

Ciência de Dados para Todos (Data Science For All) - 2018.2 - Análise da Produção Científica e Acadêmica da Universidade de Brasília

Engenharia Biomédica, Engenharia de Sistemas Eletrônicos e de Automação, Engenharia Elétrica P7 e Engenharia Elétrica P8.

Marcos Vinicius Prescendo Tonin, Lucas Nascimento Santos Souza e Carlos Aragão

02/12/2018

I - Introdução

O presente trabalho tem por finalidade entender, correlacionar e analisar os dados da produção acadêmica e científica dos seguintes cursos de pós-graduação: Engenharia Biomédica, Engenharia Elétrica em seus dois ramos (53001010059P8 e 53001010080P7) e Engenharia de Sistemas Eletrônicos e de Automação.

II - Metodologia

Para um melhor resultado do trabalho buscou-se seguir e adaptar-se a metodologia CRISP-DM, para isso baseou-se no ciclo de projeto da metodologia CRISP-DM.

O ciclo é definido pelas seguintes fases:

- **Entendimento do negócio:** Primeiramente deve-se entender o que se busca encontrar, haja visto que não faz sentido fazer uma análise de dados sem saber o que se busca.
- **Entendimento dos dados:** Busca-se entender os dados de forma mais superficial inicialmente.
- **Preparação de dados:** Então os dados são tratados para que se possa facilitar o processamento dos mesmos.
- **Modelagem:** Realiza a modelagem do dado.
- **Avaliação:** Avalia-se o resultado da modelagem.
- **Implantação:** Utilização dos dados.

Para melhor visualização das fases e a comunicação entre elas, deve-se observar a Figura 1. Por esta figura consegue-se perceber que, por exemplo, o entendimento do negócio e do dado pode muitas vezes ser se alternar no ciclo da metodologia adotada.

III - Delimitações iniciais

III.1 - Domínio de Aplicação do projeto

O domínio de Aplicação do projeto é a produção científica e acadêmica dos docentes e discentes dos seguintes cursos de pós-graduação da Universidade de Brasília: **Engenharia Biomédica, Engenharia Elétrica P7 e P8 e Engenharia de Sistemas Eletrônicos e de Automação**. Os dados para análise foram obtidos a partir da plataforma elattes e da plataforma sucupira.

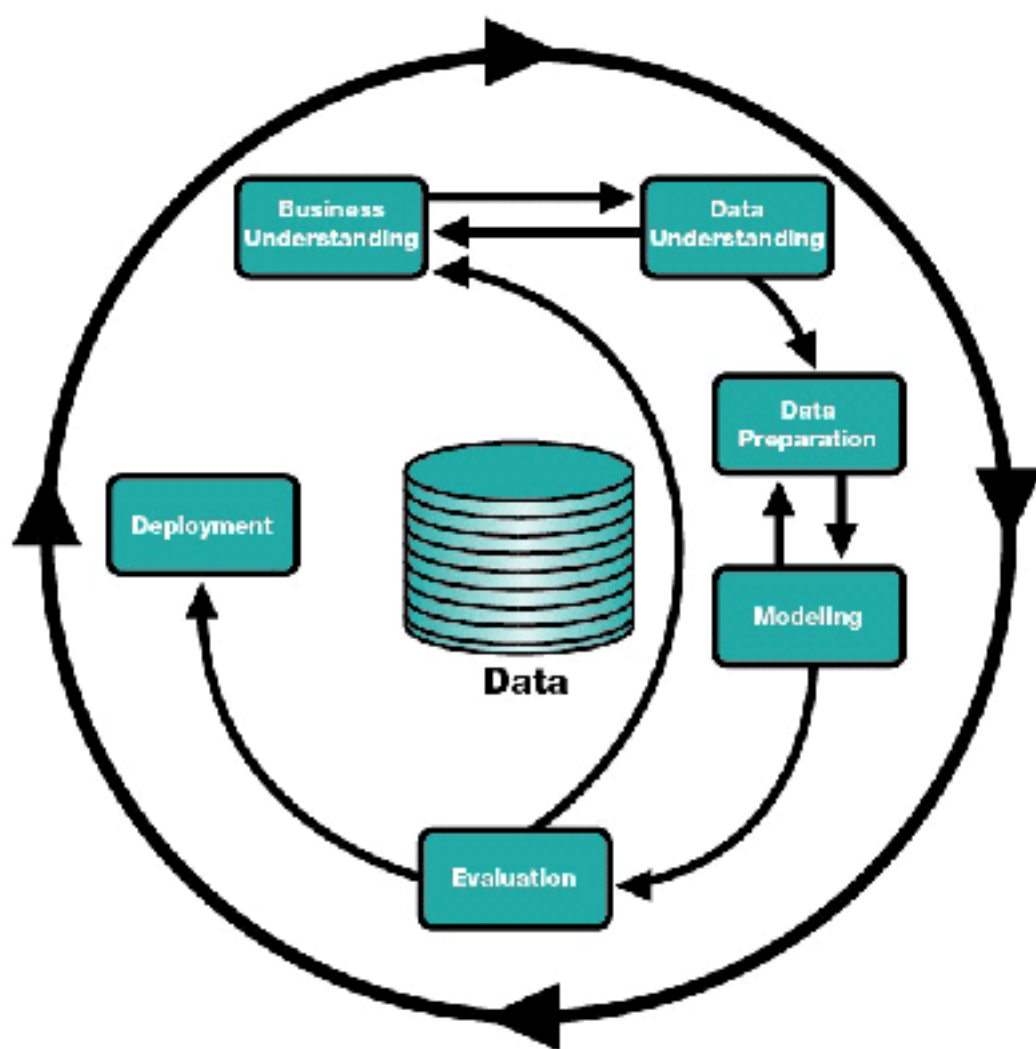


Figure 1: Ciclo do CRISP-DM que será usado como base

III.2 - Problema abordado

O problema tem por finalidade obter dados de forma descritiva, quantitativa e de modelagem computacional ou estatística, que permitam caracterizar como, porque e também para que ocorre a produção científica e acadêmica nos cursos de pós-graduação descritos anteriormente.

Engenharia Biomédica

O programa de pós-graduação da engenharia biomédica - possui as informações: Código: 53001010083P6; Área-base: 31300006.

O corpo discente é composto de 15 membros, já o docente é de 673, conforme scripts.

Numero de disciplina retirado do site <https://fga.unb.br/pgengbiomedica/lista-de-oferta> : 9 disciplinas.

Na área de concetração engenharia biomédica, há 4 áreas(linhas) de pesquisa:

- **Análise e Desenvolvimento de Sistemas Inteligentes e de Saúde :** " Relaciona os aspectos ligados à análise, desenvolvimento, aquisição, representação e descoberta do conhecimento para a produção de sistemas inteligentes e de saúde. É uma área multidisciplinar, que apoia a tomada de decisão, diagnóstico, tratamento e desenvolvimento de técnicas para prática na saúde. Aborda os principais avanços teóricos e práticos de metodologias e técnicas utilizadas no desenvolvimento de Sistemas Inteligentes e de Saúde. As principais áreas de investigação são: Análise e modelagem de dados epidemiológicos; Análise de sistemas de saúde; Gestão e avaliação de tecnologias em saúde; Engenharia Clínica; Inteligência Artificial; Aprendizado de Máquina; Sistemas Baseados em Conhecimento; Paradigmas Simbólicos; Conexionistas, Evolutivos e Híbridos; Inteligência Computacional; Nanotecnologia; entre outras."
- **Física Médica e Bioengenharia.** " Relaciona a aplicação e as consequências das radiações ionizantes e não ionizantes na área da saúde, que a cada dia se tornam mais complexas e desafiadoras. Isso exige que os profissionais da área possuam uma formação específica de grande qualidade em diferentes campos, tais como física, biologia e química. Ademais, estes profissionais devem dominar os fundamentos de proteção radiológica e conhecer normas nacionais e recomendações internacionais a ela relacionadas, visando a sua própria proteção, bem como à proteção do paciente, do público geral e do meio ambiente. "
- **Sistemas Eletrônicos e Instrumentação Biomédica :** " Abrange a pesquisa e o desenvolvimento de transdutores, equipamentos, aparelhos ou sistemas microprocessados de uso biológico ou aplicação médica, odontológica ou laboratorial. A linha aborda o desenvolvimento e análise tanto de software como de hardware, abrangendo componentes e técnicas de microeletrônica, circuitos integrados e dispositivos programáveis. A ênfase diz respeito a aplicações em prevenção, diagnóstico, tratamento ou reabilitação. Em comparação com outros produtos de engenharia, a instrumentação biomédica se distingue por obter seus sinais de entrada a partir de seres vivos ou seus extratos, e ainda por aplicar energia (eletromagnética, mecânica, térmica) diretamente a tecidos vivos. As principais áreas de investigação são: Desenvolvimento de eletrodos ativos; Aquisição de sinais biológicos e biomédicos; Estimulação de tecidos ou sistemas biológicos e biomédicos; Desenvolvimento e análise de dispositivos microeletrônicos; Projetos e configuração de dispositivos programáveis; Projeto de circuitos integrados. "
- **Processamento de Sinais e Imagens :** " Sinais e imagens são representações da realidade, na forma de funções matemáticas a serem submetidas a métodos rigorosos e eficientes de extração e tratamento de informação. Pode-se definir um sinal como um conjunto físico, abstrato ou simbólico de dados de interesse específico (i.e., informação) que muda ao longo do tempo ou do espaço. Uma imagem, especificamente, é conceituada como uma representação visual, real ou pictórica, parcial ou total, de uma estrutura, um ser ou até mesmo de ideias. É possível agrupar sinais e imagens como um mesmo tipo de objeto matemático, sendo que as imagens constituem sinais em domínios de duas ou mais dimensões. Processamento de sinais e imagens, portanto, consiste em utilizar métodos científicos e técnicas de engenharia para trabalhar com toda a informação contida nestas representações, conforme

definido pela Sociedade de Processamento de Sinais do IEEE em 2009. Isto significa o uso de representações, formalismos e técnicas matemáticas, estatísticas, computacionais, heurísticas e/ou linguísticas na informação existente para fins de representação, modelagem, análise, síntese, revelação, recuperação, percepção, identificação, aquisição, extração, aprendizagem, segurança e/ou análise forense. Na engenharia biomédica, a linha de processamento de sinais é fundamental para aplicações como desenvolvimento de sistemas de auxílio a diagnóstico, formação de imagens médicas, controle de próteses e outros dispositivos médicos, desenvolvimento de métodos eficientes de tratamento e condicionamento físico, análise científica de dados relacionados à saúde etc.”

Informações retiradas do próprio site, disponível em :https://fga.unb.br/articles/0001/4233/Linhas_de_Pesquisa_2016.pdf .

Há muitas projetos de pesquisas, no total 41 projetos. Sendo descrito no site <https://fga.unb.br/pgengbiomedica/projetos-de-pesquisas>:

Projetos de Pesquisa	Labs	Período	Financiador	Linha	Situação
AUTOBAC - Plataforma Embarcada para a Identificação de Bactérias por Processamento de Imagens	LIPIS	2011- Atual	Financiadora de Estudos e Projetos		Andamento
Biomaterial Látex	BioEng	2011- 2016	Capes		Andamento
Acompanhamento do crescimento e desenvolvimento de recém-nascidos pré-termo		2009- Atual		Instrumentação e Modelagem Biomédica	Andamento
Arte e Tecnociência: interações expandidas e condição híbrida em software art	LART	2010- 2014	Capes/CNPq	Processamento de Sinais Biomédicos e Imagens Médicas	Andamento
Atlas Anatômico 3D Aplicado à Mama	LIS	2009- 2010	FINATEC	Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
Biosensores Aplicados no Controle da Obesidade e Diabete Mellitus	BioEng	2010- 2015	CNPq	Instrumentação e Modelagem Biomédica	Andamento
Bolsa de Produtividade em Pesquisa - Construção de um Ambiente Web de Ensino e Aprendizagem para a Área de Saúde	LIS	2010- 2012	CNPq	Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
CIBERCOMUNICAÇÃO: interatividade, imersão, autonomia e mobilidade em software art - PQ 1B	LART	2009- 2010	CNPq	Projeto Isolado	Andamento
Colchão Inteligente Derivado de Latex anti-escara	BioEng	2009- 2014		Instrumentação e Modelagem Biomédica	Andamento
Computação Massivamente Paralela Utilizando Hardware Gráfico		2008- Atual	FAP/DF	Instrumentação e Modelagem Biomédica	Andamento

Projetos de Pesquisa	Labs	Período	Financiador	Linha	Situação
Consultório Médico Virtual (CMV)	LIS	2009- Atual	UnB/Gama	Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
Desenvolvimento de Dispositivo de Infusão Contínua de Insulina (Bomba de Insulina)		2010- 2011		Instrumentação e Modelagem Biomédica	Andamento
Desenvolvimento de Modelos e Busca de Parâmetros para Minimização da Energia usada na Reversão Elétrica da Fibrilação Cardíaca		2009- Atual	FINATEC	Processamento de Sinais Biomédicos e Imagens Médicas	Andamento
Desenvolvimento de um oclisor infantil, no formato de lente de contato, feito à base de látex natural aplicado a pacientes com tendência a desenvolver o olho ambliope.	BioEng	2009- 2015		Instrumentação e Modelagem Biomédica	Andamento
Desenvolvimento de um sistema físico de controle de fluxo esofágico para o tratamento da diabetes mellitus	BioEng	2010- 2012	FAP/DF	Instrumentação e Modelagem Biomédica	Andamento
Desenvolvimento de uma palmilha para pé diabético derivada de látex natural com controle de pressão plantar e com indução de neoformação tecidual.	BioEng	2009- 2012	CAPE	Instrumentação e Modelagem Biomédica	Andamento
Dinâmica de Sistemas Complexos com Aplicações em Biologia e Meio Ambiente		2007- 2010		Instrumentação e Modelagem Biomédica	Andamento
Estudo da Exequibilidade, Tolerabilidade e Eficácia do Módulo Controlador de Fluxo Esofágico	BioEng	2009- 2010		Instrumentação e Modelagem Biomédica	Andamento
Estudo de sistemas embarcáveis para controle de módulo de fluxo esofágico em aplicação para tratamento de obesidade em humanos	BioEng	2010- 2013	INBD – Engen- haria Eletrônica	Instrumentação e Modelagem Biomédica	Andamento
Estudo experimental para desenvolvimento de sistema de monitoração clínico e informatizado de fatores de risco para distúrbios cardiovasculares, por meio de avaliação de diversas variáveis do Sistema Nervoso Autônomo.		2010- 2011		Instrumentação e Modelagem Biomédica	Andamento
IEA - Instrumentação Embarcada em Aeronave	LIPIS	2009- 2010		Projeto Isolado	Andamento
MINIVANT - Desenvolvimento de um sistema de Mini-VANT para imageamento giro-estabilizado	LIPIS	2010- Atual	FINEP	Projeto Isolado	Andamento
Modelagem e caracterização de motores de ciclo Otto alimentados com etanol e Hidrogênio		2009- 2011	FAP/DF	Projeto Isolado	Andamento
Modelo Matematico da Força de Resistencia do Osso Humano a Perfurações em Cirurgias Ortopédicas - Aplicação Osteotomia de Tibia		2010- 2015		Instrumentação e Modelagem Biomédica	Andamento
Obesidade: Busca de Novo Método Tratamento, novo Índice de Diagnostico e Controle	BioEng	2010- 2020	Capes	Instrumentação e Modelagem Biomédica	Andamento

Projetos de Pesquisa	Labs	Período	Financiador	Linha	Situação
Preditores de morbi/mortalidade em portadores de doenças crônico-degenerativas em regime de internação domiciliar: coorte prospectiva		2010-Atual		Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
Prevalência de tabagismo nos adolescentes e jovens do ensino médio da escola pública em Ceilândia		2010-Atual		Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
ProDC: Promover o Desempenho Cognitivo	LIS	2010-Atual		Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
Propriedades físicas de nanopartículas fora do equilíbrio, modelagem e projeto automatizado de nanocircuitos eletrônicos		2010-2012	CNPq	Instrumentação e Modelagem Biomédica	Andamento
Prótese de pé ativa para reabilitação e assistência a locomoção humana	BioEng	2009-2015		Instrumentação e Modelagem Biomédica	Andamento
Pêndulo Invertido	BioEng	2009-2011		Instrumentação e Modelagem Biomédica	Andamento
Re-estruturação de um dispositivo para UTI neonatal e Home Care - CPAP	BioEng	2010-2012	Brakko - Cooperação	Instrumentação e Modelagem Biomédica	Andamento
SAPIO - Sistema de Aquisição e Processamento de Imagens Ovitrapas	LIPIS	2008-2010	FINEP	Processamento de Sinais Biomédicos e Imagens Médicas	Andamento
Sistema de aquisição de sinais de temperatura e vibração em cirurgia de acesso ao nervo facial		2010-2012		Instrumentação e Modelagem Biomédica	Andamento
Sistema de Controle de Velocidade, Temperatura e Força em Cirurgias Ortopédicas		2009-2014		Instrumentação e Modelagem Biomédica	Andamento
Transtorno de Déficit de Atenção e Hiperatividade (TDAH)	LIS	2010-2010		Sistemas Inteligentes Aplicados à Eng. Biomédica	Andamento
Monitoramento eletrônico da variação de temperatura durante a cirurgia descompressiva do nervo facial estudo experimental em gatos	BioEng	2011-2012	FPDF		Andamento
Biblioteca digital de imagens e sinais biomédicos	LabGS	2011-2012	UNPIS		Andamento
Determinação do fator Gama para Cálculo de Taxa de Exposição em Medicina Nuclear	FCE	2010-atual	UnB	Física Médica	Andamento

A organização original em uma única área de concentração (Engenharia Biomédica) com quatro linhas de

pesquisa (Física Médica; Instrumentação e Modelagem Biomédica; Processamento de Sinais Biomédicos e Imagens Médicas; Sistemas Inteligentes Aplicados à Engenharia Biomédica) é coerente, consistente e em consonância com a proposta aprovada. O número de projetos (39 no total) é elevado para a dimensão do corpo docente, parecendo corresponder aos temas das dissertações dos alunos. Recomenda-se concentrar os estudos em projetos consistentes. A proposta de programa aponta para possíveis mudanças no futuro próximo, em função de processo de credenciamento e credenciamento de novos docentes, visando aumentar a produção e melhorar sua distribuição. A proposta curricular é adequada: 24 créditos, sendo 16 gerais, obrigatórias e 8 optativas, das áreas. Admite dispensa de cursos com teste de proficiência e há iniciativa de ensino à distância. A proposta valoriza o estabelecimento de cooperações com grupos de pesquisa nacionais (12), internacionais (13) e empresas (4). Aponta ainda ação de interesse social/sanitário (controle da dengue) e nucleação com o IFMT (Cuiabá). A autoavaliação reflete um esforço de busca de amadurecimento e consolidação. O Programa conta com 10 laboratórios com aparato instrumental adequado para a realização das atividades de ensino e pesquisa. Ao longo do período, parte das atividades e Laboratórios funcionaram em instalações provisórias. Mais ao final, houve mudança para um novo edifício, o que interferiu no próprio desempenho acadêmico/científico do programa. Há disponibilidade de recursos para um edifício adicional, previsto para o próximo quadriênio.

Sobre a última avaliação do curso engenharia biomédica, serve: <https://sucupira.capes.gov.br/sucupira/public/consultas/avaliacao/viewPreenchimentoFicha.jsf?idFicha=5739&popup=true>. Verifica-se que as notas desde de 2010 até 2017 é 3 e que ainda continua em funcionamento.

Sobre a última avaliação de 2017 temos que :

Nos quesitos de avaliação :

- Proposta do programa : Bom
- Corpo Docente : Regular
- Corpo Discente, Teses e Dissertações : Bom
- Produção Intelectual : Regular
- Inserção Social : Muito Bom

Especificamente sobre a proposta do programa em 2017, tem-se que

Nos quesitos de avaliação :

- Proposta do programa : Bom
- Corpo Docente : Regular
- Corpo Discente, Teses e Dissertações : Bom
- Produção Intelectual : Regular
- Inserção Social : Muito Bom

“A organização original em uma única área de concentração (Engenharia Biomédica) com quatro linhas de pesquisa (Física Médica; Instrumentação e Modelagem Biomédica; Processamento de Sinais Biomédicos e Imagens Médicas; Sistemas Inteligentes Aplicados à Engenharia Biomédica) é coerente, consistente e em consonância com a proposta aprovada. O número de projetos (39 no total) é elevado para a dimensão do corpo docente, parecendo corresponder aos temas das dissertações dos alunos. Recomenda-se concentrar os estudos em projetos consistentes. A proposta de programa aponta para possíveis mudanças no futuro próximo, em função de processo de credenciamento e credenciamento de novos docentes, visando aumentar a produção e melhorar sua distribuição. A proposta curricular é adequada: 24 créditos, sendo 16 gerais, obrigatórias e 8 optativas, das áreas. Admite dispensa de cursos com teste de proficiência e há iniciativa de ensino à distância. A proposta valoriza o estabelecimento de cooperações com grupos de pesquisa nacionais (12), internacionais (13) e empresas (4). Aponta ainda ação de interesse social/sanitário (controle da dengue) e nucleação com o IFMT (Cuiabá). A autoavaliação reflete um esforço de busca de amadurecimento e consolidação. O Programa conta com 10 laboratórios com aparato instrumental adequado para a realização das atividades

de ensino e pesquisa. Ao longo do período, parte das atividades e Laboratórios funcionaram em instalações provisórias. Mais ao final, houve mudança para um novo edifício, o que interferiu no próprio desempenho acadêmico/científico do programa. Há disponibilidade de recursos para um edifício adicional, previsto para o próximo quadriênio.” Informação retirada de : <https://sucupira.capes.gov.br/sucupira/public/consultas/avaliacao/viewPreenchimentoFicha.jsf?idFicha=5739&popup=true>

Sobre o corpo docente: “O número informado de docentes do Programa é estável (8 a 9 no período), mas compreende várias entradas e saídas de docentes no quadriênio, e alternâncias de classificação de docentes entre permanentes e colaboradores, sendo constituído por professores em tempo integral e dedicação exclusiva. O Programa conta com apenas um bolsista de Produtividade em Pesquisa do CNPq. A Proposta aponta para a necessidade de credenciamentos e descredenciamentos de docentes, deixando claro que o núcleo de docentes permanentes do Programa ainda não está consolidado. A carga letiva dos docentes permanentes encontra-se um pouco abaixo da média, mas está com boa distribuição. No caso da graduação, alcança” muito bom” em termos de carga didática, e o número de alunos de iniciação científica é de pouco mais de um por docente.” Informação retirada de : <https://sucupira.capes.gov.br/sucupira/public/consultas/avaliacao/viewPreenchimentoFicha.jsf?idFicha=5739&popup=true>

O corpo docente é descrito por “O número de dissertações concluídas por docente e o tempo médio de titulação são muito bons, enquanto o percentual de docentes responsável pelas titulações ano a ano é bom. Entretanto, resultou regular o indicador de qualidade das dissertações, que admite como produção intelectual mínima, relacionada à dissertação, um artigo completo em congresso ou periódico.” Informação retirada de : <https://sucupira.capes.gov.br/sucupira/public/consultas/avaliacao/viewPreenchimentoFicha.jsf?idFicha=5739&popup=true>

Já sobre produção intelectual :

Itens de Avaliação	Peso	Avaliação
4.1. Publicações qualificadas do Programa por docente permanente.	50.0	Fraco
4.2. Distribuição de publicações qualificadas em relação ao corpo docente permanente do Programa.	30.0	Regular
4.3. Produção técnica, patentes e outras produções consideradas relevantes.	20.0	Bom

“O programa apresentou como produção intelectual um total de 63 artigos completos no quadriênio. Entretanto, após a distribuição nos diferentes estratos do QUALIS, apenas 16 foram qualificados como produção relevante (extratos A1, A2 e B1). Tal montante corresponde a meio artigo relevante por docente por ano, aquém da média dos programas da área. Naturalmente, a autoria desses artigos concentra-se em uma fração inadequada do corpo docente. Cabe salientar que alguns docentes permanentes têm investido em solicitação de patentes. A lista de produção intelectual tem um alto número de trabalhos listados como capítulos de livros, e mesmo livros, mas não há justificativa da importância das obras conforme orientação do documento de área das Engenharias IV. Em vários casos, é claro tratar-se de comunicações em congressos (p/ex da série IFMBE Proceedings).” Informação retirada de : <https://sucupira.capes.gov.br/sucupira/public/consultas/avaliacao/viewPreenchimentoFicha.jsf?idFicha=5739&popup=true>

Por fim, sobre inserção social :

Itens de Avaliação	Peso	Avaliação
5.1. Inserção e impacto regional e (ou) nacional do programa.	40.0	Muito Bom
5.2. Integração e cooperação com outros programas e centros de pesquisa e desenvolvimento profissional relacionados à área de conhecimento do programa, com vistas ao desenvolvimento da pesquisa e da pós-graduação.	40.0	Muito Bom
5.3 - Visibilidade ou transparência dada pelo programa a sua atuação.	20.0	Muito Bom

Engenharia de Sistemas Eletrônicos e de Automação

O programa surgiu por meio do desmembramento do Programa de Pós Graduação em Engenharia Elétrica da UnB em dois novos Programas, ocorrido em 2010. O objetivo do programa foi julgado como coerente e consistente em termos de projetos, áreas de concentração e linhas de pesquisa; A estrutura curricular do programa foi avaliada como adequada ao perfil dos docentes. Em 2015, o regimento interno do programa foi reformado devido a uma preocupação sobre o futuro do mesmo, sendo esta preocupação a manutenção dos docentes em atividades de pós-doutorado no exterior, fortalecendo relações internacionais, esta estratégia é muito bem utilizada pelo programa. A infraestrutura disponível para a realização das atividades do programa é bem distribuída entre os docentes e discentes e considerada muito boa.

Nos últimos anos, o programa obteve 206 publicações no total, possuindo 94 discentes, 21 docentes e 53 disciplinas desde sua criação, assim como 4 linhas de pesquisa, sendo elas: CONTROLE, IDENTIFICAÇÃO E FILTRAGEM; MICROELETRÔNICA; PROCESSAMENTO DE ÁUDIO, IMAGEM E VÍDEO e ENGENHARIA BIOMÉDICA.

Resumindo os projetos de pesquisa do programa, os objetivos dos mesmos são: Processamento de sinais e imagem para análise de sinais SEMG; Avaliar o impacto do uso de ambas constelações de satélites GPS e GLONASS, ao invés de soluções baseadas em uma única constelação (GPS ou GLONASS) para cálculo de posicionamento com alta precisão; Estudar os problemas de estabilização via controle realimentado e estimação de estados via filtragem da saída de sistemas dinâmicos envolvendo deslocamentos rígidos; Desenvolvimento de diversos sistemas de instrumentação biomédicas, para a monitoração, o diagnóstico e o tratamento de doenças e/ou para a melhoria da qualidade de vida de seres humanos; Desenvolvimento de medidor de transmissão e distribuição de energia e desenvolvimento da ferramenta computacional na análise de grandezas relacionadas à qualidade da energia elétrica, proporcionando possibilidades de avaliações de conjuntos de dados, considerando-se metodologias estatísticas e normas brasileiras e internacionais que versam sobre o assunto; Projetar e construir um veículo aéreo não tripulado (VANT) de pequeno porte movido à energia solar para voos de longa duração (até 8 horas ininterruptas) e de baixo custo; Propor um método NR para aferição da qualidade de vídeo que seja genérico (não restrito a aplicações e degradações específicas), rápido (execução em tempo real) e preciso (boa correlação com medidas subjetivas); Investigar, desenvolver e implementar em plataforma experimental algoritmos para análise, controle e filtragem de sistemas dinâmicos com vibrações; Analisar e caracterizar os diversos artefatos digitais, suas combinações e interferências; Construção de órteses ativas transparentes para o auxílio à locomoção de pessoas com deficiência; Implantação do Laboratório de Certificação de Equipamentos Eletromédicos (LabCert) da Universidade de Brasília (UnB); Mostrar o valor da Floresta Amazônica, aprendendo com a natureza e em seguida ousar, na produção de novos materiais, na forma de óleos modificados pela inserção de nanoestruturas, filmes e blendas poliméricas, tendo como dopantes óleos ou substâncias presentes neles; Transferências e posturas assistidas por eletroestimulação poderiam ser uma alternativa para opções mais portáteis de tecnologias que contribuiriam para uma verticalização mais funcional e com maior participação do sujeito, baseada na coordenação entre o controle voluntário do usuário e suas respostas periféricas influenciadas pela eletroestimulação (controle artificial).

De acordo com a avaliação quadrienal, as linhas de pesquisa possuem membros do corpo docente suficientes para atuação de forma exclusiva nas mesmas. O indicador de consolidação do corpo docente, que é determinado pelo número de docentes permanentes que possuem bolsa de produtividade do CNPq é bom, porém precisa de atenção. A dedicação do corpo docente nos projetos de pesquisa e nas disciplinas é adequado, entretanto necessita de atenção para o próximo período de avaliação. A distribuição das orientações dos projetos de pesquisa possui concentração em um grupo de docentes permanentes, também merecendo atenção por parte do programa para melhorar tal distribuição. O número de trabalhos de conclusão defendidos é adequado de acordo com o tamanho do corpo docente, porém a avaliação alerta sobre a tendência de diminuição deste número de dissertações, apontando como uma possível melhora ao programa a qualificação da produção científica associada com estas teses e dissertações. O tempo médio de formação dos discentes foi avaliado como regular pela última avaliação. A produção intelectual do programa possui diversos trabalhos de conclusão publicados em meios que possuem Qualis alto, sendo muito boa em quantidade e qualidade. Os discentes e docentes possuem um alto volume de publicações de trabalhos completos em eventos. O programa é relevante na formação de discentes em nível de mestrado, cooperando com instituições nacionais e inter-

nacionais para promover o desenvolvimento da região, pode-se perceber que a inserção social do programa é boa devido as oportunidades de publicações internacionais e ao website do programa que proporciona boa visibilidade ao mesmo. Para aumentar a inserção social, o programa deve publicar mais em português para aumentar a participação em revistas nacionais de Qualis de estratos superiores.

O programa só possui uma publicação na revista brasileira “Revista Brasileira de Engenharia Biomédica”, de Qualis B2, no entanto, o programa possui diversas publicações em revistas internacionais de Qualis A1, A2 e B2 como por exemplo: “IEEE Signal Processing Letters”, “IEEE Transactions on Circuits and Systems for Video Technology”, “Journal of Control, Automation and Electrical Systems”, “Microelectronics” e “IEEE TRANSACTIONS ON NANOTECHNOLOGY”.

Todos os dados utilizados acima foram retirados da Plataforma Sucupira.

Engenharia Elétrica P8

Engenharia Elétrica P7

IV - Entendimento do Negócio

A Universidade de Brasília é uma das grandes universidades federais, tendo enfoque na produção de artigos acadêmicos. Os programas a seguir são todos da área de avaliação ENGENHARIAS IV e serão analisados pelo trabalho como já exposto anteriormente:

IV.1 - Programas de pós-graduação Grupo 15

Nome do Programa	ME	DO	MP
ENGENHARIA ELÉTRICA (53001010059P8)	-	-	3
ENGENHARIA ELÉTRICA (53001010080P7)	4	4	-
ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO (53001010081P3)	4	4	-
ENGENHARIA BIOMÉDICA (53001010083P6)	3	-	-

Sendo o enfoque principal os dados referente a Engenharia Biomédica, que por sua vez possui muitas de suas pesquisas voltadas a novas soluções na área de Biomédica e demais áreas representadas na tabela acima.

IV.2 - Avaliação das Circunstâncias

Este trabalho terá como principal circunstância o limite que os próprios dados obtidos por meio das plataformas elattes e sucupira impõem, delimitando a análise dos dados de acordo com a disponibilidade de informações disponíveis nesta plataforma.

V - Entendimento dos Dados

V.1 - Coleta inicial dos dados

Todos os arquivos com dados iniciais a seguir foram fornecidos pelos professores responsáveis pela disciplina, através da plataforma elattes. Os dados foram gerados no mês de setembro de 2018, data em que o presente grupo começou a realizar o trabalho e compilam informações entre os anos de 2010 e 2017, das áreas expostas na seção anterior. Todos os arquivos se encontram no formato JSON.

Perfil profissional dos docentes vinculados às pós-graduações

```
json.perfil.biomed <- "dados-2018-2/engenharia-biomedica/279.profile.json"
file.info(json.perfil.biomed)
```

```
##                                size isdir mode
## dados-2018-2/engenharia-biomedica/279.profile.json 770014 FALSE 666
##                                                                mtime
## dados-2018-2/engenharia-biomedica/279.profile.json 2018-09-20 19:26:09
##                                                                ctime
## dados-2018-2/engenharia-biomedica/279.profile.json 2018-09-20 19:26:09
##                                                                atime exe
## dados-2018-2/engenharia-biomedica/279.profile.json 2018-12-02 21:24:51 no
```

```
json.perfil.auto <- "dados-2018-2/engenharia-eletronicos-automacao/280.profile.json"
file.info(json.perfil.auto)
```

```
##                                size
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json 828607
##                                                                isdir mode
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json FALSE 666
##                                                                mtime
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json 2018-09-20 19:26:09
##                                                                ctime
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json 2018-09-20 19:26:09
##                                                                atime
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json 2018-12-02 21:24:52
##                                                                exe
## dados-2018-2/engenharia-eletronicos-automacao/280.profile.json no
```

```
json.perfil.ele_p8 <- "dados-2018-2/engenharia-eletrica-p8/281.profile.json"
file.info(json.perfil.ele_p8)
```

```
##                                size isdir mode
## dados-2018-2/engenharia-eletrica-p8/281.profile.json 559453 FALSE 666
##                                                                mtime
## dados-2018-2/engenharia-eletrica-p8/281.profile.json 2018-09-20 19:26:09
##                                                                ctime
## dados-2018-2/engenharia-eletrica-p8/281.profile.json 2018-09-20 19:26:09
##                                                                atime
## dados-2018-2/engenharia-eletrica-p8/281.profile.json 2018-12-02 21:24:52
##                                                                exe
## dados-2018-2/engenharia-eletrica-p8/281.profile.json no
```

```
json.perfil.ele_p7 <- "dados-2018-2/engenharia-eletrica-p7/282.profile.json"
file.info(json.perfil.ele_p7)
```

```
##                                size isdir mode
## dados-2018-2/engenharia-eletrica-p7/282.profile.json 711039 FALSE 666
##                                                                mtime
## dados-2018-2/engenharia-eletrica-p7/282.profile.json 2018-09-20 19:26:09
##                                                                ctime
## dados-2018-2/engenharia-eletrica-p7/282.profile.json 2018-09-20 19:26:09
##                                                                atime
## dados-2018-2/engenharia-eletrica-p7/282.profile.json 2018-12-02 21:24:52
##                                                                exe
## dados-2018-2/engenharia-eletrica-p7/282.profile.json no
```

Os arquivos dados-2018-2/engenharia-biomedica/279.profile.json, dados-2018-2/engenharia-eletronicos-automacao/280.profile.json, dados-2018-2/engenharia-eletrica-p7/282.profile.json, dados-2018-2/engenharia-eletrica-p8/281.profile.json apresentam dados sobre o perfil de todos os docentes vinculados a programas de pós-graduação da UnB alvos do trabalho, entre os anos de 2010 e 2017.

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

```
json.advise.biomed <- "dados-2018-2/engenharia-biomedica/279.advise.json"
file.info(json.advise.biomed)
json.advise.auto <- "dados-2018-2/engenharia-eletronicos-automacao/280.advise.json"
file.info(json.advise.auto)
json.advise.ele_p8 <- "dados-2018-2/engenharia-eletrica-p8/281.advise.json"
file.info(json.advise.ele_p8)
json.advise.ele_p7 <- "dados-2018-2/engenharia-eletrica-p7/282.advise.json"
file.info(json.advise.ele_p7)
```

Os arquivos dados-2018-2/engenharia-biomedica/279.advise.json, dados-2018-2/engenharia-eletronicos-automacao/280.advise.json, dados-2018-2/engenharia-eletrica-p7/282.advise.json, dados-2018-2/engenharia-eletrica-p8/281.advise.json apresenta dados sobre o orientações de mestrado e doutorado feitas por todos os docentes vinculados a programas de pós-graduação da UnB alvos do trabalho, entre os anos de 2010 e 2017.

Produção bibliográfica gerada pelos docentes vinculados às pós-graduações

```
json.producao.bibliografica.biomed <- "dados-2018-2/engenharia-biomedica/279.publication.json"
file.info(json.producao.bibliografica.biomed)
json.producao.bibliografica.auto <- "dados-2018-2/engenharia-eletronicos-automacao/280.publication.json"
file.info(json.producao.bibliografica.auto)
json.producao.bibliografica.ele_p8 <- "dados-2018-2/engenharia-eletrica-p8/281.publication.json"
file.info(json.producao.bibliografica.ele_p8)
json.producao.bibliografica.ele_p7 <- "dados-2018-2/engenharia-eletrica-p7/282.publication.json"
file.info(json.producao.bibliografica.ele_p7)
```

Os arquivos dados-2018-2/engenharia-biomedica/279.publication.json, dados-2018-2/engenharia-eletronicos-automacao/280.publication.json, dados-2018-2/engenharia-eletrica-p7/282.publication.json e dados-2018-2/engenharia-eletrica-p8/281.publication.json apresenta dados sobre a Produção bibliográfica gerada por todos os docentes vinculados a programas de pós-graduação da UnB alvos do trabalho, entre os anos de 2010 e 2017.

ID's dos docentes participantes e o que contém o arquivo list.json

```
json.list.biomed <- "dados-2018-2/engenharia-biomedica/279.list.json"
file.info(json.list.biomed)
json.list.auto <- "dados-2018-2/engenharia-eletronicos-automacao/280.list.json"
file.info(json.list.auto)
json.list.ele_p8 <- "dados-2018-2/engenharia-eletrica-p8/281.list.json"
file.info(json.list.ele_p8)
json.list.ele_p7 <- "dados-2018-2/engenharia-eletrica-p7/282.list.json"
file.info(json.list.ele_p7)
```

Os arquivos dados-2018-2/engenharia-biomedica/279.list.json apresenta o id de todos docentes vinculados a programas de pós-graduação da UnB alvos do trabalho, entre os anos de 2010 e 2017, porém tal arquivo se mostra inútil, haja visto que só contém o id e que as outras variáveis como “**nome**” sempre tem seu valor igual a “”.

Redes de colaboração entre docentes

```
json.graph.biomed<- 'dados-2018-2/engenharia-biomedica/279.graph.json'
file.info(json.graph.biomed)
json.graph.auto <- "dados-2018-2/engenharia-eletronicos-automacao/280.graph.json"
file.info(json.graph.auto)
json.graph.ele_p8 <- "dados-2018-2/engenharia-eletrica-p8/281.graph.json"
file.info(json.graph.ele_p8)
json.graph.ele_p7 <- "dados-2018-2/engenharia-eletrica-p7/282.graph.json"
file.info(json.graph.ele_p7)
```

Os arquivos dados-2018-2/engenharia-biomedica/279.graph.json, dados-2018-2/engenharia-eletrica-p7/282.graph.json, dados-2018-2/engenharia-eletrica-p8/281.graph.json e dados-2018-2/engenharia-eletronicos-automacao/280.graph.json apresenta redes de colaboração na co-autoria de artigos científicos, feitas entre os docentes vinculados a programas de pós-graduação da UnB alvos do trabalho, entre os anos de 2010 e 2017.

V.2 - Descrição dos Dados

Para ler, manipular, analisar e visualizar estes dados, serão utilizadas as seguintes bibliotecas:

```
library(jsonlite)
library(listviewer)
library(readxl)
library(readr)
library(readtext)
library(ggplot2)
library(tidyverse)
library(stringr)
library(igraph)
library(tm)
library(wordcloud)
library(RColorBrewer)
library(SnowballC)
```

Com estas bibliotecas seremos capazes de responder e determinar qual o volume de dados, a estrutura dos dados (tipos), codificações usadas, entre outras atividades importantes para análise dos dados.

Descrição dos Dados do Perfil

Engenharia Biomédica

```
unb.prof.biomed <- fromJSON("dados-2018-2/engenharia-biomedica/279.profile.json")
```

Quantidade de docentes sob análise:

```
length(unb.prof.biomed)
```

```
## [1] 15
```

Para um melhor entendimento é importante saber como os dados estão dispostos tomando um docente como exemplo, a seguir o primeiro docente no arquivo foi utilizado para análise da estrutura do arquivo.

```
### Usando glimpse
glimpse(unb.prof.biomed[[1]])
```

```
## List of 7
## $ nome           : chr "Leandro Xavier Cardoso"
```

```

## $ resumo_cv          : chr "Possui graduação em Física Bacharelado e Licenciatura pela Universid
## $ areas_de_atuacao    : 'data.frame':  1 obs. of  4 variables:
##   ..$ grande_area    : chr ""
##   ..$ area           : chr ""
##   ..$ sub_area       : chr ""
##   ..$ especialidade  : chr ""
## $ endereco_profissional :List of 8
##   ..$ instituicao: chr "UnB - UNIVERSIDADE DE BRASÃ?LIA - FGA - FACULDADE GAMA"
##   ..$ orgao      : chr ""
##   ..$ unidade    : chr ""
##   ..$ DDD        : chr "61"
##   ..$ telefone   : chr "31078903"
##   ..$ bairro     : chr "Setor Leste (Gama)"
##   ..$ cep        : chr "72444240"
##   ..$ cidade     : chr "Brasília"
## $ producao_bibliografica :List of 2
##   ..$ EVENTO      : 'data.frame':   3 obs. of  11 variables:
##     .. ..$ natureza      : chr [1:3] "COMPLETO" "COMPLETO" "COMPLETO"
##     .. ..$ titulo        : chr [1:3] "As Tic's na Educação: Mudnça ou Modernização" "A Educação no C
##     .. ..$ nome_do_evento : chr [1:3] "IV Colóquio Internacional Educação e Contemporaneidade" "IV Co
##     .. ..$ ano_do_trabalho : chr [1:3] "2010" "2010" "2014"
##     .. ..$ pais_do_evento : chr [1:3] "Brasil" "Brasil" "Brasil"
##     .. ..$ cidade_do_evento: chr [1:3] "São Cristovão" "São Cristovão" "Uberlândia"
##     .. ..$ doi           : chr [1:3] "" "" ""
##     .. ..$ classificacao  : chr [1:3] "INTERNACIONAL" "INTERNACIONAL" "NACIONAL"
##     .. ..$ paginas       : chr [1:3] " - " " - " "2636 - 2638"
##     .. ..$ autores       :List of 3
##     .. ..$ autores-endogeno:List of 3
##   ..$ PERIODICO: 'data.frame':   10 obs. of  10 variables:
##     .. ..$ natureza      : chr [1:10] "COMPLETO" "COMPLETO" "COMPLETO" "COMPLETO" ...
##     .. ..$ titulo        : chr [1:10] "Thermoluminescent dose reconstruction using quartz extracted
##     .. ..$ periodico     : chr [1:10] "Journal of Physics. Conference Series (Online)" "Scientia Plen
##     .. ..$ ano           : chr [1:10] "2010" "2011" "2011" "2013" ...
##     .. ..$ volume       : chr [1:10] "249" "7" "80" "477" ...
##     .. ..$ issn          : chr [1:10] "17426596" "18082793" "20103778" "17426588" ...
##     .. ..$ paginas       : chr [1:10] "012031 - " "014101 - " "285 - 290" "012011 - " ...
##     .. ..$ doi          : chr [1:10] "10.1088/1742-6596/249/1/012031" "" "" "10.1088/1742-6596/477/
##     .. ..$ autores       :List of 10
##     .. ..$ autores-endogeno:List of 10
## $ orientacoes_academicas:List of 2
##   ..$ ORIENTACAO_EM_ANDAMENTO_MESTRADO: 'data.frame':  5 obs. of  13 variables:
##     .. ..$ natureza      : chr [1:5] "Dissertação de mestrado" "Dissertação de mestrado"
##     .. ..$ titulo        : chr [1:5] "Utilização de dosímetros termoluminescentes comerci
##     .. ..$ ano           : chr [1:5] "2014" "2014" "2015" "2015" ...
##     .. ..$ id_lattes_aluno : chr [1:5] "" "" "" "" ...
##     .. ..$ nome_aluno     : chr [1:5] "Rafael Assunção Gomes de Souza" "Marcelo Oppermann"
##     .. ..$ instituicao     : chr [1:5] "Universidade de Brasília" "Universidade de Brasília
##     .. ..$ curso         : chr [1:5] "Engenharia Biomédica" "Engenharia Biomédica" "Engenl
##     .. ..$ codigo_do_curso : chr [1:5] "60059672" "90000006" "90000006" "60059672" ...
##     .. ..$ bolsa         : chr [1:5] "NAO" "NAO" "NAO" "NAO" ...
##     .. ..$ agencia_financiadora : chr [1:5] "" "" "" "" ...
##     .. ..$ codigo_agencia_financiadora: chr [1:5] "" "" "" "" ...
##     .. ..$ nome_orientadores :List of 5
##     .. ..$ id_lattes_orientadores :List of 5

```

```
## ..$ OUTRAS_ORIENTACOES_CONCLUIDAS : 'data.frame': 5 obs. of 13 variables:
## .. ..$ natureza : chr [1:5] "MONOGRAFIA_DE_CONCLUSAO_DE_CURSO_APERFEICOAMENTO_E_
## .. ..$ titulo : chr [1:5] "Insuficiência na Aprendizagem de Matemática do 9º an
## .. ..$ ano : chr [1:5] "2011" "2011" "2011" "2011" ...
## .. ..$ id_lattes_aluno : chr [1:5] "" "" "" "" ...
## .. ..$ nome_aluno : chr [1:5] "Décio Luiz Alves Barreto e outros" "Izaque dos Santo
## .. ..$ instituicao : chr [1:5] "Faculdade Serigy" "Faculdade Serigy" "Faculdade Ser
## .. ..$ curso : chr [1:5] "Fundamentos e Métodos do ensino da Matemática" "Fun
## .. ..$ codigo_do_curso : chr [1:5] "900000005" "900000005" "900000005" "900000005" ...
## .. ..$ bolsa : chr [1:5] "NAO" "NAO" "NAO" "NAO" ...
## .. ..$ agencia_financiadora : chr [1:5] "" "" "" "" ...
## .. ..$ codigo_agencia_financiadora: chr [1:5] "" "" "" "" ...
## .. ..$ nome_orientadores :List of 5
## .. ..$ id_lattes_orientadores :List of 5
## $ senioridade : chr "9"
```

Podemos inferir que: O professor não é da área da engenharia elétrica por formação, mas acabou aderindo a subárea engenharia Biomédica; É formado pela UFG; Atualmente trabalha no campus Gama da UnB. Sua senioridade é de 9.

