find the healthiest employees for the bonus
select * from Absenteeism_at_work
where Social_drinker = 0 and Social_smoker = 0 and Body_mass_index < 25
we also want people who have absences

-- that are lower than average so I am going to use a SUBQUERY

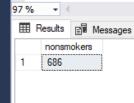
WHERE Social\_drinker = 0 and Social\_smoker = 0 and Body\_mass\_index < 25 and

absenteeism\_time\_in\_hours < (SELECT AVG(Absenteeism\_time\_in\_hours) FROM Absenteeism\_at\_work)

98 % + ∢ ID Reason\_for\_absence Month\_of\_absence Day of the week Distance from Residence to V Seasons Transportation expense 66 23 68 23 70 23 <

Query executed successfully.

-- compensation\_rate increase for non-smokers, knowing that the budget cap from HR is \$983,221 so 0.68 increase per hour select count(\*) AS nonsmokers from Absenteeism\_at\_work a left join compensation c on a.id = c.id where social\_smoker = 0  $\dot{\bar{\mathbb{S}}}$  -- we know that we work 40 hours a week and the total number of weeks in a year is 52. -- 40 \* 52 = 2080 total number of hours worked in a year -- 2080 \* 0.68 = \$ 1,414.4 -- the company is being able to give each employee \$1,414.4 per year if they are non\_smokers



```
-- optimize this query to bring in to the dashboard in Power BI. I want to import into Power BI the following columns:
   🖆 select a.ID, r.reason, a.Body_mass_index, a.Month_of_absence, a.Day_of_the_week, a.Transportation_expense, a.education, a.son, a.social_drinker, a.social_smoker, a.pet,
    a.Disciplinary_failure, a.Age, a.Work_load_Average_day, a.Absenteeism_time_in_hours
    from absenteeism_at_work a
    left join compensation b
    on a.ID = b.ID
    left join Reasons r
    on a.Reason_for_absence = r.Number
    🖆-- As part of optimizing I also want to be able to make some categories with Body_mass_index column and make some categories with the Month_of_absence column.
    -- I could do this job both in Power BI but also in SQL by using CASE statements
    select a.ID, r.reason, a.Body_mass_index, a.Month_of_absence, a.Day_of_the_week, a.Transportation_expense, a.education, a.son, a.social_drinker, a.social_smoker, a.pet,
    a.Disciplinary_failure, a.Age, a.Work_load_Average_day, a.Absenteeism_time_in_hours,
    CASE WHEN Body_mass_index < 18.5 THEN 'Underweight'
         WHEN Body_mass_index BETWEEN 18.5 AND 25 THEN 'Healthy'
         WHEN Body_mass_index BETWEEN 25 AND 30 THEN 'Overweight'
         WHEN Body_mass_index > 18.5 THEN 'Obese'
    ELSE 'Unknown' END AS BMI_Category,
     CASE WHEN month_of_absence IN (12, 1, 2) THEN 'Winter'
         WHEN month_of_absence IN (3, 4, 5) THEN 'Spring'
          WHEN month_of_absence IN (6,7,8) THEN 'Summer
          WHEN month_of_absence IN (9, 10, 11) THEN 'Fall'
    ELSE 'Unknown' END AS Season_Names
     from absenteeism_at_work a
    left join compensation b
     on a.ID = b.ID
     left join Reasons r
     on a.Reason_for_absence = r.Number
97 %
 Results Messages
     ID reason Body_mass_index Month_of_absence Day_of_the_week Transportation_expense education son social_drinker social_smoker pet Disciplinary_failure Age Work_load_Average_day Absenteeism_time_in_hours BMI_Cate(
```