UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM POLÍTICA SOCIAL

Curso na quarentena:

Introdução ao R para Política Social

CARGA HORÁRIA: 60H

R é um software cada vez mais usado no fluxo de trabalho de pesquisadores que procuram aproveitar ao máximo a disponibilidade de dados.

Ao mesmo tempo, dada a complexidade da tarefa associada ao uso de grande quantidade de dados e a complexidade do software, pesquisadores se encontram com urgência e dificuldade no aprendizado do novo ferramental.

Para o caso da Política Social, além disso, parece oportuno procurar mostrar grandes fontes de informação disponíveis publicamente, e, até o momento, relativamente pouco aproveitadas em estudos críticos em profundidade.

Diante disso, propõe-se trabalhar com método "aplicado" e "em espiral".

O motivo do desenvolvimento e escolha desse método é acostumar os aprendizes às etapas de fluxo de trabalho e mantê-los motivados ao fazer avançar em várias dimensões de uma maneira holística.

Isso se traduz a que a cada semana, após introduções geral e específica, vamos executar todo o fluxo de trabalho com relação a projetos concretos dos mais simples. Qual é o fluxo de trabalho ? *1)Obter dados, 2) importar dados, 3)limpar dados, 4) manipular dados, criar indicadores, obter informação necessária, 5)produzir resultados seja em tabela seja em gráfico, seja em mapa, e exportar*. Com o objetivo de manter a motivação de produzir resultados o mais rápido alcançável, os aprendizes são chamados a trazer problemas que precisem resolver em suas pesquisas em andamento, havendo flexibilidade na definição de projetos.

Para o presente curso, pretende-se estrutura na seguinte forma:

- O fluxo de trabalho e as ferramentas (introdução, instalação, e exploração geral das interfaces) (Semana 1)
- R e RStudio - visão geral (Semana 2)
- Explorando Trabalho Intermitente e Parcial - Projeto 1(Semana 3)

- População e municípios - Projeto 2 (Semana 4)
- Assistência (SUAS) - Projeto 3 (Semana 5)
- Orçamento, Transparência e Assistência Estudantil - Projeto 4 (Semana 6)
- Informações de Educação – Censos – Projeto 5 (Semana 7)
- Saúde (SUS) – Projeto 6 (Semana 8)
- CadÚnico – Projeto 7 (Semana 9)
- Rendimento e Ocupação via PNAD Contínua – Projeto 8 (Semana 10)
- Cesta de Consumo de trabalhadores (POF) – Projeto 9 (Semana 11)
- Previdência Social – Projeto Final (Semana 10)

- Método e forma de trabalho:
  - 2 Sessões de 2H/aula - reuniões virtuais (Plata
  - Exercícios / acesso a servidor com RStudio em caso de dificuldades de instalação, tutoria estimulada e disponível com prioridade

Inscrições em: https://politicasocialcomr.ddns.net

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM POLÍTICA SOCIAL

Curso/atividade na quarentena:

Introdução ao R para Política Social

CARGA HORÁRIA: 60H

Ementa detalhada:

1. O fluxo de trabalho e as ferramentas (Sem 1)
   - Tópicos cobertos nas aulas da semana:
     - Qual é o fluxo de trabalho do trabalho com dados dentro do trabalho científico?
     - O trabalho sério e crítico com dados para avanço científico
     - Manual x máquina -> qual parte importa/o que deixar para a máquina
     - Noções de Algoritmos
     - O *script* como uma receita

2. R e RStudio - visão geral (Sem. 2)
   - Tópicos cobertos nas aulas da semana:
     - Por que o R? O que é o R? O que é o RStudio?
     - Conhecendo o RStudio -
       - painéis e elementos principais
       - Script x Console
       - A modularidade do R e os "pacotes"/ "bibliotecas" (CRAN)
     - Tipos básicos de objetos no R
       - Comandos básicos
       - leitura de dados
       - aritmética
       - filtragem

