



**UNIVERSIDADE FEDERAL DE CAMPINA GRANDE
CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA
UNIDADE ACADÊMICA DE SISTEMAS E COMPUTAÇÃO
CURSO DE BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

MATHEUS HENRIQUE ALVES LEAL

**O COMPORTAMENTO DOS ALUNOS SOB AVALIAÇÃO EM
PROGRAMAÇÃO ORIENTADA A OBJETOS**

CAMPINA GRANDE - PB

2019

MATHEUS HENRIQUE ALVES LEAL

**O COMPORTAMENTO DOS ALUNOS SOB AVALIAÇÃO EM
PROGRAMAÇÃO ORIENTADA A OBJETOS**

**Trabalho de Conclusão Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharela em
Ciência da Computação.**

Orientadora: Professora Dra. Livia Maria Sampaio Campos.

CAMPINA GRANDE - PB

2019



L435c Leal, Matheus Henrique Alves.

O comportamento dos alunos sob avaliação em programação orientada a objetos. / Matheus Henrique Alves Leal. - 2019.

12 f.

Orientadora: Profa. Dra. Livia Maria Rodrigues Sampaio Campos.

Trabalho de Conclusão de Curso - Artigo (Curso de Bacharelado em Ciência da Computação) - Universidade Federal de Campina Grande; Centro de Engenharia Elétrica e Informática.

1. Programação orientada a objetos. 2. Disciplina Laboratório de Programação II - UFCG. 3. Avaliação de alunos - disciplina de programação - UFCG. 4. Regressão linear. I. Campos, Livia Maria Rodrigues Sampaio. II. Título.

CDU:004(045)

Elaboração da Ficha Catalográfica:

Johnny Rodrigues Barbosa
Bibliotecário-Documentalista
CRB-15/626

MATHEUS HENRIQUE ALVES LEAL

**O COMPORTAMENTO DOS ALUNOS SOB AVALIAÇÃO EM
PROGRAMAÇÃO ORIENTADA A OBJETOS**

**Trabalho de Conclusão Curso
apresentado ao Curso Bacharelado em
Ciência da Computação do Centro de
Engenharia Elétrica e Informática da
Universidade Federal de Campina
Grande, como requisito parcial para
obtenção do título de Bacharel em Ciência
da Computação.**

BANCA EXAMINADORA:

**Professora Dra. Lívia Maria Rodrigues Sampaio Campos
Orientadora – UASC/CEEI/UFCG**

**Professora Dra. Joseana Macêdo Fachine Regis de Araújo
Examinadora – UASC/CEEI/UFCG**

**Professor Dr. Tiago Lima Massoni
Disciplina TCC – UASC/CEEI/UFCG**

Trabalho aprovado em: 25 de novembro de 2019.

CAMPINA GRANDE - PB

O comportamento dos alunos sob avaliação em Programação Orientada a Objetos

Uma análise estatística do desempenho dos estudantes da disciplina Laboratório de Programação II (LPII) do curso de Ciência da Computação da UFCG
Trabalho de Conclusão de Curso

Matheus Henrique Alves Leal
matheus.leal@ccc.ufcg.edu.br
Universidade Federal de Campina Grande
Campina Grande, Paraíba

Lívia Maria Rodrigues Sampaio Campos *
livia@computacao.ufcg.edu.br
Universidade Federal de Campina Grande
Campina Grande, Paraíba

RESUMO

A disciplina Laboratório de Programação II (LPII) foi concebida para desenvolver nos estudantes habilidades em programação Orientada a Objetos e é oferecida no segundo período do curso para uma quantidade média de 90 alunos. A metodologia de avaliação da disciplina envolve diferentes tipos de atividades, como provas, atividades práticas em sala e projeto, porém, não é fácil identificar o impacto de cada avaliação no desempenho dos alunos nem o quanto elas estão relacionadas ou se estão relacionadas. Sem essas informações, não é possível direcionar bem os alunos ao longo da disciplina ou realizar ajustes adequados no sistema de avaliação de modo a contribuir de forma mais efetiva para a melhoria do ensino em LPII. Neste trabalho, realizamos um estudo sobre o desempenho dos alunos nas diferentes atividades avaliativas para entender a sua relação e o perfil dos alunos nesse contexto. Nele, observou-se que embora as avaliações não mostrem uma correlação expressiva, foi possível identificar grupos de alunos com comportamentos específicos, assim como observamos que, mesmo que alunos apresentem notas finais semelhantes, suas trajetórias ao longo da disciplina podem ter sido bem diferentes.

1 INTRODUÇÃO

Os processos de avaliação de aprendizado vêm sendo objeto de estudo de pesquisadores com certa frequência, seja na educação básica ou ensino superior. Em ambos os casos eles esbarram na mesma problemática: os professores estão avaliando os estudantes de maneira eficaz? O quão bem a metodologia de avaliação consegue medir o nível de aprendizado real do aluno?

Pensando em responder esses questionamentos, é discutido na literatura como moldar os métodos de avaliação para melhor se adequar aos alunos. Salles e Melo dizem que “...ao ter acesso ao perfil delineado pelos discentes, os professores puderam verificar que características já conheciam e o que ainda faltava conhecer. Puderam também avaliar a relevância de ter ciência do perfil da turma logo no início do ano.” [1]. Esse estudo sugere que ter conhecimento sobre o perfil do aluno pode ajudar no processo de ensino-aprendizagem.

Identificar os perfis de alunos quando se consideram turmas volumosas, tendo esses alunos diferentes experiências, não é uma tarefa

trivial. Buscar medir o aprendizado do aluno continuamente, a fim de identificar alunos em situação de risco a tempo de intervenção é uma das alternativas disponíveis para mitigar o problema, como aludido por Hernández (2012) [3].

No contexto do nosso trabalho, a metodologia de avaliação da disciplina Laboratório de Programação II (LPII) se dá por meio de atividades práticas individuais no formato de mini-projetos, denominados de laboratórios; provas práticas individuais; e projeto em grupo. Além disso, em alguns períodos letivos, foi também considerado o desempenho do aluno na disciplina teórica ministrada em conjunto.

