

## پروژه پایانی درس داده کاوی

موضوع: تشخیص الگو های خواب برای بهبود سلامت روان

دیتاست: [Sleep Health and Lifestyle Dataset](#)

1. درباره دیتاست:

دیتا شامل 374 سطر و 13 ستون می باشد.

اطلاعات ستون ها:

Person ID: یک شناسه برای هر فرد

Gender: جنسیت فرد (مرد/زن)

Age: سن هر فرد بر اساس سال

Occupation: شغل یا حرفه شخص

Sleep Duration: مدت زمان خواب (بر اساس ساعت)، مدت زمانی که فرد در روز میخوابد

Quality of Sleep: (مقیاس: 1-10) یک رتبه بندی از کیفیت خواب

Physical Activity: سطح فعالیت بدنی (مقیاس: روز/دقیقه) تعداد دقایقی که فرد روزانه به فعالیت بدنی می پردازد.

Stress Level: (مقیاس: 1-10) رتبه بندی از سطح استرس تجربه شده توسط فرد.

BMI Category: دسته BMI برای هر فرد (به عنوان مثال: کم وزن، نرمال، اضافه وزن)

Blood Pressure: (مقیاس: systolic/diastolic) اندازه گیری فشار خون هر فرد که به صورت فشار سیستولیک بر فشار دیاستولیک نشان داده میشود.

Heart Rate: (مقیاس bpm) ضربان قلب هر فرد در حالت استراحت بر حسب ضربان در دقیقه.

Daily Steps: تعداد قدم هایی که فرد در روز بر می دارد.

Sleep Disorder: وجود یا عدم وجود اختلال خواب در فرد (هیچکدام، بی خوابی، آپنه).

نمایش اطلاعات ستون ها:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 374 entries, 0 to 373
Data columns (total 13 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   Person ID                            374 non-null    int64
1   Gender                               374 non-null    object
2   Age                                  374 non-null    int64
3   Occupation                           374 non-null    object
4   Sleep Duration                       374 non-null    float64
5   Quality of Sleep                     374 non-null    int64
6   Physical Activity Level               374 non-null    int64
7   Stress Level                         374 non-null    int64
8   BMI Category                         374 non-null    object
9   Blood Pressure                       374 non-null    object
10  Heart Rate                           374 non-null    int64
11  Daily Steps                          374 non-null    int64
12  Sleep Disorder                       155 non-null    object
dtypes: float64(1), int64(7), object(5)
memory usage: 38.1+ KB
```

نمایش داده:

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Blood Pressure	Heart Rate	Daily Steps	Sleep Disorder
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	126/83	77	4200	NaN
1	2	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	NaN
2	3	Male	28	Doctor	6.2	6	60	8	Normal	125/80	75	10000	NaN
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	140/90	85	3000	Sleep Apnea
...	...	...	...	...	...	...	...	...	...	...	...	...	...
369	370	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
370	371	Female	59	Nurse	8.0	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
371	372	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
372	373	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea
373	374	Female	59	Nurse	8.1	9	75	3	Overweight	140/95	68	7000	Sleep Apnea

374 rows × 13 columns

بررسی متغیر های متنی داده:

	Gender	Occupation	BMI Category	Blood Pressure	Sleep Disorder
count	374	374	374	374	155
unique	2	11	4	25	2
top	Male	Nurse	Normal	130/85	Sleep Apnea
freq	189	73	195	99	78

بررسی متغیر های عددی داده:

	count	mean	std	min	25%	50%	75%	max
Person ID	374.0	187.500000	108.108742	1.0	94.25	187.5	280.75	374.0
Age	374.0	42.184492	8.673133	27.0	35.25	43.0	50.00	59.0
Sleep Duration	374.0	7.132086	0.795657	5.8	6.40	7.2	7.80	8.5
Quality of Sleep	374.0	7.312834	1.196956	4.0	6.00	7.0	8.00	9.0
Physical Activity Level	374.0	59.171123	20.830804	30.0	45.00	60.0	75.00	90.0
Stress Level	374.0	5.385027	1.774526	3.0	4.00	5.0	7.00	8.0
Heart Rate	374.0	70.165775	4.135676	65.0	68.00	70.0	72.00	86.0
Daily Steps	374.0	6816.844920	1617.915679	3000.0	5600.00	7000.0	8000.00	10000.0

## 2. پردازش داده و آماده سازی آن برای تحلیل

1. اولین گام برای پردازش داده ها این است که داده های غیر عادی و دارای نویز را کشف کنیم. به همین منظور از تابعی استفاده کرده ایم که میزان که فرکانس هر مقدار منحصر به فرد را در یک ستون می شمارد.

بخشی از خروجی این تابع به شکل زیر است:

```

Person ID      -----
374  1          Occupation
1    1          Nurse
2    1          Doctor
3    1          Engineer
4    1          Lawyer
..    ..
11   1          Teacher
10   1          Accountant
9    1          Salesperson
8    1          Scientist
7    1          Software Engineer
Name: count, Length: 374, dtype: int64
-----
Gender
Male      189
Female    185
Name: count, dtype: int64
-----
Age
43  34
44  30
37  20
38  20
50  20
31  18
32  17
53  17
59  16
39  15

2.2  36
6.0  31
7.8  28
6.5  26
6.1  25
7.7  24
6.6  20
7.1  19
8.1  15
7.3  14

Occupation
Nurse
Doctor
Engineer
Lawyer
Teacher
Accountant
Salesperson
Scientist
Software Engineer
Sales Representative
Manager
Name: count, dtype: int64
-----
Sleep Duration
7.2  36
6.0  31
7.8  28
6.5  26
6.1  25
7.7  24
6.6  20
7.1  19
8.1  15
7.3  14

Stress Level
3  71
8  70
4  70
5  67
7  50
6  46
Name: count, dtype: int64
-----
BMI Category
Normal      195
Overweight  148
Normal Weight  21
Obese       10
Name: count, dtype: int64
-----
Blood Pressure
130/85  99
125/80  65
120/80  45
115/75  32
135/90  27
140/90  4
125/82  4

Quality of Sleep
8  109
6  105
7  77
9  71
5  7
4  5
Name: count, dtype: int64
-----
Physical Activity Level
60  70
30  68
45  68
75  67
90  67
40  6
55  6
35  4
50  4
70  3

```

در این مرحله نکته جالب توجهی که یافته شد به این شکل است که مقدار نتایج برای ستون BMI به این شکل می باشد:

```
BMI Category
Normal      195
Overweight  148
Normal Weight  21
Obese       10
Name: count, dtype: int64
```

با توجه به شکل بالا برای BMI دو لیبل با نام های Normal, Normal Weight ثبت شده است که این دو اسم در معنا یکی هستند منتها مقداری که نشان میدهند معلوم نمی‌کند که دقیقا این دو لیبل نشان دهنده چه چیزی هستند، برای این منظور احتمالاتی می‌تواند وجود داشته باشد مثل اینکه داده ها از دو دیتا ست متفاوت گردآوری شده اند و یا اینکه یکی از اینها به معنی Underweight می باشد که به اشتباه با لیبل های Normal مشخص شده است. از طرفی ب دلیل اینکه مقدار عددی برای BMI نداریم نمی‌توانیم که تحلیل آماری بر روی داده ها با برچسب های Normal, Normal weight انجام دهیم به همین منظور در گام اول مقادیر Normal Weight را با مقدار Normal جایگزین میکنیم.

2. در مرحله بعدی لازم است که بررسی شود آیا ستون ها مقدار null دارند یا خیر برای این منظور باید از تابع isnull استفاده شود، که مقدار null را با TRUE جایگزین میکند و مقدار غیر NULL را با False. سپس یک loop ایجاد میکنیم که در هر سطر چرخش کند و تعداد TRUE و False را برای هر ستون بدست آورد. با توجه به نتایج بدست آمده صرفا در ستون Sleep Disorder مقدار Null به تعداد 219 وجود دارد. اما این ستون مهم است چون ستون target می‌باشد پس نمیتوان ان را حذف کرد پس به جای حذف آن، مقادیر Null را برابر با Normal قرار میدهیم که یعنی بیمار اختلال خواب ندارد.

