

In [25]:

```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv("kobe.csv")

# Show an augmented dataframe
pd.set_option('display.max_columns', 30)

print(df)
```

	action_type	combined_shot_type	game_event_id	game_id \
0		Jump Shot	Jump Shot	10 200000012
1		Jump Shot	Jump Shot	12 200000012
2		Jump Shot	Jump Shot	25 200000012
3		Jump Shot	Jump Shot	43 200000012
4	Driving Dunk Shot	Dunk	Dunk	155 200000012
5		Jump Shot	Jump Shot	244 200000012
6	Layup Shot	Layup	Layup	251 200000012
7		Jump Shot	Jump Shot	254 200000012
8		Jump Shot	Jump Shot	265 200000012
9	Running Jump Shot	Jump Shot	Jump Shot	294 200000012
10		Jump Shot	Jump Shot	309 200000012
11		Jump Shot	Jump Shot	4 200000010
12	Running Jump Shot	Jump Shot	Jump Shot	27 200000010
13		Jump Shot	Jump Shot	387 200000010
14		Jump Shot	Jump Shot	86 200000010
15		Jump Shot	Jump Shot	86 200000010
16	Driving Layup Shot	Layup	Layup	6 200000010
17		Jump Shot	Jump Shot	138 200000010
18		Jump Shot	Jump Shot	200 200000010
19	Driving Layup Shot	Layup	Layup	240 200000010
20		Jump Shot	Jump Shot	255 200000010
21		Jump Shot	Jump Shot	265 200000010
22	Running Jump Shot	Jump Shot	Jump Shot	274 200000010
23	Running Jump Shot	Jump Shot	Jump Shot	299 200000010
24	Running Jump Shot	Jump Shot	Jump Shot	307 200000010
25	Layup Shot	Layup	Layup	332 200000010
26		Jump Shot	Jump Shot	389 200000010
27		Jump Shot	Jump Shot	369 200000010
28		Jump Shot	Jump Shot	400 200000010
29		Jump Shot	Jump Shot	400 200000010
...	...	...	...	...
30667		Jump Shot	Jump Shot	368 499000008
30668		Jump Shot	Jump Shot	386 499000007
30669		Jump Shot	Jump Shot	425 499000007
30670	Running Jump Shot	Jump Shot	Jump Shot	15 499000008
30671	Driving Layup Shot	Layup	Layup	25 499000008
30672		Jump Shot	Jump Shot	29 499000008
30673		Jump Shot	Jump Shot	36 499000008
30674		Jump Shot	Jump Shot	81 499000008
30675		Jump Shot	Jump Shot	84 499000008
30676	Running Jump Shot	Jump Shot	Jump Shot	121 499000008
30677		Jump Shot	Jump Shot	101 499000008
30678	Driving Layup Shot	Layup	Layup	189 499000008
30679		Jump Shot	Jump Shot	212 499000008
30680	Tip Shot	Tip Shot	Tip Shot	213 499000008
30681		Jump Shot	Jump Shot	218 499000008
30682	Running Jump Shot	Jump Shot	Jump Shot	228 499000008
30683		Jump Shot	Jump Shot	228 499000008
30684		Jump Shot	Jump Shot	231 499000008
30685		Jump Shot	Jump Shot	249 499000008
30686	Running Jump Shot	Jump Shot	Jump Shot	269 499000008
30687		Jump Shot	Jump Shot	284 499000008
30688		Jump Shot	Jump Shot	308 499000008
30689		Jump Shot	Jump Shot	326 499000008
30690		Jump Shot	Jump Shot	331 499000008
30691	Driving Layup Shot	Layup	Layup	382 499000008
30692		Jump Shot	Jump Shot	387 499000008
30693		Tip Shot	Tip Shot	398 499000008
30694	Running Jump Shot	Jump Shot	Jump Shot	459 499000008
30695		Jump Shot	Jump Shot	446 499000008
30696		Jump Shot	Jump Shot	471 499000008

```
lat loc_x loc_y lon minutes_remaining period playoffs \
0 33.9723 167 72 -118.3028 10 1 0
1 34.0443 -157 0 -118.4208 10 1 0
2 33.9993 -101 135 -118.3708 7 1 0
3 33.8693 138 175 -118.1318 6 1 0
