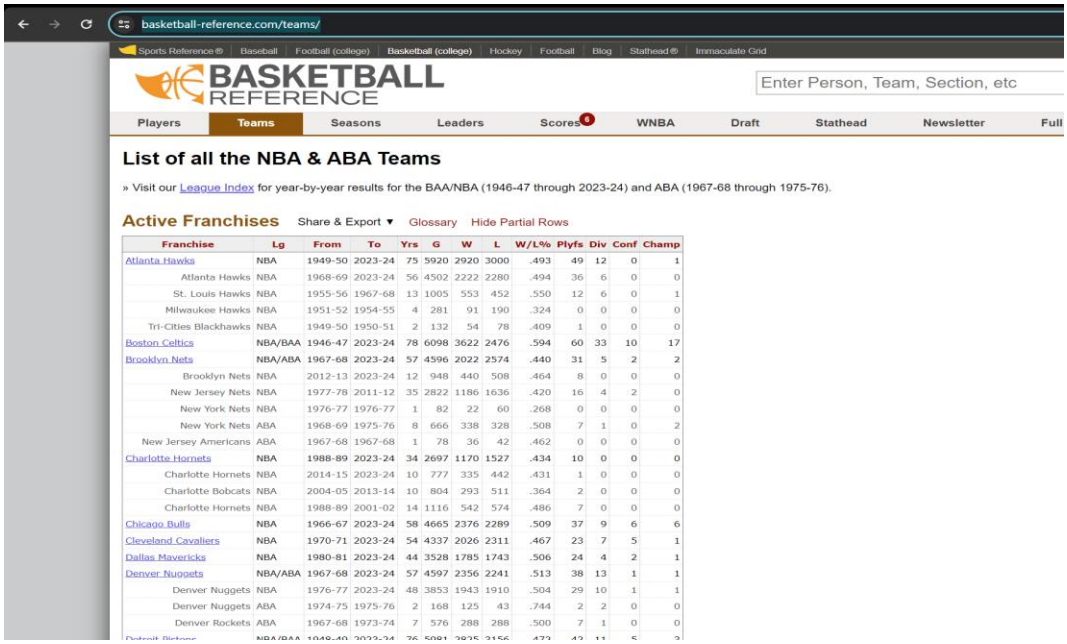
Assignment 4

Web Scrapping with Power BI

1. Select a website page/URL of your own choice (look for table format) which has some numeric/measurable fields and values (10 points). Provide screenshot of the website page you want to scrape.

