



# Challenges and Opportunities in Data Science in Pharmaceutical industry

---

EFSPI MEETING, 1<sup>ST</sup> JULY 2021

BY JULIA CHERNOVA (BAYER)

# Disclaimer

---

Opinions expressed herein are solely my own and might not reflect the views of my employer

# Overview

---

// Data Science and Statistics

// Data Science SIG Introduction

// Survey Results Snapshot

// Current Challenges in Data Science in Pharma

// Opportunities

# Data Science and Statistics

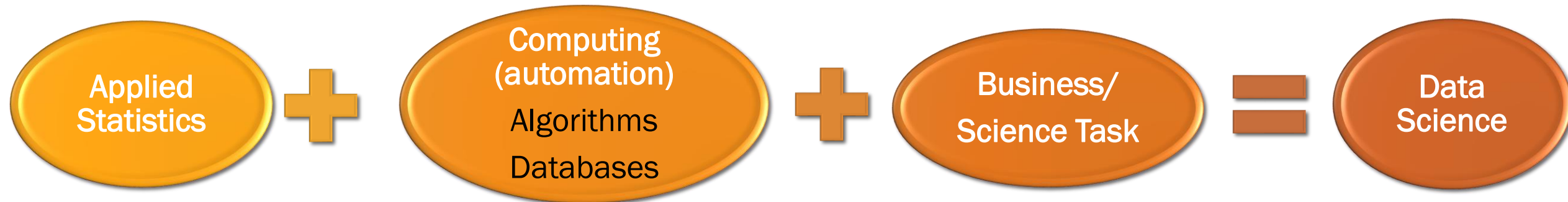
Peter Diggle (2015), JRSS (A), *Statistics: a data science for the 21st century*

## Statistics

- the science of [collecting, analyzing, presenting, and interpreting data](#) (Britannica)
- 

## Data Science

- is the science [of the analysis and interpretation](#) of large amounts of [data](#), esp customer data held by a company, in order to inform future practice (Collins)
- is an interdisciplinary field that uses [scientific methods](#), processes, algorithms and systems to [extract knowledge and insights from structured and unstructured data](#), and [apply knowledge and actionable insights from data](#) across a broad range of [application domains](#). (Wikipedia)
- uses techniques and theories drawn from [mathematics, statistics, computer science and information science](#) (Cambridge)



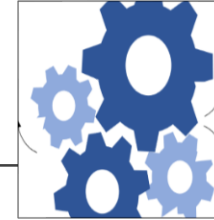
# Data Science SIG at a Glance *(formed Oct 2019)*

## Members



- Carsten Henneges (Syneos Health)
- Domingo Salazar (AZ)
- Federico Concas (GSK)
- Imran Hossain (Cytel)
- Jennifer Bradford (Phastar)
- Julia Chernova (Bayer)
- Milan Geybels (Novo Nordisk)
- Peter Krusche (Novartis)
- Vlad Anisimov (Amgen)

## Directions



### Communication

- Open conversations
- Building bridges
- Integration
- Sharing examples
- Best Practice

## Activities



- ✓ Presented at PSI 2021
- Blog Posts
- Webinar



*Survey*

# Survey Framework

## AIM

// Learn about current status of Data Science community: tools, involvement areas, challenges

## WHEN

// 5 MAR 2021 and 18 June 2021

## HOW

// Distributed within our own organisations, PSI distributed to EFSPi leaders, shared on LinkedIn

