**PRESIDÊNCIA DA REPÚBLICA**

**SECRETARIA DE DIREITOS HUMANOS**

**Secretaria Executiva**

**Coordenação-Geral de Informações e Indicadores em Direitos Humanos**

Resumo Executivo nº 2/2014

1. **Apresentação**

Este resumo visa a apresentar a proposta de realização de cursos de métodos quantitativos em planejamento, monitoramento e avaliação de políticas públicas, visando a qualificar a equipe da Secretaria de Direitos Humanos da Presidência da República e de ministérios parceiros.

1. **Público**

Com 30 vagas e aberto a todos os servidores da SDH/PR, o curso será ofertado também a servidores do Ministério do Desenvolvimento Social (MDS), da Secretaria de Políticas para a Mulher (SPM), da Secretaria de Políticas de Promoção da Igualdade Racial (SEPPIR), Secretaria de Assuntos Estratégicos (SAE), Ministério do Trabalho e Emprego (MET), Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e Departamento de Informática do SUS (DATASUS).

1. **Conteúdo Programático**

Totaliza 44 horas aula, divididas em 11 aulas de quatro horas cada estruturadas em três sessões semanais.

Conteúdo: O curso busca familiarizar o aluno com a manipulação de bases de dados governamentais e aplicação de métodos estatísticos. O aluno será apresentado à manipulação de bases de dado no Software R e ao uso das principais bases de dados governamentais neste software. Adicionalmente, serão apresentados instrumentos analíticos para análise e visualização destes dados, incluindo estatística básica (distribuições, inferência estatística e regressão).

1. **Detalhamento dos Conteúdos**

**Aula 1: arquitetura de dados, principais fontes no Brasil e escolha de software estatístico**

* R em 30 minutos: tour rápido do material do curso
* Tipos e Estruturas de dados: dados estruturados e desestruturados. Registros administrativos e pesquisas amostrais e censitárias
* O que são dados abertos?
* Fontes de informação: principais pesquisas e bancos de dados governamentais;
* Alternativas disponíveis para trabalhar com dados (softwares estatísticos, linguagens de programação, DBMS)
* Arquitetura de dados: dados tabulares. Noções de bancos de dados relacionais (DBMS, Data WareHouse, ferramentas de BI)
* *Big Data* e dados governamentais

**Aula 2: Iniciação no R**

* Instalação R, Rstudio e Notepad++
* Obter e instalar pacotes do R
* Escolher diretório e criar as subpastas
* Ambientação: R e Rstudio, interativo vs. R file
* Criar objetos, atribuição, listar objetos
* Tipos de dados do R (char, num, logical, int) e tipos de Variáveis (categórica, cat. Ordinal, Numérica (continua, discreta)
* Tipos de Objetos: vetores, matrizes, listas, dataframe, factors
* Objetos e seus atributos
* Manipulação básica de objetos

**Aula 3: Manipulação de bases de dados**

* Manipulação de bases de dados:
  + Extrair subconjunto dos dados
  + Ordenar um banco de dados
  + Eliminar variáveis existentes e acrescentar novas
  + Criar variável categórica a partir de variável numérica
  + Recodificar variáveis
  + Variáveis categóricas (labels)
* Importação de dados, diversos formatos:
  + formato tabular (.csv, .txt)
  + .txt colunado
  + Excel
  + Outros formatos “binários” (SPSS, SAS, STATA)

**Aula 4: Manipulação de bases de dados e dados governametais**

* Importação de dados Governamentais:
  + Pactores: IBGE.pesq, le.pesquisa e survey
  + dados da PNAD
  + dados da PNAD-contínua
  + dataSUS
  + RAIS
* Junção de bases de dados:
  + Vertical: rbind
  + Horizontal: merge
* Reformatar banco de dados: reshape

**Aula 5: Manipulação de bases de dados**

* Manipulação de variáveis textuais
* Manipulação de variáveis de datas
* Aplicações:
  + Transformar o indicador de proporção para percentual
  + Uso de deflator

**Aula 6: Funções, debuggin e tabulações**

* Noções de Programação de funções
  + Argumentos, erros,
* Digas para debuggin
* Programação orientada a objeto
* Dúvidas e atividades práticas.
* Tabulações
  + Exemplos práticos com base na PNAD

**Aula 7: Tabulações e Visualizações**

* Introdução aos gráficos e visualizações
* Pacote ggplot2
  + Exemplos práticos com base na PNAD

**Aula 8: Estatística Básica**

* Estatísticas Descritivas: População; amostra; Distribuição de Frequência; Medidas de tendência central; Medidas de variabilidade; Medidas de posição
  + Exercícios no R: cálculo de estatísticas descritivas
* Estatística Inferencial; Distribuição normal de probabilidade;
  + Simulando e visualizando distribuições no R

**Aula 9: Inferência e regressões:**

* Teoria da estimação estatística
* Testes de hipóteses
* Análise de associações e correlações estatísticas
* Regressão linear simples e múltipla

**Aula 10: Amostragem e ponderação de bases de dados**

* Estimativas de tamanhos de amostras
  + Simulação de uma amostra simples e complexa
* Uso dos pesos nas bases públicas governamentais
  + Uso dos pesos na PNAD

**Aula 11 – 4 horas: Extensões e dúvidas**

* Noções de Sistemas Geográficos de Informação (SIG/GIS)
  + Aplicações no R e outros pacotes
* Atividades práticas, dúvidas e extensões.

1. **Cronograma**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Semana | Dia/Mês | Segunda | Terça | Quarta | Quinta | Sexta |
| 1ª | 17/set |  |  | 1ª aula |  |  |
| 19/set |  |  |  |  | 2ª aula |
| 2ª | 22/set | 3ª aula |  |  |  |  |
| 24/set |  |  | 4ª aula |  |  |
| 26/set |  |  |  |  | 5ª aula |
| 3ª | 29/set | 6ª aula |  |  |  |  |
| 01/out |  |  | 7ª aula |  |  |
| 03/out |  |  |  |  | 8ª aula |
| 4ª | 06/out | 9ª aula |  |  |  |  |
| 08/out |  |  | 10ª aula |  |  |
| 10/out |  |  |  |  | 11ª aula |
| (\*) As aulas serão ministradas na SDH/PR, das 8:30 às 12:30 | | | | | |  |

**Breve bibliografia pra as aulas 8 a 10**

**Aula 8: Estatística Básica**

* Livro de referência: Estatística Básica, Bussab e Moretin, parte III (cap. 10, 11 e 12)
* Curso on-line (em inglês) Statistical Inference ([Johns Hopkins Bloomberg School of Public Health](http://www.jhsph.edu/)). Link: <https://class.coursera.org/statinference-005>

**Aula 9: Inferência e regressão**

* Livro de referência: Introdução à Econometria: Uma abordagem moderna, Jeffrey M Wooldridge, mais especificamente a Parte 1: Análise de Regressão com Dados de Corte Transversal
* Curso on-line (em inglês): Regression Models ([Johns Hopkins Bloomberg School of Public Health](http://www.jhsph.edu/)). Link: <https://class.coursera.org/regmods-006>

**Aula 10: Amostragem e ponderação de bases de dados**

* Análise de Dados Amostrais Complexos, Pessoa e Silva, IBGE, 1998. Link: <http://www.ernestoamaral.com/docs/mq13reg/Pessoa1998.pdf>
* Introdução à análise de dados amostrais complexos, Pessoa e Costa, Seminário de Metodologia do IBGE de 2013. Link:<http://eventos.ibge.gov.br/images/smi2013/downloads/MC2/CursoMC2.pdf>