STUDENT ACADEMIC AND CAREER TRENDS ANALYSIS

Insights using SQL

BY Dipali Patel

Abstract:

This project leverages SQL to analyse and interpret student data from two interconnected datasets: one detailing academic and personal characteristics and another focusing on career aspirations and technical skills. By integrating these datasets, the project aims to uncover meaningful relationships between academic performance, family background, extracurricular activities, and future career goals.

Key objectives include identifying top-performing students, evaluating the impact of study time and family size on GPA, and exploring the alignment between technical skills (e.g., Python and SQL proficiency) and desired career paths. The project also investigates factors influencing academic success, such as internet access, romantic relationships, and parental education levels.

The analysis utilizes advanced SQL techniques to group, filter, and rank students based on various metrics, providing actionable insights for educators and policymakers. The findings highlight patterns in academic achievement and skill acquisition, supporting data-driven decisions to enhance student learning and career readiness.

Introduction:

In the modern educational landscape, understanding the factors contributing to academic success and career preparation is essential. This project examines the relationships between various student attributes, such as family background, study habits, technical skills, and career goals. Using SQL as the primary analytical tool and Excel for data visualization, the project seeks to:

- Identify top-performing students across different categories.
- Evaluate the influence of family size, romantic relationships, and study time on academic performance.
- Explore the alignment between technical skills (e.g., SQL and Java proficiency) and career aspirations.
- Provide actionable recommendations for stakeholders to improve student outcomes.

Objectives:

1. To identify top-performing students based on GPA and final grades.
2. To evaluate the relationship between study habits, family size, and academic performance.
3. To analyse the alignment between technical skills and career aspirations.
4. To provide actionable recommendations for enhancing academic and career outcomes.

Methodology:

1. Data Integration: The two datasets were merged using the common field StudentID to establish a comprehensive view of each student.

2. SQL Queries:

- Ranking students based on GPA and final grades.
- Grouping and filtering data by attributes like family size, study time, and technical skills.
- Calculating averages and identifying top performers.

3. Insights Extraction: Patterns and trends were derived using SQL functions like aggregation, ranking, and conditional filtering.

4. Visualization: Excel was utilized to create visual representations of the data,to simplify trends analysis.

Tools and Technologies:

The analysis for this report was conducted using SQL and Excel, serving distinct roles in data processing and evaluation.

- **SQL**: SQL was employed for data querying and extraction, enabling detailed analysis of various factors such as academic performance, family dynamics, and technical proficiency. Its ability to handle complex queries helped uncover valuable insights from the datasets.
- **Excel**: Excel played a key role in the initial stages of data exploration and visualization. It allowed for quick data reviews, identification of trends or irregularities, and summarization of findings into concise tables and charts.

These tools ensured a streamlined approach to extracting, analyzing, and presenting data in a meaningful way.

Interpretation and Business Recommendations:

How can I retrieve the final grades and GPAs of all students along with their interested domains?

- The data provides insights into student performance based on their final grades, GPAs, and areas of interest. It highlights diverse interests like Artificial Intelligence, Cybersecurity, Web Development, and Software Engineering, with GPAs ranging from 3.2 to 3.9 and final grades varying widely. Students interested in domains like Machine Learning and Cloud Computing often show high GPAs, reflecting strong academic performance. Conversely, some areas like Data Science and Blockchain Technology have relatively lower final grades, indicating potential challenges or a need for targeted academic support.
- **Recommendations:** Focus on strengthening curricula for domains with lower grades to address potential knowledge gaps. Provide tailored guidance for students in high-demand fields like AI and Machine Learning to maintain performance levels and encourage career advancement. Offer workshops or mentoring for fields with mixed outcomes, such as Web Development and Cybersecurity, to bridge the gap between high and low performers.

Who are the students with a GPA greater than 3.5, and what are their final grades?