```
Sleep Disorder
True      219
False     155
Name: count, dtype: int64
```

3. یکی از مراحل preprocessing این است که داده هایی که فرمت string دارند تبدیل به عدد شوند تا کار با انها راحت تر باشد(مثلا برای استفاده در شبکه های عصبی و یا ساخت درخت تصمیم لازم است چنین کاری صورت گیرد) به همین منظور از روی داده ها clean شده یک کپی میگیریم و سپس ستون هایی را که string، datatype دارند را به فرمت numeric تبدیل میکنیم. مثلا male تبدیل به صفر و female تبدیل به یک خواهد شد.

### 3. تحلیل داده

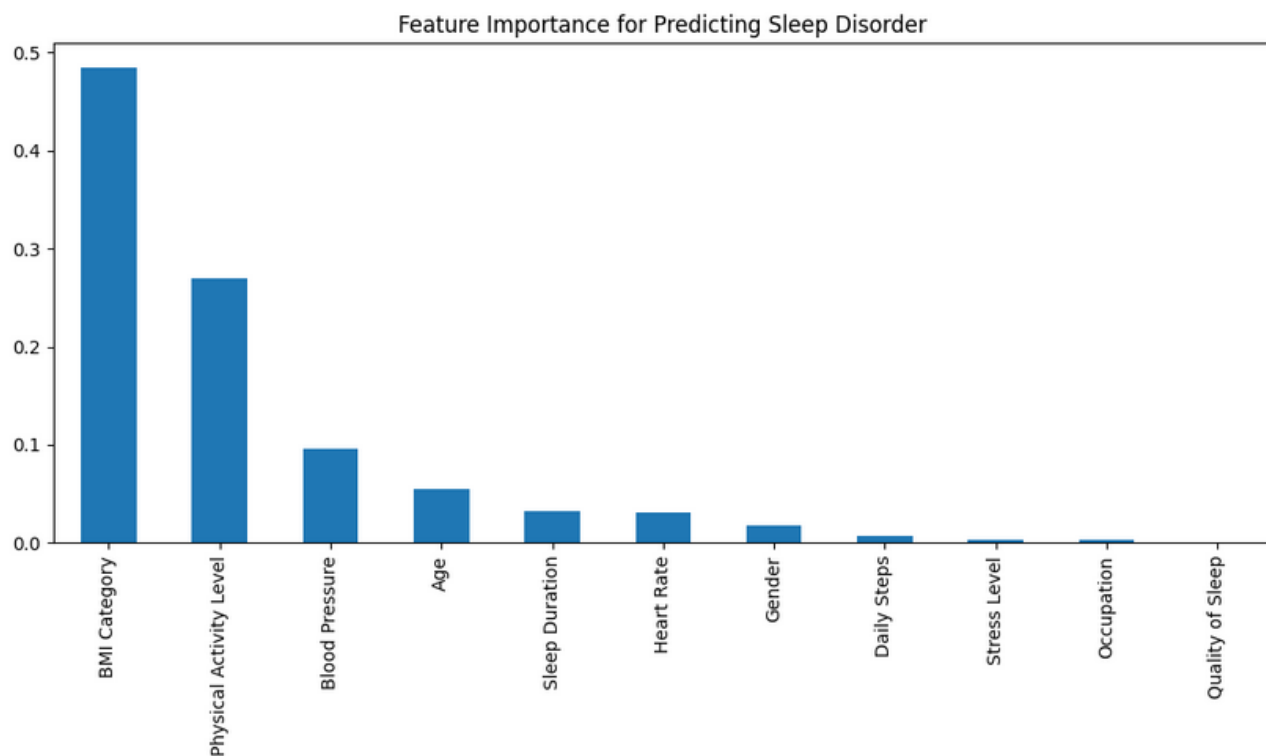
1. در ابتدا از درخت تصمیم استفاده میکنیم برای اینکه بدست آوریم کدام یکی از feature ها میتواند در ساخت درخت تصمیم به ما کمک کند .

این کار می‌تواند از جهات بسیاری به ما کمک کند، به عنوان مثال میتوان فهمید که کدام یکی از feature ها بر روی قدرت پیش بینی مدل تاثیر بیشتری دارد و به همین جهت می‌توان داده هایی با اهمیت بیشتر را پیدا کرد و ارتباط آنها را بر روی target بدست آورد.

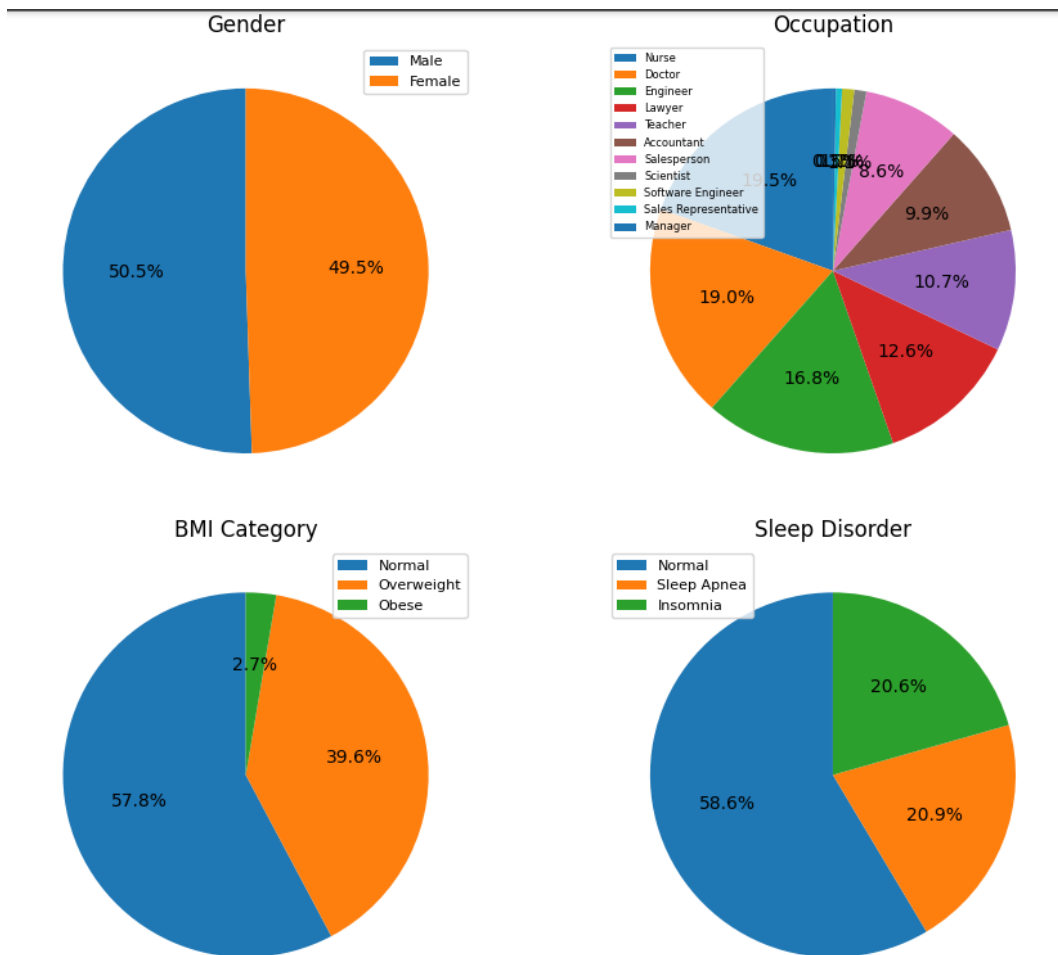
به همین منظور پس از فراخوانی کتابخانه های مورد نیاز همچون sklearn باید ویژگی هایی همچون Person ID حذف شود(چون منجر به overfitting درخت خواهد شد) و سپس داده ها به دو مجموعه train, test تقسیم شوند

و مدل درخت تصمیم بر روی داده های **train** آموزش داده شود و بر روی داده های تست، ارزیابی شود. سپس **feature** های مهم درخت تصمیم استخراج میشود.

	Feature	Importance
0	BMI Category	0.485229
1	Physical Activity Level	0.269911
2	Blood Pressure	0.096091
3	Age	0.054532
4	Sleep Duration	0.032708
5	Heart Rate	0.030960
6	Gender	0.017431
7	Daily Steps	0.006855
8	Stress Level	0.003418
9	Occupation	0.002867
10	Quality of Sleep	0.000000



در این قسمت با استفاده از تکنیک های **visualization** تلاش میکنیم تا توزیع داده ها را بدست آوریم.  
برای این کار ، برای داده های که تایپ **string** هستند از نمودار دایره ای استفاده میکنیم:



- همان طور که از شکل بالا پیدا است بین ویژگی های، **Gender** از توزیع متوازن تری برخوردار است. پس کار بر روی این داده می تواند نتایج قابل اتکا تری به ما بدهد.
- از طرفی نمودار **Occupation** بسیار نامتوازن می باشد و هنگامی که تعداد هر **category** در این ستون را بدست آوردیم دریافتیم که شغل های **software engineer, scientist, sales representative, manager** کمترین تعداد را دارند در نتیجه شاید اتکا بر این توزیع مارا دچار گمراهی کند و در حین ساختن مدل هم باید برای هر دسته یک دسته بند متفاوت انتخاب کند که ممکن است با **overfitting** مواجه شویم.

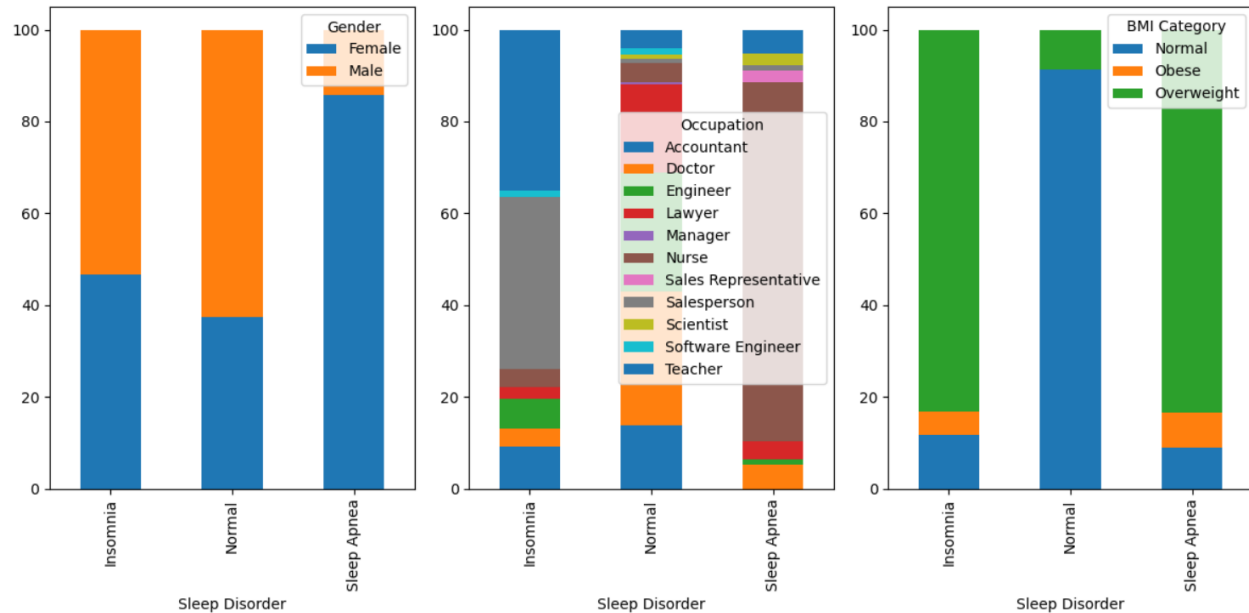
Software Engineer	4