4 34.0443 0 0 -118.2698 10 1 0
5 34.0553 -145 -11 -118.4148 9 3 0
6 34.0443 0 0 -118.2698 8 3 0
7 34.0163 1 28 -118.2688 8 3 0
8 33.9363 -65 108 -118.3348 6 3 0
9 33.9193 -33 125 -118.3828 10 1 0
10 33.8863 -94 238 -118.3638 1 3 0
11 33.9173 121 127 -118.1488 11 1 0
12 33.9243 -67 119 -118.3368 7 1 0
13 34.0403 -94 4 -118.3638 2 1 0
14 33.9973 -23 47 -118.2928 0 1 0
15 33.8523 62 192 -118.2078 0 1 0
16 34.0443 0 0 -118.2698 0 1 0
17 33.8183 -117 226 -118.3888 8 3 0
18 33.9473 -132 97 -118.4018 11 3 0
19 34.0443 0 0 -118.2698 10 3 0
20 33.9003 3 144 -118.2698 10 1 0
21 33.9173 134 127 -118.1358 9 3 0
22 33.9343 -16 119 -118.2858 10 1 0
23 33.8943 -109 150 -118.3788 5 3 0
24 33.8813 -46 63 -118.3158 5 3 0
25 34.0443 0 0 -118.2698 6 1 0
26 33.8483 -58 196 -118.3278 2 3 0
27 33.8583 -183 108 -118.4528 0 3 0
28 33.8713 85 173 -118.1848 8 4 0
29 33.9573 3 87 -118.2668 6 4 0
...
```

```
season seconds_remaining shot_distance shot_made_flag \
0 2000-01 27 18 NaN
1 2000-01 22 15 0.0
2 2000-01 45 16 1.0
3 2000-01 52 22 0.0
4 2000-01 19 0 1.0
5 2000-01 32 14 0.0
6 2000-01 52 0 1.0
7 2000-01 5 2 NaN
8 2000-01 12 12 1.0
9 2000-01 36 12 0.0
10 2000-01 56 25 0.0
11 2000-01 0 17 1.0
12 2000-01 9 12 1.0
13 2000-01 1 9 0.0
14 2000-01 16 5 0.0
15 2000-01 48 20 0.0
16 2000-01 1 1 NaN
17 2000-01 50 25 1.0
18 2000-01 29 16 0.0
19 2000-01 46 0 NaN
20 2000-01 8 14 0.0
21 2000-01 8 13 1.0
22 2000-01 57 11 1.0
23 2000-01 47 18 1.0
24 2000-01 11 7 0.0
25 2000-01 36 0 0.0
26 2000-01 4 20 0.0
27 2000-01 30 26 0.0
28 2000-01 19 19 0.0
29 2000-01 1 8 0.0
...
```

```
shot_type shot_zone_area shot_zone_basic \
0 2PT Field Goal Right Side(R) Mid-Range
1 2PT Field Goal Left Side(L) Mid-Range
2 2PT Field Goal Left Side Center(LC) Mid-Range
3 2PT Field Goal Right Side Center(RC) Mid-Range
4 2PT Field Goal Center(C) Restricted Area
5 2PT Field Goal Left Side(L) Mid-Range
6 2PT Field Goal Center(C) Restricted Area
7 2PT Field Goal Center(C) Restricted Area
8 2PT Field Goal Left Side(L) In The Paint (Non-RA)
9 2PT Field Goal Center(C) In The Paint (Non-RA)
10 3PT Field Goal Left Side Center(LC) Above the Break 3
11 2PT Field Goal Right Side Center(RC) Mid-Range
12 2PT Field Goal Left Side(L) In The Paint (Non-RA)
13 2PT Field Goal Left Side(L) Mid-Range
14 2PT Field Goal Center(C) In The Paint (Non-RA)
15 2PT Field Goal Center(C) Mid-Range
16 2PT Field Goal Center(C) Restricted Area
17 3PT Field Goal Left Side Center(LC) Above the Break 3
18 2PT Field Goal Left Side Center(LC) Mid-Range
19 2PT Field Goal Center(C) Restricted Area
20 2PT Field Goal Center(C) Restricted Area
21 2PT Field Goal Right Side Center(RC) Mid-Range
22 2PT Field Goal Center(C) In The Paint (Non-RA)
23 2PT Field Goal Left Side Center(LC) Mid-Range
24 2PT Field Goal Center(C) In The Paint (Non-RA)
25 2PT Field Goal Center(C) Restricted Area
26 2PT Field Goal Center(C) Mid-Range
27 2PT Field Goal Left Side Center(LC) Above the Break 3
28 2PT Field Goal Right Side Center(RC) Mid-Range
29 2PT Field Goal Center(C) In The Paint (Non-RA)
...
