Datu vizualizācija: kādēļ, kad un kā?

21.09.2011
Jurģis Bērziņš
jurgis.berzins@gmail.com

IEGULDĪJUMS TAVĀ NĀKOTNĒ!
AN INFOGRAPHIC IS

DATA

SORTED

ARRANGED

PRESENTED VISUALLY
Datu vizualizācija

Wikipedia: visual representation of data
Infografika

Wikipedia: graphic visual representations of information, data or knowledge
- Espresso [ess-pres-oh]
- Espresso Macchiato [ess-pres-oh mak-ee-ah-toe]
- Espresso con Panna [ess-pres-oh kon pahn-nah]
- Milk Foam
- Steamed Milk
- Espresso
- Steamed Milk
- Steamed Milk
- Espresso
- Steamed Milk
- Steamed Milk
- Espresso
- Steamed Half-and-Half
- Espresso
- Milk Foam
- Steamed Milk
- Espresso
- Whipped Cream
- Steamed Milk
- Chocolate Syrup
- Espresso
- Water
- Espresso
- Cappuccino [kepp-oo-chee-noh]
- Caffe Mocha [ca-fay-moh-kah]
- Americano [uh-mer-oh-kee-oh]
Vizualizācija

strukturēts,
grafisks
datu
attēlojums

Infografika

vizuāls stāsts:
dati+viedoklis
All About the Junk Mail

- A LOOK AT THE NUMBERS:
  Junk mail in the U.S. accounts for over 100,000,000,000 pieces of mail each year. (About 20% of all the mail delivered in the world.) The response rate is less than 2%.

- ACCORDING TO THE USPS:
  Direct mail volume over the years has snowballed.

- IS IT SUCCESSFUL?
  Approximately 44% of junk mail goes to landfills unopened.

- Households only average one personal correspondence each week compared to almost 18 pieces of junk mail.

- A response rate of < 0.25% is considered acceptable for the 500 million U.S. credit card solicitations that are mailed monthly.

- WHAT IS THE COST?
  It costs $550 million yearly to transport junk mail & $370 million annually to dispose of junk mail that doesn’t get recycled.

- 100 million trees are cut down to make junk mail.

- About 28,000,000,000 gallons of water are wasted to produce and transport junk mail.
THE INTERNET

PRIVACY

A HELPFUL VENN DIAGRAM
It doesn’t take a genius.

<table>
<thead>
<tr>
<th>IPHONE 5</th>
<th>GALAXY S III</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0” screen</td>
<td>4.8” screen</td>
</tr>
<tr>
<td>4G LTE</td>
<td>4G LTE</td>
</tr>
<tr>
<td>Retina Display</td>
<td>HD Super AMOLED™ Display</td>
</tr>
<tr>
<td>1136 X 640 resolution</td>
<td>1280 X 720 HD resolution</td>
</tr>
<tr>
<td>Up to 205 hrs standby time</td>
<td>Up to 720 hrs standby time</td>
</tr>
<tr>
<td>Up to 9 hrs talk time</td>
<td>Up to 11.4 hrs talk time</td>
</tr>
<tr>
<td>Full HD 1080p video recording</td>
<td>Full HD 1080p video recording</td>
</tr>
<tr>
<td>3.95 oz weight</td>
<td>4.7 oz weight</td>
</tr>
<tr>
<td>1GB RAM</td>
<td>2GB RAM</td>
</tr>
<tr>
<td>Siri</td>
<td>S Voice</td>
</tr>
<tr>
<td>16, 32 or 64GB fixed internal storage</td>
<td>16 or 32GB fixed + up to 94GB microSD storage</td>
</tr>
<tr>
<td>iOS 6.0 OS</td>
<td>Android 4.0 OS</td>
</tr>
<tr>
<td>A totally different plug</td>
<td>Standard micro USB plug</td>
</tr>
<tr>
<td></td>
<td>NFC</td>
</tr>
<tr>
<td></td>
<td>Smart Stay</td>
</tr>
<tr>
<td></td>
<td>S-Beam</td>
</tr>
<tr>
<td></td>
<td>ShareShot</td>
</tr>
<tr>
<td></td>
<td>Group Cast</td>
</tr>
<tr>
<td></td>
<td>Direct Call</td>
</tr>
<tr>
<td></td>
<td>Smart Alert</td>
</tr>
<tr>
<td></td>
<td>Tilt to Zoom</td>
</tr>
<tr>
<td></td>
<td>Palm Swipe Capture</td>
</tr>
<tr>
<td></td>
<td>Palm Touch Mute Pause</td>
</tr>
<tr>
<td></td>
<td>Picture In Picture</td>
</tr>
<tr>
<td></td>
<td>Team Over To Mute</td>
</tr>
<tr>
<td></td>
<td>Shake To Update</td>
</tr>
<tr>
<td></td>
<td>Removable battery</td>
</tr>
</tbody>
</table>


