**How to join the related tables of absenteeism ?**

**Find the healthiest person from the database?**

**What is the compensation rate increase for non-smokers?**

**Optimize the query ?**

**Optimize the query creating seasons column along with the reason column?**

**Optimize the query retreive the data of seasons and the person BMI Categories?**

**-- Creating join table**

select \* from absenteeism\_work a

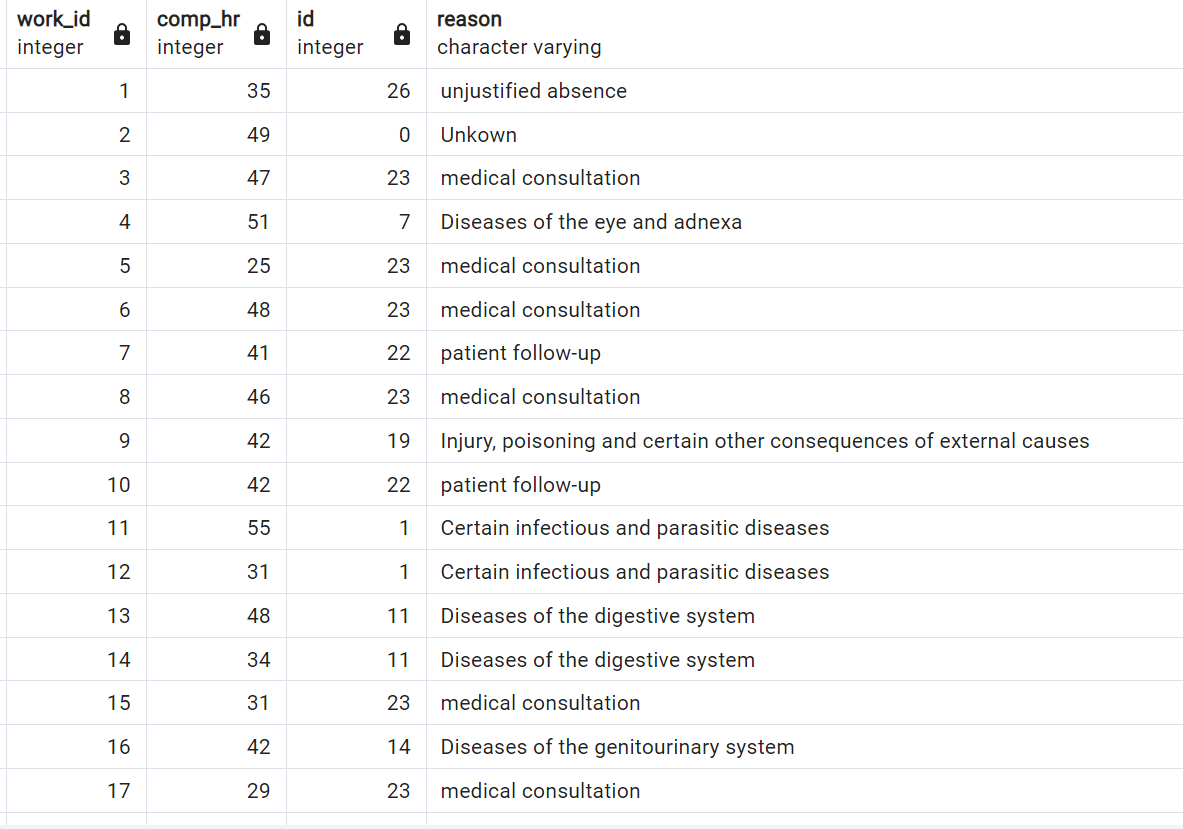
left join work\_compensation b

on a.id = b. work\_id

left join work\_reasons c

on a.reason\_absence = c.id;

We can see the main table named absenteeism\_work has been joined and the below is the image of displayed columns of two sub tables in absenteeism\_work.



**--Finding the healthiest**

select \* from absenteeism\_work

where social\_drinker = false and social\_smoker = false

and body\_mass\_index <25 and

absenteeism\_time\_in\_hours < (select avg(absenteeism\_time\_in\_hours) from absenteeism\_work)

LIMIT 5;

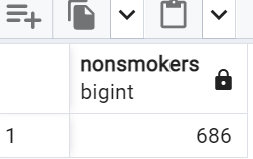
Retrieve the top 5 healthiest individuals based on criteria from four columns: `social\_drinker`, `social\_smoker`, `body\_mass\_index`, and `absenteeism\_time\_in\_hours`. A person is deemed healthier if they don't drink or smoke socially, have a body mass index less than 25, and their absenteeism time is below the average.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| id | reason\_absence | month\_absence | day\_of\_week | seasons | transportation\_expense | distance\_to\_work | service\_time | age | workload\_per\_day | hit\_target | disciplinary\_failure | education | son | social\_drinker | social\_smoker | pet | weight | height | body\_mass\_index | absenteeism\_time\_in\_hours |  |
| 41 | 23 | 9 | 3 | 1 | 184 | 42 | 7 | 27 | 241476 | 92 | FALSE | 1 | 0 | FALSE | FALSE | 0 | 58 | 167 | 21 | 2 |  |
| 52 | 0 | 9 | 2 | 4 | 225 | 26 | 9 | 28 | 241476 | 92 | TRUE | 1 | 1 | FALSE | FALSE | 2 | 69 | 169 | 24 | 0 |  |
| 53 | 23 | 9 | 3 | 4 | 225 | 26 | 9 | 28 | 241476 | 92 | FALSE | 1 | 1 | FALSE | FALSE | 2 | 69 | 169 | 24 | 2 |  |
| 57 | 18 | 9 | 4 | 4 | 225 | 26 | 9 | 28 | 241476 | 92 | FALSE | 1 | 1 | FALSE | FALSE | 2 | 69 | 169 | 24 | 3 |  |
| 66 | 23 | 10 | 5 | 4 | 179 | 26 | 9 | 30 | 253465 | 93 | FALSE | 3 | 0 | FALSE | FALSE | 0 | 56 | 171 | 19 | 1 |  |

**---compensation rate increase for non-smokers/budget $983,221 so .68 increase per hour/$1,414 per year**

select count(\*) as nonsmokers from absenteeism\_work

where social\_smoker = false;



**--optimize this query**

select

a.id,

r.reason

from absenteeism\_work a

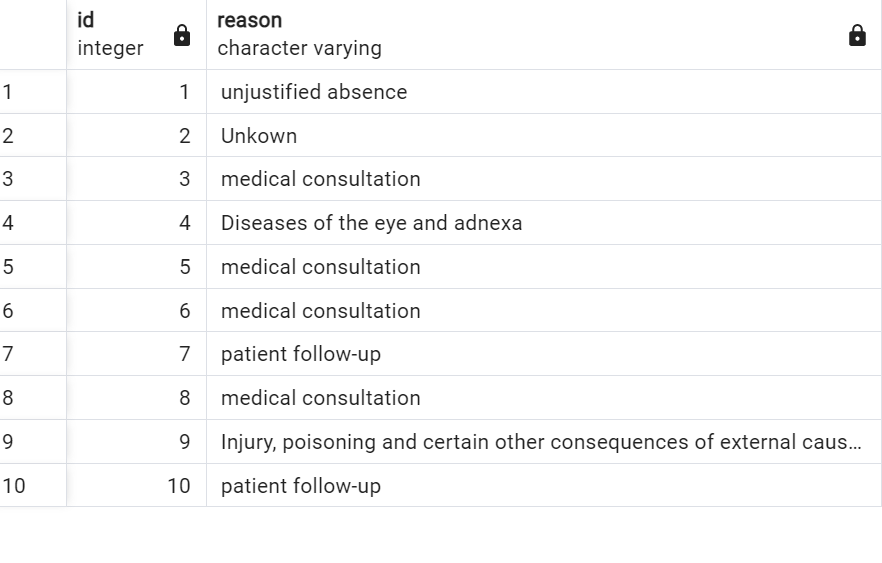
left join work\_compensation b

on a.id = b.work\_id

left join work\_reasons r on

a.reason\_absence = r.id

LIMIT 10;



-**-optimize this query (creating a seasons column with case statement)**

select

a.id,

r.reason,

CASE WHEN Month\_absence IN (12,1,2) THEN 'Winter'

WHEN Month\_absence IN (3,4,5) THEN 'spring'

WHEN Month\_absence IN (6,7,8) THEN 'summer'

WHEN Month\_absence IN (9,10,11) THEN 'Fall'

ELSE 'Unknown' END as season\_names

from absenteeism\_work a

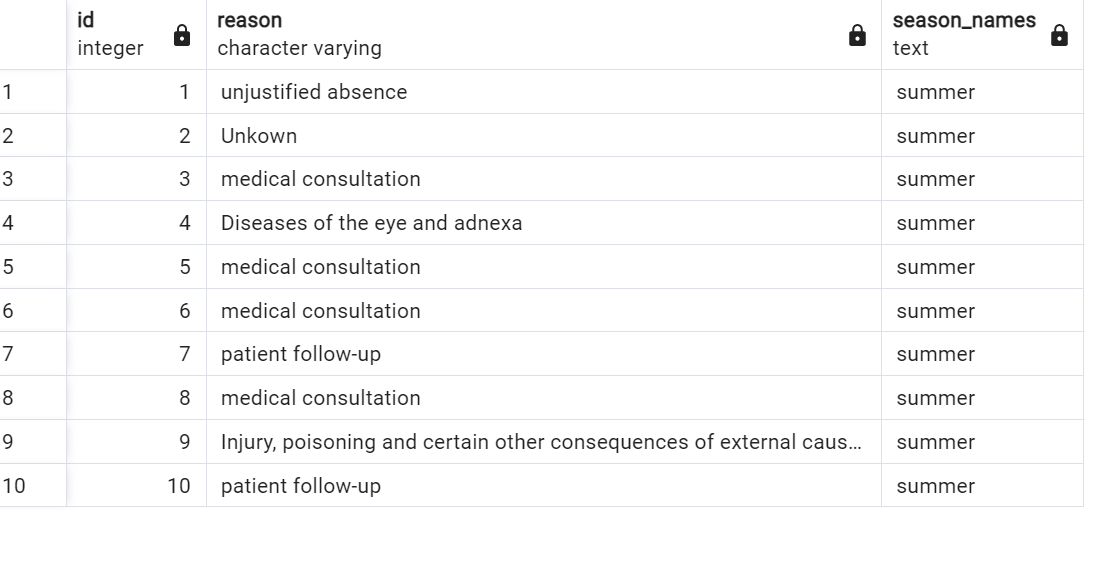
left join work\_compensation b

on a.id = b.work\_id

left join work\_reasons r on

1. reason\_absence = r.id

Limit 10;



**--optimize this query(with BMI Category and seasons column using Case statements)**

select

a.id,

r.reason,

month\_absence,

body\_mass\_index,

CASE WHEN body\_mass\_index <18.5 THEN 'Under Weight'

WHEN body\_mass\_index between 18.5 and 25 THEN 'Normal Weight'

WHEN body\_mass\_index between 25 and 30 THEN 'Over Weight'

WHEN body\_mass\_index >18.5 THEN 'Obesity'

ELSE 'Unknown' END as BMI\_Category,

CASE WHEN Month\_absence IN (12,1,2) THEN 'Winter'

WHEN Month\_absence IN (3,4,5) THEN 'spring'

WHEN Month\_absence IN (6,7,8) THEN 'summer'

WHEN Month\_absence IN (9,10,11) THEN 'Fall'

ELSE 'Unknown' END as season\_names

from absenteeism\_work a

left join work\_compensation b

on a.id = b.work\_id

left join work\_reasons r on

a.reason\_absence = r.id

LIMIT 10;



The query with added fields to use it to represent in powerbi -

select

a.id,

r.reason,

month\_absence,

body\_mass\_index,

CASE WHEN body\_mass\_index <18.5 THEN 'Under Weight'

WHEN body\_mass\_index between 18.5 and 25 THEN 'Normal Weight'

WHEN body\_mass\_index between 25 and 30 THEN 'Over Weight'

WHEN body\_mass\_index >18.5 THEN 'Obesity'

ELSE 'Unknown' END as BMI\_Category,

CASE WHEN Month\_absence IN (12,1,2) THEN 'Winter'

WHEN Month\_absence IN (3,4,5) THEN 'spring'

WHEN Month\_absence IN (6,7,8) THEN 'summer'

WHEN Month\_absence IN (9,10,11) THEN 'Fall'

ELSE 'Unknown' END as season\_names,

seasons,

month\_absence,

day\_of\_week,

transportation\_expense,

education,

son,

social\_drinker,

social\_smoker,

pet,

disciplinary\_failure,

age

from absenteeism\_work a

left join work\_compensation b

on a.id = b.work\_id

left join work\_reasons r on

a.reason\_absence = r.id

LIMIT 10;

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| id | reason | month\_absence | body\_mass\_index | bmi\_category | season\_names | seasons | month\_absence-2 | day\_of\_week | transportation\_expense | education | son | social\_drinker | social\_smoker | pet | disciplinary\_failure | age |
| 1 | unjustified absence | 7 | 30 | Over Weight | summer | 1 | 7 | 3 | 289 | 1 | 2 | TRUE | FALSE | 1 | FALSE | 33 |
| 2 | Unkown | 7 | 31 | Obesity | summer | 1 | 7 | 3 | 118 | 1 | 1 | TRUE | FALSE | 0 | TRUE | 50 |
| 3 | medical consultation | 7 | 31 | Obesity | summer | 1 | 7 | 4 | 179 | 1 | 0 | TRUE | FALSE | 0 | FALSE | 38 |
| 4 | Diseases of the eye and adnexa | 7 | 24 | Normal Weight | summer | 1 | 7 | 5 | 279 | 1 | 2 | TRUE | TRUE | 0 | FALSE | 39 |
| 5 | medical consultation | 7 | 30 | Over Weight | summer | 1 | 7 | 5 | 289 | 1 | 2 | TRUE | FALSE | 1 | FALSE | 33 |
| 6 | medical consultation | 7 | 31 | Obesity | summer | 1 | 7 | 6 | 179 | 1 | 0 | TRUE | FALSE | 0 | FALSE | 38 |
| 7 | patient follow-up | 7 | 27 | Over Weight | summer | 1 | 7 | 6 | 361 | 1 | 1 | TRUE | FALSE | 4 | FALSE | 28 |
| 8 | medical consultation | 7 | 23 | Normal Weight | summer | 1 | 7 | 6 | 260 | 1 | 4 | TRUE | FALSE | 0 | FALSE | 36 |
| 9 | Injury, poisoning and certain other consequences of external causes | 7 | 25 | Normal Weight | summer | 1 | 7 | 2 | 155 | 1 | 2 | TRUE | FALSE | 0 | FALSE | 34 |
| 10 | patient follow-up | 7 | 29 | Over Weight | summer | 1 | 7 | 2 | 235 | 3 | 1 | FALSE | FALSE | 1 | FALSE | 37 |