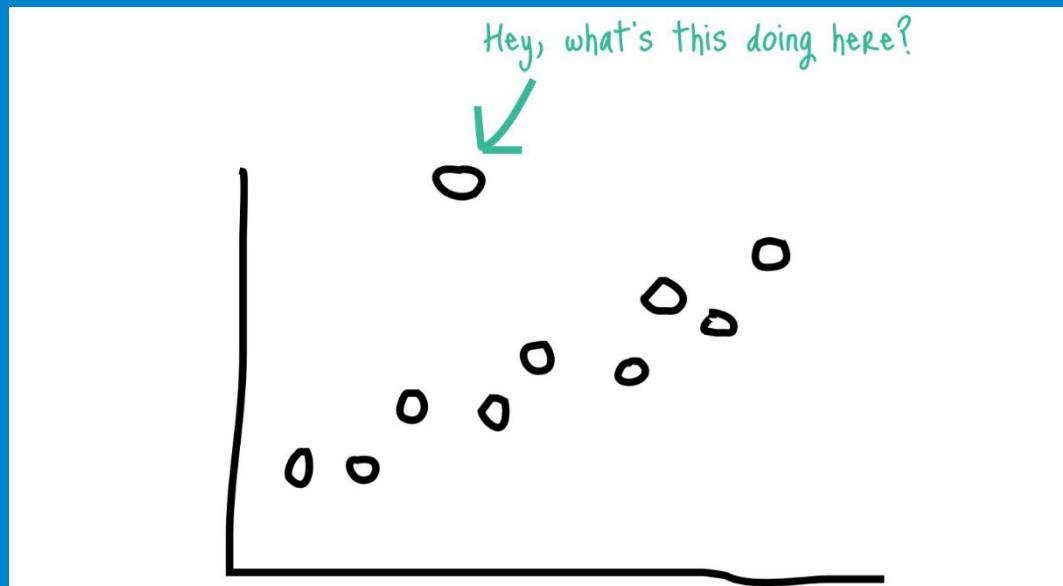


# Cómo ser un outlier en el mundo de los Data Scientists?



**Andrea Villanes**  
**@andreasrr**



2,583 trabajos  
relacionados a data science



Prior experience in statistical programming with tools such as [R](#), [SAS](#), [SPSS](#) etc.

Expertise in at least one scripting language, and proficiency in [SQL/Hive](#)

