Process Book (CS171 Week 10)

Project Title: Ukraine Improvised Explosive Devices

Project Team

Online Studio 3 Group 3

Valérie Lavigne
valelavi@gmail.com

Marius Panga
marius.c.panga@gmail.com

Jayaram Shivas Vadakumpu-ram
shivasj@gmail.com

Team Roles

Team Coordinator: Valérie

- Producing a tasks list in Asana from the assignements for each week and making sure all the work is assigned

Code Collaborator: Shivas

- Setting up and organizing the Github repository
- Overall web app layout and designer

Data Stewart: Marius

- Identifying potentially relevant data from various sources, extracting, translating and transforming it to a format that is consistent and easy to merge with the main dataset.

Alternating Responsibilities

Role shared across team members, one team member volunteers each week:

- Team Submitter: Packaging the week's work and submitting it
- Updating the process book
- Tasks for each week are listed in Asana and each team member volunteers for the tasks he/she wants to do, tasks assignment is also discussed at the weekly meting
- Updating the various supporting documents is done in a collaborative manner, with each member contributing in an agile way with relevant input.

Project Description

Background and Motivation

Valérie works as a defence scientist for Defence R&D Canada and is the Canadian representative on the NATO Research Task Group IST-141 Exploratory Visual Analytics. Through her work, she was exposed to a dataset and presentation about the Ukraine Improvised Explosive Devices (IED) situation produced by the NATO Counter-IED Center of Excellence (NATO C-IED COE) which is an International Military Organization, multinationally manned and funded by contributions from 9 sponsoring NATO nations (http://www.coec-ied.es/). Figures 1 and 2 below show current static visualization employed by the C-IED COE to visualize this data.

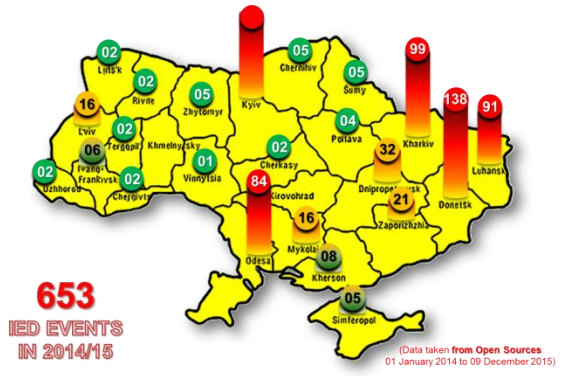


Figure 1: Map of Ukraine IED incidents in 2014-2015.

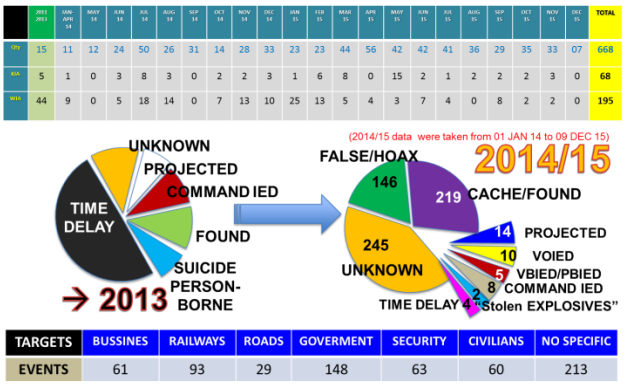


Figure 2: Statistical data about Ukraine IED incidents in 2014-2015.

Valérie, Marius and Shivas agreed that these visualizations could be improved upon using what they learned in the CS171 Visualization class. They decided to take the challenge of creating an interactive visualization for this data. The dataset is interesting because it contains many datatypes (quantitative, geographical, temporal, textual) and it can be augmented using additional data from the web. We intend to share the end result of our project with the NATO C-IED COE.

Project Objectives and Goals

This project will allow the exploration of over 600 Improvised Explosive Devices incidents in Ukraine mostly over the past 2 years for the purpose of highlighting and better understanding the temporal, geographical and political patterns in that data. We will consider additional census data and election statistics to uncover potential regional patterns, as well as publicly available data about the existing conflict in Eastern Ukraine. We intend to employ interactive visualization to generate better insights about the Ukraine IED situation.

Questions

By answering some or all of the questions below, our project aims to provide better insights into the Ukraine IDE situation:

- Is there a temporal pattern in the number of IED incidents?
- Where are the IED incidents located within Ukraine?

- How do the incidents relate to the conflict in Eastern Ukraine? Is there a relation between he number of incidents and the distance to the conflict zone and the political situation?
- Do the ratios of incidents remain stable over time between the different regions? Do the incidents seem to move from one region to another?
- How does the rate of IDE incidents relate to the total number of reported casualties and injuries?
- Do the different types of IEDs have interesting geo-temporal patterns?
- Is the number of incidents correlated with regional census data, more specifically:
    - Are there less or more incidents in Russian speaking regions?
    - Are there less or more incidents in poorer regions?
    - Is there a connections between the incidents and the political allegiance of the regions, according to the 2010 Presidential election data?
- Can we correlate any spikes or patterns in the data with political developments in the conflict (Crimea Annexation, agreed Ceasefire, Ukraine elections …)
- Is there any additional insight that can be obtained from the free-text incident description field in the main data source?

Tasks

Following is a list of tasks that could be identified in this early phase of the project. As the project progresses, this list will evolve with more tasks being added and some becoming obsolete or irrelevant.

- Data analysis: go over all the candidate data sources and selecting the bits of data that will be relevant to the project
- Data gathering: while our main data source is already in a structured format, the additional sources are not(news articles, PDF files, images). This step involves getting all the relevant data into a structured format (txt or cvs files)
- Data cleaning: make sure the data that has been identified as relevant is in a consistent format. Standardize the handling of incorrect or missing data, as well as any formatting issues
- Data filtering: based on the data analysis outcome, remove any data that is needed for the visualization
- Data model design: identify the JavaScript entities that will contain / reference the source data, and for each such entity define:
    - Relationship with other data (how will it merge with other data)
    - Properties (name and data type) and methods
    - Default sort order

Data

Data Description

NATO Ukraine IED Incidents Data

This is a NATO Unclassified IED events spreadsheet from the NATO C-IED COE. It contains 665 events, with 15 in 2001-2013, about 230 in 2014 and the rest in 2015. It contains the following column headers: Date, Type, KIA (Killed in Action), WIA (Wounded in Action), City, Region, Country, Details, Group, Remarks. There are missing values in the dataset. Table 1 below provides a sample of the data for December 2015.

Table 1: Sample of the Ukraine IED incidents data for December 2015.

DATE	TYPE	KIA	WIA	CITY	REGION	COUNTRY	DETAILS	GROUP	REMARKS
2015 DEC 09	CACHE/FOUND	0	0	Kramatorsk	DONETSK	UKRAINE	IEDs were found and disposed by Combat Engineers		TBC if they were just landmines
2015 DEC 09	UNKNOWN	0	0	Kharkiv	KHARKIV	UKRAINE	An IED was blown up in front of a ROSHEN shop		Shopping Center
2015 DEC 08	UNKNOWN	0	0	Kiev	KIEV	UKRAINE	A device was detonated against a restaurant (L'Kafa)		Boulevard Lesi Ukrainian
2015 DEC 07	HOAX/FALSE	0	0	Ivano-Frankivsk	IVANO-FRANKIVSK	UKRAINE	Call reporting IED - First Responders action - no explosive		Central Metro/bus station
2015 DEC 02	CACHE/FOUND	0	0	Krasnoarmiisk	DONETSK	UKRAINE	An IED was found and disposed by EOD		
2015 DEC 02	CACHE/FOUND	0	0	Avdeevka	DONETSK	UKRAINE	A cache with 3 IEDs with TNT was found and cleared		Inside an abandoned house
2015 DEC 02	UNKNOWN	0	0	Uzhgorod	ZAKARPATS'KA	UKRAINE	An IED was detonated against a store in Franko		

							Street		

Ukraine Census Data

We also want to include various statistics about Ukraine to see if we can find patterns between the IED events and these statistics. The data was cleansed and transformed in order to have the same grain and use the same common attributes as our primary data set (the NATO IED dataset). Transformations applied in order to facilitate the data merge with the other datasets: removed irrelevant data from the data set, converted all fields to the correct type, translated the region names to the ones used by the primary data set. The tool used to facilitate these transformations was Microsoft Power BI Desktop Designer.

1) Ukraine population by region and settlement type, as per January 1, 2013. The data was obtained from the Ukraine Census website http://database.ukrcensus.gov.ua/MULT/Database/Census/databasetree_en.asp):

Region	Rural Population	Urban Population	Total Population	Rural Pct	Urban Pct
CHERKASY	649300	753600	1402900	46%	54%
CHERNIHIV	518100	727200	1245300	42%	58%
CHERNIVTSI	549300	373500	922800	60%	40%
CRIMEA	759400	1274300	2033700	37%	63%
DNIPROPETROVSK	607300	2960300	3567600	17%	83%
DONETSK	477500	4363600	4841100	10%	90%
IVANO-FRANKIVSK	816800	593000	1409800	58%	42%
KHARKIV	625500	2288700	2914200	21%	79%
KHERSON	468900	706200	1175100	40%	60%

2) Percentage of the Ukraine population that have Russian as their native tongue, as of 2001, from the Ukraine Census website.