// 66 responses

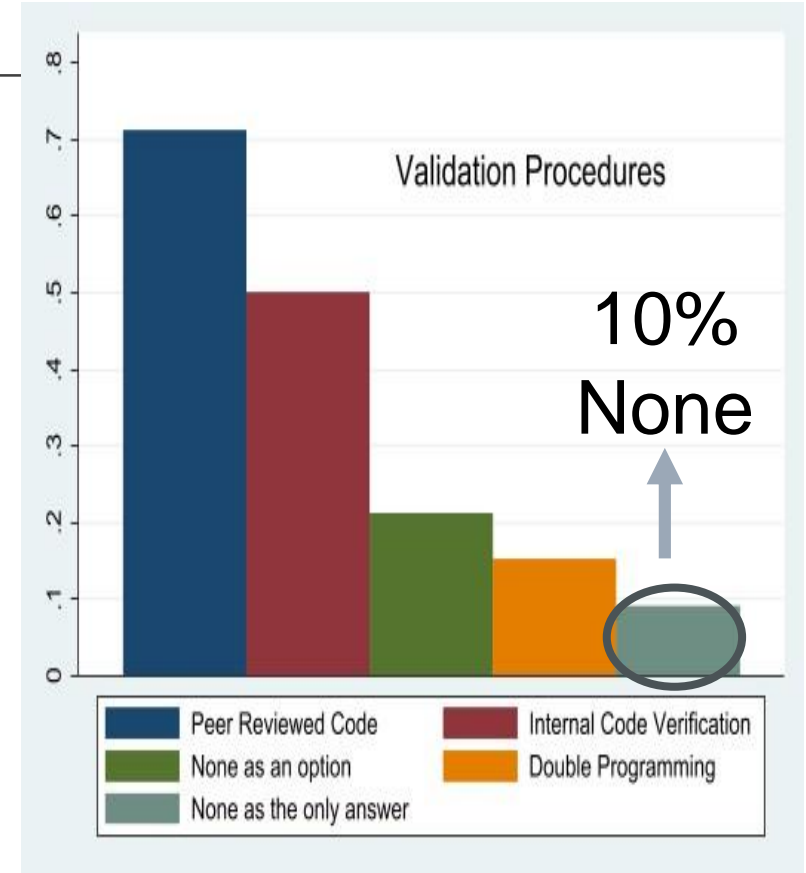
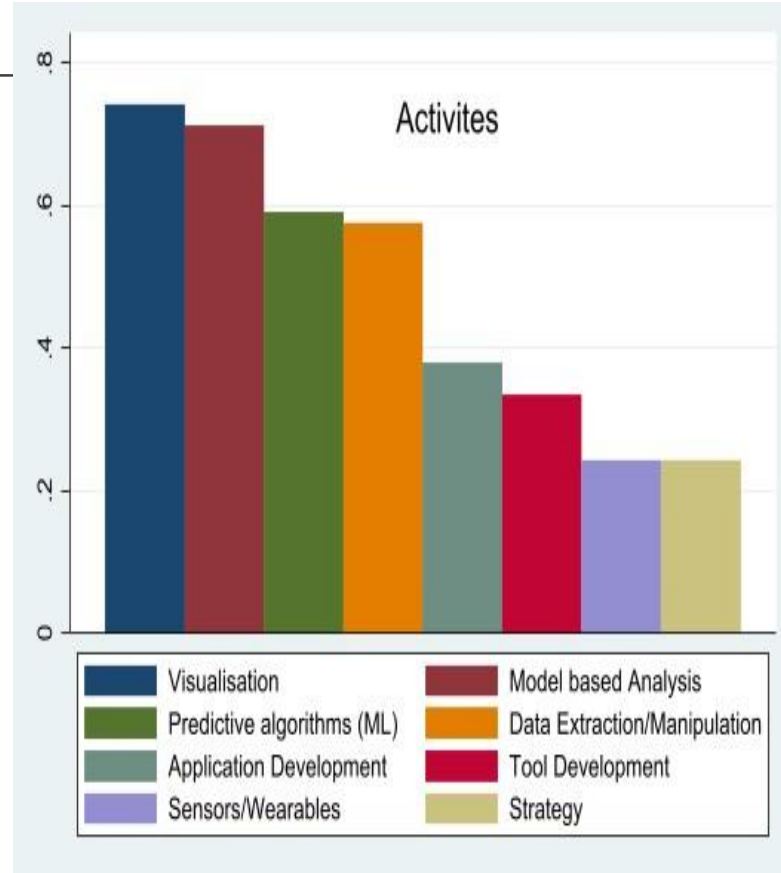
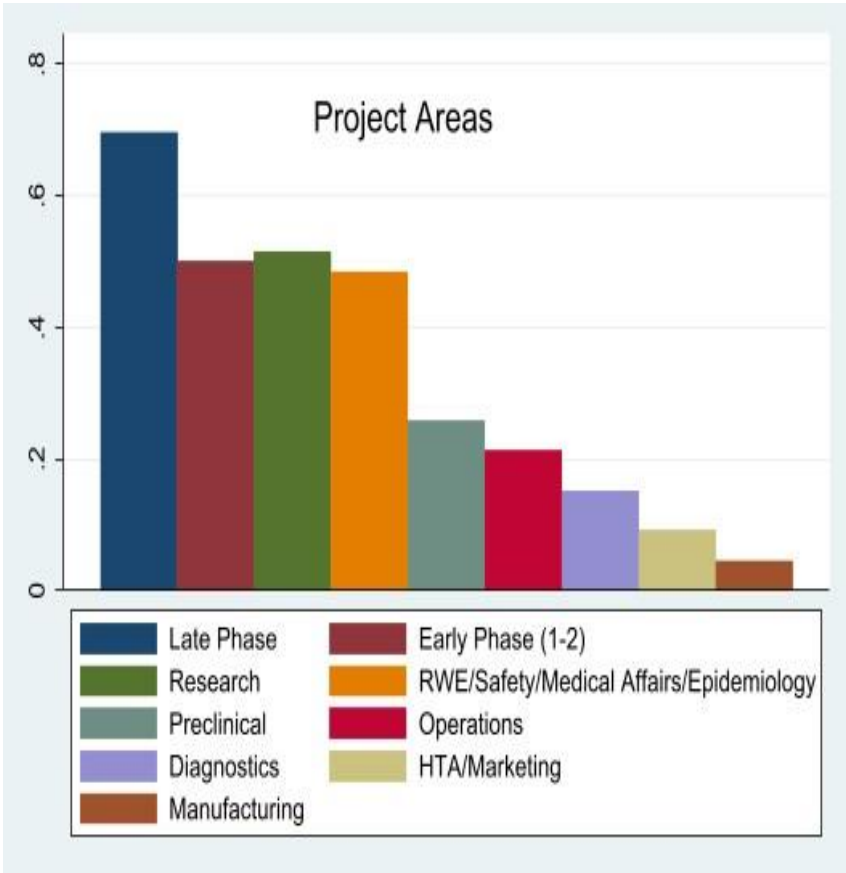
## WHO

