

OpenBudgets.eu: Fighting Corruption with Fiscal Transparency

Project Number: 645833

Start Date of Project: 01.05.2015

Duration: 30 months

Deliverable D5.3

User Requirement Reports Final Cycle

Dissemination Level	Public
Due Date of Deliverable	Month 20, 31.12.2016
Actual Submission Date	
Work Package	WP5, Test Beds and Evaluation - Journalism
Task	T5.1, T5.2, T5.3
Type	Report
Approval Status	pending
Version	
Number of Pages	42
Filename	Deliverable 5.3 H2020 OpenBudgets.eu

Abstract: This document describes several gaps, problems and roadblocks experienced by journalists when accessing budget data. Uses cases, workshops and a curriculum are examined in order to resolve those gaps.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability.

History

Version	Date	Reason	Revised by
1	16 December 2016		

Author List

Organisation	Name	Contact Information
J++	Nicolas Kayser-Bril	nkb@jplusplus.org

Reviewer

Organisation	Name	Contact Information
Fraunhofer IAIS	Alexandra Garatzogianni	Alexandra.Garatzogianni@ iais.fraunhofer.de

Executive Summary

This work package researched how journalists use budget data in their daily routines and in investigations. We interviewed dozens of journalists, carried out several case studies and ran an experiment in the perception of large amounts.

The findings shed light on how journalists use and could improve their use of budget data. We showed consistently and convincingly that budget data cannot be automatically analyzed in a journalistic setting. We showed, in an experiment, that readers understand large amounts better when presented in plain text, as opposed to comparisons, which is of tremendous importance for tools that aim at automating comparisons. We showed that what journalists need is expert knowledge in accounting and the budgeting process, which is why this work package increased its focus on trainings, tutorials and tools to better grasp how public budgets are made and used.

Table of Contents

Executive Summary	3
Table of Contents	4
Introduction	5
Working with large amounts	6
Problem description	6
Experimental setup	8
Results	12
A professional issue	16
The relationship to the norm	17
Implications for tool-centric approaches	19
How budget data can help journalists	21
The very small role of budget data	21
Going from data to the story	21
Finding the data involved in a story	23
The exception of direct payments	24
The unbridgeable gaps	25
Non-availability	25
Obscurity	26
Conclusion	27
References	29
Annex 1: Answers to the Order of Magnitude Guesser	30

1. Introduction

Newsrooms the world over have trouble making sense of numerical information. Matt Waite, a journalism professor formerly senior news technologist for the St. Petersburg Times of Florida, says that being bad at math is not only frequent among journalists, it is “a badge of honor”, “woven into the very fabric of identity as a journalist” (Waite, 2013).

This state of affairs is of utmost concern when covering stories that involve public budgets. Not only do budgets deal with large amounts, their understanding also requires deep technical knowledge of public accounting practices. In such conditions, how can journalists determine the newsworthiness of a budget-related story? How can they hierarchize information based on numbers that very few of them understand?

While journalists are the focus of this work package, let us not forget that they are one group among many others where knowledge of basic arithmetic is greatly lacking (Pro Bono Economics, 2014). Politics, public relations and police work, to name a few, could also benefit from improved numeracy, but analyzing their relationship to public budget data goes beyond the scope of this work.

Examples abound of the objective failure of journalists to accurately determine the newsworthiness of stories involving budget data. In France, for instance, hundreds of news articles were published on the expenditures of the head of a public organization who claimed €40,000 in taxi costs in just a few months (see for instance Laubacher, 2015). Even more were written about what came to be known as the “Coiffeurgate”, when it became known that the French president had a personal hairdresser paid nearly €10,000 a month (Begley, 2016). Very few articles were written over stories that involved much larger mismanagement of public funds. An embezzlement scheme worth over €10 million (250 times more than the misspending in taxi costs) was unveiled at one French university but received almost no media attention (Delaporte, 2014). A search in the archives of Le Monde, one of the top French dailies, shows that the taxi story was mentioned in 35 articles, against 3 for the university story.

French journalists are by no means alone in prioritizing stories involving small amounts over stories involving larger ones. The Premier Minister of New South Wales, Australia, following outrage from journalists, resigned over a bottle of wine worth €2,000 he received as a gift and failed to declare (Whitbourn, 2014). In Sweden, Deputy Prime Minister Mona Sahlin resigned after it was made public that she bought a bar of Toblerone chocolate for personal use with her professional credit card (Svensson, 2007)¹. The interest in these stories and the impact they provoked were far higher than any of the other corruption stories ranging in the millions of euros that took place before, during or after these scandals in the same countries.

This discrepancy between media attention and the behavior of politicians undermines the trusts citizens can have in the institutions that provide the framework for public life. Because politicians who find themselves accused of minor misdemeanors fall under heavy media criticism while others who embezzle millions - sometimes billions - keep a pristine reputation when seen through the prism of most media outlets, citizens who are aware of the larger stories cannot trust the media to play a role of counterweight to political power. One of the goals of this work package being to “augment citizens’ trust in public administration”, understanding the reasons that lead to journalists’ failing to provide a proportionate coverage of stories involving budget data is of utmost importance.

This work package explored why journalists fail to hierarchize information about public budgets and spending (part 2 - Working with large amounts) and how open budget data could help in improving their practice (part 3 - What budget data can do for journalists). Over the course of one and a half years, we conducted several case studies, ran experiments, analyzed hundreds of stories involving mismanagement of public funds and interviewed dozens of journalists and experts. This report summarizes our findings.

2. Working with large amounts

2.1. Problem description

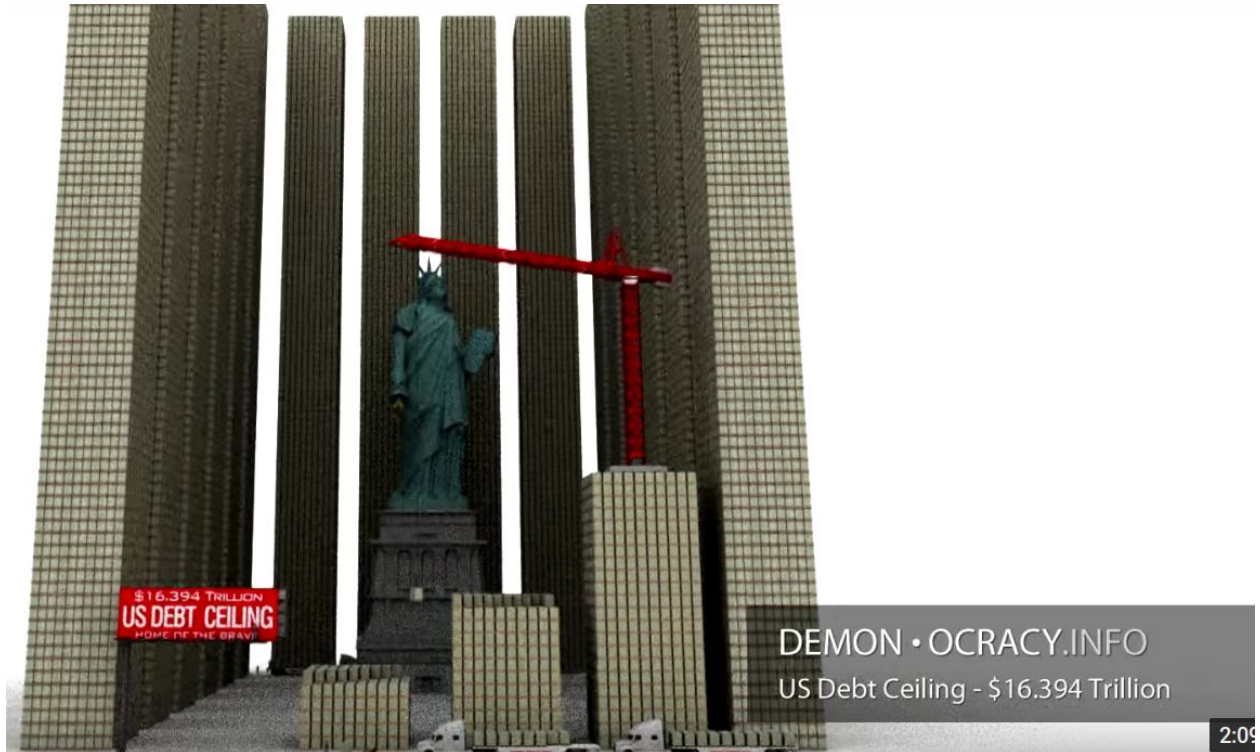
Budgets are made of large amounts, usually much more than what one person might encounter in her life, and use concepts unknown to most people. From this starting point, the question of how to communicate large amounts is of primary importance if one is to publish a story that involves budget data. That the City of London incurred a £18.7 million charge for the

¹ Truth be said, she bought more than Toblerone. The total amount to be repaid was about 5,000€.

amortization and depreciation related to its committees in 2015 will probably mean nothing to a vast majority of readers, who do not know what an amount close to 20 million pounds represents, nor what amortization stands for.

A common practice in journalism to communicate large amounts is to express them in another unit, which can be football fields for area or olympic swimming pools for volume. An Associated Press headline from 2015 informed us for instance that “Enough water to fill 1 billion Olympic-sized swimming pools has melted off of Alaska's glaciers over the past 5 decades” (Borenstein, 2015). While the origin of this practice has not been researched, examples abound. In 2002, a writer from the New-York Times told its readers that a farm in Uzbekistan used enough water “to fill 1,300 Olympic swimming pools” (Wines, 2002). Four decades before, his colleagues were already counting in football fields to describe the area of an office building (“10 football fields”, in New-York Times, 1962). While this work package could not research the archive of European newspapers for practical reasons, there is no reason to believe that European journalists did not apply the same techniques as their counterparts in the United States.

The development of datavisualization in the early 21st century kept the same trick with new tools. Demonocracy, an infographics studio specialized in economics, tried to make sense of the public debt of the government of the United States by visualizing it in \$100-dollar bills, for instance.



A screenshot from *Demonocracy* (2012).

