

Ciência de Dados para Todos (Data Science For All) - 2018.1 - Análise da Produção Científica e Acadêmica da Universidade de Brasília - Relatório Final da Disciplina - Departamento de Ciência da Computação da UnB

Gustavo Henrique Fernandes Carvalho, Heron Trindade Fonseca e Iure Vieira Brandão

07/07/2018

Introdução

A ciência de dados, cada vez mais, ganha espaço e importância no âmbito tecnológico por ser uma área que pode agregar uma quantidade de informações à qualquer área muito interessante. É como disse o Jeff Weiner, CEO do LinkedIn, “Os dados realmente impulsionam tudo o que fazemos”.

A enorme quantidade de dados que temos a nossa disposição atualmente torna difícil a compreensão e extração de informações úteis, como é o caso dos dados de pesquisadores que estão espalhados em enormes bases.

E devido à grande quantidade de informação que se é obtida nos dias atuais, surgiu um campo da ciência responsável por tratar esse conhecimento que é adquirido e armazenado. Essa área responsável pelo tratamento de dados é conhecida como Ciência de Dados, que combina Big Data com mineração de dados, de modo que grandes quantidade de informações sejam tratadas a fim de identificar padrões, regras, relacionamentos ou apenas para poder analisar esses dados de uma maneira mais clara.

Então, a partir de toda essa base que a ciência de dados nos dá e nos proporciona, foi analisado dados de currículos lattes dos professores da Universidade de Brasília, mais especificamente do curso de Comunicação, afim de avaliar a produção científica deles, contextualizar e realizar uma primeira análise sobre o programa de pós-graduação dos cursos de Direito, Ciência da Informação e Comunicação.

CRISP-DM

A metodologia utilizada no processamento dos dados foi a Cross Industry Standard Process for Data Mining (CRISP-DM), que consiste em um fluxo de processamento de dados que está sendo muito utilizado no processamento de Big Data por normalmente ajudar na geração de bons resultados de mineração de dados.

As ferramentas utilizadas na análise e tratamento dos dados foram a linguagem R e o RStudio, com as bibliotecas seguintes bibliotecas:

```
library(jsonlite)
library(dplyr)
library(ggplot2)
library(tibble)
library(knitr)
library(kableExtra)
```

Como podemos observar na Figura 1, o fluxo de mineração de dados seguindo a metodologia CRISP-DM é dividido em seis etapas, descritas individualmente a seguir assim como elas foram aplicadas nesse projeto.

A. *Entendimento do negócio*

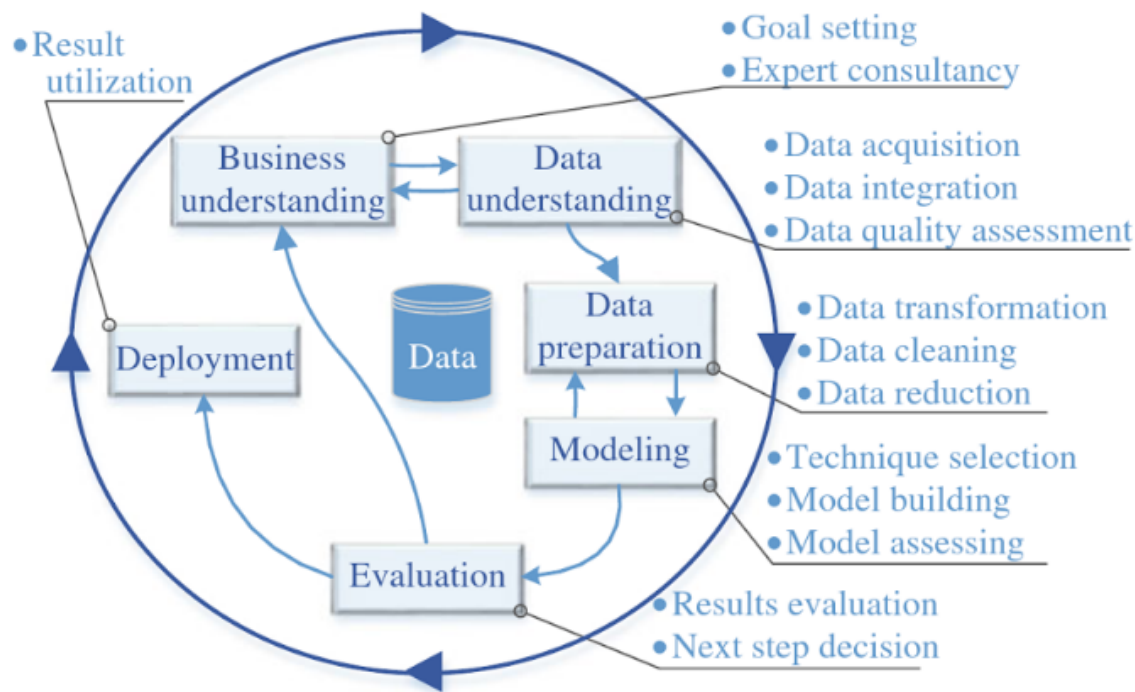


Figure 1: Fluxo de processo do CRISP-DM

É nessa etapa que o profissional que estará analisando os dados deve procurar conhecer mais a respeito do ambiente no qual os dados foram obtidos para entender melhor o real significado dos dados que serão trabalhados.

Nesse projeto essa etapa consistiu em uma pesquisa a respeito dos cursos de Direito, Ciência da Informação e Comunicação na universidade de Brasília e a conversa com por e-mail com professores dos cursos.

B. *Compreensão dos dados*

Nessa segunda etapa o objetivo é inspecionar, organizar e conhecer quais dados temos a nossa disposição para análise.

A bases de dados que serão utilizadas nesse projeto são a OASIS, BDTD e de publicações dos cursos mencionados acima.

C. *Preparação dos dados*

Nessa etapa realizamos a limpeza dos dados para ser possível trabalharmos com eles nas próximas etapas, seguindo o padrão de 'tidy data'.

D. *Modelagem*

É nessa etapa que a mineração dos dados ocorre de fato, iniciamos com uma análise preditiva das possíveis relações dos dados e com uma análise mais profunda as previsões iniciais podem ser validadas, além de termos novos insights a respeito dos dados durante esse processo de validação.

E. *Avaliação*

A avaliação é uma validação mais qualitativa que quantitativa, isto é, uma validação se os resultados obtidos na modelagem são realmente úteis. Essa etapa é essencial quando a análise dos dados está sendo feita para

um cliente, pois com ela os cientistas de dados podem ter uma ideia melhor se os resultados das análises estão de acordo com o que o cliente espera.

F. *Desenvolvimento*

Agora com todas as validações e ultimos ajustes nos algoritmos de análise dos dados, o modelo de mineração finalmente pode começar a ser aplicado em soluções de problemas reais.

CRISP-DM Fase 1 - Entendimento do Negócio

O que é o Sistema Nacional de Pós-Graduação? (Contextualização)

A produção do conhecimento científico, no Brasil, é predominantemente efetuada por meio do Sistema Nacional de Pós-Graduação - SNPG, e mais fortemente relacionada com a formação de doutores nesse sistema (Pátaro e Mezzomo, 2013), por meio de cursos de pós-graduação *strictu sensu*.

Fernandes e Sampaio (2017) já indicaram que a ciência é reconhecidamente um elemento essencial para o desenvolvimento social e econômico de qualquer nação. Assim sendo, faz-se mister aprimorar o SNPG como forma de promoção desse crescimento, visando maximizar o retorno decorrente do emprego dos recursos nele aplicados. A promoção do crescimento do SNPG se dá predominantemente por meio de avaliações regulares de seus programas de pós-graduação, sob responsabilidade da CAPES, que realiza a cada quatro anos um complexo (Leite, 2018, p. 13) e custoso processo de coleta de dados, análise e deliberação sobre as pós-graduações *strictu sensu*, em coerência com o estabelecido no Plano Nacional de Pós-Graduação (PNPG) 2012-2020 (CAPES, 2010) e nos diversos documentos que definem os critérios de organização da pós-graduação em cada área do conhecimento (CAPES, 2018). Leite (2018) faz uma apresentação geral de como se organizam e são avaliadas as pós-graduações no Brasil.

O Plano Nacional de Pós-Graduação (PNPG), por outro lado, define diretrizes estratégicas para desenvolvimento da pós-graduação brasileira, que deve abordar prioritariamente grandes temas de interesse nacional, tais como a redução das assimetrias de desenvolvimento entre as regiões do Brasil, a formação de professores para a educação básica, a formação de recursos humanos para as empresas, a resposta aos grandes desafios brasileiros sobre Água, Energia, Transporte, Controle de Fronteiras, Agronegócio, Amazônia, Amazônia Azul (Mar), Saúde, Defesa, Programa Espacial, além de Justiça, Segurança Pública, Criminologia e Desequilíbrio Regional. O PNPG também traça as diretrizes para financiamento da pós-graduação e sua internacionalização, apresentando conclusões e recomendações.

As avaliações do SNPG, ao atribuírem mensurações de desempenho às diversas pós-graduações que dele fazem parte, geram incentivos e penalidades aos programas, tendo em vista a limitada disponibilidade de recursos para investimento em bolsas, taxas de bancada etc. Embora o sistema seja altamente sofisticado ele é também altamente criticado (Azevedo et al., 2016), sobretudo porque há percalços na busca por um equilíbrio entre as diferentes concepções de finalidade da ciência. Se de um lado a promoção do conhecimento gerado predominantemente nas ditas ciências *hard* contribui para criar fluxos econômicos mais intensos, isso não significa que essa promoção possa ocorrer em detrimento da menor promoção na geração de conhecimento sobre problemas sociais, predominantemente gerado nas ditas ciências *soft*, especialmente das áreas de humanidades, sob pena de ampliação de desigualdades (Azevedo et al., 2016).

Não há solução simples, mas postula-se, nesta disciplina, que uma maior agilidade na avaliação e a utilização de critérios mais objetivos, poderá facilitar a melhoria do sistema.

A UnB dentro do Sistema Nacional de Pós-Graduação (Contextualização)

O que é a UnB?

A Universidade de Brasília (UnB) foi inaugurada, em 21 de abril de 1962, após dois anos da criação de Brasília. A criação da Universidade foi fruto do sonho e do trabalho de educadores como Darcy Ribeiro e Anísio Teixeira.

As regras, a estrutura e concepção da Universidade foram definidas pelo Plano Orientador, uma espécie de Carta Magna, datada de 1962, e ainda hoje em vigor. O Plano foi a primeira publicação da Editora UnB e mostra o espírito inovador da instituição. “Só uma universidade nova, inteiramente planejada, estruturada em bases mais flexíveis, poderá abrir perspectivas de pronta renovação do nosso ensino superior”, diz o Plano Orientador.

Hoje a Universidade de Brasília é formada por 4 campi, a Faculdade UnB Planaltina (FUP), a primeira fora do Plano Piloto, a Faculdade de Ceilândia (FCE) e a Faculdade do Gama (FGA) e o Campus Universitário Darcy Ribeiro.

A missão da Universidade é ser uma instituição inovadora, comprometida com a excelência acadêmica, científica e tecnológica formando cidadãos conscientes do seu papel transformador na sociedade, respeitadas a ética e a valorização de identidades e culturas com responsabilidade social.

Descrição das pós-graduações da UnB

O programa de pós-graduação da faculdade de comunicação (FAC) é dividido em dois grupos: mestrado e doutorado. Além disso, o Programa de Pós-Graduação em Comunicação da Faculdade de Comunicação da Universidade de Brasília (PPG/FAC/UnB) completou, em 2018, 44 anos de atividades ininterruptas de formação e pesquisa de excelência na área. Com um Mestrado consolidado, em atividade desde 1974, e um Doutorado de alto padrão, instalado em 2003, o PPG/FAC/UnB, por sua localização central, tem marcado sua atuação por uma forte inserção regional, nacional e mesmo internacional e pela formação de novos quadros qualificados para o ensino, a pesquisa e a extensão principalmente para as regiões Centro-Oeste e Norte.

O Programa, a partir de suas quatro linhas de pesquisa, tem reconhecido destaque nos estudos da Comunicação e da Sociedade, com ênfase em epistemologia, teorias e tecnologias de comunicação, políticas de comunicação e de cultura (das quais é pioneiro no Brasil), jornalismo (com atualizado enfoque para as transformações estruturais) e estudos de linguagens e narrativas. Os docentes são líderes de grupos de pesquisa e têm se dedicado a um intenso esforço de publicação em conjunto com os discentes do Programa e parceiros de pesquisa (o que pode ser atestado pelos indicadores quantitativos e qualitativos da produção intelectual).

Ao longo dos 43 anos de atividade do Mestrado e 14 anos de existência do Doutorado, o PPG/FAC/UnB formou 522 Mestres e 112 Doutores.

Já o curso de Direito da Universidade de Brasília foi criado junto com a Universidade em 1962, dentro do Instituto Central de Ciências Humanas (ICCH) até a criação da Faculdade de Ciências Jurídicas e Sociais (FCJS), em 1967.

O curso de mestrado em Direito foi criado apenas em 1975. O curso de pós graduação em direito na Universidade de Brasília possui um grande foco em Direito e Estado, sendo referencia nacional nos estudos de Direito Público.

Atualmente o curso de graduação é oferecido tanto no período diurno como noturno e os coordenadores dos cursos de cada período são, respectivamente, Professor Henrique Araújo Costa e a Professora Inez Lopes. O coordenador dos cursos de pós-graduação é o Professor Valcir Gassen.

A Faculdade de Direito da Universidade de Brasília também possui um periódico anual chamado Revista Direito.UnB no qual são publicados artigos de alunos e professores.

E, por fim, o Programa de Pós-graduação em Ciência da Informação (PPGCIInf), da Faculdade de Ciência da Informação (FCI), da Universidade de Brasília (UnB), visa aprofundar os conhecimentos adquiridos por graduados e pós-graduados em cursos superiores, regulamentados pelos órgãos competentes no Brasil, com interesses afins ao Programa, dando-lhes oportunidade de desenvolver competência científica e capacidade profissional e criadora em Ciência da Informação, formando pesquisadores, professores e profissionais de alto nível, aptos a desenvolver pesquisas e realizar inovações nesta área do saber. O PPGCIInf tem como área de concentração: Gestão da Informação. O PPGCIInf compreende os cursos de Mestrado Acadêmico e Doutorado (stricto sensu) na área de Ciência da Informação.

O Mestrado Acadêmico objetiva promover a competência científica de graduados, contribuindo para a formação de docentes e pesquisadores na área da informação. O Doutorado visa formar profissionais de alto nível, que possam atuar como pesquisadores autônomos e como docentes na área de Ciência da Informação, buscando: propiciar visão abrangente da área; desenvolver sólida base teórica relativa à área; estimular o desenvolvimento da produção científica da área, com projetos de pesquisa inovadores e socialmente relevantes.

A Linha de Pesquisa Organização da Informação busca propor conhecimentos nos níveis epistemológico, científico e prático relativos à origem, coleta, organização, estocagem, recuperação, interpretação, transmissão, transformação e uso da informação. Neste contexto, relaciona-se com a natureza da informação, a terminologia e modelos de tratamento e recuperação de informações; as necessidades dos usuários de informação e suas implicações; a identificação dos recursos necessários a partir dos tipos e formatos; a identificação, o tratamento e a recuperação de informações adequadas para o usuário; a formulação de políticas, estratégias, planejamentos, normas e processos relacionados a diferentes espaços de informação.

A Linha de Pesquisa Comunicação e Mediação da Informação reúne estudos que buscam conhecimentos nos níveis epistemológico, científico e prático sobre processos de comunicação em diversos contextos e setores da sociedade. Os estudos desenvolvidos sob esta linha privilegiam a busca pelo entendimento dos fenômenos relacionados ao fluxo da informação, os atores que deles participam e os canais utilizados. Os estudos dessa linha contemplam ainda fatores internos e externos que influenciam tais fluxos e a produção e aplicação de indicadores para sua avaliação. A linha inclui também estudos sobre políticas, estratégias e planejamento dos processos de comunicação nos diversos contextos e setores da sociedade e desdobramentos sociais, políticos, culturais e econômicos da comunicação e acesso à informação, e ainda estudos relacionados às profissões ligadas a esses processos.

O que a Organização precisa realmente alcançar?

Vários stakeholders estão envolvidos no projeto em curso, e poderíamos considerar cada um deles como distintas organizações que possuem interesses distintos e complementares. Elas são: * A Disciplina Ciência de Dados para Todos 2018.1, que quer comprovar que seus alunos dominam ferramentas e técnicas de ciência de dados, para fins de avaliação de rendimento da disciplina. * A UnB, representada pelos decanatos de pós-graduação (DPG) e de pesquisa e inovação (DPI), que querem dispor de instrumentos para realização de avaliações contínuas de suas pós-graduações. * O SNPG, que assim com o DPG e DPI, também pode se beneficiar do uso de instrumentos para realização de avaliações contínuas de suas pós-graduações. * Os interessados em melhor conhecer o que é produzido pelo Sistema Nacional de Pós-graduação, como empresas privadas, que querem desfrutar dos benefícios gerados pela ciência brasileira.