Selected website URL- https://www.basketball-reference.com/teams/

- Screen shot of the selected website

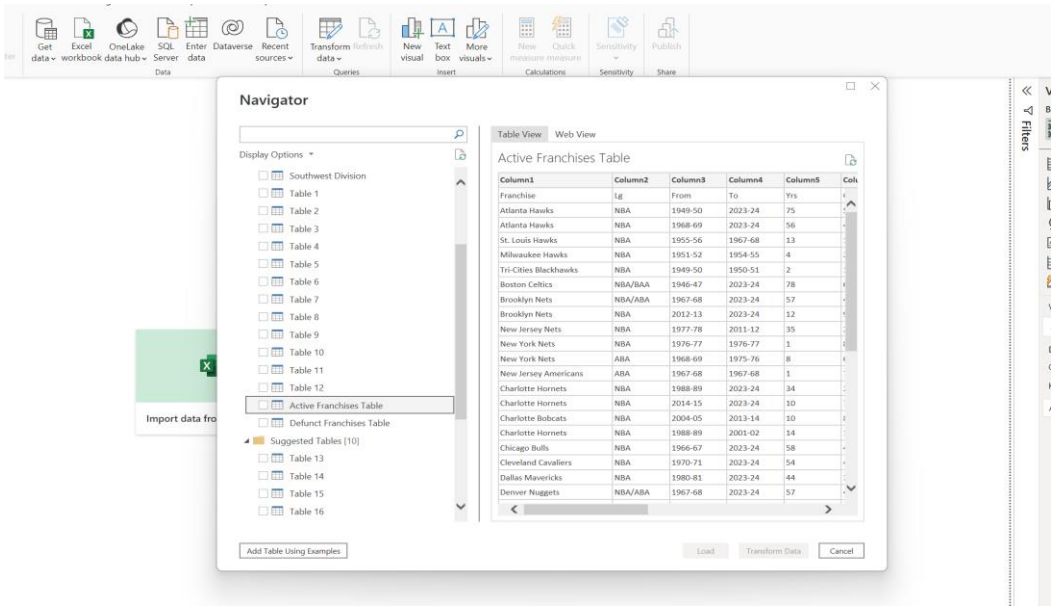


Franchise	Lg	From	To	Yrs	G	W	L	W/L%	Plyfs	Div	Conf	Champ
Atlanta Hawks	NBA	1949-50	2023-24	75	5920	2920	3000	.493	49	12	0	1
Atlanta Hawks	NBA	1968-69	2023-24	56	4502	2222	2280	.494	36	6	0	0
St. Louis Hawks	NBA	1955-56	1967-68	13	1005	553	452	.550	12	6	0	1
Milwaukee Hawks	NBA	1951-52	1954-55	4	281	91	190	.324	0	0	0	0
Tri-Cities Blackhawks	NBA	1949-50	1950-51	2	132	54	78	.409	1	0	0	0
Boston Celtics	NBA/BAA	1946-47	2023-24	78	6098	3622	2476	.594	60	33	10	17
Brooklyn Nets	NBA/ABA	1967-68	2023-24	57	4596	2022	2574	.440	31	5	2	2
Brooklyn Nets	NBA	2012-13	2023-24	12	948	440	508	.464	8	0	0	0
New Jersey Nets	NBA	1977-78	2011-12	35	2822	1186	1636	.420	16	4	2	0
New York Nets	NBA	1976-77	1976-77	1	82	22	60	.268	0	0	0	0
New York Nets	ABA	1968-69	1975-76	8	666	338	328	.508	7	1	0	2
New Jersey Americans	ABA	1967-68	1967-68	1	78	36	42	.462	0	0	0	0
Charlotte Hornets	NBA	1988-89	2023-24	34	2697	1170	1527	.434	10	0	0	0
Charlotte Hornets	NBA	2014-15	2023-24	10	777	335	442	.431	1	0	0	0
Charlotte Bobcats	NBA	2004-05	2013-14	10	804	293	511	.364	2	0	0	0
Charlotte Hornets	NBA	1988-89	2001-02	14	1116	542	574	.486	7	0	0	0
Chicago Bulls	NBA	1966-67	2023-24	58	4665	2376	2289	.509	37	9	6	6
Cleveland Cavaliers	NBA	1970-71	2023-24	54	4337	2026	2311	.467	23	7	5	1
Dallas Mavericks	NBA	1980-81	2023-24	44	3528	1785	1743	.506	24	4	2	1
Denver Nuggets	NBA/ABA	1967-68	2023-24	57	4597	2356	2241	.513	38	13	1	1
Denver Nuggets	NBA	1976-77	2023-24	48	3853	1943	1910	.504	29	10	1	1
Denver Nuggets	ABA	1974-75	1975-76	2	168	125	43	.744	2	2	0	0
Denver Rockets	ABA	1967-68	1973-74	7	576	288	288	.500	7	1	0	0

- Using the selected website in Power Bi

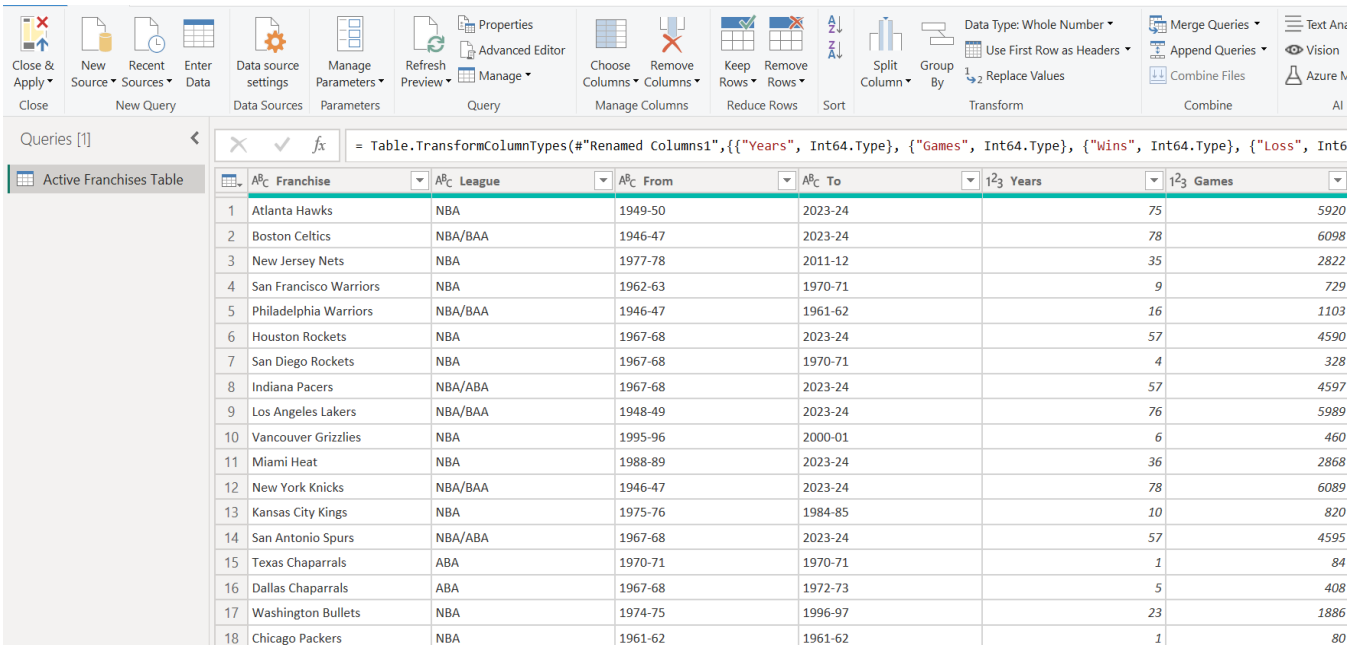
2. Scrape it using Power BI desktop (see Tutorial below for steps). Provide screenshot of the final cleaned up table in the Power Query Editor before applying it to the report (30 points).

