

In Search of Emerald City

Kirill Semenov, Daria Shaporenko
DH minor, HSE
Moscow, 2018



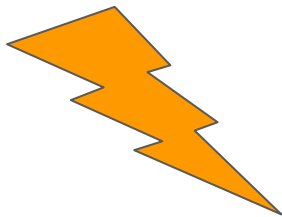
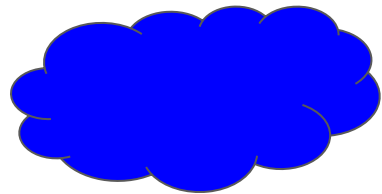
Sample: Colors

BLUE

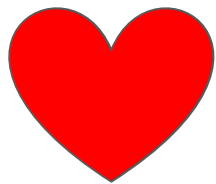
RED

YELLOW

GREEN



WHITE



BLACK

Sample: Colours (now serious)

- red
- blue
- green
- yellow
- white
- black

Sample: Languages

- English:
 - the UK, the US, Canada, Australia, South Africa
- German:
 - Germany, Austria, Switzerland, Liechtenstein, Luxembourg
- Spanish:
 - Spain, 10 countries of Latin America, Equatorial Guinea, the US
- Italian:
 - Italy, Switzerland, San-Marino
- Russian:
 - Russia, Belarus, Transnistria (partially)
- Chinese:
 - China (PRC with Hong Kong and Macau), Taiwan, Singapore

青色 is for this colour...



...and for this one. Feel difference?



And even for both of these!



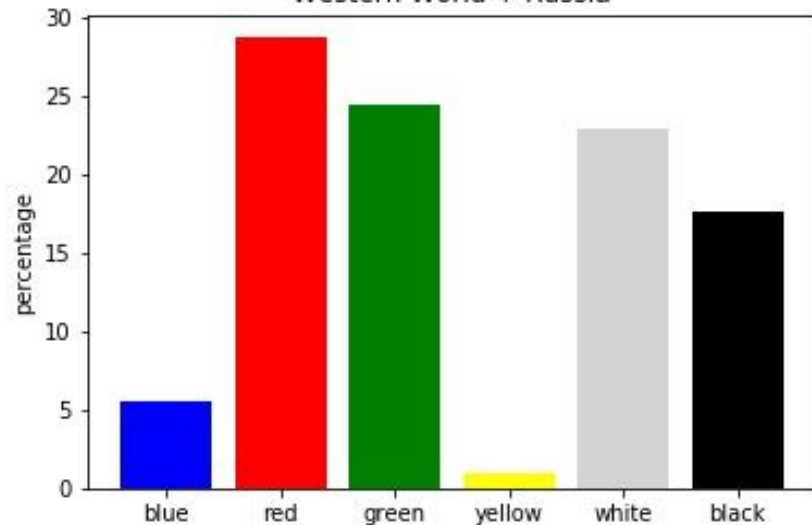
no, that's a joke

Comparison: color-to-language

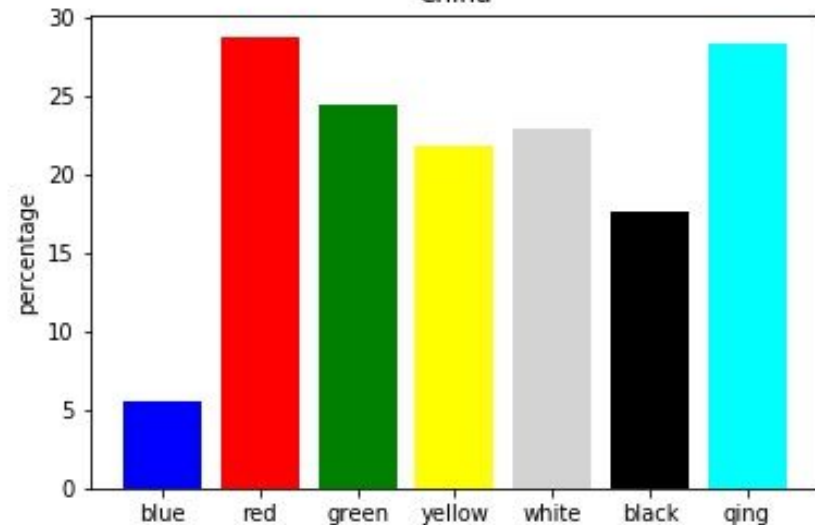
	red	blue	yellow	green	white	black	total
Russian	1626	153	19	144	814	509	3265
English	431	252	40	1492	553	467	3235
German	129	8	6	228	159	175	705
Spanish	16	31	10	70	173	85	385
Italian	164	11	5	74	162	209	625
Chinese (- qing)	25	3	23	1	21	4	77
total	2391	458	103	2009	1882	1449	8292

Comparison: Chinese VS other languages

Western World + Russia



China



Observations

Spanish world: the least number of 'colored' settlements in our sample (excluding China)

Spanish & Italian: very fond of BLACK and WHITE

Spanish & English: almost no BLUE in metropolies

Russian: Belarus is more fond of BLACK, WHITE and BLUE

English: the biggest number of GREEN across the world

Italian & German: the least difference between RED, GREEN, WHITE and BLACK

Observations (2)

- Russian: the biggest number of RED across the world
- RED is the most popular colour in the sample (probably due to communism in Russia and the former USSR)
- YELLOW is the least popular one
 - is yellow less 'basic'?..