A fim de dar maior fidelidade e homogeneidade ao exercício realizado na disciplina, focaremos em atendimento aos interesses comuns das organizações DPI, DPG e CAPES, que desejam dispor de instrumentos ágeis para avaliação contínua da pós-graduação brasileira.

Com base no exposto, o objetivo do trabalho final a ser alcançado pelos produtos de mineração de dados desenvolvido pelos alunos da disciplina Ciência de Dados para Todos é produzir, tomando por base inicial os dados fornecidos pelos professores responsáveis pela disciplina, ferramentas para análise e avaliação contínuas e de baixo custo, do desempenho de um conjunto de pós-graduações que estão vinculadas a uma mesma subárea ou grupo de conhecimento. Cada área de pós-graduação apresenta suas características peculiares, assim como cada um dos programas vinculados a essas áreas. Como já informado, características peculiares de cada programa podem ser obtidas a partir de visita ao site da CAPES (2018).

Avaliação das Circunstâncias

Este documento serve como base para a realização dos trabalhos dos alunos. apresenta limitações no tocante à quantidade pequena de dados que serão empregados para análises e avaliações, tendo em vista sua finalidade maior que é a didática, de permitir aos alunos demonstrarem a capacidade de aplicação das técnicas e ferramentas apreendidas durante o semestre.

CRISP-DM Fase 2 - Entendimento dos Dados

Doravante, a fim de facilitar aos alunos seguirem a metodologia CRISP-DM, os nomes das seções e subseções de texto serão prefixadas com o número e nome da fase e atividade genérica do CRISP-DM. Fica facultado aos grupos seguir ou não a sequência prevista, tendo em vista que se pode retornar às fases anteriores, bem como podem haver atividades que não foram adequadas às características do problema específico sob análise.

CRISP-DM Fase.Atividade 2.1 - Coleta inicial dos dados

Todos os arquivos com dados iniciais a seguir apresentados foram fornecidos pelos professores responsáveis pela disciplina. Os dados foram gerados no mês de maio de 2018, e compilam informações entre os anos de 2010 e 2017. Os arquivos estão no formato JSON, e seus atributos iniciais e conteúdos são apresentados a seguir.

Perfil profissional dos docentes vinculados às pós-graduações do curso de direito, comunicação e ciência da informação

```
json.perfil_ci <- "data/CI.profile.json"
file.info(json.perfil_ci)
```

```
##                size isdir mode                mtime
## data/CI.profile.json 1126228 FALSE  666 2018-06-21 01:20:00
##                                ctime                atime uid  gid
## data/CI.profile.json 2018-07-07 17:21:30 2018-07-07 17:21:55 1000 1000
##                                uname grname
## data/CI.profile.json gustavo gustavo
```

```
json.perfil_comunicacao <- "data/Comunicacao.profile.json"
file.info(json.perfil_comunicacao)
```

```
##                size isdir mode                mtime
## data/Comunicacao.profile.json 1247316 FALSE  666 2018-06-21 01:35:36
##                                ctime                atime uid
## data/Comunicacao.profile.json 2018-07-07 17:20:18 2018-07-07 17:21:55 1000
##                                gid  uname grname
## data/Comunicacao.profile.json 1000 gustavo gustavo
```

```
json.perfil_direito <- "data/Direito.profile.json"
file.info(json.perfil_direito)
```

```
##                size isdir mode                mtime
## data/Direito.profile.json 1782980 FALSE  666 2018-06-21 01:40:18
##                                ctime                atime uid
## data/Direito.profile.json 2018-07-07 17:18:23 2018-07-07 17:21:55 1000
##                                gid  uname grname
```

```
## data/Direito.profile.json 1000 gustavo gustavo
```

Os arquivos “CI.profile.json”, “Comunicacao.profile.json” e “Direito.profile.json” apresentam dados sobre o perfil de todos os docentes vinculados a programas de pós-graduação da UnB, entre 2010 e 2017, nos cursos citados anteriormente.

Esse arquivo foi fornecido pelos docentes responsáveis pela disciplina.

Orientações de mestrado e doutorado realizadas pelos docentes vinculados às pós-graduações

```
json.advise_ci <- "data/CI.advise.json"
file.info(json.advise_ci)
```

```
##              size isdir mode              mtime
## data/CI.advise.json 526741 FALSE  666 2018-06-21 01:20:00
##              ctime              atime uid  gid
## data/CI.advise.json 2018-07-07 17:21:42 2018-07-07 17:21:55 1000 1000
##              uname  grname
## data/CI.advise.json gustavo gustavo
```

```
json.advise_comunicacao <- "data/Comunicacao.advise.json"
file.info(json.advise_comunicacao)
```

```
##              size isdir mode              mtime
## data/Comunicacao.advise.json 569439 FALSE  666 2018-06-21 01:35:36
##              ctime              atime uid
## data/Comunicacao.advise.json 2018-07-07 17:20:11 2018-07-07 17:21:55 1000
##              gid  uname  grname
## data/Comunicacao.advise.json 1000 gustavo gustavo
```

```
json.advise_direito <- "data/Direito.advise.json"
file.info(json.advise_direito)
```

```
##              size isdir mode              mtime
## data/Direito.advise.json 934056 FALSE  666 2018-06-21 01:40:18
##              ctime              atime uid  gid
## data/Direito.advise.json 2018-07-07 17:18:10 2018-07-07 17:21:55 1000 1000
##              uname  grname
## data/Direito.advise.json gustavo gustavo
```

Os arquivos “CI.advise.json”, “Comunicacao.advise.json” e “Direito.advise.json” apresentam dados sobre o orientações de mestrado e doutorado feitas por todos os docentes vinculados a programas de pós-graduação da UnB, entre 2010 e 2017, dos cursos citados na seção anterior.

Esse arquivo foi fornecido pelos docentes responsáveis pela disciplina.

Produção bibliográfica gerada pelos docentes vinculados às pós-graduações

```
json.producao.bibliografica_ci <- "data/CI.publication.json"
file.info(json.producao.bibliografica_ci)
```

```
##              size isdir mode              mtime
## data/CI.publication.json 465994 FALSE  666 2018-06-21 01:20:00
##              ctime              atime uid  gid
## data/CI.publication.json 2018-07-07 17:21:24 2018-07-07 17:21:55 1000 1000
```

```
##                               uname  grname
## data/CI.publication.json gustavo gustavo

json.producao.bibliografica_comunicacao <- "data/Comunicacao.publication.json"
file.info(json.producao.bibliografica_comunicacao)
```

```
##                               size isdir mode                               mtime
## data/Comunicacao.publication.json 539523 FALSE  666 2018-06-21 01:35:36
##                               ctime                               atime
## data/Comunicacao.publication.json 2018-07-07 17:20:14 2018-07-07 17:21:55
##                               uid  gid  uname  grname
## data/Comunicacao.publication.json 1000 1000 gustavo gustavo

json.producao.bibliografica_direito <- "data/Direito.publication.json"
file.info(json.producao.bibliografica_direito)
```

```
##                               size isdir mode                               mtime
## data/Direito.publication.json 675135 FALSE  666 2018-06-21 01:40:18
##                               ctime                               atime  uid
## data/Direito.publication.json 2018-07-07 17:18:19 2018-07-07 17:21:55 1000
##                               gid  uname  grname
## data/Direito.publication.json 1000 gustavo gustavo
```

Os arquivos “CI.advise.json”, “Comunicacao.advise.json” e “Direito.advise.json” apresentam dados sobre a produção bibliográfica gerada por todos os docentes vinculados a programas de pós-graduação da UnB, entre 2010 e 2017, dos cursos citados na seção anterior.

CRISP-DM Fase.Atividade 2.2 - Descrição dos Dados

Como já informado, a descrição dos dados verifica se os dados sendo acessados terão potencial para responder às questões de *data mining*. Além disso, deve-se avaliar qual o volume de dados, a estrutura dos dados (tipos), codificações usadas, etc. Neste projeto, a descrição dos dados é responsabilidade parcial dos alunos, tendo em vista que esta seção já oferece uma descrição inicial simplificada. O relatório final deve conter descrições significativas e aprofundadas dos dados.

Descrição dos dados do perfil

OS arquivos “CI.profile.json”, “Comunicacao.profile.json” e “Direito.profile.json” que contém dados que caracterizam o perfil profissional de todos os docentes do grupo, dos 3 cursos, sob análise, podem ser lido por meio do comando seguinte.

```
unb.prof_ci <- fromJSON("data/CI.profile.json")
unb.prof_comunicacao <- fromJSON("data/Comunicacao.profile.json")
unb.prof_direito <- fromJSON("data/Direito.profile.json")
```

A quantidade de docentes de cada curso, sob análise, é apresentada a seguir.

```
length(unb.prof_ci)
```

```
## [1] 27
```

```
length(unb.prof_comunicacao)
```

```
## [1] 31
```

```
length(unb.prof_direito)
```



```
## [1] 37
```

Para gerar uma apresentação inicial dos dados que estão contido nos dados de perfil dos docentes de cada curso, pode-se usar a função `glimpse`, da biblioteca `dplyr`, como ilustra o código seguinte, que apresenta os atributos típicos que podem ser obtidos relativamente a um pesquisador específico, o mais antigo docente ainda em exercício na UnB a ter criado seu registro na plataforma Lattes.

```
library(dplyr)
glimpse(unb.prof_ci[[1]], width = 30)
```

```
## List of 7
## $ nome : chr "Fernando César Lima Leite"
## $ resumo_cv : chr "Graduado em Biblioteconomia. Mestre em Ciência da Informação. Doutor
## $ areas_de_atuacao : 'data.frame': 4 obs. of 4 variables:
## ..$ grande_area : chr [1:4] "CIENCIAS_SOCIAIS_APLICADAS" "CIENCIAS_SOCIAIS_APLICADAS" "CIENCIAS_S
## ..$ area : chr [1:4] "Ciência da Informação" "Ciência da Informação" "Ciência da Informaçã
## ..$ sub_area : chr [1:4] "" "Comunicação Científica" "Biblioteconomia" "Gestão do Conhecimento
## ..$ especialidade: chr [1:4] "" "" "" "Gestão da informação e do conhecimento"
## $ endereco_profissional :List of 8
## ..$ instituicao: chr "Universidade de Brasília"
## ..$ orgao : chr "Faculdade de Ciência da Informação"
## ..$ unidade : chr ""
## ..$ DDD : chr "61"
## ..$ telefone : chr "33072641"
## ..$ bairro : chr "Asa Norte"
## ..$ cep : chr "70919900"
## ..$ cidade : chr "Brasília"
## $ producao_bibliografica :List of 5
## ..$ CAPITULO_DE_LIVRO : 'data.frame': 3 obs. of 13 variables:
## .. ..$ tipo : chr [1:3] "Capítulo de livro publicado" "Capítulo de livro publico
## .. ..$ titulo_do_capitulo : chr [1:3] "Insumos conceituais e práticos para iniciativas de rep
## .. ..$ titulo_do_livro : chr [1:3] "Implantação e gestão de repositórios institucionais: p
## .. ..$ ano : chr [1:3] "2010" "2013" "2017"
## .. ..$ doi : chr [1:3] "" "" ""
## .. ..$ pais_de_publicacao : chr [1:3] "Brasil" "Brasil" "Brasil"
## .. ..$ isbn : chr [1:3] "9788523206550" "9789899870406" "9788570141972"
## .. ..$ nome_da_editora : chr [1:3] "EdUFBA" "Universidade do Minho, Serviços de Documentaç
## .. ..$ numero_da_edicao_revisao: chr [1:3] "" "1" "1"
## .. ..$ organizadores : chr [1:3] "MARCONDES, C. H.; SAYÃO, L. F.; TOUTAIN, L. B.; ROSA, M
## .. ..$ paginas : chr [1:3] "163 - 202" "133 - 150" "35 - 65"
## .. ..$ autores :List of 3
## .. ..$ autores-endogeno :List of 3
## ..$ DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA: 'data.frame': 1 obs. of 9 variables:
## .. ..$ natureza : chr "Cartilha"
## .. ..$ titulo : chr "Boas práticas para a construção de repositórios institucionais da p
## .. ..$ ano : chr "2012"
## .. ..$ pais_de_publicacao: chr "Brasil"
## .. ..$ editora : chr "IBICT"
## .. ..$ doi : chr ""
## .. ..$ numero_de_paginas : chr "34"
## .. ..$ autores :List of 1
## .. ..$ autores-endogeno :List of 1
## ..$ EVENTO : 'data.frame': 5 obs. of 11 variables:
## .. ..$ natureza : chr [1:5] "COMPLETO" "COMPLETO" "RESUMO" "RESUMO" ...
## .. ..$ titulo : chr [1:5] "Software livre para implementação de repositórios digitais e p
```

```

## ..$ nome_do_evento : chr [1:5] "Jornadas Argentinas de Informática (Jornadas de Software Libre"
## ..$ ano_do_trabalho : chr [1:5] "2010" "2012" "2013" "2013" ...
## ..$ pais_do_evento : chr [1:5] "Brasil" "Brasil" "Brasil" "Brasil" ...
## ..$ cidade_do_evento: chr [1:5] "Caba, Argentina" "Rio de Janeiro" "Mexico City" "São Paulo" ..
## ..$ doi : chr [1:5] "" "" "" "" ...
## ..$ classificacao : chr [1:5] "NACIONAL" "NACIONAL" "INTERNACIONAL" "INTERNACIONAL" ...
## ..$ paginas : chr [1:5] "1 - 9" " - " " - " " - " ...
## ..$ autores :List of 5
## ..$ autores-endogeno:List of 5
## ..$ LIVRO : 'data.frame': 1 obs. of 13 variables:
## ..$ titulo : chr "Repositórios institucionais da América Latina e o acesso abe"
## ..$ ano : chr "2017"
## ..$ tipo : chr "LIVRO_PUBLICADO"
## ..$ natureza : chr "TEXTO_INTEGRAL"
## ..$ pais_de_publicacao : chr "Brasil"
## ..$ isbn : chr "9788570131201"
## ..$ doi : chr ""
## ..$ nome_da_editora : chr "IBICT"
## ..$ numero_da_edicao_revisao: chr "1"
## ..$ numero_de_paginas : chr ""
## ..$ numero_de_volumes : chr ""
## ..$ autores :List of 1
## ..$ autores-endogeno :List of 1
## ..$ PERIODICO : 'data.frame': 13 obs. of 10 variables:
## ..$ natureza : chr [1:13] "COMPLETO" "COMPLETO" "COMPLETO" "COMPLETO" ...
## ..$ titulo : chr [1:13] "An Open Access Approach to Scientific Information Management a"
## ..$ periodico : chr [1:13] "Scholarly and Research Communication" "Bibliotecas Universitá"
## ..$ ano : chr [1:13] "2010" "2011" "2012" "2015" ...
## ..$ volume : chr [1:13] "1" "1" "8" "" ...
## ..$ issn : chr [1:13] "03163849" "21799504" "18083536" "15624730" ...
## ..$ paginas : chr [1:13] "1 - 12" "71 - 80" "414 - 430" "22 - 42" ...
## ..$ doi : chr [1:13] "" "" "10.18225/liinc.v8i2.503" "10.5195/biblios.2014.195" ...
## ..$ autores :List of 13
## ..$ autores-endogeno:List of 13
## $ orientacoes_academicas:List of 4
## ..$ ORIENTACAO_CONCLUIDA_DOUTORADO : 'data.frame': 3 obs. of 13 variables:
## ..$ natureza : chr [1:3] "Tese de doutorado" "Tese de doutorado" "Tese de dou"
## ..$ titulo : chr [1:3] "Depósito da produção científica em repositórios ins"
## ..$ ano : chr [1:3] "2015" "2016" "2017"
## ..$ id_lattes_aluno : chr [1:3] "8636654005487585" "8743793460674396" "5676915146365"
## ..$ nome_aluno : chr [1:3] "Marília Augusta de Freitas" "José Antônio Machado d"
## ..$ instituicao : chr [1:3] "Universidade de Brasília" "Universidade de Brasília"
## ..$ curso : chr [1:3] "Ciência da Informação" "Ciências da Informação" "Ci"
## ..$ codigo_do_curso : chr [1:3] "90000011" "90000010" "90000005"
## ..$ bolsa : chr [1:3] "NAO" "NAO" "SIM"
## ..$ agencia_financiadora : chr [1:3] "" "" "Conselho Nacional de Desenvolvimento Científi"
## ..$ codigo_agencia_financiadora: chr [1:3] "" "" "002200000000"
## ..$ nome_orientadores :List of 3
## ..$ id_lattes_orientadores :List of 3
## ..$ ORIENTACAO_CONCLUIDA_MESTRADO : 'data.frame': 7 obs. of 13 variables:
## ..$ natureza : chr [1:7] "Dissertação de mestrado" "Dissertação de mestrado" "
## ..$ titulo : chr [1:7] "Acesso aberto à informação científica: estudo sobre"
## ..$ ano : chr [1:7] "2010" "2013" "2014" "2014" ...
## ..$ id_lattes_aluno : chr [1:7] "" "8018556894529318" "5676915146365703" "1937356712"