- Selecting the 'Active Franchises table' for data transformation in power query editor.

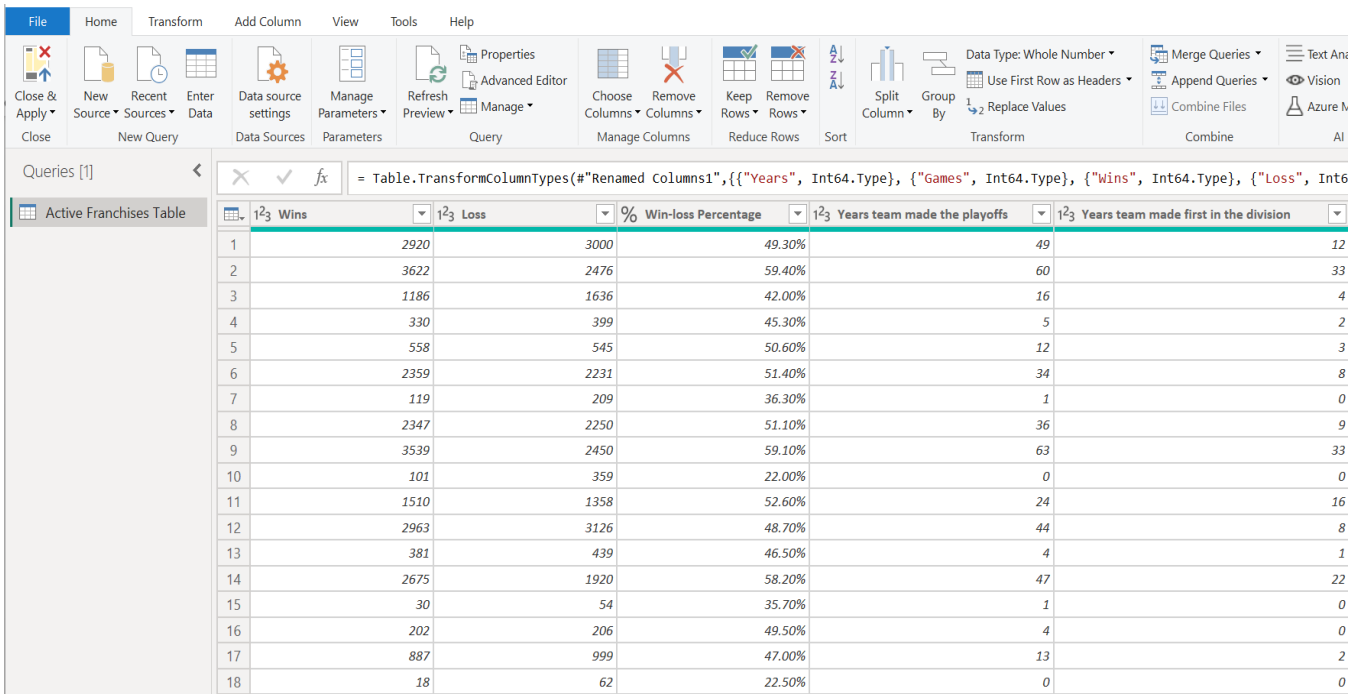


- Transformed data using power query editor.

