



Pearson

What drumming taught me about leading data science teams

Kacper Łodzikowski

29 Sep 2017 :: [Why R?](#)



Which one is
the drummer?



1

2

3

4



1

2

3

4

READY TO JAM



Hidden / 9GAG

EVERYONE ELSE STILL TUNING



1

They set the beat

Flite

Cinematic Orchestra

♩ = 136

A musical score for Flute (Cinematic Orchestra) in 8/8 time. The key signature is A major (no sharps or flats). The tempo is indicated as ♩ = 136. The score consists of eight staves of music, numbered 9 through 33. Each staff begins with a dynamic marking of *mf*. The music features a continuous pattern of eighth-note pairs and sixteenth-note groups, primarily consisting of X_x and X_x symbols, with occasional X_x and X_x symbols. Measure 9 starts with a X_x symbol. Measures 13 and 17 start with a X_x symbol. Measures 21 and 25 start with a X_x symbol. Measures 29 and 33 start with a X_x symbol.

1

They set the beat

- goals
- tools
- processes

Flite

Cinematic Orchestra

♩ = 136

A

9 +

13

17

21

25

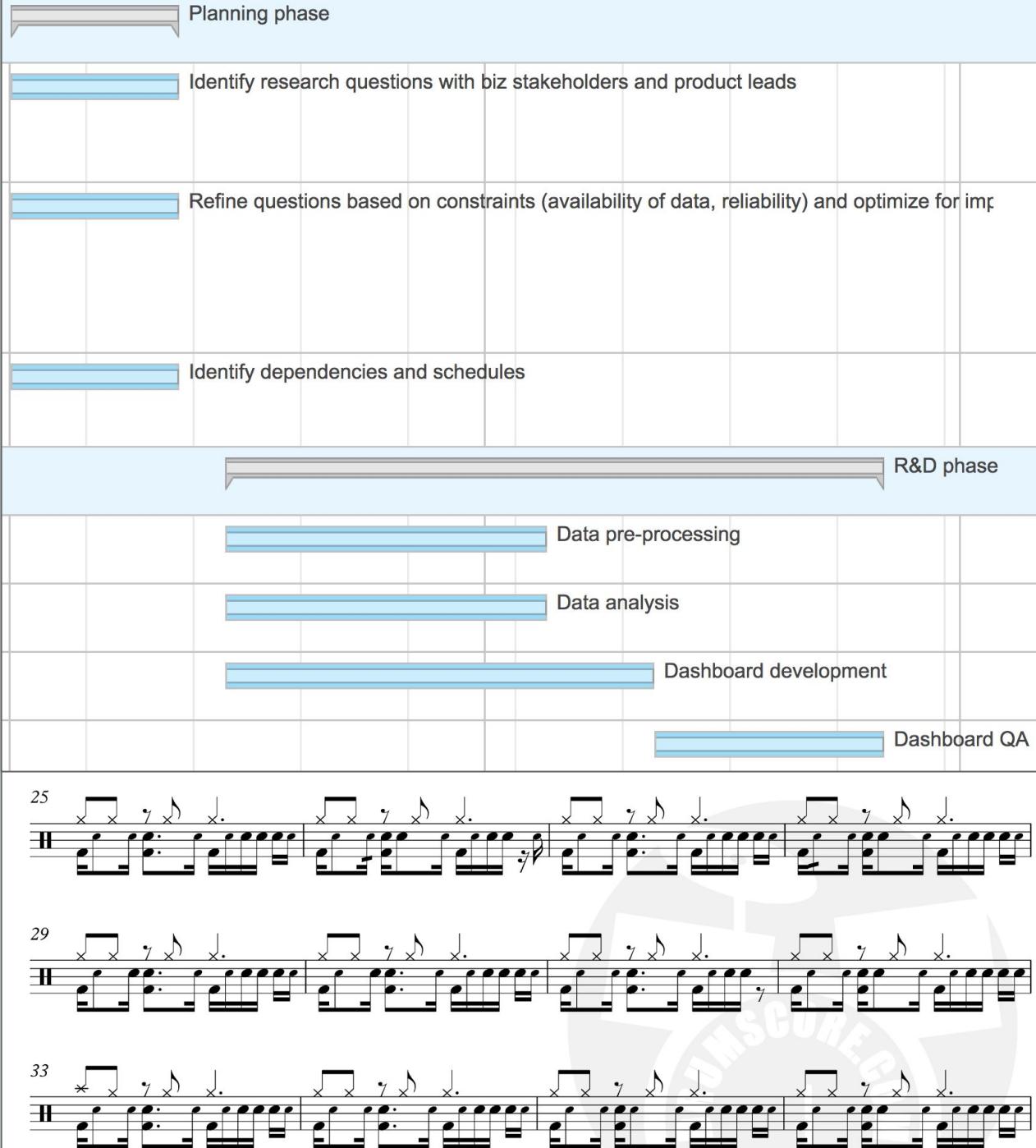
29

33

1

They set the beat

- goals
- tools
- processes



1

They set the beat

- goals
- tools
- processes

document ID: P15

reviewer: Kacper Łodzikowski

approver: Kacper Łodzikowski (process owner)

status: draft/ approved

1. PURPOSE

The goals of this process is to provide product and content teams in Global Product with analytics services that provide them with internal evidence as to how learners and teachers use their products, so that the teams can make informed decisions about the development of their products based on data.

2. SCOPE

[This high-level process](#) has three sub-processes:

2.1. Managing requests

2.2. Obtaining data

2.3. DSLC (Data Science Life Cycle)

which are described in [the Efficacy Analytics Handbook](#).

3. MAIN RESPONSIBILITIES

3.1. project manager

- Responsible for making sure the Efficacy Analytics team follows the Efficacy Analytics process.
- Responsible for making sure the Efficacy Analytics process is

