



UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação
Departamento de Ciências de Computação
SCC0252 - Visualização Computacional

Projeto de Exploração e Análise Visual de Dados: Entrega 1

Relatório de processamento, elaboração e visualização de dados de desempenho e percurso acadêmico de alunos da graduação por meio de exploração e análise de visualizações.

Clara Ernesto de Carvalho - 14559479
Felipe Carneiro Machado - 14569373
Renan Parpinelli Scarpin - 14712188
Gabriel Barbosa dos Santos - 14613991
Lourenço de Salles Roselino - 11796805

USP – São Carlos
Outubro de 2025

Introdução.....	3
Conjunto de Dados.....	4
Tabela dos Alunos.....	4
Tabela das Disciplinas.....	4
Análises e questões preliminares.....	5
Pré-processamento.....	6
Discussão.....	7
Limitações e críticas.....	8
Próximos passos.....	9

Introdução

Durante o processo de graduação é comum o estudante perder a perspectiva da sua progressão no curso, especialmente após a conclusão das disciplinas obrigatórias é produtivo refletir sobre o progresso realizado durante os semestres fundamentais, entender seus próprios pontos fortes e fracos e identificar as tendências em seu desempenho para poder tomar decisões racionais no restante de sua trajetória acadêmica e profissional.

Do ponto de vista da gestão e coordenação do curso é de interesse compreender os padrões entre os seus alunos, suas preferências e comportamentos durante os semestres, entender que padrões são relacionados a problemas e quais caminhos estão sendo seguidos permite estruturar o currículo e a oferta de disciplinas de forma a aumentar a qualidade de ensino do instituto.

Para estes dois problemas e perspectivas este trabalho se propõe a processar dados e explorar técnicas para melhor compreensão do percurso escolar de estudantes da graduação, buscando auxiliar estudantes e coordenação na visualização de dados e tomada de decisões pertinentes.

Conjunto de Dados

Para este projeto foram utilizados dados coletados do histórico escolar dos membros do grupo totalizando 5 históricos extraídos. Eles foram gerados pela plataforma Jupiterweb e separados em 2 tabelas, 1 para informações gerais dos estudantes e outra para as disciplinas cursadas pelos semestres anteriores.

Tabela dos Alunos

A tabela de alunos contém uma linha para cada histórico extraído, possuindo as informações de idade, curso, ano de ingresso, média ponderada, classificação na turma, ano de conclusão estimado, carga horária em AAC e carga horária em AEX

Importante destacar que 4 dos 5 estudantes ingressaram no ano de 2023 e todos os 5 são estudantes do Bacharelado de Ciências da Computação.

Tabela das Disciplinas

A tabela de disciplinas contém 224 linhas no total, ela possui informações de todos os históricos extraídos contendo colunas para a Sigla, para o nome da disciplina, o ano e semestre que ela foi cursada, seus créditos aula, crédito trabalho e a soma dos dois, também possui a carga horária EXT, a frequência e nota final além de um identificador do aluno que cursou a disciplina.

Não foram inseridos dados das disciplinas sendo cursadas no segundo semestre de 2025 por ainda não conterem as informações de nota e frequência.

Análises e questões preliminares

Este trabalho se propõe em guiar a interpretação destes dados por três questões principais:

- Existe relação de frequência e nota com o departamento da disciplina?
- Existe uma preferência por departamentos entre os estudantes?
- Como as escolhas de grade mudam ao longo da graduação?

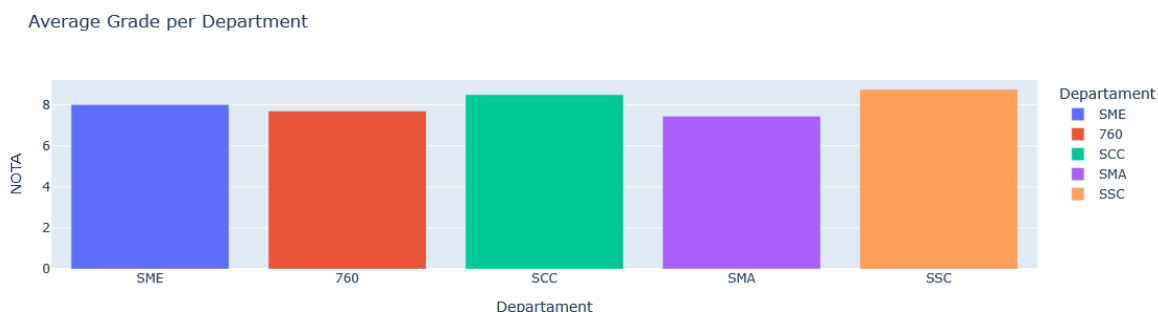
Essas questões se mostram relevantes para estudantes e gestores e vão ser exploradas e desenvolvidas ao longo deste relatório.

