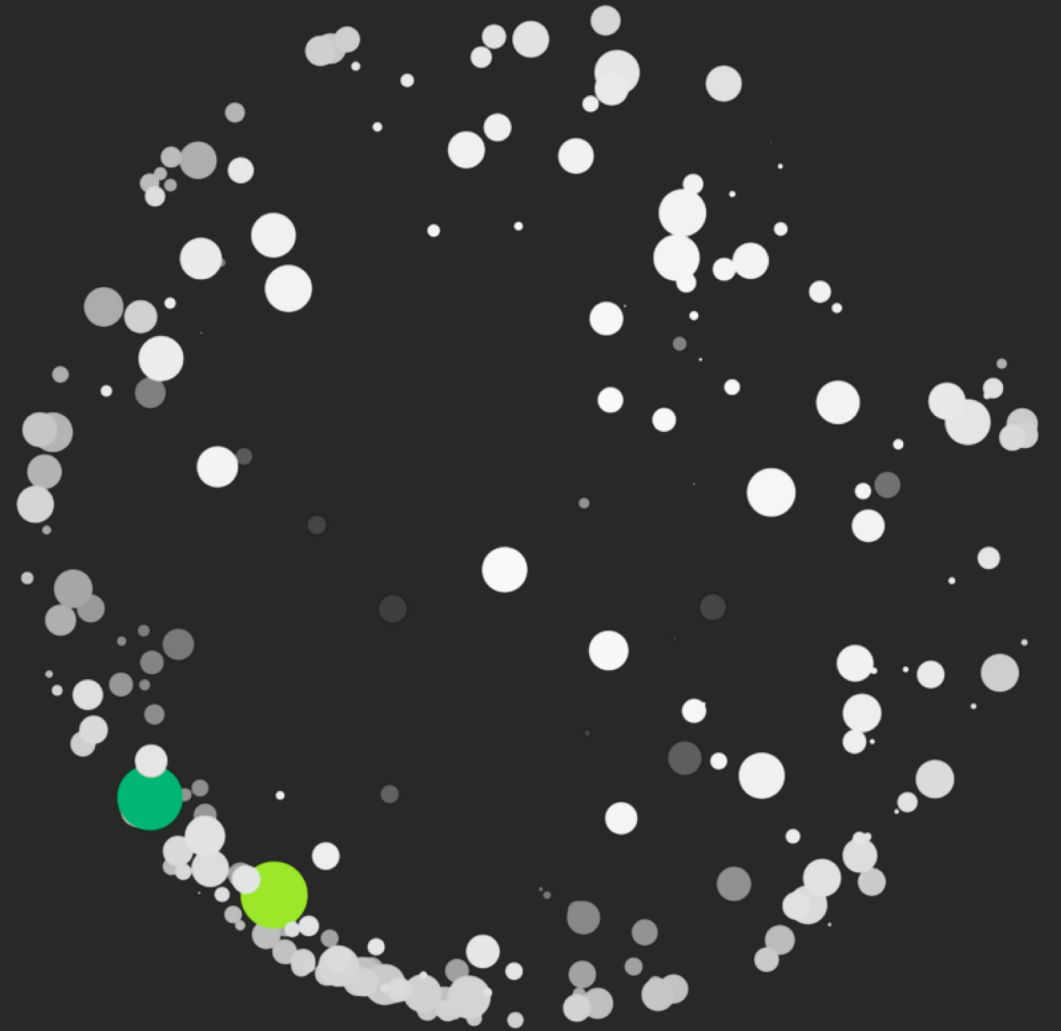


Grünenthal

Mine the GAP:

How the Role of the Data
Scientist Fills a Need in the
Pharmaceutical Industry



Michael S Rimler, FMD K&L Inc.
Jorine Putter, Grünenthal GmbH

WHAT IS DATA SCIENCE

“interdisciplinary field about scientific methods, processes, and systems to extract knowledge or insights from data in various forms, either structured or unstructured, similar to data mining.”

