

# Eltecon Data Science Course by Emarsys

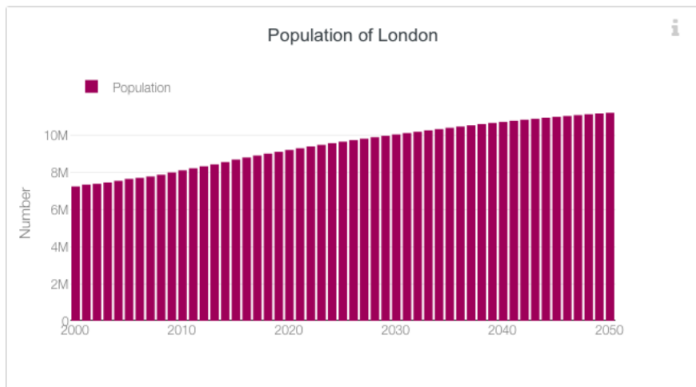
## Communication Tools for Data Science

Tamás Koncz

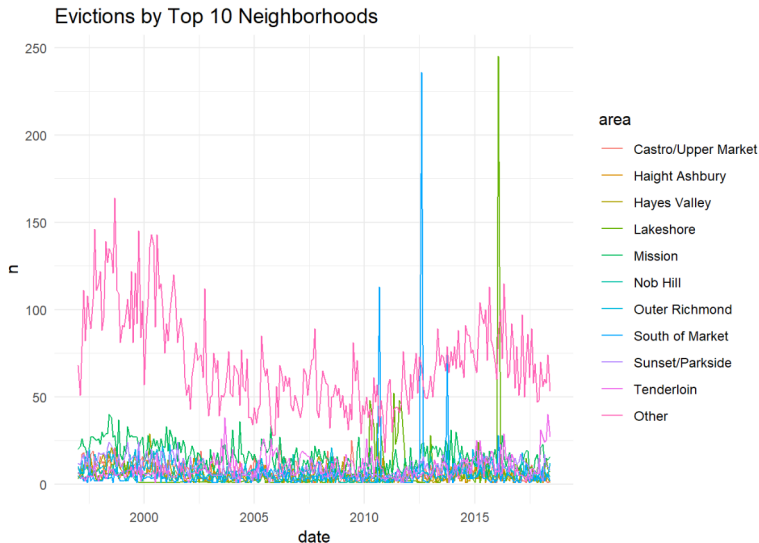
October 2, 2019

# Recap - List the issues & propose improvements!

# London's population



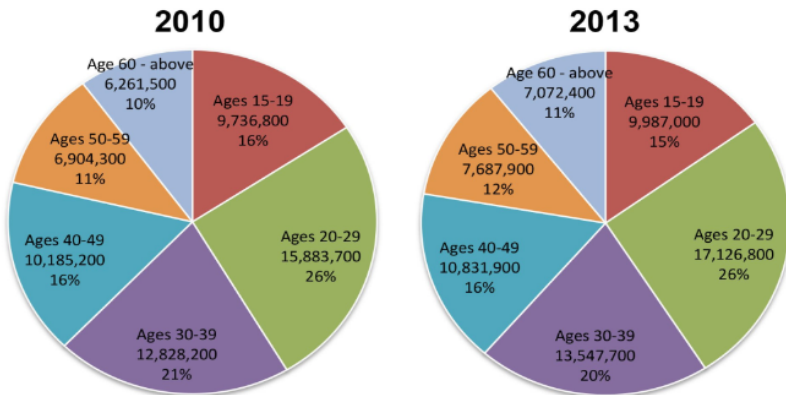
# San Francisco evictions



# Population by age group

Before

2010 and 2013 Population by Age Group



Source: Charted 2010 Census data and charted PSA estimated 2013 population data

# QUIZ TIME

**.Rmd**

# What is R Markdown?

- Authoring framework for data science
  - You can use a single R Markdown file to both
    - save and execute code
    - generate high quality reports that can be shared with an audience
  - R Markdown documents are fully reproducible
  - They support dozens of static and dynamic output formats
- Source: [rmarkdown.rstudio.com](https://rmarkdown.rstudio.com)



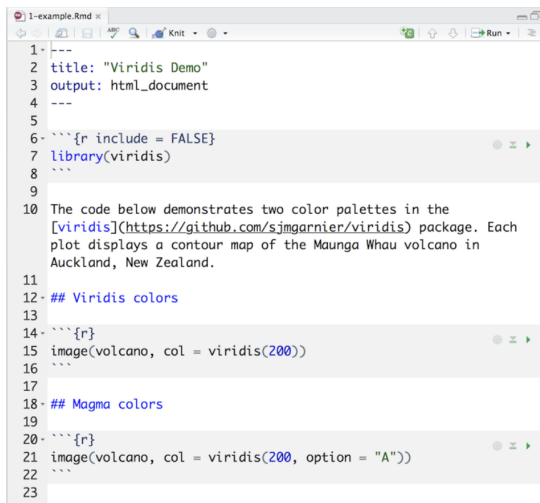
# Other uses of (R) Markdown

- Readme documents
- R pkg vignettes
- Blogposts
- Books

# Parts of an .Rmd file

- An (optional) YAML header surrounded by `---`
- Code chunks surrounded by `````s
- Text mixed with simple text formatting

# Parts of an .Rmd file



```
1 ---
2 title: "Viridis Demo"
3 output: html_document
4 ---
5
6 ```{r include = FALSE}
7 library(viridis)
8 ```
9
10 The code below demonstrates two color palettes in the
11 [viridis](https://github.com/sjmgarnier/viridis) package. Each
12 plot displays a contour map of the Maunga Whau volcano in
13 Auckland, New Zealand.
14
15 ## Viridis colors
16
17 ```{r}
18 image(volcano, col = viridis(200))
19 ```
20
21 ## Magma colors
22
23 ```{r}
24 image(volcano, col = viridis(200, option = "A"))
25 ```
```

# Basics of text formatting

```
- *italics*
- **bold**
- `code`
- \[links\]\(rmarkdown.rstudio.com\)
```

- *italics*
- **bold**
- `code`
- [links](https://rmarkdown.rstudio.com)

# For more, please see

- <https://rmarkdown.rstudio.com/lesson-1.html>
- <https://bookdown.org/yihui/rmarkdown/>
- Cheatsheet
- Recommended reading: How R Helps Airbnb Make the Most of Its Data

# Cons

- The longer the analysis, the more messy it becomes (*scroll-scroll-scroll*)
- Does not facilitate good coding style
- Knit fails on error
- Reporting results is not always parallel with the code

# Interactive visualizations – Plotly & Shiny DEMO TIME

# Today's dataset



# Kaggle

The Kaggle logo, featuring the word "kaggle" in a lowercase, rounded, blue sans-serif font, followed by a small "TM" trademark symbol.