Scientist	4
Sales Representative	2
Manager	1

- داده های **BMI** هم توزیع نامتناسب دارند اما با این حال توازن داده بین **Normal, Overweight** بیشتر است.

3. در این مرحله قصد داریم که ارتباط هریک از **feature** ها را با **target** متوجه شویم، در اصل می خواهیم متوجه شویم که هر گروه از **sleep disorder** از نظر ویژگی ها چطور می توانند از هم متمایز شوند.

پس تحلیل خود را از داده های به فرم **caegorical** شروع میکنیم. و نسبت توزیع **category** های مختلف در هر کدام از سه ستون (**Gender, Occupation, BMI Category**) در انواع مختلف **sleep disorder** بدست می آوریم.

خروجی در نهایت به این شکل خواهد شد:



پس از نمایش میتوان از شکل توزیع جنسیتی مشاهده کرد:

- مردان بیشتر از زنان از بی خوابی (Insomnia) رنج می‌برند.
- همچنین به نظر می‌رسد که مردان الگوهای خواب نرمال بیشتری نسبت به زنان دارند.
- از طرفی به نظر می‌رسد که زنان نسبت به مردان به طور قابل توجهی بیشتر احتمال دارد که آپنه خواب (sleep Apnea) را تجربه کنند.

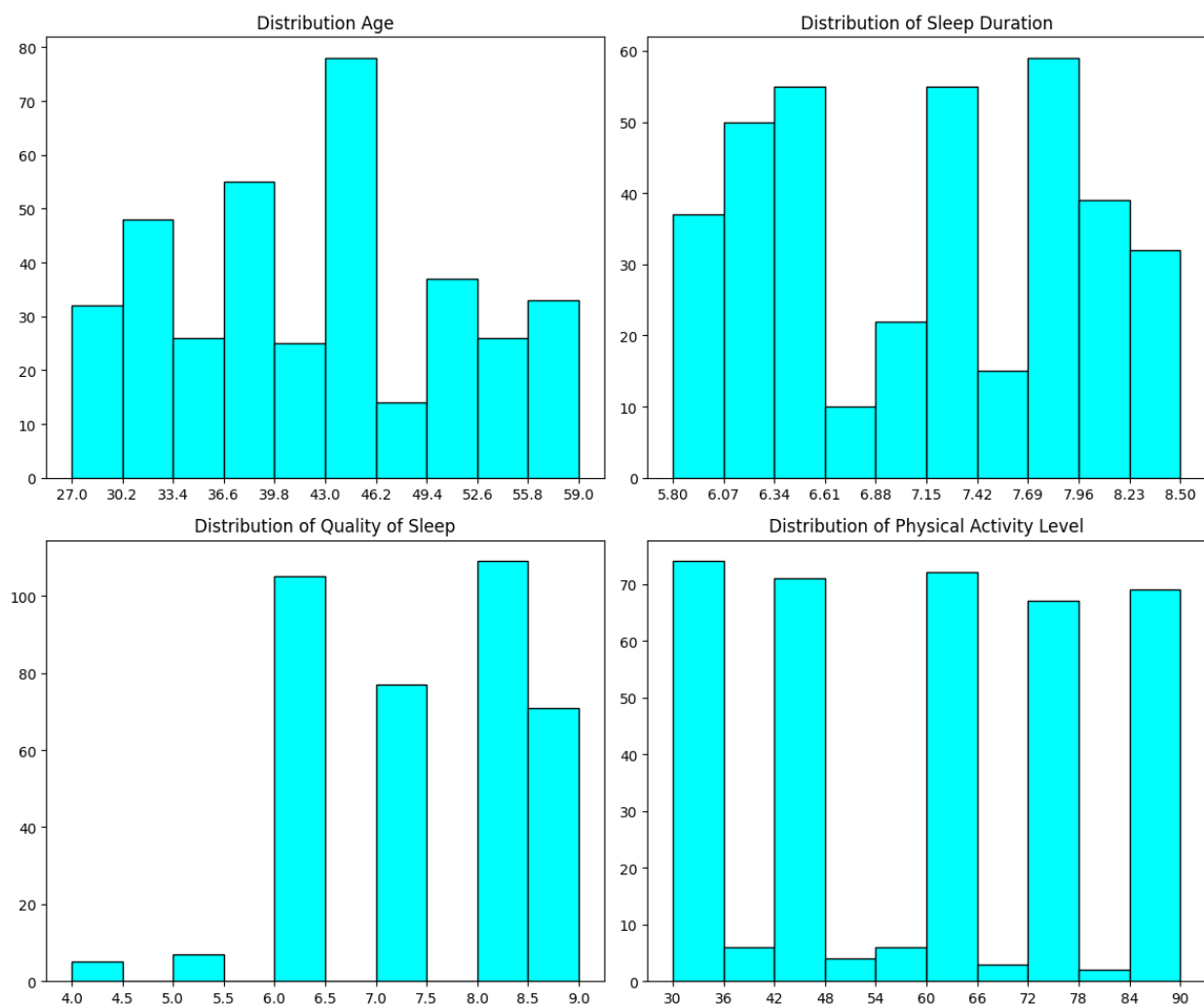
از نمودار Occupation میتوان نتایج زیر را گرفت:

- اول از همه همانطور که از توزیع این feature در مرحله قبل بدست آمد، دریافتیم که این ویژگی متوازن نیست و نمیتوان بیش از حد به داده های آن اکتفا کرد اما نکته قابل ملاحظه این است که Nurse بیشترین تعداد داده را در این ستون دارد و همانطور که از شکل بالا مشخص است تعداد قابل توجهی از nurse ها از sleep Apnea رنج می‌برند.
- Doctor هم به چون از لحاظ تعداد دومین مرتبه را دارد میتوان به تحلیل آن پرداخت، که همانطور که از شکل پیداست، اغلب دکتر ها در وضعیت نرمال قرار دارند.

از نمودار BMI Category هم میتوان نتایج زیر را گرفت:

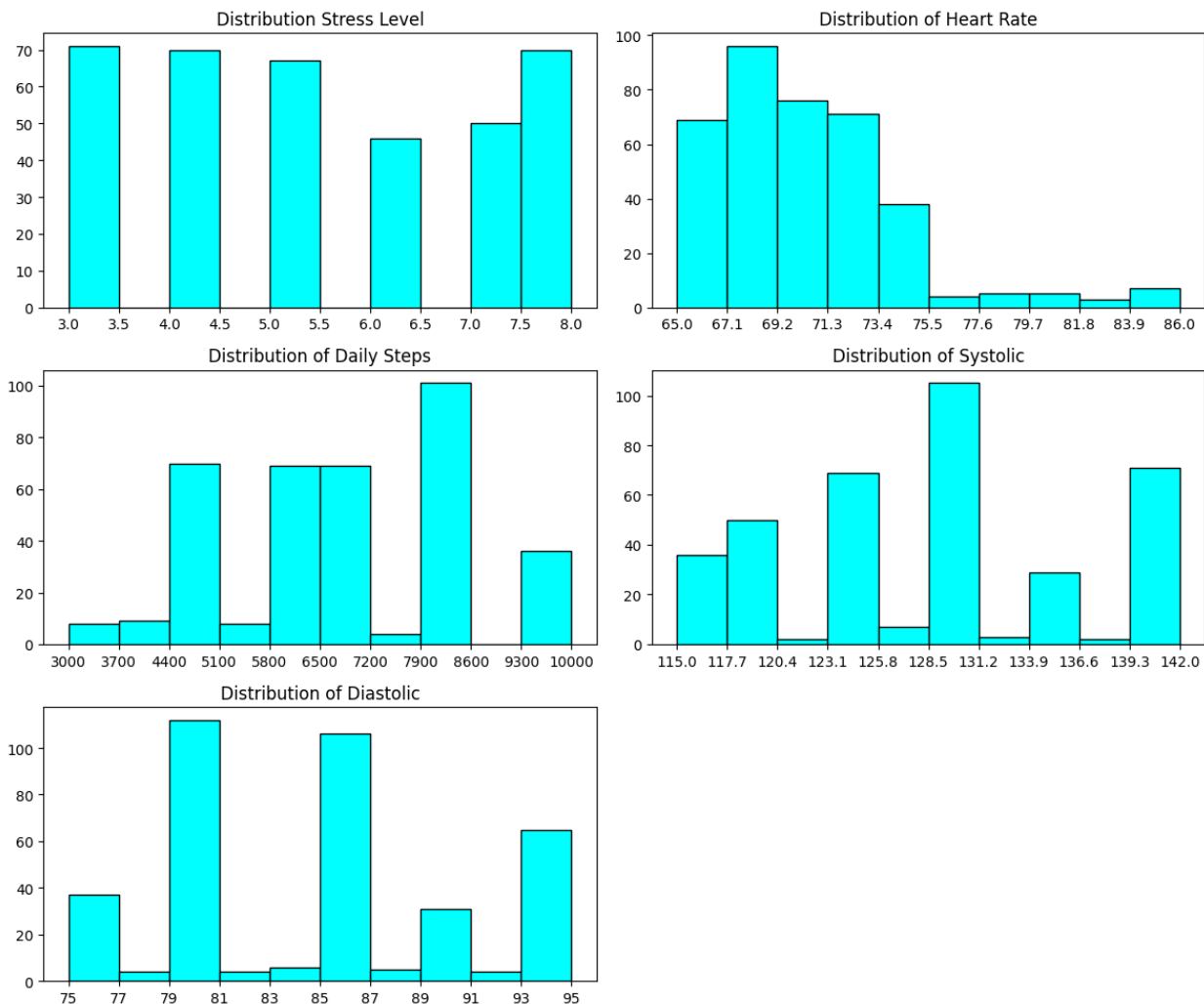
- توزیع داده ها هم در این feature نامتوازن هست و تعداد داده های normal, Overweight به نسبت داده ها Obese بیشتر میباشد.
- از شکل مشخص است که موارد Obese , Overweight هر دو به طور متوازی میتوانند هم از Insomnia رنج ببرند و هم از sleep Apnea.

ویژگی‌های عددی: ما از هیستوگرام‌ها برای تجسم توزیع آنها استفاده خواهیم کرد، که به ما در درک پراکندگی، گرایش مرکزی و هرگونه داده پرت بالقوه در داده‌ها کمک می‌کند.



پس از تجسم چهار ویژگی عددی اول، می‌توانیم ببینیم که توزیع آنها از یک الگوی ثابت پیروی نمی‌کند و فراوانی در هر ستون در محدوده‌های مختلف متفاوت است. با این حال، برای ستون کیفیت خواب، یک روند واضح مشاهده می‌کنیم: توزیع به سمت انتهای بالاتر مقیاس متمایل است، که نشان می‌دهد اکثر افراد کیفیت خواب بهتری را گزارش می‌کنند.



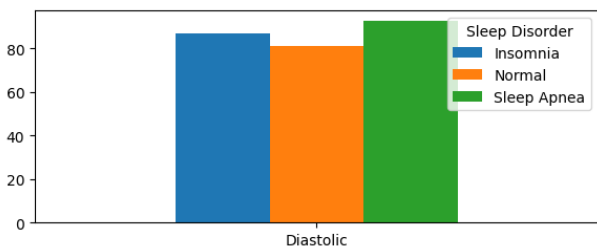
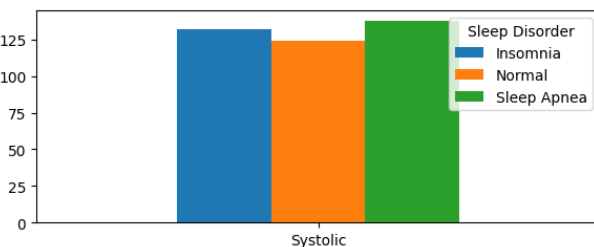
