

DSS
@Berkeley

R for Data Science
Workshop Series #4

Instructors + TA's



Created by Rockicon
from Noun Project

Instructors



Alex Nakagawa

alex.nakagawa@berkeley.edu

Sophomore

Computer Science & Cognitive Science



Jun Seo Park

junseopark@berkeley.edu

Junior

Statistics & Economics

TA's



Shalini Kunapuli



Fayyaz Ahamed



Pooja Doctor

Lesson Plan



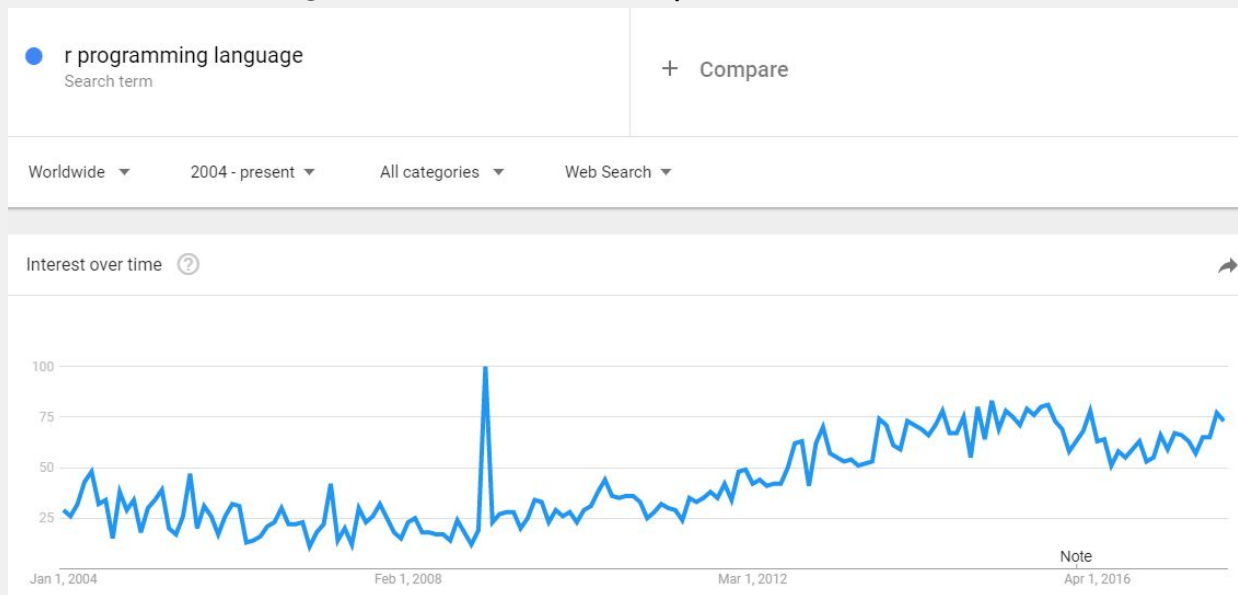
Lesson Plan

1. Overview
 - 1.1. History of R
 - 1.2. RStudio
 - 1.3. Tidyverse
2. Using R
 - 2.1. Basic syntax/RStudio
 - 2.2. Vectors/Matrices/Dataframes
 - 2.3. readr/dplyr/ggplot2
3. Mini-Project

Overview

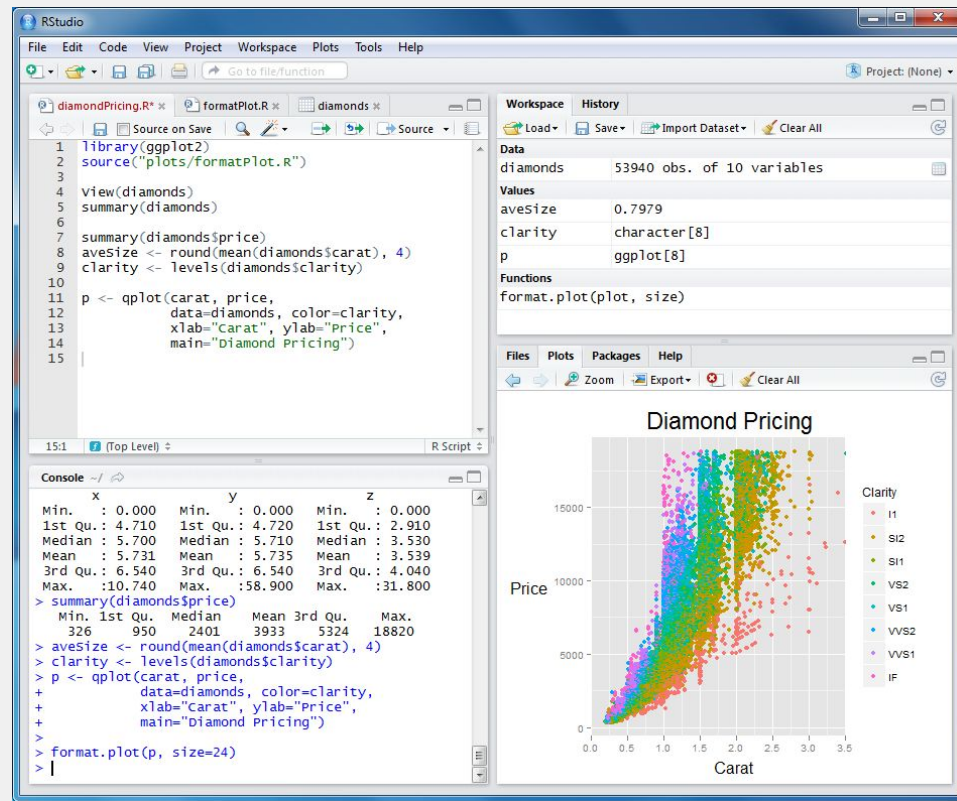
1.1 History of R

- Origins in S (developed by Bell Labs)
- Ported to open-source software in 1990s
- Gained traction following 2008 recession for open-source, OOP, well-documented environment



1.2 RStudio

- Primary IDE for R
- Built-in support for plots, packages, documentation, environment details, etc.
- Build dashboards with R Markdown files



1.3 Tidyverse

- Hadley Wickham
- ggplot2: Plotting data
- dplyr: Data manipulation
- readr: Reading data files
- tidyr: Create “tidy” data
 - Each observation is a row
 - Each variable is a column
- shiny: Interactive data dashboards



Using R!!

<http://tinyurl.com/f17-dss-r>

<https://piazza.com/class/j7s01y165odq5>

Data Science FAQs