

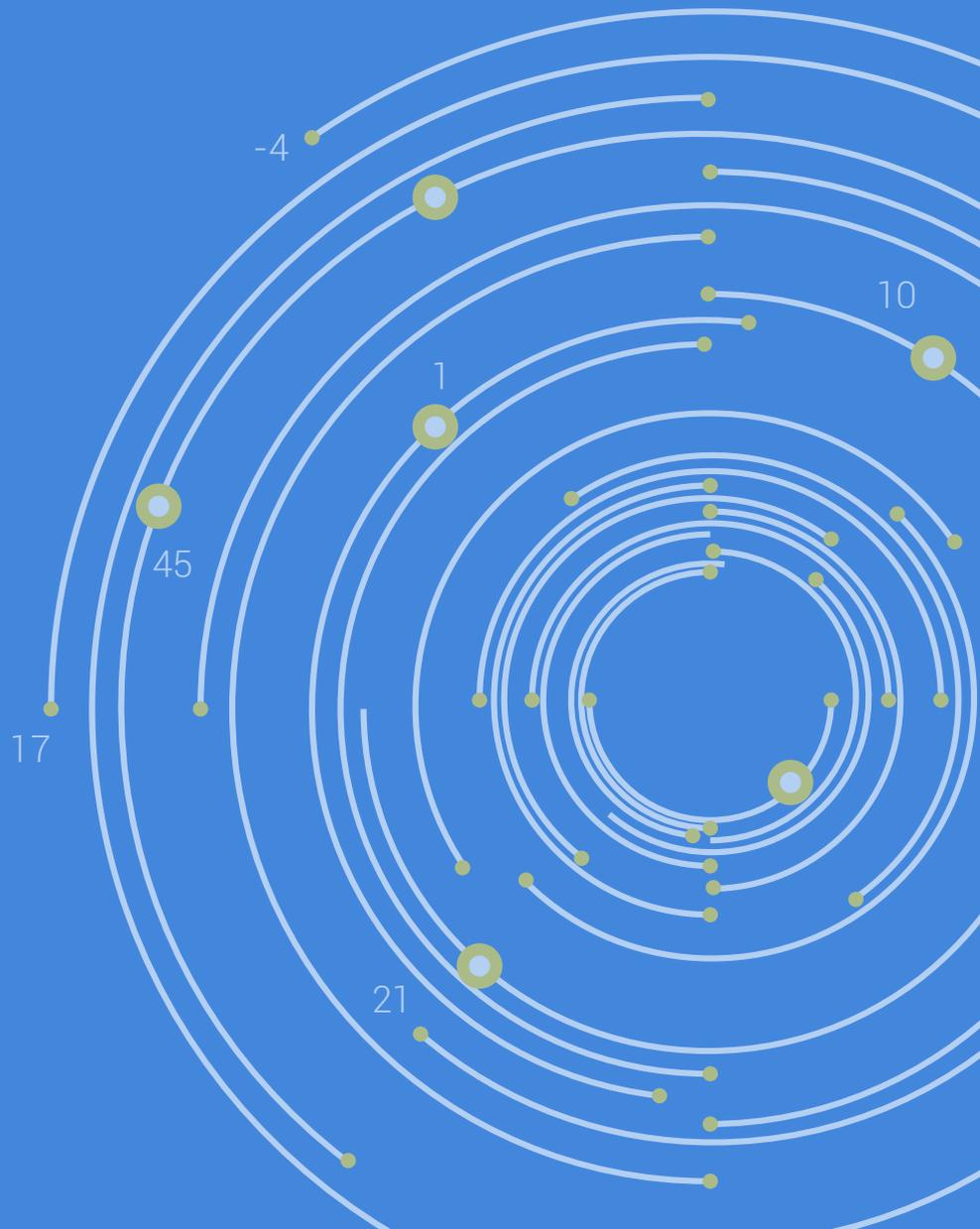
Arab  
Data  
Journalists'  
Network

# Data Journalism

## In The Arab World

Survey results

July - October 2017



Published in 2017 by the Arab Data Journalists' Network,  
21 Dareh Saed, Al Inshaa WA Al Munirah, El-Sayeda Zainab, Cairo  
Governorate, Cairo, Egypt  
© ArabDJN, 2017



Original title: *Data Journalism in the Arab World*

The ideas and opinions expressed in this publication are those of the participants; they are not necessarily those of ArabDJN and do not commit the company.

