GROUP J –PROJECT DELIVERABLE 3 - CS425

Henrique Magalhaes Rio –hmagalhaesrio@hawk.iit.edu-20%
Suresh Muppaneni -smuppaneni@hawk.iit.edu-20%
Pranav Sistla -psistla@hawk.iit.educ--20%
Ramchandra Reddy Sathu -rsathu@hawk.iit.edu-20%
Gaurav Nagesh -gnagesh@hawk.iit.edu-20%

1) Retrieving the total budget for each team in each season, including subtotals for individual teams and seasons, and the grand total, ordered by team and season.

Query: SELECT season_id,team,SUM(team_budget) As Total_Budget from team_budget
GROUP BY CUBE(season_id,team)
ORDER BY (team,season_id);

Output:

	season_id	team	total_budget
1	2021	ATL	$19,431,009.00
2	2022	ATL	$19,470,837.00
3	2023	ATL	$17,109,014.00
4	<null>	ATL	$56,010,860.00
5	2021	ATX	$11,912,568.00
6	2022	ATX	$15,526,338.00
7	2023	ATX	$17,200,030.00
8	<null>	ATX	$44,638,936.00
9	2021	CHI	$12,995,659.00
10	2022	CHI	$19,423,055.00
11	2023	CHI	$19,395,506.00
12	<null>	CHI	$51,814,220.00
13	2021	CIN	$15,759,050.00
14	2022	CIN	$15,266,634.00
15	2023	CIN	$14,727,694.00
16	<null>	CIN	$45,753,378.00
17	2021	CLB	$13,619,212.00
18	2022	CLB	$17,165,486.00
19	2023	CLB	$14,939,367.00
20	<null>	CLB	$45,724,065.00
21	2022	CLT	$11,921,442.00
22	2023	CLT	$12,436,737.00
23	<null>	CLT	$24,358,179.00
24	2021	COL	$9,026,656.00
25	2022	COL	$8,362,930.00

100	2021	SKC	$13,778,555.00
101	2022	SKC	$11,947,409.00
102	2023	SKC	$15,195,891.00
103	<null>	SKC	$40,921,855.00
104	2023	STL	$10,706,495.00
105	<null>	STL	$10,706,495.00
106	2021	TOR	$18,788,014.00
107	2022	TOR	$32,872,845.00
108	2023	TOR	$21,324,374.00
109	<null>	TOR	$72,985,233.00
110	2021	VAN	$8,810,559.00
111	2022	VAN	$13,010,648.00
112	2023	VAN	$10,911,267.00
113	<null>	VAN	$32,732,474.00
114	2021	<null>	$343,687,579.00
115	2022	<null>	$431,619,628.00
116	2023	<null>	$417,607,000.00
117	<null>	<null>	$1,192,914,207.00

2) Query to calculate the total points for the 'ATL' and 'MIA' teams in each season, including subtotals for individual teams and seasons, ordered by team and season.

Query: SELECT Team,Season_ID,SUM(Points) FROM teams WHERE team IN ('ATL','MIA')
GROUP BY ROLLUP(Team,Season_ID)
ORDER BY Team,Season_ID;

Output:

	team	season_id	sum
1	ATL	2021	51
2	ATL	2022	40
3	ATL	2023	50
4	ATL	<null>	141
5	MIA	2021	41
6	MIA	2022	48
7	MIA	2023	34
8	MIA	<null>	123
9	<null>	<null>	264

3) Creating a view called 'Team_Rosters_By_Season' that contains the information about team rosters for each season and then retrieving the data from the view for the team 'ATL' and season '2021'.