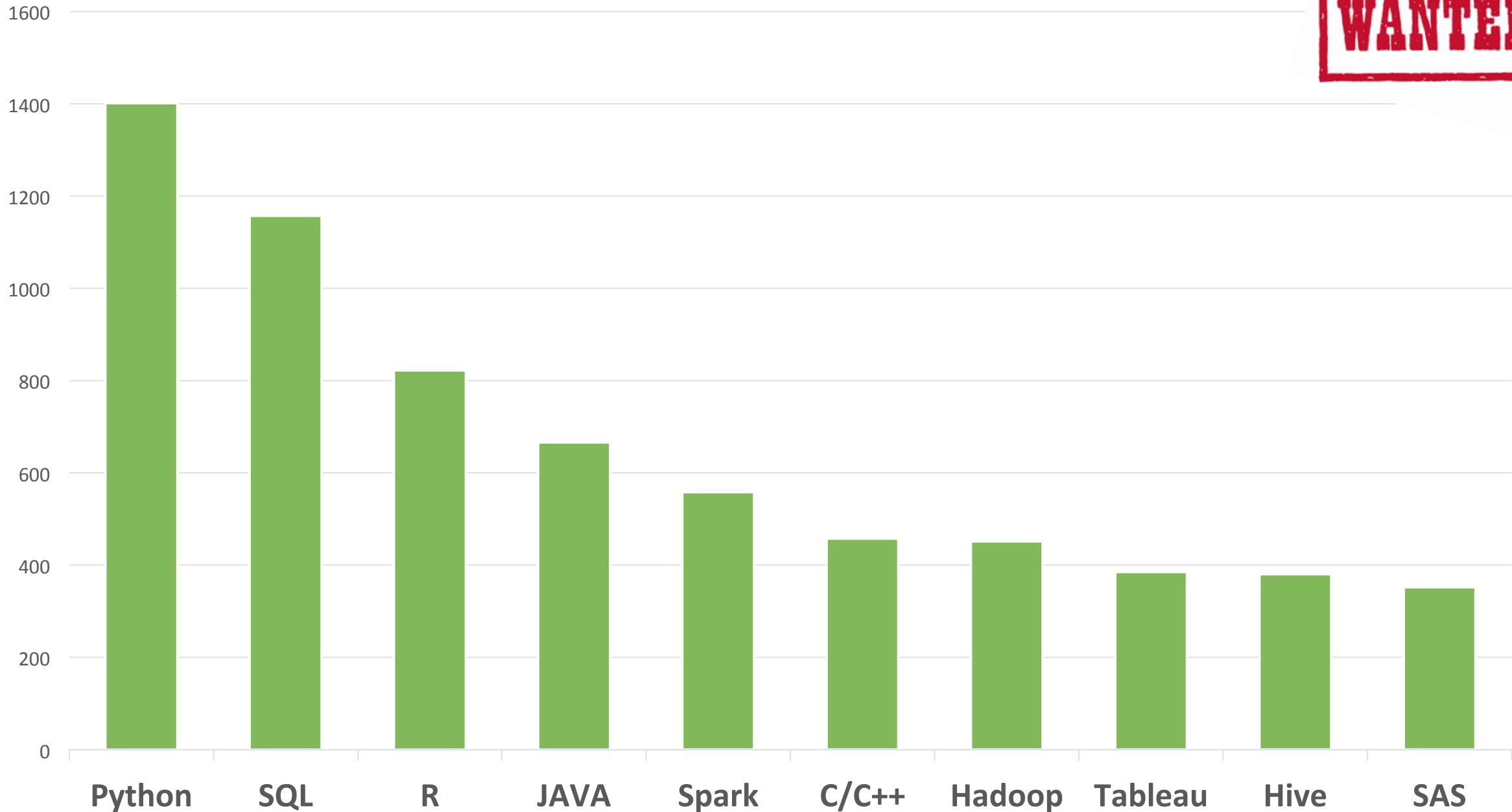
Experience with [Pandas](#), [numpy](#), [scipy](#), [scikit-learn](#)

Fluency in scripting languages ([R](#) or [Python](#)).

Experience with relational databases / [SQL](#).

Experience with [Spark](#), [TensorFlow](#) and [Keras](#)

Experience with A/B Testing, Cohort Analysis, and Distributed Systems ([Spark](#), [Hive](#), [Pig](#), [Hadoop](#), or [Map-Reduce](#)).





**Collaborate** with Data Scientist and Software Developer colleagues to research, debate, experiment and devise highly effective solutions

**Communicate** insights to company management with high-quality presentations

Think and work independently, demonstrate creativity and deliver quick results, yet demonstrate a **team-player attitude**

**Lead** the company in best practices for data and analytics.

**team**

**collaborate**

**multidisciplinary  
consultant**

lead  
passionate  
**support**  
influence      help      plan  
design

**Lo que quieren las empresas  
como perfil del data scientist...**

## MATH & STATISTICS

- ★ Machine learning
- ★ Statistical modeling
- ★ Experiment design
- ★ Bayesian inference
- ★ Supervised learning: decision trees, random forests, logistic regression
- ★ Unsupervised learning: clustering, dimensionality reduction
- ★ Optimization: gradient descent and variants

## DOMAIN KNOWLEDGE & SOFT SKILLS

- ★ Passionate about the business
- ★ Curious about data
- ★ Influence without authority
- ★ Hacker mindset
- ★ Problem solver
- ★ Strategic, proactive, creative, innovative and collaborative



## PROGRAMMING & DATABASE

- ★ Computer science fundamentals
- ★ Scripting language e.g. Python
- ★ Statistical computing package e.g. R
- ★ Databases SQL and NoSQL
- ★ Relational algebra
- ★ Parallel databases and parallel query processing
- ★ MapReduce concepts
- ★ Hadoop and Hive/Pig
- ★ Custom reducers
- ★ Experience with xaaS like AWS

## COMMUNICATION & VISUALIZATION

- ★ Able to engage with senior management
- ★ Story telling skills
- ★ Translate data-driven insights into decisions and actions
- ★ Visual art design
- ★ R packages like ggplot or lattice
- ★ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

**Lo que enseñan las  
universidades...**

## MATH & STATISTICS

- ★ Machine learning
- ★ Statistical modeling
- ★ Experiment design
- ★ Bayesian inference
- ★ Supervised learning: decision trees, random forests, logistic regression
- ★ Unsupervised learning: clustering, dimensionality reduction
- ★ Optimization: gradient descent and variants