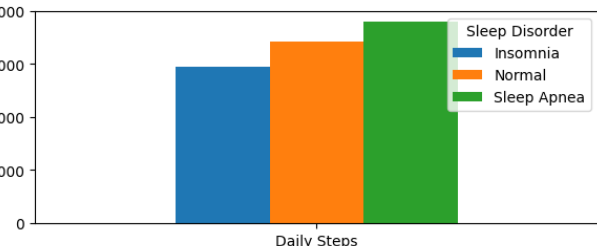
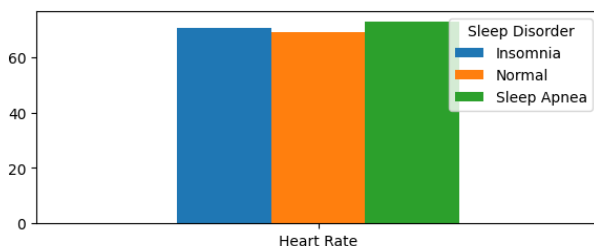
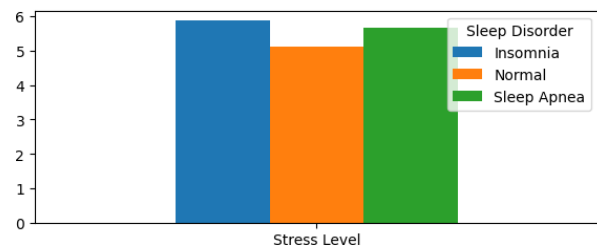
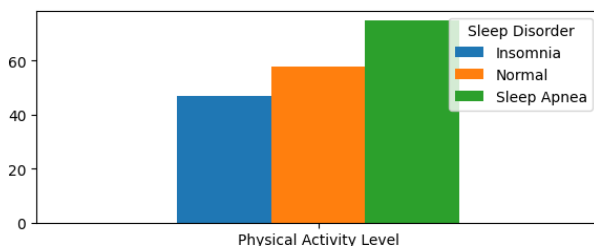
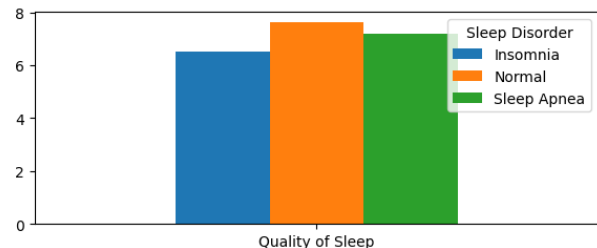
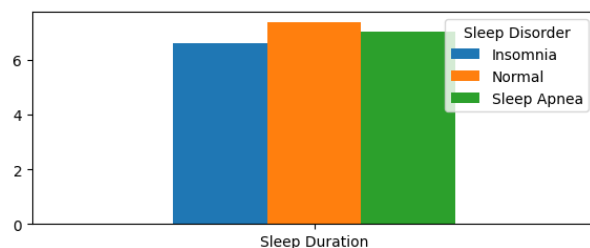


پس از بررسی بقیه ستون‌های عددی، متوجه روند مشابهی شدیم - اکثر آنها از یک الگوی توزیع واضح یا ثابت پیروی نمی‌کنند. با این حال، ستون ضربان قلب برجسته است. مقادیر آن عمدتاً در محدوده پایین تا متوسط، به ویژه بین ۶۵ تا ۷۵ ضربه در دقیقه متمرکز شده‌اند. سپس فرکانس به تدریج کاهش می‌یابد و در انتهای بالاتر مقیاس تقریباً ناچیز می‌شود.

بررسی ویژگی‌های عددی:

بررسی اینکه مقادیر ویژگی‌های عددی در دسته‌های مختلف اختلالات خواب چگونه تغییر می‌کند. ما از نمودارهای میله‌ای برای مقایسه میانگین مقادیر هر ویژگی عددی برای هر دسته در ستون اختلال خواب استفاده کردیم.

	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	Heart Rate	Daily Steps	Systolic	Diastolic
Sleep Disorder								
Insomnia	6.589610	6.532468	46.818182	5.870130	70.467532	5901.298701	132.038961	86.857143
Normal	7.358447	7.625571	57.949772	5.114155	69.018265	6852.968037	124.045662	81.000000
Sleep Apnea	7.032051	7.205128	74.794872	5.666667	73.089744	7619.230769	137.769231	92.717949



همانطور که نمودارها را بررسی می‌کنید، تفاوت بین طبقات هدف از نظر بصری واضح می‌شود. با این حال، توجه به این نکته مهم است که برای اکثر ویژگی‌های عددی، محدوده مقادیر نسبتاً باریک است و اغلب بین 0 تا 10 است. این بدان معناست که مقادیر میانگین (میانگین) در بین دسته‌های مختلف اختلالات خواب نیز نزدیک به هم به نظر می‌رسند. با وجود این تفاوت‌های کوچک، حتی تغییرات جزئی در میانگین می‌تواند بینش‌های معناداری را به همراه داشته باشد، به خصوص در داده‌های پزشکی یا رفتاری مشاهده می‌شود.

## 1. Sleep Duration (مدت زمان خواب):

- افراد **Insomnia** کمترین خواب را دارند.

- افراد **Normal** بیشترین میزان خواب را دارند.
- افراد با **Sleep Apnea** در میانه قرار دارند.

نتیجه: بی‌خوابی با کاهش مدت خواب همراه است، همان‌طور که انتظار می‌رود.

## 2. Quality of Sleep (کیفیت خواب):

- بالاترین کیفیت برای **Normal**.
- **Sleep Apnea** کیفیتی کمتر از **Normal** دارد.
- **Insomnia** کمترین کیفیت خواب را دارد.

نتیجه: کیفیت خواب با نوع اختلال رابطه مستقیم دارد. بی‌خوابی و وقفه تنفسی هر دو بر کیفیت خواب تأثیر منفی دارند.

## 3. Physical Activity Level (سطح فعالیت بدنی):

- **Sleep Apnea** بالاترین فعالیت بدنی را دارد!
- **Normal** در وسط است.
- **Insomnia** کمترین فعالیت بدنی را دارد.

