

# European Data Science Salary Survey

Tools, Trends, What Pays (and What Doesn't) for Data Professionals in Europe



John King & Roger Magoulas



San Jose



London



Beijing



New York



Singapore

# Strata + Hadoop

---

WORLD

Make Data Work  
[strataconf.com](http://strataconf.com)

Presented by O'Reilly and Cloudera, Strata + Hadoop World helps you put big data, cutting-edge data science, and new business fundamentals to work.

- Learn new business applications of data technologies
- Develop new skills through trainings and in-depth tutorials
- Connect with an international community of thousands who work with data

# Take the Data Science Salary Survey

As data analysts and engineers—as professionals who like nothing better than petabytes of rich data—we find ourselves in a strange spot: we know very little about ourselves. But that's changing. This salary and tools survey is the third in an annual series. To keep the insights flowing, we need one thing: **PEOPLE LIKE YOU TO TAKE THE SURVEY.**

Anonymous and secure, the survey will continue to provide insight into the demographics, work environments, tools, and compensation of practitioners in our field. We hope you'll consider it a civic service. We hope you'll participate today.

[Take the Survey](#)

# 2017 European Data Science Salary Survey

Tools, Trends, What Pays (and What Doesn't)  
for Data Professionals in Europe

John King and Roger Magoulas

O'REILLY

## **2017 EUROPEAN DATA SCIENCE SALARY SURVEY**

by John King and Roger Magoulas

Editor: Shannon Cutt

Designer: Ellie Volckhausen

Production Editor: Shiny Kalapurakkal

Copyright © 2016 O'Reilly Media, Inc. All rights reserved.

Printed in Canada.

Published by O'Reilly Media, Inc., 1005 Gravenstein Highway North,  
Sebastopol, CA 95472.

O'Reilly books may be purchased for educational, business, or sales promotional use. Online editions are also available for most titles (<http://safaribooksonline.com>). For more information, contact our corporate/institutional sales department: 800-998-9938 or [corporate@oreilly.com](mailto:corporate@oreilly.com).

2017-02-10. First Edition

ISBN: 978-1-491-97750-7

## **REVISION HISTORY FOR THE FIRST EDITION**

2017-02-10: First Release

While the publisher and the authors have used good faith efforts to ensure that the information and instructions contained in this work are accurate, the publisher and the authors disclaim all responsibility for errors or omissions, including without limitation responsibility for damages resulting from the use of or reliance on this work. Use of the information and instructions contained in this work is at your own risk. If any code samples or other technology this work contains or describes is subject to open source licenses or the intellectual property rights of others, it is your responsibility to ensure that your use thereof complies with such licenses and/or rights.

# Table of Contents

2017 European Data Science Salary Survey .....	i
Executive Summary.....	1
Introduction .....	2
Countries .....	4
Salary Versus GDP.....	8
Company Size.....	10
Industry .....	12
Tools .....	14
Tasks.....	18
Coding and Meetings.....	22
Salary Change .....	24
Conclusion .....	26

## 2017 EUROPEAN DATA SCIENCE SALARY SURVEY

HERE WE TAKE A DEEP DIVE  
INTO THE RESULTS FROM  
RESPONDENTS BASED IN  
EUROPE, EXPLORING CAREER  
DETAILS AND FACTORS THAT  
INFLUENCE SALARY

**YOU CAN PRESS ACTUAL BUTTONS** (and earn our sincere gratitude) by taking the 2017 survey—it only takes about 5 to 10 minutes, and is essential for us to continue to provide this kind of research.  
[oreilly.com/ideas/take-the-2017-data-science-salary-survey](http://oreilly.com/ideas/take-the-2017-data-science-salary-survey)



# Executive Summary

## IN 2016, O'REILLY MEDIA CONDUCTED A DATA SCIENCE SALARY SURVEY ONLINE.

The survey contained 40 questions about the respondents' roles, tools, compensation, and demographic backgrounds. About 1,000 data scientists, analysts, engineers, and other professionals working in Data participated in the survey—359 of them from European countries. Here, we take a deep dive into the results from respondents based in Europe, exploring career details and factors that influence salary. Some key findings include:

- Most of the variation in salaries can be attributed to differences in the local economy
- Data professionals who use Hadoop and Spark earn more

Respondents who use Hadoop, Spark, or Python were twice as likely to have a major increase in salary over the last three years.

- Among those who use R or Python, users of both have the highest salaries
- A few technical tasks correlate with higher salaries: developing prototype models, setting up/maintaining data platforms, and developing products that depend on real-time analytics
- Respondents who use Hadoop, Spark, or Python were twice as likely to have a major increase in salary over the last three years, compared with those whose stack consists of Excel and relational databases

We hope that these findings will be useful as you develop your career in data science.

# Introduction

**SINCE 2013, WE HAVE CONDUCTED AN ONLINE SALARY SURVEY FOR DATA PROFESSIONALS** and published a report on our findings. US respondents typically dominate the sample, at about 60%–70%. Although many of the findings do appear to apply to people across the globe, we thought it would be useful to show results specific to Europe, looking at finer geographical details and identifying any patterns that seem to only apply to Europe. In this report, we pool all 359 European respondents from the Data Salary Survey over a 13-month period: September 2015 to October 2016.

The median salary of European respondents was €48K, but the spread was huge. For example, the top third earned almost four times on average as the bottom third. Such a large variance is not surprising due to the differences in the per capita income of countries represented.

A note on currency: we requested responses about salaries and other monetary amounts in US dollars. In this report, we have converted all amounts into euros, though many European

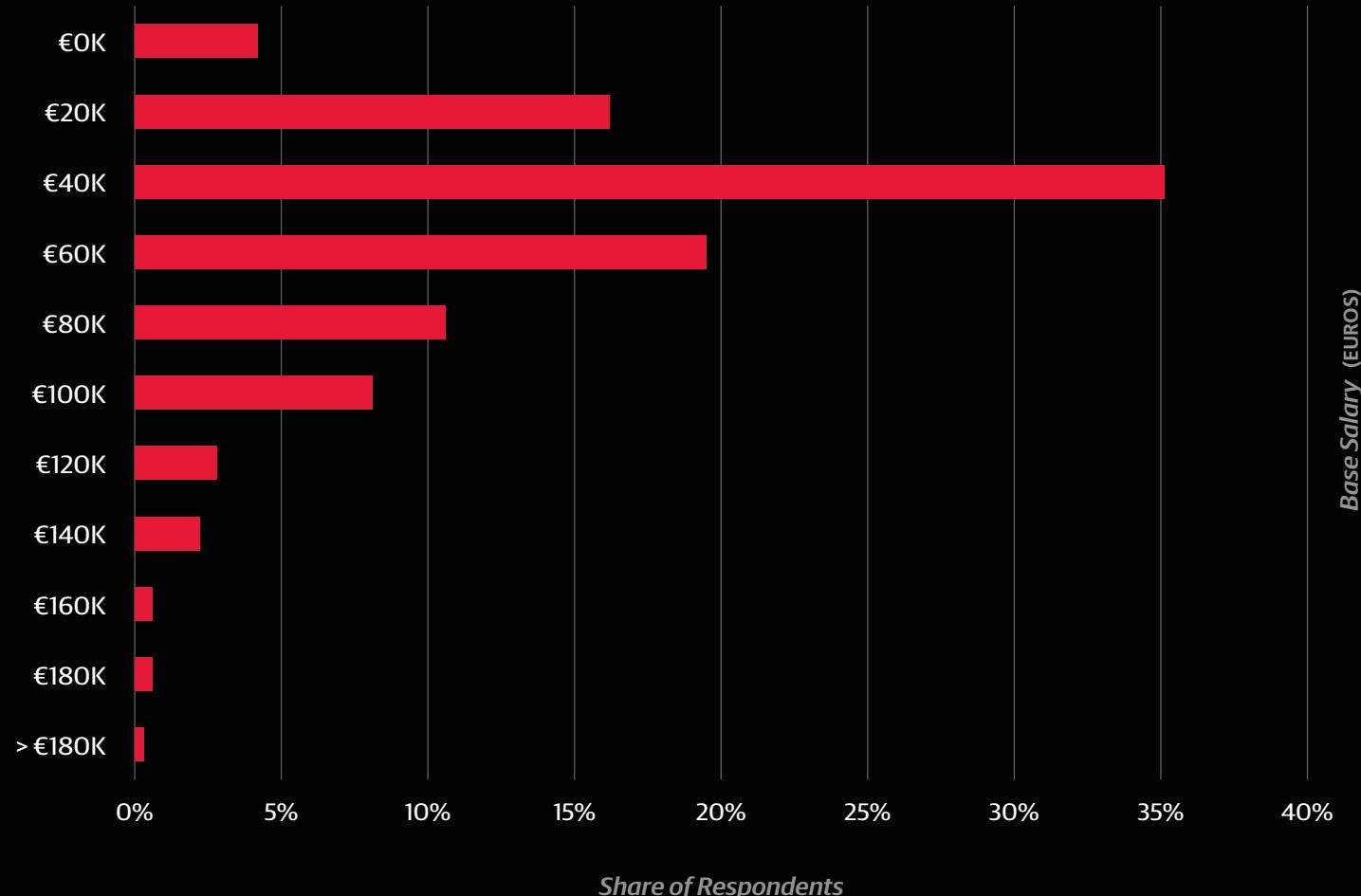
respondents are paid in other currencies, such as pounds or rubles. Over the period in which responses were collected, there were some important shifts in exchange rates, most notably the fall of the pound after Brexit. However, the geographical distribution of responses did not correlate in any meaningful way with any period of collection (e.g., when the pound was high or low), so these currency fluctuations likely translate into noise rather than bias.