- The list identifies students with a GPA greater than 3.5 and their corresponding final grades. It showcases a wide range of academic performance, with final grades spanning from 0 to 17. Notably, students achieving GPAs of 3.8 and 3.9 often tend to score higher final grades, reflecting their academic consistency.
- **Recommendations:**

- High Performers: Students with GPAs of 3.8 or 3.9 and high final grades (e.g., IDs 15, 33, 58, 110) could be encouraged for advanced learning opportunities, scholarships, or leadership roles.
- Low Final Grades: Students with low grades but high GPAs (e.g., IDs 19, 53, 126) may need personalized support to align practical performance with theoretical strengths. This could involve targeted skill-building workshops or tutoring.

What is the alignment between students' Python proficiency and their academic performance?

- The alignment between students' Python proficiency and academic performance shows a varied relationship. Students with strong Python skills generally align well with higher academic performance, as seen in students with strong proficiency and good grades (e.g., IDs 3, 6, 15, 33). However, there are also students with strong Python skills who show partial alignment or need improvement in academic results, indicating that proficiency in Python doesn't always correlate with top academic performance.
- On the other hand, students with weaker Python skills often show lower academic performance, though there are exceptions where students with weak proficiency still achieve decent grades. This suggests that while Python proficiency is an asset, other factors, like study habits and subject comprehension, also play a critical role.
- **Recommendation:**
  - For students with strong Python skills but poor alignment (e.g., IDs 4, 19, 53), focus on identifying and addressing other potential academic challenges, such as time management or test-taking strategies.
  - For students with weak Python skills, consider providing targeted support in Python to bridge the gap, while also improving their overall study techniques and subject understanding to raise their academic performance.

Which students are interested in Machine Learning, and what are their grades and technical skills?

- The students interested in Machine Learning, as identified by their strong Python proficiency, include Student 6, 17, 28, 39, 87, 121, 138, 145, 151, 158, 165, 171, and 178. Their final grades range from 8 to 14, with GPAs mostly clustered around 3.9, except for Students 158 and 178, who have GPAs of 3.4. These students demonstrate strong Python skills, which are essential for Machine Learning, but their SQL proficiency is generally average, and their Java proficiency is consistently weak. For instance, Student 17 has an excellent final grade of 14 and a GPA of 3.9, with strong Python skills but only average SQL skills and weak Java skills. Similarly, Student 121, with a final grade of 14 and a GPA of 3.8, exhibits strong Python capabilities but weaker SQL and Java skills. This pattern indicates a strong alignment with Python, which is critical for Machine Learning, while SQL and Java appear to be secondary priorities for these students.

- **Recommendations:**
  To strengthen their technical foundation for Machine Learning, these students should focus on improving their SQL skills, as database querying and management are integral to handling large datasets. Additionally, enhancing Java proficiency can be beneficial for projects requiring platform-independent applications or integrating with existing enterprise systems. Providing targeted workshops or hands-on projects in SQL and Java, along with more advanced Python-based Machine Learning modules, will ensure a well-rounded skillset. This will not only align their technical capabilities with industry expectations but also increase their overall competency and employability in Machine Learning domains.

What is the future career preference of students with high grades (G3 >= 14) and a GPA above 3.7?

- Students with high grades (FinalGrade ≥ 14) and a GPA above 3.7 include Student 17 (FinalGrade 14, GPA 3.9) and Student 121 (FinalGrade 14, GPA 3.8). Both students demonstrate strong Python skills, which are essential for technical and data-oriented careers. However, their proficiency in SQL and Java is either average or weak, suggesting that their technical foundation is not yet well-rounded. This highlights a likely preference for careers in fields like Data Science, Machine Learning, or Artificial Intelligence, where Python is a core skill and SQL/Java may be supplementary.
- **Recommendations:**
  To support their career aspirations, these students should prioritize improving their SQL skills to handle databases effectively and consider strengthening their Java proficiency for broader versatility. Offering targeted training in these areas, along with advanced Python-based projects and exposure to real-world applications, will better prepare them for competitive roles in the tech industry.

How many students live in urban areas and are studying Computer Science?