Engenharia de Sistemas Eletrônicos e de Automação

```
unb.prof.auto <- fromJSON("dados-2018-2/engenharia-eletronicos-automacao/280.profile.json")
```

Quantidade de docentes sob análise:

```
length(unb.prof.auto)
```

```
## [1] 21
```

Engenharia Elétrica P8

```
unb.prof.ele_p8 <- fromJSON("dados-2018-2/engenharia-eletrica-p8/281.profile.json")
```

Quantidade de docentes sob análise:

```
length(unb.prof.ele_p8)
```

```
## [1] 11
```

Engenharia Elétrica P7

```
unb.prof.ele_p7 <- fromJSON("dados-2018-2/engenharia-eletrica-p7/282.profile.json")
```

Quantidade de docentes sob análise:

```
length(unb.prof.ele_p7)
```

```
## [1] 14
```

Descrevendo Dados de Orientação

Engenharia Biomédica

```
###Descrição dos dados de orientações
```

```
unb.adv.biomed <- fromJSON("dados-2018-2/engenharia-biomedica/279.advise.json")
```

```
# Mostrando as listas presentes neste arquivo.
```

```
names(unb.adv.biomed)
```

```

## [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"

# Explorando um nível de detalhe de Orientações de doutorados concluídas
names(unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO)

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

##### DOUTORADO 2017
#Buscando cursos que mais produziram doutorados.
head(sort(table(unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)

##
## ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO
##                                     1
##           Pós-graduação em Ciências Médicas
##                                     1

#Sabendo suas instituições
head(sort(table(unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$instituicao), decreasing = TRUE), 10)

##
## Universidade de Brasília - Campus Darcy Ribeiro
##                                     1
##           Universidade de Brasília
##                                     1

#Sabendo seus orientadores
data_orienM <- capture.output(str(unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$nome_orientadores))
unique(data_orienM)

## [1] "List of 2"
## [2] " $ : chr \"Marcus Vinícius Chaffim Costa\""
## [3] " $ : chr \"Marilia Miranda Forte Gomes\""

#Sabendo os nomes dos trabalhos
data_orienM <- capture.output(str(unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$titulo))
unique(data_orienM)

## [1] " chr [1:2] \"COMPRESSÃO DE SINAIS DE S-EMG EM ABORDAGENS 1D E 2D\" ..."

##### MESTRADO 2017
#Buscando cursos que mais produziram mestrados.
head(sort(table(unb.adv.biomed$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)

##
##           Engenharia Biomédica Mestrado em Engenharia Biomédica
##                                     9                                     4
##           ENGENHARIA ELÉTRICA                                     Física
##                                     1                                     1
##           PPDSCI/CEAM
##                                     1

```



```
#Sabendo suas instituições
```

```
head(sort(table(unb.adv.biomed$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$instituicao), decreasing = TRUE), 1
```

```
##
##                               Faculdade do Gama da UnB
##                               6
##                               Universidade de Brasília
##                               5
## Universidade de Brasília - Faculdade UnB-Gama
##                               3
##                               Faculdade UNB Gama - FGA
##                               1
## Universidade de Brasília - Campus Darcy Ribeiro
##                               1
```

```
#Sabendo seus orientadores
```

```
data_orien3 <- capture.output(str(unb.adv.biomed$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_orientadores))
unique(data_orien3)
```

```
## [1] "List of 16"
## [2] " $ : chr \"Georges Daniel Amvame Nze\""
## [3] " $ : chr \"Jose Felicio da Silva\""
## [4] " $ : chr \"Lourdes Mattos Brasil\""
## [5] " $ : chr \"Marcelino Monteiro de Andrade\""
## [6] " $ : chr \"Marilia Miranda Forte Gomes\""
## [7] " $ : chr \"Ronni Geraldo Gomes de Amorim\""
```

```
#Sabendo seus orientados
```

```
data_orien2 <- capture.output(str(unb.adv.biomed$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_aluno))
unique(data_orien2)
```

```
## [1] " chr [1:16] \"Larinni Malheiros\" \"CARLO ALESSANDRO MELO NOCE\" ..."
```

Como se pode perceber apenas um, professor que orientou doutorado concluído em 2017 (“Marilia Miranda Forte Gomes”) também fez parte dos professores que orientaram no mestrado: “Georges Daniel Amvame Nze”, “Jose Felicio da Silva”, “Lourdes Mattos Brasil”, “Marcelino Monteiro de Andrade”, “Marilia Miranda Forte Gomes”, “Ronni Geraldo Gomes de Amorim”. Sendo o professor “Marcus Vinícius Chaffim Costa” responsável apenas por orientações de doutorado em 2017.

Engenharia de Sistemas Eletrônicos e de Automação

```
###Descrição dos dados de orientações
```

```
unb.adv.auto <- fromJSON("dados-2018-2/engenharia-eletronicos-automacao/280.advise.json")
```

```
# Mostrando as listas presentes neste arquivo.
```

```
names(unb.adv.auto)
```

```
## [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"
```

```

# Explorando um nível de detalhe de Orientações de doutorados concluídas
names(unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO)

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

##### DOUTORADO 2017
#Buscando cursos que mais produziram doutorados.
head(sort(table(unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)

##
## ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO
##                                     5
## Engenharia de Sistemas Eletrônicos e de Automação
##                                     1
##                                     Informática
##                                     1

#Sabendo suas instituições
head(sort(table(unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$instituicao), decreasing = TRUE), 10)

## Universidade de Brasília
##                                     7

#Sabendo seus orientadores
data_orienM <- capture.output(str(unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$nome_orientadores))
unique(data_orienM)

## [1] "List of 7"
## [2] " $ : chr \"Adolfo Bauchspiess\""
## [3] " $ : chr \"Francisco Assis de Oliveira Nascimento\""
## [4] " $ : chr \"João Yoshiyuki Ishihara\""
## [5] " $ : chr \"Mylene Christine Queiroz de Farias\""
## [6] " $ : chr \"Renato Mariz de Moraes\""
## [7] " $ : chr \"Sandro Augusto Pavlik Haddad\""

#Sabendo os nomes dos trabalhos
data_orienM <- capture.output(str(unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$titulo))
unique(data_orienM)

## [1] " chr [1:7] \"Novas Metodologias AQM e TCP Visando a E ciência de Fluxos de Controle UDP e Dados"

##### MESTRADO 2017
#Buscando cursos que mais produziram mestrados.
head(sort(table(unb.adv.auto$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)

##
## ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO
##                                     12
## Engenharia de Sistemas Eletrônicos e de Automação
##                                     1

#Sabendo suas instituições
head(sort(table(unb.adv.auto$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$instituicao), decreasing = TRUE), 10)

## Universidade de Brasília
##                                     13

#Sabendo seus orientadores
data_orien3 <- capture.output(str(unb.adv.auto$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_orientadores))

```

```
unique(data_orien3)

## [1] "List of 13"
## [2] " $ : chr \"Antonio Padilha Lanari Bo\""
## [3] " $ : chr \"Flavia Maria Guerra de Sousa Aranha Oliveira\""
## [4] " $ : chr \"Janaina Goncalves Guimaraes\""
## [5] " $ : chr \"Marcelo Menezes de Carvalho\""
## [6] " $ : chr \"Mylene Christine Queiroz de Farias\""
## [7] " $ : chr \"Renato Alves Borges\""
## [8] " $ : chr \"Renato Mariz de Moraes\""
## [9] " $ : chr \"Sandro Augusto Pavlik Haddad\""

#Sabendo seus orientados
data_orien2 <- capture.output(str(unb.adv.auto$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_aluno))
unique(data_orien2)
```

```
## [1] " chr [1:13] \"Marien Cristina Narvaez Dorado\" \"Lucas Fernandes Aguiar\" ..."
```

Observando o resultado, pode-se perceber que foram realizados 7 doutorados e 13 mestrados no ano de 2017, todos realizados na Universidade de Brasília.

Engenharia Elétrica P8

```
###Descrição dos dados de orientações
unb.adv.ele_p8 <- fromJSON("dados-2018-2/engenharia-eletrica-p8/281.advise.json")
##### DOUTORADO 2017
#Buscando cursos que mais produziram doutorados.
head(sort(table(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)

##
## ENGENHARIA ELÉTRICA          Doutorado
##                2                1

#Sabendo suas instituições
head(sort(table(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$instituicao), decreasing = TRUE), 10)

##
##                                Universidade de Brasília
##                                2
## Instituto de Computação - Universidade Estadual de Campinas
##                                1

#Sabendo seus orientadores
data_orienM <- capture.output(str(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$nome_orientadores))
unique(data_orienM)

## [1] "List of 3"
## [2] " $ : chr \"Anderson de Rezende Rocha\""
## [3] " $ : chr \"Joao Paulo Carvalho Lustosa da Costa\""
## [4] " $ : chr \"Rafael Timóteo de Sousa Jr\""

#Sabendo os nomes dos trabalhos
data_orienM <- capture.output(str(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$titulo))
unique(data_orienM)

## [1] " chr [1:3] \"iSuper-resolution in low-quality videos for forensics, surveillance, and mobile ap"
```

```
##### MESTRADO 2017
#Buscando cursos que mais produziram mestrados.
head(sort(table(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso), decreasing = TRUE), 10)

##
##                               Engenharia Elétrica
##                               6
##                               ENGENHARIA ELÉTRICA
##                               3
## Electrical Engineering and Information Technology
##                               1
##                               Engenharia Biomédica
##                               1
##                               Mestrado
##                               1
##                               Sistemas Mecatrônicos
##                               1

#Sabendo suas instituições
head(sort(table(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$instituicao), decreasing = TRUE), 10)

##
##                               Universidade de Brasília
##                               11
## Instituto de Computação - Universidade Estadual de Campinas
##                               1
##                               Technische Universität Ilmenau
##                               1

#Sabendo seus orientadores
data_orien3 <- capture.output(str(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_orientadores))
unique(data_orien3)

## [1] "List of 13"
## [2] " $ : chr \"Anderson de Rezende Rocha\""
## [3] " $ : chr \"Daniel Guerreiro e Silva\""
## [4] " $ : chr \"Díbio Leandro Borges\""
## [5] " $ : chr \"Georges Daniel Amvame Nze\""
## [6] " $ : chr \"Joao Paulo Carvalho Lustosa da Costa\""
## [7] " $ : chr \"Rafael Timóteo de Sousa Jr\""

#Sabendo seus orientados
data_orien2 <- capture.output(str(unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_aluno))
unique(data_orien2)

## [1] " chr [1:13] \"Geise Kelly da Silva Santos\" ..."
```

Mais uma vez, os dados mais atualizados são de 2017. Há apenas 3 doutorados concluídos em 2017, sendo duas na UnB e uma em conjunto com o Instituto de Computação. Os orientadores são: “Anderson de Rezende Rocha”, “Joao Paulo Carvalho Lustosa da Costa”, “Rafael Timóteo de Sousa Jr”. Um dos trabalhos tem o título: “iSuper-resolution in low-quality videos for forensics, surveillance, and mobile applications”. Já no mestrado, podemos ver um grande problema da ciência de dados, o tratamento dos mesmos.

Engenharia Elétrica P7

```
###Descrição dos dados de orientações
unb.adv.ele_p7 <- fromJSON("dados-2018-2/engenharia-eletrica-p7/282.advise.json")
```

```
##### DOUTORADO 2017
#Buscando cursos que mais produziram doutorados.
head(sort(table(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$curso), decreasing = TRUE), 10)

## ENGENHARIA ELÉTRICA
##                2
#Sabendo suas instituições
head(sort(table(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$instituicao), decreasing = TRUE), 10)

## Universidade de Brasília
##                2
#Sabendo seus orientadores
data_orienM <- capture.output(str(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$nome_orientadores))
unique(data_orienM)

## [1] "List of 2"
## [2] " $ : chr \"Joao Paulo Carvalho Lustosa da Costa\""
## [3] " $ : chr \"Rafael Timóteo de Sousa Jr\""
#Sabendo os nomes dos trabalhos
data_orienM <- capture.output(str(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_DOUTORADO$`2017`$titulo))
unique(data_orienM)

## [1] " chr [1:2] \"Beamforming Techniques for Next Generation Communication Systems\" ..."
##### MESTRADO 2017
#Buscando cursos que mais produziram mestrados.
head(sort(table(toupper(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$curso)), decreasing = TRUE), 10)

##
##                ENGENHARIA ELÉTRICA
##                12
## ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY
##                1
## PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
##                1
head(sort(table(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$instituicao), decreasing = TRUE), 10)

##
##                Universidade de Brasília Technische Universität Ilmenau
##                13                1
#Sabendo seus orientadores
data_orien3 <- capture.output(str(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_orientadores))
unique(data_orien3)

## [1] "List of 14"
## [2] " $ : chr \"Andre Noll Barreto\""
## [3] " $ : chr \"Daniel Guerreiro e Silva\""
## [4] " $ : chr \"Francisco Damasceno Freitas\""
## [5] " $ : chr \"Joao Paulo Carvalho Lustosa da Costa\""
## [6] " $ : chr \"Kleber Melo e Silva\""
## [7] " $ : chr \"Rafael Timóteo de Sousa Jr\""
#Sabendo seus orientados
data_orien2 <- capture.output(str(unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_MESTRADO$`2017`$nome_aluno))
```

```
unique(data_orien2)
```

```
## [1] " chr [1:14] \"Sávio Oliveira de Oliveira Neves\" \"Lucas Sousa e Silva\" ..."
```

Há apenas três doutorados concluídos em 2017 e 14 mestados neste mesmo ano. Os professores Rafael Timóteo e João da Costa orientaram mestrados e doutorados concluídos em 2017. Houve um aluno da Universitat Ilmenau - Binghua Guo, “Low Complexity Schemes to Semi-Blind Joint Channel and Symbol Estimation in MIMO-OFDM” .

Descrição dos Dados presentes no list.json

Engenharia Biomédica

```
unb.list.biomed <- fromJSON("dados-2018-2/engenharia-biomedica/279.list.json")  
#analizando a quantidade de elementos presente em list.json  
length(unb.list.biomed$fiocruz$id)
```

```
## [1] 15
```

```
# Mostrando alguns id, verificar que eles são diferentes
```

```
d_listID <- unb.list.biomed$fiocruz$id  
unique(d_listID)
```

```
## [1] "1141716826787805" "0535100751136568" "5330755818114960"  
## [4] "5810353896294133" "2957228356035337" "0201204222182378"  
## [7] "9190489069187153" "4739013535126469" "4839052902231824"  
## [10] "1524924375222848" "9169095482512290" "4086384842130773"  
## [13] "5928104758017036" "1154673226500318" "7294738832905991"
```

```
#Mostrando nome sempre igual ""
```

```
d_listNO <- unb.list.biomed$fiocruz$nome  
unique(d_listNO)
```

```
## [1] ""
```

```
#Mostrando periodo sempre igual 2010-2017
```

```
d_listDT <- unb.list.biomed$fiocruz$periodo  
unique(d_listDT)
```

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

```
## [1] "2010" "2017"
```

Feito a análise, percebe-se que apenas o campo id que muda, tanto “**nome**” (obtendo sempre o valor “”) e “**periodo**”(obtendo sempre o valor[“2010”,“2017”]), por isso considera-se este arquivo JSON como inútil, haja visto que não há matemática para obter quaisquer dado plausível, somente os ids.

Engenharia de Sistemas Eletrônicos e de Automação

```
unb.list.auto <- fromJSON("dados-2018-2/engenharia-eletronicos-automacao/280.list.json")  
#analizando a quantidade de elementos presente em list.json  
length(unb.list.auto$fiocruz$id)
```

```
## [1] 21
```

```
# Mostrando alguns id, verificar que eles são diferentes
```

```
d_listID <- unb.list.auto$fiocruz$id  
unique(d_listID)
```

```
## [1] "0301021863146083" "1141716826787805" "5202540363896199"
```

```
## [4] "6959718976106778" "2065104260659969" "2443108673822686"
## [7] "4442238042952234" "3269842674142738" "6103264336506171"
## [10] "8850603018835560" "8032871065166582" "6509594616230305"
## [13] "5224996059295555" "1641340120936086" "0335058635088347"
## [16] "4465619366143200" "9986885788660516" "7284320756124645"
## [19] "9271885452744905" "7445354319753187" "3740757562716147"
```

```
#Mostrando nome sempre igual ""
d_listNO <- unb.list.auto$fiocruz$nome
unique(d_listNO)
```

```
## [1] ""
```

```
#Mostrando periodo sempre igual 2010-2017
d_listDT <- unb.list.auto$fiocruz$periodo
unique(d_listDT)
```

```
## [[1]]
## [1] "2010" "2017"
```

Novamente, considera-se este arquivo JSON como inútil, haja visto que não tem matemática para obter quaisquer dado plausível, somente os ids.

Engenharia Elétrica P8

```
unb.list.ele_p8 <- fromJSON("dados-2018-2/engenharia-eletrica-p8/281.list.json")
#analizando a quantidade de elementos presente em list.json
length(unb.list.ele_p8$fiocruz$id)
```

```
## [1] 11
```

```
# Mostrando alguns id, verificar que eles são diferentes
d_listID <- unb.list.ele_p8$fiocruz$id
unique(d_listID)
```

```
## [1] "4438697154081039" "1521815731111261" "5453814076903614"
## [4] "3636638127781726" "1808368824068028" "9788199690491510"
## [7] "2217050272348357" "5810353896294133" "1786889674911887"
## [10] "3196088341529197" "3408305339297459"
```

```
#Mostrando nome sempre igual ""
d_listNO <- unb.list.ele_p8$fiocruz$nome
unique(d_listNO)
```

```
## [1] ""
```

```
#Mostrando periodo sempre igual 2010-2017
d_listDT <- unb.list.ele_p8$fiocruz$periodo
unique(d_listDT)
```

```
## [[1]]
## [1] "2010" "2017"
```

Novamente, considera-se este arquivo JSON como inútil, haja visto que não tem matemática para obter quaisquer dado plausível, somente os ids.

Engenharia Elétrica P7

```
unb.list.ele_p7 <- fromJSON("dados-2018-2/engenharia-eletrica-p7/282.list.json")
#analizando a quantidade de elementos presente em list.json
length(unb.list.ele_p7$fiocruz$id)
```

```
## [1] 14
```

```
# Mostrando alguns id, verificar que eles são diferentes
```

```
d_listID <- unb.list.ele_p7$fiocruz$id
unique(d_listID)
```

```
## [1] "8999676470821349" "1573568521807396" "3636638127781726"
## [4] "3155646870666658" "7997341055520362" "1786889674911887"
## [7] "8795500242651581" "9279262545449206" "0997629063334821"
## [10] "8709856465061922" "9868233765380165" "3196088341529197"
## [13] "7293346445799114" "8701122785988842"
```

```
#Mostrando nome sempre igual ""
```

```
d_listNO <- unb.list.ele_p7$fiocruz$nome
unique(d_listNO)
```

```
## [1] ""
```

```
#Mostrando periodo sempre igual 2010-2017
```

```
d_listDT <- unb.list.ele_p7$fiocruz$periodo
unique(d_listDT)
```

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

```
## [1] "2010" "2017"
```

Novamente, considera-se este arquivo JSON como inútil, haja visto que não tem matemática para obter quaisquer dado plausível, somente os ids.

Descrição dos Dados de Produção bibliográfica

Engenharia Biomédica

```
unb.pub.biomed <- fromJSON("dados-2018-2/engenharia-biomedica/279.publication.json")
# Verificando os tipos de produções que existe.
names(unb.pub.biomed)
```

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

```
#Analizando o qual tipo de informação se tem em periódicos no ano 2012.
```

```
names(unb.pub.biomed$PERIODICO$`2012`)
```

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

```
#Nomes dos periódicos em que tiveram mais publicações na área de engenharia Biomédica.
```

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



```
##
## International Journal of Biosensors & Bioelectronics
## 2
## Revista Brasileira de Ensino de Física (Online)
## 2
## ACTA PAUL DE ENFERM
## 1
## ADOLESCÊNCIA & SAÚDE
## 1
## Advanced Materials Letters
## 1
## Advances in High Energy Physics
## 1
## AFRICAN JOURNAL OF BIOTECHNOLOGY
## 1
## ANNALS OF PHYSICS
## 1
## ARTEFACTUM (RIO DE JANEIRO)
## 1
## CADERNO DE FÍSICA DA UEFS
## 1
```

```
#Nomes dos periódicos em que tiveram mais publicações de artigos aceitos na área de engenharia Biomédica
head(sort(table(unb.pub.biomed$ARTIGO_ACEITO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
## Acta Paulista de Enfermagem
## 1
## Advanced Materials Letters
## 1
## Ciência e Saúde Coletiva (Impresso)
## 1
## IEEE Journal of Biomedical and Health Informatics
## 1
## IEEE Latin America Transactions
## 1
## JOURNAL OF NANOSCIENCE AND NANOTECHNOLOGY
## 1
```

```
#Nomes dos autores que produziram um tipo de Produção que não estava contemplada .
head(sort(table(unb.pub.biomed$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$`2011`$autores), decreasing = TRUE), 10)
```

```
##
##
##
## GOMES, Marília Miranda Forte;GOMES, MARÍLIA MIRANDA FORTE;GOMES, MARÍLIA M F;GOMES, Marília Miranda Forte
##
##
##
##
##
##
##
```

Engenharia de Sistemas Eletrônicos e de Automação

```
unb.pub.auto <- fromJSON("dados-2018-2/engenharia-eletronicos-automacao/280.publication.json")
# Verificando os tipos de produções que existe.
names(unb.pub.auto)
```

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

```
#Analisando o qual tipo de informação se tem em periódicos no ano 2012.
```

```
names(unb.pub.auto$PERIODICO$`2012`)
```

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

```
#Nomes dos periódicos em que tiveram mais publicações na área de engenharia Biomédica.
```

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

```
##
##               IEEE Journal of Biomedical and Health Informatics
##                                     2
##               IEEE TRANSACTIONS ON IMAGE PROCESSING
##                                     2
##               JOURNAL OF MOLECULAR MODELING
##                                     2
## JOURNAL OF THE FRANKLIN INSTITUTE-ENGINEERING AND APPLIED MATHEMATICS
##                                     2
##               ACTA PAUL DE ENFERM
##                                     1
##               ADVANCES IN SPACE RESEARCH
##                                     1
##               AMERICAN JOURNAL OF PHYSICAL MEDICINE & REHABILITATION
##                                     1
##               ARTIFICIAL ORGANS
##                                     1
##               European Journal of Translational Myology
##                                     1
##               IEEE Journal of the Electron Devices Society
##                                     1
```

```
#Nomes dos periódicos em que tiveram mais publicações de artigos aceitos na área de engenharia Biomédica.
```

```
head(sort(table(unb.pub.auto$ARTIGO_ACEITO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
## IEEE TRANSACTIONS ON AUTOMATIC CONTROL
##                                     1
##               INTERNATIONAL JOURNAL OF CONTROL
##                                     1
```

```
#Nomes dos autores que produziram um tipo de Produção que não estava contemplada .
head(sort(table(unb.pub.auto$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$`2011`$autores), decreasing = TRUE))
```

```
## integer(0)
```

Engenharia Elétrica P8

```
unb.pub.ele_p8 <- fromJSON("dados-2018-2/engenharia-eletrica-p8/281.publication.json")
# Verificando os tipos de produções que existe.
names(unb.pub.ele_p8)
```

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

```
#Analisando o qual tipo de informação se tem em periódicos no ano 2012.
names(unb.pub.ele_p8$PERIODICO$`2012`)
```

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

```
#Nomes dos periódicos em que tiveram mais publicações na área de engenharia Biomédica.
head(sort(table(unb.pub.ele_p8$PERIODICO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
##                                     IEEE Access
##                                     2
##                               IEEE Transactions on Information Forensics and Security
##                                     2
##                               JOURNAL OF MOLECULAR MODELING
##                                     2
##                               Applied Sciences-Basel
##                                     1
##                               CLEI ELECTRONIC JOURNAL
##                                     1
##                               Cluster Computing
##                                     1
## Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization
##                                     1
##                               Concurrency and Computation
##                                     1
##                               Digital Signal Processing (Print)
##                                     1
##                               IEEE Journal of Biomedical and Health Informatics
##                                     1
```

```
#Nomes dos periódicos em que tiveram mais publicações de artigos aceitos na área de engenharia Biomédica.
head(sort(table(unb.pub.ele_p8$ARTIGO_ACEITO$`2017`$periodico), decreasing = TRUE), 10)
```

```
## IEEE TRANSACTIONS ON INFORMATION THEORY
##                                     1
```

```
#Nomes dos autores que produziram um tipo de Produção que não estava contemplada em 2011 .
head(sort(table(unb.pub.ele_p8$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$`2011`$autores), decreasing = TRUE), 10)
```

```
## ARANHA, D. F.;ARANHA, DIEGO F.
## 1
```

Engenharia Elétrica P7

```
unb.pub.ele_p7 <- fromJSON("dados-2018-2/engenharia-eletrica-p7/282.publication.json")
# Verificando os tipos de produções que existe.
names(unb.pub.ele_p7)
```

```
## [1] "PERIODICO"
## [2] "LIVRO"
## [3] "CAPITULO_DE_LIVRO"
## [4] "TEXT0_EM_JORNAIS"
## [5] "EVENTO"
## [6] "ARTIGO_ACEITO"
## [7] "DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA"
```

```
#Analisando o qual tipo de informação se tem em periódicos no ano 2012.
names(unb.pub.ele_p7$PERIODICO$`2012`)
```

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

```
#Nomes dos periódicos em que tiveram mais publicações na área de engenharia Biomédica.
head(sort(table(unb.pub.ele_p7$PERIODICO$`2017`$periodico), decreasing = TRUE), 10)
```

```
##
## JOURNAL OF MOLECULAR MODELING
## 2
## Applied Sciences-Basel
## 1
## CLEI ELECTRONIC JOURNAL
## 1
## Cluster Computing
## 1
## Digital Signal Processing (Print)
## 1
## IEEE Access
## 1
## IEEE ANTENNAS AND PROPAGATION MAGAZINE
## 1
## IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION
## 1
## IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT
## 1
## IEEE Transactions on Power Delivery
## 1
```

```
#Nomes dos periódicos em que tiveram mais publicações de artigos aceitos na área de engenharia Biomédica.
head(sort(table(unb.pub.ele_p7$ARTIGO_ACEITO$`2017`$periodico), decreasing = TRUE), 10)
```

```
## IEEE TRANSACTIONS ON COMMUNICATIONS
```

```
## 1
#Nomes dos autores que produziram um tipo de Produção que não estava contemplada em 2014.
head(sort(table(unb.pub.ele_p7$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$`2011`$autores), decreasing = TRUE))
```

```
## integer(0)
```

Descrição dos Dados de Redes de Colaboração

Engenharia Biomédica

```
unb.graph.biomed <- fromJSON("dados-2018-2/engenharia-biomedica/279.graph.json")
# Suas variáveis
names(unb.graph.biomed)
```

```
## [1] "label" "nodes" "links"
```

```
# Quantidade de nós
```

```
length(unb.graph.biomed$nodes$id)
```

```
## [1] 15
```

```
# Quantidade de links de fonte
```

```
length(unb.graph.biomed$links$source)
```

```
## [1] 28
```

```
# Quantidade de links de chegada
```

```
length(unb.graph.biomed$links$target)
```

```
## [1] 28
```

```
#exemplos de pesos da aresta
```

```
str(unb.graph.biomed$links$weight)
```

```
## chr [1:28] "1" "3" "9" "1" "2" "3" "1" "1" "5" "15" "1" "3" "9" "1" ...
```

Engenharia de Sistemas Eletrônicos e de Automação

```
unb.graph.auto <- fromJSON("dados-2018-2/engenharia-eletronicos-automacao/280.graph.json")
# Suas variáveis
names(unb.graph.auto)
```

```
## [1] "label" "nodes" "links"
```

```
# Quantidade de nós
```

```
length(unb.graph.auto$nodes$id)
```

```
## [1] 21
```

```
# Quantidade de links de fonte
```

```
length(unb.graph.auto$links$source)
```

```
## [1] 21
```

```
# Quantidade de links de chegada
```

```
length(unb.graph.auto$links$target)
```

```
## [1] 21
```

```
#exemplos de pesos da aresta
```

```
str(unb.graph.auto$links$weight)
```

```
## chr [1:21] "1" "6" "7" "6" "3" "2" "4" "1" "2" "2" "1" "1" "2" "1" ...
```

Engenharia Elétrica P8

```
unb.graph.ele_p8 <- fromJSON("dados-2018-2/engenharia-eletrica-p8/281.graph.json")  
# Suas variáveis  
names(unb.graph.ele_p8)
```

```
## [1] "label" "nodes" "links"
```

```
# Quantidade de nós
```

```
length(unb.graph.ele_p8$nodes$id)
```

```
## [1] 11
```

```
# Quantidade de links de fonte
```

```
length(unb.graph.ele_p8$links$source)
```

```
## [1] 6
```

```
# Quantidade de links de chegada
```

```
length(unb.graph.ele_p8$links$target)
```

```
## [1] 6
```

```
#exemplos de pesos da aresta
```

```
str(unb.graph.ele_p8$links$weight)
```

```
## chr [1:6] "27" "2" "2" "3" "4" "1"
```

Engenharia Elétrica P7

```
unb.graph.ele_p7 <- fromJSON("dados-2018-2/engenharia-eletrica-p7/282.graph.json")  
# Suas variáveis  
names(unb.graph.ele_p7)
```

```
## [1] "label" "nodes" "links"
```

```
# Quantidade de nós
```

```
length(unb.graph.ele_p7$nodes$id)
```

```
## [1] 14
```

```
# Quantidade de links de fonte
```

```
length(unb.graph.ele_p7$links$source)
```

```
## [1] 5
```

```
# Quantidade de links de chegada
```

```
length(unb.graph.ele_p7$links$target)
```

```
## [1] 5
```

```
#exemplos de pesos da aresta
```

```
str(unb.graph.ele_p7$links$weight)
```

```
## chr [1:5] "27" "2" "13" "7" "1"
```

V.3 - Análise Exploratória dos Dados

Arquivo Profile

Engenharia Biomédica

```
## PUBLICAÇÕES
# Total de áreas de atuação de todos profissionais
sum(sapply(unb.prof.biomed, function(x) nrow(x$areas_de_atuacao)))

## [1] 57

# Número de pessoas por grande area
table(unlist(sapply(unb.prof.biomed, function(x) (x$areas_de_atuacao$grande_area))))

##
##                                CIENCIAS_BIOLOGICAS
##                                1                      4
##          CIENCIAS_DA_SAUDE CIENCIAS_EXATAS_E_DA_TERRA
##                                7                      9
## CIENCIAS_SOCIAIS_APLICADAS                                ENGENHARIAS
##                                1                      34
##                                OUTROS
##                                1

# Número de pessoas que produziram os tipos de Produção específico
table(unlist(sapply(unb.prof.biomed, function(x) names(x$producao_bibliografica))))

##
##                                ARTIGO_ACEITO
##                                7
##                                CAPITULO_DE_LIVRO
##                                9
## DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA
##                                1
##                                EVENTO
##                                14
##                                LIVRO
##                                3
##                                PERIODICO
##                                15
##                                TEXTO_EM_JORNAIS
##                                1

# Número de publicações por tipo
#### ARTIGO ACEITO
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano)))

## [1] 7

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
## 0 1
## 8 7

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$ARTIGO_ACEITO$ano))))
```

```
##
## 2015 2017
##    1    6

#### CAPITULO DE LIVRO
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano)))

## [1] 47

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
##  0  1  3  5  9 17
##  6  4  1  1  2  1

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    1    3    8    2    2   21    5    5

#### LIVRO
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$LIVRO$ano)))

## [1] 3

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$LIVRO$ano))))

##
##  0  1
## 12  3

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$LIVRO$ano))))

##
## 2015 2016 2017
##    1    1    1

#### PERIÓDICO
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$PERIODICO$ano)))

## [1] 172

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$PERIODICO$ano))))

##
##  1  2  3  4  6  8  9 10 11 13 16 22 29 36
##  1  2  1  1  1  1  1  1  1  1  1  1  1  1

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$PERIODICO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   14    9   10   29   18   39   21   32
```



```

#### TEXTO EM JORNAIS
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano)))

## [1] 1

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 0 1
## 14 1

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 2010
## 1

#### DEMAIS TIPOS
sum(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG)))

## [1] 1

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG))))

##
## 0 1
## 14 1

# Número de produções por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG$ano))))

##
## 2011
## 1

## ORIENTAÇÕES
# Número de pessoas que realizaram diferentes tipos de orientações
length(unlist(sapply(unb.prof.biomed, function(x) names(x$orientacoes_academicas))))

## [1] 60

# Número de pessoas por tipo de orientação
table(unlist(sapply(unb.prof.biomed, function(x) names(x$orientacoes_academicas))))

##
## ORIENTACAO_CONCLUIDA_DOUTORADO
## 6
## ORIENTACAO_CONCLUIDA_MESTRADO
## 13
## ORIENTACAO_CONCLUIDA_POS_DOUTORADO
## 1
## ORIENTACAO_EM_ANDAMENTO_DOUTORADO
## 2
## ORIENTACAO_EM_ANDAMENTO_GRADUACAO
## 3
## ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA

```

```

##                                7
##          ORIENTACAO_EM_ANDAMENTO_MESTRADO
##                                13
##          OUTRAS_ORIENTACOES_CONCLUIDAS
##                                15

# Número de publicações por tipo
#### MESTRADO
# Número de orientações concluídas
sum(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$ar

## [1] 109

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_M

##
##  0  1  3  6  7  9 11 18 19
##  2  1  2  1  3  3  1  1  1

# Número de orientações por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRAD

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    7    4   10   12   23   24   13   16

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_MESTR

##
## 2013 2014 2015 2016 2017
##    1    2    6   12   37

#### DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$ar

## [1] 14

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_D

##
## 0 1 2 3 4
## 9 2 1 2 1

# Número de orientações por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORAI

##
## 2011 2012 2013 2014 2015 2017
##    2    2    6    1    1    2

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_DOUT

##
## 2015 2017
##    3    1

```

```
#### POS-DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)))

## [1] 2

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.biomed, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))))

##
## 0 2
## 14 1

# Número de orientações por ano
table(unlist(sapply(unb.prof.biomed, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))))

##
## 2014 2017
## 1 1
```

Pelos dados, a maioria dos profissionais atuam na área de engenharia (34). A maioria da produção científica é dada pela publicações em periódicos. Um fato interessante é que em 2015 houve um grande número de capítulos de livro publicado, fugindo do padrão de 4. Há apenas um texto em jornal publicado em 2010. Houve apenas dois pós-doutorado feito em biomédica na UnB em 2014 e 2017.

Engenharia de Sistemas Eletrônicos e de Automação

```
## PUBLICAÇÕES
# Total de áreas de atuação de todos profissionais
sum(sapply(unb.prof.auto, function(x) nrow(x$areas_de_atuacao)))

## [1] 68

# Número de pessoas por grande area
table(unlist(sapply(unb.prof.auto, function(x) (x$areas_de_atuacao$grande_area))))

##
##          CIENCIAS_DA_SAUDE CIENCIAS_EXATAS_E_DA_TERRA
##                1                8
##          ENGENHARIAS                OUTROS
##                57                2

# Número de pessoas que produziram os tipos de Produção específico
table(unlist(sapply(unb.prof.auto, function(x) names(x$producao_bibliografica))))

##
##          ARTIGO_ACEITO CAPITULO_DE_LIVRO          EVENTO          LIVRO
##                4                10                20                6
##          PERIODICO  TEXTO_EM_JORNAIS
##                21                2

# Número de publicações por tipo
#### ARTIGO ACEITO
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))

## [1] 7

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))
```

```
##
## 0 1 4
## 17 3 1

# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
## 2015 2016 2017
## 3 2 2

#### CAPITULO DE LIVRO
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

## [1] 28

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 0 1 2 3 6 9
## 11 5 1 2 1 1

# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
## 7 6 6 2 3 2 1 1

#### LIVRO
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$LIVRO$ano))))

## [1] 6

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$LIVRO$ano))))

##
## 0 1
## 15 6

# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$LIVRO$ano))))

##
## 2010 2011 2014 2015
## 1 2 1 2

#### PERIÓDICO
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$PERIODICO$ano))))

## [1] 226

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$PERIODICO$ano))))

##
## 2 4 5 6 7 9 11 12 13 17 18 22 25
## 1 2 1 3 1 3 1 3 1 2 1 1 1
```

```

# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$PERIODICO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   24   15   18   30   28   36   37   38
#### TEXTO EM JORNAIS
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

## [1] 2
# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
##   0   1
## 19   2
# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 2014
##    2
#### DEMAIS TIPOS
sum(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$ano))))

## [1] 0
# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA$ano))))

##
##   0
## 21
# Número de produções por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO$ano))))

## < table of extent 0 >
## ORIENTAÇÕES
# Número de pessoas que realizaram diferentes tipos de orientações
length(unlist(sapply(unb.prof.auto, function(x) names(x$orientacoes_academicas))))

## [1] 99
# Número de pessoas por tipo de orientação
table(unlist(sapply(unb.prof.auto, function(x) names(x$orientacoes_academicas))))

##
##          ORIENTACAO_CONCLUIDA_DOUTORADO
##                                13
##          ORIENTACAO_CONCLUIDA_MESTRADO
##                                19
##          ORIENTACAO_CONCLUIDA_POS_DOUTORADO
##                                3
##          ORIENTACAO_EM_ANDAMENTO_DOUTORADO

```

```

##                                15
##          ORIENTACAO_EM_ANDAMENTO_GRADUACAO
##                                6
## ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA
##                                9
##          ORIENTACAO_EM_ANDAMENTO_MESTRADO
##                                14
##          OUTRAS_ORIENTACOES_CONCLUIDAS
##                                20

# Número de publicações por tipo
#### MESTRADO
# Número de orientações concluídas
sum(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$ano))

## [1] 107

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$ano))

##
##  0  1  2  4  5  6  7  8  9 11
##  2  2  2  2  3  3  1  3  2  1

# Número de orientações por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$ano))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    9   13   12   16   22    9   13   13

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_MESTRADO$ano))

##
## 2014 2015 2016 2017
##    1    5   14   13

#### DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$ano))

## [1] 43

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$ano))

##
## 0 1 2 3 4 5 8
## 8 4 2 2 2 1 2

# Número de orientações por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$ano))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    3    3    5    4    5    9    7    7

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_DOUTORADO$ano))

```

```
##
## 2012 2013 2014 2015 2016 2017
##    1    3    8   10   14   10

#### POS-DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)))

## [1] 7

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.auto, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))), 0:4)

##
##  0  1  2  4
## 18  1  1  1

# Número de orientações por ano
table(unlist(sapply(unb.prof.auto, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))), 2010:2017)

##
## 2010 2013 2014 2015 2016 2017
##    1    1    1    1    2    1

todo
```

Engenharia Elétrica P8

```
## PUBLICAÇÕES
# Total de áreas de atuação de todos profissionais
sum(sapply(unb.prof.ele_p8, function(x) nrow(x$areas_de_atuacao)))

## [1] 45

# Número de pessoas por grande area
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$areas_de_atuacao$grande_area))))

##
## CIENCIAS_EXATAS_E_DA_TERRA          CIENCIAS_HUMANAS
##                      20                      1
##                      ENGENHARIAS
##                      24

# Número de pessoas que produziram os tipos de Produção específico
table(unlist(sapply(unb.prof.ele_p8, function(x) names(x$producao_bibliografica))))

##
##                      ARTIGO_ACEITO
##                      1
##                      CAPITULO_DE_LIVRO
##                      7
## DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA
##                      1
##                      EVENTO
##                      10
##                      LIVRO
##                      6
##                      PERIODICO
##                      10
```

```

##                                TEXTO_EM_JORNAIS
##                                3

# Número de publicações por tipo
#### ARTIGO ACEITO
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano)))

## [1] 1

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
## 0 1
## 10 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
## 2017
## 1

#### CAPITULO DE LIVRO
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano)))

## [1] 26

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 0 1 4 5 6
## 4 2 2 2 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
## 2 5 4 2 2 5 4 2

#### LIVRO
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$LIVRO$ano)))

## [1] 20

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$LIVRO$ano))))

##
## 0 1 2 3 4 9
## 5 2 1 1 1 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$LIVRO$ano))))

##
## 2010 2011 2013 2014 2015 2016 2017
## 1 3 7 2 2 3 2

```



```
#### PERIÓDICO
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$PERIODICO$ano)))

## [1] 179

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$PERIODICO$ano))))

##
## 0 1 2 3 9 16 28 45 63
## 1 1 1 2 2 1 1 1 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$PERIODICO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
## 8 18 19 18 18 21 39 38

#### TEXTO EM JORNAIS
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano)))

## [1] 14

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 0 2 4 8
## 8 1 1 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 2011 2012 2013 2015 2017
## 3 6 1 2 2

#### DEMAIS TIPOS
sum(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG)))

## [1] 3

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG))))

##
## 0 3
## 10 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG$ano))))

##
## 2011 2012
## 1 2

## ORIENTAÇÕES
# Número de pessoas que realizaram diferentes tipos de orientações
length(unlist(sapply(unb.prof.ele_p8, function(x) names(x$orientacoes_academicas))))
```

```
## [1] 45
```

```
# Número de pessoas por tipo de orientação
table(unlist(sapply(unb.prof.ele_p8, function(x) names(x$orientacoes_academicas))))
```

```
##
##          ORIENTACAO_CONCLUIDA_DOUTORADO
##                                5
##          ORIENTACAO_CONCLUIDA_MESTRADO
##                                10
##          ORIENTACAO_CONCLUIDA_POS_DOUTORADO
##                                3
##          ORIENTACAO_EM_ANDAMENTO_DOUTORADO
##                                5
##          ORIENTACAO_EM_ANDAMENTO_GRADUACAO
##                                2
## ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA
##                                5
##          ORIENTACAO_EM_ANDAMENTO_MESTRADO
##                                7
##          OUTRAS_ORIENTACOES_CONCLUIDAS
##                                8
```

```
# Número de publicações por tipo
```

```
#### MESTRADO
```

```
# Número de orientações concluídas
```

```
sum(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$MESTRADO)))
```

```
## [1] 108
```

```
# Número de pessoas por quantitativo de orientações por pessoa
```

```
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$MESTRADO))))
```

```
##
##  0  4  5  6  7 13 15 20 28
##  1  2  1  2  1  1  1  1  1
```

```
# Número de orientações por ano
```

```
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$MESTRADO))))
```

```
##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    9   10   26   12   11    5   22   13
```

```
# Número de orientações não concluídas por ano
```

```
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_MESTRADO$MESTRADO))))
```

```
##
## 2015 2016 2017
##    2    4   12
```

```
#### DOUTORADO
```

```
# Número de orientações concluídas
```

```
sum(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$DOUTORADO)))
```

```
## [1] 17
```

```
# Número de pessoas por quantitativo de orientações por pessoa
```

```
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$DOUTORADO))))
```

```
##
## 0 1 2 5 8
## 6 2 1 1 1

# Número de orientações por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO))),
      ##
      ## 2012 2013 2014 2015 2016 2017
      ##    3    3    3    2    3    3

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_DOUTORADO))),
      ##
      ## 2013 2014 2015 2016 2017
      ##    1    5    5    6   15

#### POS-DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))),
    ## [1] 12

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.ele_p8, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))),
      ##
      ## 0 1 10
      ## 8 2 1

# Número de orientações por ano
table(unlist(sapply(unb.prof.ele_p8, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO))),
      ##
      ## 2012 2013 2014 2015 2016 2017
      ##    1    1    1    4    3    2
```

Em Engenharia Elétrica P8, tem uma boa divisão das grandes áreas dos orientadores, 20 para ciências exatas e da terra e 24 para engenharias, que é maioria de todas as pesquisas feitas. Houve 108 orientações concluídas de mestrado e apenas 17 doutorado, que não é tão diferente de pós-doutorados: 12. Há um doutorado de 2013 que não foi concluído ainda.

