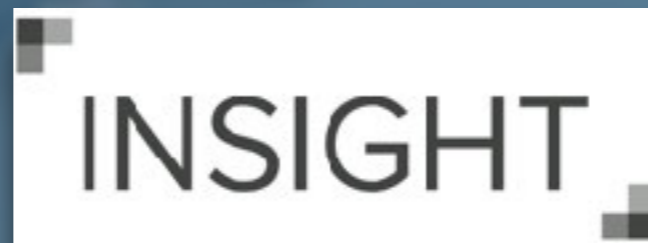


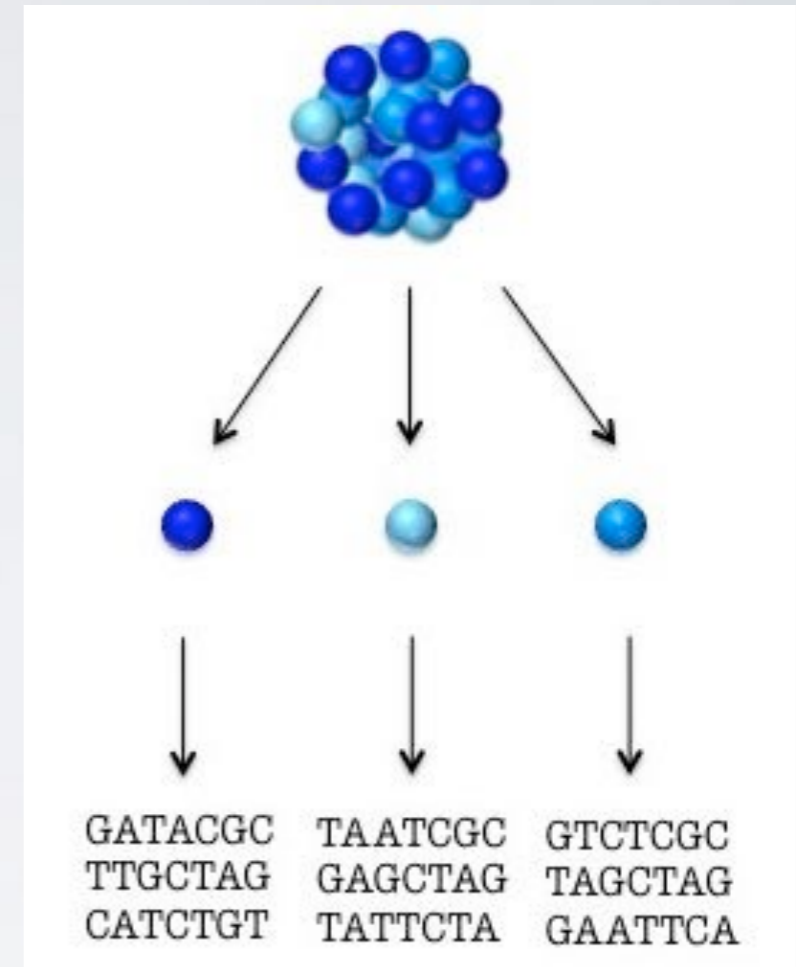
# Academia to Data Science

Faye Zheng  
Program Director  
Insight Data Science





# Northwestern University



PURDUE UNIVERSITY  
CENTER FOR CANCER RESEARCH





**Business  
Analytics**

**Genomics**

**Artificial  
Intelligence**

**Data  
Engineering**

**HEALTH**

Memorial Sloan Kettering  
Flatiron Health  
ZocDoc

**MEDIA**

The New York Times  
MTV, Netflix,  
Nielsen

**SOCIAL**

Facebook  
Twitter  
LinkedIn

**FINANCE**

Bloomberg  
Capital One Labs  
AmEx

**ENERGY**

Bright Power  
EnergyHub

**TRAVEL**

AirBnB  
TripAdvisor

**ADVERTISING**

Sailthru  
Tapad  
Dstillery

**APPAREL**

Rent the Runway  
StitchFix  
Bonobos

**GOVERNMENT/NON-PROFIT**

The City of NYC  
The Census Bureau  
DataKind

## WHITE HOUSE NAMES DJ PATIL AS THE FIRST US CHIEF DATA SCIENTIST



“The best data scientists tend to be hard scientists, rather than computer science majors.”

