MANUAL PËR GAZETARËT

E TË DHËNAVE

GAZETARIA



Albanian Media Institute Instituti Shqiptar i Medias

GAZETARIA E TË DHËNAVE

Manual për gazetarët





Tiranë, 2018

Botues: Friedrich-Ebert-Stiftung Office Tirana Rr. Abdi Toptani Torre Drin, Kati i 3-të Kutia Postare 1418 Tiranë, ALBANIA

Autor: Rrapo Zguri

Redaktor: Lutfi Dervishi

Kopertina: Bujar Karoshi

Opinionet, gjetjet, konluzionet dhe rekomandimet e shprehura në këtë botim janë të autorëve dhe nuk reflektojnë domosdoshmërisht ato të Fondacionit Friedrich Ebert.

Publikimet e Fondacionit Friedrich Ebert nuk mund të përdoren për arsye komerciale pa miratim me shkrim.

1 Ç'JANË TË DHËNAT DHE Ç'ËSHTË GAZETARIA E TË DHËNAVE

Të dhënat janë një pjesë e rëndësishme e informacionit të pafund që gjenerohet, qarkullon dhe ruhet gjatë veprimtarisë së përditshme të institucioneve. organizatave apo individëve. Të dhëna krijohen në të gjitha hallkat e menaxhimit të të ardhurave apo shpenzimeve të buxhetit të shtetit apo të kompanive, kur organizohen zgjedhje krijohen lokale apo parlamentare, kur ndodhin ngjarje apo fatkeqësi, kur ndërmerren sondazhe apo anketime, kur regjistrohen mjaft tregues të jetës shoqërore apo kur maten rezultatet e një aksioni civil, etj, etj. Lindjet, vdekjet, fatkeqësitë, arritjet, korrupsioni, aksidentet - të gjitha këto dhe shumë e shumë të tjera - prodhojnë të dhëna. Niveli i cilësisë dhe i përgjegjësisë shoqërore për mbledhjen, menaxhimin, përpunimin, ruajtjen dhe aksesimin e të dhënave, bën diferencën mes vendeve dhe sistemeve të ndryshme shoqërore. Sa më i lartë ky nivel, aq më demokratike dhe më e hapur është një shoqëri dhe një vend i caktuar.

Zhvillimet e dekadave të fundit në fushën e teknologjisë së informacionit kanë rinxjerrë në pah rëndësinë e të dhënave për shoqërinë në përgjithësi dhe për gazetarinë në veçanti. Njëherësh, këto zhvillime kanë ndikuar edhe në një ridimensionim dhe një zgjerim të konceptit mbi "të dhënat". Deri para 20 vjetësh, të dhënat ishin kryesisht koleksione numrash të ruajtura në tabela dhe të printuara në letër. Kurse sot, në epokën dixhitale, është bërë e mundur që pothuajse gjithçka të shprehet me numra dhe gjithnjë e më shumë këto numra të krijohen, zhvillohen e ruhen në format dixhital. Në thelb, çdo lloj informacioni dixhital, qoftë ky tekst, video, foto, audio apo animacion, etj., formohet përmes kodit binar, si një kombinim i numrave 1 dhe 0.

Disa nga përparësitë e dixhitalizimit të të dhënave janë:

- 1. Të dhënat janë bërë lehtësisht të përpunueshme
- 2. Të dhënat janë bërë lehtësisht të riprodhueshme
- 3. Të dhënat janë bërë lehtësisht të aksesueshme

Të dhënat janë gjithnjë e më të rëndësishme jo vetëm për funksionimin demokratik të shoqërisë në përgjithësi, por edhe për gazetarinë dhe përmbushjen e misionit publik të saj. Veçanërisht në kohën e sotme, kur ka një prani të shtuar të "fake news" dhe të gazetarisë së bazuar tek opinioni, të dhënat marrin vlerë jetike dhe mund të kontribuojnë në përmbushjen e standarteve të gazetarisë dhe në rikthimin e besimit të publikut në të.

Ç'janë të dhënat?

Të dhënat janë një përmbledhje e pjesëve të veçanta të infomacionit, të cilat janë regjistruar dhe strukturuar në një mënyrë që e bën të lehtë për t'u analizuar. Të dhënat zakonisht gjenerohen kur diçka matet apo regjistrohet. Të dhënat duhet të shihen si një material bruto që duhet të përpunohet për të prodhuar informacion kuptimplotë.¹

Të dhënat zakonisht gjenden në tabela. Përmbajtja e një tabele të vetme me të dhëna formon një **dataset**. Tërësia e dataseteve që lidhen me një çështje apo tematikë të caktuar formon një **databazë**.

Përveçse ka shtuar sasinë e të dhënave në dispozicion, dixhitalizimi ka rritur edhe mundësinë për një transparencë gjithnjë e më të madhe të këtyre të dhënave. Kërkesa për "*open data*" (të dhëna të hapura) vjen gjithnjë e më shumë në rritje në vendet demokratike dhe bën diferencën në lidhje me nivelin dhe cilësinë e demokracisë. Të dhënat e hapura janë ai lloj të dhënash që mund të përdoren, modifikohen apo ndahen lirshëm nga çdokush dhe për çdo qëllim.

Në vendet perëndimore dhe në një pjesë të vendeve në proces demokratizimi janë krijuar tashmë portalet e të dhënave të hapura (open data portals). Edhe disa organizata ndërkombëtare, si Banka Botërore, OKB, Organizata Botërore e Shëndetësisë, etj., kanë hapur portale të tilla. Këto portale janë një burim i rëndësishëm dhe i pasur informacioni për gazetarët si dhe për organizatat e shoqërisë civile.

Gjithsesi, pavarësisht nga këto inisiativa, në vende të tilla si Shqipëria mbetet ende shumë për të bërë në fushën e të dhënave. Pjesa më e madhe e informacionit shtetëror është ende e padixhitalizuar. Ka gjithashtu shumë probleme lidhur me aksesin e të dhënave dhe me transparencën e tyre.

¹ http://www.odecanet.org/wp-content/uploads/2016/10/ Module-1-UNDP-student.pdf

Ç'është gazetaria e të dhënave?

Një përgjigjje e thjeshtë për këtë pyetje mund të ishte: është gazetaria që bëhet me të dhëna. Por në fakt gazetaria e të dhënave është më shumë se kaq. Gazetaria e të dhënave përfshin përdorimin e të dhënave për të raportuar për ngjarje dhe dukuri në mënyrën më të mirë të mundshme duke përdorur teknikat më të mira të gazetarisë: përfshi vizualizimin, paraqitjen konçize dhe teknologjitë më të fundit.²

Për ekspertin e gazetarisë së të dhënave Jonathan Gray, ajo që e bën të ndryshme gazetarinë e të dhënave është ndoshta prania e mundësive të reja që shpërfaqen kur ju kombinoni "nuhatjen" tradicionale për lajmin dhe aftësinë për të treguar një histori bindëse, me gamën e gjerë të informacionit dixhital në dispozicion.

Sipas Gray, të dhënat mund të jenë burim për gazetarinë, por mund të jenë edhe një mjet për të treguar për një ngjarje, ose të dyja bashkë. Nga ana tjetër, si çdo burim, të dhënat duhet të trajtohen me skepticizëm dhe si çdo mjet ne duhet të jemi të vetëdijshëm lidhur me mënyrën se si ai i formëson apo i kufizon ngjarjet që paraqiten nëpërmjet tij.³

Të dhënat e ndihmojnë gazetarin për të provuar hipotezat që ai ngre kur nis të merret me një ngjarje apo dukuri. Ato gjithashtu ndihmojnë për të evidentuar prirjet e ndryshme në fusha të ndryshme, për të ballafaquar raportimet zyrtare me realitetin, për

² Po aty

³ http://datajournalismhandbook.org/

të zbuluar kontrastet mes vendeve apo grupeve të ndryshme shoqërore, për të zbuluar informacione të fshehura qëllimisht, për të nxjerrë në dritë lidhjet e padukshme të problemeve me individë të caktuar në funksione publike, etj.

Të dhënat e pasurojnë raportimin e një gazetari dhe e bëjnë atë më interesant dhe më të vlefshëm për lexuesit. Ato i mundësojnë gazetarëve që të shkojnë në rrënjët e problemeve dhe krahas ekspozimit të kontribuojnë edhe në zgjidhjen e tyre.