Pré-processamento

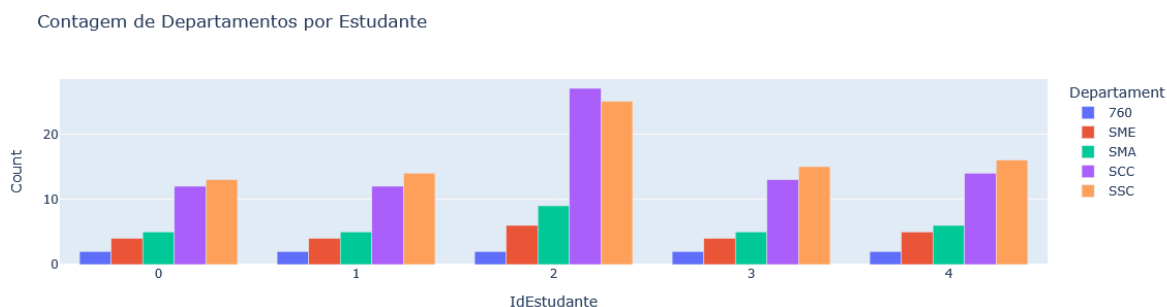
As disciplinas sendo cursadas no momento da coleta foram removidas por terem valor nulo em nota e frequência e não participarem dessa análise. As tabelas de disciplinas individuais dos alunos foram unificadas, tendo suas colunas reordenadas e seus tipos padronizados. Os valores nulos não referentes a nota ou presença foram transformados em 0, por ser condizente com o sentido semântico.

Discussão

No gráfico de nota média por departamento podemos identificar uma maior média entre os departamentos SCC e SSC e um valor significativamente menor para o departamento SMA e 760, esse comportamento é paralelo ao que pode ser visto na contagem de departamentos por estudante, houve uma tendência de uma quantidade maior de disciplinas dos departamentos SCC e SSC do que dos SMA e 760.



Nesta mesma visualização podemos ver uma proporção a normal entre as disciplinas dos departamentos SCC e SSC com o resto no estudante 2, sendo uma diferença maior até do que a apresentada entre os outros estudantes, uma interpretação possível é de que os alunos teriam uma preferência por escolher optativas desses departamentos e o estudante 2 sendo o único ingressante de 2020 teria tido mais semestres para demonstrar este comportamento, mas para verificarmos essa hipótese seria necessário coletar mais dados de outros anos de ingresso.



Olhando para o gráfico de linha da média de crédito aula por semestre vemos um crescimento do primeiro ao segundo semestre seguido de uma forte queda, isso sugere uma estratégia de experimentação entre os alunos nos primeiros semestres, sendo o segundo a primeira oportunidade do estudante escolher sua grade é possível que ele se matricule em mais créditos em um semestre compensando em uma quantidade menor no seguinte.