- There are 150 students who live in urban areas and are studying Computer Science. This suggests a significant interest in Computer Science among urban students, likely due to better access to educational resources, technology, and career opportunities in urban regions.
- **Recommendations:**
  To further support these students, institutions should focus on providing specialized Computer Science workshops, access to advanced technology labs, and internships with urban tech companies. Additionally, fostering partnerships with industry leaders can enhance their practical knowledge and employability in competitive fields like software development, AI, and data analytics**.**

What is the average GPA and final grade for students who have "Strong" SQL proficiency?

- The average GPA for students with "Strong" SQL proficiency is 3.62, and their average final grade is 11.89. This indicates that students with strong SQL skills tend to perform

well academically, suggesting a positive correlation between SQL proficiency and overall academic success.

- **Recommendations:**
  To further improve academic outcomes, institutions should encourage students to enhance their SQL skills through targeted training, real-world database projects, and certifications. Providing additional resources and mentorship in SQL can help students achieve higher proficiency, which may positively impact their overall performance.

What are the future career aspirations of students whose mothers work "at home" and who have achieved a final grade above 12?

- Students whose mothers work "at home" and have achieved a final grade above 12 aspire to future careers in cutting-edge technology fields. Their career preferences include becoming a DevOps Engineer, Machine Learning Engineer, Data Scientist, AI Researcher, and Cloud Solutions Architect. These roles indicate a strong interest in advanced computing, artificial intelligence, and cloud technologies.
- **Recommendations:**
  To support these students in achieving their aspirations, institutions should provide access to relevant resources, such as specialized courses, hands-on projects, and mentorship in their desired fields. Encouraging participation in internships, tech boot camps, and industry networking events can also help them build the skills and connections necessary for these high-demand roles.

Which students with "Strong" Python skills have fathers working as teachers and achieved a final grade above 13?

- Students with "Strong" Python skills whose fathers work as teachers and who have achieved a final grade above 13 are interested in pursuing careers in technology. Their career aspirations include becoming a UX Designer, AI Researcher, and Software Engineer.
- **Recommendations:**
  To help these students achieve their career goals, providing specialized training in user experience design, artificial intelligence, and software development is crucial. Additionally, encouraging involvement in relevant extracurricular activities like coding competitions, hackathons, and internships can help them gain practical experience and networking opportunities in their chosen fields.

What is the average GPA of students who participate in extracurricular activities and have a final grade greater than 10?

- The average GPA of students who participate in extracurricular activities and have a final grade greater than 10 is 3.63.
- **Recommendation:**
  Encouraging more students to engage in extracurricular activities could be beneficial, as it is associated with a higher average GPA. Schools should continue promoting

activities that help students develop time management, teamwork, and leadership skills, which could positively impact both their academic performance and personal growth.

What are the technical skills of students who live in urban areas and are interested in Artificial Intelligence?

- The students who live in urban areas and are interested in Artificial Intelligence have the following technical skills:
    - Python: All of them have "Strong" Python skills.
    - SQL: Most of the students have "Average" SQL skills, with a few having "Strong" or "Weak" skills.
    - Java: All students have "Weak" Java skills, except for one student who has "Average" Java skills.
- **Recommendation:**
  To enhance their potential in Artificial Intelligence, these students should focus on strengthening their SQL and Java skills. Building proficiency in these areas will make them more well-rounded and better prepared for a wider range of AI-related roles.

Which students with strong Java skills have a family relationship score greater than 4, and what are their GPAs?

- The students with strong Java skills and a family relationship score greater than 4 have the following GPAs:
    - GPAs range from 3.3 to 3.9 with the majority of students having GPAs around 3.5 to 3.8.
    - These students have demonstrated strong Java skills, with a consistent Family Relationship Score of 5, indicating a positive family environment.
- **Recommendation:**
  While these students have strong Java skills, their GPAs indicate room for improvement in other areas. It may be beneficial for them to focus on strengthening their overall academic performance, particularly in subjects outside of Java. Leveraging their strong family support, they could also benefit from seeking additional resources such as tutoring or mentorship in other technical skills to further enhance their capabilities and achieve a more balanced GPA.

What are the future career goals and final grades of students who have no academic failures and whose primary guardian is their mother?

