

Análise de Padrões de Sono dos Estudantes

Daniel Luan Lourenço de Lima
Maria Eduarda Lima da Luz

20 de Dezembro de 2024

Este relatório técnico apresenta a análise de padrões de sono dos estudantes com base no conjunto de dados disponibilizado pelo Kaggle. Apesar das análises realizadas, não foram identificadas correlações ou conclusões claras devido à natureza sintética e limpa do conjunto de dados, que não apresentava valores faltantes, linhas duplicadas ou outliers significativos. O trabalho discute as limitações encontradas e sugere caminhos para estudos futuros com dados mais representativos.

1 Introdução

O sono desempenha um papel fundamental na saúde e no desempenho acadêmico dos estudantes. O objetivo deste trabalho é identificar fatores que afetam a qualidade e a duração do sono com base no conjunto de dados *Student Sleep Patterns*, contendo informações sobre estilo de vida, atividades diárias e hábitos de sono de estudantes universitários. A escolha do tema justifica-se pela relevância do sono para o bem-estar e pela possibilidade de identificar padrões e insights úteis para melhorias na rotina dos estudantes.

1.1 Objetivos

O presente estudo busca analisar a qualidade do sono em função de fatores como atividade física, tempo de tela e duração do sono, além de identificar padrões gerais nos hábitos dos estudantes. Também se pretende discutir as limitações encontradas no trabalho e propor possíveis melhorias para estudos futuros, considerando a importância de dados mais variados e representativos para análises robustas.

2 Descrição da Base de Dados

2.1 Origem e Características Gerais

A base de dados foi obtida no Kaggle e contém 14 variáveis principais, com informações de 500 estudantes. As variáveis incluem características demográficas (idade, gênero), hábitos de sono (duração, qualidade) e estilo de vida (atividade física, uso de telas).

2.2 Qualidade dos Dados

Por se tratar de um dataset sintético, não foram encontrados valores faltantes, linhas duplicadas ou outliers significativos. Essa característica garantiu a consistência dos dados,

mas também limitou a identificação de padrões complexos e mais profundos, devido à ausência de variabilidade real nos registros analisados.

2.3 Atributos Mais Relevantes

- **Sleep_Quality**: Qualidade subjetiva do sono (escala de 1 a 10).
- **Sleep_Duration**: Duração do sono em horas.
- **Physical_Activity**: Tempo médio diário de atividades físicas.
- **Screen_Time**: Tempo médio diário em dispositivos.

3 Metodologia

3.1 Etapas

- Carregamento e limpeza dos dados.
- Análise exploratória com visualizações gráficas.

3.2 Ferramentas Utilizadas

- Python (bibliotecas: pandas, seaborn, matplotlib).
- Jupyter Notebook para desenvolvimento iterativo.

4 Resultados da Análise Exploratória

4.1 Análise Univariada

A análise exploratória começa com a visualização das distribuições básicas dos dados demográficos presentes no conjunto. Isso ajuda a entender o perfil dos estudantes analisados.

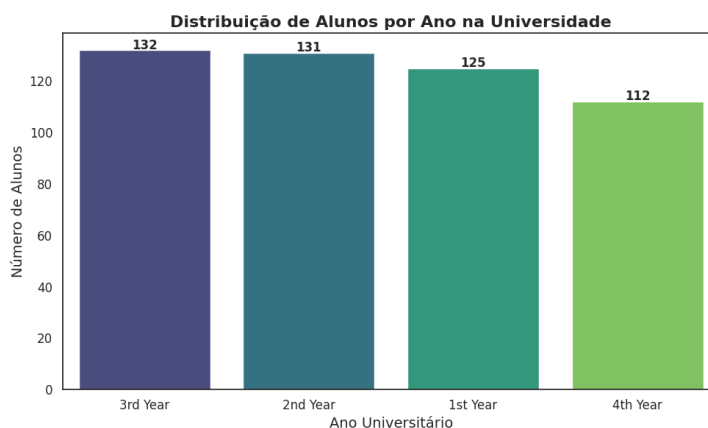


Figure 1: Distribuição de Alunos por Ano na Universidade.

A Figura 1 apresenta a distribuição dos estudantes por ano na universidade, mostrando uma maior concentração de alunos nos anos iniciais.

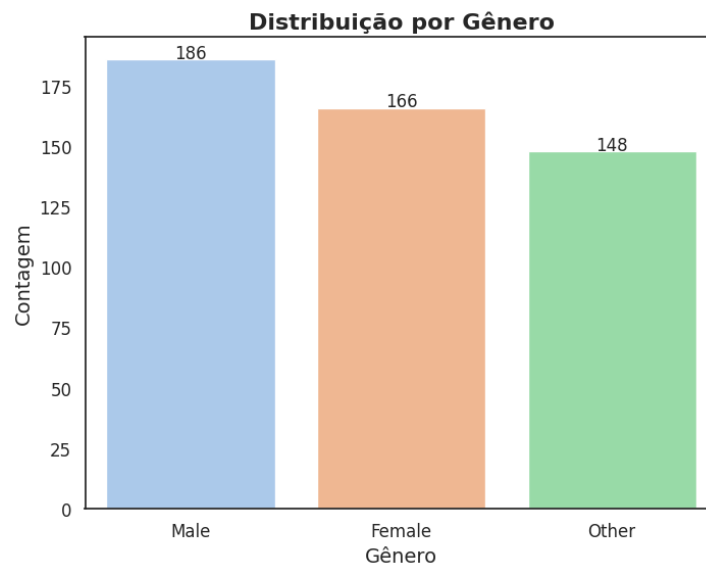


Figure 2: Distribuição por Gênero.

A Figura 2 apresenta a proporção de estudantes por gênero, permitindo observar uma distribuição equilibrada entre os grupos.

