

Phase 6 (Approach):

The approach phase has a moderate fuel flow rate, indicating potential for optimization to reduce fuel consumption. Possible optimization strategies may include optimizing the descent rate and approach speed.

Phase 7 (Rollout):

Rollout phase has a relatively low fuel flow rate, indicating that this phase has a minimal impact on fuel consumption. Possible optimization strategies may include optimizing the taxi route to reduce the time spent on the runway.

Analysis 1 - MACH Speed vs FF

The fuel flow rate of an airliner can be affected by various factors, and one of them is the oil temperature of the engines.

1. **Engine efficiency: The oil temperature of an aircraft engine is a critical parameter that affects its efficiency. If the oil temperature is too low, the engine may not be able to operate at its peak performance, which could result in higher fuel consumption. On the other hand, if the oil temperature is too high, it could cause the engine to overheat, leading to increased fuel consumption or even engine failure.**

2. **Fuel heating: The fuel flow rate of an airliner can also be affected by the temperature of the fuel. The colder the fuel, the higher the viscosity, which can cause fuel flow restrictions and ultimately lead to higher fuel consumption. Therefore, the warmer oil temperature of the engines could help to heat up the fuel, making it less viscous and easier to flow, resulting in a lower fuel flow rate.**

3. **Altitude: Another factor that can affect fuel flow rate is the altitude of the aircraft. During the climb phase of a flight, the aircraft is ascending to a higher altitude, which means that the air pressure and temperature are decreasing. The decrease in air pressure and temperature can cause the engine to operate less efficiently, leading to higher fuel consumption. However, the warmer oil temperature of the engines could help to offset the decrease in engine efficiency caused by the lower air pressure and temperature, resulting in a lower fuel flow rate.**

Overall, the average oil temperature of all four engines during the climb phase of a flight can have a significant impact on the fuel flow rate of an airliner. The higher oil temperature can improve engine efficiency, heat up the fuel, and offset the decrease in engine efficiency caused by the lower air pressure and temperature, resulting in a lower fuel flow rate.