*In the horizontal bar charts throughout this report, we include the interquartile range (IQR) to show the middle 50% of respondents' answers to questions such as salary. One quarter of the respondents have a salary below the displayed range, and one quarter have a salary above the displayed range.*

*The IQRs are represented by colored, horizontal bars. On each of these colored bars, the white vertical band represents the median value.*

## BASE SALARY (EURO)

SHARE OF RESPONDENTS



# Countries

**THE UK WAS THE MOST WELL-REPRESENTED EUROPEAN COUNTRY**, with about a quarter of the sample, followed by Germany, Spain, and the Netherlands. By far, the highest salaries were in Switzerland, with a median salary of €117K, followed by Norway with €96K, although the latter figure is only based on five respondents. Among countries represented by more than just a handful of respondents, the UK had the second-highest median salary: €63k (£53).

Even within Western Europe, there was significant variation in salary. While UK, Swiss, and Scandinavian salaries were significantly higher than the Western European median of

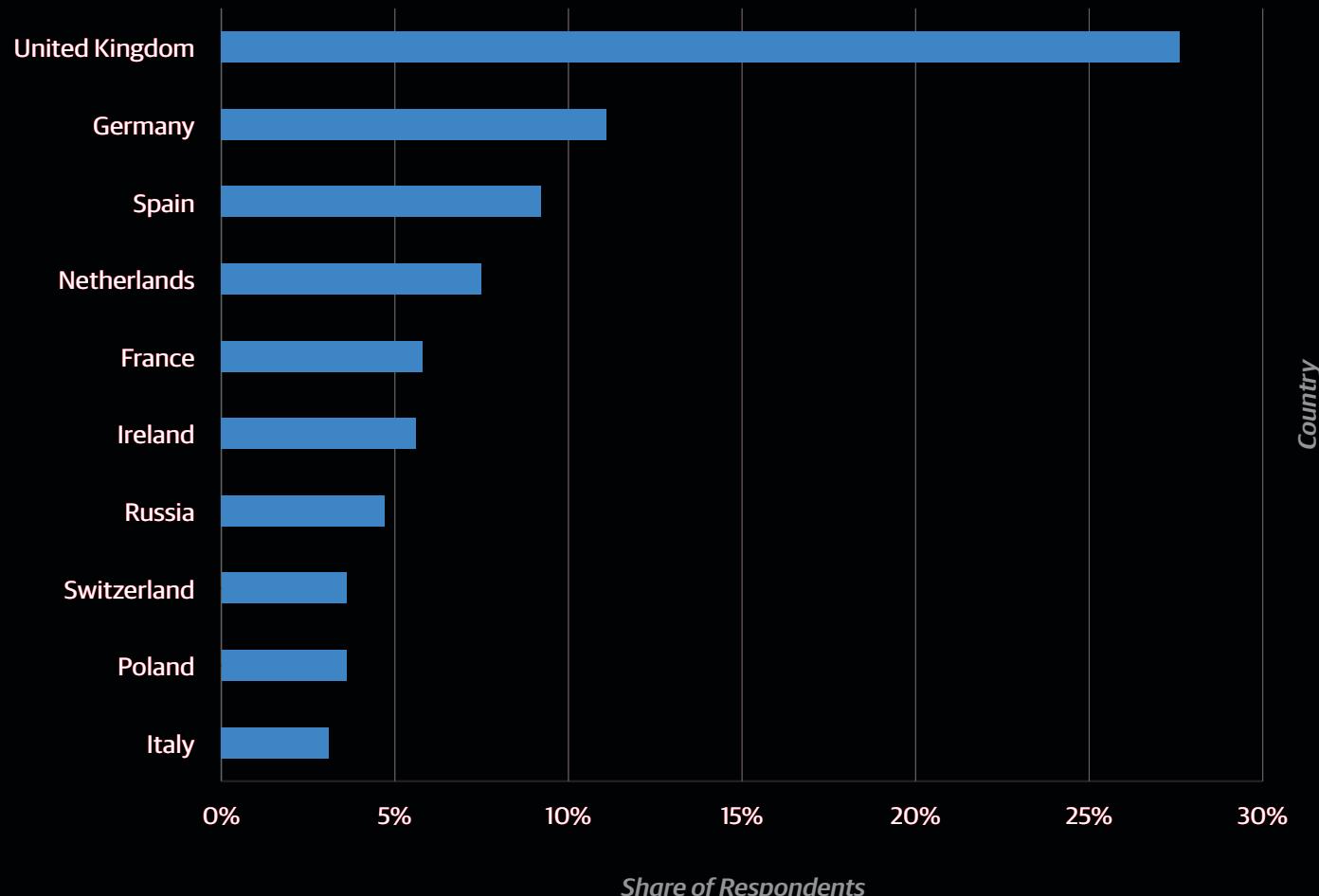
€54K, Spanish and Italian respondents tended to have much lower salaries (€35K). Portugal was somewhat of an outlier in Western Europe, with a median of €22K. The median salaries of Germany, the Netherlands, and France were close to the regional median (about €53K).

Unlike in the west, Eastern European salaries appeared to be fairly consistent, even across national borders.

Salaries drop dramatically as we move south and east. The median salary of respondents from Central and Eastern Europe was €17K. Russia and Poland, the two most well-represented countries in this half of the continent, also had median salaries of €17K: unlike in the west, Eastern European salaries appeared to be fairly consistent, even across national borders.

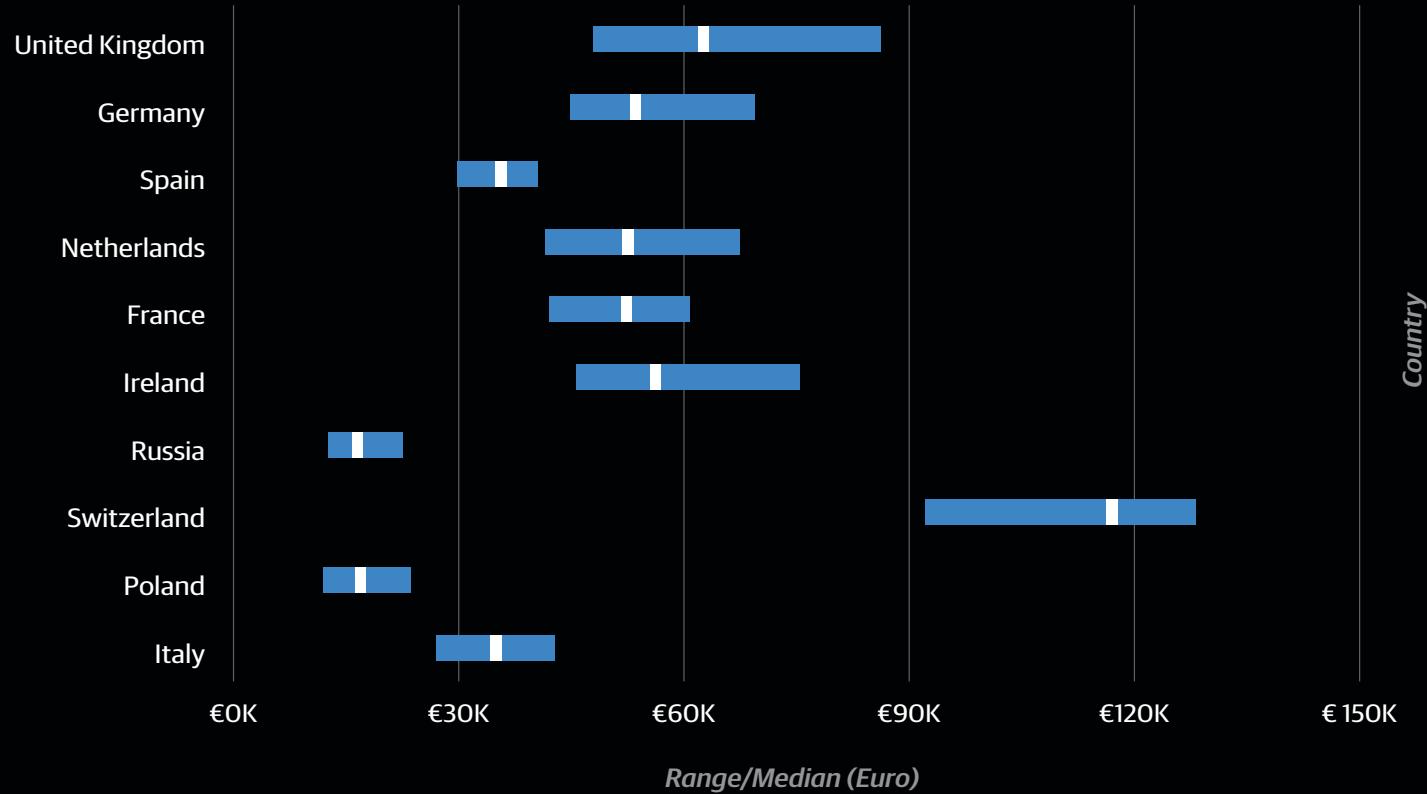
## COUNTRIES

### SHARE OF RESPONDENTS



# COUNTRIES

## SALARY MEDIAN AND IQR\* (EURO)



\*The interquartile range (IQR) is the middle 50% of respondents' salaries. One quarter of respondents have a salary below this range, one quarter have a salary above this range.



# Salary Versus GDP

## NATIONAL MEDIAN SALARIES SHOULD BE EXPECTED

**TO VARY** according to the economic conditions of the country, so the question becomes: given a country's economy (in particular, its per capita GDP), do the salaries of data scientists and engineers vary? Here, we plot per capita GDP and median salary of each country in the sample. The resulting graph is remarkably linear, with outliers largely explained by small sample size: Greece, for example, has a higher-than-expected median salary given a relatively low per capita GDP, but this is based on just one respondent.

The question becomes,  
given a country's  
economy (in particular,  
its per capita GDP),  
do the salaries of  
data scientists and  
engineers vary?