Editorial coordination: Jihad Al Shebeni  
Graphic design, Infographics: Henar Khater



This publication was supported by Arab Data Journalists' Network



## Survey Overview

This report summarizes the results of the survey Arab Data Journalists Network has conducted about data journalism in the Arab world, during the first three months after its launching.

Arab Data Journalists Network is a platform about all that is related to Data Journalism all over the world. It aims to provide Data Journalists in the Arab world with all that they might need, old and new.

The survey aims to measure how much Arab data journalists use data in their stories and how easy or difficult it is for them to get data from official sources, and to be conducted again in the future for a deeper and wider look at the situation.



**60** Applications



**08** Countries



**03** Months

# How the Survey was Conducted?

The survey was administered by the ArabDJN from July to October 2017. The form for participating in the survey was published via the official page of the network on Facebook.

The questions were grouped by the country of the participants, years of experience, the means participants work in whether it was printed or online, how do they get data, what are the tools they use to extract data, whether they have ever used to the right to access to Information law, and other.

Created by



Published on



Shared via



## Description of Sample

All data journalists in the Arab world were welcomed to participate in the survey, 60 people from Egypt, Morocco, Jordan, Tunisia, Iraq, Syria, Lebanon and Amman participated in the survey.



## The Panel of Experts

Upon reviewing the results of the survey, both Eva Constantaras and Amr Eleraqi, two journalists specialized in data, proposed some recommendations for both professional and aspiring data journalists.

Eva Constantaras is an investigative data journalist and consultant specialized in data journalism projects to combat corruption and foster transparency. She has run data journalism activities for USAID, UNDP, the World Bank, Hivos and Google. Her reporting has appeared in media outlets including the Seattle Times, El Mundo and El Confidencial.

Amr Eleraqi is a data journalist, and a trainer specialized in data journalism and data visualization. Amr is the CEO of InfoTimes and the founder of Arab Data Journalists Network.

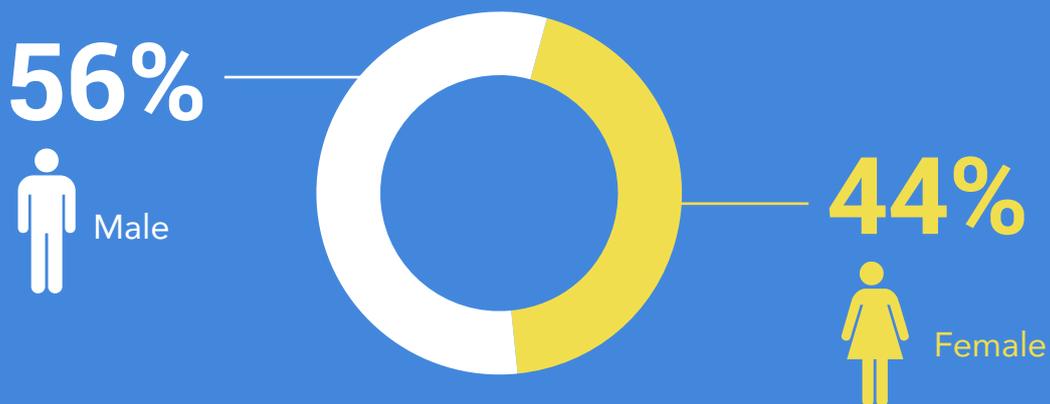
Amr and Eva during the discussion of data journalism survey

