Kennesaw State University

Department of Information Technology

College of Computing and Software Engineering

**Development Project Report**

**Death by Revolution? The Economic and Socioeconomic Impact of the Arab Uprising**

URL: http://almuthmerkel.com/ArabUprising/

Submitted by: Almuth Dorothea Merkel

Course: IT 7113 - Data Visualization

Advisor: Dr. Jack Zheng

Submission Date: 12/08/2019

**Executive Summary**

As Hiram Johnson once said: “In times of war truth is the first victim” (quoted in Bagdikian 2004, p. xi). There is truth in this quote for the Arab Uprising, the context of this development project. Especially socioeconomic data is scarce. Thus, the objectives of this development project are to (1) raise awareness for the consequences of conflict and therefrom stemming losses in future earnings, i.e. income per capita, (2) take advantage of recent innovations in data analysis, data analytical technologies, and socioeconomic modelling and the opportunities for burden of conflict estimates, and (3) transfer a prospective estimation strategy and simulation framework into a data visualization tool that is useful beyond academia and may contribute to a more informed public debate, appeal to potential humanitarian assistance organizations and enhance human assistance programming during crises. In order to especially achieve the latter, data storytelling is incorporated into interactive dashboard design. The final product may be found under the following link: <http://almuthmerkel.com/ArabUprising/>.

**Background**

In 2010, a series of anti-government protests, uprisings, and armed rebellions spread across Middle Eastern and North African countries. Because the Arab Uprising evolved into many armed conflicts, such as in Syria, Yemen and Libya, the focus of the burden of conflict is often limited to a dead body count. Dead body counts, however, do not provide an accurate picture of the burden of conflict. Conflict leads to destruction of economic assets (capital like machines, roads, and infrastructure), the diversion of scarce resources from a peace to a war economy (in economics famously known as the guns and butter trade-off), and the disruption of economic activities (through military actions, such as blockades, sieges, and embargos)(Anderton & Carter, 2009).

Yet, a full picture of the impact of the socioeconomic consequences of conflict is not provided by a mere dead body count. Many more factors need to be considered. As for the Arab Uprising, despite the almost real time coverage in new forms of communication such as social media, available data from the ground still provides only a partial picture of the socioeconomic consequences of the Arab Uprising. Moreover, the literature investigates merely the causes and consequences, but fails to address imperative questions, such as: How many children die before the age of five as a result of the Third Army of war? Such a question, however, is important for a more informed public debate and humanitarian crisis management.

In order to bridge the gap between academic research and the private and public sector, these questions not only need to be addressed - a prospective estimation strategy and simulation framework needs to be transferred into a data visualization tool that is useful beyond academia. Thus, the goal of this project is to develop an interactive data storytelling dashboard, which aims to answer the following questions:

* How much GDP per capita is lost?
* How many children’s lives are lost?
* How much weight gain is lost?
* How many school years are lost?

**Data, Methodology, Technology**

*Data and Methodology*

First, I estimate the impact of conflict on economic and socioeconomic indicators using a panel fixed-effect model of the following kind:

$Income\_{it}= β\_{0}+β\_{1}×Conflict\_{it}$ (1)

$Socioeconimic Indicator\_{it}= γ\_{0}+γ\_{1}×Income\_{it}+γ\_{2}×Conflict\_{it}$ (2)

For these estimations I use a panel dataset with variables, described in more detail in the below table 1. This dataset covers the time period from 1960 to 2019 and all 217 countries available in the World Development Indicator Database. I show the results of the panel fixed effect regression in Appendix table 1.

**Table 1: Data and Sources**

|  |  |  |  |
| --- | --- | --- | --- |
| Variable | Variable | Source | Description |
| GDP per capita | GDP per capita | UN National Accounts (2019) | GDP per capita (constant 2010 US$) |
| Socioeconomic Indicator | Infant Mortality | IHME (2019) | Under-5 mortality rate (probability of dying before the age of 5 per 1,000 livebirths) |
| Wasting | IHME (2019) | Prevalence of wasting in children under 5 (%) |
| Primary School Enrolment | UNESCO (2019) | Adjusted net enrolment rate, primary, both sexes (%) |
| Conflict | Armed Conflict Total | INSCR (2018) | Total summed magnitudes of all (societal and interstate) major episodes of political violence |

I then use the coefficients, obtained from the employed panel fixed-effects regression, to calculate two different economic (socioeconomic) trajectories, one showing a country’s development with conflict, and one without conflict. To do so, I calculate the average growth rate from 2001 to 2010 for GDP per capita and use this value to simulate the “no conflict”-trajectory.