Engenharia Elétrica P7

```
## PUBLICAÇÕES
# Total de áreas de atuação de todos profissionais
sum(sapply(unb.prof.ele_p7, function(x) nrow(x$areas_de_atuacao)))

## [1] 55

# Número de pessoas por grande area
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$areas_de_atuacao$grande_area))))

##
##                CIENCIAS_EXATAS_E_DA_TERRA
##                1                15
##      CIENCIAS_HUMANAS                ENGENHARIAS
##                1                38
```

```

# Número de pessoas que produziram os tipos de Produção específico
table(unlist(sapply(unb.prof.ele_p7, function(x) names(x$producao_bibliografica))))

##
##      ARTIGO_ACEITO  CAPITULO_DE_LIVRO      EVENTO      LIVRO
##              1              8              14              4
##      PERIODICO  TEXTO_EM_JORNAIS
##              14              4

# Número de publicações por tipo
#### ARTIGO ACEITO
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))

## [1] 1

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
##  0  1
## 13  1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$ARTIGO_ACEITO$ano))))

##
## 2017
##    1

#### CAPITULO DE LIVRO
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

## [1] 30

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 0 1 4 5 6 8
## 6 3 2 1 1 1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$CAPITULO_DE_LIVRO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   10    2    5    3    3    4    2    1

#### LIVRO
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$LIVRO$ano))))

## [1] 14

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$LIVRO$ano))))

##
##  0  1  2  9
## 10  1  2  1

```

```

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$LIVRO$ano))))

##
## 2010 2011 2013 2014 2015 2017
##    4    3    3    1    1    2

#### PERIÓDICO
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$PERIODICO$ano)))

## [1] 178

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$PERIODICO$ano))))

##
##  3  5  7  8  9 10 13 17 28 45
##  1  1  1  4  2  1  1  1  1  1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$PERIODICO$ano))))

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   15   18   15   23   16   17   43   31

#### TEXTO EM JORNAIS
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano)))

## [1] 13

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
##  0  1  3  8
## 10  2  1  1

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$TEXTO_EM_JORNAIS$ano))))

##
## 2010 2011 2012 2013 2014 2015
##    1    2    7    1    1    1

#### DEMAIS TIPOS
sum(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG)))

## [1] 0

# Número de pessoas por quantitativo de produções por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG))))

##
##  0
## 14

# Número de produções por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$producao_bibliografica$DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOG$ano))))

## < table of extent 0 >

```

```
## ORIENTAÇÕES
# Número de pessoas que realizaram diferentes tipos de orientações
length(unlist(sapply(unb.prof.ele_p7, function(x) names(x$orientacoes_academicas))))
```

```
## [1] 66
```

```
# Número de pessoas por tipo de orientação
table(unlist(sapply(unb.prof.ele_p7, function(x) names(x$orientacoes_academicas))))
```

```
##
##          ORIENTACAO_CONCLUIDA_DOUTORADO
##                               8
##          ORIENTACAO_CONCLUIDA_MESTRADO
##                               13
##          ORIENTACAO_EM_ANDAMENTO_DOUTORADO
##                               10
##          ORIENTACAO_EM_ANDAMENTO_GRADUACAO
##                               5
## ORIENTACAO_EM_ANDAMENTO_INICIACAO_CIENTIFICA
##                               7
##          ORIENTACAO_EM_ANDAMENTO_MESTRADO
##                               11
##          OUTRAS_ORIENTACOES_CONCLUIDAS
##                               12
```

```
# Número de publicações por tipo
```

```
#### MESTRADO
```

```
# Número de orientações concluídas
```

```
sum(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRADO$ar
```

```
## [1] 140
```

```
# Número de pessoas por quantitativo de orientações por pessoa
```

```
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_M
```

```
##
##  0  1  4  5  6  8 10 12 14 15 20 28
##  1  1  1  2  1  1  1  2  1  1  1  1
```

```
# Número de orientações por ano
```

```
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_MESTRAD
```

```
##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    6   16   25   19   18   16   26   14
```

```
# Número de orientações não concluídas por ano
```

```
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_MESTR
```

```
##
## 2015 2016 2017
##    4    9   22
```

```
#### DOUTORADO
```

```
# Número de orientações concluídas
```

```
sum(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO$ar
```

```
## [1] 22
```

```

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO)),
##
## 0 1 2 3 5 6
## 6 2 3 1 1 1

# Número de orientações por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_DOUTORADO)),
##
## 2010 2012 2013 2014 2015 2016 2017
##      4    3    5    2    4    2    2

# Número de orientações não concluídas por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$orientacoes_academicas$ORIENTACAO_EM_ANDAMENTO_DOUTORADO)),
##
## 2013 2014 2015 2016 2017
##      2    5    7    7    15

#### POS-DOUTORADO
# Número de orientações concluídas
sum(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)),
## [1] 0

# Número de pessoas por quantitativo de orientações por pessoa
table(unlist(sapply(unb.prof.ele_p7, function(x) length(x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)),
##
## 0
## 14

# Número de orientações por ano
table(unlist(sapply(unb.prof.ele_p7, function(x) (x$orientacoes_academicas$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)),
## < table of extent 0 >

```

Por fim, na área de Engenharia Elétrica P7, vê-se, novamente, um domínio da grande área de engenharias, como era de se esperar. Em 2016, é possível observar que se teve o maior numero de publicações em periódicos, 42 no total.

Sobre as orientações, dos temas desta pesquisa, foi a que teve mais equilíbrio entre o número de profissionais de mestrados e os que trabalham com doutorados concluídos, tendo uma diferença de 5. Especificamente, sobre as orientações vê-se um enorme número de mestrados concluídos desde 2011, com valor baixo em 2010 (6 no total, número similar a de doutorados feitos: 4, naquele mesmo ano). Um dado que pode ser alarmante é que ainda existe têm dois doutorados iniciados em 2013 que não estão concluídos ainda, estando em seu último ano, pois o tempo máximo do mestrado é de 5 anos, já do mestrado há 4 trabalhos que estão no último ano possível.

Arquivo Publication

Engenharia Biomédica

```

# Criando um data-frame com todos os anos
unb.pub.biomed.df <- data.frame()
for (i in 1:length(unb.pub.biomed[[1]]))
  unb.pub.biomed.df <- rbind(unb.pub.biomed.df, unb.pub.biomed$PERIODICO[[i]])

```

```

# Limpando o data-frame de listas
unb.pub.biomed.df$autores <- gsub("\\", "\\|\\", "\\|", ";", unb.pub.biomed.df$autores)
unb.pub.biomed.df$autores <- gsub("\\|c\\(|\\)", "", unb.pub.biomed.df$autores)
unb.pub.biomed.df$`autores-endogeno` <- gsub(",", ";", unb.pub.biomed.df$`autores-endogeno`)
unb.pub.biomed.df$`autores-endogeno` <- gsub("\\|c\\(|\\)", "", unb.pub.biomed.df$`autores-endogeno`)
#### ARQUIVO PROCESSADO E LIMPO
## Publicações por natureza
table(unb.pub.biomed.df$natureza)

##
## COMPLETO
##      147

## Publicações 2010 até 2017
table(unb.pub.biomed.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   12    8    9   27   14   31   20   26

## Publicações por periódico, sendo os 5 periódicos com masi publicação
d <- table(unb.pub.biomed.df$periodico)
head(sort(d,decreasing = TRUE),n=5)

##
##          Revista Brasileira de Engenharia Biomédica (Impresso)
##                                                                7
## Global Journal of Engineering Science and Research Management
##                                                                6
##          Biomedical Engineering Online (Online)
##                                                                4
##          Revista Brasileira de Ensino de Física (Online)
##                                                                4
##          Revista Brasileira de Inovação Tecnológica em Saúde
##                                                                4

# Autores que tiveram mais publicações
head(sort(table(toupper(unlist(strsplit(unb.pub.biomed.df$autores,";")))),decreasing = TRUE),n=10)

##
##          AMORIM, R. G. G.
##          17
##          AMORIM, R. G. G.
##          16
##          MARÃES, V. R. F. S.
##          14
##          BRASIL, L. M.
##          7
##          MARÃES, V. R. F. S.
##          7
##          SILVA, W. B.
##          7
##          GOMES, MARÍLIA MIRANDA FORTE
##          7
## ROSA, SUÉLIA DE SIQUEIRA RODRIGUES FLEURY
##          7

```



```
##                      DA ROCHA, A. F.
##                      6
##          GOMES, MARÍLIA MIRANDA FORTE
##                      6
```

Há 147 publicações completas, sendo a maioria começando a ocorrer a partir de 2013. A maioria da publicação Revista Brasileira de Engenharia Biomédica (Impresso), 7, e Global Journal of Engineering Science and Research Management, 6 publicações. Amorim é o autor com mais publicações.

Engenharia de Sistemas Eletrônicos e de Automação

```
# Criando um data-frame com todos os anos
unb.pub.auto.df <- data.frame()
for (i in 1:length(unb.pub.auto[[1]]))
  unb.pub.auto.df <- rbind(unb.pub.auto.df, unb.pub.auto$PERIODICO[[i]])
# Limpando o data-frame de listas
unb.pub.auto.df$autores <- gsub("\\", "\\|\\", "\\|", "; ", unb.pub.auto.df$autores)
unb.pub.auto.df$autores <- gsub("\\|c\\|(\\|\\|)", "", unb.pub.auto.df$autores)
unb.pub.auto.df$`autores-endogeno` <- gsub(",", ";", unb.pub.auto.df$`autores-endogeno`)
unb.pub.auto.df$`autores-endogeno` <- gsub("\\|c\\|(\\|\\|)", "", unb.pub.auto.df$`autores-endogeno`)
#### ARQUIVO PROCESSADO E LIMPO
## Publicações por natureza
table(unb.pub.auto.df$natureza)
```

```
##
## COMPLETO
##      206
```

```
## Publicações 2010 até 2017
table(unb.pub.auto.df$ano)
```

```
##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   20   14   18   25   28   31   36   34
```

```
## Publicações por periódico, sendo os 5 periódicos com masi publicação
d <- table(unb.pub.auto.df$periodico)
head(sort(d,decreasing = TRUE),n=5)
```

```
##
## Revista Brasileira de Engenharia Biomédica (Impresso)
##                      7
##          IEEE Signal Processing Letters
##                      5
## Journal of Control, Automation and Electrical Systems
##                      5
##                      Electronics Letters
##                      4
## INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
##                      4
```

```
# Autores que tiveram mais publicações
head(sort(table(toupper(unlist(strsplit(unb.pub.auto.df$autores,";")))),decreasing = TRUE),n=10)
```

```
##
##                      BOTTARO, MARTIM
##                      12
```

```
##                ISHIHARA, J. Y.
##                12
##                DE QUEIROZ, RICARDO
##                12
##                DE QUEIROZ, RICARDO L.
##                11
##                CLAUS, MARTIN
##                10
##                PERES, P. L. D.
##                10
##                QUEIROZ, R.L.
##                10
##                DE QUEIROZ, R. L.
##                9
##                ISHIHARA, JOÃO Y.
##                9
##  NASCIMENTO, FRANCISCO ASSIS DE OLIVEIRA
##                9
```

todo

Engenharia Elétrica P8

```
# Criando um data-frame com todos os anos
unb.pub.ele_p8.df <- data.frame()
for (i in 1:length(unb.pub.auto[[1]]))
  unb.pub.ele_p8.df <- rbind(unb.pub.ele_p8.df, unb.pub.ele_p8$PERIODICO[[i]])
# Limpando o data-frame de listas
unb.pub.ele_p8.df$autores <- gsub("\\", "\\|", unb.pub.ele_p8.df$autores)
unb.pub.ele_p8.df$autores <- gsub("\\|c\\(|\\)", "", unb.pub.ele_p8.df$autores)
unb.pub.ele_p8.df$`autores-endogeno` <- gsub(",", ";", unb.pub.ele_p8.df$`autores-endogeno`)
unb.pub.ele_p8.df$`autores-endogeno` <- gsub("\\|c\\(|\\)", "", unb.pub.ele_p8.df$`autores-endogeno`)
#### ARQUIVO PROCESSADO E LIMPO
## Publicações por natureza
table(unb.pub.ele_p8.df$natureza)

##
## COMPLETO
##      168

## Publicações 2010 até 2017
table(unb.pub.ele_p8.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    8   15   17   17   18   21   36   36

## Publicações por periódico, sendo os 5 periódicos com masi publicação
d <- table(unb.pub.ele_p8.df$periodico)
head(sort(d,decreasing = TRUE),n=5)

##
##      IEEE Transactions on Information Forensics and Security
##                                14
##      Digital Signal Processing (Print)
##                                7
##      Journal of Communication and Information Systems (Online)
```

```
## 6
## Journal of Visual Communication and Image Representation (Print)
## 6
## Pattern Recognition Letters
## 5

# Autores que tiveram mais publicações
head(sort(table(toupper(unlist(strsplit(unb.pub.ele_p8.df$autores,";")))),decreasing = TRUE),n=10)

##
##          ROCHA, ANDERSON          ROCHA, ANDERSON
##                52                21
##    GOLDENSTEIN, SIOME    DA COSTA, J. P. C. L.
##                17                15
##    SOUSA JR, R. T. DE REZENDE ROCHA, ANDERSON
##                14                13
##          ROCHA, A          ROCHA, A.
##                13                13
## ROCHA, ANDERSON DE REZENDE    WAINER, JACQUES
##                13                9
```

Processando os dados, pode-se concluir que todas as publicações foram completas, a maioria foi publicada na revista “IEEE Transactions on Information Forensics and Security”, 14 no total. O autor que teve mais publicação foi Anderson de Rezendo Rocha.

```
# Criando um data-frame com todos os anos
unb.pub.ele_p7.df <- data.frame()
for (i in 1:length(unb.pub.auto[[1]]))
  unb.pub.ele_p7.df <- rbind(unb.pub.ele_p7.df, unb.pub.ele_p7$PERIODICO[[i]])
# Limpando o data-frame de listas
unb.pub.ele_p7.df$autores <- gsub("\",\"|\", \"\", \";\", \"\", unb.pub.ele_p7.df$autores)
unb.pub.ele_p7.df$autores <- gsub("\"|c\\(|\\\\)", "\"", unb.pub.ele_p7.df$autores)
unb.pub.ele_p7.df$`autores-endogeno` <- gsub(",", " ", unb.pub.ele_p7.df$`autores-endogeno`)
unb.pub.ele_p7.df$`autores-endogeno` <- gsub("\"|c\\(|\\\\)", "\"", unb.pub.ele_p7.df$`autores-endogeno`)
#### ARQUIVO PROCESSADO E LIMPO
## Publicações por natureza
table(unb.pub.ele_p7.df$natureza)

##
## COMPLETO
##      161

## Publicações 2010 até 2017
table(unb.pub.ele_p7.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    15   16   14   22   15   16   37   26

## Publicações por periódico, sendo os 5 periódicos com masi publicação
d <- table(unb.pub.ele_p7.df$periodico)
head(sort(d,decreasing = TRUE),n=5)

##
## Journal of Communication and Information Systems (Online)
```

```
##                                     10
##                               Revista IEEE América Latina
##                                     8
##                               Digital Signal Processing (Print)
##                                     7
##                               IEEE Transactions on Power Delivery
##                                     6
##                               Electric Power Systems Research (Print)
##                                     5

# Autores que tiveram mais publicações
head(sort(table(toupper(unlist(strsplit(unb.pub.ele_p7.df$autores,";")))),decreasing = TRUE),n=10)

##
##          DA COSTA, J. P. C. L.          SOUSA JR, R. T.
##                               14                               14
##          MOLINARO, L. F. R.          ATTUX, ROMIS
##                               13                               6
## DE SOUSA JÚNIOR, RAFAEL TIMÓTEO      DA COSTA, JOÃO PAULO C.L.
##                               6                               5
##          DIAS, U. S.          FRAZAO, A.
##                               5                               5
##          GALPERIN, H.          SILVA, K. M.
##                               5                               5
```

Há um total de 161 publicações em engenharia elétrica p7, espaçado desde 2010 até 2017. Duas revistas 4 revistas se destacam, quando analisamos as revistas que foram enviadas: “Journal of Communication and Information Systems (Online)”, “Revista IEEE América Latina”, “Digital Signal Processing (Print)”, “IEEE Transactions on Power Delivery”.

O professor com mais publicação é o professor Da Costa (professor Lustosa), seguido por Molina. E mais uma vez é verificado a dificuldade que se encontra ao analisar nomes de autores que são os mesmo, mas que foram preenchido de foram diferente.

Arquivo Advise

Engenharia Biomédica

```
# Reunindo todos os anos e orientações concluídas em um mesmo dataframe
unb.adv.biomed.tipo.df <- data.frame(); unb.adv.biomed.df <- data.frame()
for (i in 1:length(unb.adv.biomed[[1]]))
  unb.adv.biomed.tipo.df <- rbind(unb.adv.biomed.tipo.df, unb.adv.biomed$ORIENTACAO_CONCLUIDA_POS_DOUTORADO[i])
unb.adv.biomed.df <- rbind(unb.adv.biomed.df, unb.adv.biomed.tipo.df); unb.adv.biomed.tipo.df <- data.frame()
for (i in 1:length(unb.adv.biomed[[1]]))
  unb.adv.biomed.tipo.df <- rbind(unb.adv.biomed.tipo.df, unb.adv.biomed$ORIENTACAO_CONCLUIDA_DOUTORADO[i])
unb.adv.biomed.df <- rbind(unb.adv.biomed.df, unb.adv.biomed.tipo.df); unb.adv.biomed.tipo.df <- data.frame()
for (i in 1:length(unb.adv.biomed[[1]]))
  unb.adv.biomed.tipo.df <- rbind(unb.adv.biomed.tipo.df, unb.adv.biomed$ORIENTACAO_CONCLUIDA_MESTRADO[i])
unb.adv.biomed.df <- rbind(unb.adv.biomed.df, unb.adv.biomed.tipo.df)
# Transformando as colunas de listas em caracteres eliminando c("")
unb.adv.biomed.df$nome_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.biomed.df$nome_orientadores)
unb.adv.biomed.df$id_lattes_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.biomed.df$id_lattes_orientadores)
# Separação das colunas com dois orientadores
unb.adv.biomed.df <- separate(unb.adv.biomed.df, nome_orientadores, into = c("ori1", "ori2"), sep = ",")

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 123 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, ...].
```

```

unb.adv.biomed.df <- separate(unb.adv.biomed.df, id_lattes_orientadores, into = c("idLattes1", "idLattes2"))

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 123 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, ...].

# Orientações com pelo menos 1 orientador
length(unb.adv.biomed.df$ori1)

## [1] 124

# Orientações com apenas 1 orientador
sum(is.na(unb.adv.biomed.df$idLattes2))

## [1] 123

# Número de orientacoes por ano
table(unb.adv.biomed.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##    7    6   12   17   25   25   13   19

# Instituições
sort(table(unb.adv.biomed.df$instituicao),decreasing = TRUE)

##
##                               Universidade de Brasília
##                                           82
##                               Universidade de Brasília - Faculdade UnB-Gama
##                                           21
##                               Faculdade do Gama da UnB
##                                           7
##                               Faculdade UNB Gama - FGA
##                                           7
##                               Universidade de Brasília - Campus Darcy Ribeiro
##                                           3
## Centro de Desenvolvimento e Planejamento Regional - Cedeplar/FACE/UFMG
##                                           1
##                               Faculdade do Gama/ UnB
##                                           1
##                               Universidade do Estado do Rio Grande do Norte
##                                           1
##                               Universidade Nacional de Brasília
##                                           1

# Cursos
cursos_d <- sort(table(unb.adv.biomed.df$curso), decreasing = TRUE)
cursos_d5<-head(cursos_d,5)
cursos_d

##
##                               Engenharia Biomédica
##                                           41
##                               Mestrado em Engenharia Biomédica
##                                           26
## Programa de Pós-Graduação em Engenharia Biomédica
##                                           7
##                               Engenharia Elétrica

```

```

##                                     6
##                                     Física
##                                     6
##                                     Ciências Médicas
##                                     5
## ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO
##                                     5
##             Pós-Graduação em Engenharia Biomédica
##                                     5
##                                     ENGENHARIA ELÉTRICA
##                                     3
##
##                                     2
##             ciências de materiais
##                                     2
##             PPDSCI/CEAM
##                                     2
##             Sistemas Mecatrônicos
##                                     2
##             Ciência da Computação - Uern - Ufersa
##                                     1
##             Ciências da Saúde
##                                     1
##             Ciências e Tecnologias em Saúde
##                                     1
##             Ciências Mecânicas
##                                     1
##             Demografia
##                                     1
##             Educação Física
##                                     1
##             Informática
##                                     1
##             Medicina (Clínica Médica)
##                                     1
##             Nanociência e Nanobiotecnologia
##                                     1
##             Pós-graduação em Ciências Médicas
##                                     1
##             Psicologia
##                                     1
##             Psicologia Clínica e Cultura
##                                     1
##
# Natureza das orientações
table(unb.adv.biomed.df$natureza)

##
##             Dissertação de mestrado Supervisão de pós-doutorado
##                                     109                                     2
##             Tese de doutorado
##                                     13
##
# 10 Professores que mais possuem orientações
head(sort(table(rbind(unb.adv.biomed.df$ori1, unb.adv.biomed.df$ori2)), decreasing = TRUE), 10)

```

```
##
##          Lourdes Mattos Brasil
##          24
## Suélia de Siqueira Rodrigues Fleury Rosa
##          20
##          Adson Ferreira da Rocha
##          13
##          Jose Felicio da Silva
##          11
##          Marilia Miranda Forte Gomes
##          11
## Cristiano Jacques Miosso Rodrigues Mendes
##          10
##          Marcelino Monteiro de Andrade
##          7
##          Ronni Geraldo Gomes de Amorim
##          7
## Vera Regina Fernandes da Silva Marães
##          7
##          Georges Daniel Amvame Nze
##          6
```

```
# 10 Alunos que possuem mais orientações
```

```
head(sort(table(toupper(unb.adv.biomed.df$nome_aluno)),decreasing = TRUE),n=10)
```

```
##
## AMILTON DOS REIS CAPISTRANO          ANTONIO DOMINGUES NETO
##          3          2
## CAMILA CADENA DE ALMEIDA  CRISTINA AKEMI SHIMODA UECHI
##          2          2
## LEINA ADRIANA BARBOSA PIMENTA          LUIZ ALBER LEMOS
##          2          2
## MARIA DO CARMO DOS REIS          ROBERTO AGUIAR LIMA
##          2          2
## ROOZBEH TAHMASEBI          SIMONE BEZERRA FRANCO
##          2          2
```

```
# Quantidade de alunos que não são bolsistas
```

```
sum(unb.adv.biomed.df$bolsa == 'NAO')
```

```
## [1] 78
```

```
# Quantidade de alunos que são bolsistas
```

```
sum(unb.adv.biomed.df$bolsa == 'SIM')
```

```
## [1] 46
```

```
# 10 Agências Financiadoras que mais financiam
```

```
head(sort(table(unb.adv.biomed.df$agencia_financiadora),decreasing = TRUE),n=10)
```

```
##
##
##          80
## Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
##          32
## Conselho Nacional de Desenvolvimento Científico e Tecnológico
##          10
```

```
##                Centro de Apoio ao Desenvolvimento Tecnológico
##                1
##                Fundação de Apoio à Pesquisa do Distrito Federal
##                1
```

Das 123 orientações, apenas uma delas contém um segundo orientador. A maioria das orientações ocorrem na UnB - Darcy Ribeiro depois é unidade do Gama, na UnB. Das 123, a maioria é dissertação de mestrado 109, Lourdes Matos Brasil é a professora com mais orientações. 46 alunos são bolsistas.

Engenharia de Sistemas Eletrônicos e de Automação

```
# Reunindo todos os anos e orientações concluídas em um mesmo dataframe
unb.adv.auto.tipo.df <- data.frame(); unb.adv.auto.df <- data.frame()
for (i in 1:length(unb.adv.auto[[1]]))
  unb.adv.auto.tipo.df <- rbind(unb.adv.auto.tipo.df, unb.adv.auto$ORIENTACAO_CONCLUIDA_POS_DOUTORADO[[i]])
unb.adv.auto.df <- rbind(unb.adv.auto.df, unb.adv.auto.tipo.df); unb.adv.auto.tipo.df <- data.frame()
for (i in 1:length(unb.adv.auto[[1]]))
  unb.adv.auto.tipo.df <- rbind(unb.adv.auto.tipo.df, unb.adv.auto$ORIENTACAO_CONCLUIDA_DOUTORADO[[i]])
unb.adv.auto.df <- rbind(unb.adv.auto.df, unb.adv.auto.tipo.df); unb.adv.auto.tipo.df <- data.frame()
for (i in 1:length(unb.adv.auto[[1]]))
  unb.adv.auto.tipo.df <- rbind(unb.adv.auto.tipo.df, unb.adv.auto$ORIENTACAO_CONCLUIDA_MESTRADO[[i]])
unb.adv.auto.df <- rbind(unb.adv.auto.df, unb.adv.auto.tipo.df)
# Transformando as colunas de listas em caracteres eliminando c("")
unb.adv.auto.df$nome_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.auto.df$nome_orientadores)
unb.adv.auto.df$id_lattes_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.auto.df$id_lattes_orientadores)
# Separação das colunas com dois orientadores
unb.adv.auto.df <- separate(unb.adv.auto.df, nome_orientadores, into = c("ori1", "ori2"), sep = ",")
```

```
## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 141 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].
```

```
unb.adv.auto.df <- separate(unb.adv.auto.df, id_lattes_orientadores, into = c("idLattes1", "idLattes2"))
```

```
## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 141 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].
```

```
# Orientações com pelo menos 1 orientador
length(unb.adv.auto.df$ori1)
```

```
## [1] 149
```

```
# Orientações com apenas 1 orientador
sum(is.na(unb.adv.auto.df$idLattes2))
```

```
## [1] 141
```

```
# Número de orientacoes por ano
table(unb.adv.auto.df$ano)
```

```
##
## 2010 2011 2012 2013 2014 2015 2016 2017
## 13 16 17 20 25 17 20 21
```

```
# Instituições
sort(table(unb.adv.auto.df$instituicao),decreasing = TRUE)
```

```
##
## Universidade de Brasília Universidade de Pernambuco
## 141 5
```



```

##      Universidade de São Paulo Technische Universität Dresden
##                                     2                               1

# Cursos
cursos_d <- sort(table(unb.adv.auto.df$curso), decreasing = TRUE)
cursos_d5<-head(cursos_d,5)
cursos_d

##
## ENGENHARIA DE SISTEMAS ELETRÔNICOS E DE AUTOMAÇÃO
##                                     74
##                               Engenharia Elétrica
##                                     23
## Engenharia de Sistemas Eletrônicos e de Automação
##                                     9
##
##                                     7
##                               Ciências Médicas
##                                     5
##                               Engenharia da Computação
##                                     5
##                               ENGENHARIA ELÉTRICA
##                                     5
##                               Informática
##                                     5
##                               Engenharia Biomédica
##                                     4
##                               Ciência da Computação
##                                     2
##                               Educação Física
##                                     2
##                               Ciências e Tecnologias em Saúde
##                                     1
##                               Electrical Engineering
##                                     1
##                               Eng. de Sist. Eletrônicos e de Automação
##                                     1
##                               Engenharia Eletrônica e de Automação
##                                     1
## Pós-Grad. em Sistemas Eletrônicos e de Automação
##                                     1
##                               Química
##                                     1
##                               Sistemas Mecatrônicos
##                                     1
##                               Tecnologias Química e Biológica
##                                     1

# Natureza das orientações
table(unb.adv.auto.df$natureza)

##
##      Dissertação de mestrado Supervisão de pós-doutorado
##                                     102                               7
##      Tese de doutorado
##                                     40

```

```
# 10 Professoes que mais possuem orientações
head(sort(table(rbind(unb.adv.auto.df$ori1, unb.adv.auto.df$ori2)), decreasing = TRUE), 10)
```

```
##
## Francisco Assis de Oliveira Nascimento
## 16
## João Yoshiyuki Ishihara
## 15
## Adson Ferreira da Rocha
## 13
## Antonio Padilha Lanari Bo
## 12
## Renato Mariz de Moraes
## 12
## Ricardo Lopes de Queiroz
## 11
## Geovany Araujo Borges
## 10
## Artemis Marti Ceschin
## 8
## Janaina Goncalves Guimaraes
## 8
## Mylene Christine Queiroz de Farias
## 8
```

```
# 10 Alunos que possuem mais orientações
head(sort(table(toupper(unb.adv.auto.df$nome_aluno)),decreasing = TRUE),n=10)
```

```
##
## HENRIQUE MARRA MENEGAZ LUIS FELIPE DA CRUZ FIGUEREDO
## 3 3
## ROBERTO DE SOUZA BAPTISTA ELDER OROSKI
## 3 2
## GABRIEL FIGUEIRÓ DE OLIVEIRA JUAN CAMILO MONTEALEGRE RIVERA
## 2 2
## LUCIANO ALMEIDA LEAL LUCIANO MAURO ARLEY SUP
## 2 2
## MARCEL HENRIQUE TRABUCO MOEMA DE BARROS E SILVA BOTELHO
## 2 2
```

```
# Quantidade de alunos que não são bolsistas
sum(unb.adv.auto.df$bolsa == 'NAO')
```

```
## [1] 56
```

```
# Quantidade de alunos que são bolsistas
sum(unb.adv.auto.df$bolsa == 'SIM')
```

```
## [1] 93
```

```
# 10 Agências Financiadoras que mais financiam
head(sort(table(unb.adv.auto.df$agencia_financiadora),decreasing = TRUE),n=10)
```

```
##
```

```
##
```

```
##
```

```
## Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
```

```
## 57
## Conselho Nacional de Desenvolvimento Científico e Tecnológico
## 25
## Escola Politécnica de Pernambuco
## 2
## Fundação de Apoio à Pesquisa do Distrito Federal
## 2
## PNPd/CAPES
## 2
## Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco
## 1
## Fundação de Amparo à Pesquisa do Estado de São Paulo
## 1
```

todo

Engenharia Elétrica P8

```
# Reunindo todos os anos e orientações concluídas em um mesmo dataframe
unb.adv.ele_p8.tipo.df <- data.frame(); unb.adv.ele_p8.df <- data.frame()
for (i in 1:length(unb.adv.ele_p8[[1]]))
  unb.adv.ele_p8.tipo.df <- rbind(unb.adv.ele_p8.tipo.df, unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)
unb.adv.ele_p8.df <- rbind(unb.adv.ele_p8.df, unb.adv.ele_p8.tipo.df); unb.adv.ele_p8.tipo.df <- data.frame()
for (i in 1:length(unb.adv.ele_p8[[1]]))
  unb.adv.ele_p8.tipo.df <- rbind(unb.adv.ele_p8.tipo.df, unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_DOUTORADO)
unb.adv.ele_p8.df <- rbind(unb.adv.ele_p8.df, unb.adv.ele_p8.tipo.df); unb.adv.ele_p8.tipo.df <- data.frame()
for (i in 1:length(unb.adv.ele_p8[[1]]))
  unb.adv.ele_p8.tipo.df <- rbind(unb.adv.ele_p8.tipo.df, unb.adv.ele_p8$ORIENTACAO_CONCLUIDA_MESTRADO)
unb.adv.ele_p8.df <- rbind(unb.adv.ele_p8.df, unb.adv.ele_p8.tipo.df)
# Transformando as colunas de listas em caracteres eliminando c("")
unb.adv.ele_p8.df$nome_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.ele_p8.df$nome_orientadores)
unb.adv.ele_p8.df$id_lattes_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.ele_p8.df$id_lattes_orientadores)
# Separação das colunas com dois orientadores
unb.adv.ele_p8.df <- separate(unb.adv.ele_p8.df, nome_orientadores, into = c("ori1", "ori2"), sep = ",")

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 133 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].
unb.adv.ele_p8.df <- separate(unb.adv.ele_p8.df, id_lattes_orientadores, into = c("idLattes1", "idLattes2"), sep = ",")

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 133 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].

# Orientações com pelo menos 1 orientador
length(unb.adv.ele_p8.df$ori1)

## [1] 135

# Orientações com apenas 1 orientador
sum(is.na(unb.adv.ele_p8.df$idLattes2))

## [1] 133

# Número de orientações por ano
table(unb.adv.ele_p8.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
```

```
##      9      10      29      15      15      11      28      18
```

```
# Instituições
```

```
sort(table(unb.adv.ele_p8.df$instituicao),decreasing = TRUE)
```

```
##
##                               Universidade de Brasília
##                               90
## Instituto de Computação - Universidade Estadual de Campinas
##                               29
##                               Universidade Estadual de Campinas
##                               7
##                               Universidade Salvador
##                               3
##                               City University of Hong Kong
##                               1
## Faculté des Sciences et Techniques - Universtité de Limoges
##                               1
##                               Technische Universität Ilmenau
##                               1
##      Universidade Federal da Bahia/Universidade Salvador/UEFS
##                               1
##                               Universidade Federal de Goiás
##                               1
##                               Universitat Politècnica de Catanlunya
##                               1
```

```
# Cursos
```

```
cursos_d <- sort(table(unb.adv.ele_p8.df$curso), decreasing = TRUE)
```

```
cursos_d5<-head(cursos_d,5)
```

```
cursos_d
```

```
##
##                               Engenharia Elétrica
##                               46
##                               ENGENHARIA ELÉTRICA
##                               30
##                               Mestrado
##                               14
##                               12
##                               Doutorado
##                               8
##      Engenharia Elétrica (53001010059F0) (F)
##                               5
##                               Ciência da Computação
##                               4
##                               Engenharia Biomédica
##                               3
##      Mestrado em Engenharia Elétrica
##                               3
##                               Sistemas e Computação
##                               2
##                               Sistemas Mecatrônicos
##                               2
##      Doutorado em Engenharia Elétrica
```

```

##                                     1
## Doutorado Multiinstitucional Ciência da Computação
##                                     1
## Electrical Engineering and Information Technology
##                                     1
##                                     Informática
##                                     1
##                                     Mestrado Em Redes
##                                     1
## Telecommunication Engineering & Management
##                                     1

# Natureza das orientações
table(unb.adv.ele_p8.df$natureza)

##
## Dissertação de mestrado Supervisão de pós-doutorado
##                                     106                                     12
## Tese de doutorado
##                                     17

# 10 Professores que mais possuem orientações
head(sort(table(rbind(unb.adv.ele_p8.df$ori1, unb.adv.ele_p8.df$ori2)), decreasing = TRUE), 10)

##
## Anderson de Rezende Rocha Rafael Timóteo de Sousa Jr
##                                     33                                     31
## Joao Paulo Carvalho Lustosa da Costa Anderson Clayton Alves Nascimento
##                                     22                                     13
## Dúbio Leandro Borges Bruno Werneck Pinto Hoelz
##                                     8                                     6
## Diego de Freitas Aranha Georges Daniel Amvame Nze
##                                     6                                     6
## Daniel Guerreiro e Silva William Ferreira Giozza
##                                     5                                     5

# 10 Alunos que possuem mais orientações
head(sort(table(toupper(unb.adv.ele_p8.df$nome_aluno)),decreasing = TRUE),n=10)

##
## ANTONIO MANUEL RUBIO SERRANO DANILO FERNANDES TENÓRIO
##                                     2                                     2
## FILIPE DE OLIVEIRA COSTA GILVAN MARTINS DURÃES
##                                     2                                     2
## GIOVANI CHIACHIA MARINA ATSUMI OIKAWA
##                                     2                                     2
## RICARDO KEHRLE MIRANDA VINÍCIUS COUTINHO GUIMARÃES COELHO
##                                     2                                     2
## ADRIANA BASTOS PINTO ADRIANA SILVA NEIVA
##                                     1                                     1

# Quantidade de alunos que não são bolsistas
sum(unb.adv.ele_p8.df$bolsa == 'NAO')

## [1] 83

```

```
# Quantidade de alunos que são bolsistas
sum(unb.adv.ele_p8.df$bolsa == 'SIM')
```

```
## [1] 52
```

```
# 10 Agências Financiadoras que mais financiam
```

```
head(sort(table(unb.adv.ele_p8.df$agencia_financiadora),decreasing = TRUE),n=10)
```

```
##
##
##                                     83
##  Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
##                                     25
##      Fundação de Amparo à Pesquisa do Estado de São Paulo
##                                     8
## Conselho Nacional de Desenvolvimento Científico e Tecnológico
##                                     5
##                                     Samsung Eletrônica da Amazônia
##                                     5
##                                     Intel Labs
##                                     2
##                                     LG Electronics
##                                     2
##      Fundação de Amparo a Pesquisa do Estado de Goiás
##                                     1
##                                     Google Inc.
##                                     1
##                                     Hong Kong City University
##                                     1
```

Das 135 orientações apenas 2 tem mais de um orientador. 2012 é o ano com mais orientação. A principal instituição é, sem novidades, UnB, seguindo pelo instituto de Computação - Universidade Estadual de Campinas. A maioria é dissertação de mestrado 106. Anderson Rocha e Rafael Sousa são os maiores colaboradores. A maioria dos alunos não são bolsistas 83 de 135.

Engenharia Elétrica P7

```
# Reunindo todos os anos e orientações concluídas em um mesmo dataframe
```

```
unb.adv.ele_p7.tipo.df <- data.frame(); unb.adv.ele_p7.df <- data.frame()
```

```
for (i in 1:length(unb.adv.ele_p7[[1]]))
```

```
  unb.adv.ele_p7.tipo.df <- rbind(unb.adv.ele_p7.tipo.df, unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_POS_DOUTORADO)
```

```
  unb.adv.ele_p7.df <- rbind(unb.adv.ele_p7.df, unb.adv.ele_p7.tipo.df); unb.adv.ele_p7.tipo.df <- data.frame()
```

```
for (i in 1:length(unb.adv.ele_p7[[1]]))
```

```
  unb.adv.ele_p7.tipo.df <- rbind(unb.adv.ele_p7.tipo.df, unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_DOUTORADO)
```

```
  unb.adv.ele_p7.df <- rbind(unb.adv.ele_p7.df, unb.adv.ele_p7.tipo.df); unb.adv.ele_p7.tipo.df <- data.frame()
```

```
for (i in 1:length(unb.adv.ele_p7[[1]]))
```

```
  unb.adv.ele_p7.tipo.df <- rbind(unb.adv.ele_p7.tipo.df, unb.adv.ele_p7$ORIENTACAO_CONCLUIDA_MESTRADO)
```

```
  unb.adv.ele_p7.df <- rbind(unb.adv.ele_p7.df, unb.adv.ele_p7.tipo.df)
```

```
# Transformando as colunas de listas em caracteres eliminando c("")
```

```
unb.adv.ele_p7.df$nome_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.ele_p7.df$nome_orientadores)
```

```
unb.adv.ele_p7.df$id_lattes_orientadores <- gsub("\\|c\\(|\\)", "", unb.adv.ele_p7.df$id_lattes_orientadores)
```

```
# Separação das colunas com dois orientadores
```

```
unb.adv.ele_p7.df <- separate(unb.adv.ele_p7.df, nome_orientadores, into = c("ori1", "ori2"), sep = ",")
```

```
## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 154 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].
```

```

unb.adv.ele_p7.df <- separate(unb.adv.ele_p7.df, id_lattes_orientadores, into = c("idLattes1", "idLattes2"))

## Warning: Expected 2 pieces. Missing pieces filled with `NA` in 154 rows [1,
## 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, ...].

# Orientações com pelo menos 1 orientador
length(unb.adv.ele_p7.df$ori1)

## [1] 158

# Orientações com apenas 1 orientador
sum(is.na(unb.adv.ele_p7.df$idLattes2))

## [1] 154

# Número de orientacoes por ano
table(unb.adv.ele_p7.df$ano)

##
## 2010 2011 2012 2013 2014 2015 2016 2017
##   10   16   27   23   18   20   28   16

# Instituições
sort(table(unb.adv.ele_p7.df$instituicao),decreasing = TRUE)

##
##               Universidade de Brasília
##                               146
##               Universidade de Brasilia
##                               7
##               City University of Hong Kong
##                               1
##               Technische Universität Ilmenau
##                               1
##               Universidade Estadual de Campinas
##                               1
## Universidade Federal de Campina Grande
##                               1
## Universitat Politècnica de Catanlunya
##                               1

# Cursos
cursos_d <- sort(table(unb.adv.ele_p7.df$curso), decreasing = TRUE)
cursos_d5<-head(cursos_d,5)
cursos_d

##
##                               Engenharia Elétrica
##                               77
##                               ENGENHARIA ELÉTRICA
##                               55
## Programa de Pós-Graduação em Engenharia Elétrica
##                               12
##                               Mestrado em Engenharia Elétrica
##                               7
##                               Doutorado em Engenharia Elétrica
##                               4
## Electrical Engineering and Information Technology

```

```

##                                     1
##           Mestrado em Engenharia Elétrica - PPGEE
##                                     1
##           Telecommunication Engineering & Management
##                                     1

# Natureza das orientações
table(unb.adv.ele_p7.df$natureza)

##
## Dissertação de mestrado           Tese de doutorado
##           136                       22

# 10 Professores que mais possuem orientações
head(sort(table(rbind(unb.adv.ele_p7.df$ori1, unb.adv.ele_p7.df$ori2)), decreasing = TRUE), 10)

##
##           Rafael Timóteo de Sousa Jr Joao Paulo Carvalho Lustosa da Costa
##                                     31                                     22
##           Kleber Melo e Silva           Paulo Roberto de Lira Gondim
##                                     16                                     16
##           Andre Noll Barreto   Marco Aurelio Gonçalves de Oliveira
##                                     12                                     11
##           Anésio de Leles Ferreira Filho           Francisco Damasceno Freitas
##                                     10                                     10
##           Ugo Silva Dias           Luis Fernando Ramos Molinaro
##                                     10                                     9

# 10 Alunos que possuem mais orientações
head(sort(table(toupper(unb.adv.ele_p7.df$nome_aluno)),decreasing = TRUE),n=10)

##
##           ALFREDO MIGUEL PIANETA ESCUDERO           ANTONIO MANUEL RUBIO SERRANO
##                                     2                                     2
##           DANIEL VIEIRA           DANILO FERNANDES TENÓRIO
##                                     2                                     2
##           EDUARDO JORGE SILVA LEITE JUNIOR           MARIA LEONOR SILVA DE ALMEIDA
##                                     2                                     2
##           RICARDO KEHRLE MIRANDA VINÍCIUS COUTINHO GUIMARÃES COELHO
##                                     2                                     2
##           ABIEZER AMARILIA FERNANDES           ABRAHAM ELIAS ORTEGA PAREDES
##                                     1                                     1

# Quantidade de alunos que não são bolsistas
sum(unb.adv.ele_p7.df$bolsa == 'NAO')

## [1] 106

# Quantidade de alunos que são bolsistas
sum(unb.adv.ele_p7.df$bolsa == 'SIM')

## [1] 52

# 10 Agências Financiadoras que mais financiam
head(sort(table(unb.adv.ele_p7.df$agencia_financiadora),decreasing = TRUE),n=10)

##
##
##

```



```
## Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
## 42
## Conselho Nacional de Desenvolvimento Científico e Tecnológico
## 7
## Hong Kong City University
## 1
```

Dos 158 trabalhos orientados apenas 4 possuem 2 orientadores, sendo 154 possuindo apenas um orientador. Apenas um curso não é especificamente engenharia elétrica - Telecommunication Engineering & Management. Das universidades que não são UnB tem as seguintes : City University of Hong Kong; Technische Universität Ilmenau; Universidade Estadual de Campinas; Universidade Federal de Campina Grande e Universität Politècnica de Catalunya.