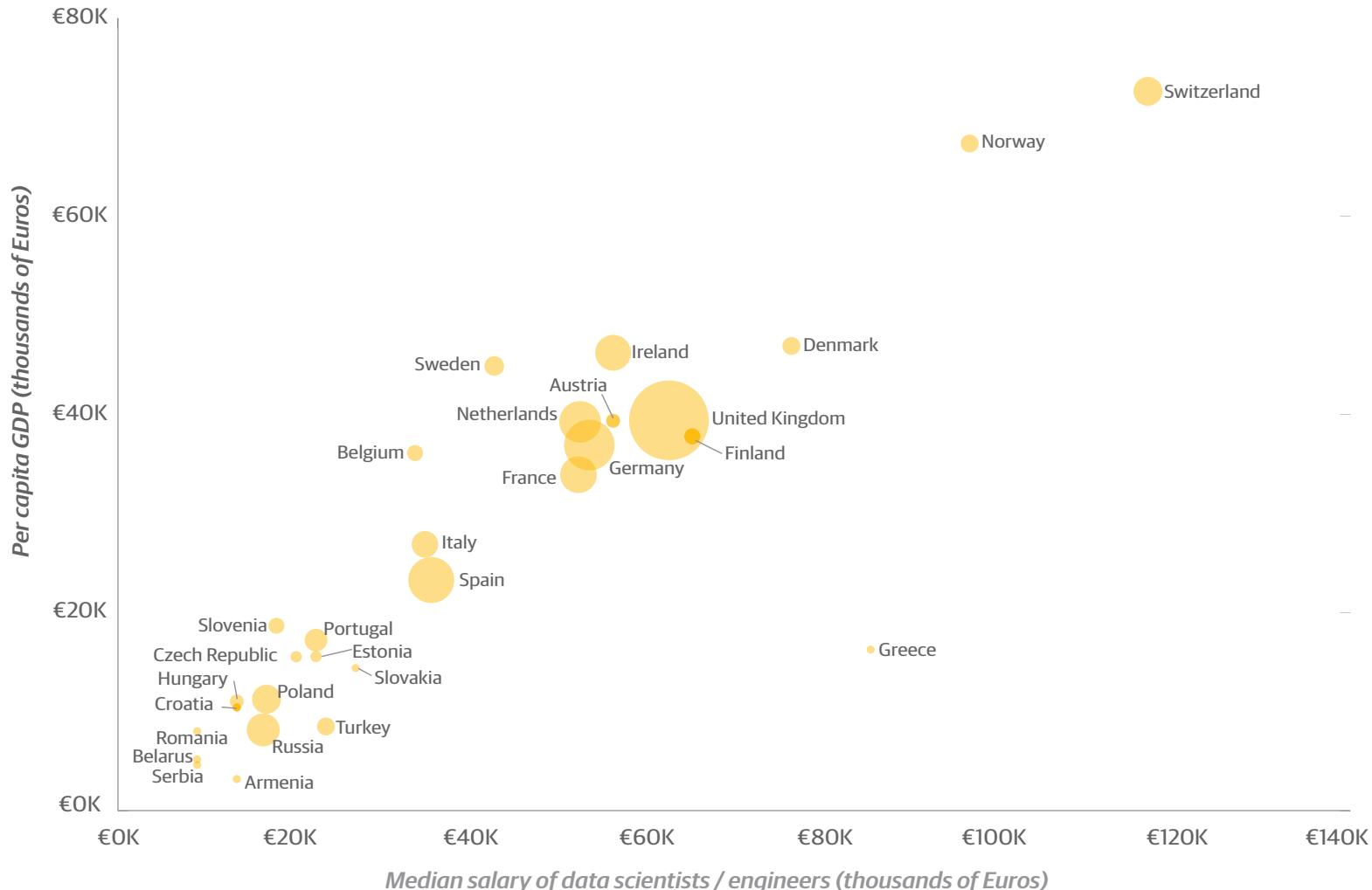
One shortcoming of this plot is that it does not take into account years of experience, which turns out to be very uneven in the sample among different countries. In particular, respondents from Western Europe tended to be much more experienced (with an average of seven years) than respondents from Eastern Europe (with an average of four years). Since experience correlates with salary, the West-East salary difference is exaggerated due to this experience differential.

# SALARY VERSUS GDP

The size of each circle represents the number of respondents from the country in the sample.

## MEDIAN SALARY VERSUS PER CAPITA GDP

Source for per capita GDP: [https://en.wikipedia.org/wiki/List\\_of\\_countries\\_by\\_GDP\\_\(nominal\)\\_per\\_capita](https://en.wikipedia.org/wiki/List_of_countries_by_GDP_(nominal)_per_capita)