نتیجه: این مورد شاید غیرمنتظره باشد، اما می‌تواند نشان دهد که افراد با بی‌خوابی تمایل به کم‌تحرکی دارند.

## 4. Stress Level (سطح استرس):

- **Insomnia** بیشترین سطح استرس را دارد.

- Sleep Apnea کمی کمتر است.

- Normal پایین‌ترین سطح استرس را دارد.

نتیجه: استرس با اختلالات خواب رابطه مستقیم دارد.

## 5. Heart Rate (ضربان قلب):

- Sleep Apnea بیشترین ضربان قلب را دارد.

- Normal و Insomnia نزدیک به هم و کمی کمترند.

نتیجه: وقفه تنفسی ممکن است موجب افزایش فشار روی قلب شود.

## 6. Daily Steps (تعداد قدم روزانه):

- Sleep Apnea بیشترین قدم، Normal دوم، Insomnia کمترین.

نتیجه: باز هم الگوی کم‌حرکی در افراد با بی‌خوابی تأیید می‌شود.

## 7. Systolic (فشار خون سیستولیک):

- Sleep Apnea بیشترین فشار سیستولیک را دارد.

- Insomnia دوم و Normal کمترین.

نتیجه: وقفه تنفسی ممکن است منجر به افزایش فشار خون شود.

## 8. Diastolic (فشار خون دیاستولیک):

- Sleep Apnea بیشترین، Insomnia وسط، Normal کمترین.

نتیجه: همانند فشار سیستولیک، فشار دیاستولیک هم با Sleep Apnea بالاتر است.

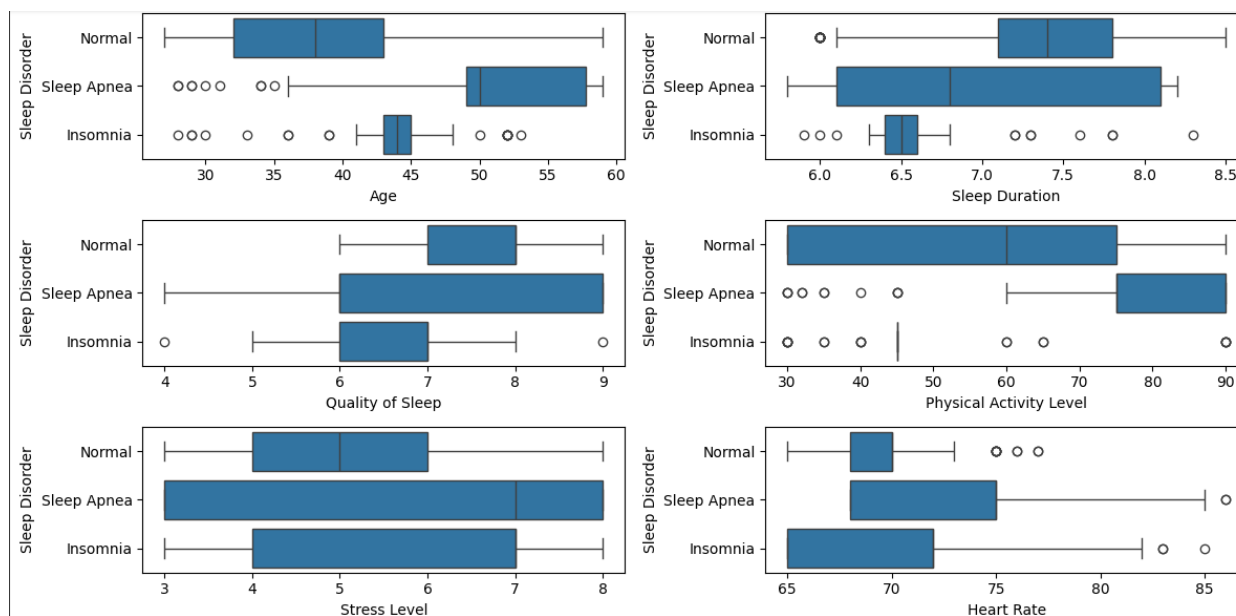
جمع‌بندی:

1. **Insomnia** با: خواب کم، کیفیت خواب پایین، فعالیت پایین، استرس بالا.

2. **Sleep Apnea** با: ضربان قلب بالا، فشار خون بالا، فعالیت نسبتاً بالا اما کیفیت خواب پایین.

3. **Normal** با: شرایط میانه یا مطلوب در اکثر متغیرها.

تحلیل باکس پلات ویژگی‌های مختلف نسبت به **target**:



برای چهار ویژگی اول می‌توان گفت:

برای AGE:

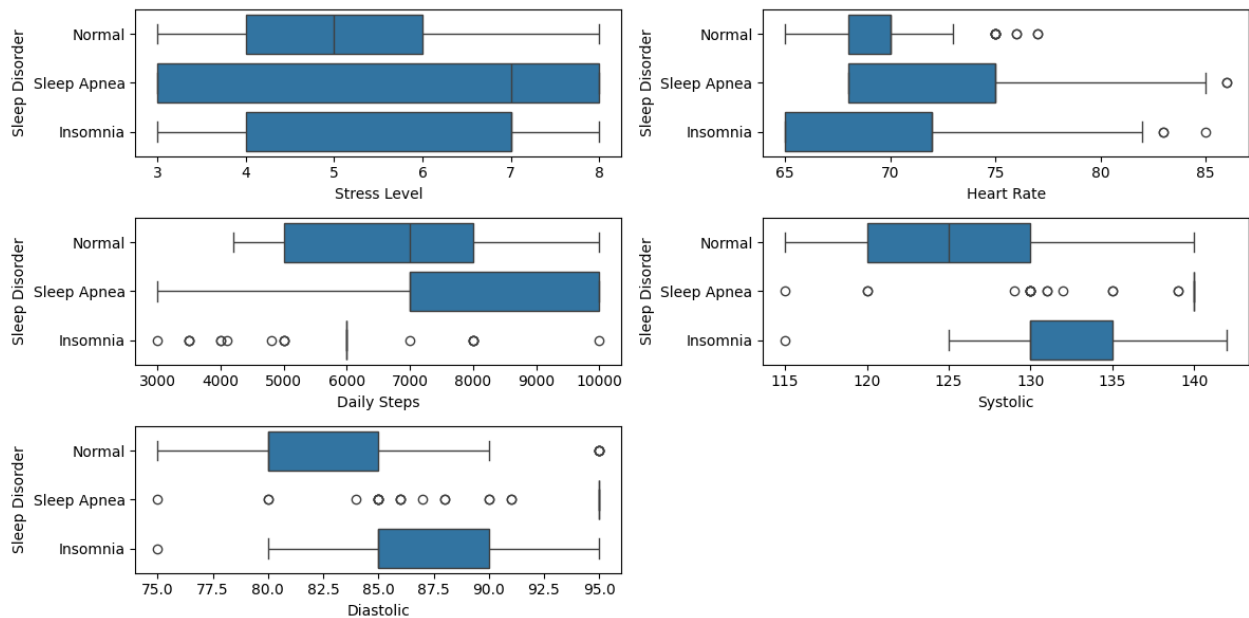
- برای AGE افرادی که در بازه سنی بین حدودا 32 تا 43 سال هستند خواب Normal دارند.
- بازه سنی بین حدودا 48 تا 56 سال معمولا از مشکلات sleep apnea رنج می‌برند.

برای Physical Activity Level:

- بازه بین 30 تا 75 خواب normal دارند در حالی که اگر فعالی بدنی از حدی بیشتر شود و بین 75 تا 90 باشد مشکلاتی از قبیل آپنه خواب ایجاد خواهد کرد.

برای باقی feature ها چون بایکدیگر همپوشانی دارند نمی‌توان با اتکا به باکس پلات به نتیجه خاصی رسید.

در زیر box plot را برای باقی feature ها رسم شده است:



اگر فشار خون را بر اساس systolic , diastolic جدا کنیم همان طور که از شکل پیدا است می‌توان نتایج زیر را بدست آورد.

برای systolic:

- بازه بین 120 تا 130 خواب normal به همراه دارد.
- بازه بین 130 تا 135 مبتلا به insomnia خواهند شد.

برای diastolic:

- بازه بین 80 تا 85 خواب Normal.
- بازه بین 85 تا 90 دچار Insomnia خواهند شد.

با توجه به نتایج Box Plot می‌توان فهمید که در حال حاضر feature هایی همچون Age, Physical Activity Level, systolic, diastolic می‌تواند برای تحلیل نتیجه خوب عمل کنند.

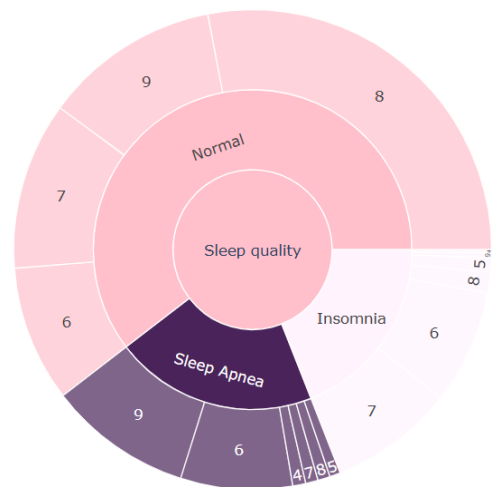