*Photo Credit: Tahrir Lounge*

# Demographics:

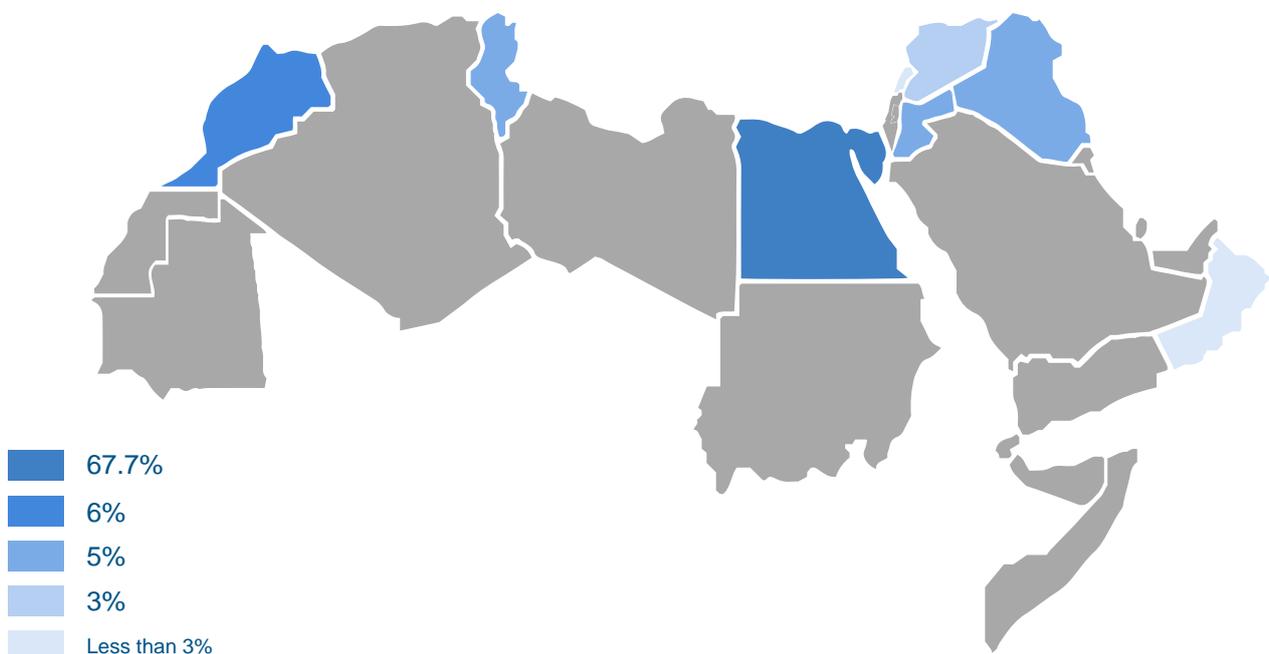
## ▲ Gender of The Participants

56% of the participants were males, while 44% of them were females



## ▲ Nationalities of Participants

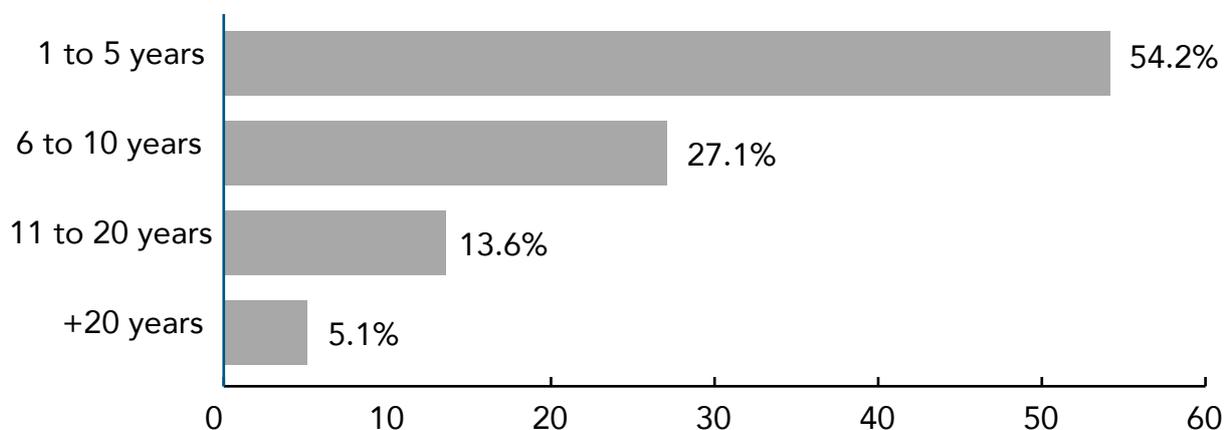
The participants were 67.7% Egyptians, 6% were from Morocco, 5% were from Jordan, 5% were from Tunisia, 5% were from Iraq, 3% were from Syria, while the rest were from Lebanon and Amman.