- The students whose primary guardian is their mother and who have no academic failures have varied career aspirations. These students have expressed interest in a wide range of fields, including Machine Learning, Software Engineering, Web Development, Data Science, Cloud Solutions, and AI Research. Their final grades vary from 8 to 18,

with a significant number of students achieving final grades above 12. Many students with final grades over 12 aim for high-tech, emerging career paths such as Machine Learning Engineer, Data Scientist, Mobile App Developer, and AI Researcher.

- **Recommendation:** For students with high final grades (above 12), continued focus on developing technical skills and deepening knowledge in their chosen fields will be key to success. It's essential for these students to leverage their strong academic records and seek internships, workshops, and networking opportunities to gain hands-on experience. For students aiming for careers in AI, Machine Learning, and Cloud Solutions, further specialization in programming languages, data analysis tools, and system architecture will be valuable. Students with lower grades should consider additional academic support to improve their foundational knowledge, which will help them align with their career aspirations.

Which students with GPAs above 3.7 are not in a romantic relationship and also live in a small family?

- The students with GPAs above 3.7 who are not in a romantic relationship and live in small families (LE3) include those with GPAs ranging from 3.8 to 3.9. These students demonstrate strong academic performance and, given their family dynamics, may benefit from personalized academic support. Notably, they are all not involved in romantic relationships, which might allow them to dedicate more time to their studies or extracurricular activities.
- **Recommendation:** These students should consider using their academic focus to explore research opportunities, internships, or advanced learning in their fields of interest. Since they are not currently in romantic relationships, they have more time to engage in extracurriculars, network with professionals, or seek out mentorship opportunities that could advance their careers. Additionally, fostering well-being and balancing academic pursuits with personal development can help them maintain motivation and a healthy study-life balance.

How many students from small families, study time is 3 or less than 3 and have each GPA?

- The majority of students in this group have GPAs between 3.3 and 3.8, suggesting that while a few students achieve higher academic performance (GPA $\geq$ 3.9), many others fall within the average range of 3.3 to 3.8. This indicates that although these students are from small families and have limited study hours, there is still a significant number with relatively good academic standing.
- **Recommendation:** To enhance academic performance, these students might benefit from improving their study habits and time management. Even with limited study hours, students could adopt more efficient study techniques, such as active learning or focused study sessions, to boost their GPA. Additionally, students with lower GPAs may need additional support in time management or access to resources such as tutoring or study groups to improve their results.

What is the relationship between study time and GPA for students interested in Web Development?

- The relationship between study time and GPA for students interested in Web Development indicates a positive trend. Students who devote more study time tend to achieve slightly higher GPAs. For example:
    - A student with the most study time scored a GPA of approximately 3.67.
    - Those with moderate and lower study times achieved GPAs of 3.63 and 3.61, respectively.
    - While the differences in GPA are relatively small, they suggest that consistent study time positively impacts academic performance for Web Development students.
- **Recommendation:** Encourage students to dedicate regular and focused study time to reinforce their learning. Additionally, monitor their engagement in practical Web Development activities to ensure theoretical study is complemented by hands-on experience, which can enhance both understanding and performance.

What are the GPAs and final grades of students who do not have internet access at home but are pursuing careers in Data Science?

- The data indicates that students without internet access at home, who are pursuing careers in Data Science, may face challenges in achieving higher academic outcomes. For instance, the student identified in the dataset has a GPA of 3.6 and a final grade of 10, which are moderate but may reflect the constraints of limited resources affecting their performance.
- **Recommendation:** To support such students, institutions should consider providing alternative resources, such as access to on-campus internet, offline learning materials, or mentorship programs. This can help bridge the digital divide and enhance their ability to excel in competitive fields like Data Science.

What is the average GPA of students grouped by family size and romantic relationship status?