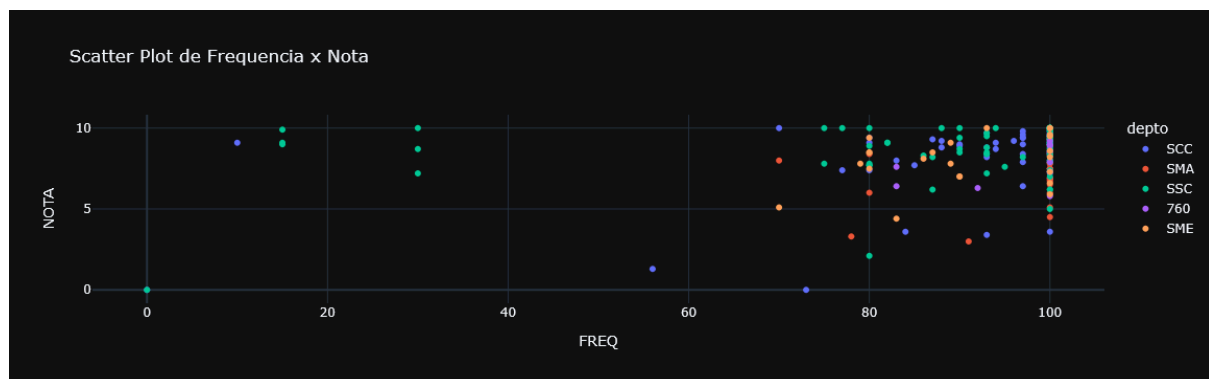


Um padrão paralelo ocorre no gráfico de linha de nota por semestre, existe uma queda do primeiro ao segundo semestre seguido de uma melhora do segundo ao terceiro, sugerindo que a estratégia de compensar a quantidade de créditos no semestre é acompanhada de mudanças na nota média.

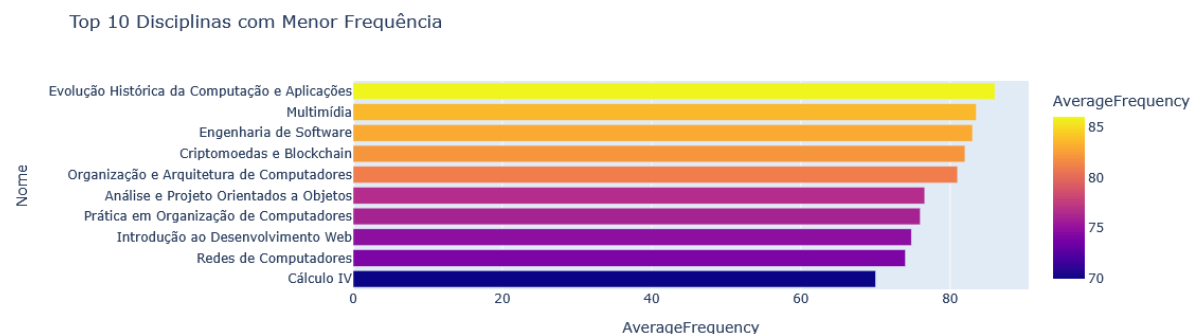


No Scatter Plot de Frequência x Nota podemos notar uma falta de disciplinas na área entre 30% a 70%, sendo 70% o limite inferior de presença para poder ser aprovado em uma disciplina, estudantes não têm incentivo para continuarem frequentando as aulas após faltarem além deste limite, logo ficam acima dela ou próximos de 0%. No mesmo gráfico também não conseguimos identificar

visualmente uma relação entre nota e faltas em nenhum dos departamentos, mas podemos ver uma maioria na marca de 100% de presença, especialmente para o departamento SMA.



Observando as disciplinas com maior quantidade de faltas temos um grande número de disciplinas do departamento SSC, o que se é esperado considerada a maior quantidade de disciplinas cursadas neste departamento, mas não se vê a mesma representação do departamento de SCC, que também tem uma contagem geral alta.



Limitações e críticas

A baixa quantidade de amostras não permite fazer análises mais confiantes dos padrões apresentados, além da pouca diversidade de cursos e anos de ingresso, apesar de estudar apenas alunos de um único curso seja interessante para entender as tendências dentro dele também seria valioso observar as relações entre os estudantes de outros cursos do mesmo instituto.

Também é importante ressaltar que, além de existir pouca variedade de ano de ingresso, o único estudante de antes de 2023 é também o único que teve semestres durante a pandemia do COVID, uma análise mais complexa seria possível com dados de mais alunos que fizeram graduação durante este período.

Outro ponto a ser comentado é a falta de dados individuais sobre as disciplinas, as visualizações utilizaram apenas as informações apresentadas nos históricos escolares, como quantidade de créditos e departamento, seria interessante expandir neste ponto e extrair dados como ementa, requisitos e contribuição para especialização no curso, abrindo possibilidades para outras visualizações.

Próximos passos

Na próxima entrega iremos incluir informações extras sobre as disciplinas, buscando estudar relações entre disciplinas com requisitos e escolhas de optativas, com a inclusão de outras opções de interação com as visualizações vamos expandir o uso e possibilidades para estudantes buscando compreender as opções e oportunidades dentro do curso.