# Demographics:

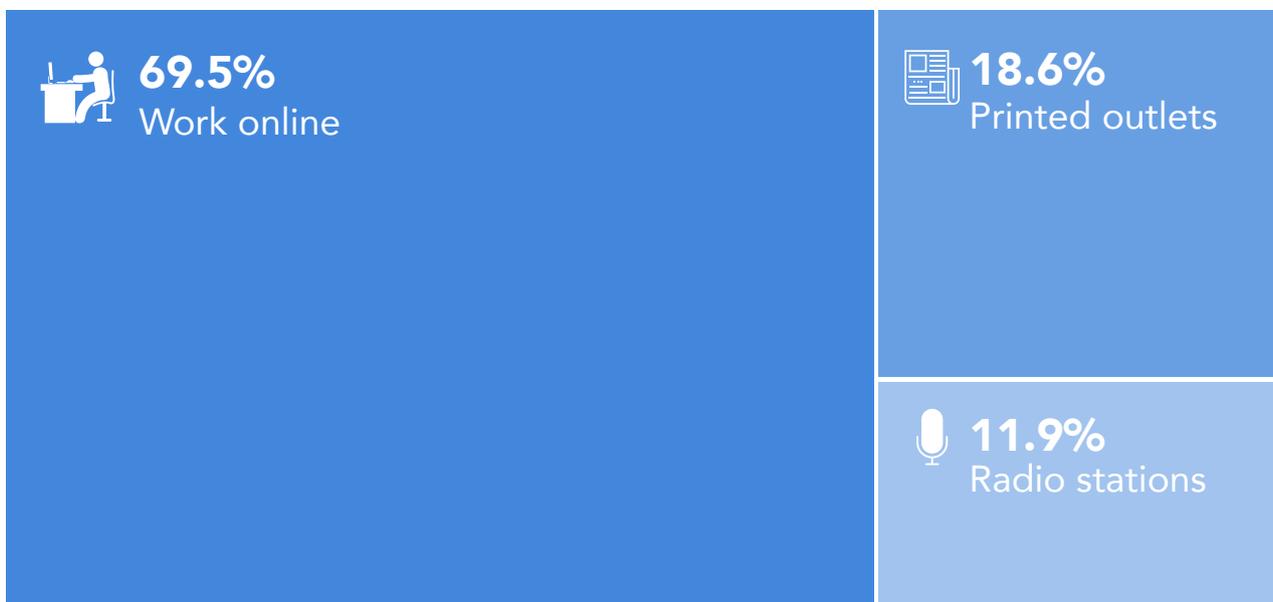
## ▲ Participants Years of Experience

54.2% of the participants have a work experience between one and 5 years, while 27.1% of them have an experience ranging between 6 and 10 years. Those who have an experience from 11 to 20 years experience had a share of 13.6%, while the rest were more than two-decade experienced.



## ▲ Means of Communication

69.5% of the participants indicated that they basically work online, while 18.6% pointed out that they work in printed outlets. The radio had a share of 11.9% of the participants.

