## Important Job Roles in Data Science Industry Today – Who Does What ?

## TRANSFORM YOUR CAREER WITH BIG DATA ANALYTICS CERTIFICATION FROM

Tech Mahindra



#### Introduction

One evening, I was catching up with a friend over a few drinks – let's call him Jon (name changed). He seemed determined to become a data scientist and was charting out his career plan accordingly.

I quizzed him around his awareness of what a data scientist does and sniffed that he wasn't sure. So I threw this puzzle to him:

There are 4 people A, B, C and D, each with one of the these designations: A Data Scientist, A Data Engineer, A Data Analyst and a Data Architect. I also had job descriptions of their roles, which I passed on to Jon. He had to match the description to the right designation.

To penalize him for his errors, he had to take an alcohol shot for every wrong match he did. Can you guess the number of shots he took that evening? How many would you end up taking, if you were in this situation?

I would just say that he was lucky that I didn't throw this puzzle at him with all the 8 designations I had JDs of that evening.

#### The problem

While the demand for various data science roles is increasing by the day, people in industry have used the designations and descriptions a bit loosely. Hence, there is a lot of confusion around who does what in the industry!

We hope this infographic "The Data Science Industry: Who Does What designed by <u>DataCamp</u>" helps you to find out ideal job role. We have provided a brief description of these roles and an infographic for you to take away.

If you have any questions related to designations and roles, feel free to ask them through comments below. I will be as helpful as my time can permit. Enjoy!

#### **Description of various data science roles:**

#### The Data Scientist

Data Scientist is most likely one of the most sizzling job titles that you can have these days, and with a yearly average salary of \$118,70, they are amongst the top earners in the data science industry. The data scientist has the ability to handle the crude data using the latest technologies and techniques, can perform the necessary analysis, and can present the acquired knowledge to his associates in an informative way.

#### The Data Analyst

Languages like R, Python and SQL are part of the data analyst's basic knowledge. Much like the data scientist role, a broad skillset is also required for the data analyst role, which combines technical and analytical knowledge with ingenuity. This profile is often looked for by companies such as HP and IBM.

#### The Data Architect

Data is (being collected) everywhere and as a consequence, more and more organizations are in need of a data architect. Industries like banking and FMCG use data architects to integrate, centralize, protect and maintain their data sources. These architects often work with the latest technologies such as Spark and always need to be on top of the game to stay relevant.

#### The Statistician

The title of statistician is regularly overlooked by or replaced by fancier sounding job titles. This is a bit of a pity, given that the statistician, with his solid foundation in statistical theories and methodologies, can be seen as the pioneer of the data science field. It is often he who reaps the information from the data and transforms it into actionable insights. To end, although the title can sound a bit dull compared to the others in this list, modern statisticians are always ready to rapidly ace new advancements and utilize these to benefit their research.

#### The Database Administrator

As a database administrator, you ensure that the database is accessible to every stakeholder in the organizations, is performing legitimately and that the necessary safety measures are in place to keep the stored data save. You need to master different technologies going from SQL and XML up to a more general programming language like Java.

#### The Business Analyst

This is probably the least technical profile mentioned on the infographic. However, the

business analyst compensates for this lack of technical knowhow with a profound understanding of the various business processes that are in place. A business analyst therefore often performs the role of the middle person between the business folks and the techies. Organizations searching for business analysts are companies like Uber, Dell and Oracle.

#### **Data and Analytics Manager**

The data and analytics manager steers the direction of the data science team. This individual consolidates strong and specialized skills in a various arrangement of advancements (SQL, R, SAS, ... ) with the social aptitudes required to deal with a group. It's a hard employment but if you feel up for the challenge, make sure to have a look at offerings from companies such as Coursera, Slack. Luckily, with a yearly average salary of \$116k, the financial compensation is in line with the high requirements.