```

```
## ..$ nome_aluno : chr [1:7] "Raísa Mendes Fernandes de Souza" "Tainá Batista de A
## ..$ instituicao : chr [1:7] "Universidade Federal de Minas Gerais" "Universidade
## ..$ curso : chr [1:7] "Ciências da Informação" "Ciências da Informação" "C
## ..$ codigo_do_curso : chr [1:7] "32000286" "51500183" "90000002" "51500183" ...
## ..$ bolsa : chr [1:7] "NAO" "NAO" "NAO" "SIM" ...
## ..$ agencia_financiadora : chr [1:7] "" "" "" "Conselho Nacional de Desenvolvimento Cient
## ..$ codigo_agencia_financiadora: chr [1:7] "" "" "" "002200000000" ...
## ..$ nome_orientadores :List of 7
## ..$ id_lattes_orientadores :List of 7
## ..$ ORIENTACAO_EM_ANDAMENTO_MESTRADO:'data.frame': 2 obs. of 13 variables:
## ..$ natureza : chr [1:2] "Dissertação de mestrado" "Dissertação de mestrado"
## ..$ titulo : chr [1:2] "Diretrizes internacionais para bibliotecas prisiona
## ..$ ano : chr [1:2] "2016" "2016"
## ..$ id_lattes_aluno : chr [1:2] "2403536925847453" "1563407482755076"
## ..$ nome_aluno : chr [1:2] "Luanna Vieira Leite" "Danyelle Mayara da Silva"
## ..$ instituicao : chr [1:2] "Universidade de Brasília" "Universidade de Brasília
## ..$ curso : chr [1:2] "Ciência da Informação" "Ciência da Informação"
## ..$ codigo_do_curso : chr [1:2] "90000012" "90000012"
## ..$ bolsa : chr [1:2] "NAO" "NAO"
## ..$ agencia_financiadora : chr [1:2] "" ""
## ..$ codigo_agencia_financiadora: chr [1:2] "" ""
## ..$ nome_orientadores :List of 2
## ..$ id_lattes_orientadores :List of 2
## ..$ OUTRAS_ORIENTACOES_CONCLUIDAS : 'data.frame': 19 obs. of 13 variables:
## ..$ natureza : chr [1:19] "TRABALHO_DE_CONCLUSAO_DE_CURSO_GRADUACAO" "TRABALH
## ..$ titulo : chr [1:19] "Acesso aberto à informação científica: análise bib
## ..$ ano : chr [1:19] "2010" "2010" "2010" "2011" ...
## ..$ id_lattes_aluno : chr [1:19] "" "" "" "" ...
## ..$ nome_aluno : chr [1:19] "Sílvia Regina Batista Mendonça" "Michelli Costa" ".
## ..$ instituicao : chr [1:19] "Universidade de Brasília" "Universidade de Brasília
## ..$ curso : chr [1:19] "Biblioteconomia" "Biblioteconomia" "Biblioteconomia
## ..$ codigo_do_curso : chr [1:19] "90000000" "90000000" "90000000" "90000000" ...
## ..$ bolsa : chr [1:19] "NAO" "NAO" "NAO" "NAO" ...
## ..$ agencia_financiadora : chr [1:19] "" "" "" "" ...
## ..$ codigo_agencia_financiadora: chr [1:19] "" "" "" "" ...
## ..$ nome_orientadores :List of 19
## ..$ id_lattes_orientadores :List of 19
## $ senioridade : chr "9"
```

Uma breve inspeção visual dos atributos do curso de ciência da informação anteriormente apresentados permite inferir que o pesquisador sob análise:

- Atua predominantemente nas áreas de arquivologia, museologia e biblioteconomia.
- Trabalha no Instituto de Ciência de Informação da UnB.

```
glimpse(unb.prof_comunicacao[[1]], width = 30)
```

```
## List of 7
## $ nome : chr "Dione Oliveira Moura"
## $ resumo_cv : chr "Professora e pesquisadora em Comunicação. Possui graduação em Comuni
## $ areas_de_atuacao : 'data.frame': 3 obs. of 4 variables:
## ..$ grande_area : chr [1:3] "CIENCIAS_SOCIAIS_APLICADAS" "CIENCIAS_SOCIAIS_APLICADAS" "CIENCIAS_S
## ..$ area : chr [1:3] "Comunicação" "Comunicação" "Ciência da Informação"
## ..$ sub_area : chr [1:3] "jornalismo científico e ambiental" "Jornalismo e Sociedade" "comunica
## ..$ especialidade: chr [1:3] "" "" ""
## $ endereco_profissional :List of 8
```

```

## ..$ instituicao: chr "Universidade de Brasília"
## ..$ orgao      : chr "Faculdade de Comunicação"
## ..$ unidade    : chr ""
## ..$ DDD        : chr "61"
## ..$ telefone   : chr "33072295"
## ..$ bairro     : chr "Asa Norte"
## ..$ cep        : chr "70900-000"
## ..$ cidade     : chr "BRASILIA"
## $ producao_bibliografica :List of 5
## ..$ CAPITULO_DE_LIVRO:'data.frame': 14 obs. of 13 variables:
## .. ..$ tipo : chr [1:14] "Capítulo de livro publicado" "Capítulo de livro publi
## .. ..$ titulo_do_capitulo : chr [1:14] "Trabalhar a pauta socioambiental diante da imperativa
## .. ..$ titulo_do_livro : chr [1:14] "Comunicação e Cidadania. Conceitos e processos" "Univ
## .. ..$ ano : chr [1:14] "2011" "2011" "2012" "2012" ...
## .. ..$ doi : chr [1:14] "" "" "" "" ...
## .. ..$ pais_de_publicacao : chr [1:14] "Brasil" "Brasil" "Brasil" "Brasil" ...
## .. ..$ isbn : chr [1:14] "9788580030102" "9788565088008" "9788574746296" "97885
## .. ..$ nome_da_editora : chr [1:14] "Francis" "Editora UnB" "Insular" "Editora UnB" ...
## .. ..$ numero_da_edicao_revisao: chr [1:14] "1" "" "1" "3" ...
## .. ..$ organizadores : chr [1:14] "MOURA, Dione O.; GERALDES, Elen; PEREIRA, Fábio; CALA
## .. ..$ paginas : chr [1:14] "37 - 46" "303 - 310" "15 - 38" "309 - 320" ...
## .. ..$ autores :List of 14
## .. ..$ autores-endogeno :List of 14
## ..$ EVENTO : 'data.frame': 19 obs. of 11 variables:
## .. ..$ natureza : chr [1:19] "COMPLETO" "COMPLETO" "COMPLETO" "RESUMO" ...
## .. ..$ titulo : chr [1:19] "Novas demandas para os estudos de jornalismo e temática socio
## .. ..$ nome_do_evento : chr [1:19] "IX Congresso Regional Intercom Centro-Oeste" "9. Encontro Nac
## .. ..$ ano_do_trabalho : chr [1:19] "2010" "2011" "2011" "2011" ...
## .. ..$ pais_do_evento : chr [1:19] "Brasil" "Brasil" "Brasil" "Brasil" ...
## .. ..$ cidade_do_evento: chr [1:19] "Goiânia" "Rio de Janeiro" "Brasília" "Rio de Janeiro" ...
## .. ..$ doi : chr [1:19] "" "" "" "" ...
## .. ..$ classificacao : chr [1:19] "REGIONAL" "NACIONAL" "INTERNACIONAL" "NACIONAL" ...
## .. ..$ paginas : chr [1:19] "1 - 12" "1 - 12" "1 - 12" "20 - 20" ...
## .. ..$ autores :List of 19
## .. ..$ autores-endogeno:List of 19
## ..$ LIVRO : 'data.frame': 9 obs. of 13 variables:
## .. ..$ titulo : chr [1:9] "Comunicação e Cidadania: conceitos e processos" "Jorna
## .. ..$ ano : chr [1:9] "2011" "2012" "2012" "2014" ...
## .. ..$ tipo : chr [1:9] "LIVRO_ORGANIZADO_OU_EDICAO" "LIVRO_ORGANIZADO_OU_EDICAO
## .. ..$ natureza : chr [1:9] "LIVRO" "LIVRO" "LIVRO" "LIVRO" ...
## .. ..$ pais_de_publicacao : chr [1:9] "Brasil" "Brasil" "Brasil" "Brasil" ...
## .. ..$ isbn : chr [1:9] "9788580030105" "9788574746296" "9789898600080" "978857
## .. ..$ doi : chr [1:9] "" "" "" "" ...
## .. ..$ nome_da_editora : chr [1:9] "Francis" "Insular" "Editora do Centro de Estudos de Cor
## .. ..$ numero_da_edicao_revisao: chr [1:9] "1" "1" "1" "1" ...
## .. ..$ numero_de_paginas : chr [1:9] "304" "264" "224" "280" ...
## .. ..$ numero_de_volumes : chr [1:9] "1" "" "1" "1" ...
## .. ..$ autores :List of 9
## .. ..$ autores-endogeno :List of 9
## ..$ PERIODICO : 'data.frame': 13 obs. of 10 variables:
## .. ..$ natureza : chr [1:13] "COMPLETO" "COMPLETO" "COMPLETO" "COMPLETO" ...
## .. ..$ titulo : chr [1:13] "Corresponsales internacionales: problematización em torno a la
## .. ..$ periodico : chr [1:13] "Revista Latinoamericana de Ciencias de la Comunicación" "Revi
## .. ..$ ano : chr [1:13] "2012" "2012" "2012" "2012" ...

```

```

## .. ..$ volume      : chr [1:13] "9" "9" "11" "17" ...
## .. ..$ issn         : chr [1:13] "18073026" "18073026" "22374531" "22381694" ...
## .. ..$ paginas      : chr [1:13] "278 - " "278 - " "27 - " "279 - 289" ...
## .. ..$ doi          : chr [1:13] "" "" "" "" ...
## .. ..$ autores      :List of 13
## .. ..$ autores-endogeno:List of 13
## ..$ TEXTO_EM_JORNAIS :'data.frame': 2 obs. of 9 variables:
## .. ..$ natureza     : chr [1:2] "JORNAL_DE_NOTICIAS" "JORNAL_DE_NOTICIAS"
## .. ..$ titulo       : chr [1:2] "Brasil Submerso (Artigo de Opinião)" "Mídia e educação na era c
## .. ..$ periodico    : chr [1:2] "Correio Braziliense" "Observatório da Imprensa (São Paulo)"
## .. ..$ ano          : chr [1:2] "2011" "2012"
## .. ..$ issn         : chr [1:2] "" "15197670"
## .. ..$ paginas      : chr [1:2] "2 - 2" "1 - 2"
## .. ..$ doi          : chr [1:2] "" ""
## .. ..$ autores      :List of 2
## .. ..$ autores-endogeno:List of 2
## $ orientacoes_academicas:List of 5
## ..$ ORIENTACAO_CONCLUIDA_DOUTORADO      :'data.frame': 8 obs. of 13 variables:
## .. ..$ natureza     : chr [1:8] "Tese de doutorado" "Tese de doutorado" "Tese de dou
## .. ..$ titulo       : chr [1:8] "Controle Social e a Narrativa Moralizante do Telejo
## .. ..$ ano          : chr [1:8] "2011" "2013" "2014" "2014" ...
## .. ..$ id_lattes_aluno : chr [1:8] "0107922722452146" "9022043664068608" "0198365351847
## .. ..$ nome_aluno    : chr [1:8] "Elizena de Jesus Barbosa Rossy" "Ana Lúcia de Mede
## .. ..$ instituicao    : chr [1:8] "Universidade de Brasília" "Universidade de Brasília
## .. ..$ curso         : chr [1:8] "Comunicação" "Doutorado em Comunicação" "Doutorado c
## .. ..$ codigo_do_curso : chr [1:8] "51500191" "90000024" "90000024" "90000024" ...
## .. ..$ bolsa        : chr [1:8] "SIM" "SIM" "NAO" "NAO" ...
## .. ..$ agencia_financiadora : chr [1:8] "Coordenação de Aperfeiçoamento de Pessoal de Nível
## .. ..$ codigo_agencia_financiadora: chr [1:8] "045000000000" "045000000000" "" "" ...
## .. ..$ nome_orientadores :List of 8
## .. ..$ id_lattes_orientadores :List of 8
## ..$ ORIENTACAO_CONCLUIDA_MESTRADO      :'data.frame': 2 obs. of 13 variables:
## .. ..$ natureza     : chr [1:2] "Dissertação de mestrado" "Dissertação de mestrado"
## .. ..$ titulo       : chr [1:2] "Racismo em pauta: a pluralidade confrontada no noti
## .. ..$ ano          : chr [1:2] "2011" "2013"
## .. ..$ id_lattes_aluno : chr [1:2] "5925686589614354" "6722668433844386"
## .. ..$ nome_aluno    : chr [1:2] "Isabel Cristina Clavelin da Rosa" "Juliana César Nur
## .. ..$ instituicao    : chr [1:2] "Universidade de Brasília" "Universidade de Brasília
## .. ..$ curso         : chr [1:2] "Mestrado em Comunicação" "Mestrado em Comunicação"
## .. ..$ codigo_do_curso : chr [1:2] "90000023" "90000023"
## .. ..$ bolsa        : chr [1:2] "NAO" "NAO"
## .. ..$ agencia_financiadora : chr [1:2] "" ""
## .. ..$ codigo_agencia_financiadora: chr [1:2] "" ""
## .. ..$ nome_orientadores :List of 2
## .. ..$ id_lattes_orientadores :List of 2
## ..$ ORIENTACAO_CONCLUIDA_POS_DOUTORADO      :'data.frame': 2 obs. of 13 variables:
## .. ..$ natureza     : chr [1:2] "Supervisão de pós-doutorado" "Supervisão de pós-dou
## .. ..$ titulo       : chr [1:2] "Pesquisa UnBClick tv: midiamorfose da televisão par
## .. ..$ ano          : chr [1:2] "2011" "2014"
## .. ..$ id_lattes_aluno : chr [1:2] "" ""
## .. ..$ nome_aluno    : chr [1:2] "Alzimar Ramalho" "Helder Prior"
## .. ..$ instituicao    : chr [1:2] "Universidade de Brasília" "UnB"
## .. ..$ curso         : chr [1:2] "" ""
## .. ..$ codigo_do_curso : chr [1:2] "" ""

```

```
## ..$ bolsa : chr [1:2] "NAO" "SIM"
## ..$ agencia_financiadora : chr [1:2] "" "Coordenação de Aperfeiçoamento de Pessoal de Nível Superior"
## ..$ codigo_agencia_financiadora: chr [1:2] "" "045000000000"
## ..$ nome_orientadores :List of 2
## ..$ id_lattes_orientadores :List of 2
## ..$ ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA:'data.frame': 2 obs. of 13 variables:
## ..$ natureza : chr [1:2] "Iniciação Científica" "Iniciação Científica"
## ..$ titulo : chr [1:2] "Revista Capricho: representações sociais do feminino" "Revista Capricho: representações sociais do feminino"
## ..$ ano : chr [1:2] "2017" "2017"
## ..$ id_lattes_aluno : chr [1:2] "" ""
## ..$ nome_aluno : chr [1:2] "Melissa Maria de Oliveira Duarte" "Lis Gabriela de Azevedo"
## ..$ instituicao : chr [1:2] "UnB" "UnB"
## ..$ curso : chr [1:2] "" "Comunicação"
## ..$ codigo_do_curso : chr [1:2] "" "90000051"
## ..$ bolsa : chr [1:2] "SIM" "NAO"
## ..$ agencia_financiadora : chr [1:2] "Conselho Nacional de Desenvolvimento Científico e Tecnológico" "Conselho Nacional de Desenvolvimento Científico e Tecnológico"
## ..$ codigo_agencia_financiadora: chr [1:2] "002200000000" ""
## ..$ nome_orientadores :List of 2
## ..$ id_lattes_orientadores :List of 2
## ..$ OUTRAS_ORIENTACOES_CONCLUIDAS :'data.frame': 44 obs. of 13 variables:
## ..$ natureza : chr [1:44] "TRABALHO_DE_CONCLUSAO_DE_CURSO_GRADUACAO" "TRABALHO_DE_CONCLUSAO_DE_CURSO_GRADUACAO" ...
## ..$ titulo : chr [1:44] "Documentário Entre Vãos (Entre Vãos, um documentário sobre a vida de)" "Documentário Entre Vãos (Entre Vãos, um documentário sobre a vida de)" ...
## ..$ ano : chr [1:44] "2010" "2010" "2010" "2010" ...
## ..$ id_lattes_aluno : chr [1:44] "" "" "" "" ...
## ..$ nome_aluno : chr [1:44] "Luisa Ramos Caetano" "Mônica Beatriz Piñeda Godoy" ...
## ..$ instituicao : chr [1:44] "Universidade de Brasília" "Universidade de Brasília" ...
## ..$ curso : chr [1:44] "Comunicação" "Comunicação" "Comunicação" "Comunicação" ...
## ..$ codigo_do_curso : chr [1:44] "90000000" "90000022" "90000022" "90000022" ...
## ..$ bolsa : chr [1:44] "NAO" "NAO" "SIM" "SIM" ...
## ..$ agencia_financiadora : chr [1:44] "" "" "UnB" "UnB" ...
## ..$ codigo_agencia_financiadora: chr [1:44] "" "" "002700000999" "002700000999" ...
## ..$ nome_orientadores :List of 44
## ..$ id_lattes_orientadores :List of 44
## $ senioridade : chr "7"
```