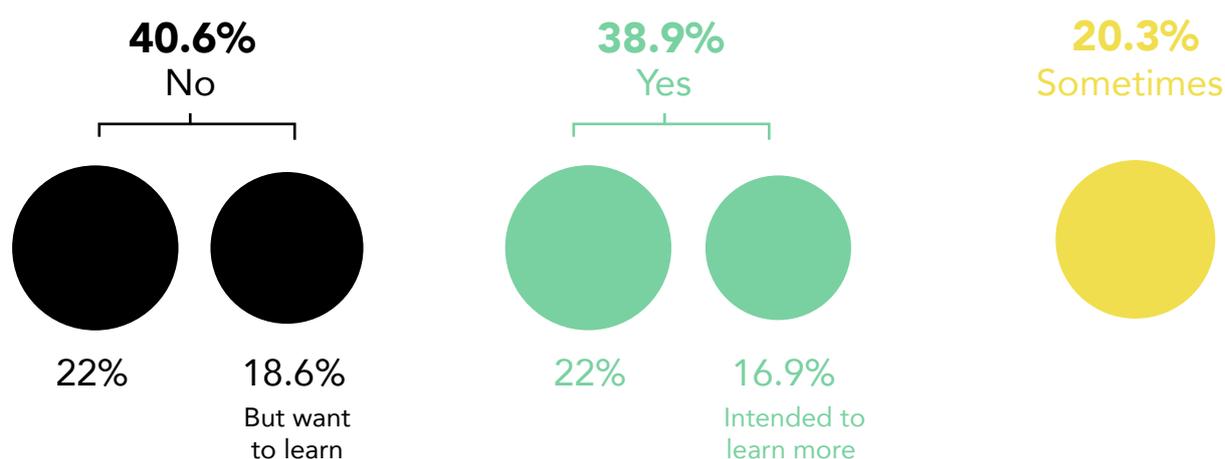




# Key Findings

## 1 Do you use Excel in the preliminary analysis?

The answers “yes” and “no” got an equal percentage of 22% for each. 18.6% of the participants indicated that they want to learn how they can use it in the analysis, while 16.9% illustrated that they use it in the analysis but they need to learn more. As for the rest 20.3%, they answered that they “sometimes” use Excel in the analysis.



*“Learning to use some tools is not so important as learning to use Excel in analyzing data; the most important brick, on which the story is built,” she continued.*

*“The longest part in the process of producing a data-driven story is the analysis part. A big part of this analysis can be done using Excel, as it helps in showing the data in an organized way, and in identifying the similarities and the differences between the multiple sets of data,” Amr Eleraqi indicated.*

*“Excel also allows journalists to analyze relations and networks, as it is the best method, by which journalists can dig into the data,” Amr added. “Uploading data to Excel can be very similar to meeting a source for an interview”.*

*“On the other hand”, Amr stated, “If journalists feel that Excel is hard to be dealt with, they can use Google Spreadsheets instead. It is pretty much the same as Excel, but it does the same job with a much easier interface.”*

# Key Findings

## 2 What is the most common file type, through which you get the data?

Asking the participants about the most common file type, through which they get the data, we find that the PDFs appeared among the answers with a percentage of 55.9%, followed by Word with a 22%, then Excel with a 20.3%, while the rest went for images.



55.9%  
Portable Document Format



22%  
Microsoft Word Files



20.3%  
Microsoft Excel Files



1.7%  
Images and Scanned Papers

### FACT BOX

*Government releasing data in PDF tends to be catastrophic for Open Government advocates, journalists and our readers because of the amount of overhead it takes to get data out of it.*

*When a government agency publishes its data and documents as PDFs, it makes us Open Government advocates and developers cringe, tear our hair out.*



### Nothing was created as PDF in the first place

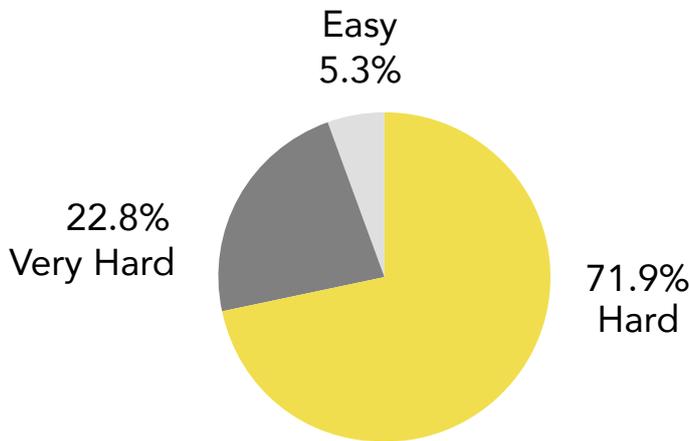
*Regarding the PDF file type, through which journalists usually get data, Constantaras commented that journalists should determine the file type, through which they want to get the data, in their request to obtaining data. "Any PDF file was initially created through a machine readable file type," she stressed.*

*"In case journalists failed to get the data in the format they wished, they could contact some of those who are behind some online free tools as Tabula, asking them for help in extracting the data and in modifying the files in a way that makes it easier to be dealt with. Those who are working on these sites want to help," the expert added.*

# Key Findings

## 3 How hard it is for you to find available official data in your country?

The answer of "hard" had the biggest share, "very hard" had a share of 22.8%, while "easy" had the rest percentage of participations. The answer **"very easy" didn't get one vote.**