Kādēļ?
koncentrēts stāstījums
4. pielikums „Valsts pamatbudžeta ieņēmumu un izdevumu atšifrējums pa programmām un apakšprogrammām”

<table>
<thead>
<tr>
<th>Programmas, apakšprogrammas kods</th>
<th>Funkciju klasifikācijas kods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ieņēmumi – kopā</td>
<td>3 269 756 726</td>
</tr>
<tr>
<td>Resursi izdevumu segšanai</td>
<td>3 248 411 991</td>
</tr>
<tr>
<td>Ieņēmumi no maksas pakalpojumiem un citi pašu ieņēmumi – kopā</td>
<td>64 573 329</td>
</tr>
<tr>
<td>Ārvalstu finanšu palīdzība iestādes ieņēmumos</td>
<td>58 533 342</td>
</tr>
<tr>
<td>Transferti</td>
<td>113 917</td>
</tr>
<tr>
<td>Pašvaldību budžeta transferti</td>
<td>113 917</td>
</tr>
<tr>
<td>Valsts budžeta iestāžu saņemtie transferti no pašvaldībām</td>
<td>113 917</td>
</tr>
<tr>
<td>Valsts budžeta iestāžu saņemtie transferti (izņemot atmaksas) no pašvaldībām</td>
<td>113 917</td>
</tr>
<tr>
<td>Dotācija no vispārējiem ieņēmumiem</td>
<td>3 125 191 403</td>
</tr>
<tr>
<td>Vispārējā kārtībā sadaļmā dotācija no vispārējiem ieņēmumiem</td>
<td>3 125 191 403</td>
</tr>
<tr>
<td>Izdevumi – kopā</td>
<td>3 257 060 183</td>
</tr>
<tr>
<td>Uzturēšanas izdevumi</td>
<td>2 904 090 139</td>
</tr>
<tr>
<td>Kārtējie izdevumi</td>
<td>832 268 347</td>
</tr>
<tr>
<td>Atļaidība</td>
<td>479 253 539</td>
</tr>
<tr>
<td>Atalgojums</td>
<td>359 813 706</td>
</tr>
<tr>
<td>Preces un pakalpojumi</td>
<td>353 014 808</td>
</tr>
<tr>
<td>Procentu izdevumi</td>
<td>233 782 825</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Likuma "Par valsts budžetu 2012.gadam"
saprotams stāstījums
Espresso [ess-press-oh]
Espresso Macchiato [ess-press-oh mock-e-ah-toe]
Espresso con Panna [ess-press-oh kon pawn-nah]
Milk Foam [milk foem]
Steamed Milk [stee-mird milk]
Espresso [ess-press-oh]
Caffé Latte [cafi lah-tey]
Flat White [flat wht]
Cafe Breve [cafi brev]
Milk Foam [milk foem]
Steeamed Milk [stee-mird milk]
Espresso [ess-press-oh]
Whipped Cream [wif-peed cram]
Espresso [ess-press-oh]
Cappuccino [kapp-oh chee-noh]
Caffé Mocha [cafi moh-kah]
Americano [uh-mer-oh mahn-oh]
kontekstualizēts stāstījums
Internets Latvijā

Trīs apmeklētākie Latvijas interneta resursi kopā apkalpo 1 129 917*

lietotāju mēnesī jeb 93% no visiem Latvijas interneta lietotājiem. Kā savu auditoriju savā starpā sadala šīs vietnes?

Interneta lietotāji Latvijā, kopā 1 273 559 lietotāju mēnesī

- inbox.lv
  822 363 lietotāju mēnesī jeb 64.57%
  no visiem Latvijas interneta lietotājiem

- draugiem.lv
  772 724 lietotāju mēnesī jeb 60.67%

- delfi.lv
  652 610 lietotāju mēnesī jeb 51.24%

*Dati: 2010. gada decembris, no pētījuma GemiusAudience, vecuma grupa: 15-74 gadi
Kad?
liels datu apjoms
struktūras, procesi, notikumi
"Congress shall have Power...to make all Laws which shall be necessary and proper."