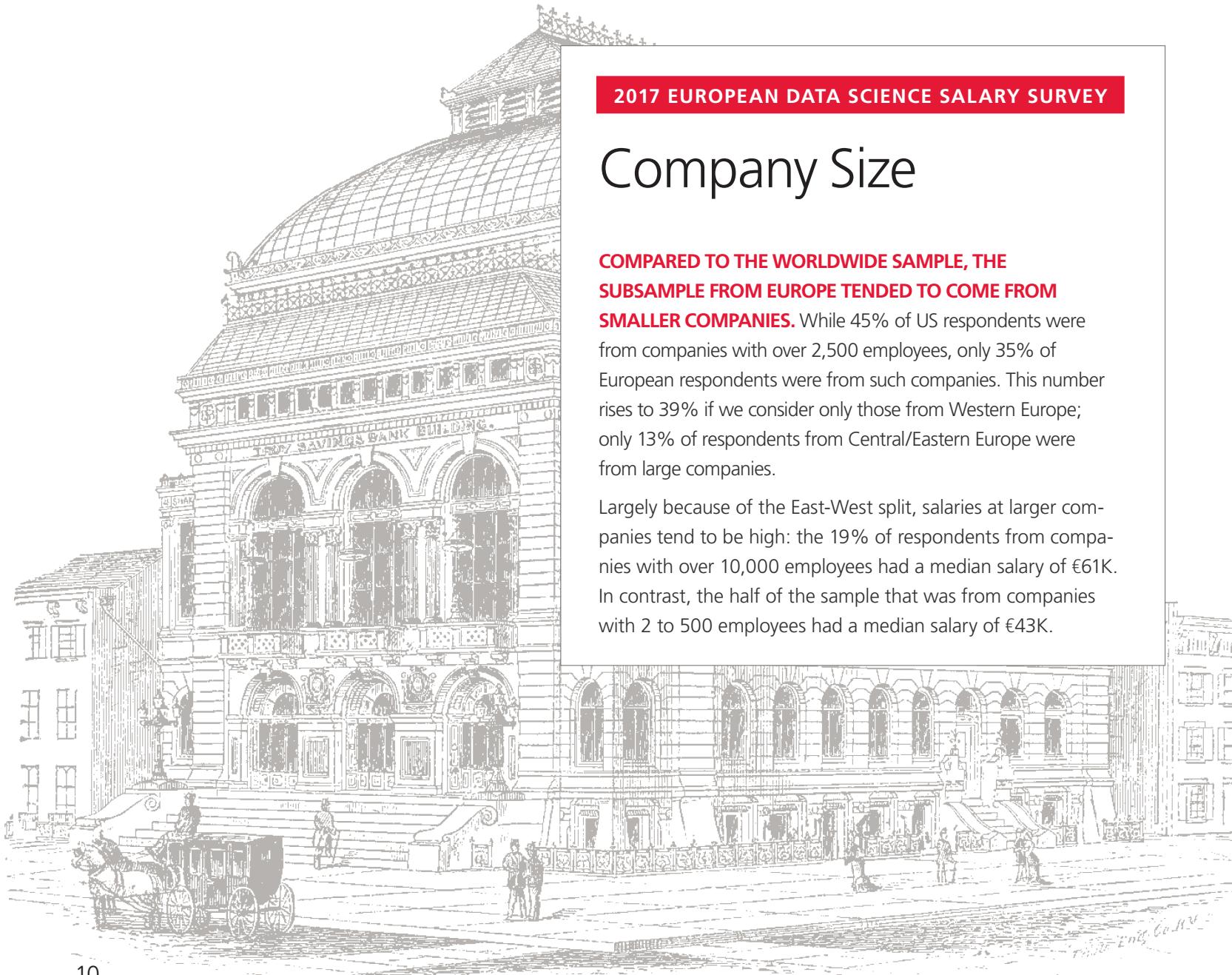


# Company Size

## COMPARED TO THE WORLDWIDE SAMPLE, THE SUBSAMPLE FROM EUROPE TENDED TO COME FROM SMALLER COMPANIES.

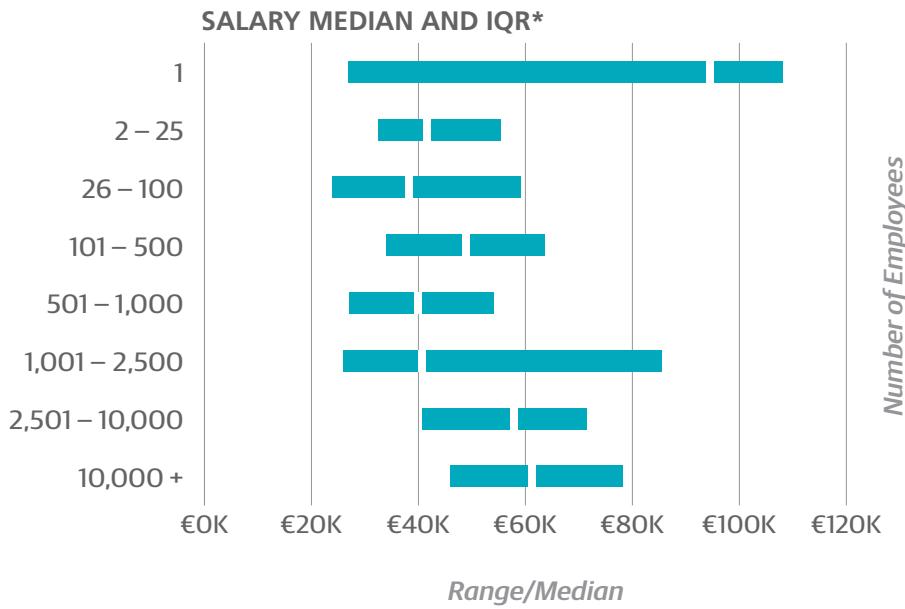
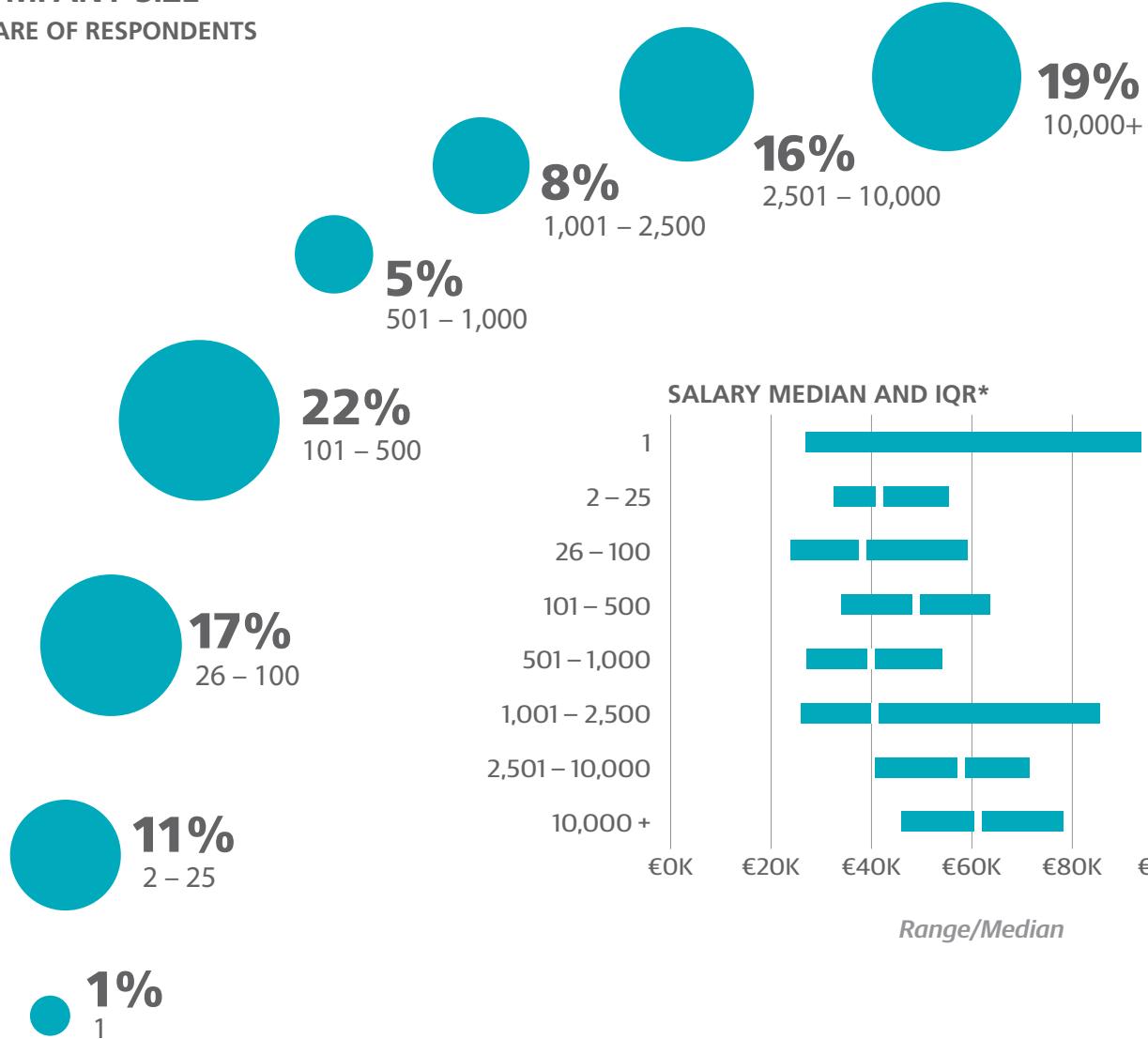
While 45% of US respondents were from companies with over 2,500 employees, only 35% of European respondents were from such companies. This number rises to 39% if we consider only those from Western Europe; only 13% of respondents from Central/Eastern Europe were from large companies.

Largely because of the East-West split, salaries at larger companies tend to be high: the 19% of respondents from companies with over 10,000 employees had a median salary of €61K. In contrast, the half of the sample that was from companies with 2 to 500 employees had a median salary of €43K.



## COMPANY SIZE

SHARE OF RESPONDENTS



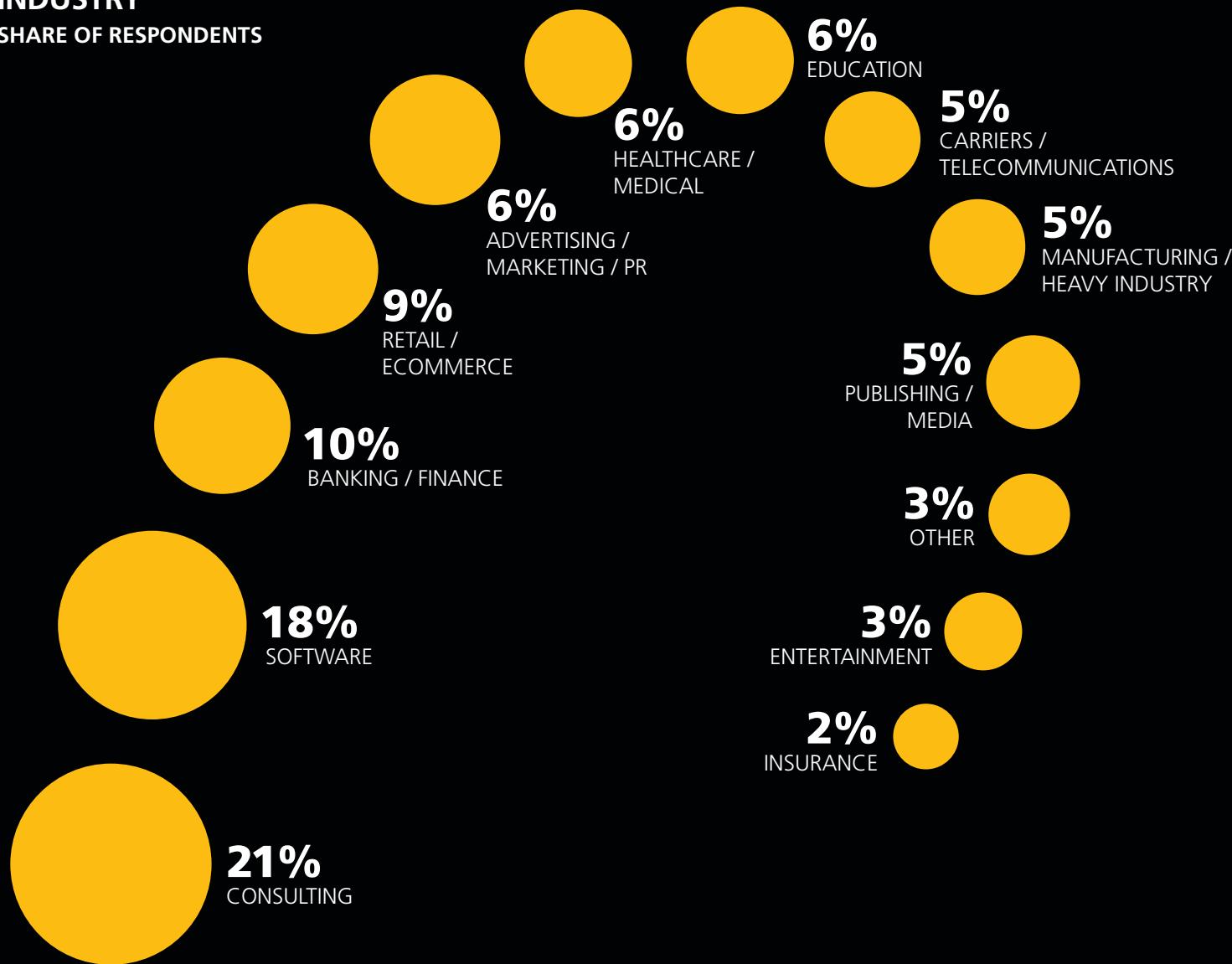
# Industry

**A PLURALITY OF RESPONDENTS (20%) WORKED IN CONSULTING**, after which the top industries were software (18%), banking/finance (10%), and retail/ecommerce (9%). These figures are very similar to those of the worldwide sample.

As with company size, the differences in salaries among industries was largely attributable to geography. Manufacturing, insurance, and publishing/media were all overrepresented by countries with higher salaries. One exception to this was banking/finance, which had a high median salary of €58K and did not correlate with a particular country or region: data professionals in banking do appear to earn more.

# INDUSTRY

## SHARE OF RESPONDENTS



# Tools

## THE TOP FOUR TOOLS FROM EUROPEAN RESPONDENTS WERE EXCEL, SQL, R, AND PYTHON,

each used by over half of all respondents. These four tools have kept their top positions in every Data Salary Survey we have conducted, and there does not appear to be any sign of this changing. Almost every respondent reported using at least one, and about half the sample used three or all four.