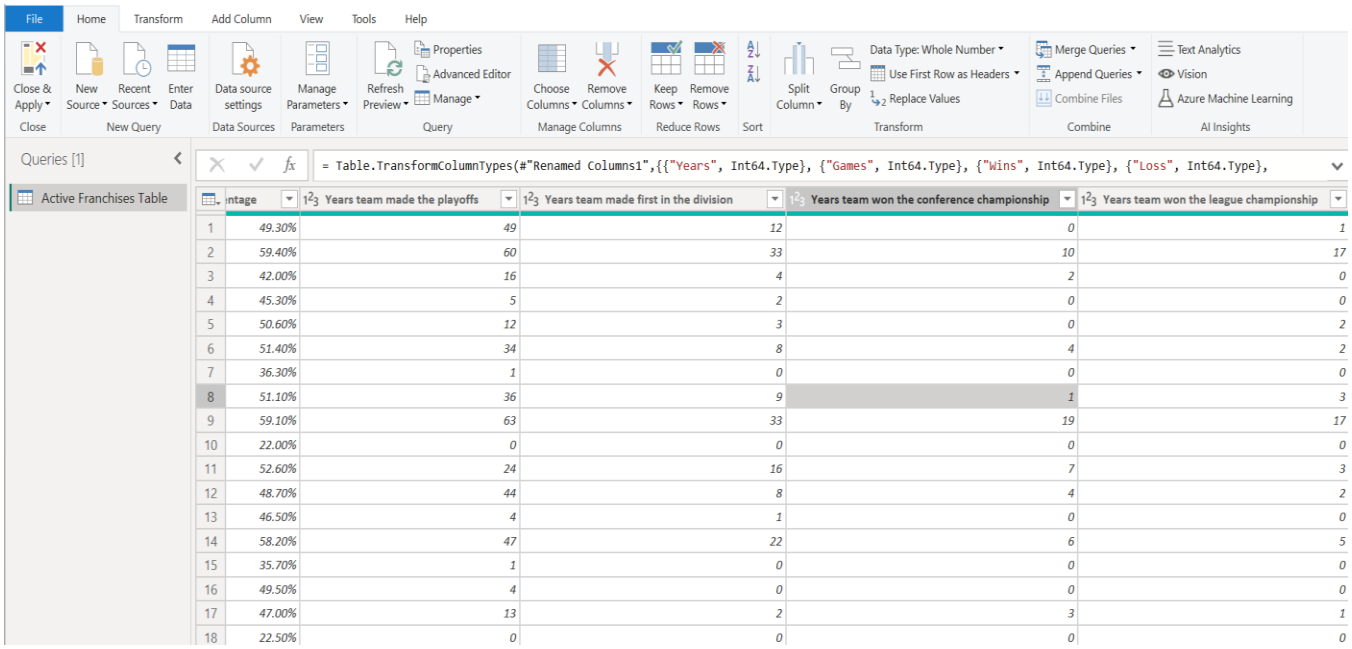
Data after transformation



	Franchise	League	From	To	Years	Games
1	Atlanta Hawks	NBA	1949-50	2023-24	75	5920
2	Boston Celtics	NBA/BAA	1946-47	2023-24	78	6098
3	New Jersey Nets	NBA	1977-78	2011-12	35	2822
4	San Francisco Warriors	NBA	1962-63	1970-71	9	729
5	Philadelphia Warriors	NBA/BAA	1946-47	1961-62	16	1103
6	Houston Rockets	NBA	1967-68	2023-24	57	4590
7	San Diego Rockets	NBA	1967-68	1970-71	4	328
8	Indiana Pacers	NBA/ABA	1967-68	2023-24	57	4597
9	Los Angeles Lakers	NBA/BAA	1948-49	2023-24	76	5989
10	Vancouver Grizzlies	NBA	1995-96	2000-01	6	460
11	Miami Heat	NBA	1988-89	2023-24	36	2868
12	New York Knicks	NBA/BAA	1946-47	2023-24	78	6089
13	Kansas City Kings	NBA	1975-76	1984-85	10	820
14	San Antonio Spurs	NBA/ABA	1967-68	2023-24	57	4595
15	Texas Chaparrals	ABA	1970-71	1970-71	1	84
16	Dallas Chaparrals	ABA	1967-68	1972-73	5	408
17	Washington Bullets	NBA	1974-75	1996-97	23	1886
18	Chicago Packers	NBA	1961-62	1961-62	1	80