# THE DATA SCIENCE INDUSTRY WHO DOES WHAT

A LOOK AT THE KEY ROLES IN DATA SCIENCE

The vast range of careers in the data science industry is accompanied by an avalanche of job postings. DataCamp took a closer look at recent data science job postings to learn more about the differences between various roles in the industry and to unravel the meaning behind many cool-sounding job titles.

#### Languages

R, SAS, Python, Matlab, SQL, Hive, Pig, Spark

#### **Skills & Talents**

- ✓ *Distributed computing*
- ✓ *Predictive modeling*
- ✓ Story-telling and visualizing
- ✓ Math, Stats, Machine Learning

# HIRED BY

Google Microsoft

## DATA SCIENTIST 'AS RARE AS UNICORNS'

#### Role

Cleans, massages and organizes (big) data

Mindset
Curious data wizard

## DATA ANAL DATA DETECTIV

#### Role

Collects, processes and performs statistical data analyses

#### Mindset

Intuitive data junkie with high "figure-it-out" quotient



Languages

*R*, *Python*, *HTML*, *Javascript*, *C/C++*, SQL

#### **Skills & Talents**

- ✓ *Spreadsheet tools* (*e.g. Excel*)
- ✓ Database systems (SQL and NO SQL based)
- ✓ Communication & visualization
- ✓ Math, Stats, Machine Learning

#### Languages

SQL, XML, Hive, Pig, Spark

#### **Skills & Talents**

- ✓ Data warehousing solutions
- ✓ In-depth knowledge of database architecture
- ✓ Extraction Transformation and Load (ETL), spreadsheet and BI tools
- ✓ Data modeling
- ✓ Systems development



HIRED BY

VISA Coca Cola logitech

#### DATA ARCHITEC THE CONTEMPORARY DATA MODELLER

#### Role:

Creates blueprints for data management systems to integrate, centralize, protect and maintain data sources

#### **Mindset:**

*Inquiring ninja with a love for* data architecture design patterns

## DATA ENGINEER SOFTWARE ENGINEERS BY TRADE

#### Role

Develops, constructs, tests and maintains architectures (such as databases and large-scale *processing systems)* 

#### **Mindset**

All-purpose everyman



#### Languages

SQL, Hive, Pig, R, Matlab, SAS, SPSS, Python, Java, Ruby, C++, Perl

#### **Skills & Talents**

- ✓ Database systems (SQL & NO SQL
- ✓ Data modeling & ETL tools
- ✓ Data APIs
- ✓ Data warehousing solutions

#### Languages

R, SAS, SPSS, Matlab, Stata, Python, Perl, Hive, Pig, Spark, SQL

#### **Skills & Talents**

- ✓ Statistical theories & methodology
- Data mining on machina lagraina

#### Role

Collects, analyzes and interpretsqualitative as well as quantitive data with statistical theories and mothode

Data mining o machine warning

✓ *Distributed Computing (Hadoop)* 

✓ Database systems (SQL and NO SQL based)

✓ Cloud tools



**Mindset** 

Logical and enthusiastic stats genius Linked in Johnson Pepsico

## DATABASE ADMINISTRATOR 'DATABASE CARETAKER'

#### Role

Ensures that the database is available to all relevant users, is performing properly and is being kept safe

#### Mindset

Master of Disaster Prevention



💠 + a b l e a v 👸 reddit 🔰

#### Languages

SQL, Java, Ruby on Rails, XML, C#, Python

#### **Skills & Talents**

- ✓ Backup & recovery
- ✓ Data modeling and design
- ✓ *Distributed Computing (Hadoop)*
- ✓ Database systems (SQL and NO SQL based)
- ✓ Data security
- ✓ ERP & business knowledge

#### Languages SQL

#### **Skills & Talents**

- ✓ Basic tools (e.g. MS Office)
- ✓ Data visualization tools (e.g. Tableau)
- ✓ Conscious listening and storytelling
- ✓ Business Intelligence understanding
- ✓ Data modeling