Observando o desempenho de cada aluno nas diferentes atividades que compõem o sistema de avaliação de LPII, é possível identificar situações bem controversas. Por exemplo, espera-se que a atividade “laboratórios” tenha uma forte correlação com a prova, porém, considerando uma turma em um determinado semestre letivo, para uma parte dos alunos o desempenho nos laboratórios é semelhante ao desempenho na prova, já para outra parte não.

Existe um esforço dos professores de LPII em fazer ajustes na metodologia a cada semestre letivo para melhor avaliar os alunos na disciplina. Tais ajustes não consideram um melhor entendimento da relação entre as atividades e seu impacto sobre os alunos, podendo comprometer a eficácia do mecanismo. Nesse sentido, nosso trabalho contribui para evitar tal situação.

2 TRABALHOS RELACIONADOS

Saber como avaliar o aprendizado de alunos sempre foi um desafio enfrentado por professores nas mais diversas áreas. Avaliar estudantes no ensino superior se mostra uma tarefa ainda mais complicada, considerando as diferentes experiências dos alunos de uma mesma turma. Sanders e Horn (1995) [4] dizem que o comportamento e a aparência do aluno também podem prejudicar a avaliação. Por outro lado, atenção cuidadosa às informações fornecidas pelas ações dos alunos podem levar a um entendimento profundo e revelador da compreensão de conteúdo e obtenção de habilidades pelos alunos.

A percepção do aluno quanto à metodologia de avaliação também já se mostrou um fator importante que deve ser considerado no processo de ensino-aprendizagem. Muitos alunos reagem negativamente quando discutem sobre métodos tradicionais de avaliação, alegando que tais métodos tiveram um efeito gravemente prejudicial no processo de aprendizagem. A situação se inverte quando se discute sobre novos métodos de avaliação, como debatido por Struyven, Dochy, e Janssens (2005)[5].

*Os autores retêm os direitos, ao abrigo de uma licença Creative Commons Atribuição CC BY, sobre todo o conteúdo deste artigo (incluindo todos os elementos que possam conter, tais como figuras, desenhos, tabelas), bem como sobre todos os materiais produzidos pelos autores que estejam relacionados ao trabalho relatado e que estejam referenciados no artigo (tais como códigos fonte e bases de dados). Essa licença permite que outros distribuam, adaptem e evoluam seu trabalho, mesmo comercialmente, desde que os autores sejam creditados pela criação original.

Não sendo exceção, o ensino de programação também tem seus obstáculos quanto ao método de avaliação. Defendendo uma metodologia construtivista de ensino de programação, em especial Programação Orientada a Objetos (POO), Said (1999), propôs atividades avaliativas com características específicas para garantir qualidade no ensino de POO[2].

Os trabalhos citados mostram, assim como o nosso, uma preocupação da comunidade acadêmica com o processo de avaliação de alunos, e também com medir a eficácia e eficiência dos métodos.

3 METODOLOGIA

A nossa pesquisa teve natureza exploratória e quali-quantitativa, na qual buscamos enxergar relações entre as avaliações e padrões no comportamento dos alunos na perspectiva de seu desempenho na disciplina LPII.

Sendo uma disciplina do primeiro ano do curso de Ciência da Computação, os professores de LPII têm o desafio de ensinar alunos recém ingressos e, em sua maioria, com pouca experiência em programação, conceitos de Orientação a Objetos, como herança, composição, polimorfismo e tipos abstratos de dados. O aluno também é apresentado à sua segunda linguagem de programação na graduação — Java —, bem como tecnologias de controle de versão e testes. Ao mesmo tempo, estão aprendendo conceitos básicos sobre padrões de projeto e modelagem de sistemas.

A disciplina possui um sistema contínuo de avaliação, com atividades ao longo de todo o semestre letivo, práticas e teóricas, individuais e em grupo, com ou sem consulta, cada uma com seus respectivos pesos.

3.1 Objetivos de Pesquisa

Os instrumentos de avaliação utilizados na disciplina foram concebidos para que existam intersecções entre os conteúdos trabalhados, e, portanto, espera-se que essas avaliações tenham certo nível de relação, sendo sua variação dependente do formato de aplicação. Sabendo disso, queremos entender o relacionamento entre elas.

Em simultâneo, é importante sabermos a natureza do comportamento dos alunos na perspectiva de seu desempenho, seja este um comportamento mais homogêneo ou heterogêneo, dedicando atenção para os grupos mais peculiares de alunos, tendo em vista que é de nosso interesse conhecer os perfis destes.

Também é de nosso interesse, ainda que sem tanto enfoque, rastrear as reações dos alunos quando realizadas mudanças no instrumento de avaliação para aumentar nossa compreensão sobre o modo como eles se comportam.

3.2 Os dados

Os dados que utilizamos na nossa análise são referentes ao desempenho de 441 alunos de LPII distribuídos em cinco períodos letivos: 2017.1, 2017.2, 2018.1, 2018.2 e 2019.1. Temos o desempenho dos alunos em cada avaliação individualmente e suas médias ponderadas. A Tabela 1 possui uma descrição mais detalhada das variáveis e quais avaliações elas representam.

Como os nossos dados são referentes a períodos letivos distintos, direcionamos nossa análise dependendo do quanto expressiva foi a mudança no sistema de avaliação entre eles. Dentre essas mudanças, podemos destacar duas:

Tabela 1: Variáveis (Avaliações)

Variável	Descrição
lab_n	Atividade prática individual com consulta e prazo de entrega de 1 semana
projeto	Atividade prática em grupo com consulta e prazo de entrega de 1 mês
prova_n, reposicao e prova_final	Atividade prática individual sem consulta e prazo de entrega de 2 horas
*primeira_prova	prova_1
*ultima_prova	prova_2 ou prova_3 caso exista

* Em alguns momentos utilizamos diferentes nomenclaturas para as variáveis, pois em alguns períodos não existem algumas provas, mas ainda desejamos saber a ordem em que ocorrem.

- Só houve duas provas no período 2019.1;
- No período 2019.1 a nota do projeto depende mais do desempenho individual.

Além dessas, durante a evolução do mecanismo de avaliação, os professores modificaram os pesos das avaliações, tanto para suas médias específicas para tipos de avaliação, quanto seus pesos na média final.

