2Outline: Identifying Key Attributes Correlated with High Academic Performance

Objective

* What is the main goal of the analysis?
  + By analyzing various factors, we hope to uncover insights that can enhance educational strategies and support student success.

Research Questions

**How does the time allocated to different types of activities (study, leisure, work, etc.) influence student success?**

* This would help uncover the balance between study time and other commitments, and its impact on academic results

**Independent variables** -> Extracurricular activities (categorical; yes/no), tutoring sessions (numerical), physical activity (numerical)

**Dependent variables** -> Exam score

**How does a student’s attendance and motivation affect student success and learning outcomes? r2**

* Investigating the role of mental capacity and performance can lead to recommendations on student lifestyle adjustments.

**Independent variables** -> Attendance, Motivation level

**Dependent variables** -> Exam score

**How does the distance from home to school impact student performance, and what measures can mitigate potential negative effects (e.g., fatigue or reduced study time)?**

* This can lead to insights on how transportation or proximity to school affects academic success.

**Independent variables** -> Distance from home, amount of sleep, and hours studied

**Dependent variables** -> Exam score

Insight:

1. The closer you are to campus -> the better exam score (average)

* Other factors affected this: number of hours studied has a positive correlation with average exam score
* Amount of sleep has a negative correlation with average exam score

### **Team Role Assignments**

1. **Steven & Nick**
   * **Responsibilities**:
     + **Cleaning Data**: Handle data preprocessing tasks, including removing duplicates, managing missing values, and converting data types.
2. **Shiv**
   * **Responsibilities**:
     + **Presents Research Questions**: Prepare a brief presentation on the research questions and their relevance to the analysis.
3. **Will**
   * **Responsibilities**:
     + **Visualizations**: Collaborate with Alex to create visual representations of the data using Tableau.
     + **Assist with EDA**: Help analyze the data and identify key insights that will inform the visuals.
4. **Alex**
   * **Responsibilities**:
     + **Exploratory Data Analysis (EDA)**: Conduct the EDA to uncover patterns and relationships within the data.
     + **Create Visuals**: Work with Will to design and finalize visualizations in Tableau, ensuring they effectively convey insights.
5. **Nick & Steven**
   * **Responsibilities**:
     + **Design Slides**: Create presentation slides that summarize the findings, including visualizations and key insights from the analysis.
     + **Final Presentation Coordination**: Ensure that the slides flow well and support the overall narrative of the presentation.

Questions for Sylas

* How accurate (%) should our decision model be?

Things We Learned On The Way

* We realized that having exam scores messes up with our implemented models because of the popularity of exam scores being on the lower side with the data set
  + Therefore, we implemented an ExamCode variable that helps us differentiate between a good score and a bad score. In doing so we can alter the fluctuation of the data being one sided

Presentation Outline BHDAC

The purpose of this presentation is to help develop a skill set for success in any professional

environment: the ability to effectively communicate the results of your analysis to stakeholders,

especially those without a technical background. Mastering the art of storytelling is key to

translating complex data insights into compelling narratives that resonate, engage, and drive

informed decision-making.

1. Introduction &amp; Problem Statement

- Begin by clearly explaining the purpose of your analysis. Describe why analyzing this

dataset is valuable and what specific problem you aim to solve or research.

- Identify and define stakeholders. Think beyond the traditional business definition when

necessary. Stakeholders could include decision makers, policy influencers, target

customers, or communities impacted by the data. Consider who would benefit from your

findings and why they would be invested in this analysis.

2. Dataset Context

- Provide a brief description of the dataset including the source of the data (open-source

database, Kaggle, etc.), number of observations (rows) and variables (columns), types of

variables (categorical, numerical, text, boolean).

- Highlight the main variables you analyzed and explain why they are relevant to your

research questions.

- Mention any limitations of the dataset (e.g. missing values) and briefly discuss how you

handled them.

3. Process of Answering Research Questions

- Clearly list the research questions you sought to answer and explain why these questions

are important in the context of your dataset.

- Outline the process you used to address your research questions. This should include data

cleaning and preprocessing steps, exploratory data analysis, and tools used.

4. Research Question Answers

- For each research question present the key findings supported by visuals (charts and4

graphs)

- Highlight any significant patterns, correlations, or trends that address the research

questions

- Introduce the implications of your findings. What do these results mean in the context of

your problem statement and for your stakeholders?

5. Modeling (OPTIONAL)

- If applicable, describe the ML model(s) you used.

Presentation Outline BHDAC

- Explain why you chose this specific model for your analysis. How does it align with your

dataset and research questions?