Query: CREATE VIEW Team_Rosters_By_Season AS SELECT t.team,p.season_id,p.Player_Name FROM player p, salaries s, teams t WHERE s.player_id=p.player_id AND p.team_id=t.team_id ORDER BY Team,Season_ID;
select * from team_rosters_by_season where team='ATL' AND season_id=2021;

Output:

	team	season_id	player_name
1	ATL	2021	Marcelino-Moreno
2	ATL	2021	Josef-Martínez
3	ATL	2021	Ezequiel-Barco
4	ATL	2021	Brooks-Lennon
5	ATL	2021	Luiz-Araújo
6	ATL	2021	George-Bello
7	ATL	2021	Jake-Mulraney
8	ATL	2021	Erik-López
9	ATL	2021	Santiago-Sosa
10	ATL	2021	Erick-Torres
11	ATL	2021	Miles-Robinson
12	ATL	2021	Alan-Franco
13	ATL	2021	Jackson-Conway
14	ATL	2021	Anton-Walkes
15	ATL	2021	Jürgen-Damm
16	ATL	2021	Emerson-Hyndman
17	ATL	2021	Ronald-Hernández
18	ATL	2021	Matheus-Rossetto
19	ATL	2021	Machop-Chol
20	ATL	2021	Franco-Ibarra
21	ATL	2021	George-Campbell
22	ATL	2021	Brad-Guzan
23	ATL	2021	Lisandro-López
24	ATL	2021	Alex-De-John
25	ATL	2021	Tyler-Wolff
26	ATL	2021	Mohammed-Adams
27	ATL	2021	Alec-Kann
28	ATL	2021	Mikey-Ambrose
29	ATL	2021	Jack-Gurr

4) Query to retrieve the player names and their dense rank based on guaranteed compensation in descending order for each team in the 2022 season, ordered by team.

Query: SELECT t.Team, p.Player_Name,dense_rank()
OVER (PARTITION BY t.Team ORDER BY s.GuaranteedCompensation DESC) as Dense_Rank
FROM Teams t ,Salaries s ,Player p
WHERE s.Season_ID=2022 AND p.Player_ID=S.Player_ID AND p.Team_ID=t.team_id
order by Team;

Output:

	team	player_name	dense_rank
1	ATL	Luiz-Araújo	1
2	ATL	Josef-Martínez	2
3	ATL	Thiago-Almada	3
4	ATL	Miles-Robinson	4
5	ATL	Alan-Franco	5
6	ATL	Matheus-Rossetto	6
7	ATL	Emerson-Hyndman	7
8	ATL	Santiago-Sosa	8
9	ATL	Marcelino-Moreno	9
10	ATL	Franco-Ibarra	10
11	ATL	Brooks-Lennon	11
12	ATL	Juan-José-Purata	12
13	ATL	Brad-Guzan	13
14	ATL	Edwin-Mosquera	14
15	ATL	Andrew-Gutman	15
16	ATL	Ronald-Hernández	16
17	ATL	Raúl-Gudiño	17
18	ATL	Ronaldo-Cisneros	18
19	ATL	Bobby-Shuttleworth	19
20	ATL	Rocco-Rios-Novo	20
21	ATL	Tyler-Wolff	21
22	ATL	George-Campbell	22
23	ATL	Amar-Sejdic	23
24	ATL	Alex-De-John	23
25	ATL	Mikey-Ambrose	23
26	ATL	Dom-Dwyer	24
27	ATL	Jackson-Conway	24
28	ATL	Osvaldo-Alonso	24
29	ATL	Caleb-Wiley	25
30	ATL	Aiden-McFadden	26
31	ATL	Machop-Chol	26
32	ATX	Sebastián-Driussi	1
33	ATX	Emiliano-Rigoni	2
34	ATX	Cecilio-Domínguez	3
35	ATX	Alexander-Ring	4
36	ATX	Ruben-Gabrielsen	5
37	ATX	Danny-Hoesen	6
38	ATX	Maximiliano-Urruti	7
39	ATX	Moussa-Djitté	8
40	ATX	Diego-Fagundez	9

5) calculate the home and away goal averages along with their average goal difference between home and away games for each team in a specific season and ordered by team and season. (HomexAway Perfomance per season for all teams)

Query: SELECT x.Team,x.Season_ID as Season,x.HomeAVG,y.AwayAvg, X.HomeAVG-y.AwayAvg as AVG_GOAL_DIFFERENTIAl
FROM (SELECT t.team,ROUND(AVG(g.Homegoals),4) as HomeAVG,Season_ID,g.home_team_id FROM teams t,games g WHERE t.team_id=g.home_team_id
GROUP BY t.team_id,g.home_team_id ORDER BY Team_ID) as x,
(SELECT t.team,ROUND(AVG(g.AwayGoals),4) as AwayAvg,Season_ID,g.away_team_id
FROM teams t,games g WHERE t.team_id=g.away_team_id GROUP BY t.team_id,g.away_team_id ORDER BY Team_ID) as y
WHERE x.Home_Team_ID=y.Away_team_ID;

