

Questão 1: Dentre os usuários cadastrados em Nov/2017 que assinaram o Plano Premium, qual a probabilidade do usuário virar Premium após o cadastro em ranges de dias? A escolha dos ranges deve ser feita por você, tendo em vista os insights que podemos tirar para o negócio.

A probabilidade do usuário virar Premium após o cadastro, em ranges de dias, pode ser vista na Figura 1.

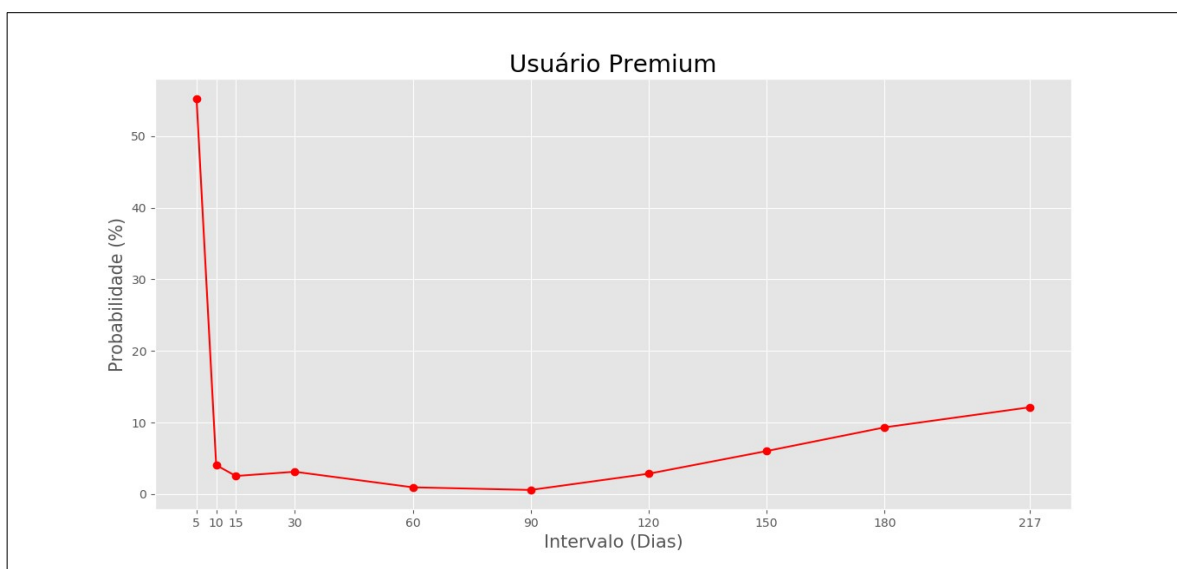


Figura 1 – Período de Assinaturas Premium após cadastro

É possível inferir da figura que existe uma probabilidade maior de um usuário se tornar premium nos primeiros 5 dias, 55%, período em que o usuário está buscando o conteúdo de interesse. Do quinto para o décimo dia há uma diminuição significativa nesta probabilidade e que se mantém entre 4% e 2% até o intervalo de 120 dias. Provavelmente o usuário consumiu o conteúdo que buscava não-premium ou buscou outra forma de obter o conteúdo.

A partir do intervalo de 120 dias há um leve crescimento até alcançar 12% para o intervalo maior do que 180 dias. Este leve crescimento pode ter relação com a transição entre semestres, onde um usuário cadastrado não-premium aumenta o seu interesse em consumir conteúdos com as vantagens do plano premium.

Questão 2: Faça uma análise geral dos dados e apresente as informações que julgar mais relevantes dessa base.

- Análise do dataset students.json

Esta amostra de dados contém 60 mil registros no período de 2012-05-29 a 2017-11-30 contendo 8 atributos do usuário:

1. Id
2. Data de registro
3. Universidade
4. Curso
5. Estado
6. Cidade
7. Forma de login
8. Meio de login

Como informações relevantes vale destacar que a principal forma de acesso é feita pelo Facebook, seguida de Google e Email, como pode ser visto na Figura 2.

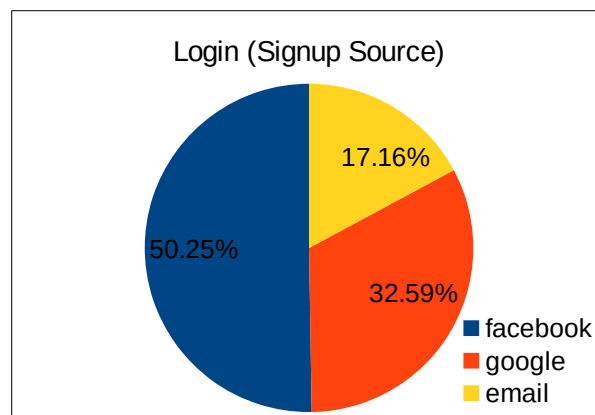


Figura 2 - Formas de login

As 5 universidades com maior número de usuários são Estácio, UNIP, UNOPAR, UNINTER e Estácio EAD (Figura 3).

