Analysis of Sleep Health and Lifestyle Habits

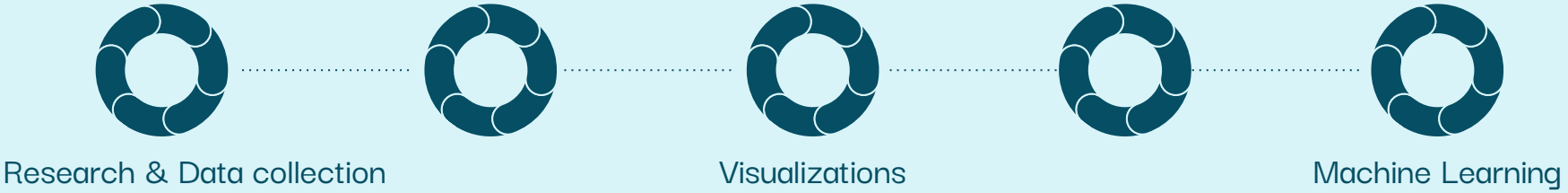
Réda RERIZANI

My goal is

- Analyze trends in the quality of sleep across individuals' lifestyles.
- Identify key indicators of sleep quality and duration.
- Create a prediction model for stress levels based on available data.

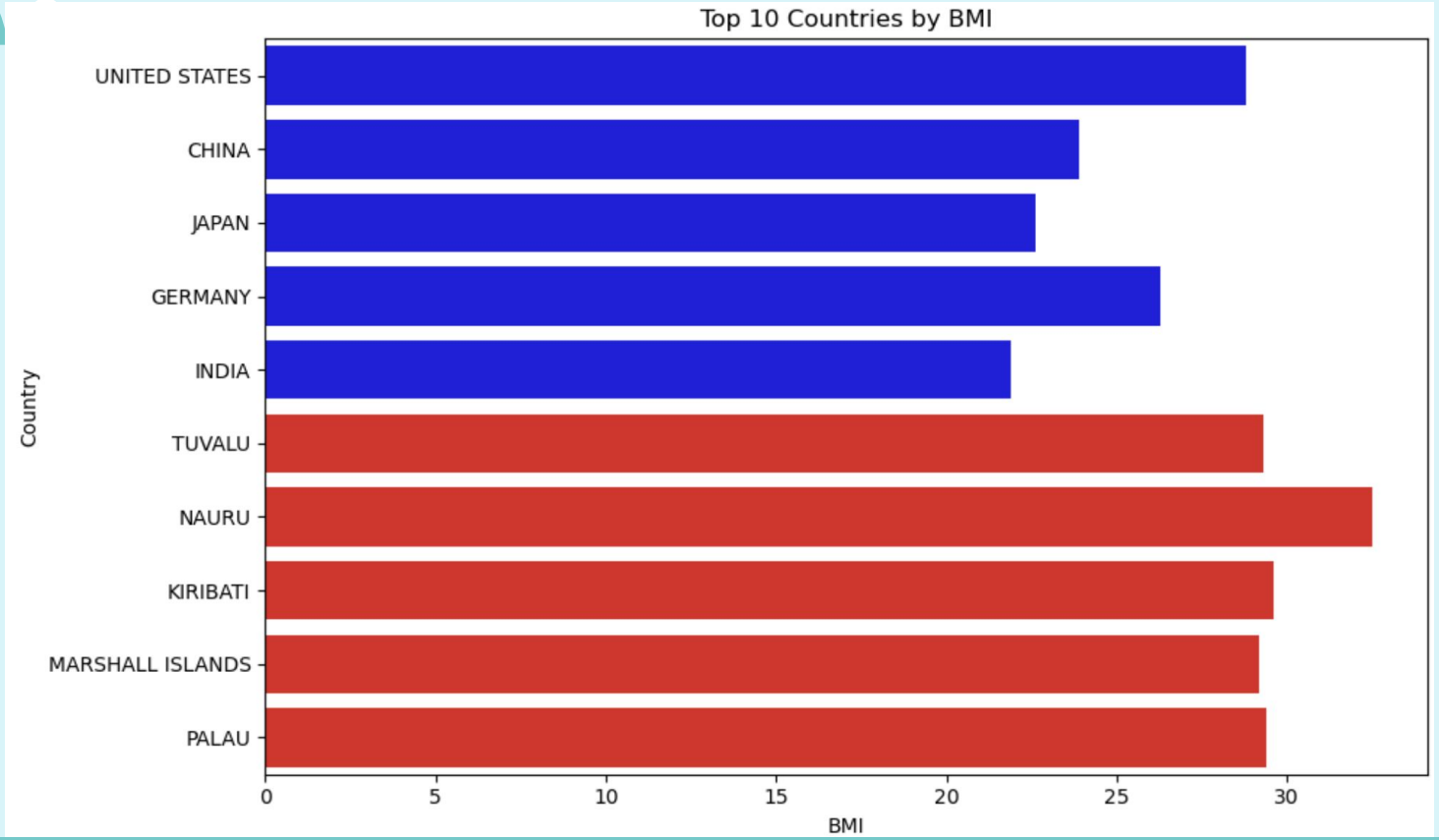
Project Management

On Trello :



Data gathering

- **Flat flies:** Two datasets “Sleep Health and Lifestyle” from Kaggle
- **API:** GDP by country from World Bank Data API
- **Web scraping:** Two datasets :
  1. Body mass index (BMI) categorie
  2. Lists countries by their average BMI



Note: We cannot observe any significant relationship between the GDP of rich and poor countries and BMI (Body Mass Index). Therefore, for the remainder of my data analysis, I will not take them into consideration

Dataset Overview

- ❏ Person ID
- ❏ Gender
- ❏ Age
- ❏ Occupation/profession of the person.

Personal information

- ❏ Sleep Duration (hours)
- ❏ Quality of Sleep (scale: 1-10)
- ❏ Sleep Disorder

Sleep information

- ❏ Physical Activity Level (minutes/day)
- ❏ Stress Level (scale: 1-10)
- ❏ BMI Category: The BMI category of the person
- ❏ Blood Pressure (systolic/diastolic)
- ❏ Heart Rate (bpm)
- ❏ Daily Steps

factors influencing sleep patterns.

Data cleaning

- Merge main tables into a single dataframe.
- Checking for data shapes & data types.
- Column Formatting (lowercase, removing special characters...)
- Handle Null Values: Address null values using **"isna.sum()"** method.