Region	Russian Speakers pct
CHERKASY	6.66
CHERNIHIV	10.26
CHERNIVTSI	5.27
CRIMEA	76.55
DNIPROPETROVSK	31.91
DONETSK	74.92
IVANO-FRANKIVSK	1.78

3) Ukraine population by region and education level, as of 1989. The data was obtained from the Ukraine Census website:

Region	Have no primary education	Have primary education	Have incomplete secondary education	Have secondary general education	have incomplete higher education	With higher education	Have vocational education	I
CHERKASY	6115	49616	112234	303899	6449	83259	174145	
CHERNIHIV	5829	55631	119554	274914	5641	69288	139274	
CHERNIVTSI	8066	39214	77075	190552	4912	48489	88589	
DNIPROPETROVSK	13593	96264	244973	758944	24173	301055	537767	
DONETSK	23123	158101	365255	1043418	28983	347106	747433	
IVANO-FRANKIVSK	10128	58080	103008	277113	6953	80204	145228	
				617124	27500	313490	365257	

4) Ukraine population by region and source of income, as of 1989. The data was obtained from the Ukraine Census website:

Region	work in enterprise, organization	work in farm	pension, public assistance	dependant of individuals
CHERKASY	560309	171758	382642	361194
CHERNIHIV	494631	172511	388702	315126
CHERNIVTSI	345132	108218	179730	258798
DNIPROPETROVSK	1811304	151674	768528	978164
DONETSK	2583755	115083	1087434	1328773
IVANO-FRANKIVSK	544690	128203	253116	402271
KHARKIV	1518784	108594	634470	753443

fellowship	work in cooperative enterprise	work on individual labor contract	working for individual employees
28353	2506	1007	145
25614	1955	938	96
22139	1980	1315	252
113925	10803	2592	405
133155	10594	3486	520

another kind of public support	another source	personal subsidiary economy
10935	1273	7231
7024	783	5390
9028	1181	13028
18800	3937	9726
33494	6351	9136

5) Ukraine population by region and nationality of income, as of 1989. For this exercise, only the Ukrainian and Russian nationalities were considered. The data was obtained from the Ukraine Census website:

Region	Russians	Ukrainians
CHERKASY	122308	1381742
CHERNIHIV	96562	1292106
CHERNIVTSI	63066	666095
CRIMEA	1629542	625919

Ukraine 2010 Presidential Elections Data

Since the Ukraine started with the removal from power of the previously elected president Viktor Yanukovych, we wish to provide the user with an underlying political map, based on the 2010 presidential elections. The source data has the same geographical grain as our main IED data set (region). Source: https://commons.wikimedia.org/wiki/User:DemocracyATwork

Region	Viktor Yanukovyc pct	Yulia Tymoshenko pct
CHERKASY	0.79	1.8
CHERNIHIV	0.48	1.56
CHERNIVTSI	0.48	1.15
CRIMEA	3.22	0.71
DNIPROPETROVSK	4.53	2.1
DONETSK	9.55	0.68
IVANO-FRANKIVSK	0.23	2.87
KHARKIV	4.22	1.33
KHERSON	1.26	0.71
KHMELNYTSKYI	0.75	2.1
KIROVOHRAD	0.8	1.1

Ukraine Map Data

We will a geojson/topojson data about the various regions of Ukraine as we expect some of this information might be displayed on a map. We have found two potentially suitable Ukraine map data files.

Ukraine Conflict Casualties and Injuries Data

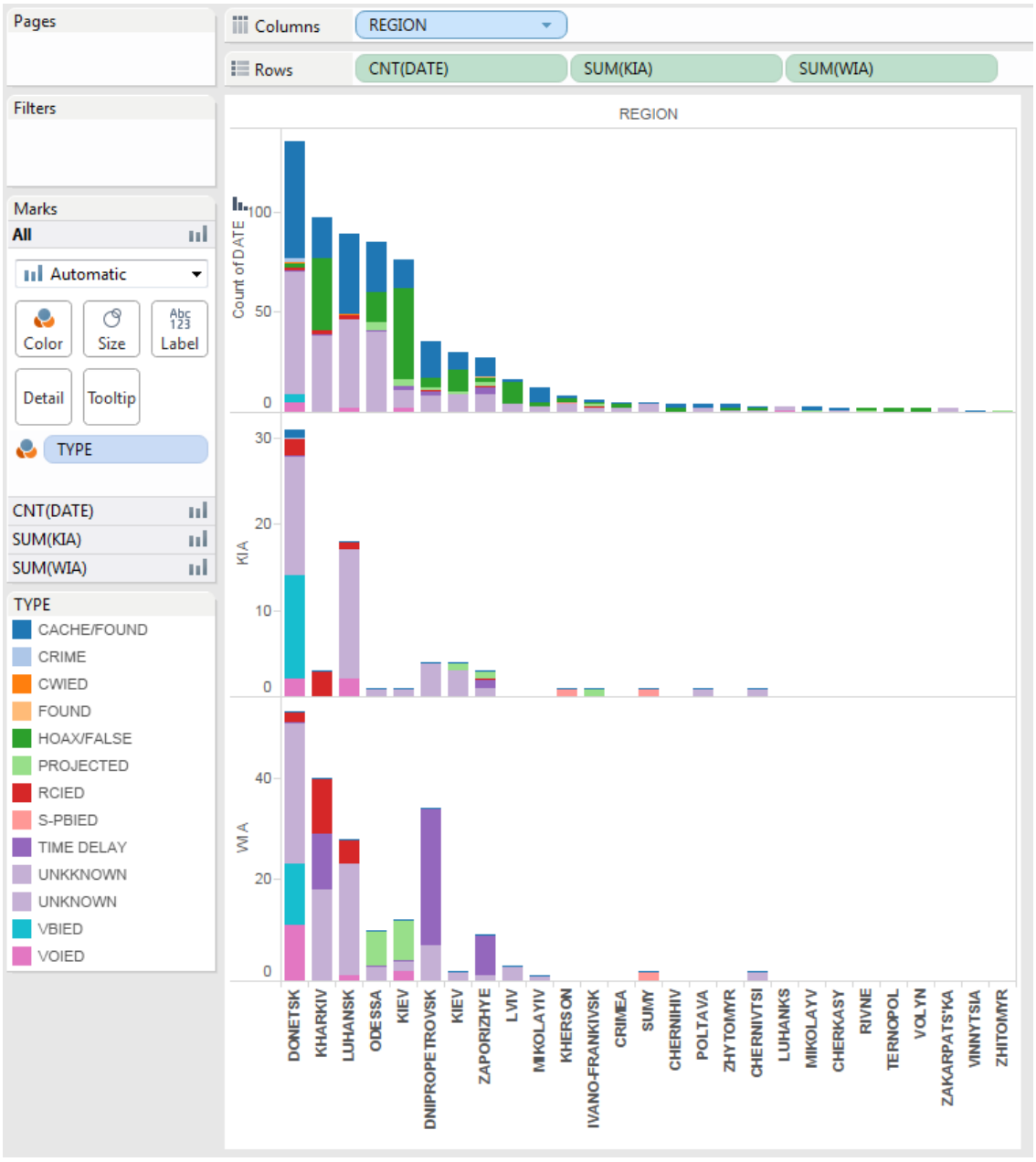
We will consider using this UN report in order to provide additional key facts about the Ukraine conflict (number of casualties and injuries by month), as well as to try to establish a relation between the number of casualties and injuries and the number if IED explosions. The data was available between 16 Feb 2015 and 15 Nov 2015. Source: http://www.ohchr.org/Documents/Countries/UA/12thOHCHRreportUkraine.pdf

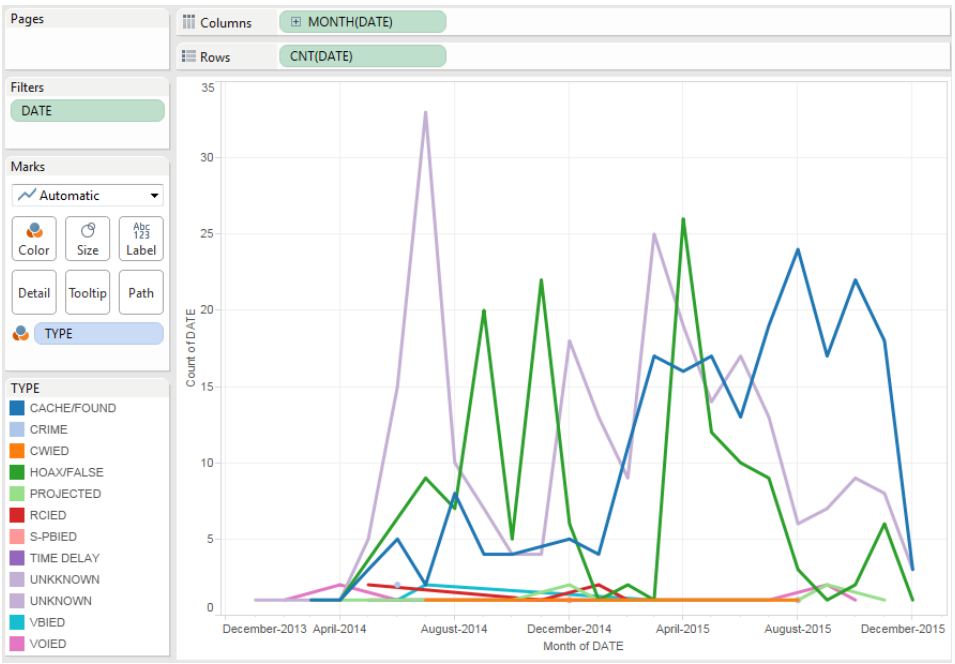
Month	Killed	Injured	Total
Feb-15	17	17	34
Mar-15	24	37	61
Apr-15	9	22	31

