

# Data Science

The Intersection of Mathematics  
and Programming

Joseph Ferrara

[mathcodelife.com](http://mathcodelife.com)

February 27, 2018

# Mathematics

- Develop **critical thinking** and **problem solving** skills.
- Refine your ability to **reason** and utilize **logical arguments**.
- Hone your sense of **numerical thinking**.
- Learn to write technical documentation.

# Mathematics

- Develop **critical thinking** and **problem solving** skills.
- Refine your ability to **reason** and utilize **logical arguments**.
- Hone your sense of **numerical thinking**.
- Learn to write technical documentation.

# Programming

- Requires **critical thinking** and **problem solving** skills.
- Requires **reason** and **logical thinking**.
- A certain extent of **numerical thinking** is required.
- Need to write pseudocode to explain concepts to non-programmers.

# Programming

- Requires **critical thinking** and **problem solving** skills.
- Requires **reason** and **logical thinking**.
- A certain extent of **numerical thinking** is required.
- Need to write pseudocode to explain concepts to non-programmers.

# Data Analyst

- Familiarity with data warehousing and business intelligence concepts.
- Knowledge of **SQL** (ex: MySql) or **NoSQL** (ex: MongoDB).
- Understanding of **ETL: Extract, Transform, and Load**.
- Ability to produce reports that explain data to business leaders.
- Have working knowledge of simple **statistics**.

# Data Analyst

- Familiarity with data warehousing and business intelligence concepts.
- Knowledge of **SQL** (ex: MySQL) or **NoSQL** (ex: MongoDB).
- Understanding of **ETL**: **E**xtract, **T**ransform, and **L**oad.
- Ability to produce reports that explain data to business leaders.
- Have working knowledge of simple **statistics**.

# Data Analyst

- Familiarity with data warehousing and business intelligence concepts.
- Knowledge of **SQL** (ex: MySql) or **NoSQL** (ex: MongoDB).
- Understanding of **ETL: Extract, Transform, and Load**.
- Ability to produce reports that explain data to business leaders.
- Have working knowledge of simple **statistics**.

# Data Analyst

- Familiarity with data warehousing and business intelligence concepts.
- Knowledge of **SQL** (ex: MySql) or **NoSQL** (ex: MongoDB).
- Understanding of **ETL: Extract, Transform, and Load**.
- Ability to produce reports that explain data to business leaders.
- Have working knowledge of simple **statistics**.

# Data Analyst

- Familiarity with data warehousing and business intelligence concepts.
- Knowledge of **SQL** (ex: MySql) or **NoSQL** (ex: MongoDB).
- Understanding of **ETL: Extract, Transform, and Load**.
- Ability to produce reports that explain data to business leaders.
- Have working knowledge of simple **statistics**.

# Data Scientist

- Needs to have a full understanding of a **Data Analyst** role.
- Advanced **mathematics** and **statistics** is needed to understand more complicated algorithms.
- Extensive knowledge of a **data science friendly programming language** is a must. (ex: R, Python, or SAS)
- Ability to translate **business problems** into automated systems via programming.
- Must be a **generalist** and **specialist** at the same time.

# Data Scientist

- Needs to have a full understanding of a Data Analyst role.
- Advanced **mathematics** and **statistics** is needed to understand more complicated algorithms.
- Extensive knowledge of a **data science friendly programming language** is a must. (ex: R, Python, or SAS)
- Ability to translate **business problems** into automated systems via programming.
- Must be a **generalist** and **specialist** at the same time.

# Data Scientist

- Needs to have a full understanding of a Data Analyst role.
- Advanced **mathematics** and **statistics** is needed to understand more complicated algorithms.
- Extensive knowledge of a **data science friendly programming language** is a must. (ex: R, Python, or SAS)
- Ability to translate **business problems** into automated systems via programming.
- Must be a **generalist** and **specialist** at the same time.

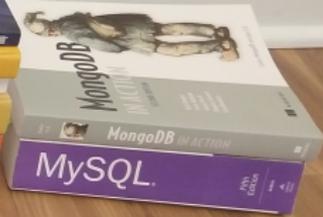
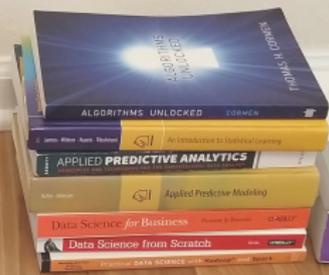
# Data Scientist

- Needs to have a full understanding of a Data Analyst role.
- Advanced **mathematics** and **statistics** is needed to understand more complicated algorithms.
- Extensive knowledge of a **data science friendly programming language** is a must. (ex: R, Python, or SAS)
- Ability to translate **business problems** into automated systems via programming.
- Must be a **generalist** and **specialist** at the same time.