Output:

	team	season	homeavg	awayavg	avg_goal_differential
1	ATL	2021	1.2941	1.1765	0.1176
2	ATL	2022	1.8235	1	0.8235
3	ATL	2023	2.3529	1.375	0.9779
4	ATX	2021	1.5294	0.4706	1.0588
5	ATX	2022	2.4118	1.3529	1.0589
6	ATX	2023	1.5882	1.25	0.3382
7	CHI	2021	1.2353	0.7647	0.4706
8	CHI	2022	1.0588	1.2353	-0.1765
9	CHI	2023	1.4118	0.75	0.6618
10	CIN	2021	0.9412	1.1765	-0.2353
11	CIN	2022	2.1176	1.4118	0.7058
12	CIN	2023	1.875	1.4118	0.4632
13	CLB	2021	1.7059	0.7059	1
14	CLB	2022	1.2941	1.4118	-0.1177
15	CLB	2023	2.375	1.3529	1.0221
16	CLT	2022	1.5882	0.8824	0.7058
17	CLT	2023	1.3125	1.2353	0.0772
18	COL	2021	1.7647	1.1765	0.5882
19	COL	2022	1.7059	1	0.7059
20	COL	2023	0.8125	0.7647	0.0478
21	DCU	2021	1.8235	1.4118	0.4117
22	DCU	2022	1.2941	0.8235	0.4706

6) query to retrieve the top 10 goalkeepers with the highest ratio of goals saved to shots faced (in percentage) in seasons where they faced more than 100 shots, along with their names and season IDs, ordered by the ratio in descending order.

Query: SELECT GK_Name,Season_ID,concat(round(cast(Saves as DECIMAL)/cast(ShotsFaced as decimal)*100),'%') as RATIO_GOALSSAVED from goalkeepers WHERE ShotsFaced>100
GROUP BY Season_ID, concat(round(cast(Saves as DECIMAL)/cast(ShotsFaced as decimal)*100),'%'), GK_Name
ORDER BY RATIO_GOALSSAVED DESC LIMIT 10;

Output:

	gk_name	season_id	ratio_goalssaved
1	Maarten-Paes	2023	78%
2	Đorđe-Petrović	2023	78%
3	Brad-Guzan	2021	77%
4	Maxime-Crépeau	2021	77%
5	Andre-Blake	2022	77%
6	Joe-Willis	2021	76%
7	Đorđe-Petrović	2022	76%
8	Tyler-Miller	2021	75%
9	William-Yarbrough	2021	75%
10	Joe-Willis	2023	75%

7) Query to retrieve the top 10 STRIKERS with the highest probability of scoring a goal, given it is a shot on goal in the 2022 season, where they played more than 1000 minutesordered by the ratio in descending order.

Query: SELECT Player_Name,concat(round(cast(goals as decimal)/cast(shotsongoal as decimal)*100,2),'%') AS Ratio_GoalsVSShotsonGoal
FROM player
WHERE Season_ID=2022 AND minutes>1000 AND Position='ST'
order by Ratio_GoalsVSShotsonGoal DESC
limit 10;

Output:

	player_name	shotsongoal	goals	ratio_goalsvsshotsongoal
1	Gonzalo-Higuaín	20	16	80.00%
2	Ercan-Kara	19	11	57.89%
3	Héber	15	8	53.33%
4	Kei-Kamara	17	9	52.94%
5	Brenner	34	18	52.94%
6	Gyasi-Zardes	19	10	52.63%
7	Dejan-Joveljic	22	11	50.00%
8	Daniel-Ríos	14	7	50.00%
9	Carlos-Ferreira	26	13	50.00%
10	Jaroslaw-Niezgoda	18	9	50.00%

8) Query to calculate the percent rank of each position based on the total number of goals scored by players in that position, excluding goalkeepers, and then display the result as a percentage, ordered by position.