- Present key performance metrics (e.g. accuracy, precision, recall, F1-score, R 2 , etc.)

- Summarize key akeaways from your model’s results and how they contribute to

answering your research questions

6. Recommendations

- Based on your findings, provide actionable recommendations that align with the context

of the dataset.

- Consider how your recommendations can be implemented. This is your chance to

showcase critical thinking by translating your analysis into real-world solutions.

- Encourage further investigation or suggest areas where additional data might be needed

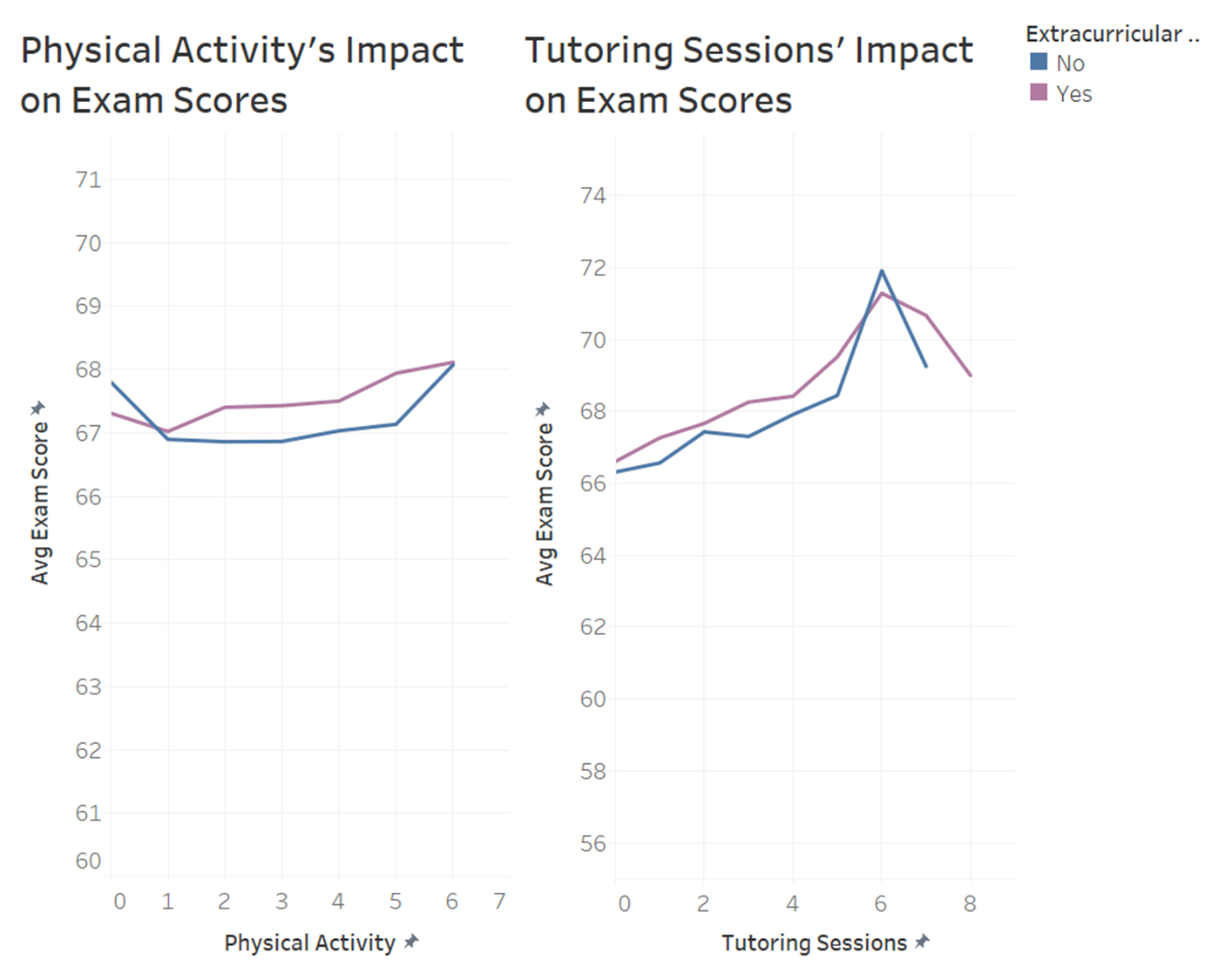
to refine your recommendations.