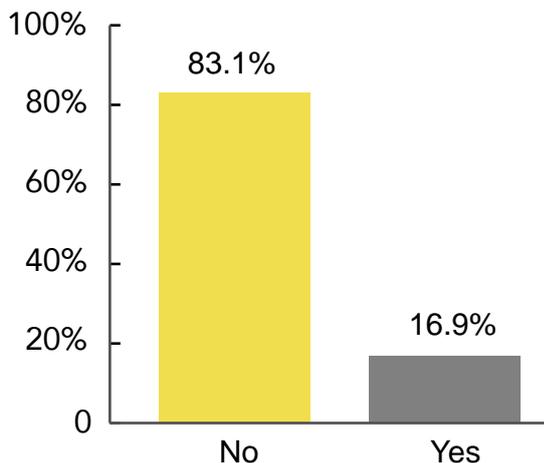


### FACT BOX

Nearly 100 countries have laws on the books granting individuals a general right to access information held by public bodies, and imposing an obligation on public bodies to proactively disclose key types of information.

Jordan had passed a FOI law in 2007, a Decree-Law on freedom of information was enacted in 2011 in Tunisia. In turn, Yemen passed an FOI law in July 2012. Moreover, the new Constitution adopted in Morocco in July 2011 includes a provision guaranteeing access to public information.

## 4 Have you ever used the Right to Access to Information law?



83.1% of the participants answered that they never used it, while 16.9% indicated that they did.

# Key Findings



## As for the lack of data

Constantaras demonstrated that journalists have a role to play regarding this matter, which is asking for the data they want. “If journalists don’t ask for data, they won’t get it. If they ask for it and actually get it, they will help in making this data available, journalists shouldn’t wait until the data is released, they should work for it.” Constantaras stated.

Regarding some of the stories that might lack data, Constantaras pointed out that journalists should work on the stories, about which they have data, not the ones they wish they had had data about. She also stressed that journalists should be realistic when collecting data and pay attention to the amount of data they are hoping to collect.

Answering a question about alternative methods to collecting data, in case the needed data is not available, the expert illustrated that “surveys might be tricky, as it might not be very accurate sometimes.” Constantaras also stated that taking a look at the surveys carried by universities might be very important.

## From information into data

“Lacking to an effective law that guarantees getting first-hand data, we should, as journalists, turn to secondary sources. We live in the time of leaks, big data and social media, and I believe that taking the other direction, from information into data, is possible.” Amr indicated.

“Collecting news on sexual harassment, for instance, with details about the time, the place, when and where the incidents took place, in addition to some information about the type of harassment and the age of the girl being harassed, might lead to creating a database about sexual harassment in Egypt.”

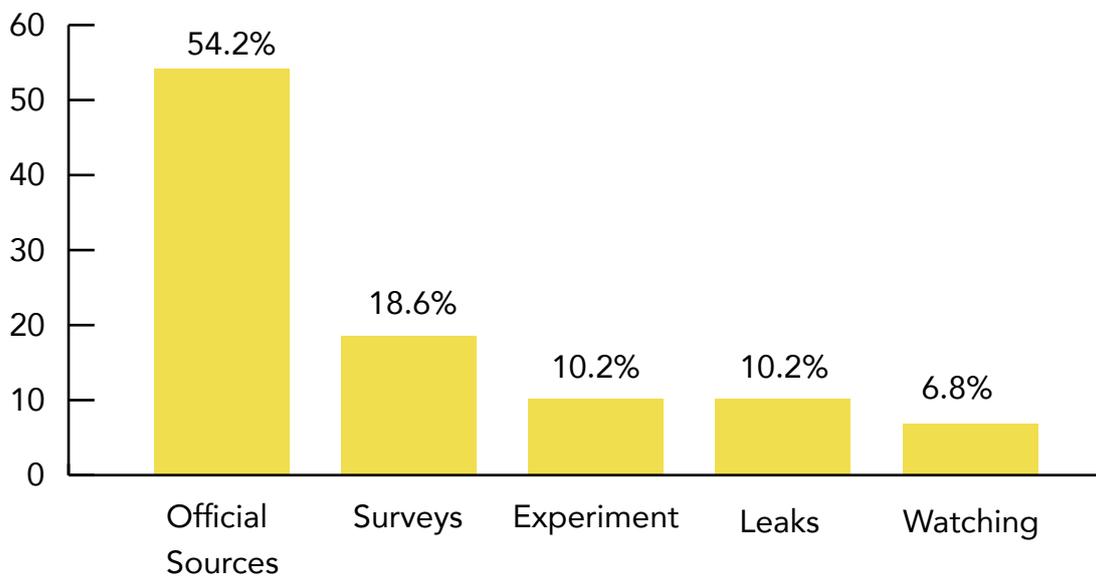
Eleraqi added that through such a database, journalists can do some analysis and reach conclusions about sexual harassment in Egypt. “The number of cases collected is not going to be inclusive for sure, but it is going to be somehow inclusive regarding what was published in the news.”