$No Conflict Income\_{2011}=Income\_{2010}×\left(1+2001 to 2010 Average Income Growth\right)$ (3)

$No Conflict Income\_{t>2011}=No Conflict Income\_{t-1}×\left(1+2001 to 2010 Average Income Growth\right)$ (4)

The Conflict-trajectory is calculated similarly, using the coefficient instead of the moving average.

$Conflict Income\_{2011}=Income\_{2010}×(1+β\_{1})$ (5)

$Conflict Income\_{t>2011}=Conflict Income\_{t-1}×(1+β\_{1})$ (6)

As for the socioeconomic development, I use the change in income to estimate the conflict and no-conflict trajectories of the socioeconomic indicators (SI):

$No Conflict SI\_{2011}=SI\_{2010}×(1+2001 to 2010 Average Income Growth×\frac{γ\_{2}}{shock absorption time})$ (7)

$No Conflict SI\_{t>2011}=No Conflict SI\_{t-1}×(1+2001 to 2010 Average Income Growth×\frac{γ\_{2}}{shock absorption time})$ (8)

$Conflict SI\_{2011}=SI\_{2010}×(1+\frac{γ\_{2}}{shock absorption time})×(1+\left(\frac{Income\_{2011}}{Income\_{2010}}-1\right)×γ\_{1}$ (9)

$Conflict SI\_{t>2011}=Conflict SI\_{t-1}×(1+\frac{γ\_{2}}{shock absorption time})×(1+\left(\frac{Income\_{t}}{Income\_{t-1}}-1\right)×γ\_{1}$ (10)

In my calculations I assume an immediate shock absorption for income and wasting (less than a year). I furthermore assume shock absorption times of 5 years for school enrolment, and 2 years for infant mortality.

Thereafter, I calculate lost income, school years, weight gain and children lives as the cumulative difference between the conflict and no-conflict trajectory per age group. I use population data until 2015 (UN Population Division, 2019). In order to cumulate the losses beyond 2015, I use the 2006 to 2015 average growth rate to simulate the population. Also, as for school enrolment and wasting, I assume that each child forgoes on average 3 years of schooling and 4.35kg of weight. Since primary school consists of 6 years a randomly selected student may either already be in her sixth year and loose no school years, or in her first year and lose 6 school years, averaging to 3 forgone school years as a best estimate. The 4.35kg corresponds to two standard deviations from the average five-year old child’s normal weight, according to World Health Organization tables.

The development project, however, focuses on the 22 countries in Middle East and North Africa region (author’s regional definition). A JSON file holds all necessary information for these 22 countries.

*Technology*

Running the regression, I use the open source statistics software gretl (2018).

When developing a dashboard, a wide selection of technologies may be considered. I program my dashboard using conventional client-side technologies. I write a markup using html, format by using CSS and add functionality through programming in JavaScript and jQuery. All code is written using the Open Source text editor Atom (2019). Furthermore, I make use of the open source Google Charts API to create the graphs for my dashboard (Google Developers, 2019).

**Dashboard Design**

The purpose of my dashboard is to provide condensed information about the Arab Uprising. Specifically, it provides answers to the following questions:

* Which country faced conflict in 2011 after the Arab Uprising in 2010? This question is answered using a geochart.
* How did a selected country perform before the Arab Uprising compared to another country of the region, the regional average and the world average? This question is answered with user input and illustrated in a bar chart.
* How does (could have) conflict impact(ed) economic and socioeconomic development in a particular country? The right pane of my dashboard is dedicated to answer this question.

Creating the dashboard, I follow the steps for developing a data story, incorporating interactive dashboard design principles where necessary and appropriate. Figure 1 shows a screenshot of my dashboard.

**Figure 1: Final Dashboard “Death by Revolution? The Economic and Socioeconomic Impact of the Arab Uprising**

****

In the following I will address each of the data storytelling steps as proposed by Nussbaumer (2015), pointing out dashboard and visualization design principles.