# Data Scientist

- Needs to have a full understanding of a Data Analyst role.
- Advanced **mathematics** and **statistics** is needed to understand more complicated algorithms.
- Extensive knowledge of a **data science friendly programming language** is a must. (ex: R, Python, or SAS)
- Ability to translate **business problems** into automated systems via programming.
- Must be a **generalist** and **specialist** at the same time.

**THAT'S A LOT!  
HOW DO I GET  
STARTED?**



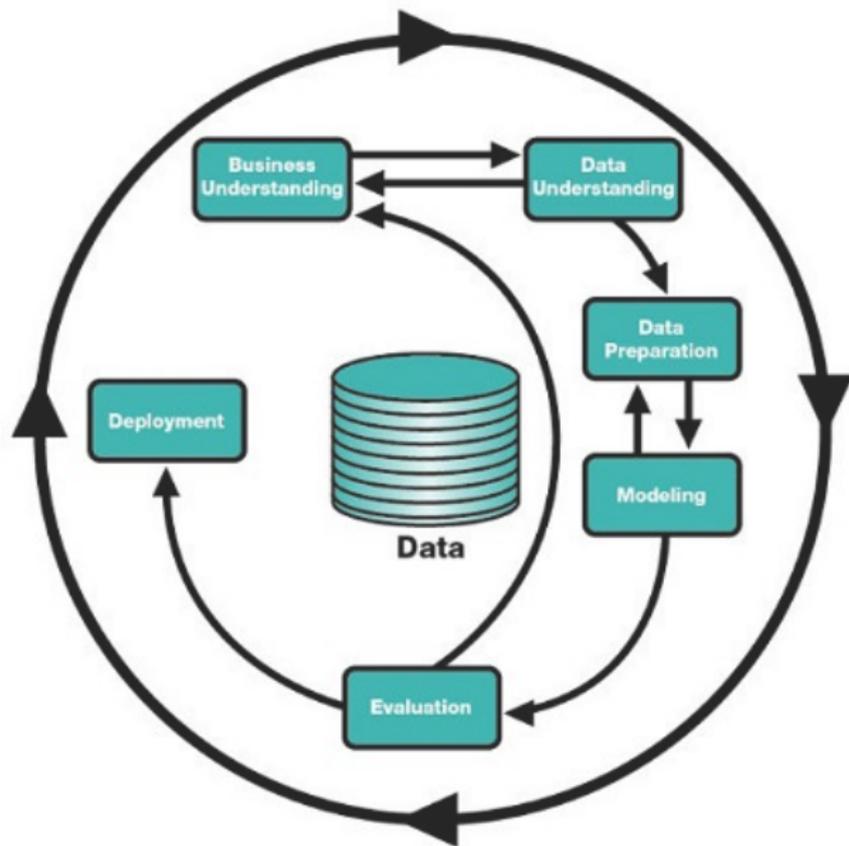
Books!

# The Churn Problem

- The business leaders at your company want to reduce their customer churn rate, and they want your help doing it.
- You apply the **CRISP-DM** method (**Cross-industry standard process for data mining**).

# The Churn Problem

- The business leaders at your company want to reduce their customer churn rate, and they want your help doing it.
- You apply the **CRISP-DM** method (**C**ross-**i**ndustry **s**tandard **p**rocess for data mining).



# The Churn Problem

- Gather and transform the appropriate data.
- Try different models on the data and compare their performance.
- Validate your model results (for example use cross-validation).
- Walk through your newly created **black box** with your business leaders.
- Deploy the model.

# The Churn Problem

- Gather and transform the appropriate data.
- Try different models on the data and compare their performance.
- Validate your model results (for example use cross-validation).
- Walk through your newly created **black box** with your business leaders.
- Deploy the model.

# The Churn Problem

- Gather and transform the appropriate data.
- Try different models on the data and compare their performance.
- **Validate your model results (for example use cross-validation).**
- Walk through your newly created **black box** with your business leaders.
- Deploy the model.

