Day 32

Exploratory Data Analysis (day 2)

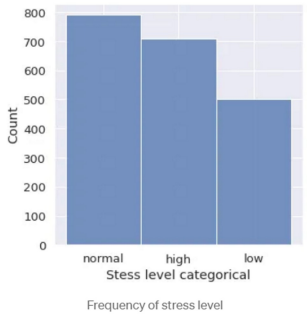
Graphs for categorical and numerical Data:

Categorical:

Graphs used can be:

- Histogram

```
sns.histplot(data=df, x='Stess level categorical')
```



Frequency of stress level

Numerical:

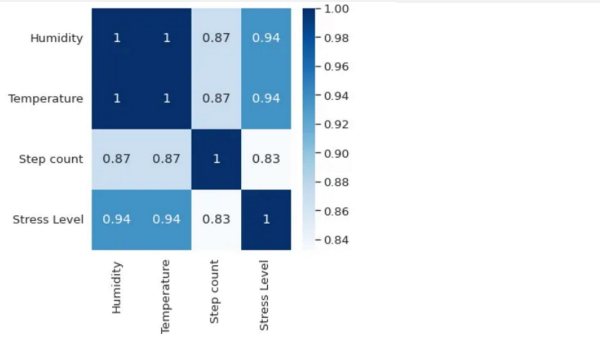
Numerical values can be either discrete or continuous.

Graphs used can be:

- Histogram
- Heatmap:

  The diagonal line describes the correlation among the same feature. The rest describe the correlation between the other features. For example, this heatmap shows that temperature is highly correlated with humidity. Also, humidity and temperature are highly correlated to the target stress level.

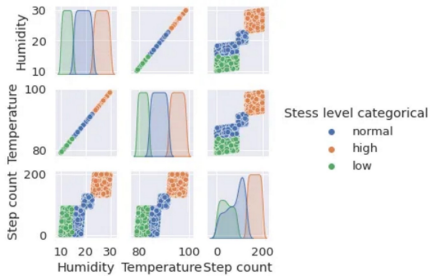
```
sns.heatmap(data=df[['Humidity', 'Temperature', 'Step count', 'Stress
Level']].corr(), cmap='Blues', annot=True)
```



- Pairplot:

  visualises the distribution of data for each target class.

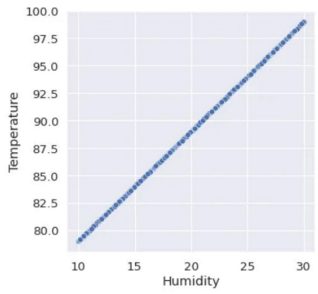
```
sns.pairplot(data=df[['Humidity', 'Temperature', 'Step count', 'Stess
level categorical']], hue='Stess level categorical', height=1.5)
```



Numerical vs Numerical:

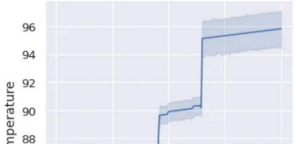
- Scatter plot:

```
sns.scatterplot(data=df, x='Humidity', y='Temperature')
```



- LinePlot:

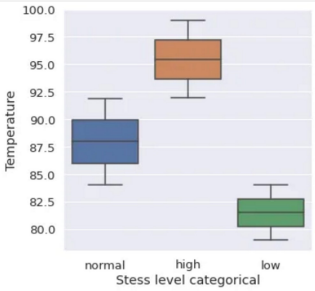
```
sns.lineplot(data=df, x='Step count', y='Temperature')
```



Categorical vs categorical:

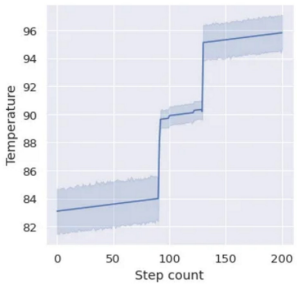
- Scatter plot
- Lineplot
- Boxplot:

```
sns.boxplot(data=df, x='Stess level categorical', y='Temperature')
```



- Barplot:

```
sns.barplot(data=df, x='Stess level categorical', y='Temperature')
```



- Barplot:

```
sns.barplot(data=df, x='Stess level categorical', y='Temperature')
```