2 FORMATET E TË DHËNAVE DHE UNIFIKIMI I TYRE

Të dhënat mund të gjenden në format dixhital, por një pjesë e madhe e tyre mund të jenë akoma në format analog apo konvencional, përkatësisht të ruajtura në letër, film apo kaseta.

Nga ana e tyre të dhënat dixhitale gjenden në dy lloje të përgjithshme formatesh:

- *a. Formatet e strukturuara*, të aksesueshme nga kompjuteri dhe të përpunueshme nga aplikacionet përkatëse, siç janë formatet XLS dhe CSV
- *b. Formatet e pastrukturuara*, siç janë PDF, Word, JPG, PNG, etj

Duhet të themi që në fillim se formati më i mirë për përpunimin dhe analizën e të dhënave është formati burimor i programit Excel, XLS. Për këtë arsye, duhen bërë të gjitha përpjekjet dhe të ndërmerren operacionet e duhura që të dhënat që gjenden në formatet e tjera të konvertohen apo të sillen në këtë format. Aktualisht shumica e gazetarëve të kësaj fushe përdorin pikërisht programin Excel për përpunimin dhe analizën e të dhënave. Më poshtë po paraqesim rrugët dhe mjetet për sjelljen në Excel të formateve të tjera të të dhënave.

Hapja e dhe ruajtja e dokumenteve të formatit CSV në Excel

Nëse do të hyni në databazat online të të dhënave,

shumica e dataseteve të tyre gjenden në formatet XLS dhe CSV. Formati CSV (*comma-separated values*) është një format i thjeshtëzuar teksti që përmban vlera të cilat janë të ndara nga njera-tjetra me presje. Në mënyrë që të lehtësohet përpunimi dhe analiza e të dhënave që janë në CSV, duhen konvertuar apo hapur dhe më pas ruajtur ato në Excel në formatin XLS.

Për hapjen dhe ruajtjen e dokumenteve të formatit CSV në Excel, ndiqet kjo rrugë:

1. Hapni programin Excel dhe hapni një *spreadsheet* të ri. Më pas tek menutë klikoni tek "Data" e më tej "From text", si në figurën më poshtë.



2. Në dritaren që hapet, gjeni *file*-in tuaj në CSV dhe më pas klikoni tek butoni "Import" si në figurë



- 3. Në dritaren në vijim, pasi siguroheni që keni zgjedhur "Delimited" dhe "Windows (ANSI)", klikoni tek butoni "Next".
- 4. Në dritaren tjetër, bëni kujdes që të keni selektuar "Tab" dhe "Comma" dhe klikoni tek "Finish".

Dokumenti që në origjinë ishte në format CSV tani hapet në Excel dhe vlerat apo kategoritë vendosen secila në një qelizë më vete. Ju tani mund ta ruani më tej dokumentin në format Excel, dhe të vijoni punën me përpunimin dhe analizën e të dhënave.

Konvertimi i formatit PDF në Excel

Shpeshherë ndodh që tabelat me të dhëna gjenden vetëm në format PDF. Ju mund t'i hidhni në Excel një e nga nga një dhe manualisht këto të dhëna, por kjo gjë do t'ju merrte një kohë të gjatë dhe do të ishte e ekspozuar ndaj gabimeve të mundshme. Për të dalë nga kjo situatë, ju mund të përdorni disa rrugë ekzistuese për konvertimin e PDF në Excel. Për këtë mund t'ju shërbejnë fare mirë disa platforma në web, mes të cilavë njëra nga më të përdorurat është www.cometdocs.com. Ja rruga që duhet të ndiqni

- 1. Hyni tek <u>www.cometdocs.com</u>. Në dritaren që hapet klikoni tek "Go to web app" dhe më pas tek "Upload" dhe ngarkoni *file*-in tuaj PDF
- 2. Klikoni tek "Convert" dhe më pas tërhiqni *file*-in PDF në dritaren që hapet poshtë "Convert"
- 3. Më pas klikoni tek "to Excel", shkruani në të djathtë e-mailin tuaj, klikoni tek butoni "Convert" dhe në e-malin tuaj do t'ju dërgohet dokumenti i konvertuar në Excel.

Upload Import files from			
shkaqet-e-vdekjev			
Drag files to t	he left field, select a conversion type, enter ye	our email addr	ess and click Convert.
Drag files to t	he left field, select a conversion type, enter yo Transfer	our email addr Store	ess and click Convert. Host
Drag files to t Convert	he left field, select a conversion type, enter ye Transfer to Word (docx)	our email addre	Enter your email
Convert Convert	Transfer To Excel (xlsx) to Excel (xlsx)	Store	Host Interpretation Interpre
Convert Convert shkaqet-e-vdekjev	Transfer	Store	Enter your email I'd like to receive more information about PDF news, tips, and offers
Convert Convert	Transfer Transfer to Word (docx) to Excel (xlsx) to DWG (dwg) to Powerpoint (pptx)	Store	Host Host Interpret Provide the second secon
Convert Convert	Transfer Transfer to Word (docx) to Excel (xisx) to DWG (dwg) to Powerpoint (pptx) to DXF (dxt)	Store	ess and click Convert. Host Enter your email I'd like to receive more information about PDF news, tips, and offers Convert

Scraping si rrugë për nxjerrjen e të dhënave nga faqet e webit

Në internet ka një pafundësi tabelash me të dhëna, të cilat duhen sjellë në Excel për t'u përpunuar e analizuar më tej. Si edhe për të dhënat në format PDF, hedhja manuale dhe një e nga një e të dhënave do t'ju merrte një kohë të gjatë dhe do të ishte e ekspozuar ndaj gabimeve të mundshme. A ka ndonjë rrugë më të shkurtër? Ka disa, por më e përdorura është ajo përmes aplikacioneve që mundësojnë "*data scraping*", që nënkupton nxjerrjen e të dhënave.

Për të bërë të mundur "*scraping*", duhet instaluar në kompjuterin tuaj një aplikacion i cili është një lloj zgjatimi i *browser*-it. Për ata që përdorin browser-in Google Chrome, ky aplikacion gjendet në linkun:

https://chrome.google.com/webstore/detail/scraper/ mbigbapnjcgaffohmbkdlecaccepngjd?hl=en Pasi hyni në këtë link hapet dritarja "Scraper" në të cilën ju klikoni tek butoni + ADD TO CHROME. M**ë** pas klikoni tek butoni "Add extension" dhe aplikacioni **ë**sht**ë** instaluar n**ë** kompjuterin tuaj.

Për të aplikuar *"scraping"* në një rast konkret ndiqni këtë rrugë:

- Shkoni tek: <u>http://data.un.org/CountryProfile.</u> <u>aspx?crName=Albania</u>, dhe më pas klikoni tek "Economic Indicators"
- 2. Selektoni një pjesë të tabeles që hapet

C data.un.org/en/iso/al.html			\$ G
Albania			
Economic indicators			
	2005	2010	2017
GDP: Gross domestic product (million current US\$)	8 052	11 927	11 541 ^b
GDP growth rate (annual %, const. 2005 prices)	5.5	Copy Ctri+C	2.6 ^b
GDP per capita (current US\$)	2 612	Search Google for "2005 2010 2017 GDP: Gross domestic product (" Print Ctrl+P	3 984.2 ^b
Economy: Agriculture ^d (% of GVA)	21 👌	Scrape similar	22.4 ^b
Economy: Industry ^d (% of GVA)	28	Inspect Ctrl+Shift+1	26.0 ^b
Economy: Services and other activity ^d (% of GVA)	49.8	50.7	51.7 ^b
Employment: Agriculture ^e (% of employed)	49.3	42.6	41.4
Employment: Industry ^e (% of employed)	17.8	19.9	18.3
Employment: Services ^e (% of employed)	32.9	37.5	40.3
Unemployment (% of labour force)	13.8 ^e	14.2	15.8 ^e

3. Bëni një klikim me të djathtën dhe klikoni tek "Scrape similar". Do t'ju hapet një dritare tjetër me të dhëna të organizuara në tabelë. Klikoni tek "Export to Google Docs" dhe tabela me të dhëna do të ruhet në profilin tuaj në Google Drive. Ju tani mund ta hapni atë në Google Speadsheets ose në Excel.

