

Fatores de Estresse em Estudantes

Análise Exploratória de Dados

Lucas Linhares
novembro/23





OS DADOS

Os dados foram coletados de estudantes do Ensino Médio e Universitários da cidade de Dharan no Nepal. Universidade: Tribhuvan University.
Ano: 2022

Fonte: Kaggle

Dataset: Student Stress Factors: A Comprehensive Analysis



OBJETIVO

Explorar os dados disponibilizados, aplicando os conceitos de Estatística Descritiva de forma a entender a Distribuição e a Correlação dos dados.

Introdução



ETAPAS DO PROJETO



1- Escolha dos Dados

Por não possuir uma base de dados própria e não poder utilizar dados do trabalho, busquei um dataset no Kaggle.

2 - Entendimento dos dados

Após a escolha do dataset, li a descrição das colunas, os comentários do autor e as perguntas realizadas por outras pessoas previamente.

3 - Limpeza

Haviam mais de 20 de colunas, escolhi 8 para elaboração do projeto.

4 - Análise

Fiz o download do arquivo .csv e tratei ele com Python no Google Colab. Utilizei a biblioteca Pandas para estatística descritiva e as bibliotecas Matplot e Seaborn para elaboração dos gráficos e Correlação.

5 - Resultados

Após a análise realizei a interpretação dos resultados e levantei possíveis insights.

A BASE DE DADOS

Autoestima	Depressão	Histórico de Saúde Mental	Nível de Ansiedade
<ul style="list-style-type: none"> • Escala de 0 a 30. <ul style="list-style-type: none"> ◦ 0 a 15 - baixo; ◦ 16 a 25 - normal; ◦ 25 a 30 - alto. • Escala de Autoestima de Rosenberg. 	<ul style="list-style-type: none"> • Escala de 0 a 27. <ul style="list-style-type: none"> ◦ 0 - sem indícios de depressão; ◦ 1 a 4 - depressão mínima; ◦ 5 a 9 - depressão leve; ◦ 10 a 14 - depressão moderada; ◦ 15 a 19 - depressão moderadamente severa; ◦ 20 a 27 - depressão severa. • PHQ-9 Patient Depression Questionnaire. 	<ul style="list-style-type: none"> • 1 se existe histórico de questões com saúde mental. • 0 se não há histórico. 	<ul style="list-style-type: none"> • Escala de 0 a 21. <ul style="list-style-type: none"> ◦ 0 a 4 - mínimo; ◦ 5 a 9 - leve; ◦ 10 a 14 - moderada; ◦ 15 a 21 - severa. • General Anxiety Disorder-7.

Nível de Estresse	Performance Acadêmica	Preocupação com a Carreira	Qualidade do Sono
<ul style="list-style-type: none"> • Escala de 0 a 5. <ul style="list-style-type: none"> ◦ 0 a 1 - baixo ◦ 2 a 3 - moderado ◦ 4 a 5 - alto 	<ul style="list-style-type: none"> • Escala de 0 a 5. <ul style="list-style-type: none"> ◦ 0 a 1 - baixo ◦ 2 a 3 - moderado ◦ 4 a 5 - alto 	<ul style="list-style-type: none"> • Escala de 0 a 5. <ul style="list-style-type: none"> ◦ 0 a 1 - baixo ◦ 2 a 3 - moderado ◦ 4 a 5 - alto 	<ul style="list-style-type: none"> • Escala de 0 a 5. <ul style="list-style-type: none"> ◦ 0 a 1 - baixo ◦ 2 a 3 - moderado ◦ 4 a 5 - alto