- Wikipedia

ACADEMIC TRAINING

3 primary threads of master's degrees in data science

Business

Using analytics in business and industry

How to understand, frame, and solve problems in marketing, operations, finance, management of information technology, human resources, and accounting

Computational

The practice of dealing with so-called “big data”

How to acquire, preprocess, store, manage, analyze, and visualize data arriving at high volume, velocity, and variety

Analytical

The quantitative methodologies of descriptive, predictive, and prescriptive analytics

How to select, build, solve, and analyze models using methodologies such as parametric and non-parametric statistics, regression, forecasting, data mining, machine learning, optimization, stochastics, and simulation

DATA SCIENCE ORGANIZATIONS & CONFERENCES

Strata

<https://conferences.oreilly.com/strata>

Predictive Analytics World (PAW)

<https://www.predictiveanalyticsworld.com>

KDD (Knowledge Discovery and Data Mining)

<http://www.kdd.org>

The Data Science Conference

<https://www.thedatascienceconference.com>

IEEE International Conference on Data Mining

<http://icdm2018.org>

Open Data Science Conference

<https://odsc.com>

THE INNOVATION ENTERPRISE

<https://theinnovationenterprise.com/summits/>

- ✓ London (March), [Philadelphia \(June\)](#), San Francisco (October), New York (December)
- ✓ Patient-Centricity: Utilizing data from a variety of different/disparate sources to personalize treatment to the patient
- ✓ Use of Real World Data/Evidence: Data to identify gaps in patient care, improving quality of care, and patient outcomes
- ✓ Real World Data Usage in Clinical Trials: What trials may look like in the future
- ✓ Data Opportunities for Commercial Analytics: Customer targeting/segmentation and customer engagement

WHAT IS THE DATA SCIENTIST

“conventionally understood to elicit insights from data, using highly technical methods and ultimately presenting them to the layperson in an understandable way”

- ✓ often excels in analytical fields
- ✓ must exhibit a high level of curiosity to identify and investigate problems
- ✓ must be creative in generating solutions using various sources and formats
- ✓ must be able to use sophisticated technical algorithms
- ✓ must be able to communicate non-technically

WHERE DOES DATA SCIENCE FIT IN THE PHARMACEUTICAL INDUSTRY

HISTORICALLY



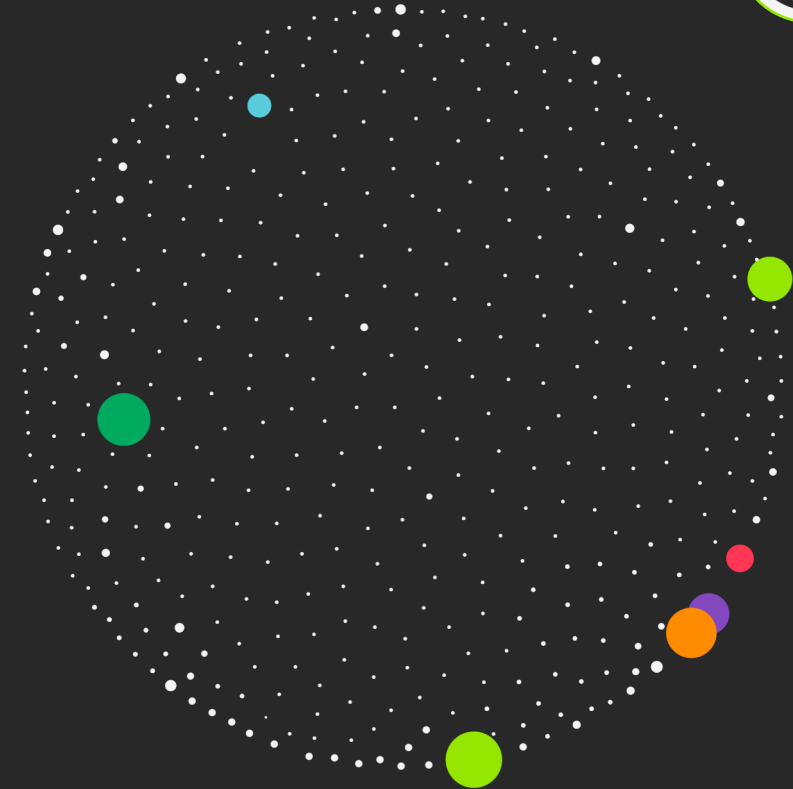
Design trials to answer limited set of questions