Houveram 22 doutorados e 136 mestrados. Os professores que mais orientaram são: Rafael Timóteo de Sousa Jr com 31 orientações, Joao Paulo Carvalho Lustosa da Costa com 22 e Kleber Melo e Silva com 16.

Apenas 52 alunos são bolistas.

V.4 - Verificação da Qualidade dos Dados

Para o levantamento de informação os dados em geral são precisos, como por exemplo saber quais os professores mais envolvidos em orientações, quantos artigos foram concluídos, informações agrupadas por ano. Mas houve dificuldade quando o dado que se queria analisar dependia de como a pessoa que cadastrou escreveu, como por exemplo os autores nos artigos, haja visto que não foi encontrado um padrão, dificultou uma análise mais precisa dos dados.

VI - Preparação dos Dados

VI.1 - Seleção dos Dados

Para a principal utilização dos dados foram definidos os dataframes *adv*, que neste caso contém informações de qual a natureza da pesquisa produzida, qual aluno produziu, seus respectivos orientadores, o ano e etc. também foi definido os dados importantes do arquivo profile.json, gerando o mais dataframes, chamados de *prof*, que já está limpo, com as principais colunas definidas.

VI.2 - Limpeza dos Dados

Esta etapa se encorporea com a próxima etapa, que é a construção dos dados. Durante a construção, a limpeza dos dados também foi realizada.

VI.3 - Construção dos Dados

Engenharia Biomédica

```
## Começo da Preparação dos dados referente a profile
unb.prof.biomed.json <- read_file("dados-2018-2/engenharia-biomedica/279.profile.json")
unb.prof.biomed.df.capes <- read_csv("dados-2018-2/PesqPosCapes.csv",
                                     sep = ";", header = TRUE, colClasses = "character")
unb.prof.biomed <- fromJSON(unb.prof.biomed.json)
length(unb.prof.biomed)
```

```
## [1] 15
```

```
# Extrai perfis dos professores
unb.prof.biomed.df.professores <- extrai_perfis(unb.prof.biomed)
# Extrai producao bibliografica de todos os professores
unb.prof.biomed.df.publicacoes <- extrai_producoes(unb.prof.biomed)
```

```

# Extraí orientacoes
unb.prof.biomed.df.orientacoes <- extrai_orientacoes(unb.prof.biomed)
# Extraí areas de atuacao
unb.prof.biomed.df.areas.de.atuacao <- extrai_areas_atuacao(unb.prof.biomed)
# Salva os dataframes
save(unb.prof.biomed.df.professores, unb.prof.biomed.df.publicacoes,
     unb.prof.biomed.df.orientacoes, unb.prof.biomed.df.areas.de.atuacao, file = "dataframes.Rda")
# Cria arquivo para análise
unb.prof.biomed.df <- data.frame()
unb.prof.biomed.df <- unb.prof.biomed.df.professores %>%
  select(idLattes, nome, resumo_cv, senioridade) %>%
  left_join(
    unb.prof.biomed.df.orientacoes %>%
      select(orientacao, idLattes) %>%
      filter(!grepl("EM_ANDAMENTO", orientacao)) %>%
      group_by(idLattes) %>%
      count(orientacao) %>%
      spread(key = orientacao, value = n),
    by = "idLattes") %>%
  left_join(
    unb.prof.biomed.df.publicacoes %>%
      select(tipo_producao, idLattes) %>%
      filter(!grepl("ARTIGO_ACEITO", tipo_producao)) %>%
      group_by(idLattes) %>%
      count(tipo_producao) %>%
      spread(key = tipo_producao, value = n),
    by = "idLattes") %>%
  left_join(
    unb.prof.biomed.df.areas.de.atuacao %>%
      select(area, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(area)),
    by = "idLattes") %>%
  left_join(
    unb.prof.biomed.df.capes %>%
      select(AreaPos, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(AreaPos)),
    by = "idLattes")
## Mostrando dados processados referente ao arquivo profile
head(unb.prof.biomed.df,3)

```

```

##           idLattes                                nome
## 1 0201204222182378                                Leandro Xavier Cardoso
## 2 0535100751136568 Cristiano Jacques Miosso Rodrigues Mendes
## 3 1141716826787805                                Adson Ferreira da Rocha
##
## 1
## 2
## 3 Engenheiro Eletricista pela Universidade de Brasília (1988), Mestre em Engenharia Elétrica pela Un
##   senioridade ORIENTACAO_CONCLUIDA_DOUTORADO ORIENTACAO_CONCLUIDA_MESTRADO
## 1           9                                NA                                NA
## 2           6                                1                                9
## 3           9                                4                                9

```

```
## ORIENTACAO_CONCLUIDA_POS_DOUTORADO OUTRAS_ORIENTACOES_CONCLUIDAS
## 1 NA 5
## 2 NA 2
## 3 NA 6
## CAPITULO_DE_LIVRO DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA EVENTO LIVRO
## 1 NA NA 3 NA
## 2 1 NA 23 NA
## 3 9 NA 51 NA
## PERIODICO TEXTO_EM_JORNAIS n_distinct(area) n_distinct(AreaPos)
## 1 10 NA 1 1
## 2 8 NA 3 1
## 3 22 NA 2 3
```

```
## Mostrando dados processados referente ao arquivo publication
head(unb.pub.biomed.df,3)
```

```
## natureza
## 1 COMPLETO
## 2 COMPLETO
## 3 COMPLETO
##
## 1 Prótese para controle de fluxo esofagiano como nova técnica para o tratamento da obesidade (QUALIS
## 2 Thermoluminescent dose reconstruction using quartz extracted from unfired build
## 3 Breast cancer image assessment using an adaptative network-based fuzzy inference sy
##
## periodo ano volume
## 1 Revista Brasileira de Engenharia Biomédica (Impresso) 2010 26
## 2 Journal of Physics. Conference Series (Online) 2010 249
## 3 Pattern Recognition and Image Analysis 2010 20
##
## issn paginas doi
## 1 15173151 49 - 54
## 2 17426596 012031 - 10.1088/1742-6596/249/1/012031
## 3 10546618 192 - 200 10.1134/S1054661810020112
##
## 1
## 2 Campos, Simara S; Almeida, Geângela M; CARDOSO, L. X.;Cardoso, Leandro X;CARDOSO, L X;XAVIER CARDOS
## 3 Fernandes, F. C.; BRASIL, L. M.;Brasil, L. M.;Brasil, L
##
## autores-endogeno
## 1 1141716826787805
## 2 0201204222182378
## 3 9190489069187153
```

```
## Mostrando dados processados referente ao arquivo advise
head(unb.adv.biomed.df,3)
```

```
## natureza
## 1 Supervisão de pós-doutorado
## 2 Supervisão de pós-doutorado
## 3 Tese de doutorado
##
## 1
## 2
## 3 Influência da Eletroestimulação Neuromuscular de Baixa Frequência nas Variáveis Eletromiográficas
## ano id_lattes_aluno nome_aluno
## 1 2014 Leandro Xavier Cardoso
## 2 2017 Glécia Virgolino da Silva Luz
## 3 2011 7129464687368571 Kenia Fonseca Pires
```

```
##                               instituicao                curso
## 1 Universidade de Brasília - Faculdade UnB-Gama
## 2 Universidade de Brasília - Faculdade UnB-Gama
## 3                               Universidade de Brasília Ciências Médicas
##   codigo_do_curso bolsa
## 1                               SIM
## 2                               SIM
## 3           60021152   SIM
##                               agencia_financiadora
## 1   Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
## 2   Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
## 3 Conselho Nacional de Desenvolvimento Científico e Tecnológico
##   codigo_agencia_financiadora        ori1 ori2
## 1           045000000000    Lourdes Mattos Brasil <NA>
## 2           045000000000    Lourdes Mattos Brasil <NA>
## 3           002200000000 Adson Ferreira da Rocha <NA>
##       idLattes1 idLattes2
## 1 9190489069187153    <NA>
## 2 9190489069187153    <NA>
## 3 1141716826787805    <NA>
```

```
# Numero de docentes
length(unb.prof.ele_p8)
```

```
## [1] 11
```

```
# Numero de discentes
unique(length(unb.prof.biomed.df.orientacoes$nome_aluno))
```

```
## [1] 673
```

Engenharia de Sistemas Eletrônicos e de Automação

```
## Começo da Preparação dos dados referente a profile
unb.prof.auto.json <- read_file("dados-2018-2/engenharia-eletronicos-automacao/280.profile.json")
unb.prof.auto.df.capes <- read.csv("dados-2018-2/PesqPosCapes.csv",
                                   sep = ";", header = TRUE, colClasses = "character")
unb.prof.auto <- fromJSON(unb.prof.auto.json)
length(unb.prof.auto)
```

```
## [1] 21
```

```
# Extrai perfis dos professores
unb.prof.auto.df.professores <- extrai_perfis(unb.prof.auto)
# Extrai producao bibliografica de todos os professores
unb.prof.auto.df.publicacoes <- extrai_producoes(unb.prof.auto)
# Extrai orientacoes
unb.prof.auto.df.orientacoes <- extrai_orientacoes(unb.prof.auto)
# Extrai areas de atuacao
unb.prof.auto.df.areas.de.atuacao <- extrai_areas_atuacao(unb.prof.auto)
# Salva os dataframes
save(unb.prof.auto.df.professores, unb.prof.auto.df.publicacoes,
      unb.prof.auto.df.orientacoes, unb.prof.auto.df.areas.de.atuacao, file = "dataframes.Rda")
# Cria arquivo para análise
unb.prof.auto.df <- data.frame()
unb.prof.auto.df <- unb.prof.auto.df.professores %>%
  select(idLattes, nome, resumo_cv, senioridade) %>%
```

```

left_join(
  unb.prof.auto.df.orientacoes %>%
    select(orientacao, idLattes) %>%
    filter(!grepl("EM_ANDAMENTO", orientacao)) %>%
    group_by(idLattes) %>%
    count(orientacao) %>%
    spread(key = orientacao, value = n),
  by = "idLattes") %>%
left_join(
  unb.prof.auto.df.publicacoes %>%
    select(tipo_producao, idLattes) %>%
    filter(!grepl("ARTIGO_ACEITO", tipo_producao)) %>%
    group_by(idLattes) %>%
    count(tipo_producao) %>%
    spread(key = tipo_producao, value = n),
  by = "idLattes") %>%
left_join(
  unb.prof.auto.df.areas.de.atuacao %>%
    select(area, idLattes) %>%
    group_by(idLattes) %>%
    summarise(n_distinct(area)),
  by = "idLattes") %>%
left_join(
  unb.prof.auto.df.capes %>%
    select(AreaPos, idLattes) %>%
    group_by(idLattes) %>%
    summarise(n_distinct(AreaPos)),
  by = "idLattes")
## Mostrando dados processados referente ao arquivo profile
head(unb.prof.auto.df,3)

```

```

##           idLattes           nome
## 1 0301021863146083      Adolfo Bauchspiess
## 2 0335058635088347          Martin Claus
## 3 1141716826787805 Adson Ferreira da Rocha
##
## 1
## 2
## 3 Engenheiro Eletricista pela Universidade de Brasília (1988), Mestre em Engenharia Elétrica pela Un
##   senioridade ORIENTACAO_CONCLUIDA_DOUTORADO ORIENTACAO_CONCLUIDA_MESTRADO
## 1           6                2                5
## 2           5                NA                NA
## 3           9                4                9
##   ORIENTACAO_CONCLUIDA_POS_DOUTORADO OUTRAS_ORIENTACOES_CONCLUIDAS
## 1                NA                39
## 2                NA                NA
## 3                NA                6
##   CAPITULO_DE_LIVRO EVENTO LIVRO PERIODICO TEXTO_EM_JORNAIS
## 1           1       36       1           4                NA
## 2          NA       NA       NA           9                NA
## 3           9       49       NA          22                NA
##   n_distinct(area) n_distinct(AreaPos)
## 1           4           1
## 2           1           1

```

```

## 3                2                3
## Mostrando dados processados referente ao arquivo publication
head(unb.pub.auto.df,3)

## natureza
## 1 COMPLETO
## 2 COMPLETO
## 3 COMPLETO
##
## 1                Prótese para controle de fluxo esofagiano como nova técnica para o tratamento da obesidade
## 2                Sudy of Organic MIM Junctions for use in Photocatalysis
## 3 Electrical and humidity characterization of humic substances (HS) from peat for a possible use in
##
##                periodico ano volume
## 1 Revista Brasileira de Engenharia Biomédica (Impresso) 2010 26
## 2 ECS transactions 2010 31
## 3 ECS transactions 2010 31
## issn paginas doi
## 1 15173151 49 - 54
## 2 19385862 219 - 224 10.1149/1.3474162
## 3 19385862 419 - 424 10.1149/1.3474187
##
##                autores
## 1 Rosa, S. S. R. F.; da Rocha, A F; Brasil, L. M.; Oliveira, E. C.; de Paula, P. M. C.
## 2 E.R.da Silva; A.R.S.Romariz; CESCHIN, A. M.
## 3 G.L.Sandri; N.C.Ribeiro; A.R.S.Romariz; CESCHIN, A. M.
##
##                autores-endogeno
## 1 1141716826787805
## 2 5202540363896199; 2065104260659969
## 3 5202540363896199; 2065104260659969
## Mostrando dados processados referente ao arquivo advise
head(unb.adv.auto.df,3)

## natureza titulo ano id_lattes_aluno
## 1 Supervisão de pós-doutorado 2010
## 2 Supervisão de pós-doutorado 2013
## 3 Supervisão de pós-doutorado 2014
## nome_aluno instituicao curso
## 1 Camilo C. Dorea Universidade de Brasília
## 2 Moema de Barros e Silva Botelho Universidade de Brasília
## 3 MOEMA DE BARROS E SILVA BOTELHO Universidade de Brasília
## codigo_do_curso bolsa agencia_financiadora codigo_agencia_financiadora
## 1 NAO
## 2 SIM PNPD/CAPES 000600000990
## 3 SIM PNPD/CAPES 000600000990
## ori1 ori2 idLattes1 idLattes2
## 1 Ricardo Lopes de Queiroz <NA> 9271885452744905 <NA>
## 2 Artemis Marti Ceschin <NA> 2065104260659969 <NA>
## 3 Artemis Marti Ceschin <NA> 2065104260659969 <NA>
# Numero de docentes
length(unb.prof.ele_p8)

## [1] 11
# Numero de discentes
unique(length(unb.prof.auto.df.orientacoes$nome_aluno))

```

```
## [1] 655
```

Engenharia Elétrica P8

```
## Começo da Preparação dos dados referente a profile
unb.prof.ele_p8.json <- read_file("dados-2018-2/engenharia-eletrica-p8/281.profile.json")
unb.prof.ele_p8.df.capes <- read_csv("dados-2018-2/PesqPosCapes.csv",
                                     sep = ";", header = TRUE, colClasses = "character")
unb.prof.ele_p8 <- fromJSON(unb.prof.ele_p8.json)
length(unb.prof.ele_p8)
```

```
## [1] 11
```

```
# Extrai perfis dos professores
unb.prof.ele_p8.df.professores <- extrai_perfis(unb.prof.ele_p8)
# Extrai producao bibliografica de todos os professores
unb.prof.ele_p8.df.publicacoes <- extrai_producoes(unb.prof.ele_p8)
# Extrai orientacoes
unb.prof.ele_p8.df.orientacoes <- extrai_orientacoes(unb.prof.ele_p8)
# Extrai areas de atuacao
unb.prof.ele_p8.df.areas.de.atuacao <- extrai_areas_atuacao(unb.prof.ele_p8)
# Salva os dataframes
save(unb.prof.ele_p8.df.professores, unb.prof.ele_p8.df.publicacoes,
     unb.prof.ele_p8.df.orientacoes, unb.prof.ele_p8.df.areas.de.atuacao, file = "dataframes.Rda")
# Cria arquivo para análise
unb.prof.ele_p8.df <- data.frame()
unb.prof.ele_p8.df <- unb.prof.ele_p8.df.professores %>%
  select(idLattes, nome, resumo_cv, senioridade) %>%
  left_join(
    unb.prof.ele_p8.df.orientacoes %>%
      select(orientacao, idLattes) %>%
      filter(!grepl("EM_ANDAMENTO", orientacao)) %>%
      group_by(idLattes) %>%
      count(orientacao) %>%
      spread(key = orientacao, value = n),
    by = "idLattes" %>%
  left_join(
    unb.prof.ele_p8.df.publicacoes %>%
      select(tipo_producao, idLattes) %>%
      filter(!grepl("ARTIGO_ACEITO", tipo_producao)) %>%
      group_by(idLattes) %>%
      count(tipo_producao) %>%
      spread(key = tipo_producao, value = n),
    by = "idLattes" %>%
  left_join(
    unb.prof.ele_p8.df.areas.de.atuacao %>%
      select(area, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(area)),
    by = "idLattes" %>%
  left_join(
    unb.prof.ele_p8.df.capes %>%
      select(AreaPos, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(AreaPos)),
```

```
by = "idLattes")
## Mostrando dados processados referente ao arquivo profile
head(unb.prof.ele_p8.df,3)
```

```
##          idLattes                                nome
## 1 1521815731111261          Anderson de Rezende Rocha
## 2 1786889674911887 Joao Paulo Carvalho Lustosa da Costa
## 3 1808368824068028          DÍbio Leandro Borges
##
## 1 Anderson de Rezende Rocha é professor associado (Livre-Docente) do Instituto de Computação, Univer
## 2
## 3
##      senioridade ORIENTACAO_CONCLUIDA_DOUTORADO ORIENTACAO_CONCLUIDA_MESTRADO
## 1              9                             8                          15
## 2              8                             2                          20
## 3              8                             NA                          7
##      ORIENTACAO_CONCLUIDA_POS_DOUTORADO OUTRAS_ORIENTACOES_CONCLUIDAS
## 1                                10                          23
## 2                                NA                          29
## 3                                1                           9
##      CAPITULO_DE_LIVRO DEMAIS_TIPOS_DE_PRODUCAO_BIBLIOGRAFICA EVENTO LIVRO
## 1              4                                NA      66      NA
## 2              4                                NA      99      1
## 3              NA                                NA      12      NA
##      PERIODICO TEXTO_EM_JORNAIS n_distinct(area) n_distinct(AreaPos)
## 1          63              2              1              1
## 2          28              8              1              2
## 3           3              NA              2              2
```

```
## Mostrando dados processados referente ao arquivo publication
head(unb.pub.ele_p8.df,3)
```

```
##      natureza
## 1 COMPLETO
## 2 COMPLETO
## 3 COMPLETO
##
##                                     titulo
## 1          Automatic fruit and vegetable classification from images
## 2          Progressive randomization: Seeing the unseen
## 3 Efficient implementation of elliptic curve cryptography in wireless sensors
##
##      periodico ano volume issn
## 1      Computers and Electronics in Agriculture 2010      70 01681699
## 2 Computer Vision and Image Understanding (Print) 2010    114 10773142
## 3      Advances in Mathematics of Communications 2010      4 19305346
##
##      paginas doi
## 1 96 - 104 10.1016/j.compag.2009.09.002
## 2 349 - 362 10.1016/j.cviu.2009.10.002
## 3 169 - 187 10.3934/amc.2010.4.169
##
## 1 ROCHA, Anderson;ROCHA, ANDERSON;ROCHA, A.;DE REZENDE ROCHA, ANDERSON;ROCHA, ANDERSON DE REZENDE;RO
## 2          ROCHA, Anderson;ROCHA, ANDERSON;ROCHA, A.;DE REZENDE ROCHA, A
## 3          ARANHA
##
##      autores-endogeno
## 1 1521815731111261
```



```
## 2 1521815731111261
## 3 9788199690491510

## Mostrando dados processados referente ao arquivo advise
head(unb.adv.ele_p8.df,3)

##              natureza
## 1 Supervisão de pós-doutorado
## 2 Supervisão de pós-doutorado
## 3 Supervisão de pós-doutorado
##
## 1 Reconhecimento e Classificação de Padrões por Meio de Técnicas de Engenharia de Características, F
## 2
## 3
##      ano id_lattes_aluno              nome_aluno
## 1 2012              Otávio Augusto Bizetto Penatti
## 2 2013              Marina Atsumi Oikawa
## 3 2014      Leandro Luís Galdino de Oliveira
##
##              instituicao curso
## 1 Instituto de Computação - Universidade Estadual de Campinas
## 2 Instituto de Computação - Universidade Estadual de Campinas
## 3      Universidade Federal de Goiás
##      codigo_do_curso bolsa              agencia_financiadora
## 1              SIM              Samsung Eletrônica da Amazônia
## 2              SIM              União Européia
## 3      SIM Fundação de Amparo a Pesquisa do Estado de Goiás
##      codigo_agencia_financiadora              ori1 ori2
## 1      002400000993 Anderson de Rezende Rocha <NA>
## 2      002600000997 Anderson de Rezende Rocha <NA>
## 3      IXOT00000002      Díbio Leandro Borges <NA>
##      idLattes1 idLattes2
## 1 1521815731111261      <NA>
## 2 1521815731111261      <NA>
## 3 1808368824068028      <NA>

# Numero de docentes
length(unb.prof.ele_p8)

## [1] 11

# Numero de discentes
unique(length(unb.prof.ele_p8.df.orientacoes$nome_aluno))

## [1] 340
```

Engenharia Elétrica P7

```
## Começo da Preparação dos dados referente a profile
unb.prof.ele_p7.json <- read_file("dados-2018-2/engenharia-eletrica-p7/282.profile.json")
unb.prof.ele_p7.df.capes <- read_csv("dados-2018-2/PesqPosCapes.csv",
                                     sep = ";", header = TRUE, colClasses = "character")
unb.prof.ele_p7 <- fromJSON(unb.prof.ele_p7.json)
length(unb.prof.ele_p7)

## [1] 14
```

```

# Extraí perfis dos professores
unb.prof.ele_p7.df.professores <- extrai_perfis(unb.prof.ele_p7)
# Extraí producao bibliografica de todos os professores
unb.prof.ele_p7.df.publicacoes <- extrai_producoes(unb.prof.ele_p7)
# Extraí orientacoes
unb.prof.ele_p7.df.orientacoes <- extrai_orientacoes(unb.prof.ele_p7)
# Extraí areas de atuacao
unb.prof.ele_p7.df.areas.de.atuacao <- extrai_areas_atuacao(unb.prof.ele_p7)
# Salva os dataframes
save(unb.prof.ele_p7.df.professores, unb.prof.ele_p7.df.publicacoes,
      unb.prof.ele_p7.df.orientacoes, unb.prof.ele_p7.df.areas.de.atuacao, file = "dataframes.Rda")
# Cria arquivo para análise
unb.prof.ele_p7.df <- data.frame()
unb.prof.ele_p7.df <- unb.prof.ele_p7.df.professores %>%
  select(idLattes, nome, resumo_cv, senioridade) %>%
  left_join(
    unb.prof.ele_p7.df.orientacoes %>%
      select(orientacao, idLattes) %>%
      filter(!grepl("EM_ANDAMENTO", orientacao)) %>%
      group_by(idLattes) %>%
      count(orientacao) %>%
      spread(key = orientacao, value = n),
    by = "idLattes") %>%
  left_join(
    unb.prof.ele_p7.df.publicacoes %>%
      select(tipo_producao, idLattes) %>%
      filter(!grepl("ARTIGO_ACEITO", tipo_producao)) %>%
      group_by(idLattes) %>%
      count(tipo_producao) %>%
      spread(key = tipo_producao, value = n),
    by = "idLattes") %>%
  left_join(
    unb.prof.ele_p7.df.areas.de.atuacao %>%
      select(area, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(area)),
    by = "idLattes") %>%
  left_join(
    unb.prof.ele_p7.df.capes %>%
      select(AreaPos, idLattes) %>%
      group_by(idLattes) %>%
      summarise(n_distinct(AreaPos)),
    by = "idLattes")
## Mostrando dados processados referente ao arquivo profile
head(unb.prof.ele_p7.df,3)

```

```

##           idLattes                                nome
## 1 0997629063334821      Luis Fernando Ramos Molinaro
## 2 1573568521807396      Anésio de Leles Ferreira Filho
## 3 1786889674911887 Joao Paulo Carvalho Lustosa da Costa
##

```

```

## 1 Possui graduação em Engenharia Elétrica pela Universidade de Brasília (1979), mestrado em Engenhar
## 2
## 3

```

```
##      senioridade ORIENTACAO_CONCLUIDA_DOUTORADO ORIENTACAO_CONCLUIDA_MESTRADO
## 1           9                3                6
## 2           7                NA               12
## 3           8                2               20
##      OUTRAS_ORIENTACOES_CONCLUIDAS CAPITULO_DE_LIVRO EVENTO LIVRO PERIODICO
## 1                NA                8      17      2      17
## 2                38                1      36     NA      9
## 3                29                4      99      1     28
##      TEXTO_EM_JORNAIS n_distinct(area) n_distinct(AreaPos)
## 1                NA                1                1
## 2                 1                1                1
## 3                 8                1                2
```

```
## Mostrando dados processados referente ao arquivo publication
head(unb.pub.ele_p7.df,3)
```

```
##      natureza
## 1 COMPLETO
## 2 COMPLETO
## 3 COMPLETO
##
```

```
## 1                                     Computing Rightmost Eigenvalues for Small-Sig
## 2                                     Distance
## 3 Modeling and Characterization for Microstrip Filters in the Manufacturing Process through the Unsc
##      periodoico ano volume issn paginas
## 1 IEEE Transactions on Power Systems 2010      25 15580679 929 - 938
## 2 Electric Power Systems Research (Print) 2010      80 03787796 84 - 90
## 3 Modelling and Simulation in Engineering 2010 2010 16875591 1 - 5
##      doi
## 1 10.1109/tpwrs.2009.2036822
## 2 10.1016/j.epsr.2009.08.012
## 3 10.1155/2010/691241
```

```
##      autores
## 1 Rommes, Joost; MARTINS, NELSON; FREITAS, F. D.;Freitas, Francisco Damasceno
## 2 SILVA, Kleber Melo; NEVES, W. L. A.; SOUZA, B. A.
## 3 Ortega Paredes, Abraham E.; de Menezes, Leonardo R. A. X.; Abdalla, Humberto; Romani, Ivan N. A.
##      autores-endogeno
## 1 7997341055520362
## 2 8795500242651581
## 3 9279262545449206
```

```
## Mostrando dados processados referente ao arquivo advise
head(unb.adv.ele_p7.df,3)
```

```
##      natureza
## 1 Tese de doutorado
## 2 Tese de doutorado
## 3 Tese de doutorado
##
```

```
## 1                                     Aplicação da Transformada da Incerteza para Projetos de
## 2 Análise do Perfil do Profissional de Tecnologia da Informação e Comunicação a partir de Percep
## 3 Método para determinação do limite de penetração da geração distribuída fotovoltaica em sistemas r
##      ano id_lattes_aluno nome_aluno
## 1 2010 6817085573778215 Abraham Elias Ortega Paredes
## 2 2010 9354456664541013 Marcelo Stehling de Castro
```

```
## 3 2010 6471578140876893          RAFAEL AMARAL SHAYANI
##              instituicao              curso
## 1 Universidade de Brasília          Engenharia Elétrica
## 2 Universidade de Brasília Doutorado em Engenharia Elétrica
## 3 Universidade de Brasília          Engenharia Elétrica
##      codigo_do_curso bolsa
## 1      60030232      SIM
## 2      90000003      NAO
## 3      51500140      NAO
##              agencia_financiadora
## 1 Conselho Nacional de Desenvolvimento Científico e Tecnológico
## 2
## 3
##      codigo_agencia_financiadora              ori1
## 1      002200000000 Leonardo Rodrigues Araujo Xavier de Menezes
## 2              Luis Fernando Ramos Molinaro
## 3              Marco Aurelio Gonçalves de Oliveira
##      ori2      idLattes1 idLattes2
## 1 <NA> 9279262545449206      <NA>
## 2 <NA> 0997629063334821      <NA>
## 3 <NA> 8709856465061922      <NA>
```

```
# Numero de docentes
length(unb.prof.ele_p7)
```

```
## [1] 14
```

```
# Numero de discentes
unique(length(unb.prof.ele_p7.df.orientacoes$nome_aluno))
```

```
## [1] 511
```

VI.4 - Integração dos dados

Neste presente trabalho, não se viu a necessidade de fazer integração entre os dados, sendo somente realizada a construção dos dataframes expostos acima.

VI.5 - Formatação dos dados

As formatações de dados necessárias já foram feitas, como por exemplo “orientadores” se tornou “orie1” e “orie2”, fazendo com que cada variável contenha apenas um elemento, e não mais uma lista de orientadores.

VII - Modelagem

Após a realização da limpeza e da preparação dos dados, a ferramenta que será utilizada para modelagem será computacional, para fins de visualização gráfica dos dados obtidos a partir dos dataframes criados.

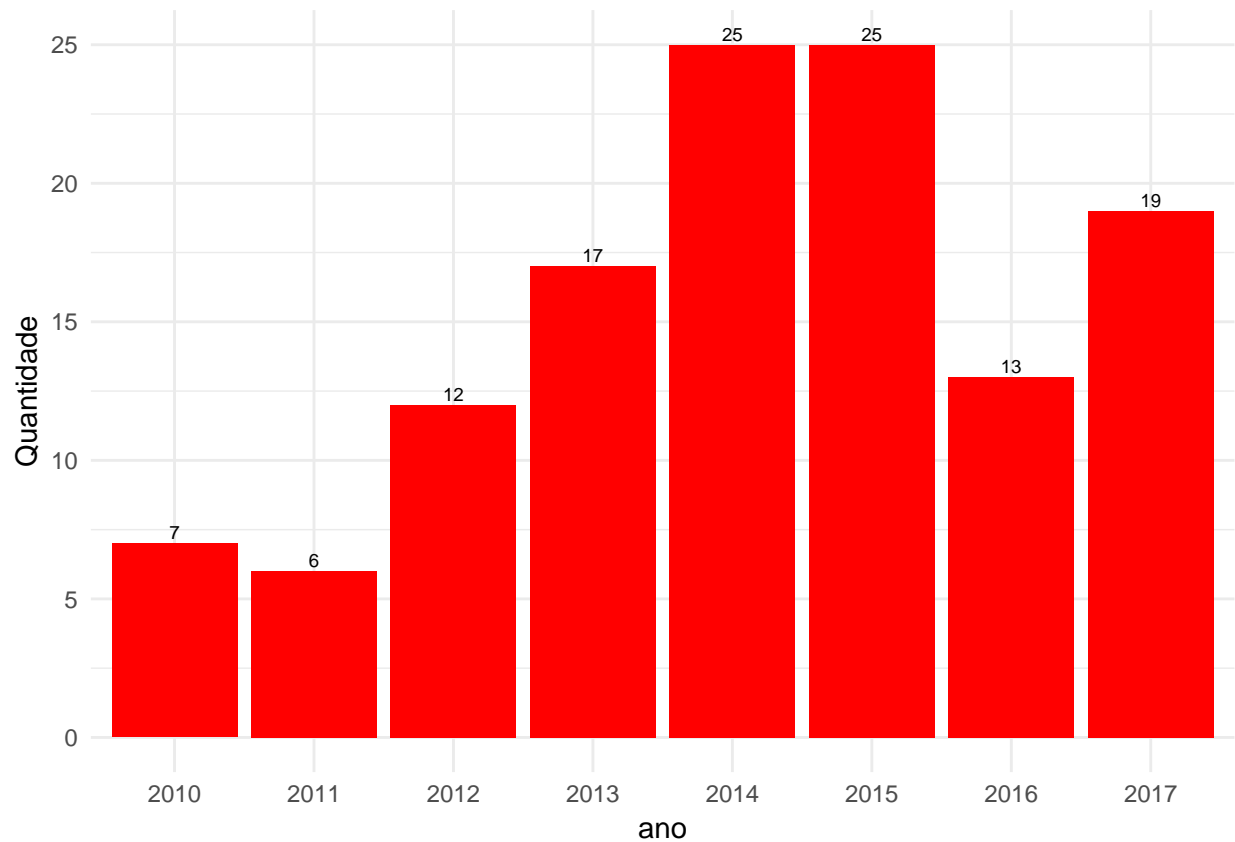
VIII - Avaliação

A avaliação será feita de forma gráfica, sendo estes gráficos exibidos a seguir:

Engenharia Biomédica

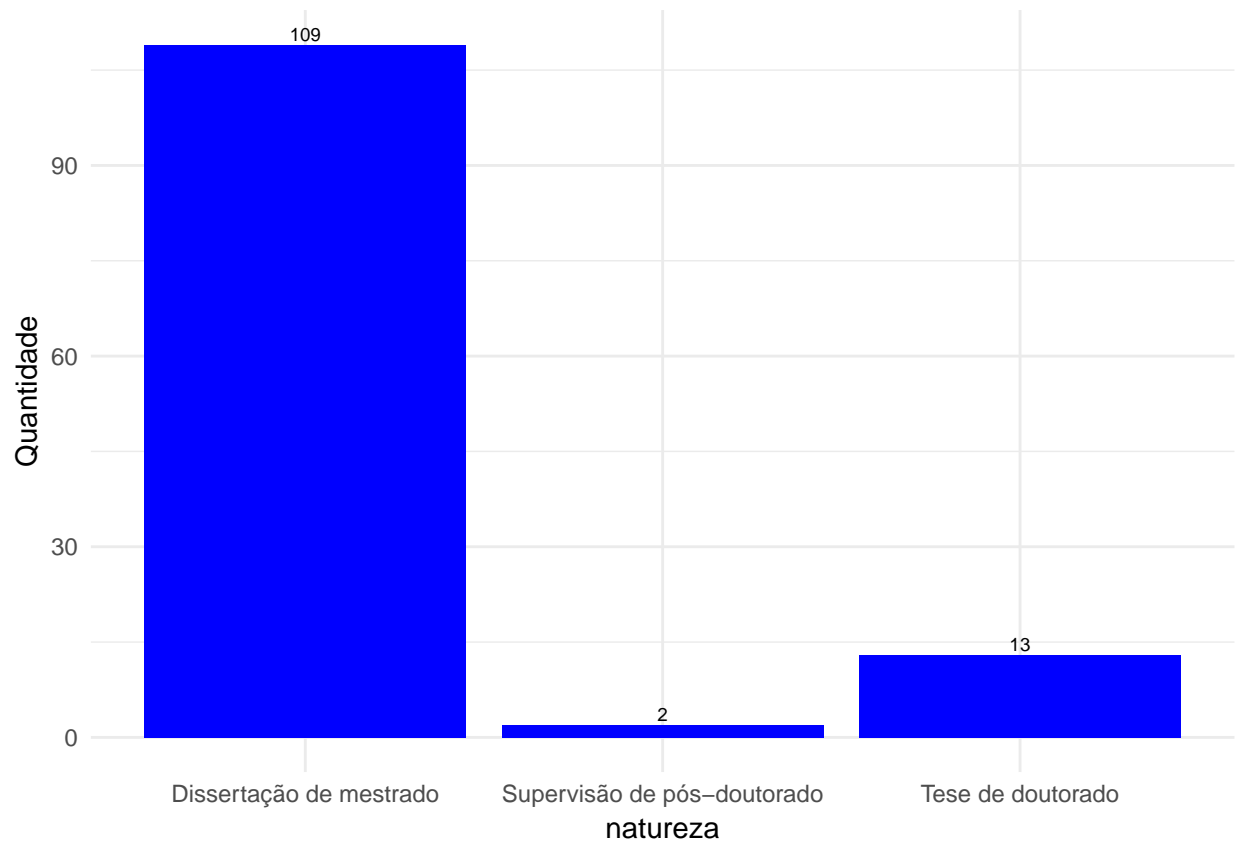
Quantidade de orientações feitas por ano:

```
unb.adv.biomed.df %>%
  group_by(ano) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = ano, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "red")+
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5)+
  theme_minimal()
```



Quantidade de orientações separadas por tipo de publicação:

```
unb.adv.biomed.df %>%
  group_by(natureza) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = natureza, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "blue")+
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5)+
  theme_minimal()
```



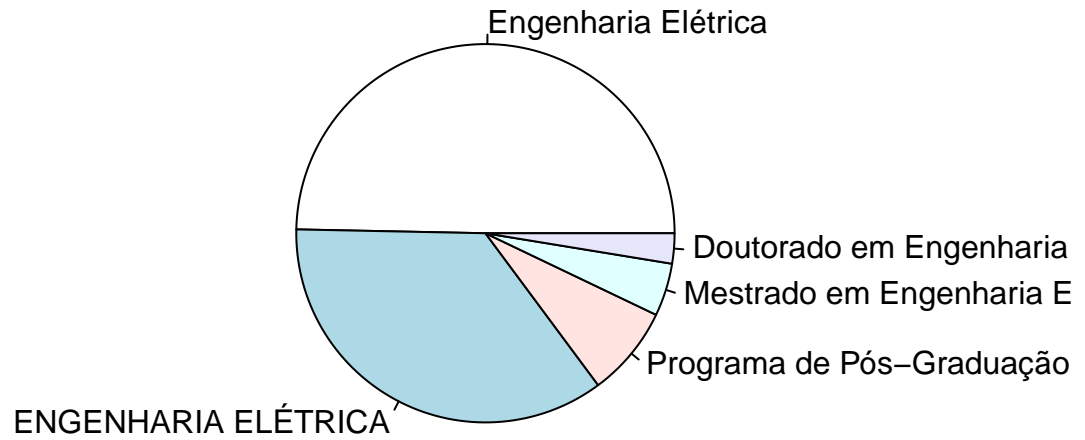
Cursos mais presentes:

`cursos_d5`

```
##
##          Engenharia Elétrica
##                77
##          ENGENHARIA ELÉTRICA
##                55
## Programa de Pós-Graduação em Engenharia Elétrica
##                12
##          Mestrado em Engenharia Elétrica
##                7
##          Doutorado em Engenharia Elétrica
##                4
```

```
pie(cursos_d5,clockwise=F,main="Cursos mais presentes")
```

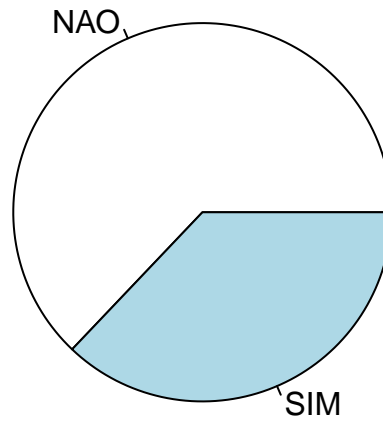
Cursos mais presentes



Quantidade de bolsistas e não bolsistas:

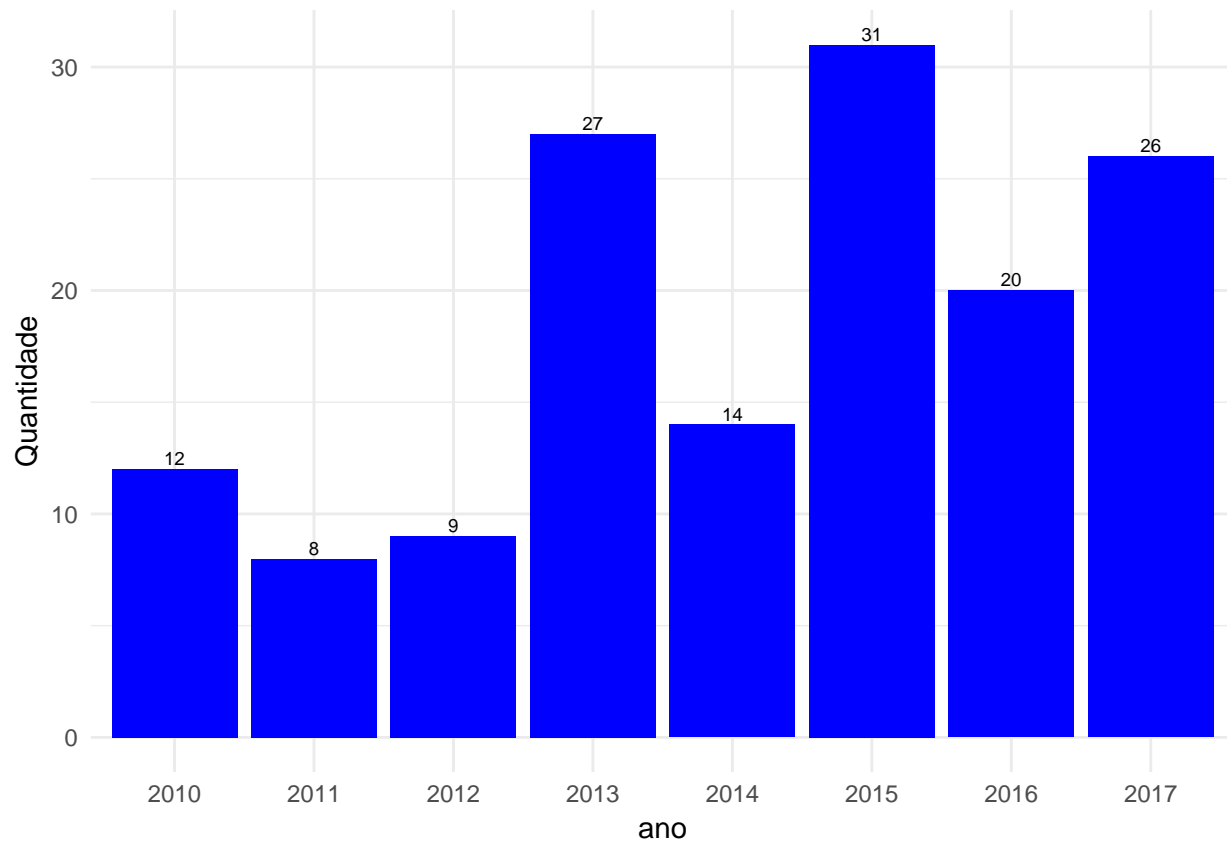
```
pie(table(unb.adv.biomed.df$bolsa),main="Bolsas")
```

Bolsas



Quantidade de publicações feitas por ano:

```
unb.pub.biomed.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```

Periódicos com mais publicações:

```
d <- table(unb.pub.biomed.df$periodico)
head(sort(d,decreasing = TRUE),n=5)
```

```
##
##          Revista Brasileira de Engenharia Biomédica (Impresso)
##                                     7
## Global Journal of Engineering Science and Research Management
##                                     6
##          Biomedical Engineering Online (Online)
##                                     4
##          Revista Brasileira de Ensino de Física (Online)
##                                     4
##          Revista Brasileira de Inovação Tecnológica em Saúde
##                                     4
```

Autores com mais participações em publicações:

```
head(sort(table(toupper(unlist(strsplit(unb.pub.biomed.df$autores,";")))),decreasing = TRUE),n=10)
```

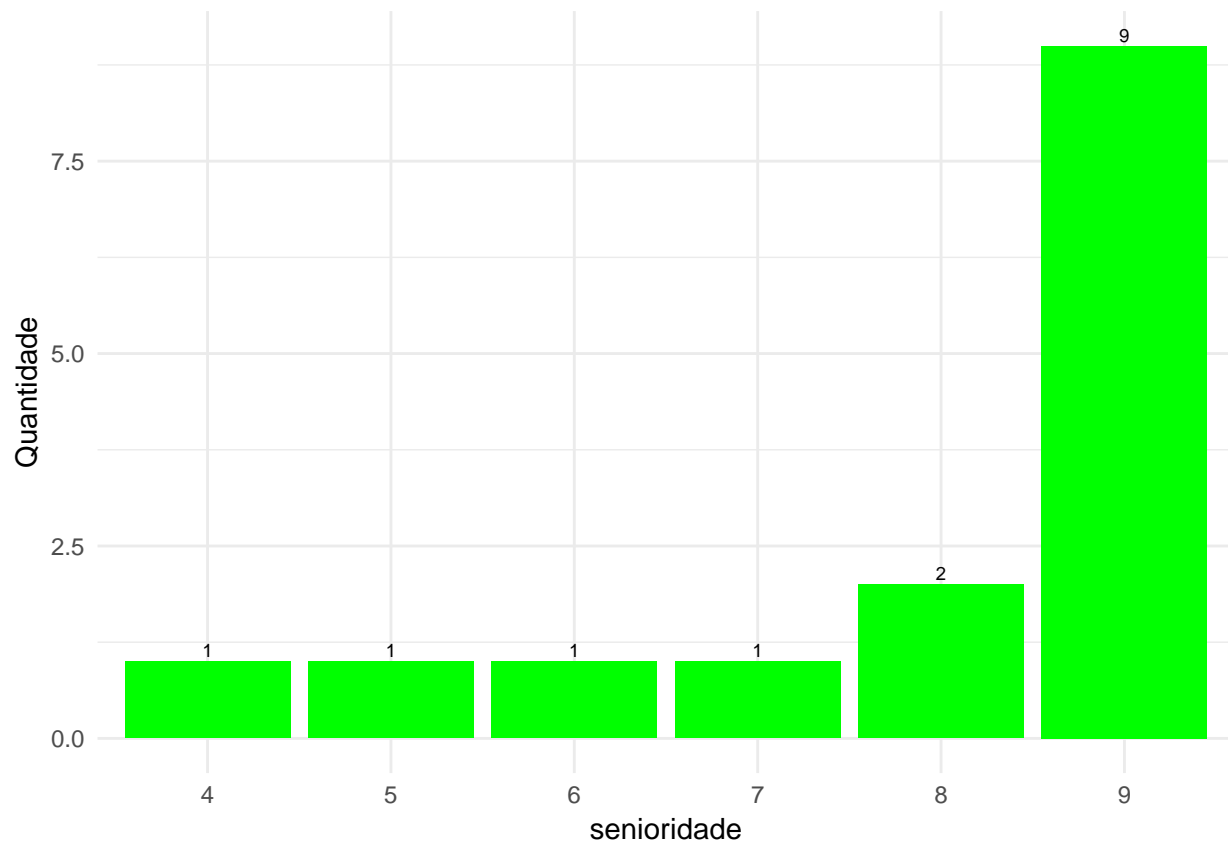
```
##
##          AMORIM, R. G. G.
##                                     17
##          AMORIM, R. G. G.
##                                     16
##          MARÃES, V. R. F. S.
##                                     14
```

```
##                BRASIL, L. M.
##                7
##            MARÃES, V. R. F. S.
##                7
##                SILVA, W. B.
##                7
##            GOMES, MARÍLIA MIRANDA FORTE
##                7
## ROSA, SUÉLIA DE SIQUEIRA RODRIGUES FLEURY
##                7
##                DA ROCHA, A. F.
##                6
##            GOMES, MARÍLIA MIRANDA FORTE
##                6
```

Vemos que há algumas incoerência neste dado acima, isto se dá pela falta de uniformidade dos dados.

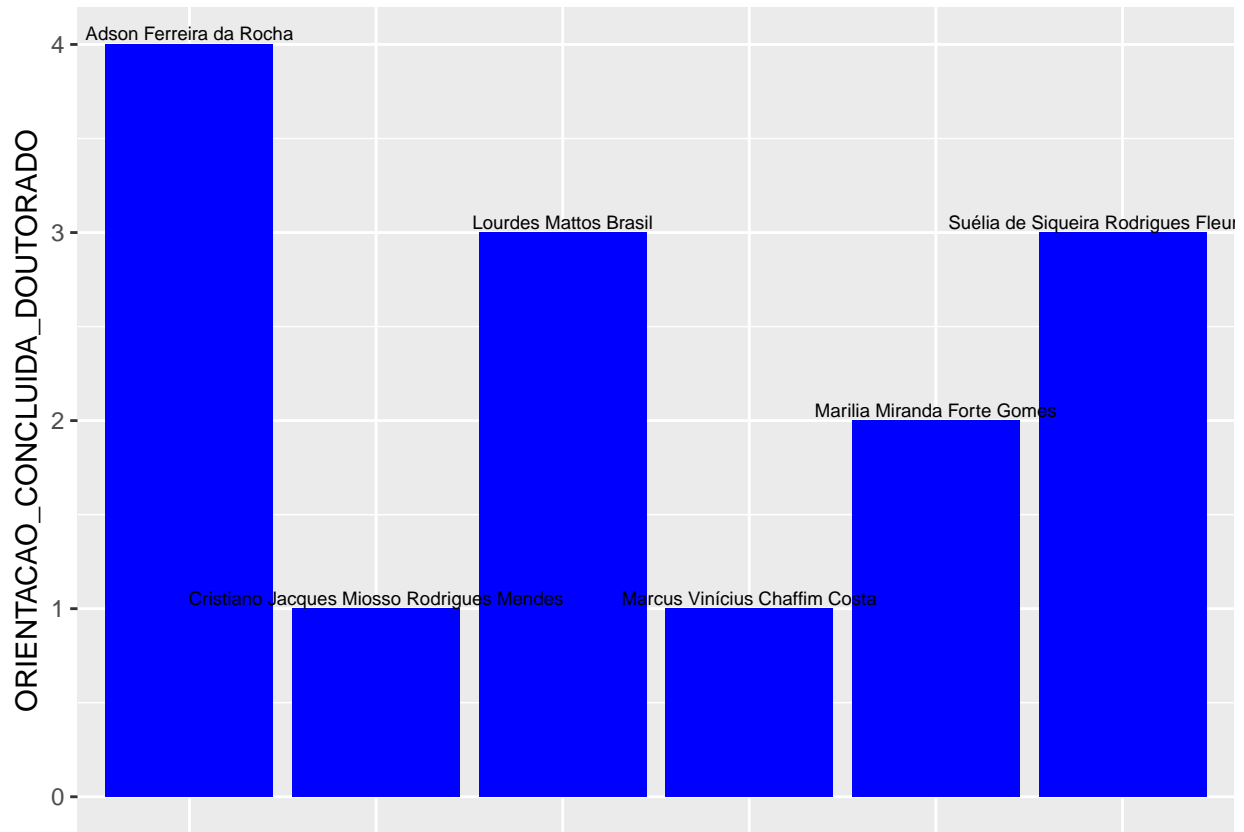
Perfis de docentes por senioridade:

```
unb.prof.biomed.df %>%
  group_by(senioridade) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = senioridade, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "green") +
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +
  theme_minimal()
```



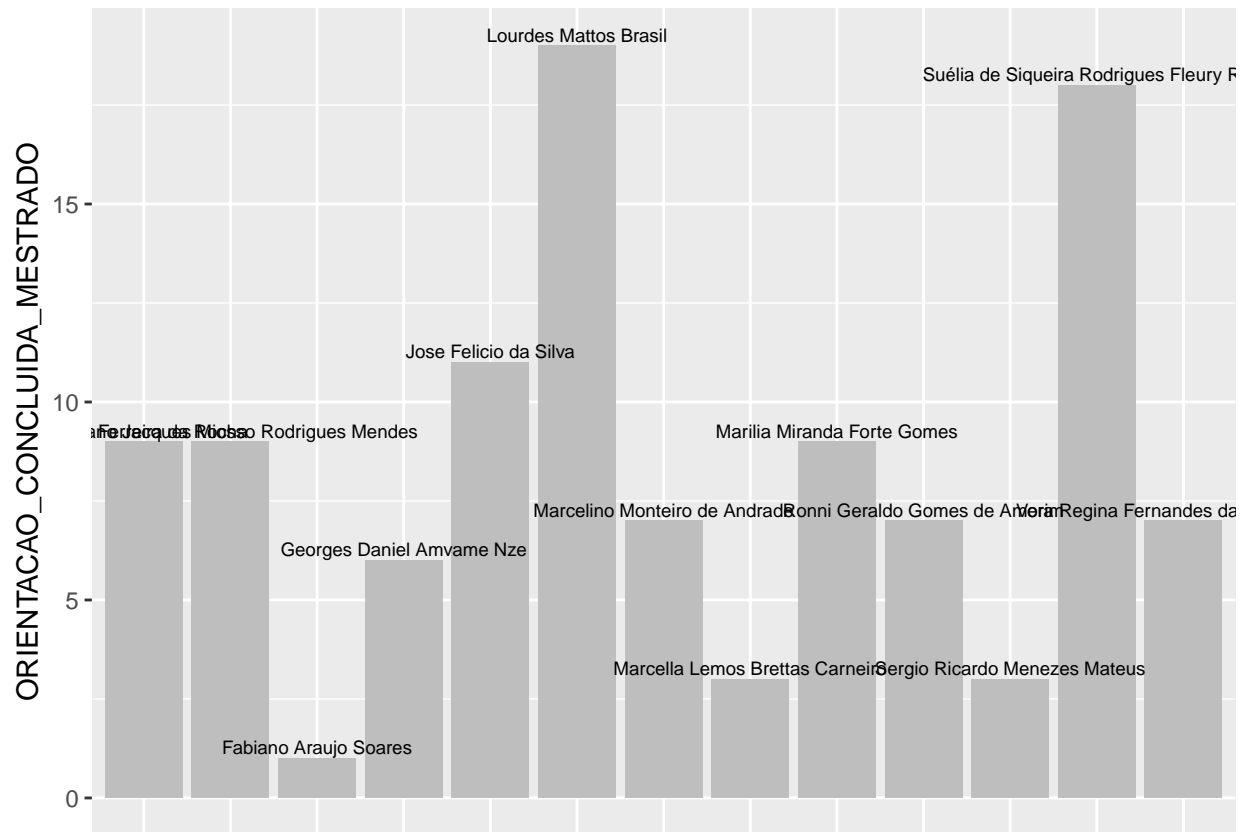
Quantidade de orientações concluídas no doutorado por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_DOUTORADO))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_DOUTORADO)) +
geom_bar(position = "stack",stat = "identity", fill = "blue")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```



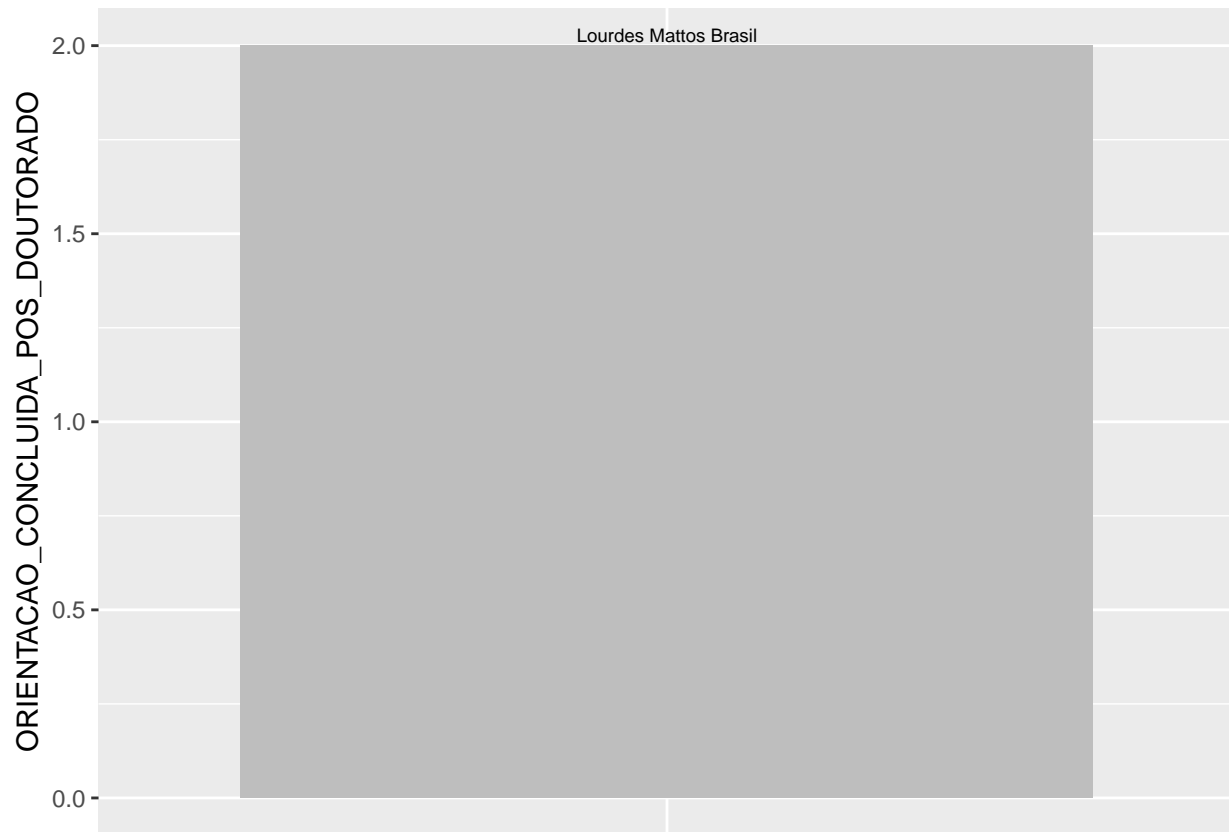
Quantidade de orientações concluídas no mestrado por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_MESTRADO ))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_MESTRADO)) +
geom_bar(position = "stack",stat = "identity", fill = "gray")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```



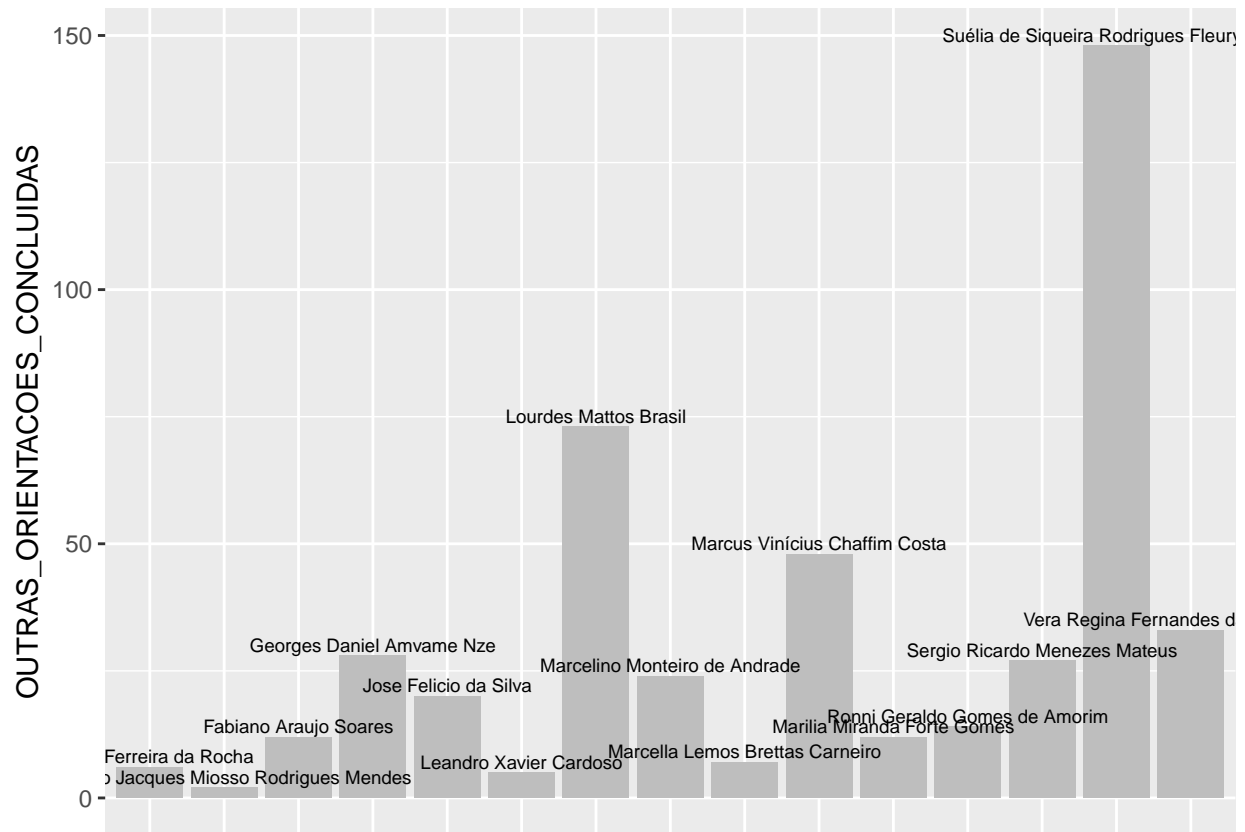
Quantidade de orientações concluídas no pós-doutorado por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_POS_DOUTORADO))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_POS_DOUTORADO)) +
geom_bar(position = "stack",stat = "identity", fill = "gray")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```



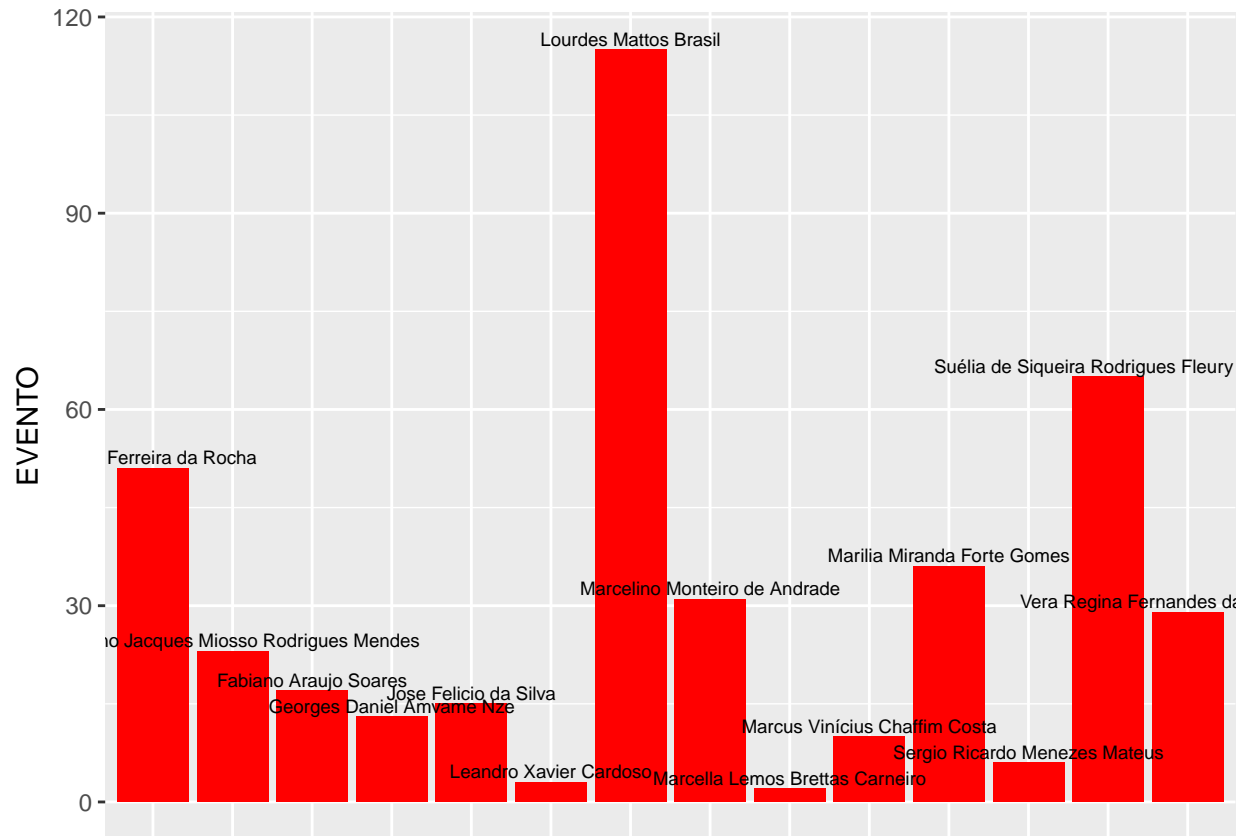
Quantidade de orientações concluídas em outros tipos por professor:

```
unb.prof.biomed.df %>%
  group_by(nome) %>%
  filter(!is.na(OUTRAS_ORIENTACOES_CONCLUIDAS))%>%
  ggplot(aes(x = nome, y = OUTRAS_ORIENTACOES_CONCLUIDAS)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



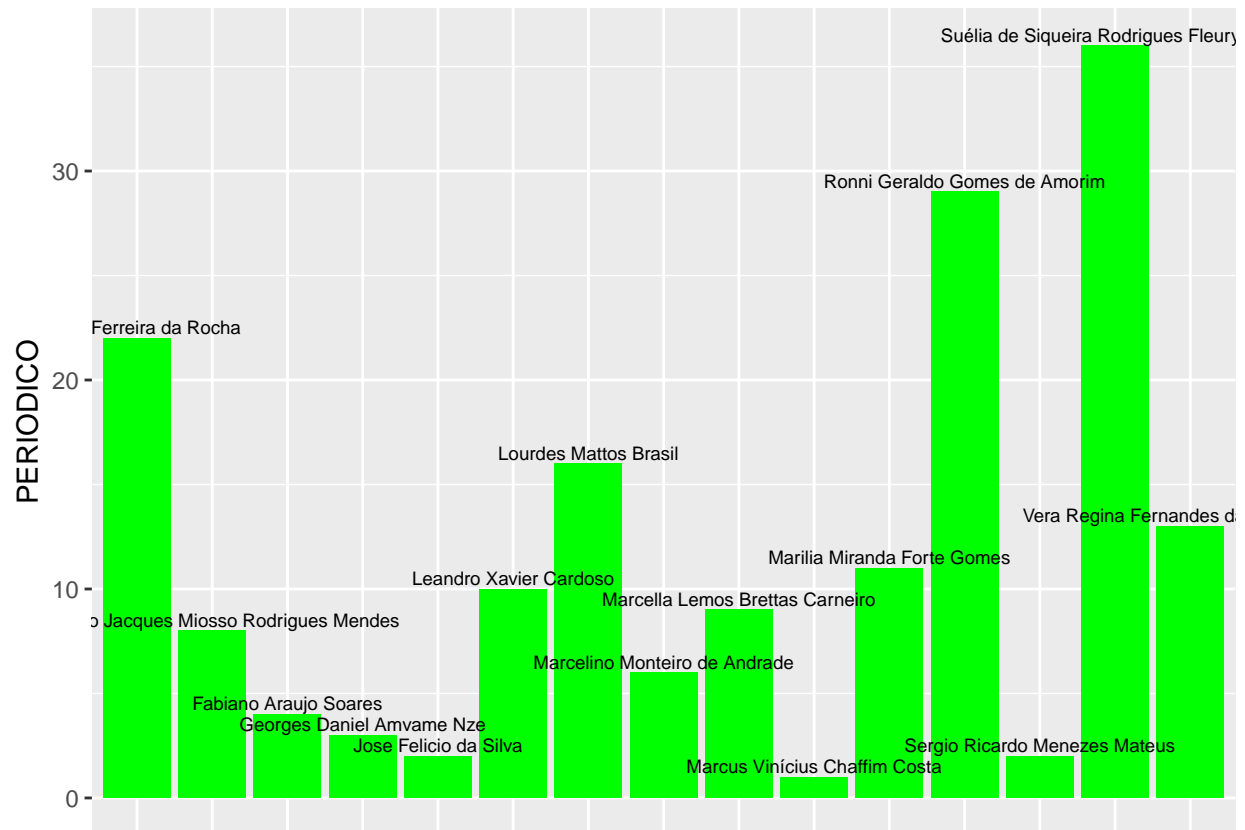
Quantidade de eventos por professor:

```
unb.prof.biomed.df %>%
  group_by(nome) %>%
  filter(!is.na(EVENTO))%>%
  ggplot(aes(x = nome, y = EVENTO)) +
  geom_bar(position = "stack", stat = "identity", fill = "red")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



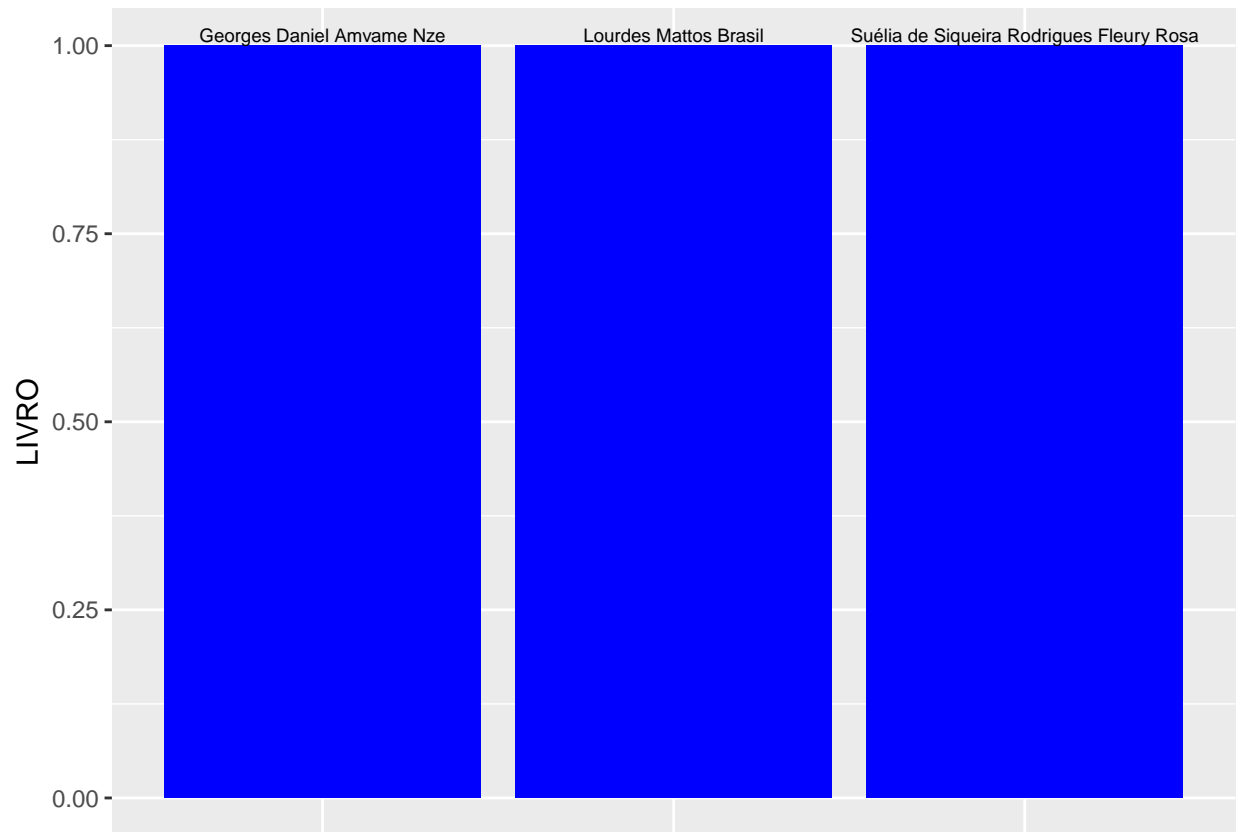
Quantidade de periódicos por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(PERIODICO))%>%
ggplot(aes(x = nome, y = PERIODICO)) +
geom_bar(position = "stack",stat = "identity", fill = "green")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank())
```



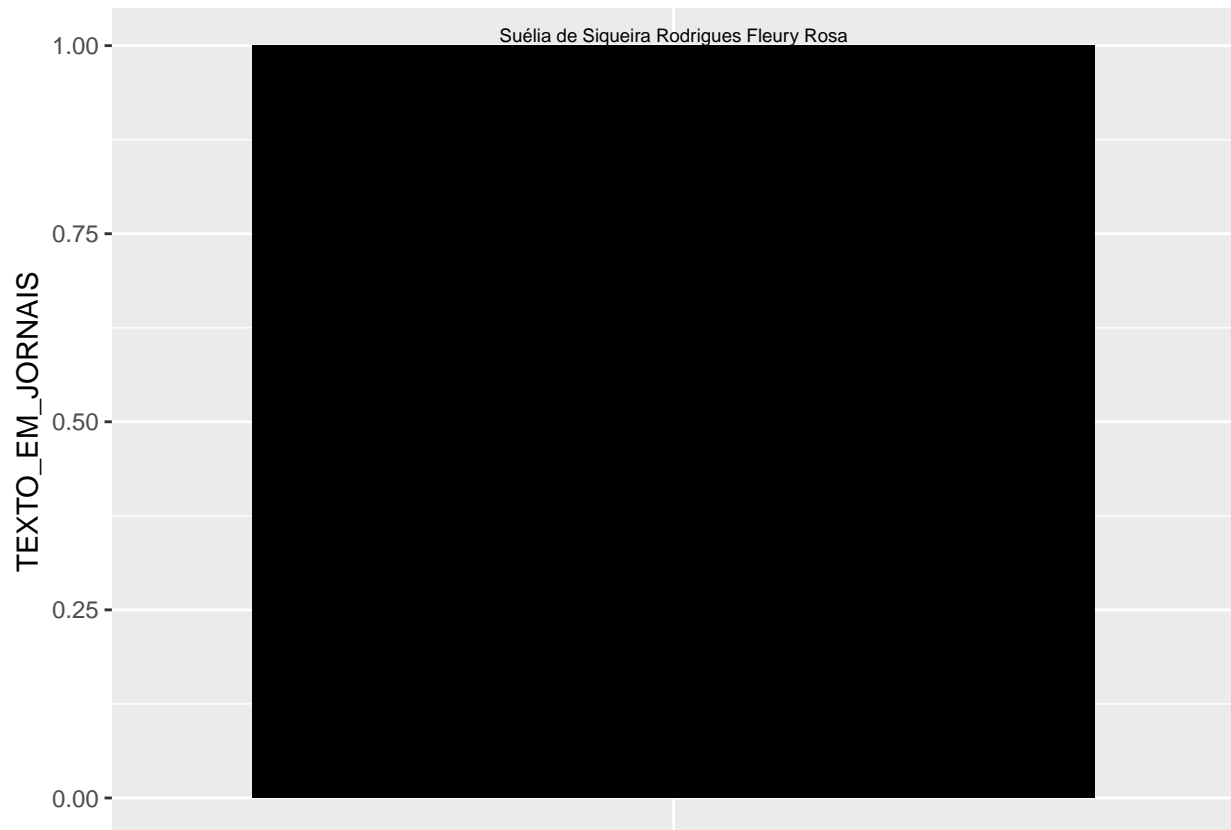
Quantidade de livros por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(LIVRO))%>%
ggplot(aes(x = nome, y = LIVRO)) +
geom_bar(position = "stack",stat = "identity", fill = "blue")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank())
```

Quantidade de texto em jornais por professor:

```
unb.prof.biomed.df %>%
group_by(nome) %>%
filter(!is.na(TEXTO_EM_JORNAIS))%>%
ggplot(aes(x = nome, y = TEXTO_EM_JORNAIS)) +
geom_bar(position = "stack",stat = "identity", fill = "black")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```



Relações entre professores:

```
list <- c(rbind(unb.graph.biomed$links$source, unb.graph.biomed$links$target))
idx <- match(list, unb.graph.biomed$nodes$id)
names <- unb.graph.biomed$nodes$properties$name
graph_names <- names[idx]
g <- graph(graph_names, directed=TRUE)
plot(g, layout = layout_(g, nicely()),
     edge.width = 1,
     edge.arrow.width = 1,
     vertex.size = 5,
     edge.arrow.size = 1,
     vertex.size2 = 3,
     vertex.label.cex = 1,
     asp = 0.35,
     margin = -0.1)
```



Nuvem de palavras baseadas no resumo do currículo dos professores:

```
unb.prof.biomed.df$resumo_cv <- iconv(unb.prof.biomed.df$resumo_cv,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.prof.biomed.df$idLattes, text = unb.prof.biomed.df$resumo_cv, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words = (d$word), freq = d$freq, min.freq = 3,
          max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
```



Nuvem de palavras baseadas no título das orientações:

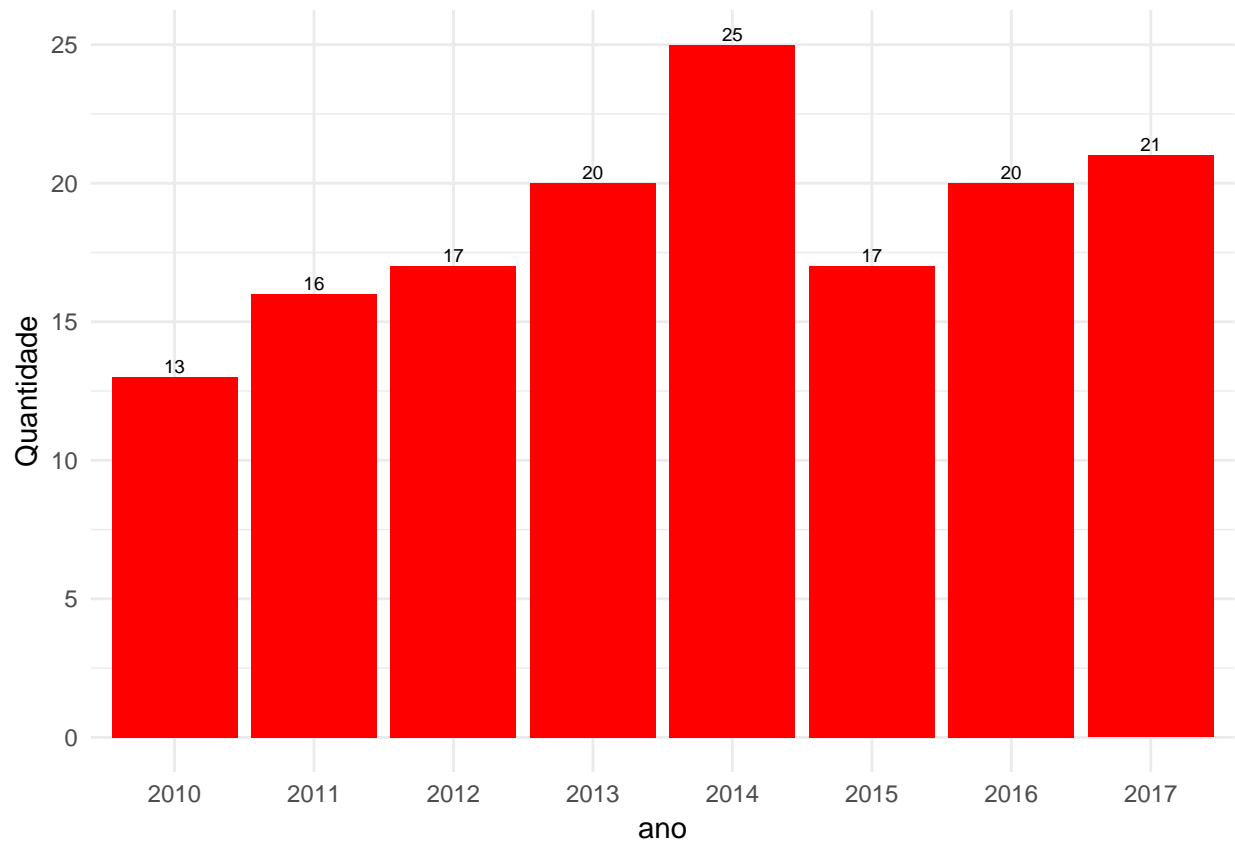
```
unb.adv.biomed.df$titulo <- iconv(unb.adv.biomed.df$titulo,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.adv.biomed.df$idLattes1, text = unb.adv.biomed.df$titulo, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, stopwords("english"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words = (d$word), freq = d$freq, min.freq = 3,
          max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
```



Engenharia de Sistemas Eletrônicos e de Automação

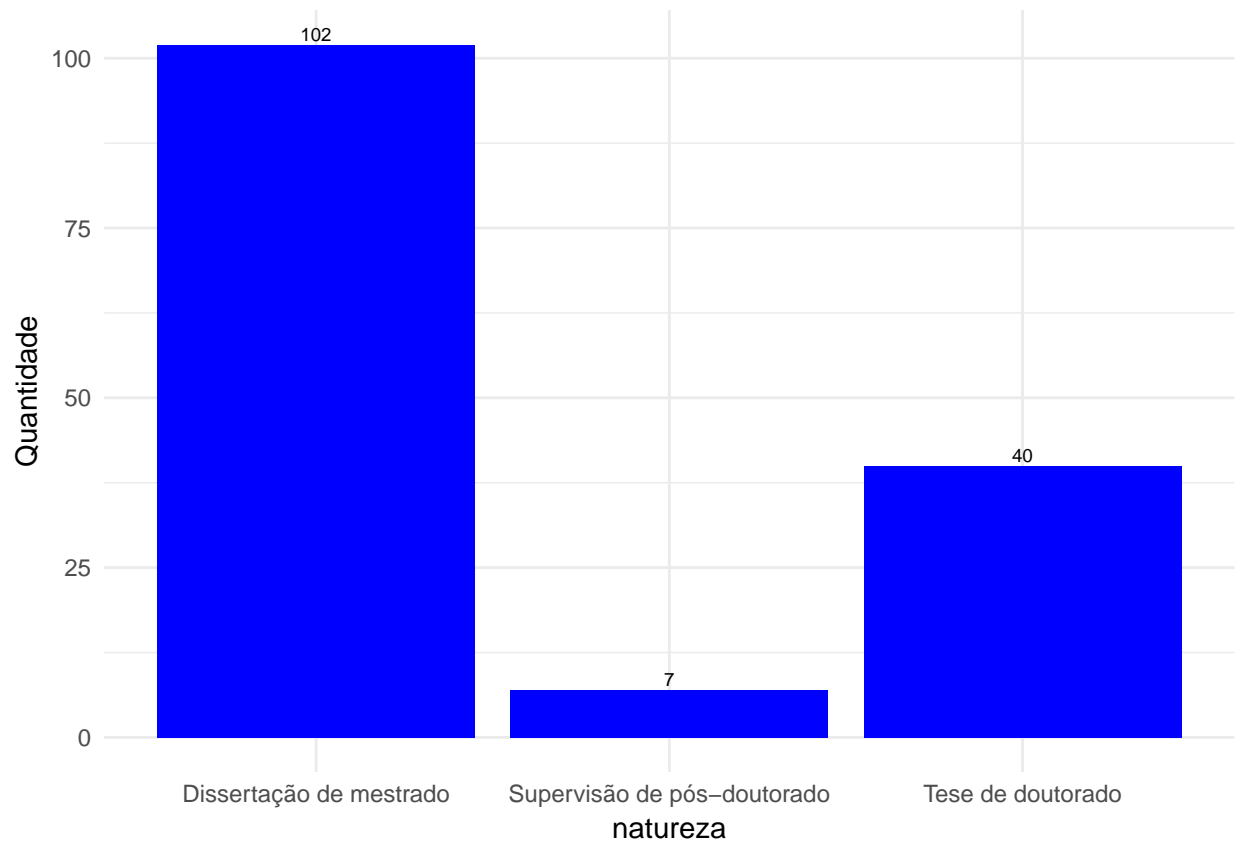
Quantidade de orientações feitas por ano:

```
unb.adv.auto.df %>%
  group_by(ano) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = ano, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "red") +
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +
  theme_minimal()
```



Quantidade de orientações separadas por tipo de publicação:

```
unb.adv.auto.df %>%  
group_by(natureza) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = natureza, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



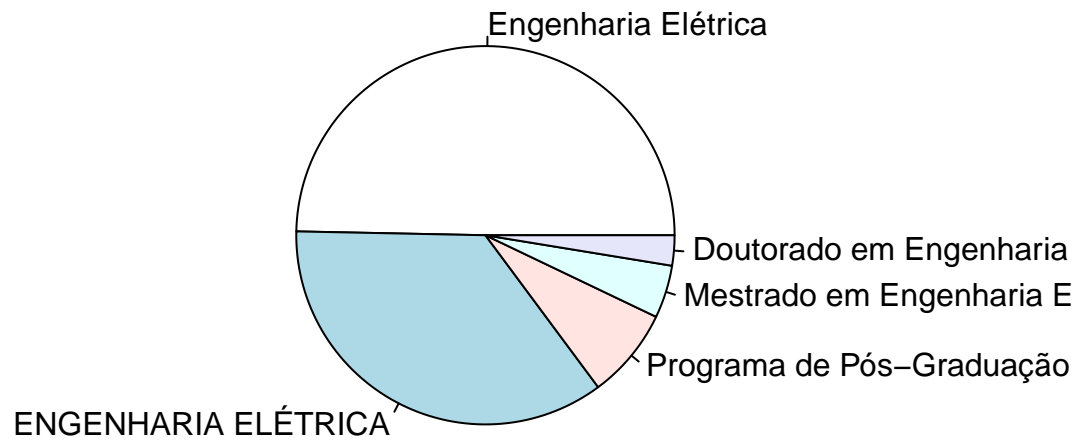
Cursos mais presentes:

`cursos_d5`

```
##
##          Engenharia Elétrica
##                77
##          ENGENHARIA ELÉTRICA
##                55
## Programa de Pós-Graduação em Engenharia Elétrica
##                12
##          Mestrado em Engenharia Elétrica
##                7
##          Doutorado em Engenharia Elétrica
##                4
```