1

They set the beat

- goals
- tools
- processes

This document describes the processes and best practices in the Efficacy Analytics team within the Efficacy & Research unit. This is a living document. Please feel free to update it accordingly. Please note this document contains confidential business information.

About our work

[Team structure](#)

[Goals for 2017](#)

[Other analytics teams](#)

Processes

[Efficacy & Research](#)

[Efficacy Analytics](#)

[Pearson IOKI](#)

Best practices

[How to work in an asynchronous team](#)

[Deep work](#)

[MacOS setup](#)

[ENR analytics R package](#)

[Preparing visualisations](#)

[Preparing slide decks](#)

[Naming convention](#)

Docker

[Basic info](#)

[Docker image creation](#)

Packrat

1

They set the beat

- goals
- tools
- processes

Docker

[Basic info](#)

[Docker image creation](#)

Packrat

[New project - steps](#)

[Existing project - steps](#)

Case study

[Other important issues](#)

[Potential weird behaviours](#)

Security

[Pearson and IOKI security regulations](#)

[Data privacy](#)

[Data retention](#)

Professional development

[On-the-job resources](#)

[Conferences and training](#)

Reproducibility

Access

[Pearson products \(databases, demo accounts, etc.\)](#)

[Pearson internal tools / services](#)

[E&R Analytics local servers](#)

Always learning

Contents Notebook

Search This Course 

Introduction

- Welcome 1m 23s
- What do you already know?

1. Define Data Science Success

- Review a typical project 3m 52s
- Work on a data science project 4m 21s
- Compare project challenges 4m 31s
- Set success criteria 4m 8s

2. Use a Data Science Lifecycle (DSLC)

- Use a DSLC 4m 51s
- See how to use a DSLC 4m 18s
- Loop through questions 4m 7s
- Iterate through DSLC sprints 4m 13s

3. Work in Sprints

Data Science Life Cycle

Six steps

- ① Identify
- ② Question
- ③ Research
- ④ Results
- ⑤ Insights
- ⑥ Learn

LinkedIn  1x CC

[Overview](#) [Transcript](#) [View Offline](#)

