

## DATA STORY ONE

As the country's flagship university, the National University of Tropicorn (NUT) can only admit a limited number of students to its undergraduate programs each year. Recently, NUT Management has introduced a strategic initiative to better understand the role the university undergraduate programs have on the academic and professional development of their students.

NUT Management called for an investigation starting with the most recent graduating cohort. The aim of the investigation is to determine whether students' past academic performance during their junior college (JC) or polytechnic years is highly associated with their academic performance at the university. If the relationship is found to be the case, high achieving JC and polytechnic students at NUT would likely be expected to continue their high-achieving ways.

NUT Management reached out to the NUT Baccalaureate Analysis and Reporting Unit (NUT-BAR) to assist the investigation by providing statistical support and strategic recommendations. To kick off the call, a representative of NUT Management invited NUT-BAR to an in-person meeting held in a NUT Management conference room.

Management: "Good morning, everyone. It's a pleasure to have the team here. With the 2023 academic year officially closed, I am proud to announce that 5576 NUT undergraduates completed their programmes this year."

Team: "Congratulations to them!"

Management: "While we are pleased with the number of graduates and proud of how they've persevered over the years, NUT Management is mindful that to better plan for the future, the university needs to understand its past and present. The purpose of today's meeting is to explore students' academic performance, specifically their past and current academic records. As you are aware, our student selection process involves the thorough review of pre-entry tests and other academic qualifications. Past academic success may have led students to NUT; however, we strive to ensure that learning at NUT is fundamentally different from learning at JC and polytechnics. We want NUT to provide a different experience with different kinds of outcomes for our students."

With that in mind, NUT Management wants to know how strongly students' past academic achievement is related to their achievements at the university. This might help them rebalance their admission criterion and realign their four-year undergraduate programs going forward.

Management ceremoniously handed over the storage device containing the relevant files and dataset to NUT-BAR's team leader, Angela.

Management: "Everything you need to get started on is in this thumb drive. We know you'll keep it safe, Angela. We're hoping to get your team's conclusions and recommendations in short order."

Angela: "Thanks for trusting us. We will definitely do a full investigation into your question and share our considered recommendations."

Angela, as the team leader, was well-aware of NUT-BAR's capabilities. She had no reservations the team could quickly deliver on NUT Management's call. Before lunch, she relayed her interpretation of NUT Management's announcement and quickly scheduled a meeting for that afternoon.

Angela: "Good afternoon, team. As you've now heard twice, we received a task from the bosses this morning. In our possession, we have a dataset containing information on **5576** newly minted graduates. Our mission is to analyse the data and provide statistical support for higher-level decision-making. We know that we depend largely on students' pre-entry test scores to select qualified students as they are one important criterion in academic performance. Now we want to find out if students' past academic records look to have a significant influence on their academic performance at NUT. The focus of our analysis will be to examine the data and draw implications regarding the relationship between students' pre-entry scores and their 4.0-scale Grade Point Average (GPA). Could each of you please look at the data I just sent and provide some initial feedback on how we should propose a focused question or two?"

Jayce: "I think based on what I am looking at now, the dataset contains several items that are related to students' academic performance. We have **pre-entry test scores, GPAs in Terms 1 and 2, and their final cumulative GPAs.**"

Jennifer: "Yes, I think the rest of the items might not be as directly relevant to the learning experience to affect students' academic performance. Maybe we should look at the test scores and the GPAs?"

Luke: "I also agree with, Jenn. I think if the bosses want to know if students' past academic performance is influential or not, we can look at the relationship between their pre-entry test scores and their GPAs. See if there is any pattern."

Angela: "Good points, all. But do we want to focus on both of the early terms and the final one? Students went through four years, which was eight terms in total."

Luke: "I doubt that. Since we only have three GPA-related figures, maybe it's best for us to only look at the first and final. The second-term GPA may not be so different from the first-term one."

Jennifer: "Agreed. If we look at the first term, we can know the initial impact past academic performance may have on students. I assume maybe there are some effects?"

Jayce: "And if we look at the final term and the pre-entry scores, we can know if the past performance is significant to students after the four-year education."

Angela: "Good job, team! That's what I've been thinking of as well. So, my proposed two questions would be:

1. **How strongly are the past academic performances of students related to their subsequent academic performances at NUT?**
2. **How strongly are the past academic performances of students related to their cumulative academic performance at NUT?"**

The team took the initiative to look through the dataset. Jayce volunteered to do some preliminary data exploration.

Jayce: “I’m going to first explore the data a bit. We can start with something simple. Since the key variables are all numerical, we can calculate the means to see the averages of the pre-entry scores and GPAs. If we’re calculating means, we might as well do the standard deviations as well to see the spread of data. Maybe that will set us up to observe something interesting later.”

