```
import pandas as pd
```

```
import pandas as pd
from google.colab import drive
drive.mount('/content/drive')
```

```
Drive already mounted at /content/drive; to attempt to forcibly remount, call
```

1.Read the data set and replace dashes with 0 to make sure you can perform arithmetic operations on the data. (

```
file = '/content/drive/My Drive/PGML/Laliga.csv'
df = pd.read_csv(file,skiprows=[0])
df.replace('-',0,inplace=True)
df
```

	Pos	Team	Seasons	Points	GamesPlayed	GamesWon	GamesDrawn	GamesLo
0	1	Real Madrid	86	4385	2762	1647	552	5
1	2	Barcelona	86	4262	2762	1581	573	6
2	3	Atletico Madrid	80	3442	2614	1241	598	7
3	4	Valencia	82	3386	2664	1187	616	8
4	5	Athletic Bilbao	86	3368	2762	1209	633	9
5	6	Sevilla	73	2819	2408	990	531	8
6	7	Espanyol	82	2792	2626	948	608	10
7	8	Real Sociedad	70	2573	2302	864	577	8
8	9	Zaragoza	58	2109	1986	698	522	7
9	10	Real Betis	51	1884	1728	606	440	6
10	11	Deportivo La Coruna	45	1814	1530	563	392	5
11	12	Celta Vigo	51	1789	1698	586	389	7
12	13	Valladolid	42	1471	1466	463	384	6
13	14	Racing Santander	44	1416	1428	453	336	6
14	15	Sporting Gijon	43	1389	1458	471	358	6
15	16	Osasuna	37	1351	1318	426	327	5
16	17	Malaga	36	1314	1255	390	330	5
17	18	Oviedo	38	1174	1192	408	292	4
18	19	Mallorca	27	1148	988	333	256	3
19	20	Las Palmas	33	1020	1096	367	242	4
20	21	Villarreal	17	970	646	266	172	2
21	22	Granada	23	667	742	218	175	3

22	23	Rayo Vallecano	17	662	652	189	148	3
23	24	Elche	21	606	678	203	180	2
24	25	Getafe	12	553	456	147	112	1
25	26	Hercules	20	538	628	184	149	2
26	27	Tenerife	13	510	494	155	128	2
27	28	Murcia	18	445	586	145	143	2
28	29	Alaves	12	421	380	125	81	1
29	30	Levante	11	416	402	113	95	1
...	...	...	...	...	...	...	...	
31	32	Sabadell	14	353	426	129	95	2
32	33	Cadiz	12	343	448	104	127	2
33	34	Logrones	9	293	346	96	92	1
34	35	Castellon	11	285	334	103	79	1
35	36	Albacete	7	277	270	76	76	1
36	37	Almeria	6	242	228	62	56	1
37	38	Cordoba	9	230	282	82	63	1
38	39	Compostela	4	190	160	52	45	
39	40	Recreativo	5	188	186	50	46	
40	41	Burgos CF	6	168	204	59	50	
41	42	Pontevedra	6	150	180	53	44	
42	43	Numancia	4	148	152	37	37	
43	44	Eibar	3	132	114	35	27	

44	45	Arenas	7	107	130	43	21
45	46	Real Burgos	3	96	114	26	44
46	47	Gimnastic	4	91	116	34	16
47	48	Extremadura	2	83	80	20	23
48	49	Merida	2	81	80	19	24
49	50	Alcoyano	4	76	108	30	16
50	51	Jaen	3	71	90	29	13
51	52	Real Union	4	56	72	21	14
52	53	AD Almeria	2	52	68	17	18
53	54	Europa	3	42	54	18	6
54	55	Lleida	2	40	68	13	14
55	56	Leganes	1	35	38	8	11
56	57	Xerez	1	34	38	8	10
57	58	Condal	1	22	30	7	8
58	59	Atletico Tetuan	1	19	30	7	5
59	60	Cultural Leonesa	1	14	30	5	4
60	61	Girona	1	0	0	0	0