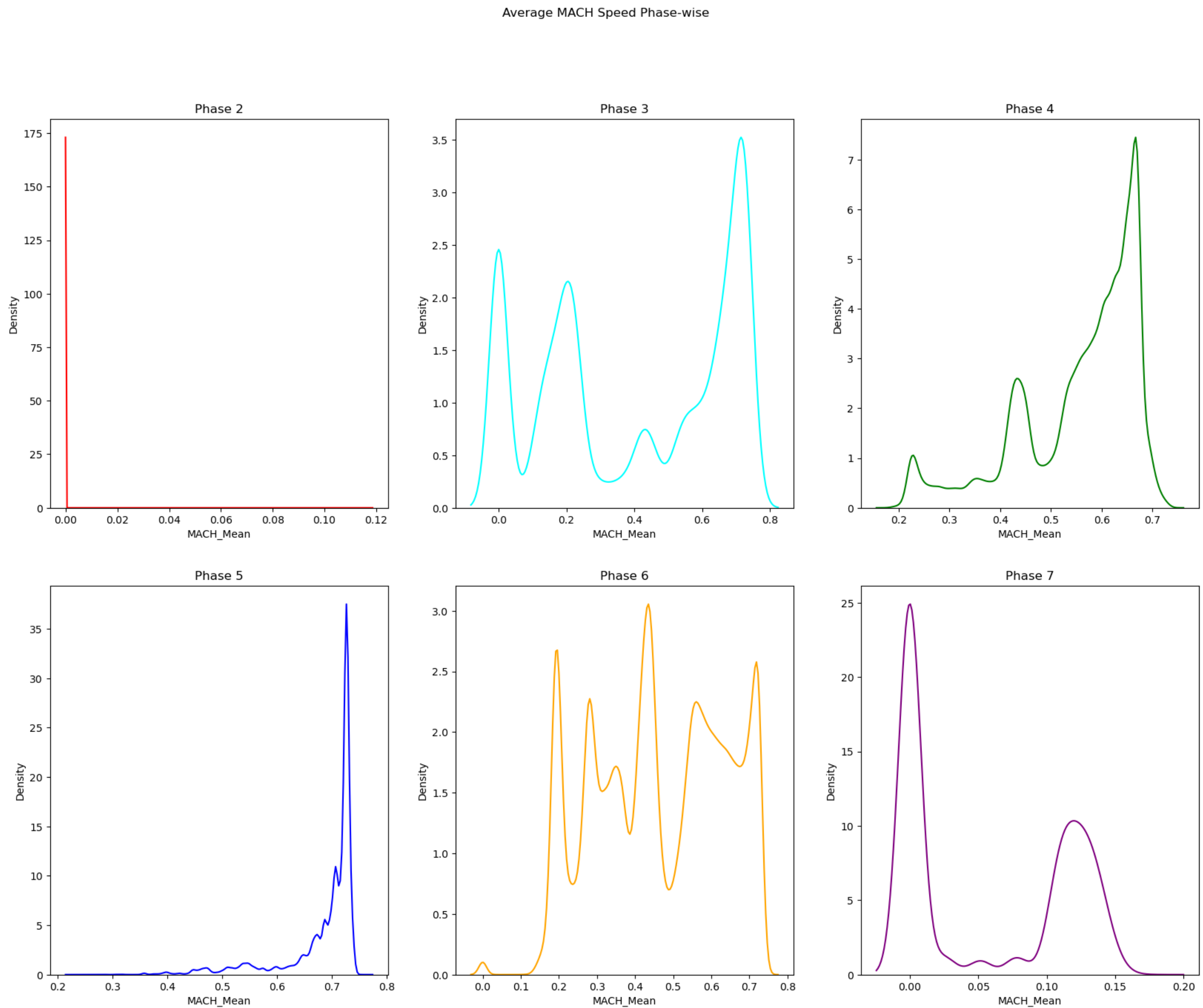
In [42]:

```
fig, axes = plt.subplots(2, 3, figsize=(20, 15))
sns.kdeplot(ax=axes[0,0],x="MACH_Mean",data=df_PH2,color='red')
axes[0,0].set_title('Phase 2')
sns.kdeplot(ax=axes[0,1],x="MACH_Mean",data=df_PH3,color='cyan')
axes[0,1].set_title('Phase 3')
sns.kdeplot(ax=axes[0,2],x="MACH_Mean",data=df_PH4,color='green')
```

```
axes[0,2].set_title('Phase 4')
sns.kdeplot(ax=axes[1,0],x="MACH_Mean",data=df_PH5,color='blue')
axes[1,0].set_title('Phase 5')
sns.kdeplot(ax=axes[1,1],x="MACH_Mean",data=df_PH6,color='orange')
axes[1,1].set_title('Phase 6')
sns.kdeplot(ax=axes[1,2],x="MACH_Mean",data=df_PH7,color='purple')
axes[1,2].set_title('Phase 7')
fig.suptitle('Average MACH Speed Phase-wise')
```

Out[42]:

```
Text(0.5, 0.98, 'Average MACH Speed Phase-wise')
```



From the above plots, we can infer the following:

Phase 1 (Pre-flight):

The MACH speed is 0, as expected, since the aircraft is not yet in motion. No inference can be drawn for optimization of fuel flow rate from this phase.

Phase 2 (Taxi):

The MACH speed is relatively low, indicating that the aircraft is moving slowly on the ground. The operator can optimize the fuel flow rate by reducing engine power during taxiing to minimize fuel consumption.

Phase 3 (Takeoff):

The MACH speed increases significantly during takeoff, indicating that the aircraft is using more fuel. To optimize fuel flow rate during takeoff, the operator should ensure that the aircraft is properly loaded, use the most efficient engine thrust setting, and minimize takeoff distance.

Phase 4 (Climb):

The MACH speed continues to increase during climb, indicating that the aircraft is using more fuel to gain altitude. To optimize fuel flow rate during climb, the operator should use the most efficient engine thrust setting and consider adjusting the Bleed Air Valve (BAV) setting based on the current MACH speed.