## DOMAIN KNOWLEDGE & SOFT SKILLS

- ★ Passionate about the business
- ★ Curious about data
- ★ Influence without authority
- ★ Hacker mindset
- ★ Problem solver
- ★ Strategic, proactive, creative, innovative and collaborative



## PROGRAMMING & DATABASE

- ★ Computer science fundamentals
- ★ Scripting language e.g. Python
- ★ Statistical computing package e.g. R
- ★ Databases SQL and NoSQL
- ★ Relational algebra
- ★ Parallel databases and parallel query processing
- ★ MapReduce concepts
- ★ Hadoop and Hive/Pig
- ★ Custom reducers
- ★ Experience with xaaS like AWS

80%

## COMMUNICATION & VISUALIZATION

- ★ Able to engage with senior management
- ★ Story telling skills
- ★ Translate data-driven insights into decisions and actions
- ★ Visual art design
- ★ R packages like ggplot or lattice
- ★ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

20%

**Cómo aprender los skills que  
nos hacen falta?**

**Matemática y Estadística**

**Programación y Base de Datos**

**Soft Skills**

**Comunicación y Visualización**

## **Matemática y Estadística**

- Fundamentos de estadística
- Educación formal (si es posible)
- Diplomado, maestría
- Cursos en línea (MOOC)
- edX, Udemy, Udacity, Coursera:  
Foundations of Data Analysis,  
Inferential Statistics

## **Programación y Base de Datos**

**Soft Skills**

**Comunicación y Visualización**

<h3>Matemática y Estadística</h3> <ul style="list-style-type: none"><li>• Fundamentos de estadística</li><li>• Educación formal (si es posible)</li><li>• Diplomado, maestría</li><li>• Cursos en línea (MOOC)</li><li>• edX, Udemy, Udacity, Coursera: Foundations of Data Analysis, Inferential Statistics</li></ul>	<h3>Programación y Base de Datos</h3> <ul style="list-style-type: none"><li>• Python, SQL, R, SAS</li><li>• DataCamp: Learn R, Python &amp; Data Science Online</li><li>• Concentrarse en que herramientas piden la industria deseada</li><li>• Practicar, practicar: Proyectos!</li></ul>
<h3>Soft Skills</h3>	<h3>Comunicación y Visualización</h3>

<p><b>Matemática y Estadística</b></p> <ul style="list-style-type: none"><li>• Fundamentos de estadística</li><li>• Educación formal (si es posible)</li><li>• Diplomado, maestría</li><li>• Cursos en línea (MOOC)</li><li>• edX, Udemy, Udacity, Coursera: Foundations of Data Analysis, Inferential Statistics</li></ul>	<p><b>Programación y Base de Datos</b></p> <ul style="list-style-type: none"><li>• Python, SQL, R, SAS</li><li>• DataCamp: Learn R, Python &amp; Data Science Online</li><li>• Concentrarse en qué herramientas piden la industria deseada</li><li>• Practicar, practicar: Proyectos!</li></ul>
<p><b>Soft Skills</b></p> <ul style="list-style-type: none"><li>• Pedir feedback para saber cuáles son tus debilidades</li><li>• No sólo a tus amigos, si no también a tu jefe, colegas, incluso en casa</li><li>• Ser proactivo</li></ul>	<p><b>Comunicación y Visualización</b></p>

<h3>Matemática y Estadística</h3> <ul style="list-style-type: none"> <li>• Fundamentos de estadística</li> <li>• Educación formal (si es posible)</li> <li>• Diplomado, maestría</li> <li>• Cursos en línea (MOOC)</li> <li>• edX, Udemy, Udacity, Coursera: Foundations of Data Analysis, Inferential Statistics</li> </ul>	<h3>Programación y Base de Datos</h3> <ul style="list-style-type: none"> <li>• Python, SQL, R, SAS</li> <li>• DataCamp: Learn R, Python &amp; Data Science Online</li> <li>• Concentrarse en que herramientas piden la industria deseada</li> <li>• Practicar, practicar: Proyectos!</li> </ul>
<h3>Soft Skills</h3> <ul style="list-style-type: none"> <li>• Pedir feedback para saber cuales son tus debilidades</li> <li>• No sólo a tus amigos, si no también a tu jefe, colegas, incluso en casa</li> <li>• Ser proactivo</li> </ul>	<h3>Comunicación y Visualización</h3> <ul style="list-style-type: none"> <li>• Conoce a tu audiencia!</li> <li>• Practicar, practicar, practicar</li> <li>• Pedir feedback antes y después de presentar</li> <li>• Udemy: Tableau for Beginners - Get Certified Accelerate Your Career</li> </ul>