Uma breve inspeção visual dos atributos do curso de Comunicação anteriormente apresentados permite inferir que o pesquisador sob análise:

- Atua predominantemente nas áreas de comunicação visual e jornalismo.
- Trabalha no Instituto de Comunicação da UnB.

```
glimpse(unb.prof_direito[[1]], width = 30)
```

```
## List of 7
## $ nome : chr "Alexandre Araújo Costa"
## $ resumo_cv : chr "Professor Adjunto da Faculdade de Direito da Universidade de Brasília"
## $ areas_de_atuacao :'data.frame': 6 obs. of 4 variables:
## ..$ grande_area : chr [1:6] "CIENCIAS_HUMANAS" "CIENCIAS_SOCIAIS_APLICADAS" "CIENCIAS_SOCIAIS_APLICADAS" ...
## ..$ area : chr [1:6] "Ciência Política" "Direito" "Direito" "Direito" ...
## ..$ sub_area : chr [1:6] "Estado e Governo" "Direito e Tecnologia" "Filosofia do Direito" "Ética" ...
## ..$ especialidade: chr [1:6] "Política e Direito" "" "" "" ...
## $ endereco_profissional :List of 8
## ..$ instituicao: chr "Universidade de Brasília"
## ..$ orgao : chr ""
## ..$ unidade : chr ""
## ..$ DDD : chr "61"
```

```

## ..$ telefone : chr "33072347"
## ..$ bairro : chr "Asa Norte"
## ..$ cep : chr "70000000"
## ..$ cidade : chr "Brasília"
## $ producao_bibliografica :List of 6
## ..$ CAPITULO_DE_LIVRO :'data.frame': 6 obs. of 13 variables:
## .. ..$ tipo : chr [1:6] "Capítulo de livro publicado" "Capítulo de livro publica
## .. ..$ titulo_do_capitulo : chr [1:6] "Transconstitucionalismo?" "Um Perfil das Decisões de P
## .. ..$ titulo_do_livro : chr [1:6] "Anais do XIX Congresso Nacional do CONPEDI" "Anais do 1
## .. ..$ ano : chr [1:6] "2010" "2010" "2014" "2014" ...
## .. ..$ doi : chr [1:6] "" "" "" "" ...
## .. ..$ pais_de_publicacao : chr [1:6] "Brasil" "Brasil" "Brasil" "Brasil" ...
## .. ..$ isbn : chr [1:6] "9788578400477" "9788578400477" "9788555050213" "978856
## .. ..$ nome_da_editora : chr [1:6] "Boiteux" "Fundação Boiteux" "CONPEDI" "CONPEDI" ...
## .. ..$ numero_da_edicao_revisao: chr [1:6] "" "" "1" "1" ...
## .. ..$ organizadores : chr [1:6] "CONPEDI" "CONPEDI" "CONPEDI, UFPB" "CONPEDI; UFPB" ...
## .. ..$ paginas : chr [1:6] "5432 - 5448" "2013 - 2040" "227 - 256" "209 - 227" ...
## .. ..$ autores :List of 6
## .. ..$ autores-endogeno :List of 6
## ..$ DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA:'data.frame': 14 obs. of 9 variables:
## .. ..$ natureza : chr [1:14] "Entrevista realizada" "Artigo publicado em página na intern
## .. ..$ titulo : chr [1:14] "Derrubando as cercas do saber jurídico: a abertura da Univer
## .. ..$ ano : chr [1:14] "2010" "2011" "2011" "2011" ...
## .. ..$ pais_de_publicacao: chr [1:14] "Brasil" "Brasil" "Brasil" "Brasil" ...
## .. ..$ editora : chr [1:14] "Universidade Federal de Santa Catarina" "" "" "" ...
## .. ..$ doi : chr [1:14] "" "" "" "" ...
## .. ..$ numero_de_paginas : chr [1:14] "18" "" "" "" ...
## .. ..$ autores :List of 14
## .. ..$ autores-endogeno :List of 14
## ..$ EVENTO :'data.frame': 2 obs. of 11 variables:
## .. ..$ natureza : chr [1:2] "RESUMO_EXPANDIDO" "COMPLETO"
## .. ..$ titulo : chr [1:2] "AS SENTENÇAS ADITIVAS À LUZ DA HERMENÊUTICA FILOSÓFICA" "Supren
## .. ..$ nome_do_evento : chr [1:2] "V Jornada Brasileira de Filosofia do Direito" "8o Encontro da A
## .. ..$ ano_do_trabalho : chr [1:2] "2011" "2012"
## .. ..$ pais_do_evento : chr [1:2] "Brasil" "Brasil"
## .. ..$ cidade_do_evento: chr [1:2] "Belo Horizonte" "Gramado-RS"
## .. ..$ doi : chr [1:2] "" ""
## .. ..$ classificacao : chr [1:2] "NACIONAL" "NACIONAL"
## .. ..$ paginas : chr [1:2] "53 - 58" " - "
## .. ..$ autores :List of 2
## .. ..$ autores-endogeno:List of 2
## ..$ LIVRO :'data.frame': 1 obs. of 13 variables:
## .. ..$ titulo : chr "Teoria Dialética do Direito: A filosofia jurídica de Roberto
## .. ..$ ano : chr "2017"
## .. ..$ tipo : chr "LIVRO_PUBLICADO"
## .. ..$ natureza : chr "NAO_INFORMADO"
## .. ..$ pais_de_publicacao : chr "Brasil"
## .. ..$ isbn : chr "9788587999078"
## .. ..$ doi : chr ""
## .. ..$ nome_da_editora : chr "Faculdade de Direito - UnB"
## .. ..$ numero_da_edicao_revisao: chr "1"
## .. ..$ numero_de_paginas : chr "110"
## .. ..$ numero_de_volumes : chr "1"
## .. ..$ autores :List of 1

```

```

## ..$ autores-endogeno :List of 1
## ..$ PERIODICO :'data.frame': 19 obs. of 10 variables:
## ..$ natureza : chr [1:19] "COMPLETO" "COMPLETO" "COMPLETO" "COMPLETO" ...
## ..$ titulo : chr [1:19] "Instrumentalismo x Neoinstitucionalismo: uma avaliação das cr
## ..$ periodico : chr [1:19] "Revista Brasileira de Direito Processual (Impresso)" "Constit
## ..$ ano : chr [1:19] "2010" "2010" "2011" "2011" ...
## ..$ volume : chr [1:19] "72" "35" "192" "19" ...
## ..$ issn : chr [1:19] "01002589" "19838646" "01001981" "15184471" ...
## ..$ paginas : chr [1:19] "127 - 141" "6 - " "397 - 418" "198 - 227" ...
## ..$ doi : chr [1:19] "" "" "" "" ...
## ..$ autores :List of 19
## ..$ autores-endogeno:List of 19
## ..$ TEXTO_EM_JORNAIS :'data.frame': 1 obs. of 9 variables:
## ..$ natureza : chr "JORNAL_DE_NOTICIAS"
## ..$ titulo : chr "Intervenção Federal no DF"
## ..$ periodico : chr "UnB Agência"
## ..$ ano : chr "2010"
## ..$ issn : chr ""
## ..$ paginas : chr " - "
## ..$ doi : chr ""
## ..$ autores :List of 1
## ..$ autores-endogeno:List of 1
## $ orientacoes_academicas:List of 4
## ..$ ORIENTACAO_CONCLUIDA_DOUTORADO :'data.frame': 1 obs. of 13 variables:
## ..$ natureza : chr "Tese de doutorado"
## ..$ titulo : chr "Jurisdição, racionalidade e construção de sentido: a forç
## ..$ ano : chr "2016"
## ..$ id_lattes_aluno : chr "9329102524031487"
## ..$ nome_aluno : chr "Alexandre Douglas Zaidan de Carvalho"
## ..$ instituicao : chr "Universidade de Brasília"
## ..$ curso : chr "Direito"
## ..$ codigo_do_curso : chr "51500175"
## ..$ bolsa : chr "NAO"
## ..$ agencia_financiadora : chr ""
## ..$ codigo_agencia_financiadora: chr ""
## ..$ nome_orientadores :List of 1
## ..$ id_lattes_orientadores :List of 1
## ..$ ORIENTACAO_CONCLUIDA_MESTRADO :'data.frame': 9 obs. of 13 variables:
## ..$ natureza : chr [1:9] "Dissertação de mestrado" "Dissertação de mestrado"
## ..$ titulo : chr [1:9] "Ativismo Digital: um estudo sobre blogs ativistas"
## ..$ ano : chr [1:9] "2012" "2013" "2014" "2014" ...
## ..$ id_lattes_aluno : chr [1:9] "5991363955648924" "6397489628008813" "5647095651363
## ..$ nome_aluno : chr [1:9] "Marina Cruz Vieira Villela" "Erika Neves Oliveira"
## ..$ instituicao : chr [1:9] "Universidade de Brasília" "Universidade de Brasília
## ..$ curso : chr [1:9] "Ciência Política" "Ciência Política" "Direito" "Ciê
## ..$ codigo_do_curso : chr [1:9] "51500264" "51500264" "51500175" "51500264" ...
## ..$ bolsa : chr [1:9] "NAO" "NAO" "NAO" "SIM" ...
## ..$ agencia_financiadora : chr [1:9] "" "" "" "Coordenação de Aperfeiçoamento de Pessoal
## ..$ codigo_agencia_financiadora: chr [1:9] "" "" "" "045000000000" ...
## ..$ nome_orientadores :List of 9
## ..$ id_lattes_orientadores :List of 9
## ..$ ORIENTACAO_EM_ANDAMENTO_DOUTORADO:'data.frame': 4 obs. of 13 variables:
## ..$ natureza : chr [1:4] "Tese de doutorado" "Tese de doutorado" "Tese de dou
## ..$ titulo : chr [1:4] "Política e Governamentalidade" "Direito e Racionali

```



```
## .. ..$ ano : chr [1:4] "2014" "2015" "2015" "2015"
## .. ..$ id_lattes_aluno : chr [1:4] "9662096422086996" "1044955971496851" "8047384065910"
## .. ..$ nome_aluno : chr [1:4] "Fernando Modelli" "Ricardo de Lins e Horta" "Kelton"
## .. ..$ instituicao : chr [1:4] "Universidade de Brasília" "Universidade de Brasília"
## .. ..$ curso : chr [1:4] "Ciência Política" "Direito" "Direito" "Direito"
## .. ..$ codigo_do_curso : chr [1:4] "51500264" "51500175" "51500175" "51500175"
## .. ..$ bolsa : chr [1:4] "SIM" "NAO" "NAO" "NAO"
## .. ..$ agencia_financiadora : chr [1:4] "Coordenação de Aperfeiçoamento de Pessoal de Nível S"
## .. ..$ codigo_agencia_financiadora: chr [1:4] "045000000000" "" "" ""
## .. ..$ nome_orientadores :List of 4
## .. ..$ id_lattes_orientadores :List of 4
## ..$ OUTRAS_ORIENTACOES_CONCLUIDAS :'data.frame': 30 obs. of 13 variables:
## .. ..$ natureza : chr [1:30] "TRABALHO_DE_CONCLUSAO_DE_CURSO_GRADUACAO" "TRABALHO"
## .. ..$ titulo : chr [1:30] "Força de Lei em Grande Sertão: Veredas" "O sistema"
## .. ..$ ano : chr [1:30] "2010" "2011" "2011" "2011" ...
## .. ..$ id_lattes_aluno : chr [1:30] "" "" "" "" ...
## .. ..$ nome_aluno : chr [1:30] "Bruno Amâncio Vial" "Marconi Borges" "Hércules Cez"
## .. ..$ instituicao : chr [1:30] "Universidade de Brasília" "Universidade de Brasília"
## .. ..$ curso : chr [1:30] "Bacharelado em Direito" "Ciência Política" "Ciência"
## .. ..$ codigo_do_curso : chr [1:30] "90000000" "90000016" "90000016" "90000018" ...
## .. ..$ bolsa : chr [1:30] "NAO" "NAO" "NAO" "NAO" ...
## .. ..$ agencia_financiadora : chr [1:30] "" "" "" "" ...
## .. ..$ codigo_agencia_financiadora: chr [1:30] "" "" "" "" ...
## .. ..$ nome_orientadores :List of 30
## .. ..$ id_lattes_orientadores :List of 30
## $ senioridade : chr "9"
```

Uma breve inspeção visual dos atributos do curso de Direito anteriormente apresentados permite inferir que o pesquisador sob análise:

- Atua predominantemente na área de direito.
- Trabalha na Faculdade de Direito da UnB.

Descrição dos dados de orientações de Ciência da Informação

```
unb.adv_ci <- fromJSON("data/CI.advise.json")
names(unb.adv_ci)

## [1] "ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO"
## [2] "ORIENTACAO_EM_ANDAMENTO_DOUTORADO"
## [3] "ORIENTACAO_EM_ANDAMENTO_MESTRADO"
## [4] "ORIENTACAO_EM_ANDAMENTO_GRADUACAO"
## [5] "ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA"
## [6] "ORIENTACAO_CONCLUIDA_POS_DOUTORADO"
## [7] "ORIENTACAO_CONCLUIDA_DOUTORADO"
## [8] "ORIENTACAO_CONCLUIDA_MESTRADO"
## [9] "OUTRAS_ORIENTACOES_CONCLUIDAS"

names(unb.adv_ci$ORIENTACAO_CONCLUIDA_DOUTORADO)

## [1] "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017"

length(unb.adv_ci$ORIENTACAO_CONCLUIDA_DOUTORADO$`2016`$natureza)

## [1] 11
```

```
head(sort(table(unb.adv_ci$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)
```

```
##
##                               Ciências da Informação
##                               10
## Programa de Pós-Graduação em Ciência da Informação
##                               2
##                               Ciência da Informação
##                               1
##                               Ciências Sociais
##                               1
##                               Educação
##                               1
##                               História da Ciência / Museologia
##                               1
##                               História e Filosofia da Ciência Museologia
##                               1
```

```
head(sort(table(unb.adv_ci$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)
```

```
##
##                               Ciências da Informação
##                               4
##                               Comunicação
##                               2
##                               Ciência da Informação
##                               1
## Programa de Pós-Graduação em Ciência da Informação
##                               1
```

Descrição dos dados de orientações de Comunicação

```
unb.adv_comunicacao <- fromJSON("data/Comunicacao.advise.json")
names(unb.adv_comunicacao)
```

```
## [1] "ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO"
## [2] "ORIENTACAO_EM_ANDAMENTO_DOUTORADO"
## [3] "ORIENTACAO_EM_ANDAMENTO_MESTRADO"
## [4] "ORIENTACAO_EM_ANDAMENTO_GRADUACAO"
## [5] "ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA"
## [6] "ORIENTACAO_CONCLUIDA_POS_DOUTORADO"
## [7] "ORIENTACAO_CONCLUIDA_DOUTORADO"
## [8] "ORIENTACAO_CONCLUIDA_MESTRADO"
## [9] "OUTRAS_ORIENTACOES_CONCLUIDAS"
```

```
names(unb.adv_comunicacao$ORIENTACAO_CONCLUIDA_DOUTORADO)
```

```
## [1] "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017"
```

```
length(unb.adv_comunicacao$ORIENTACAO_CONCLUIDA_DOUTORADO$`2016`$natureza)
```

```
## [1] 8
```

```
head(sort(table(unb.adv_comunicacao$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)
```

```
##
```

```
##                               Comunicação
##                               9
##          PROGRAMA DE PÓS-GRADUAÇÃO
##                               1
## Programa de Pós-Graduação em Comunicação
##                               1

head(sort(table(unb.adv_comunicacao$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)

##
##                               Comunicação
##                               15
## Programa de Pós-Graduação em Comunicação
##                               2
##                               Mestrado
##                               1
##                               PPG/FAC/UnB
##                               1
```