Phase 5 (Cruise):

The MACH speed is relatively high and stable during cruise, indicating that the aircraft is operating efficiently. To optimize fuel flow rate during cruise, the operator should maintain a steady altitude and airspeed, minimize drag, and use the most efficient engine thrust setting.

Phase 6 (Approach):

The MACH speed is variable during approach, indicating that the aircraft is adjusting its speed to prepare for landing. To optimize fuel flow rate during approach, the operator should use the most efficient engine thrust setting and consider adjusting the BAV setting based on the current MACH speed.

Phase 7 (Rollout):

The MACH speed is low during rollout, indicating that the aircraft is slowing down after landing. The operator can optimize the fuel flow rate by reducing engine power during rollout to minimize fuel consumption. Overall, to optimize fuel flow rate throughout all phases of flight, the operator should consider adjusting engine thrust and BAV settings based on the current MACH speed, maintaining a steady altitude and airspeed, minimizing drag, and ensuring proper aircraft loading.

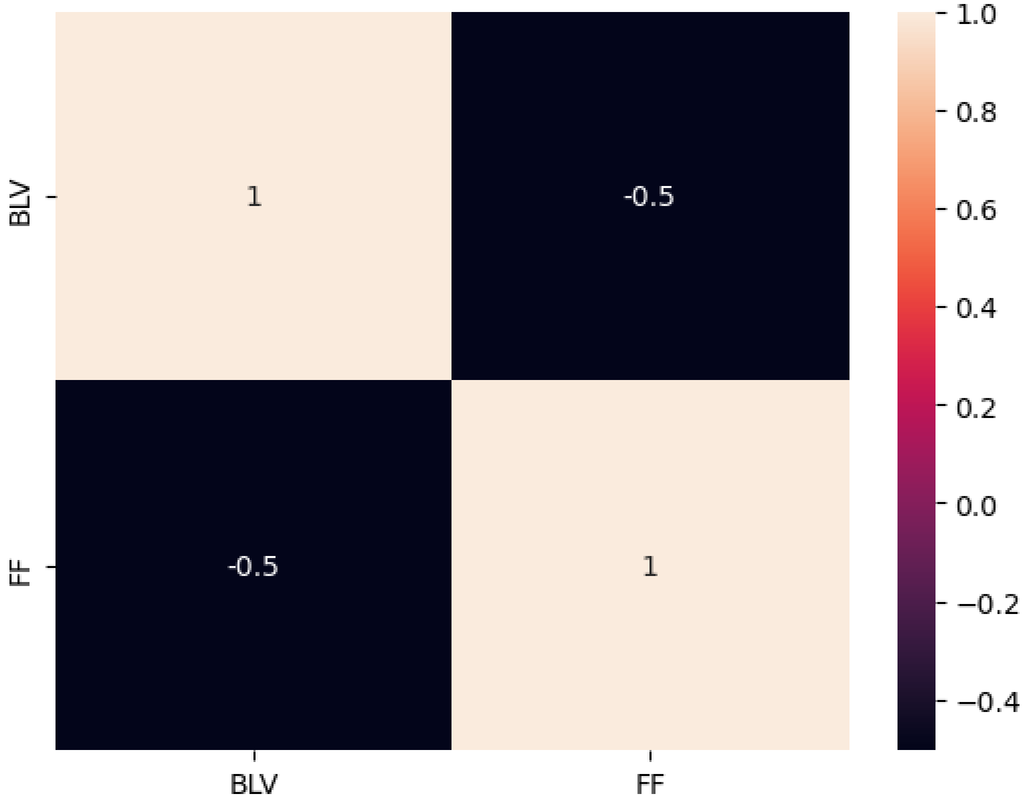
Analysis 2 - Bleed Air Valves vs FF

The Bleed Air System is an integral part of modern aircraft and serves a critical role in ensuring the safe and efficient operation of various aircraft systems. The Bleed Air System draws compressed air from the aircraft's engines and uses it to power a range of systems, including air conditioning, anti-icing, and hydraulic power. The Bleed Air Valves (BAV) are a key component of the Bleed Air System and are responsible for regulating the flow of compressed air to different systems throughout the aircraft. Proper management of the BAV is essential to ensure that aircraft systems receive the appropriate amount of compressed air, which ultimately impacts the aircraft's performance, fuel efficiency, and overall safety.