1. *The Importance of context*

Context provides the framework for both, data storytelling and a dashboard. In dashboard design, however, there is not much space for long and elaborate text paragraphs. Yet, a well-formulated and catchy heading introduces the audience to the topic and helps it connect to already available knowledge. In my dashboard I show the economic and socioeconomic consequences of the Arab Uprising. Thus, my heading includes a short but concise question and a hint to the content of the dashboard, illustrated in Figure 1.

**Figure 2: Dashboard Heading**



1. *Choosing effective visuals*

Although undoubtedly sad, the story line of the Arab Uprising is quite simple. In 2010, anti-government demonstrations, riots, and civil unrest erupted in the Arab world, in some countries leading to conflict in 2011. This allowed for two potential economic and socioeconomic development trajectories: (1) development with conflict and (2) development without conflict. Developments over time are usually accompanied by three questions: What was before, what is now, and what will be? My dashboard aims to capture these dynamics and appropriate visuals.

I address the question “Which country faced conflict in 2011?” with a geochart, coloring only the relevant countries in a color-scale from green (no conflict) to red (high conflict intensity). Figure 3 illustrates the geochart. For further information I include the country name and conflict intensity in 2011 in the tooltip that shows when hovering over the map.

Figure : Geochart to illustrate the Conflict Intensity in different countries



But, how was the economic and socioeconomic performance in 2010 before the Arab Uprising began? Clicking a country on the geomap updates the below comparison graph, comparing the selected country to a second country of the user’s choice, the region’s average, and the world average in terms of economic or socioeconomic development. User selection allows the audience to interact with the data and information, including the individual interests of the user. Figure 4 illustrates the comparison graph. Comparisons of few items within one category are most commonly illustrated using a column chart (Abela, 2009). For my dashboard, however, I choose a bar chart to direct the user’s focus towards the right-hand side of the dashboard. Furthermore, I chose the same green that I already applied to the geochart, connecting the comparison with the idea of “no conflict”. Again, a tooltip appears on hover and reveals the “category”, selected measure, and 2010 value.

**Figure 4: Comparison Section with Connecting Passage**



Since 2010, two different development trajectories are possible for each country in the Arab World. Non-cyclical data over time, as is each development trajectory, is most effectively compared using a line chart (Abela, 2009). I color-code the “without conflict” trajectory in the same green, the “conflict” trajectory in the same red that I already used for the geochart. The four graphs automatically update their content whenever a country on the geochart is clicked or the first country from the comparison section is changed. This background and back-up information of the call for action is illustrated in Figure 5.

**Figure 5: Economic and Socioeconomic Development Since 2010 in Selected Country**



The ending of the story, or “the call for action” consists of a live calculation as explained in the data, methodology, and technology section. For the selected country, a short summary section is updated (see orange box), telling the user whether the selected country faced conflict. Furthermore, the short section indicates how the below numbers are to be interpreted: either as an actual loss, or hypothetical loss in the case of conflict. Since the calculations are country specific and the ranges are large, there is no visual that could illustrate the numbers adequately. Indexing and categorizing the calculation results would potentially allow for the use of gauge charts. From a personal perspective, I feel that this would lessen the significance of the call to action. Thus, I display lost income, lost school years. Lost weight gain, and lost children lives in plain numbers. This is illustrated in Figure 6. Additionally, I add a hover effect to the “loss boxes,” which contains a short explanation of the importance of this measure.

**Figure 6: Call for Action**



The punchline of the data story is clear. Due to conflict (hypothetical conflict), the selected country lost (could have lost) this much income, school years, weight gain, and children’s lives.

The main character of data storytelling, however, is not the researcher or developer but the audience (Nussbaumer, 2015). So far, the audience was presented with the following information: In 2010 a country had some socioeconomic performance and in 2011 eventually (no) conflict. By today, the economic and socioeconomic impact can be expressed by the calculations. Yet, what if the conflict had ended before 2019? What if the current situation would continue? What if the conflict were to end this year, how would economic and socioeconomic development look like 2, 5, or 10 years from now? In order to provide the audience with as many answers as possible, I include a filter console, which is illustrated in Figure 7.

**Figure 7: User Console - Adding Interactivity**



1. *Clutter is your Enemy*

For the individual pieces of my data story, I keep the balance in the data-ink-ratio. The data-ink-ratio “is the proportion of ink that is used to present actual data compared to the total amount of ink (or pixels) used in the entire display” (Tufte, 1983). For each graph, I only display essential information, and if necessary, “hide” more in tooltips and on hover. For example, I remove gridlines, avoid axis distortions, 3-D, and acronyms, and use consistent intervals on axis (EEA, 2019). As for the bigger picture, I only include the data information necessary to tell the story. With the data available alone a lot more graphs could be generated; however, following the law of diminishing returns, more is not necessarily better (Curtis, 2017).