### BUSINESS ANALYST

#### Role

Improves business processes as intermediary between business and IT

#### **Mindset**

Resilient project juggler

## DATA AND ANALYTICS MANAGER DATA SCIENCE TEAM LEADER

UBER

#### Role

Manages a team of analysts and data scientists

#### Mindset

Data Wizards' Cheerleader



#### Languages

SQL, R, SAS, Python, Matlab, Iava

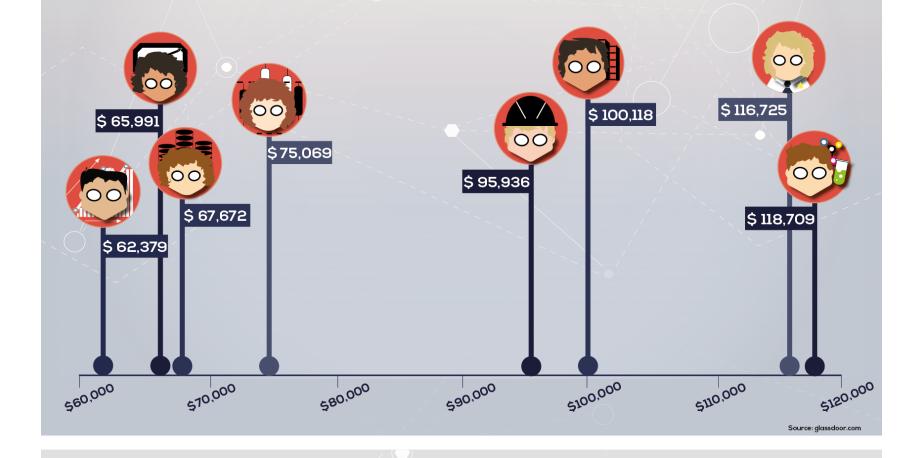
#### **Skills & Talents**

- ✓ Database systems (SQL and NO SQL based)
- ✓ Leadership & project management
- ✓ Interpersonal communication
- ✓ Data mining & predictive modeling

COURSER #Slack MOTOROLA SOLUTIONS

## WHO EARNS WHAT SALARY

NATIONAL AVERAGE SALARIES (US)



## FIND YOUR DATA SCIENCE DREAM JOB TODAY!

#### **BUT REMEMBER...**

- ✓ There isn't a generally accepted definition of data science and its roles
  (yet)
- ✓ Make sure to ask about details on what projects and technologies you'll be working with to make sure that there is a fit for you and your skills!

## WANT TO BECOME A DATA SCIENCE PROFESSIONAL? Join over 230,000 data science enthusiasts at DataCamp!

Learn R and data analysis at your own pace and in the comfort of your own browser with DataCamp's fun video lessons and interactive coding challenges.



www.DataCamp.com

#### REFERENCES

http://www.datasciencecentral.com/profiles/blogs/what-roles-do-you-need-in-your-data-science-team http://blog.udacity.com/2014/11/data-science-job-skills.html https://www.linkedin.com/pulse/big-data-what-key-jobs-salaries-available-bernard-marr http://skillcrush.com/2015/03/05/41-tech-job-titles/

http://blog.hackerrank.com/the-risky-eclipse-of-statisticians/
http://www.analyticsvidhya.com/blog/2014/11/data-science-manager-lead-team/
https://datajobs.com/Data-Science-Jobs
http://blog.udacity.com/2014/12/9-essential-skills-data-analyst-job.html
http://study.com/articles/How\_to\_Become\_a\_Data\_Architect\_Education\_and\_Career\_Roadmap.

http://insights.dice.com/2014/08/21/key-skills-needed-big-data-engineers/ html



#### **Author:**



<u>Martijn Theuwissen</u> is co-founder of <u>DataCamp</u>, which helps people learn R and Data Science from their browser and their comfort through fun videos and interactive coding challenges.

If you like what you just read & want to continue your analytics learning, <u>subscribe to our emails</u>, <u>follow us on twitter</u> or like our <u>facebook page</u>.