Let us assume:
H0(Null Hypothesis): BAV does not affect the Fuel flow rate of an aircraft.
H1(Alternate Hypothesis): BAV has an impact on the Fuel flow rate of an aircraft.

In [40]:
```
sns.heatmap(data=df_PH4[['BLV','FF']].corr(),annot=True,cmap='rocket')
```

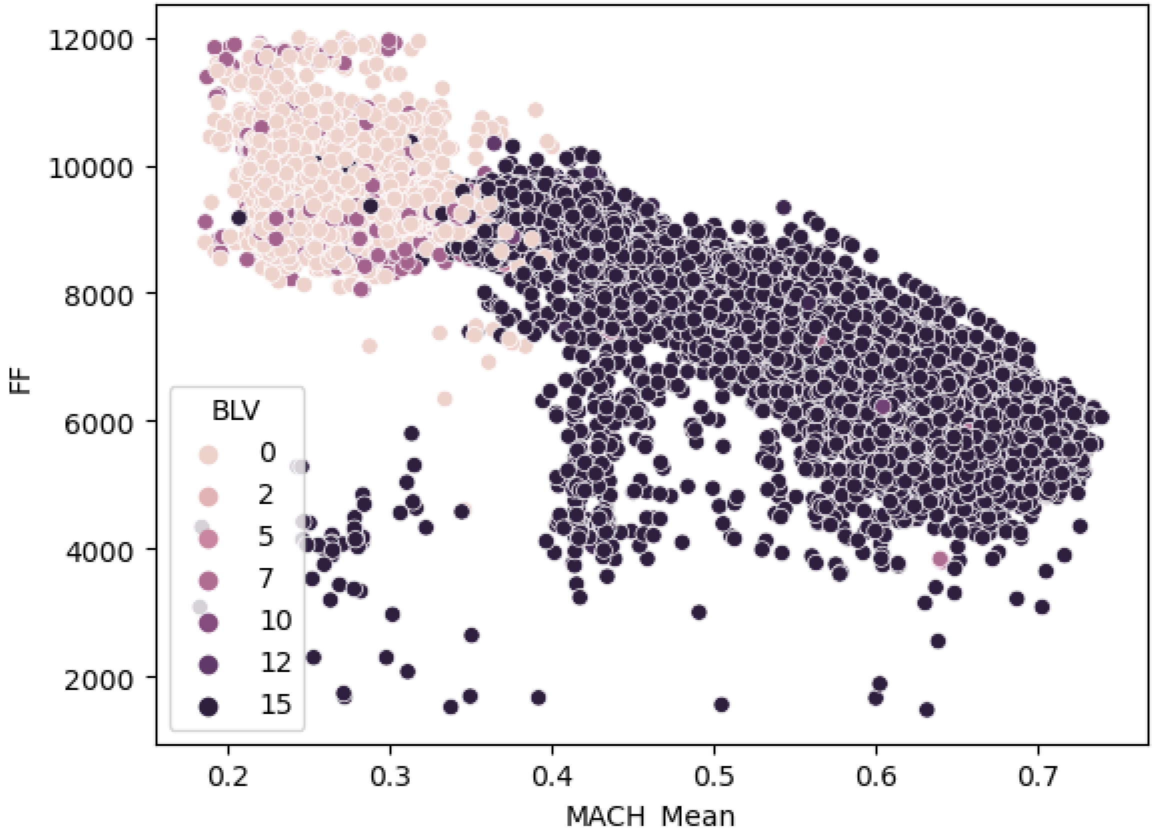
Out[40]: <AxesSubplot:>



The above plot indicates a negative correlation between BAV and FF. We can study this more during the climb phase of the flight.