Transnistria: 8 “red” settlements of 110 as a whole

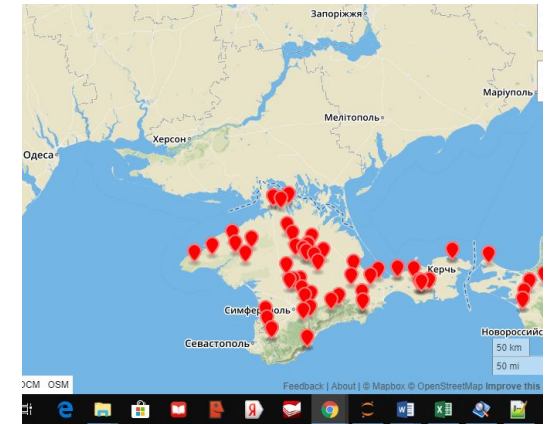
Similar to Kaliningrad Oblast and Crimea



Kaliningrad Oblast

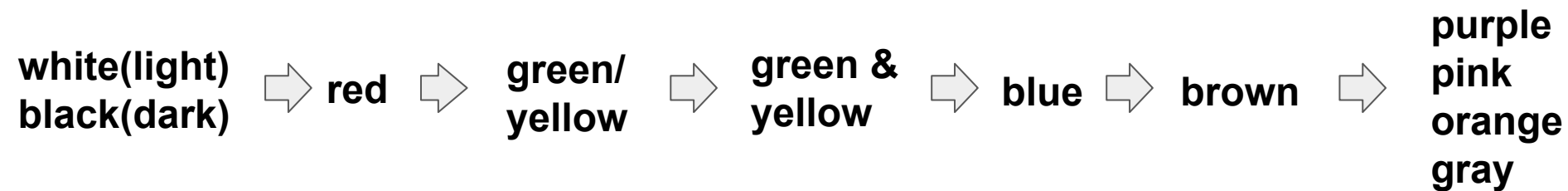


Transnistria

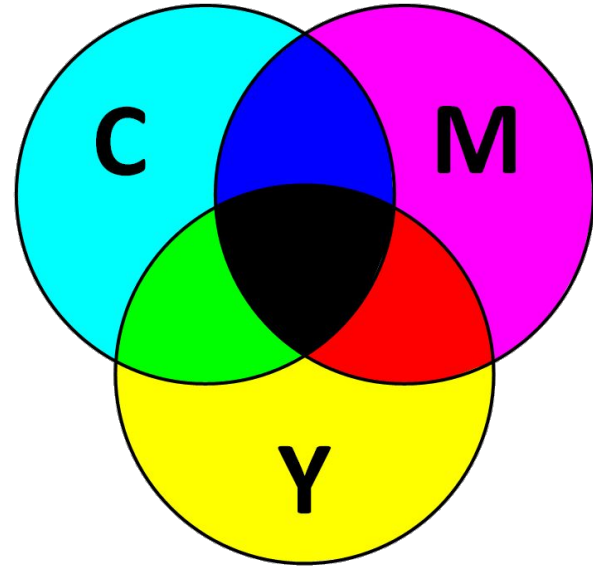
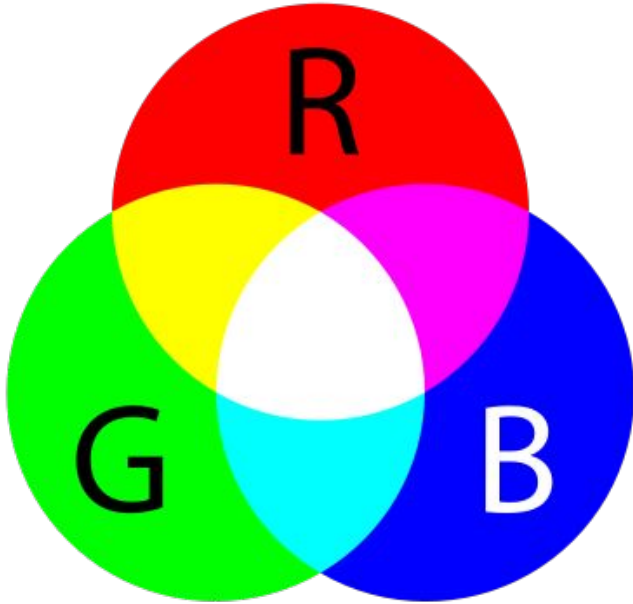


Crimea

Classical approach - Berlin & Kay 1969



RGB-languages VS CMYK-languages?



Problems

- Europe+Russia: homonyms for some colours in some countries
 - more exact regex
 - manual check
 - multilingual sample
- China: little dataset
 - use of English, Chinese traditional and simplified regex
 - manual adding of cities which could be found via Google.maps
 - China has its own analogue of Wikipedia with no access to its databases
- Settlements with unclear etymology
 - Vienna - a theory that the name is from Vindobona which is from a Celtic word for 'white' but still not clear

Future Horizons

- more precise samples with local languages:
 - Russia: “Красный” + “Кызыл” + “Улан” + “Ал...” +etc.
- Disambiguation of words with the same root:
 - Russia: “Синицы” - not blue, but a bird
 - China: 黄林 can be for 黄河+林
- Study on symbolism of each colour in different countries
 - Is BLUE a color of christianity?
 - Is RED really a color of communists?
- Sample of Muslim countries
 - will GREEN become as popular as other colors?

Thank you for your attention!