Commonly used tools with above-average salaries include Scikit-learn (whose users have a median salary of €52K), Spark (€55K), Hive (€57K), and Scala (€70K). Readers may notice that most tools have a higher median salary than the sample-wide median salary of €48K. This is because respondents who use lots of tools tend to earn more (and they are counted in a large number of tool salary medians). The 43% of respondents who used no more than 10 tools had a median salary of €43K, while

Commonly used tools with above-average salaries include Scikit-learn (whose users have a median salary of (€52K), Spark (€55K), Hive (€57K), and Scala (€70K).

those who used more than 10 tools had a median salary of €53K.

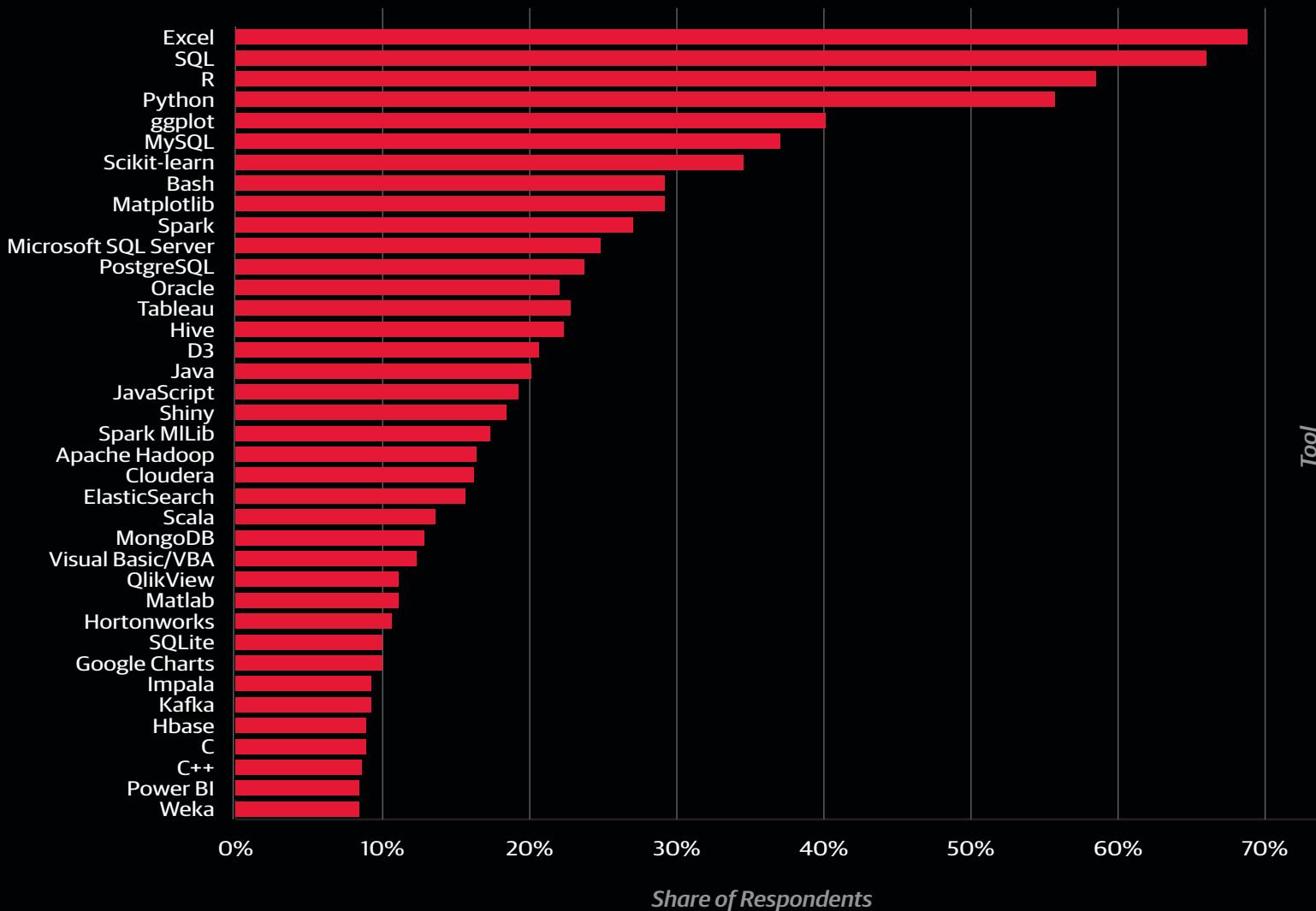
Since there is significant overlap between users of individual tools, it is useful to consider mutually exclusive groups of respondents based on tool usage. The groups we will define here are based on a simple set of rules, but using a clustering algorithm would produce very similar results. The rules are:

- 1) If someone used Spark or Hadoop, we call them "Hadoop"
- 2) If someone (not in the Hadoop group) uses R and/or Python, they are labeled "R+Python," "R-only," or "Python-only," as appropriate
- 3) Everyone who uses SQL and/or Excel (usually both), we call "SQL/Excel"

The five resulting groups each contain between 13% and 26% of the sample. The Hadoop group reported the