Descrição dos dados de orientações de Direito

```
unb.adv_direito <- fromJSON("data/Direito.advise.json")
names(unb.adv_direito)

## [1] "ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO"
## [2] "ORIENTACAO_EM_ANDAMENTO_DOUTORADO"
## [3] "ORIENTACAO_EM_ANDAMENTO_MESTRADO"
## [4] "ORIENTACAO_EM_ANDAMENTO_GRADUACAO"
## [5] "ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA"
## [6] "ORIENTACAO_CONCLUIDA_POS_DOUTORADO"
## [7] "ORIENTACAO_CONCLUIDA_DOUTORADO"
## [8] "ORIENTACAO_CONCLUIDA_MESTRADO"
## [9] "OUTRAS_ORIENTACOES_CONCLUIDAS"

names(unb.adv_direito$ORIENTACAO_CONCLUIDA_DOUTORADO)

## [1] "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017"

length(unb.adv_direito$ORIENTACAO_CONCLUIDA_DOUTORADO$`2016`$natureza)

## [1] 23

head(sort(table(unb.adv_direito$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)

##
##                               Direito
##                               8
##          Direito, Estado e Constituição
##                               3
## Programa de Pós-Graduação Faculdade de Direito
##                               3
##    Doutorado em Direito, Estado e Constituição
##                               1

head(sort(table(unb.adv_direito$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)

##
```

```
##                               Direito
##                               34
##      Direito, Estado e Constituição
##                               3
##      Direitos Humanos e Cidadania
##                               3
##      Mestrado em Direito
##                               2
##      Programa de Pós-Graduação em Direito
##                               2
##      Antropologia
##                               1
##      Ciências da Saúde
##                               1
##      Design
##                               1
##      Programa de Pós-Graduação em Direito - UnB
##                               1
##      Transportes
##                               1
```

Descrição dos dados de produção bibliográfica de Ciência da Informação

```
unb.pub_ci <- fromJSON("data/CI.publication.json")
names(unb.pub_ci)
```

```
## [1] "PERIODICO"
## [2] "LIVRO"
## [3] "CAPITULO_DE_LIVRO"
## [4] "TEXTO_EM_JORNAIS"
## [5] "EVENTO"
## [6] "ARTIGO_ACEITO"
## [7] "DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA"
```

```
names(unb.pub_ci$PERIODICO$`2012`)
```

```
## [1] "natureza"      "titulo"        "periodico"
## [4] "ano"           "volume"        "issn"
## [7] "paginas"       "doi"           "autores"
## [10] "autores-endogeno"
```

```
head(sort(table(unb.pub_ci$PERIODICO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
##      REVISTA IBERO-AMERICANA DE CIÊNCIA DA INFORMAÇÃO
##                               6
##      Informação & Informação (Online)
##                               5
##      INFORMAÇÃO & SOCIEDADE (UFPB. ONLINE)
##                               5
##      PESQUISA BRASILEIRA EM CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA
##                               4
##      RECIIS - Revista Eletrônica de Comunicação, Informação & Inovação em Saúde
##                               4
```

```
## BIBLIOS (LIMA)
## 3
## BRAZILIAN JOURNAL OF INFORMATION SCIENCE
## 3
## Em Questão
## 2
## EM QUESTÃO
## 2
## Perspectivas em Ciência da Informação (on line)
## 2
```

```
head(sort(table(unb.pub_ci$LIVRO$`2015`$nome_da_editora), decreasing = TRUE), 10)
```

```
##
## Atlas
## 1
## Brasil. Ministério da Cultura.
## 1
## Cebraspe
## 1
## Editora Thesaurus
## 1
## IGI Global
## 1
## Novas Edições Acadêmicas/ OmniScriptum GmbH & Co. Kg
## 1
## OmniScriptum GmbH & Co. KG
## 1
## Thesaurus
## 1
```

Descrição dos dados de produção bibliográfica de Comunicação

```
unb.pub_comunicacao <- fromJSON("data/Comunicacao.publication.json")
names(unb.pub_comunicacao)
```

```
## [1] "PERIODICO"
## [2] "LIVRO"
## [3] "CAPITULO_DE_LIVRO"
## [4] "TEXTOS_EM_JORNAIS"
## [5] "EVENTO"
## [6] "ARTIGO_ACEITO"
## [7] "DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA"
```

```
names(unb.pub_comunicacao$PERIODICO$`2012`)
```

```
## [1] "natureza" "titulo" "periodico"
## [4] "ano" "volume" "issn"
## [7] "paginas" "doi" "autores"
## [10] "autores-endogeno"
```

```
head(sort(table(unb.pub_comunicacao$PERIODICO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
## REVISTA COMPOLITICA
```

```

##                                     2
##                                REVISTA COMUNICAÇÃO MIDIÁTICA (ONLINE)
##                                     2
##                                AMERIKA
##                                     1
##                                BRAZILIAN JOURNALISM RESEARCH (ONLINE)
##                                     1
##                                Communication Technologies et Developpement
##                                     1
##                                COMUNICAÇÃO & INOVAÇÃO (ONLINE)
##                                     1
##                                COMUNICAÇÃO & SOCIEDADE (ONLINE)
##                                     1
##                                COMUNICOLOGIA (BRASÍLIA)
##                                     1
##                                Contextos (Santiago)
##                                     1
## CONTRATEXTO - REVISTA DE LA FACULTAD DE COMUNICACIÓN DE LA UNIVERSIDAD DE LIMA
##                                     1
head(sort(table(unb.pub_comunicacao$LIVRO$`2015`$nome_da_editora), decreasing = TRUE), 10)

##
##                                Editora Atlas S.A.
##                                     2
##                                Insular
##                                     2
## Centro de Estudos da Mídia Alternativa ?Barão de Itararé
##                                     1
##                                EDISE
##                                     1
##                                Editora da UFBA
##                                     1
##                                Editora Kiron
##                                     1
##                                EDUFBA
##                                     1
##                                Friedrich Ebert Stiftung Brasil
##                                     1
##                                Hedra
##                                     1
##                                PALGRAVE MACMILLAN
##                                     1

```

Descrição dos dados de produção bibliográfica de Direito

```

unb.pub_direito <- fromJSON("data/Direito.publication.json")
names(unb.pub_direito)

## [1] "PERIODICO"
## [2] "LIVRO"
## [3] "CAPITULO_DE_LIVRO"
## [4] "TEXTO_EM_JORNAIS"
## [5] "EVENTO"

```

```
## [6] "ARTIGO_ACEITO"
## [7] "DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA"

names(unb.pub_direito$PERIODICO$`2012`)

## [1] "natureza"      "titulo"        "periodico"
## [4] "ano"           "volume"        "issn"
## [7] "paginas"       "doi"           "autores"
## [10] "autores-endogeno"

head(sort(table(unb.pub_direito$PERIODICO$`2017`$periodico), decreasing = TRUE), 10)

##
##          REVISTA BRASILEIRA DE CIÊNCIAS CRIMINAIS
##                                     4
##          CADERNOS ADENAUER (SÃO PAULO)
##                                     3
##          CADERNOS IBERO-AMERICANOS DE DIREITO SANITÁRIO
##                                     2
##          Developing World Bioethics
##                                     2
##          ICON-INTERNATIONAL JOURNAL OF CONSTITUTIONAL LAW
##                                     2
## Publicações da Escola da AGU: Direito, Gestão e Democracia
##                                     2
##          REVISTA DE DIREITO, ESTADO E TELECOMUNICAÇÕES (IMPRESSO)
##                                     2
##          REVISTA DE ESTUDOS CRIMINAIS
##                                     2
##          REVISTA DIREITO E PRÁXIS
##                                     2
##          REVISTA DO INSTITUTO HISTÓRICO E GEOGRAPHICO BRAZILEIRO
##                                     2

head(sort(table(unb.pub_direito$LIVRO$`2015`$nome_da_editora), decreasing = TRUE), 10)

##
## Laccademia Publishing          Lumen Juris          LTr
##          8          6          5
##          Arraes Editores Civilização Brasileira          CONPEDI
##          1          1          1
##          Editora Lumen Juris          Editora UnB          editorafi
##          1          1          1
##          Escolar Editora
##          1
```

CRISP-DM Fase.Atividade 2.3 - Análise exploratória dos dados

Como já informado, a análise exploratória dos dados possibilita um entendimento mais profundo da relação estatística existente entre os dados dos *datasets* para um melhor entendimento da qualidade daqueles dados para os objetivos do projeto.

Como já informado, a análise exploratória dos dados é responsabilidade parcial dos alunos, tendo em vista que este relatório apresenta uma análise exploratória preliminar. O relatório final deve conter análises exploratórias dos dados que sejam significativas e aprofundadas.

Arquivo Profile

Com esta função abaixo é possível visualizar a estrutura do arquivo de perfis dos 3 cursos e assim possibilitando um entendimento mais profundo da relação estatística existente entre os dados dos datasets para melhor entender a qualidade desses dados e alcançar o objetivo do projeto.

```
# Estrutura do arquivo de perfis de ciências da informação
names(unb.prof_ci$`1923828283438124`)
```

```
## [1] "nome"                "resumo_cv"
## [3] "areas_de_atuacao"    "endereco_profissional"
## [5] "producao_bibliografica" "orientacoes_academicas"
## [7] "senioridade"
```

```
# Estrutura do arquivo de perfis de comunicação
names(unb.prof_comunicacao$`0796700383761856`)
```

```
## [1] "nome"                "resumo_cv"
## [3] "areas_de_atuacao"    "endereco_profissional"
## [5] "producao_bibliografica" "orientacoes_academicas"
## [7] "senioridade"
```

```
# Estrutura do arquivo de perfis de direito
names(unb.prof_direito$`1540773562032795`)
```

```
## [1] "nome"                "resumo_cv"
## [3] "areas_de_atuacao"    "endereco_profissional"
## [5] "producao_bibliografica" "orientacoes_academicas"
## [7] "senioridade"
```

Arquivo Publicação

A estrutura dos arquivos de publicações é a mesma para os 3 cursos, com a função abaixo é possível visualizar um exemplo dessa estrutura para o arquivo de publicações do curso de Ciência da Informação.

```
# Estrutura do arquivo de perfis de ciências da informação
names(unb.pub_ci)
```

```
## [1] "PERIODICO"
## [2] "LIVRO"
## [3] "CAPITULO_DE_LIVRO"
## [4] "TEXTO_EM_JORNAIS"
## [5] "EVENTO"
## [6] "ARTIGO_ACEITO"
## [7] "DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA"
```

```
names(unb.pub_ci$PERIODICO)
```

```
## [1] "2010" "2011" "2012" "2013" "2014" "2015" "2016" "2017"
```

Arquivo Orientação

A estrutura dos arquivos de publicações é a mesma para os 3 cursos, com a função abaixo é possível visualizar um exemplo dessa estrutura para o arquivo de orientações do curso de Ciência da Informação.

```
names(unb.adv_ci)
```



```
## [1] "ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO"
## [2] "ORIENTACAO_EM_ANDAMENTO_DOUTORADO"
## [3] "ORIENTACAO_EM_ANDAMENTO_MESTRADO"
## [4] "ORIENTACAO_EM_ANDAMENTO_GRADUACAO"
## [5] "ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA"
## [6] "ORIENTACAO_CONCLUIDA_POS_DOUTORADO"
## [7] "ORIENTACAO_CONCLUIDA_DOUTORADO"
## [8] "ORIENTACAO_CONCLUIDA_MESTRADO"
## [9] "OUTRAS_ORIENTACOES_CONCLUIDAS"

names(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_MESTRADO$`2017`)

## [1] "natureza" "titulo"
## [3] "ano" "id_lattes_aluno"
## [5] "nome_aluno" "instituicao"
## [7] "curso" "codigo_do_curso"
## [9] "bolsa" "agencia_financiadora"
## [11] "codigo_agencia_financiadora" "nome_orientadores"
## [13] "id_lattes_orientadores"
```

CRISP-DM Fase.Atividade 2.4 - Verificação da qualidade dos dados.

Como já informado, a verificação da qualidade dos dados envolve responder se os dados disponíveis estão realmente completos.

Com relação à qualidade dos dados é possível concluir que os dados não têm uma estrutura boa e adequada para análises estatísticas, porém esses dados possuem informações relevantes sobre os programas de pós graduação dos 3 cursos. Além disso, esses dados possuem integridade sem duplicação, o que facilita o pré processamento desses dados.

CRISP-DM Fase 3 - Preparação dos Dados

Como já informado, na fase de **Preparação dos Dados** os *datasets* que serão utilizados em todo o trabalho são construídos a partir dos dados brutos. Aqui os dados são “filtrados” retirando-se partes que não interessam e selecionando-se os “campos” necessários para o trabalho de mineração.

São 5 as atividades genéricas nesta fase de preparação dos dados, a seguir divididas em subseções

CRISP-DM Fase.Atividade 3.1 - Seleção e Limpeza dos dados.

Como já informado, a seleção dos dados envolve identificar quais dados, da nossa “montanha de dados”, serão realmente utilizados.

```
# Limpeza dos dados de publicações
periodicos_ci <- tibble()
for (i in 1:length(unb.pub_ci$PERIODICO))
  periodicos_ci <- rbind(periodicos_ci, unb.pub_ci$PERIODICO[[i]])
livros_ci <- tibble()
for (i in 1:length(unb.pub_ci$LIVRO))
  livros_ci <- rbind(livros_ci, unb.pub_ci$LIVRO[[i]])
capitulo_de_livros_ci <- tibble()
for (i in 1:length(unb.pub_ci$CAPITULO_DE_LIVRO))
  capitulo_de_livros_ci <- rbind(capitulo_de_livros_ci,
```

```

        unb.pub_ci$CAPITULO_DE_LIVRO[[i]])
texto_em_jornais_ci <- tibble()
for (i in 1:length(unb.pub_ci$TEXTO_EM_JORNAIS))
  texto_em_jornais_ci <- rbind(texto_em_jornais_ci,
                               unb.pub_ci$TEXTO_EM_JORNAIS[[i]])
eventos_ci <- tibble()
for (i in 1:length(unb.pub_ci$EVENTO))
  eventos_ci <- rbind(eventos_ci, unb.pub_ci$EVENTO[[i]])
artigos_aceitos_ci <- tibble()
for (i in 1:length(unb.pub_ci$ARTIGO_ACEITO))
  artigos_aceitos_ci <- rbind(artigos_aceitos_ci, unb.pub_ci$ARTIGO_ACEITO[[i]])
demais_publicacoes_ci <- tibble()
for (i in 1:length(unb.pub_ci$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA))
  demais_publicacoes_ci <- rbind(demais_publicacoes_ci,
                                   unb.pub_ci$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA[[i]])

periodicos_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$PERIODICO))
  periodicos_comunicacao <- rbind(periodicos_comunicacao, unb.pub_comunicacao$PERIODICO[[i]])
livros_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$LIVRO))
  livros_comunicacao <- rbind(livros_comunicacao, unb.pub_comunicacao$LIVRO[[i]])
capitulo_de_livros_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$CAPITULO_DE_LIVRO))
  capitulo_de_livros_comunicacao <- rbind(capitulo_de_livros_comunicacao,
                                           unb.pub_comunicacao$CAPITULO_DE_LIVRO[[i]])
texto_em_jornais_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$TEXTO_EM_JORNAIS))
  texto_em_jornais_comunicacao <- rbind(texto_em_jornais_comunicacao,
                                           unb.pub_comunicacao$TEXTO_EM_JORNAIS[[i]])
eventos_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$EVENTO))
  eventos_comunicacao <- rbind(eventos_comunicacao, unb.pub_comunicacao$EVENTO[[i]])
artigos_aceitos_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$ARTIGO_ACEITO))
  artigos_aceitos_comunicacao <- rbind(artigos_aceitos_comunicacao, unb.pub_comunicacao$ARTIGO_ACEITO[[i]])
demais_publicacoes_comunicacao <- tibble()
for (i in 1:length(unb.pub_comunicacao$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA))
  demais_publicacoes_comunicacao <- rbind(demais_publicacoes_comunicacao,
                                           unb.pub_comunicacao$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA[[i]])

periodicos_direito <- tibble()
for (i in 1:length(unb.pub_direito$PERIODICO))
  periodicos_direito <- rbind(periodicos_direito, unb.pub_direito$PERIODICO[[i]])
livros_direito <- tibble()
for (i in 1:length(unb.pub_direito$LIVRO))
  livros_direito <- rbind(livros_direito, unb.pub_direito$LIVRO[[i]])
capitulo_de_livros_direito <- tibble()
for (i in 1:length(unb.pub_direito$CAPITULO_DE_LIVRO))

```

```

capitulo_de_livros_direito <- rbind(capitulo_de_livros_direito,
                                   unb.pub_direito$CAPITULO_DE_LIVRO[[i]])
texto_em_jornais_direito <- tibble()
for (i in 1:length(unb.pub_direito$TEXTO_EM_JORNAIS))
  texto_em_jornais_direito <- rbind(texto_em_jornais_direito,
                                   unb.pub_direito$TEXTO_EM_JORNAIS[[i]])
eventos_direito <- tibble()
for (i in 1:length(unb.pub_direito$EVENTO))
  eventos_direito <- rbind(eventos_direito, unb.pub_direito$EVENTO[[i]])
artigos_aceitos_direito <- tibble()
for (i in 1:length(unb.pub_direito$ARTIGO_ACEITO))
  artigos_aceitos_direito <- rbind(artigos_aceitos_direito, unb.pub_direito$ARTIGO_ACEITO[[i]])
demais_publicacoes_direito <- tibble()
for (i in 1:length(unb.pub_direito$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA))
  demais_publicacoes_direito <- rbind(demais_publicacoes_direito,
                                       unb.pub_direito$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA[[i]])

# Limpeza dos dados de orientações
nao_concluido_ci <- data.frame()
for (i in 1:length(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO))
  nao_concluido_ci <- rbind(nao_concluido_ci, unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_DE_POS_DOUTORADO[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_DOUTORADO))
  nao_concluido_ci <- rbind(nao_concluido_ci, unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_DOUTORADO[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_MESTRADO))
  nao_concluido_ci <- rbind(nao_concluido_ci, unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_MESTRADO[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_GRADUACAO))
  nao_concluido_ci <- rbind(nao_concluido_ci, unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_GRADUACAO[[i]])

concluido_ci <- data.frame()

for (i in 1:length(unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA))
  concluido_ci <- rbind(concluido_ci, unb.adv_ci$ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))
  concluido_ci <- rbind(concluido_ci, unb.adv_ci$ORIENTACAO_CONCLUIDA_POS_DOUTORADO[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_CONCLUIDA_DOUTORADO))
  concluido_ci <- rbind(concluido_ci, unb.adv_ci$ORIENTACAO_CONCLUIDA_DOUTORADO[[i]])

for (i in 1:length(unb.adv_ci$ORIENTACAO_CONCLUIDA_MESTRADO))
  concluido_ci <- rbind(concluido_ci, unb.adv_ci$ORIENTACAO_CONCLUIDA_MESTRADO[[i]])

```

CRISP-DM Fase 4 - Resultados

Nessa seção são apresentados os resultados obtidos nessa etapa do projeto assim como uma explicação mais profunda e demonstrativa dos dados utilizados.