- **REGION:** 77% Europe, 17% US, 6% ROW
- **EDUCATION:** 58% PhD, 38% MSc, 5% BSc
- **BACKGROUND:** 42% Stats, 9% Maths, 12% Maths+Stats, 10% Computer Science, 8% Bioinformatics, 8% Other, 11% Multiple Choices
- **AFFILIATION:** 90% Pharma, 6% CRO, 4% Biotech



# Projects

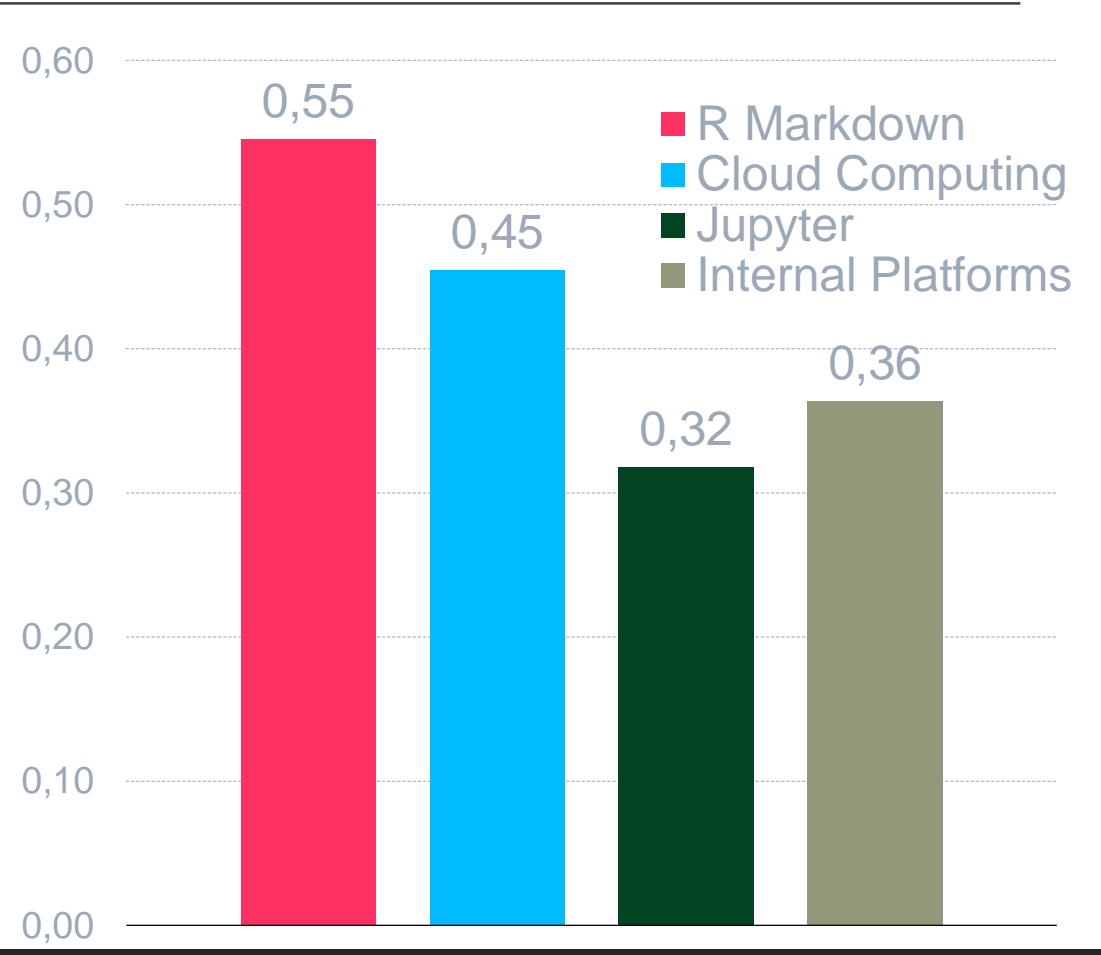
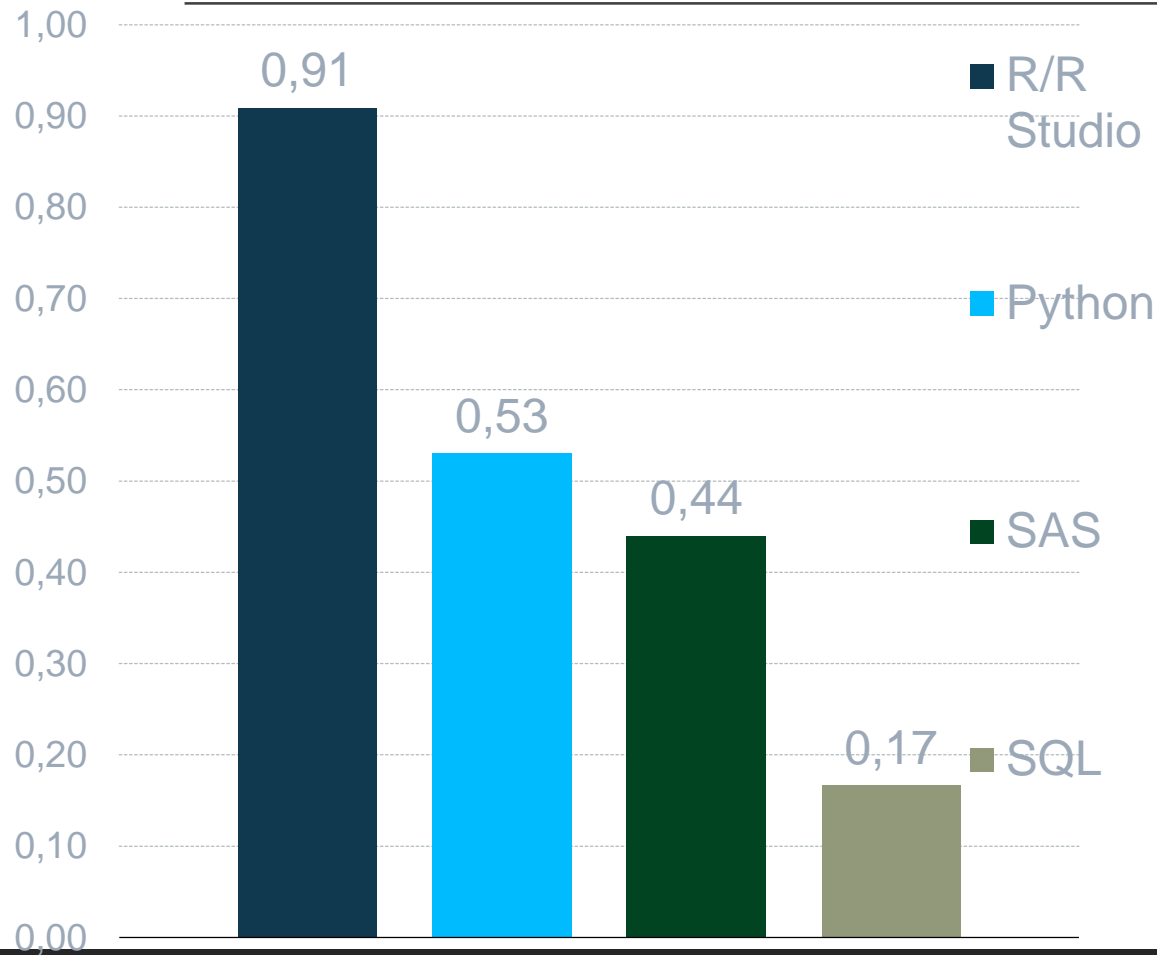
Multiple choices were possible







# Software and Tools



# *Challenges*



# Strategic

---

- No clear vision and strategy on organisational and team level
- Each function has its own vision/strategy
- People/skills resources
- Identification of the right tasks for the right data for the right models
- Data acquisition/sharing/access/infrastructure

86%

# Cross-functional/Operational

---

- Integration of colleagues with different backgrounds
- Building bridges between functions and knowledge exchange
- Establishing trust
- Teams' structure

# Collaborative

---

- Best practices not available
- Regulatory input limited
- Limited understanding of IP rights in this area
- Unrealistic expectations
- Limited External Collaborations
- Processes specification (survey: “too much” 22%; “little” 12%)
- Onboarding
- Manager’s support



*Opportunities*

---

*Thank you!*