```

```
30667 3PT Field Goal Center(C) Above the Break 3
30668 2PT Field Goal Center(C) Mid-Range
30669 2PT Field Goal Right Side(R) Mid-Range
30670 2PT Field Goal Center(C) In The Paint (Non-RA)
30671 2PT Field Goal Center(C) Restricted Area
30672 2PT Field Goal Right Side(R) Mid-Range
30673 2PT Field Goal Right Side Center(RC) Above the Break 3
30674 2PT Field Goal Left Side Center(LC) Above the Break 3
30675 2PT Field Goal Left Side Center(LC) Above the Break 3
30676 2PT Field Goal Left Side(L) Mid-Range
30677 2PT Field Goal Left Side Center(LC) Mid-Range
30678 2PT Field Goal Center(C) Restricted Area
30679 2PT Field Goal Center(C) Restricted Area
30680 2PT Field Goal Center(C) Restricted Area
30681 3PT Field Goal Center(C) Above the Break 3
30682 2PT Field Goal Left Side(L) In The Paint (Non-RA)
30683 2PT Field Goal Center(C) Mid-Range
30684 2PT Field Goal Left Side(L) Mid-Range
30685 2PT Field Goal Center(C) Above the Break 3
30686 2PT Field Goal Center(C) In The Paint (Non-RA)
30687 2PT Field Goal Center(C) In The Paint (Non-RA)
30688 2PT Field Goal Left Side(L) Mid-Range
30689 3PT Field Goal Back Court(BC) Backcourt
30690 2PT Field Goal Left Side(L) Mid-Range
30691 2PT Field Goal Center(C) Restricted Area
30692 2PT Field Goal Center(C) In The Paint (Non-RA)
30693 2PT Field Goal Center(C) Restricted Area
30694 2PT Field Goal Left Side Center(LC) Mid-Range
30695 3PT Field Goal Center(C) Above the Break 3
30696 2PT Field Goal Center(C) In The Paint (Non-RA)
...
```

```
shot_zone_range team_id team_name game_date \
0 16-24 ft. 1610612747 Los Angeles Lakers 2000-10-31
1 8-16 ft. 1610612747 Los Angeles Lakers 2000-10-31
2 16-24 ft. 1610612747 Los Angeles Lakers 2000-10-31
3 16-24 ft. 1610612747 Los Angeles Lakers 2000-10-31
4 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-10-31
5 8-16 ft. 1610612747 Los Angeles Lakers 2000-10-31
6 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-10-31
7 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-10-31
8 8-16 ft. 1610612747 Los Angeles Lakers 2000-10-31
9 8-16 ft. 1610612747 Los Angeles Lakers 2000-10-31
10 24+ ft. 1610612747 Los Angeles Lakers 2000-10-31
11 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
12 8-16 ft. 1610612747 Los Angeles Lakers 2000-11-01
13 8-16 ft. 1610612747 Los Angeles Lakers 2000-11-01
14 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-11-01
15 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
16 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-11-01
17 24+ ft. 1610612747 Los Angeles Lakers 2000-11-01
18 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
19 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-11-01
20 8-16 ft. 1610612747 Los Angeles Lakers 2000-11-01
21 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
22 8-16 ft. 1610612747 Los Angeles Lakers 2000-11-01
23 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
24 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-11-01
25 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-11-01
26 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
27 24+ ft. 1610612747 Los Angeles Lakers 2000-11-01
28 16-24 ft. 1610612747 Los Angeles Lakers 2000-11-01
29 8-16 ft. 1610612747 Los Angeles Lakers 2000-11-01
...