A primeira observação que podemos fazer é que os professores de direito tiveram apenas 4 artigos aceitos como mostrado a seguir.

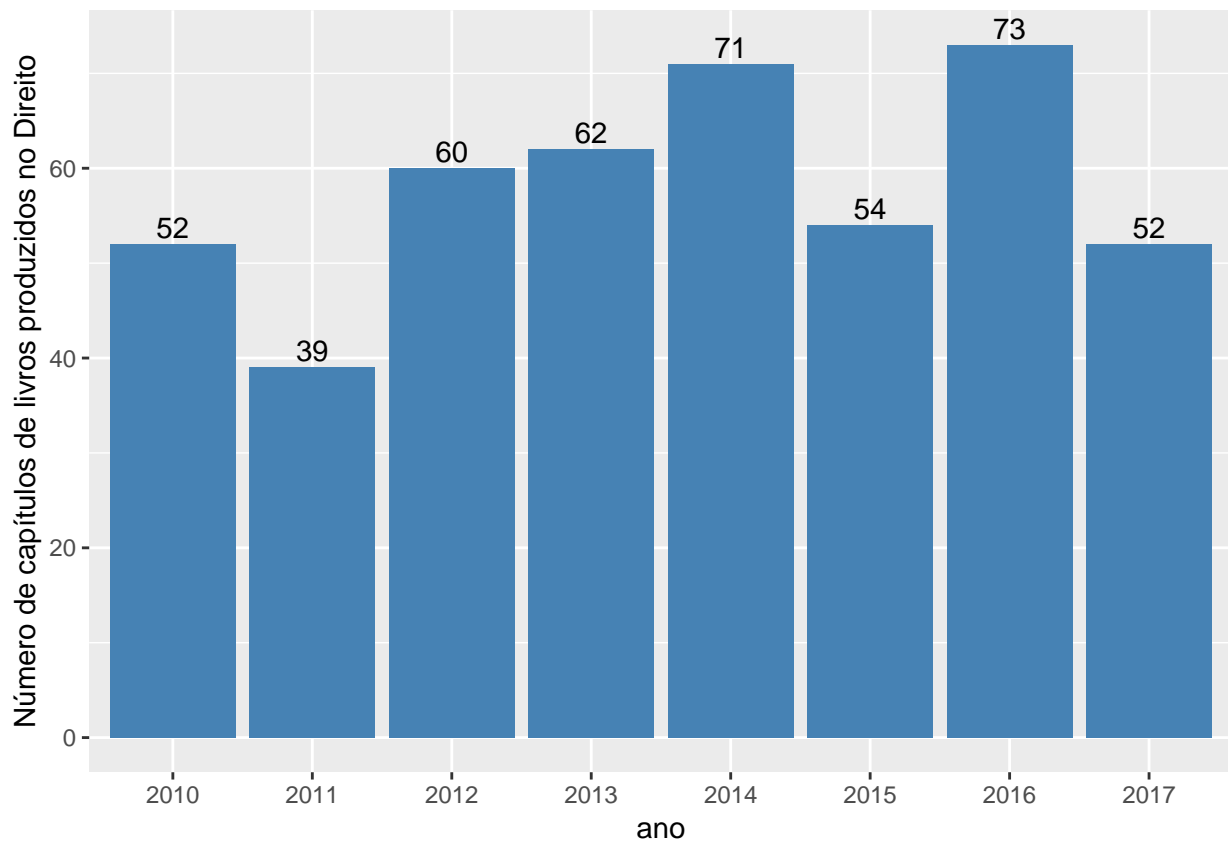
```
subset(artigos_aceitos_direito, select=c("ano", "titulo", "periodico")) %>%
  kable() %>%
  kable_styling() %>%
  column_spec(2:3, width = "17em")
```

ano	titulo	periodico
2013	A argumentação do Supremo Tribunal Federal e do Superior Tribunal de Justiça acerca da periculosidade de agentes inimputáveis e semi-imputáveis	Novos Estudos Jurídicos (UNIVALI) (Cessou em 2007. Cont. ISSN 2175-0491 Novos Estudos Jurídicos (Online))
2015	Os limites dos termos bem público mundial, patrimônio comum da humanidade e bens comuns para delimitar as obrigações de preservação dos recursos marinhos	Revista de Direito Internacional
2015	Perspectives on human dignity (on judicial rulings regarding contemporary slavery in Brazil)	Quaderni Fiorentini per la Storia del Pensiero Giuridico Moderno
2016	Some New Ideas on the Role of Legal Analysis Applied to the Regulation of Telecommunications Services in Brazil	Revista de Direito, Estado e Telecomunicações (Impresso)

Isso se deve a produção científica do Direito ser mais focada em publicações e principalmente e na produção de capítulos de livros, se diferenciando da produção em outros cursos das áreas de Engenharia e Infomática que publicam muitos artigos.

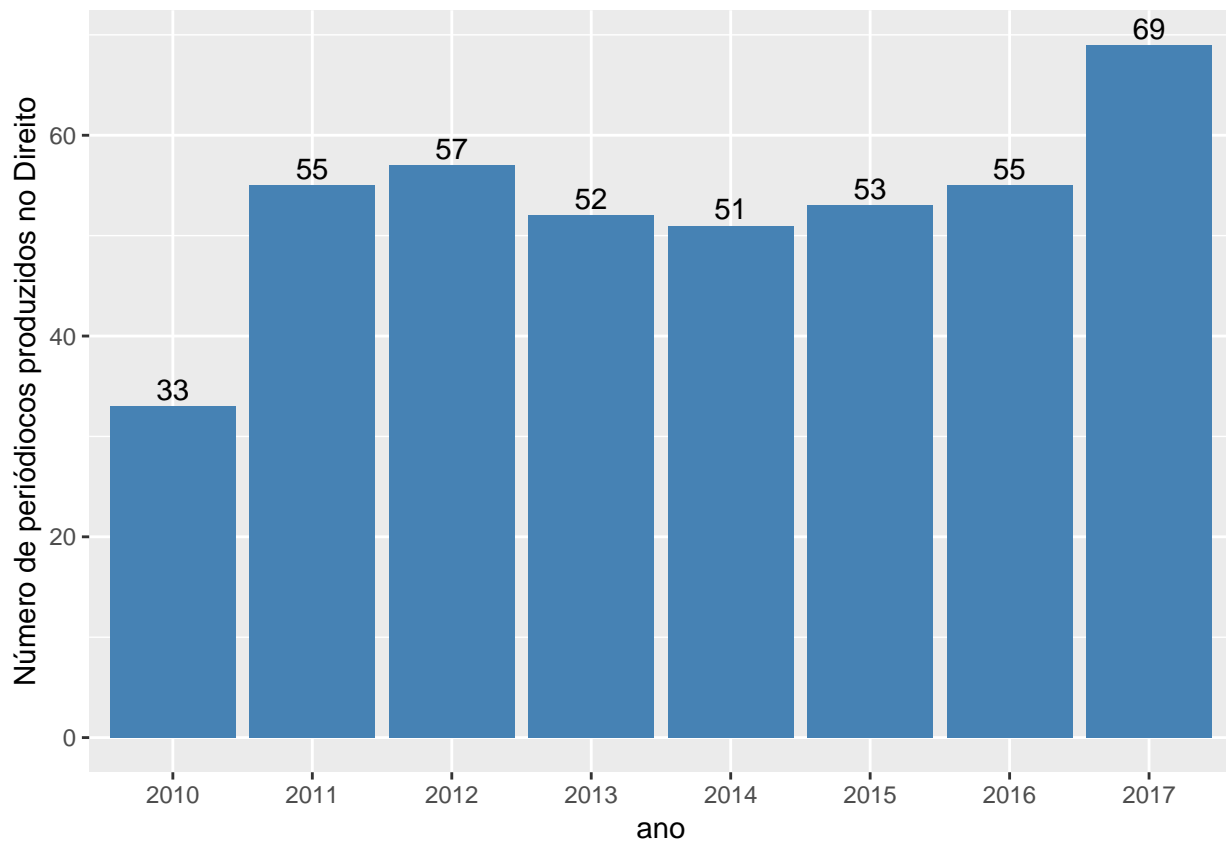
Agora focando mais nos capítulos de livros produzidos, a primeira análise inicial que temos é na quantidade de capítulos de livros produzidos por ano entre 2010 e 2017.

```
capitulo_de_livros_direito %>%
  group_by(ano) %>%
  summarise(n = n()) %>%
  ggplot(aes(x = ano, y = n)) +
  geom_bar(position = "stack", stat = "identity", fill = "steelblue") +
  geom_text(aes(label=n), vjust=-0.3) +
  ylab("Número de capítulos de livros produzidos no Direito")
```



O curso de direito também produz muitas publicações em periódicos, então realizando a mesma análise por ano como feita com os capítulos de livros, para podermos comparar os dois tipos de produção, temos o seguinte resultado.

```
periodicos_direito %>%  
  group_by(ano) %>%  
  summarise(n = n()) %>%  
  ggplot(aes(x = ano, y = n)) +  
  geom_bar(position = "stack", stat = "identity", fill = "steelblue") +  
  geom_text(aes(label=n), vjust=-0.3) +  
  ylab("Número de periódicos produzidos no Direito")
```



Por esse gráfico podemos observar que a produção anual de periódicos e capítulos de livros são bem próximas, com exceção apenas do ano de 2010 e 2011, quando a produção de periódicos e capítulos de livros foram bem menores que nos outros anos.

Voltando nossa análise para os capítulos de livros produzidos, podemos observar o nível de internacionalização dessas produções, isto é, quantos livros foram publicados em cada país.

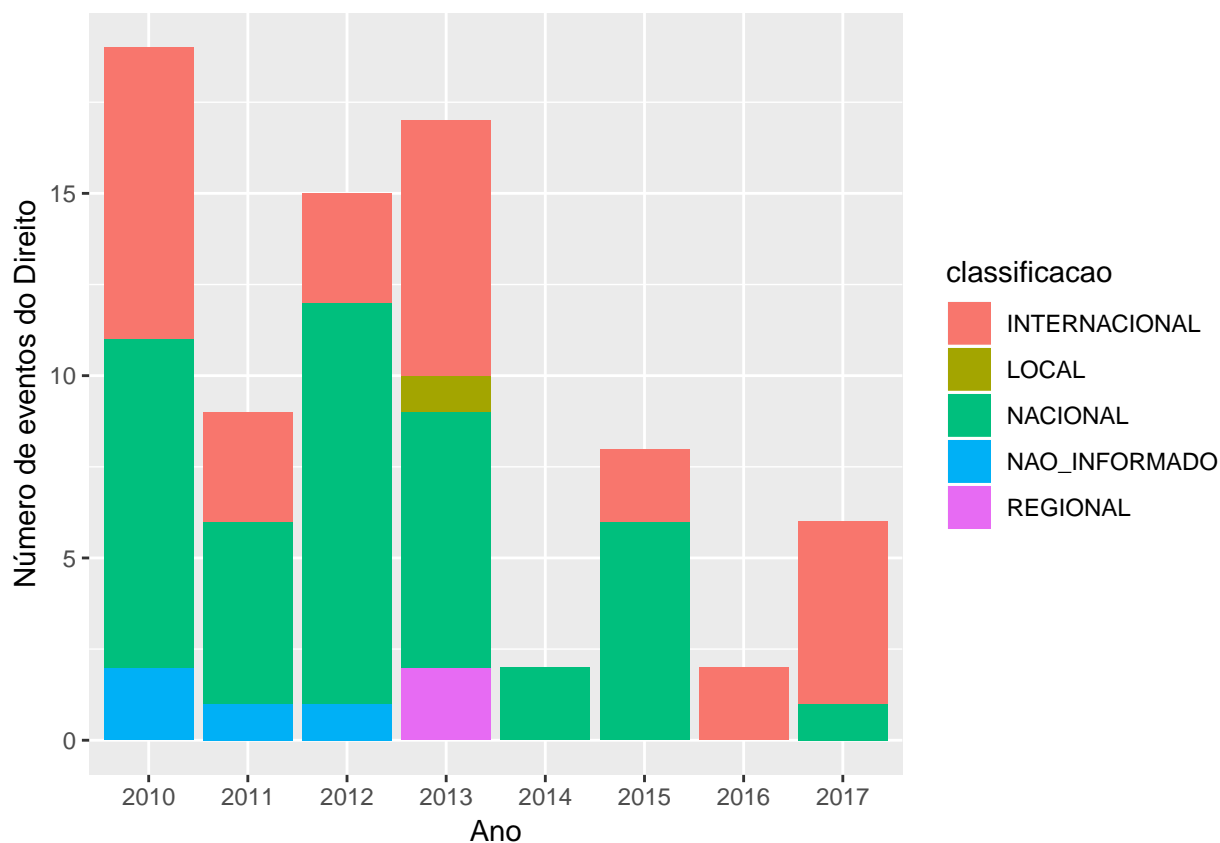
```
subset(capitulo_de_livros_direito, select=c("pais_de_publicacao", "ano")) %>%  
  group_by(pais_de_publicacao) %>%  
  summarise(n = n()) %>%  
  data.frame %>%  
  arrange(desc(n)) %>%  
  kable(col.names=c("País de publicação", "Número de capítulos de livro")) %>%  
  kable_styling()
```

País de publicação	Número de capítulos de livro
Brasil	444
Bélgica	3
Alemanha	2
Grã-Bretanha	2
Holanda	2
Inglaterra	2
México	2
Argentina	1
Espanha	1
Estados Unidos	1
França	1
Portugal	1
Sérvia	1

Podemos observar que temos poucas publicações internacionais do curso de Direito. Após pesquisar sobre o assunto e conversar com alguns alunos do curso de Direito para tentar entender porque isso ocorre, o principal motivo encontrado é que produções como capítulos de livro e publicações normalmente estão muito atrelados ao sistema Jurídico de um país específico.

Continuando a análise do nível de internacionalização, mas agora analisando as produções que foram para eventos temos o seguinte resultado.

```
ggplot(data = eventos_direito) +
  geom_bar(mapping = aes(x = ano_do_trabalho, fill = classificacao)) +
  xlab("Ano") +
  ylab("Número de eventos do Direito")
```



Nessa análise podemos observar que diferentemente das publicações internacionais de capítulos de livros, uma quantidade considerável das participações em eventos de professores do Direito são em eventos internacionais.

Agora são feitas as análises estatísticas do curso de Comunicação.