U.S. Constitution Article I Section 8
tabulas
Kukuļu apjomi (saskaitot visus vienā krimināllīrietā faktiski dotos vai ķemtos kukuļus)

*Kukuļošanas (ķemšana, piesavināšanās, starpniecība, došana) krimināllīetu skaits, kur uz 31.10.2010 pēdējais nolēmums notiesājošs

© PROVIDUS Korupcijas F. Jurģis Bērziņš
Kā?
Dažādi vizualizāciju veidi
In 2009 there were 9,219 reported crashes involving 126,717 drivers:

195 crashes occurred each day
8 crashes occurred every hour

100 people were injured each day
1 crash every 7 minutes

Of all reported motor vehicle accidents on Oklahoma state highways:

72% occurred in daylight
9% occurred between the hours of 5 p.m. and 6 p.m.

48% involved two vehicles
17% occurred on Friday

COMMON CAR ACCIDENT INJURIES:
PERCENT OF RISK TO THE BODY

FRONTAL COLLISIONS

Chest 53%
Abdomen 19%
Pelvis 17%
Head 35%
The Costs and Benefits of a Green Revolution

Organic refers to the way agricultural products are grown and processed. It includes a system of production, processing, distribution and sales that assures consumers that the products maintain the organic integrity that begins on the farm.

- **Newly Organic families** have only just begun purchasing organic products within the past 2 years.
- **20% Experienced Organic** first bought organic products up to 5 years ago.
- **21% Seasoned Organic** have been buying organic products for more than 5 years, and in some cases, longer than 15 years.

73% of US families have purchased organic products.

US Organic Food Market Value: $ Billion

- Compound Annual Growth Rate 2005-2009: 16.8%
- Projected Compound Annual Growth Rate 2009-2014: 10.3%
Data visualization is a popular new way of sharing research. Here is a look at some of the visual devices, informational elements, and general trends found in the modern day infographic.

**Design**

**Chart Style**
- Percentage of infographics with the following charts:
  - Pie Chart: 22%
  - Pictorial Chart: 24%
  - Line Chart: 24%
  - Bar Chart: 32%

**Font**
- Sans Serif: 85%
- Condensed Sans Serif: 15%

**Key Info**
- Percentage of infographics with key: 33%
- Average number of symbols per key: 5.1

**Base Color**
- Green: 29%
- Yellow: 18%
- Red: 18%
- Pink: 10%
- Orange: 10%

**Navigational Iconography**
- Frequency of arrows & connecting lines in infographics:
  - Arrows: 13%
  - Lines: 38%
  - Both: 13%

**Content**

**Countries Featured**
- United States: 88%
- China: 22%
- United Kingdom: 12%
- Australia: 10%
- Canada: 10%
- India: 10%
- France: 10%
- Mexico: 8%

**Sections**
- Average number of sections per infographic: 2.12

**Credited Sources**
- Average number of sources per infographic: 2.29

**Title**
"Richest and Poorest American Neigh..."
Ouch?
The biggest commercial fine in US history

- Pfizer, the world's largest drug company, maker of Viagra.
- Fined $2.3 bn for promoting withdrawn painkiller Bextra for uses not approved by medical regulators.

$173.2 billion Pfizer Revenue 2002-2005 (years of Bextra's release)

$32.3 Profit 02-05

$2.3 Fine

David McCandless // InformationIsBeautiful.net
source: Wikipedia / The Guardian
Geometrija: izmēru salīdzinājums
WHERE IS THE WORLD’S LABOR?

Brazil
- Industry: 19%
- Agriculture: 19%
- Services: 52%
- Unemployed: 10%

South Africa
- Industry: 18%
- Agriculture: 8%
- Services: 47%
- Unemployed: 27%

China
- Industry: 22%
- Agriculture: 45%
- Services: 29%
- Unemployed: 4%

Geomērija: izmēru salīdzinājums
History of Social Media

Social media has become an integral part of modern society. These are general social networks with user bases larger than the population of most countries. There are niche sites for virtually every special interest out there. These are sites to share photos, videos, status updates, sites for meeting new people and sites to connect with old friends. There are social solutions to just about every need.