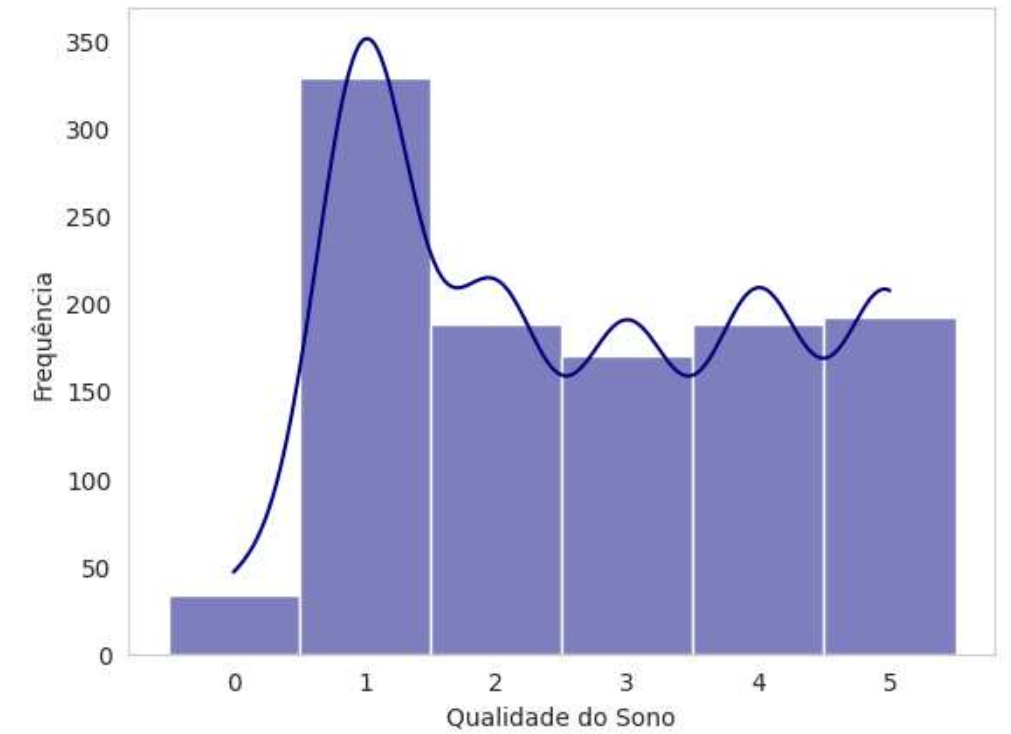
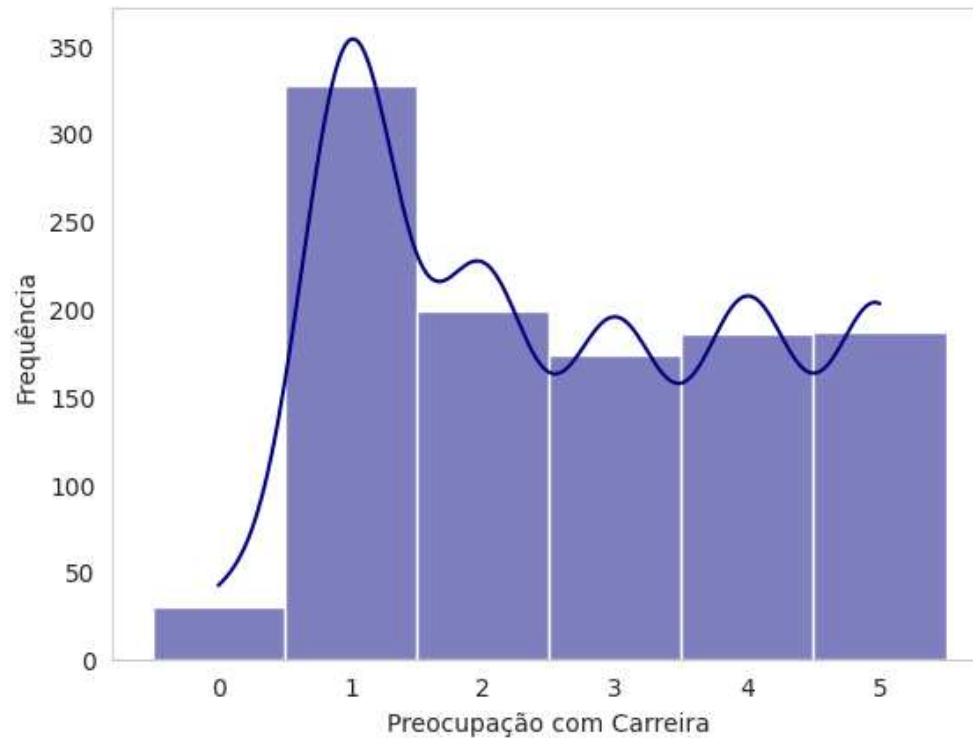