3. Explorando Trabalho Intermitente e Parcial - Projeto 1(Sem 3)
    - Objetivos da semanal: reproduzir um fluxo completo de trabalho: importação, limpeza, exploração e processamento, formatação e exportação de dados. O objetivo da semana será importar arquivos csv com dados de trabalho intermitente e parcial, criar objetos no R a partir desses arquivos, fazer filtragens básicas, e produzir um gráfico
    - Nessa semana, os objetivos de conhecimento dentro do ambiente RStudio serão conhecer:
        - importação de dados (csv) – carregar módulos/bibliotecas
        - objetos – data.frame e list
        - filtragem(subset)
        - visualizando objetos – str e class
        - plotagem simples

4. População e tabelas - Projeto 2 (Semana 4)
    - Objetivos da semana – importar e processar dados de xls, modificar leiaute dos mesmos, calcular novas colunas/variáveis, exportar tabela formatada em csv.
    - Elementos a conhecer do R:
        - importar planilhas
        - visualizando objetos - summary
        - pacote dplyr – junção e reformatação (rbind, cbind, bind_rows)
        - criação de tabelas formatadas
        - Exportação para csv e xls

5. Dados de Assistência SUAS - Projeto 3 (Semana 5)
    - Objetivos da semana – acesso básico a dados do censo suas – estimar indicadores municipais vinculados às populações locais
    - A conhecer no RStudio:
        - download.file e unzip

    - utilizar pacotes não disponíveis no CRAN, ou com versões mais atualizadas fora do mesmo
    - scripts – microdadospoliticasocialbrasil voltados ao CensoSUAS
    - pivot_longer / pivot wider
    - plotagem com ggplot
    - introdução a funções

6. Orçamento, Transparência e Assistência Estudantil - Projeto 4 (Semana 6)
    - Fluxo completo com dados retirados do portal da transparência
    - A conhecer no R:
        - laços ou loops (lapply, sapply, apply, for)
        - pacote portransp
        - plotly (gráficos interativos)
        - exportação em html interativo a partir do plotly

7. Informações de Educação - Censos - Projeto 5 (Semana 7)
    - Fluxo completo com dados de Censo Escolar ou Censo da Educação Superior
    - A conhecer no R:
        - Pacote microdadosbrasil
        - funções de sumarização – (ex. aggregate , colsums, rowsums)
        - introdução a mapeamento

8. Informações de Saúde (tabnet/DATASUS) – Projeto 6 (Semana 8)
    - Fluxo completo de trabalho com dados do datasus
    - A conhecer no R:
        - pacote datasus (dados populacionais e de saúde)
        - pacotes de informação econômica complementar (sidrar e ipeadata)
        - função source
        - introdução a funções de predição simples (predict)

9. CadÚnico – trabalhando com seus microdados – Projeto 7 (Semana 9)
    - Fluxo completo de trabalho com dados do CadÚnico (foco em BPC, Bolsa Família e primeira infância)
    - A conhecer no R:
        - aproveitando exemplos compartilhados - scripts cadunico (JACOB, G.)
        - scripts microdadospoliticasocial (indicadores) relativos ao cadunico
        - introdução ao trabalho com dados amostras complexos – pacote survey

10. Rendimento e Ocupação via PNADc – Projeto 8 (Semana 10)
    - Fluxo de trabalho completo com dados da PNADc, com foco nas informações de trabalhadores informais
    - A conhecer no R
        - pacote geobr (mapeamento por detalhe até bairro
        - pacote PNADcIBGE
        - introdução a séries temporais e dados encadeados
11. Cesta de consumo de trabalhadores (POF) – Projeto 9 (Semana 11)
    - Fluxo de trabalho completo com dados da Pesquisa de Orçamentos Familiares, com foco em reclassificar famílias de acordo com critérios de classe social, ao invés de faixa de renda, comparando POF 2008-2009 e POF 2017-2018
    - A conhecer no R:
        - importando microdados de maneira direta, sem pacotes
        - scripts pofesferas
        - compatibilizando diferentes agregações – da POF para os setores do SCN
        - extrapolações simples da cesta de consumo para o Consumo das Famílias nas Contas Nacionais
12. Previdência Social – Projeto Final (Semana 12)

- Projeto final obtendo dados de previdência social (script previdência social e de portais de dados abertos com sistema CKAN), procurando informações para o Espírito Santo
- A conhecer do R:
  - CKAN, ipeadata, sidrar
  - brazilmaps
  - fixação do uso de funções de loop para procedimentos repetidos de:
    - importação
    - tratamento
    - plotagem
  - informação georreferenciada
  - mapeamento e deflacionamento.