Query: SELECT Position,CONCAT(ROUND(cast(percent_rank()
OVER (ORDER BY SUM(Goals)) as DECIMAL),2)*100,'%') as Position_Rank
FROM Player WHERE NOT Position='GK'
GROUP BY Position;

Output:

	position	position_rank
1	DM	0.00%
2	FB	17.00%
3	CB	33.00%
4	CM	50.00%
5	AM	67.00%
6	W	83.00%
7	ST	100.00%

9) Query to calculate the percentage of players who did not score any goals in the 2021 season and display the result.

Query: Select concat(ROUND(cast(count(*) as DECIMAL)*100/(Select cast(count(*) as decimal) from Player WHERE Season_ID=2021),2),'%') as NONSCORERS
from Player where Goals=0 AND Season_ID=2021;

Output:

	nonscorers
1	51.64%

10) Query to retrieve top team in each season with the highest number of points.

Query: SELECT Season_ID,*MAX*(Points) as Max,Team
FROM teams
GROUP BY Season_ID,Team_ID
order by Max DESC
LIMIT 3;

Output:

	season_id	max	team
1	2021	73	NER
2	2023	68	CIN
3	2022	67	PHI

11) Running Total budget of a team by season.

Query: SELECT team,season_id,Team_Budget,
SUM(Team_Budget) OVER (partition by team ORDER BY season_id) as RunningTotalBudget
FROM team_budget;

Output:

	team	season_id	team_budget	runningtotalbudget
1	ATL	2021	$19,431,009.00	$19,431,009.00
2	ATL	2022	$19,470,837.00	$38,901,846.00
3	ATL	2023	$17,109,014.00	$56,010,860.00
4	ATX	2021	$11,912,568.00	$11,912,568.00
5	ATX	2022	$15,526,338.00	$27,438,906.00
6	ATX	2023	$17,200,030.00	$44,638,936.00
7	CHI	2021	$12,995,659.00	$12,995,659.00
8	CHI	2022	$19,423,055.00	$32,418,714.00
9	CHI	2023	$19,395,506.00	$51,814,220.00
10	CIN	2021	$15,759,050.00	$15,759,050.00
11	CIN	2022	$15,266,634.00	$31,025,684.00
12	CIN	2023	$14,727,694.00	$45,753,378.00
13	CLB	2021	$13,619,212.00	$13,619,212.00
14	CLB	2022	$17,165,486.00	$30,784,698.00
15	CLB	2023	$14,939,367.00	$45,724,065.00

12) Query to calculate the average cumulative distribution of passing scores for players based on their positions.

QUERY: SELECT Position, cume_dist() OVER (ORDER BY AVG(Score)) as AVG_CUME_DIST_PASSING_SCORE FROM Player,Passes where Passes.Player_ID=Player.Player_ID group by Position;

Output:

	position	avg_cume_dist_passing_score
1	ST	0.125
2	W	0.25
3	GK	0.375
4	AM	0.5
5	FB	0.625
6	CM	0.75
7	DM	0.875
8	CB	1

13) Query to retrieve the TOP 15 games with the highest aggregate score. (Aggregate score is total number of goals scored by both teams in that game)

QUERY:

Query: SELECT Home_Team_ID,Away_team_ID,*MAX*(HomeGoals+AwayGoals) as Total_Goals_Scored FROM Games
GROUP BY Home_Team_ID, Away_team_ID
ORDER BY Total_Goals_Scored DESC LIMIT 15;

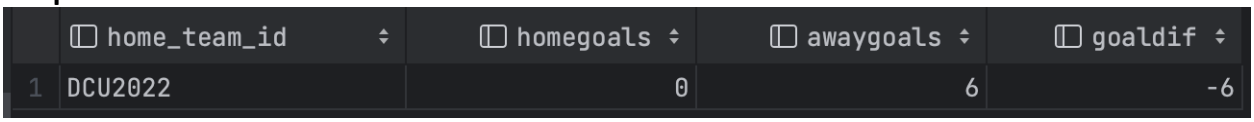
Output:

	home_team_id	away_team_id	total_goals_scored
1	MTL2021	CIN2021	9
2	NYC2022	TOR2022	9
3	POR2022	SKC2022	9
4	NYRB2022	COL2022	9
5	STL2023	ATX2023	9
6	MIA2022	CIN2022	8
7	DCU2023	NYRB2023	8
8	VAN2023	HOU2023	8
9	ORL2022	DCU2022	8
10	POR2021	SEA2021	8
11	ATX2021	FCD2021	8
12	CIN2022	NYC2022	8
13	CIN2021	NSH2021	8
14	RSL2021	SJE2021	7
15	DCU2021	TOR2021	7