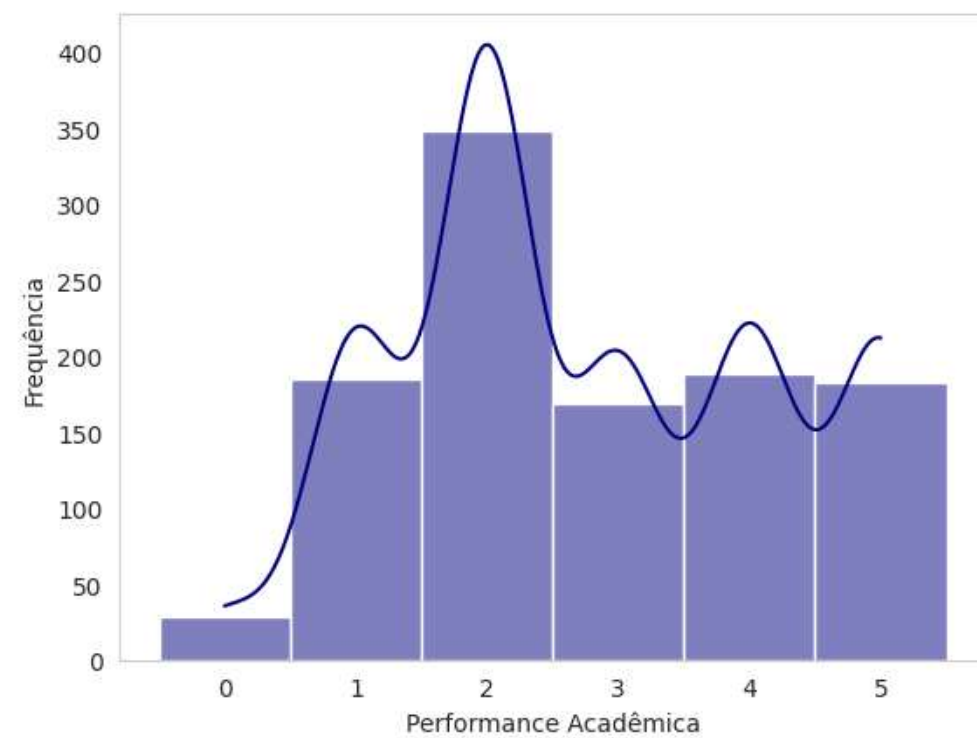
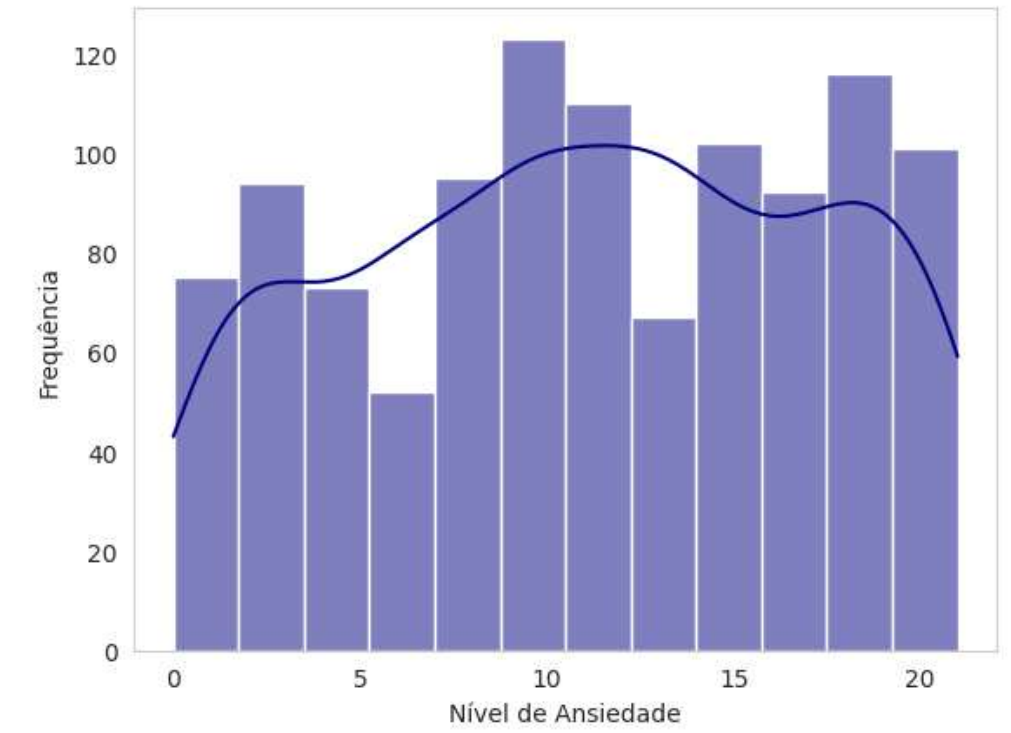