BIBLIOGRAFIA

ALCOFORADO, L. F.; LEVY, A. **Visualização de dados com o software R**. Niterói: LFA, 2017.

BOB RUDIS. Making waffle charts in R (with the new "waffle" package) | R-bloggers. **R-Bloggers**. Recuperado fevereiro 21, 2019, de https://www.r-bloggers.com/making-waffle-charts-in-r-with-the-new-waffle-package/, 2015.

BORGES, R. Microdados de Política Social do Brasil (scripts para R). Disponível em: https://gitlab.com/rodrigoesborges/microdadosbrasilpoliticasocial . 2020

BORGES, R. Curso prático crítico de introdução ao R (scripts para R). Disponível em: https://gitlab.com/rodrigoesborges/conjunturacurso . 2020

CHANG, W. **R Graphics Cookbook**. Recuperado de http://www.it-ebooks.info/book/1316/%5Cnhttp://ebooks.cambridge.org/ref/id/CBO9781107415324A009%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/8366360, 2013.

DE OLIVEIRA, P. F.; GUERRA, S.; MCDONNELL, R. **Ciência de Dados com R - Introdução**. Recuperado janeiro 1, 2019, de https://cdr.ibpad.com.br/cdr-intro.pdf, 2018.

JĀNIS GULBIS. Data Visualization – How to Pick the Right Chart Type? **Eazybi.com Blog**. Recuperado fevereiro 21, 2019, de https://eazybi.com/blog/data_visualization_and_chart_types/, 2016.

NEIL SAUNDERS. Infographic-style charts using the R waffle package | What You're Doing Is Rather Desperate. **WHAT YOU'RE DOING IS RATHER DESPERATE Notes from the life of a [data] scientist**. Recuperado fevereiro 21, 2019, de https://nsaunders.wordpress.com/2017/09/08/infographic-style-charts-using-the-r-waffle-package/, 2017.

PARADIS, E. **R for Beginners**, 2005.

PENG, R. D. **R Programming for Data Science**. v. 1. Leanpub, 2015.

PENG, R. D.; MATSUI, E. **The Art of Data Science: A Guide for Anyone Who Works with Data**. v. 53. Leanpub. Recuperado de http://leanpub.com/artofdatascience, 2017.

R Cookbook, 2nd Edition. . Recuperado maio 5, 2020, de https://rc2e.com/.

RIGGS, S.; KROSING, H. **PostgreSQL 9 Administration Cookbook: Solve real-world PostgreSQL problems with over 100 simple, yet incredibly effective recipesBirmingham: Packt Publishing**, 2010.

ROBERT KOSARA. Women in IT – Squaring the Pie? **Eagereyes**. Recuperado fevereiro 21, 2019, de https://eagereyes.org/techniques/square-pie-charts, 2006.

SÄRNDAL, C. E.; SWENSSON, B.; WRETMAN, J. **Model assisted survey sampling**. New York: Springer, 2003.

SONASSI. Simple daily git workflow. . Recuperado fevereiro 22, 2019, , 2012.

VINCENT DRIESSEN. A successful Git branching model. **NVie.com**. Recuperado fevereiro 22, 2019, de https://nvie.com/posts/a-successful-git-branching-model/, 2010.

WICKHAM, H. **ggplot2: Elegant Graphics for Data Analysis**. Springer. doi: 10.1007/978-0-387-98141-3, 2015.

WICKHAM, H. **Advanced R**. Recuperado agosto 5, 2019, de https://adv-r.hadley.nz/.

XIE, Y. **Dynamic Documents with R and knitr**. Recuperado de http://www.crcpress.com/product/isbn/9781482203530, 2013.