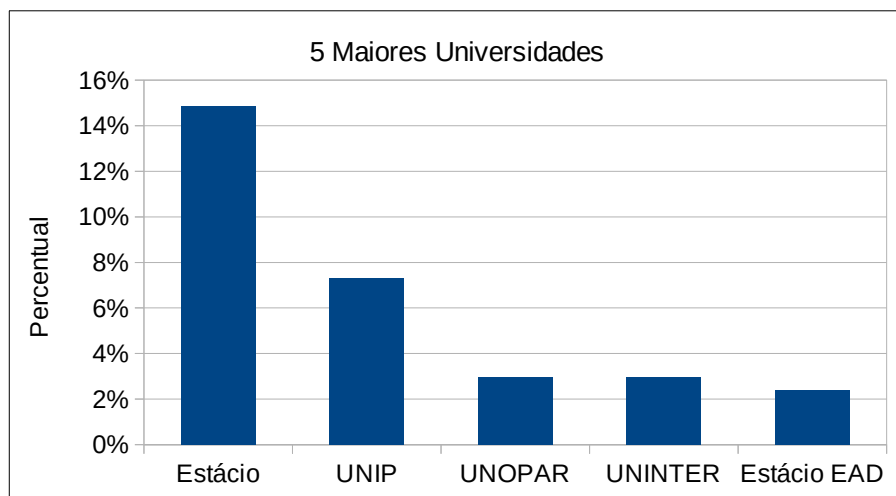


Figura 3- Universidades com maiores acessos

Sobre a origem dos usuários, vale ressaltar o Rio de Janeiro como o Estado com maior número de usuários, 8.67%, seguido de São Paulo, 6.46% (resultado surpreendente por se tratar do Estado mais rico do país, com as principais universidades, recursos e maior população), Minas Gerais, 3.94%, Paraná, 2.33% e Bahia, 2.19% completam os cinco maiores Estados. Do ponto de vista das cidades com maior número de usuários, aparecem Rio de Janeiro, 4.10%, São Paulo, 1.94%, Salvador, 1.04%, Belo Horizonte 0.92% e Fortaleza 0.76% como as cinco maiores. Os gráficos dos Estados e cidades podem ser visto, respectivamente, nas Figuras 4 e 5.

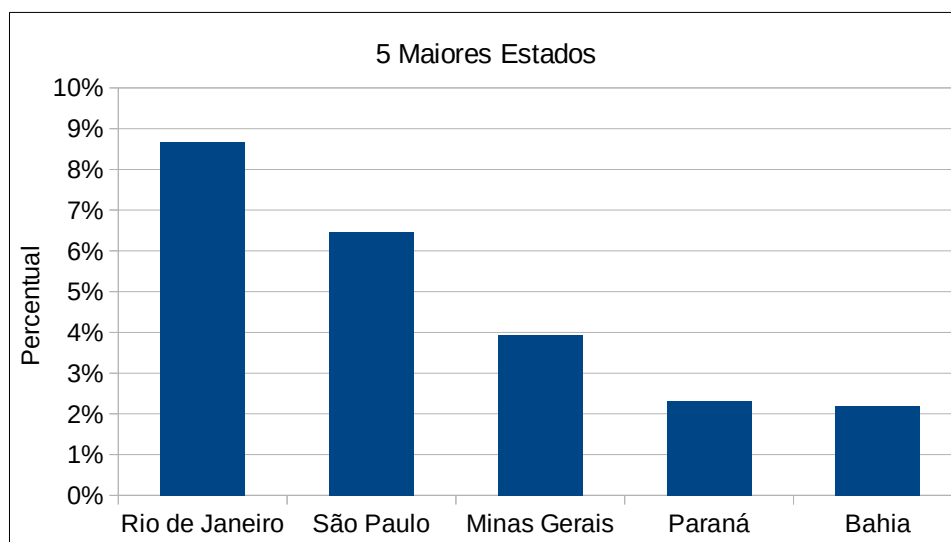


Figura 4 - Estados com maior acesso

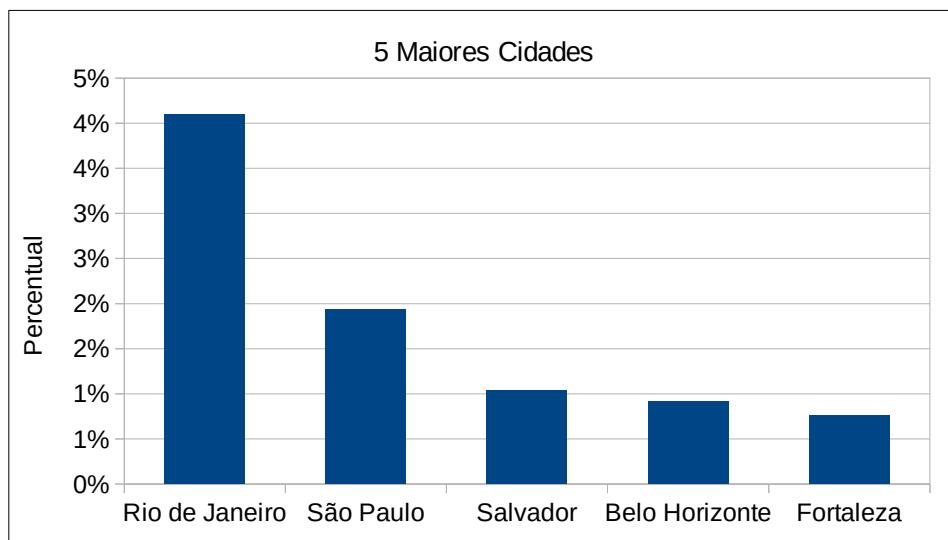


Figura 5 - Cidades com maior acesso

Sobre os 5 cursos com maior número de usuários, destacam-se: Direito (16.44%), Administração (6.83%), Engenharia Civil (5.98%), Contabilidade (4.57%) e Pedagogia (4.38%). A figura 6 apresenta este resultado.

