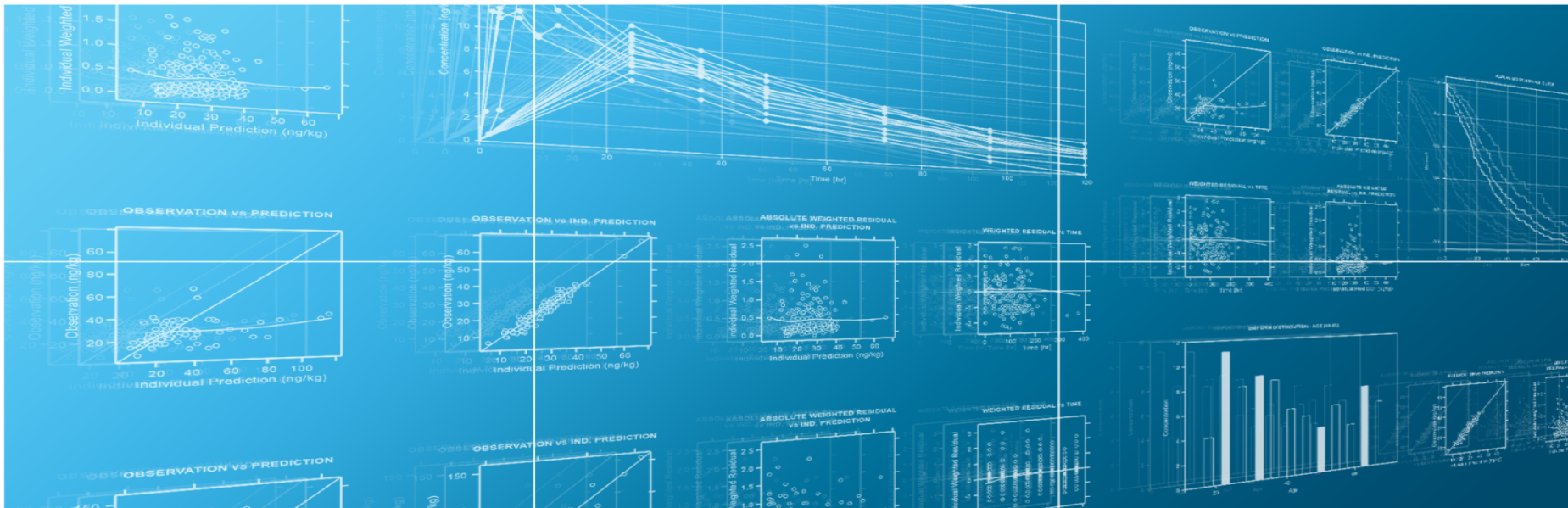


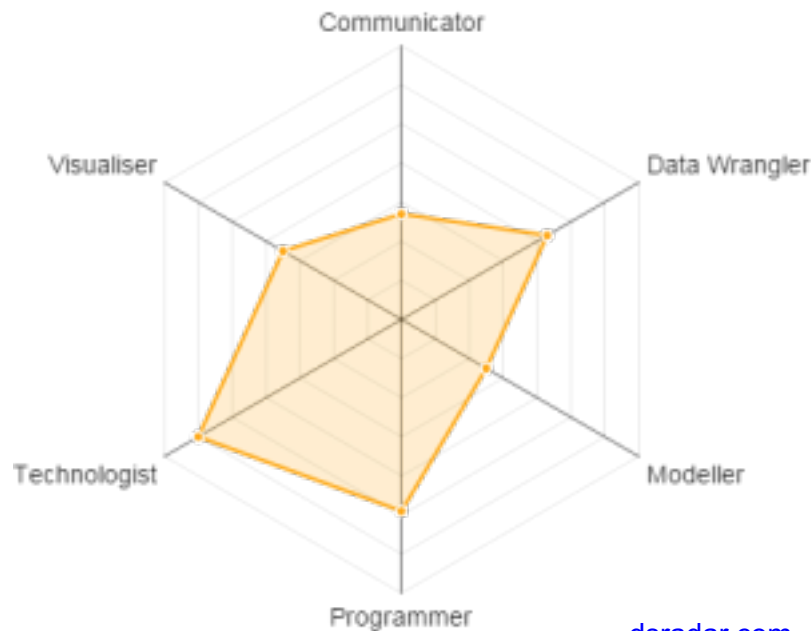
R is Production-Safe



Mark Sellors
Head of Data Engineering - Mango Solutions

About Me

- Head of Data Engineering
- Run Mango Data Labs
- Architecture
- Automation
- Industrialisation
- DevOps/DataOps
- @sellorm



dsradar.com

What is production?

- Many different things to many people
- Could be:
 - A web-scale, public facing server environment
 - Could just be a simple script
- In general it's just something that gets run more than once, and is important to your organisation

**“You can’t use R in
Production”**
-- Anonymous

**“You can’t use R in
Production”**

-- Anonymous

Things We Know About R...

- Mature
- Well documented
- Well understood
- Commercial versions and/or support
- Dynamic – fast moving ecosystem
- Amazing community

Things *other people* think about R...