- **Modifications:**

  _ Fill null values in **"sleep_disorder"** column with "**no disorder"**

  _ Classify "blood_pressure" column: convert les valeurs **"126/83"** to categorical variable **(hypotension ,Normal, elevated and hypertension)**

Final Sleep Dataframe After Cleaning

person_id	gender	age	occupation	sleep_duration	quality_of_sleep	physical_activity_level	stress_level	id_bmi	heart_rate	daily_steps	sleep_disorder	blood_pressure_class
1	Male	27	Software Engineer	6.1	6	42	6	5	77	4200	no disorder	Normal
2	Male	28	Doctor	6.2	6	60	8	4	75	10000	no disorder	Normal
3	Male	28	Doctor	6.2	6	60	8	4	75	10000	no disorder	Normal
4	Male	28	Sales Representative	5.9	4	30	8	6	85	3000	Sleep Apnea	Hypertension
5	Male	28	Sales Representative	5.9	4	30	8	6	85	3000	Sleep Apnea	Hypertension
6	Male	28	Software Engineer	5.9	4	30	8	6	85	3000	Insomnia	Hypertension
7	Male	29	Teacher	6.3	6	40	7	6	82	3500	Insomnia	Hypertension
8	Male	29	Doctor	7.8	7	75	6	4	70	8000	no disorder	Hypotension
9	Male	29	Doctor	7.8	7	75	6	4	70	8000	no disorder	Hypotension
10	Male	29	Doctor	7.8	7	75	6	4	70	8000	no disorder	Hypotension
11	Male	29	Doctor	6.1	6	30	8	4	70	8000	no disorder	Hypotension
12	Male	29	Doctor	7.8	7	75	6	4	70	8000	no disorder	Hypotension
13	Male	29	Doctor	6.1	6	30	8	4	70	8000	no disorder	Hypotension
14	Male	29	Doctor	6.0	6	30	8	4	70	8000	no disorder	Hypotension
15	Male	29	Doctor	6.0	6	30	8	4	70	8000	no disorder	Hypotension
16	Male	29	Doctor	6.0	6	30	8	4	70	8000	no disorder	Hypotension
17	Female	29	Nurse	6.5	5	40	7	4	80	4000	Sleep Apnea	Elevated