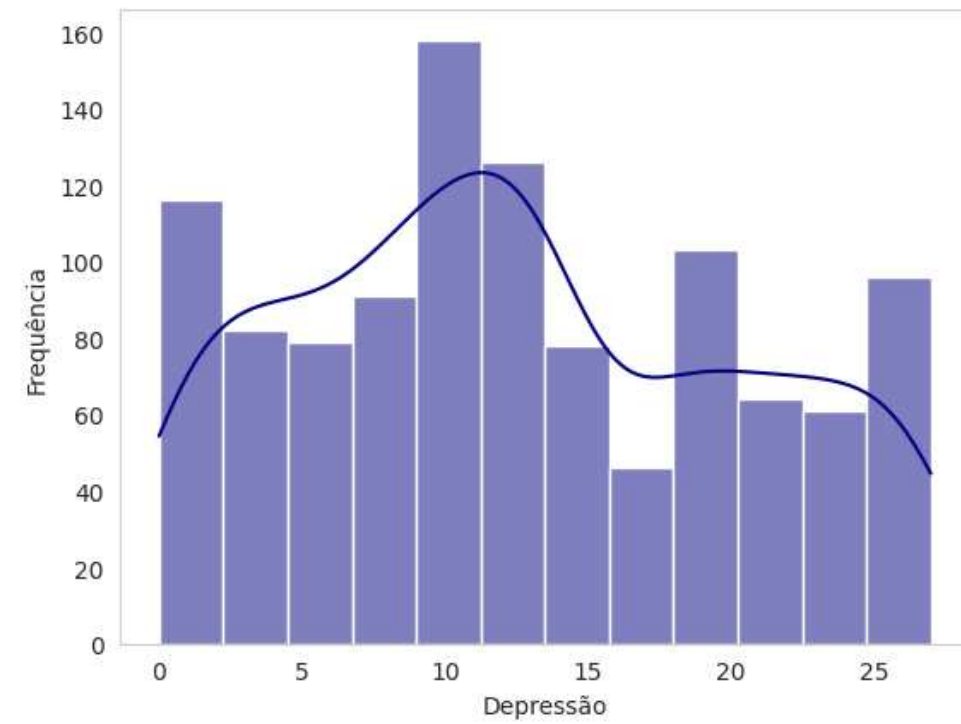
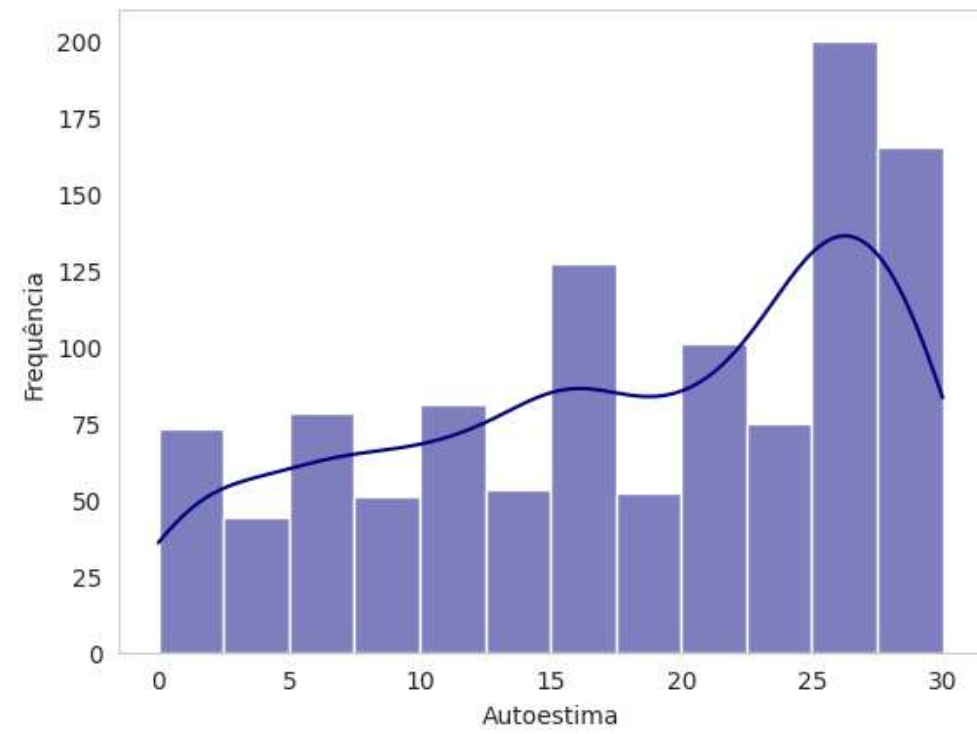
Estatística Descritiva