🖋 Scraper - data.un.org/en/iso/al.html	Scraper - dataun.org/en/iso/al.html						
data.un.org/en/iso/al.html			Column 1	Column 2	Column 3	Column 4 🔺	
9	1	1	GDP: Gross domestic product (million current US\$)	8 052	11 927	11 541b	
Selector	1	2	GDP growth rate (annual %, const. 2005 prices) 5	5.5	3.7	2.6b	
YDath • //dataila[3]/tabla/ikash/htfsd]	1	3	GDP per capita (current USS) 2	2 612.5	4 110.1	3 984.2b	
() XPath Vietensz (Rabiertood (Rabiertood (Rabiertood Vietensz (Rabiertood (Rabiertood (Rabie	1	4	Economy: Agricultured (% of GVA) 2	21.5	20.7	22.4b	
	1	5	Economy: Industryd (% of GVA)	28.7	28.7	26.0b	
Columns	1	6	Economy: Services and other activityd (% of GVA)	49.8	50.7	51.7b	
XPath Name	1	7	Employment: Agriculturee (% of employed)	49.3	42.6	41.4	
≝ *[1] Column 1 ● •	1	8	Employment: Industrye (% of employed)	17.8	19.9	18.3	
≝ *[2] Column 2 ● ●	1	9	Employment: Servicese (% of employed) 3	32.9	37.5	40.3	
:: *[3] Column 3 • •	1	10	Unemployment (% of labour force)	13.8e	14.2	15.8e	
	1	11	Labour force participatione (female/male pop. %)	47.6 / 68.5	44.1 / 63.6	40.2 / 61.0	
	1	12	CPI: Consumer Price Index (2000=100)	117	135	147f	
Filters	1	13	Agricultural production index (2004-2006=100)	98	119	134f	
Exclude empty results	1	14	Food production index (2004-2006=100)	98	119	134f	
	1	15	Index of industrial production (2005=100)	100	200	3551	
Presets Reset Scrape				Copy to cl	ipboard Expor	t to Google Docs	

<u>Nxjerrja e të dhënave nga imazhet e skanuara</u>

Si veprohet në rastet kur të dhënat ne i kemi vetëm në një dokument të skanuar?

Pas skanimit të një dokumenti, ajo që prodhohet është praktikisht një imazh. Tabelat dhe të dhënat në një dokument të skanuar është e pamundur të selektohen apo të aksesohen drejtpërdrejt nga Excel.

Gjithsesi, si në mjaft situata në botën dixhitale, edhe për këtë situatë ekziston një zgjidhje.

Të dhënat nga imazhet mund të nxirren përmes aplikacioneve OCR (*Optical Character Recognition*). Gjithsesi, produkti final nuk ka gjithnjë cilësinë që kërkohet, për këtë arsye duhet verifikuar rezultati final me dokumentin origjinal.

Disa nga aplikacionet OCR më të përdorura janë:

- Adobe Acrobat Professional: Ky është një program me pagesë.
- *Google Docs:* Aksesohet përmes G-mail dhe është falas.
- *Document Cloud:* Ky është një shërbim falas online që gjendet në: <u>https://www.documentcloud.org</u>

Për të përdorur *Document Cloud*, media juaj duhet të bëjë një kërkesë për të hapur llogari në këtë platformë.

3 BURIMET E TË DHËNAVE DHE RRUGËT PËR PËRFTIMIN E TYRE

Ku mund t'i gjejmë të dhënat? Cilat janë rrugët apo mjetet për përftimin e tyre?

Natyrisht një burim i rëndësishëm i të dhënave janë statistikat qeveritare dhe i gjithë dokumentacioni shtetëror i arkivuar, qoftë ky në format dixhital apo jo. Çdo vend ka Institutin e vet të statistikave, siç është në Shqipëri INSTAT. Por një burim autentik të dhënash është Arkivi Kombëtar si dhe arkivat lokale apo ato të institucioneve. Ndërkohë mjaft të dhëna në format dixhital janë publikuar apo mund të aksesohen edhe përmes websajteve të dikastereve qendrore, të institucioneve të ndryshme si dhe të kompanive apo bizneseve private.

Së fundmi një burim i rëndësishëm të dhënash po bëhen edhe portalet e të dhënave të hapura, siç është Open Data Albania në Shqipëri si dhe disa websajte të organizatave të shoqërisë civile të fokusuara në këtë fushë apo në fusha të përafërta.

Një burim tjetër i besueshëm janë statistikat e organizatave ndërkombëtare apo rajonale.

Për gazetarët, krahas praktikave *off line* të sigurimit të të dhënave, janë në dispozicion edhe një numër i konsiderueshëm i kanaleve online të përftimit të tyre. Më poshtë po paraqesim burimet më të rëndësishme online si dhe mënyrat dhe specifikat e aksesimit të tyre.

Databazat kombëtare të të dhënave

Një nga burimet më të rëndësishme të të dhënave në nivel kombëtar për gazetarët shqiptarë është databaza statistikore e Institutit të Statistikave (INSTAT). Të dhënat në portalin e INSTAT (<u>http://www.instat.gov.</u> <u>al/</u>) janë të organizuara sipas temave, siç e shihni edhe në figurën më poshtë, por përveç kësaj ju mund të hyni edhe tek databaza statistikore e plotë, duke klikuar tek linku me të njëjtin emër.



Databaza statistikore e INSTAT është e strukturuar në mënyrë të ngjashme me disa prej databazave ndërkombëtare të njohura siç ajo e Data Bank e Bankës Botërore. Kur klikoni tek "Hap databazën", hapet një faqe që ka një listë tematike të databazave. Për të aksesuar një dataset në një databazë të caktuar, si përshembull në atë të Mjedisit, ndiqet kjo rrugë:

- Klikoni tek shenja + që ndodhet përpara fjalës "Mjedisi". Ky klikim bën të mundur hapjen e dataseteve që gjenden brenda databazës "Mjedisi"
- Klikojmë tek "Importi i mbetjeve sipas llojit 2012-2016", për të hapur datasetin që përmban të dhënat

për importin e mbetjeve në Shqipëri. Dritarja që hapet ka pamjen e mëposhtme:

ar * duhet të zg	jidhni të paktën një vlerë * Tipi i mbetjeve	* Viti
	* Tipi i mbetjeve	* Viti
ir 1	Gjithsej 31 Përzgjedhur 31	Gjithsej 5 Përzgjedhur 5
	 Gjithsej Ngjitje prej gome plastike, përfshirë Mbetje të lëngjeve pastruese të meta Plehra, copa dhe mbetje, të polimere Mbetje, mbeturina e kthime gome, p 	2016 2015 2014 2013 2012
	Kërko 💽 🚬 🛛 Fillimi i rreshtit	Kërko 💽 🔀
	ır 1	rr 1 Gjithsej 31 Përzgjedhur 31 Gjithsej 31 Përzgjedhur 31 Gjithsej mbetje të lëngjeve pastruese të meta Plehra, copa dhe mbetje, të polimere Plehra, copa dhe mbetje, të polimere Mbetje, mbeturina e kthime gome, p ▼ Kërko Fillimi i rreshtit

- 3. Në kolonën majtas ne duhet të zgjedhim variablin që na intereson. Në kolonën e mesit zgjedhim tipin e mbetjeve. Në këtë rast janë selektuar të gjitha tipet, gjithsej 31. Në kolonën djathtas zgjedhim vitet. Në këta rast janë përzgjedhur të gjitha vitet për të cilat ka informacion që nga 2012 deri në 2016.
- 4. Pas kësaj klikojmë në fund tek butoni "Vazhdo". Hapet një faqe e re që përmban tabelën me të dhënat përkatëse. Këtë tabelë ne e shkarkojmë në formatin Excel (xlsx), duke klikuar XLSX tek butonat e formateve, të cilin e shihni edhe në figurën më poshtë.

Shfaq tabelën	Rreth tabelës						
Modifiko dhe Përllogarit V Ruaj si V Tabela - Paraqitja 2 V V 1							
Importi i mbetjeve sipas Tipi i mbetjeve, Variabla dhe Viti							
Importi i i	mbetjeve si	oas Tipi i mbetjeve	, Variabla d	he Viti			
Importi i i	mbetjeve sij	oas Tipi i mbetjeve,	, Variabla d Importi	he Viti i mbetjeve			
Importi i i	mbetjeve sij	oas Tipi i mbetjeve,	Variabla d Importi 2012	he Viti i mbetjeve 2013	2014	2015	2016
Importi i i Gjithsej	mbetjeve sij	pas Tipi i mbetjeve,	, Variabla d Importi 2012 4.192,0	he Viti i mbetjeve 2013 207.689,6	2014 333.553,0	2015 60.194,5	2016 140.228,0
Importi i i Gjithsej Ngjitje prej goo	m betjeve sij me plastike, përfs	pas Tipi i mbetjeve,	Variabla d Importi 2012 4.192,0 534,0	he Viti i mbetjeve 2013 207.689,6 730,6	2014 333.553,0 	2015 60.194,5 	2016 140.228,0 782,9

Pas klikimit tek butoni `Vazhdo`, menjëherë tabela shkarkohet në kompjuterin tonë dhe ne mund ta hapim dhe të punojmë me të dhënat në Excel.