Such approaches assume that, by comparing an unknown amount to a one supposedly known from the reader, the reader will be able to interpret the unknown amount correctly. This assumption - which was never checked - led to the creation of tools to speed up and multiply such comparisons, in the hope that readers would be able to make sense of large, unknown amounts. (The tools are discussed below.)

We designed an experiment to test this assumption and showed that people actually understand amounts much less well when they are expressed as a proportion of another amount.

2.2. Experimental setup

Users are given an interactive questionnaire, presented as a game called “Order of Magnitude Guesser”. This experimental setup is inspired by GeoGuessr and EarthQuiz (De Paor, Whitmeyer, Bentley, and Dordevic 2014), which were used to assess place location knowledge (Zhu et al., 2015).

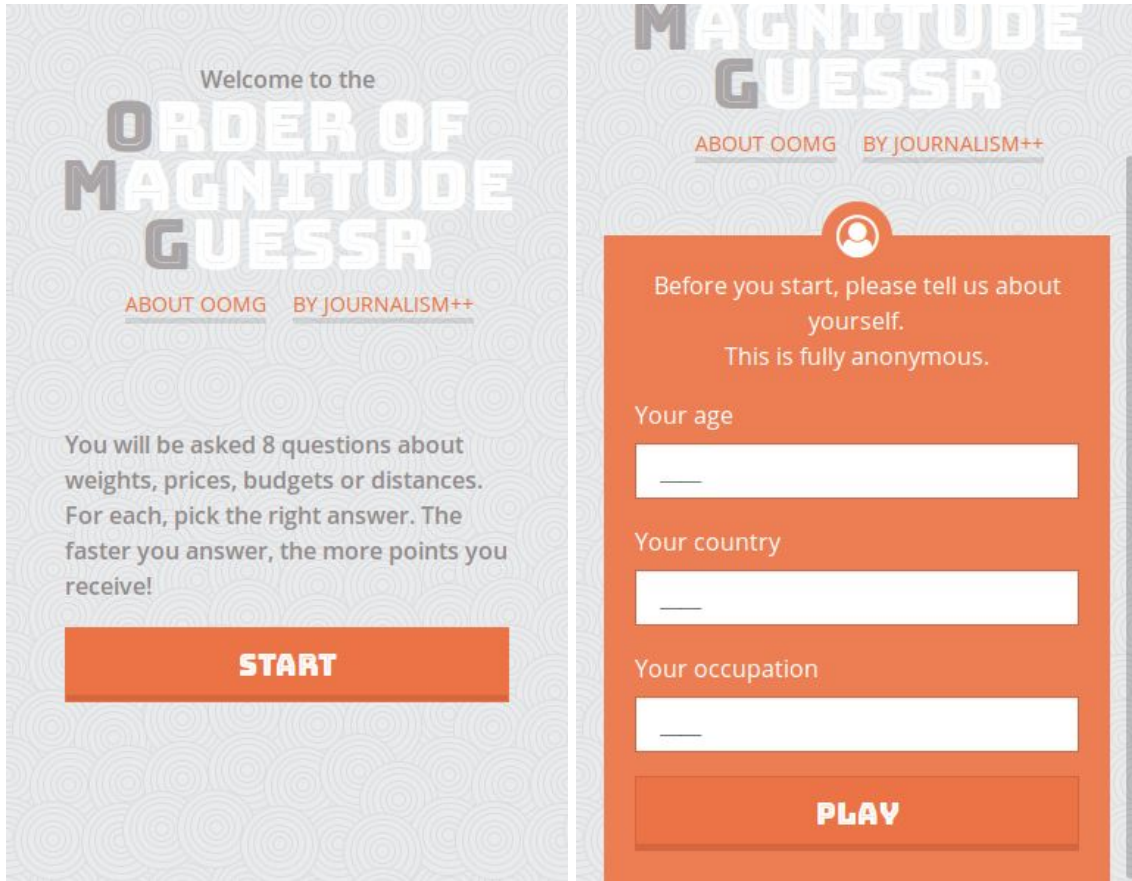
The goal is to correctly guess the order of magnitude of 8 questions. Questions ask orders of magnitude of daily life items and larger ones and provide three choices, the wrong ones being at

least one order of magnitude above or below the right answer. The answers can be phrased in three ways: Numerical (“€1,000,000”), written as text (“1 million”) or relative to a daily-life item (“the lifetime earning of an average employee”). The form of the question was selected at random. The amounts being compared to were a selection of 31 unique items. Out of the 41 questions, 32 of the comparative answers use items from daily life, such as the weight of an apple, the length of a car or the minimum wage.

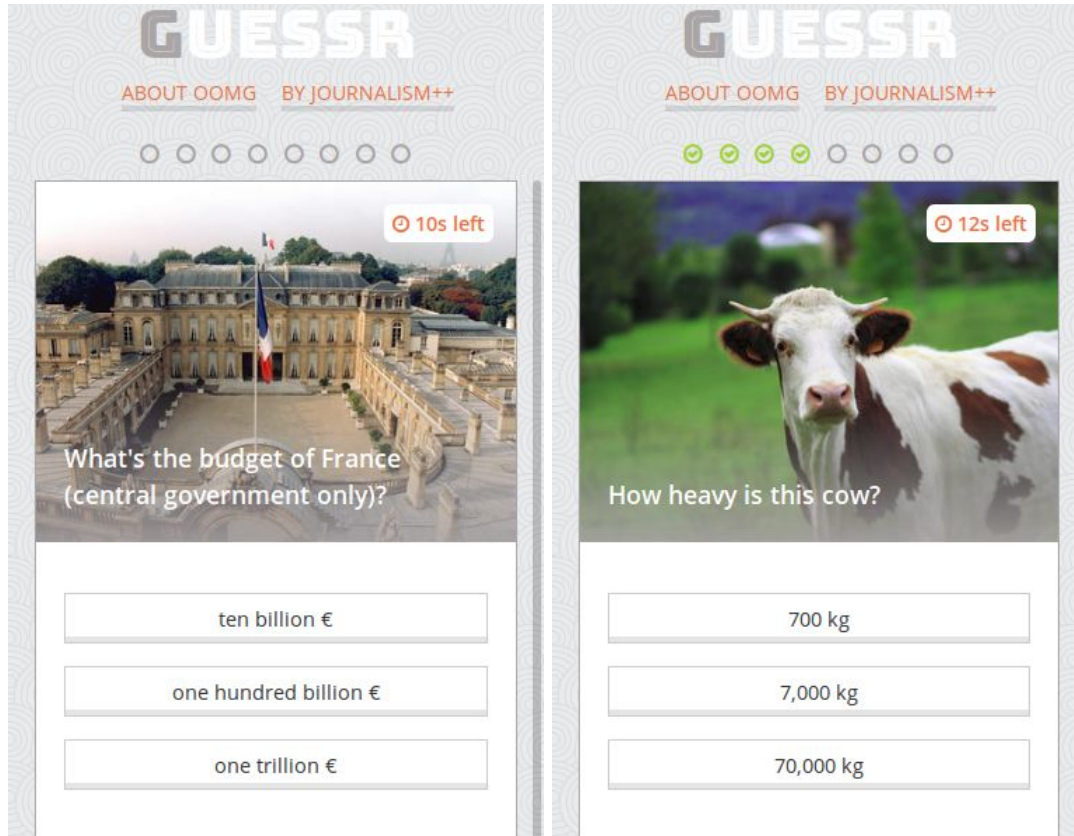
Each question must be answered in less than 15 seconds, to prevent users from using a search engine to find the answer and add excitement to the game. The game can be found at oomg.jplsplus.org.

At the beginning of the game, users are asked about age, country and occupation. The questions are optional and no personal information is stored (no information is asked or collected, such as cookies or IP addresses, that would let us connect the answers to an individual).

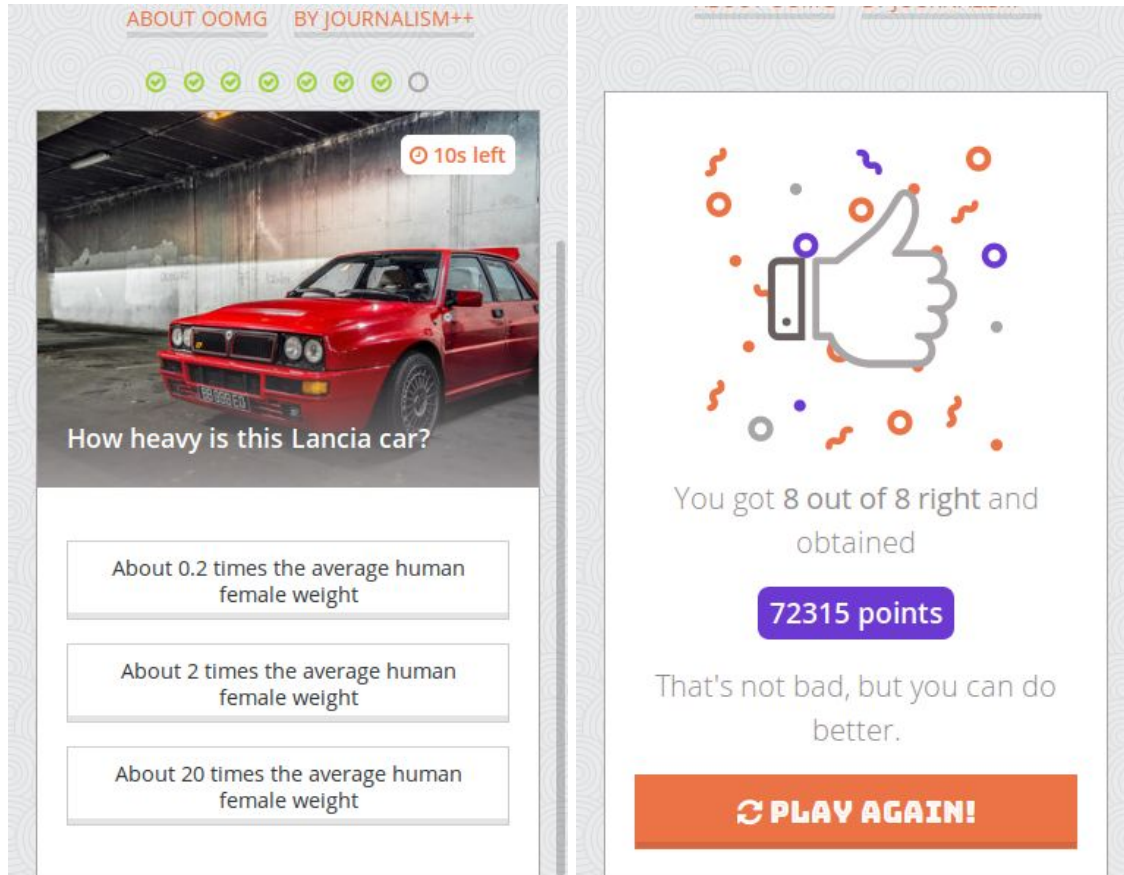
In total, 41 questions were asked, including 14 that dealt with daily life items. Given the large difference between the possible choices (a factor of 10), most questions could be answered correctly using deductive reasoning while the others could be answered with only a modicum of topic-specific knowledge. The list of questions is displayed in Annex 1.