## Matemática y Estadística

- Fundamentos de estadística
- Educación formal (si es posible)
- Diplomado, maestría
- Cursos en línea (MOOC)
- edX, Udemy, Udacity, Coursera:  
Foundations of Data Analysis,  
Inferential Statistics

## Programación y Base de Datos

- Python, SQL, R, SAS
- DataCamp: Learn R, Python & Data Science Online
- Concentrarse en que herramientas piden la industria deseada  
Practicar, practicar: Proyectos!



## Soft Skills

- Pedir feedback para saber cuales son tus debilidades
- No sólo a tus amigos, si no también a tu jefe, colegas, incluso en casa
- Ser proactivo

## Comunicación y Visualización

- Conoce a tu audiencia!
- Practicar, practicar, practicar
- Pedir feedback antes y después de presentar
- Udemy: Tableau for Beginners - Get Certified Accelerate Your Career

# MASTER OF SCIENCE IN ANALYTICS

INTENSIVE | INTERACTIVE | IMMERSIVE

[analytics.ncsu.edu](http://analytics.ncsu.edu)



# Master of Science in Analytics

- Founded in 2007 as the nation's first full-fledged professional master's degree that focuses on the tools, methods and applications of advanced analytics
- Integrated and dynamic curriculum: a unique blend of statistics, applied mathematics, computer science and business disciplines that continually evolves to meet industry needs



Housed within NC State's  
Institute for Advanced Analytics



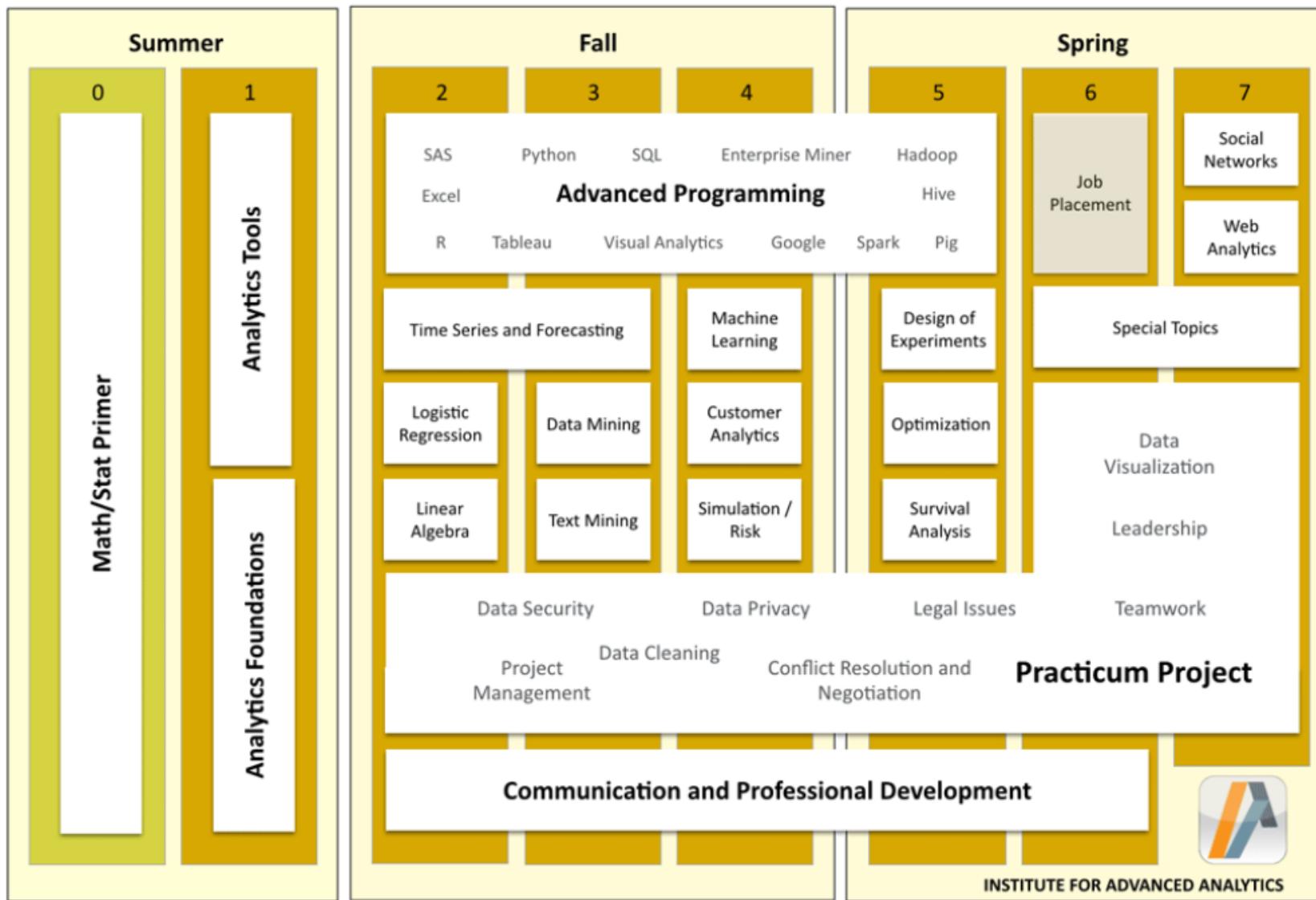
# Diseñado para ser...