Të dhëna dhe raportime mund të gjenden edhe tek portali i Open Data Albania (<u>http://</u><u>www.open.data.al/</u>). Ky portal bën një punë të mirë për të rritur transparencën e të



dhënave dhe për të sjellë në axhendën e diskutimit publik çështje që lidhen me to. Gjithsesi portali nuk krijon vetë të dhëna, por i merr ato nga institucionet shtetërore. Në një masë të konsiderueshme këtë gjë bën edhe INSTAT. Kjo do të thotë se në instancë të fundit burimi parësor i të dhënat janë institucionet shtetërore. Por në websajtet e këtyre institucioneve nuk ka databaza statistikore të mirëfillta dhe gjen shumë pak të dhëna, kryesisht sporadike që paraqiten në raporte apo materiale të ndryshme.

Databazat rajonale të të dhënave

Me financime ndërkombëtare janë krijuar dhe funksionojnë edhe disa databaza rajonale të cilat sjellin të dhëna për fusha specifike për të gjithë rajonin e Ballkanit. Një nga databazat më të njohura në këtë grup databazash është ajo e **Balkans Opinion Barometer** (http://www. rcc.int/seeds/results/2/balkan-opinion-barometer). Në të gjenden rezultatet e mjaft anketimeve në nivel rajonal që masin perceptimin e popullsisë për aspekte të ndryshme duke nisur që nga niveli i kënaqësisë së jetës (life satisfaction) e deri tek situata ekonomike, ajo e sigurisë apo situata e shëndetit publik.

Një dataset tjetër që përmban informacion në nivel rajonal është edhe **Balkans Data Pool** që aksesohet në web në adresën <u>http://bit.ly/2sPodfB</u>. Por më së shumti në këtë faqe gjen informacione lidhur me burimet e të dhënave, adresat e institucioneve që merren me to, personat e mundshëm të kontaktit, etj.

Databazat ndërkombëtare të të dhënave

Një pjesë e mirë e organizatave ndërkombëtare, si përshembull Banka Botërore, OKB, UNICEF, Organizata Botërore e Shëndetësisë, OECD, etj. kanë krijuar databaza ndërkombëtare të dhënash për fusha të ndryshme të cilat janë të natyrës së *open data* si dhe janë të aksesueshme online. Përparësia e këtyre databazave është se ato janë ndërtuar duke u bazuar në metodologji të avancuara shkencore prej nga ku ka rezultuar edhe një nivel i lartë besueshmërie. Njëherësh, duke sjellë të dhëna nga të gjithë vendet e botës, ato krijojnë mundësinë për të krahasur zhvillimet e një vendi në një fushë të caktuar me zhvillimet në vende të tjera. Kjo mundësi u ofron gazetarëve një bazë të mirë për të provuar hipotezat e tyre si dhe për të arritur në përfundime më të plota dhe më të argumentuara. Më poshtë po hedhim një vështrim të shpejtë në portalin e të dhënave të Bankës Botërore meqënëse është një nga portalet më të përdorura nga gazetarët e gazetarisë së të dhënave.

Kërkimi i të dhënave në Data Bank, siç quhet banka e të dhënave e BB, mund të bëhet duke zgjedhur vendin apo vendet për të cilat jeni të interesuar, mund të bëhet sipas tematikës së fushës së kërkimit por mund të bëhet edhe përmes fjalëkyçeve. Kërkimi përmes fjalëkyçeve bëhet tek "Explore databases". Kurse kërkimi sipas vendeve bëhet duke klikuar tek "Countries".

Një seksion i rëndësishëm i Data Bank është ai i emërtuar "World Development Indicators" (Treguesit e zhvillimit botëror), në të cilin gjenden një pafundësi të dhënash nga fusha të ndryshme. Ja hapat që duhen ndjekur për të aksesuar të dhënat e këtij seksioni:

 Hyni tek <u>http://databank.worldbank.org/data/</u> <u>home.aspx</u> dhe në homepage klikoni tek "World Development Indicators". Do të hapet një faqe e re e cila ka dritaren që shihni.

This page is in	s in English Español Français حربی 中文						
DataBank World Development Indicators						ators	
Variables	Layout	Styles	Save	Share	Embed		
 Database 				Availa	able	Selected	1
Country				Available	264	Selected	1
 Series 				Available	1591	Selected	0
 Time 				Available	58	Selected	0

- 2. Fillimisht tek variablat duhet të zgjedhim databazën që kërkohet duke klikuar tek "Database". Tek Database janë listuar fushat e përgjithshme të kërkimit.
- Pasi kemi zgjedhur databazën, klikojmë tek "Country" për të përzgjedhur edhe vendin apo vendet për të cilat jemi të interesuar.
- Hapi tjetër është përzgjedhja e serisë që mundësohet duke klikuar tek "Series". Tek seritë janë listuar fushat e detajuara të kërkimit të cilat përmbajnë datasetet e tyre.
- 5. Pas zgjedhjes së serisë i vjen radha përzgjedhjes së periudhës apo të viteve për të cilat jeni të i interesuar.
- Pasi keni bërë përzgjedhjet e të gjitha variablave të kërkuar, hapet një dritare e vogël *pop-up*. Klikoni tek *"Apply changes"*, dhe hapet dataseti me të dhënat e kërkuara.
- 7. Për të shkarkuar të dhënat shkoni tek "*Download option*" dhe zgjidhni "Excel". Të dhënat e kërkuara do të shkarkohen në kompjuterin tuaj dhe ju mund të nisni punën për përpunimin dhe analizën e tyre.

Kërkimi i avancuar në Google



Në disa raste, të dhëna të rëndësishmë mund të zbulohen edhe përmes kërkimit në Google. Në këtë rast rekomandhet që të përdoret kërkimi i avancuar (Advanced Search).

Për të aksesuar opsionet e kërkimit të avancuar ju duhet të shkoni tek <u>http://www.google.com/advanced_search</u>.

Në faqen që hapet, gjenden shumë opsione të kërkimit të avancuar në Google që e lehtësojnë dhe e bëjnë më të fokusuar dhe më të detajuar kërkimin tuaj. Kështu ju mund të bëni kërkim sipas fjalëve apo numrave, mund të përzgjidhni gjuhën, rajonin, periudhën kohore, llojin e formatit, etj.

Natyrisht rezultatet e kërkimit në Google duhet të kalohen në një filtër të mirë verifikimi. Nëse të dhënat gjende apo vijnë nga portale të panjohura, besueshmëria tek to është shumë e ulët. Por nëse të dhënat që ju gjeni, ndodhen në portale zyrtare institucionesh të njohura kombëtare apo ndërkombëtare, ato mund të merren në konsideratë.

Burime alternative të të dhënave

Krahas burimeve online, gazetarët e gazetarisë së të dhënave duhet të konsiderojnë edhe burimet alternative tradicionale të të dhënave. Ndodh jo rrallë që të dhëna të rëndësishme të mos publikohen online, kështu që mbetet të ndiqen rrugët e tjera alternative. Një prej këtyre rrugëve është edhe **kërkesa për informacion ndaj institucioneve** që zotërojnë të dhënat.

Duke ndjekur modelin e vendeve europiane, edhe Shqipëria ka tashmë një ligj të posaçëm për të drejtën e informimit. Është e domosdoshme që gazetarët të njohin të drejat dhe procedurat ligjore të parashikuara në këtë ligj, për t'i përdorur ato për sigurimin e të dhënave.