	Wins	Loss	Win-loss Percentage	Years team made the playoffs	Years team made first in the division
1	2920	3000	49.30%	49	12
2	3622	2476	59.40%	60	33
3	1186	1636	42.00%	16	4
4	330	399	45.30%	5	2
5	558	545	50.60%	12	3
6	2359	2231	51.40%	34	8
7	119	209	36.30%	1	0
8	2347	2250	51.10%	36	9
9	3539	2450	59.10%	63	33
10	101	359	22.00%	0	0
11	1510	1358	52.60%	24	16
12	2963	3126	48.70%	44	8
13	381	439	46.50%	4	1
14	2675	1920	58.20%	47	22
15	30	54	35.70%	1	0
16	202	206	49.50%	4	0
17	887	999	47.00%	13	2
18	18	62	22.50%	0	0

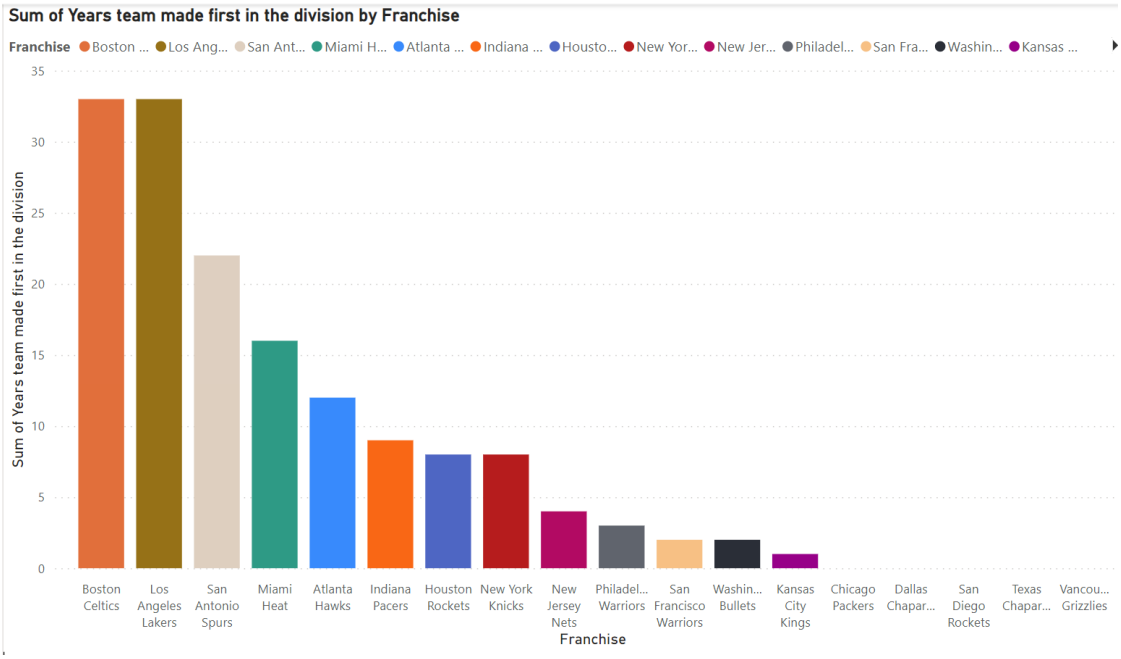


	ntage	Years team made the playoffs	Years team made first in the division	Years team won the conference championship	Years team won the league championship
1	49.30%	49	12	0	1
2	59.40%	60	33	10	17
3	42.00%	16	4	2	0
4	45.30%	5	2	0	0
5	50.60%	12	3	0	2
6	51.40%	34	8	4	2
7	36.30%	1	0	0	0
8	51.10%	36	9	1	3
9	59.10%	63	33	19	17
10	22.00%	0	0	0	0
11	52.60%	24	16	7	3
12	48.70%	44	8	4	2
13	46.50%	4	1	0	0
14	58.20%	47	22	6	5
15	35.70%	1	0	0	0
16	49.50%	4	0	0	0
17	47.00%	13	2	3	1
18	22.50%	0	0	0	0