EDA & Visualization

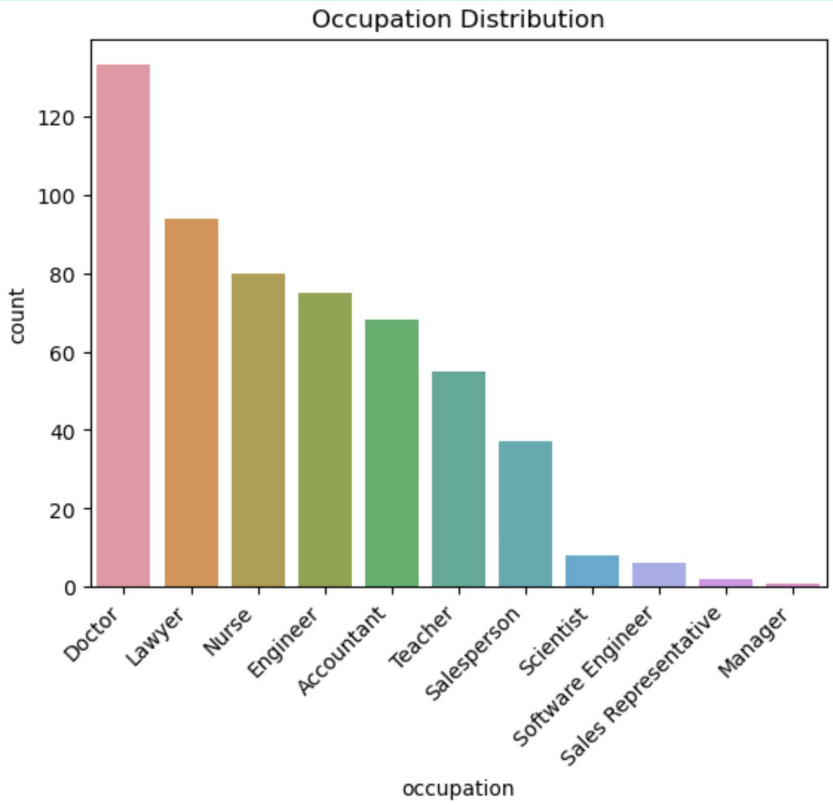
Identify categorical and numerical variables

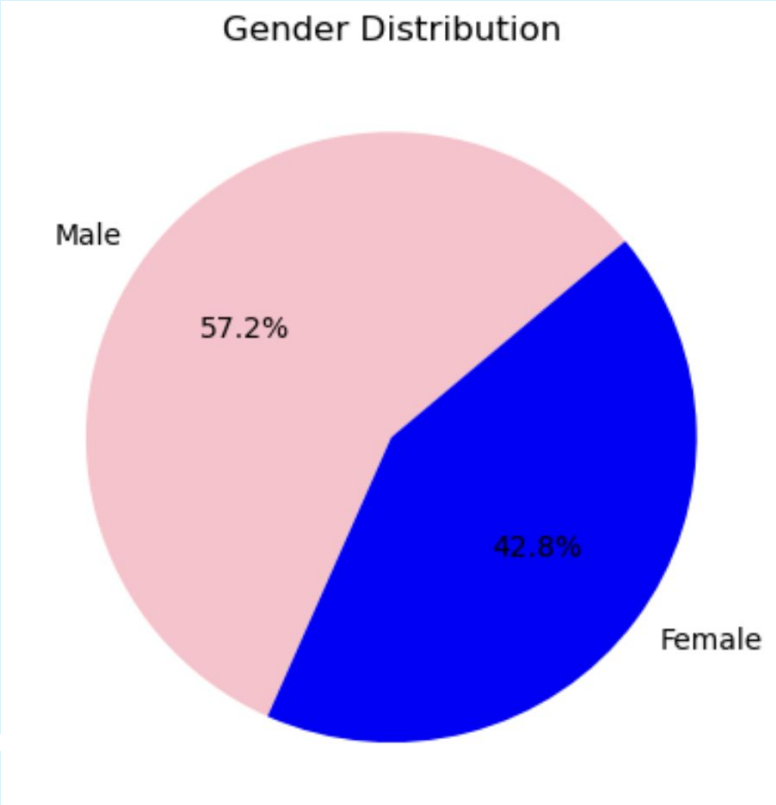
- Categorical values: gender, occupation, sleep disorder, blood pressure, bmi