Sipas "Ligjit nr. 119/2014 Për të drejtën e informimit" çdo person, përfshi edhe gazetarët, gëzon të drejtën e njohjes me informacionin publik, pa u detyruar të shpjegojë motivet. Autoriteti publik është i detyruar të informojë kërkuesin nëse ka ose jo në zotërim informacionin e kërkuar. Gazetarët duhet të jenë të informuar edhe për afatet ligjore dhe të dinë se "autoriteti publik trajton kërkesën për informim, duke parashtruar informacionin e kërkuar sa më shpejt që të jetë e mundur, por jo më vonë se 10 ditë pune nga dita e dorëzimit të saj".⁴

Një burim tradicional i rëndësishëm për të dhënat mbeten edhe **individët**. Për këtë arsye gazetarët e gazetarisë së të dhënave do të duhet të ndërtojnë një databazë sa më të gjerë me emrat dhe kontaktet e njerëzve që potencialisht mund të ofrojnë apo të ndihmojnë në sigurimin e të dhënave.

Si burim të dhënash mund të përdoren edhe **kontributet e gazetarisë qytetare** por natyrisht duke i kaluar në një filtër të kujdesshëm verifikimi,. Gjithashtu të dhëna të rëndësishme për ngjarje apo zhvillime të ndryshme mund të mblidhen edhe nga **mediat**, por gjithnjë duke aplikuar filtrimin që përmendëm edhe më lart.

⁴ Ligji nr. 119/2014 Për të drejtën e informimit"

4 PYETJE TË RËNDËSISHME QË DUHEN BËRË KUR GJENDEMI PËRPARA TË DHËNAVE

Skepticizmi është një premisë e domosdoshme që duhet ta shoqërojë gazetarin gjatë gjithë punës së tij për mbledhjen dhe raportimin e të dhënave. Të sigurosh të dhënat nuk mjafton. Krahas kësaj është e domosdoshme që t'iu përgjigjesh edhe një numri pyetjesh që duhen bërë kur gjendemi përpara të dhënave. Sipas Jonathan Stray, më kryesoret prej këtyre pyetjeve mund të ishin⁵:

Nga vijnë të dhënat?

Pra cili është institucioni që i ka publikuar të dhënat dhe sa i besueshëm është ai dhe të dhënat e tij.

Kush i ka mbledhur të dhënat?

Janë mbledhur nga vetë institucioni apo nga një kompani e pajtuar prej tij? A u trajnuan ata që mblodhën të dhënat dhe ç'lloj trajnimi u ndërmor?

Si u mblodhën të dhënat?

U mblodhën duke shkuar direkt në burim apo nga raporte të ndryshme?

Për ç'qëllim u mblodhën të dhënat?

A u mblodhën për të raportuar tek një donator se objektivat ishin arritur? Apo si pjesë normale e një procedure zyrtare të mbledhjes së të dhënave?

⁵ Jonathan Stray, https://source.opennews.org/articles/statistically-sound-data-journalism/

Si ta dimë që të dhënat janë të plota?

A mund të pyesim mbledhësit e të dhënave për këtë? A ka ndonjë shpjegim për kufizime lidhur me të dhënat?

Cila është demografia?

A u përfaqësuan të gjitha shtresat, rajonet, grupmoshat, grupet gjinore, etj?

A është kjo mënyra e duhur për të matur diçka?

A është zgjedhur rruga e duhur apo ka mënyra apo rrugë të tjera që mund të ishin më efektive?

Kush nuk përfshihet tek shifrat?

Janë përfshirë të gjithë aktorët dhe faktorët apo kanë mbetur jashtë disa prej tyre, përshembull për mungesë të aksesit?

Kush mund të duket keq apo të humbasë parà nga këto të dhëna?

Mos ishte një punë e porositur nga një organizatë për ta përdorur për nevoja të PR-it të vet?

A janë të dhënat konsistente ditë pas dite apo janë mbledhur nga grupe të ndryshme në periudha të ndryshme?

A u mblodhën nga i njëjti grup dhe me të njëjtën metodologji?

A ka patur zgjedhje arbitrare në procesin e gjenerimit të të dhënave?

Përshembull, në përcaktimin e kampionit të anketimit, etj?

A janë të dhënat konsistente edhe me burimet e tjera? Kush tjetër i ka analizuar ato?

A ka datasete të tjera për të njëtën çështje dhe a përputhen të dhënat?

5 PASTRIMI DHE STANDARTIZIMI I TË DHËNAVE

Pastrimi dhe standartizimi i të dhënave janë procedura të rëndësishme në punën e gazetarit të gazetarisë së të dhënave. Mbledhja e tyre është vetëm fillimi. Për të qenë të sigurt në cilësinë e shkrimit që do të prodhojmë, do të duhet më parë të jemi të sigurt për cilësinë e të dhënave. Dhe një rol në këtë cilësi ka procesi i pastrimit të të dhënave.

Pastrimi është një proces që i paraprin analizës kompjuterike të të dhënave. Ai përfshin zakonisht eleminimin e gabimeve njerëzore që mund të jenë bërë në tabelat e të dhënave, heqjen e të dhënave apo të informacionit të panevojshëm, kthimin apo unifikimin e të dhënave në formatin e duhur, etj.

Kjo është një punë për të cilën do të duhen programet kompjuterike përkatëse, dhe siç e kemi thënë më lart, programi më i mirë dhe më i përdorur për këtë qëllim është Excel.

Le të shohim një rast të thjeshtë të pastrimit të të dhënave.

Shkojmë në portalin e INSTAT në adresën <u>http://www.instat.gov.al/al/temat/kushtet-sociale/sh%C3%ABndet%C3%ABsia/#tab2</u>. Në këtë faqe gjendet tabela e mëposhtme:

s 🔒	hëndetë	sia		
Të Rejat	Shifrat	Publikimet	Metodologji	
Aktivite	ti i konsultoi	rit të gruas për 2	2012 – 2016	<u>*</u>
Aktivite	ti i konsultoi	rit të fëmijëve pë	er 2012 – 2016	*
Aktivite	ti i qëndrave	e shëndetesore,	ambulancave dhe poliklinikave për 2012 – 2016	*
Aktivite	ti i institucio	neve me shtretë	ər 2011 – 2015	*
Përqino	lja e vdekje	ve sipas shkaqe	ve kryesore, 2015-2016	*
Se Dat	abaza stati	stikore		

- Shkarkojmë datasetin me titull "Përqindja e vdekjeve sipas shkaqeve kryesore 2015-2016", duke klikuar te butoni i shkarkimit në të djathtë. Tabela e të dhënave që shkarkojmë është në format Ecxel-i, pra XLS.
- 3. Hapim tabelën në Excel. Ja pamja e saj përpara se të aplikojmë pastrimin:

	А	В	С	D				
1	Përqindja e vdekjeve sipas sh	irqindja e vdekjeve sipas shkaqeve kryesore, 2015-2016						
2	Percentage of Deaths by Main	Causes, 2015-2016						
3								
4	Shkaqet e vdekjeve	2015	2016	Cause of Death				
5	Natyrale	86.8	87,4	Natural				
6	Aksident	2,2	2,1	Accident				
7	Vetvrasje	1,0	1,1	Suicide				
8	Vrasje	0,4	0,4	Homicide				
9	Aksident në punë	0,3	0,2	Accidentwork				
10	E panjohur	9,3	8,8	No information				
11	Gjithsej	100,0	100,0	Total				
10								

4. Në fazën e parë të procesit të pastrimit, në dokumentin e hapur duhet krijojmë një kopje të worksheet-it (faqes së punës) që ka të dhënat. Kjo bëhet për të ruajtur të paprekur origjinalin dhe për të punuar tek kopja. Duhet të shtojmë edhe një worksheet tjetër bosh ku do të vendosim burimin e të dhënave si dhe informacione të tjera që mund të kemi rreth datasetit. Pra do të kemi tre tab-e aktive si më poshtë:

38	
39	
$H \rightarrow H$	Origjinali Burimi Kopja 🦓
Ready	

- 5. Tek *worksheet*-i "Burimi" hedhim linkun e webit nga kemi shkarkuar datasetin, për të patur në çdo rast burimin nga kemi marrë të dhënat. Vendosim përkufizimin për variablin "shkaku i vdekjes" që gjendet tek Metodologjia, në faqen e webit ku lokalizohej dataseti. Këtu mund të vendosim edhe informacione të tjera shpjeguese apo plotësuese që lidhen me kontekstin e të dhënave.
- 6. Shkojmë në faqen e punës që ka kopjen e dataset.it. Shikojmë me kujdes të dhënat dhe informacionin. Fillojmë pastrimin duke fshirë tre radhët e para, përkatësisht 1,2 dhe 3 - dy që kanë titullin në shqip dhe anglisht dhe radhën poshtë poshtë tyre që është bosh.
- Fshijmë edhe kolonën D që ka treguesit në anglisht, pasi i kemi ato në shqip në kolonën A.
- 8. Shohim me kujdes vlerat e shprehura në numra. Ndarja dhjetore e tyre është bërë me presje. Por në këtë rast program Excel i trajton si tekst. Për t'i kthyer në format numri, duhet të zëvendësojmë presjet me pikë përmes butoni *Replace*. Pas kësaj vlerat janë numra dhe mund të përpunohen si të tilla. Me këtë ne kemi përfunduar pastrimin e të dhënave dhe kthimin e tyre në formatin e duhur. Ja pamja që ka tabela pas pastrimit:

А	В	С
Shkaqet e vdekjeve	2015	2016
Natyrale	86.8	87.4
Aksident	2.2	2.1
Vetvrasje	1	1.1
Vrasje	0.4	0.4
Aksident në punë	0.3	0.2
E panjohur	9.3	8.8
Gjithsej	100	100
	A Shkaqet e vdekjeve Natyrale Aksident Vetvrasje Vrasje Vrasje Aksident në punë E panjohur Gjithsej	ABShkaqet e vdekjeve2015Natyrale86.8Aksident2.2Vetvrasje1Vrasje0.4Aksident në punë0.3E panjohur9.3Gjithsej100

Tani të dhënat janë gati për t'u përpunuar dhe analizuar përmes programit Excel.

Në rastet kur kemi informacion si data, adresa, monedha të pagesës, etj duhet të bëjmë kujdes që ato të jenë të unifikuara. Për këtë na vjen në ndihmë standartizimi i tyre. Kështu datën 12 qershor 2018 ne mund ta shkruajmë edhe në variantet 12 qershor 2018, 12/6/2018, 6/12/2018 , 12/6/18, 12/qersh/2018. Standartizim do të thotë zgjedhje dhe përdorim i një formati të vetëm përgjatë gjithë datasetit.

Po kështu ne mund ta shënojmë me fjalë monedhën dollar, por edhe me shenjën \$. Por rregulli e kërkon që të përdorim vetëm një mënyrë, dhe konkretisht simbolin që e njeh edhe programi Excel. Në disa raste duhet unifikuar emri i një vendi. Kështu ne e gjejmë Shqipëri në disa datasete, por mund ta gjejmë edhe Albania, ne e gjejmë Amerikë, por mund ta gjejmë edhe SHBA, ashtu siç mund ta gjejmë BE-në në formën e gjatë Bashkimi Europian. Gjithashtu duhen standartizuar emrat dhe mbiemrat e njerëzve duke vendosur standartin nëse do të jetë emri përpara apo mbiemri i ndarë me presje, etj. Të gjitha këto duhen unifikuar gjatë procesit të standartizimit.



Janë të njohura epërsitë e komunikimit vizual ndaj komunikimit tekstual apo verbal dhe këto epërsi janë theksuar edhe më shumë në epokën dixhitale. Zhvillimet e fundit në teknologjinë e informacionit dhe të komunikimit kanë ndikuar dukshëm në një rivlerësim dhe rritje të vizualizimit të të dhënave.

Vizualizimi e pasuron raportimin e bazuar tek të dhënat. Por shumë ekspertë theksojnë se ai nuk duhet kthyer në një qëllim në vetvete. Dhe synimi final nuk është loja me ngjyrat apo zgjidhjet e bujshme vizuale. Sipas Alberto Cairo, ekspert i vizualizimit të të dhënave në Univesitetin e Majamit, infografikat dhe vizualizimet në radhë të parë duhet të synojnë të tërheqin mendjen dhe jo syrin.⁶

Moment i rëndësishëm e procesit të vizualizimit të të dhënave është përzgjedhja e llojit të duhur të vizualizimit për përmbajtjen e duhur të të dhënave. Kështu për të vizualizuar të dhëna që paraqesin rritjen e një dukurie nga viti në vit, ne nuk mund të përdorim vizualizimin përmes hartës. Por nëse do të duhet të krahasojmë vendet të ndryshme ne mund ta përdorim hartën kur kjo gjë ka sens.

⁶ https://visual.ly/blog/infographics-and-visualizations-as-tools-for-the-mind/

Mjetet kryesore të vizualizimit të të dhënave janë

Bar charts: krahasojnë të dhënat përgjatë kategorive

Line Graphs: krahasojnë të dhënat përgjatë kohës

Scatter plot: krahason të dhënat mes dy variableve

Piktogramet: paraqesin apo krahasojnë aspekte të të dhënave përmes ikonave grafike

Harta: krahason apo sjell të dhëna specifike që lidhen me lokalizimin

Nga ana e tyre hartat, po përdoren gjithnjë e më shumë si mjete vizualizimi duke sjellë edhe lindjen e një lloj gazetarie të veçantë brenda gazetarisë së të dhënave që njihet me emrin "*map journalism*" (gazetaria e hartës). Edhe në rastin e vizualizimit nëpërmjet hartave duhet të zgjedhim llojin e duhur për përmbajtjen e duhur. Llojet kryesore të hartave që përdorem aktualisht për vizualizim janë tre:

Point map: Përdoret nëse keni të dhëna specifike që lidhen me lokalizimin

Bubble map: Përdoret për të treguar të dhënat për një pikë lokalizimi.

Choropleth map: Përmes k**ë**tij lloji harte ne paraqesim dhe krahasojmë të dhënat që lidhen me njësi të ndryshme administrative apo me vende të ndryshme.

Më poshtë po listojmë disa nga platformat më të njohura të vizualizimit në web të cilat aplikojnë të gjitha ose disa prej llojeve të vizualizimit:

Carto, që gjendet në adresën: <u>https://carto.com/</u>

Data Wrapper: <u>https://www.datawrapper.de/</u>

 Tableau Public: <u>http://www.tableausoftware.com/</u>

 public/_

Gap Minder: <u>http://www.gapminder.org/</u>

Google Fusion Tables: <u>http://www.google.com/drive/</u> <u>apps.html#fusiontables</u>

Piktochart: http://piktochart.com/

Many Eyes: http://www-958.ibm.com/software/analytics/manyeyes/

Në shumë prej këtyre platformave mjafton të importosh tabelën e pastruar dhe të standartizuar të të dhënave apo ti sjellësh këto të dhëna direkt në platformë me *copypaste*, dhe më pas mund të zgjedhësh dhe të aplikosh llojin e vizualizimit që dëshiron. Secila nga platformat ka faqet e ndihmës si dhe videot shpjeguese për gjithçka që ato bëjnë. Duhet vetëm pak ditë për një gazetar me aftësi mesatare dixhitale për të përvetësuar teknikat dhe mënyrat kryesore të vizualizimit përmes këtyre platformave.

Më poshtë po paraqesim një shembull vizualizimi përmes platformës **Carto** (<u>https://carto.com/</u>). Nga databaza e OECD janë gjetur të dhënat lidhur me numrin e vdekjeve nga aksidentet rrugore në Ballkan për vitin 2015 (për Kosovën mungonin të dhënat) dhe pas përpunimit të tyre rezultojnë shifrat në tabelën e mëposhtme:

	А	В
1	Vendet	Vdekjet për 100 mijë banorë: 2015
2	Albania	9.35
3	Macedonia	7.12
4	Montenegro	10.29
5	Serbia	8.47

Për të krahasuar vendet e Ballkanit lidhur me numrin e vdekjeve nga aksidentet rrugore, natyrisht mënyra më e mirë është që të krahasojmë numrin e vdekjeve për çdo 100 mijë banorë. Për të dhënat e tabelës së mësipërme një mënyrë e mirë vizualizimi do të ishte ajo përmes hartës, dhe konkretisht përmes Choropleth Map. Ja si bëhet vizualizimi në platformën Carto:

- 1. Hapni Carto duke shkuar tek <u>https://carto.com/</u> dhe krijoni një llogari në të.
- Në zonën e punës të Carto (Dashboard), shkoni tek Datasets dhe më pas klikoni tek New dataset. Tek dritarja që hapet, klikojmë tek Browse dhe ngarkojmë file-in e Excel që ka tabelën e mësipërme.
- 3. Pasi e kemi ngarkuar klikojmë tek *Connect Dataset*. Shfaqet në Carto tabela e mëposhtme:

DAIIKAN_VQEKJET_NgA_AKSIQENTET : PUBLC Updated a few seconds ago					
cartodb_id \uparrow : number	the_geom	cartodb_georef_status :	vendet :	vdekjet_per_100_mije :	
1	Polygon	true	Albania	9.34525418572	
2	Polygon	true	Macedonia	7.1206806216	
3	Polygon	true	Montenegro	10.286759494	
4	Polygon	true	Serbia	8.47029681132	

 Klikojmë tek butoni *Create Map*, në cepin poshtë djathtas të dritares së punës.