Exploratory Data Analysis

Tableau Data Analysis

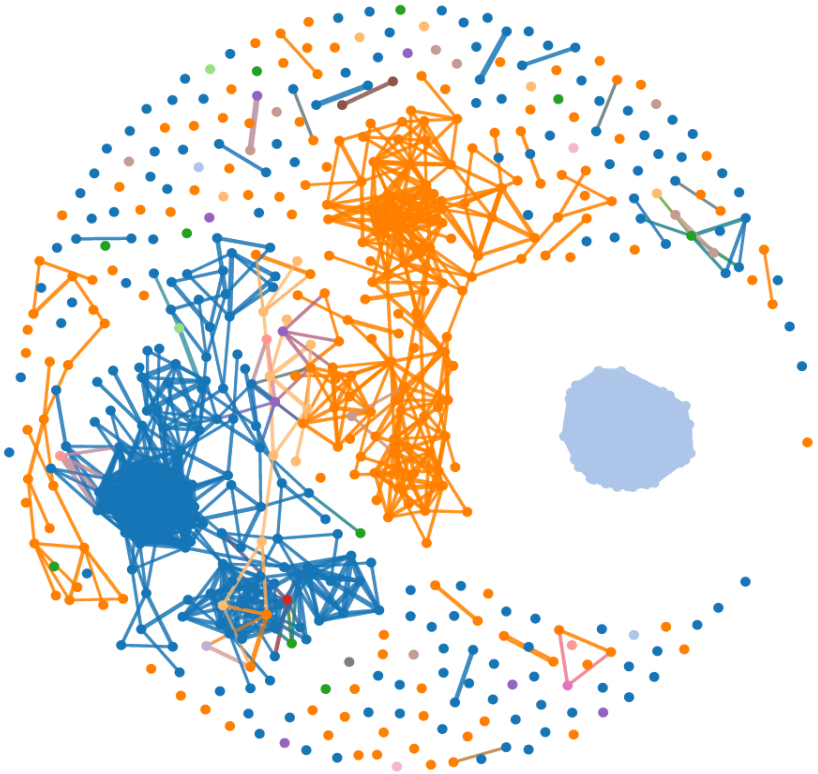
Using Tableau, we explored the number of events, of persons killed and of persons wounded by region. The events are colored by type, which also highlights the presence of data errors like the “UNKKNOWN” label. We also explored the number of events happening in the 2014-2015 time period for each event type.





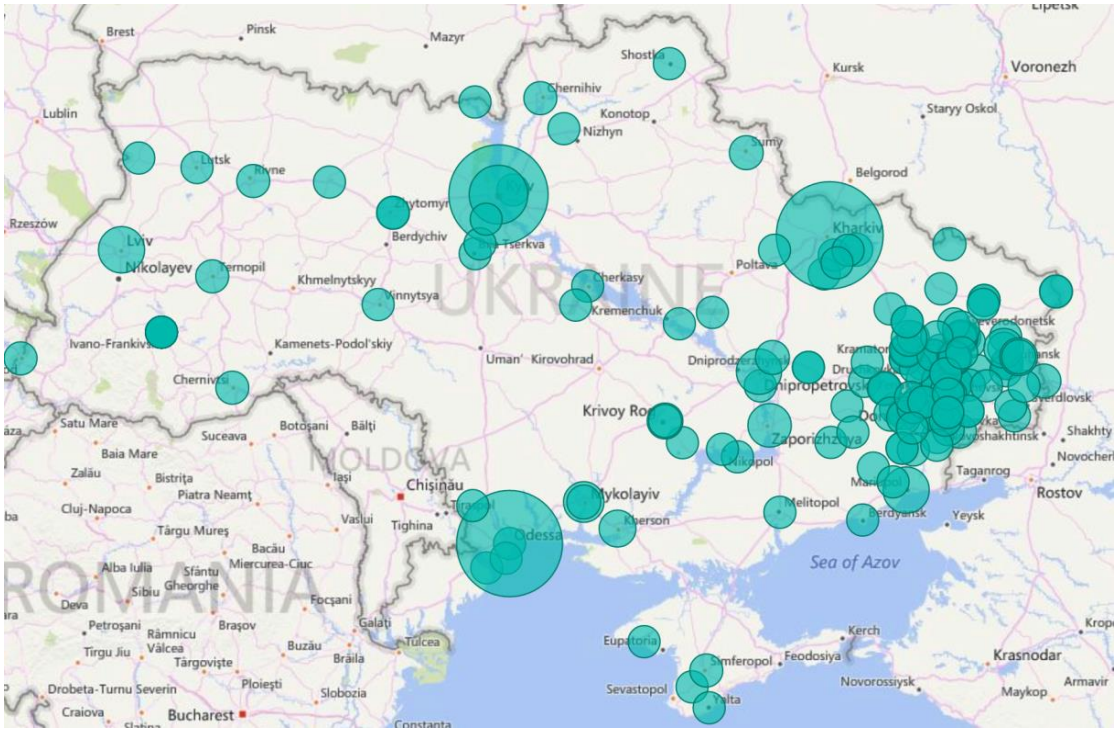
Text Exploration

We produced a force layout to find relationships between the various IED event texts (details and remarks fields) and see the ones which have most unusual words in them.

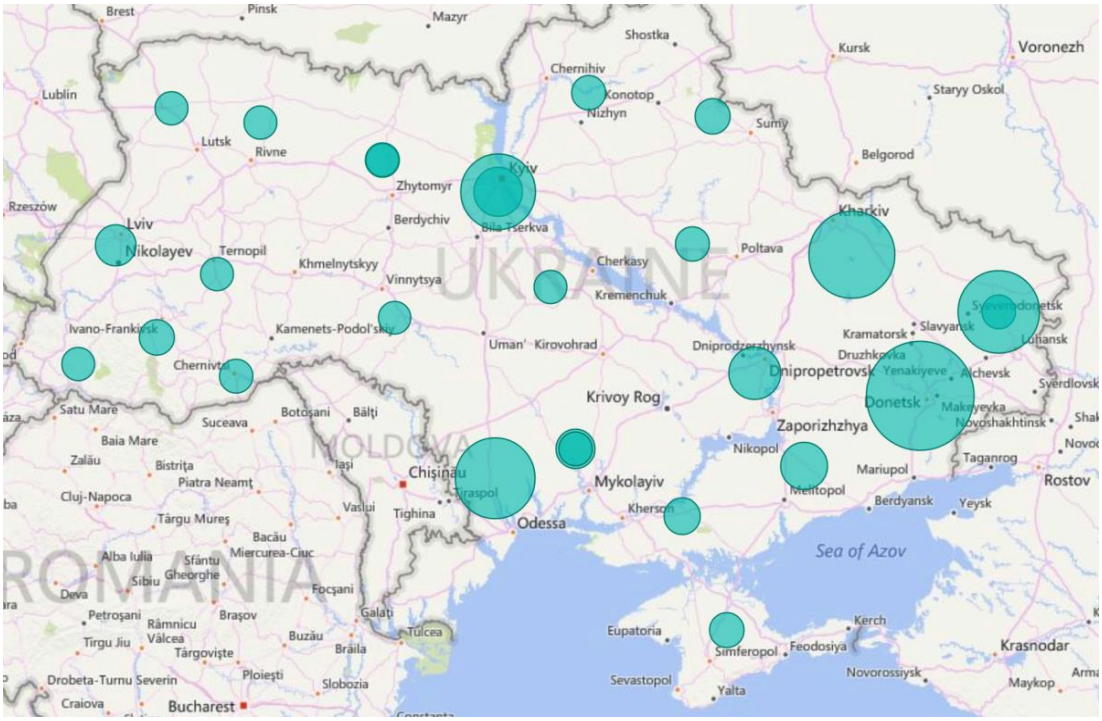


Power BI Map

Using Power BI, we visualized the number of IED events by city and by regions.