Alguns outros metadados foram levados em conta durante nosso processo de análise:

- Peso maior atribuído às últimas avaliações quando comparadas com as primeiras;
- Mudança dos professores;
- Mudança do local físico da realização das avaliações;
- As atividades "laboratórios" são as únicas não corrigidas pelos professores e sim por monitores.

A Tabela 2 possui a descrição das variáveis que representam as médias ponderadas.

3.3 Análise exploratória e descritiva

Na primeira parte da análise, identificamos os comportamentos médios e discrepantes, comparamos esses comportamentos, investigamos a interdependência entre as variáveis, buscando padrões e tendências. Iniciamos a análise a partir das variáveis compostas (médias), em seguida partimos para as variáveis simples. Aqui, também consideramos as distâncias entre as avaliações para rastrear anomalias, como, por exemplo, alunos que tiveram uma diferença absoluta muito alta entre as avaliações.

3.4 Regressão linear e correlação

Com o resultado obtido na nossa análise exploratória, aprofundamos nossa investigação para algumas avaliações, considerando o quanto importante elas são para a nota final do aluno, a prévia suspeita dos professores da disciplina sobre essa relação e o comportamento dos dados quanto a elas. Para isso, utilizamos os métodos de *Pearson*, *Kendall* e *Spearman* para calcular a correlação entre as avaliações desejadas e rodamos modelos de regressão linear simples e múltipla.

Tabela 2: Variáveis (Médias)

Variável	Descrição
media_labs	Média ponderada das atividades práticas individuais com consulta e prazo de entrega de 1 semana
media_provas	Média ponderada das atividades práticas individuais sem consulta e prazo de entrega de 2 horas
media_final	Média ponderada da media parcial com a prova final
media_parcial_minitests	Média de todas as atividades práticas individuais sem consulta e prazo de entrega de 15 minutos
p2	Média das atividades teóricas individuais sem consulta e prazo de entrega de 2 horas

3.5 Clusterização

Segmentar objetos para nós seres humanos pode ser uma tarefa relativamente fácil, conseguimos identificar padrões levando em consideração vários atributos como cor, forma, tamanho e peso de forma rápida e simples. Para as máquinas, o conceito é parecido só que bem mais desafiador.

O processo para agrupamento de dados se baseia no conceito de similaridade, ou seja, a ideia principal é encontrar itens semelhantes de acordo com seus atributos. O fator que determina essa similaridade entre os registros é o cálculo de distância.

Para identificar melhor os perfis dos alunos, tentamos agrupá-los quanto ao seu desempenho. Foram feitos agrupamentos considerando, primeiramente, as avaliações que ocorreram em todos os períodos observados. Após isso, agrupamos os alunos a partir de suas médias e, por último, consideramos as distâncias entre pares de avaliação na formação dos grupos. Utilizamos, então, alguns métodos bem conhecidos para esse tipo de problema.

3.5.1 K-means. Um dos algoritmos mais comuns de clusterização é o *k-means*, ilustrado na figura 1. O *k-means* é um algoritmo do tipo não supervisionado. O objetivo do algoritmo é encontrar similaridades entre os dados e agrupá-los conforme o número de *clusters* passado pelo argumento *k*. Esse algoritmo utiliza um método simples baseado no conceito de distância, neste caso, a distância Euclidiana.

Resumidamente, o algoritmo escolhe, de forma iterativa, centroides aleatórios e calcula a distância de todos os pontos para cada um dos centroides, atribuindo cada ponto ao grupo do centroide de menor distância. Em seguida calcula-se a média de cada grupo, que será o novo centroide. Esses passos são repetidos até que os *clusters* se tornem estáticos ou algum critério de parada tenha sido atingido.

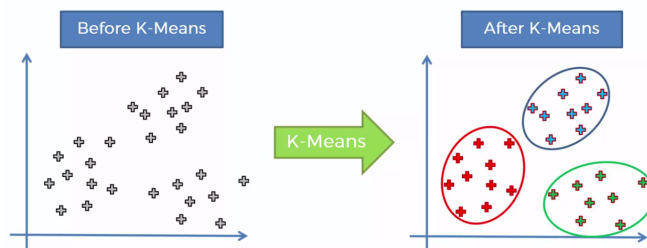


Figura 1: Agrupamento com k-means

3.5.2 Betweenss/totss ou Elbow method. Mas como escolher um bom valor para *k*?

Uma medida comumente usada no *k-means* é comparar a distância entre os centroides e o centro dos dados com a distância entre todos os pontos e o centro dos dados, sendo este um ponto imaginário na média de todas as variáveis. Em uma situação onde cada ponto é um grupo, a soma das distâncias dos grupos para o centro dos dados é igual a soma da distância dos pontos para o centro dos dados. Se houver estrutura de grupos e ela estiver capturada pelo agrupamento, o somatório da distância do centro de cada grupo para o centro geral dos dados será um valor alto.

Para medir para quais valores de *k* isso acontece, calculamos a distância do centro de cada *cluster* para o centro dos dados e multiplicamos pelo número de pontos nesse *cluster*. Somando esse valor para todos os *clusters*, temos *betweenss*.

Se esse valor for próximo do somatório total das distâncias dos pontos para o centro dos dados (*totss*), os pontos estão próximos do centro de seu *cluster*. Essa proporção pode ser usada para definir um bom valor de *k*. Quando ela para de crescer, para de valer à pena aumentar *k*, afinal, não queremos grupos com dois pontos, por exemplo.

Na figura 2 o *k* escolhido está no *elbow*, ou cotovelo. Não ganhamos muito quando aumentamos o valor de *k* e especializamos demais os grupos.

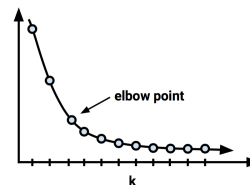


Figura 2: Elbow method

3.5.3 PCA. Pode ser simples visualizar a forma como vários pontos dispersos em um gráfico de uma ou duas dimensões se agrupam, mas, em certos casos, quando muitas variáveis estão presentes, não é fácil visualizar os dados em sua forma bruta, tornando difícil a tarefa de detectar tendências. O *PCA* (*Principal Component Analysis*) permite que vejamos o "formato" geral dos dados, tornando possível a identificação dos grupos e a descoberta de qual variável faz um grupo diferente de outro.