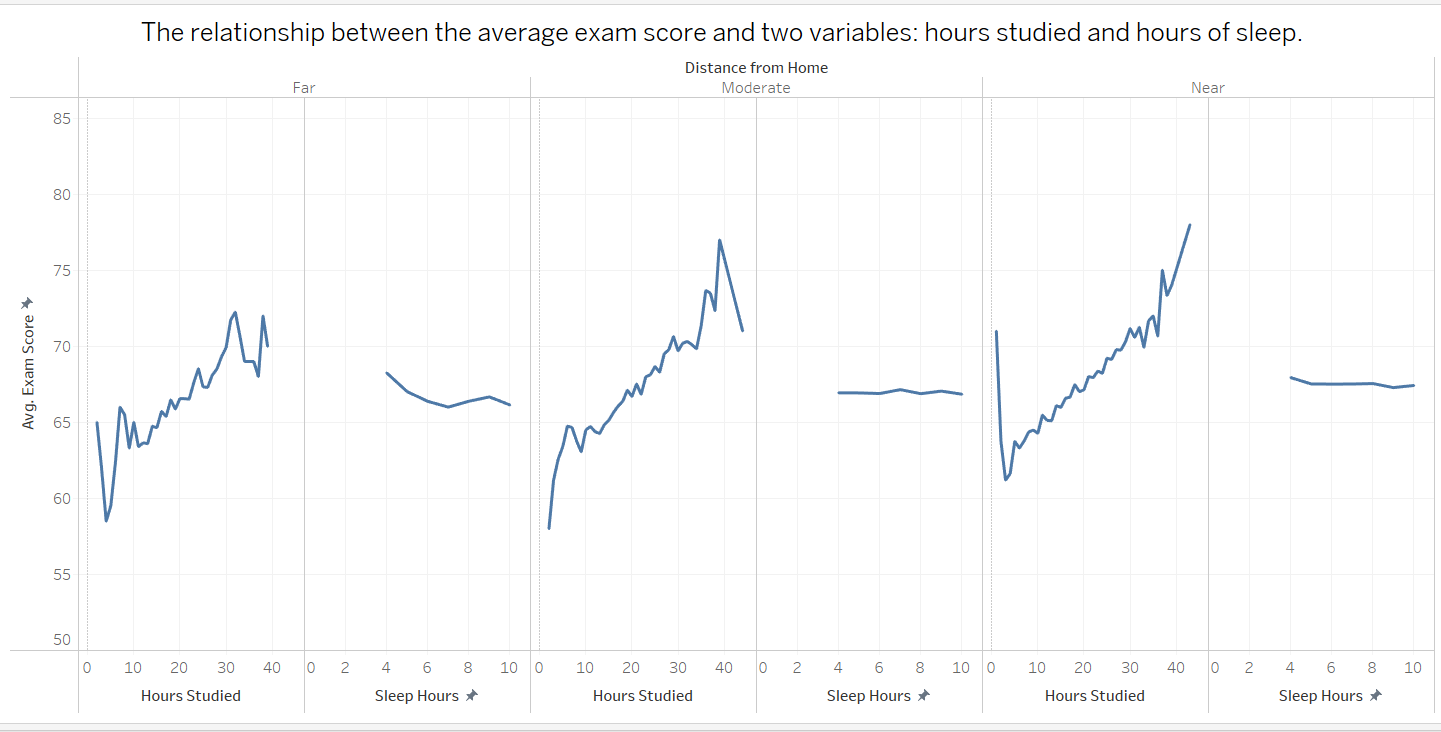
- If your project is not tied to a specific company, think creatively how your analysis could

be applied in industry settings, policy-making, or societal contexts.

At the end of the day, BE A STORYTELLER.

VISUALIZATIONS (TABLEAU)

1. Question 1:
   1. 
   2. Divided data into 2 categories, those who participated in ECs and those who didn’t.
   3. Physical Activity: it actually appears the more time spent on exercising actually boosts Exam scores.
      1. Could be due to the mental benefits from the stress relief of exercising
   4. Tutoring sessions:
      1. More = Higher Exam Scores
         1. \*To a certain extent. More than 6 sessions can cause burnout
      2. Tutoring Sessions w/ EC’s:
         1. Too much for students to handle if they have >5 tutoring sessions. Can see a large drop off in scores
   5. Comparison
      1. Student who aims to raise their test scores -> prioritize tutoring sessions as impact is higher
      2. Try not to get burnt out by having too many things on your plate
      3. In both charts, those who participate in EC’s normally score higher than those who don’t



The graph provides the display of the relationship between the average exam score and two variables: hours studied and hours of sleep. It is segmented into three categories based on the distance from home (far, moderate, near).