```
pie(cursos_d5,clockwise=F,main="Cursos mais presentes")
```

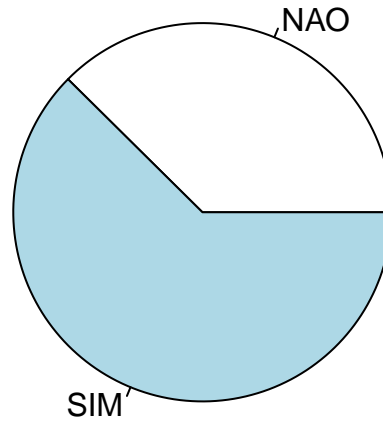
Cursos mais presentes



Quantidade de bolsistas e não bolsistas:

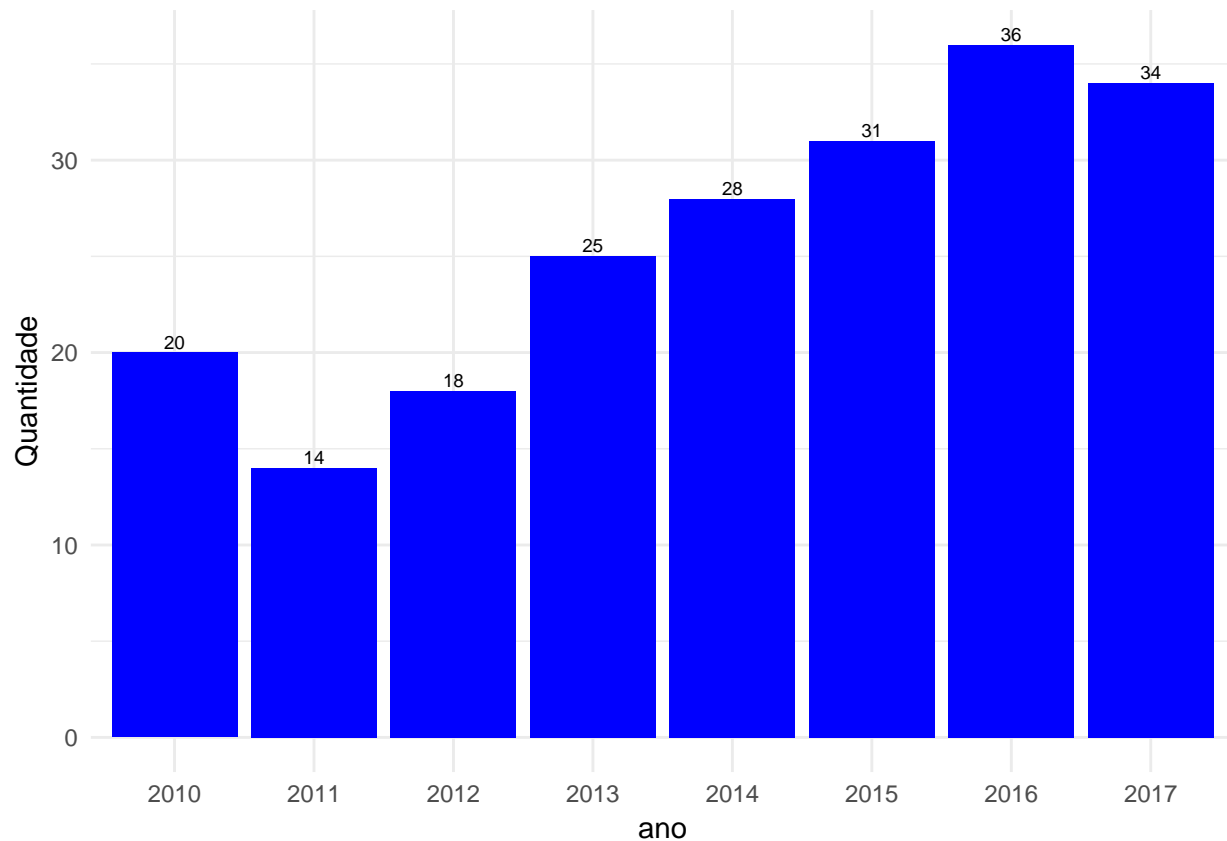
```
pie(table(unb.adv.auto.df$bolsa),main="Bolsas")
```


Bolsas



Quantidade de publicações feitas por ano:

```
unb.pub.auto.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



Periódicos com mais publicações:

```
d <- table(unb.pub.auto.df$periodico)
head(sort(d,decreasing = TRUE),n=5)
```

```
##
## Revista Brasileira de Engenharia Biomédica (Impresso)
## 7
## IEEE Signal Processing Letters
## 5
## Journal of Control, Automation and Electrical Systems
## 5
## Electronics Letters
## 4
## INTERNATIONAL JOURNAL OF ROBUST AND NONLINEAR CONTROL
## 4
```

Autores com mais participações em publicações:

```
head(sort(table(toupper(unlist(strsplit(unb.pub.auto.df$autores,";")))),decreasing = TRUE),n=10)
```

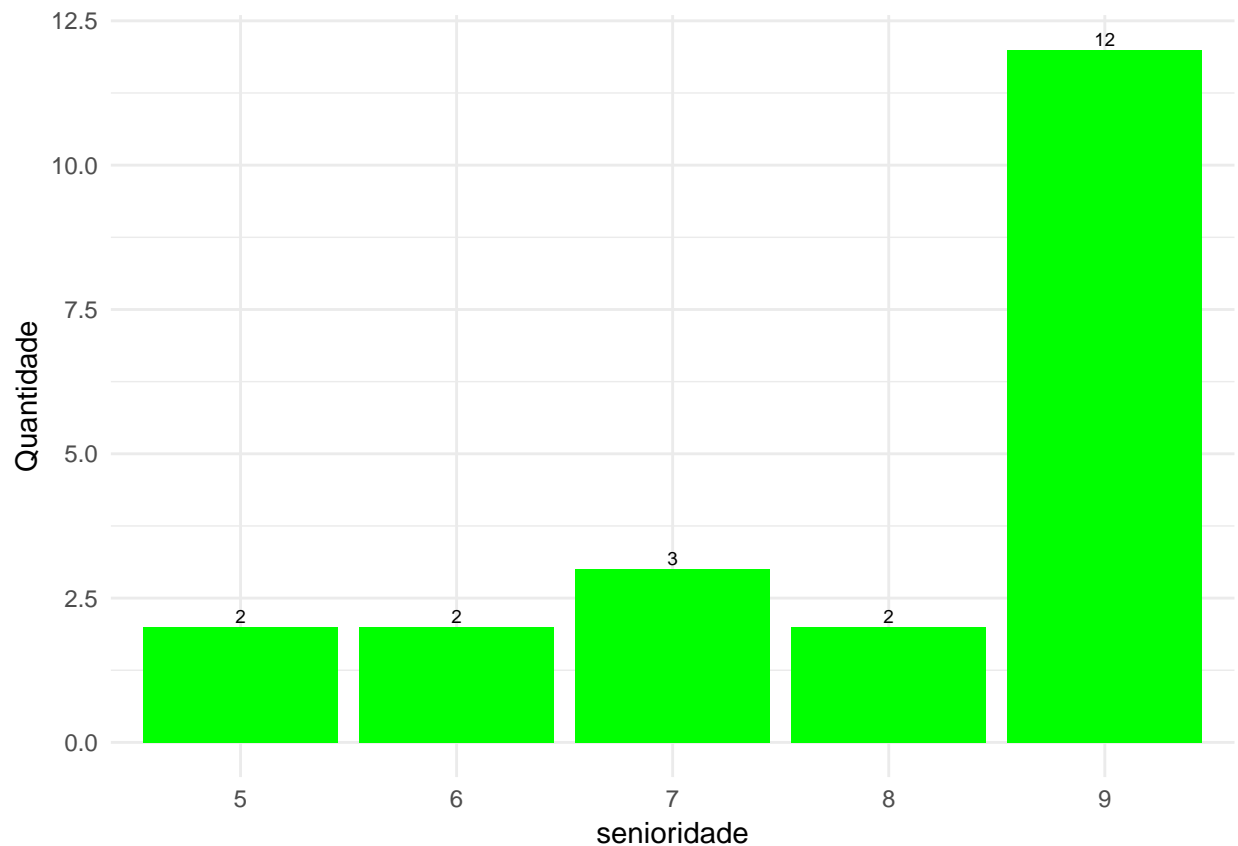
```
##
## BOTTARO, MARTIM
## 12
## ISHIHARA, J. Y.
## 12
## DE QUEIROZ, RICARDO
## 12
```

```
##          DE QUEIROZ, RICARDO L.
##                      11
##          CLAUS, MARTIN
##                      10
##          PERES, P. L. D.
##                      10
##          QUEIROZ, R.L.
##                      10
##          DE QUEIROZ, R. L.
##                      9
##          ISHIHARA, JOÃO Y.
##                      9
##  NASCIMENTO, FRANCISCO ASSIS DE OLIVEIRA
##                      9
```

Vemos que há algumas incoerência neste dado acima, isto se dá pela falta de uniformidade dos dados.

Perfis de docentes por senioridade:

```
unb.prof.auto.df %>%
  group_by(senioridade) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = senioridade, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "green")+
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5)+
  theme_minimal()
```

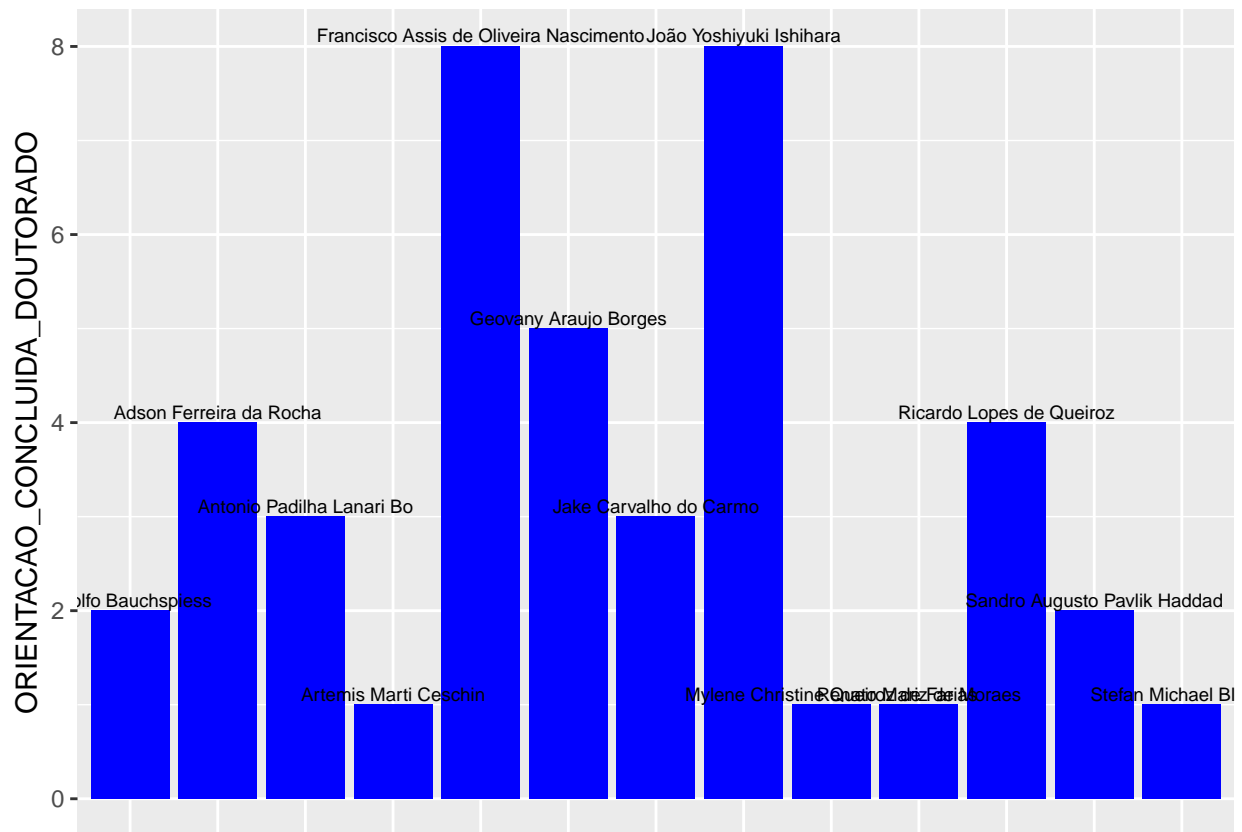


Quantidade de orientações concluídas no doutorado por professor:

```

unb.prof.auto.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_DOUTORADO))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_DOUTORADO)) +
geom_bar(position = "stack",stat = "identity", fill = "blue")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())

```

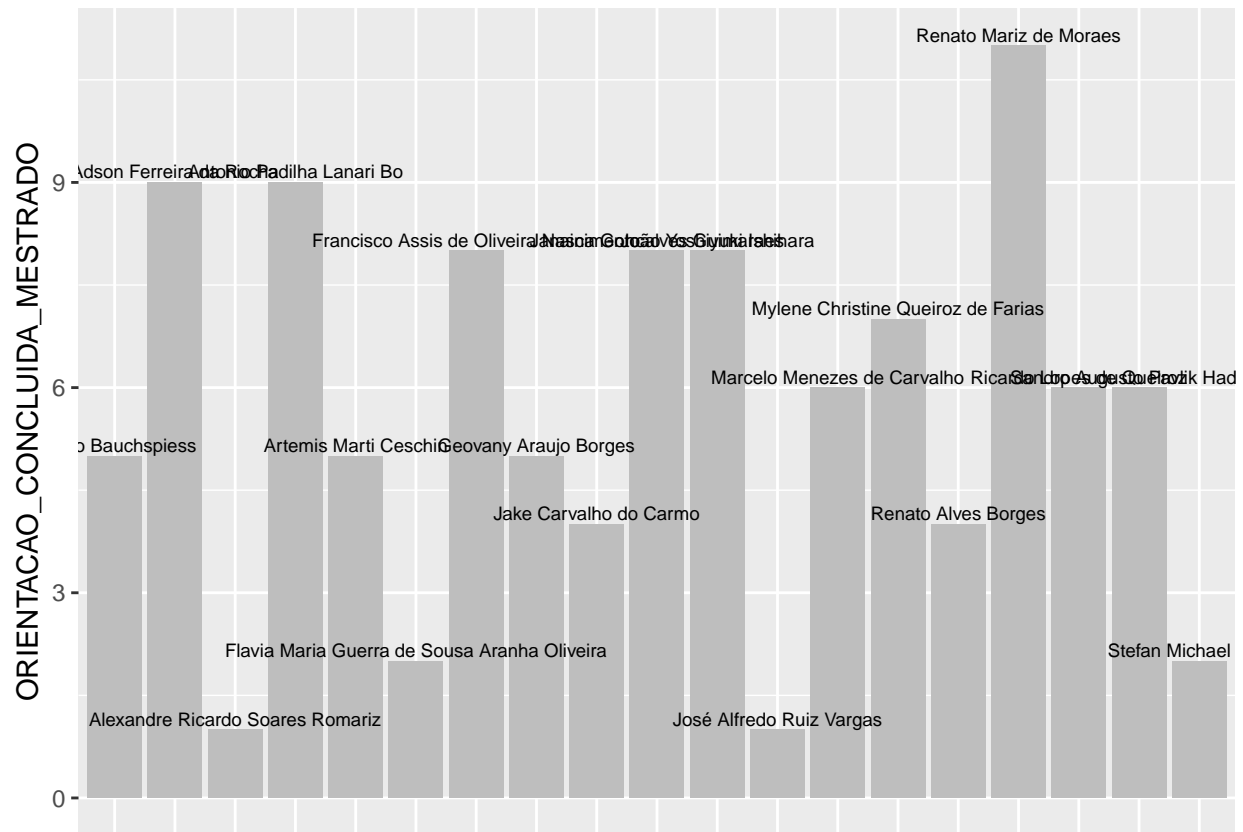


Quantidade de orientações concluídas no mestrado por professor:

```

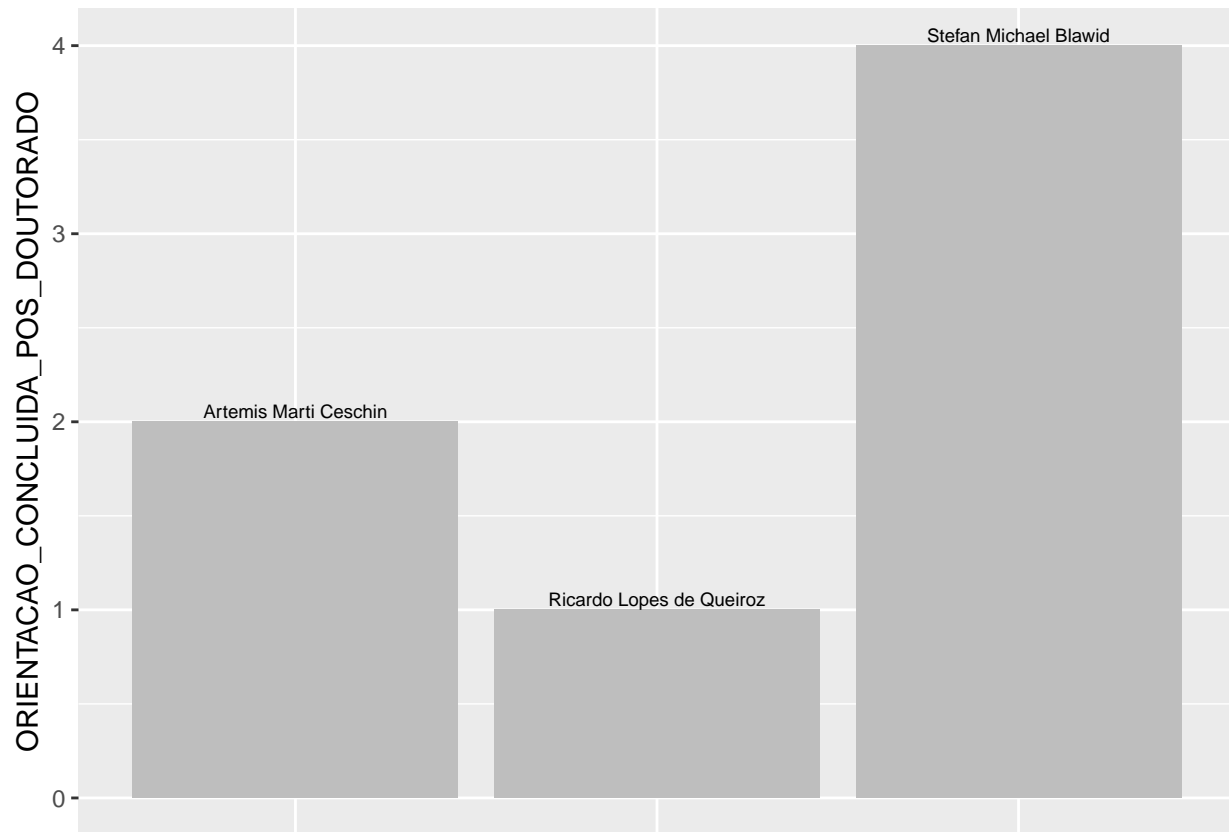
unb.prof.auto.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_MESTRADO ))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_MESTRADO)) +
geom_bar(position = "stack",stat = "identity", fill = "gray")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())

```



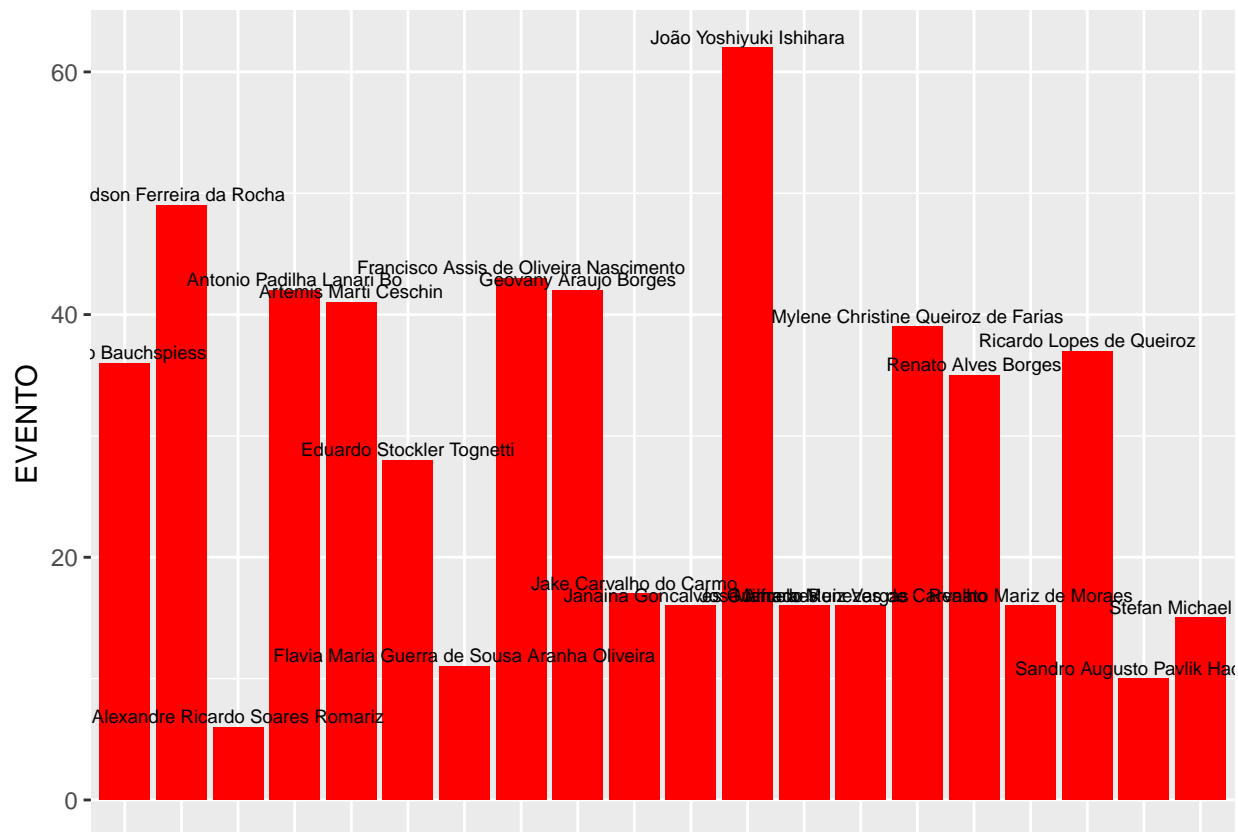
Quantidade de orientações concluídas no pós-doutorado por professor:

```
unb.prof.auto.df %>%
group_by(nome) %>%
filter(!is.na(ORIENTACAO_CONCLUIDA_POS_DOUTORADO))%>%
ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_POS_DOUTORADO)) +
geom_bar(position = "stack",stat = "identity", fill = "gray")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
      axis.text.x=element_blank(),
      axis.ticks.x=element_blank())
```



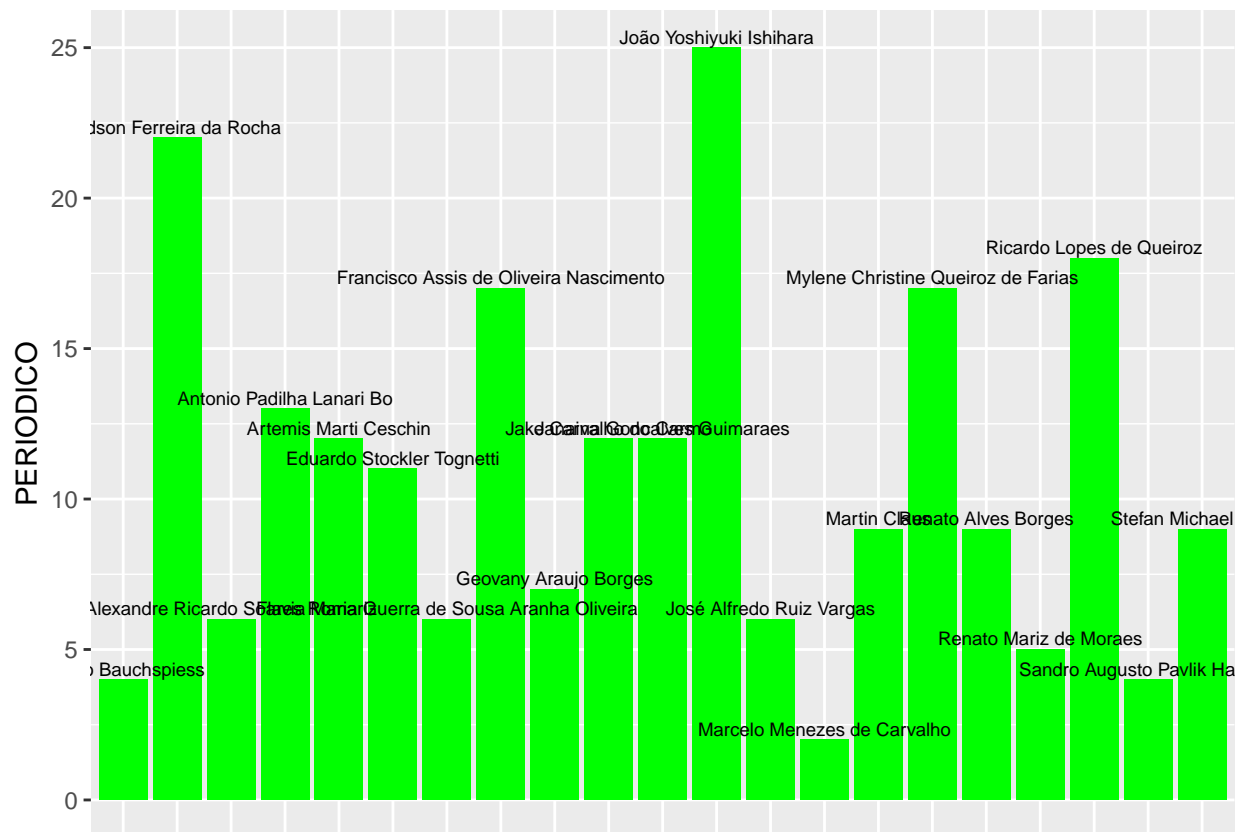
Quantidade de orientações concluídas em outros tipos por professor:

```
unb.prof.auto.df %>%
  group_by(nome) %>%
  filter(!is.na(OUTRAS_ORIENTACOES_CONCLUIDAS))%>%
  ggplot(aes(x = nome, y = OUTRAS_ORIENTACOES_CONCLUIDAS)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```

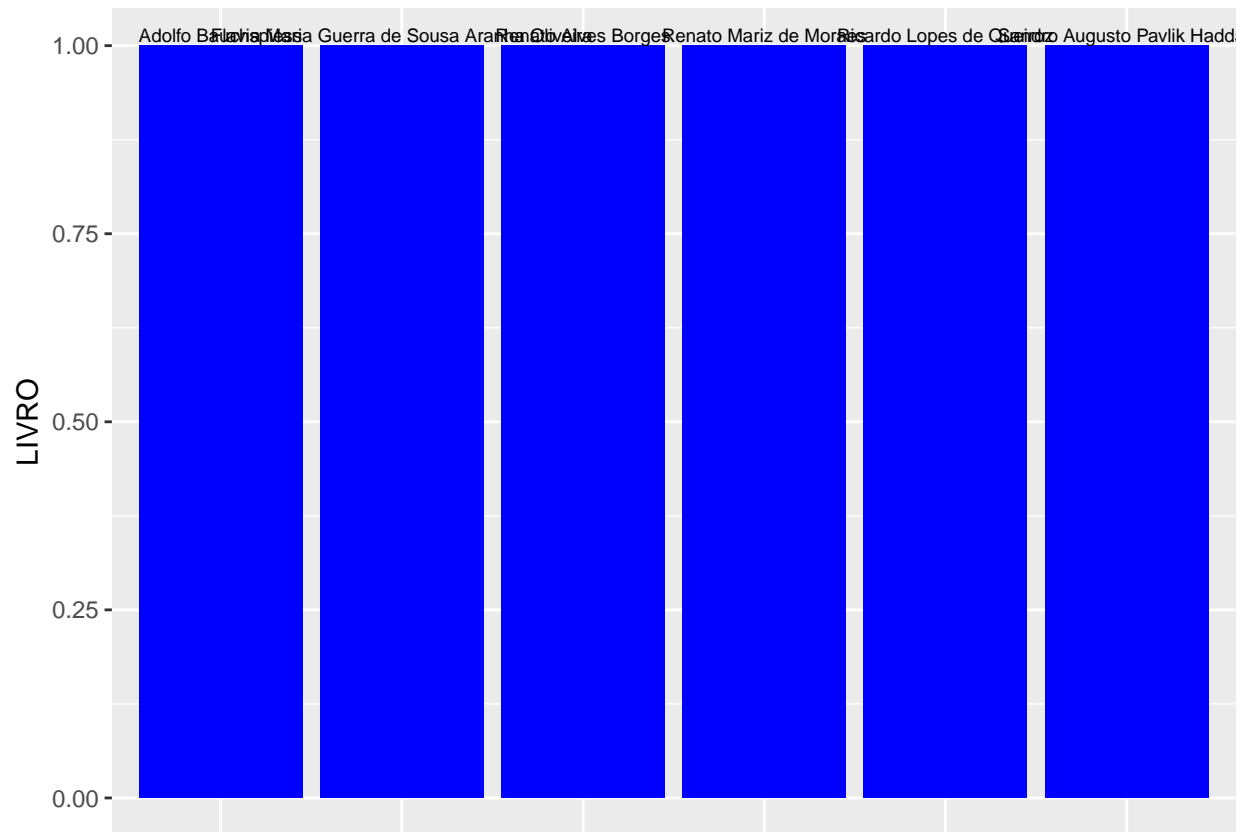
Quantidade de periódicos por professor:

```
unb.prof.auto.df %>%
  group_by(nome) %>%
  filter(!is.na(PERIODICO))%>%
  ggplot(aes(x = nome, y = PERIODICO)) +
  geom_bar(position = "stack",stat = "identity", fill = "green")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```

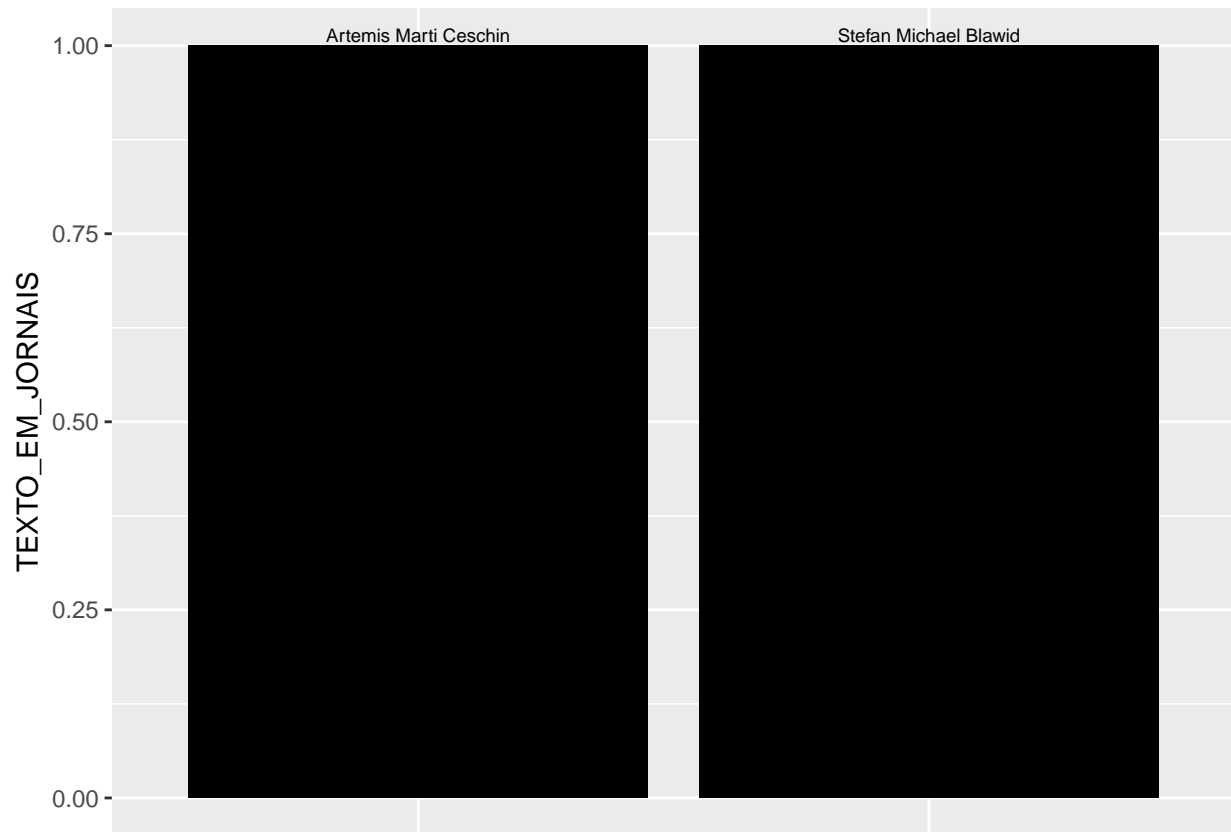
Quantidade de livros por professor:

```
unb.prof.auto.df %>%
group_by(nome) %>%
filter(!is.na(LIVRO))%>%
ggplot(aes(x = nome, y = LIVRO)) +
geom_bar(position = "stack",stat = "identity", fill = "blue")+
geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
theme(axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank())
```



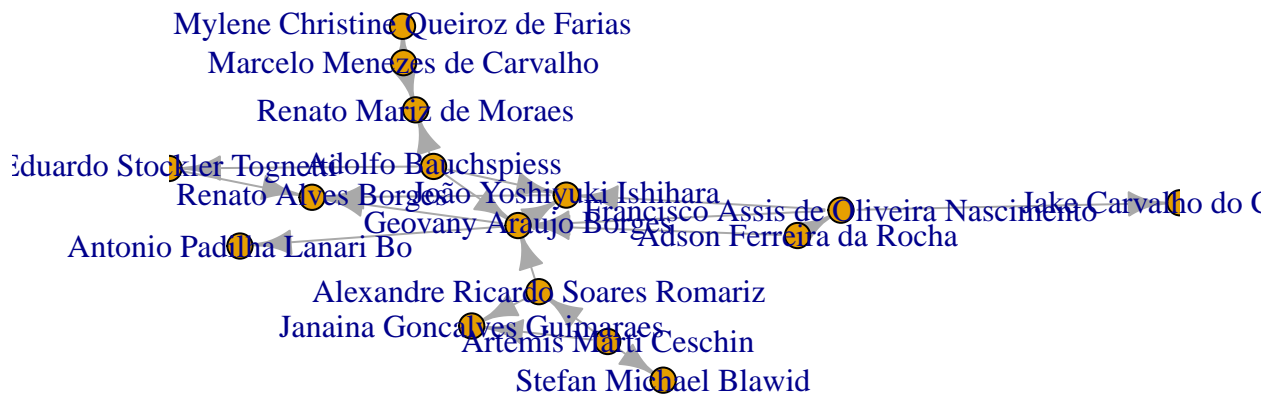
Quantidade de texto em jornais por professor:

```
unb.prof.auto.df %>%
  group_by(nome) %>%
  filter(!is.na(TEXTO_EM_JORNAIS))%>%
  ggplot(aes(x = nome, y = TEXTO_EM_JORNAIS)) +
  geom_bar(position = "stack", stat = "identity", fill = "black")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



Relações entre professores:

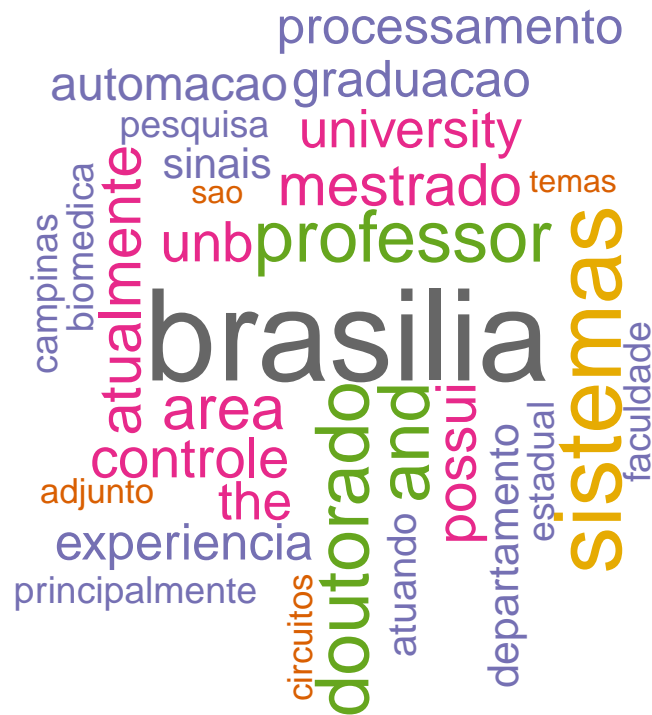
```
list <- c(rbind(unb.graph.auto$links$source, unb.graph.auto$links$target))
idx <- match(list, unb.graph.auto$nodes$id)
names <- unb.graph.auto$nodes$properties$name
graph_names <- names[idx]
g <- graph(graph_names, directed=TRUE)
plot(g, layout = layout_(g, nicely()),
     edge.width = 1,
     edge.arrow.width = 1,
     vertex.size = 5,
     edge.arrow.size = 1,
     vertex.size2 = 3,
     vertex.label.cex = 1,
     asp = 0.35,
     margin = -0.1)
```



Nuvem de palavras baseadas no resumo do currículo dos professores:

```

unb.prof.auto.df$resumo_cv <- iconv(unb.prof.auto.df$resumo_cv,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.prof.auto.df$idLattes, text = unb.prof.auto.df$resumo_cv, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words =(d$word), freq = d$freq, min.freq = 3,
          max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
  
```



Nuvem de palavras baseadas no título das orientações:

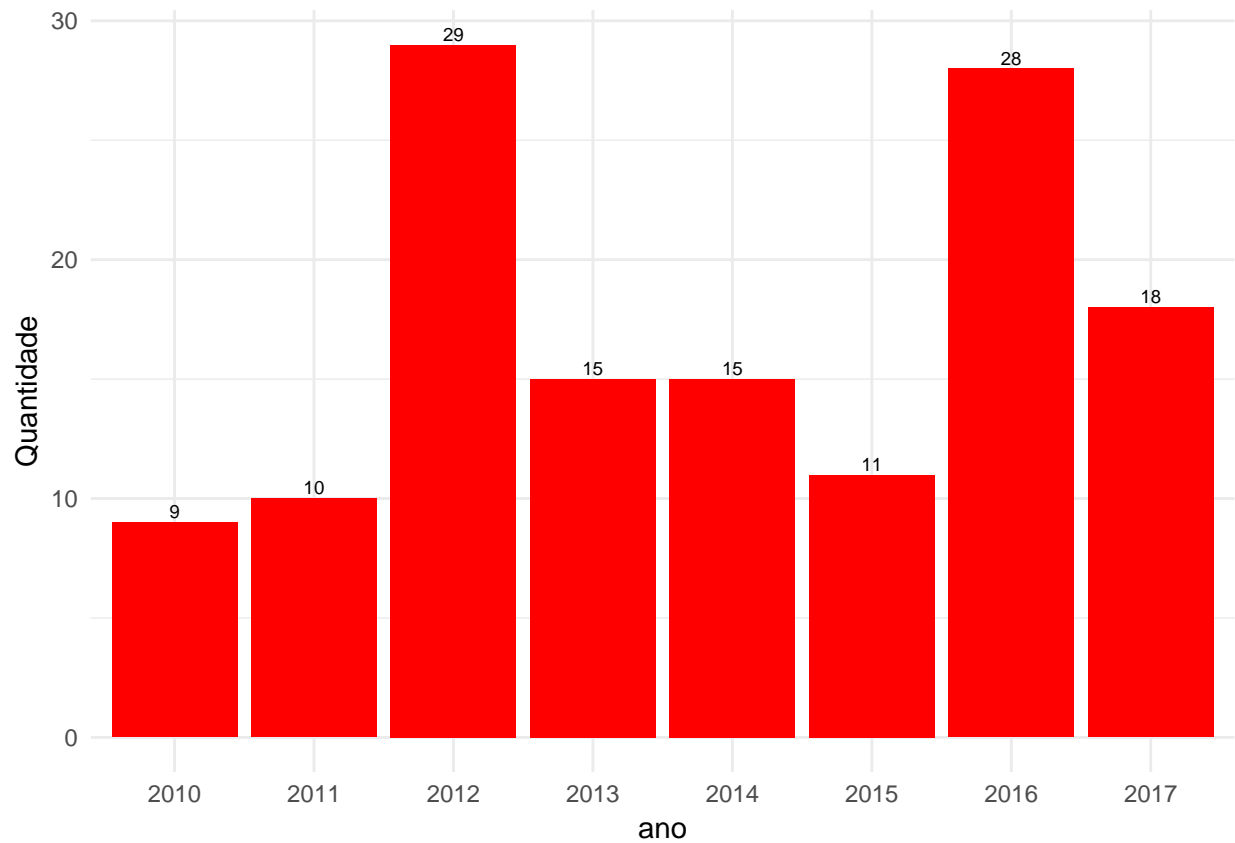
```
unb.adv.auto.df$titulo <- iconv(unb.adv.auto.df$titulo,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.adv.auto.df$idLattes1, text = unb.adv.auto.df$titulo, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, stopwords("english"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words = (d$word), freq = d$freq, min.freq = 3,
           max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
           colors=brewer.pal(8, "Dark2"))
```



Engenharia Elétrica P8

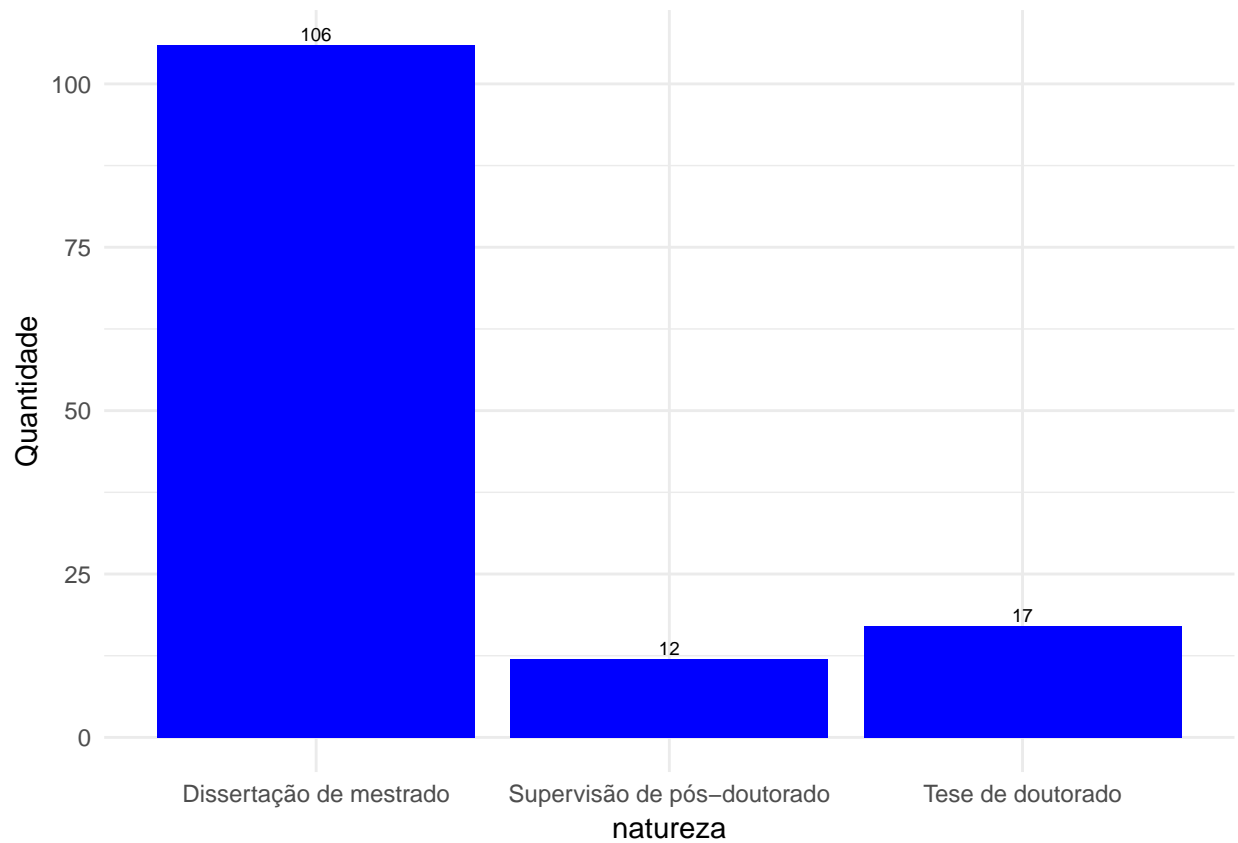
Quantidade de orientações feitas por ano:

```
unb.adv.ele_p8.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "red") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



Quantidade de orientações separadas por tipo de publicação:

```
unb.adv.ele_p8.df %>%  
group_by(natureza) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = natureza, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



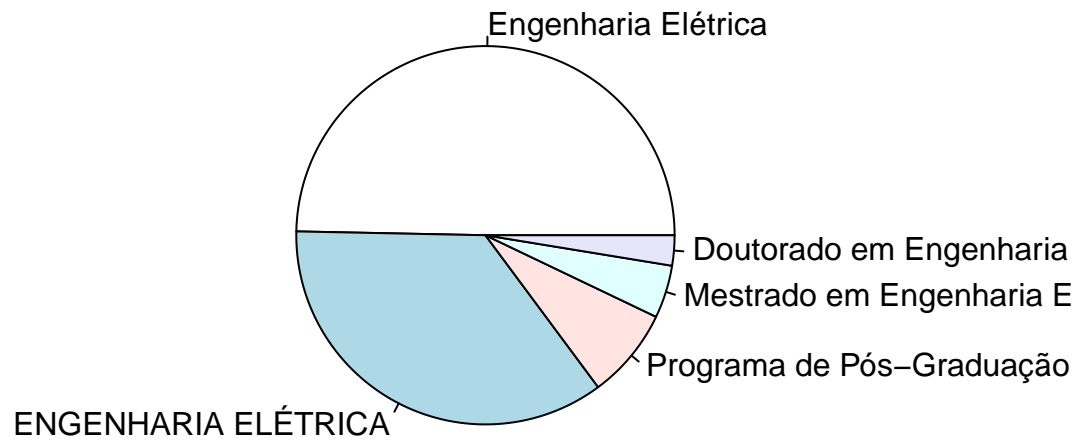
Cursos mais presentes:

`cursos_d5`