Os componentes principais são a estrutura subjacente nos dados, são as direções em que os dados mostram melhor sua variação, onde estão mais dispersos. Tentamos encontrar uma linha reta que melhor descreve a dispersão dos dados quando a projetamos sobre eles, esse é o primeiro componente principal (PC1).

A figura 3 ilustra dois componentes que traçamos para explicar a variação dos dados dispersos em duas dimensões. O componente 2 parece mostrar melhor essa dispersão que o componente 1. Essa linha seria, então, a representação em uma dimensão dos dados anteriormente mostrados em duas dimensões, nosso PC1.

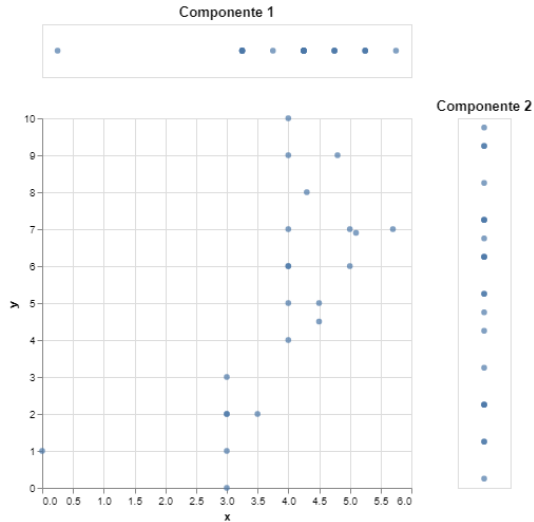


Figura 3: Ilustração do PCA

No contexto desse trabalho, temos diversas atividades, e, como queremos encontrar grupos em torno dessas atividades, utilizamos o PCA para reduzir a dimensionalidade dos nossos dados e nos ajudar na tarefa de identificar esses grupos.

4 RESULTADOS E DISCUSSÕES

Nesta seção discorremos sobre os resultados obtidos e filtrados durante o processo. As subseções abaixo seguem, para melhor entendimento, a mesma ordem lógica em que delineamos a seção que descreve o método.

4.1 Da análise exploratória

Estando cientes que os dados tratados em nossa pesquisa foram pouco explorados anteriormente, nosso primeiro passo, como dito na seção 3.3, foi uma análise exploratória e descritiva.

A média parcial dos alunos nos pareceu um bom ponto de partida, já que se trata de uma média ponderada composta pela maior parte das outras variáveis que nos interessa. A figura 4 nos mostra a dispersão das médias parciais dos alunos separadas por período, nos dando, também, uma noção do que aconteceu ao longo do tempo com a disciplina.

Note que o período 2017.2 mostrou um desempenho mediano razoavelmente acima dos outros, levando em conta que também parece ser o menos disperso. Os períodos 2017.2 e 2018.1, embora

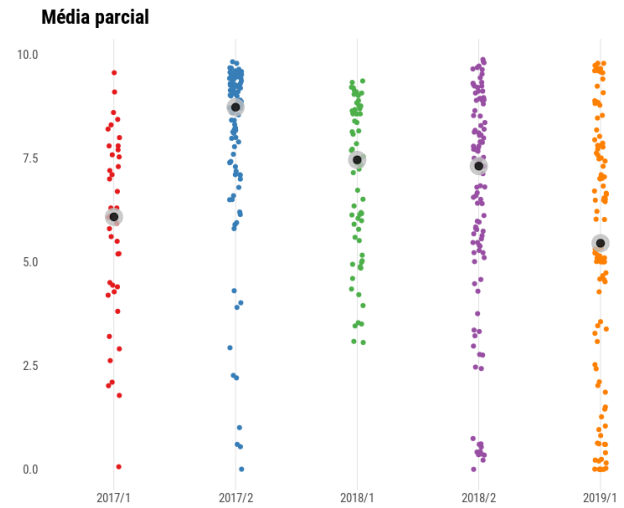


Figura 4: Média parcial dos alunos por período.

tenham a mediana bem semelhante, mostram dispersões consideravelmente diferente, dando indícios de que resumir demais nossos dados pode esconder informações importantes, como o grupo de alunos do 2018.2 com o desempenho bem abaixo do restante da turma. Por fim, o 2019.1 tem a menor mediana dos períodos analisados e também a maior dispersão, sendo este o mais recente, refletindo melhor a atual situação da disciplina.

Vejam agora o desempenho dos alunos considerando as diferentes atividades avaliativas realizadas na disciplina. Primeiro, na figura 5, a dispersão dos alunos na atividade "laboratórios". Esta atividade não recebeu nenhuma alteração expressiva em nenhum dos períodos analisados, sendo aplicada da mesma forma em todos estes.

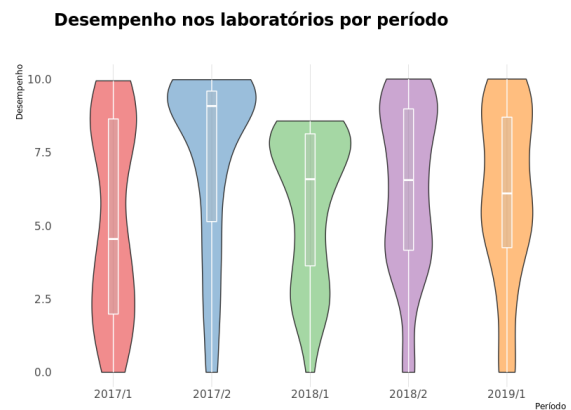


Figura 5: Média dos laboratórios por período.

Repetindo o que notamos na figura 4, o período 2017.2 teve um desempenho mediano maior que os outros, além de uma maior concentração em torno desta mediana. Os períodos 2018.2 e 2019.1 mostraram comportamentos bem semelhantes. A mediana das turmas se manteve parecida nos três últimos períodos.