Figures 1 & 2. The home screen of the game and the questions about the user.



Figures 3 & 4. Examples of questions in “written” form (left) and “numerical” form (right).



Figures 5 & 6. Example of question in “comparative” form (left) and final screen.

2.3. Results

The game was played 600 times between November 8 and November 18, 2016, by users from 51 countries. Over 10,000 questions were answered, an average of two games by user. The link to the game was shared on social networks such as Twitter and Facebook and via email. A small (€30) budget was used to advertise the link on Facebook.

Users found the correct answers two times out of three, on average. Random chance would have returned the correct answer in once every three questions. The detailed list of questions and their success rate can be found in Annex 1.

Users performed significantly better when the answers were displayed in numerical or written formats, as opposed to a comparison (see figure 7). Even for items from daily life, which admitted a correct answer inferior to one million euros, 20 kilograms or 10 meters, the format users could answer most easily was numerical ('1,000,000') with a success rate of 90.5%

followed by written ('one million') with 88.4% and comparative ('a lifetime's worth of earnings') with 75.1% (see figure 8). The difference between numerical and written out for small amounts is largely irrelevant as they are equivalent for small amounts from 0 to 99. (The numerical format for 100 is "100" against "one hundred" for the written format.)

Correct answers by question format

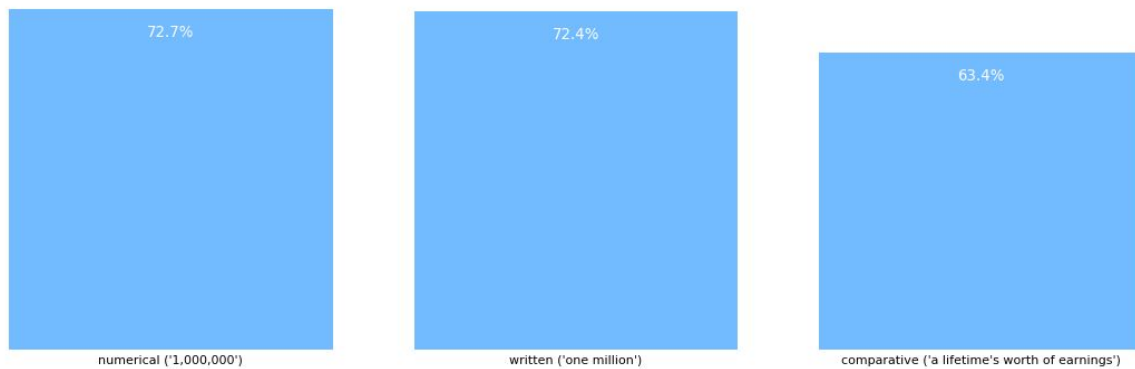


Figure 7. Correct answers by format.

Correct answers by question format for daily life items

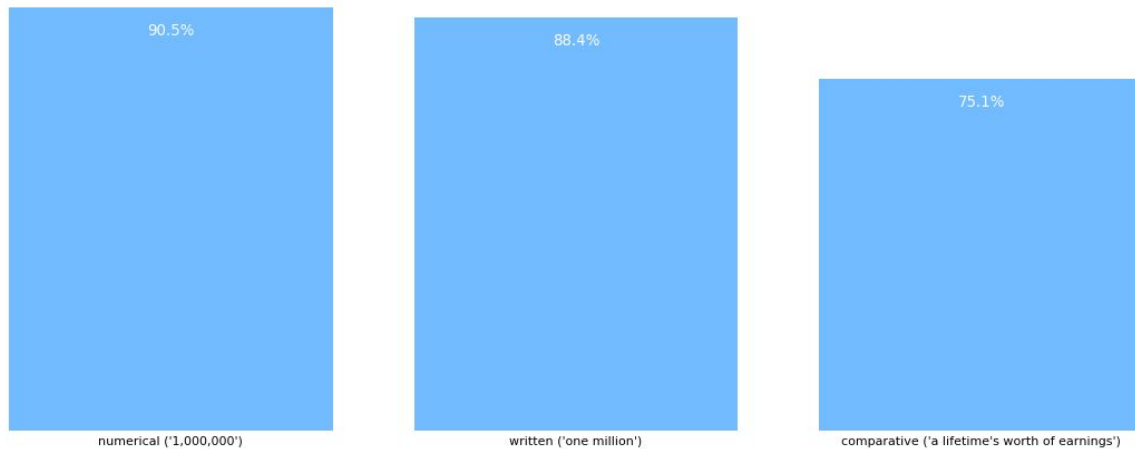


Figure 8. Correct answers by format for daily life items.

Whatever the amount and the category, the comparative format was the one least likely to be understood by users. Figures 9 to 13 show the success rate of the different formats for the four categories used in the game: distances, weights, prices and budgets.

Correct answers by format for questions of the price category

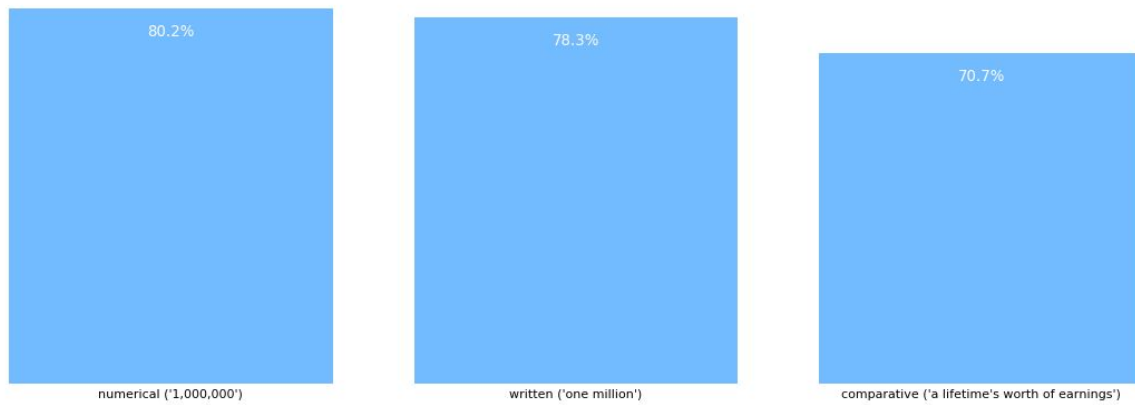


Figure 9. Correct answers by format in the price category.

Correct answers by format for questions of the weight category

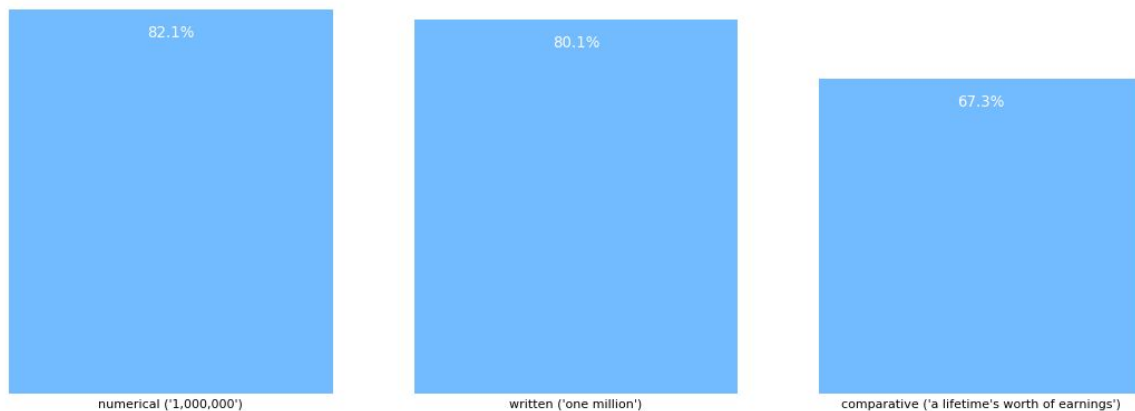


Figure 10. Correct answers by format in the weight category.

We ran a variety of tests to find out if comparisons amount could be superior to the numerical or written out format in some cases. It could have been, for instance, that comparing numbers that are not round (e.g. “The tower of Pisa is 31 times the height of an average female”) is harder than comparing round ones (e.g. “A Bentley Mulsanne costs 10 times a Volkswagen Golf”). However, as figure 13 shows, this is not the case.