IED events by city

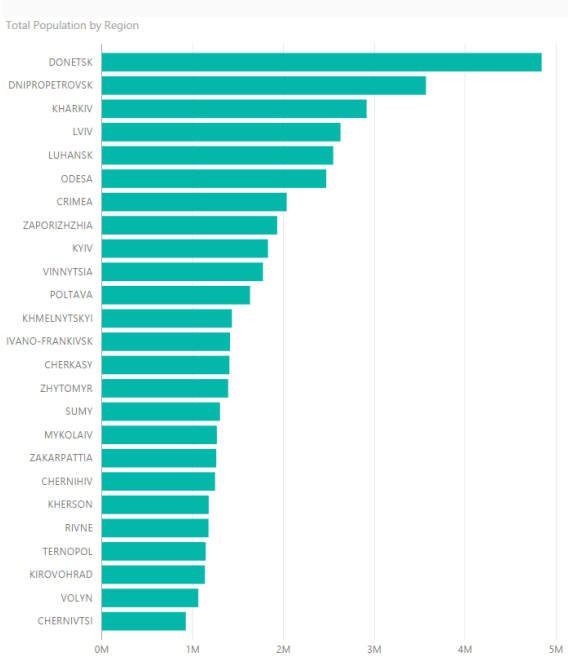


IED events by region

Power BI Statistics

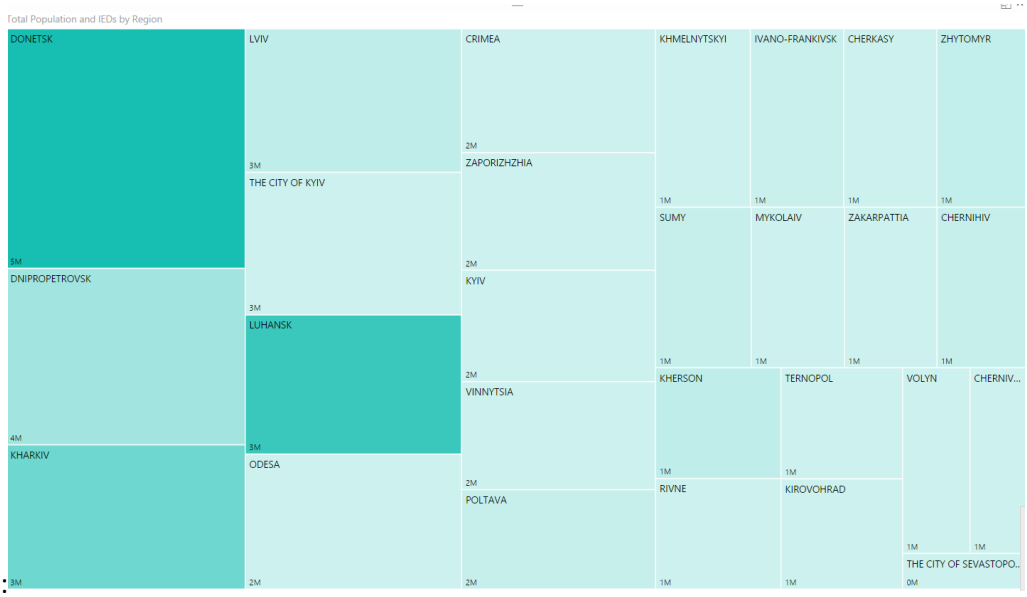
During the data collection exercise we used some graphics to explore and visualize the new dataset, as well as identify some early correlation between this dataset and the IDE dataset:

Graphical overview of the population data:

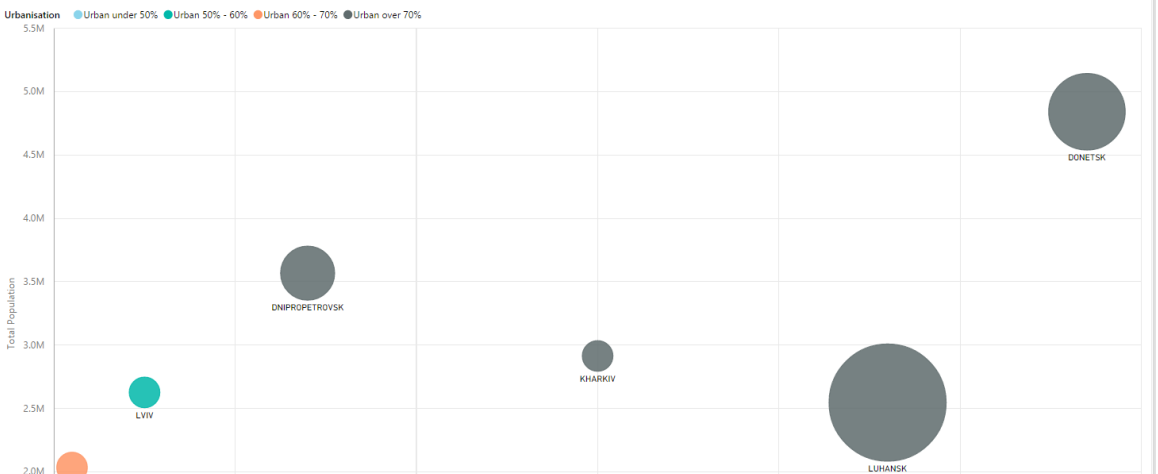


Region	Rural Population	Urban Population	Total Population	Rural Pct	Urban Pct
CHERKASY	649,300	753,600	1,402,900	46%	54%
CHERNIHIV	518,100	727,200	1,245,300	42%	58%
CHERNIVTSI	549,300	373,500	922,800	60%	40%
CRIMEA	759,400	1,274,300	2,033,700	37%	63%
DNIPROPETROVSK	607,300	2,960,300	3,567,600	17%	83%
DONETSK	477,500	4,363,600	4,841,100	10%	90%
IVANO-FRANKIVSK	816,800	593,000	1,409,800	58%	42%
KHARKIV	625,500	2,288,700	2,914,200	21%	79%
KHERSON	468,900	706,200	1,175,100	40%	60%
KHMELNYTSKYI	701,200	729,600	1,430,800	49%	51%
KIROVOHRAD	451,100	682,000	1,133,100	40%	60%
KYIV	774,400	1,053,500	1,827,900	42%	58%
LUHANSK	355,400	2,190,800	2,546,200	14%	86%
LVIV	1,067,800	1,558,700	2,626,500	41%	59%
MYKOLAIV	425,900	838,800	1,264,700	34%	66%
ODESA	844,400	1,624,600	2,469,000	34%	66%
POLTAVA	673,300	956,800	1,630,100	41%	59%
RIVNE	623,600	549,700	1,173,300	53%	47%
SUMY	456,800	842,900	1,299,700	35%	65%
TERNOPOL	656,800	485,600	1,142,400	57%	43%
THE CITY OF KYIV	0	2,611,300	2,611,300	0%	100%
THE CITY OF SEVASTOPOL' X)	21,400	358,100	379,500	6%	94%
VINNYTSIA	953,500	818,900	1,772,400	54%	46%
VOLYN	527,500	533,200	1,060,700	50%	50%
ZAKARPATTIA	792,300	466,000	1,258,300	63%	37%
ZAPORIZHZHIA	471,000	1,458,200	1,929,200	24%	76%
ZHYTOMYR	614,100	775,400	1,389,500	44%	56%

Heat map showing a potential correlation between the population of a region and the number of incidents:



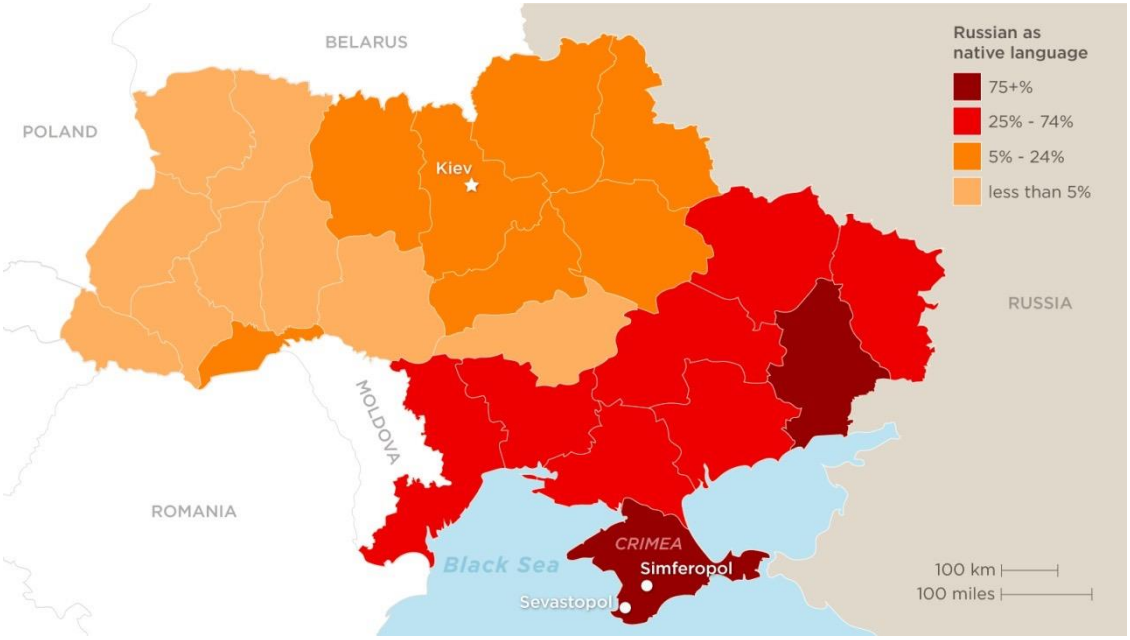
Bubble graph showing regions by number of IDEs, Population and degree of urbanization. The size of the bubbles is the fatalities resulting from the incidents.