Já na figura 6 podemos ver o desempenho dos alunos nas provas práticas, estas, por sua vez, tiveram uma mudança em como foram aplicadas. No período 2019.1 só houve duas provas práticas, enquanto nos demais foram três.

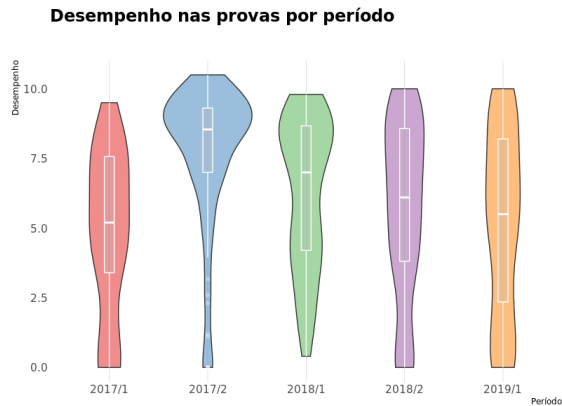


Figura 6: Média das provas por período.

Igualmente como vimos com os laboratórios, os períodos 2018.2 e 2019.1 tiveram comportamentos bem similares, sendo a mediana deste último um pouco abaixo, destacando-se o fato de não notarmos nenhuma alteração aparente no comportamento após a remoção de uma prova prática do sistema de avaliação. Já o período 2017.2 segue com um desempenho acima dos outros e alta concentração nas maiores notas.

Fechando as três principais notas do aluno, a nota dos projetos está dispersa na figura 7 e nela encontramos algo interessante, que parece refletir uma alteração no sistema de avaliação:

No período 2019.1 percebemos uma mudança representativa na dispersão das notas do projeto quando comparado com os outros períodos. Isso pode ser explicado pela mudança metodológica citada em 3.2, onde a nota do aluno na atividade se tornou mais dependente do desempenho individual. Essa mudança parece ser também a responsável pela maior dispersão das médias parciais, tendo em vista que as outras duas médias não apresentaram quase nenhuma alteração no seu espalhamento.

Finalmente, concluindo nossa análise exploratória, na figura 8 apresentando o correlograma baseado no método de *pearson*.

É importante lembrar que as variáveis *media_provas*, *media_labs*, *media_parcial* e *media_final* são constituídas por outras variáveis que também estão no correlograma, portanto, não devemos levar em conta a correlação entre elas. A tabela 2 descreve como são constituídas as variáveis de médias.

Sendo assim, podemos destacar a correlação entre *p2*, que representa a média das provas teóricas, e as variáveis que representam

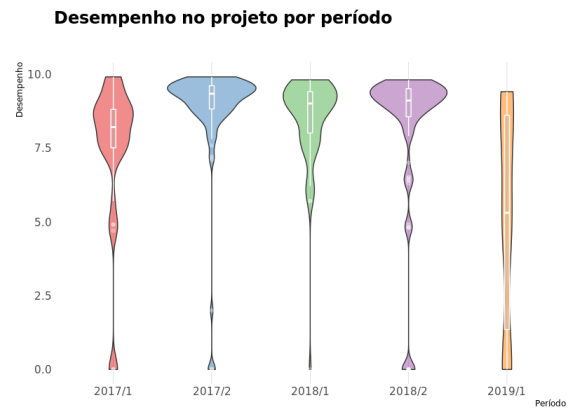


Figura 7: Nota dos projetos por período.

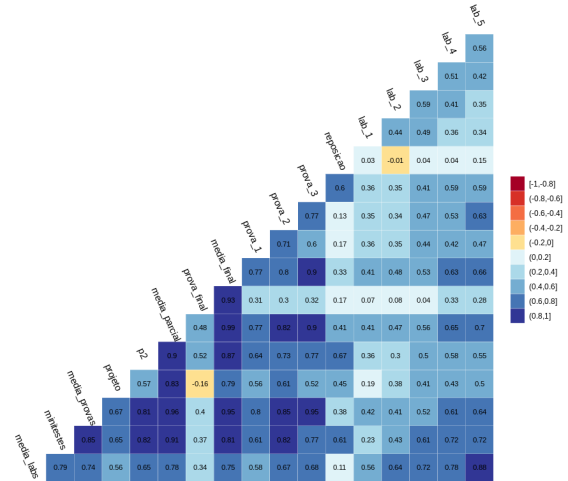


Figura 8: Matriz de correlações.

as provas práticas, em especial *prova_3* (0.77), assim como a variável que as compõe *media_provas* (0.81). A variável *minitestes* tem uma correlação consideravelmente alta com quase todas as outras variáveis, destacando-se *media_provas* (0.85). A *media_labs* possui correlação entre 0.6 e 0.8 com diversas variáveis, destacaremos com *minitestes* (0.79), pois a correlação entre *minitestes* e *media_labs* deveria ser maior que *minitestes* e *media_provas*, tendo em vista que os *minitestes* e laboratórios são avaliações elaboradas para serem correspondentes. Por fim, a variável *projeto* não mostra uma correlação expressiva com nenhuma outra variável.

4.2 Da análise de regressão

Partindo do resultado obtido com o correlograma, utilizamos dois parâmetros para escolha das variáveis para aprofundarmos a análise: aquelas variáveis independentes com maior correlação exibida na matriz de correlação e aquelas em que existia uma intuição sobre sua relação por parte dos professores que ministram a disciplina.

O par de variáveis em que nos debruçamos foi aquele em que as variáveis representam os laboratórios e as provas práticas: *media_labs* e *media_provas* respectivamente.

Como nosso correlograma foi construído a partir do método de *pearson* que considera os valores absoluto das variáveis, resolvemos descobrir qual seria essa correlação caso utilizássemos os métodos de ranking: *spearman* e *kendall*. O resultados estão na figura 9.

pearson	spearman	kendall
0.7382207	0.7382201	0.5547705

Figura 9: Correlação entre *media_labs* e *media_provas*.

Onde obtivemos um resultado bem semelhante para o método de *spearman*. Em seguida vejamos na figura 10 como se distribuem os alunos em torno dessas duas variáveis.