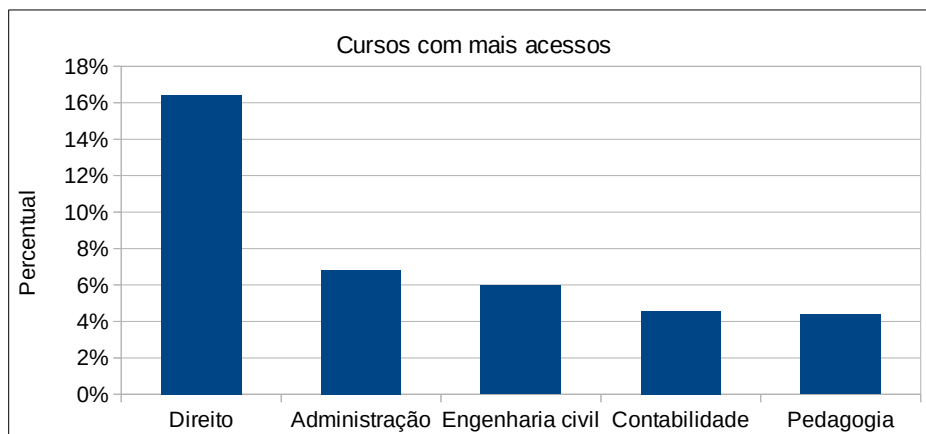


Figura 6 - Cursos mais acessados

- Análise do dataset sessions.json

Esta amostra de dados contém aproximadamente 1.4 milhão de registros no período de 2017-02-07 a 2017-11-30 contendo 3 atributos do usuário:

1. Id
2. Data do início dos acessos
3. Meio de login.

Quanto aos acessos de usuários destaca-se o percentual ao longo dos meses, onde observa-se o mês de Novembro de 2017 com 20% dos acessos no dataset. Também é possível observar o comportamento dos usuários para diferentes épocas do ano, como diminuições significativas dos acessos nos períodos de férias nos meses de Janeiro, Fevereiro e Julho, bem como aumentos gradativos entre estes períodos. Estes resultados podem ser vistos na Figura 7.

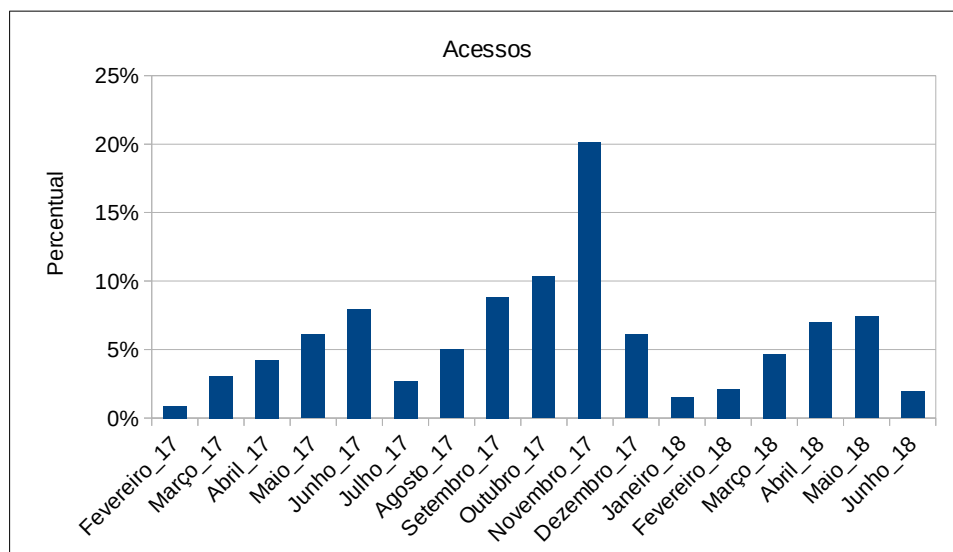


Figura 7 - Acessos ao longo do tempo

Em relação aos meios de login observa-se um destaque para acessos via Website (61.1%), seguido de Android (25.76%), IOS (7.64%) e Webapp (5.44%). A Figura 8 apresenta estes resultados.

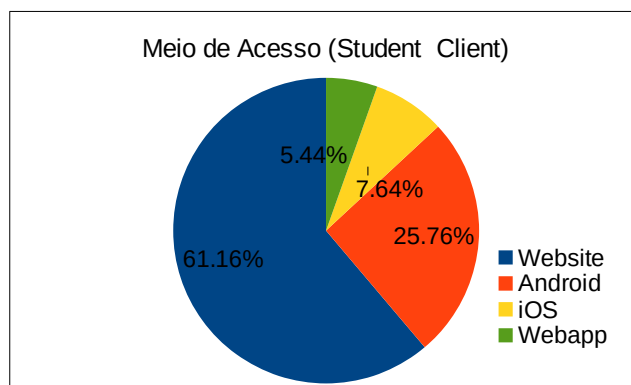


Figura 8 - Meios de Acesso

- Análise do dataset subjects.json

Esta amostra de dados contém aproximadamente 316 mil registros no período de 2017-08-31 a 2018-06-08 contendo 3 atributos do usuário:

1. Id

2. Data de Seguidas
3. Disciplinas

Vale destacar Bioquímica (7.05%) como a disciplina mais acessada, seguida de Cálculo I (6.88%), Anatomia Humana (6.65%), Direito Constitucional (6.55%), Matemática Financeira (5.70%) e Física I (5.15%) como as 6 mais acessadas. A Figura 9 mostra além deste resultado, outras disciplinas mais seguidas.