- Foram coletados dados de **1100** estudantes.
- O Nível de Estresse não chegou ao máximo possível, o valor mais alto foi **2**. Logo, esta variável não será considerada nas correlações.
- Todas as outras variáveis atingiram o máximo possível nas escalas.
- A média do Nível de Ansiedade é de **11,1** (moderada) e a mediana **11** (moderada).
- A média do Grau de Depressão é **12,6** (moderada) e a mediana **12** (moderada).
- A média de Autoestima está em **17,8** (normal) e a mediana **19** (normal).

Medida	Autoestima	Depressão	Nível de Ansiedade	Nível de Estresse	Performance Acadêmica	Preocupação com Carreira	Qualidade do Sono
Contagem	1100.0	1100.0	1100.0	1100.0	1100.0	1100.0	1100.0
Média	17.8	12.6	11.1	1.0	2.8	2.6	2.7
Desvio Padrão	8.9	7.7	6.1	0.8	1.4	1.5	1.5
Mínimo	0.0	0.0	0.0	0.0	0.0	0.0	0.0
25%	11.0	6.0	6.0	0.0	2.0	1.0	1.0
50%	19.0	12.0	11.0	1.0	2.0	2.0	2.5
75%	26.0	19.0	16.0	2.0	4.0	4.0	4.0
Máximo	30.0	27.0	21.0	2.0	5.0	5.0	5.0