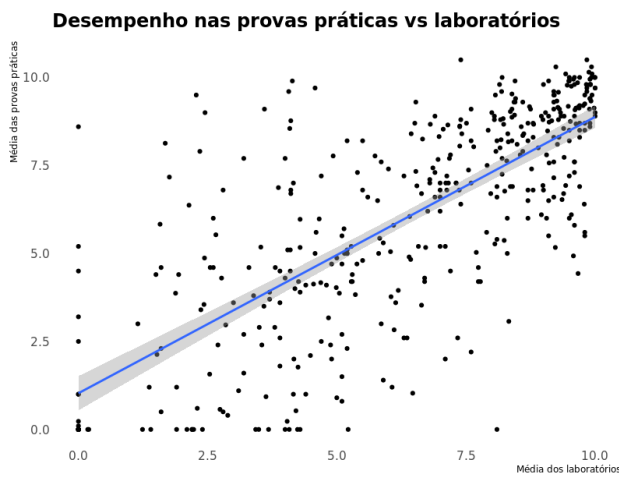


Figura 10: Dispersão dos alunos nas médias dos laboratórios e provas.

Existe, então, uma correlação linear positiva razoavelmente alta entre as duas variáveis, também demonstrada pelo gráfico de dispersão como a linha do modelo de regressão linear projetado sobre os pontos, sendo esse modelo no formato: $media_provas = 1.02 + 0.78 * (media_labs)$, que explica 54% da variável resposta ($R^2 = 0.5449699$).

Além disso, também rodamos um modelo de regressão linear múltipla com as notas individuais dos laboratórios, excluindo o *lab_1* pois em apenas um período temos o desempenho dos alunos. Este modelo teve o formato: $media_provas: 1.53 + 0.022 * (lab_2) + 0.14 * (lab_3) + 0.24835 * (lab_4) + 0.30517 * (lab_5)$, que explica 53% da variável resposta ($R^2 = 0.5316434$).

Podemos ver que, mesmo quando desconsideramos os pesos de cada laboratório, o *lab_5* possui maior influência na variável resposta *media_provas*.

Devemos lembrar que, embora teoricamente cobrem o mesmo conteúdo que as provas, os laboratórios possuem um tempo maior de entrega e é permitida consulta de material de apoio, as correções desses são feitas por monitores, além de diferenças subjetivas, como o nervosismo dos alunos durante as provas.

Voltando para a figura 10, mesmo que boa parte dos alunos estejam concentrados próximo à linha do modelo, a alta dispersão nos deixa atentos para a existência de grupos especiais de alunos que não possuem desempenho semelhante entre as avaliações.

4.3 Da clusterização

Observando a dispersão dos alunos na figura 10, vemos que existe uma quantidade considerável de alunos com desempenhos bem diferentes entre as avaliações. Essas discrepâncias foram motivação para nossa investigação de grupos de alunos que demonstram esse comportamento.

4.3.1 *Com as diferenças.* Continuamos analisando a relação entre os laboratórios e as provas práticas para nos ajudar a identificar os grupos ou perfis de alunos.

A figura 11 mostra a dispersão dos alunos quanto a diferença entre seu desempenho nas provas práticas e nos laboratórios. Ainda, rodamos o *k-means* para $k = 3$ – encontrado pelo *elbow method*.

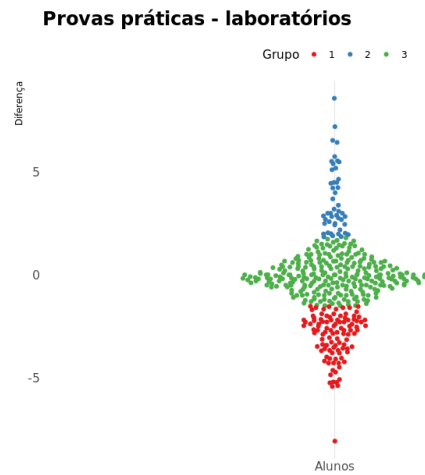


Figura 11: Dispersão das diferenças entre *media_provas* e *media_labs*.

A maior parte desses alunos se concentram em torno do eixo 0, o que significa que tiveram desempenho semelhante em ambas as avaliações, são o Grupo 3. Aqueles com maior nota nas provas são o grupo 2, já aqueles que foram melhor nos laboratórios são o grupo 1. A tabela 3 mostra algumas estatísticas sobre cada grupo para que possamos discutir melhor.

Como o esperado é um desempenho semelhante nas atividades, não discutiremos sobre o grupo 3, mesmo que ele seja o mais populoso. Os grupos 1 e 2 possuem, somados, 151 alunos, 36% das observações válidas e demonstram um comportamento não tão esperado. O grupo 2 possui uma diferença mediana de 3 pontos positivos

Tabela 3: Estatísticas dos grupos formados com a diferença entre $media_provas$ e $media_labs$

Grupo	Desvio Padrão	Mediana	Média	Variância	Alunos
1	1.12	-2.7	-3.01	1.26	104
2	1.62	3.0	3.65	2.64	47
3	0.76	-0.1	-0.02	0.58	267

entre as variáveis $media_provas$ e $media_labs$, enquanto o grupo 3 mostra uma diferença mediana de 2.7 pontos negativos. Sintetizando, existem grupos expressivos de alunos que possuem uma diferença razoavelmente alta — cerca de 3 pontos, para mais ou para menos — no seu desempenho nas atividades.

Realizamos ainda essa mesma análise, mas considerando outras duas avaliações que são $projeto$ e $media_labs$, pois os projetos são mais parecidos com os laboratórios na sua forma de aplicação do que quando comparamos as provas com os projetos. A figura 12 mostra as diferenças de desempenho dos alunos nas avaliações.

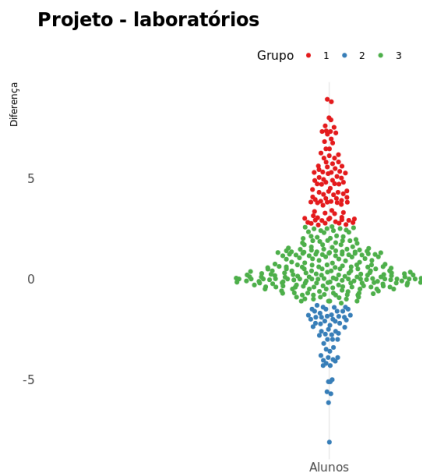


Figura 12: Dispersão das diferenças entre projeto e $media_labs$.