5. Hapet dritarja që ka hartën, ku Carto ka lokalizuar edhe vendet përkatëse të Ballkanit. Theksojmë se që

të njihen nga Carto vendet, emërtimi i tyre duhet të jetë në anglisht. Klikojmë sërish tek emri i *file*-it të të dhënave, për të hapur opsionet e hartëzimit.

6. Klikojmë tek STYLE, dhe ndër opsionet përzgjedhim *By value*. Kjo bën të mundur që secila vlerë që ne

kemi për çdo vend të marrë ngjyrën e vetë brenda spektrit të së njëjtës ngjyrë. Vlera më e madhe ka tonin më të errët të ngjyrës. Vlera më e vogël, ka tonin më të çelët.

 Klikojmë më pas tek POP-UP. Më tej klikojmë tek *Hover* dhe më tej tek *Pop Up Light,* në mënyrë që kur

•	← Back / Layer options ▲ ballkan_vdekjet ** : ballkan_vdekjet_nga_aksidentet
	DATA ANALYSIS STYLE POP-UP LEGEND
<u>0</u> -	1 Polygons style CHANGE THE VISUALIZATION
	POLYGON COLOR 🔷 Solid 🧿 By value
	vdekjet_per_10
	STROKE SIZE -O
	STROKE COLOR
	BLENDING none
	LABELS

të afrojmë *mouse*-in tek çdo vend të shfaqen vlerat përkatëse. Përzgjedhim "Vendet" dhe "Vdekjet për 100 mijë banorë: 2015".

- 8. Klikojmë tek LEGEND dhe përzgjedhim *Choropleth Legend*. Riemërtojmë legjendën me titullin "Ballkan vdekjet nga aksidentet".
- 9. Klikojmë tek butoni *Publish*, për të përfunduar me editimin dhe publikimin e vizualizimit. Nëse duam që ta publikojmë vizualizimin në websajtin tonë, shkojmë tek *Get Link*, kopjojmë linkun dhe e

integrojmë tek websajti ynë.



Ja si do të duket vizualizimi përfundimtar:

Siç rezulton edhe nga ky vizualizim, vendi me tonin më të errët të ngjyrës ka numrin më të madh të vdekjeve nga aksidentet për 100 mijë banorë, kurse vendi me tonin më të çelët të ngjyrës ka numrin më të vogël. Kur afrojmë mouse-in, siç kemi bërë në rastin e Maqedonisë, shfaqen vlerat përkatëse për secilin vend.

34

PARAQITJA E TË DHËNAVE GJATË TË SHKRUARIT PËR GAZETARINË E TË DHËNAVE

Të dhënat si rregull shprehen përmes numrave. Por numrat në vetvete janë "të ftohtë" dhe mund të mos përcjellin emocionin e duhur tek publiku. Për këtë arsye, në procesin e të shkruarit për gazetarinë e të dhënave shpesh rekomandohet "humanizmi" i numrave apo i përqindjeve që sillen në shkrim.

Humanizoni shifrat

Le ta ilustrojmë këtë me një shembull:

Nga kërkimi juaj në një databazë apo raport, ju zbuloni që numri i vdekjeve nga aksidentet rrugore gjatë vitit të kaluar është dyfishuar. Nëse këtë situatë do ta përshkruanit me fjalinë "Siç rezulton nga Raporti X, numri i vdekjeve nga aksidentet rrugore gjatë vitit të kaluar është rritur me 100 % krahasuar me një vit më parë", ju keni rënë pre e një dehumanizimi të fakteve. Bëhet fjalë për jetë njerëzish të humbura dhe jo për thjesht për numra apo përqindje. Kështu, për të krijuar edhe emocionin e duhur tek publiku i gjerë dhe tek aktorët e përfshirë, situata do të duhet të përshkruhej më fjalët: "Siç rezulton nga Raporti X, numri i njerëzve që kanë vdekur nga aksidentet rrugore gjatë vitit të kaluar është dyfishuar në krahasim me një vit më parë".

Në raportime të tjera edhe më delikate kur bëhet fjalë, përshembull, për humbje jetësh të fëmijëve, ne nuk mund të raportojmë me terma të ftohtë si, psh "niveli i vdekshmërisë foshnjore është rritur me x % krahasur me 5 vjet më parë". Sepse bëhet fjalë për jetë foshnjesh të humbura dhe jo thjesht për "vdekshmëri foshnjore".

Tregojeni siç do t'ua tregonit prindërve tuaj

Një rekomandim tjetër i dobishëm që bëhet për paraqitjen e të dhënave gjatë të shkruarit për këtë lloj gazetarie, ka të bëjë me thjeshtëzimin sa më të mundshëm të tyre, natyrisht pa i banalizuar aspak. Shprehja "tregoje siç do t'ia tregoje prindërve të tu" aplikohet më së miri edhe në rastin e të shkruarit për gazetarinë e të dhënave.

Në databazat e portaleve të të dhënave ndeshen jo rrallë terma shumë teknikë, që mund edhe të mos kuptohen nga audienca e gjerë. Kështu, për shembull, në databazën e INSTAT gjenden të dhëna edhe për një variabël që quhet "Incidenca e sëmundjeve infektive⁷". Nëse do të ndërmerrnit inisiativën për të bërë një shkrim lidhur me incidencën e sëmundjeve infektive në Shqipëri gjatë viteve të fundit, atëhere do të duhet të bënit një lloj thjeshtëzimi të termave. Kështu në vend që të shkruanit përshembull "Sipas INSTAT, incidenca e sëmundjeve infektive në Shqipëri është shtuar me 25 % gjatë vitit 2017 në krahasim me vitin 2016", do të ishte shumë më mirë që të shkruanit: "Sipas INSTAT, numri i rasteve të reja të sëmundjeve infektive në Shqipëri është shtuar

⁷ http://databaza.instat.gov.al/pxweb/sq/DST/START_HE_KSH/ HE0007/?rxid=eee1d774-cce6-4fa9-a0b3-d27c0a4a4833

me 25 % gjatë vitit 2017 në krahasim me vitin 2016". Në këtë rast audienca e gjerë e kupton se për çfarë bëhet fjalë, kurse po të përdornim termin "incidencë" zor se do ta kuptonte njeri.

Madje do të ishte mirë që edhe vlera "25 %" të mos shprehej si shifër, por të shprehej në mënyrë të thjeshtëzuar. Në këtë rast një rritje me 25 % do të thotë që për çdo 4 raste të reja të sëmundjeve infektive për 1000 mijë banorë gjatë vitit 2016, në 2017-ën kemi një rast më shumë. Në situata të tjera, e njëjta vlerë mund të përshkruhej edhe me fjalët "një e katërta", ose "çereku i…" për të shmangur shprehjen e thatë të të dhënave nëpërmjet përqindjeve. Ashtu siç mund të thoshim "shitjet u trefishuan" në vend të "shitjet e rritën me 200 %".

Rrumbullakosni dhe përmblidhni të dhënat

Gjatë të shkruarit për gazetarinë e të dhënave, rekomandohet që të aplikohet edhe rrumbullakosja e të dhënave, natyrisht deri në masën që nuk shtrembërohet mesazhi. Po të shpreheshim thjesht me shifra të plota, mund të krijohej situata kur audienca nuk do të mund të arrinte t'i perceptonte menjëherë përmasat e një objekti, të një situate apo të një zhvillimi.