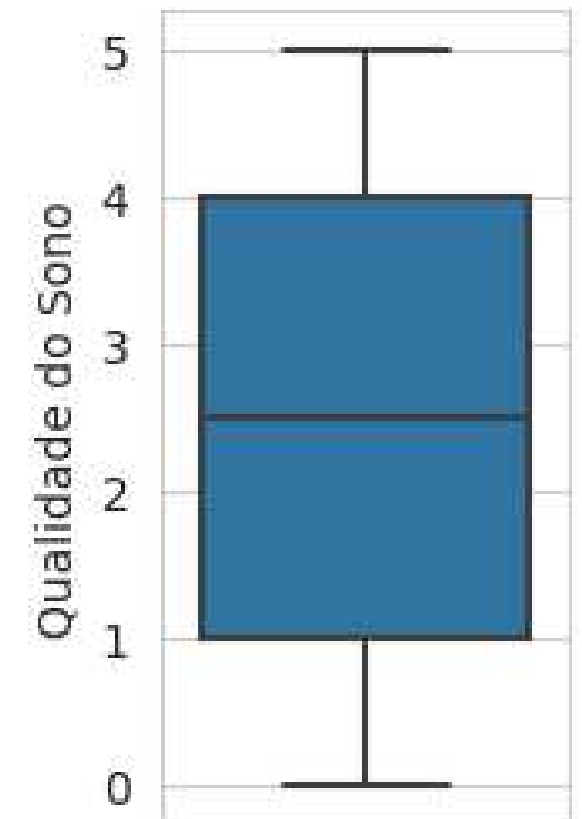
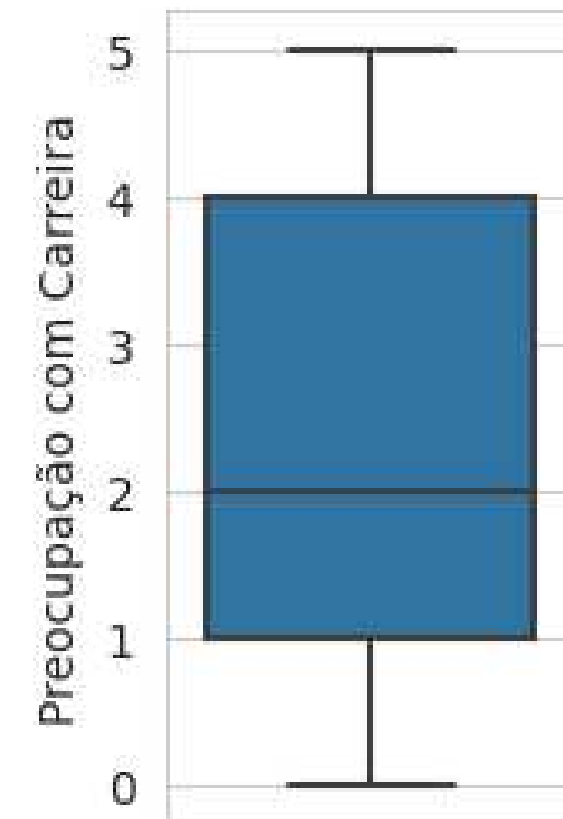
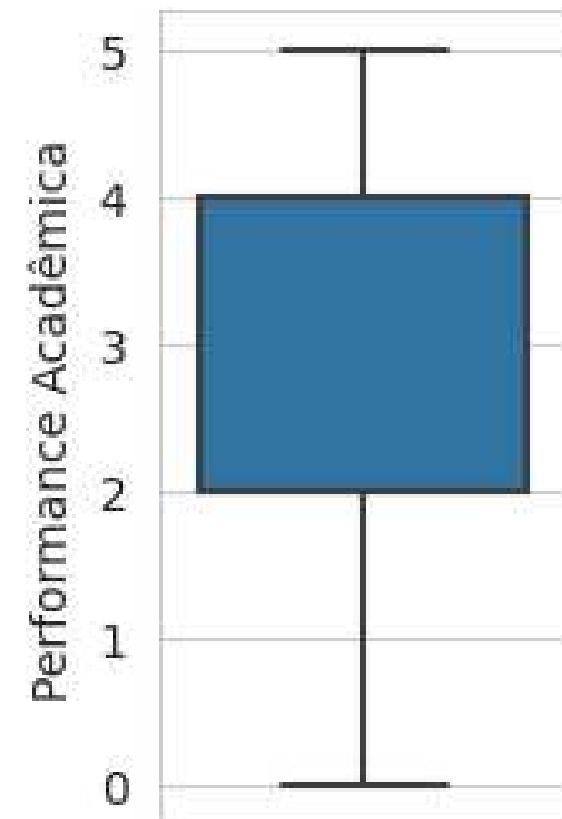
