

From Zero to

A Hands-On Approach to R Stats

Pierre Lafortune

pierre.a.lafortune@gmail.com

Agenda

- R Introduction
- Problem Statement
- Live Coding
- Participant Challenges
- Learn More

R Quick Launch

- Statistical software environment
- Built on S programming language
- Named R after originators Robert Gentleman and Ross Ihaka
- Created in 1993
- Open-source under General Public License
- Source-code written in C, Fortran, and R

Congratulations!

- You are the newest data scientist at Tesla Motors
- First task: Fuse two data sets
- Skills Required:
- Data cleansing, Missing Value imputation, Classification Modeling, Clustering, Regression Analysis, Data Manipulation

Dataset Link



Dataset A: Tesla
Customers

Dataset B: USA Consumer
Database



That's easy.
Use a common
id!

What if no
common
link exists?

Link Details

Dataset A

| ID | Gender M | Gender F | Indivd_Age | Income | Environ mentalis t |
|------|----------|----------|------------|--------|--------------------------|
| 1043 | "" | female | 34 | 9 | 1 |
| 5021 | MALE | "" | 72 | 12 | 1 |
| 4863 | "" | FEMALE | 81 | 6.5 | 1 |
| 4651 | male | "" | NA | 3.2 | 0 |
| 1002 | 99 | 99 | 49 | 21 | 0 |
| 7523 | Male | "" | 21 | -7.9 | 0 |

Dataset B

| Age | Gender | Income | Environmental ist |
|-----|--------|--------|----------------------|
| 28 | M | 90 | Yes |
| 35 | F | 81 | No |
| 14 | F | 65 | No |
| 65 | M | 42 | Yes |
| 44 | M | 110 | No |
| 39 | M | 49 | No |

Learn More

- R Cookbook - Many helpful examples and topics. The “Graphs” section helps tremendously with ggplot2 <http://www.cookbook-r.com/>
- Concise Guide to R – Covers all major R functionality <http://www.cis.upenn.edu/~matuszek/Concise%20Guides/Concise%20R.html>
- Stack Overflow – The best help is available when you post a question here. <http://stackoverflow.com/questions/tagged/r>
- Cross Validated – For statistics and modeling help. <http://stats.stackexchange.com/>
- Introduction to Statistical Learning – The BEST way to learn statistics and machine learning. And it’s free!
 - Book: <http://www-bcf.usc.edu/~gareth/ISL/>
 - Video Lectures: <http://www.r-bloggers.com/in-depth-introduction-to-machine-learning-in-15-hours-of-expert-videos/>