FONTES DE DADOS PRINCIPAIS :

DATASUS (MS)

IBGE

IPEA

DADOS ABERTOS (DATAPREV)

PORTAL DA TRANSPARÊNCIA (BRASIL)

CENSOS ESCOLARES (MEC)

CENSO SUAS (Ministério da Cidadania)

MICRODADOS DO CADUNICO (Ministério da Cidadania)

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM POLÍTICA SOCIAL

ANEXO – DA PLATAFORMA

A plataforma utilizada será um ambiente virtual de aprendizado atualizado de Moodle, com acesso a nuvem de arquivos (Nextcloud) e salas virtuais (vídeo chamadas com Jitsi embutido). Privilegia-se a interação direta com alunos, ao invés de formato vertical e com barreira ao docente. Será possível acessar às aulas gravadas por meio do ambiente virtual, para aqueles que não puderem participar presencialmente em alguma sessão.
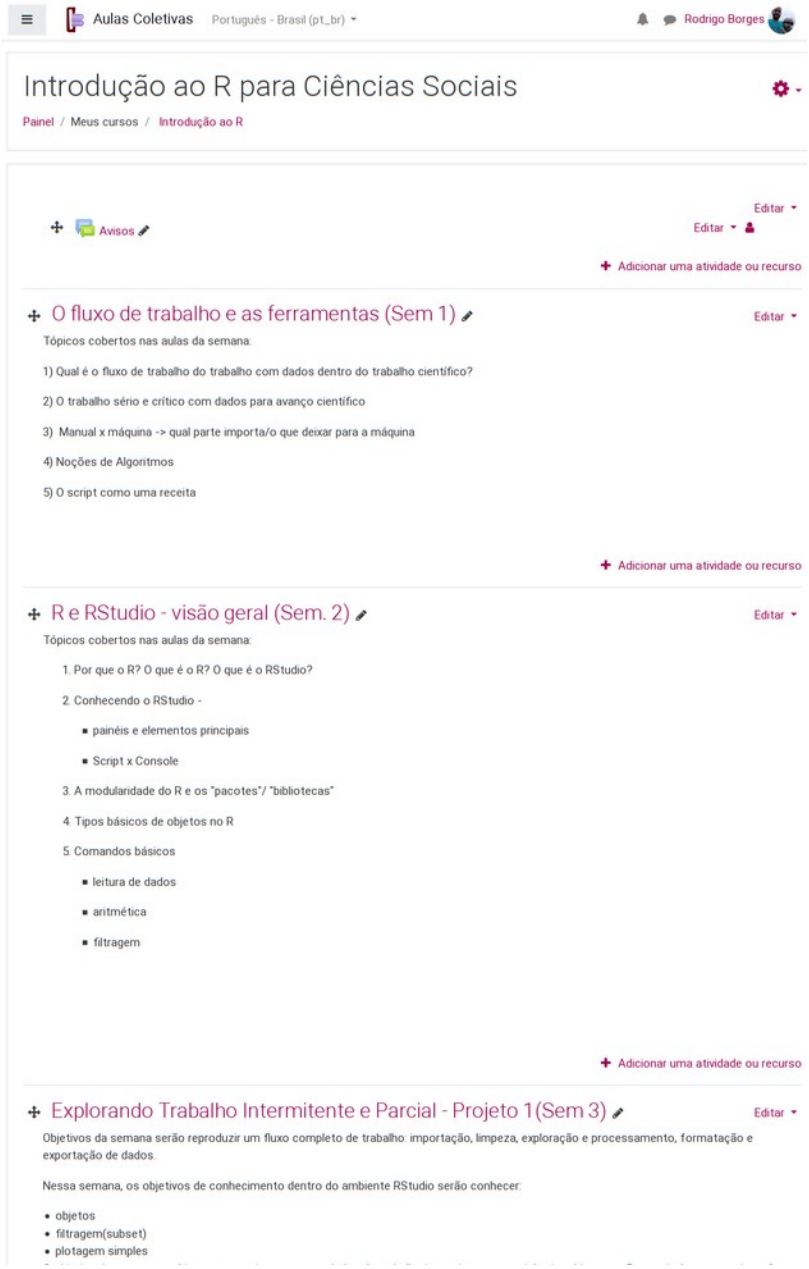
Figura 1: Ambiente Virtual de aprendizado