```

```
30667 24+ ft. 1610612747 Los Angeles Lakers 2000-06-16
30668 16-24 ft. 1610612747 Los Angeles Lakers 2000-06-16
30669 16-24 ft. 1610612747 Los Angeles Lakers 2000-06-16
30670 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-16
30671 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-16
30672 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30673 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30674 24+ ft. 1610612747 Los Angeles Lakers 2000-06-19
30675 24+ ft. 1610612747 Los Angeles Lakers 2000-06-19
30676 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30677 16-24 ft. 1610612747 Los Angeles Lakers 2000-06-19
30678 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30679 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30680 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30681 24+ ft. 1610612747 Los Angeles Lakers 2000-06-19
30682 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30683 16-24 ft. 1610612747 Los Angeles Lakers 2000-06-19
30684 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30685 24+ ft. 1610612747 Los Angeles Lakers 2000-06-19
30686 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30687 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30688 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30689 Back Court Shot 1610612747 Los Angeles Lakers 2000-06-19
30690 8-16 ft. 1610612747 Los Angeles Lakers 2000-06-19
30691 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30692 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30693 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
30694 16-24 ft. 1610612747 Los Angeles Lakers 2000-06-19
30695 24+ ft. 1610612747 Los Angeles Lakers 2000-06-19
30696 Less Than 8 ft. 1610612747 Los Angeles Lakers 2000-06-19
...
```

```
matchup opponent shot_id
0 LAL @ POR POR 1
1 LAL @ POR POR 2
2 LAL @ POR POR 3
3 LAL @ POR POR 4
4 LAL @ POR POR 5
5 LAL @ POR POR 6
6 LAL @ POR POR 7
7 LAL @ POR POR 8
8 LAL @ POR POR 9
9 LAL @ POR POR 10
10 LAL @ POR POR 11
11 LAL vs. UTA UTA 12
12 LAL vs. UTA UTA 13
13 LAL vs. UTA UTA 14
14 LAL vs. UTA UTA 15
15 LAL vs. UTA UTA 16
16 LAL vs. UTA UTA 17
17 LAL vs. UTA UTA 18
18 LAL vs. UTA UTA 19
19 LAL vs. UTA UTA 20
20 LAL vs. UTA UTA 21
21 LAL vs. UTA UTA 22
22 LAL vs. UTA UTA 23
23 LAL vs. UTA UTA 24
24 LAL vs. UTA UTA 25
25 LAL vs. UTA UTA 26
26 LAL vs. UTA UTA 27
27 LAL vs. UTA UTA 28
28 LAL vs. UTA UTA 29
29 LAL vs. UTA UTA 30
...
```

```
30667 LAL @ IND IND 30668
30668 LAL @ IND IND 30669
30669 LAL @ IND IND 30670
30670 LAL vs. IND IND 30671
30671 LAL vs. IND IND 30672
30672 LAL vs. IND IND 30673
30673 LAL vs. IND IND 30674
30674 LAL vs. IND IND 30675
30675 LAL vs. IND IND 30676
30676 LAL vs. IND IND 30677
30677 LAL vs. IND IND 30678
30678 LAL vs. IND IND 30679
30679 LAL vs. IND IND 30680
30680 LAL vs. IND IND 30681
30681 LAL vs. IND IND 30682
30682 LAL vs. IND IND 30683
30683 LAL vs. IND IND 30684
30684 LAL vs. IND IND 30685
30685 LAL vs. IND IND 30686
30686 LAL vs. IND IND 30687
30687 LAL vs. IND IND 30688
30688 LAL vs. IND IND 30689
30689 LAL vs. IND IND 30690
30690 LAL vs. IND IND 30691
30691 LAL vs. IND IND 30692
30692 LAL vs. IND IND 30693
30693 LAL vs. IND IND 30694
30694 LAL vs. IND IND 30695
30695 LAL vs. IND IND 30696
30696 LAL vs. IND IND 30697
Check rows x 25 columns