- Numérical variable: age, sleep duration, quality of sleep, physical activity level, stress level, heart rate, daily steps

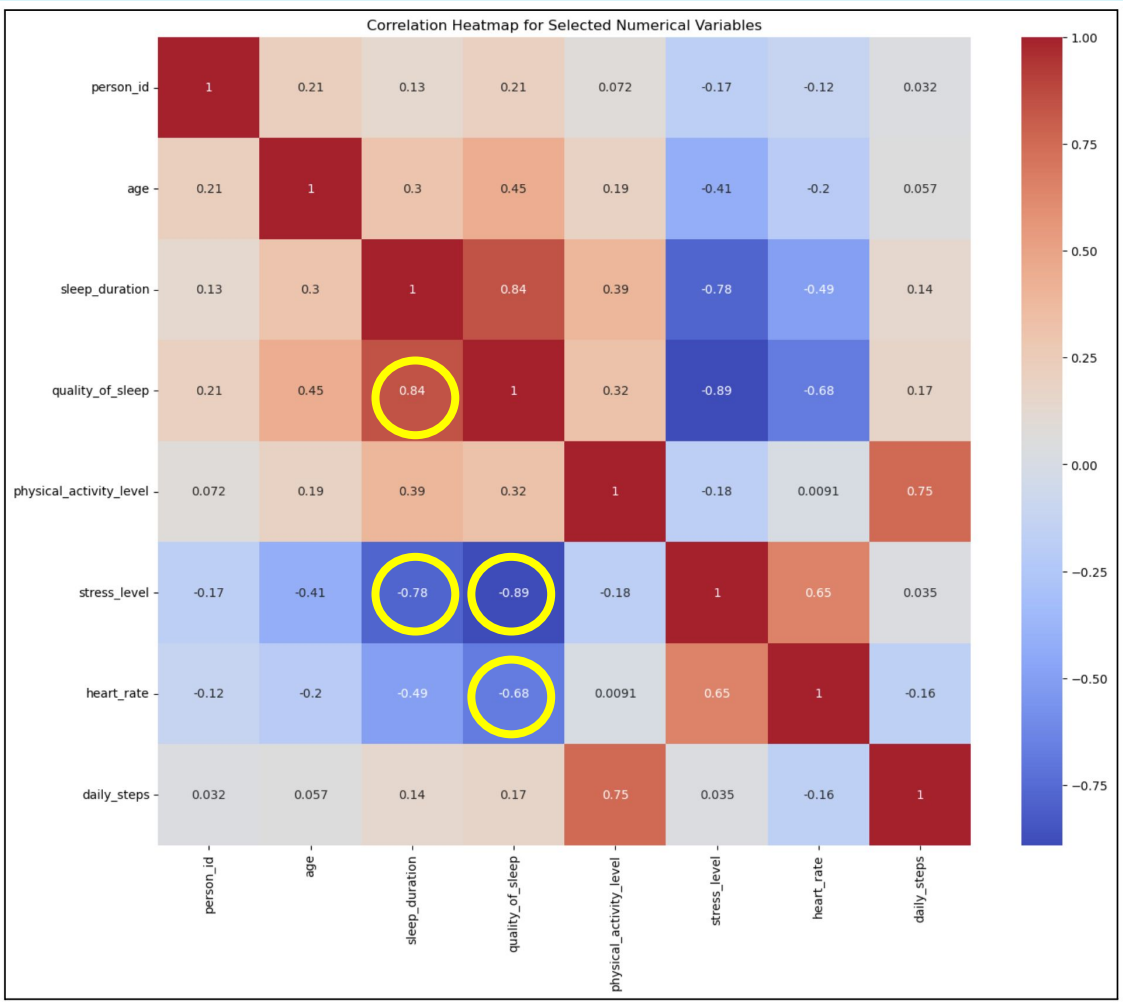
EDA & Visualization

Distribution

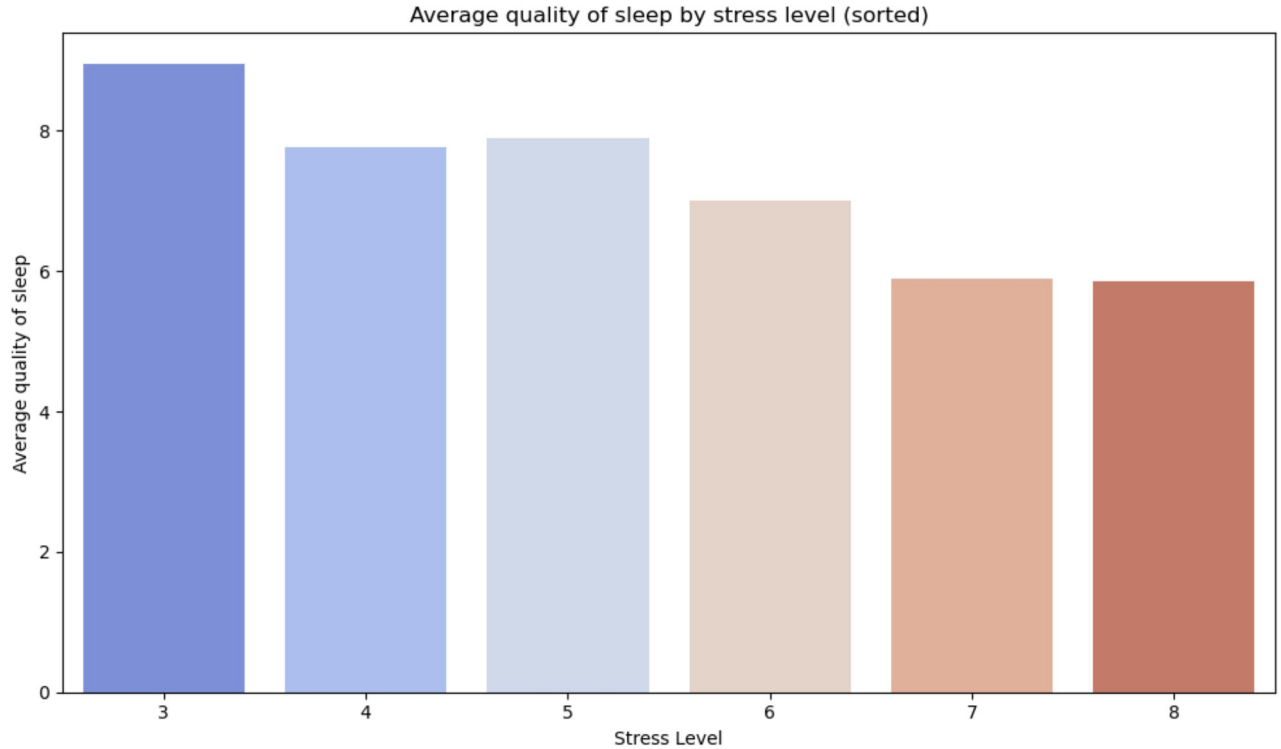




Correlation

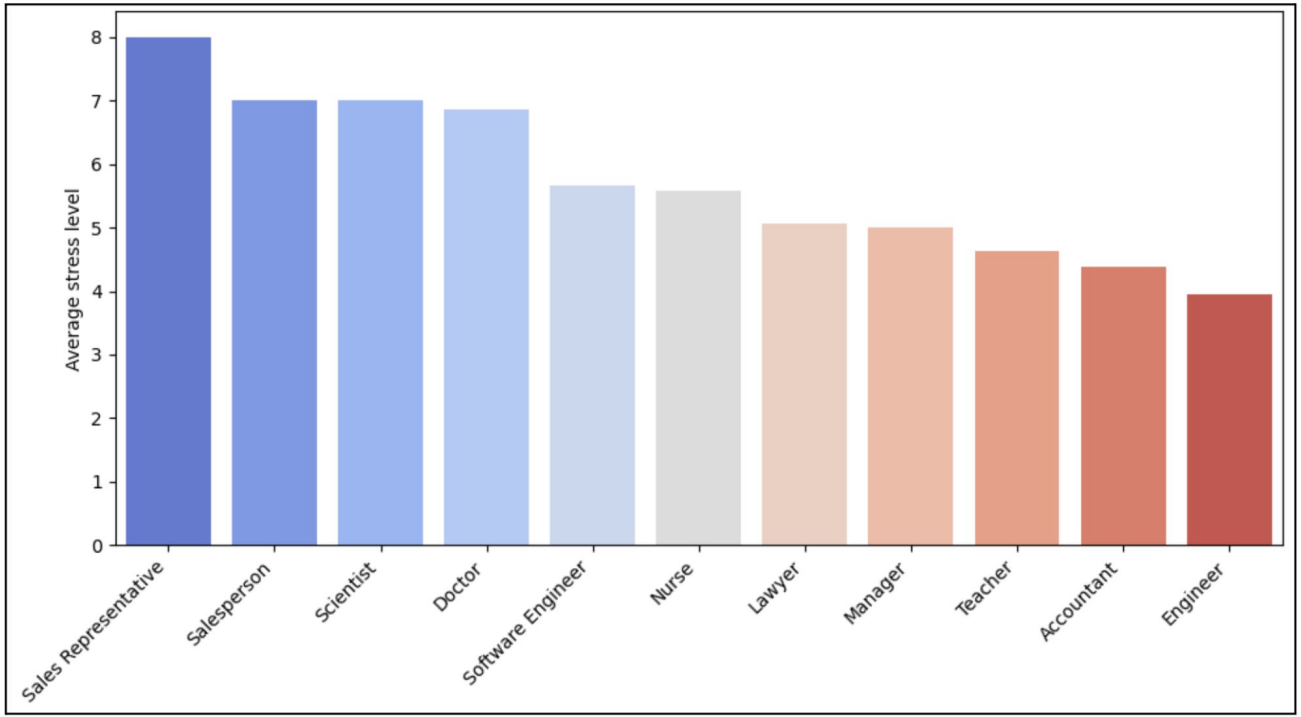


EDA & Visualization

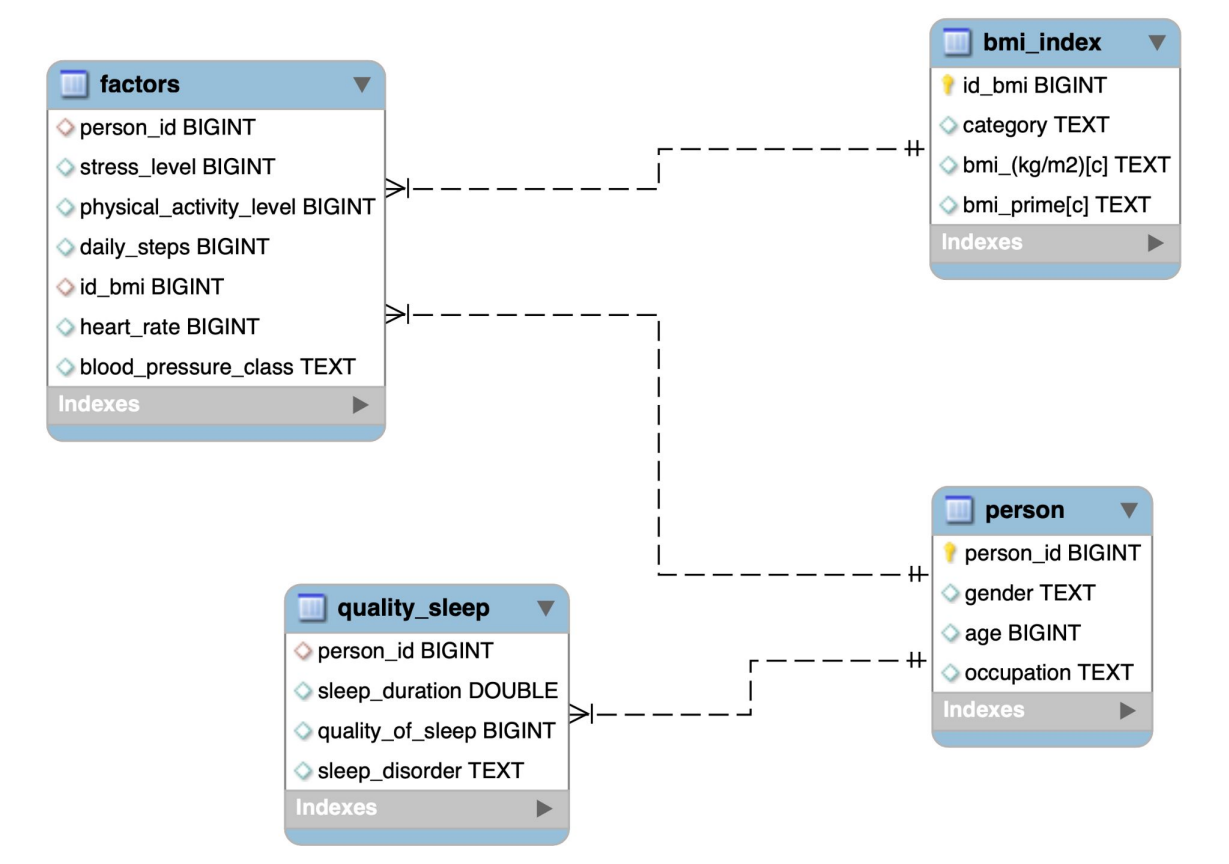


EDA & Visualization

- **Average stress level by profession**



Entity Relationship Diagram(ERD)



Examples of queries in MySQL

1. Create view quality and hour of sleep according to occupation and stress level

```
41 •  CREATE VIEW quality_sleep_by_occupation_and_stress_level AS
42    SELECT p.occupation, ROUND(AVG(f.stress_level),2) as avg_stress_level, ROUND(AVG(qs.quality_of_sleep),2)