DJ Patil

# Hard scientists tend to already do the work of data scientists

Obtain → Clean → Explore → Model → Interpret

- Data intuition
- Know the right questions to ask
- Statistics knowledge
- Strong coding skills
- Communicate clearly\*

\*results may vary

# INSIGHT BY THE NUMBERS

The Alumni Split



PHYSICAL SCIENCES



BIOSCIENCES



ENGINEERING



MATHEMATICS



SOCIAL SCIENCES



EARTH SCIENCES  
AND ENERGY



COMPUTER  
SCIENCE



OTHER



# Job of an academic scientist

- Collect and clean data
- Use programming and statistical methods to discover meaningful patterns
- Convey results to scientific community

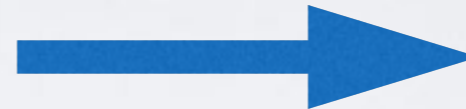
# Job of a scientist **in industry**

- Collect and clean data
- Use programming and statistical methods to discover meaningful patterns
- Convey results to the **team/company/stakeholders**
- **Make data-informed decisions that directly impact the product and ultimately the business**

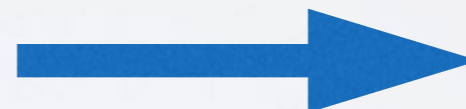
# The gap from academia to industry



**Tools**



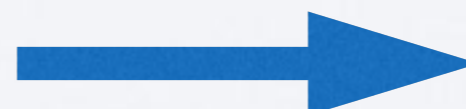
**Pace**



Results published  
*once a year*

Results delivered  
*weekly*

**Product**



Analysis for  
*scientific community*

Analysis for  
*stakeholders*

Smart scientists  
with the right skills

**INSIGHT**

High-impact  
data teams

**7 weeks**

**full-time**

**tuition-  
free**

# INSIGHT IN A NUTSHELL













# Insight Fellows are Data Scientists and Data Engineers at...



Silicon Valley • New York • Boston • Seattle  
+ many others...