Provide informative or confirmatory results

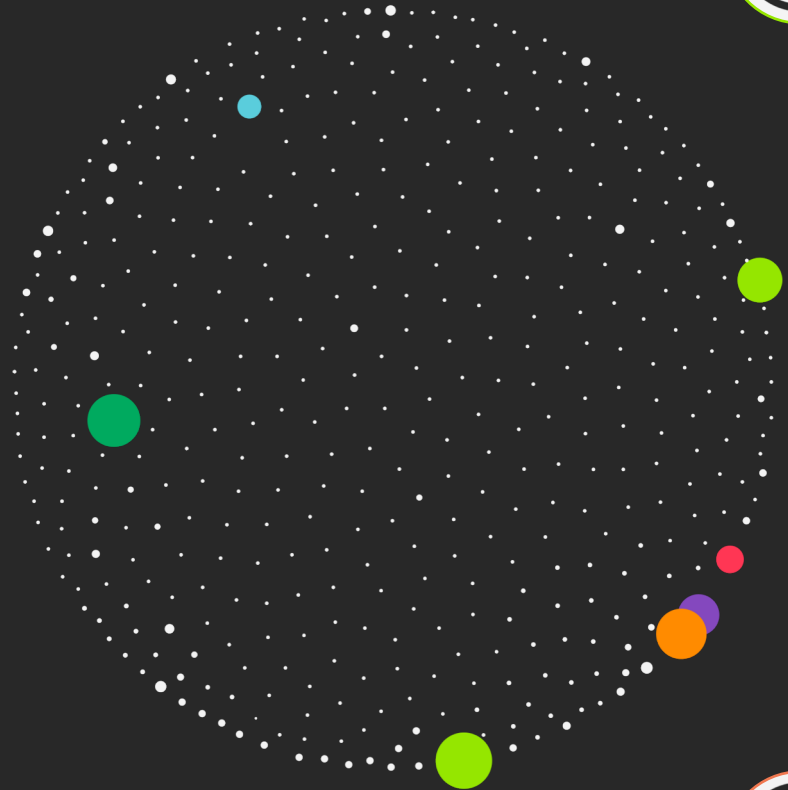


Mainly biostatisticians, programmers & pharmacometricians



WHERE DOES DATA SCIENCE FIT IN THE PHARMACEUTICAL INDUSTRY

CURRENT ERA



Increased regulatory, payer & patient pressure



Need to develop agile CDP



Need innovative adaptive methodologies



Explore RWD



Gain insights from disparate data sources

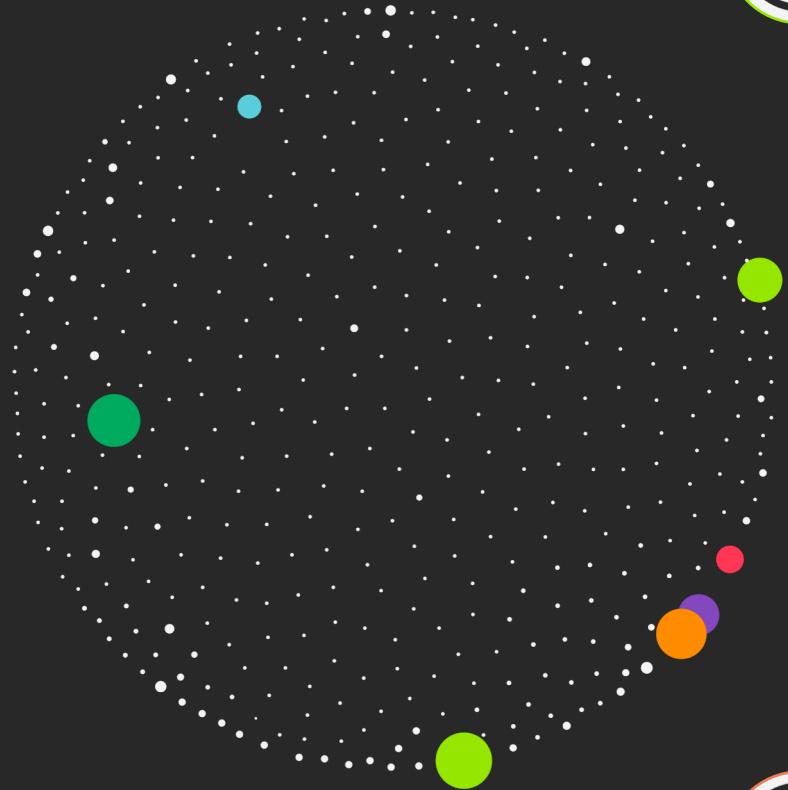


Need to be faster, cheaper & increase quality

**CHALLENGE: HOW CAN WE DO
THIS?**

ADDRESSING A PROBLEM WITH DATA SCIENCE

Bandyopadhyay's Framework



Frame the problem



Collect the required raw data



Process the data



Explore the data



Perform in-depth analysis

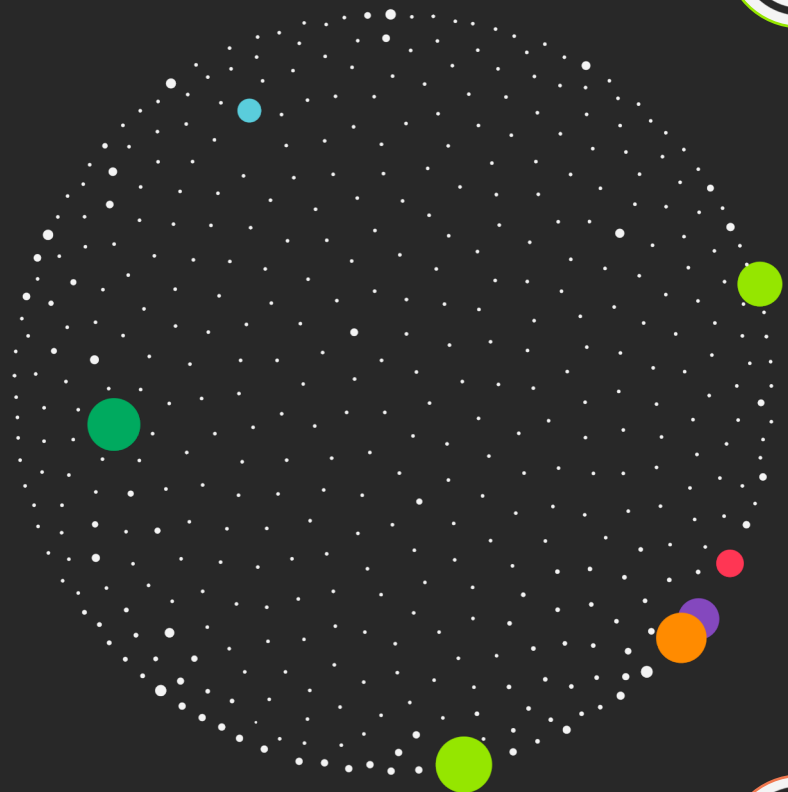


Communicate the results

APPLYING THE **BANDYOPADHYAY'S** FRAMEWORK

DETECTION OF EFFICACIOUS SUBPOPULATIONS