Other Sources

Directly on the web, we found a Ukraine map showing the percentage of Russian speakers by regions. This should correspond to the data on language that is available in the Ukraine census data.

http://www.cnn.com/interactive/2014/02/world/ukraine-divided/



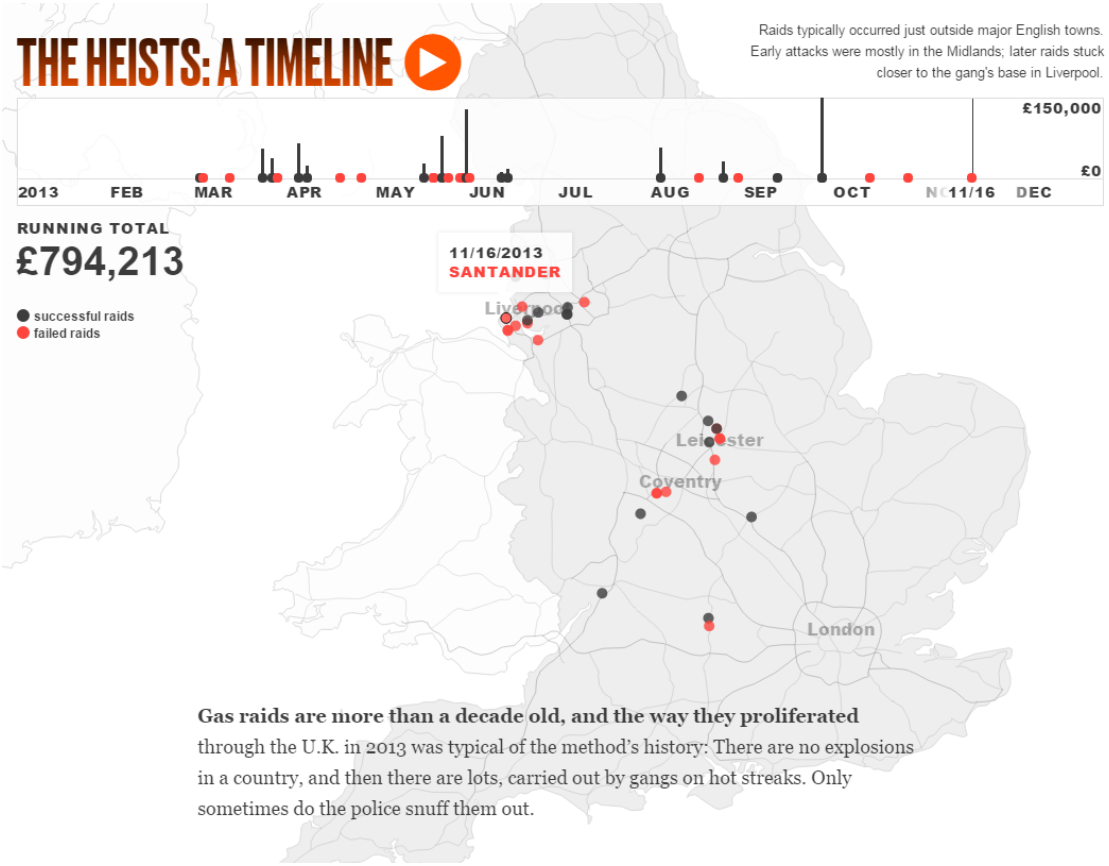
Visualization Design

Inspirations

Map and Timeline

http://www.bloomberg.com/graphics/2015-atm-bombers/

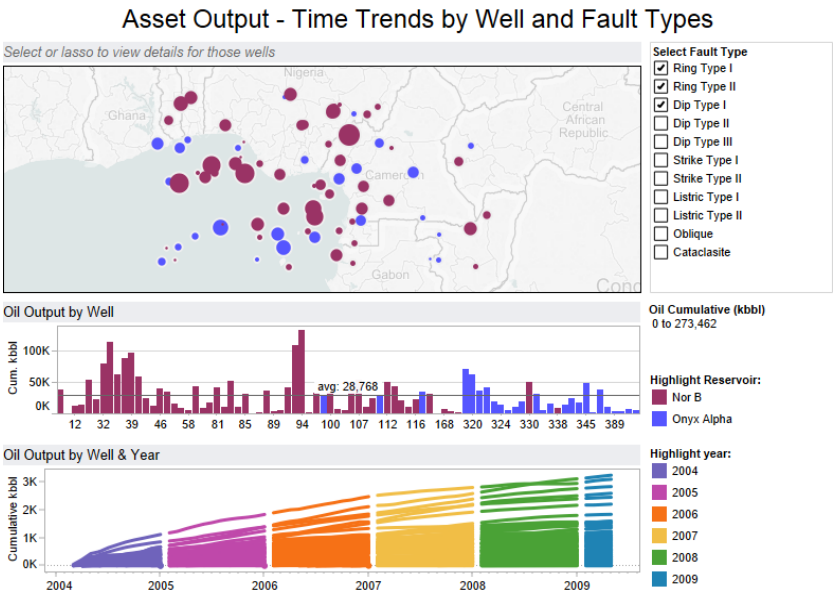
This visualization shows the geo-temporal aspects of U.K. ATM bombing incidents. The webpage itself has a very strong storytelling design.



Linked Views

http://www.kms-world.com/solutions/industries/oil-gas

This visualization links multiple data aspects together with an interesting use of colors.

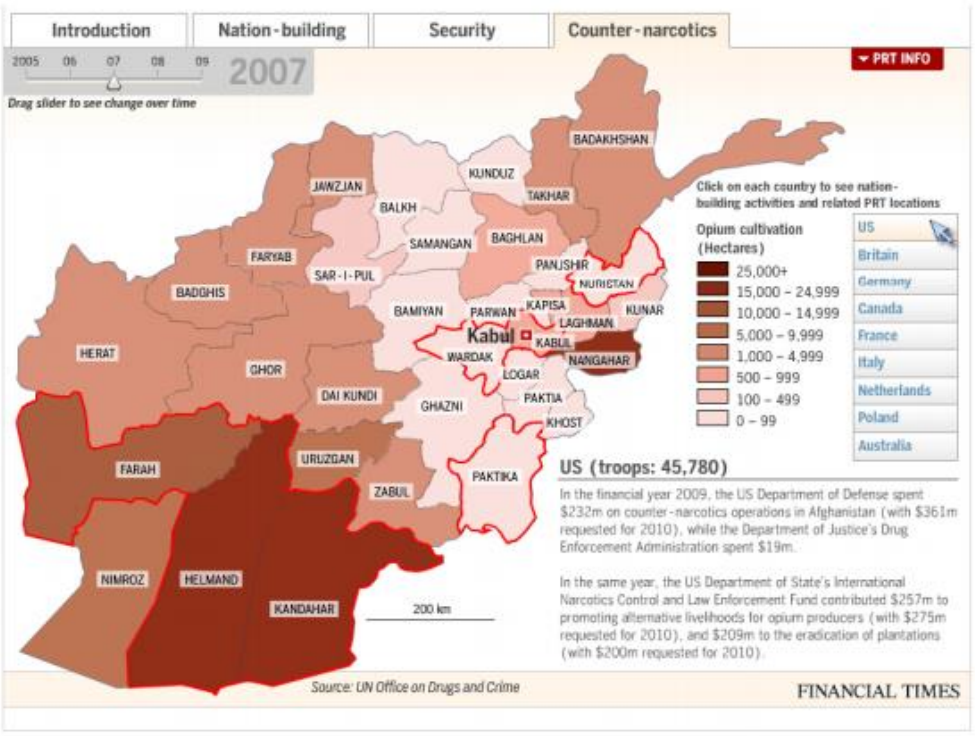


Interactive Storytelling

http://vis.stanford.edu/files/2010-Narrative-InfoVis.pdf

This visual does a great job at guiding the reader through the visualization and highlighting key events, while at the same time introducing more context and bridging the gap between story and interactive visual.

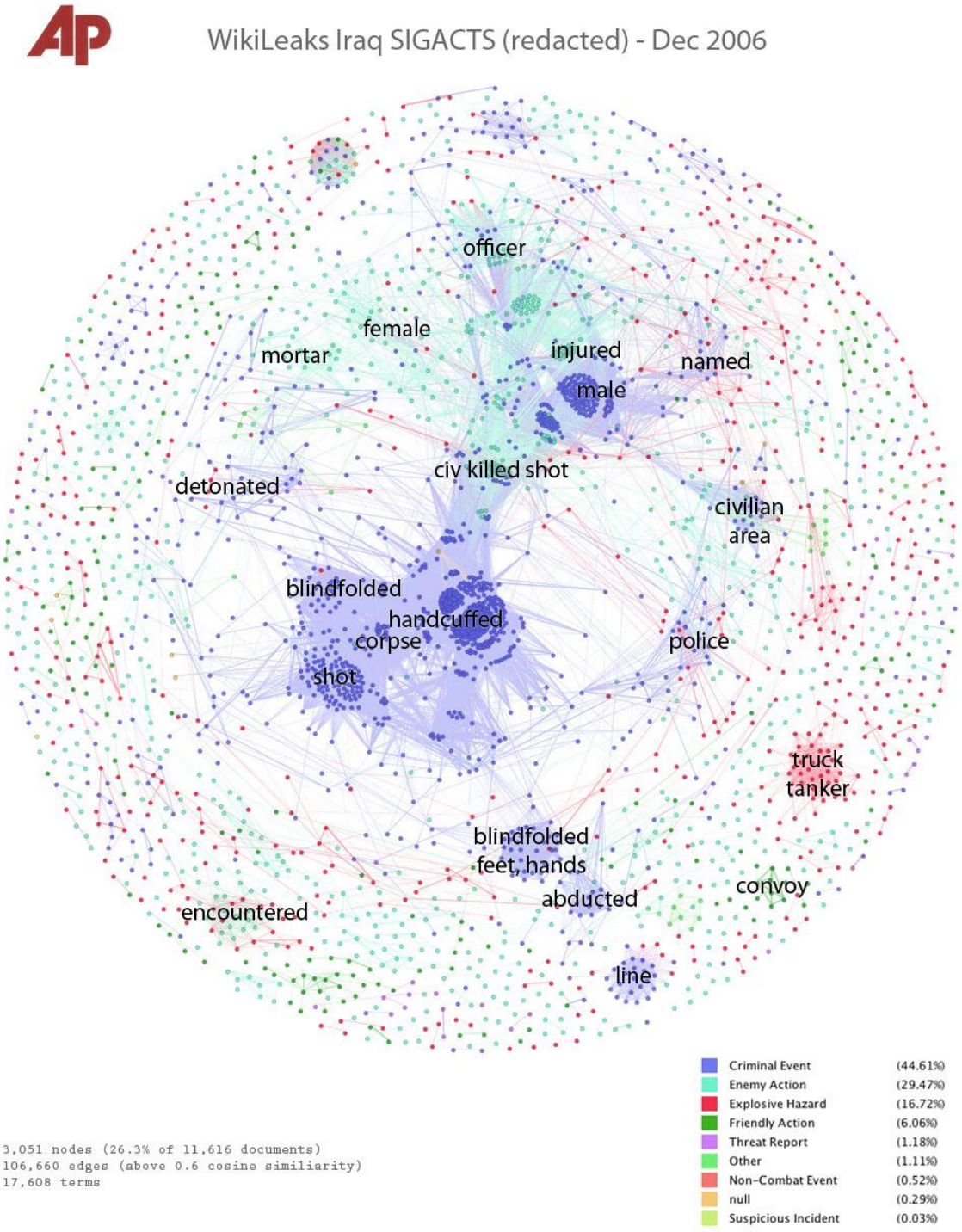
Afghanistan: Behind the Front End. Financial Times