```
##
##          Engenharia Elétrica
##                77
##          ENGENHARIA ELÉTRICA
##                55
## Programa de Pós-Graduação em Engenharia Elétrica
##                12
##          Mestrado em Engenharia Elétrica
##                7
##          Doutorado em Engenharia Elétrica
##                4
```

```
pie(cursos_d5,clockwise=F,main="Cursos mais presentes")
```

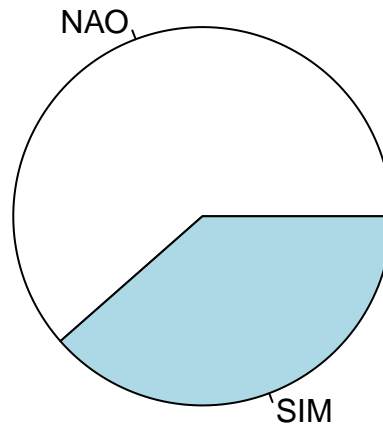

Cursos mais presentes



Quantidade de bolsistas e não bolsistas:

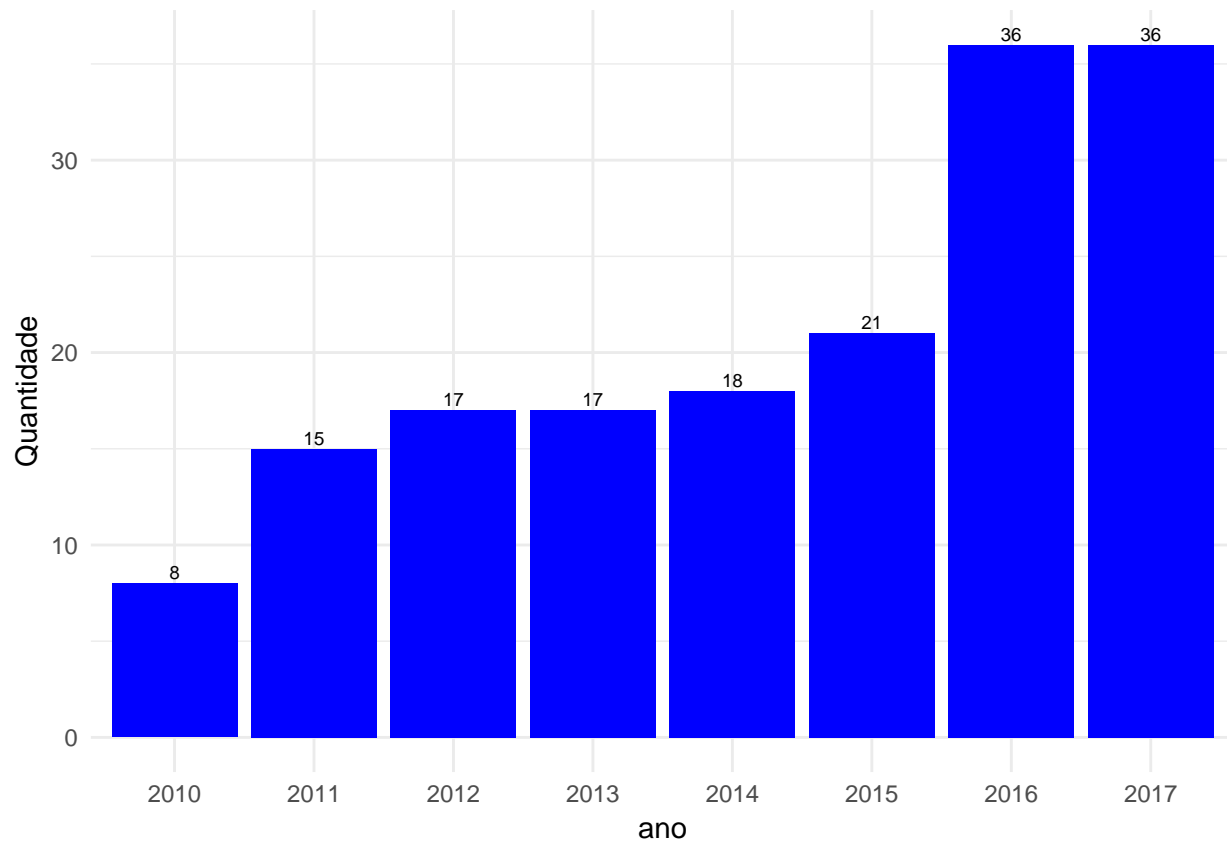
```
pie(table(unb.adv.ele_p8.df$bolsa),main="Bolsas")
```

Bolsas



Quantidade de publicações feitas por ano:

```
unb.pub.ele_p8.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



Periódicos com mais publicações:

```
d <- table(unb.pub.ele_p8.df$periodico)
head(sort(d,decreasing = TRUE),n=5)
```

```
##
##      IEEE Transactions on Information Forensics and Security
##                                     14
##      Digital Signal Processing (Print)
##                                     7
##      Journal of Communication and Information Systems (Online)
##                                     6
## Journal of Visual Communication and Image Representation (Print)
##                                     6
##      Pattern Recognition Letters
##                                     5
```

Autores com mais participações em publicações:

```
head(sort(table(toupper(unlist(strsplit(unb.pub.ele_p8.df$autores,";")))),decreasing = TRUE),n=10)
```

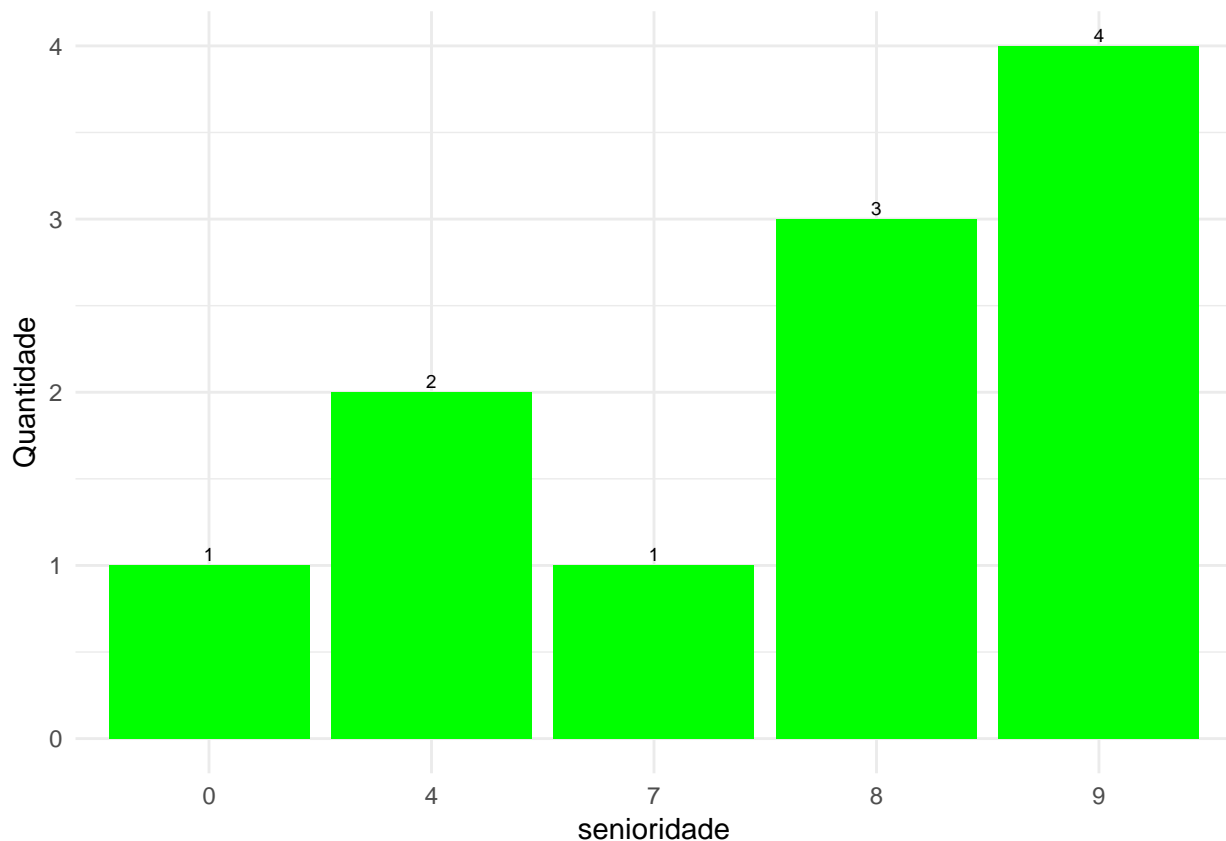
```
##
##      ROCHA, ANDERSON      ROCHA, ANDERSON
##          52              21
##      GOLDENSTEIN, SIOME  DA COSTA, J. P. C. L.
##          17              15
##      SOUSA JR, R. T. DE REZENDE ROCHA, ANDERSON
##          14              13
```

##	ROCHA, A	ROCHA, A.
##	13	13
##	ROCHA, ANDERSON DE REZENDE	WAINER, JACQUES
##	13	9

Vemos que há algumas incoerência neste dado acima, isto se dá pela falta de uniformidade dos dados.

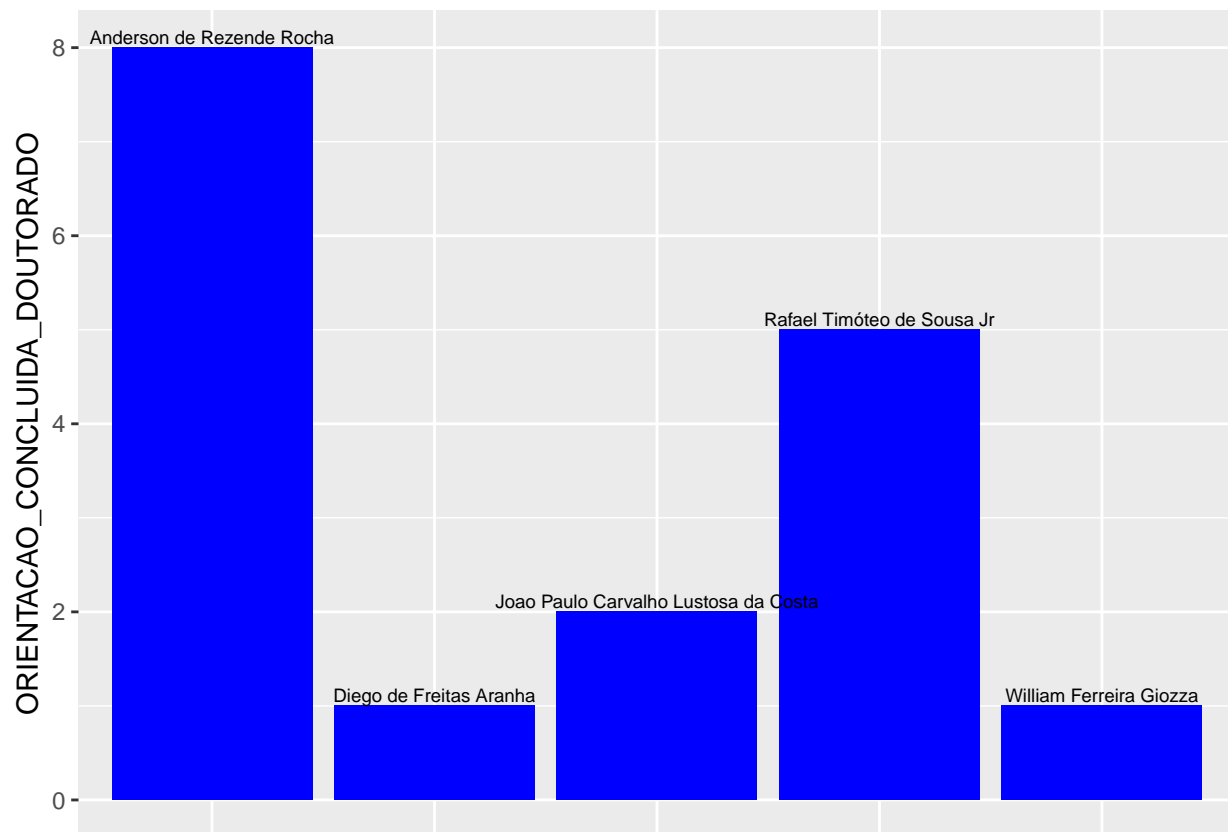
Perfis de docentes por senioridade:

```
unb.prof.ele_p8.df %>%
  group_by(senioridade) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = senioridade, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "green")+
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5)+
  theme_minimal()
```



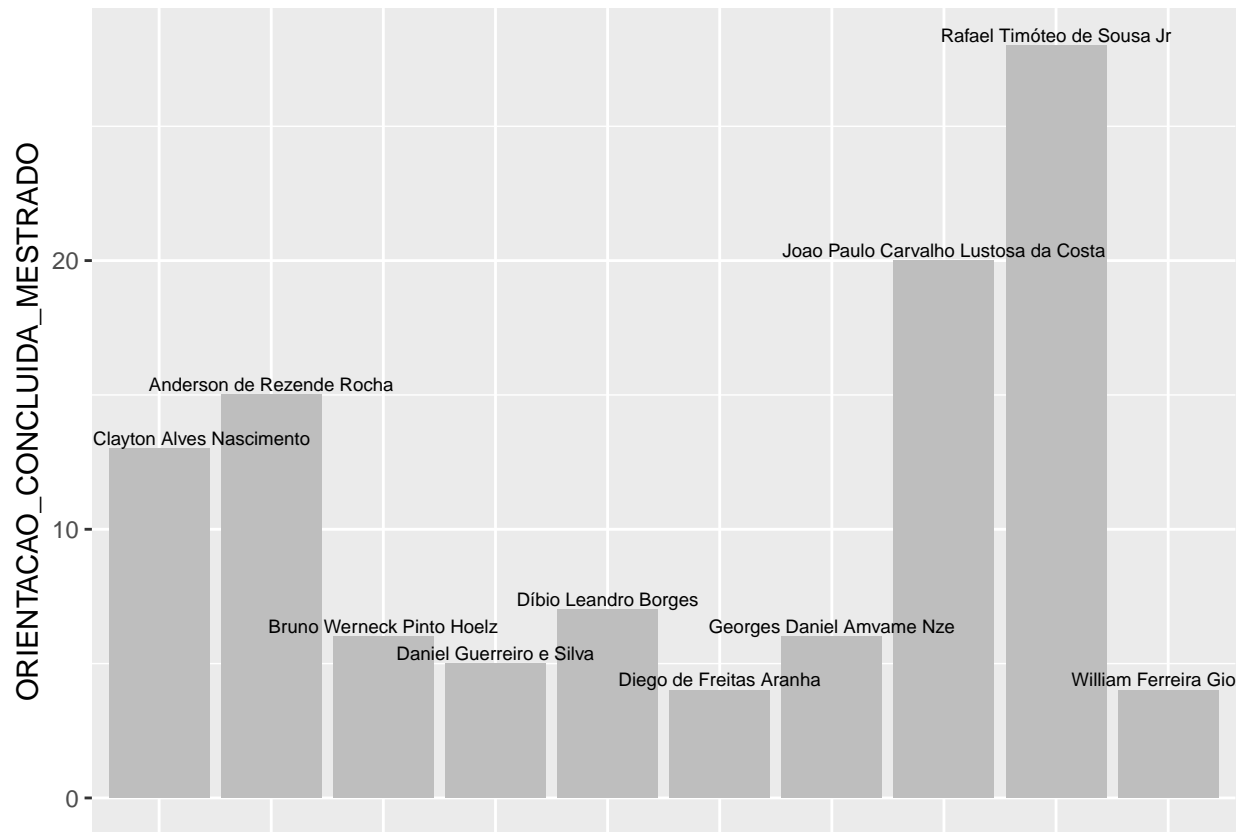
Quantidade de orientações concluídas no doutorado por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(ORIENTACAO_CONCLUIDA_DOUTORADO))%>%
  ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_DOUTORADO)) +
  geom_bar(position = "stack", stat = "identity", fill = "blue")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



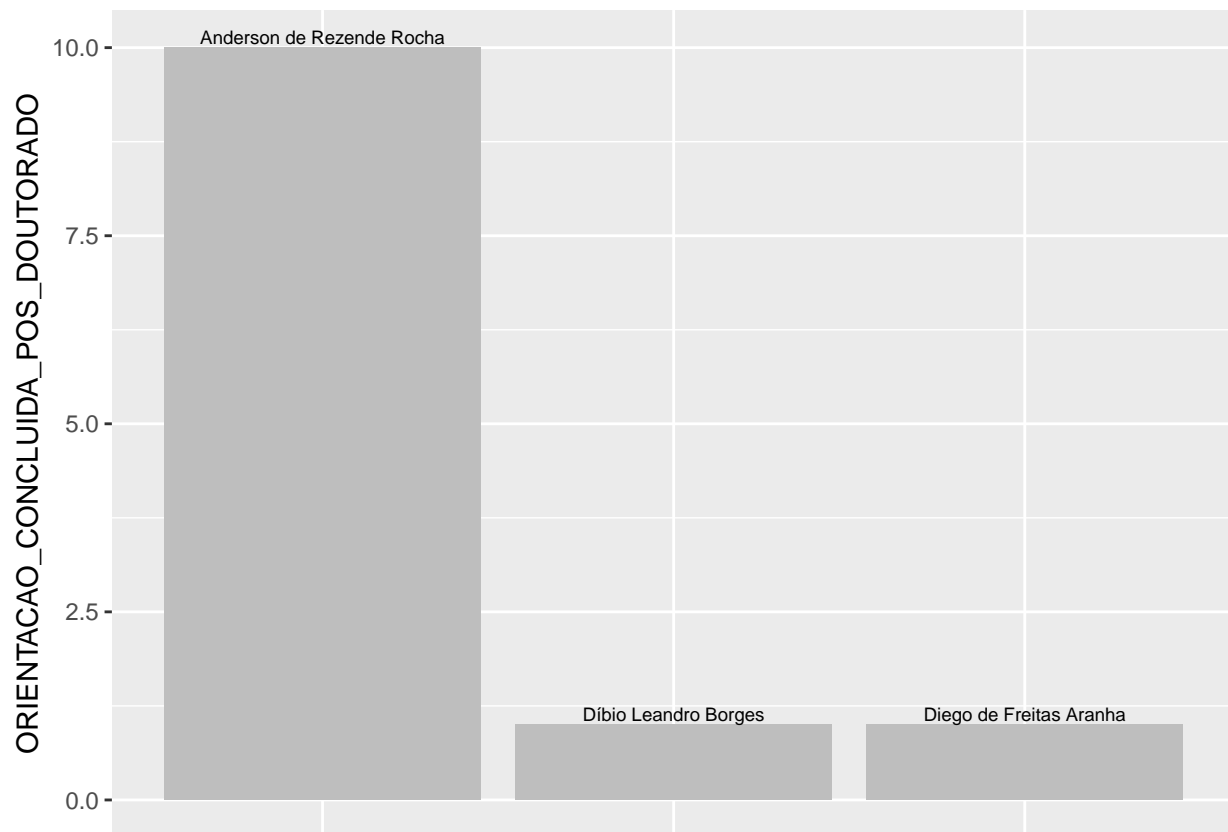
Quantidade de orientações concluídas no mestrado por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(ORIENTACAO_CONCLUIDA_MESTRADO)) %>%
  ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_MESTRADO)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray") +
  geom_text(aes(label=nome), vjust=-0.3, size=2.5) +
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



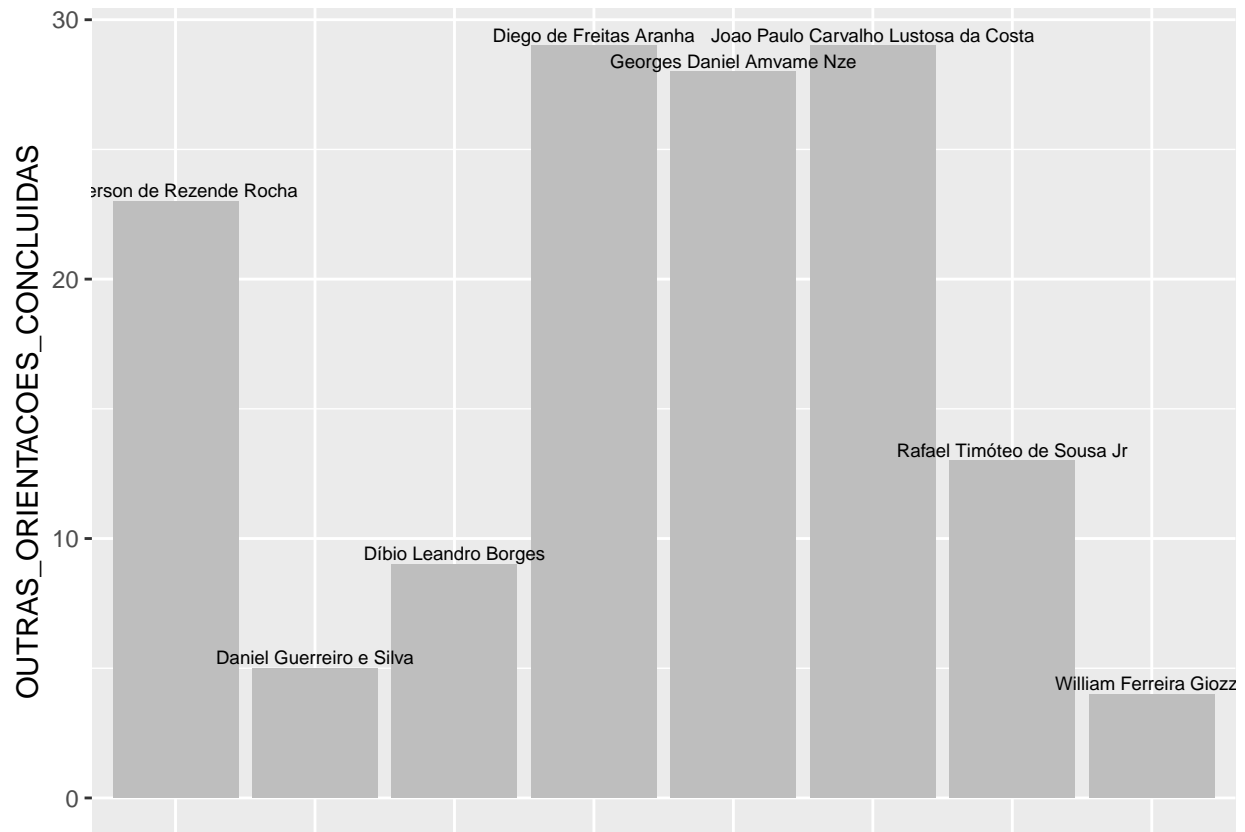
Quantidade de orientações concluídas no pós-doutorado por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(ORIENTACAO_CONCLUIDA_POS_DOUTORADO))%>%
  ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_POS_DOUTORADO)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



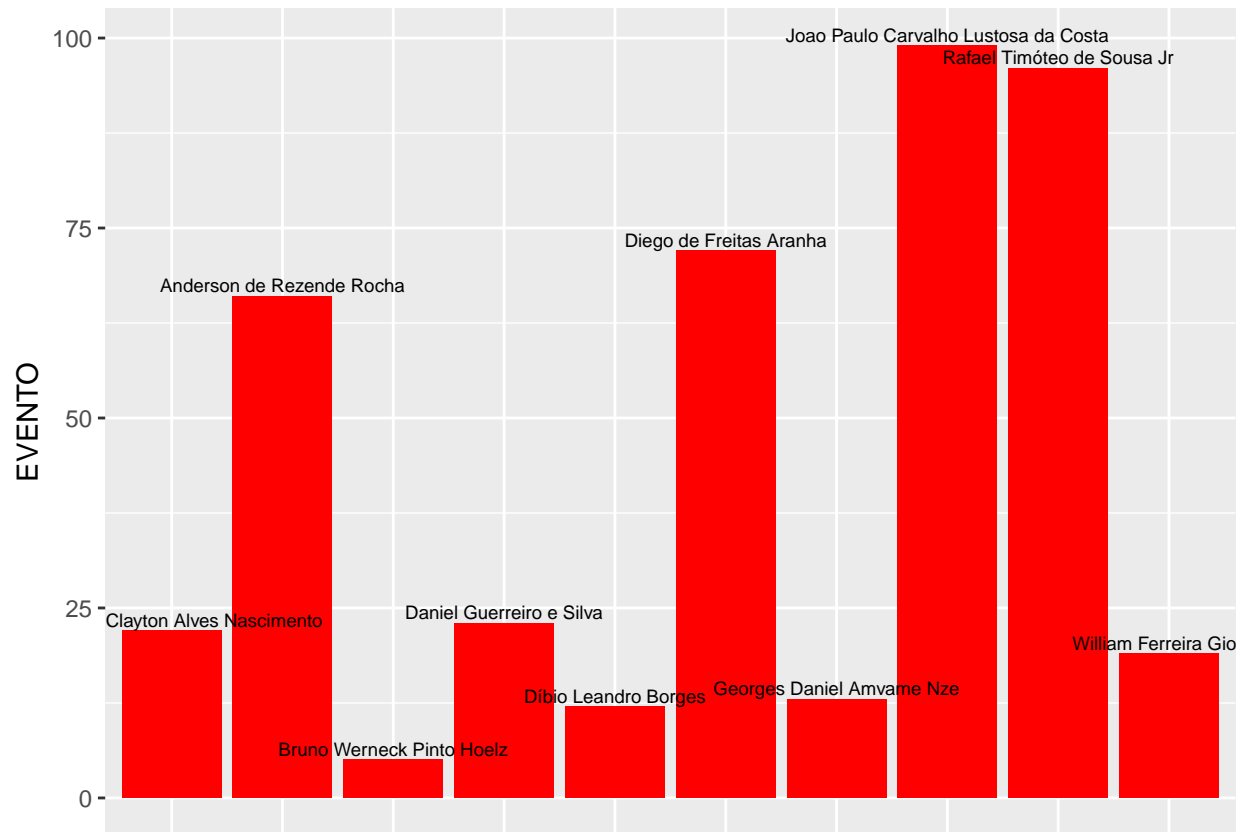
Quantidade de orientações concluídas em outros tipos por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(OUTRAS_ORIENTACOES_CONCLUIDAS))%>%
  ggplot(aes(x = nome, y = OUTRAS_ORIENTACOES_CONCLUIDAS)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



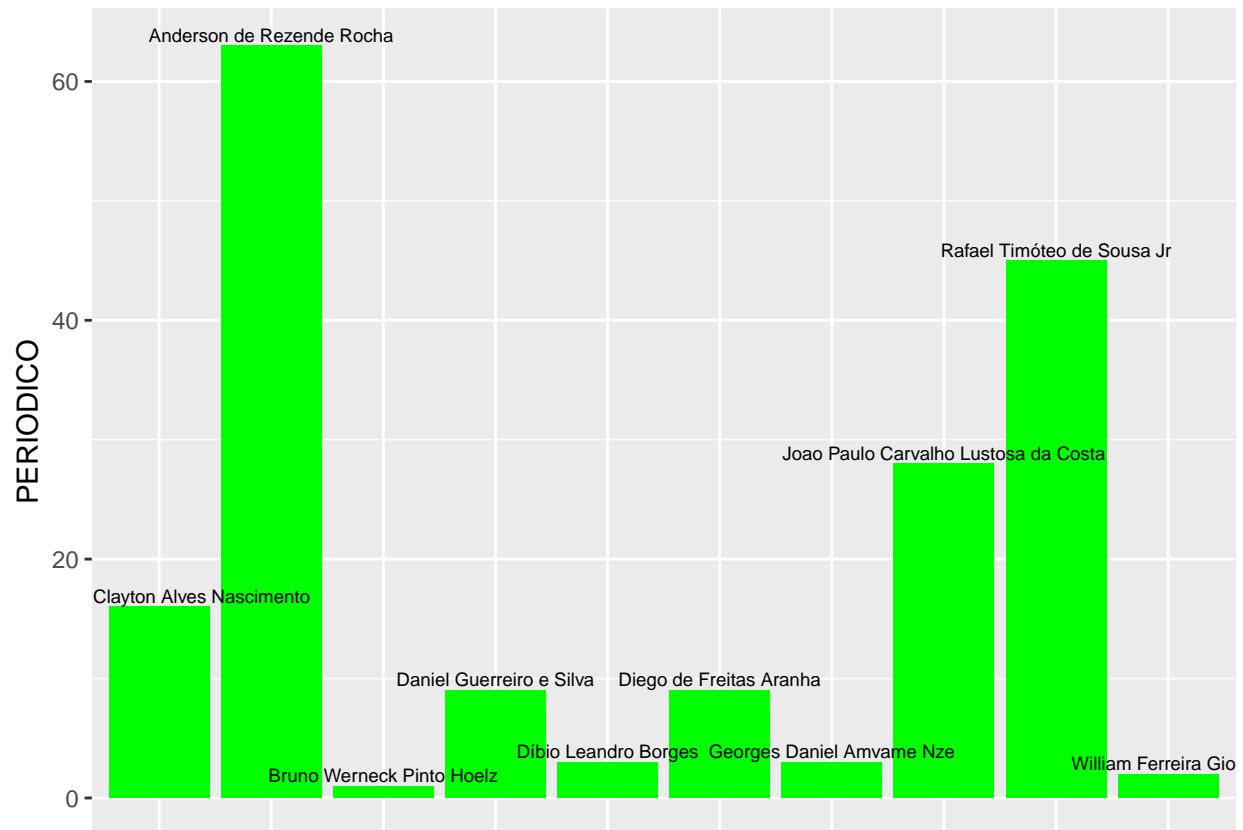
Quantidade de eventos por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(EVENTO))%>%
  ggplot(aes(x = nome, y = EVENTO)) +
  geom_bar(position = "stack", stat = "identity", fill = "red")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```

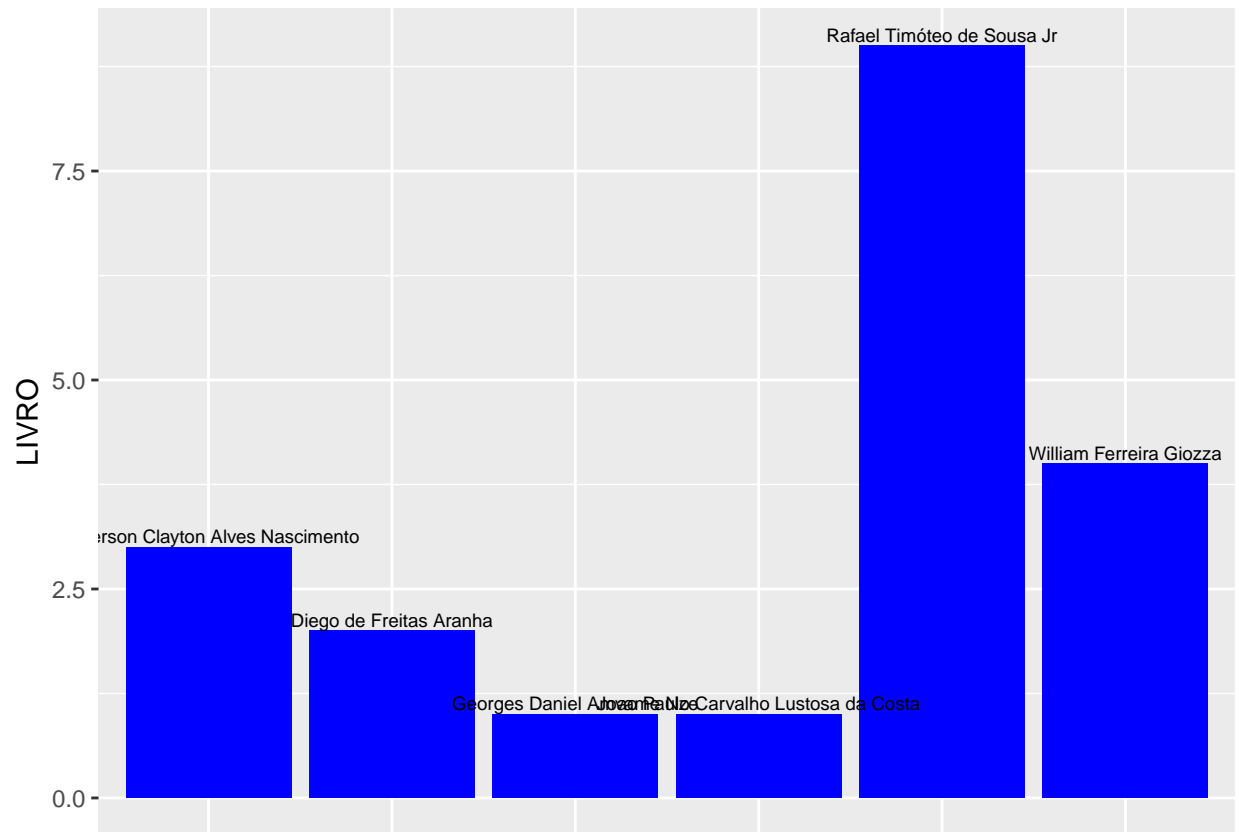
Quantidade de periódicos por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(PERIODICO))%>%
  ggplot(aes(x = nome, y = PERIODICO)) +
  geom_bar(position = "stack", stat = "identity", fill = "green")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



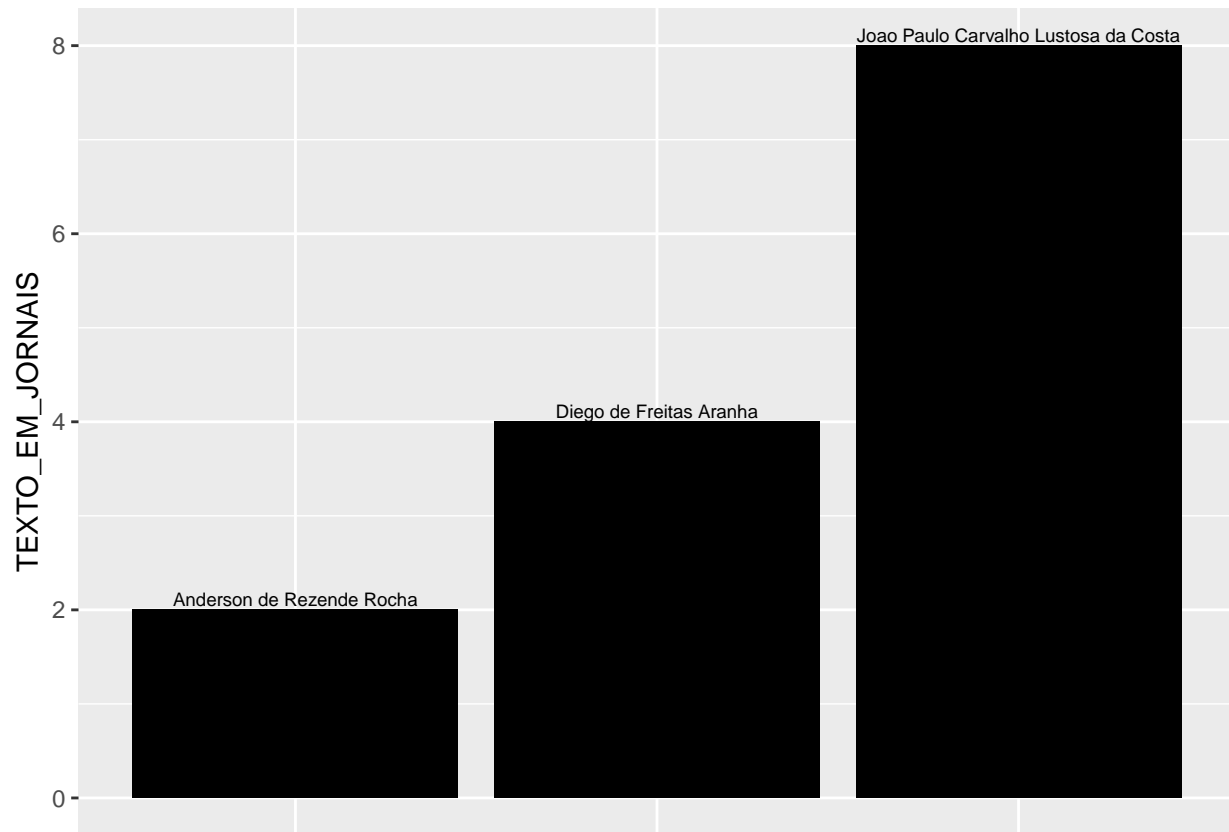
Quantidade de livros por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(LIVRO))%>%
  ggplot(aes(x = nome, y = LIVRO)) +
  geom_bar(position = "stack", stat = "identity", fill = "blue")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



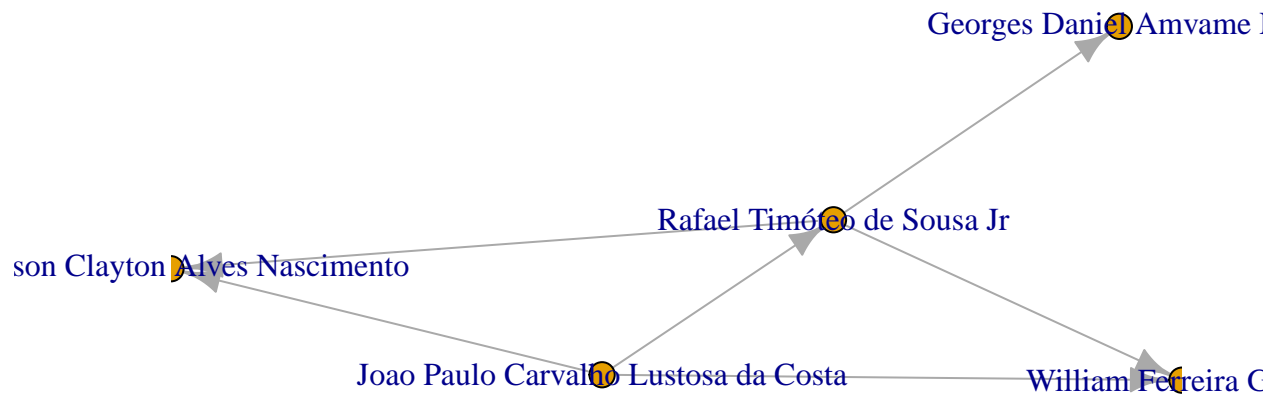
Quantidade de texto em jornais por professor:

```
unb.prof.ele_p8.df %>%
  group_by(nome) %>%
  filter(!is.na(TEXTO_EM_JORNAIS))%>%
  ggplot(aes(x = nome, y = TEXTO_EM_JORNAIS)) +
  geom_bar(position = "stack",stat = "identity", fill = "black")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



Relações entre professores:

```
list <- c(rbind(unb.graph.ele_p8$links$source, unb.graph.ele_p8$links$target))
idx <- match(list, unb.graph.ele_p8$nodes$id)
names <- unb.graph.ele_p8$nodes$properties$name
graph_names <- names[idx]
g <- graph(graph_names, directed=TRUE)
plot(g, layout = layout_(g, nicely()),
     edge.width = 1,
     edge.arrow.width = 1,
     vertex.size = 5,
     edge.arrow.size = 1,
     vertex.size2 = 3,
     vertex.label.cex = 1,
     asp = 0.35,
     margin = -0.1)
```



Nuvem de palavras baseadas no resumo do currículo dos professores:

```

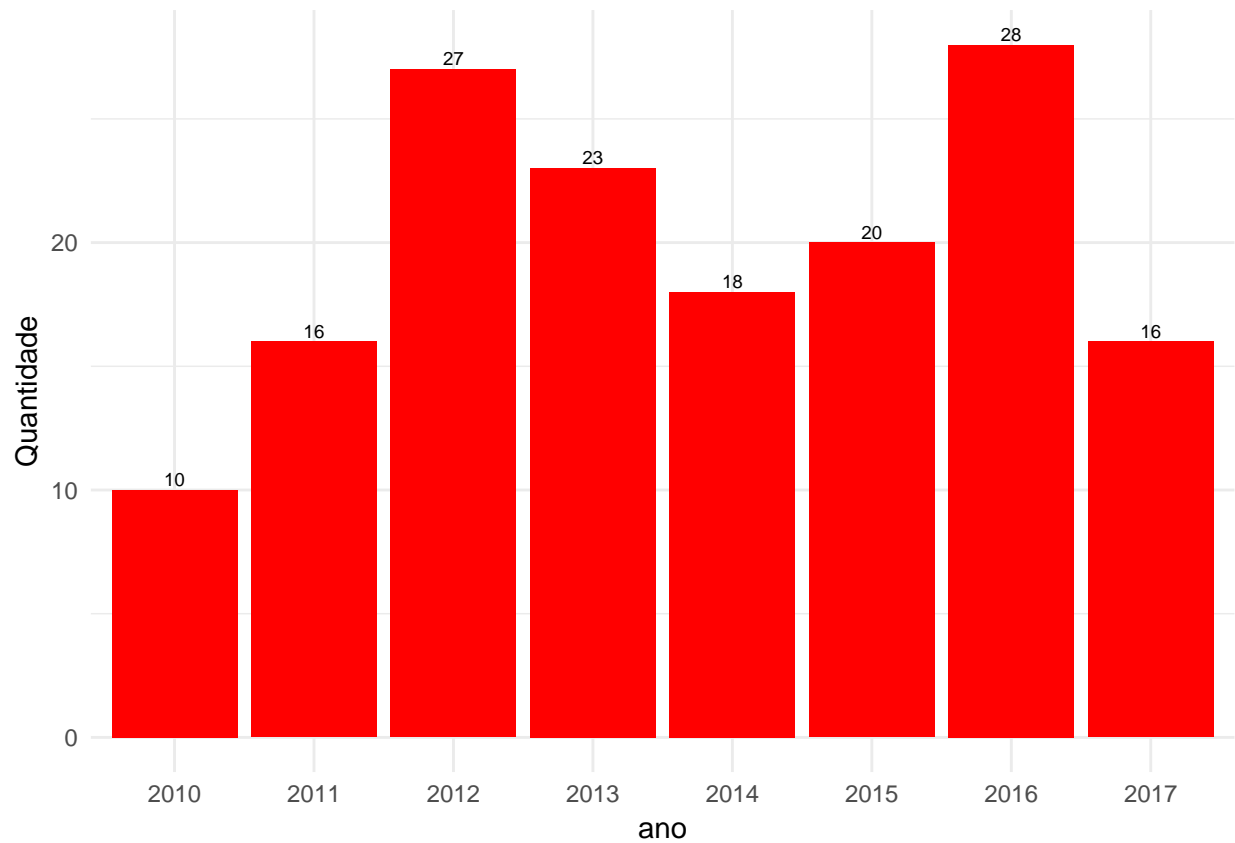
unb.prof.ele_p8.df$resumo_cv <- iconv(unb.prof.ele_p8.df$resumo_cv,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.prof.ele_p8.df$idLattes, text = unb.prof.ele_p8.df$resumo_cv, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words =(d$word), freq = d$freq, min.freq = 3,
          max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
  
```




Engenharia Elétrica P7

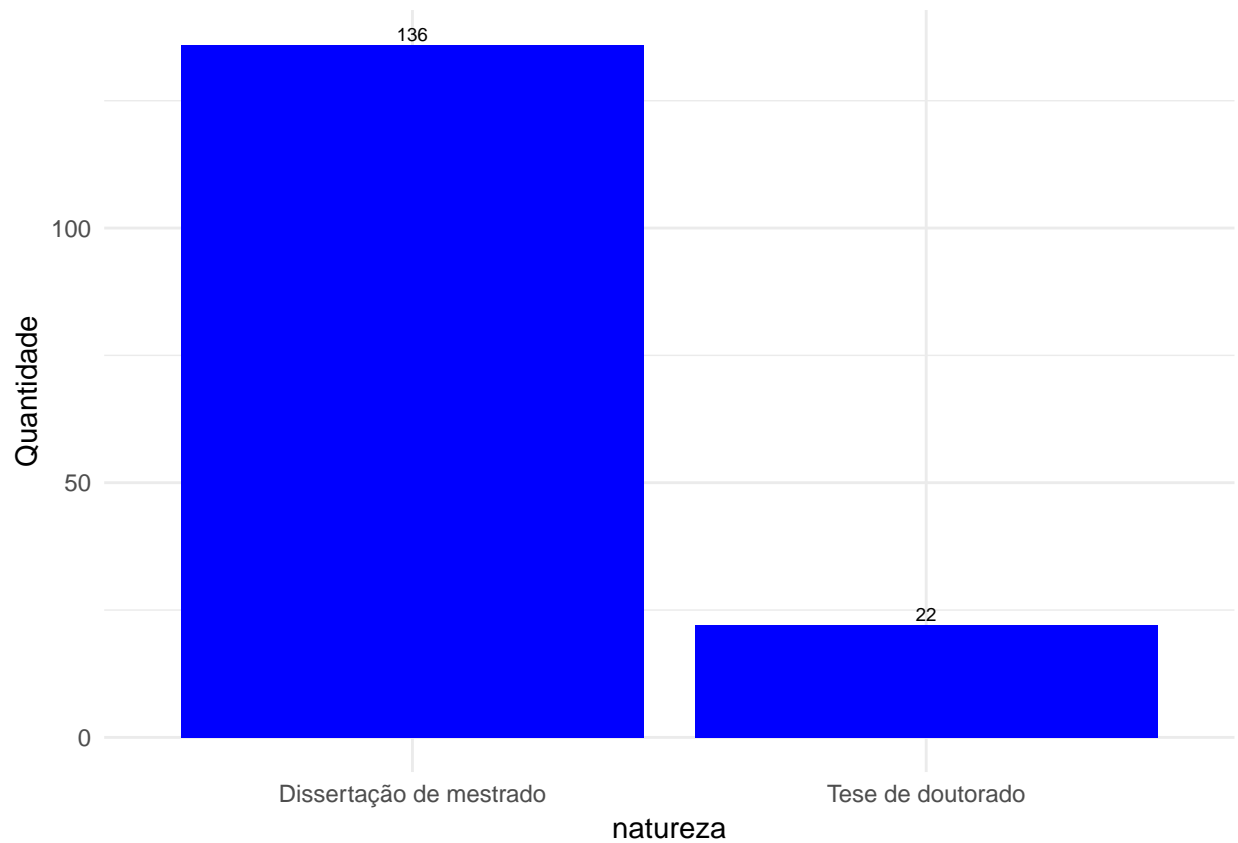
Quantidade de orientações feitas por ano:

```
unb.adv.ele_p7.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "red") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



Quantidade de orientações separadas por tipo de publicação:

```
unb.adv.ele_p7.df %>%  
group_by(natureza) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = natureza, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```

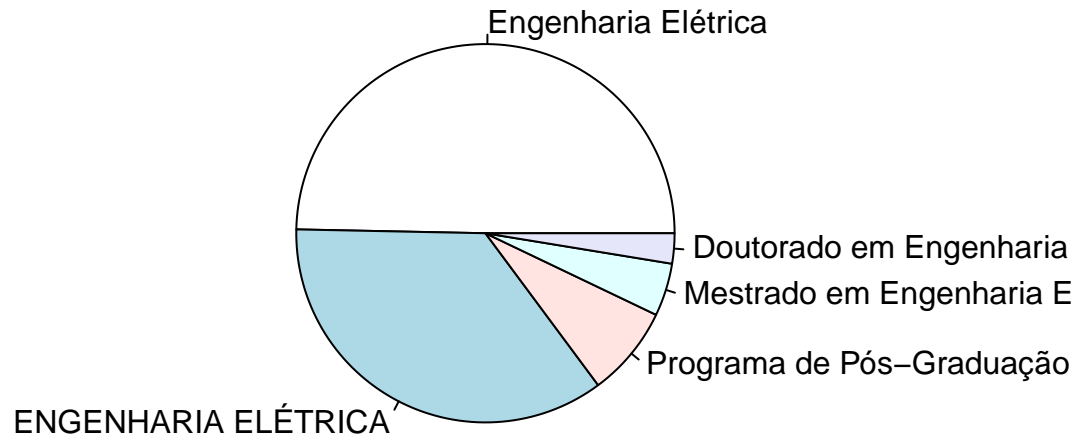
Cursos mais presentes:

`cursos_d5`

```
##
##          Engenharia Elétrica
##                77
##          ENGENHARIA ELÉTRICA
##                55
## Programa de Pós-Graduação em Engenharia Elétrica
##                12
##          Mestrado em Engenharia Elétrica
##                7
##          Doutorado em Engenharia Elétrica
##                4
```

```
pie(cursos_d5,clockwise=F,main="Cursos mais presentes")
```

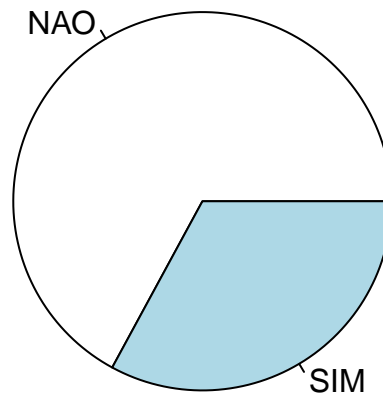
Cursos mais presentes



Quantidade de bolsistas e não bolsistas:

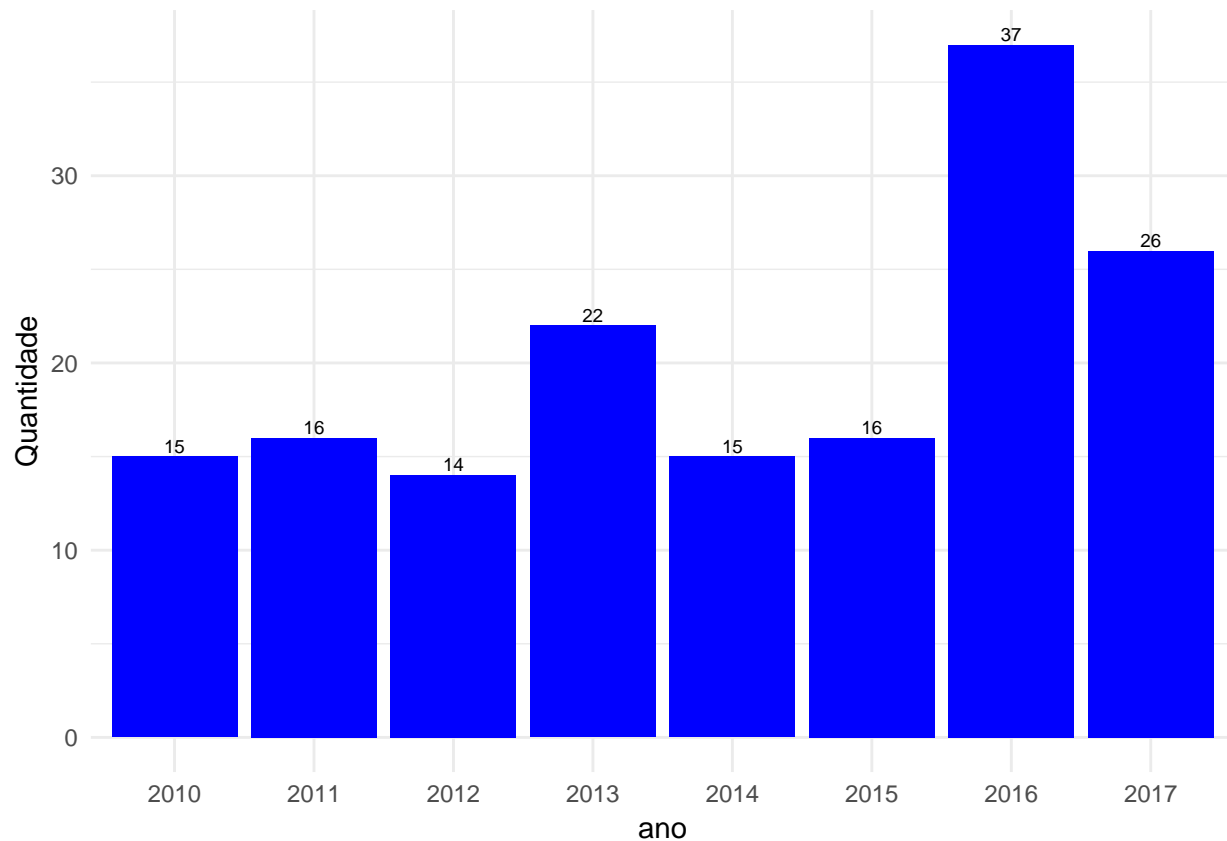
```
pie(table(unb.adv.ele_p7.df$bolsa),main="Bolsas")
```

Bolsas



Quantidade de publicações feitas por ano:

```
unb.pub.ele_p7.df %>%  
group_by(ano) %>%  
summarise(Quantidade = n()) %>%  
ggplot(aes(x = ano, y = Quantidade)) +  
geom_bar(position = "stack", stat = "identity", fill = "blue") +  
geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +  
theme_minimal()
```



Periódicos com mais publicações:

```
d <- table(unb.pub.ele_p7.df$periodico)
head(sort(d,decreasing = TRUE),n=5)
```

```
##
## Journal of Communication and Information Systems (Online)
##                               10
##               Revista IEEE América Latina
##                               8
##               Digital Signal Processing (Print)
##                               7
##               IEEE Transactions on Power Delivery
##                               6
##               Electric Power Systems Research (Print)
##                               5
```

Autores com mais participações em publicações:

```
head(sort(table(toupper(unlist(strsplit(unb.pub.ele_p7.df$autores,";")))),decreasing = TRUE),n=10)
```

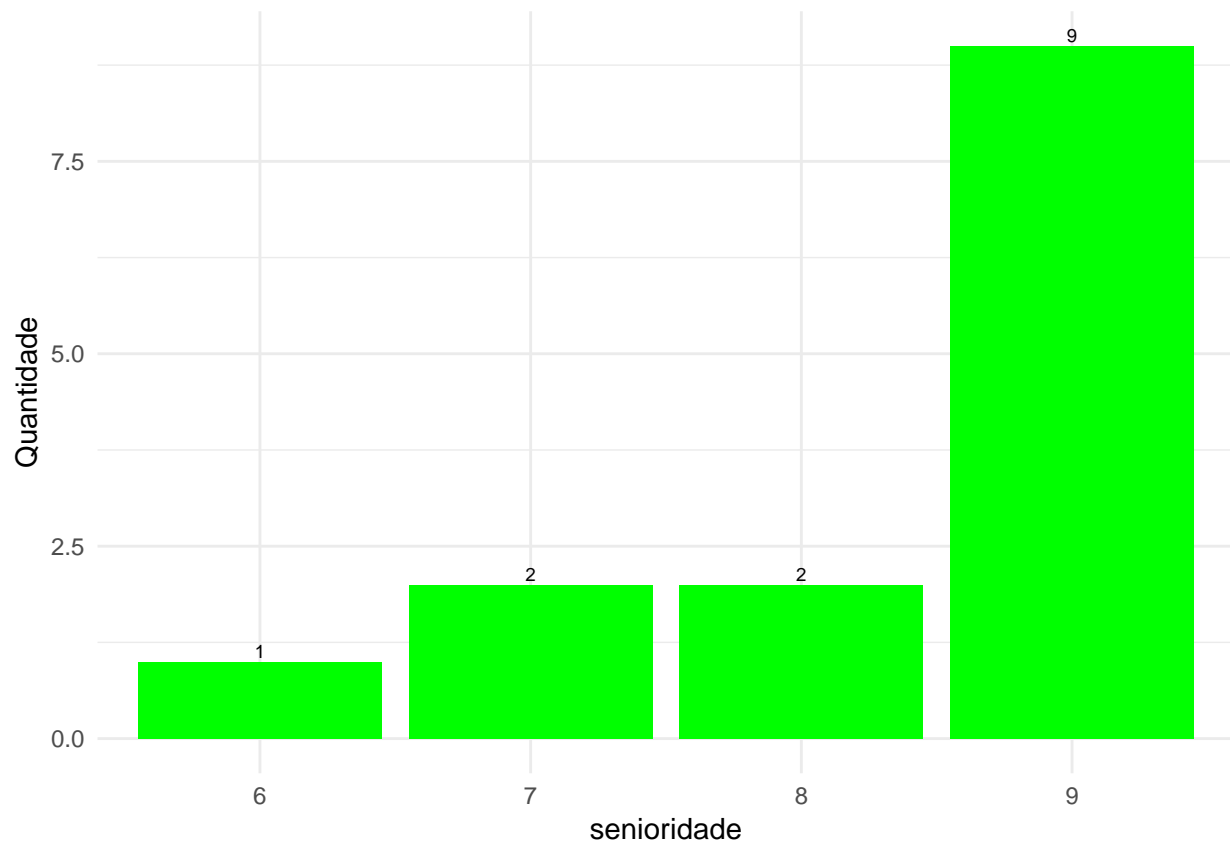
```
##
##          DA COSTA, J. P. C. L.          SOUSA JR, R. T.
##                               14          14
##          MOLINARO, L. F. R.          ATTUX, ROMIS
##                               13          6
## DE SOUSA JÚNIOR, RAFAEL TIMÓTEO    DA COSTA, JOÃO PAULO C.L.
##                               6          5
```

##	DIAS, U. S.	FRAZAO, A.
##	5	5
##	GALPERIN, H.	SILVA, K. M.
##	5	5

Vemos que há algumas incoerência neste dado acima, isto se dá pela falta de uniformidade dos dados.

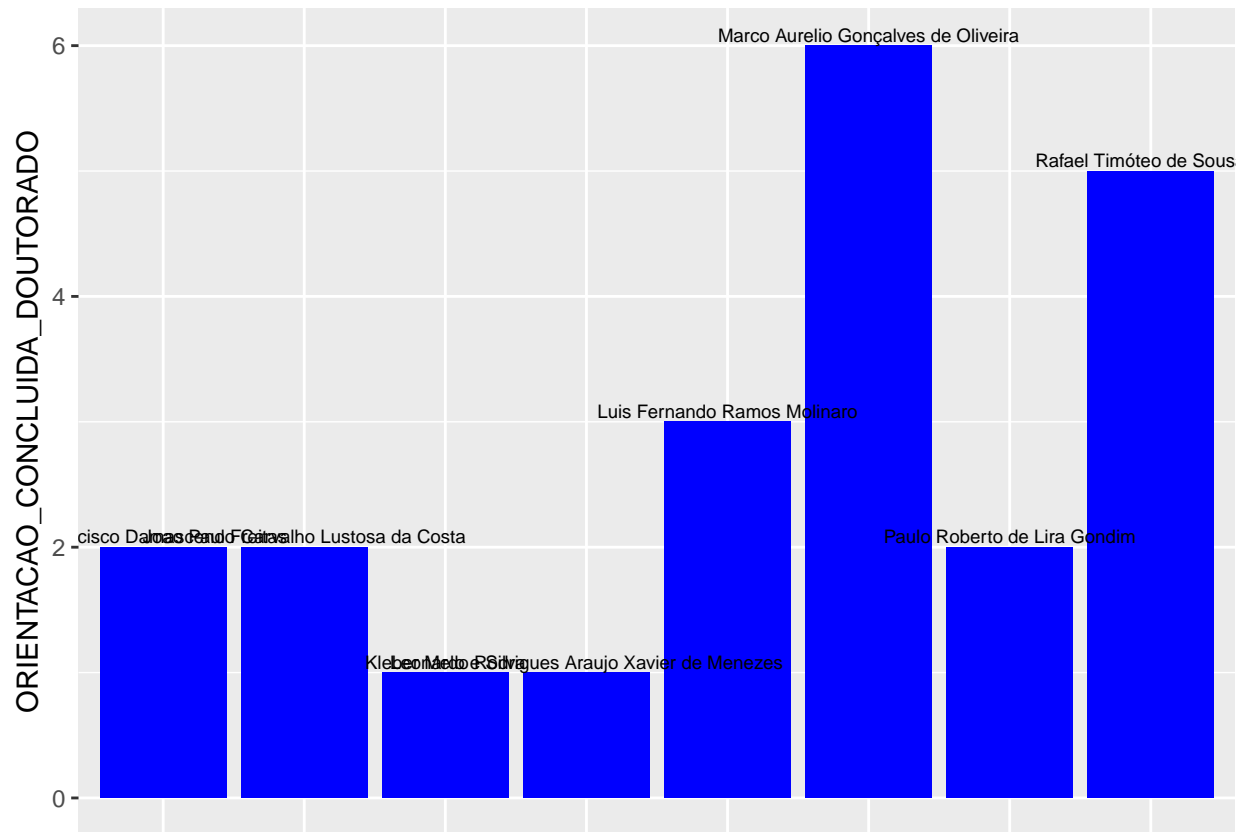
Perfis de docentes por senioridade:

```
unb.prof.ele_p7.df %>%
  group_by(senioridade) %>%
  summarise(Quantidade = n()) %>%
  ggplot(aes(x = senioridade, y = Quantidade)) +
  geom_bar(position = "stack", stat = "identity", fill = "green") +
  geom_text(aes(label=Quantidade), vjust=-0.3, size=2.5) +
  theme_minimal()
```



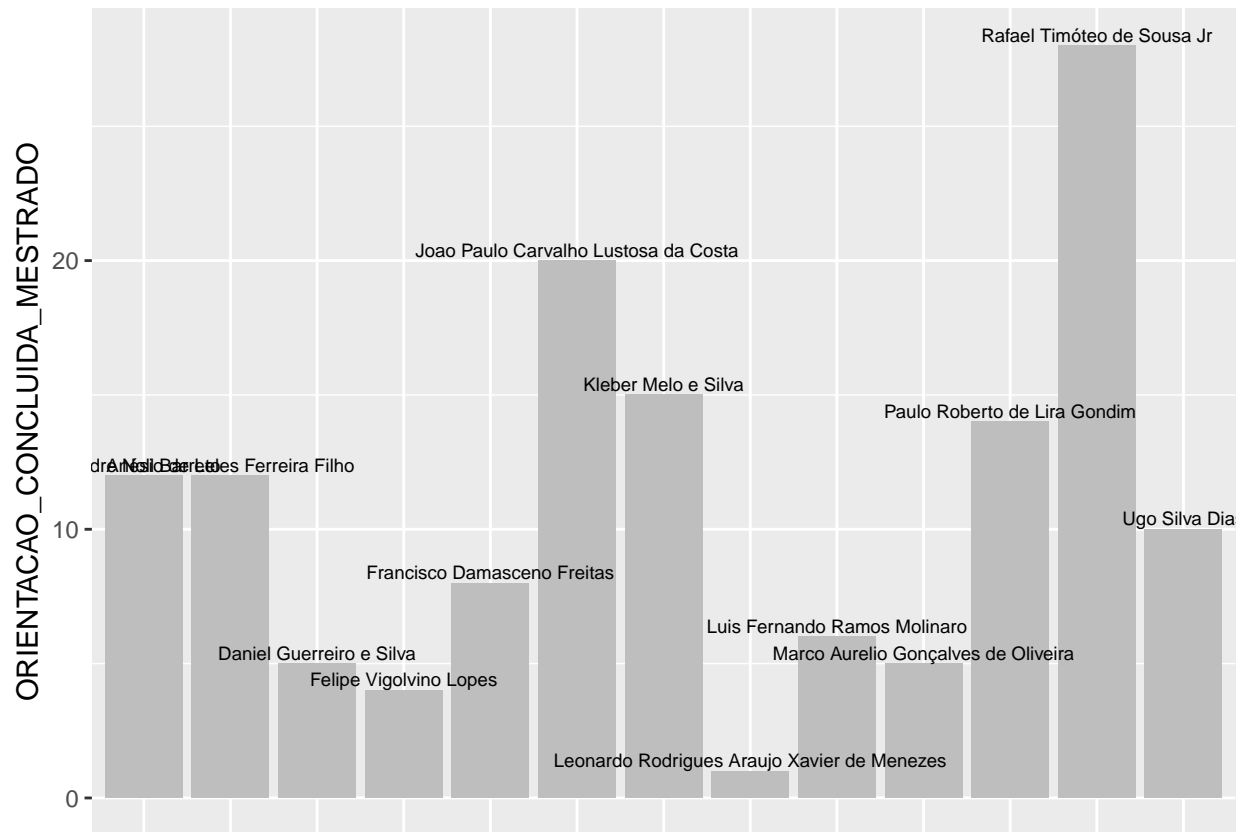
Quantidade de orientações concluídas no doutorado por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(ORIENTACAO_CONCLUIDA_DOUTORADO)) %>%
  ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_DOUTORADO)) +
  geom_bar(position = "stack", stat = "identity", fill = "blue") +
  geom_text(aes(label=nome), vjust=-0.3, size=2.5) +
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



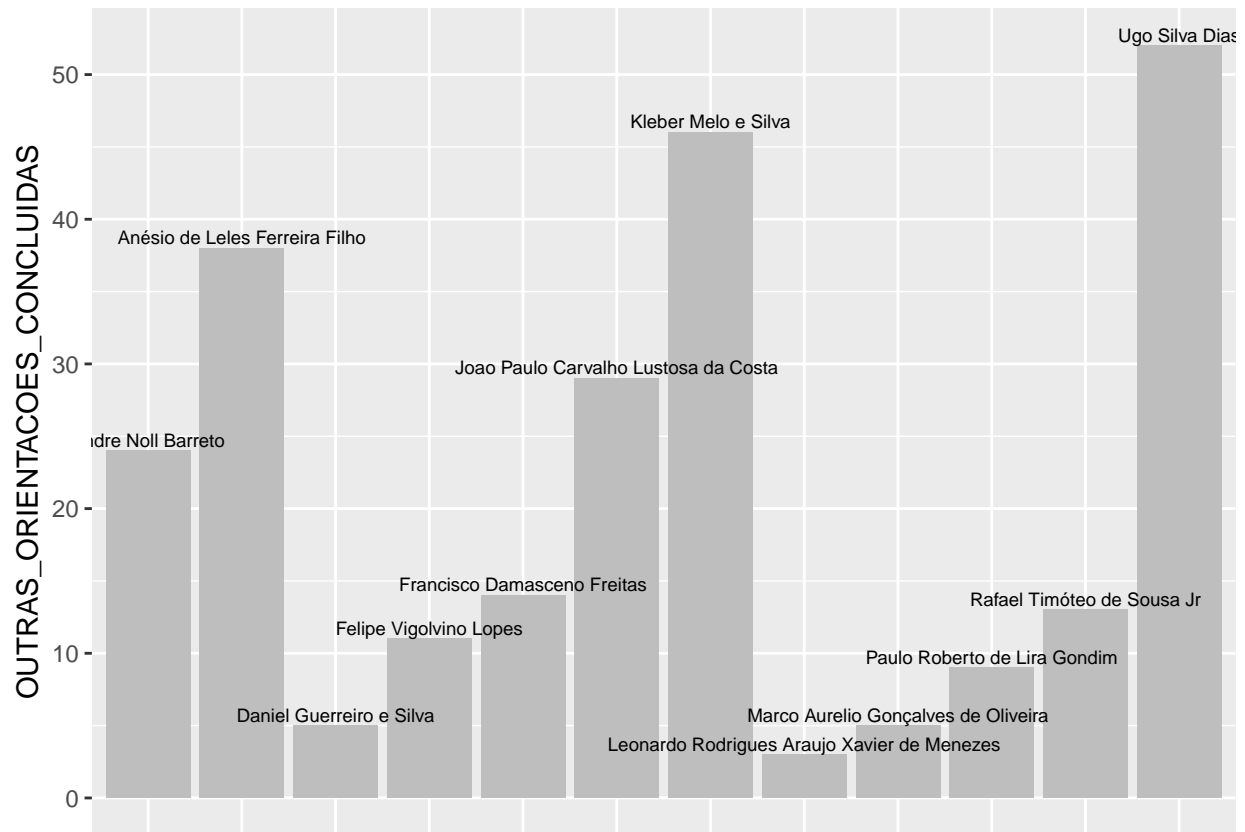
Quantidade de orientações concluídas no mestrado por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(ORIENTACAO_CONCLUIDA_MESTRADO)) %>%
  ggplot(aes(x = nome, y = ORIENTACAO_CONCLUIDA_MESTRADO)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray") +
  geom_text(aes(label=nome), vjust=-0.3, size=2.5) +
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



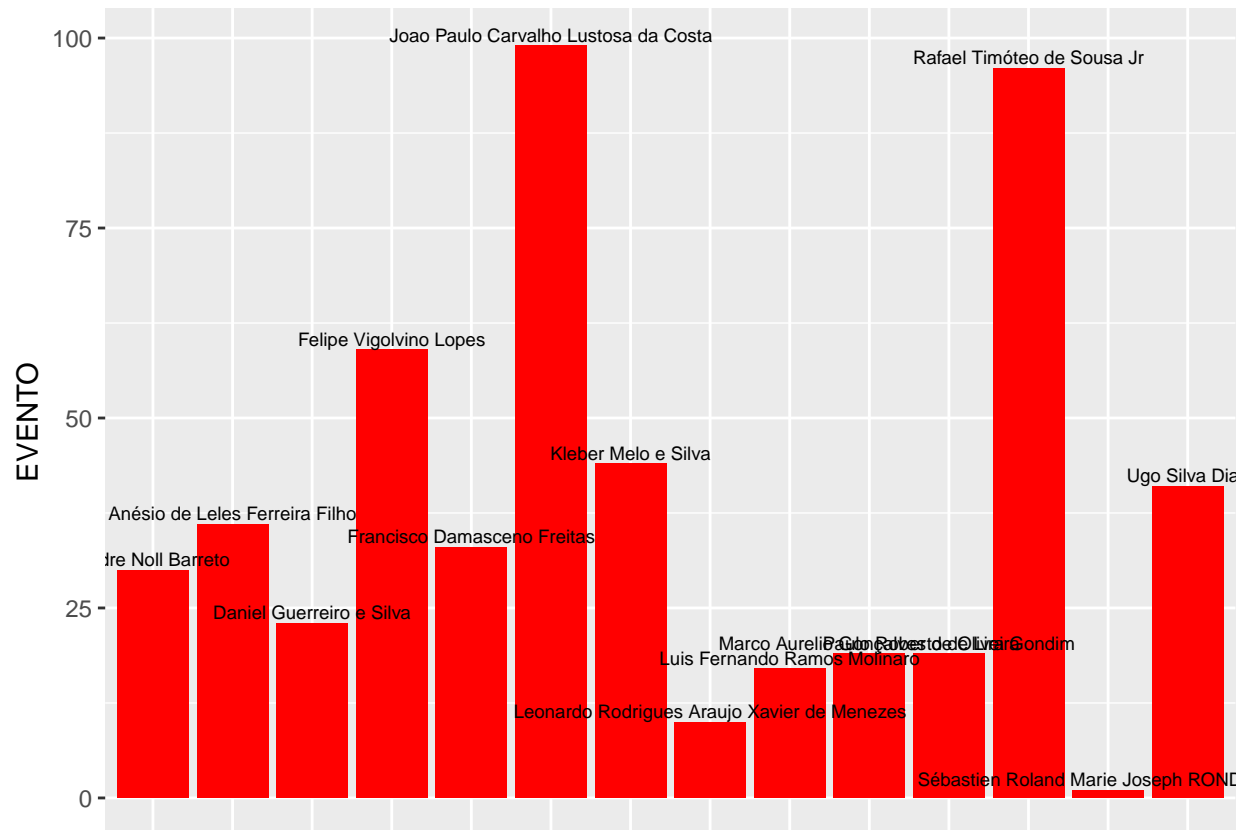
Quantidade de orientações concluídas em outros tipos por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(OUTRAS_ORIENTACOES_CONCLUIDAS))%>%
  ggplot(aes(x = nome, y = OUTRAS_ORIENTACOES_CONCLUIDAS)) +
  geom_bar(position = "stack", stat = "identity", fill = "gray")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



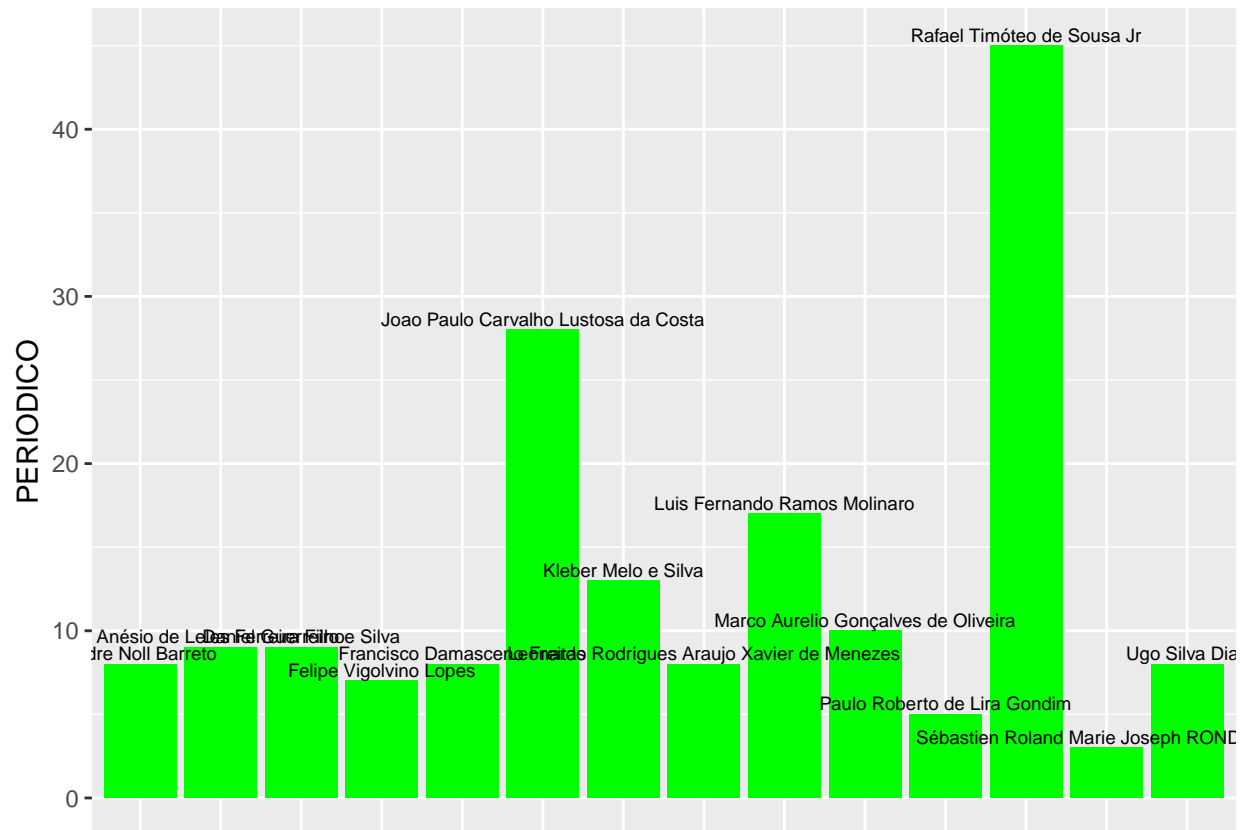
Quantidade de eventos por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(EVENTO))%>%
  ggplot(aes(x = nome, y = EVENTO)) +
  geom_bar(position = "stack", stat = "identity", fill = "red")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```

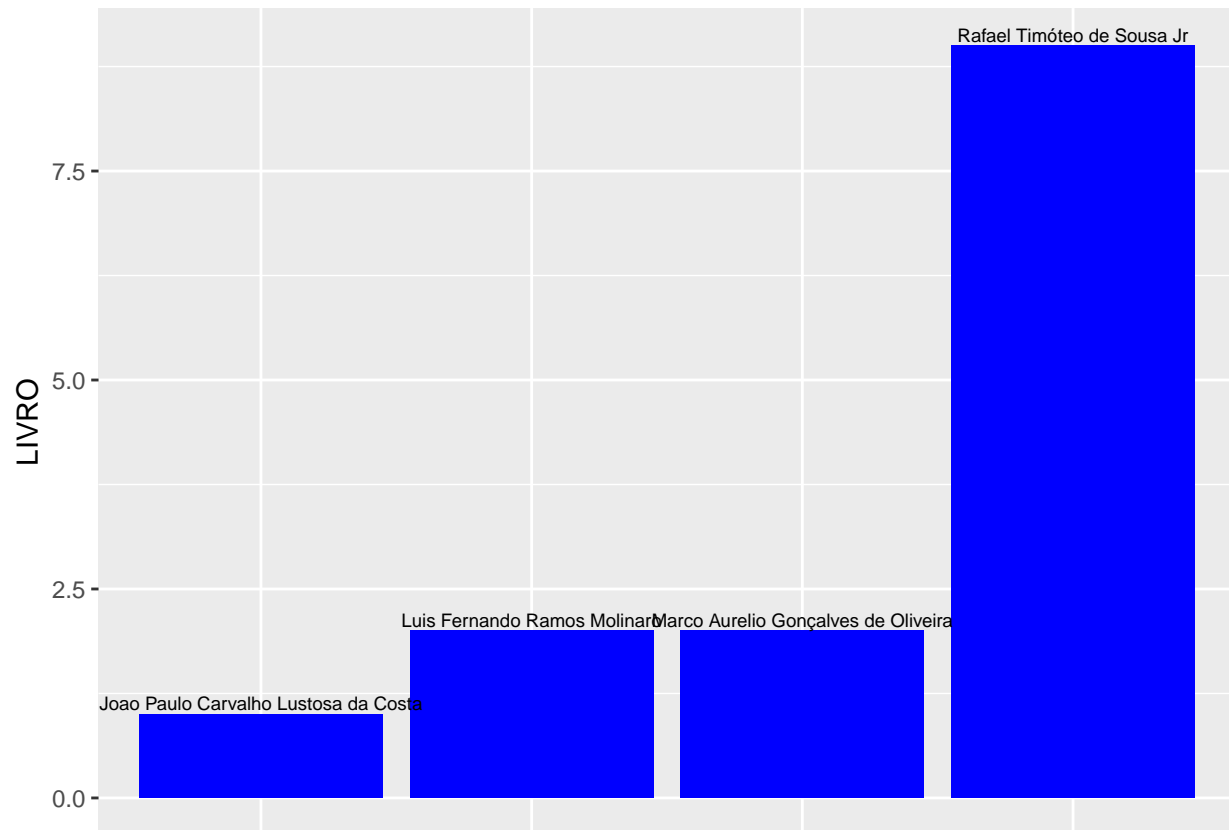
Quantidade de periódicos por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(PERIODICO))%>%
  ggplot(aes(x = nome, y = PERIODICO)) +
  geom_bar(position = "stack", stat = "identity", fill = "green")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



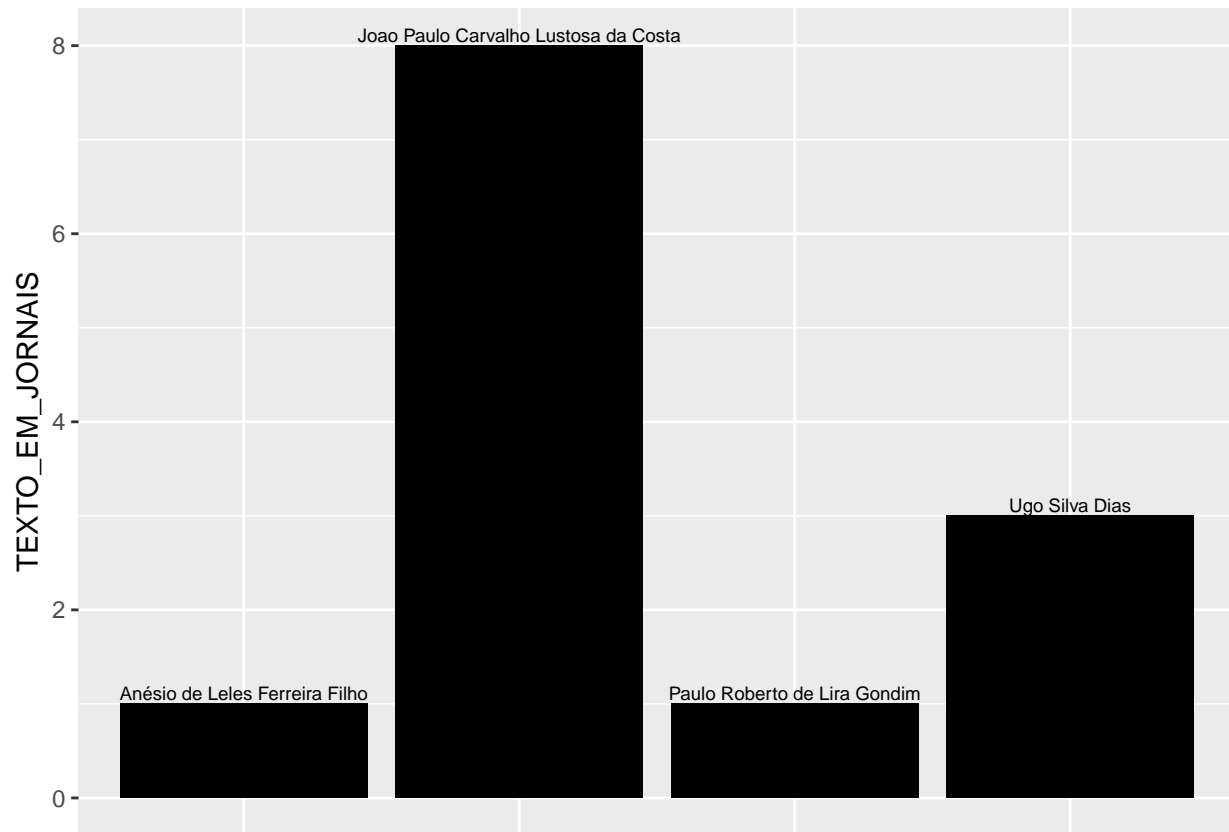
Quantidade de livros por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(LIVRO))%>%
  ggplot(aes(x = nome, y = LIVRO)) +
  geom_bar(position = "stack", stat = "identity", fill = "blue")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



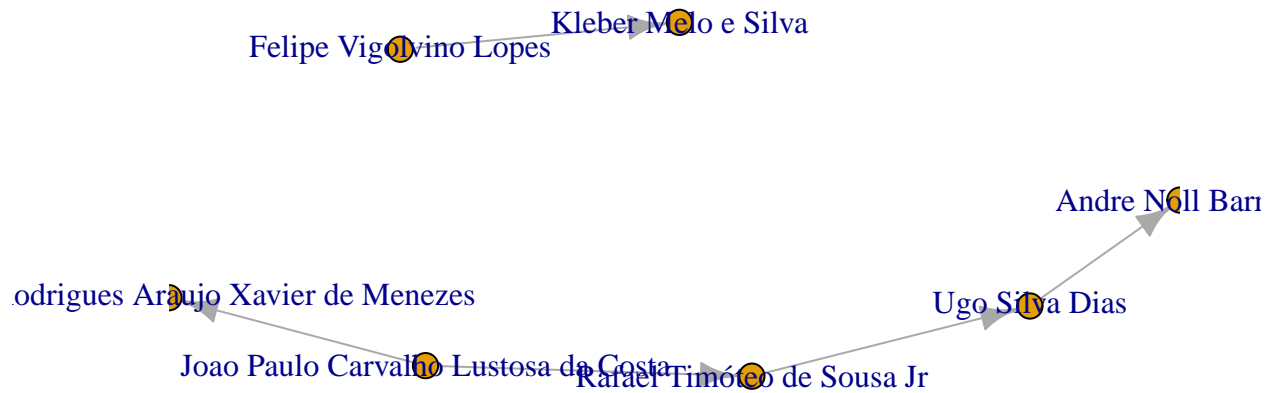
Quantidade de texto em jornais por professor:

```
unb.prof.ele_p7.df %>%
  group_by(nome) %>%
  filter(!is.na(TEXTO_EM_JORNAIS))%>%
  ggplot(aes(x = nome, y = TEXTO_EM_JORNAIS)) +
  geom_bar(position = "stack", stat = "identity", fill = "black")+
  geom_text(aes(label=nome), vjust=-0.3, size=2.5)+
  theme(axis.title.x=element_blank(),
        axis.text.x=element_blank(),
        axis.ticks.x=element_blank())
```



Relações entre professores:

```
list <- c(rbind(unb.graph.ele_p7$links$source, unb.graph.ele_p7$links$target))
idx <- match(list, unb.graph.ele_p7$nodes$id)
names <- unb.graph.ele_p7$nodes$properties$name
graph_names <- names[idx]
g <- graph(graph_names, directed=TRUE)
plot(g, layout = layout_(g, nicely()),
     edge.width = 1,
     edge.arrow.width = 1,
     vertex.size = 5,
     edge.arrow.size = 1,
     vertex.size2 = 3,
     vertex.label.cex = 1,
     asp = 0.35,
     margin = -0.1)
```



Nuvem de palavras baseadas no resumo do currículo dos professores:

```

unb.prof.ele_p7.df$resumo_cv <- iconv(unb.prof.ele_p7.df$resumo_cv,from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.prof.ele_p7.df$idLattes, text = unb.prof.ele_p7.df$resumo_cv, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "e"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m),decreasing=TRUE)
d <- data.frame(word = names(v),freq=v)
set.seed(1234)
wordcloud(words =(d$word), freq = d$freq, min.freq = 3,
          max.words=30,scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
  
```



Nuvem de palavras baseadas no título das orientações:

```
unb.adv.ele_p7.df$titulo <- iconv(unb.adv.ele_p7.df$titulo, from="UTF-8", to = "ASCII//TRANSLIT")
df_wc <- data.frame(doc_id = unb.adv.ele_p7.df$idLattes1, text = unb.adv.ele_p7.df$titulo, stringsAsFactors=FALSE)
docs <- Corpus(DataframeSource(df_wc))
docs <- tm_map(docs, content_transformer(tolower))
docs <- tm_map(docs, removeNumbers)
docs <- tm_map(docs, removeWords, stopwords("pt"))
docs <- tm_map(docs, removeWords, stopwords("english"))
docs <- tm_map(docs, removeWords, c("universidade", "federal", "eletrica", "posgraduacao", "engenheiro", "engenharia"))
docs <- tm_map(docs, removePunctuation)
docs <- tm_map(docs, stripWhitespace)
dtm <- TermDocumentMatrix(docs)
m <- as.matrix(dtm)
v <- sort(rowSums(m), decreasing=TRUE)
d <- data.frame(word = names(v), freq=v)
set.seed(1234)
wordcloud(words =(d$word), freq = d$freq, min.freq = 3,
          max.words=30, scale = c(4, 0.2), random.order=FALSE, rot.per=0.35,
          colors=brewer.pal(8, "Dark2"))
```



IX - Implatação

No caso deste trabalho, a implantação foi realizada por meio da implementação dos scripts apresentados no decorrer do documento.

X - Conclusão

Ao final deste trabalho, podemos chegar a conclusão que com a metodologia CRISP-DM foi possível realizar uma análise de dados a partir de vários arquivos **JSON**, assim, foi possível perceber como funciona - pelo menos de maneira simplória - a vida de cientista de dados, como se deve preparar os dados, como se deve buscar informações úteis, analisar determinadas informações, plotar gráficos para um efeito mais visual, ver a dificuldade que falta de padronização causa, ainda mais quando se tem os mais diversos arquivos para analisar das mais diversas fontes.

O trabalho usou como base o arquivo disponibilizado pelo professor, no qual continha as explicações de todas as fases do CRISP-DM, seguindo o modelo passado pelo professor, tentou-se ir percorrendo sobre o script em R que se vinha fazendo e correlacionando com alguma fase do CRISP-DM, mas tal divisão não pode ser considerada totalmente fidedigna, haja visto que muitas vezes há mistura de fases e não apenas fases totalmente isoladas, sem influência de outras, procurou-se explicar de maneira simples como cada etapa se encaixava cada fase.

Por fim, podemos chegar a alguns resultados valorosos, mas nem tudo são flores, algumas dificuldades vieram em decorrência da falta de uniformidade, porém podemos aprender bastante com este trabalho e nos interar de como começar a fazer análise de dados.

Referências