In [25]:
```
df_PH4_eda=df_PH4.sample(50000,random_state=1)
sns.scatterplot(x='MACH_Mean',y='FF',hue='BLV',data=df_PH4_eda) #done
```

Out[25]: <AxesSubplot:xlabel='MACH_Mean', ylabel='FF'>



- **Analysis of the flight data recorder data reveals a correlation between the aircraft's MACH speed and BAV setting during phase 4 (climb).**
- **As MACH speed increases from 0.2 to 0.8, the BAV setting also increases from 0 to 15. This relationship can be studied further to optimize the fuel flow rate during climb.**
- **Optimizing the fuel flow rate during climb involves adjusting the BAV setting based on the current MACH speed, evaluating the BAV setting, and considering altitude and engine power settings.**
- **The decrease in air pressure and temperature during climb results in lower engine efficiency and higher fuel consumption. The use of bleed air can help to offset the decrease in engine efficiency caused by the lower air pressure and temperature, resulting in a lower fuel flow rate.**
- **Bleed air usage can also help to cool the engine and reduce the impact of the lower air pressure and temperature on engine efficiency.**
- **An important point to note is that the above plots also indicates towards our alternate hypothesis. Thus, we can reject the null hypothesis.**
- **Overall, the relationship between MACH speed, BAV setting, and fuel flow rate is a complex one, with multiple factors contributing to the overall efficiency of the aircraft during climb.**

Analysis 3 - Thrust Mode vs FF during Climb Phase

Thrust mode is an important aspect of modern aircraft design and operation. It refers to the method used by aircraft engines to produce the necessary thrust to propel the aircraft forward. The mode of thrust used by an aircraft engine can have a significant impact on the aircraft's performance,

efficiency, and fuel consumption. Modern aircraft engines use several different thrust modes, including continuous, climb, cruise, and takeoff thrust. These different modes are selected based on the aircraft's current phase of flight and the required performance parameters, such as altitude, airspeed, and fuel consumption.

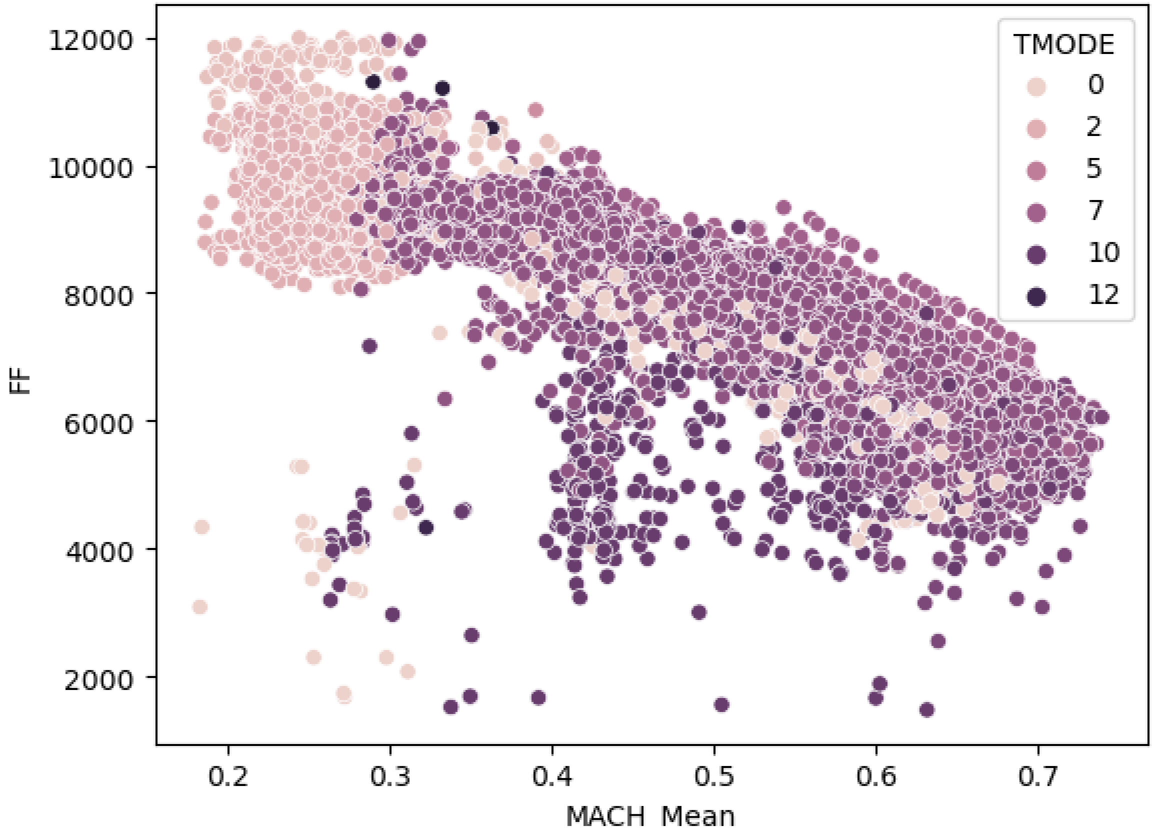
In general, FDR data records Thrust mode in the following 12 categories:

1. **Idle: This is the minimum thrust setting for the engine and is used when the aircraft is stationary or taxiing on the ground.**
2. **Approach: This is a low thrust setting used during the approach and landing phases of flight.**
3. **Climb: This is a higher thrust setting used during the climb phase of flight to gain altitude.**
4. **Takeoff: This is the maximum thrust setting used during the takeoff phase to achieve the necessary acceleration for lift-off.**
5. **Flex: This is a takeoff thrust setting that allows the engine to operate below its maximum power while still providing sufficient takeoff performance.**
6. **Derate: This is a takeoff thrust setting that reduces the engine's maximum power output to save fuel and increase engine life.**
7. **Go-Around: This is a high thrust setting used when a landing approach is aborted and the aircraft must quickly gain altitude.**
8. **Maximum Continuous Thrust (MCT): This is the highest thrust setting that the engine can sustain indefinitely.**
9. **Maximum Climb Thrust (MCT): This is the maximum thrust setting that can be used during the climb phase of flight.**
10. **Maximum Takeoff Thrust (MTO): This is the maximum thrust setting that can be used during takeoff.**
11. **Maximum Cruise Thrust (MCT): This is the maximum thrust setting that can be used during cruise flight.**
12. **Reverser: This is a thrust mode used to reverse the direction of the engine's thrust, which is used to help slow the aircraft down after landing.**

We will continue our analysis based on these thrust modes.

In [26]:
```
sns.scatterplot(x='MACH_Mean',y='FF',hue='TMODE',data=df_PH4_eda) #done
```

Out[26]:
```
<AxesSubplot:xlabel='MACH_Mean', ylabel='FF'>
```



- **The analysis of flight data recorder data indicates a relationship between MACH speed, thrust mode, and fuel flow rate during the climb phase of flight.**
- **As MACH speed increases from 0.2 to 0.8, thrust mode also increases from 1 to 11, indicating a need for more engine power to maintain altitude.**
- **Despite the increased engine power, the fuel flow rate decreases as thrust mode and MACH speed increase, indicating more efficient engine operation at higher thrust settings.**
- **This relationship can be studied further to optimize fuel flow rate during climb by adjusting the thrust mode based on the current MACH speed, altitude, and engine power settings.**
- **Derate and Flex thrust modes can be used during takeoff to reduce engine power output and save fuel without sacrificing takeoff performance.**
- **Approach thrust mode can be used during landing to reduce fuel consumption during the descent and approach phases. Maximum Continuous Thrust (MCT) and Maximum Climb Thrust (MCT) can be used during the climb phase to achieve optimal engine efficiency while still maintaining a safe ascent rate.**
- **Maximum Cruise Thrust (MCT) can be used during cruise flight to balance fuel efficiency and speed. Reverser thrust mode can be used during landing to help slow the aircraft down and reduce the need for excessive braking, which can also save fuel.**

End of Notebook