      AS quality_sleep, ROUND(AVG(qs.sleep_duration), 2) AS sleep_duration
43    FROM person p
44    JOIN quality_sleep qs ON p.person_id = qs.person_id
45    JOIN factors f ON p.person_id = f.person_id
46    GROUP BY p.occupation;
47
```

100% 44:45

Result Grid | Filter Rows: Search | Export:

occupation	avg_stress_le...	quality_sleep	sleep_duration
Software Engineer	5.67	7.00	7
Doctor	6.86	6.58	6.94
Sales Representative	8.00	4.00	5.9
Teacher	4.64	6.98	6.74
Nurse	5.59	7.25	7.04
Engineer	3.95	8.31	7.88
Accountant	4.38	7.97	7.16
Scientist	7.00	5.00	6
Lawyer	5.06	7.89	7.41
Salesperson	7.00	6.00	6.41
Manager	5.00	7.00	6.9

Result Grid | Form Editor | Field Types

Examples of queries in MySQL

2. The average quality of sleep and sleep duration for doctors who have a stress level less than 5

```
71 •  SELECT p.occupation, ROUND(AVG(qs.quality_of_sleep), 2) AS avg_quality_of_sleep, ROUND(AVG(qs.
      sleep_duration), 2) AS sleep_duration
72    FROM person p
73    JOIN quality_sleep qs ON p.person_id = qs.person_id
74    JOIN factors f ON p.person_id = f.person_id
75    WHERE p.occupation = 'Doctor' AND f.stress_level < 5
76    GROUP BY p.occupation;
77
78
```

100% 1:77

Result Grid Filter Rows: Search Export:

occupation	avg_quality_of_sleep	sleep_duration
Doctor	9.00	8.15

Flask API

- **Ressources from MySql query**
- **5 Endpoints :**

  /person?include_details=0 or 1

  /person/<int:person_id>

  /quality_sleep

  /quality_sleep/bmi

  /quality_sleep/occupation

Flask API Documentation

Sample /pe


```
{
  "last_page": "/person?page=19&page_size=30&include_details=1",
  "next_page": "/person?page=1&page_size=30&include_details=1",
  "person": [
    {
      "age": 27,
      "gender": "Male",
      "occupation": "Software Engineer",
      "person_id": 1,
      "quality_sleep": [
        {
          "quality_of_sleep": 6,
          "sleep_disorder": "no disorder",
          "sleep_duration": 6.1
        }
      ]
    },
    {
      "age": 28,
      "gender": "Male",
      "occupation": "Doctor",
      "person_id": 2,
      "quality_sleep": [
        {
          "quality_of_sleep": 6,
          "sleep_disorder": "no disorder",
          "sleep_duration": 6.2
        }
      ]
    },
    {
```