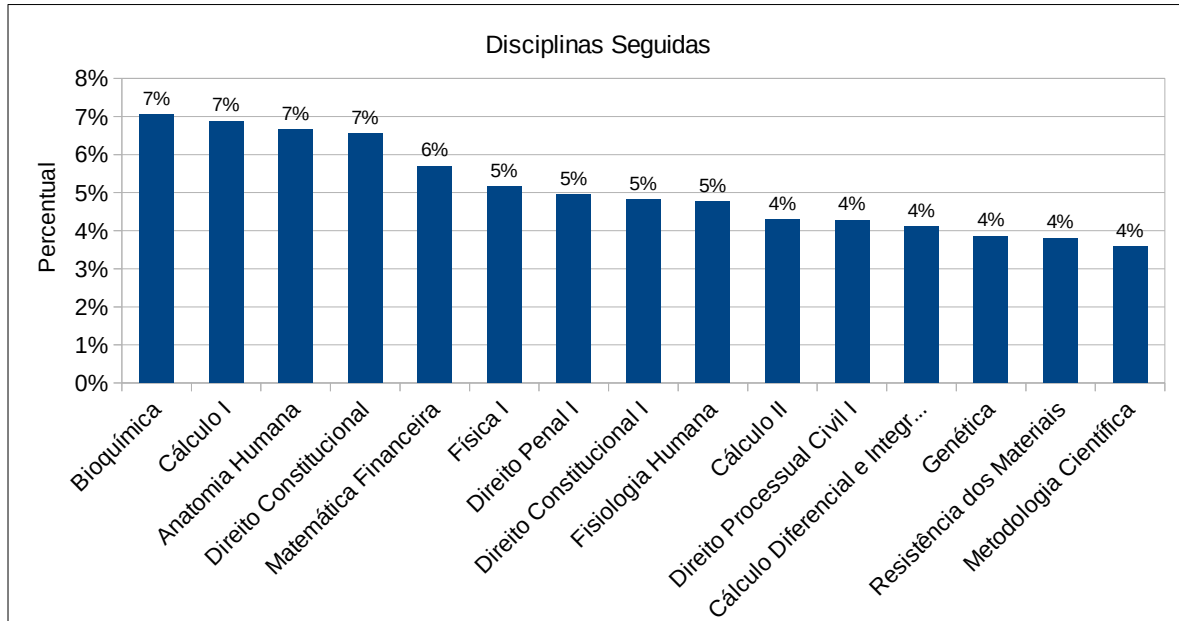


Figura 9 - Disciplinas mais seguidas

- Análise do dataset questions.json e answers.json

A amostra de perguntas contém aproximadamente 3.9 mil registros no período de 2013-09-04 a 2018-06-07 e a de respostas, aproximadamente 7.5 mil registros no período de 2013-08-30 a 2018-06-07, ambas contendo 4 atributos do usuário:

1. Id
2. Data das Perguntas/Respostas
3. Perguntas/Respostas
4. Meio de Login

Observa-se que no período da amostra de dados analisada, o percentual de perguntas e respostas dos usuários apresentam os maiores valores ao longo dos meses de Outubro e Novembro de 2017, destacando-se o mês de Novembro de 2017 com mais de 20%, tanto para perguntas como respostas. É

possível observar o comportamento dos usuários em diferentes épocas do ano, com diminuições significativas em períodos de férias nos meses de Janeiro, Fevereiro e Julho, bem como aumentos gradativos entre estes períodos e que ambos os atributos possuem um comportamento similar. Estes resultados podem ser vistos na Figura 10.

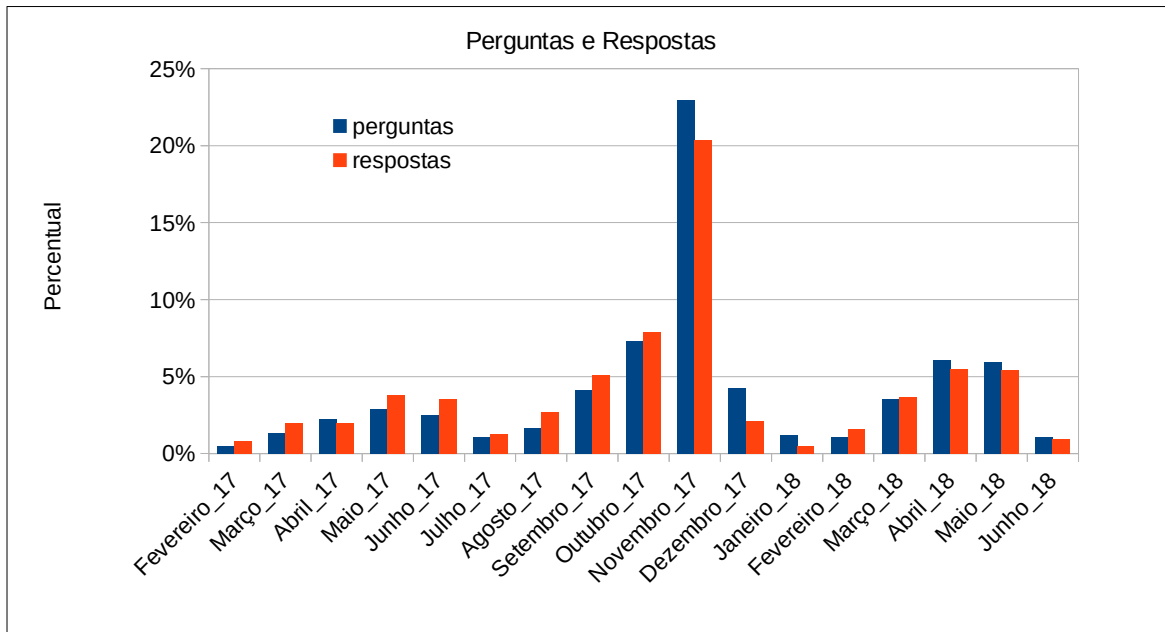


Figura 10 - Perguntas e Respostas ao longo tempo

- Análise do dataset fileViews.json

Esta amostra de dados contém aproximadamente 3 milhões de registros no período de 2016-11-24 a 2018-06-08 contendo 4 atributos do usuário:

1. Id
2. Data de Acesso
3. Arquivo
4. Meio de Acesso

Nesta amostra observa-se que os arquivos de avaliações, tais como AV1 e AV2, apresentam os maiores número de visualizações dentre os usuários, como pode ser visto na Tabela 1, e que o número de acesso aos arquivos foi maior nos meses de Outubro e Novembro de 2017, como descrito na Tabela 2.

