**Data Science for Business**

**Data Visualization Final Project**

**Anna Tsaturyan, Sargis Petrosyan, Artashes Harutyunyan**

**Terrorism around the world!**

**Story**

***What is the problem?***

Խնդիրը գլոբալ ահաբեկչության տվյալների մանրամասն, մատչելի և ինտերակտիվ վերլուծության անհրաժեշտությունն է՝ օրինաչափությունները և ազդեցությունները հասկանալու համար: Այս նախագծի հիմնական նպատակն է ստեղծել ինտերակտիվ և համապարփակ տվյալների վիզուալիզացիա, որը թույլ կտա օգտվողներին ուսումնասիրել և հասկանալ համաշխարհային ահաբեկչության կառուցվածքը: Տվյալների վերլուծության և վիզուալիզացիայի կիրառմամբ՝ նախագիծը փորձում է պատասխանել տարիների ընթացքում ահաբեկչության տարածման, առավել տուժած շրջանների, ամենամահաբեր ահաբեկչական հարձակումների տեսակների և ամենաակտիվ ահաբեկչական խմբերի վերաբերյալ հիմնական հարցերին:

***What question do you want to answer?***

Ուզում ենք պատասխանել հետևյալ հարցերին․

* Ինչպե՞ս են փոխվել ահաբեկչական հարձակումների հաճախականությունն ու տարածումը տարիների ընթացքում:
* Ո՞ր շրջաններն են ամենաշատը տուժել ահաբեկչությունից, և ինչպե՞ս են տարբերվում հարձակումների տեսակներն այս շրջաններում:
* Ահաբեկչական հարձակումների հետևանքով զոհերի թվի ինչպիսի՞ միտումներ կան:
* Ովքե՞ր են ամենաակտիվ ահաբեկչական խմբավորումները, և որո՞նք են հարձակման նախընտրելի մեթոդները:
* Կա՞ն արդյոք նկատելի օրինաչափություններ ահաբեկչական միջադեպերի վերաբերյալ:

***Why did you choose it?***

Մենք սա ընտրեցինք, քանի որ ահաբեկչության միտումների ըմբռնումը կարևոր է անվտանգային գործողություններ կազմակերպելու համար, արդյունավետ քաղաքականություն մշակելու և հանրային իրազեկվածության բարձրացման համար:

***What is the context?***

Պրոյեկտում ընդգրկված են մի քանի տասնամյակների տվյալների բազա ահաբեկչության դեպքերի վերաբերյալ: Այն ներառում է տեղեկություններ ամսաթվերի, վայրերի, հարձակման տեսակների, զոհերի թվի, կազմակերպչական խմբերի վերաբերյալ՝ ապահովելով խորը պատկերացում՝ օգնելու մեղմել ահաբեկչության ազդեցությունը։

**Audience**

***Who is going to use the viz?***

Վիզուալիզացիաները առաջին հերթին անհրաժեշտ են քաղաքականություն մշակողներին, ազգային անվտանգության ծառայության աշխատակիցներին, տարբեր հետազոտողներին, ահաբեկչության ֆինանսավորման դեմ պայքարող կազմակերպություններին, ինչպիսիք են բանկերը, մի խոսքով՝ ցանկացած սուբյեկտ, որ ցանկանում է պաշտպանվել ահաբեկչությունից։

**Data**

***What type of data do you have?***

Ունենք թվային, կատեգորիկ և տեքստային տվյալներ։ Տվյալները ներառում են տեղեկատվություն ամբողջ աշխարհում ահաբեկչական հարձակումների մասին, ինչպիսիք են ամսաթվերը, վայրերը, հարձակման տեսակները, զոհերը, հանցագործների խմբերը:

***Where is it stored?***

Ունենք 1970-2017թթ. տվյալներ՝ վերցրած Kaggle-ից։ Տվյալները պահված են CSV ձևաչափով։

***Is it static or changing?***

Մեր ունեցած տվյալները ստատիկ են, որոնք ներկայացնում են ամբողջ ժամանակաշրջանի համար բոլոր երկրների պատմական տվյալները:

**Tools**

***What tool(s) are you going to use?***

Օգտագործել ենք Dash՝ ինտերակտիվ վիզուալիզացիաներ ստեղծելու համար:

***Why did you choose it?***

Dash-ն ապահովում է ճկուն և հզոր հարթակ տվյալների ինտերակտիվ վիզուալիզացիաներ ստեղծելու համար:

**Charts**

***What types of charts are choosen?***

Ընտրել ենք հետևյալ չարտերը․ Line charts, maps (interactive), bar charts, histogram, pie chart։

***Why are they the best to answer the question posted?***

Line chart -> լավագույնս է արտահայտում ժամանակի ընթացքում ահաբեկչական հարձակումների միտումները:

Map -> պատկերացում է տալիս տարբեր տարածաշրջաններում հարձակումների տեսակների հաճախականությունների վերաբերյալ:

Bar Charts -> հարձակումների, զոհերի և այլ կատեգորիկ տվյալների քանակի վիզուալ պատկերացման ամենահարմար chart-ն է:

Pie chart -> Սպանված մարդկանց քանակը ըստ տարածաշրջանների TOP-10-ը պատկերելու համար ենք ընտրել։

Հիստոգրամ -> ցույց է տալիս բաշխվածությունը։

**”Who did what”**

Ամբողջ աշխատանքը կատարվել է միասին՝ zoom-ի միջոցով, և ունենք բոլոր տեսակի ապացույցները՝ սկսած zoom-ի տեսանկարահանումներից, վերջացրած զանգերով (ֆայլը այնքան էր ծանրաբեռնում համակարգիչը, որ ստիպված էինք լինում zoom-ը անջատել և հեռախոսով զանգահարել միմյանց)։ Աշխատանքները կատարվել են երեկոյան ժամերին՝ մոտավորապես 20:30-03:35: Մեր ընտրած դատան բավականին ծավալուն է (մոտավորապես 180 հազ․ տող) և սկզբնական շրջանում մտածում էինք կրճատել և դիտարկել մեր տարածաշրջանի համար (այդ ժամանակ ունենում էինք մոտավոր 30 հազ տող), սակայն հասկացանք, որ իմաստը կորցնում է և ռեալ պատկերացում չի տալիս, ու, ամեն դեպքում, որոշեցինք սկզբնական բազան էլ պահել։ Ամենաշատը չարչարվել ենք Dash-ի ժամանակ callback-երը ստանալիս, բայց ստացանք։) Մեկ էլ ամեն մի տողից հետո ստիպված էինք անընդհատ restart անել, որպեսզի վերջին գրած կոդը հասկանար: Ունեցել ենք կարծիքների փոխանակում՝ կապված գծագրերի տեսքի, գույնների, ինչ-որ վիզուալիզացիաներ ներառելու կամ հանելու հետ կապված, ինչպես նաև ամենաշատ կարծիքները դատա ընտրելու հետ կապված էին։