- Weird syntax
- Not many users
- Only used by statisticians and Data Scientists
- Performance problems
- Not very versatile
- No good for large data sets

Some *myths* about R...

- Weird syntax
- Not many users
- Only used by statisticians and Data Scientists
- Performance problems
- Not very versatile
- No good for large data sets

Disconnect, what disconnect?

- We all know that R is great
- "All the people I follow on twitter are using R in production"
- There are conferences, like this one (!), where even more people are using R

Why the disconnect?

- Great communities can also create powerful feedback loops and echo chambers
- R's domain specific nature makes it somewhat unapproachable for many outside the community
- Fear – A thing that I don't understand must surely be a bad thing



FEAR IS A LIAR.

How do we combat fear?

- Education
- Outreach
- Collaboration
- Empathy



**“Nobody else is using R
in Production”**
-- Anonymous

**“Nobody else is using R
in Production”**

-- Anonymous

Allianz

- What: Insurance Claim scoring
- How: API's in R
- Using: Rserve, Java

Worldpay

- What: Call Centre CRM tool and x-sell pipeline
- How: Web app written in R
- Using: Shiny

ONS

- What: Migrate from SAS
- How: Migrate SAS based services to R
- Using: Plumber & command line

Hedge Fund

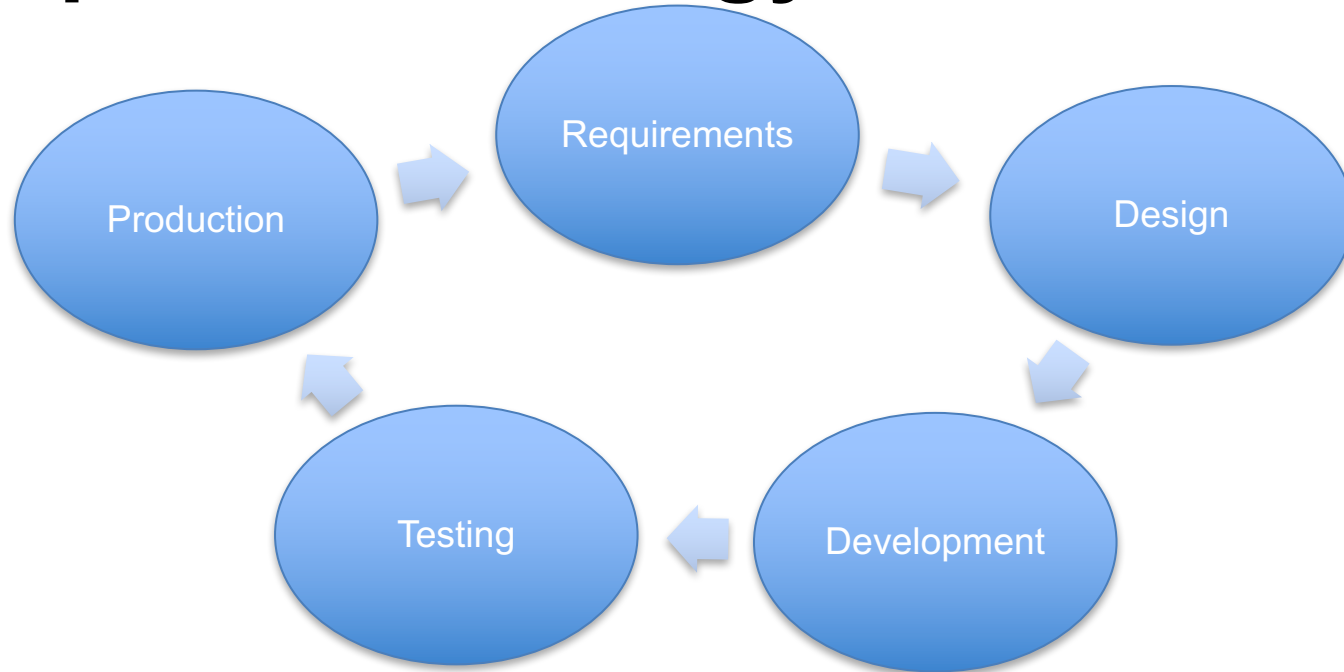
- What: Batch based scoring system
- How: Sophisticated command line application
- Using: R and various databases

How are they doing it?

It's not that hard!

- Collaboration
- Thinking about what they're doing
- Implementing more formal methodologies
- Rigour around processes
- Well managed release strategy
- Isolated environments

Example Methodology



Example Environments





This all sounds like a lot of effort

How much rigour do I need?

Rigour investment checklist

- What's the impact of what you're doing?
- Who's affected if it's not working?
- What is the cost of an outage to the business?
- Why is the new solution even needed?
- Where will the code be run?
- When will it be run?



Think about
what you're
doing!

Why does any of this matter?

- Reduce time to insight
- Lower costs
- Improved quality

**“I’ll never be able to do any of
this in my organisation”**

Don't despair, there is hope

- Find a sympathetic person on the ops side
- Learn to do it yourself
- Start small
- Amaze people!

> summary(talk)

- You can use R in production environments
 - If you think about what you're doing
 - and approach the task with care
 - Find sympathetic collaborators
- If you are doing it already, tell people about it
- If you aren't, what are you waiting for?

Any Questions?