# The Churn Problem

- Gather and transform the appropriate data.
- Try different models on the data and compare their performance.
- Validate your model results (for example use cross-validation).
- Walk through your newly created **black box** with your business leaders.
- Deploy the model.

# The Churn Problem

- Gather and transform the appropriate data.
- Try different models on the data and compare their performance.
- Validate your model results (for example use cross-validation).
- Walk through your newly created **black box** with your business leaders.
- **Deploy the model.**

Landing your first  
job as a Data  
Analyst.

# Landing your first job

- Take free or paid courses. Try [coursera.com](https://www.coursera.com).
- Get certifications. For example, SAS and cloudera exams.
- Build a profile on [upwork.com](https://www.upwork.com) (freelancing) or [wyzant.com](https://www.wyzant.com) (tutoring).
- Learn a programming language. Python is a good one to start with.
- Make [stackoverflow.com](https://stackoverflow.com) your best friend. Submit to [kaggle.com](https://www.kaggle.com).

# Landing your first job

- Take free or paid courses. Try [coursera.com](https://www.coursera.com).
- **Get certifications. For example, SAS and cloudera exams.**
- Build a profile on [upwork.com](https://www.upwork.com) (freelancing) or [wyzant.com](https://www.wyzant.com) (tutoring).
- Learn a programming language. Python is a good one to start with.
- Make [stackoverflow.com](https://stackoverflow.com) your best friend. Submit to [kaggle.com](https://www.kaggle.com).

# Landing your first job

- Take free or paid courses. Try [coursera.com](https://www.coursera.com).
- Get certifications. For example, SAS and cloudera exams.
- **Build a profile on [upwork.com](https://www.upwork.com) (freelancing) or [wyzant.com](https://www.wyzant.com) (tutoring).**
- Learn a programming language. Python is a good one to start with.
- Make [stackoverflow.com](https://stackoverflow.com) your best friend. Submit to [kaggle.com](https://www.kaggle.com).

# Landing your first job

- Take free or paid courses. Try [coursera.com](https://www.coursera.com).
- Get certifications. For example, SAS and cloudera exams.
- Build a profile on [upwork.com](https://www.upwork.com) (freelancing) or [wyzant.com](https://www.wyzant.com) (tutoring).
- Learn a programming language. Python is a good one to start with.
- Make [stackoverflow.com](https://stackoverflow.com) your best friend. Submit to [kaggle.com](https://www.kaggle.com).

# Landing your first job

- Take free or paid courses. Try [coursera.com](https://www.coursera.com).
- Get certifications. For example, SAS and cloudera exams.
- Build a profile on [upwork.com](https://www.upwork.com) (freelancing) or [wyzant.com](https://www.wyzant.com) (tutoring).
- Learn a programming language. Python is a good one to start with.
- Make [stackoverflow.com](https://stackoverflow.com) your best friend. Submit to [kaggle.com](https://www.kaggle.com).

# Landing your first job

- Look for junior or entry level roles.
- Take a 1 year contract role.
- Apply even if you don't meet all of the requirements.
- If your GPA is good put it on your resume.
- Network and stay close to your school friends.

# Landing your first job

- Look for junior or entry level roles.
- **Take a 1 year contract role.**
- Apply even if you don't meet all of the requirements.
- If your GPA is good put it on your resume.
- Network and stay close to your school friends.

# Landing your first job

- Look for junior or entry level roles.
- Take a 1 year contract role.
- **Apply even if you don't meet all of the requirements.**
- If your GPA is good put it on your resume.
- Network and stay close to your school friends.

# Landing your first job

- Look for junior or entry level roles.
- Take a 1 year contract role.
- Apply even if you don't meet all of the requirements.
- **If your GPA is good put it on your resume.**
- Network and stay close to your school friends.

# Landing your first job

- Look for junior or entry level roles.
- Take a 1 year contract role.
- Apply even if you don't meet all of the requirements.
- If your GPA is good put it on your resume.
- Network and stay close to your school friends.

Some example  
job listings.

---

## About the Job

### **Job Profile Summary:**

As a Data Analyst you will be expected to work with all aspects of data: import data from various sources using multiple methods, be able to transform and clean the data and prepare for analysis, as well as deliver a desired outcome in a form of a report, a data visualization dashboard or a data-driven insight.