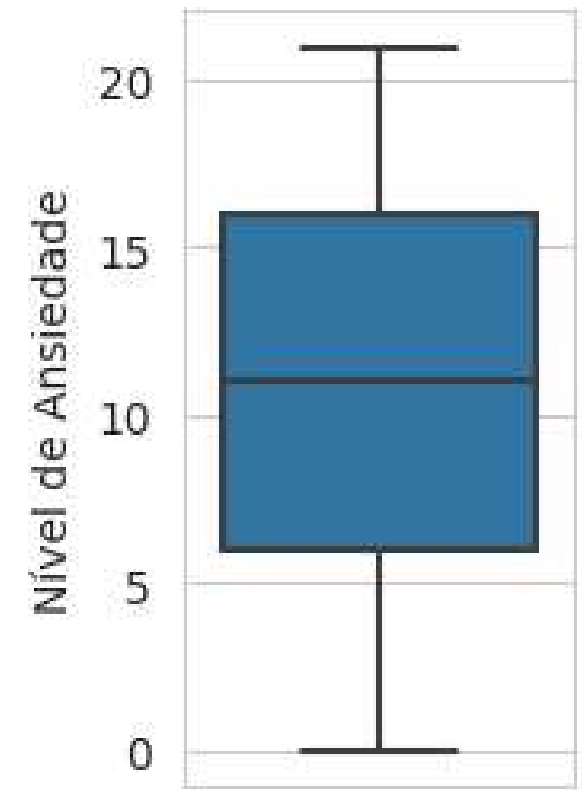
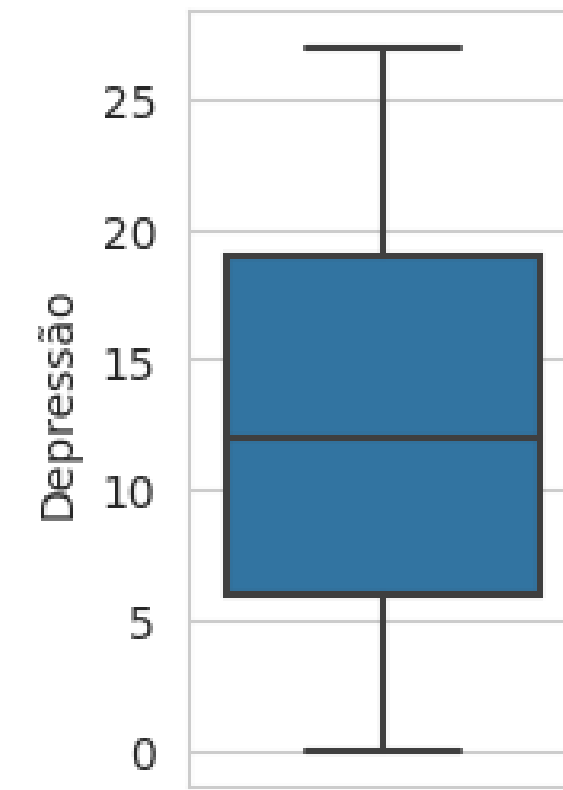
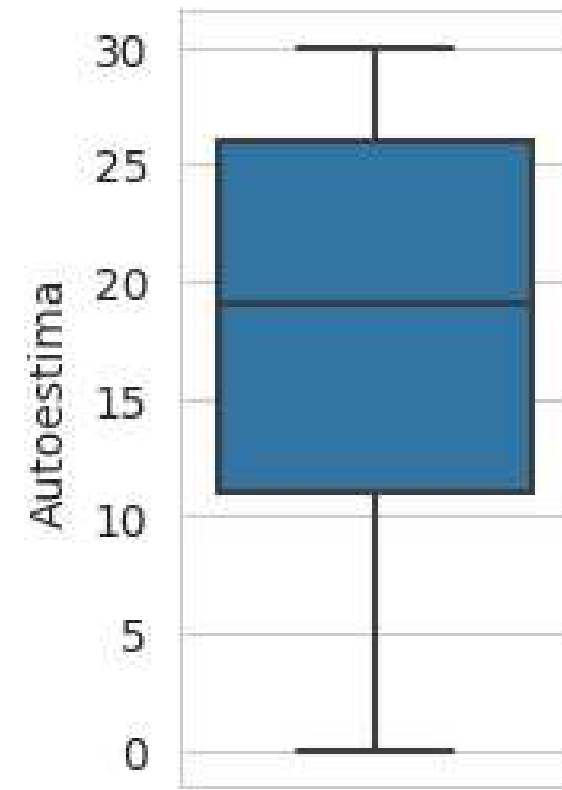
Histogramas