یکی از سوالاتی که می‌توان مطرح کرد این است که آیا واقعا افزایش کیفیت خواب می‌تواند باعث کاهش اختلالات خواب شود یا خیر.

- برای پاسخ به این سوال از نمودار **ecdf** استفاده کرده ایم.
- دایره مرکزی نشان می‌دهد که تمرکز کلی بر روی **sleep quality** می‌باشد.
- دایره های اطراف هم **category** های مربوط به **target** هستند.
- طول هر کدام از کمان های دایره نشان دهنده تعداد داده در آن **category** خاص می‌باشد. به عنوان مثال برای دایره میانی می‌توان گفت بیشترین داده ها مربوط به دسته **Normal** می‌باشد.
- دایره سوم هم نشان می‌دهد که برای هر **category** از **target** از هر درجه **sleep quality** چه مقدار داریم.

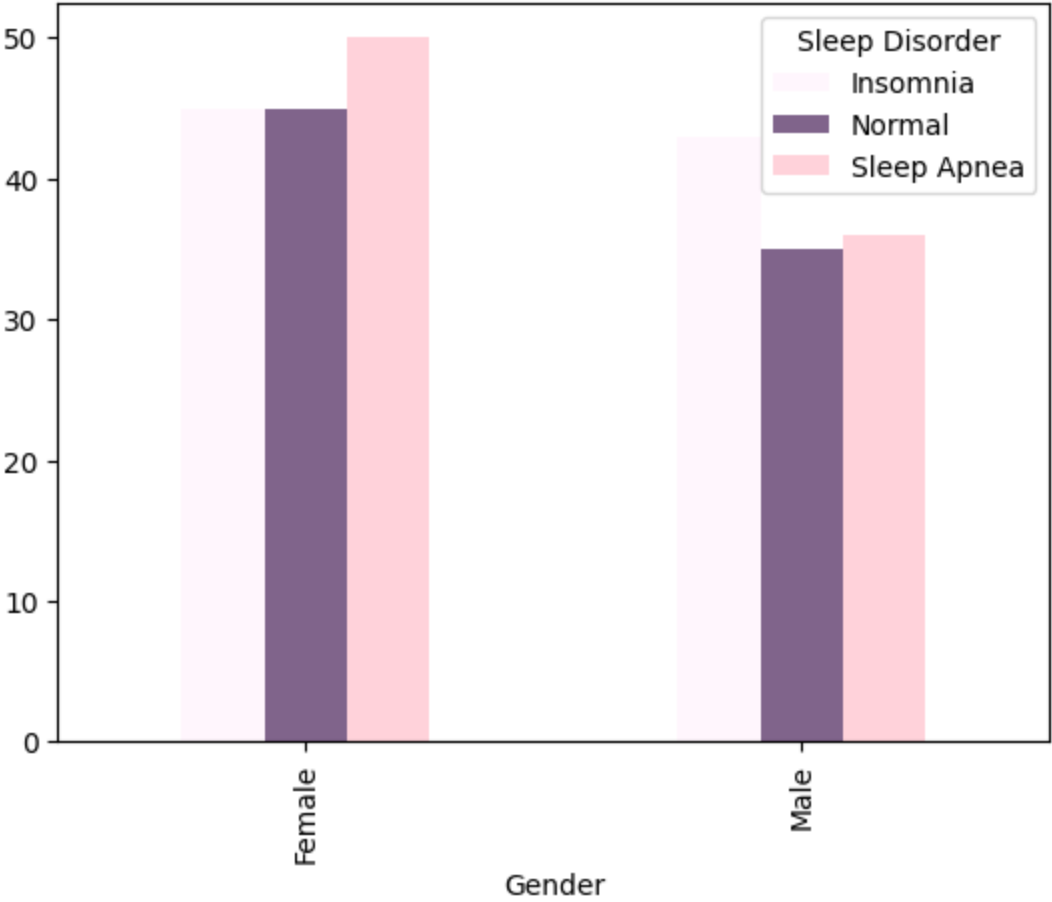
تحلیل نتایج:

- برای داده های نرمال می‌توان گفت **sleep quality** بالا می‌توان باعث تجربه خواب **Normal** شود.
- اما مورد مشکوکی که در میان داده ها وجود دارد این است که با این حالت که کیفیت خواب بالا هست اما فرد می‌توان دچار اختلال خواب می‌شوند . به عنوان مثال برای **sleep Apnea** درصد قابل توجهی از داده ها مربوط به **sleep quality** برابر با 9 است.
- در **Insomnia** هم بیشتر افراد کیفیت خواب بالایی دارند منتها با بی خوابی درگیر هستند.

## The effect of quality of sleep on sleep ..

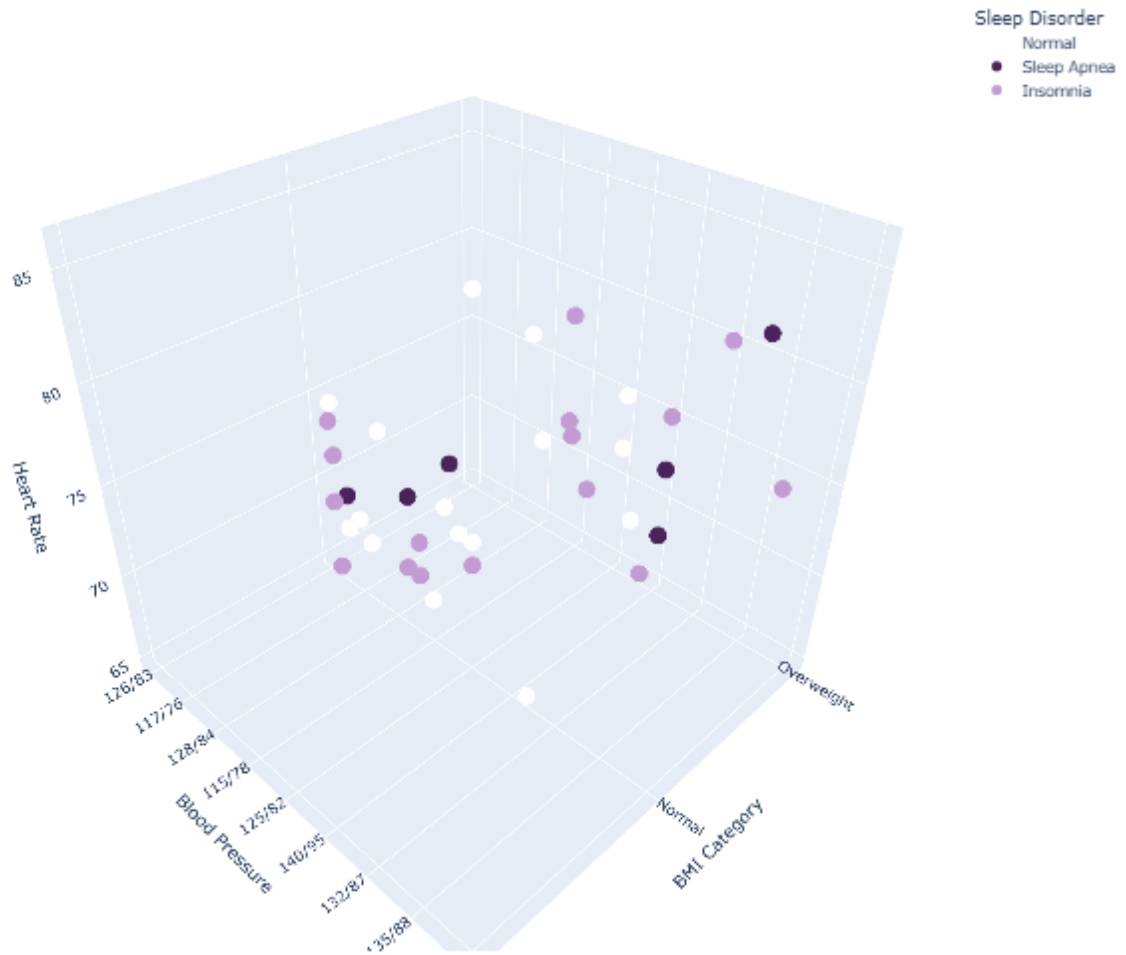


Most affected ages in each type of Sleep Disorder

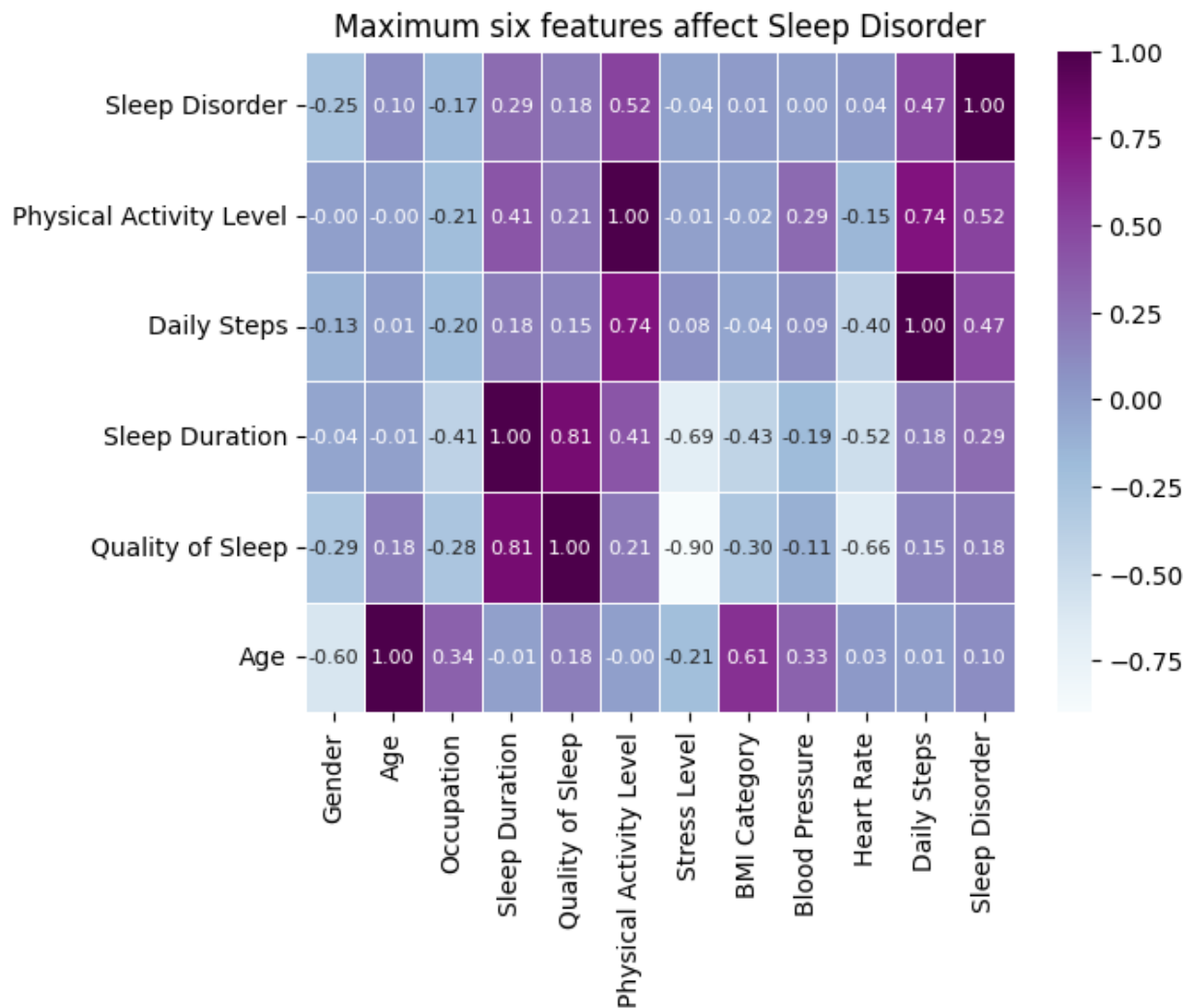




### The relationship between (BMI Category , Blood Pressure and Heart Rate) and their effect on Sleep Disorder ..



بررسی همبستگی میان داده ها، می‌توان از نمودار **heatmap** استفاده کرد ، در این نمودار هر دو ویژگی که همبستگی بیشتری باهم دارند رنگ پر رنگ و شماره آنها به یک نزدیک تر می‌باشد.



برای شکل بالا می‌توان گفت:

- دو ویژگی **Quality of sleep** و **sleep Duratio**

## تحلیل جامع اختلالات خواب بر اساس شغل: کشف علل و ارائه راهکارهای بهبود سلامت

هدف این تحلیل، شناسایی شغل‌هایی است که بیشترین اختلالات خواب را دارند و یافتن ویژگی‌های فردی مؤثر در بروز این اختلالات برای هر شغل به‌صورت جداگانه، به‌منظور ارائه پیشنهادهای سلامت‌محور و شخصی‌سازی‌شده می‌باشد. با استفاده از مدل درخت تصمیم (Decision Tree)، برای هر شغل بررسی می‌کنیم که کدام ویژگی‌ها بیشترین تأثیر را بر بروز اختلال خواب دارند.

Occupation: Doctor

Daily Steps 0.685088

Physical Activity Level 0.250205

Sleep Duration 0.033432

Age 0.031276

Quality of Sleep 0.000000

Stress Level 0.000000

Heart Rate 0.000000

dtype: float64

Occupation: Teacher

Heart Rate 0.568432

Quality of Sleep 0.253151