# Growing network of active Insight alumni





**Adam Jenkins**

Data Scientist, Biogen  
Insight Fellow Winter 2015



**Jing Tong**

Bioinformatics Scientist, Foundation  
Medicine  
Insight Fellow Summer 2015



**Péter Kómár**

Scientist, Seven Bridges  
Insight Fellow Summer 2015



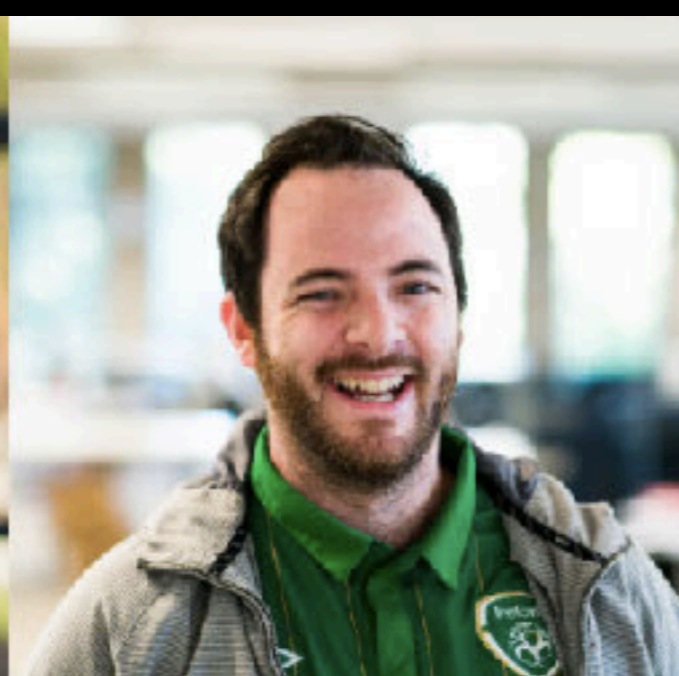
**Sara Gettel**

Data Engineer Associate, Athena Health  
Insight Fellow Winter 2015



**Jenelle Bray**

Head of Anti-Abuse Modeling,  
LinkedIn  
Insight Fellow Summer 2013



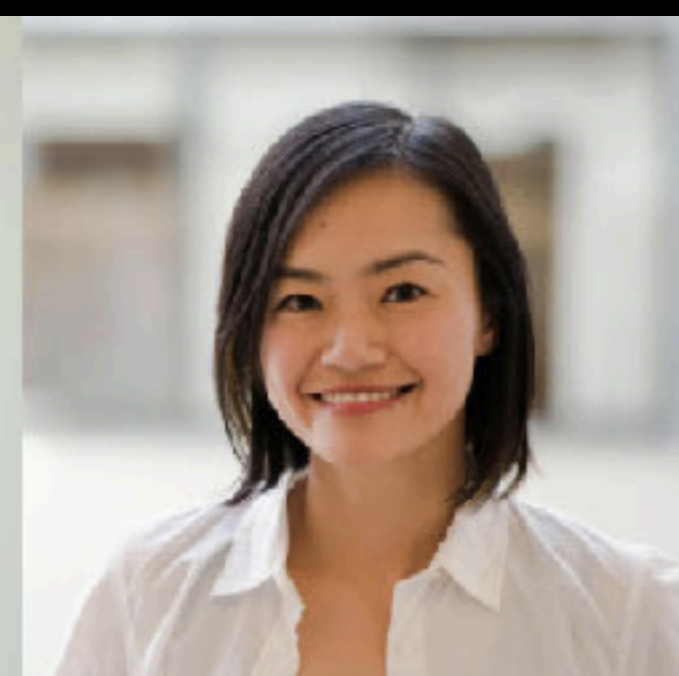
**Joe Gallagher**

Data Lead, reddit  
Insight Fellow Winter 2013



**Anne Bauer**

Data Scientist, The New York Times  
Insight Fellow Winter 2015



**Lisa Qian**

Data Science Manager, Airbnb  
Insight Fellow Summer 2013

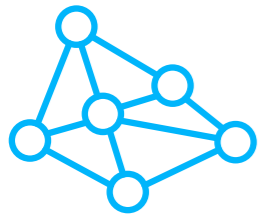
---

---

## PROGRAMS

---

---



**DATA SCIENCE**  
*All Locations*



**DATA ENGINEERING**  
*SV, NYC*



**HEALTH DATA**  
*SV, Boston*



**ARTIFICIAL INTELLIGENCE**  
*SV*

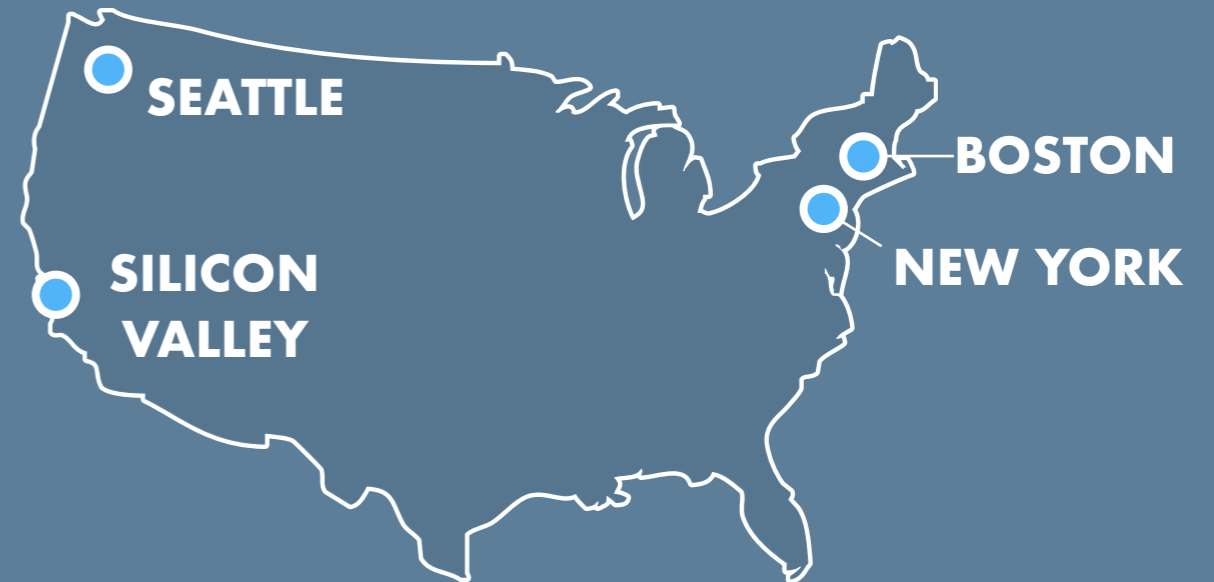
---

---

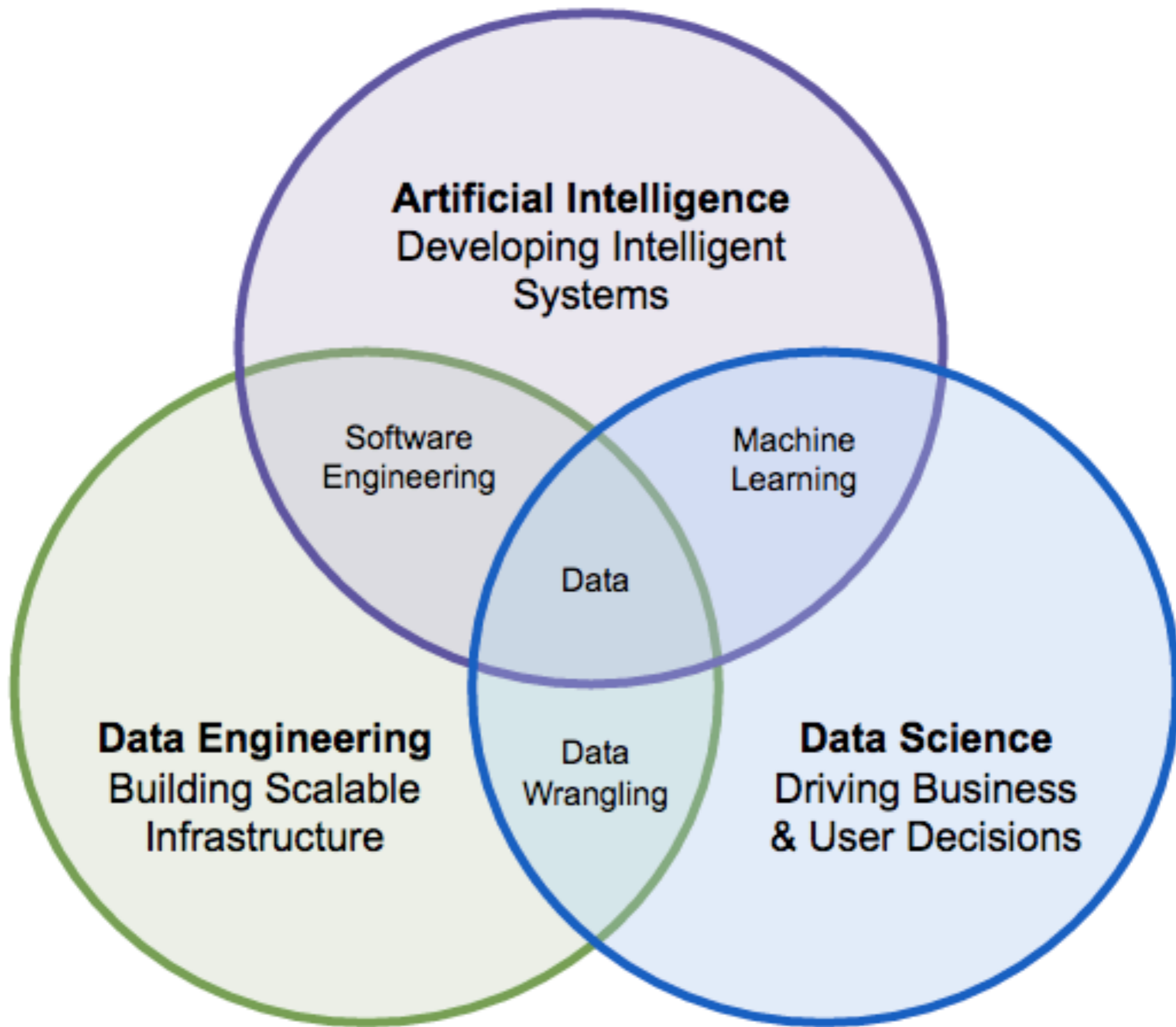
## LOCATIONS

---

---



**+ REMOTE**



# INSIGHT DATA ENGINEERING

**Locations:** *Silicon Valley, New York*

“Data engineers are specialized software engineers that enable others to answer questions on datasets, within latency constraints”