Amr indicated that relying on such source is not going to provide accurate results, but it is going to be an indicator. “We can use this method with different issues, such as: terrorism, murder and human rights violations. The absence of official data must not hold us back, it should push us forward finding other ways to produce more stories. This might put governments in a delicate position, pushing for more available data.”

# Key Findings

## 5 What source of information do you mostly depend on for getting data?

Among five, official sources were on top with a 54.2% as an answer to: “what source do you mostly depend on to obtain data?” Surveys came next with a 18.6%, followed by “experiment” with a percentage of 10.2%, while “leaks” came in fourth place before “watching” that came last.



### Data is about people

Replying to what some say about data journalism that it is mainly about numbers, Constantaras explained that data journalism is about people not about numbers. “Producing a data-driven story should be relying on what the journalist believes that readers need to know about,” she illustrated.

“For example, readers might want to know which school is better for their children? Public or private? Hence, journalists can decide what kind of information they might need, and then they can cluster, analyze and reach outcomes that could help readers make a decision about which school they should enroll their children in. Readers don’t need to read numbers, they need to see how these numbers are affecting their lives.”

# Key Findings

## 6 Can you extract data from websites?

83.1% of those who participated in the survey indicated that they can't extract data from a website, while 16.9% answered that they can.



## 7 What are the used tools in extracting data?

Those who answered that "yes," they can extract data from websites illustrated that the tools they can use for this matter are Google Sheet, Tabula, Excel, Openrefine.



*\*The answers showed that these tools were mentioned one time only.*



Amr noted that "some of these tools are not for data extraction, something that shows that participants who used these tools may not be aware of the purposes, which for these tools are use. Mentioning Open Refine and Excel, as an answer to this question might be reflecting that the definition of extracting data is not clear for those who participated in the survey."

# Key Findings

## 8 Have you ever used a free tool in visualizing?

50.8% of those who participated use free tools to visualize, while the rest that forms a percentage of 49.2% pointed out that don't use any free tools.

## 9 Mention some of these free tools.

Piktochart, Infogram and Google Sheets were mentioned with a percentage of 33% among other tools, with a percentage of 11% for each. As for Canva, it had a 5% of the suggestions, while Excel and Tableau were suggested with a percentage of 3% for each.

|                   |                 |                  |                         |                  |
|-------------------|-----------------|------------------|-------------------------|------------------|
| Google sheet<br>7 | Piktochart<br>5 | photoshop<br>1   | Google data studio<br>1 | Timeline.js<br>1 |
|                   |                 | Datawrapper<br>1 | Atlas<br>1              | Word<br>1        |
| Infogram<br>6     | Canva<br>3      | Tableau<br>1     | Power point<br>1        | Easel.ly<br>1    |
|                   | Excel<br>2      | RAW<br>1         |                         |                  |

# Key Findings



## Not a big problem

*Lacking the skills that might be considered an obstacle to becoming a data journalist is not a big problem from Constantaras's perspective. "Journalists should start thinking about the problem they want to tackle, not the tools they want to use in producing the story and solving the problem," Constantaras recommended.*

*Constantaras explained that the reluctance by some journalists to work in data journalism goes back to the assumption that it might be more difficult than traditional journalism. "The first data-driven story is the longest, as journalists wouldn't be aware of the tools they should use to produce the story," she indicated.*

*"It gets easier with each passing time, as journalists become aware of what they might need to produce data-driven stories. Producing stories repeatedly helps journalists in becoming professional over time," Constantaras added.*

## 10 Do you know of a place that offers data journalism courses?

Asking the participants whether they know of a place that offers trainings on Data Journalism in their countries, 62.7% answered that "yes," while 37.3% answered that "no".

Asking them about the places that offer trainings on Data Journalism, InfoTimes name was mentioned 22 times, Tahrir Lounge was mentioned 9 times, while the Center For Defending Freedom of Journalists, Arab Data Journalists Network, Friedrich Naumann Foundation and the African Centre for the Training of Journalists and Communicators were mentioned once.