### **Functions and Responsibilities:**

- Extract, transform, model data and visualize and deploy final results via Microsoft Power BI
- Use SQL, R, Power Query (M) and DAX to efficiently import, transform and model the data
- Explore and utilize other Microsoft Office 365 products (SharePoint, Flow, Power Apps, Excel) necessary to efficiently solve existing tasks
- Understand key business drivers and data capture methods
- Handle confidential information appropriately

### **Our Ideal Candidate**

- Has one (1) year of experience in a similar role
- Has a Bachelor's degree or an additional one (1) year of relevant experience performing similar duties
- Has experience using Microsoft Power BI
- Is comfortable navigating ambiguity and demonstrate the ability to manage multiple initiatives simultaneously
- Is a self-starter, comfortable working with little supervision
- Is a creative problem solver
- Is an extremely effective communicator
- Is detail oriented

## About the Job

Alluvion Staffing is seeking a Data analyst for an immediate long term contract with a client in Jacksonville, FL.

*MUST BE ABLE TO WORK ON W2- ABSOLUTELY NO THIRD PARTIES!*

### OVERVIEW:

The Data Analyst will be working and supporting several business units including Sales and finance from a reporting standpoint. Great communication, working with business to identify gaps in reporting, build new reporting process, heavy excel user and experience, sharepoint, sql is a plus. This job is heavy in documentation and attention to detail. Will essentially go to the business, identify new reporting needs, create a process in excel, in put to sharepoint, the devs will make the changes and this person will then go back and train the business on the new reporting process and the data output involved.

### PRIMARY RESPONSIBILITIES:

- Use statistical methods to analyze data and generate useful business reports
- Work with management team to create a prioritized list of needs for each business segment
- Identify and recommend new ways to save money by streamlining business processes
- Use data to create models that depict trends in the customer base and the consumer population as a whole
- Work with departmental managers to outline the specific data needs for each business method analysis project
- skills

### REQUIRED:

- Strong attention to detail
- Bachelor's Degree in Mathematics or Computer Engineering
- 2+ years' Data mining experience
- 4+ years in a data analyst role
- SQL experience preferred
- Extensive experience in reporting
- Advance excel experience
- Business and system process improvement experience required
- Strong communication and collaboration skills

## About the Job

	Data Scientist
<b>Relevant Experience (in Yrs)</b>	<p>5+ years of experience in Hadoop Development</p> <ul style="list-style-type: none"><li>· Candidate should have experience in some form of SQL (e.g. HIVE, PostgreSQL, MSSQL, etc.) and either Python or R for statistical application development. Experience in a Hadoop environment and / or statistical development tool (e.g. RapidMiner, Knime) a plus</li><li>· 5+ years with relevant Data Science experience</li></ul>
<b>Technical/Functional Skills</b>	<ul style="list-style-type: none"><li>· Machine learning, Training and Scoring on real time basis</li><li>· Proven record of successful statistical product delivery</li><li>· Ability to execute health insurance use cases via statistical applications such as R or Python (Sci-Py, Orange, etc.)</li><li>· Significant experience with SQL and working with large datasets required</li><li>· Strong verbal and written communication skills</li><li>· Proficient understanding of distributed computing principles.</li></ul>
<b>Experience Required</b>	<p>8+ years</p> <ul style="list-style-type: none"><li>· Develop and communicate a deep understanding of client needs, perform analytical deep-dives to identify problems, opportunities and specific actions required</li><li>· Develop reproducible and deployable statistical applications on platform such as R/Python /Spark using techniques such as Regression, SVM, and Neural Networks</li></ul>
<b>Roles &amp; Responsibilities</b>	<ul style="list-style-type: none"><li>· Efficiently access data via multiple vectors (e.g. NFS, FTP, SSH, SQL, Sqoop, Flume, Spark)</li><li>· Design experiments to maximize insights while minimizing error</li><li>· Work with cross-functional teams (including Product Management, Engineering, and senior executives) to rapidly execute and iterate</li><li>· Develops business relationships and integrates activities with other IT departments, i.e., Security, Database Administrators, etc., to ensure successful implementation and support of project efforts.</li></ul>
<b>Generic Managerial Skills</b>	<p>Potential opportunities for exploring news ideas and innovations to anchor outcome based achievements</p> <ul style="list-style-type: none"><li>· Coordinate Activities with Others on All Phases of SDLC</li><li>· Experience in Health Care (preferable)</li></ul>

**Above all else:  
Don't give up.**