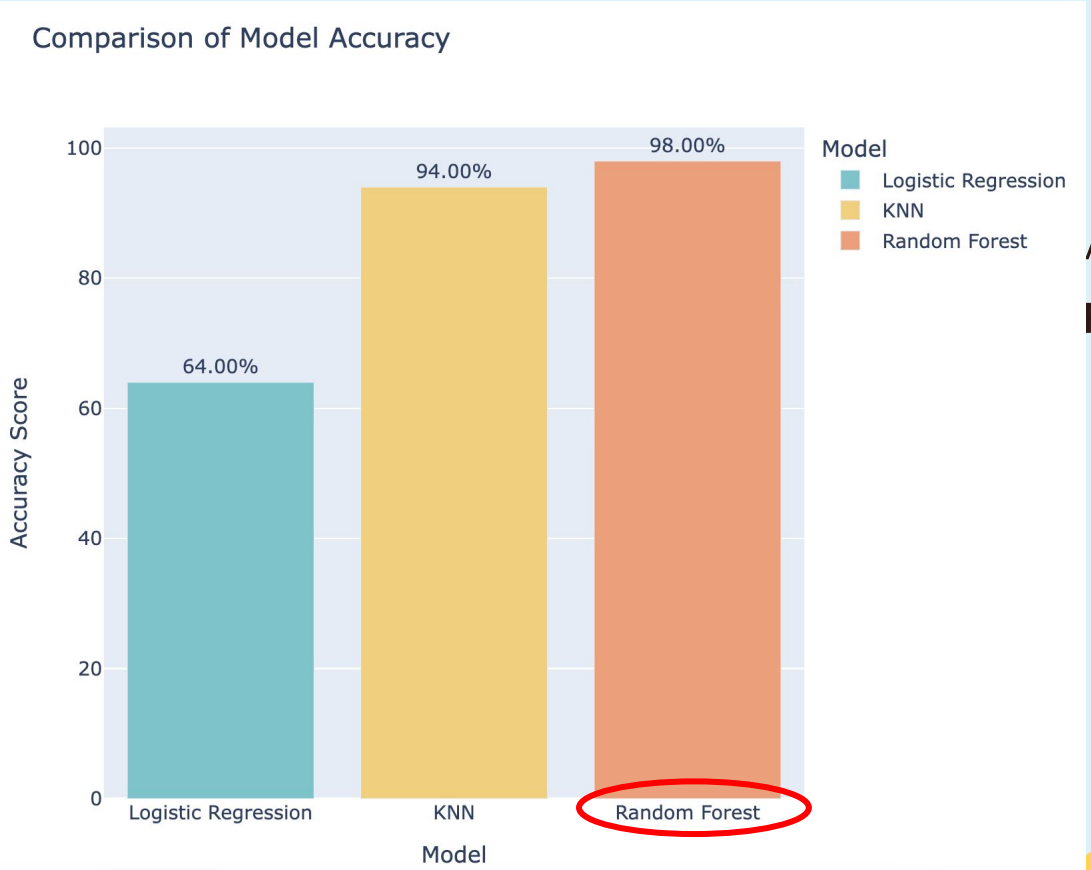
Machine learning

Steps in C

- ❑ De
- ❑ Do
- ❑ Try
- ❑ Ch

variable.

Forest'.



Main Results

- There's a strong connection between how long you sleep and how good your sleep is.
- People with healthier hearts tend to sleep better
- when people are more stressed, their sleep tends to be worse and shorter. This shows how stress can really mess with your sleep.

Highlight

Managing stress, staying active, and caring for our heart health can all improve our sleep. It's not just about bedtime; it's about making healthy choices all day for better rest.

NEXT STEP

As companies in the nutritional supplement industry, this information can help us create products that improve sleep by focusing on stress reduction

Challenges

- Find the right datasets
- Struggling to find the right API sources to gather data that matches my study.
- Insufficient data. I would prefer more extensive datasets to conduct a comprehensive analysis.
- Write the report and prepare the presentation in English

THANK YOU