Tabela 1 - Arquivos mais visualizados

Arquivos Visualizados	Nº de visualizações	Percentual
AV1	2303	0.08%
AV2	2169	0.07%
Avaliando o aprendizado	1804	0.06%
BDQ Prova	1541	0.05%
PLANEJAMENTO DE CARREIRA E SUCESSO PROFISSIONAL	1441	0.05%

Tabela 2 - Período de Visualizações

Período	Nº de visualizações	Percentual
Set_17	291051	9.61%
Out_17	373772	12.34%
Nov_17	470164	15.52%
Dez_17	103643	3.42%
Jan_18	18098	0.60%

- Análise do dataset studyPlanViews.json

Esta amostra de dados contém 7.7 mil registros no período de 2017-08-01 a 2018-06-07 contendo 4 atributos do usuário:

1. Id
2. Data de Acesso
3. Tópico
4. Disciplina.

Nesta amostra observa-se que os tópicos mais visualizados apresentam campos com valores nulos (N/A), representando 23.41% dos dados. Destacam-se tópicos tais como Resumos de Química, Bioquímica Estrutural, Conceitos Gerais de Contabilidade, Introdução ao Estudo de Direito Civil e Dimensão do Direito. Estes resultados podem ser vistos na Tabela 3.

Tabela 3 - Tópicos mais Visualizados

Tópicos	Nº de visualizações	Percentual
N/A	1814	23.41%
Resumos-de-quimica	241	3.11%
Bioquimica-estrutural-parte-i	233	3.01%
Conceitos-gerais-de-contabilidade	213	2.75%
Introducao-ao-estudo-do-direito-civil	166	2.14%
Dimensão-do-direito	165	2.13%

Sobre as disciplinas mais visualizadas, destacam-se as de Direito Penal I, Bioquímica, Cálculo I, Direito Civil I e Contabilidade I. O campo nulo (N/A) aparece na sequência com 3.72% de percentual sobre os dados. Os percentuais podem ser vistos na Tabela 4.

Tabela 4 - Disciplinas mais Visualizadas

Disciplinas	Nº de visualizações	Percentual
Direito Penal I	622	8.03%
Bioquímica	554	7.15%
Cálculo I	551	7.11%
Direito Civil I	542	6.99%
Contabilidade I	394	5.08%
N/A	288	3.72%

- Análise do dataset textBookSolutionViews.json

Esta amostra de dados contém 23 mil registros no período de 2017-08-01 a 2018-06-08 contendo 3 atributos do usuário:

1. Id
2. Data de Acesso
3. Nome do Livro.

Desta amostra observa-se na Tabela 5 que os livros mais visualizados são: Resistência dos Materiais 7º Edição 2010, Mecânica Vetorial Para Engenheiros – Estática 9º Edição 2009, Cálculo Vol. 2 7º Edição 2013, Mecânica dos Materiais 7º Edição 2015 e Princípios de Bioquímica de Lehninger 6º Edição 2014.

Tabela 5 - Livros mais visualizados

Título do Livro	Nº de visualizações	Percentual
Resistência Dos Materiais – 7 Ed. 2010	1012	4.38%
Mecânica Vetorial Para Engenheiros - Estática – 9 Ed. 2011	941	4.08%
Cálculo - Vol. 2 – 7 Ed. 2013	712	3.08%
Mecânica Dos Materiais - 7ª Ed. 2015	690	2.99%
Princípios de Bioquímica de Lehninger – 6 Ed. 2014	594	2.57%

- Análise do dataset evaluations.json

Esta amostra de dados contém 106.8 mil registros aproximadamente, o período não é especificado, e ela contém 4 atributos do usuário:

1. Id
2. Avaliação
3. Objeto da avaliação
4. Meio de acesso.

Destaca-se o alto nível de aceitação dos usuários com apenas 4% de “dislikes” (Figura 11) nesta amostra. É importante também notar que dentre as avaliações negativas a Tabela 6 mostra que o objeto da avaliação “arquivos” representa a maior parte com 79.82%.

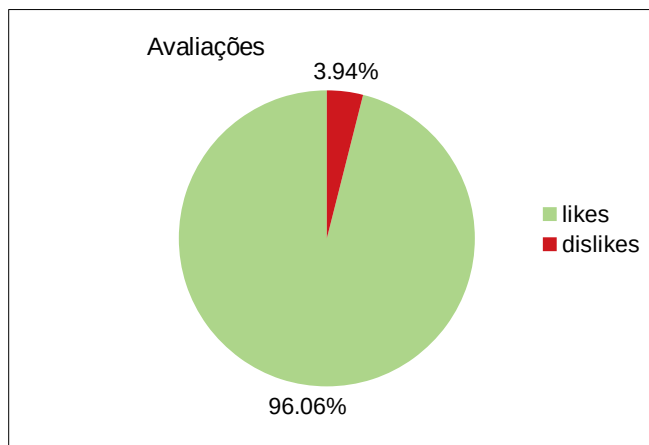


Figura 11 - Avaliações recebidas

Tabela 6 - Avaliações referentes aos "dislikes"

Item avaliado	Nº de avaliações	Percentual
File	3362	79.82
Question	470	11.16
Answer	293	6.96
Premium Content	34	0.81
Video	25	0.59
Comment	25	0.59
BookSolution	3	0.07

- Análise do dataset premium_payments.json e premium_cancellations.json

A amostra de pagamentos contém 7276 registros no período de 2015-08-08 a 2018-06-08 contendo 3 atributos de usuário (Id, Data de Pagamento e Tipo de Plano) e a de cancelamentos, 844 registros no período de 2016-05-05 a 2018-06-07, contendo 2 atributos do usuário: Id e Data de Cancelamento.

A Figura 12 mostra que o plano de assinatura mais realizado pelos usuários é o mensal, com mais de 95% das assinaturas.