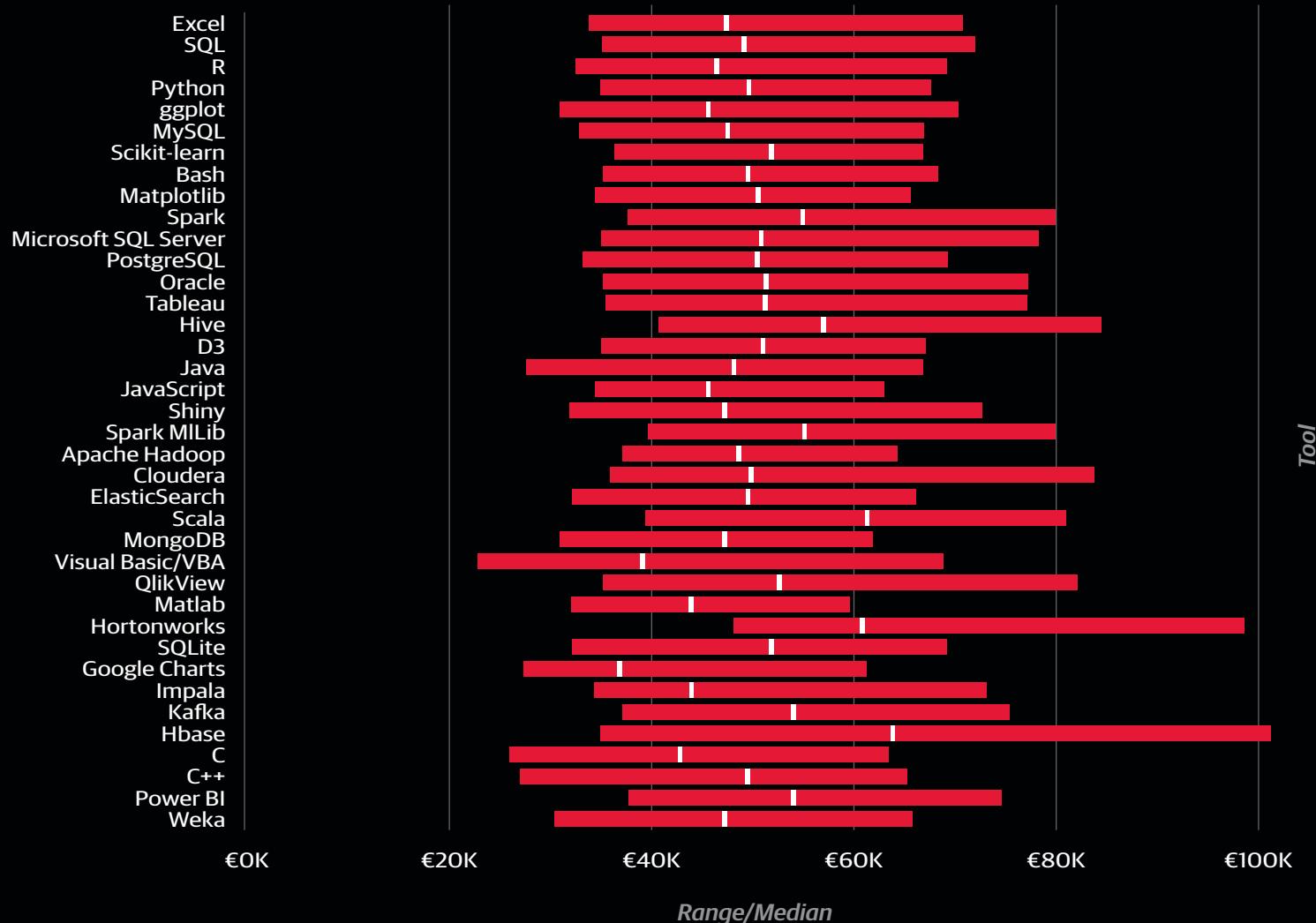
# TOOLS

## SHARE OF RESPONDENTS



# TOOLS

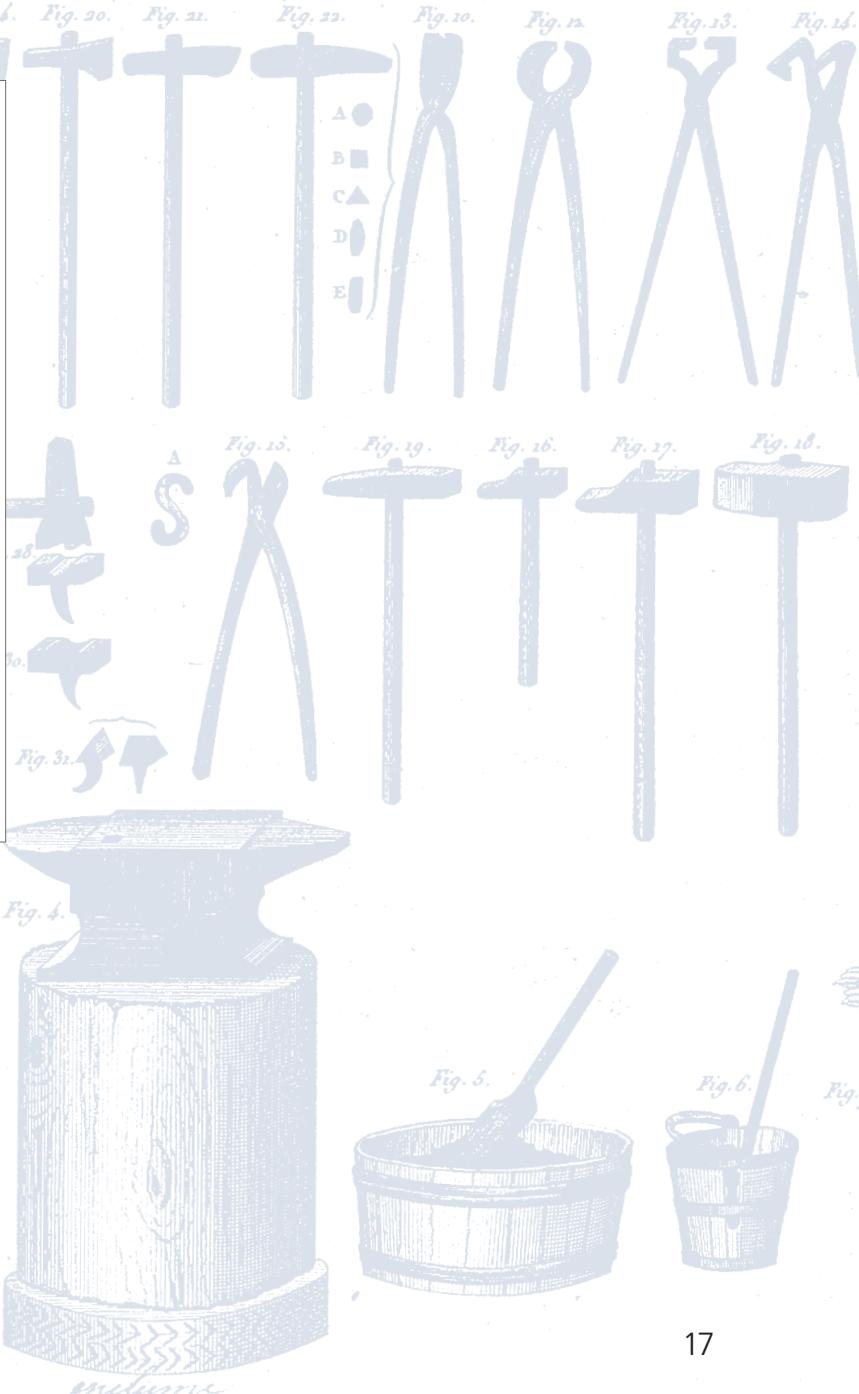
## SALARY MEDIAN AND IQR\*



## 2017 EUROPEAN DATA SCIENCE SALARY SURVEY

highest salaries (median: €56K), while the R-only group had the lowest (€42K). However, this doesn't mean that knowing R means less pay: respondents using Python and R earned slightly more than those using Python and not R.

Aside from salary, one important difference between the groups is experience. The SQL/Excel group—in other words, those who don't use Python, R, Spark, or Hadoop—was more experienced than the other groups (8.3 years on average), followed by the R-only (7.3 years), Hadoop (6.3 years), Python-only (6 years), and Python+R groups (5.2 years). Since we expect more-experienced data professionals to earn higher salaries, the median salary of €46K for the SQL/Excel group is actually quite low, while the €48K of the Python-R group is high.



# Tasks

## WE ALSO ASKED FOR INFORMATION ABOUT WORK

**TASKS:** this is meant to dig a little deeper than what we can glean from a job title. Respondents could say they had “major” or “minor” involvement in each task. For the most part, tasks that correlate positively with salary also correlate positively with years of experience (and often are clearly associated with being a manager).

Among the most common tasks were “basic exploratory data analysis,” “data cleaning,” “creating visualizations,” and “conducting data analysis to answer research questions,” each with 85%–93% of the sample

as a major or minor task. Data cleaning has the unfavorable distinction of being the only task for which each level of involvement means less pay: those with major involvement earn less than those with minor involvement, who in turn earn less than those who never clean data. However, this may have more to do with the fact that more-experienced data professionals (who we know earn more) tend to do less data cleaning.

Tasks that correlate most strongly with high salaries are those that involve management and business decisions.

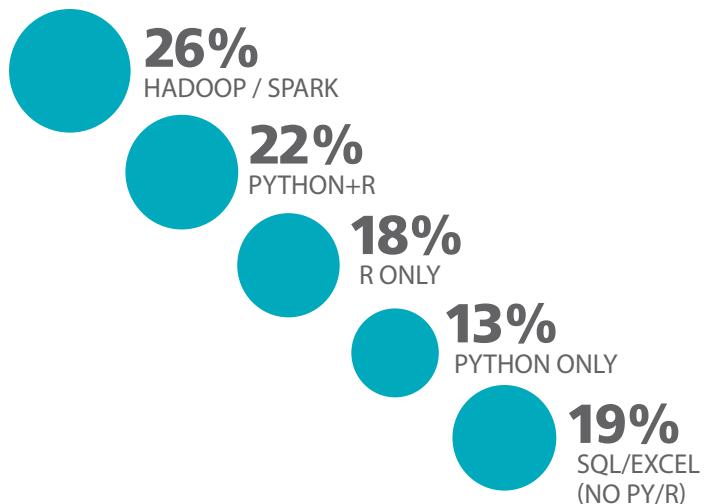
Tasks that correlate most strongly with high salaries are those that involve management and business decisions, such as “communicating findings to business decision-makers,” “identifying business problems to be solved with analytics,” “organizing and guiding team projects,” and “communicating with people outside of your company”. The median salaries of respondents who reported major involvement in these tasks were €54K, €56K, €66K, and €55K, respectively.

Aside from management and business strategy, several technical tasks stood out for above-average salaries:

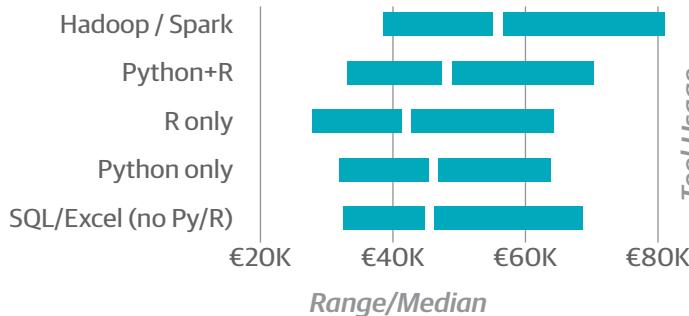
“developing prototype models” (major involvement: €52K), “setting up/maintaining data platforms” (€50K), and “developing products that depend on real-time analytics” (€62K). For each of these tasks, respondents who reported major involvement earned more than those who reported minor involvement, and those who reported minor involvement earned more than those who did not engage in these tasks at all.