Visualizing Text

http://jonathanstray.com/a-full-text-visualization-of-the-iraq-war-logs

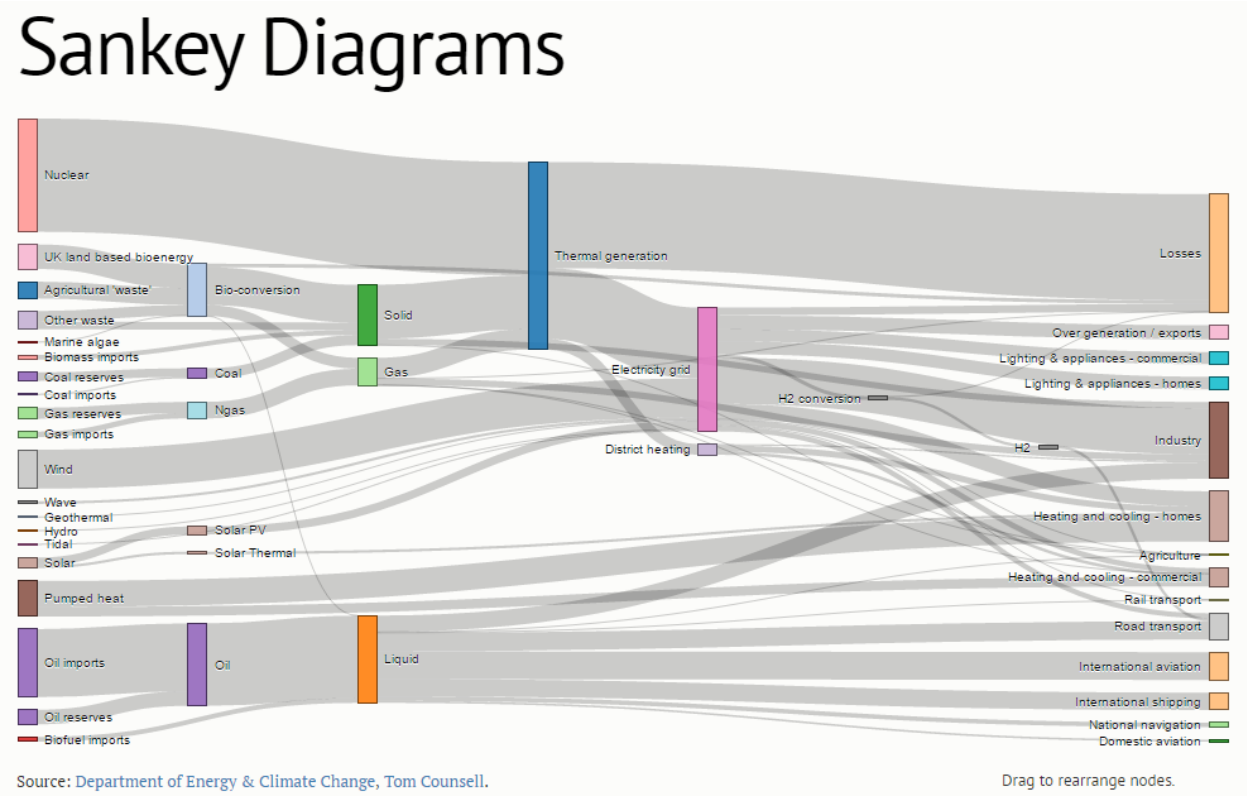
This visualization is a great way to look into the logs of text for the various incidents and get insights from textual data. It shows how various incidents could be related.



Categorical Data

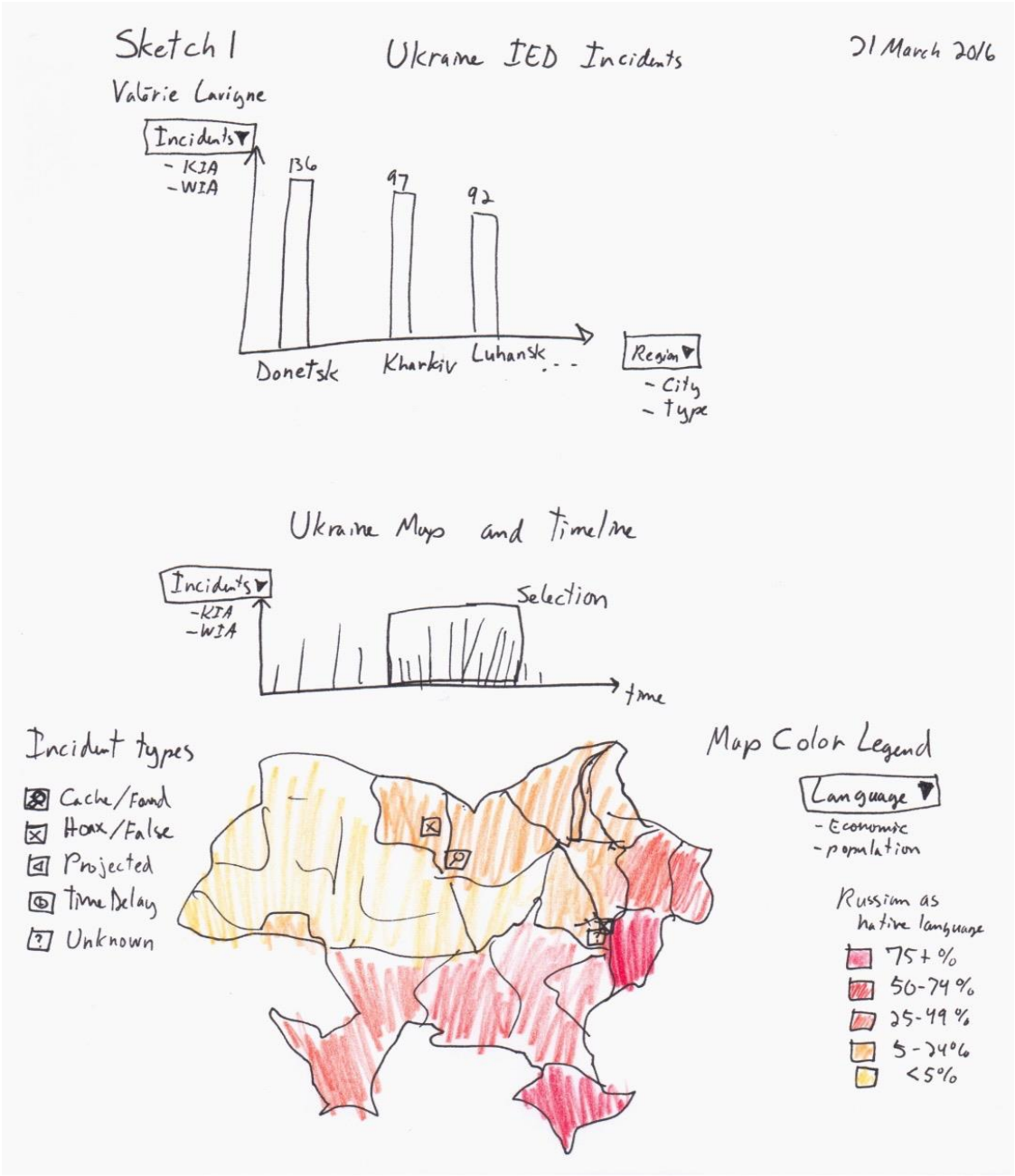
https://bost.ocks.org/mike/sankey/

A Sankey diagram of IED events types, regions, cities, numbers of killed victims and wounded persons could show patterns in the data.