Correct answers by format for questions of the distance category

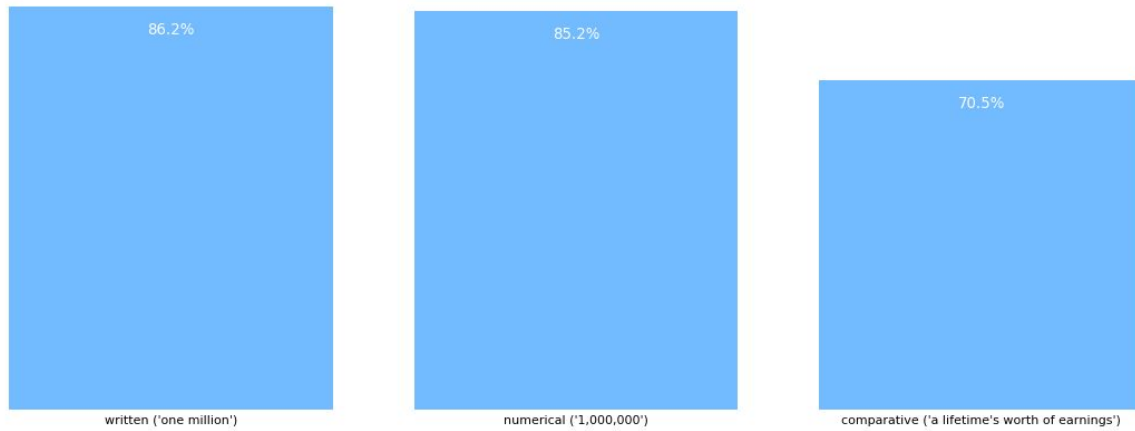


Figure 11. Correct answers by format in the distance category.

Correct answers by format for questions of the budget category

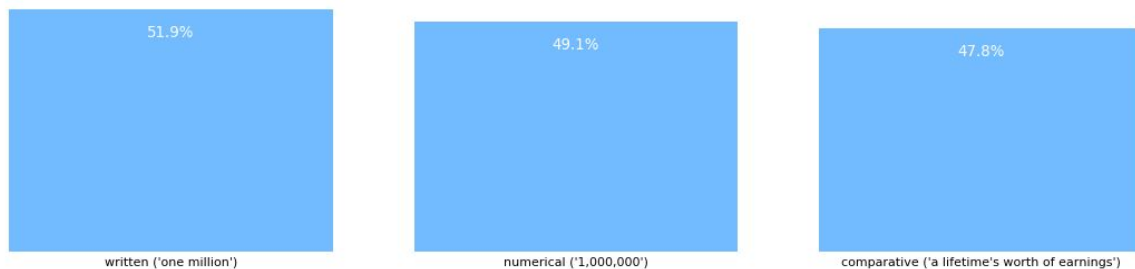


Figure 12. Correct answers by format in the budget category.

In all analyses, the comparative format performs a good 10 percentage points lower than other formats, except in the budget category (figure 12), where the difference is barely 3 percentage points. The reason for this curious result is due to the answers to two questions about the budget of the city of Berlin and the budget of the central government of France. Paragraph 2.5 - *The relationship to the norm* explores in details why this is the case.

Correct answers by format for questions that admitted a multiple of ten as a comparison amount

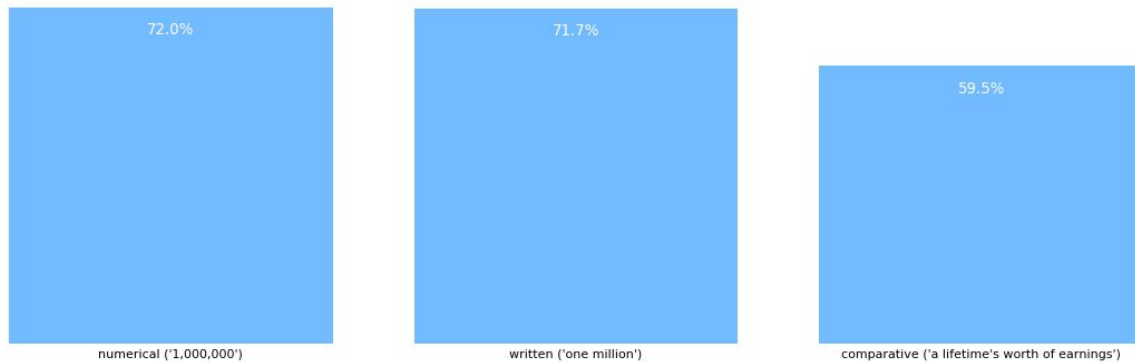


Figure 13. Correct answers for questions that admitted a multiple of 10 as a comparison amount.

2.4. A professional issue

While few would argue, like Matt Waite, quoted in the introduction, that journalists have trouble with numbers, little experimental evidence exists to demonstrate it. Harrison (2016) ran a questionnaire in basic numeracy to journalists and statistics students. Unsurprisingly, statistics students performed better.

The results of the Order of Magnitude Guesser abound in this direction. Users who self-identified as journalists performed less well than average (one percentage point or 1.5% worse than average), as can be seen in figure 14. On one question (distance between the Earth and the Sun), the rate of correct answers for journalists was 23%, much lower than a random guess.

This results shows that specific efforts must be made when conveying numerical information to journalists and that greater care must be taken when communicating it with them than with the general population.



Figure 14. Correct answers by self-declared occupation.

2.5. The relationship to the norm

The results of the Order of Magnitude Guessr show that users performed best when they could relate the amount under consideration with something they knew (hence the good results regarding items from daily life) or something they assumed to be normal.

Some questions yielded results that differed from the trends exposed above. When asked what the budget of the central government of France was, users answered wrongly 6 times out of 10 when presented with the numerical (€1,000,000,000,000) or written form (one trillion euros). This result, barely above random choice, is in line with other findings. However, when presented with the comparative form (twice the budget of Spain), users found the correct answer four times out of five. Given that users probably had no idea about the size of the budget of Spain, the informational value of the statement was close to nil. However, users could correctly guess that two countries they knew to be somewhat similar in size had budgets that were of the same order of magnitude. In other words, they assumed that it would be *normal* for two countries of similar size to have budgets that are not too far apart.

Other examples point in the same direction. Users were twice as likely to correctly guess the budget of the city of Berlin in the comparative format (twice the budget of Paris) than in the numerical or written formats (twenty billion euros). Again, users could guess that Berlin and Paris, two cities they knew to be somehow similar in size (2 million inhabitants in Paris and 3.4 in Berlin) had budgets of the same order of magnitude. Conversely, when asked to compare

budgets of entities users could not relate to any sense of normality, users performed worse than random chance. When asked to compare the budget of the European Institutions against the budget of Sweden, for instance, users found the correct answer 28 times out of 100 (random chance would have yielded a correct results 33 times out of 100). Similarly, users were wholly unable to correctly guess the budget of the European Space Agency. If the comparative form had been a comparison with the budget of NASA (which is 4 times higher than ESA's), users would probably have been able to correctly guess, because it would be *normal* that NASA and ESA have budgets in the same range.

That users relate to normality when estimating correct amounts explains a lot of the choices journalists make when covering stories that involve monetary amounts. A story begins when something out of the ordinary happens. In other words, it begins when an *abnormal* situation develops. A situation can only be correctly reported as abnormal if the person who reports it knows what normality is on a particular issue. This is why stories that involve a large discrepancy between what the journalist considers to be normal and the situation at hand can get a lot of coverage. These include the “CoiffeurGate” (François Hollande personal hairdresser being paid €10,000 a month) and the taxi story (a public official raking €40,000 in taxi expenditures in a few months), where journalists could easily understand the difference between a normal hairdresser and a normal taxi expenditure for them. It does not matter that normality, in this case, could not be the same for heads of state and heads of public institutions than for journalists. French presidents enjoy privileges ranging in the millions of euros per year in personal benefits during and after their presidencies (Dosières, 2016). In particular, former president Sarkozy employed a full-time make-up artist for €8,000 per month (Des Déserts, 2016). The hairdresser story became popular not because journalists correctly identified an abnormal situation, but because they used a wrong normality to reach their conclusion. They used their normality vs. the French president's instead of the French president's vs. other French presidents' and public officials'.

In cases when journalists cannot relate to any normality, stories do not appear as abnormal. A story on the laundering of 20 billion euros from Russia to the United Kingdom, for instance (OCCRP, 2014), received no coverage at all from most large British outlets. Following our theory, it can be explained by the fact that journalists have no clue about how money *normally* transits from Russia to London and could not identify a major story when it appeared. A counter argument could be that the topic requires expert knowledge to be understood. However, few

would argue that readers and journalists should not read or cover a trial because they do not know how the judicial system works. Most topics are covered while neither reader nor journalist has comprehensive knowledge of it.

This theory on journalists covering only stories that involve amounts they consider to be not normal can be contested on several fronts. First, many factors come into play when deciding to pursue a story or not. Self-censorship is most obvious, especially in newsrooms where clients (i.e. advertisers) or patrons have strong demands regarding the nature of the stories being covered. The reason *The Daily Telegraph* did not cover the “Swiss Leaks” stories was not because journalists there could not understand the abnormality of tax evasion, but because one of their clients - HSBC - asked them not to (Plunkett et al., 2015). Similarly, it could be argued that *Le Figaro* does not cover the illegal financing of Member of French Parliament Serge Dassault’s electoral campaign not because their journalists do not understand it, but because Mr. Dassault owns the newspaper.

Another argument against our theory lies in the embodiment of a story. A story that involves well-known characters will get more coverage and receive more interest from readers than a story that does not. Such an argument aims to explain why the €100,000 “CoiffeurGate” story received as much attention from journalists, for instance (it featured French president Hollande), and why the €20 billion laundering scheme did not (it featured no known characters). This argument is very weak. Indeed, the role of the journalist is precisely to find and describe characters that can carry a story. Several investigations relied on personification to make the story more enjoyable for the audience. “Swiss Leaks”, for instance, a 2013 investigation into how HSBC organized tax evasion from Switzerland, used the person of leaker Hervé Falciani to support the narrative. The taxi story exposed above used the person of Agnès Saal, who was unknown to the public prior to the story. Any schemes that involve abnormal sums of money also involve colorful personalities that can be turned into public persons and serve as a backbone for articles.