*- Nathan März, Inventor of Apache Spark*

- **Data engineers** build the infrastructure and pipelines for effectively **collecting, processing, and accessing data**
- *Requisite skills:* Fundamentals in computer science, databases, data structures, and algorithms (no educational requirement)

# INSIGHT ARTIFICIAL INTELLIGENCE

**Locations:** *Silicon Valley*

“AI is the new electricity. Just as 100 years ago electricity transformed industry after industry, AI will now do the same.”

*- Andrew Ng, Chief Scientist of Baidu*

- AI focuses on **understanding core human abilities** such as vision, speech, language, decision making, and other complex tasks, and **designing machines and software** to emulate these processes
- *Requisite skills:* Advanced machine learning and strong software engineering (no educational requirement)



# INSIGHT HEALTH DATA

**Locations:** *Silicon Valley, Boston*

“The number of challenges that one sees at the intersection of the life sciences and technology is quite vast.”

*- Anthony Philipakis, Chief Data Officer of the Broad Institute*

- Health data focuses on the application of data science and bioinformatics to the fields of **healthcare, genomics, biotech, and drug development.**
- PhD necessary, but domain knowledge not required

# Thinking about data science?

*For undergraduates / early grad students...*

- Start working with open source tools
  - e.g. Python, R, SQL - great time to start using GitHub
- Practice explaining your research to non-experts
- Start practicing statistical analysis and machine learning fundamentals
  - Kaggle is a great place to start!

# Thinking about data science?

## *For senior grad students...*

- Work on data-related side project beyond research
  - Time to move beyond Kaggle! Start by determining a question you may be interested in that can be answered by data.
- Gain experience gathering real, messy datasets from various sources
- Think about where you want to fit into the data analysis-engineering spectrum

# Thinking about data science?

## *For everyone...*

- Attend data science meet-ups, conferences, and hackathons
- Read newsletters and blogs to keep abreast of latest techniques and industry trends
- Follow thought leaders in the online data science community
- Talk to people and learn from them!

**Next Session Starts: September 5, 2017**  
**Application Deadline: June 26, 2017**

**Website** [insightdatascience.com](http://insightdatascience.com)

**Blog** [blog.insightdatascience.com](http://blog.insightdatascience.com)



@InsightDataSci



Insight Data Science



[info@insightdatascience.com](mailto:info@insightdatascience.com)

[faye@insightdatascience.com](mailto:faye@insightdatascience.com)