# Key Findings



## **Self-taught data journalists**

*Eva Constantaras pointed out that many educational programs specialized in data journalism started to be available over the past few years, in consistence with the spread of data journalism.*

*“However,” she added, “Earliest data journalists taught themselves, and so can anyone wants to become a data journalist. Tutorials are open and knowledge is available.”*

General

# Recommendations

The recommendations mentioned below are advices, given by data journalists experts, that were collected by the editorial team from different websites such as GIGN, Poynter, Ojo Publico... and others.

## Text of Austria-Hungary's Declaration of War.

VIENNA, July 23.—Austria-Hungary's declaration of war against Serbia was passed here late this afternoon. The text is as follows:—  
The Royal Government of Serbia not having replied in a satisfactory manner to the note received in it by the Austro-Hungarian Government on July 23, 1914, the Imperial and Royal Government itself compelled to proceed itself to safeguard its rights and to have recourse for this purpose to force of arms. Austria-Hungary considers itself, therefore, from this moment at war with Serbia.  
Count "COUNT BERCHTOLD,  
Minister of Foreign Affairs of Austria-Hungary."

## Russian Government Announces It Wish to Remain at Peace Yet Is Determined to Guard Its Interests.

MOSCOW, July 23.—The Russian Government tonight announced the following official communication:—  
The Russian Government desires to remain at peace and to continue its friendly relations with all nations. It is, however, determined to guard its interests and to defend its rights. It is, therefore, determined to remain at peace, but it is also determined to defend its rights and to guard its interests. It is, therefore, determined to remain at peace, but it is also determined to defend its rights and to guard its interests.

## CEAR'S FORCES MASS ON EASTERN BORDER

His Capital Expects War and Counts Confidently on England's Aid.

## MOBILIZATION ORDER READY

Former Official Says It Would Be Issued in Case of War.

Data Journalism  
In The Arab World

# General Recommendations

- ▲ Data on its own doesn't tell a story: We still need journalists to decide how to arrange, analyze and visualize data, as well as to provide appropriate context and interpretation.
- ▲ Whatever your background is, you'll need to consider it in your quest to acquiring the three skills you need to become a data journalist, that are journalism skills, design talent and coding acumen. Mainly, it's just about teaching yourself.
- ▲ Great data projects don't generally begin with great data sets. They begin with great questions and the desire to find the hardest evidence available to answer those questions.
- ▲ You can get the data just by asking, and they will email them to you. Else you can download them from the web, scrape them or get them using FOI-requests (Freedom Of Information). Develop and tune these methods, just as you do with interviews. All are journalistic methods.
- ▲ Government agencies are slowly adapting to providing data in excel format. Always try to obtain information in a readable format, if not, you may have to resort to entering in the data manually.
- ▲ Much data journalism is too much data and too little journalism. We do this for stories. If there isn't a story, don't tell it. Find the stories and tell them one by one. Using interactive graphics is the same. Break it down to one clear story for each graphic or map. And tell it, so it is possible to get it. Don't mix everything together in longreads, with too many angles.

# General Recommendations

- ▲ There are nearly always error in data. Also when you get it from estimated authorities. You have the responsibility to get rid of the errors, before presenting it to the audience. Count totals – check if all is included or you miss something.
- ▲ Staying updated on new available digital tools and the most innovative ways people are using the technology to enhance journalistic research, training and ongoing exploration is part of the team's daily routine. Following data blogs, such as these from The Guardian, Nación Data and ProPublica, is as necessary as participation in hackathons, online courses and local and international workshops.
- ▲ If you don't know spreadsheets, learn. Search for help/tips online. Give yourself plenty of time when first getting started. Don't do it on deadline. You will make mistakes.
- ▲ Data Journalism is most often about humans. Get rid of the numbers and find the humans, who are the best examples of your data. Actually sometimes zero numbers are the best way to do a data journalism story. You then might only include the data in a graphic. Or keep them for later. If you can't find humans to explain your findings, they might be wrong.
- ▲ There is a great history of sharing by the people in the data journalism-community. This is the only way we can keep up with the speed of the development of tools and methods. This way of sharing is more common here than in perhaps any other area in journalism.