More research would be needed to estimate the amount of coverage a story receives based on its actual and perceived abnormality. However, our theory, which states that stories receive attention from journalists when journalists can relate them to a situation they know to be normal, already helps in explaining the questions that we set out to answer in this work package and provides very useful insights regarding the approach to be taken to improve the coverage of

budget data by journalists.

2.6. Implications for tool-centric approaches

Several tools have been built based on the assumption that large numbers that do not make sense can be explained by providing automated contextualization. The Dictionary of Numbers, for instance, is a browser extension that automatically finds comparisons for amounts found in articles.

A winner can take the payout over a span of 29 years or take a more immediate but smaller, lump sum payment.

Mega Millions players have a 1 in 292,201,338 chance of winning the top prize. To do so they must match all five of the white ball numbers and the yellow mega ball.

The biggest Mega Millions prize to date was a \$656 million [≈ box office sales of The Graduate, 1967] jackpot in March 2012 that was split among three tickets.

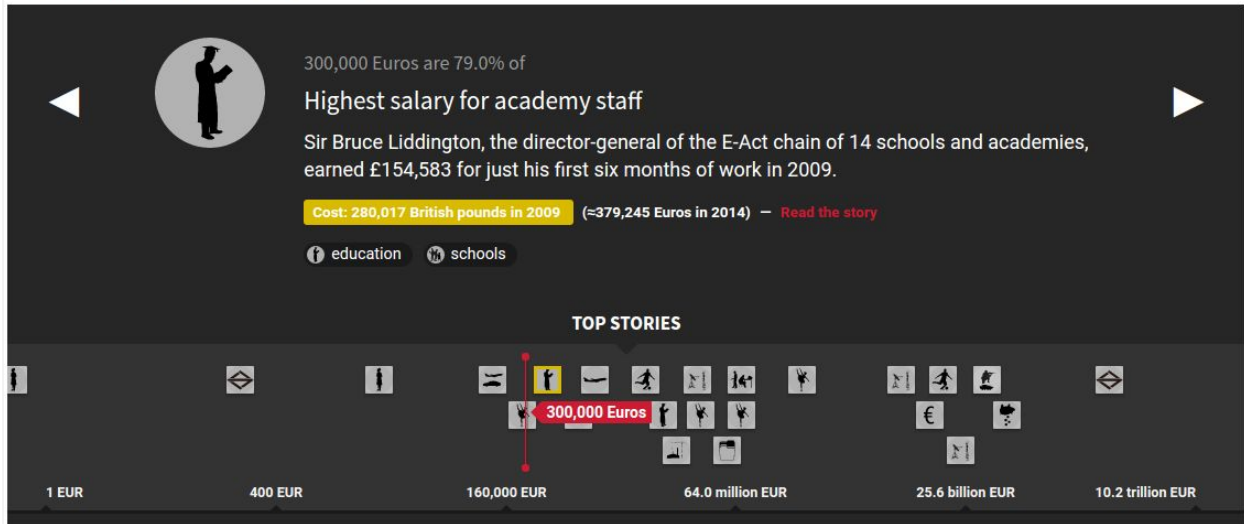
Related: [Meet the couple who sold the winning Powerball ticket](#)

Powerball -- the other major multi-state lottery -- has yielded the biggest jackpots this year. It paid out a \$1.6 billion [≈ total US hockey salaries for all teams, 2011] pot in January, and in May there was a \$430 million [≈ 2008 presidential election contributions to John McCain] prize.

An illustration of Dictionary of Numbers activated in a CNN article.

Another project, developed by Journalism++ for Open Knowledge Foundation, Spending Stories, was based on the same assumption and offered the same kind of contextualization. Users could type in any amount and see how it related to other amounts in a database.

Your query for **300,000 Euros** represents:



A query for “€300,000” in Spending Stories.

Both the Dictionary of Numbers and Spending Stories assumed that users can better understand an amount when it is compared to another amount, which they know. However, our experiment with the Order of Magnitude Guesser showed that the opposite was true: users are less able to understand amounts when they are compared to other amounts.

This explains the relative failure of both tools. Even though they were positively received upon launch, neither saw much usage - because, as the experiment above showed, they do not help users better understand numbers. To do so would require an automated appreciation of normality, e.g., a program that could automatically detect if an amount is within normal bounds. While this might be possible for some domains (a human being who is 2.4m tall is not normal, a plane that carries more than 200 tons of cargo is not normal²), our research showed that it was impossible to automate the contextualization of public budget data.

3. How budget data can help journalists

This work package focused on the elements ‘fostering debate’ and the ‘creation of narratives’ around the numbers contained in budget and spending data. To this end, we investigated several dozens of stories that involved public data and carried out case studies based on public

² The record for cargo on a plane is 250 tons in an Antonov 125, of which there is only one aircraft. An An-124, the largest common aircraft used for air cargo, carries a maximum of 150 tons.

budget data. Both approaches show that budget data is never enough in itself for journalistic work. Furthermore, the most important gaps for journalists are access to documents and cannot be bridged as part of the OpenBudgets.eu project.

3.1. The very small role of budget data

To assess the needs of the users (journalists), we looked at how budget data was used in journalistic work in two ways. First, we created case studies (see deliverables 5.1 and 5.2) for which we tried to answer research questions using budget data. Second, we looked at stories about corruption and other uses of public funds and looked how journalists came to their conclusions. A third paragraph explains how budget data can be helpful for journalists when it takes the form of direct payment lists.

3.1.1. Going from data to the story

We conducted several case studies of different scale, following the same principle: Asking a question which admitted an answer that involved budget data, usually at different administrative levels and in different member states. The first case study concerned the cost of the policies of the member states of the European Union that prevented men, women and children to access their territories or deport them therefrom (“The Migrants Files”). The second case study, “The Football Tax”, aimed at measuring the cost for taxpayers of professional football. The third case study, smaller in scope, looked at the cost of hosting people fleeing war in the city of Bonn, Germany, and its impact on its financial imbalances. The first two case studies were discussed in details in previous deliverables and will not be extensively covered here. Details of the third one can be found [in an article available online](#).

In all case studies, we and partner journalists found that, even when budget data was available, it, in itself, could not answer the research question. Knowing how much a policy costs, which is what journalists and their readers need an answer to, requires analytical, or mission-based, accounting. Most of the categories used in regular accounting cannot be linked to a specific mission. Instead, they are linked to paying units, which makes most sense for the administration itself.

Knowing how much the maintenance of a stadium costs, for instance, is impossible if the personnel doing the maintenance works only part-time on the stadium. A gardener paid by the city hall, for instance, might work on Mondays in the stadium and other days in city gardens.

Without this information (how much of the salary of the gardener was used for the stadium), it is impossible to break down costs by destination. It goes without saying that this example is an extreme simplification of an issue that can be much more complex.

Comparing across cities is similarly complex. Going back to the example of the gardener above, it can be the case that a city sends gardeners on its payroll to the stadium, while another outsources this work to a private company. Such situations can quickly escalate in complexity. A public mission can be carried out by several institutions at once (the city, the region and the state, for instance), and these institutions might have set up private companies belonging to public institutions to carry these out. The extraordinary variety of public-private partnerships which, until the early 2010's, were barely regulated, gave birth to a series of complex financial set ups where the responsibilities and duties of public and private parties can remain obscure. Finally, in-kind benefits given to private parties (e.g., free or cheap rent, sale of land under market prices, loan guarantees) are all but impossible to measure, were it only because there does not exist a market for some of these transactions. The sale of a very particular plot of land at a given price might be considered underpriced by local experts, but, absent a clear benchmark, such an estimation is impossible without detailed knowledge of the deal.

Experience from the field shows that doing analytical accounting can be, in itself, extremely complex. It is sometimes possible. In the Bonn case study, for instance, journalists managed to reach the same conclusions as the city comptroller, using public data (though no budget data ; the data in question was from a previous exercise in analytical accounting by the city and updated using other sources of information). In some other cases, analytical accounting is impossible to carry out because the documents that would allow it (mostly contracts between parties to the realization of a public service) are not released.

3.1.2. Finding the data involved in a story

This work package analyzed several dozen stories involving corruption or mismanagement of public funds. Stories from over twenty countries were analyzed, ranging in size from a few hundred euros to several dozen billion and, in time, from the 1980's to today. A rapid typology of the stories can be drawn from these resources. All the information in this paragraph is derived from the tutorial website of this work package, which can be found online at cookingbudgets.com.

Bribes. A public official requires a direct payment from a private person or organization to allocate a specific resource or to make a certain decision. Bribing is common in public housing in all countries we looked for it (France, Germany and Sweden), where payments can be required to obtain a flat more rapidly. At a higher level, members of parliaments can be offered 6-figure sums, usually in the form of consulting fees to a private company, to vote in a certain direction. In all of these cases, any misuse of fund that follows the bribe does not show in public budget data. Journalists pursuing these stories mostly rely on investigations from anti-corruption police or on leaks and testimonies.

Kickbacks. A public official awards a public contract to a company that is not the most qualified to carry out the task or deliberately fails to notice fraud from the contracting company in exchange for cash or favors. This is, by far, the most common form of corruption in Europe, excluding nepotism. It is common in public works, where cartels are set up by private companies to coordinate the answer to a public tender and thereafter pay part of the contract money they receive back to the contracting authority in exchange for their acceptance of the practice. This form of mismanagement does not show in public budget data³. Investigations on such schemes usually start from the concrete outcome of the task carried out by the contractor (e.g., the airport commissioned by the public authority is not up to the contracted standards), from appeals by unsuccessful bidders or from whistleblowers. Journalists usually rely on the work of anti-corruption police or transparency NGOs to cover such stories.

Indirect appropriation (conflict of interest). A public official awards contracts or makes direct payments to an organization to which she is linked. Conflicts of interest occur when a public official contracts a private company she (or someone close to her) owns to realize a task, even if this company is not the best to do so. Such schemes can be uncovered by journalists themselves if they have access to the list of payments made by a public administration *and* to the beneficial ownership of the private entities at the receiving end. Some member states of the European Union do make it easy for journalists to access this information but, in most countries, access to business, property and land registers is off-limits.