14) Worst Home Game performance for all seasons by a Team.

--- Biggest Home Supporter Disapointment
Query: SELECT Home_Team_ID, HomeGoals, AwayGoals, Games.HomeGoals-Games.AwayGoals as GoalDif
FROM games WHERE AwayGoals>HomeGoals
order by GoalDif
LIMIT 1;

Output:

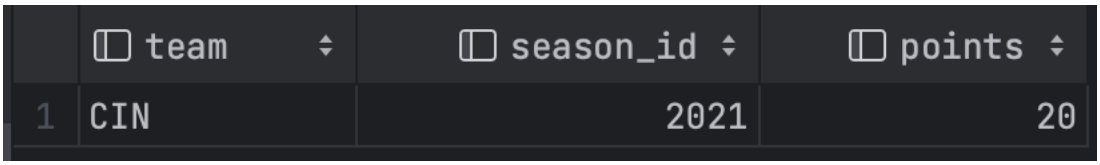


	home_team_id	homegoals	awaygoals	goaldif
1	DCU2022	0	6	-6

15) Lowest Points Scored in a Season by a Team.

Query: SELECT Team,Season_ID,MIN(Points) as points
FROM Teams
GROUP BY Team, Season_ID
order by points
LIMIT 1;

Output:

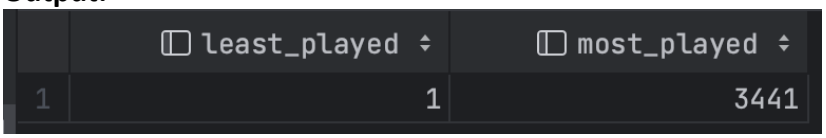


	team	season_id	points
1	CIN	2021	20

16) Players with the Least Minutes played and Player with the most minutes played.

Query: SELECT MIN(Minutes) as Least_Played, MAX(minutes) as Most_Played
From player;

Output:



	least_played	most_played
1	1	3441

17) Player with the Lowest Salary and Player with the Highest salary.

Query: SELECT cast(MIN(Base_Salary) as money) as Lowest_Salary, cast(MAX(Base_Salary) as money) as Highest_Salary From salaries;

Output:

	lowest_salary	highest_salary
1	$63,547.00	$14,000,000.00

18) 7-Day Moving Average of totals goals scored in Games.

Query: SELECT date,homegoals+games.awaygoals as Total_Goals,
AVG(homegoals+games.awaygoals) OVER (ORDER BY Date ROWS BETWEEN 3 PRECEDING AND 3 FOLLOWING) as Moving_AVG_7_Day
FROM Games LIMIT 20;

Output:

	date	total_goals	moving_avg_7_day
1	2021-04-16	4	2.25
2	2021-04-16	3	2.4
3	2021-04-17	0	2
4	2021-04-17	2	2.1428571428571429
5	2021-04-17	3	2.1428571428571429
6	2021-04-17	0	2.2857142857142857
7	2021-04-17	3	3.1428571428571429
8	2021-04-17	4	3
9	2021-04-17	4	3.2857142857142857
10	2021-04-17	6	3.2857142857142857
11	2021-04-18	1	3.1428571428571429
12	2021-04-18	5	2.8571428571428571
13	2021-04-18	0	2.8571428571428571
14	2021-04-23	2	2.5714285714285714
15	2021-04-24	2	3
16	2021-04-24	4	2.8571428571428571
17	2021-04-24	4	3.1428571428571429
18	2021-04-24	4	2.8571428571428571
19	2021-04-24	4	3
20	2021-04-24	2	2.8571428571428571