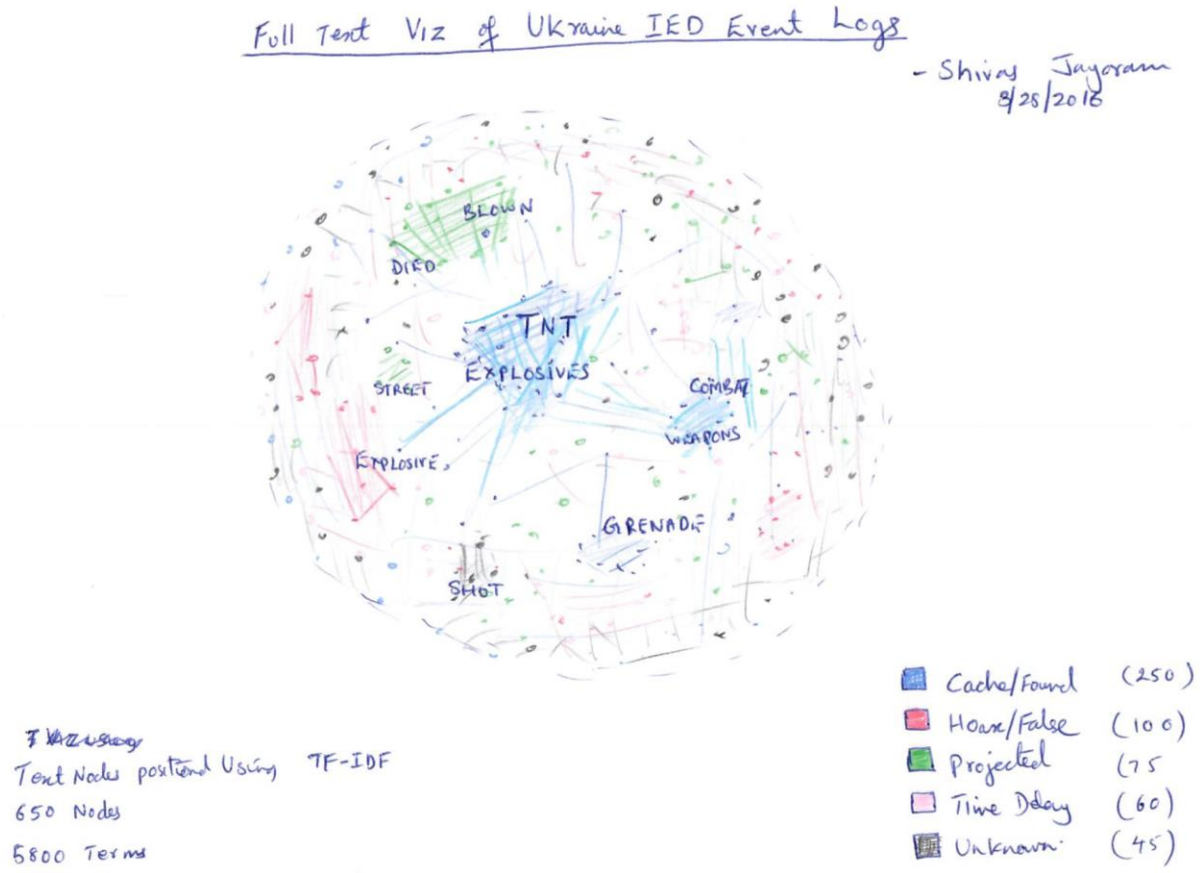
Design Evolution – First Iteration

First Iteration Sketch 1



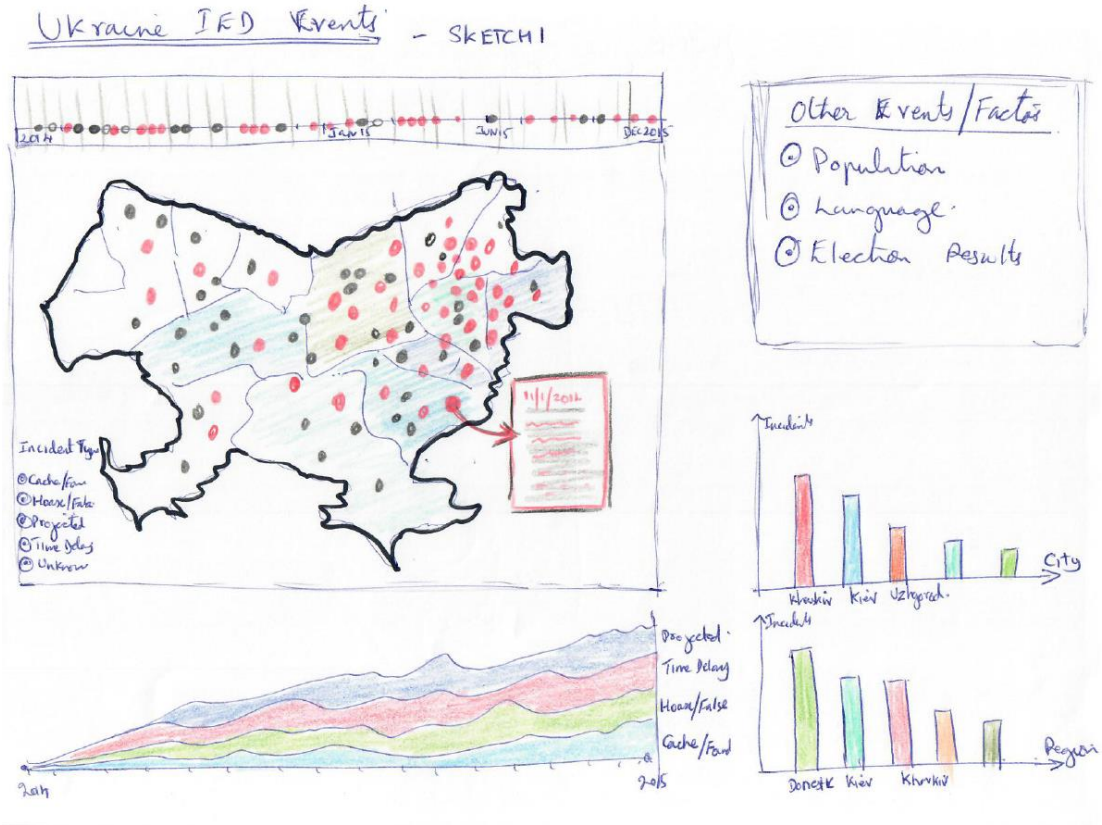
First Iteration Sketch 2

First Iteration Sketch 3



Design Evolution – Second Iteration

Sketches



Feature List

Introduction page

- Produce basic page
- Visualization static map screenshot in the background
- Add high-level KPIs / counters
- Animate high-level KPIs / counters

Map and timeline

- Find a suitable Ukraine geojson/topojson
- Create a basic Ukraine map visualization
- Find the latitude and longitude to associate with each region and city
- Display IED events as color-encoded dots on the map
- Merge data sets about regions information
- Implement map-shaped buttons to allow the user to select a specific overlay for the map (population, education …)

- Color map regions according to the user selection
- Create a IDE timeline visualization and use that to brush the main map
- Create a total IED events over time bar chart that will also act as a filter for the map
- Add IED events tooltips with more information about the event
- Add region tooltips with more information about the region
- Insert mini version of the secondary visualizations
- Prepare and create a text visualization to help with the guided navigation

Text analysis

- Parse text
- Display a word network
- Add a details view with samples of text with the selected word
- Create the log timeline