Na análise abaixo, foi feito um gráfico com o quantitativo de orientandos atuais e que já finalizaram. Para isso, foi feita uma contagem do número de orientações concluídas de cada tipo (mestrado, doutorado etc) e de orientandos que ainda não concluíram, isto é, que são atuais.

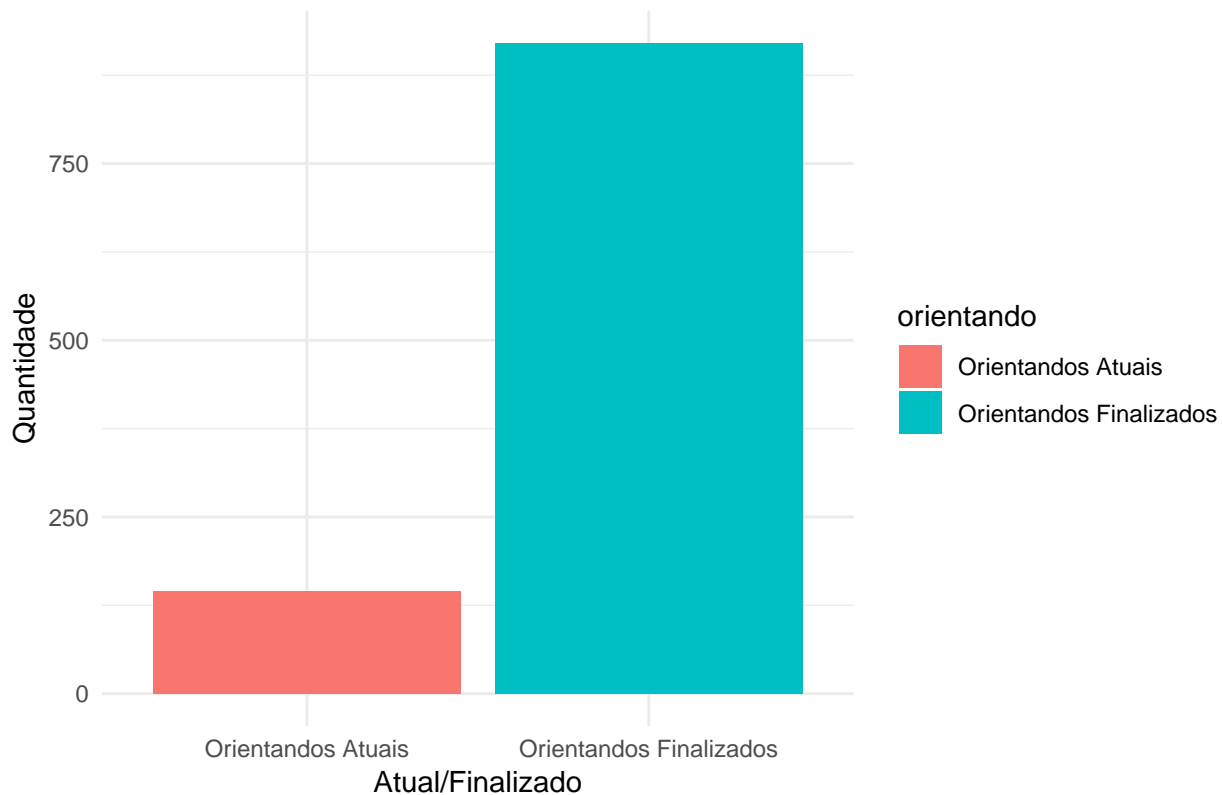
```
conc_doutorado <- sum(unlist(sapply(unb.prof_comunicacao,
  function(x)nrow(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO))))
conc_mestrado <- sum(unlist(sapply(unb.prof_comunicacao,
  function(x)nrow(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO))))
conc_pos_doc <- sum(unlist(sapply(unb.prof_comunicacao,
  function(x)nrow(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))))
conc_outras <- sum(unlist(sapply(unb.prof_comunicacao,
  function(x)nrow(x$orientacoes_academicas$OUTRAS_ORIENTACOES_CONCLUIDAS))))
orientandos_concluidos <- conc_doutorado + conc_mestrado + conc_pos_doc + conc_outras

atual_cientifica <- sum(unlist(sapply(unb.prof_comunicacao, function(x)
  nrow(x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA))))
atual_doutorado <- sum(unlist(sapply(unb.prof_comunicacao, function(x)
  nrow(x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_DOUTORADO))))
atual_graduacao <- sum(unlist(sapply(unb.prof_comunicacao, function(x)
  nrow(x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_GRADUACAO))))
atual_mestrado <- sum(unlist(sapply(unb.prof_comunicacao, function(x)
  nrow(x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_MESTRADO))))
orientandos_atuais <- atual_cientifica + atual_doutorado + atual_graduacao +
  atual_mestrado

data <- data.frame(quantidade=c(orientandos_concluidos, orientandos_atuais),
  orientando=c("Orientandos Finalizados", "Orientandos Atuais"))

ggplot(data, aes(x = orientando,y=quantidade,fill = orientando)) +
  xlab("Atual/Finalizado") +
  ylab("Quantidade") +
  labs(title = "Quantidade de orientandos de Comunicação que já finalizaram e atuais") +
  geom_bar(stat="identity")+theme_minimal()
```


Quantidade de orientandos de Comunicação que já finalizaram e atuais



```
orientandos_atuais + orientandos_concluidos + length(unb.prof_comunicacao)
```

```
## [1] 1096
```

Percebe-se, pelo gráfico, que o número de orientandos que já finalizaram é bem superior aos atuais. O que leva a uma conclusão realista pois, há dados de vários anos anteriores, logo, a tendência é realmente ter mais orientandos concluídos do que atuais.

Para medir o nível de internacionalização, foi necessário analisar a produção bibliográfica de cada professor, mais especificamente a classificação do evento que eles tiveram essa produção bibliográfica, que pode variar entre local, regional, nacional e internacional.

```
internacional <- sapply(unb.prof_comunicacao,
  function(x) x$producao_bibliografica$EVENTO$classificacao == "INTERNACIONAL")
num_internacional <- 0
for (i in 1:length(internacional))
  num_internacional <- num_internacional + sum(internacional[[i]], na.rm=TRUE)

nacional <- sapply(unb.prof_comunicacao,
  function(x) x$producao_bibliografica$EVENTO$classificacao == "NACIONAL")
num_nacional <- 0
for (i in 1:length(nacional))
  num_nacional <- num_nacional + sum(nacional[[i]], na.rm=TRUE)

regional <- sapply(unb.prof_comunicacao,
  function(x) x$producao_bibliografica$EVENTO$classificacao == "REGIONAL")
num_regional <- 0
for (i in 1:length(regional))
  num_regional <- num_regional + sum(regional[[i]], na.rm=TRUE)
```

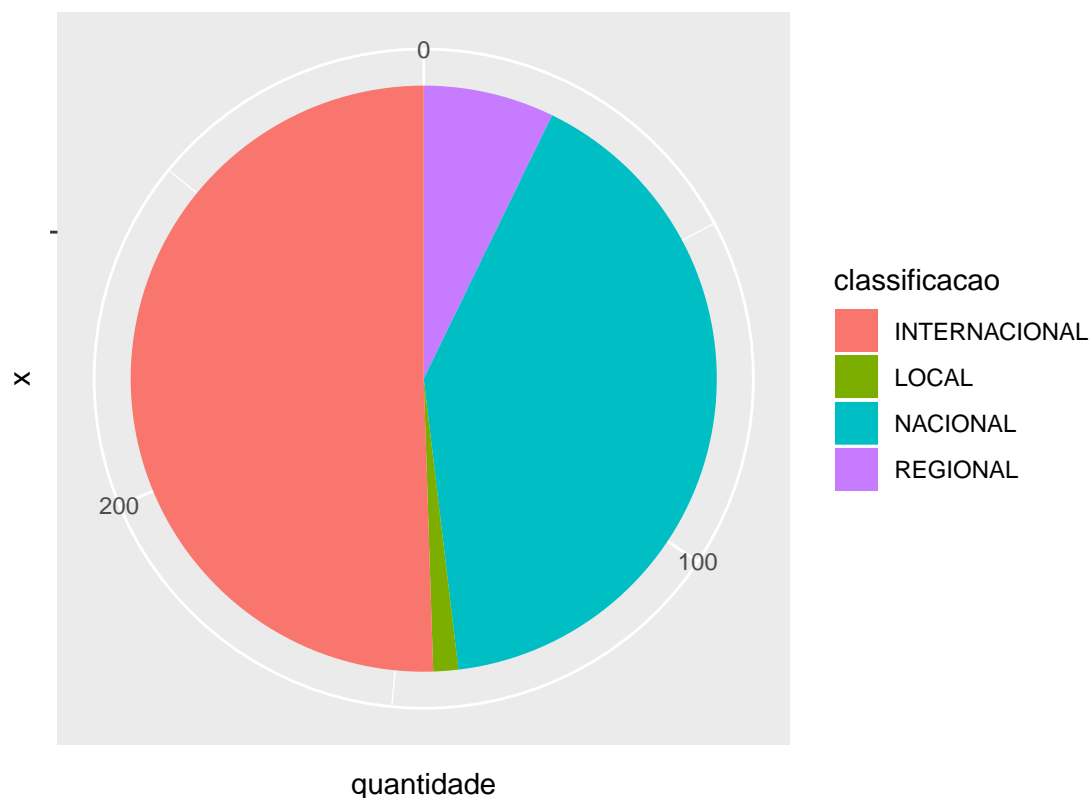
```

local <- sapply(unb.prof_comunicacao,
  function(x) x$producao_bibliografica$EVENTO$classificacao == "LOCAL")
num_local <- 0
for (i in 1:length(local))
  num_local <- num_local + sum(local[[i]], na.rm=TRUE)

df <- data.frame(classificacao=c("INTERNACIONAL", "NACIONAL", "REGIONAL", "LOCAL"),
  quantidade=c(num_internacional, num_nacional, num_regional, num_local))
bp<- ggplot(df, aes(x="", y=quantidade, fill=classificacao)) +
  labs(title = "Internacionalização do curso de Comunicação") +
  geom_bar(width = 1, stat = "identity")
pie <- bp + coord_polar("y", start=0)
pie

```

Internacionalização do curso de Comunicação



A partir desse gráfico de pizza gerado é possível concluir que há um bom nível de internacionalização do ramo acadêmico do curso de Comunicação da Universidade de Brasília.

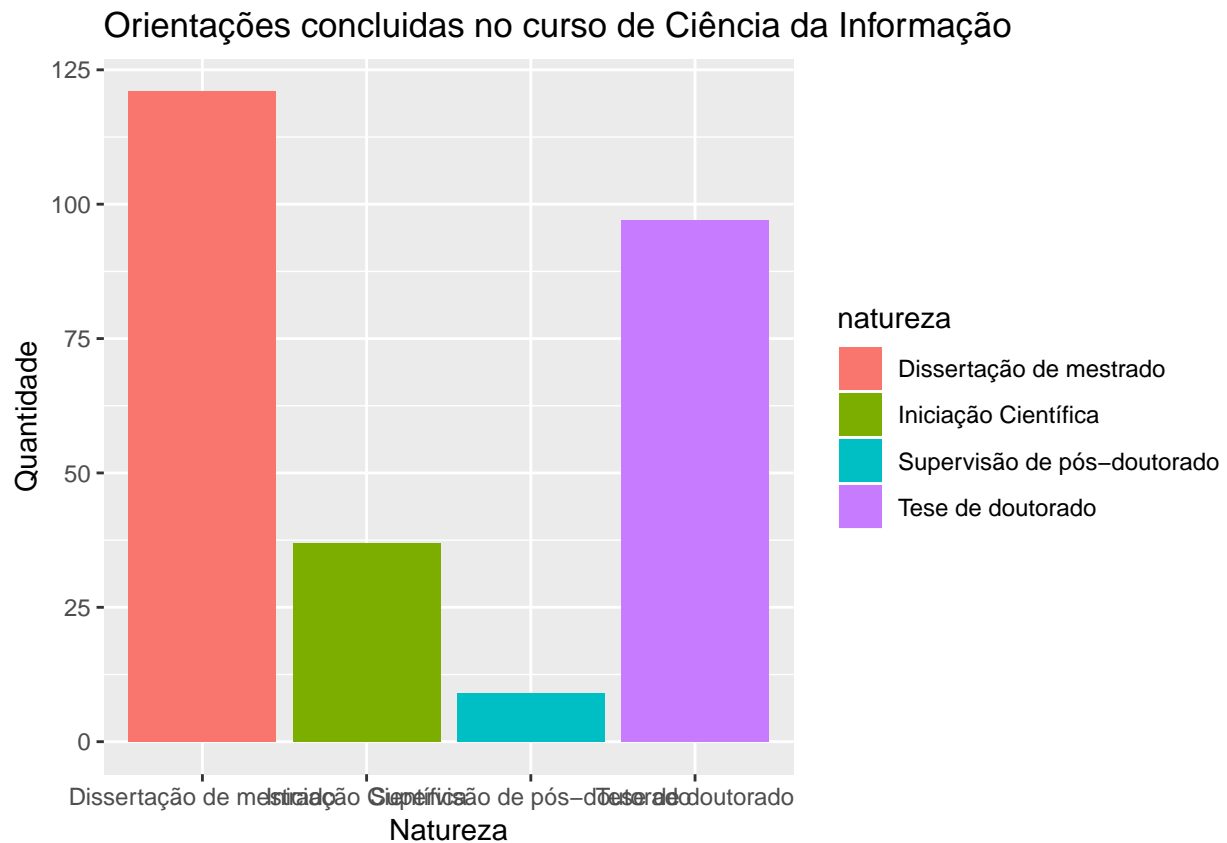
Neste momento são feitas as análises estatísticas do curso de Ciência da Informação.

Como pode-se observar na imagem abaixo, ao pegar os dados das orientações concluídas, a maior quantidade de orientações concluídas veio dos trabalho de conclusão de curso de graduação e em segundo lugar as orientações de outras naturezas

```

ggplot(concluido_ci, aes(x = natureza, fill = natureza)) +
  geom_bar() +
  xlab("Natureza") +
  ylab("Quantidade") +
  labs(title = "Orientações concluídas no curso de Ciência da Informação")

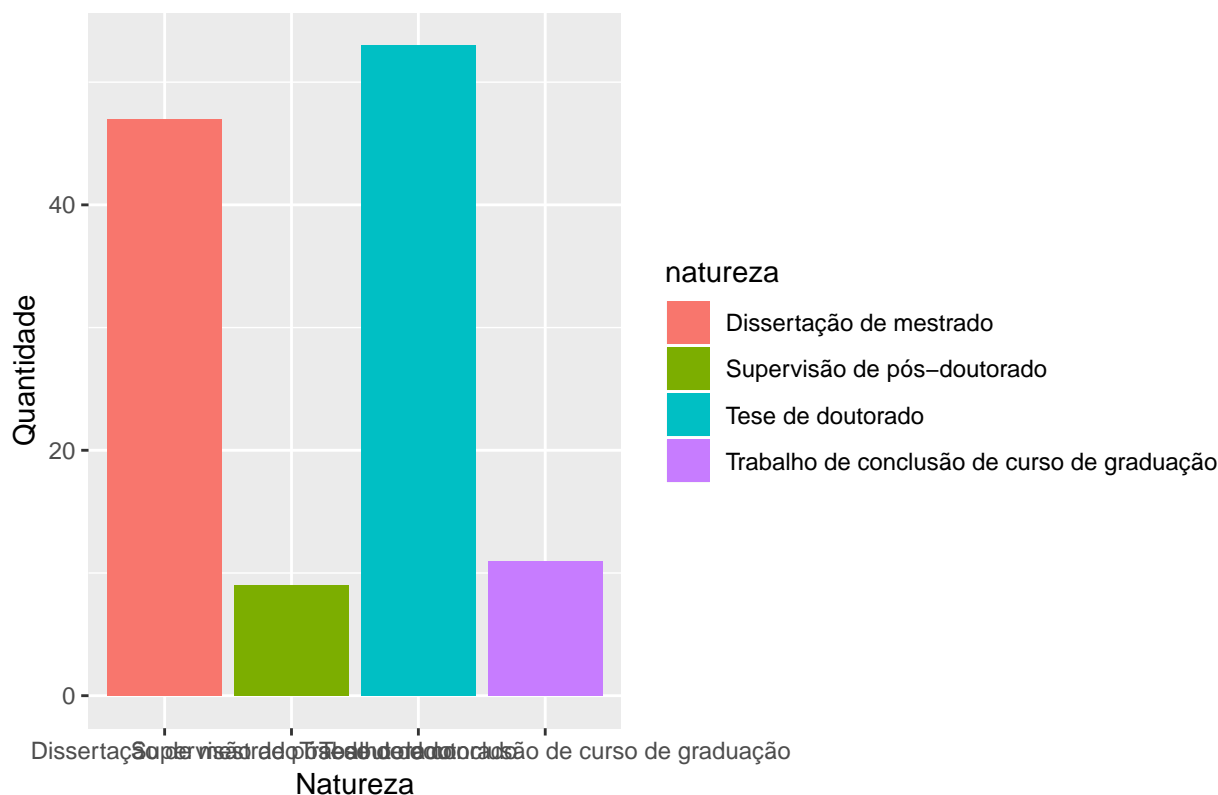
```