4.2 Análise Univariada

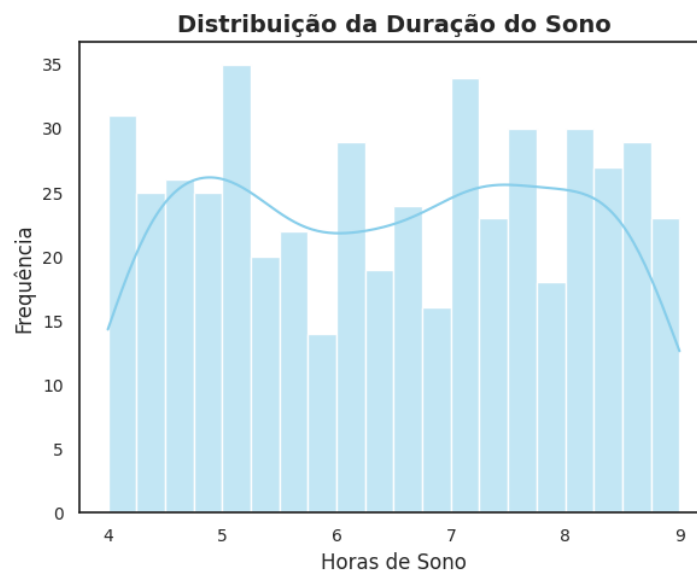


Figure 3: Distribuição da duração do sono.

A duração média do sono está abaixo das 7-9 horas recomendadas, com variação significativa entre os estudantes.

4.3 Análise Bivariada

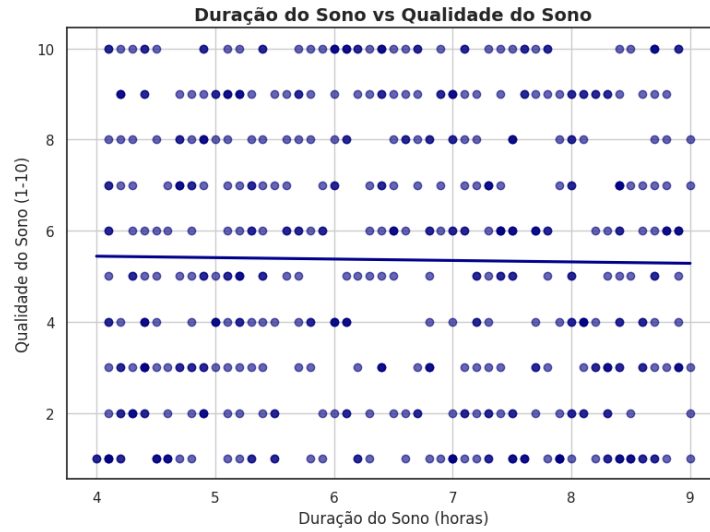


Figure 4: Correlação entre duração do sono e qualidade do sono.

Apesar de esperada, a correlação entre duração e qualidade do sono é baixa. Não foram observados padrões relevantes, reforçando as limitações impostas pelo caráter sintético dos dados.

4.4 Análise Multivariada

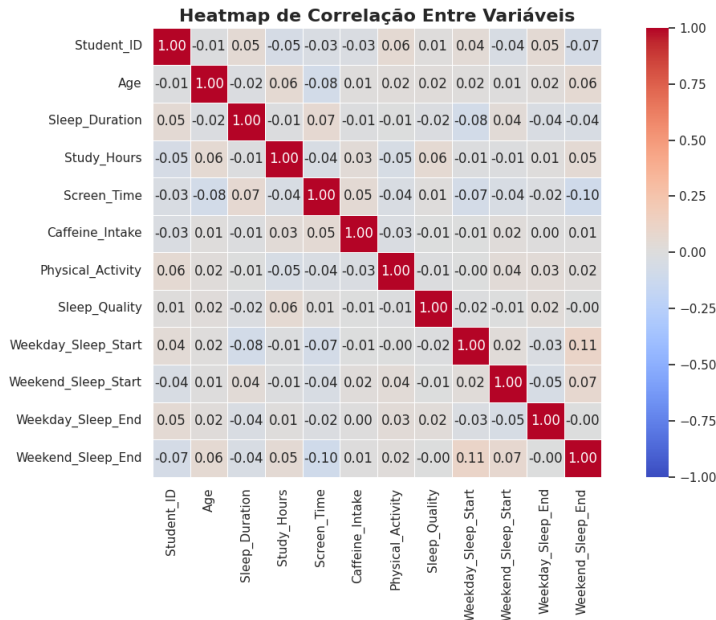


Figure 5: Matriz de correlação entre variáveis numéricas.

O heatmap de correlação revela relações fracas entre as variáveis. A ausência de variabilidade significativa no conjunto de dados limitou a identificação de insights mais profundos.

5 Discussão dos Insights

Os resultados indicam que a natureza já tratada e limpa dos dados dificultou a identificação de padrões ou insights significativos. A duração do sono, isoladamente, não demonstrou correlações claras com a qualidade do sono, e fatores como tempo de tela e atividade física apresentaram apenas relações fracas. Estudos futuros com dados mais variados e representativos poderiam oferecer perspectivas mais sólidas.

6 Conclusão

O estudo revelou limitações impostas pela natureza do conjunto de dados utilizado, que não permitiu a identificação de padrões claros ou conclusões robustas. Intervenções visando a coleta de dados mais ricos, com maior variabilidade e representatividade, são recomendadas para análises futuras.

7 Referências

- Dataset disponível em // <https://www.kaggle.com/datasets/arsalanjamal002/student-sleep-patterns>
- Documentação e bibliotecas utilizadas: pandas, seaborn, matplotlib.