1. *Focus your Audience’s Attention*

In order to focus my audience’s attention, I first apply the Gestalt Principles (Interactive Design Foundation, na), and second follow the F-pattern (Pernice, 2019). Specifically, I apply the Gestalt Principle of proximity and similarity, by arranging objects that form one informational unit closely together and formatting these objects in the same style, for example the loss calculations or the interactivity console. Furthermore, I align objects to create invisible (guide)lines for orientation (see black lines in Figure 8). To facilitate the organization of the individual objects I use a grid design, which I implement using html classes and CSS formatting.

Eye-tracking studies have shown that visitors approach websites in an F-Pattern (Pernice, 2019). This means that visitors start from the upper left corner to the right, then go back and downwards, to move slightly upwards and right again (see orange arrows in Figure 8). Therefore, I arrange the individual units of my data story in this specific order. Thus, (and as the audience would expect it), I write the heading on the very top of the dashboard to tune the audience into the topic and add the conflict information with a geochart to the top-right to emphasize the conflict as main character of my story. Below I then show the “before conflict comparison.” Since the conflict/no conflict trajectories are back-up information for the calculations, I show them to the bottom-right, and the calculation and call for action in the upper-middle of the screen. By putting the console to the top-right, or “end” of the F-pattern, I hope to keep the audience and revisit or change the story as they wish.

Additionally, I use the Gestalt principle of enclosure (using colors or boundaries to highlight or contrast information) to pick up and support the F-Pattern by adding grey and white backgrounds. The grey background guides the user the first three orange arrows. With the white background I intend to highlight the call for action and, furthermore, generate a sense of continuity (Gestalt principle) among the development trajectory graphs and calculation. Slightly darker grey lines accentuate the grey and white areas.

Additionally, the calculation results are surrounded by a border to highlight their importance.

**Figure 8: Focusing Attention throughout the Dashboard**

****

1. *Think like a Developer*

All prior steps stem from the considering the developer principles of affordance, accessibility, aesthetics, and acceptance (Nussbaumer, 2015). Including the F-Pattern in the design supports the user in grasping the basic idea of the dashboard. That is, to learn about the economic and socioeconomic impact of the Arab Uprising. Additionally, I use graphs that are commonly perceived as appropriate for the type of data I show. To address accessibility, I add text where needed, but provide the audience with more information in tooltips and on hover. Furthermore, by choosing moderate and well-combining colors, aligning objects and using white space thoughtfully I create a clean design to address aesthetics. In order to increase acceptance, I do not “break common design rules,” which would cause discomfort and resistance among the audience.

1. *Tell a Story*

Nussbaumer (2015) suggests to structure the narrative either chronologically or reversed and to apply either horizontal or vertical logic. Since the final product is a dashboard, vertical logic is the only way to go. Horizontal logic refers to communicating the overarching story in snippets, in other words, presentation-slide style. Vertical logic means to compose an interplay between words and visual, for which dashboards are predestinated. Yet, while I arrange the story units following the F-Pattern, thus arranging them chronologically, I assign the end (call to action), the most space, emphasizing this particular unit.

**Reflection**

Creating a dashboard is challenging. Trying to tell a story within one page limited to visuals and some text is particularly challenging. Working through this project, however, I learned a lot. Especially challenging was the comparison section: How can several, yet content-wise different drop-down menus be incorporated into the design? I choose to use different sizes and add build a sentence around the drop-down menus (Figure 9). Yet, this area still leaves room for future improvements.

**Figure 9: Incorporating several Drop-down Menus with different Content**



From a technical perspective, using Google Charts API’s poses limitation. First, Google geochart does not allow for custom regions. Since the focus countries of my dashboard do not form a default google region, I manipulated the chart height and width and the html div height and width. Then I created an overflow through CSS so that only the section of the map shows that I need for the dashboard. Second, while pulling the appropriate data from a JSON file for calculations in JavaScript is an easy task, creating each Google chart in an individual function adds a layer of complication. In order to capture all user changes immediately, all chart functions are nested into one overarching function (onchange/onclick). Hence, storing the data in global variables does not work as needed. For example, if the data were stored in a global function, user input would add more information to the global variable instead of updating the data. This requires to individually pull the data from the JSON file for each chart and calculation making the JavaScript file unnecessarily long. Yet, with more experience and practice, I believe, this problem can be solved.

