In [1]: ```python
import pandas as pd
```

In [2]: ```python
path = "data/IPL Player Stats/"


d_16 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2016.csv")
d_17=  pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2017.csv")
d_18 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2018.csv")
d_19 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2019.csv")
d_20 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2020.csv")
d_21 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2021.csv")
d_22 = pd.read_csv(path + "Batting Stats/BATTING STATS - IPL_2022.csv")
```

In [3]: ```python
df = pd.concat([d_16,d_17,d_18,d_19,d_20,d_21,d_22], ignore_index=True)
```

In [4]: ```python
df.shape
dfs = [d_16,d_17,d_18,d_19,d_20,d_21,d_22]
```

In [5]: ```python
df.head()
```

Out[5]:

|   | POS | Player | Mat | Inns | NO | Runs | HS | Avg | BF | SR | 100 | 50 | 4s | 6s |
|---|-----|--------|-----|------|----|----|----|------|-----|--------|-----|----|----|----|
| 0 | 1 | Virat Kohli | 16.0 | 16 | 4.0 | 973 | 113 | 81.08 | 640 | 152.03 | 4 | 7 | 83 | 38 |
| 1 | 2 | David Warner | 17.0 | 17 | 3.0 | 848 | 93* | 60.57 | 560 | 151.42 | 0 | 9 | 88 | 31 |
| 2 | 3 | AB de Villiers | 16.0 | 16 | 3.0 | 687 | 129* | 52.84 | 407 | 168.79 | 1 | 6 | 57 | 37 |
| 3 | 4 | Gautam Gambhir | 15.0 | 15 | 2.0 | 501 | 90* | 38.53 | 411 | 121.89 | 0 | 5 | 54 | 6 |
| 4 | 5 | Shikhar Dhawan | 17.0 | 17 | 4.0 | 501 | 82* | 38.53 | 429 | 116.78 | 0 | 4 | 51 | 8 |

In [6]: ```python
df.head()
```

Out[6]:

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	1	Virat Kohli	16.0	16	4.0	973	113	81.08	640	152.03	4	7	83	38
1	2	David Warner	17.0	17	3.0	848	93*	60.57	560	151.42	0	9	88	31
2	3	AB de Villiers	16.0	16	3.0	687	129*	52.84	407	168.79	1	6	57	37
3	4	Gautam Gambhir	15.0	15	2.0	501	90*	38.53	411	121.89	0	5	54	6
4	5	Shikhar Dhawan	17.0	17	4.0	501	82*	38.53	429	116.78	0	4	51	8

In [7]: ```python
df["Avg"] = df["Avg"].replace("-", '-1')
df['Avg'] = df.Avg.astype("float64")
```

In [8]: ```python
all_batsmen =  df.groupby(by="Player").agg({"Inns": 'sum', 'Runs':'sum', 'Avg':
                              "100": 'sum', '4s': 'sum', '6s':"sum"}).reset_in
```

In [9]: ```python
all_batsmen.shape
```

Out[9]: ```
(338, 9)
```

```
In [10]:   # writing into 1 file
           all_batsmen.to_csv("data/IPL Player Stats/Batting Stats/all_batsmen_data_2016_2
```

```
In [ ]:
```