Utilizando-se uma função semelhante à utilizada acima, obteve-se o gráfico abaixo que mostra que as teses de doutorado estão em maior número, sendo seguidas pelas dissertações de mestrado, mostrando que quanto maior o aprofundamento da natureza de uma especialização, maior a chance de não se concluir a dissertação, como pode ser visto abaixo.

```
ggplot(nao_concluido_ci, aes(x = natureza, fill = natureza)) +
  geom_bar() +
  xlab("Natureza") +
  ylab("Quantidade") +
  labs(title = "Orientações não concluídas no curso de Ciência da Informação")
```

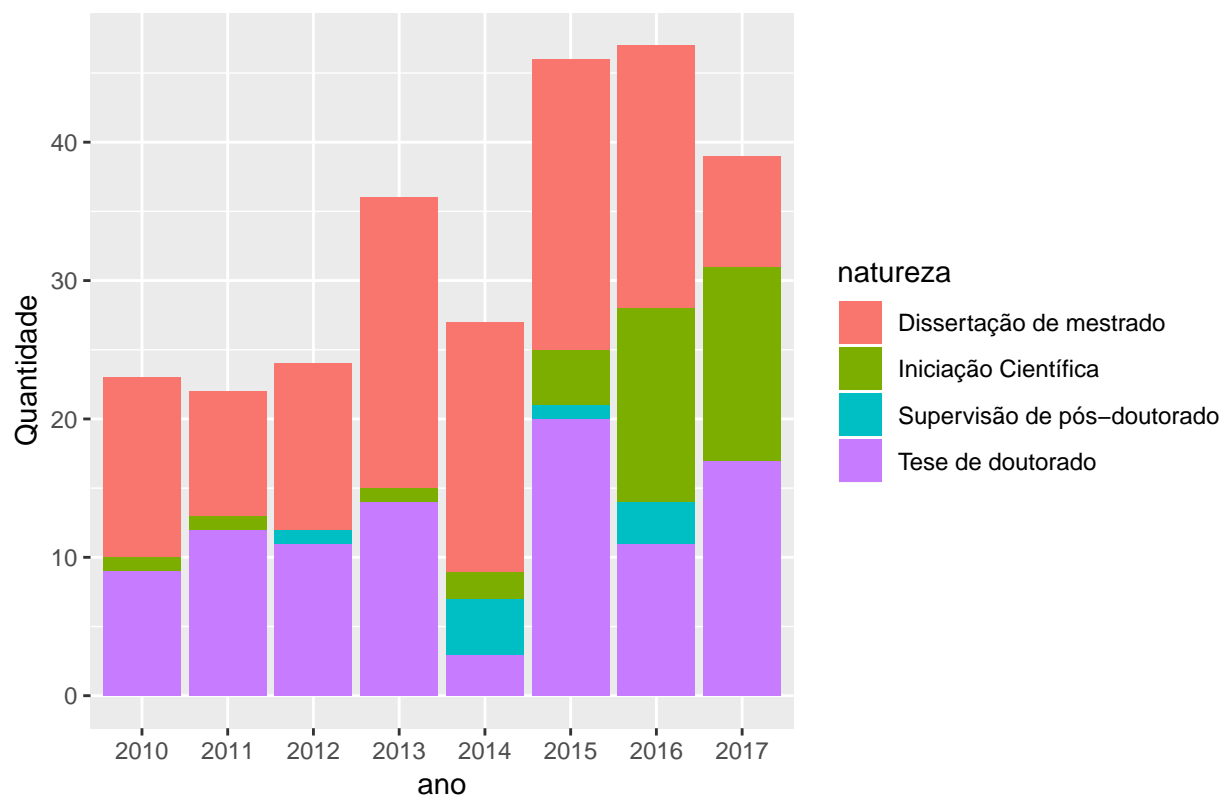
Orientações não concluídas no curso de Ciência da Informação



Após verificar a quantidade total de orientações, baseando-se na natureza dessas orientações, foi analisado essas quantidades em cada ano compreendido entre o período de 2010 a 2017. Ao fazer o confronto, na área de orientações concluídas, entre o ano x quantidade x natureza. Obteve-se um gráfico mostrando que a conclusão das dissertações de mestrado foi bastante expressiva ao longo dos anos. Quanto às orientações não concluídas, percebe-se que na maioria dos anos, poucos trabalhos de conclusão do curso de graduação não foram finalizados. Isso tudo pode-se ser observado utilizando-se a função `ggplot()` da seguinte maneira:

```
ggplot(concluido_ci, aes(x = ano, fill = natureza)) +
  geom_bar() +
  ylab("Quantidade") +
  labs(title = "Orientações concluídas por ano no curso de Ciência da Informação")
```

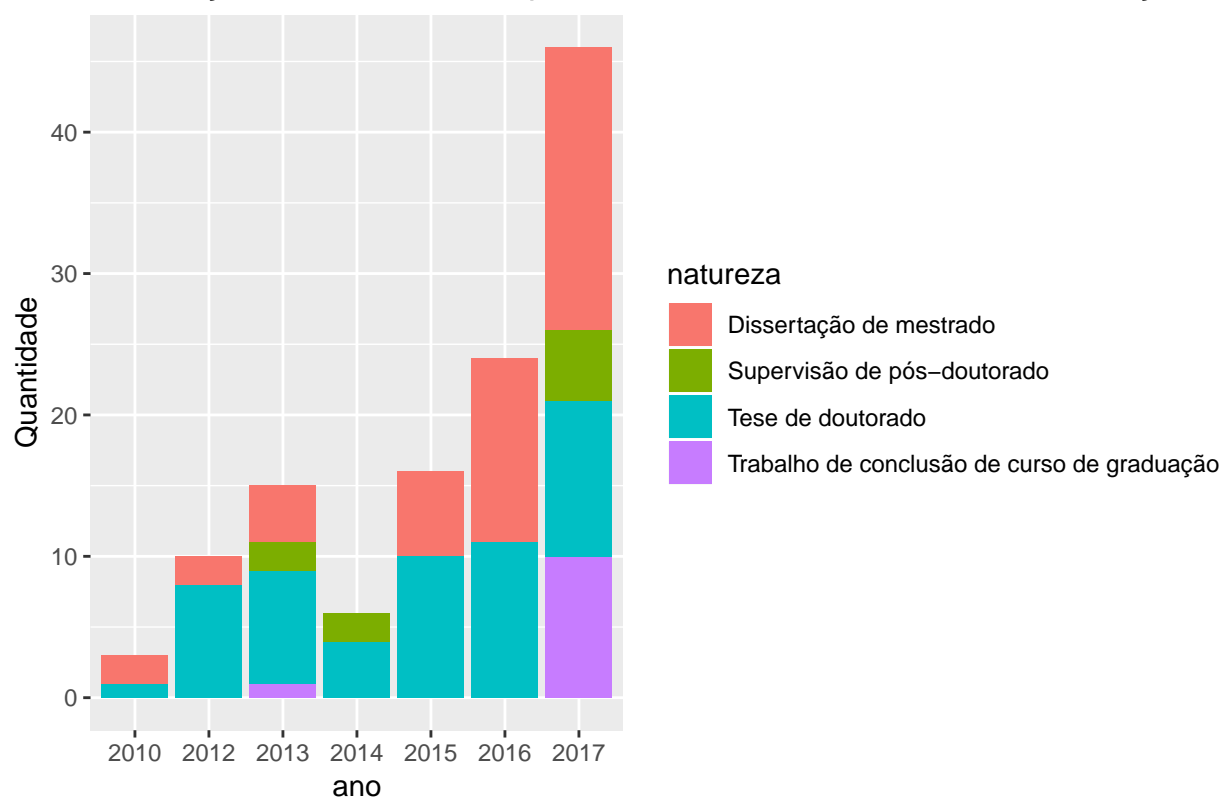
Orientações concluídas por ano no curso de Ciência da Informação



A mesma função foi utilizada para as Orientações não concluídas, produzindo os dois gráficos abaixo, que confirmam a análise feita acima:

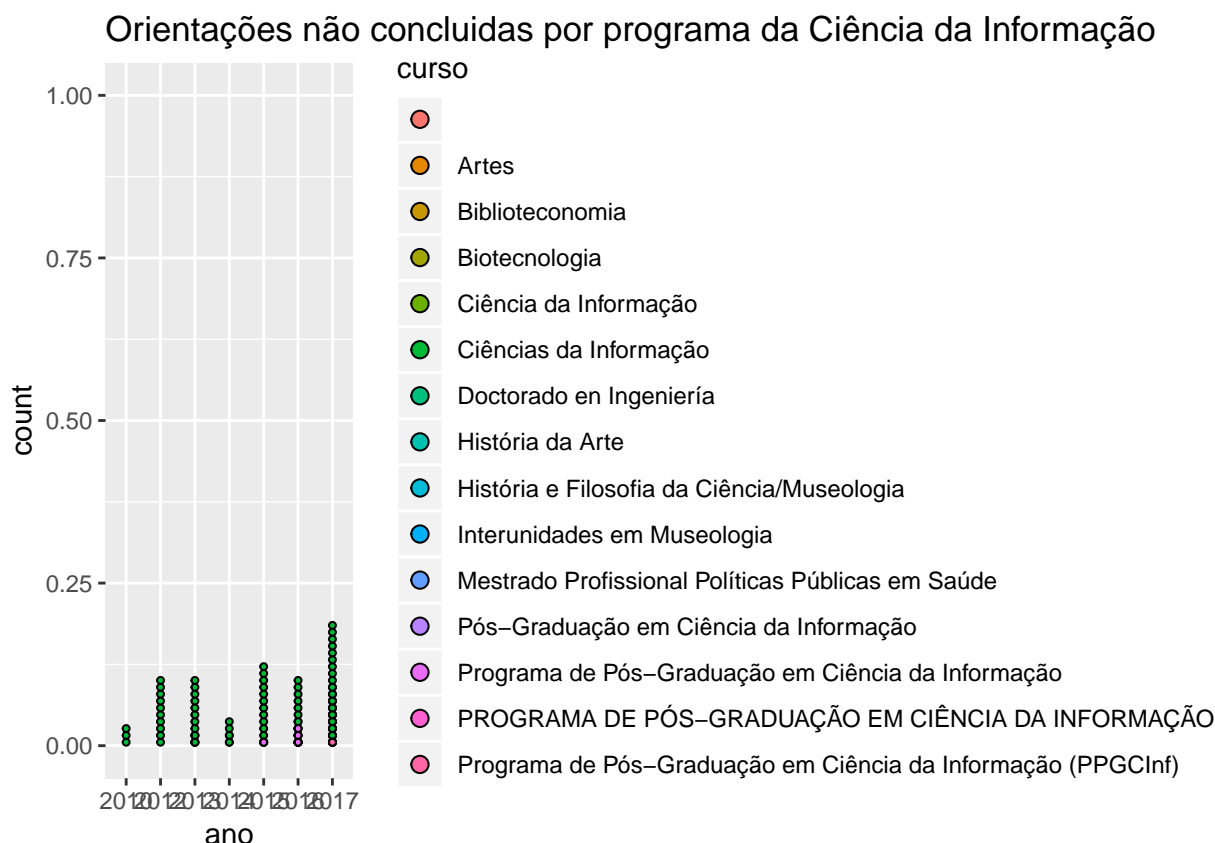
```
ggplot(nao_concluido_ci, aes(x = ano, fill = natureza)) +  
  geom_bar() +  
  ylab("Quantidade") +  
  labs(title = "Orientações não concluídas por ano no curso de Ciência da Informação")
```

Orientações não concluídas por ano no curso de Ciência da Informação



Além desse dados obtidos, foi verificado quais e quantos cursos não concluíram as suas orientações. Isso foi feito, utilizando a função `ggplot()` da seguinte maneira:

```
ggplot(nao_concluido_ci, aes(x = ano, fill = curso)) +  
  geom_dotplot() +  
  labs(title = "Orientações não concluídas por programa da Ciência da Informação")
```



CRISP-DM Fase 5 - Avaliação

Portanto, a análise desses dados permitiu ter uma base de entendimento interessante sobre os currículos de professores da UnB dos cursos de Ciência da Informação, Direito e Comunicação. Além de descobrir informações sobre orientações em vários âmbitos acadêmicos, de publicações em periódicos, livros, artigos, textos em jornais, de linhas de pesquisa e como esses dados são armazenados e como podem ser relacionados entre si.

Referências

- Can, Fazli, Tansel Özyer, e Faruk Polat, orgs. State of the Art Applications of Social Network Analysis. Lecture Notes in Social Networks. Switzerland: Springer International Publishing, 2014.
- CAPES. “Documentos de Área”. CAPES.gov.br. Acessado 12 de junho de 2018. <http://avaliacaoquadrienal.capes.gov.br/documentos-de-area>.
- . “Plano Nacional de Pós-Graduação - PNPG 2011/2020 Vol. 1”. Brasília - DF, dezembro de 2010. <http://www.capes.gov.br/images/stories/download/Livros-PNPG-Volume-I-Mont.pdf>.
- . “Plano Nacional de Pós-Graduação - PNPG 2011/2020 Vol. 2”. Brasília - DF, dezembro de 2010. http://www.capes.gov.br/images/stories/download/PNPG_Miolo_V2.pdf.
- . “Sucupira: coleta de dados, docentes de pós-graduação stricto sensu no Brasil 2015”. CAPES - Banco de Metadados, 16 de março de 2016. <http://metadados.capes.gov.br/index.php/catalog/63>.
- Chapman, Pete, Julian Clinton, Randy Kerber, Thomas Khabaza, Thomas Reinartz, Colin Shearer, e Rüdiger Wirth. “CRISP-DM 1.0: Step-by-Step Data Mining Guide”. USA: CRISP-DM Consortium, 2000. <https://www.the-modeling-agency.com/crisp-dm.pdf>.

- Datacamp. “Machine Learning with R (Skill Track)”. Datacamp, 2018. <https://www.datacamp.com/tracks/machine-learning>.
- Fernandes, Jorge H C, e Ricardo Barros Sampaio. “DataScienceForAll”. Zotero, 13 de junho de 2018. <https://www.zotero.org/groups/2197167/datascienceforall>.
- ———. “Especificação do Trabalho Final da Disciplina de Ciência de Dados para Todos 2017.2: Estudo sobre a visibilidade internacional da produção científica das pós-graduações vinculadas às áreas de conhecimento da CAPES, na Universidade de Brasília (Comunicação Interna)”. Disciplina 116297 - Tópicos Avançados em Computadores, turma D, do semestre 2017.2, do Departamento de Ciência da Computação do Instituto de Ciências Exatas da Universidade de Brasília, 28 de novembro de 2017. https://aprender.ead.unb.br/pluginfile.php/474549/mod_resource/content/1/Estudo%20da%20Cie%CC%82ncia.pdf.
- Fernandes, Jorge H C, Ricardo Barros Sampaio, e João Ribas de Moura. “Ciência de Dados para Todos (Data Science For All) - 2018.1 - Análise da Produção Científica e Acadêmica da Universidade de Brasília - Modelo de Relatório Final da Disciplina - Departamento de Ciência da Computação da UnB”. Disciplina 116297 - Tópicos Avançados em Computadores, turma D, do semestre 2018.1, do Departamento de Ciência da Computação do Instituto de Ciências Exatas da Universidade de Brasília, 13 de junho de 2018.
- Frickel, Scott, e Kelly Moore. *The New Political Sociology of Science: Institutions, Networks, and Power*. Science and technology in society. USA: The University of Wisconsin Press, 2006.
- Graduate Prospects Ltd. “Job profile: Higher education lecturer”, 2018. <https://www.prospects.ac.uk/job-profiles/higher-education-lecturer>.
- Kalpazidou Schmidt, Evanthia, e Ebbe Krogh Graversen. “Persistent factors facilitating excellence in research environments”. *Higher Education* 75, n^o 2 (1^o de fevereiro de 2018): 341–63. <https://doi.org/10.1007/s10734-017-0142-0>.
- Kilduff, Martin, e Wenpin Tsai. *Social Networks and Organizations*. UK: Sage Publications, 2003.
- Kolaczyk, Eric D., e Gábor Csárdi. *Statistical Analysis of Network Data with R*. USA: Springer, 2014.
- Kuhn, Max, Jed Wing, Steve Weston, Andre Williams, Chris Keefer, Allan Engelhardt, Tony Cooper, et al. “Package ‘Caret’ - Classification and Regression Training”, 27 de maio de 2018. <https://cran.r-project.org/web/packages/caret/caret.pdf>.