Direct appropriation. A public official makes a direct transfer of public funds or property to himself. An example of direct appropriation are expenses claims, through which public officials

³ It can show in public procurement data. The Redflags.eu project, for instance, tries to identify bogus procurements.

can ask an administration to pay for personal shopping or travel. Such cases can be uncovered by journalists when the lists of direct payments from an administration are available.

This short typology of abnormal situations shows that, except in a very few occasions, public budget data is not used by journalists working on a story, who rely, instead, on information from the police and justice departments and on their contacts therein, or, more rarely, on whistleblowers.

Our findings are in line with the TACOD report, which explored the impact of open data in the fight against corruption. It found that open data had, at best, a minor role in assisting investigative journalists (RiSSC, 2015).

3.1.3. The exception of direct payments

A notable exception to this relative uselessness of budget data are direct payments. The list of direct payments can be used effectively for two reasons. Unlike budget data, which only gives an indication of how money might be spent, spending are actual expenditures and need not be checked against disbursements for validation. More importantly, they are concrete transactions where money changed hands, as opposed to the accounting lines which rarely deal with actual exchanges of funds⁴. Such transactions help journalists insofar that they can use the information to follow up on existing stories or find leads for new ones.

Unfortunately, direct payment data is rarely made available. The European Commission is the one of few public administrations in Europe to release it systematically. Other administrations sometimes publish lists of subsidies, which is already helpful - if only partial.

The potency of the publication of direct payments was best illustrated at the European Parliament, where some members changed their behavior after it was made clear that some of their expenditures would be made public. In this case, the transparency of budget data produced tangible results.

However, such results can only be attained if the public officials whose expenses are made public know that journalists or citizens will search through the data to find interesting stories. At

⁴ Simple math shows that less than half of all analytical writings deal with money transfers. Double-entry bookkeeping ensures that any cash transaction produces one non-cash accounting line. Add all the other non-cash accounting writings (amortizations, provisions etc.) and you see that money transfers are only a fraction of all accounting lines.

the European Parliament, British journalists are known to stand guard over any piece of information that could be used to run an anti-European Union headline. Other examples show that publication alone, without sustained outrage from journalists and voters, does nothing to change habits. The French parliament published the details of the “réserve parlementaire”, a €150,000 yearly sum that members of parliament receive and can give to any organization they wish. One news outlet used this newly published information to show the many ways MPs used this source of cash to advance personal projects, such as repairs in the village where they own a holiday home or donations to organizations run by members of their families (Equy et al. 2015; Léchenet, 2016). Absent a wider coverage, such stories had no impact on corruption practices in the French parliament.

3.2. The unbridgeable gaps

So far, we have seen that, while budget data could rarely be of use for journalists, detailed spending data and other documents (chiefly contracts with private entities) could help journalists when working on stories that involve public spending. Such documents are almost never accessible. The detailed list of gaps has been explained comprehensively in deliverables 5.1 and 5.2, the following paragraphs summarize these findings and include additional research carried out in months 12 to 20.

3.2.1. Non-availability

Data and documents that citizens and journalists are legally entitled to have access to are not accessible. Some administrations, especially at the local level, refuse to abide by their obligation to publish the deliberations of their councils. Request to publish these are almost always turned down. The case studies carried out in this work package produced dozens of requests to communicate public data. Not once was the request answered in a timely manner. Given that most journalists work on very tight deadlines, such refusals amount to preventing a journalist from working on a story. Appeals can be filed against such decisions, but the time and resources required prevent any journalist working in normal conditions to do so.

These findings, detailed in previous deliverables after research in six member states, were confirmed by new results from Spain. A 2004 investigation by the Spanish Court of Auditors found many irregularities in the subsidies given by public administrations to football clubs. Journalists tried to follow-up and update the investigation as part of the “Football Tax” case

study. They found it impossible to access the data. They filed several freedom of information requests under the newly-implemented Freedom of Information legislation in Spain but were not given a positive answer by any of the contacted administrations.

On top of this blatant disregard for the law, public administrations also engage in the destruction of online archives, making it impossible for journalists to access budget data from previous years (see deliverable 5.2).

3.2.2. Obscurity

When information is available to journalists, it is often unusable, or barely. Format is the most immediate concern. Information is often available in PDF files, which makes it very cumbersome to work with. The format of the PDF files makes any automated parsing impossible. Not one administration contacted during the case studies provided us with structured data. (The case study in Bonn is the exception, but the decision to do the case study stemmed from Bonn city data being available in a machine-readable format). The issue is not only technical. In some cases, the text of the deliberation is the only information source on a precise transfer of public money. A subsidy can be voted in a municipal council, for instance, and disbursed in installments (always the case for most activities that follow the academic year, not the fiscal one), so that it would be impossible to track when looking at the banking statements of the organization. In accounting documents, the subsidy would appear, but its beneficiary might not.

Finer issues ensure that public budget documents are extremely hard to interpret correctly. Public administrations usually abide by specific VAT regulations. Knowing whether an amount paid by an administration includes VAT or not makes a difference, usually 20% of the total amount. Despite the crucial importance of this aspect, public administrations never specify if a spending includes VAT and, if it does, if the VAT is later reimbursed by the entity that collects it.

Co-financing also proves to be an unsolvable problem. Money can be transferred from one public entity to the other to the point when it becomes impossible to disentangle the source of financing. This is especially problematic when the decision power an entity in the chain has is unclear. Some public institutions receive money that they are bound by law to transfer forward, but the precise amount of leeway they have is rarely clear-cut. This again was shown to be true in the work on the "Football Tax" in Spain, where football clubs have to report direct subsidies from public bodies. However, most subsidies to football clubs are funneled through intermediate

vehicles, which do not have to disclose information.

Some special cases arise, too, where private entities are tasked by law with the collection of money and the provision of public services (the Swiss national lottery, for instance).

Finally, obscurity comes from the accounting and legal terms used in budget data, which can require thorough training.

These limitations to public budget data explains why the journalists we contacted for in-depth interviews never quoted public budget data as a source for their work. Instead, they rely on direct contacts with officials, on leaks and on testimonies.

4. Conclusion

The gap analysis of this work package had to (1) take stock of the current situation, (2) determine how budget data could be used by journalists, (3) define the gaps to be filled and (4) assess the urgency of individual measures.

We showed in deliverables 5.1, 5.2 and this one that journalists very rarely use open budget data (1) and for good reason. Budget data does not let journalists investigate or write stories, either because it is incomplete in itself (e.g., when the interesting information is to be found elsewhere), because it would require extensive work (analytical accounting) or because, as happens most of the time, it is inaccessible (2). Of the many gaps to be filled, we showed that access to documents, especially contracts and detailed spending data, was the most pressing need (3). However, because the vast majority of public administrations we looked at do not respect the law, these gaps are not bridgeable. The measures to be taken mostly concern the ability for journalists to engage in serious training (4). The case studies, workshops and the experiment on orders of magnitudes, showed that journalists in particular needed to upgrade their skills in numeracy, in accounting and in the basic knowledge of corruption practices.

The analysis done under task 5.1 also showed that budget data was not comparable across time and space, even within the same administration. This has profound implications for other partners in this project. It impacts work package 1 insofar as the code lists for a budget line cannot be mapped across administrations or across time. Administrations have different duties, defined by law, different accounting standards (which can be defined by law, as in France, or

left to the discretion of the administration itself, as in Germany and Austria) and different practices that ensures that the underlying reality of a single budget line, even if it shares a name, will not be comparable. It impacts work package 2 insofar as data analysis that aims at automating the finding of insights in budget data is unrealistic. Not only does the impossibility to match the namespaces of different budgets make comparisons impossible, but most interesting findings regarding public monies are made off-budget, as we showed in this and previous deliverables. These findings have been communicated to the leaders of the respective work packages throughout the duration of the project.

This work package helped advance our understanding of how journalists work with budget data. Extensive research, be it with journalists during workshops, in interviews with practitioners at all levels, from local journalists to coordinators of world-wide consortia or a controlled experiment, let us show convincingly that journalists rely on their perception of normality when tackling stories involving budget data and that this perception of normality may not be obtained from budget data. Instead, sector-specific expertise is required to be able to correctly spot abnormal, and hence newsworthy, situations.

We showed that the assumptions that we used on previous projects, and in OpenBudgets.eu, were false. In particular, it is not true that comparisons provide a benefit to the user by helping her contextualize an amount. This was the basis for tools such as Spending Stories and the Billion-dollar-o-gram. The realization of this error will help ensure that it is not repeated and that practitioners can learn from our mistakes.

Other assumptions used in OpenBudgets.eu have been proven wrong by the research in this work package (others found concurring results). In particular, it is false that budgets can be compared effectively across time and space, because budget data is much too heterogeneous. It is false to assume that administrations will publish detailed budget data when asked, even if the law requires them to. It is false that transparency in budget data alone can reduce corruption, because most corruption happens outside of what budget data covers. Such findings will no doubt be precious for other partners in OpenBudgets.eu and for the Commission itself in preparing further calls on the topic of transparency.

5. References

Begley, Sarah. "François Hollande Draws Scorn for Paying Hairdresser Nearly \$11,000 Per Month", TIME Magazine, 13 July 2016.

Borenstein, Seth, "Enough water to fill 1 billion Olympic-sized swimming pools has melted off of Alaska's glaciers over the past 5 decades", Associated Press, 2015.

Delaporte, Lucie, "Scandale de l'université des Antilles: deux responsables suspendus", Mediapart, 3 June 2014.

Demonocracy, "US Debt - Visualized in physical \$100 bills", Youtube, 2012.

De Paor, D. G., S. J. Whitmeyer, C. Bentley, and M. M. Dordevic. 2014. "EarthQuiz: A Crowd-Sourced Resource for Teaching and Learning Geoscience"

Des Déserts, Sophie. "Le pouvoir dans la peau. La maquilleuse de Sarkozy livre ses secrets", Vanity Fair, 21 mars 2016

Dosière, René. "Anciens présidents de la République: toujours plus", 16 April 2016.