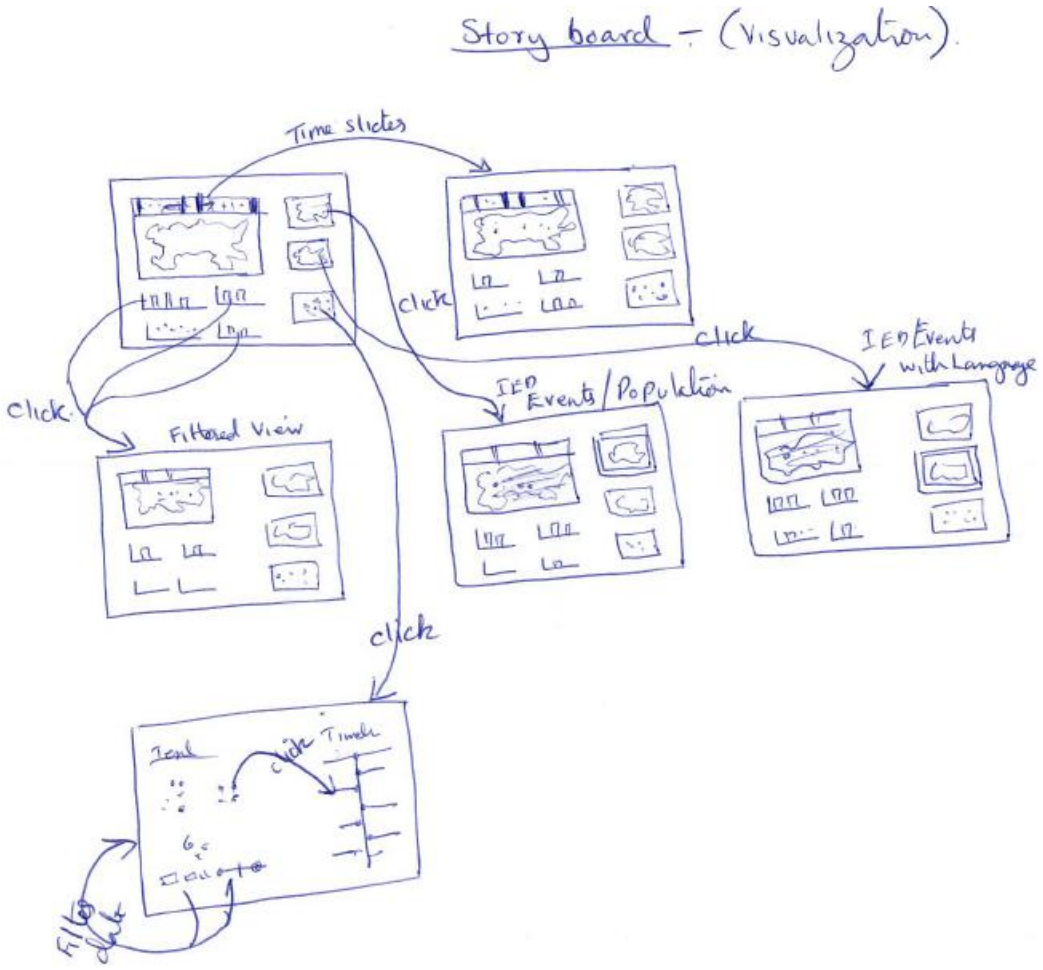
Sankey Analysis

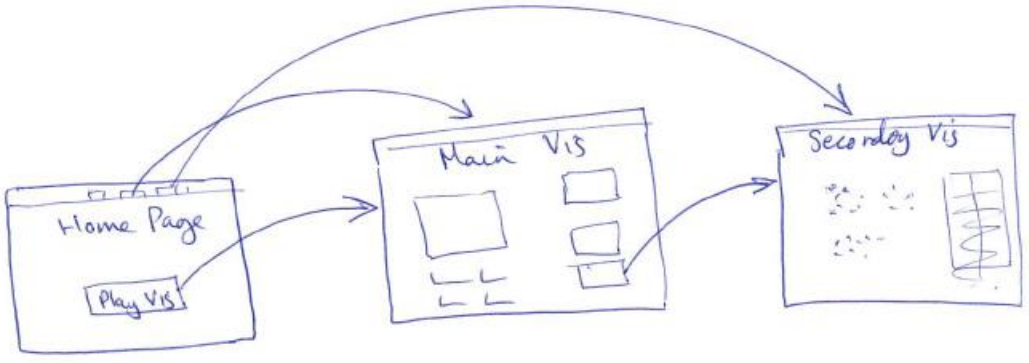
- Produce a Sankey diagram with the IED events data
- Produce additional Sankey diagrams showing other relations within the data

Storytelling

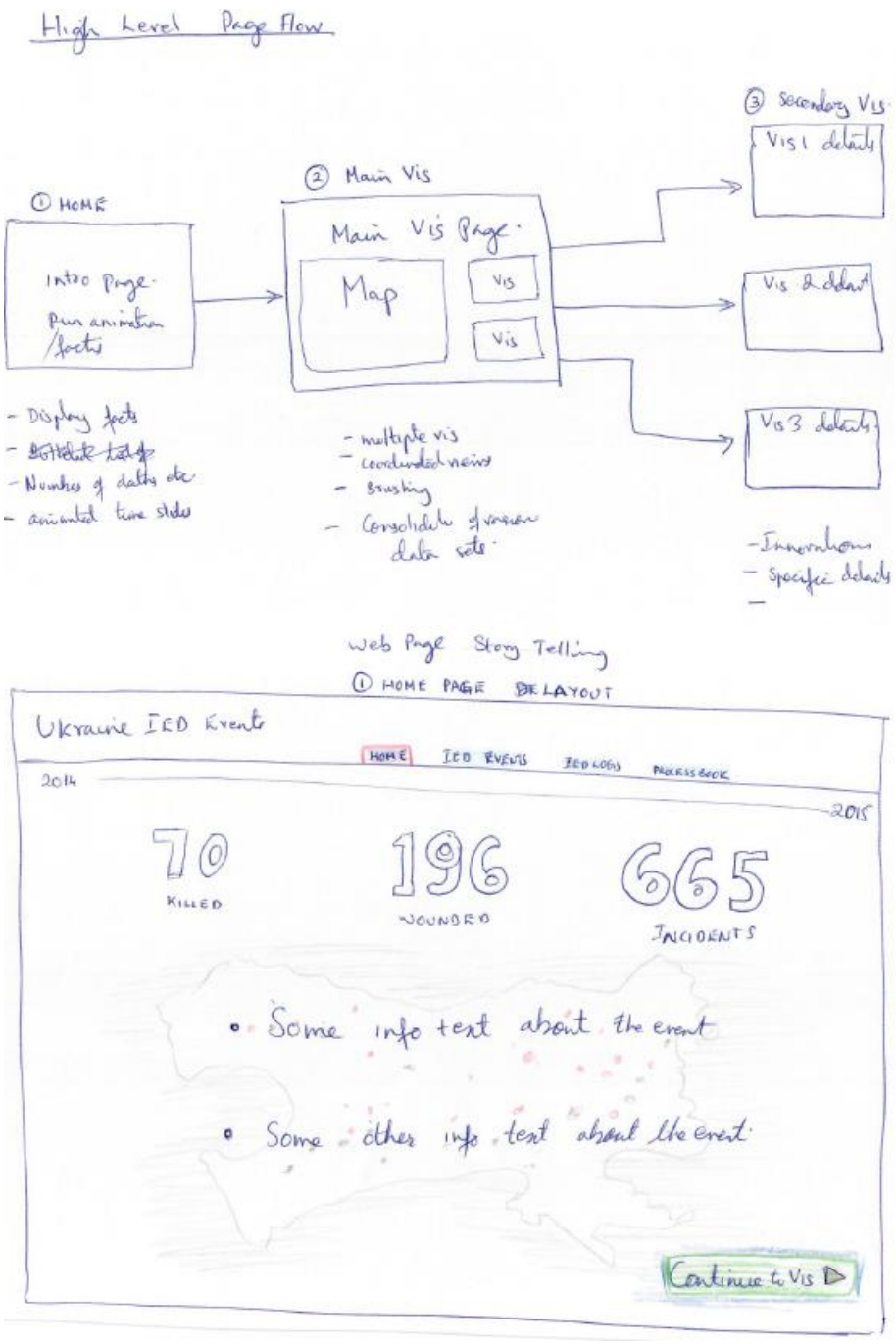
- Produce the overall webpage layout frame
- Decide on a guided exploration flow

Storyboard


Story board - (Visualization).
Time slices
click
click
IED Events with Language
click
IED Events/Population
Filtered View
click
Timeline
Filter


Story boared - (web site).
Home Page
Play Vis
Main Vis
Secondary Vis

Storytelling

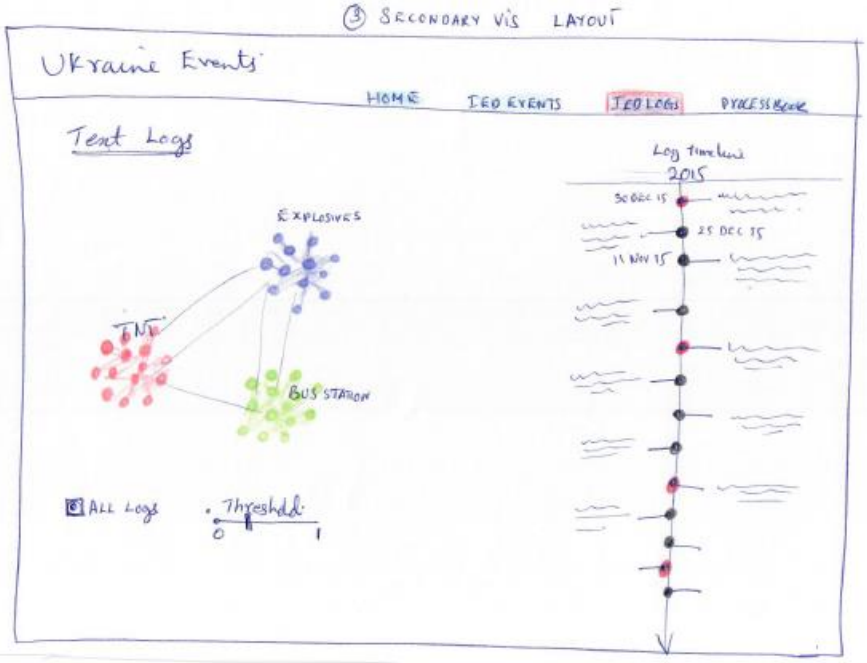


Web Page Story Telling

② Main Vis LAYOUT

Web page story Telling

③ SECONDARY VIS LAYOUT



Implementation

Project Timeline

Visualization Component	Feature	Apr 11 Final Design	Apr 18 Project V1	Apr 25 Project V2	May 2 Final Project	Must Have
Introduction Page	Produce basic page			x		Yes
	Visualization static map screenshot in the background			x		No
	Add high-level KPIs / counters			x		Yes
	Animate high-level KPIs / counters			x		No
Map and time-line	Find a suitable Ukraine geojson/topojson	x				Yes
	Create a basic Ukraine map visualization	x				Yes
	Find the latitude and longitude to asso-ciate with each region and city	x				Yes
	Display IED events as color-encoded dots on the map		x			Yes
	Merge data sets about regions infor-mation		x			Yes
	Implement map-shaped buttons to allow the user to select a specific overlay for the map (population, education …)		x			Yes
	Color map regions according to the user selection		x			Yes
	Create a IDE timeline visualization and use that to brush the main map			x		Yes
	Create a total IED events over time bar chart that will also act as a filter for the map			x		No
	Add IED events tooltips with more in-formation about the event		x			Yes
	Add region tooltips with more infor-mation about the region			x		No

	Insert mini version of the secondary visualizations		x			Yes
	Prepare and create a text visualization to help with the guided navigation				x	No
Text analysis	Parse text	x				Yes
	Display a word network		x			Yes
	Add a details view with samples of text with the selected word		x			No
	Create the log timeline		x			No
Sankey Analy-sis	Produce a Sankey diagram with the IED events data		x			Yes
	Produce additional Sankey diagrams showing other relations within the data			x		No
Storytelling	Produce the overall webpage layout frame	x				Yes
	Decide on a guided exploration flow	x				Yes