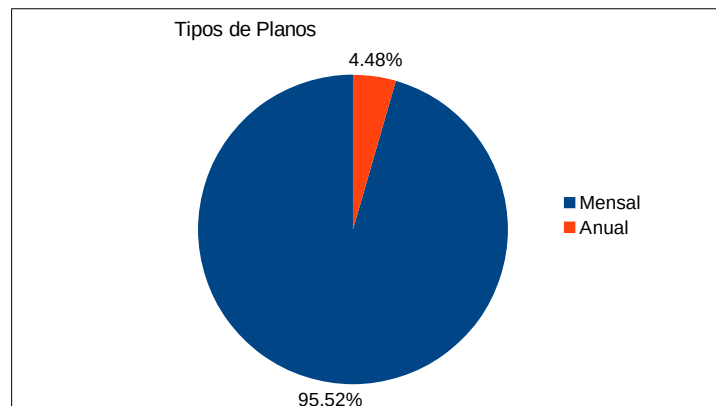


Figura 12 - Tipos de Pagamentos

A Figura 13 mostra a comparação entre o percentual de pagamentos, cancelamentos e a taxa de “churn” no período de maio de 2016 a junho de 2018. É possível observar um leve aumento tanto do número de pagamentos quanto de cancelamentos no decorrer dos meses. Entretanto, observa-se um aumento elevado a partir do mês de novembro de 2017. Observa-se uma inversão na relação de cancelamentos e pagamentos nos períodos de férias, onde ocorre maior inatividade dos usuários. A taxa de “churn” foi calculada pela definição simples, divisão do número total de cancelamentos dentro de

um intervalo pelo número de assinantes que haviam no primeiro dia do intervalo e apresenta um comportamento correlacionado aos cancelamentos.

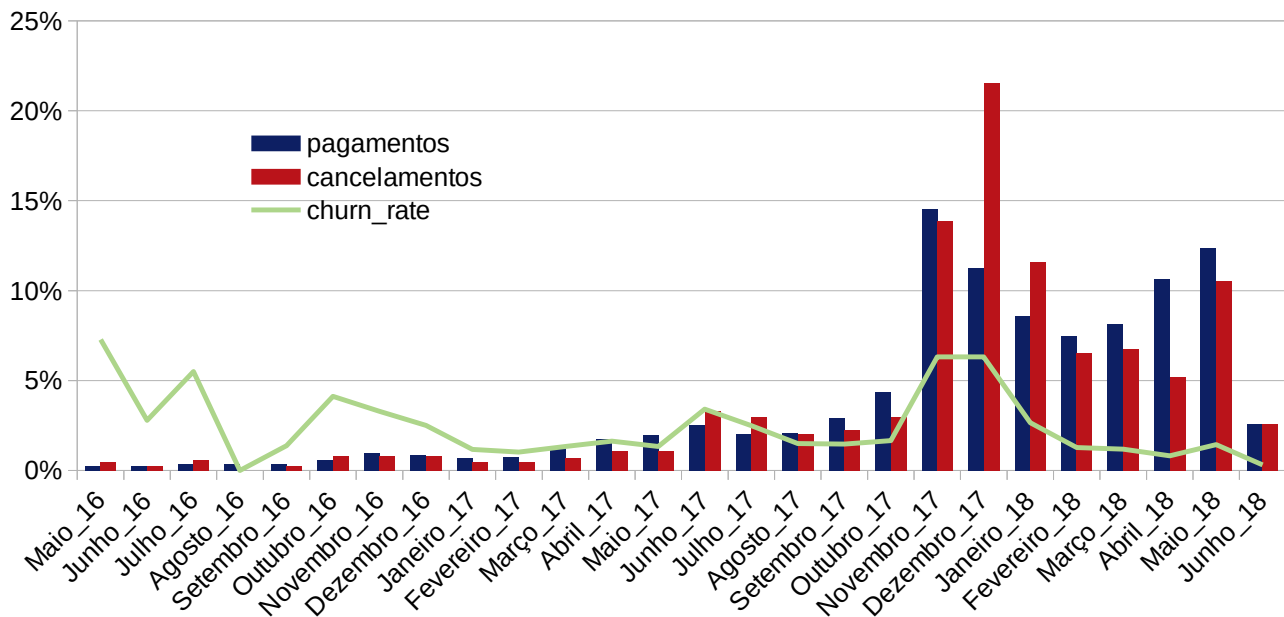


Figura 13 - Pagamentos, Cancelamentos e Churn Rate ao longo do período

Questão 3: Faça uma análise comparativa do comportamento dos usuários não Premium e dos usuários Premium. Que tipos de ações podemos direcionar para usuários não Premium fazerem com o objetivo de termos um maior número de assinantes?

Fazendo uma análise comparativa da amostra de dados `students.json`, `premium_payments.json` e `evaluations.json`, verifica-se que o comportamento de usuários premium e não premium são semelhantes em relação a origem, Estados e cidades, a universidade e a avaliações.

Algumas ações podem direcionar usuários para o plano premium:

- Aplicação de novas restrições, levando em conta que houve um aumento do número de usuários premium devido as restrições de conteúdo feitas em novembro de 2017.
- Investir em conteúdos.
- Criar relacionamento e atendimento personalizados, ações de engajamento e interações.
- Buscando novas estratégias de marketing, como o marketing digital, diversidade de planos e novas tecnologias.

Questão 4: Em Novembro de 2017 fizemos uma grande mudança no PD: o Content Restriction. Os usuários não Premium passaram a poder consumir no máximo 3 arquivos diferentes por mês. Diante dessa mudança, qual passou a ser o Lifetime Value (LTV) dos usuários Premium a partir de Novembro de 2017?

O Lifetime Value (LTV) é definido como o Tempo de Retenção (TR) do usuário vezes o Ticket médio que neste caso seria de 29.90. O TR é calculado como o número de cancelamentos vezes a quantidade de meses dividido pelo total de cancelamentos. Desta forma pode-se separar a amostra em dois períodos: Antes de Novembro de 2017 (AN) e Depois Novembro de 2017 (DN). Para o período antes de Novembro de 2017, o TR e o LTV obtidos são iguais a 13.11 e 388.70, respectivamente. Para o período depois de Novembro de 2017, o TR e o LTV obtidos são iguais a 21.53 e 657.80, respectivamente. A mudança representou um percentual de quase 70% no LTV.