## RESPONDENT CATEGORIES BASED ON TOOL USAGE

### SHARE OF RESPONDENTS

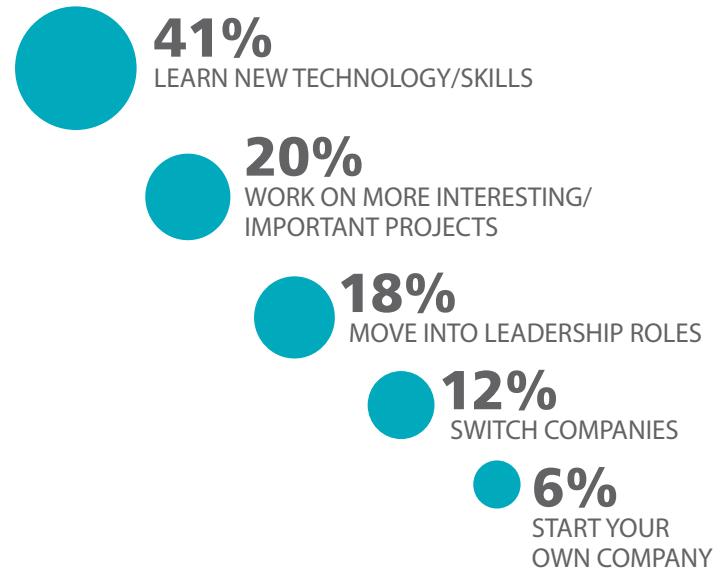


### SALARY MEDIAN AND IQR (EUROS)

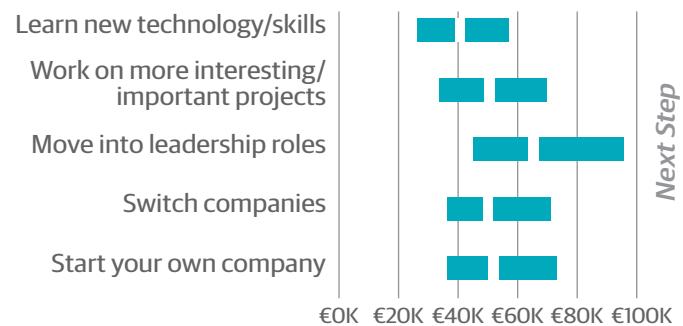


## WHICH OF THE FOLLOWING MOST ACCURATELY DESCRIBES THE NEXT STEP YOU WOULD LIKE TO TAKE TO ADVANCE YOUR CAREER?

### SHARE OF RESPONDENTS

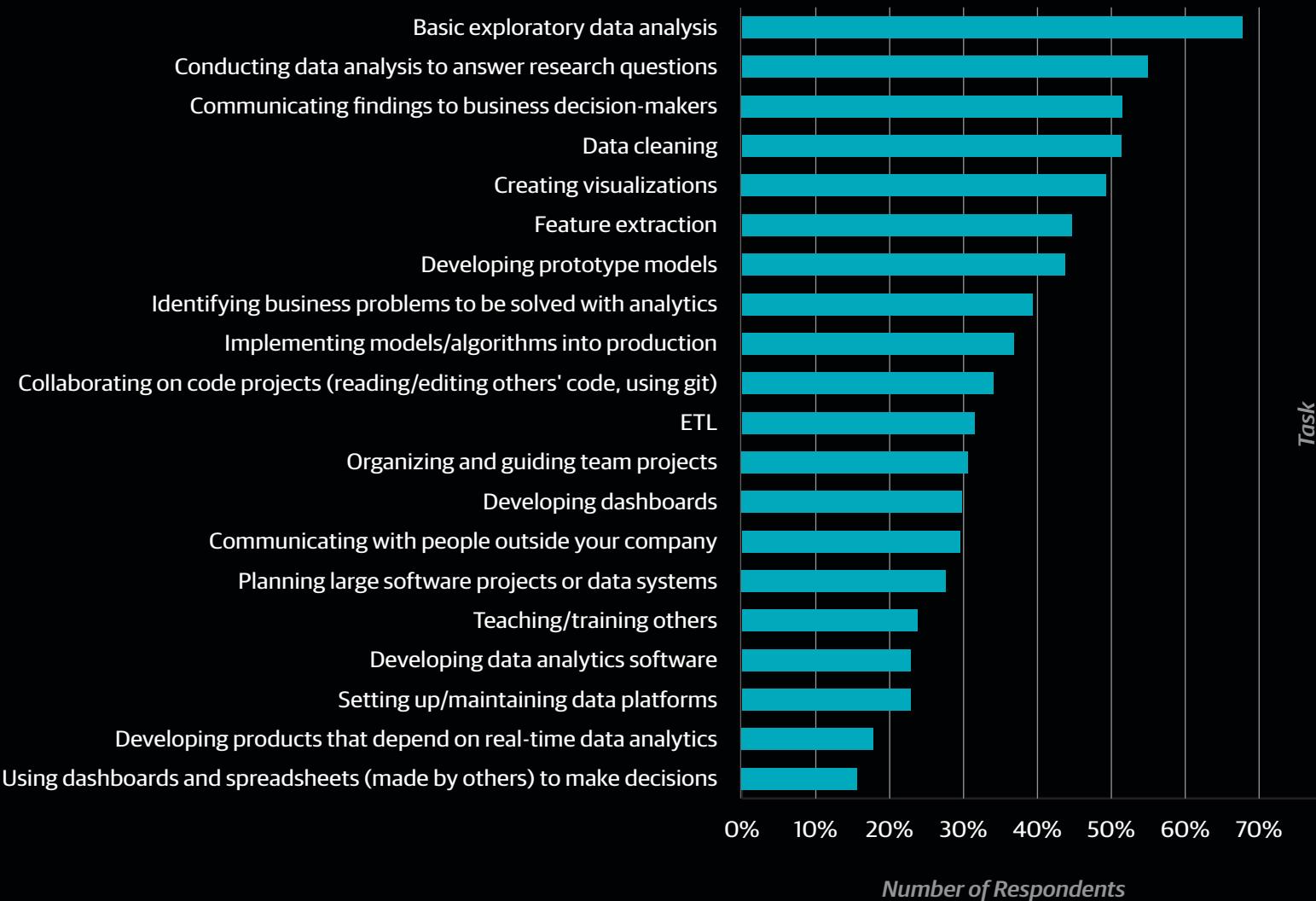


### SALARY MEDIAN AND IQR (EUROS)



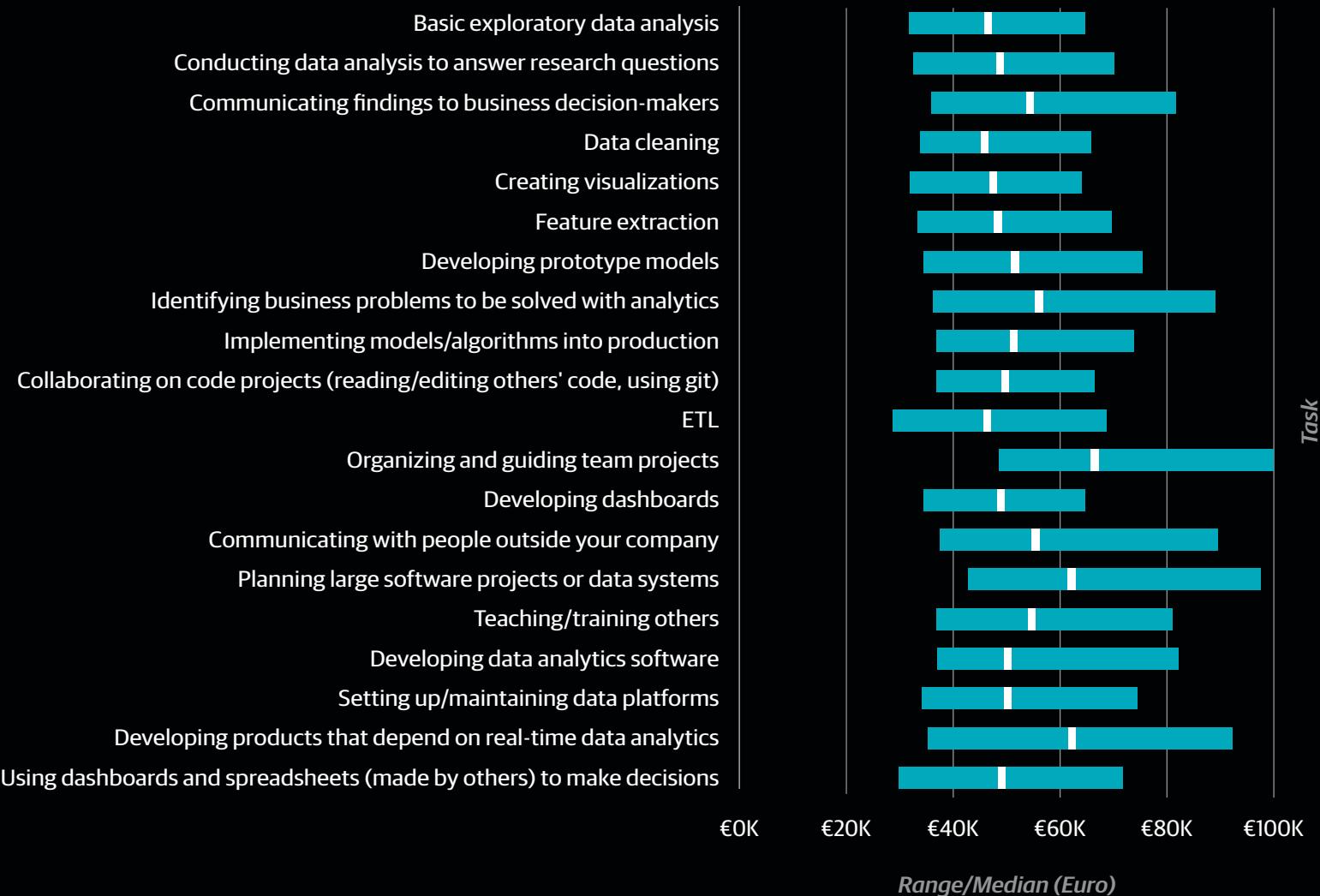
# TASKS

RESPONDENTS COUNTED IF THEY SAID THEY HAVE "MAJOR INVOLVEMENT" IN THIS TASK



# TASKS

## SALARY MEDIAN AND IQR\*



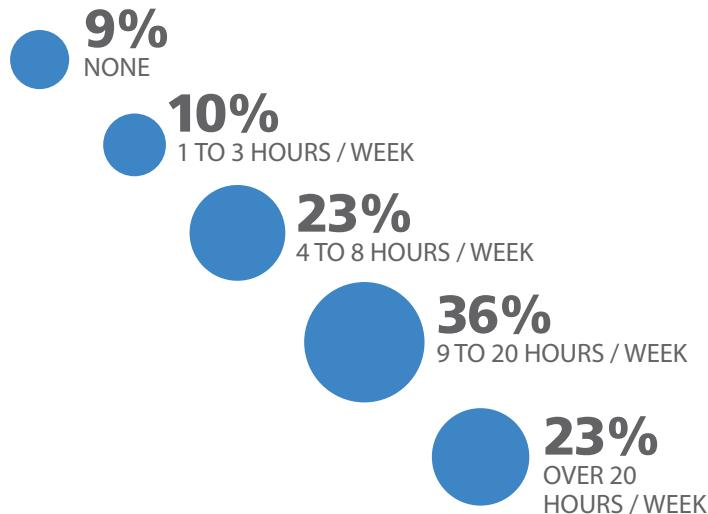
# Coding and Meetings

**FOR TWO BROADER TASKS,** coding and attending meetings, we asked respondents for more detail: namely, how much time they spend on them. As we have consistently seen, attending meetings correlates with salary: respondents who spend over 20 hours per week in meetings earn more than those who spend 9–20 hours, who in turn earn more than those whose spend 4–8 hours per week in meetings, and so on. This is unlikely to be a direct causal relationship, but rather both are effects of a shared cause (such as working in management).

As for coding, the highest earners were those who don't code at all, but that's because they tended to be managers. There is a dip in salaries among respondents who code over 20 hours per week, but this is explained by the fact that this group was, on average, less experienced than the rest of the sample. Within the middle groups—those who code 1–20 hours per week—there was not much variation in pay.