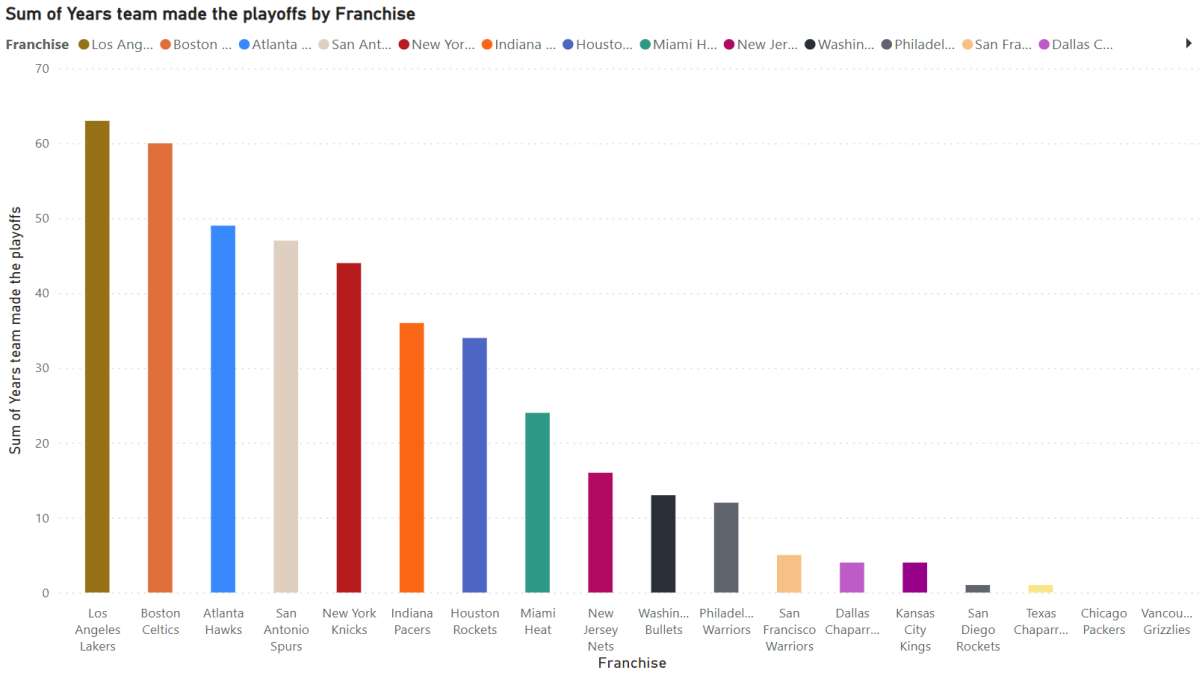
3. Create at least 4 visualizations and then Publish your report to Power BI cloud service. Give 'Titles' to each of the visualizations, they should have appropriate font size for easy readibility, not too busy (use filters as applicable) (40 points).

Visualizations:

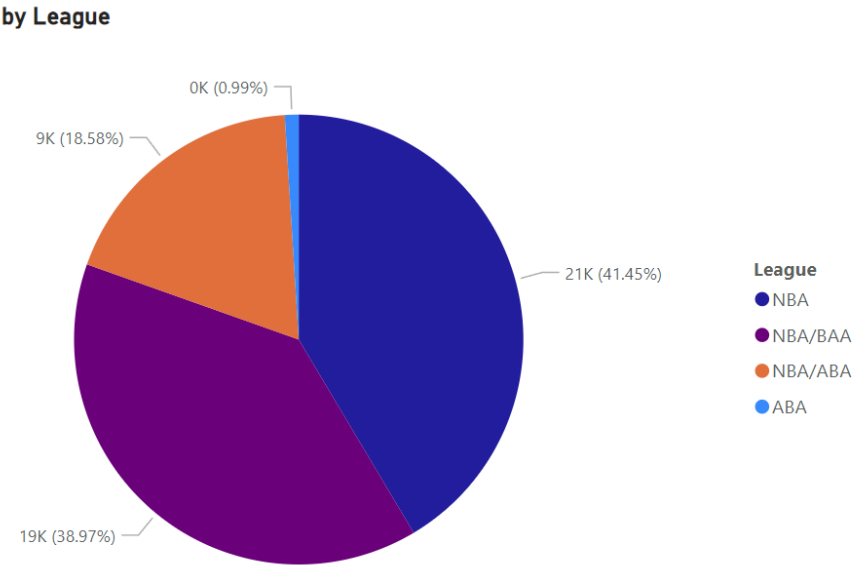
- **Stacked column chart: Sum of Years team made first in the division by Franchise**