Período	LTV
Antes de Novembro/17	388.70
Depois de Novembro/17	657.80

 69.23%

Questão 5: Entre os usuários que "Churnaram" solicitando ativamente o cancelamento do Plano Premium, o que está fortemente correlacionado com o cancelamento? (Ex: uso das features, tempo de inatividade, etc)

Na Figura 13 é possível observar um aumento significativo no percentual de cancelamentos dos usuários a partir de Novembro de 2017 que está fortemente correlacionado com o aumento no percentual de pagamentos. Ou seja, quanto mais assinaturas, mais cancelamentos. A taxa de "churn" também possui um comportamento recíproco ao percentual de cancelamentos. É possível observar também uma forte correlação com períodos de inatividade, neste caso, períodos de férias em Dezembro, em Janeiro, em Fevereiro e em Julho.

Questão 6: Quais são as 5 maiores universidades no Passei Direto? E quais são os principais tipos de Arquivos consumidos em cada uma delas?

As cinco maiores universidades no Passei Direto estão apresentadas na Tabela 7, que mostra a quantidade de usuários por universidade e o respectivo percentual. Já os principais tipos de arquivos

consumidos em cada universidade estão apresentados na Tabela 8 e o percentual de utilização desses arquivos pelas 5 principais universidades são apresentados nas Figuras 14 a 18.

Tabela 7 - Cinco universidades com maiores acesso

Universidade	Quantidade	Percentual
Estácio	8919	14.87%
UNIP	4392	7.32%
UNOPAR	1784	2.97%
UNINTER	1780	2.97%
Estácio EAD	1446	2.41%

Tabela 8 - Tipos de arquivos consumidos pelas principais Universidades

Universidades	Arquivos mais consumidos
1 Estácio	Provas, resumo, planejamentos e exercícios
2 UNIP	Apanhados, resumos, disciplinas online, provas, atividades práticas supervisionadas (APS) e exercícios
3 UNOPAR	Provas presenciais, relatórios de estágio curricular obrigatório, produção textual interdisciplinar e resumos
4 UNINTER	Atividade pedagógica on-line (APOL), provas e metodologias
5 Estácio EAD	Avaliações (AV), planejamentos, exercícios, provas e resumos. O gráfico mostra o percentual de utilização

