











<div><div>MOTIVATION</div><div>Target group is young people / young adults interested in improving their own health and sleep</div><div>Goals:</div><div>Offer individuals personal advice on how to improve their well-being</div><div>Improve public health</div></div>	<div><div>DATA COLLECTION</div><div>Data from the user is not saved, so no privacy concerns. All data gathered online is anonymous and publicly available</div><div>Data to be used:</div><div>Sleep data from kaggle.com</div><div>Possibly other sources</div></div>	<div><div>PREPROCESSING</div><div>Cleaning the data:</div><div>Remove data from old individuals, since the project is aimed for young people. More filters can be applied</div><div>Remove other unnecessary data, if there is any</div></div>	<div><div>EXPLORATORY DATA ANALYSIS (EDA)</div><div>Get a general sense of the data: which factors affect the quality of sleep, for example</div><div>Data possibly from different sources: how to combine?</div><div>Check that for example all ages and genders are well represented in the data</div></div>	<div><div>VISUALIZATIONS</div><div>Create a Jupyter notebook which asks for info from the user. Based on this create plots comparing the given info to the general input data</div><div>Show visually how small lifestyle changes can affect sleep for the user</div><div>Communicate to the user how they can improve their sleep</div></div>
<div><div>LEARNING TASK (focus on problem definition)</div><div>Problem: poor health habits / sleep. Need to identify which factors affect sleep quality the most and give advice to users based on their input</div><div>Tools: linear regression</div><div>Input variables: age, step counts</div></div>	<div><div>LEARNING APPROACH (focus on solution implementation)</div><div>Methods: try regression methods first, van try something more complex if needed</div><div>Relevant metrics: averages, deviation from the average, variances</div></div>		<div><div>COMMUNICATION OF RESULTS</div><div>User interface: we use a Jupyter notebook as a prototype, which the users can use. In the future an app could be created</div><div>App could be advertised to the target group</div><div>Communicate that the results are personal: may be more motivating</div><div>Focus on the positive: give advice and don't scold</div></div>	<div><div>DATA PRIVACY AND ETHICAL CONSIDERATIONS (if applicable)</div><div>Private health data from the users is not saved</div><div>Input data is anonymous and already publicly available</div></div>
<div><div>ADDED VALUE</div><div>Possibility for the user to improve their health and sleep</div></div>				<div><div>LEGEND</div><div>WEEK 1: Data collection/preprocessing</div><div>WEEK 2: EDA & visualizations</div><div>WEEKS 3-4: Machine/deep learning</div><div>WEEK 5: Fairness & data privacy</div></div>