- **Intensive** - accelerating the learning curve by condensing two years of graduate study into 10 months
- **Interactive** - leveraging peer-to-peer learning through teamwork throughout the entire curriculum
- **Immersive** - learning by doing, engaging in practical problems from the real world

# Estructura del programa

- 10 months of full-time residential study (Monday - Friday / 9a-5p)
- Fully integrated curriculum with its own faculty
- Limited to 120 students split into two 60-student sections
- Taught as a cohort, with all classes restricted to students in the program
- A professional degree aimed at placing its students in employment upon graduation
- Currently priced at ~\$26,500 for NC residents and ~\$46,500 for non-residents (these figures exclude living expenses)

## MASTER OF SCIENCE IN ANALYTICS 2017–18



# Aprendizaje vivencial



- 8-month practicum project replaces traditional thesis
- Teams of 4-5 students
- Real world gigabyte- to terabyte-scale data problems provided by industry sponsor (under confidentiality)
- Culminates in final report and presentation to sponsor
- Projects selected on the basis of educational value

# 2017-18 Practicum Sponsors

Elevate



visionist<sup>INC</sup>



DCLi  
DIRECT CHASSIS LINK INC.



PROMETRIC



Chick-fil-A®

asurion ▶

Putnam

MFA H

The Museum of Fine Arts, Houston



HOME DEPOT®

M&T Bank



Museum of  
Fine Arts Boston



+R  
Revenue  
Management  
Solutions  
truth in numbers + trust in people



eBay practicum Team:  
Preston, Dough, Nicole, and Pranav, MSA'17  
with alumni Steven and Ryan (back row).



Disney practicum team on-site in  
Orlando: John, Nick, Jennifer,  
Sydney and Alex, Class of 2018

B2W practicum team in Rio de Janeiro, Brazil:  
Holden, Ryan, Catherine, Cecilia, and Kosta,  
Class of 2017



# Resultados



- For 11 consecutive years more than 90% of the candidates have received one or more employment offers by graduation
- This year, 96% had offers of employment and 95% were employed at graduation
- MSA Class of 2018 candidates averaged 12 initial interviews and 3 job offers

# Resultados

Class of 2018 average base starting salary: \$98,200

- Candidates with 3+ years of professional experience prior to entry: \$106,000+
- Candidates with little or no work experience prior to entry: \$92,000
- Base salary range: \$72,500 to \$145,000
- Median signing bonus: \$10,000
- Estimated ROI payback period: 20 months



# The MSA Approach



Students at the Institute receive:

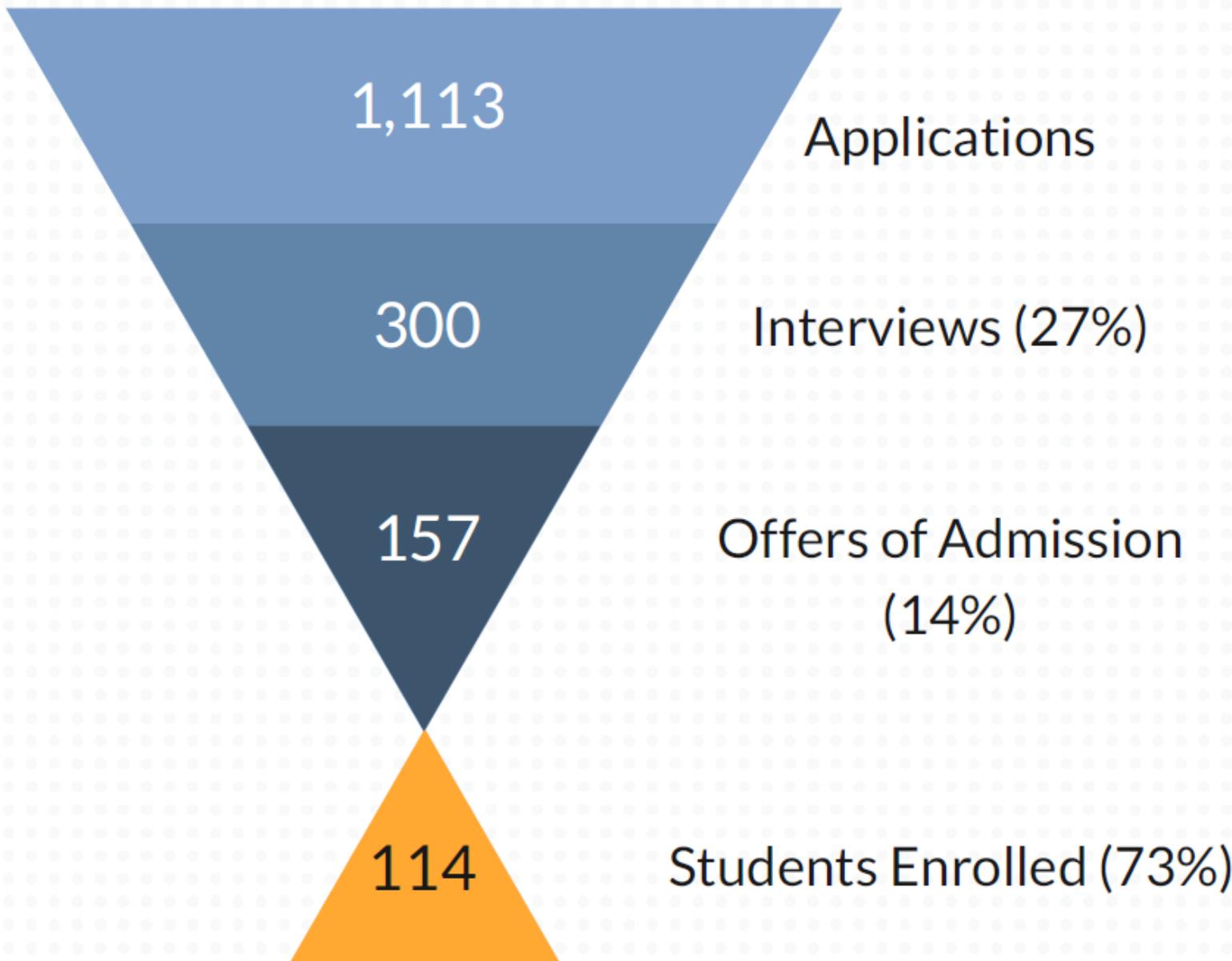
- Dedicated Career Services
- In-House Communications Training
- Regular Peer Feedback
- Professional Development Coaching
- Rigorous curriculum utilizing the latest tools and techniques from industry
- MSA Teaching Faculty
- Practicum Support
- In-House Technical Support

# Qué buscamos?



- Strong foundation (~2 college-level courses) in statistical methods, including regression analysis
- Coding experience preferred/encouraged
- Propensity for teamwork
- Effective communication skills
- Leadership potential (maturity, self-initiative, etc.)
- Career aspirations align with the program's mission

# Admissions Stats: Class of 2019



Gracias!

[andrea.villanes@gmail.com](mailto:andrea.villanes@gmail.com)

[@andreagrr](https://twitter.com/andreagrr)

[www.andreavillanes.com](http://www.andreavillanes.com)

[analytics.ncsu.edu](http://analytics.ncsu.edu)