Sleep Duration 0.079384

Physical Activity Level 0.076229

Age 0.022804

Stress Level 0.000000

Daily Steps	0.000000
-------------	----------

dtype: float64

Occupation: Nurse

Physical Activity Level	0.696214
-------------------------	----------

Age	0.229680
-----	----------

Sleep Duration	0.055405
----------------	----------

Heart Rate	0.018701
------------	----------

Quality of Sleep	0.000000
------------------	----------

Stress Level	0.000000
--------------	----------

Daily Steps	0.000000
-------------	----------

dtype: float64

Occupation: Engineer

Heart Rate	0.718812
------------	----------

Physical Activity Level	0.102998
-------------------------	----------

Sleep Duration	0.096979
----------------	----------

Age	0.081211
-----	----------

Quality of Sleep	0.000000
------------------	----------

Stress Level	0.000000
--------------	----------

Daily Steps	0.000000
-------------	----------

dtype: float64

Occupation: Accountant

Stress Level	0.909627
--------------	----------

Sleep Duration	0.048303
----------------	----------

Age	0.042070
-----	----------

Quality of Sleep	0.000000
------------------	----------

Physical Activity Level	0.000000
-------------------------	----------

Heart Rate	0.000000
------------	----------

Daily Steps	0.000000
-------------	----------

dtype: float64

Occupation: Lawyer

Daily Steps	0.800904
-------------	----------

Physical Activity Level	0.111934
-------------------------	----------

Sleep Duration	0.067364
----------------	----------

Age	0.019797
-----	----------

Quality of Sleep	0.000000
------------------	----------

Stress Level	0.000000
--------------	----------

Heart Rate	0.000000
------------	----------

dtype: float64

Occupation: Salesperson

Sleep Duration	0.705507
----------------	----------

Age	0.294493
-----	----------

Quality of Sleep	0.000000
------------------	----------

Physical Activity Level	0.000000
-------------------------	----------

Stress Level	0.000000
--------------	----------

Heart Rate	0.000000
------------	----------

Daily Steps 0.000000

dtype: float64

نتایج تحلیل برای مشاغل مختلف:

## 1. شغل: Doctor

● مهمترین عامل:

○ Daily Steps (گام‌های روزانه): 68.5% اهمیت

● عوامل بعدی:

○ Physical Activity Level: 25%

○ Sleep Duration و Age: کم‌اثر (۳٪)

● عوامل بی‌تأثیر:

○ Quality of Sleep, Stress Level, Heart Rate