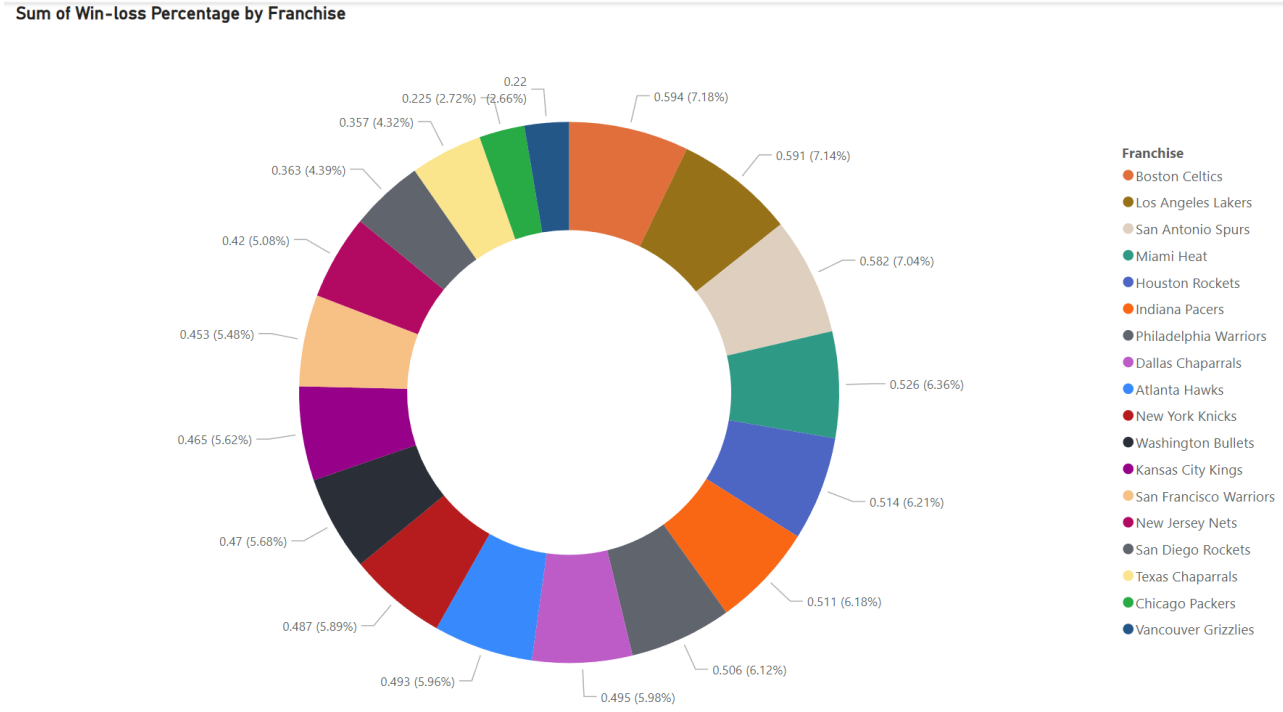
- **Ribbon chart: Sum of Years team made the playoffs by Franchise**