Kështu në një lajm mund të raportohej, përshembull, se "sipërfaqja e mbjellë me luleshtrydhe në Maminas gjatë vitit 2017 arriti në 10000 metra katrorë". Shifra 10000 metra katrorë për shumë njerëz që e kanë të vështirë perceptimin e përmasave, do ta bënte edhe më të vështirë punën. Cila mund të ishte zgjidhja? Fare mirë mund të shpreheshim se "sipërfaqja e mbjellë me luleshtrydhe në Maminas gjatë vitit 2017 ishte 1 hektar". Nëse do të kishim vlera të tilla si 32.55 %, apo 51.2%, do të ishte më mirë që të aplikonim një rrumbullakosje dhë të thoshim "rreth një e treta" ose "rreth gjysma". Lidhur me këtë, studjuesi i njohur gjerman i gazetarisë Stephan Russ-Mohl shkruan: "Kur botojnë të dhëna në përqindje, gazetarët duhet t'ia bëjnë lexuesit tablonë sa më të lehtë që të jetë e mundur".⁸

Jepuni kuptim të dhënave

Gjatë raportimit për gazetarinë e të dhënave, ndodh që vlerat absolute mund të mos kenë sens, kur krahasoni përshembull dy vende të ndryshme lidhur me një zhvillim të caktuar.

Kështu le të supozojmë se kemi dy të dhëna: një për numrin e aksidenteve rrugore në SHBA, i cili supozojmë që arriti në 158 000 gjatë vitit 2017 dhe një për numrin e aksidenteve rrugore në Shqipëri, i cili arriti në 16 000 gjatë vitit 2017. Po të shpreheshim me vlera absolute, atëhere mund të kishim një titull shkrimi të tillë "Numri i aksidenteve rrugore gjatë vitit 2017 në SHBA ishte më i lartë se numri i aksidenteve rrugore në Shqipëri".

A ka sens ky titull? Natyrisht që jo. Logjikisht, nëse do të bazoheshim tek shifrat absolute, numri i aksidenteve rrugore në SHBA kuptohet që do të jetë më i lartë, pasi popullsia është 100 herë më e madhe se ajo e Shqipërisë dhe gjithashtu edhe numri i makinave është disa qindra herë më i madh. Për këtë arsye, krahasimi me shifra absolute çalon. Cila do të ishte zgjidhja? Në këtë rast

⁸ Russ-Mohl, Stephan, Gazetaria, K&B, Tiranë, 2014

duhet të përdornim ato që në gjuhën e statistikës quhen "rates", konkretisht duhet të bënim krahasimin mes numrit të aksidenteve për 100 mijë banorë në një vend krahasuar me tjetrin.

Ka të dhëna të tjera që në disa situata marrin kuptim vetëm po të paraqiten si vlera mesatare. Kështu nëse do të krahasonim jetëgjatësinë në një vend me një vend tjetër, nuk mund të merrnim si bazë, përshembull, jetëgjatësinë e dy individëve që jetuan më gjatë. Kjo nuk do të kishte sens. Për t'i dhënë kuptim, krahasimi duhet bërë mes vlerave të jetëgjatësisë mesatare për një vit të caktuar në secilin vend.



STRUKTURA E RAPORTIMIT TË BAZUAR TEK TË DHËNAT DHE DISA REKOMANDIME PËR TË SHKRUARIT NË GAZETARINË E TË DHËNAVE

Llojet kryesore të raportimeve të bazuara tek të dhënat janë tre⁹:

- 1. Raportimi kronologjik
- 2. Raportimi i tipit "Compare and contrast" (Krahaso dhe kontrasto)
- 3. Raportimi i tipit "Step by step" (Hap pas hapi)

Ja veçantitë dhe hapat që ndiqen për realizimin e secilit prej këtyre llojeve të raportimit

1. Raportimi kronologjik

- Nis me një rast specifik
- Shpjegohet historia e problemit
- Vendoset fokusi në qëllimin e raportimit
- Analizohen arsyet pas asaj që ka ndodhur
- Analizohen ndikimi dhe pasojat
- Identifikohen reagimet dhe masat
- Paraqitet trendi në të ardhmen

⁹ http://www.odecanet.org/wp-content/uploads/2016/10/ Module-1-UNDP-student.pdf

Shembull nga përvoja botërore:

http://www.npr.org/2017/05/12/527806002/ focus-on-infants-during-childbirth-leaves-u-smoms-in-danger

2. Raportimi i tipit "Compare and contrast"

- Hyrja nis me një gjetje që paraqet kontrastin mes dy grupeve që analizohen, psh, një inisitiavë e suksesshme përballë një inisiative të dështuar në kushtet e situatave të ngjashme
- Shpjegohet historia e problemit
- Ilustrohen vendimet e ndryshme drastike që çuan në rezultate divergjente
- Eksplorohet sistematikisht impakti dhe pasojat e kundërta tek të dy grupet
- Paraqitet mënyra se si grupet do të vazhdojnë të zhvillohen veçmas nëse nuk ndërhyhet
- Eksplorohen zgjidhjet e ndryshme që mund t'i sjellin më pranë dy grupet

Shembull nga përvoja botërore:

https://www.propublica.org/article/machine-biasrisk-assessments-in-criminal-sentencing

3.Raportimi i tipit "Step by step"

- Hyrja paraqet shembuj të dështimeve të shumëfishta të sistemit që studjohet
- Paraqiten konteksti si dhe burimet e informacionit
- Ndahen dhe analizohen veçmas secili prej komponentëve, duke e shpjeguar secilin në një seksion më vete përmes shembujve ilustrues për t'i dhënë raportimit një fytyrë njerëzore, pra për ta

lidhur me njeriun

 Mbyllet me një shqyrtim të mundësive për zgjidhje gjithëpërfshirëse në mënyrë që të përmirësohet sistemi i gjërave

Shembull nga përvoja botërore:

http://www.tampabay.com/projects/2015/ investigations/pinellas-failure-factories/chartfailing-black-students/

Pas zgjedhjes së llojit të raportimit që përshtatet më shumë me historinë për të cilën ju do të raportoni, nis procesi i të shkruarit. Më poshtë po paraqesim disa rekomandime për të shkruarit për gazetarinë e të dhënave të ofruara nga Instituti Poynter¹⁰:

- ✓ Shkruaj për një audiencë të gjerë
- ✓ Tregoje ngjarjen sikur do t'ua tregoje prindërve
- ✓ Ngadalëso ritmin e informacionit
- ✓ Paraqit një e nga një karakteret e reja apo konceptet e vështira
- ✓ Konsidero vlerën e përsëritjes
- ✓ Mos krijo rrëmujë që në nisje
- ✓ Përdor fjali të thjeshta
- ✓ Ki parasysh se numrat e mpijnë lexuesin
- ✓ Mendo grafikisht
- ✓ Përkthe zhargonet
- ✓ Përdor analogjitë
- ✓ Shiko dhe gjej anën njerëzore
- ✓ Shpërble lexuesin
- ✓ Konsidero impaktin
- ✓ Zhvillo kronologji
- ✓ Hiq ujërat e tepërta

¹⁰ http://www.poynter.org/2014/a-new-explanatory-journalism-can-be-built-on-a-strong-foundation/244285/

9 DHJETË RREGULLA TË GAZETARISË SË TË DHËNAVE ¹¹

- Nëse mendon se diçka që je duke bërë është më e vështirë nga sa dukej, atëhere je në rrugën e duhur
- Ka 99.99 % shanse që dikush tjetër të ketë bërë të njëjtën analizë që ju doni të bëni. Mëso prej këtyre rasteve
- Asnjëherë mos puno në kopjen origjinale të një dataset-i
- 4. Nuk ekzistojë të dhëna të dukshme. Kërko për të metat në vend që të presësh që ato të të gjejnë ty.
- 5. Nëse rezultatet duken "*too good to be true*", ka 99.99% shanse që ashtu të jenë
- 6. Fokusohu tek ajo që është e mundur, në vend që të përpiqesh të bësh gjithçka
- 7. Asnjëherë mos mendo se e di se çdo të thotë diçka
- 8. Mbani shënime në letër, duke dokumentuar pastrimin dhe analizën e të dhënave tuaja.
- 9. Nuk e dini si bëhet diçka? Kërko në Google.
- 10. Mënyra më e mirë për të ngulitur atë që keni mësuar: t'ia mësoni atë dikujt tjetër.

¹¹ Burimi: MaryJo Webster, http://mjwebster.github.io/DataJ/

CIP Katalogimi në botim BK Tiranë

Zguri, Rrapo Gazetaria e të dhënave : manual për gazetarët / Rrapo Zguri; red. Lutfi Dervishi. - Tiranë : Kumi, 2018 ... f. : me il. ; ... cm. Bibliogr. ISBN 978-9928-247-37-7

1.Gazetaria 2.Përpunimi elektronik i të dhënave 3.Manuale 070 : 004.6 (035)