Harrison, Steven. "Journalists, Numeracy and Cultural Capital". Scholar Commons 9.2, 2016.

Laubacher, Paul. "Agnès Saal, un (grand) tour en taxi et puis s'en va (de l'INA)", L'Obs, 28 April 2015.

Léchenet, Alexandre, "La réserve parlementaire, source de subventions et de conflits d'intérêts", Libération, 27 September 2016.

Equy, Laure; Laïreche, Rachid; Bouchet-Petersen, Jonathan. "Réserve parlementaire : les députés à visage découvert", Libération, 12 February 2015.

New-York Times, "HUGE WAREHOUSE RISING IN JERSEY; Seeman Bros. Structure in Carlstadt Is 572,000 Sq. Ft. Clear Height of 21 Feet Extensive Offices", 20 May 1962.

OCCRP, "The Russian Laundromat", 22 August 2014.

Plunkett, John; Quinn, Ben. "Telegraph's Peter Osborne resigns, saying HSBC coverage a 'fraud

on readers", The Guardian, 18 February 2015.

Pro Bono Economics. "Cost of outcomes associated with low levels of adult numeracy in the UK", 2014.

RiSSC, "Revolution Delayed: The impact of open data in the fight against corruption", 2015.

Svensson, Britta, "Nej det handlade inte bara om Toblerone...", Krönikörer, 4 January 2007.

Waite, Matt. "How I faced my fears and learned to be good at math", Nieman Lab, 13 November 2013.

Whitbourn, Michaela. "NSW Premier Barry O'Farrell resigns", The Sydney Morning Herald, 16 April 2014

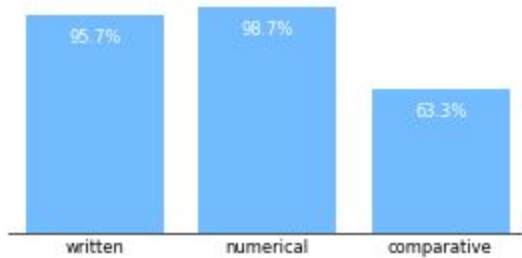
Wines, Michael, "Grand Soviet Scheme for Sharing Water in Central Asia Is Foundering", The New York Times, 9 December 2002.

Zhu, Liangfeng, Xin Pan, and Gongcheng Gao. "Assessing Place Location Knowledge Using a Virtual Globe." Journal of Geography 115.2 (2015): 72-80.

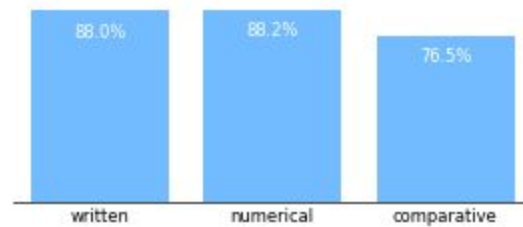
Annex 1: Answers to the Order of Magnitude Guesser

This annex shows the correct answers to each question in the Order of Magnitude Guesser, broken down by question format.

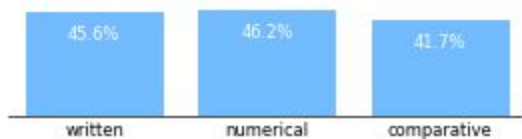
How heavy is this laptop?
Answer: 2 kg



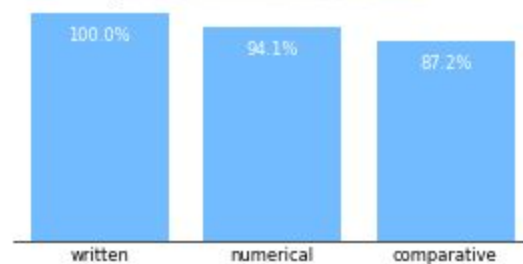
How much does this piece of bread weight?
Answer: 1 kg



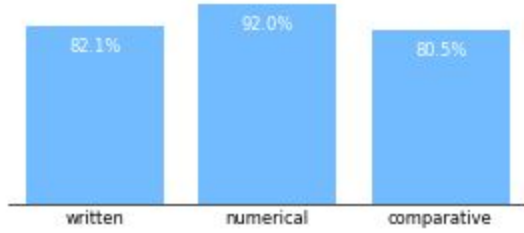
How much does this tanker weight?
Answer: 100000 tons



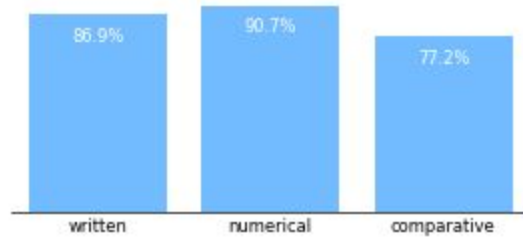
How high is the Pisa tower?
Answer: 50 m



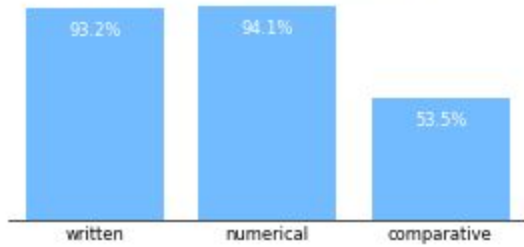
What's the length of a high-speed train?
(with 8 cars)
Answer: 200 m



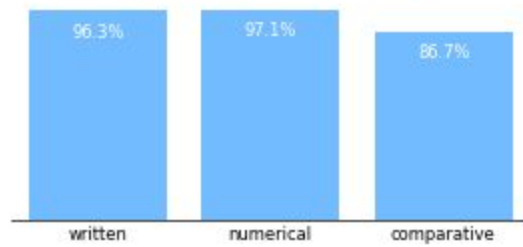
How high is the Burj Khalifa?
Answer: 800 m



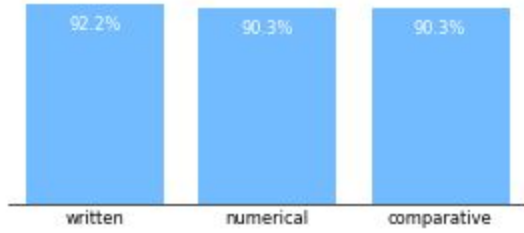
How high is the Eiffel tower?
Answer: 300 m



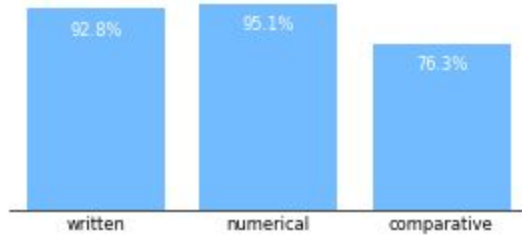
How much does this chai latte at Starbucks
cost? (normal size)
Answer: 5 EUR



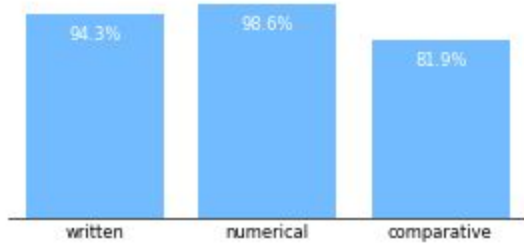
How much is the Spotify subscription (per month)?
Answer: 10 EUR



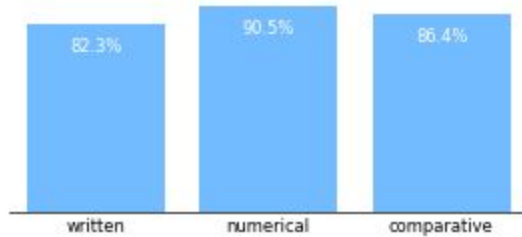
How much does a BigMac menu at MacDonald's cost?
Answer: 8 EUR



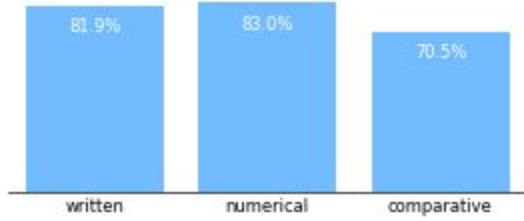
How much does a movie ticket in a large-screen theater costs?
Answer: 10 EUR



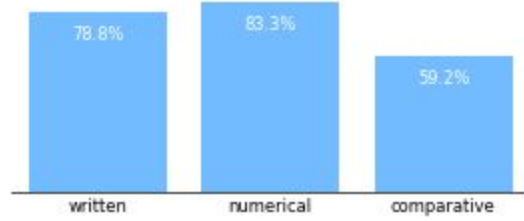
What is the price of a H&M blazer?
Answer: 80 EUR



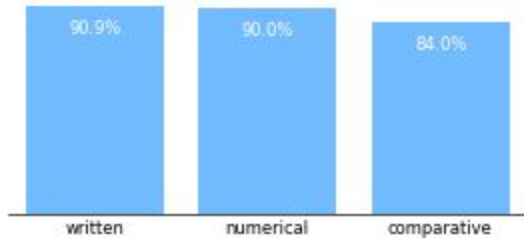
How heavy is this book?
Answer: 300 g



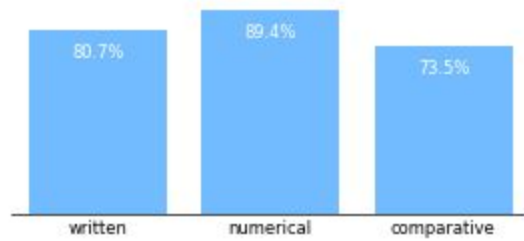
How much does this elephant weight?
Answer: 4000 kg