Box-Plot



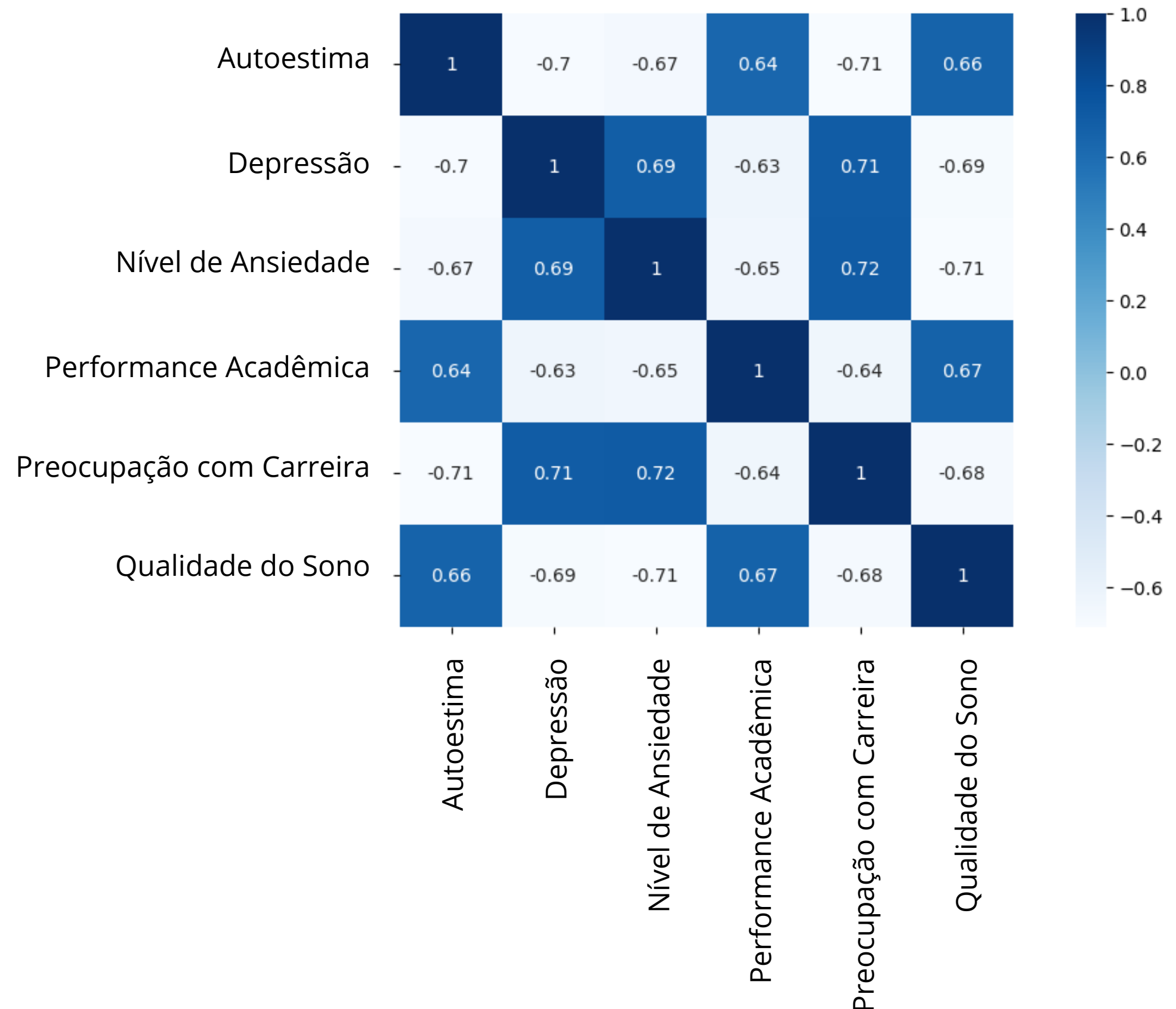
- Não há outliers em nenhuma das variáveis.
- A variável Performance Acadêmica não possui marcação de mediana, pois possui o mesmo valor do Q1 (2).



Correlação

- Podemos notar no Mapa de Calor que as variáveis possuem **Correlação linear moderada** entre elas, tanto **positivas**, quanto **negativas**.
- Analisando a Performance Acadêmica, as variáveis que demonstram mais correlação são:
 - Qualidade de Sono (**0,67**);
 - Nível de Ansiedade (**-0,65**);
- A Qualidade do sono possui uma Correlação Moderada Negativa com a Depressão (**-0,69**).
- O Nível de Ansiedade possui correlação possui Correlação Positiva Moderada com a Preocupação com Carreira (**0,72**).

Mapa de Calor - Correlação entre as Variáveis



Lucas Linhares



[linkedin.com/in/eng-lucas-linhares/](https://www.linkedin.com/in/eng-lucas-linhares/)



lucaslinhares.f@gmail.com