## TIME SPENT CODING

### SHARE OF RESPONDENTS

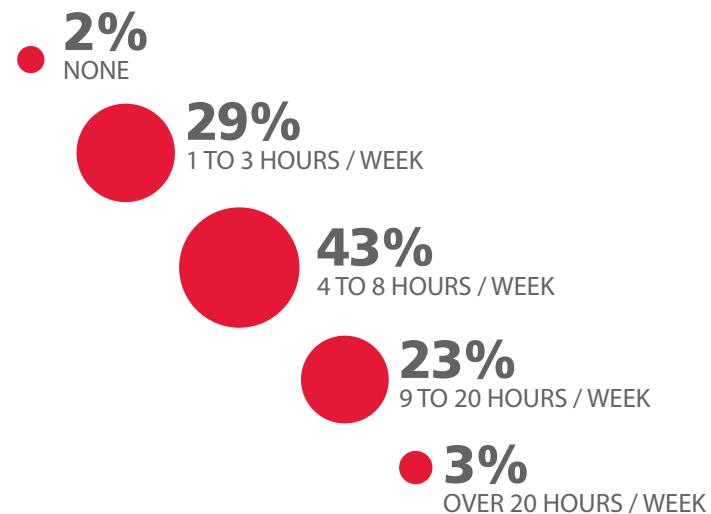


### SALARY MEDIAN AND IQR (EUROS)

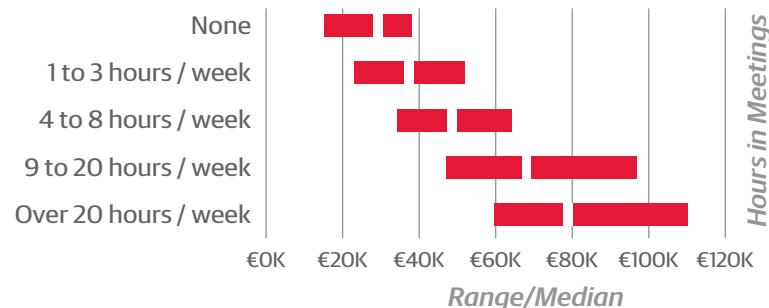


## TIME SPENT IN MEETINGS

### SHARE OF RESPONDENTS



### SALARY MEDIAN AND IQR (EUROS)



# Salary Change

**AN ALTERNATIVE METRIC TO CURRENT SALARY** is the amount that one's salary changed in the last three years. Most respondents' salaries grew at least a little in the last three years, and about a third of the sample saw their wages rise by 50% or more over this period. This latter group tended to be less experienced, with an average of 4.4 years of experience (compared to 7.6 years among those whose salaries did not grow by 50% or more).

For Spark/Hadoop and Python-only users, we use the tool-defined groups from page 8. They were most likely to have had 50% or more wage growth (40% and 44% of them did, respectively). Respondents who did not use Hadoop, Python, or R (the "SQL/Excel" group) were the least likely: only 19% of them reported a 50% rise in their salaries.

Most respondents' salaries grew at least a little in the last three years

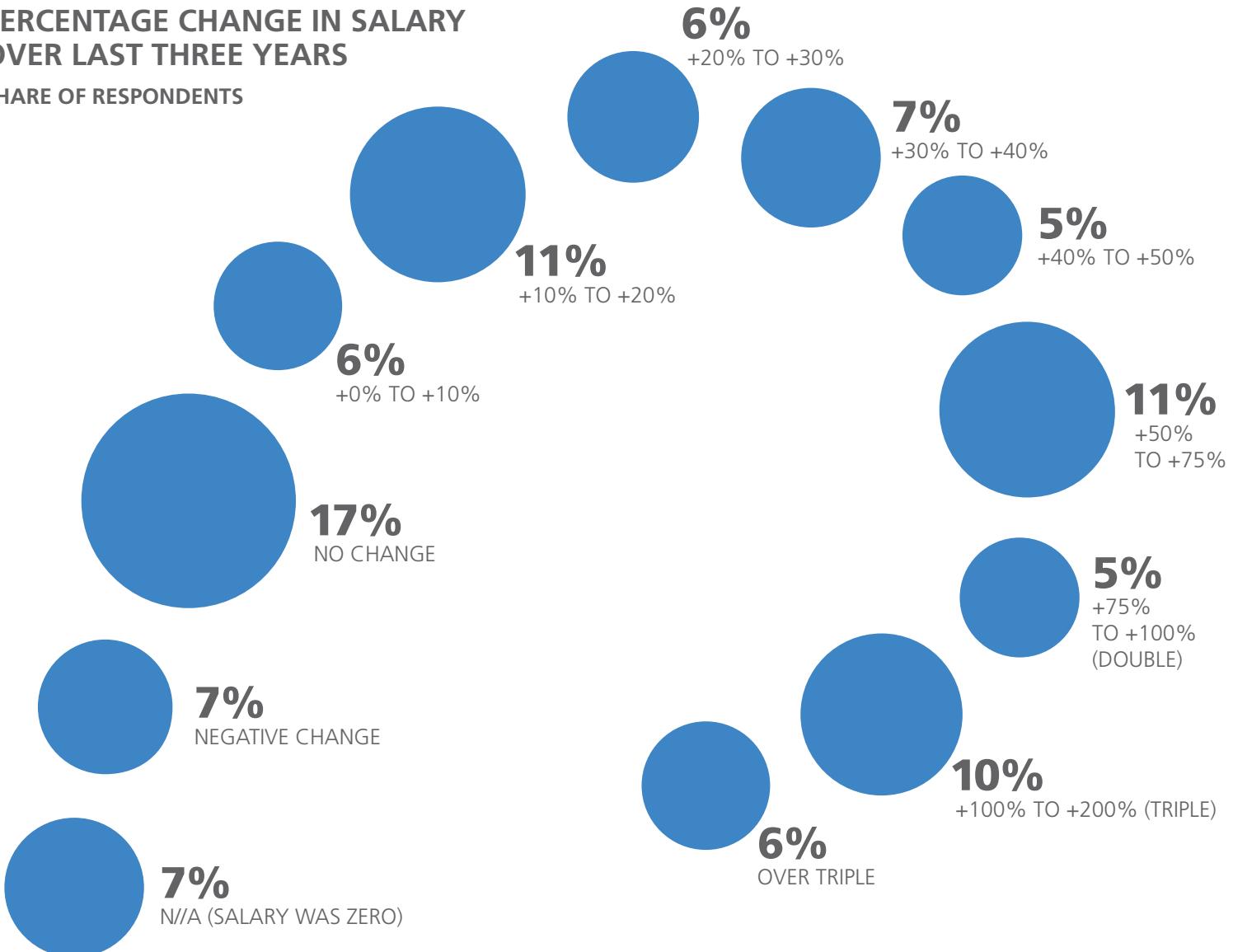
A final question asked respondents about the next step they would like to take in their career. The top response was "learn new technology/skills" and respondents who gave this answer tended to be less experienced (5.5 years on average) and have smaller salaries (€40K median) than the rest of the sample.

Respondents who said they would like to move into leadership roles had salaries far above average (€65K median). The other top

responses were "work on more interesting/important projects," "switch companies," and "start your own company". Respondents who work in the healthcare industry were far more likely to choose "switch companies" (33%) than respondents from other industries (11%).

## PERCENTAGE CHANGE IN SALARY OVER LAST THREE YEARS

SHARE OF RESPONDENTS



# Conclusion

**THE PURPOSE OF OUR SALARY SURVEYS** and the reports based on them is to provide an annual, data-driven snapshot of how much professionals in your field make, and to expose details of their work and career. There are plenty of resources out there that can give an idea of how much a data scientist can expect to earn or which software tools are on the rise, but there aren't many places where these data points are integrated into one report.

This information isn't just for employees, either. Business leaders choosing technologies need to consider not just the

software costs, but labor expenses as well. We hope that the information in this report will aid the task of building estimates for such decisions.

Business leaders choosing technologies need to consider not just the software costs, but labor expenses as well.

we will be able to explore the data science space in Europe. Thank you!

If you made use of this report, please consider taking the online survey. Every year, we work to build on the last year's report, and much of the improvement comes from increased sample sizes. This is a joint research effort, and the more interaction we have with you, the deeper

# We need *your* data.

To stay up to date on this research, your participation is critical. The survey is now open for the 2017 report, and if you can spare just 10 minutes of your time, we encourage you to take the survey.

[oreilly.com/ideas/take-the-2017-data-science-salary-survey](http://oreilly.com/ideas/take-the-2017-data-science-salary-survey)

Take the Survey 

How do data science salaries for people in Europe compare to their counterparts in the rest of the world? Among the more than 1000 people who responded to O'Reilly's 2016 Data Science Salary Survey, 359 live and work in various European countries as data scientists, analysts, engineers, and related professions.

This report takes a deep dive into the survey results from respondents in various regions of Europe, including the tools they use, the compensation they receive, and the roles they play in their respective organizations. Even if you didn't take part in the survey, you can still plug your own information into the survey's simple linear model to see where you fit.

#### With this report, you'll learn:

- **How salaries vary by country and specific regions in Europe**
- **The average size of the companies respondents work for, according to region**
- **How a respondent's salary is affected by their country's gross domestic product**
- **The type of industry they work for, including software, banking and finance, and retail and ecommerce**
- **Which tools are most commonly used vs the tools used by respondents with above-average salaries**
- **The major and minor tasks that respondents perform**

---

**John King** is a data scientist at O'Reilly Media. Having previously worked on survey-based sociolinguistic research in the Republic of Georgia, he now runs surveys at O'Reilly, using the results not just for internal use but also to share his findings with the public.

**Roger Magoulas** is Director of Research for O'Reilly Media.

ISBN: 978-1-491-97750-7



9 781491 977507

To stay up to date on this research, your participation is crucial. The survey is now open for the 2017 report; please take just 5 to 10 minutes to participate in the survey [oreilly.com/ideas/take-the-2017-data-science-salary-survey here](http://oreilly.com/ideas/take-the-2017-data-science-salary-survey).