How heavy is this cow?
Answer: 700 kg

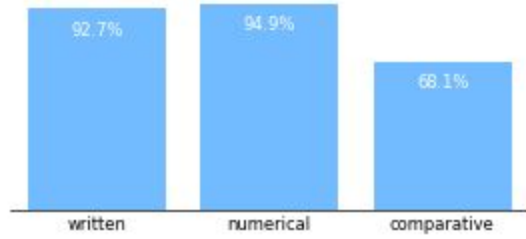


How heavy is this Lancia car?
Answer: 1200 kg



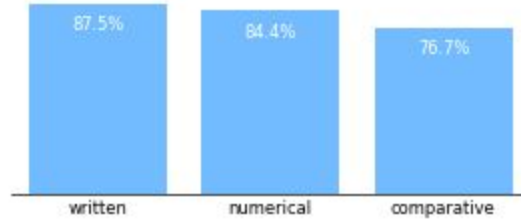
How much is this iPhone (without subscription)

?
Answer: 800 EUR



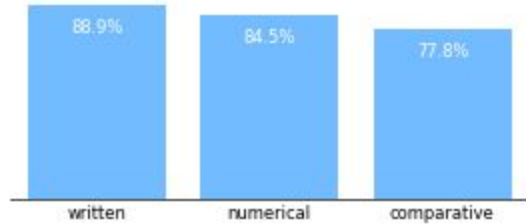
What is the distance between Paris and Berlin

?
Answer: 1000 km



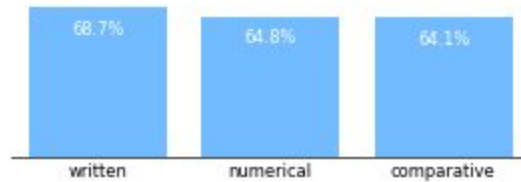
What's the distance between Munich and Berlin

?
Answer: 600 km

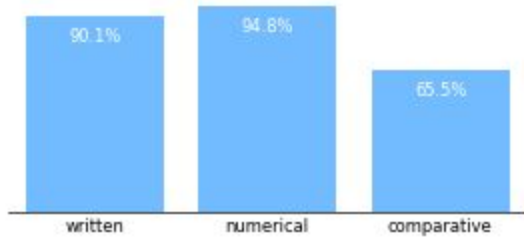


What does a herniated disc surgery (a common back surgery with a 24-hour stay in the hospital) cost?

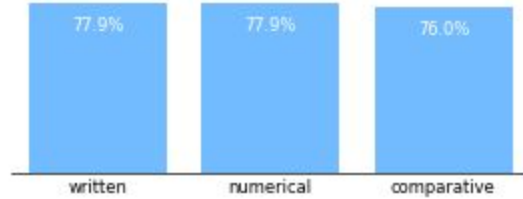
Answer: 5000 EUR



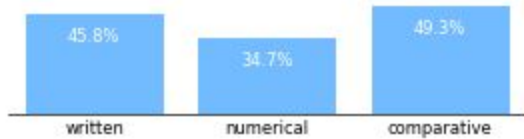
What's the monthly salary of a school teacher
(gross, in Berlin)?
Answer: 3000 EUR



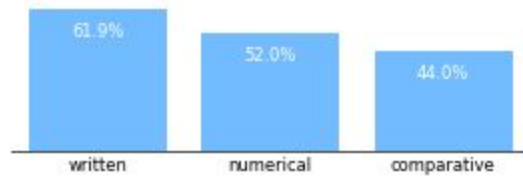
How much does this 2016 Bentley Mulsanne
cost?
Answer: 200000 EUR



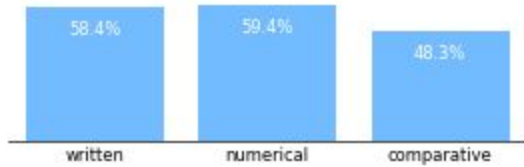
How much does this Paris opera house cost,
per year?
Answer: 100000000 EUR



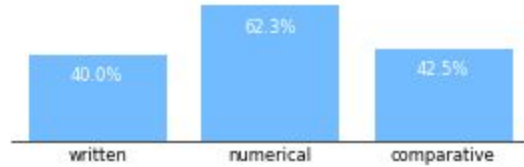
What does this A320 cost?
Answer: 100000000 EUR



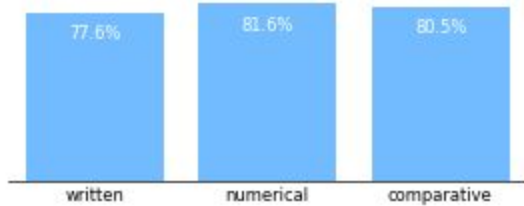
What does this SuperPuma chopper costs?
Answer: 10000000 EUR



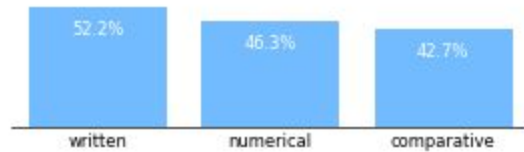
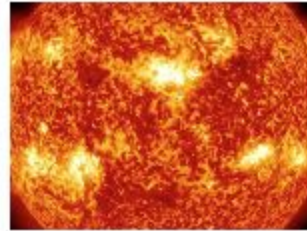
How much does a primary school cost per year?
(in Berlin, where a primary school has about 400 pupils)
Answer: 2500000 EUR



What is the price of a 3-room apartment in Paris?
Answer: 1000000 EUR

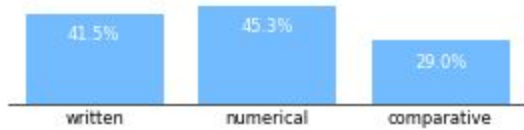


How far is the Sun from the Earth?
Answer: 150000000 km



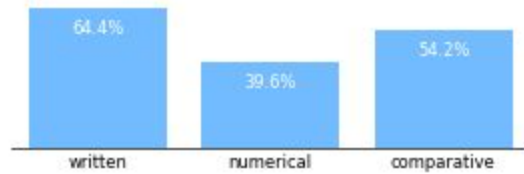
What's the budget of the European Space Agency (per year)?

Answer: 5000000000 EUR



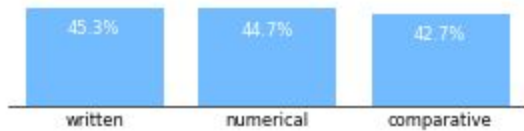
How much did the new stadium of Nice cost?

Answer: 400000000 EUR



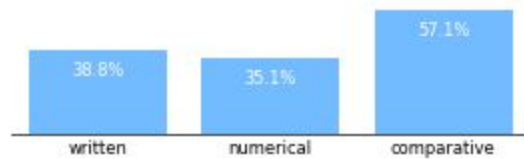
How much subsidies do Europeans give to their farmers, yearly?

Answer: 60000000000 EUR



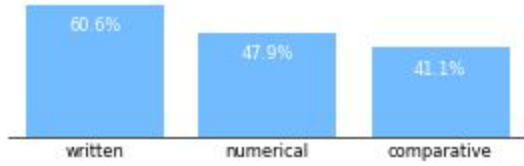
What's the annual budget of the city of Berlin?

Answer: 20000000000 EUR



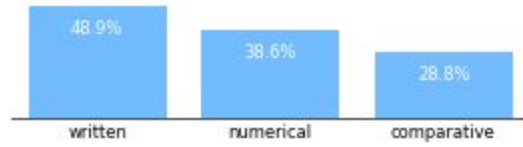
How much does Germany spend on its military (per year)?

Answer: 350000000000 EUR



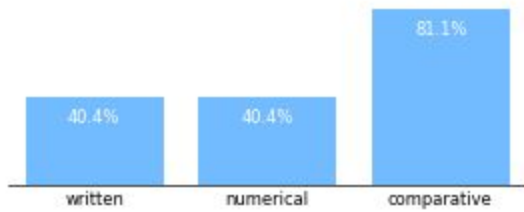
What's the budget of the EU (per year)?

Answer: 1500000000000 EUR



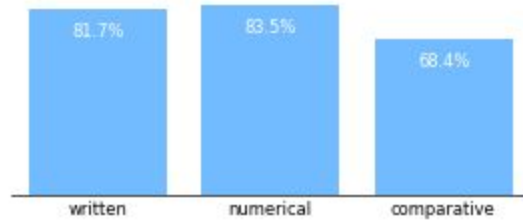
What's the budget of France (central government only)?

Answer: 10000000000000 EUR

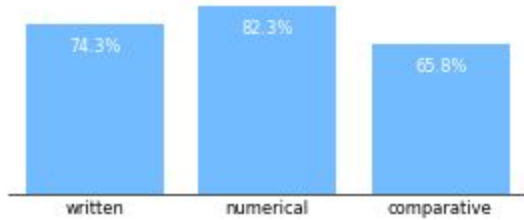


How much does this office chair weight?

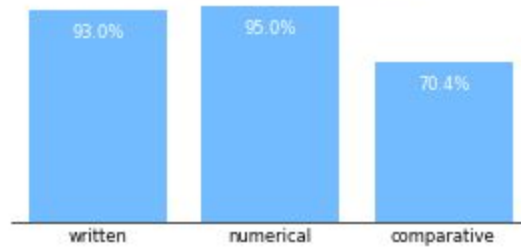
Answer: 15 kg



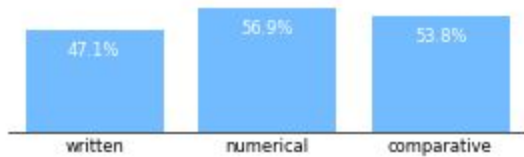
How much does this mountain bike weight?
Answer: 20 kg



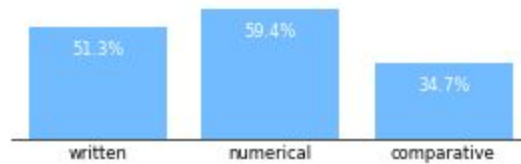
How high is this house?
Answer: 8 m



How much does this F-16 fighter jet cost?
(The F16 is not a recent plane)
Answer: 15000000 EUR



How much does Slovenia spend on its military
each year?
Answer: 400000000 EUR



How much did the new German Chancellery building cost?

Answer: 300000000 EUR