Author **Released** 12/6/2016 

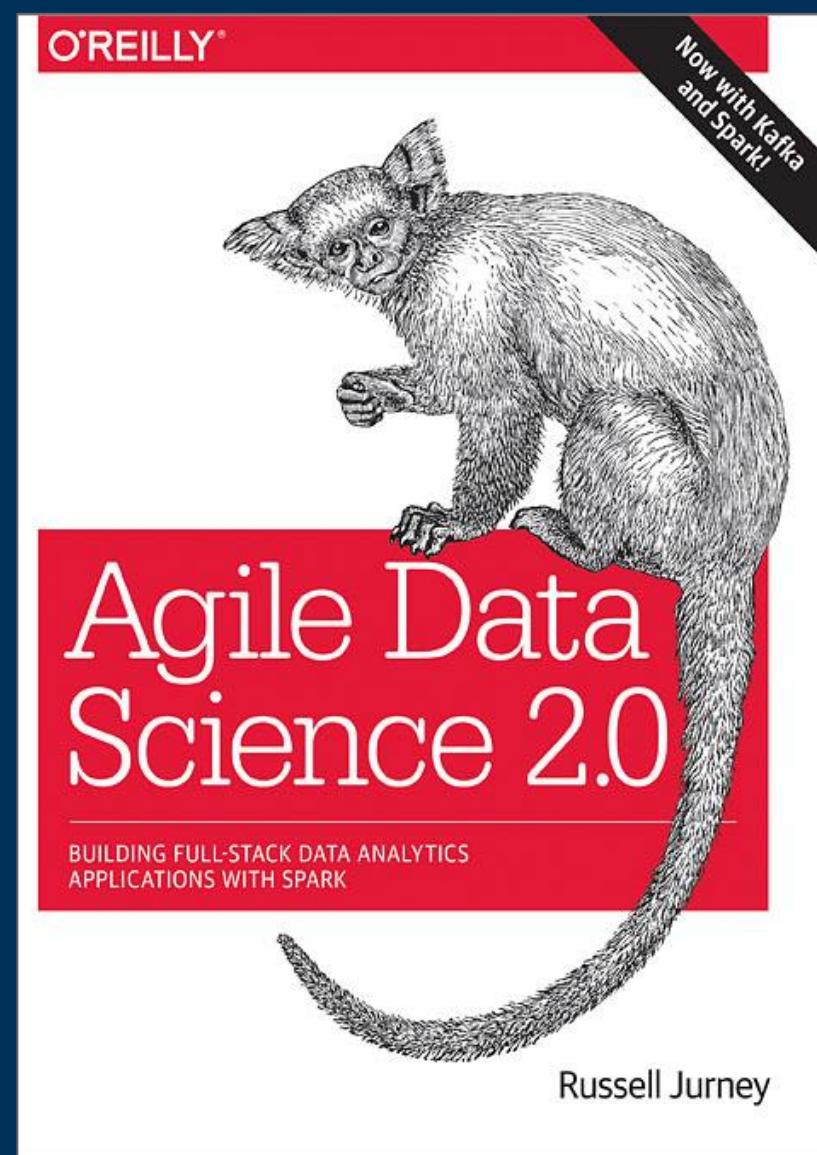


Doug Rose

A data science team asks great questions, explores the data, and delivers key insights. But at the end of the day your team needs to add real business value. Even the best data science teams won't last long in your organization if they can't generate revenue or lower expenses. The best way to generate business value is to deliver a constant stream of key insights in short two-week sprints. These short sprints give you real-time feedback to help

 Skill Level
Appropriate for all

 Duration
1h 16m



2

They set / keep
the tempo

THE TEMPO IS



**WHATEVER
I SAY IT IS**

2

They set / keep the tempo

- follow the process
(quality, deadlines)
- if a project fails,
blame the process
(and improve it)

THE TEMPO IS



**WHATEVER
I SAY IT IS**

2

They set / keep the tempo

- follow the process
(quality, deadlines)
- if a project fails,
blame the process
(and improve it)



Source: <https://gob.gl/7TND47>

Always learning

O'REILLY®

jupytercon

Brought to you by NumFOCUS Foundation
and O'Reilly Media Inc.