Bandyopadhyay's Framework



Does my compound show different results in different subpopulations?



Which individual trials possess the relevant attributes desired for exploration?



Prepare pooled datasets containing all information



Perform descriptive, inferential, and data mining analyses



Perform in-depth analysis



Communicate consistent differences among subpopulations

POTENTIAL PROJECTS

Adaptive Study
Designs

Data Mining
and
Clustering

Preclinical
Expression
Pattern Detection

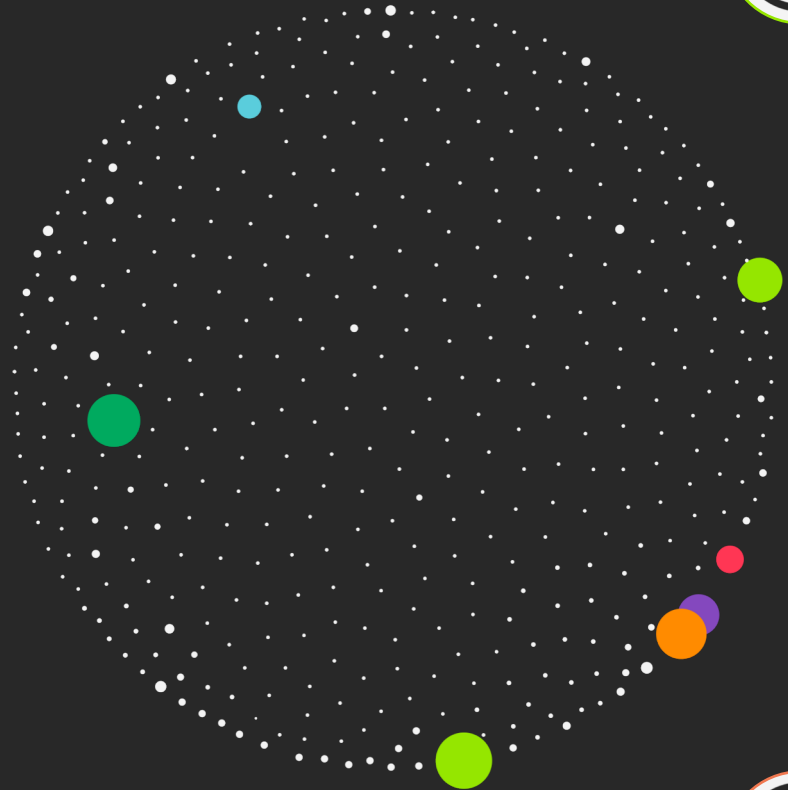
Model Selection
or Nonparametric
Methods

Safety Signal
Detection
Methods

Mining for New
Indications

ADDRESSING A PROBLEM WITH DATA SCIENCE

Bandyopadhyay's Framework



Frame the problem



Collect the required raw data



Process the data



Explore the data



Perform in-depth analysis



Communicate the results