Laika līnija

*The World of Internet: AUG/SRS*

Laika līnija

*The World of Internet: AUG/SRS*

Laika līnija

*The World of Internet: AUG/SRS*

Laika līnija

*The World of Internet: AUG/SRS*
How google works

Google bots follow links. If there are no links to your site, typically it will not get crawled deeply or regularly.

Google bots won't crawl your site if you tell them not to with a robots.txt file.

If links to your site have a nofollow tag, the google bots won't arrive from those links.

Google may also find your site by being pinged by blog software or xml sitemaps.

The more links you have from higher authority pages, the greater your own pages' authority.

As long as they are not tagged 'nofollow'.

Pages are checked against editorial policies.

Google estimates the domain and page's overall authority based on links.

Once crawled, the page is indexed within seconds.

Excursion: Pages are checked against editorial policies.

You write a blog post, tweet, update a site, or otherwise add content to the web.

Google bots crawling the web come across your post.

If there are no links to your site, typically it will not get crawled deeply or regularly.

Google bots won't crawl your site if you tell them not to with a robots.txt file.

Websites are indexed. If they have a nofollow tag, the google bots won't arrive from those links.

Google may also find your site by being pinged by blog software or xml sitemaps.

The more links you have from higher authority pages, the greater your own pages' authority.

As long as they are not tagged 'nofollow'.

Pages are checked against editorial policies.

Google estimates the domain and page's overall authority based on links.

Once crawled, the page is indexed within seconds.

You write a blog post, tweet, update a site, or otherwise add content to the web.

Google bots crawling the web come across your post.

If there are no links to your site, typically it will not get crawled deeply or regularly.

Google bots won't crawl your site if you tell them not to with a robots.txt file.

If links to your site have a nofollow tag, the google bots won't arrive from those links.

Google may also find your site by being pinged by blog software or xml sitemaps.

The more links you have from higher authority pages, the greater your own pages' authority.

As long as they are not tagged 'nofollow'.

Pages are checked against editorial policies.

Google estimates the domain and page's overall authority based on links.

Once crawled, the page is indexed within seconds.

You write a blog post, tweet, update a site, or otherwise add content to the web.

Google bots crawling the web come across your post.

If there are no links to your site, typically it will not get crawled deeply or regularly.

Google bots won't crawl your site if you tell them not to with a robots.txt file.

If links to your site have a nofollow tag, the google bots won't arrive from those links.

Google may also find your site by being pinged by blog software or xml sitemaps.

The more links you have from higher authority pages, the greater your own pages' authority.

As long as they are not tagged 'nofollow'.

Pages are checked against editorial policies.

Google estimates the domain and page's overall authority based on links.

Once crawled, the page is indexed within seconds.

You write a blog post, tweet, update a site, or otherwise add content to the web.

Google bots crawling the web come across your post.

If there are no links to your site, typically it will not get crawled deeply or regularly.

Google bots won't crawl your site if you tell them not to with a robots.txt file.

If links to your site have a nofollow tag, the google bots won't arrive from those links.

Google may also find your site by being pinged by blog software or xml sitemaps.

The more links you have from higher authority pages, the greater your own pages' authority.

As long as they are not tagged 'nofollow'.

Pages are checked against editorial policies.

Google estimates the domain and page's overall authority based on links.

Once crawled, the page is indexed within seconds.

You write a blog post, tweet, update a site, or otherwise add content to the web.
Venn diagram

Pancakes!

Flour

Pasta

Egg

Batter

Omelette

Milk

StephenWildish.co.uk
HOW WOULD YOU LIKE YOUR GRAPHIC DESIGN?

(YOU MAY PICK TWO)

Venn diagram

FAST

DIPPED IN UGLY SAUCE WITH HASTE AND CARELESSNESS

CHEAP

IMPOSSIBLE UTOPIA

JUST IN TIME TO BE TOO LATE

FREE

TRASH

GO AWAY

TRY AGAIN

GREAT

YOU GET WHAT YOU PAY FOR
Kartēšana/Scatter plot
Excel diagrammas: skalu jautājums
Figure 11: GNH index by gender

World Happiness Report, 2012
IF BUSH TAX CUTS EXPIRE

TOP TAX RATE

NOW: 35%

JAN. 1, 2013: 39.6%
Datu izvēles principi
Saprotamība
Konsekvence
Konteksts
Saprotamiba
Konsekvensce
Konteksts