Novamente, a maioria dos alunos se concentra em torno do eixo 0 (grupo 3). Existem, também, mais alunos acima do eixo 0 que abaixo (grupo 2), o que significa que mais alunos tiveram nota melhor no projeto que na média dos laboratórios. Lembrando que o projeto trata-se de uma atividade em grupo corrigida pelos professores, enquanto os laboratórios são individuais corrigidas por monitores.

Aqui, temos interesse no impacto causado pela alteração já descrita anteriormente na seção 3.2, que dá um maior peso ao desempenho individual do aluno em sua nota do projeto, tendo em vista a drástica mudança na dispersão das notas dos projetos mostrada pela 7. Por isso, na figura 13 vemos essa mesma dispersão dos alunos separados por período.

No período 2019.1, o grupo 2 se mostrou bem mais populoso que nos outros períodos, ou seja, nesse período, os alunos apresentaram

um melhor desempenho nos laboratórios do que no projeto, se destacando pois inverte o comportamento mostrado nos outros períodos.

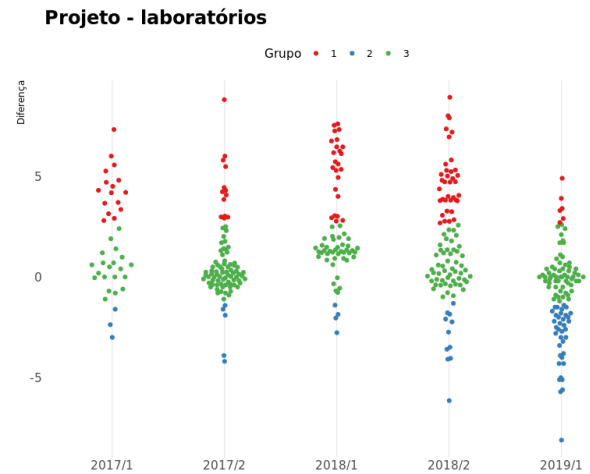


Figura 13: Dispersão das diferenças entre $media_provas$ e $media_labs$ por período.

Vejamos as estatísticas para esses grupos na tabela 4.

Tabela 4: Estatísticas dos grupos formados com a diferença entre projeto e $media_labs$

Grupo	Desvio Padrão	Mediana	Média	Variância	Alunos
1	1.55	4.5	4.76	2.43	95
2	1.40	-2.45	-2.9	1.98	60
3	0.91	0.26	0.46	0.83	245

O grupo 1, com 95 alunos, possui uma diferença mediana positiva de 4.5 entre as notas do projeto e dos laboratórios. Considerando que as notas podem variar de 0 a 10, podemos dizer que existe um grupo considerável de alunos que vai muito melhor no projeto do que nos laboratórios.

4.3.2 Com todas as avaliações. Mesmo após encontrarmos os grupos quando consideramos pares de avaliações, ainda não temos um perfil completo dos alunos, tendo em vista que a informação representada pelas médias é muito resumida. Nossa última investigação utiliza o máximo de variáveis para melhor caracterizar um aluno.

Como em alguns períodos não existiram determinadas atividades, tentamos mapeá-las para suas correspondentes em todos os períodos. As variáveis utilizadas estão também descritas na tabela 1.

Seguindo o descrito na seção 3.5.3, como temos muitas variáveis, seria impossível enxergarmos estruturas de grupos, portanto, tivemos a necessidade de reduzir a dimensionalidade dos nossos dados. Para isso utilizamos o já citado PCA.

Na figura 14 temos os alunos dispersos nos nossos dois principais componentes gerados pelo PCA. Na figura também é possível observar o as direções em que as variáveis estão deslocando os dados.

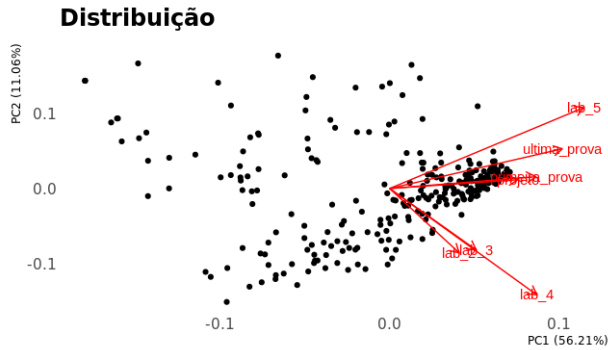


Figura 14: Dispersão dos alunos para os principais componentes.

O PC1 encontrado explica 56.21% da variação dos dados e o PC2 11.06%. Podemos considerar bons valores. Na figura 14 podemos ver uma alta concentração dos pontos mais à direita, assim como também é possível ver que as variáveis *lab_5*, *ultima_prova*, *primeira_prova*, *projeto* são aquelas que mais contribuem para essa concentração.

Sendo agora possível enxergar a variação dos pontos para todas as variáveis de interesse, vejamos na figura 15 como se distribuem os grupos encontrados pelo *k-means*, com $k = 8$ pelo *elbow method*.

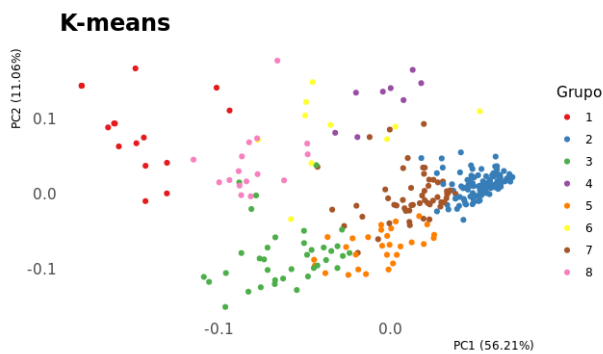


Figura 15: Alunos agrupados.

Temos, então, as estruturas de grupos. Podemos discutir mais quando analisamos algumas características dos grupos exibidas na tabela 5, segundo a variável *media_parcial*.

