

Literacia sobre dados abertos em transportes



Rosa Félix
rosamfelix@tecnico.pt

Apresentação

Rosa Félix

Investigadora e Professora convidada (CERIS, Instituto Superior Técnico)

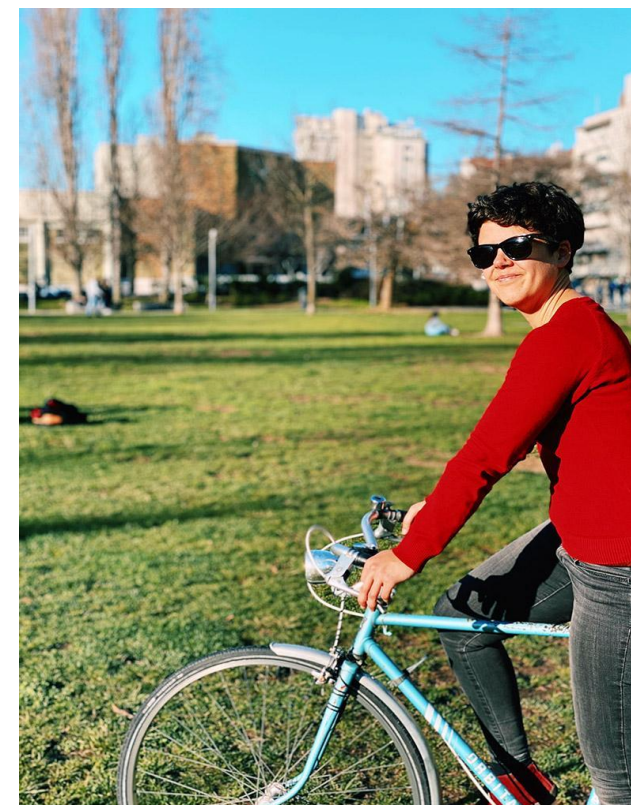
Engenheira do Território

PhD em Sistemas de Transportes (Instituto Superior Técnico, ULisboa)

- Investigação em mobilidade urbana e modos ativos
- Utiliza R e SIGs diariamente
- Entusiasta de software livre e investigação replicável
- Lecciona SIG com R em cursos de mestrado no IST
- Desenvolve ferramentas em R para análise de dados e apoio a políticas públicas



Portland State
UNIVERSITY



— Apresentações

Expectativas

Tem experiência com SIG ou com R?

LPrograma

3 módulos de 4 horas

Dia 1

- Apresentação
- Introdução ao R e RStudio: instalação de software e pacotes
- R-base: exemplos e exercícios
- Tratamento de dados: uso do pacote dplyr para select, filter, left-join, group e summarize
- Mapas interativos em R

LPrograma

3 módulos de 4 horas

Dia 2

- Dados abertos em mobilidade e transportes
- Fontes de dados abertos e redes viárias
- Exemplos nacionais e internacionais de análises com dados abertos de transportes
- Introdução aos GTFS
- GTFS: exemplos e exercícios

LPrograma

3 módulos de 4 horas

Dia 3

- Análise espacial e acessibilidade com r5r
- Boas práticas de partilha e abertura de dados de mobilidade e transportes
- Exercício prático em grupo com bases de dados de operadora de mobilidade

L Materiais

Os materiais deste módulo encontram-se online em

<https://rosamfelix.github.io/PGmob360>

Este módulo foi feito à medida para si
Aulas com abordagem *hands-on*!

Não é expectável que aprenda a programar no final do curso...

LSoftware

R e Rstudio

- Todos os dias são **recolhidos milhões e milhões de dados** de mobilidade e transpores, só na AML
 - Bilhética e portagens
 - Sensores e semáforos
 - Contadores automáticos com deteção de pessoas / veículos
 - Consumo de frotas de veículos de operadora
 - Registos de telecomunicações
 - ...
- O Microsoft Excel, até 2007, só conseguia abrir dados com um máximo de **65.536** linhas.
- Atualmente, é possível ter até **1.048.576** linhas num ficheiro excel.
- O R, Python, e outros têm capacidade consoante a memória RAM do computador (algumas excepções vão até 2^{31} linhas de dados = **2,147 mil milhões**).

Software

Black boxes vs. Código aberto



Sidney Harris, 1977

Porquê R

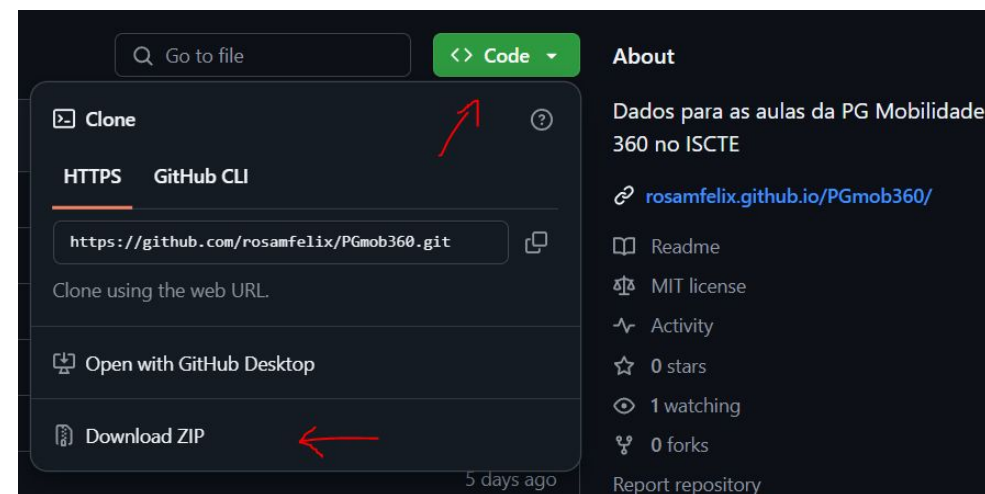
- Compreender **todos os passos**
- Enorme **comunidade online** – alguém já terá tido o mesmo problema e disponibiliza soluções. Não é necessário saber programar de cor do início ao fim
- Ferramentas poderosas, permitem vários **tipos de análise** e aceitam dados de **vários formatos**
- Usa apenas os “pacotes” necessários para as funções que precisamos, **consumindo menos memória**
- Não grava por cima dos ficheiros originais* - Podemos fazer vários testes
- **Replicabilidade e transparência** nas análises e resultados

— Instalar R e RStudio

Descarregar os materiais

<https://github.com/rosamfelix/PGmob360>

- Criar uma pasta para este modulo o seu computador
- Descarregar os dados .zip do repositório Github e exportar para a pasta criada
- Manter as pastas “code” e “data”. O resto pode eliminar
- Abrir o Rstudio (e não o R)
- Criar um novo projeto:
“File” -> “New project” -> “Existing Directory” ->
escolher a pasta e clicar em “create project”.



— Introdução ao R

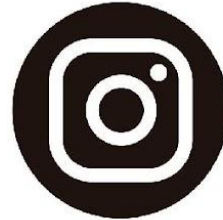
Follow us



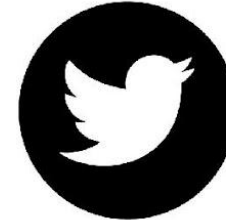
LinkedIn



Facebook



Instagram



Twitter

iscte — **Executive
Education**

Avenida Professor Aníbal Bettencourt, 1600-189 Lisboa | (+351) 217 826 100 | indeg@iscte-iul.pt