- The average GPA of students varies based on family size and romantic relationship status. Students from families larger than three members (GT3) who are in a romantic relationship have the highest average GPA of 3.65, closely followed by those not in a relationship with an average GPA of 3.62. For families with three or fewer members (LE3), students not in a relationship have a slightly higher average GPA of 3.59 compared to those in a relationship, who have the lowest average GPA of 3.55.
- **Recommendation:** The data suggests that family size and relationship status might influence academic performance. Students in smaller families or relationships may benefit from tailored academic support or counselling to balance personal and academic responsibilities, ensuring they achieve their full potential.

Who are the top-performing students (highest GPA) in small families who are not in romantic relationships?

- The top-performing students from small families (LE3) who are not in romantic relationships both achieved a GPA of **3.9**, showcasing exceptional academic performance. These students are **student 90** and **student 6**, both ranked as the highest achievers within this demographic.
- **Recommendation:** To support and motivate these top-performing students, further exploration into the factors contributing to their success is recommended. This can include their study habits, time management, or external support systems. Additionally, recognizing their achievements through awards or mentorship opportunities could inspire peers in similar demographics.

What is the GPA distribution of students not in romantic relationships, grouped by their mother's education level?

- The GPA results suggest that students whose mothers have higher education levels (levels 1 and 4) tend to perform slightly better than those with lower educational backgrounds. However, the differences in GPA are not significant, and all groups have relatively high average GPAs. This indicates that factors other than maternal education, such as individual motivation or access to resources, may also play important roles in these students' academic success.
- **Recommendation:** To support students from families with lower educational backgrounds, it would be beneficial to provide additional resources such as mentorship programs, academic counselling, or study groups. These interventions can help improve their performance, ensuring they have equal opportunities to succeed academically. Encouraging the involvement of both students and their families in educational activities may further contribute to their academic achievements, regardless of the mother's education level.

Which students in small families have the top 3 highest final grades for each GPA range?

- For students from small families, the top 3 highest final grades within each GPA range indicate their strong academic performance despite their family size. For students with a GPA of **3.9**, the highest final grade is **14** (student 110), followed by **13** (student 6) and **10** (student 90). Similarly, in the **3.8 GPA range**, the top grades are **14** (student 37), **12** (student 3), and **11** shared by students 155, 83, and 133. Students in the **3.7 GPA range** show the highest final grade of **15** (student 70), with **14** (student 23) and **13** shared by multiple students.
- In lower GPA ranges, exceptional final grades are still achieved. For instance, at a **GPA of 3.6**, students 111 and 59 scored **14**, while students with a **GPA of 3.5** achieved remarkable scores of **18** (student 114) and **16** (student 66). These results suggest that strong academic performance in specific courses can exist independently of GPA, indicating possible subject-specific strengths among these students.
- **Recommendation:** This data highlights the need to recognize and encourage high-performing students within each GPA range, particularly in small families. Academic

interventions should focus on understanding the factors that drive these performances to replicate them across broader contexts, ensuring all students achieve their potential.

Business Recommendations

- **Curriculum Development:** Incorporate technical skills training, particularly in Python and SQL, as part of the standard curriculum to enhance career readiness.
- **Parental Engagement:** Promote parental involvement in education, especially for students from small families or with stay-at-home parents.
- **Extracurricular Support:** Encourage participation in extracurricular activities to foster well-rounded development and improve academic outcomes.
- **Targeted Counselling:** Provide career counseling tailored to students' technical proficiencies and career aspirations.
- **Infrastructure Improvement:** Ensure access to resources like the internet and technology for all students, bridging the gap for those lacking such facilities.
- **Data-Driven Decision Making:** Use data analysis to continuously evaluate and refine educational strategies, ensuring alignment with student needs and industry demands.

Conclusion

The analysis provided an in-depth understanding of the factors influencing student performance and career aspirations. Key findings highlighted the role of technical skills such as Python and SQL in academic success and career alignment. Additionally, family dynamics, study time, and extracurricular participation emerged as critical factors in shaping student outcomes. Students with strong technical skills consistently displayed better academic performance, emphasizing the importance of integrating such skills into educational curricula.

The project also underscored the importance of personalized support for students based on their unique family and educational backgrounds. These insights can help educators and policymakers design more effective learning environments and career counselling programs.