Figure 1 displays the distribution of Pre-Entry Scores (Mean=90.17, Standard Deviation=5.91). The median is 90.28. From looking at the figure, one can see there are about 1000 students who entered NUT with entry scores ranging between 80 and 85 and more than 1500 with entry scores ranging from 90 to 95. The interval 90 to 95 corresponding to the peak in the figure happens to contain the median.

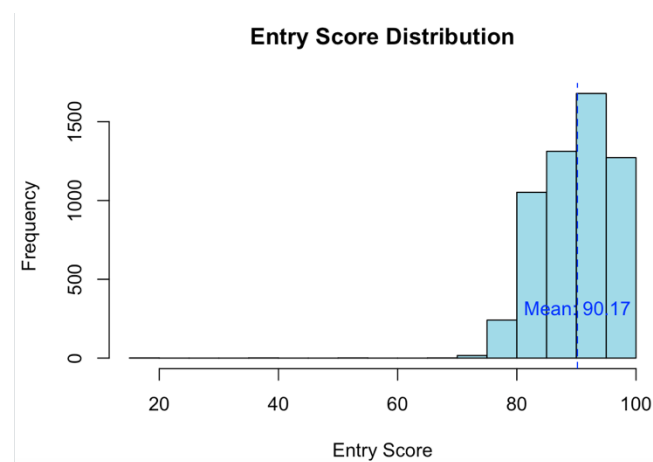


Figure 1

The distribution of Term 1 GPA (GPA\_T1) values (Mean=3.24, Standard Deviation=0.36) is shown in Figure 2. Students' Term 1 GPAs range from 1.60 to 4.00. The median is 3.28.

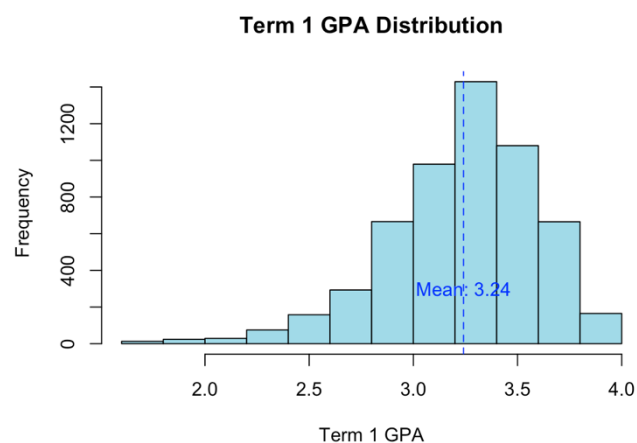


Figure 2

The distribution of Final Term Cumulative GPA values (Mean=3.30, Standard Deviation=0.27) is shown in Figure 3. Students' Final Term GPAs range between 2.17 and 3.94. The median is 3.32.

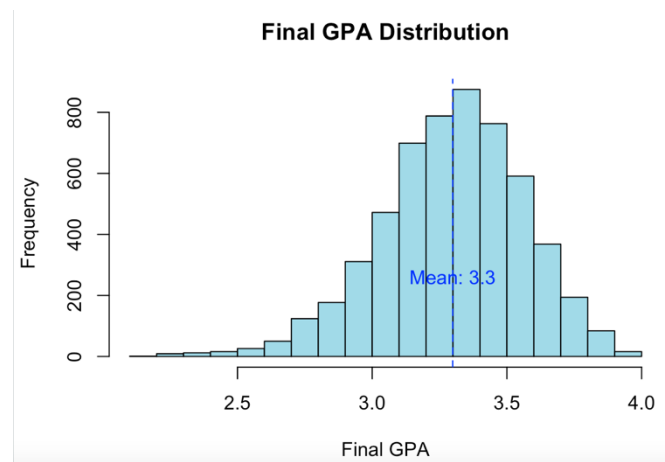


Figure 3

Jayce: “That’s a quite high mean pre-entry score, at 90.17, considering the highest pre-entry score is at 100.”

Luke: “Yes. At the same time, we can observe that there are some differences between the means of the first-term GPAs and the final-year GPAs. I see a 0.06 difference.”

Jennifer: “They may be different, but let’s avoid jumping to conclusions. A comparison of the means or medians isn’t sufficient without more context. We must look at the relationship using other methods.”

Angela: “Good work, team! As we have now vetted the dataset to the point where it looks like it can be used to answer the two questions, let’s meet tomorrow to discuss how to push forward. I’m sure you all will come with your own ideas. Until then, enjoy your evening.”