AUGUST 22-25, 2017

NEW YORK, NY

jupytercon.com

#JupyterCon

OPINIONATED ANALYSIS DEVELOPMENT

HILARY PARKER
@HSPTER

3

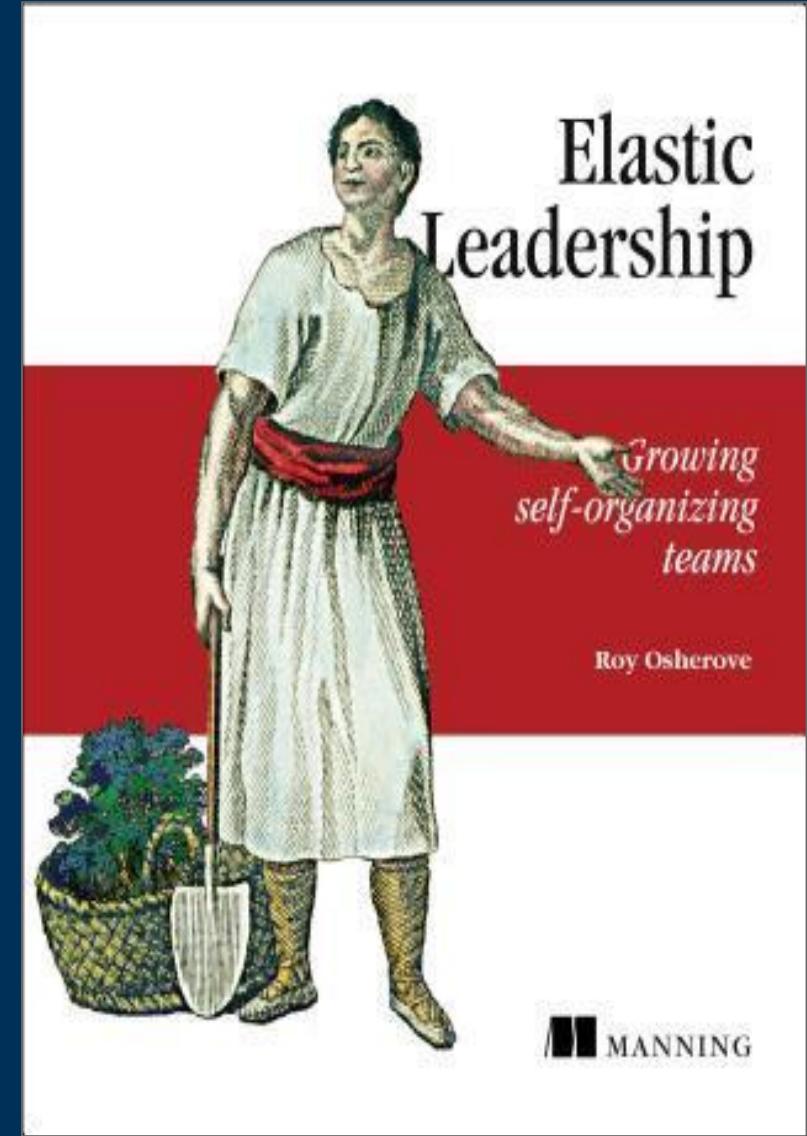
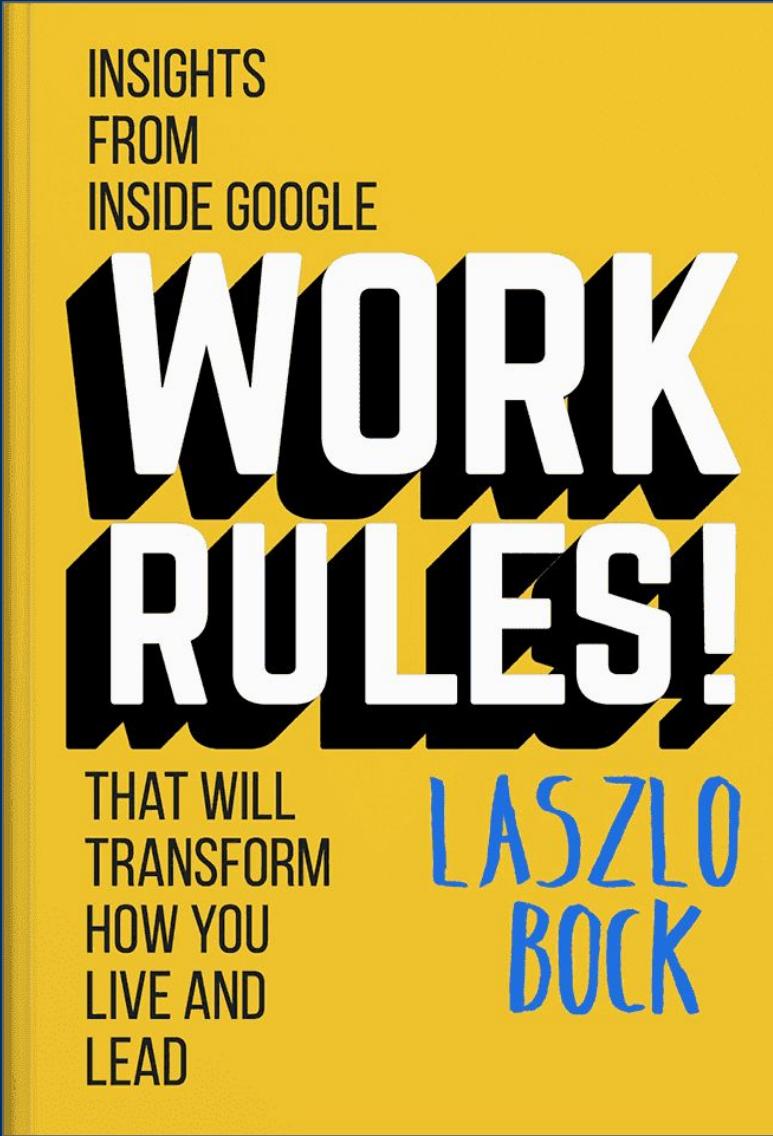
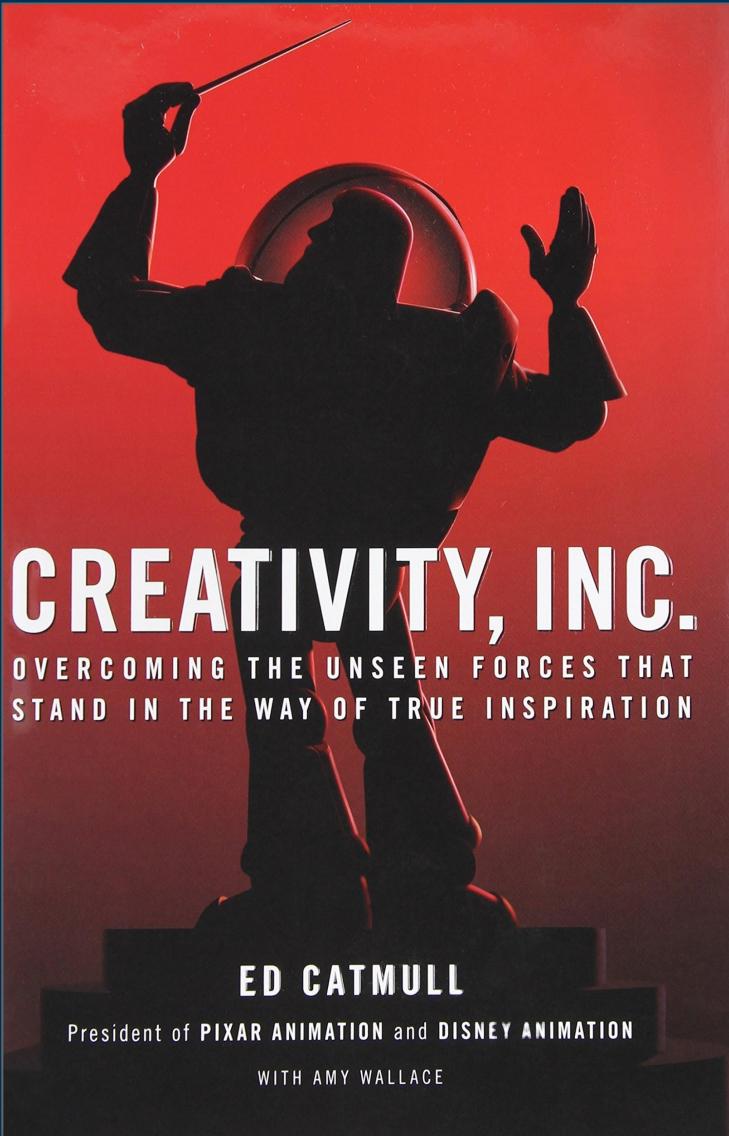
They know their place

- they are at the back
- freedom
- trust



Source: <https://giphy.com/gifs/drummer-wPVfkCFw6uA36>

Always learning



- #1 Set the beat
- #2 Set & keep the tempo
- #3 Know your place

kacper.lodzikowski@pearson.com