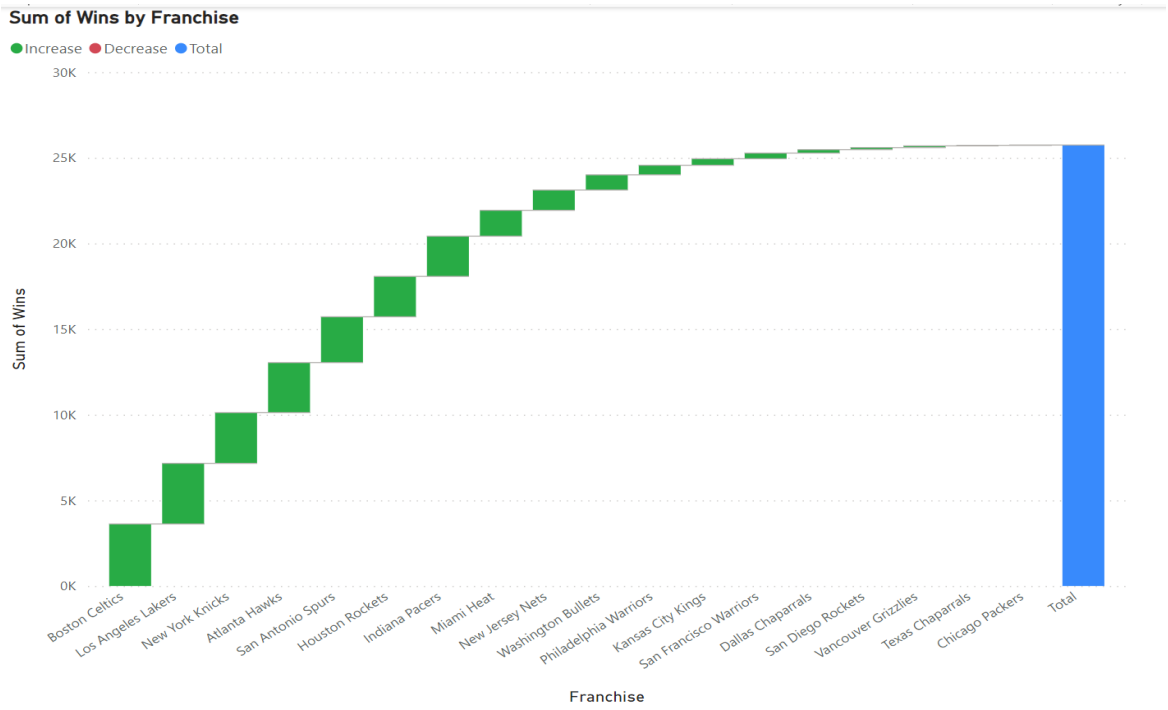
- **Pie chart: Sum of Games by League**



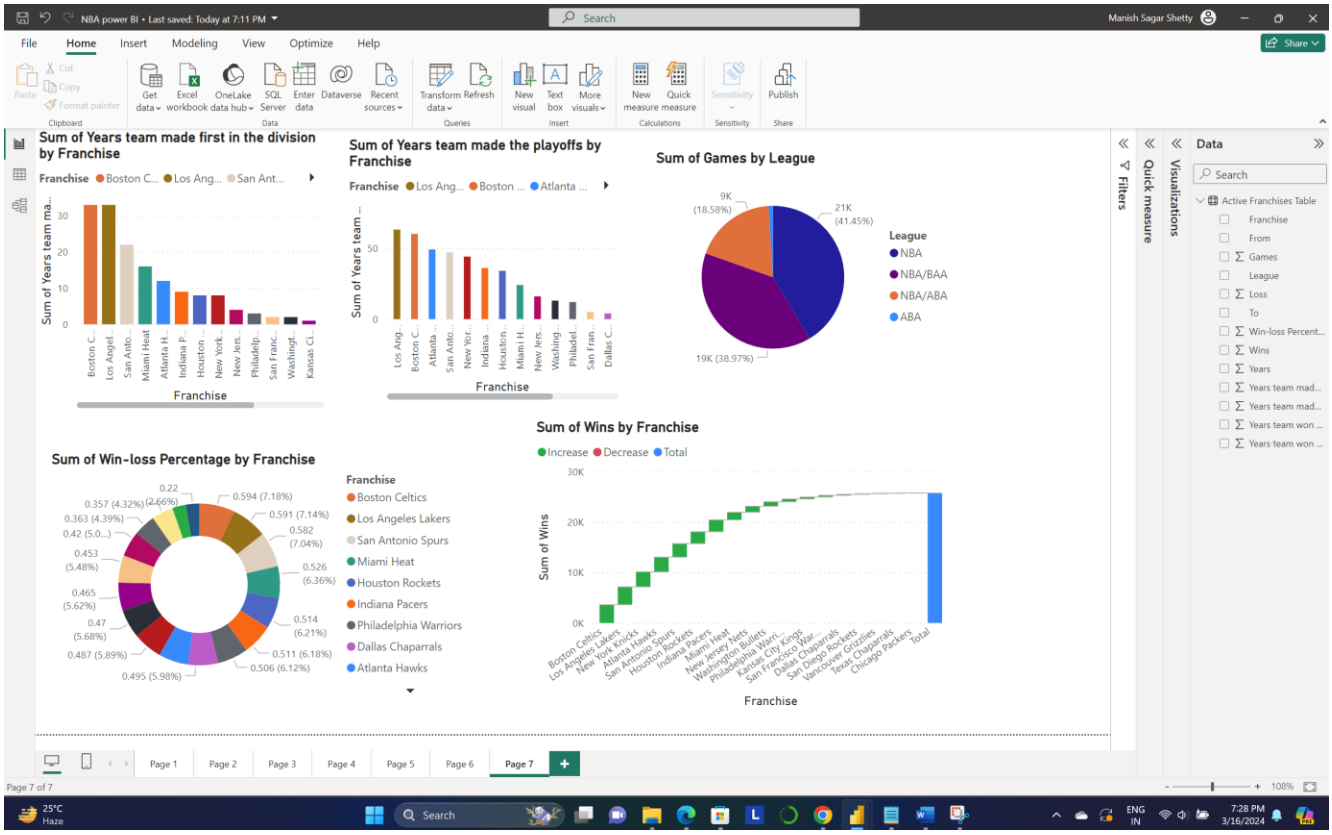
- **Donut chart: win-loss percentage by franchise**



- **Waterfall chart: Sum of wins by franchise.**



- **Report showing all the visualizations:**



4. In Power BI cloud service create a Dashboard showing the 4 visualizations. (20 points). Provide a screenshot of the full Dashboard.

- Published the report to the Power BI cloud and created a dashboard in the cloud.

Dashboard showing all the visualizations: