

SLEEP AND HEALTH INSIGHTS

REPORT



OVERVIEW

This project analyzes health and sleep data collected from a group of individuals in different occupations, focusing on their lifestyle, stress, and sleep quality.

The primary objective is to identify patterns that impact sleep health and provide insights that can guide wellness programs, workplace health initiatives, and individual lifestyle recommendations.

By leveraging SQL normalization and analytical querying, the project transforms raw health data into structured, actionable intelligence.

PROBLEM STATEMENT

The organization aims to understand the factors that influence poor sleep quality and health imbalances among working individuals.

Currently, the company lacks clear insights into:

- How occupation and stress levels impact sleep quality
- The role of physical activity in maintaining healthy BMI
- The prevalence of sleep disorders across different demographics

Without these insights, wellness programs and policy recommendations rely on assumptions rather than data-driven evidence.

What's Happening?

- The collected dataset includes personal details, sleep details and health details but was initially unstructured and unnormalized, making analysis difficult.
- No clear mapping existed between demographic details (age, gender, occupation), sleep details (sleep duration, quality of sleep, sleep disorder) and health metrics (stress, BMI, activity).
- As a result, identifying root causes of poor sleep or high stress levels was challenging.

SOLUTION GOAL

- To **clean, normalize, and structure** the dataset into relational tables — **Person**, **Sleep_details** and **Health_Info**.
- To analyze key health relationships using **SQL queries**, including:
 - How occupation affects stress and sleep quality
 - How activity and BMI influence overall health
 - Which age groups or occupations are most prone to sleep disorders
- To generate **insights for preventive health measures** and improved lifestyle recommendations.

DATA DESCRIPTION

Title: Sleep and Health Information Dataset

Sample Size: 374 Records | 12 Attributes

Data Type: Numeric, Text, and Categorical

Tables Created

- **Person** - Columns: Person_ID, Age, Gender, Occupation
- **Sleep_details** - Columns : Sleep_ID, Person_ID, Sleep_Duration, Quality_Of_Sleep, Sleep_Disorder
- **Health_Info** - Columns: Health_ID, Person_ID, Physical_Activity, Stress_Level, BMI_Category, Heart_Rate, Daily_Steps

Key Fields

- Person_ID, Sleep_ID, Health_ID
- Age, Gender, Occupation
- Sleep_Duration, Quality_Of_Sleep, Sleep_Disorder
- Stress_Level, BMI_Category, Physical_Activity, Heart_Rate, Daily_Steps

DATA CLEANING STEPS

- Removed duplicates (e.g., repeated records for same person)
- Normalized dataset into three relational tables:
 - **Person** (Demographics)
 - **Sleep** (Sleep details)
 - **Health_Info** (Health metrics)
- Standardized categorical text values (e.g., “None”, “Sleep Apnea”, “Insomnia”)
- Verified referential integrity using **foreign key** relationships (**PersonID**)
- Ensured proper data types (numeric for metrics, text for categories)

DATA MODEL

The database follows a **1-to-1 relationship pattern**:

Each record in the **Person** table (one individual) connects to one record in

Health_Info table and one record in **Sleep_Details** table

This structure allows efficient analysis of trends across demographics, occupations, sleep details and health conditions.

Category 1: Sleep Pattern

1. Which occupation tends to get less sleep?

```
--> 1. which occupation tends to get less sleep?  
select p.occupation, round(avg(s.sleep_duration),2) as avg_sleep  
from Person p  
join sleep_details s on p.person_id=s.person_id  
group by p.occupation  
order by avg_sleep asc;
```

	occupation text	avg_sleep numeric
1	Sales Representati...	5.90
2	Scientist	6.00
3	Salesperson	6.40
4	Teacher	6.69
5	Software Engineer	6.75
6	Manager	6.90
7	Doctor	6.97
8	Nurse	7.06
9	Accountant	7.11
10	Lawyer	7.41
11	Engineer	7.99

Insights:

- Sales Representatives sleep the least (5.9 hrs).
- They may face burnout due to long hours or travel.
- Company can offer flexible hours or sleep awareness sessions.

2. Which age groups are most affected by sleep disorders?

```
-- 2. which age groups are most affected by sleep disorders.  
select case  
    when age between 20 and 29 then '20-29'  
    when age between 30 and 39 then '30-39'  
    when age between 40 and 49 then '40-49'  
    else '50+'  
end as age_group,  
count(s.sleep_disorder) as disorder_count  
from Person p  
join sleep_details s on p.person_id=s.person_id  
where s.sleep_disorder is not null and s.sleep_disorder!='None'  
group by age_group  
order by age_group desc;
```

	age_group	disorder_count
1	50+	58
2	40-49	75
3	30-39	15
4	20-29	7

Insights:

- Age 40–49 has the most sleep disorder cases (75).
- Mid-career employees may need counseling or awareness programs.
- Helps prevent health problems and improve attendance.

3. Which BMI category has the lowest average quality of sleep?

```
--3. which BMI category has lowest average quality of sleep ?  
select h.bmi_category, round(avg(s.quality_of_sleep),2) as avg_quality  
from sleep_details s  
join health_info h on s.person_id = h.person_id  
group by h.bmi_category  
order by avg_quality desc;
```

	bmi_category	avg_quality
1	Normal	7.66
2	Normal Weight	7.43
3	Overweight	6.90
4	Obese	6.40

Insights:

- Obese employees have the worst sleep quality (avg 6.4).
- Company can add fitness and nutrition support.
- Gym or wellness programs can improve both sleep and health.

Category 2: Stress and Activity

4. Are males or females more stressed?

--- 4. Are males or females more stressed?

```
select p.gender,round(avg(h.stress_level),2) as avg_stress  
from Person p  
join health_info h on p.person_id=h.person_id  
group by p.gender  
order by avg_stress desc;
```

	gender 	avg_stress 
text	numeric	
1	Male	6.08
2	Female	4.68

Insights:

- Males show higher stress (6.08 vs 4.68).
- Offer stress management and mental health programs for men.
- Track stress trends regularly for better support.

5. Do male and female employees differ in their average physical activity levels?

--- 5. Do male and female employees differ in their average physical activity levels?

```
select p.Gender, round(avg(h.Physical_Activity), 2) as avg_activity_level  
from health_info h  
join Person p on h.Person_ID = p.Person_ID  
group by p.Gender  
order by avg_activity_level desc;
```

	gender 	avg_activity_level 
text	numeric	
1	Male	59.20
2	Female	59.14

Insights:

- Both are almost equal (Male 59.2, Female 59.1).
- Continue balanced fitness challenges for everyone.
- Encourage group activities to keep engagement high

6. What is the highest stress level in each occupation?

--- 6. what is the highest stress level in each occupation ?

```
select p.occupation,max(h.stress_level)
from person p
join health_info h on p.person_id=h.person_id
group by p.occupation;
```

	occupation text	max integer
1	Sales Representati...	8
2	Engineer	7
3	Doctor	8
4	Salesperson	7
5	Manager	5
6	Teacher	7
7	Software Engineer	8
8	Accountant	7
9	Lawyer	6
10	Nurse	8
11	Scientist	8

Insights:

- Doctors, Nurses, and Software Engineers have max stress (8).
- Need workload checks and relaxation programs.
- Focus counseling efforts on high-stress jobs.

7. Which occupation has the highest average daily steps?

```
--- 7. which occupation has highest average daily steps ?  
select p.occupation,round(avg(h.Daily_Steps), 2) as avg_daily_steps  
from health_info h  
join person p on p.person_id=h.person_id  
group by p.occupation  
order by avg_daily_steps desc;
```

	occupation text	avg_daily_steps numeric
1	Nurse	8057.53
2	Lawyer	7661.70
3	Accountant	6881.08
4	Doctor	6808.45
5	Salesperson	6000.00
6	Engineer	5980.95
7	Teacher	5957.50
8	Software Engineer	5800.00
9	Manager	5500.00
10	Scientist	5350.00
11	Sales Representati...	3000.00

Insights:

- Nurses walk the most (8057 steps).
- Use them to inspire others in step challenges.
- Encourage less active jobs to move more at work.

Category 3: Health Risks

8. Find people with above average heart rate

```
--- 8. find people with above average heart rate
select person_id, heart_rate
from health_info
where heart_rate > (select avg(heart_rate) from health_info);
```

	person_id integer	heart_rate integer
1	1	77
2	2	75
3	3	75
4	4	85
5	5	85
6	6	85
7	7	82
8	17	80
9	19	80
10	31	78
11	32	78
12	34	72

Insights:

- Flags 133 individuals (36% of workforce) for cardiovascular monitoring, with peaks at 86 bpm, enabling personalized alerts.
- Informs routine check-ups for elevated rates to prevent stress-related incidents.
- Guides wearable integrations for real-time tracking and early interventions in high-risk groups.

9. Who all are at high risk due to obesity, low sleep or low physical activity?

```
-- 10. who all are at high risk due to obesity,low sleep quality or low physical activity?  
select p.Person_id, p.Occupation, h.BMI_Category, s.sleep_duration, h.Physical_activity,  
case when h.BMI_category='Obese'  
      or s.sleep_duration < 6  
      or h.physical_activity < 50 then 'High Risk'  
      else 'Low Risk' end as risk_category  
from person p  
join sleep_details s on p.person_id=s.person_id  
join health_info h on s.person_id=h.person_id;
```

	person_id integer	occupation text	bmi_category text	sleep_duration numeric	physical_activity integer	risk_category text
1	1	Software Engineer	Overweight	6.1	42	High Risk
2	2	Doctor	Normal	6.2	60	Low Risk
3	3	Doctor	Normal	6.2	60	Low Risk
4	4	Sales Representati...	Obese	5.9	30	High Risk
5	5	Sales Representati...	Obese	5.9	30	High Risk
6	6	Software Engineer	Obese	5.9	30	High Risk
7	7	Teacher	Obese	6.3	40	High Risk
8	8	Doctor	Normal	7.8	75	Low Risk
9	9	Doctor	Normal	7.8	75	Low Risk
10	10	Doctor	Normal	7.8	75	Low Risk
11	11	Doctor	Normal	6.1	30	High Risk
12	12	Doctor	Normal	7.8	75	Low Risk
13	13	Doctor	Normal	6.1	30	High Risk
14	14	Doctor	Normal	6	30	High Risk
15	15	Doctor	Normal	6	30	High Risk

Insights:

- 158 high-risk employees found (mostly Nurses & Doctors).
- Focus on improving diet, sleep, and exercise.
- Plan regular wellness checks for them.

RECOMMENDATIONS

- Target high-stress males and roles like Doctors/Nurses with mindfulness apps and flexible scheduling.
- Launch inclusive activity programs bridging minor gender gaps, emphasizing Nurses as step leaders.
- Prioritize sleep training for 40-49 age group and obese employees via subsidized aids.
- Monitor 133 elevated heart rate cases with annual screenings and wearables.
- Address under-sleep in Nurses/Doctors (70 combined cases) with rest pods and audits.

CONCLUSION

This SQL-based analysis of the employee health and sleep survey provides deep insights into patterns, risks, and opportunities for intervention across 374 respondents. By normalizing the dataset and executing relational analysis, the project enables the company to shift from assumption-based wellness to data-driven decision-making.

These insights will help optimize programs, resource allocation, and employee support, fostering a healthier, more productive workforce in a demanding professional environment.