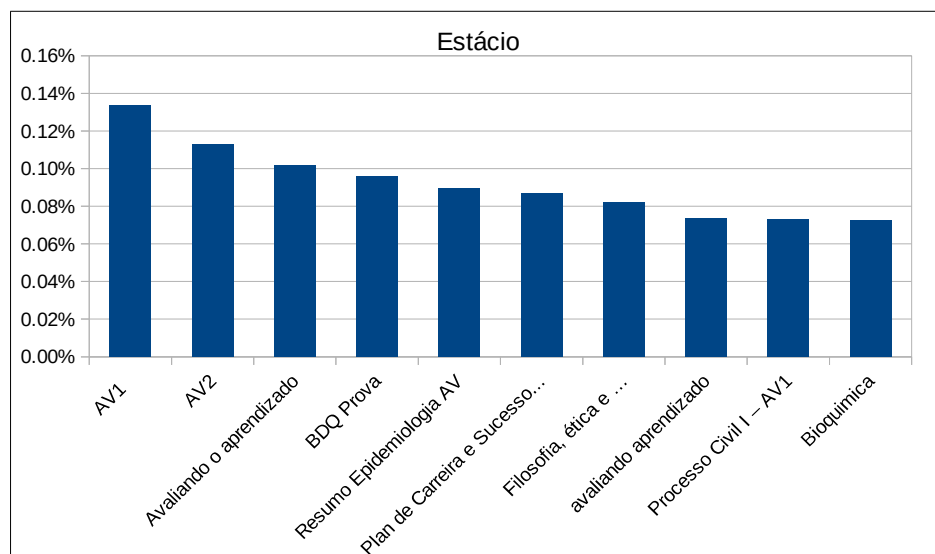


Figura 14 - Arquivos mais consumidos pela Estácio

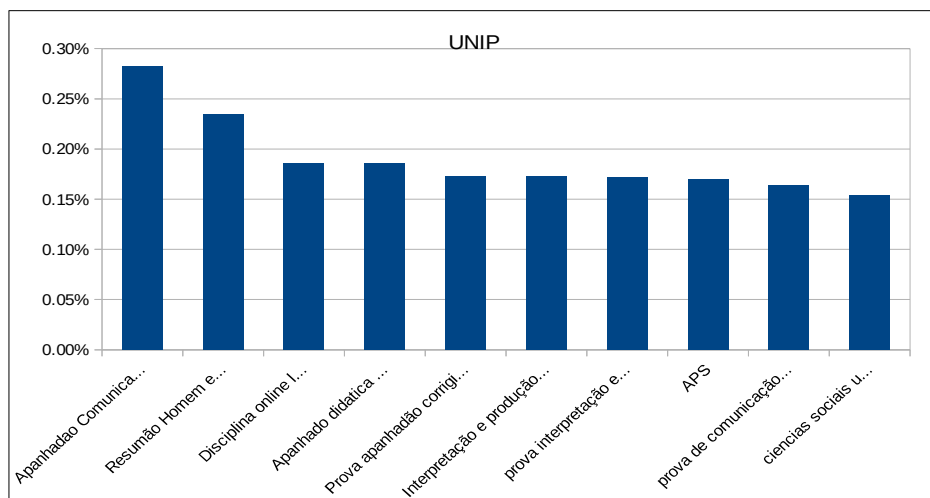


Figura 15 - Arquivos mais consumidos pela UNIP

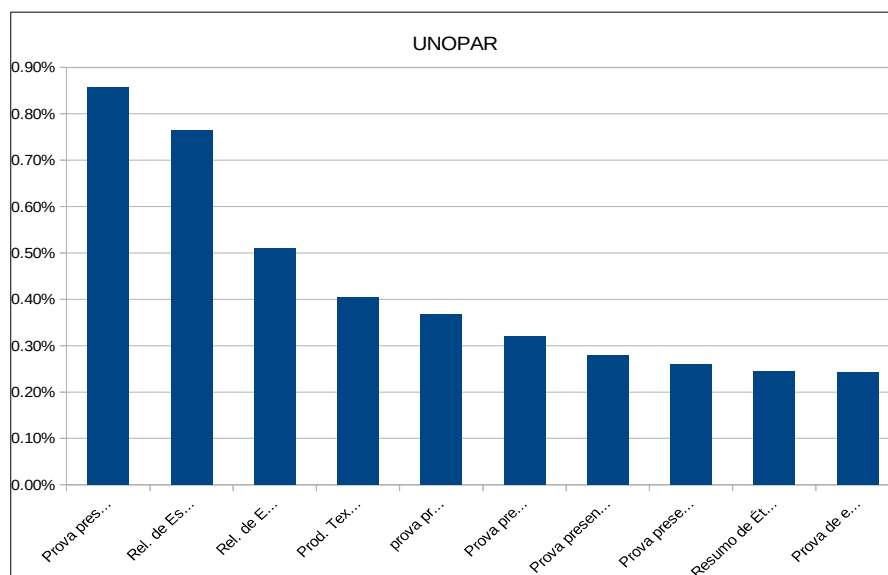


Figura 16 - Arquivos mais consumidos pela UNOPAR

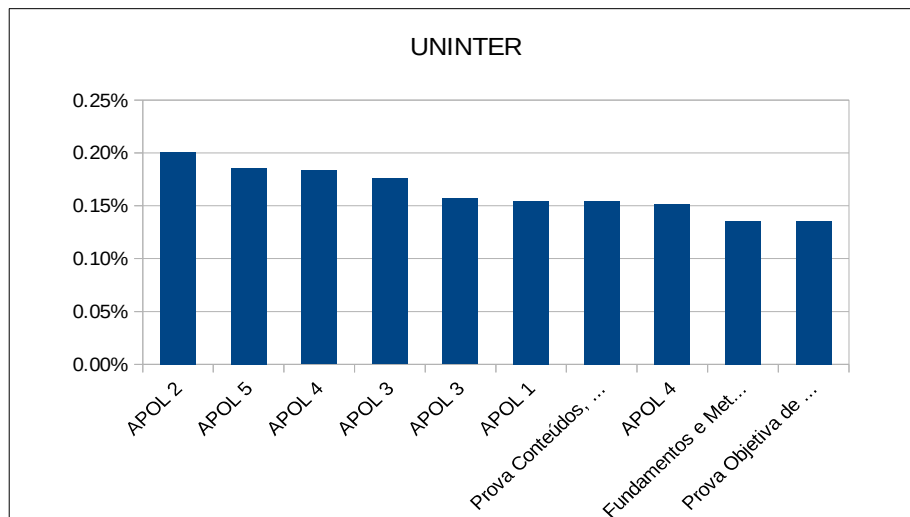


Figura 17 - Arquivos mais consumidos pela UNINTER

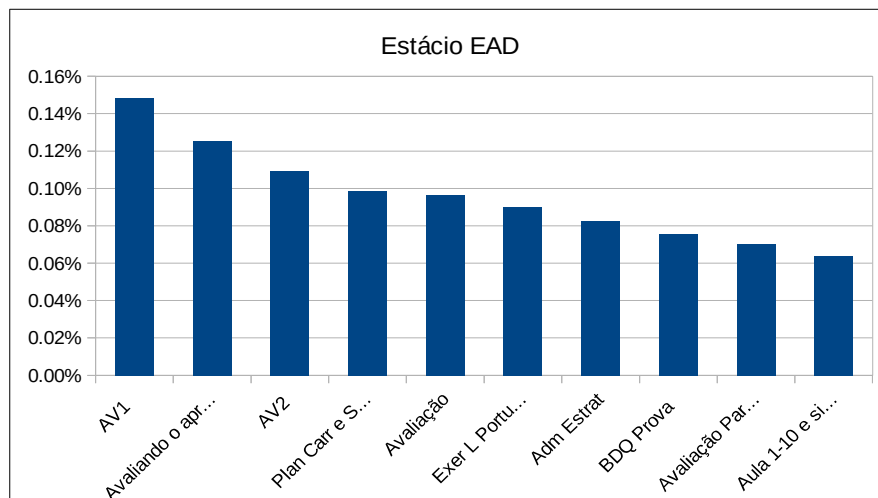


Figura 18 - Arquivos mais consumidos pela Estácio EAD

Questão 7: Considerando o seu conhecimento da plataforma do Passei Direto e a análise dos dados disponibilizados nesse teste, que oportunidades você vê para melhorarmos a experiência dos nossos usuários? Ofereça insights explorando o como e o porquê.

A partir dos resultados da questão 1 e sobre o conhecimento da plataforma, considerando do intervalo de 10 dias para o intervalo de 120 dias, onde a probabilidade do usuário virar Premium se mantém entre 4% e 2%, uma oportunidade para melhorar a experiência dos usuários seria a recomendação de conteúdos individuais por um preço específico, direcionado para este perfil de

usuário, baseando a recomendação no seu histórico de buscas e no histórico de buscas de perfis semelhantes.

A partir dos resultados e análise da questão 2, vale destacar o segundo lugar respectivamente do Estado de São Paulo e sua capital São Paulo no ranking de origem dos usuários. São Paulo é o Estado mais rico, tem as principais universidades e tem uma população maior (aproximadamente o dobro da população do RJ), apresentando um potencial muito grande a ser explorado.

Investir na retenção dos clientes, fazendo com que ele veja a plataforma como a escolha certa, investindo em conteúdos, criando relacionamento e atendimento personalizados, ações de engajamento e interações. Desta forma, diminuir as taxas de “churn” ou desistência de planos, buscando novas estratégias de marketing, como o marketing digital, diversidade de planos e novas tecnologias.