61 rows × 20 columns

2. Print all the teams which have started playing between 1930-1980.

```
df['Debut_Start_Range'] = df['Debut'].astype(str).str[:4].astype(int)

for index in df.index:
    value = df.loc[index,'Debut_Start_Range']
    if value >= 1930 and value <= 1980:
       print(df.loc[index,'Team'])
```

```
Valencia
Sevilla
Zaragoza
Real Betis
Deportivo La Coruna
Celta Vigo
Valladolid
Sporting Gijon
Osasuna
Malaga
Oviedo
Mallorca
Las Palmas
Granada
Rayo Vallecano
Elche
Hercules
Tenerife
Murcia
Alaves
Levante
Salamanca
Sabadell
Cadiz
Castellon
Cordoba
Recreativo
Burgos CF
Pontevedra
Gimnastic
Alcoyano
Jaen
AD Almeria
Lleida
Condal
Atletico Tetuan
Cultural Leonesa
```

3. Print the list of teams which came Top 5 in terms of points

```
df['Points'] = df['Points'].astype(object).astype(int)
df.sort_values(by=['Points'],ascending=False).head(5)
```

	Pos	Team	Seasons	Points	GamesPlayed	GamesWon	GamesDrawn	GamesLost
0	1	Real Madrid	86	4385	2762	1647	552	563
1	2	Barcelona	86	4262	2762	1581	573	608
2	3	Atletico Madrid	80	3442	2614	1241	598	775
3	4	Valencia	82	3386	2664	1187	616	861
4	5	Athletic Bilbao	86	3368	2762	1209	633	920

4.Write a function with name "Goal_diff_count" which should return all the teams with their Goal Differences. U maximum and minimum goal difference.

```
df['GoalsFor'] = df['GoalsFor'].astype(object).astype(int)
df['GoalsAgainst'] = df['GoalsAgainst'].astype(object).astype(int)

def Goal_diff_count(df):
    Goal_diff_count = df.groupby("Team").sum()["GoalsFor"] - df.groupby("Team").sum()["Go
    key_max = max(Goal_diff_count.keys(), key=(lambda k: Goal_diff_count[k]))
    key_min = min(Goal_diff_count.keys(), key=(lambda k: Goal_diff_count[k]))
    team = {"TopteamName":key_max,"TopTeamGolDiff":Goal_diff_count[key_max],"LowestTeam":
    return team
topTeam = Goal_diff_count(df)
topTeam
```

```
{'LowestTeam': 'Racing Santander',
 'LowestTeamGolDiff': -525,
 'TopTeamGolDiff': 2807,
 'TopteamName': 'Real Madrid'}
```

5.Create a new column with name "Winning Percent" and append it to the data set (5 points) Percentage of Wi

If there are any numerical error, replace it with 0%

Print the top 5 teams which has the highest Winning percentage

```
df['GamesWon'] = df['GamesWon'].astype(object).astype(int)
df['GamesPlayed'] = df['GamesPlayed'].astype(object).astype(int)
df['WinningPercent'] = (df['GamesWon']/df['GamesPlayed'])*100
df.sort_values(by=['WinningPercent'],ascending=False).head(5)
```

	Pos	Team	Seasons	Points	GamesPlayed	GamesWon	GamesDrawn	GamesLost
0	1	Real Madrid	86	4385	2762	1647	552	563
1	2	Barcelona	86	4262	2762	1581	573	608
2	3	Atletico Madrid	80	3442	2614	1241	598	775
3	4	Valencia	82	3386	2664	1187	616	861
4	5	Athletic Bilbao	86	3368	2762	1209	633	920

6.Group teams based on their "Best position" and print the sum of their points for all positions (5 points)

```
df['BestPosition'] = df['BestPosition'].astype(object).astype(int)
df['Points'] = df['Points'].astype(object).astype(int)
df.groupby(['BestPosition'])['Points'].sum()
```

```
BestPosition
1     27933
2      6904
3      5221
4      6563
5      1884
6      2113
7      1186
8      1134
9        96
10      450
11      445
12      511
14       71
15       14
16       81
17      266
19       81
20       34
Name: Points, dtype: int64
```