Thus, in conclusion, skills (design-wise and technical) are acquired through experience with this project certainly facilitated.

**References**

Literature

Abela, A. (2009). Chart Suggestions – A Thought-Starter [online]. Link: <https://extremepresentation.typepad.com/files/choosing-a-good-chart-09.pdf>. Retrieved: 12/6/2019.

Anderton, C. H., & Carter, J. R. (2009). Principles of conflict economics. Cambridge Books.

Bagdikian, B. (2004). The new media monopoly (Rev. ed.). Boston, MA: Beacon.

Curtis, R. (2017). Tableau Deep Dive: Dashboard Design – Visual Best Practices [online]. Link: https://interworks.com/blog/rcurtis/2017/06/20/tableau-deep-dive-dashboard-design-visual-best-practices. Retrieved: 12/6/2019.

EEA “European Environment Agency” (2019). Chart do’s and don’ts [online]. Link: https://www.eea.europa.eu/data-and-maps/daviz/learn-more/chart-dos-and-donts#toc-2. Retrieved: 12/6/2019.

Nussbaumer, C. (2015). Storytelling with data: A data visualization guide for business professionals. John Wiley & Sons.

Pernice, K. (2019). Text Scanning Patterns. Eyetracking Evidence [online]. Link: https://www.nng https://www.nngroup.com/articles/text-scanning-patterns-eyetracking/roup.com/articles/text-scanning-patterns-eyetracking/. Retrieved: 12/6/2019.

Tufte, E. (1983). The visual display of quantitative information. CT Graphics, Cheshire, 17.

Data

IHME “Institute for Health Metrics and Evaluation” (2019). Global Burden of Disease Results Tool [online]. Link: http://ghdx.healthdata.org/gbd-results-tool. Retrieved: 12/6/2019.

INSCR “Integrated Network for Societal Conflict Research” (2018). Major Episodes of Political Violence [online]. Link: http://www.systemicpeace.org/inscrdata.html. Retrieved: 12/6/2019.

UN National Accounts (2019). The Systems of National Accounts [online]. Link: https://unstats.un.org/unsd/nationalaccount/sna.asp. Retrieved: 12/6/2019.

UN Population Division (2019). World Population Prospects [online]. Link: https://population.un.org/wpp/. Retrieved: 12/6/2019.

UNESCO (2019). Sustainable Development Goals [online]. Link: http://data.uis.unesco.org/. Retrieved: 12/6/2019.

Technology

Atom (2019). A hackable text editor for the 21st Century [online]. Link: https://atom.io/. Retrieved: 12/6/2019.

Google Developers (2019). Google Charts, Interactive charts for browsers and mobile devices [online]. Link: https://developers.google.com/chart. Retrieved: 12/6/2019.

Gretl (2019). Gnu Regression, Econometrics and Time-series Library [online]. Link: http://gretl.sourceforge.net/#dl. Retrieved: 12/6/2019.

**Appendix**

**Table 1: Panel Fixed Effects Regression**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| IV | Income | School Enrolment | Wasting | Infant Mortality |
| Const | 8.111\*\*\*(0.005) | -24.722\*\*\*(4.709) | 4.176\*\*\*(0.093) | 9.857\*\*\*(0.119) |
| Income |  | 13.008\*\*\*(0.557) | -0.337\*\*\*(0.011) | -0.788\*\*\*(0.014) |
| Conflict | -0.055\*\*\*(0.003) | -0.504\*\*\*(0.151) | 0.022\*\*\*(0.003) | 0.023\*\*\*(0.004) |
| Cross-sectional units | 163 | 158 | 163 | 163 |
| MinMax | 243 | 142 | 223 | 223 |
| N | 6282 | 3136 | 3647 | 3647 |
| R-Squared | 0.960 | 0.800 | 0.972 | 0.972 |
| Log-likelihood | -1652.374 | -11186.35 | 1673.383 | 777.915 |

Standard errors in parentheses.

\*\*\*=significant at 1 %, \*\*=significant at 5 %, \*=significant at 10 %