```

In [26]:

```
# Check if there are N/A objects
print(df.info())

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30697 entries, 0 to 30696
Data columns (total 25 columns):
action_type      30697 non-null object
combined_shot_type 30697 non-null object
game_event_id    30697 non-null int64
game_id          30697 non-null float64
lat              30697 non-null float64
loc_x            30697 non-null int64
loc_y            30697 non-null int64
lon              30697 non-null float64
minutes_remaining 30697 non-null int64
period           30697 non-null int64
playoffs         30697 non-null int64
season           30697 non-null object
seconds_remaining 30697 non-null int64
shot_distance     30697 non-null float64
shot_made_flag    25697 non-null float64
shot_type         30697 non-null object
shot_zone_area    30697 non-null object
shot_zone_basic   30697 non-null object
shot_zone_range   30697 non-null object
team_id           30697 non-null int64
team_name         30697 non-null object
game_date         30697 non-null object
matchup           30697 non-null object
opponent          30697 non-null object
shot_id           30697 non-null int64
dtypes: float64(3), int64(11), object(11)
memory usage: 4.6+ MB
None
```

In [27]:

```
# We have some N/A's in the shot_made_flag column. We can remove all rows with missing values using the dropna() function.

df = df.dropna(how='any') #to drop if any value in the row has a N/A

# Show 'shot_made_flag' column in order to see that missing values were removed

print(df[['shot_made_flag']])
```

1	0.0
2	1.0
3	0.0
4	1.0
5	0.0
6	1.0
7	0.0
8	1.0
9	0.0
10	0.0
11	1.0
12	1.0
13	0.0
14	0.0
15	0.0
16	1.0
17	0.0
18	0.0
19	0.0
20	0.0
21	0.0
22	1.0
23	1.0
24	1.0
25	0.0
26	0.0
27	0.0
28	0.0
29	0.0
30	1.0
31	0.0
32	1.0
33	0.0
34	0.0
35	0.0
36	0.0
37	0.0
38	0.0
39	1.0
30661	1.0
30662	0.0
30663	0.0
30664	0.0
30665	0.0
30666	0.0
30667	0.0
30668	0.0
30669	0.0
30670	0.0
30671	0.0
30672	1.0
30673	0.0
30674	1.0
30675	0.0
30676	1.0
30677	0.0
30678	0.0
30679	0.0
30680	0.0
30681	1.0
30682	0.0
30683	0.0
30684	0.0
30685	0.0
30686	0.0
30687	0.0
30688	0.0
30689	0.0
30690	0.0
30691	0.0
30692	0.0
30693	0.0
30694	1.0
30695	0.0
30696	0.0

Name: shot\_made\_flag, Length: 25697, dtype: float64

In [28]:

```
# Scatter plot

# Use ggplot style
plt.style.use('ggplot')

plt.figure(1)

ax = plt.subplot(211)

# Break the dataframe into two cases
df1 = df[df['combined_shot_type'].str.contains("Jump Shot")]

# Add a specific color to Jump Shot categories

ax.scatter(df1['loc_x'], df1['loc_y'], c="darkslategrey", alpha=0.3, label="Jump Shot")

# Dataframe without Jump Shot
df2 = df[df['combined_shot_type'].str.contains("Jump Shot")==False]

# Group by type of shot
groups = df2.groupby('combined_shot_type')

# Build a loop to compute the groups
for name, group in groups:
    ax.plot(group['loc_x'], group['loc_y'], marker='o', linestyle='-', ms=5, label=name)

ax.legend(loc='upper right')

ax.set_title('Shot type')

ax.set_xlabel('Loc_x')
ax.set_ylabel('Loc_y')

ax2 = plt.subplot(313)

df['shot_zone_range'].value_counts().plot('bar', color='orange', rot=0)

ax2.set_ylabel('frequency')
ax2.set_title('Shot Range')

plt.savefig('ScatterClassPlot.png')

plt.rcParams['figure.figsize'] = [10, 10]

plt.show()
```