Devemos lembrar que o *k-means* não considerou a *media_parcial* na definição dos grupos, mas é de nosso interesse entender o desempenho final dos grupos, por ela foi escolhida para analisarmos os grupos.

Tabela 5: Estatísticas dos grupos formados pelo *k-means*

Grupo	Média	Mediana	Desvio Padrão	Alunos
1	1.3	0.9	1.2	16
2	9.1	9.2	0.5	109
3	4.6	5	1.7	36
4	7.9	8.1	0.8	8
5	7.4	7.6	1.0	29
6	7.1	7.1	1.0	10
7	7.2	7.3	1.0	53
8	5.3	5.3	1.1	16

Sendo um grupo bem populoso, o grupo 2, com 109 alunos, possui um alto desempenho na média parcial, que considera quase todas as atividades, excluindo a prova final. Tanto sua média, quanto sua mediana são bem altas, além de um baixo desvio padrão.

O grupo 4, bem próximo do grupo 1 no gráfico de dispersão, também mostram desempenho razoável, sendo seguido de perto pelos grupos 5, 6 e 7. Quando consideramos suas características, esses grupos (4, 5, 6 e 7) têm suas médias, medianas e desvios padrão bem semelhantes.

Os grupos 3 e 8 possuem estatísticas semelhantes, embora não estejam tão próximos no gráfico, enquanto o grupo 1 é o mais distante de todos, com baixo desempenho, mas pouco populoso.

Considerar a média parcial dos alunos nos grupos nos dá um panorama geral, mas, como já discutido anteriormente, parece esconder características importantes dos dados.

A figura 16 nos mostra como se comportaram os grupos considerando todas a variáveis utilizadas pelo *k-means*. Nesta figura, as atividades estão ordenadas cronologicamente, portanto, também é possível observar a evolução dos grupos durante a disciplina.

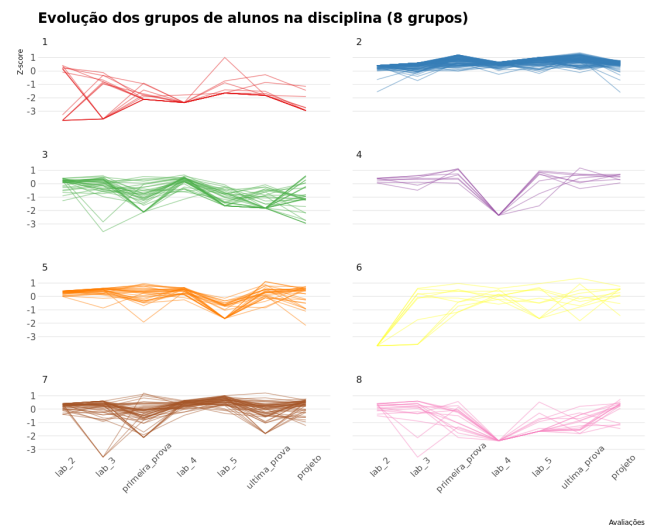


Figura 16: Evolução do desempenho dos alunos na disciplina.

Como esperado, o grupo 2 parece ter tido um bom desempenho em todas as avaliações, refletido em sua média parcial.

Destaca-se, ainda, o comportamento dos grupos 4, 5, 6 e 7, embora com médias parciais bem similares, mostram uma trajetória distinta durante a disciplina.

O grupo 3, embora tenha iniciado a disciplina com um bom desempenho nas atividades, parece ter uma leve queda próximo ao final dela, onde ocorrem as avaliações com maiores pesos, o que pode explicar sua baixa média parcial.

5 CONCLUSÃO

A disciplina Laboratório de Programação II (LPII), sendo uma disciplina prática, dispõe de diversas atividades avaliativas em variados formatos, mas que buscam medir o aprendizado dos alunos em conteúdos que se interseccionam, portanto semelhantes no que cobram. Encontramos neste trabalho uma correlação positiva entre quase todas as variáveis, que diversificam de intensidade. Das avaliações em que aprofundamos nossa análise, em especial os laboratórios e provas práticas, encontramos uma relação expressiva entre elas, mesmo que estas divirjam quanto a seu tempo de aplicação, uso ou não de material de apoio e correção. Isto sugere que estas diferenças no formato de aplicação da atividade têm impacto nos resultados obtidos por esses alunos na disciplina, podendo alterar o comportamento dos alunos quando modificado, como observado na seção 4.3.1 deste artigo.

Também foi possível notar a existência de grupos de alunos que se comportam de maneira semelhante. Nesses grupos fomos capazes de observar situações curiosas, como diferenças entre avaliações de até 5 pontos medianos para determinado grupo. Dessa forma, o mecanismo de avaliação parece não estar medindo de forma eficaz o nível de aprendizado real do aluno, uma vez que a diferença do formato de aplicação parece impactar tanto no desempenho.

Fomos ainda capazes de, com esses perfis, concluir que mesmo aqueles alunos com desempenho semelhante ao final da disciplina podem mostrar um comportamento diferente durante sua trajetória na disciplina.

REFERÊNCIAS

- [1] MESTRADO PROFISSIONAL DE SISTEMAS DE GESTÃO and CAROLINA LOBO DOS REIS MELO. [n. d.]. O PERFIL DO ALUNO E A CONSTRUÇÃO DO CONHECIMENTO: UM ESTUDO DE CASO NO CEFET MARACANÃ. ([n. d.]).
- [2] Said Hadjerrouit. 1999. A constructivist approach to object-oriented design and programming. In *ACM SIGCSE Bulletin*, Vol. 31. ACM, 171–174.
- [3] Rosario Hernández. 2012. Does continuous assessment in higher education support student learning? *Higher Education* 64, 4 (01 Oct 2012), 489–502. <https://doi.org/10.1007/s10734-012-9506-7>
- [4] William Sanders and Sandra Horn. 1995. Educational Assessment Reassessed. *education policy analysis archives* 3, 0 (1995), 6. <https://doi.org/10.14507/epaa.v3n6.1995>
- [5] Katrien Struyven, Filip Dochy, and Steven Janssens. 2005. Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education* 30, 4 (2005), 325–341.