تفسیر: پزشکان با فعالیت بدنی کم و گام‌های روزانه پایین بیشتر در معرض اختلالات خواب هستند.

پیشنهاد: افزایش گام‌های روزانه و فعالیت بدنی منظم (مثلاً پیاده‌روی قبل از خواب).

## 2. شغل: Teacher

● مهمترین عامل:

○ Heart Rate: 56.8%

- عوامل بعدی:

#### ○ Sleep Duration و Quality of Sleep

- عوامل کم اثر یا بی تأثیر:

#### ○ Stress Level, Daily Steps

تفسیر: نوسانات یا بالا بودن ضربان قلب در معلمان می تواند علامت استرس مزمن یا خواب ناکافی باشد.

پیشنهاد: آموزش تکنیک های آرام سازی (نفس عمیق، یوگا)، بهبود کیفیت خواب.

### 3. شغل: Nurse

- مهمترین عامل:

#### ○ Physical Activity Level: 69.6%

- عوامل بعدی:

#### ○ Sleep Duration و Age

- عوامل بی تأثیر:

#### ○ Stress Level, Quality of Sleep, Daily Steps

تفسیر: فعالیت بدنی پایین در پرستاران نقش اصلی در اختلال خواب ایفا می کند.

پیشنهاد: ایجاد برنامه های تناسب اندام یا حرکات ساده ورزشی در شیفت های کاری.

### 4. شغل: Engineer

- مهمترین عامل:

○ Heart Rate: 71.8%

- عوامل بعدی:

○ Sleep Duration و Physical Activity Level

- بی تأثیر:

○ Stress Level, Daily Steps

تفسیر: ضربان قلب بالا ممکن است ناشی از سبک زندگی کم‌تحرک یا اضطراب باشد.

پیشنهاد: پیاده‌روی، دوچرخه‌سواری سبک، کنترل فشار خون.

## 5. شغل: Accountant

- مهمترین عامل:

○ Stress Level: 90.9%

- عوامل بعدی:

○ Sleep Duration و Age (با فاصله زیاد)

- بی تأثیر:

○ تمام عوامل دیگر

تفسیر: استرس نقش بسیار پررنگی در بروز بی‌خوابی یا سایر اختلالات در حسابداران دارد.

پیشنهاد: آموزش مدیریت استرس، استراحت‌های کوتاه در روز، کاهش بار کاری.



## 6. شغل: Lawyer

- مهمترین عامل:

Daily Steps: 80% ○

- عوامل بعدی:

Sleep Duration و Physical Activity ○

- بی‌تأثیر:

Stress Level, Heart Rate ○

تفسیر: وکلا با سبک زندگی کم‌تحرک در معرض اختلالات خواب هستند.

پیشنهاد: اضافه کردن فعالیت فیزیکی سبک به برنامه روزانه.

## 7. شغل: Salesperson

- مهمترین عامل:

Sleep Duration: 70.5% ○

- عامل دوم:

Age: 29% ○

- بی‌تأثیر:

○ بقیه متغیرها

تفسیر: خواب ناکافی و افزایش سن، عوامل کلیدی در بروز اختلال خواب در فروشندگان هستند.

پیشنهاد: تنظیم ساعات خواب ثابت و خواب باکیفیت حداقل ۷ ساعت در شبانه‌روز.

پیشنهاد بهبود سلامت	عامل اصلی اختلال خواب	شغل
افزایش پیاده‌روی روزانه	پایین <b>Daily Steps</b>	Doctor
تکنیک‌های کاهش استرس	بالا <b>Heart Rate</b>	Teacher
ورزش منظم در شیفت‌ها	کم <b>Physical Activity</b>	Nurse
کنترل سلامت قلبی و ورزش	بالا <b>Heart Rate</b>	Engineer
کاهش استرس محیط کار	بالا <b>Stress Level</b>	Accountant
تحرك بیشتر در روز	پایین <b>Daily Steps</b>	Lawyer
تنظیم زمان خواب	خواب ناکافی	Salesperson